

Équipe d'ingénierie de l'Internet (IETF)
Request for Comments: 5670
Catégorie : Sur la voie de la normalisation
ISSN: 2070-1721

P. Eardley, éditeur, BT
novembre 2009

Traduction Claude Brière de L'Isle

Comportement de mesure et de marquage des nœuds PCN

Résumé
L'objectif de la notification de pré encombrement (PCN, *Pre-Congestion Notification*) est de protéger la qualité de service (QS) de flux non élastiques au sein d'un domaine Diffserv d'une façon simple, adaptable, et robuste. Le présent document définit les deux comportements de mesure et de marquage des nœuds PCN. La mesure et le marquage de seuil marquent tous les paquets PCN si le taux du trafic PCN est supérieur à un taux configuré ("taux de seuil de PCN"). La mesure et le marquage du trafic en excédent marquent une proportion de paquets PCN, telle que la quantité marquée égale le taux de trafic PCN en excès d'un taux configuré ("taux de PCN en excès"). Le niveau de marquage permet aux nœuds de bordure de PCN de prendre des décisions sur l'admission ou la terminaison des flux PCN.

Statut de ce mémoire
Ceci est un document de l'Internet sur la voie de la normalisation.

Le présent document a été produit par l'équipe d'ingénierie de l'Internet (IETF). Il représente le consensus de la communauté de l'IETF. Il a subi une révision publique et sa publication a été approuvée par le groupe de pilotage de l'ingénierie de l'Internet (IESG). Tous les documents approuvés par l'IESG ne sont pas candidats à devenir une norme de l'Internet ; voir la Section 2 de la RFC5741.

Les informations sur le statut actuel du présent document, tout errata, et comment fournir des réactions sur lui peuvent être obtenues à <http://www.rfc-editor.org/info/rfc5670>

Notice de droits de reproduction
Copyright (c) 2009 IETF Trust et les personnes identifiées comme auteurs du document. Tous droits réservés.

Le présent document est soumis au BCP 78 et aux dispositions légales de l'IETF Trust qui se rapportent aux documents de l'IETF (<http://trustee.ietf.org/license-info>) en vigueur à la date de publication de ce document. Prière de revoir ces documents avec attention, car ils décrivent vos droits et obligations par rapport à ce document. Les composants de code extraits du présent document doivent inclure le texte de licence simplifié de BSD comme décrit au paragraphe 4.e des dispositions légales du Trust et sont fournis sans garantie comme décrit dans la licence de BSD simplifiée.

Le présent document peut contenir des matériaux provenant de documents de l'IETF ou de contributions à l'IETF publiées ou rendues disponibles au public avant le 10 novembre 2008. La ou les personnes qui ont le contrôle des droits de reproduction sur tout ou partie de ces matériaux peuvent n'avoir pas accordé à l'IETF Trust le droit de permettre des modifications de ces matériaux en dehors du processus de normalisation de l'IETF. Sans l'obtention d'une licence adéquate de la part de la ou des personnes qui ont le contrôle des droits de reproduction de ces matériaux, le présent document ne peut pas être modifié en dehors du processus de normalisation de l'IETF, et des travaux dérivés ne peuvent pas être créés en dehors du processus de normalisation de l'IETF, excepté pour le formater en vue de sa publication comme RFC ou pour le traduire dans une autre langue que l'anglais.

Table des matières

1. Introduction.....	2
1.1 Terminologie.....	3
2. Comportements spécifiés de mesure et marquage de PCN.....	3
2.1 Fonction de classification de comportement agrégé.....	3
2.2 Fonction d'élimination.....	4
2.3 Fonction de mesure de seuil.....	4
2.4 Fonction de mesure de trafic excédentaire.....	4
2.5 Fonction de marquage.....	5
3. Considérations sur la sécurité.....	5
4. Remerciements.....	5

5. Références.....	5
5.1 Références normatives.....	5
5.2 Références pour information.....	6
Appendice A. Exemples d'algorithmes.....	6
A.1 Mesure et marquage de seuil.....	7
A.2 Mesure et marquage de trafic excédentaire.....	7
Appendice B. Notes de mise en œuvre.....	7
B.1 Trafic non PCN en compétition.....	8
B.2 Portée.....	8
B.3 Classification de comportement agrégé.....	9
B.4 Élimination.....	9
B.5 Mesure de seuil.....	10
B.6 Mesure de trafic excédentaire.....	10
B.7 Marquage.....	11
Adresse de l'éditeur.....	11

1. Introduction

L'objectif de la notification de pré encombrement (PCN, *Pre-Congestion Notification*) est de protéger la qualité de service (QS) de flux non élastiques au sein d'un domaine Diffserv d'une façon simple, adaptable, et robuste. Deux mécanismes sont utilisés : le contrôle d'admission pour décider d'admettre ou bloquer une nouvelle demande de flux, et (dans des circonstances anormales) la terminaison de flux pour décider de terminer certains des flux existants. Pour réaliser cela, le taux global du trafic de PCN est mesuré sur chaque liaison dans le domaine, et les paquets PCN sont marqués de façon appropriée quand certains taux configurés sont dépassés. Ces taux configurés sont en dessous du débit de la liaison, fournissant donc une notification aux nœuds de bordure sur les surcharges avant que l'encombrement se produise (d'où la "notification de pré encombrement"). Le niveau de marquage permet aux nœuds de bordure de prendre des décisions sur l'admission ou la terminaison. Au sein du domaine, le trafic de PCN est transmis dans une classe de trafic Diffserv prioritaire [RFC2475].

