

Groupe de travail Réseau  
Request for Comments: 3988  
Catégorie : Expérimental

B. Black  
Layer8 Networks  
K. Kompella  
Juniper Networks  
janvier 2005

## Extensions de signalisation d'unité de transmission maximum pour le protocole de distribution d'étiquettes

### Statut du présent Mémo

Le présent mémo définit un protocole expérimental pour la communauté Internet. Il ne spécifie en aucune façon une norme Internet. Il réclame une discussion et des suggestions pour son amélioration. La distribution de ce mémo n'est pas limitée.

### Notice de copyright

Copyright (C) The Internet Society (2005).

### Résumé

Le bon fonctionnement de la découverte de l'unité de transmission maximum (MTU, *Maximum Transmission Unit*) de chemin de la RFC 1191 exige que les routeurs IP aient connaissance de la MTU pour chaque liaison à laquelle ils sont connectés. Selon la spécification actuelle, le protocole de distribution d'étiquette (LDP, *Label Distribution Protocol*) n'a pas la capacité de signaler la MTU pour un chemin commuté d'étiquette (LSP, *Label Switched Path*) à l'entrée du routeur de commutation d'étiquette (LSR, *Label Switching Router*). En l'absence de cette fonctionnalité, le MTU pour chaque LSP doit être configuré de façon statique par les opérateurs de réseau ou par des mécanismes hors ligne équivalents.

Le présent document spécifie des extensions expérimentales au protocole LDP pour la prise en charge de la découverte des MTU du LSP.

## 1. Introduction

Comme il est actuellement spécifié en [2], le protocole LDP pour MPLS ne prend pas en charge la signalisation de la MTU pour les LSP à l'entrée des LSR. Cette fonctionnalité est essentielle pour le bon fonctionnement de la détection de la MTU de chemin de la RFC 1191 [3]. Sans connaissance de la MTU pour un LSP, les LSR de bordure peuvent transmettre des paquets sur ce LSP, qui seraient trop gros selon [4]. Ces paquets peuvent être éliminés en silence par les routeurs LSR le long du chemin LSP, ce qui empêche la communication effective entre certains hôtes terminaux.

La solution proposée dans le présent document permet la détermination automatique de la MTU pour un LSP par l'ajout d'un triplet Type-Longueur-Valeur (TLV) pour porter les informations de MTU à destination d'une classe d'équivalence de transmission (FEC, *Forwarding Equivalence Class*) entre les LSR adjacents dans les messages de transposition d'étiquette du protocole LDP.

Ces informations sont suffisantes pour un ensemble de LSR le long du chemin suivi par un LSP pour découvrir le MTU exact pour ce LSP, ou si une approximation ne serait pas pire que ce qui pourrait être généré avec les informations locales sur le LSR d'entrée.

### 1.1. Conventions utilisées dans le présent document

Dans le présent document, les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" doivent être interprétés comme décrit par la [RFC2119].

## 2. Signalisation de MTU

La procédure de signalisation décrite dans le présent document utilise l'ajout d'un seul TLV aux messages de transposition d'étiquette du protocole LDP et un simple algorithme pour le calcul de la MTU du LSP.

### 2.1. Définitions

**MTU de liaison :** MTU pour une liaison donnée. Cette taille inclut l'en-tête IP et les données (ou autre charge utile) et la pile d'étiquettes mais n'inclut aucun des en-têtes de couche inférieure. Une liaison peut être une interface (telle qu'Ethernet ou Paquet-sur-SONET), un tunnel (tel que GRE ou IPsec), ou un LSP.

**LSR homologues :** Pour le LSR A et la FEC F, c'est l'ensemble des LSR qui envoient une transposition d'étiquette pour la FEC F à A.

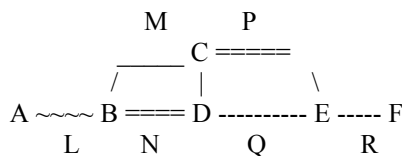
**LSR aval :** Pour le LSR A et la FEC F, c'est le sous-ensemble des LSR homologues de A pour la FEC F auxquels A va transmettre les paquets pour la FEC. Normalement, ce sous-ensemble est déterminé via le tableau d'acheminements.

**MTU de bond :** MTU d'un bond de LSP entre un LSR A, amont, et un LSR B, aval. Cette taille inclut l'en-tête IP et les données (ou autre charge utile) et la partie de la pile d'étiquettes qui est considérée comme charge utile tant que dure ce LSP. Il n'inclut aucun en-tête de niveau inférieur. (Note : S'il y a plusieurs liaisons entre A et B, le MTU de bond est le minimum de la MTU de bond parmi les liaisons utilisées pour la transmission.

