

Groupe de travail Réseau
Request for Comments : 2597
Catégorie : En cours de normalisation

J. Heinanen, Telia Finland
F. Baker, Cisco Systems
W. Weiss, Lucent Technologies
J. Wroclawski, MIT LCS
juin 1999

Traduction Claude Brière de L'Isle

Groupe de PHB Transmission assurée

Statut du présent mémoire

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

Notice de Copyright

Copyright (C) The Internet Society (1999). Tous droits réservés.

Résumé

Le présent document définit un groupe de comportement par bond (PHB, *Per-Hop-Behavior*) de services différenciés (DS, *Differentiated Services*) [RFC2475] d'utilité générale appelé Transmission assurée (AF, *Assured Forwarding*). Le groupe de PHB AF fournit la livraison des paquets IP dans quatre classes AF transmises indépendamment. Au sein de chaque classe AF, un paquet IP peut être alloué à un des trois différents niveaux de préséance d'abandon. Un nœud DS ne réarrange pas les paquets IP du même microflux si ils appartiennent à la même classe AF.

1. Objet et vue d'ensemble

Il existe une demande pour la fourniture d'une garantie de transmission des paquets IP sur l'Internet. Dans une application normale, une compagnie utilise l'Internet pour interconnecter ses sites répartis géographiquement et veut l'assurance que les paquets IP au sein de son intranet sont transmis avec une forte probabilité pour autant que le trafic agrégé provenant de chaque site n'excède pas le taux (profil) d'informations souscrit. Il est souhaitable qu'un site puisse dépasser le profil souscrit en acceptant que le trafic excédentaire ne soit pas livré avec une aussi forte probabilité que le trafic qui est dans les limites du profil. Il est aussi important que le réseau ne réarrange pas les paquets qui appartiennent au même microflux, comme défini dans la [RFC2474], qu'ils soient dans ou en dehors des limites du profil.

Le groupe de PHG Transmission assurée (AF, *Assured Forwarding*) est un moyen pour un fournisseur de domaine DS d'offrir différents niveaux d'assurances de transmission pour les paquets IP reçus d'un usager du domaine DS. Quatre classes AF sont définies, où à chaque classe AF dans chaque nœud DS est allouée une certaine quantité de ressources de transmission (espace de mémoire tampon et bande passante). Les paquets IP qui veulent utiliser les services fournis par le groupe PHB AF se voient allouer par l'usager ou par le fournisseur du domaine DS à une ou plusieurs de ces classes AF conformément aux services auxquels l'usager s'est abonné. On trouvera des précisions sur les fondements de cette facilité et certaines façons de l'utiliser dans [Clark].

Au sein de chaque classe AF, les paquets IP sont marqués (soit par l'usager, soit par le fournisseur du domaine DS) avec une des trois valeurs possibles de préséance d'abandon. En cas d'encombrement, la préséance d'abandon d'un paquet détermine l'importance relative du paquet dans la classe AF. Un nœud DS encombré essaye de protéger contre la perte les paquets ayant une valeur inférieure de préséance d'abandon en éliminant de préférence les paquets qui ont des valeurs de préséance d'abandon plus élevée.

Dans un nœud DS, le niveau d'assurance de transmission d'un paquet IP dépend donc (1) de la quantité de ressources de transmission qui ont été allouées à la classe AF à laquelle appartient le paquet, (2) de la charge actuelle de la classe AF, et (3) en cas d'encombrement au sein de la classe, de la préséance d'abandon du paquet.

Par exemple, si des actions de conditionnement du trafic à l'entrée du domaine DS du fournisseur rendent certain qu'une classe AF dans les nœuds DS est seulement modérément chargée par des paquets qui ont la plus faible valeur de préséance d'abandon et n'est pas surchargée par des paquets qui ont les deux plus faibles valeurs de préséance d'abandon, la classe AF peut offrir un haut niveau d'assurance de transmission pour les paquets qui sont dans le profil souscrit (c'est à dire, marqués avec la plus faible valeur de préséance d'abandon) et offrir jusqu'à deux niveaux inférieurs d'assurance de transmission au trafic excédentaire.