Le présent document définit les deux mesures et les comportements de marquage des nœuds de PCN. Leur but est de permettre aux nœuds à capacité PCN de donner une "annonce précoce" d'encombrement potentiel avant qu'il y ait une construction significative de paquets PCN dans leurs files d'attente. En résumé, leur objectif est :

- o Mesure et marquage de seuil : pour marquer tous les paquets PCN (avec une "marque de seuil") quand le débit binaire du trafic de PCN est supérieur à son taux de référence configuré ("taux de seuil de PCN").
- o Mesure et marquage de trafic en excès : quand le débit binaire de paquets PCN est supérieur à son taux de référence configuré ("taux d'excès de PCN") pour marquer les paquets PCN (avec une "marque d'excès de trafic") à un taux égal à la différence entre le taux de trafic de PCN et le taux d'excès de PCN.

Noter que bien que la [RFC3168] définisse un comportement de marquage d'encombrement par défaut en gros comme la détection précoce aléatoire (RED, *Random Early Detection*) elle permet de définir des solutions de remplacement ; le présent document définit une telle solution de remplacement.

La Section 2 décrit les fonctions impliquées, qui sont (voir la Figure 1) :

- o Le classement d'agrégat de comportement (BA, *Behaviour aggregate*) : décide si un paquet entrant est ou non un paquet PCN.
- o Élimination (facultatif) : élimine les paquets si la liaison est surchargée.
- o Mesure de seuil : détermine si le débit binaire du trafic PCN excède son taux de référence configuré (taux de seuil de PCN). La mesure opère sur tous les paquets PCN de la liaison, et non sur les flux individuels.
- o Mesure de trafic en excès : mesure de combien le débit du trafic de PCN excède son taux de référence configuré (taux d'excès de PCN). La mesure opère sur tous les paquets PCN sur la liaison, et non sur les flux individuels.
- o Marquage PCN : marque réellement les paquets PCN, si les fonctions de mesure indiquent de le faire.

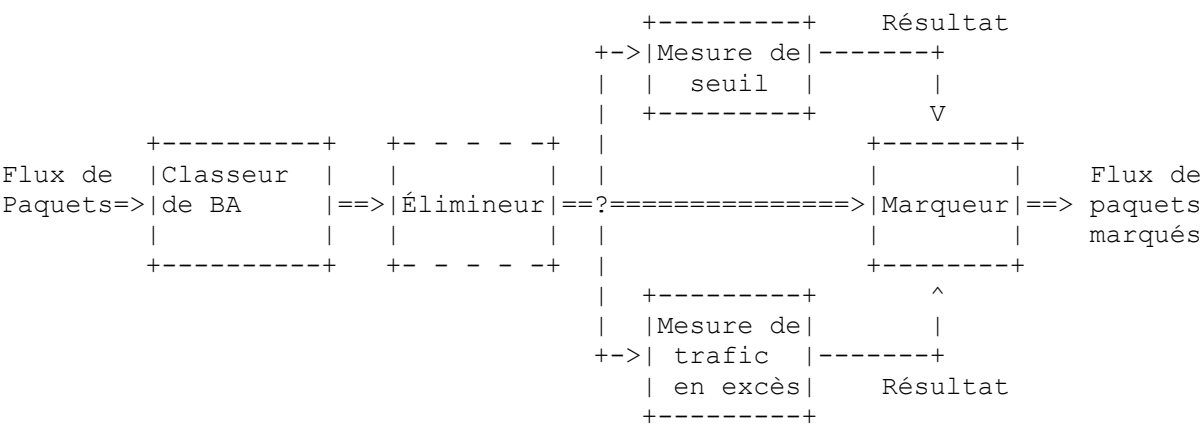


Figure 1 : Schéma des fonctions intérieures de nœud PCN

L'Appendice A donne un exemple des algorithmes qui satisfont les spécifications de la Section 2, et l'Appendice B donne des explications et des commentaires sur la Section 2. Les deux Appendices sont pour information.

L'architecture générale pour la PCN est décrite dans la [RFC5559], tandis que [Menth10] est une vue d'ensemble de PCN.

1.1 Terminologie

En plus de la terminologie définie dans les [RFC5559] et [RFC2474], les termes suivants sont définis :

- o Paquet non PCN en compétition : paquet non PCN qui partage une liaison avec des paquets PCN et est en compétition avec eux pour sa bande passante de transmission. Les paquets non PCN en compétition NE DOIVENT PAS être marqués PCN (seuls les paquets PCN peuvent être marqués PCN).

Note : En général, il n'est pas conseillé d'avoir du trafic non PCN en compétition.