**MTU de LSP :** MTU d'un LSP pour un LSR donné à la ou aux entrées, sur chaque chemin valide (de transmission). Cette taille inclut l'en-tête IP et les données (ou autres charge utile) et toute partie de la pile d'étiquettes qui a été reçue par le LSR d'entrée avant de placer le paquet dans le LSP (cette partie de la pile d'étiquette est considérée comme faisant partie de la charge utile pour ce LSP). La taille n'inclut aucun des en-têtes de niveau inférieur.

### 2.2. Exemple

Considérons les LSR A à F, interconnectés comme suit :



Disons que la MTU de liaison pour la liaison L est 9216 ; pour les liaisons M, Q, et R, 4470 ; et pour N et P, de 1500. Considérons une FEC X pour laquelle F est la sortie, et disons que tous les LSR avertissent leurs voisins de X.

Noter que bien que LDP puisse fonctionner sur la liaison C – D, il n'est pas utilisé pour transmettre (par exemple, parce qu'il a une métrique élevée). En particulier, D est un LDP voisin de C, mais D n'est pas un LSR aval de C pour la FEC X.

Les homologues de E pour la FEC X sont C, D, et F. Disons que E choisit F comme son LSR aval pour X. La MTU de bond de E pour la liaison R est 4466. Si F avertit E d'une étiquette implicite nulle, E PEUT alors établir la MTU de bond pour R à 4470.

Les homologues de C pour la FEC X sont B, D, et E. Disons que C choisit E comme son LSR aval pour X. De même (respectivement), A choisit B, B choisit C et D (chemin multiple de coût égal), D choisit E, et E choisit F, comme LSR aval.

La MTU de bond de C à E pour la FEC X est 1496. La MTU de bond de B à C est 4466 et à D de 1496. La MTU de LSP de A pour la FEC X est 1496. Si A a un autre LSP pour la FEC Y pour F (apprise via le LDP ciblé) qui passe par le LSP pour la FEC X, la MTU pour ce LSP serait 1492.

Si B a une session de LDP ciblée pour E (par exemple, sur un tunnel RSVP-TE T) et si B a reçu une transposition pour la classe de FEC X sur la session LDP ciblée, E serait alors aussi l'homologue de B, et E peut être choisi comme LSR aval pour B. Dans ce cas, la MTU de LSP de B pour la classe FEC X serait alors le plus petit de {(MTU de T - 4), MTU de LSP de E pour X}.

Le présent mémo décrit comment A détermine son MTU de LSP pour les classes de FEC X et Y.

### 2.3. Procédure de signalisation

La procédure de signalisation des MTU est effectuée bond par bond par chaque LSR L le long d'un LSP pour une classe de FEC F donnée. Les étapes sont les suivantes:

1. L calcule d'abord son unité MTU de LSP pour la classe de FEC F :

A. Si L est la sortie pour F, L établit la MTU de LSP pour F à 65535.

B. [FACULTATIF] Si le seul LSR de L est la sortie pour F (c'est-à-dire un pénultième bond pour F) et que L reçoit une étiquette nulle implicite au titre de sa transposition pour F, L peut alors établir la MTU de bond pour sa liaison aval au niveau de la MTU de liaison au lieu de (MTU de liaison – 4 octets). La MTU de LSP de L pour F est la MTU de bond.

C. Autrement (L n'est pas le LSR de sortie), L calcule la MTU de LSP pour F comme suit :

a) L détermine ses LSR aval pour la FEC F.

b) Pour chaque LSR Z aval, L calcule le minimum de la MTU de bond pour Z et de la MTU de LSP dans le TLV de MTU que Z a communiqué à L. Si Z n'a pas inclus le TLV de MTU dans sa transposition d'étiquettes, la MTU de LSP de Z est alors établie à 65535.

c) L règle sa MTU de LSP au minimum des MTU qu'il a calculées pour ses LSR aval.

2. Pour chaque LDP voisin (direct ou ciblé) de L auquel L décide d'envoyer une transposition pour la FEC F, L attache un TLV de MTU avec la MTU de LSP qu'il a calculée pour cette FEC. L PEUT (à cause de sa politique ou pour d'autres raisons) avertir d'une MTU plus petite que celle qu'il a calculée, mais L NE DOIT PAS annoncer une MTU plus grande.

3. Lorsqu'une nouvelle MTU est reçue pour la FEC F d'un LSR aval ou de l'ensemble des LSR aval pour un changement de F, L retourne à l'étape 1. Si la MTU de LSP nouvellement calculée est inchangée, L NE DEVRAIT PAS communiquer la nouvelle information à ses voisins. Autrement, L réannonce ses transpositions pour F à tous ses homologues avec un TLV de MTU mis à jour.