Le présent document décrit le groupe de PHB AF. Un nœud conforme par ailleurs à DS n'est pas obligé de mettre en œuvre ce groupe de PHB afin d'être considéré comme conforme à DS, mais lorsque un nœud conforme à DS est réputé mettre en œuvre un groupe de PHB AF, il doit se conformer aux spécifications du présent document.

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAI PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" dans ce document sont à interpréter comme décrit dans la [RFC2119].

2. Groupe de PHB AF

Le groupe de PHB Transmission assurée (AF) fournit la transmission de paquets IP dans N classes AF indépendantes. Au sein de chaque classe AF, un paquet IP est affecté à un des M différents niveaux de préséance d'abandon. Un paquet IP qui appartient à une classe AF i et a une préséance d'abandon j correspond au codet AF AF ij , où $1 \leq i \leq N$ et $1 \leq j \leq M$. Actuellement, quatre classes ($N = 4$) avec trois niveaux de préséance d'abandon dans chaque classe ($M = 3$) sont définies pour l'usage général. Plus de classes AF ou de niveaux de préséance d'abandon PEUVENT être définis pour une utilisation locale.

Un nœud DS DEVRAIT mettre en œuvre les quatre classes AF d'utilisation générale. Dans une classe AF, les paquets DOIVENT être transmis indépendamment des paquets dans une autre classe AF, c'est-à-dire, un nœud DS NE DOIT PAS agréger ensemble deux ou plus classes AF.

Un nœud DS DOIT allouer une quantité configurable minimum de ressources de transmission (espace de mémoire tampon et bande passante) à chaque classe AF mise en œuvre. Chaque classe DEVRAIT être servie de façon à réaliser le taux de service configuré (la bande passante) sur des échelles de temps à la fois brèves et longues.

Une classe AF PEUT aussi être configurable pour recevoir plus de ressources de transmission que le minimum lorsque un excédent de ressources est disponible de la part d'autres classes AF ou d'autres groupes de PHB. Le présent mémoire ne spécifie pas comment les ressources excédentaires devraient être allouées, mais les mises en œuvre DOIVENT spécifier quels algorithmes sont en fait pris en charge et comment ils peuvent être paramétrés.

Au sein d'une classe AF, un nœud DS NE DOIT PAS transmettre un paquet IP avec une plus faible probabilité si il contient une valeur de préséance d'abandon p que si il contient une valeur de préséance d'abandon q lorsque $p < q$. Noter que cette exigence peut être satisfaite sans qu'il soit besoin de sortir de la file d'attente et d'éliminer les paquets qui sont déjà dans la file d'attente.

Au sein de chaque classe AF, un nœud DS DOIT accepter tous les trois codets de préséance d'abandon et il DOIT donner au moins trois différents niveaux de probabilité de perte. Dans certains réseaux, en particulier dans les réseaux d'entreprise, où l'encombrement transitoire est rare et bref, il peut être raisonnable pour un nœud DS de mettre en œuvre seulement deux différents niveaux de probabilité de perte par classe AF. Bien que cela puisse suffire pour certains réseaux, trois différents niveaux de probabilité de perte DEVRAIENT être pris en charge dans les domaines DS où l'encombrement est un événement courant.

Si un nœud DS ne met en œuvre que deux différents niveaux de probabilité de perte pour une classe F x , le codet AF $x1$ DOIT donner la plus faible probabilité de perte et les codets AF $x2$ et AF $x3$ DOIVENT donner la plus forte probabilité de perte.

Un nœud DS NE DOIT PAS réarranger l'ordre des paquets AF du même microflux lorsque ils appartiennent à la même classe AF sans considération de leur préséance d'abandon. Il n'y a pas d'exigence temporelle quantifiable (de délai ou de variation de délai) associée à la transmission des paquets AF.

