Groupe de travail Réseau
**Request for Comments : 3618**
Catégorie : Expérimentale
Traduction Claude Brière de L'Isle

B. Fenner, éditeur
D. Meyer, éditeur
octobre 2003

# Protocole de découverte de source de diffusion groupée (MSDP)

**Statut de ce mémoire**
Le présent mémoire définit un protocole expérimental pour la communauté de l'Internet. Il ne spécifie aucune forme de norme de l'Internet. On invite à des discussions et suggestions pour son amélioration. La distribution du présent mémoire n'est soumise à aucune restriction.



**Résumé**
Le protocole de découverte de source de diffusion groupée (MSDP, *Multicast Source Discovery Protocol*) décrit un mécanisme pour connecter ensemble plusieurs domaines en mode épars de diffusion groupée indépendante du protocole (PIM-SM) IP version 4. Chaque domaine PIM-SM utilise son propre point de rendez-vous (RP, *Rendezvous Point*) indépendant et n'a pas à dépendre des RP dans d'autres domaines. Le présent document reflète les mises en œuvre MSDP existantes.

## Table des Matières

1. Introduction....2
2. Vue générale....2
3. Procédure....2
4. Mise en antémémoire....3
5. Temporisateurs....3
5.1 Temporisateur d'annonce de SA....3
5.2 Traitement du temporisateur d'annonce de SA....3
5.3 Fin de temporisation d'antémémoire de SA (temporisateur État de SA)....4
5.4 Temporisateur de garde d'homologue....4
5.5 Temporisateur Garde en vie....4
5.6 Temporisateur Essais de connexion....4
6. Homologues MSDP intermédiaires....4
7. Filtrage de SA et politique....4
8. Paquets de données encapsulés....5
9. Autres scénarios....5
10. Transmission à l'homologue MSDP sur le chemin inverse....5
10.1 Définitions....5
10.2 Sémantique de groupe MSDP maillé....6
11. Automate à états de connexion MSDP....6
11.1 Événements....7
11.2 Actions....7
11.3 Événements spécifiques de l'homologue....7
11.4 Événements indépendants de l'homologue....7
12. Formats de paquet....8
12.1 Formats de TLV MSDP....8
12.2 TLV définis....8
13. Traitement d'erreur MSDP....10
14. Encapsulation des données de SA....10
15. Déclaration d'applicabilité....10
15.1 Entre domaines PIM....10
15.2 Entre des RP en envoi à la cantonade....10
16. Propriété intellectuelle....10
17. Remerciements....11
18. Considérations sur la sécurité....11
19. Considérations relatives à l'IANA....11
19.1 Gamme de TLV alloués par l'IANA....11

19.2 Gamme expérimentale de TLV....11
20. Références....11
20.1 Références normatives....11
20.2 Références pour information....12
21. Adresses des éditeurs....12
22. Déclaration complète de droits de reproduction....12

## 1. Introduction

Le protocole de découverte de source de diffusion groupée (MSDP, *Multicast Source Discovery Protocol*) décrit un mécanisme pour connecter ensemble plusieurs domaines PIM en mode épars (PIM-SM) [RFC2362]. Chaque domaine PIM-SM utilise ses propres points de rendez-vous (RP) indépendants et n'a pas à dépendre des RP dans les autres domaines. Les avantages de cette approche incluent :

- Pas de dépendance aux ressources de tiers sur le RP d'un domaine ; les domaines PIM-SM peuvent s'appuyer seulement sur leur propre RP.
- Domaines receveurs seulement ; les domaines avec seulement des receveurs obtiennent les données sans annoncer mondialement les membres du groupe.

Noter que MSDP peut être utilisé avec des protocoles autres que PIM-SM, mais un tel usage n'est pas spécifié dans le présent mémoire.

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

## 2. Vue générale

Les routeurs à capacité MSDP dans un domaine PIM-SM ont une relation d'homologue MSDP avec leurs homologues MSDP dans un autre domaine. La relation d'homologue à homologue est constituée par une connexion TCP dans laquelle sont échangées les informations de contrôle. Chaque domaine a une ou plusieurs connexions à cette topologie virtuelle.

L'objet de cette topologie est de permettre aux domaines de découvrir les sources de diffusion groupée provenant d'autres domaines. Si les sources de diffusion groupée intéressent un domaine qui a des receveurs, le mécanisme normal de construction d'arborescence de source dans PIM-SM sera utilisé pour livrer des données en diffusion groupée sur une arborescence de distribution inter domaines.

## 3. Procédure

Lorsque un RP dans un domaine PIM-SM a connaissance pour la première fois d'un nouvel envoyeur, par exemple, via des messages Register PIM, il construit un message "Source-Active" (SA) et l'envoie à ses homologues MSDP. Tous les RP, qui ont l'intention de générer ou recevoir des messages SA doivent établir des relations d'homologue MSDP avec les autres RP, soit directement, soit via un homologue MSDP intermédiaire. Le message SA contient les champs suivants :

- Adresse de source de la source des données.
- Adresse de groupe où la source des données envoie.
- Adresse IP du RP.

Noter qu'un RP qui n'est pas un routeur désigné (DR, *Designated Router*) sur un réseau partagé NE DEVRAIT PAS générer de SA pour des sources directement connectées sur ce réseau partagé ; il devrait seulement en générer en réponse à la réception de messages Register provenant du DR.