Ce comportement est standard pour des attributs tels qu'un vecteur de chemin et un compte de bonds, et les mêmes règles s'appliquent, comme spécifié en [2].

Si la MTU de LSP diminue, L DEVRAIT réannoncer immédiatement la nouvelle MTU ; si la MTU de LSP augmente, L PEUT ne pas faire l'annonce.

### 2.4. TLV de MTU

Le TLV de MTU code les informations sur l'unité de transmission maximum pour un LSP, à partir du LSR qui annonce vers la ou les sorties sur tous les chemins valides.

Le codage pour le TLV de MTU est le suivant :

```

    0           1           2           3           0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8
9 0 1
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|1|1| TLV de MTU (0x0601) |      Longueur      |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|                MTU                |
                                         MTU

```

Ceci est un entier arithmétique de 16 bits qui représente la MTU en octets pour un LSP ou un segment d'un LSP.

Noter que les bits U et F sont établis. Un LSR qui ne reconnaît pas le TLV de MTU DOIT l'ignorer lorsqu'il traite le message de transposition d'étiquette et transmet le TLV à ses homologues. Il peut en résulter un calcul incorrect de la

MTU de LSP ; cependant, la transmission silencieuse du TLV de MTU préserve la quantité maximale d'informations sur la MTU de LSP.

### 3. Exemple de fonctionnement

Considérons l'exemple de réseau du paragraphe 2.2. Pour chaque LSR, le Tableau 1 décrit les liaisons avec ses LSR aval, la MTU de bond pour l'homologue, la MTU de LSP reçue de l'homologue, et la MTU de LSP calculée du LSR.

Considérons maintenant le même réseau avec les changements suivants : il y a un LSP T de B à E, et une session LDP ciblée de B à E. Les LSR homologues de B sont A, C, D, et E ; les LSR aval de B sont D et E ; pour atteindre E, B choisit d'aller sur T. La MTU de LSP pour le LSP T est 1496. Ces informations sont décrites au Tableau 2.

LSR	Liaison	MTU de bond	MTU de réception	MTU de LSP
F	-	65535	-	65535
E	R	4466	F : 65535	4466
D	Q	4466	E : 4466	4466
C	P	1496	E : 4466	1496
B	M	4466	C : 1496	
	N	1496	D : 4466	1496
A	L	9212	B : 1496	1496

Tableau 1

LSR	Liaison	MTU de bond	MTU de réception	MTU de LSP
F	-	65535	-	65535
E	R	4466	F : 65535	4466
D	Q	4466	E : 4466	4466
C	P	1496	E : 4466	1496
B	T	1492	E : 4466	
	N	1496	D : 4466	1492
A	L	9212	B : 1492	1492

Tableau 2

### 4. Utilisation des MTU de LSP

Un LSR d'entrée qui transmet un paquet IP dans un LSP dont il connaît la MTU DOIT soit fragmenter le paquet IP à la MTU du LSP (si le bit Don't Fragment est clair) ou abandonner le paquet et répondre par un message Destination inatteignable du protocole ICMP à la source du paquet, avec le code indiquant "fragmentation nécessaire et DF établi", et la MTU du prochain bond établie à la MTU du LSP. En d'autres termes, le LSR se comporte conformément à la RFC 1191, excepté qu'il traite le LSP comme le prochain bond "réseau".

Si la charge utile pour le LSP n'est pas un paquet IP, le LSR DOIT transmettre le paquet s'il convient (taille  $\leq$  MTU de LSP) et DEVRAIT l'abandonner sinon.

## 5. Interaction de protocole

### 5.1. Interaction avec les LSR ne prenant pas en charge la signalisation de MTU

Les changements de la MTU pour des sections d'un LSP peuvent causer la génération de messages de transposition d'étiquette non sollicitée par les LSR intermédiaires pour avertir de la nouvelle MTU. Les LSR qui ne prennent pas en charge la signalisation de MTU devront accepter, du fait des mécanismes de traitement de message et de TLV spécifiés dans la RFC3036 [2], les messages portant les TLV de MTU mais ignoreront le TLV et transmettront le TLV aux nœuds amont (voir le paragraphe 2.4).

### 5.2. Interaction avec CR-LDP et RSVP-TE

Le TLV de MTU peut être utilisé pour découvrir la MTU de chemin aussi bien des LSP de LDP que des LSP de CR-LDP. Cette proposition n'est pas influencée par la présence de LSP créés avec CR-LDP, comme spécifié en [5].

Noter que les LSP LDP/CR-LDP peuvent tunneler à travers d'autres LSP signalés comme utilisant LDP, CR-LDP, ou RSVP-TE [6] ; le mécanisme suggéré ici s'applique dans tous ces cas, principalement en traitant les LSP tunnels comme des liaisons.