Les relations entre les classes AF et les autres PHB sont décrites à la Section 7 du présent mémoire.

Le groupe de PHB AF PEUT être utilisé pour mettre en œuvre aussi bien les services de bout en bout que ceux de bord de domaine à bord de domaine.

3. Actions de conditionnement du trafic

Un domaine DS PEUT depuis la bordure d'un domaine contrôler la quantité de trafic AF qui entre ou sort du domaine aux divers niveaux de préséance d'abandon. De telles actions de conditionnement du trafic PEUVENT inclure du formatage de trafic, l'élimination de paquets, l'augmentation ou la diminution de la préséance d'abandon des paquets, et la réallocation des paquets à d'autres classes AF. Cependant, les actions de conditionnement de trafic NE DOIVENT PAS causer le réarrangement des paquets du même microflux.

4. Comportement de mise en file d'attente et d'élimination

Cette section définit les comportements de mise en file d'attente et d'élimination du groupe de PHB AF. Les autres aspects du comportement du groupe de PHB sont définis à la Section 2.

Une mise en œuvre AF DOIT essayer de minimiser l'encombrement de longue durée au sein de chaque classe, tout en permettant un encombrement de courte durée résultant de salves. Cela exige un algorithme de gestion active de file d'attente. Un exemple d'un tel algorithme est la détection précoce aléatoire (RED, *Random Early Detection*) [Floyd]. Le présent mémoire ne spécifie pas l'utilisation d'un algorithme particulier, mais exige que soient respectées plusieurs propriétés.

Une mise en œuvre AF DOIT détecter et répondre à l'encombrement de longue durée au sein de chaque classe en abandonnant des paquets, tout en traitant l'encombrement de courte durée (les salves de paquets) en mettant les paquets en file d'attente. Cela implique la présence de fonctions de lissage ou de filtrage qui surveillent le niveau d'encombrement instantané et calculent un niveau d'encombrement lissé. L'algorithme d'abandon utilise ce niveau d'encombrement lissé pour déterminer lorsque des paquets devraient être éliminés.

L'algorithme d'abandon DOIT être insensible aux caractéristiques à court terme du trafic des microflux qui utilisent une classe AF. C'est-à-dire que les flux qui ont des formes de salves de court terme différentes mais des taux de paquet à plus long terme identiques devraient avoir des éliminations de paquets de probabilités essentiellement égales. Une façon d'obtenir cela est d'utiliser l'aléation dans la fonction d'élimination.

L'algorithme d'abandon DOIT traiter de façon identique tous les paquets au sein d'une seule classe et niveau de préséance. Cela implique que pour un niveau d'encombrement lissé donné, le taux d'élimination des paquets d'un microflux particulier au sein d'un même niveau de préséance sera proportionnel au pourcentage que représente ce flux dans la quantité totale de trafic passant à travers ce niveau de préséance.

Le retour d'indication d'encombrement vers les nœuds d'extrémité, et donc le niveau d'élimination de paquet à chaque préséance d'abandon en relation avec l'encombrement, DOIT être graduel plutôt qu'abrupt, pour permettre au système global d'atteindre un point stable de fonctionnement. Une façon de le faire (RED) utilise deux seuils (configurables) de niveau d'encombrement lissé. Lorsque le niveau d'encombrement lissé est en dessous du premier seuil, aucun paquet de la préséance pertinente n'est éliminé. Lorsque le niveau d'encombrement lissé est entre le premier et le second seuil, les paquets sont éliminés avec une probabilité de croissance linéaire, allant de zéro à une valeur configurable atteinte juste avant le second seuil. Lorsque le niveau d'encombrement lissé est au dessus du second seuil, les paquets de la préséance pertinente sont éliminés avec une probabilité de 100 %.