Chaque homologue MSDP reçoit et transmet le message à partir de l'adresse de RP de la façon "arrosage d'homologue en transmission sur le chemin inverse (RPF, *reverse path forwarding*)". La notion d'arrosage d'homologue RPF est par rapport à la transmission des messages SA. La base de données d'informations d'acheminement sur le chemin inverse en diffusion groupée (MRIB, *Multicast RPF Routing Information Base*) est examinée pour déterminer quel homologue est choisi vers le RP générateur du message SA. Un tel homologue est appelé un "homologue RPF". Voir à la Section 13 les détails de la transmission d'homologue sur le chemin inverse (peer-RPF).

Si l'homologue MSDP reçoit le SA d'un homologue non RPF vers le RP générateur, il va abandonner le message. Autrement, il transmet le message à tous ses homologues MSDP (sauf celui de qui il a reçu le message SA).

Lorsque un homologue MSDP qui est aussi un RP pour son propre domaine reçoit un nouveau message SA, il détermine si il y a des membres du groupe au sein du domaine intéressés par un groupe décrit par une entrée (Source, Groupe), ou (S,G) au sein du message SA. C'est-à-dire, le RP vérifie qu'il y a une entrée (*,G) avec une liste d'interfaces sortantes non vide ; cela implique que certains systèmes dans le domaine sont intéressés par le groupe. Dans ce cas, le RP déclanche un événement Join (S,G) vers la source des données comme si un message Join/Prune avait été reçu, adressé au RP lui-même. Cela établit une branche de l'arborescence de source pour ce domaine. Les paquets de données suivants arrivent au RP via cette branche de l'arborescence, et sont retransmis à l'intérieur du domaine de long de l'arborescence partagée. Si des routeurs latéraux choisissent de se joindre à l'arborescence de source, ils ont l'option de le faire conformément aux conventions PIM-SM existantes. Finalement, si un RP dans un domaine reçoit un message PIM Join pour un nouveau groupe G, le RP DEVRAIT déclancher un événement Join (S,G) pour chaque (S,G) actif pour ce groupe dans son antémémoire de SA.

Cette procédure a été affectueusement nommée arrosage et jonction *(flood-and-join)* parce que si un RP n'est pas intéressé par le groupe, il peut ignorer le message SA. Autrement, il se joint à une arborescence de distribution.

## 4. Mise en antémémoire

Un locuteur MSDP DOIT mettre en antémémoire les messages SA. La mise en antémémoire permet de réduire le rythme des messages MSDP ainsi que la latence de jonction pour les nouveaux receveurs d'un groupe G à un RP générateur qui a déjà un état MSDP (S,G) existant. De plus, la mise en antémémoire aide beaucoup au diagnostic et au débogage de divers problèmes.

Un locuteur MSDP doit fournir un mécanisme pour réduire la transmission des nouveaux SA. L'antémémoire de SA est utilisée pour réduire les tempêtes et elle le fait en ne transmettant pas les SA sauf si ils sont dans l'antémémoire ou sont de nouveaux paquets SA que le locuteur MSDP va mettre en antémémoire pour la première fois. L'antémémoire de SA réduit aussi les tempêtes en annonçant à partir de l'antémémoire selon une périodicité de pas plus que deux fois par intervalle Tempo-Annonce-SA et pas moins que une fois par période d'annonce de SA.

## 5. Temporisateurs

Les principaux temporisateurs pour MSDP sont : Tempo-Annonce-SA, temporisateur Entrée d'antémémoire de SA, temporisateur Garde d'homologue, temporisateur Garde en vie, et temporisateur Essai de connexion. Chacun est décrit ci-dessous.

### 5.1 Temporisateur d'annonce de SA

Les RP qui génèrent des messages SA le font périodiquement pour autant que des données sont envoyées par la source. Il y a un Tempo-Annonce-SA qui couvre les sources qu'un RP peut annoncer. [Période d'annonce de SA] DOIT être de 60 secondes. Un RP NE DOIT PAS envoyer plus d'un message SA périodique pour une certaine (S,G) dans un intervalle d'annonce de SA. Générer des messages SA périodiques est nécessaire pour garder les annonces en vie dans les antémémoires. Finalement, un RP générateur DEVRAIT déclancher la transmission d'un message SA aussitôt qu'il reçoit des données d'une source interne pour la première fois. Ce message SA initial peut être en plus du message SA périodique transmis dans les 60 premières secondes pour cette (S,G).

### 5.2 Traitement du temporisateur d'annonce de SA

Un RP DOIT étaler la génération de messages SA périodiques (c'est-à-dire, messages annonçant les sources actives pour lesquelles il est le RP) sur son intervalle de rapports (c'est-à-dire, Période d'annonce de SA). Un RP lance le Tempo-Annonce-SA lorsque le processus MSDP est configuré. Lorsque le temporisateur arrive à expiration, un RP rétablit le temporisateur à [Période d'annonce de SA] secondes, et commence les annonces de ses sources actives. Les sources actives sont annoncées de la manière suivante : un RP empaquette ses sources actives dans un message SA jusqu'à ce que le plus gros paquet MSDP qui puisse être envoyé soit construit ou qu'il n'y ait plus de sources, et il envoie alors le message. Ce processus est répété périodiquement dans la période d'annonce de SA d'une façon telle que toutes les sources du RP soient annoncées. Noter que comme MSDP est un protocole périodique, une mise en œuvre DEVRAIT envoyer tous les messages SA en antémémoire lors de l'établissement d'une connexion. Finalement, le temporisateur est supprimé lorsque le processus MSDP est déconfiguré.