## 6. Références

### 6.1. Références normatives

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels" (*Mots clés à utiliser dans les RFC pour indiquer les niveaux d'exigence*), BCP 14, RFC 2119, mars 1997.

[2] Andersson, L., Doolan, P., Feldman, N., Fredette, A., et B. Thomas, "LDP Specification" (*Spécification du protocole LDP*), RFC 3036, janvier 2001.

[3] Mogul, J. et S. Deering, "Path MTU discovery" (*Découverte de la MTU de chemin*), RFC 1191, novembre 1990.

[4] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., et A. Conta, "MPLS Label Stack Encoding" (*Codage de la pile d'étiquette MPLS*), RFC 3032, janvier 2001.

[6] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., et G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels" (*RSVP-TE : extensions à RSVP pour les tunnels de LSP*), RFC 3209, décembre 2001.

### 6.2. Références informatives

[5] Jamoussi, B., Andersson, L., Callon, R., Dantu, R., Wu, L., Doolan, P., Worster, T., Feldman, N., Fredette, A., Girish, M., Gray, E., Heinanen, J., Kilty, T., et A. Malis, "Constraint-Based LSP Setup using LDP" (*Etablissement de LSP fondés sur la contrainte en utilisant le protocole LDP*) RFC 3212, janvier 2002.

## 7. Considérations de sécurité

Le présent mécanisme n'introduit aucune nouvelle faiblesse dans le protocole LDP. Il est possible d'imiter les paquets TCP qui appartiennent à une session LDP pour manipuler la MTU de LSP, mais le protocole LDP a des mécanismes pour contrecarrer ces types d'attaques. Voir la section 5 de [2] pour des informations complémentaires sur les aspects de sécurité du protocole LDP.

## 8. Considérations sur l'IANA

L'IANA a alloué 0x0601 comme nouveau type de TLV de LDP, défini au paragraphe 2.4. Voir : <http://www.iana.org/assignments/ldp-namespaces>

## 9. Remerciements

Nous remercions André Fredette pour les nombrax commentaires détaillés sur les versions précédentes du mécanisme de signalisation. Eric Gray, Giles Héron, et Mark Duffy ont contribué par de nombreuses suggestions utiles.

### Adresse des auteurs

Benjamin Black  
Layer8 Networks  
EMail: ben@layer8.net

Kireeti Kompella  
Juniper Networks  
1194 N. Mathilda Ave  
unnyvale, CA 94089  
US  
EMail: kireeti@juniper.net

### Déclaration de Copyright

Copyright (C) The Internet Society (2005).

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et sauf pour ce qui est mentionné ci-après, les auteurs conservent tous leurs droits.

Le présent document et les informations y contenues sont fournies sur une base "EN L'ÉTAT" et LE CONTRIBUTEUR, L'ORGANISATION QU'IL OU ELLE REPRÉSENTE OU QUI LE/LA FINANCE (S'IL EN EST), LA INTERNET SOCIETY ET LA INTERNET ENGINEERING TASK FORCE DÉCLINENT TOUTES GARANTIES, EXPRIMÉES OU IMPLICITES, Y COMPRIS MAIS NON LIMITÉES À TOUTE GARANTIE QUE L'UTILISATION DES INFORMATIONS CI-ENCLOSES NE VIOLENT AUCUN DROIT OU AUCUNE GARANTIE IMPLICITE DE COMMERCIALISATION OU D'APTITUDE À UN OBJET PARTICULIER.

### Propriété intellectuelle

L'IETF ne prend pas position sur la validité et la portée de tout droit de propriété intellectuelle ou autres droits qui pourrait être revendiqués au titre de la mise en œuvre ou l'utilisation de la technologie décrite dans le présent document ou sur la mesure dans laquelle toute licence sur de tels droits pourrait être ou n'être pas disponible ; pas plus qu'elle ne prétend avoir accompli aucun effort pour identifier de tels droits. Les informations sur les procédures de l'IETF au sujet des droits dans les documents de l'IETF figurent dans les BCP 78 et BCP 79.

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat de tentatives faites pour obtenir une licence ou permission générale d'utilisation de tels droits de propriété par ceux qui mettent en œuvre ou utilisent la présente spécification peuvent être obtenues sur répertoire en ligne des IPR de l'IETF à <http://www.ietf.org/ipr>.

L'IETF invite toute partie intéressée à porter son attention sur tous copyrights, licences ou applications de licence, ou autres droits de propriété qui pourraient couvrir les technologies qui peuvent être nécessaires pour mettre en œuvre la présente norme. Prière d'adresser les informations à l'IETF à [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

### Remerciement

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