Pour permettre l'utilisation du PHB AF dans de nombreux environnements de fonctionnement différents, les paramètres de commande de l'algorithme d'abandon DOIVENT être configurables indépendamment de chaque préséance d'abandon de paquet et pour chaque classe AF.

Dans les limites tracées ci-dessus, la présente spécification permet une gamme de comportements d'élimination de paquets. Les comportements d'élimination incohérents conduisent à une sémantique de service de bout en bout incohérente et limitent la gamme des utilisations possibles du PHB AF dans un environnement de concurrence. Comme l'expérience s'accumule, des versions futures de ce document pourront définir plus étroitement des aspects spécifiques du comportement souhaitable.

5. Tunnelage

Lorsque des paquets AF sont tunnelés, le PHB du tunnelage de paquet NE DOIT PAS réduire l'assurance de transmission du paquet AF tunnelé ni causer le réarrangement des paquets AF qui appartiennent au même microflux.

6. Codets recommandés

Les codets recommandés pour les quatre classes AF d'usage général sont données ci-dessous. Ces codets ne se recouvrent avec aucun autres groupe de PHB d'usage général.

Les valeurs RECOMMANDÉES des codets AF sont les suivantes :

AF11 = '001010', AF12 = '001100', AF13 = '001110', AF21 = '010010', AF22 = '010100', AF23 = '010110', AF31 = '011010', AF32 = '011100', AF33 = '011110', AF41 = '100010', AF42 = '100100', et AF43 = '100110'.

Le tableau ci-dessous résume les valeurs de codet AF recommandées.

	Classe 1	Classe 2	Classe 3	Classe 4
Préséance d'abandon faible	001010	010010	011010	100010
Préséance d'abandon moyenne	001100	010100	011100	100100
Préséance d'abandon forte	001110	010110	011110	100110

7. Interactions avec les autres groupes de PHB

Les transpositions des codets AF recommandés ci-dessus n'interfèrent pas avec les espaces d'utilisation locale ni avec les codets de sélecteur de classe recommandés dans la [RFC2474]. Les PHB choisis par ces codets de sélecteur de classe peuvent donc coexister avec le groupe de PHB AF et conserver le comportement de transmission et les relations qui étaient définies pour eux. En particulier, le codet de PHB par défaut de '000000' peut continuer d'être utilisé pour le trafic conventionnel au mieux. De même, les codets '11x000' peuvent continuer d'être utilisés pour le trafic de contrôle de réseau.

Le groupe de PHB AF, conjointement avec les actions de conditionnement de trafic de bordure qui limitent la quantité de trafic dans chaque classe AF à un pourcentage (généralement différent) des ressources allouées à la classe, peut être utilisé pour obtenir le comportement global impliqué par les PHB de sélecteur de classe. Dans ce cas, il peut être approprié d'utiliser au sein d'un domaine DS certains des , ou tous les codets de sélecteur de classe comme alias des codets AF.

En plus des PHB de sélecteur de classe, tous les autres groupes de PHB peuvent coexister avec le groupe de PHB AF au sein du même domaine DS. Cependant, toute mise en œuvre de groupe de PHB AF devrait documenter ce qui suit :

- (a) Quels autres groupes de PHB, s'il en est, peuvent préempter les ressources de transmission spécifiquement allouées à chaque classe de PHB AF. Cette préemption NE DOIT PAS survenir dans le fonctionnement normal du réseau mais peut être appropriée dans certaines situations non habituelles — par exemple, le codet '11x000' peut préempter des ressources de transmission AF, pour donner la préséance à du trafic de contrôle de réseau de haut niveau inattendu lorsque il en est besoin.
- b) De quelle façon sont allouées les ressources "en excès" entre le groupe de PHB AF et d'autres groupes de PHB mis en œuvre. Par exemple, une fois que les allocations minimales sont données à chaque classe AF, toutes les ressources restantes pourraient être allouées équitablement entre les classes AF et le PHB par défaut. Dans un autre exemple, toutes les ressources restantes pourraient être allouées pour transmettre le trafic AF excédentaire, avec les ressources dédiées au PHB par défaut seulement lorsque toute la demande AF est satisfaite.