### 5.3 Fin de temporisation d'antémémoire de SA (temporisateur État de SA)

Chaque entrée dans une antémémoire de SA a un temporisateur d'état de SA associé. Un temporisateur d'état de SA (S,G) est lancé lorsque un message SA (S,G) est reçu initialement par un homologue MSDP. Le temporisateur est remis à [Période d'état SG] si un autre message SA (S,G) est reçu avant que le temporisateur d'état de SA (S,G) arrive à expiration. [Période d'état SG] NE DOIT PAS être inférieur à [Période d'annonce de SA] + [Période de garde de SA].

### 5.4 Temporisateur de garde d'homologue

Le temporisateur de garde est initialisé à [Période de temps de garde] lors de l'établissement de la connexion de transport de l'homologue, et est remis à [Période de temps de garde] lorsque un message MSDP est reçu. Finalement, le temporisateur est supprimé lorsque la connexion de transport de l'homologue est fermée. [Période de temps de garde] DOIT être d'au moins trois secondes. La valeur recommandée pour [Période de temps de garde] est 75 secondes.

### 5.5 Temporisateur Garde en vie

Une fois qu'une connexion de transport MSDP est établie, chaque côté de la connexion envoie un message Garder en vie et établit un temporisateur de garde en vie. Si le temporisateur de garde en vie arrive à expiration, le système local envoie un message Garde en vie et relance son temporisateur de garde en vie.

Le temporisateur de garde en vie est réglé à [Période de garde en vie] lorsque l'homologue s'active. Le temporisateur est remis à [Période de garde en vie] chaque fois qu'un message MSDP est envoyé à l'homologue, et remis à zéro lorsque le temporisateur arrive à expiration.

Finalement, le temporisateur de garde en vie est supprimé lorsque la connexion de transport de l'homologue est fermée.

[Période de garde en vie] DOIT être inférieur à [Période de temps de garde], et DOIT être d'au moins une seconde. La valeur recommandée pour [Période de garde en vie] est 60 secondes.

### 5.6 Temporisateur Essais de connexion

Le temporisateur Essais de connexion est utilisé par l'homologue MSDP avec la plus faible adresse IP pour passer de l'état INACTIF à CONNEXION. Il y a un temporisateur par homologue, et la [Période d'essai de connexion] DEVRAIT être réglée à 30 secondes. Le temporisateur est initialisé à [Période d'essai de connexion] lorsque un locuteur MSDP tente activement d'ouvrir une connexion TCP avec son homologue (voir la Section 15, événement E2, action A2). Lorsque le temporisateur arrive à expiration, l'homologue réessaye la connexion et le temporisateur est remis à [Période d'essai de connexion]. Il est supprimé si la connexion passe à l'état ÉTABLI ou si l'homologue est déconfiguré.

## 6. Homologues MSDP intermédiaires

Les locuteurs MSDP intermédiaires ne génèrent pas de message SA périodique au nom des sources dans les autres domaines. En général, un RP DOIT seulement générer un SA pour une source qui va s'enregistrer auprès de lui, et SEULS les RP peuvent générer des messages SA. Les locuteurs MSDP intermédiaires PEUVENT transmettre les messages SA reçus des autres domaines.

## 7. Filtrage de SA et politique

Lorsque le nombre de paires (S,G) augmente dans l'Internet, un RP peut vouloir filtrer les sources qu'il décrit dans les messages SA. Aussi, le filtrage peut être utilisé comme moyen de politique qui peut en même temps réduire l'état. Les homologues MSDP dans les domaines de transit ne devraient pas filtrer les messages SA car le modèle d'arrosage et jonction ne peut pas garantir que les sources seront connues tout autour de l'Internet (c'est-à-dire, le filtrage de SA par les domaines de transit peut causer un manque de connexité non désiré). En général, la politique devrait être exprimée en utilisant MBGP [RFC2858]. Cela va causer autrement l'échec de l'écoulement des messages MSDP dans la direction désirée et de l'acheminement d'homologue dans la direction inverse. Une exception se produit à une limite de portée administrative [RFC2365]. En particulier, un message SA pour une (S,G) NE DOIT PAS être envoyé aux homologues qui sont de l'autre côté d'une limite de portée administrative pour G.

## 8. Paquets de données encapsulés

Le RP PEUT encapsuler des données en diffusion groupée provenant de la source. Un RP intéressé peut désencapsuler le paquet, qui DEVRAIT être transmis comme si était reçu un paquet encapsulé PIM Register. C'est-à-dire que si des paquets arrivent déjà sur l'interface avec la source, les paquets sont alors éliminés. Autrement, si la liste des interfaces sortantes n'est pas nulle, le paquet est transmis de façon appropriée. Noter que quand on fait l'encapsulation des données, une mise en œuvre DOIT limiter le temps pendant lequel les paquets sont encapsulés.