Note : Il va probablement y avoir du trafic (comme du trafic au mieux) qui est transmis à une priorité inférieure à celle du trafic PCN ; bien qu'il partage la liaison avec le trafic PCN, il n'est pas en compétition pour la bande passante de transmission, et donc il n'est pas du trafic non PCN en compétition. Voir à l'Appendice B.1 la discussion sur le trafic non PCN en compétition.

- o Paquet mesuré : paquet mesuré par les fonctions de mesure spécifiées aux paragraphes 2.3 et 2.4. Un paquet PCN DOIT être traité comme un paquet mesuré (avec l'exception mineure notée au paragraphe 2.4). Un paquet non PCN en compétition PEUT être traité comme un paquet mesuré.

1.1.1 Langage des exigences

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. Comportements spécifiés de mesure et marquage de PCN

Cette section définit les deux comportements de mesure et de marquage de PCN. Les descriptions sont fonctionnelles et ne sont pas destinées à restreindre la mise en œuvre. Les Appendices pour information complètent cette section.

2.1 Fonction de classification de comportement agrégé

Un nœud PCN DOIT classer un paquet comme paquet PCN si la valeur de ses champs Codet de services différenciés (DSCP, *Differentiated Services Code Point*) et Notification explicite d'encombrement (ECN, *Explicit Congestion Notification*) correspond à un codet à capacité PCN, comme défini dans le schéma de codage applicable au domaine PCN

(par exemple, la [RFC5696] définit le codage de base). Autrement, le paquet NE DOIT PAS être classé comme paquet PCN.

Un nœud PCN DOIT classer un paquet comme paquet non PCN en compétition si il n'est pas un paquet PCN et si il est en concurrence avec des paquets PCN pour sa bande passante de transmission sur une liaison.

2.2 Fonction d'élimination

Note : Si la file d'attente du nœud PCN déborde, alors les paquets sont naturellement éliminés. Ce paragraphe décrit une action supplémentaire.

Sur toutes les liaisons dans le domaine PCN, l'élimination PEUT être faite en mesurant d'abord tous les paquets mesurés pour déterminer si le taux de trafic mesuré sur la liaison est supérieur au taux permis pour ce trafic ; si le taux du trafic mesuré est trop élevé, on élimine alors les paquets mesurés.

Si le nœud PCN élimine les paquets PCN, alors :

- o les paquets PCN qui arrivent au nœud PCN déjà marqués comme excès de trafic DEVRAIENT être éliminés de préférence ;
- o la mesure de trafic en excès du nœud PCN NE DEVRAIT PAS mesurer les paquets PCN qu'elle élimine.

2.3 Fonction de mesure de seuil

Un nœud PCN DOIT mettre en œuvre une mesure de seuil ayant un comportement fonctionnellement équivalent à ce qui suit :

La mesure agit comme un baquet de jetons, qui est dimensionné en bits et a un taux de référence configuré (en bits par seconde). La quantité de jetons dans le baquet est appelée F_{tm} . Les jetons sont ajoutés au taux de référence (taux de seuil de PCN) à une valeur maximum BS_{tm} . Les jetons sont retirés à concurrence de la taille en bits du paquet mesuré, avec un minimum $F_{tm} = 0$. (Explication des abréviations : F pour remplissage du baquet de jetons, BS pour taille de baquet, et tm pour mesure de seuil.)

Le baquet de jetons a une profondeur intermédiaire configurée, appelée seuil. Si $F_{tm} < \text{seuil}$, alors la mesure indique à la fonction de marquage que le paquet est à marquer comme seuil ; autrement, il ne l'est pas.

2.4 Fonction de mesure de trafic excédentaire

Un paquet NE DEVRAIT PAS être mesuré (par cette fonction de mesure d'excès de trafic) dans les deux cas suivants :

- o si le paquet PCN est déjà marqué excès de trafic à l'arrivée au nœud PCN,
- o si ce nœud PCN élimine le paquet.

Autrement, le paquet PCN DOIT être traité comme un paquet mesuré -- c'est-à-dire, il est mesuré par la mesure d'excès de trafic.

Un nœud PCN DOIT mettre en œuvre une mesure d'excès de trafic. La mesure d'excès de trafic DEVRAIT indiquer les paquets à marquer comme excès de trafic, indépendamment de leur taille ("marquage indépendant de la taille de paquet") ; si le "marquage indépendant de la taille de paquet" n'est pas mis en œuvre, la mesure d'excès de trafic DOIT alors utiliser le comportement "classique" de mesure.

Pour le comportement "classique" de mesure, la mesure d'excès de trafic a une fonction de comportement équivalente à ce qui suit :

La mesure agit comme un baquet de jetons, qui est dimensionné en bits et a un taux de référence configuré (en bits par seconde). La quantité de jetons dans le baquet est appelée F_{etm} . Les jetons sont ajoutés au taux de référence (taux d'excès de PCN) à une valeur maximum BS_{etm} . Les jetons sont retirés à concurrence de la taille en bits du paquet mesuré, jusqu'à un minimum de $F_{etm} = 0$. Si le baquet de jetons est vide ($F_{etm} = 0$) alors la mesure indique à la fonction de marquage que le paquet est à marquer comme trafic en excédent. (Explication des abréviations : F pour remplissage du baquet de jetons, BS pour taille de baquet, et etm pour mesure de trafic en excédent.)