Le présent mémoire ne spécifie pas que des relations particulières devraient avoir lieu entre les groupes de PHB AF et d'autres groupes de PHB mis en œuvre ; il exige seulement que quelles que soient les relations choisies, elles soient documentées. Les mises en œuvre PEUVENT permettre que l'une et/ou l'autre de ces relations soient configurables. On pense que ce niveau de souplesse de configuration se révélera précieux pour de nombreux administrateurs de réseau.

8. Implications pour la sécurité

Afin de se protéger contre les attaques de déni de service, un fournisseur de domaine DS DEVRAIT limiter le trafic entrant dans le domaine aux profils souscrits. De plus, afin de protéger une liaison avec un domaine DS d'un usager contre des attaques de déni de service, le fournisseur de domaine DS DEVRAIT permettre au domaine DS d'usager de spécifier comment les ressources de la liaison sont allouées aux paquets AF. Si une offre de service exige que le trafic marqué avec un codet AF soit limité par des attributs tels que l'adresse de source ou de destination, il est de la responsabilité du nœud d'entrée d'un réseau de vérifier la validité de tels attributs.

D'autres considérations pour la sécurité sont traitées dans la [RFC2475] et la [RFC2474].

9. Droits de propriété intellectuelle

L'IETF a été informée de droits de propriété intellectuelle revendiqués à l'égard de certaines ou de toutes les spécifications contenues dans le présent document. Pour plus d'informations, consulter la liste en ligne des droits revendiqués.

10. Considérations pour l'IANA

Le présent document alloue douze codets, dont la liste figure à la section 6, dans le pôle 1 de l'espace des codes défini dans la [RFC2474].

Appendice : Exemple de services

Le groupe de PHB AF pourrait être utilisé pour mettre en œuvre, par exemple, ce qu'on appelle le service olympique, qui consiste en trois classes de service : bronze, argent, et or. Les paquets sont affectés à ces trois classes de telle sorte que les paquets de la classe or rencontrent la charge la plus légère (et aient donc une plus forte probabilité d'une transmission à temps que les paquets affectés à la classe argent). Le même type de relations existe entre la classe argent et la classe bronze. Si on le désire, les paquets au sein de chaque classe peuvent être séparés plus avant en leur donnant une préséance d'abandon faible, moyenne ou élevée.

Les classes de service bronze, argent, et or pourraient être transposées dans le réseau en classes AF 1, 2, et 3. De même, les préséances d'abandon faible, moyenne, et élevée peuvent être transposées en niveaux de préséance d'abandon AF de 1, 2, ou 3.

Le niveau de préséance d'abandon d'un paquet pourrait être alloué, par exemple, en utilisant un régulateur de trafic du type baquet percé, qui a comme paramètres un taux et une taille, qui est la somme de deux valeurs de salves : une taille contractuelle de salve et une taille de salve supplémentaire. Un paquet reçoit une faible préséance d'abandon si le nombre de jetons dans le baquet est supérieur à la taille de salve supplémentaire, une préséance moyenne si le nombre de jetons dans le baquet est supérieur à zéro, mais au plus de la taille de salve supplémentaire, et une préséance d'abandon élevée si le baquet est vide. Il peut être aussi nécessaire de fixer une limite supérieure à la quantité de trafic à préséance d'abandon élevée de la part d'un domaine DS d'utilisateur afin d'éviter la situation où une avalanche de paquets indélivrables à forte préséance d'abandon provenant d'un domaine DS d'utilisateur peut dénier le service à des paquets éventuellement livrables à haute préséance d'abandon d'autres domaines.