Cela permet de recevoir de petites salves avant que l'arborescence de diffusion groupée ne soit reconstruite vers le domaine de la source. Par exemple, une mise en œuvre DEVRAIT encapsuler au moins le premier paquet pour fournir le service aux sources saccadées.

## 9. Autres scénarios

MSDP ne se limite pas à un déploiement à travers différents domaines d'acheminement. Il peut être utilisé au sein d'un domaine d'acheminement lorsque on désire déployer plusieurs RP pour les mêmes gammes de groupes comme avec les RP en envoi à la cantonade. Tant que tous les RP ont une topologie MSDP interconnectée, chacun peut apprendre les sources actives ainsi que les RP dans les autres domaines.

## 10. Transmission à l'homologue MSDP sur le chemin inverse

Les règles de transmission à l'homologue MSDP sur le chemin inverse sont utilisées pour transmettre les messages SA à travers un internet à capacité MSDP. À la différence de la vérification de RPF utilisée lors de la transmission de paquets de données, qui compare généralement l'adresse de source du paquet à l'interface sur laquelle le paquet a été reçu, la vérification de transmission à l'homologue sur le chemin inverse *(Peer-RPF)* compare l'adresse de RP portée dans le message SA à l'homologue MSDP duquel le message a été reçu.

### 10.1 Définitions

Les définitions suivantes sont utilisées dans la description des règles de transmission à l'homologue sur le chemin inverse :

#### 10.1.1 Base de données d'informations d'acheminement en diffusion groupée sur le chemin inverse

La base de données d'informations d'acheminement en diffusion groupée sur le chemin inverse (MRIB, *Multicast RPF Routing Information Base*) est le tableau de la topologie de la diffusion groupée. Elle est normalement déduite du tableau d'acheminement en envoi individuel ou d'autres protocoles d'acheminement tels que BGP multi protocoles [RFC2858].

#### 10.1.2 Chemin d'homologue RPF

Le chemin d'homologue RPF est celui que la MRIB choisit pour une certaine adresse. Le chemin d'homologue RPF pour un RP générateur d'un SA est utilisé pour choisir l'homologue à partir duquel le SA est accepté.

#### 10.1.3 Règles de transmission d'homologue RPF

Un message SA généré par R et reçu par X de N est accepté si N est l'homologue RPF voisin pour X, et est éliminé autrement.

```
      MPP(R,N)              MP(N,X)
R ----------....-------> N ------------------> X
      SA(S,G,R)             SA(S,G,R)
```

MP(N,X) est une relation d'homologue MSDP entre N et X. MPP(R,N) est dans un chemin de relation d'homologue MSDP (zéro, un ou plusieurs homologues MSDP) entre R et N, par exemple, MPP(R,N) = MP(R, A) + MP(A, B) + MP(B, N). SA(S,G,R) est un message SA pour la source S sur le groupe G généré par un RP R.

Le voisin d'homologue RPF N est choisi de façon déterministe, en utilisant la première des règles suivantes qui correspond. En particulier, N est le voisin RPF de X par rapport à R si

(i) N == R (X a une relation d'homologue MSDP avec R) ;
(ii) N est le prochain bond eBGP du chemin d'homologue RPF pour R ;
(iii) le chemin d'homologue RPF pour R est appris par un protocole d'acheminement par vecteur de distance ou de chemin (par exemple, BGP, RIP, DVMRP) et N est le voisin qui annonce le chemin d'homologue RPF pour R (par exemple, N est le conseilleur iBGP du chemin pour R) ou N est le prochain bond IGP pour R si le chemin pour R est appris via un protocole d'état de liaison (par exemple, OSPF [RFC2328] ou IS-IS [RFC1142]) ;
(iv) N réside dans le plus proche AS dans le meilleur chemin vers R. Si plusieurs homologues MSDP résident dans le plus proche AS, l'homologue avec la plus forte adresse IP est l'homologue RPF ;
(v) N est configuré comme l'homologue RPF statique pour R.

Les homologues MSDP, qui NE SONT PAS dans l'état ÉTABLI (c'est-à-dire, les homologues non actifs) ne sont pas éligibles pour la prise en compte de l'homologue RPF.

### 10.2 Sémantique de groupe MSDP maillé

Un groupe MSDP maillé est un mécanisme de fonctionnement pour réduire l'arrosage de SA, normalement dans un réglage intra domaine. En particulier, lorsque un sous ensemble des locuteurs MSDP d'un domaine est complètement maillé, il peut être configuré en un groupe maillé.