Pour le "marquage indépendant de la taille de paquet", la mesure de trafic en excédent a une fonction de comportement équivalente à ce qui suit :

La mesure agit comme un baquet de jetons, qui est dimensionné en bits et a un taux de référence configuré (en bits par

seconde). La quantité de jetons dans le baquet est appelée F_{etm} . Les jetons sont ajoutés au taux de référence (taux d'excès de PCN) à une valeur maximum BS_{etm} . Si le baquet de jetons n'est pas négatif, alors les jetons sont retirés à concurrence de la taille en bits du paquet mesuré (et la mesure n'indique pas à la fonction de marquage que le paquet est à marquer comme trafic en excédent). Si le baquet de jetons est négatif ($F_{\text{etm}} < 0$) alors la mesure indique à la fonction de marquage que le paquet est à marquer comme trafic en excédent (et aucun jeton n'est retiré). (Explication des abréviations : F pour remplissage du baquet de jetons, BS pour taille de baquet, et etm pour mesure de trafic en excédent.)

Autrement, la mesure NE DOIT PAS indiquer de marquage.

2.5 Fonction de marquage

Un paquet PCN DOIT être marqué pour refléter les résultats de mesures en réglant de façon appropriée son état de codage, comme spécifié par le schéma de codage spécifique qui s'applique dans le domaine PCN. Un choix cohérent de schéma de codage DOIT être fait sur tout le domaine de PCN.

Un nœud PCN NE DOIT PAS :

- o marquer PCN un paquet qui n'est pas un paquet PCN ;
- o changer un paquet non PCN en paquet PCN ;
- o changer un paquet PCN en paquet non PCN.

Note : bien que des paquets non PCN en compétition PUISSENT être mesurés, ils NE DOIVENT PAS être marqués PCN.

3. Considérations sur la sécurité

Il est supposé que tous les nœuds PCN sont à capacité PCN et sont de confiance pour des mesures de PCN fiables et le marquage PCN. Si ce n'est pas le cas, il y a alors de nombreuses attaques potentielles. Par exemple, un nœud intérieur PCN félon pourrait marquer comme PCN tous les paquets afin qu'aucun flux ne soit admis. Une autre possibilité est qu'il ne marque aucun paquet comme PCN, même quand il y a pré-encombrement.

Noter que les nœuds intérieurs PCN ne sont pas conscients des flux. Cela empêche certaines attaques qui ciblent des flux spécifiques dans le plan des données -- par exemple, pour du déni de service (DoS) ou de l'espionnage.

En ce qui concerne les opérations et la gestion de la sécurité, PCN ajoute peu de spécificités aux bonnes pratiques générales exigées dans ce champ [RFC4778]. Par exemple, il peut être raisonnable pour un nœud PCN de déclencher une alarme si il marque de façon persistante comme PCN.

Les considérations sur la sécurité sont discutées plus en détails dans la [RFC5559].

4. Remerciements

Le présent document est le résultat d'une large collaboration au sein du groupe de travail PCN. Parmi les autres contributeurs les plus actifs au développement des idées spécifiées dans ce document ont doit citer Jozef Babiarz, Bob Briscoe, Kwok-Ho Chan, Anna Charny, Georgios Karagiannis, Michael Menth, Toby Moncaster, Daisuke Satoh, et Joy Zhang. L'Appendice A est fondé sur un texte de Michael Menth.

Le présent document est un développement de [Briscoe06-2]. Ses auteurs sont donc aussi des contributeurs du présent document : Jozef Babiarz, Attila Bader, Bob Briscoe, Kwok-Ho Chan, Anna Charny, Stephen Dudley, Philip Eardley, Georgios Karagiannis, Francois Le Faucheur, Vassilis Liatsos, Dave Songhurst, et Lars Westberg.

Merci à tous ceux qui ont fait des commentaires sur le document : Joe Babiarz, Fred Baker, David Black, Bob Briscoe, Ken Carlberg, Anna Charny, Ralph Droms, Mehmet Ersue, Adrian Farrel, Ruediger Geib, Wei Gengyu, Fortune Huang, Christian Hublet, Ingemar Johansson, Georgios Karagiannis, Alexey Melnikov, Michael Menth, Toby Moncaster, Dimitri Papadimitriou, Tim Polk, Daisuke Satoh, et Magnus Westerlund.