Une autre façon d'allouer le niveau de préséance d'abandon d'un paquet pourrait être de limiter le trafic d'utilisateur d'une classe de service olympique à un taux de crête donné et de le répartir équitablement à travers chaque niveau de préséance d'abandon. Cela donnerait un service à bande passante proportionnelle, ce qui distribue également la capacité disponible durant les périodes d'encombrement dans l'hypothèse où les usagers avec des microflux à forte bande passante ont souscrit à de plus fort taux de crête que les usagers qui ont des microflux à faible bande passante.

Le groupe de PHB AF pourrait aussi être utilisé pour mettre en œuvre un service à perte et faible latence qui utiliserait les classes AF sur provisionnées, si le taux d'arrivée maximum à cette classe est connu a priori dans chaque nœud DS. La spécification des services de contrôle d'admission requis va cependant au delà du domaine d'application du présent document. Si la faible perte n'est pas un objectif, un service à faible délai pourrait être mis en œuvre sans sur provisionnement en fixant une limite maximum faible à l'espace de mémoire tampon disponible pour une classe AF.

Références

- [RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and W. Weiss, "[Architecture pour services différenciés](#)", décembre 1998. (MàJ par [RFC3260](#)).
- [RFC2119] S. Bradner, "[Mots clés](#) à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997.
- [Clark] Clark, D. and Fang, W., Explicit Allocation of Best Effort Packet Delivery Service. IEEE/ACM Transactions on Networking, Volume 6, Number 4, août 1998, pp. 362-373.
- [Floyd] Floyd, S., and Jacobson, V., Random Early Detection gateways for Congestion Avoidance. IEEE/ACM Transactions on Networking, Volume 1, Number 4, août 1993, pp. 397-413.
- [RFC2474] Nichols, K., Blake, S., Baker, F. and D. Black, "Définition du champ [Services différenciés](#) (DS Field) dans les en-têtes IPv4 et IPv6", décembre 1998. (MàJ par [RFC3168](#), [RFC3260](#)) (P.S.)

Adresse des auteurs

Juha Heinanen
Telia Finland
Myyrmaentie 2
01600 Vantaa, Finland
mél : jh@telia.fi

Fred Baker
Cisco Systems
519 Lado Drive
Santa Barbara, California 93111
mél : fred@cisco.com

John Wroclawski	Walter Weiss
MIT Laboratory for Computer Science	Lucent Technologies
545 Technology Sq.	300 Baker Avenue, Suite 100,
Cambridge, MA 02139	Concord, MA 01742-2168
mél : jtw@lcs.mit.edu	mél : wweiss@lucent.com

Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (1999). Tous droits réservés.

Ce document et les traductions de celui-ci peuvent être copiés et diffusés, et les travaux dérivés qui commentent ou expliquent autrement ou aident à sa mise en œuvre peuvent être préparés, copiés, publiés et distribués, partiellement ou en totalité, sans restriction d'aucune sorte, à condition que l'avis de droits de reproduction ci-dessus et ce paragraphe soient inclus sur toutes ces copies et œuvres dérivées. Toutefois, ce document lui-même ne peut être modifié en aucune façon, par exemple en supprimant le droit d'auteur ou les références à l'Internet Society ou d'autres organisations Internet, sauf si c'est nécessaire à l'élaboration des normes Internet, auquel cas les procédures pour les droits de reproduction définis dans les procédures des normes d'Internet doivent être suivies, ou si nécessaire pour le traduire dans des langues autres que l'anglais.

Les permissions limitées accordées ci-dessus sont perpétuelles et ne seront pas révoquées par la Société Internet ou ses successeurs ou ayants droit.

Ce document et les renseignements qu'il contient sont fournis "TELS QUELS" et l'INTERNET SOCIETY et l'INTERNET ENGINEERING TASK FORCE déclinent toute garantie, expresse ou implicite, y compris mais sans s'y limiter, toute garantie que l'utilisation de l'information ici présente n'enfreindra aucun droit ou aucune garantie implicite de commercialisation ou d'adaptation à un objet particulier.

Remerciement

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