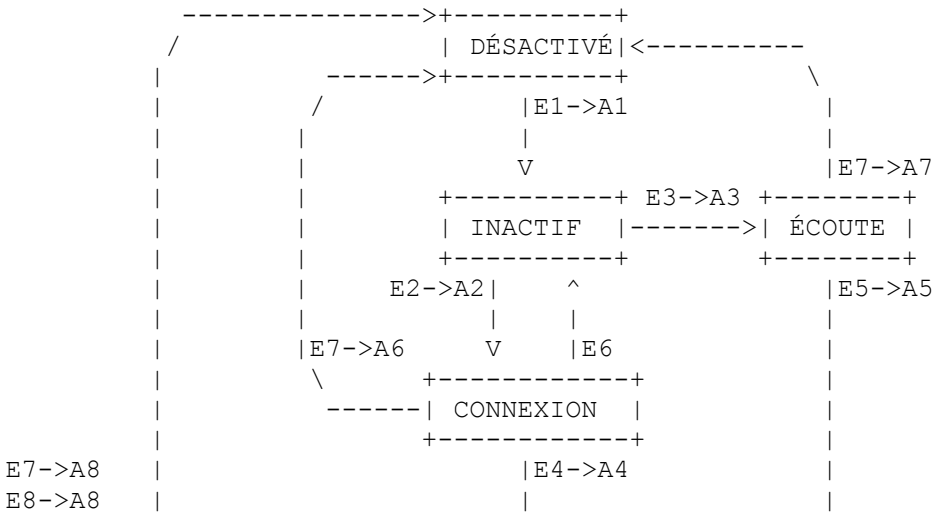
Noter que les groupes maillés supposent qu'un membre n'a pas à transmettre un SA aux autres membres du groupe maillé parce que le générateur va transmettre à tous les membres. Pour que le générateur soit capable de transmettre à tous les membres (et que chaque membre soit aussi un générateur potentiel) le groupe maillé doit être un maillage complet des homologues MSDP entre tous les membres.

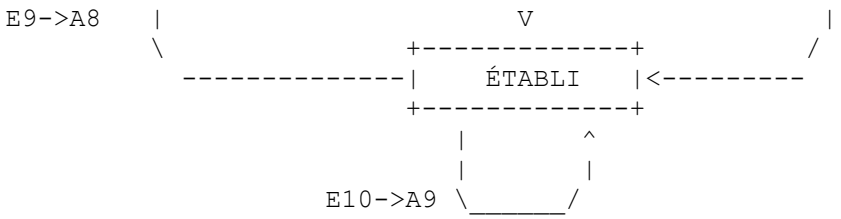
La sémantique du groupe maillé est la suivante :
(i) Si un membre R d'un groupe maillé M reçoit un message SA d'un homologue MSDP qui est aussi membre du groupe maillé M, R accepte le message SA et le transmet à tous ses homologues qui ne font pas partie du groupe maillé M. R NE DOIT PAS transmettre le message SA aux autres membres du groupe maillé M.
(ii) Si un membre R d'un groupe maillé M reçoit un message SA d'un homologue MSDP qui n'est pas un membre du groupe maillé M, et si le message SA réussit la vérification d'homologue RPF, R transmet alors le message SA à tous les membres du groupe maillé M et à tous ses autres homologues MSDP.

## 11. Automate à états de connexion MSDP

MSDP utilise TCP comme protocole de transport. Dans une relation d'homologue à homologue, un homologue MSDP écoute les nouvelles connexions TCP sur l'accès bien connu 639. L'autre côté fait une connexion active à cet accès. L'homologue qui a la plus forte adresse IP va écouter. Cet algorithme d'établissement de connexion évite les collisions d'appel. Donc, il n'y a pas besoin de procédure de collision d'appel. On notera cependant que l'inconvénient de cette approche est que le moment du démarrage dépend complètement du côté actif et de son temporisateur d'essai de connexion ; le côté passif ne peut pas causer l'établissement de la connexion.

Un homologue MSDP commence dans l'état DÉSACTIVÉ. Les homologues MSDP établissent les sessions d'homologue à homologue conformément à l'automate à états suivant :






### 11.1 Événements

E1) Permet l'établissement de la relation d'homologue MSDP avec P
E2) Adresse IP propre < adresse IP de P
E3) Adresse IP propre > adresse IP de P
E4) TCP établi (côté actif)
E5) TCP établi (côté passif)
E6) Arrivée à expiration du temporisateur Essai de connexion
E7) Désactivation de la relation d'homologue MSDP avec P (par exemple, quand sa propre adresse est changée)
E8) Arrivée à expiration du temporisateur de garde
E9) Erreur de format de TLV MSDP détectée
E10) Toute autre erreur détectée.

### 11.2 Actions

A1) Alloue des ressources pour la relation d'homologue avec P. Comparaison de l'adresse IP de l'homologue et la sienne.
A2) TCP actif OUVERT. Règle le temporisateur d'essai de connexion à [Période d'essai de connexion]
A3) TCP passif OUVERT (écoute)
A4) Supprime le temporisateur d'essai de connexion. Envoi le TLV Garde en vie. Règle le temporisateur de garde en vie à [Période de garde en vie]. Règle le temporisateur de garde à [Période de temps de garde]
A5) Envoi le TLV Garde en vie. Règle le temporisateur de garde en vie à [Période de garde en vie]. Règle le temporisateur de garde à [Période de temps de garde]
A6) Interrompt TCP actif OUVERT. Tente de libérer les ressources allouées pour la relation d'homologue avec P.
A7) Interrompt TCP passif OUVERT. Tente de libérer les ressources allouées pour la relation d'homologue avec P.
A8) Clôture de la connexion TCP. Libération des ressources allouées pour la relation d'homologue avec P.
A9) Éliminer le paquet.

### 11.3 Événements spécifiques de l'homologue

Les événements spécifiques de l'homologue suivants peuvent survenir dans l'état ÉTABLI ; ils ne causent pas de transition d'état. Les actions appropriées sont mentionnées pour chaque événement.