5. Références

5.1 Références normatives

- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. DOI 10.17487/RFC2119, (MàJ par [RFC8174](#))

5.2 Références pour information

- [Briscoe06-1] Briscoe, B., Eardley, P., Songhurst, D., Le Faucheur, F., Charny, A., Babiarz, J., Chan, K., Dudley, S., Karagiannis, G., Bader, A., and L. Westberg, "An edge-to-edge Deployment Model for Pre-Congestion Notification: Admission Control over a DiffServ Region", Travail en cours, octobre 2006.
- [Briscoe06-2] Briscoe, B., Eardley, P., Songhurst, D., Le Faucheur, F., Charny, A., Liatsos, V., Babiarz, J., Chan, K., Dudley, S., Karagiannis, G., Bader, A., and L. Westberg, "Pre-Congestion Notification marking", Travail en cours, octobre 2006..
- [Charny07] Charny, A., Babiarz, J., Menth, M., and X. Zhang, "Comparison of Proposed PCN Approaches", Travail en cours, novembre 2006..
- [Menth10] Menth, M., Lehrieder, F., Briscoe, B., Eardley, P., Moncaster, T., Babiarz, J., Chan, K., Charny, A., Karagiannis, G., Zhang, X., Taylor, T., Satoh, D., and R. Geib, "A Survey of PCN-Based Admission Control and Flow Termination", IEEE Communications Surveys and Tutorials, 2010 (third issue), <<http://www3.informatik.uni-wuerzburg.de/staff/menth/Publications/papers/Menth08-PCN-Overview.pdf>>.
- [RFC2474] K. Nichols, S. Blake, F. Baker et D. Black, "Définition du [champ Services différenciés](#) (DS) dans les en-têtes IPv4 et IPv6", DOI 10.17487/RFC2474, décembre 1998. (P.S. ; MàJ par [RFC3168](#), [RFC3260](#), [RFC8436](#))
- [RFC2475] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang et W. Weiss, "[Architecture pour services différenciés](#)", décembre 1998, DOI 10.17487/RFC2475. (MàJ par [RFC3260](#))
- [RFC3168] K. Ramakrishnan et autres, "Ajout de la [notification explicite d'encombrement](#) (ECN) à IP", DOI 10.17487/RFC3168, septembre 2001. (P.S. ; MàJ par [RFC8311](#))
- [RFC4778] M. Kaeo, "Pratiques actuelles du fonctionnement de la sécurité dans les environnements de fournisseur de service Internet", janvier 2007. (Information)
- [RFC5127] K. Chan, J. Babiarz, F. Baker, "[Agrégation de classes](#) de service DiffServ", février 2008. (Information)
- [RFC5559] P. Eardley, éd., "Architecture de notification de pré-encombrement (PCN)", juin 2009. (Information)
- [RFC5696] T. Moncaster, B. Briscoe, M. Menth, "Codage de base et transport des informations de pré encombrement", novembre 2009. (P. S. ; remplacée par la RFC6660)
- [RFC5865] F. Baker, J. Polk, M. Dolly, "Codet DS (DSCP) pour trafic à capacité admise", mai 2010. (MàJ [RFC4542](#), [RFC4594](#)). (P.S.)
- [RFC6661] A. Charny, et autres, "Comportement de nœud frontière de notification de pré encombrement (PCN) pour le mode de fonctionnement de charge contrôlée (CL)", juillet 2012. (Expérimental)
- [RFC7141] B. Briscoe, J. Manner, "Notification d'encombrement d'octet et de paquet", BCP0041, février 2014. (MàJ [RFC2309](#), [RFC2914](#))

Appendice A. Exemples d'algorithmes

Note : Le présent Appendice est pour information, et n'est pas normatif. C'est un exemple des algorithmes qui mettent en œuvre la Section 2 et il est fondé sur [Charny07] et [Menth10].

On ne tente pas d'optimiser les algorithmes. Les fonctions de mesure et de marquage sont mises en œuvre ensemble. Il est supposé que trois états de codage sont disponibles (un pour le marquage de seuil, un pour le marquage de trafic en excès, et un pour non marqué). Il est supposé que tous les paquets mesurés sont des paquets PCN et que la liaison n'est jamais surchargée. Pour le marquage de trafic en excès, le "marquage indépendant de la taille de paquet" s'applique.

A.1 Mesure et marquage de seuil

Un baquet de jetons a les paramètres suivants :

- * taux de seuil de PCN : taux de jetons du baquet de jetons (en bits/seconde)
- * BS_tm : profondeur du baquet de jetons (en bits)
- * seuil : seuil de marquage du baquet de jetons (en bits)
- * dernière mise à jour : heure de la dernière mise à jour du baquet de jetons (en secondes)
- * F_tm : quantité de jetons dans le baquet de jetons (en bits)

Un paquet PCN a les paramètres suivants :

- * taille de paquet : taille du paquet PCN (en bits)
- * marquage de paquet : état de codage PCN du paquet

De plus, il y a le paramètre :

maintenant : l'heure courante (en secondes)

Les étapes suivantes sont effectuées quand un paquet PCN arrive sur une liaison :

- * $F_{tm} = \min(BS_{tm}, F_{tm} + (\text{maintenant} - \text{dernière mise à jour}) * \text{taux de seuil de PCN})$; // ajoute des jetons au baquet de jetons
- * $F_{tm} = \max(0, F_{tm} - \text{taille de paquet})$; // retire des jetons au baquet de jetons
- * si $(F_{tm} < \text{seuil})$ ET (marquage de paquet != trafic en excès marqué) alors marquage de paquet = seuil-marqué; // faire le marquage de seuil, mais ne pas marquer à nouveau les paquets déjà marqués comme trafic en excès.
- * dernière mise à jour = maintenant // Note : "maintenant" a la même valeur qu'à l'étape 1.