*) Arrivée à expiration du temporisateur de garde en vie :
-> Envoyer le TLV Garder en vie
-> Régler le temporisateur de garde en vie à [Période de garde en vie]
*) Réception du TLV Garder en vie :
-> Régler le temporisateur de garde à [Période de temps de garde]
*) Réception du TLV Source-Active :
-> Régler le temporisateur de garde à [Période de temps de garde]
-> Faire tourner l'algorithme de transmission d'homologue RPF
-> Régler le temporisateur de garde en vie à [Période de garde en vie] pour les homologues auxquels le TLV Source-Active est transmis
-> Envoyer les informations à PIM-SM
-> Mémoriser les informations en antémémoire

### 11.4 Événements indépendants de l'homologue

Il y a aussi un certain nombre d'événements qui affectent plus d'une session d'homologue à homologue, mais exigent quand même que des actions soient effectuées homologue par homologue.

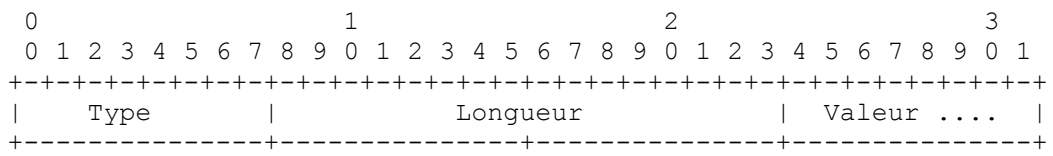
*) Arrivée à expiration de Tempo-Annonce-SA :
-> Démarrer la transmission périodique des TLV Source-Active
-> Régler le temporisateur de garde en vie à [Période de garde en vie] chaque fois qu'un TLV Source-Active est envoyé.

*) MSDP apprend une nouvelle source interne active (par exemple, un Register PIM-SM reçu pour une nouvelle source):
-> Envoyer le TLV Source-Active.
-> Régler le temporisateur de garde en vie à [Période de garde en vie].
*) Arrivée à expiration du temporisateur d'état SG (un temporisateur par entrée d'antémémoire) :
-> Spécifique de la mise en œuvre, marque normalement l'entrée d'antémémoire comme à supprimer.

## 12. Formats de paquet

Les messages MSDP sont codés en format de TLV. Si une mise en œuvre reçoit un TLV dont la longueur excède la longueur maximale de TLV spécifiée ci-dessous, le TLV DEVRAIT être accepté. Toutes les données supplémentaires, y compris les éventuels TLV suivants dans le même message, DEVRAIENT être ignorées, et la session MSDP ne devrait pas être reprise.

### 12.1 Formats de TLV MSDP




Type (8 bits) : décrit le format du champ Valeur.

Longueur (16 bits) : longueur des champs Type, Longueur et Valeur en octets. La longueur minimum requise est 4 octets, sauf pour les messages Garder en vie. La longueur maximum de TLV est 9 192.

Valeur (longueur variable) : le format se fonde sur la valeur du type. Voir ci-dessous. La longueur du champ Valeur est celle du champ Longueur moins 3. Tous les champs réservés dans le champ Valeur DOIVENT être transmis comme des zéros et ignorés à réception.

### 12.2 TLV définis

Les types de TLV suivants sont définis :

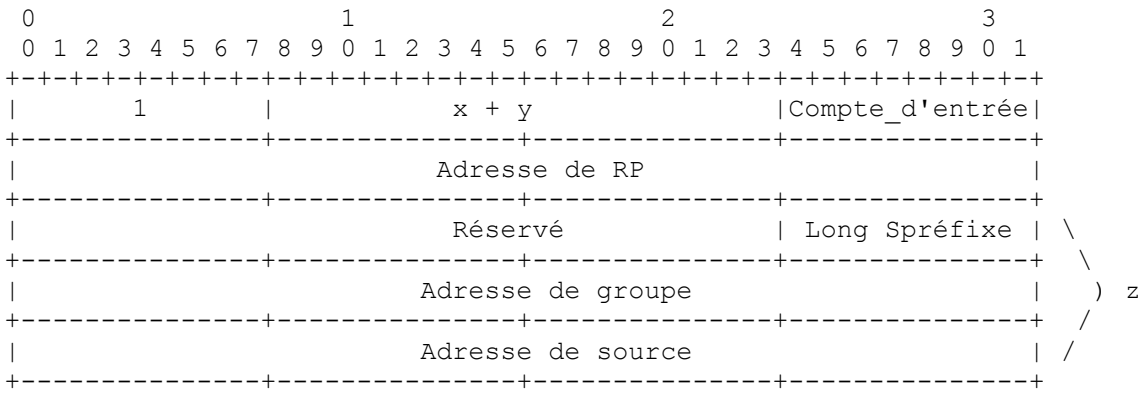
| Code | Type |
|---|---|
| 1 | Source-Active IPv4 |
| 2 | Demande Source-Active IPv4 |
| 3 | Réponse Source-Active IPv4 |
| 4 | Garder en vie |
| 5 | Réservé (antérieurement : Notification) |

Chaque TLV est décrit ci-dessous. De plus, les types de TLV suivants sont alloués mais non décrits dans ce mémoire.

| Code | Type |
|---|---|
| 6 | traceroute MSDP en cours |
| 7 | réponse traceroute MSDP |

#### 12.2.1 TLV Source-Active IPv4

La taille maximum de message SA qu'on peut envoyer est de 9 192 octets. La taille de 9192 octets n'inclut pas les en-têtes TCP, IP, et de couche 2.




Type : le TLV Source-Active IPv4 est de type 1.

Longueur x : c'est la longueur des informations de contrôle dans le message. x est 8 octets (pour les deux premières quantités de 32 bits) plus 12 fois Compte_d'entrées octets.

Longueur y : Si c'est 0, il n'y a pas de données encapsulées. Autrement un paquet IPv4 suit et y est la valeur du champ Longueur totale dans l'en-tête du paquet IP encapsulé. Si il y a plusieurs entrées (S,G) dans un message SA, seule la dernière entrée peut avoir des données encapsulées et elle doit refléter les adresses de source et de destination dans l'en-tête du paquet IP encapsulé.

Compte_d'entrée : c'est le compte des z entrées (voir la note ci-dessus) qui suivent le champ Adresse de RP. Cela permet de coder efficacement plusieurs (S,G) provenant du même domaine pour la même adresse de RP. Un message SA qui contient des données encapsulées a normalement un compte d'entrée de 1 (c'est-à-dire qu'il ne contient qu'une seule entrée, pour le (S,G) représentant le paquet encapsulé).

Adresse de RP : adresse du RP dans le domaine où la source est devenue active.

Réservé : le champ Réservé DOIT être transmis à zéro et DOIT être ignoré à réception.

Long Spréfixe : longueur du préfixe du chemin associé à l'adresse de source. Ce champ DOIT être transmis comme 32 (/32).

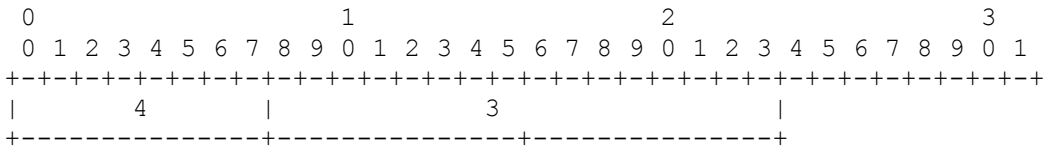
Adresse de groupe : adresse du groupe où la source active a envoyé des données.

Adresse de source : adresse IP de la source active.

Plusieurs entrées (S,G) PEUVENT apparaître dans le même SA et peuvent être regroupées en lots pour des raisons d'efficacité au prix d'une plus grande latence des données. Cela va normalement se produire sur la transmission intermédiaire des messages SA.

### 12.2.2 TLV Garder en vie

Un TLV Garder en vie est envoyé à un homologue MSDP si et seulement si il n'y a pas eu de message MSDP envoyé à l'homologue dans les [Période de garde en vie] secondes. Ce message est nécessaire pour garder la connexion MSDP en vie.




La longueur du message est 3 octets qui englobent le champ Type de un octet et le champ Longueur de deux octets.

## 13. Traitement d'erreur MSDP

Si un message MSDP est reçu avec une erreur de format de TLV, la session DEVRAIT être réinitialisée avec cet homologue. Les messages MSDP avec d'autres erreurs, comme un code de type non reconnu, reçus des homologues MSDP, DEVRAIENT être éliminés en silence et la session NE DEVRAIT PAS être réinitialisée.

## 14. Encapsulation des données de SA

Comme mentionné plus haut, l'encapsulation TCP de données dans les messages SA PEUT être prise en charge pour la rétro compatibilité avec les homologues MSDP antérieurs.

## 15. Déclaration d'applicabilité

MSDP est principalement utilisé dans deux scénarios de déploiement:

### 15.1 Entre domaines PIM

MSDP peut être utilisé entre des domaines PIM pour porter des informations sur les sources actives disponibles dans d'autres domaines. La relation d'homologue MSDP utilisée dans de tels cas est généralement d'homologue à homologue, et utilise les règles déterministes d'homologue RPF décrites dans le présent document (c'est-à-dire, n'utilisent pas de groupes maillés). Les relations d'homologues peuvent être agrégées sur un seul homologue MSDP, normalement de un à des centaines de relations d'homologues, similaires en échelle, bien que pas nécessairement cohérentes, avec les relations d'homologues BGP.

### 15.2 Entre des RP en envoi à la cantonade

MSDP est aussi utilisé entre des RP en envoi à la cantonade [RFC3446] au sein d'un domaine PIM pour synchroniser les informations sur les sources actives desservies par chaque homologue de RP en envoi à la cantonade (grâce à l'accessibilité IGP). La relation d'homologue MSDP utilisée dans ce scénario se fonde normalement sur les groupes maillés MSDP, ou tout ensemble de deux à des dizaines d'homologues peut comprendre un certain groupe maillé, bien que plus de dix ne soit pas courant. Un ou plusieurs de ces homologues de groupe maillé peuvent alors avoir des relations biunivoques supplémentaires avec des homologues MSDP en dehors de ce domaine PIM comme décrit dans le scénario A, pour la découverte de sources externes. MSDP pour RP à la cantonade sans relation d'homologue MSDP externe est une option de déploiement valide et courante.