A.2 Mesure et marquage de trafic excédentaire

Un baquet de jetons a les paramètres suivants :

- * taux de PCN en excès : taux de jetons du baquet de jetons (bits/seconde)
- * BS_ets : profondeur de TB dans le baquet de jetons (en bits)
- * dernière mise à jour : heure de la dernière mise à jour du baquet de jetons (en secondes)
- * F_ets : quantité de jetons dans le baquet de jetons (en bits)

Un paquet PCN a les paramètres suivants :

- * taille de paquet : taille du paquet PCN (en bits)
- * marquage de paquet : état de codage PCN du paquet

De plus, il y a le paramètre :

maintenant : l'heure courante (en secondes)

Les étapes suivantes sont effectuées quand un paquet PCN arrive sur une liaison :

- * $F_{ets} = \min(BS_{ets}, F_{ets} + (\text{maintenant} - \text{dernière mise à jour}) * \text{taux de PCN en excès})$; // ajoute des jetons au baquet de jetons
- * si (marquage de paquet != excès de trafic marqué) alors // ne pas mesurer les paquets qui sont déjà marqués trafic en excès
 - + si $(F_{ets} < 0)$ alors marquage de paquet = excès de trafic marqué ; // faire le marquage de trafic en excès.
 - + L'algorithme assure que c'est indépendant de la taille du paquet
 - + autrement $F_{ets} = F_{ets} - \text{taille de paquet}$; // supprime des jetons du baquet de jetons si il ne marque pas le

paquet

- * dernière mise à jour = maintenant // Note : "maintenant" a la même valeur qu'à l'étape 1.

Appendice B. Notes de mise en œuvre

Note : cet Appendice est pour information, et n'est pas normatif. Il commente la Section 2, incluant un raisonnement sur si les "DOIT" ou "DEVRAIT" sont exigés. Pour des directives sur les considérations de fonctionnement et de gestion, voir la [RFC5559].

B.1 Trafic non PCN en compétition

En général, il n'est pas conseillé d'avoir du trafic non PCN en compétition, essentiellement parce que la quantité imprévisible de trafic non PCN en compétition rend les mécanismes de PCN moins précis et donc réduit la capacité des PCN à protéger la QS des flux de PCN admis [RFC5559]. Mais si il y a du trafic non PCN en compétition, alors :

1. Il devrait y avoir un mécanisme pour le limiter, par exemple :
 - * Limiter le taux auquel le trafic non PCN en compétition peut être transmis sur chaque liaison dans le domaine PCN. Une méthode pour réaliser cela est de mettre en file d'attente les paquets non PCN en compétition séparément des paquets PCN et de limiter le taux de programmation des premiers. Une autre méthode est d'éliminer les paquets non PCN en compétition qui dépassent un certain taux.
 - * Réguler le trafic non PCN en compétition aux nœuds d'entrée de PCN, comme dans l'architecture Diffserv, par exemple. Cependant, des accords de conditionnement de trafic statiques de Diffserv risquent de concentrer une surcharge de trafic provenant de plusieurs nœuds d'entrée de PCN sur une liaison.
 - * Par conception, il est connu que le niveau de trafic non PCN en compétition est toujours très petit -- peut-être qu'il consiste seulement en messages de contrôle de l'opérateur.
2. En général, les mécanismes de PCN devraient tenir compte du trafic non PCN en compétition, afin d'améliorer la précision de la décision sur l'admission (ou la terminaison) d'un flux PCN. Par exemple :
 - * le trafic non PCN en compétition contribue aux mesures de PCN ; les paquets non PCN en compétition sont traités comme des paquets mesurés ;
 - * chaque nœud PCN, sur ses liaisons : (1) réduit les taux de référence (taux de seuil PCN et taux de PCN en excès) afin de permettre au trafic non PCN en compétition de "sortir la tête" ; (2) limite le taux de transmission maximum du trafic non PCN en compétition à être moins que de "sortir la tête". Dans ce cas, les paquets non PCN en compétition ne sont pas traités comme des paquets mesurés.
3. L'opérateur devrait décider de l'action appropriée. L'élimination est discutée plus loin dans l'Appendice B.4.

Une exemple spécifique de trafic non PCN en compétition est celui où le codet Diffserv compatible PCN est un de ceux que la [RFC5865] définit comme convenable pour utiliser avec le contrôle d'admission et il y a un tel trafic non PCN dans le domaine PCN. Un exemple similaire pourrait être donné pour des codets Diffserv d'agrégat de traitement en temps réel [RFC5127]. Dans ces cas, le trafic PCN et le trafic non PCN en compétition sont distingués par des valeurs différentes du champ ECN [RFC5696].

Un autre exemple serait si il y a plus d'un codet Diffserv compatible PCN dans un domaine PCN. Par exemple, supposons qu'il y ait deux BA PCN traités à des priorités différentes. Alors pour ce qui concerne le BA PCN de priorité inférieure, le trafic PCN de priorité supérieure doit être traité comme du trafic non PCN en compétition.

B.2 Portée

Il peut être connu, par exemple par le dessin de la topologie de réseau, que certaines liaisons ne peuvent jamais être pré-encombrées (même dans des circonstances inhabituelles, comme après la défaillance de certaines liaisons). Il n'est alors pas besoin de déployer le comportement de mesure et marquage de PCN sur ces liaisons.

Les mesures peuvent être mises en œuvre sur l'interface entrante ou sortante d'un nœud PCN. Il se peut que le matériel existant puisse seulement prendre en charge une mesure par interface entrante et une par interface sortante. Alors, par exemple, la mesure de seuil pourrait être faite sur toutes les interfaces entrantes et la mesure de trafic en excès sur toutes les interfaces sortantes ; noter que le même choix doit être fait pour toutes les liaisons dans un domaine PCN pour assurer que

les deux comportements de mesure sont appliqués exactement une fois pour toutes les liaisons.

Le codage de base [RFC5696] spécifie seulement deux états de codage (marqué PCN et non marqué). Dans ce cas, "trafic marqué en excès" signifie un paquet qui est marqué PCN par suite de la fonction de mesure du trafic en excès, et "marquage de seuil" signifie un paquet qui est marqué PCN par suite de la fonction de mesure de seuil. Pour ce qui concerne la terminologie, cette interprétation est cohérente avec celle définie dans la [RFC5559]. Noter qu'un déploiement doit faire un choix cohérent à travers le domaine PCN sur si "marqué PCN" est interprété comme trafic marqué en excès ou marquage de seuil.

Noter que même si il y a seulement deux états de codage, il est quand même exigé que les deux mesures soient mises en œuvre, afin de faciliter la compatibilité entre équipements et supprimer les options de configuration et la complexité qui y est associée. Un matériel avec une disponibilité limitée des baquets de jetons pourrait être configuré à fonctionner seulement avec une des mesures, mais il doit être possible d'activer l'une et l'autre. Cependant, dans le scénario avec deux états de codage, les indications provenant d'une des mesures sont ignorées par la fonction de marquage, elles peuvent être enregistrées ou on peut agir sur elles d'une autre façon, par exemple, par le système de gestion ou un protocole de signalisation explicite ; ces considérations sortent du domaine d'application de ce document.

B.3 Classification de comportement agrégé

La configuration des nœuds PCN va définir quelles valeurs des champs DSCP et ECN indiquent un paquet PCN dans un domaine PCN particulier. Par exemple, la [RFC5696] définit le codage de base.

La configuration va aussi définir quelles valeurs des champs DSCP et ECN indiquent un paquet non PCN en compétition dans un domaine PCN particulier.

B.4 Élimination

L'objectif de la fonction d'élimination est de minimiser le délai de mise en file d'attente subi par le trafic mesuré à un nœud PCN, dans la mesure où le trafic PCN (et peut-être le trafic non PCN en compétition) est supposé être du trafic inélastique généré par des applications en temps réel. En pratique, il serait défini comme excédant un profil de trafic spécifique, normalement fondé sur un baquet de jetons.

Si il n'y a pas de trafic non PCN en compétition, on ne s'attend pas à ce que la fonction d'élimination soit nécessaire, car les mécanismes d'admission et de terminaison de flux de PCN limitent la quantité de trafic PCN. Même là, il pourrait encore être mis en œuvre comme arrêt contre une mauvaise configuration du domaine PCN, par exemple.

Si il y a du trafic non PCN en compétition, alors les détails de la fonction d'élimination vont dépendre de la façon dont la mise en œuvre de routeur traite les deux sortes de trafic :

1. une file d'attente commune pour le trafic PCN et le trafic non PCN en compétition, avec un conditionneur de trafic pour le trafic non PCN en compétition ; ou
2. des files d'attentes séparées, et dans ce cas la quantité de trafic non PCN en compétition peut être limitée en limitant le taux auquel le programmeur (pour le trafic non PCN en compétition) transmet les paquets.

(La discussion est fondée sur celle de la [RFC5865].)

Noter que seule l'élimination des paquets est permise. Dégrader les paquets à une BA de priorité inférieure n'est pas permis (voir l'Appendice B.7) car cela conduirait à un changement de l'ordre des paquets. Le formatage (*Shaping*) ("processus de retardement des paquets" [RFC2475]) n'est pas convenable si le trafic vient d'applications en temps réel.

L'élimination préférentielle du trafic non PCN en compétition : en général, elle est raisonnable pour que le trafic non PCN en compétition obtienne un traitement plus dur que le trafic de PCN (c'est-à-dire que les paquets non PCN en compétition soient éliminés de préférence) parce que les mécanismes d'admission et terminaison de flux de PCN sont plus forts que les mécanismes qui vont probablement être appliqués au trafic non PCN en compétition. Les mécanismes de PCN signifient aussi qu'un élimineur ne devrait pas être nécessaire pour le trafic PCN.