## 16. Propriété intellectuelle

## 17. Remerciements

Les éditeurs tiennent à remercier les auteurs originaux, Dino Farinacci, Yakov Rehkter, Peter Lothberg, Hank Kilmer, et Jermey Hall de leur contribution à la spécification originale de MSDP. En plus, Bill Nickless, John Meylor, Liming Wei, Manoj Leelanivas, Mark Turner, John Zwiebel, Cristina Radulescu-Banu, Brian Edwards, Selina Priestley, IJsbrand Wijnands, Tom Pusateri, Kristofer Warell, Henning Eriksson, Thomas Eriksson, Dave Thaler, et Ravi Shekhar ont fourni d'utiles et productifs retours et commentaires. Toerless Eckert, Leonard Giuliano, Mike McBride, David Meyer, John Meylor, Pekka Savola, Ishan Wu, et Swapna Yelamanchi ont contribué à la version finale de ce document.

## 18. Considérations sur la sécurité

Une mise en œuvre MSDP DOIT utiliser MD5 chiffré [RFC2385] pour sécuriser les messages de commande, et DOIT être capable d'interopérer avec les homologues qui ne le prennent pas en charge. Cependant, si un côté de la connexion est configuré avec MD5 chiffré et pas l'autre côté, la connexion NE DEVRAIT PAS être établie.

De plus, pour atténuer l'explosion d'état durant les attaques de déni de service et autres, des filtres de SA et des limites DEVRAIENT être utilisées avec MSDP pour limiter les sources et groupes qui seront passés entre les RP [RFC4611]. Ces fonctions de filtrage et de limitation peuvent inclure, par exemple, des listes d'accès des adresses de source ou de groupes qui ne devraient pas être propagés aux autres domaines en utilisant MSDP, le nombre absolu le plus élevé d'entrées acceptable d'état de SA, ou une limite de taux de création de nouvelles entrées d'état de SA après l'établissement de la connexion.

Si des travaux suivent dans ce domaine, un mécanisme plus robuste de protection de l'intégrité, comme HMAC-SHA1 [RFC2104], [RFC2202] pourrait être employé.

## 19. Considérations relatives à l'IANA

Le présent document crée un nouvel espace de noms appelé "Valeurs de TLV MSDP" qui sera géré par l'IANA. Les sept valeurs initiales de TLV MSDP sont spécifiées au paragraphe 12.2. Les deux paragraphes qui suivent décrivent les règles d'allocation de nouvelles valeurs de TLV MSDP.

### 19.1 Gamme de TLV alloués par l'IANA

Les valeurs de TLV MSDP dans la gamme de [8 à 200] (inclus) sont à allouer en utilisant le processus d'approbation de l'IESG ou d'action de normalisation [RFC2434].

### 19.2 Gamme expérimentale de TLV

Les valeurs de TLV dans la gamme de [201 à 255] (inclus) sont allouées pour utilisation expérimentale.

## 20. Références

### 20.1 Références normatives

[RFC1142] D. Oran, "Protocole d'acheminement intra domaine IS-IS de l'OSI", janvier 1990. *(Historique, voir* RFC7142)

[RFC2119] S. Bradner, "Mots clés à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997.

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

[RFC2362] D. Estrin et autres, "Mode épars de diffusion groupée indépendante du protocole (PIM-SM) : Spécification du protocole", juin 1998. *(Obsolète, voir* RFC4601, RFC5059)

[RFC2365] D. Meyer, "Diffusion groupée sur IP limitée administrativement", juillet 1998. (BCP0023)

[RFC2385] A. Heffernan, "Protection des sessions de BGP via l'option de signature MD5 de TCP", août 1998. *(P.S.* (*MàJ*

*par la* RFC6691)) *(Remplacée par* RFC5925)

[RFC2434] T. Narten et H. Alvestrand, "Lignes directrices pour la rédaction d'une section Considérations relatives à l'IANA dans les RFC", BCP 26, octobre, 1998. *(Rendue obsolète par la* RFC5226)

[RFC2858] T. Bates et autres, "Extensions multiprotocoles pour BGP-4", juin 2000. (*Obsolète, voir* RFC4760) *(P.S.)*

[RFC3446] D. Kim et autres, "Mécanisme de point de rendez-vous (RP) en envoi à la cantonade utilisant la diffusion groupée indépendante du protocole (PIM) et le protocole de découverte de source de diffusion groupée (MSDP)", janvier 2003. *(Info.)*

### 20.2 Références pour information

[RFC2104] H. Krawczyk, M. Bellare et R. Canetti, "HMAC : Hachage de clés pour l'authentification de message", février 1997.

[RFC2202] P. Cheng et R. Glenn, "Cas d'essai pour HMAC-MD5 et HMAC-SHA-1", septembre 1997. *(Information)*

[RFC4611] M. McBride et autres, "Scénarios de développement du protocole de découverte de source de diffusion groupée (MSDP)", août 2006. (BCP0121)

## 21. Adresses des éditeurs

Bill Fenner
AT&T Labs -- Research
75 Willow Road
Menlo Park, CA 94025
mél : fenner@research.att.com

David Meyer
mél : dmm@1-4-5.net

## 22. Déclaration complète de droits de reproduction



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