L'élimination préférentielle des paquets marqués trafic en excès : le paragraphe 2.2 spécifie, "Si le nœud PCN élimine des paquets PCN, alors ... les paquets PCN qui arrivent au nœud PCN déjà marqués trafic en excès DEVRAIENT être éliminés de préférence". En bref, la raison en est que, avec le comportement bordure de "charge contrôlée" [RFC6661], cela évite une sur terminaison dans le cas de plusieurs goulets d'étranglement dans le domaine PCN [Charny07]. Une explication plus

complète est la suivante. Le comportement d'élimination optimale dépend du comportement de bordure particulier [Menth10]. Un seul comportement d'élimination est défini, car il est plus simple à normaliser, à mettre en œuvre, et faire fonctionner. Le comportement d'élimination normalisé est au moins adéquat pour tous les comportements de bordure (et bon pour certains) tandis que d'autres ne le sont pas (par exemple, avec l'élimination en queue, beaucoup trop de trafic peut être terminé avec le comportement de bordure de "charge contrôlée", dans le cas de plusieurs goulets d'étranglement dans le domaine PCN [Charny07]). Le comportement d'élimination est défini comme un "DEVRAIT", plutôt que comme un "DOIT", en reconnaissant que d'autres comportements d'élimination peuvent être préférés dans des circonstances particulières, par exemple : (1) avec le comportement bordure de terminaison de "flux marqué", l'élimination préférentielle de paquets non marqués peut être meilleure [Menth10] ; (2) l'élimination en queue peut rendre le comportement de marquage PCN plus facile à mettre en œuvre sur les routeurs courants.

Ce que signifie exactement "éliminé de préférence" est laissé à la mise en œuvre. Ainsi que ce qu'il convient de faire si il n'y a pas de paquets PCN marqués trafic en excès disponibles à un instant particulier.

Le paragraphe 2.2 spécifie aussi "la mesure de trafic en excès du nœud PCN NE DEVRAIT PAS mesurer les paquets PCN qu'elle élimine." Cela évite la sur-terminaison [Menth10]. Effectivement, cela signifie que la fonction d'élimination (si elle est présente) devrait être faite avant les fonctions de mesure -- ce qui est naturel.

B.5 Mesure de seuil

La description est en termes de "baquet de jetons avec seuil" (que [Briscoe06-1] voit comme une file d'attente virtuelle). Cependant, la description n'est pas destinée à normaliser la mise en œuvre.

Le taux de référence de la mesure de seuil (taux de seuil de PCN) est configuré à moins que le taux alloué à la classe de trafic PCN. Aussi, le taux de seuil de PCN est moins, ou éventuellement égal, que le taux de PCN en excès.

Le paragraphe 2.3 spécifie, "Si $F_{tm} < \text{seuil}$, alors la mesure indique à la fonction de marquage que le paquet est à marquer comme seuil ; autrement, il ne l'est pas." Noter qu'un paquet PCN est marqué sans biais explicite supplémentaire pour la taille du paquet.

Le comportement doit être fonctionnellement équivalent à la description du paragraphe 2.3. "Fonctionnellement équivalent" signifie que le comportement observable de "boîte noire" est le même ou très similaire, par exemple, si précisément le même ensemble de paquets est marqué ou si l'ensemble est décalé de un paquet. Il est destiné à permettre une certaine liberté de mise en œuvre sur des questions comme :

- o Si des jetons sont ajoutés au baquet de jetons à des intervalles réguliers ou seulement quand un paquet est traité.
- o Si la profondeur du nouveau baquet de jetons est calculée avant ou après qu'il est décidé de marquer le paquet comme PCN. L'effet de cela est simplement de décaler la séquence de marques de un paquet.
- o Quand le baquet de jetons est presque vide et qu'un paquet arrive plus grand que F_{tm} , alors le changement précis de F_{tm} appartient à la mise en œuvre. Par exemple :
 - * régler $F_{tm} = 0$ et indiquer le marquage de seuil à la fonction de marquage ;
 - * vérifier si $F_{tm} < \text{seuil}$ et si il l'est, indiquer alors marquage de seuil à la fonction de marquage ; puis régler $F_{tm} = 0$.
 - * laisser F_{tm} inchangé et indiquer le marquage de seuil à la fonction de marquage.
- o De façon similaire, quand le baquet de jetons est presque complet et qu'un paquet arrive plus grand que $(BS_{tm} - F_{tm})$ alors le changement précis à F_{tm} appartient à la mise en œuvre.

Noter que tous les paquets PCN, même si ils sont déjà marqués, sont mesurés par la fonction de mesure de seuil (à la différence de la fonction de mesure de trafic en excès) parce que tous les paquets devraient contribuer à la décision de si il y a de la place pour un nouveau flux.

B.6 Mesure de trafic excédentaire

La description est en termes de baquet de jetons, cependant la mise en œuvre n'est pas normalisée.

Le taux de référence de mesure de trafic en excès (taux de PCN en excès) est configuré à moins que (ou éventuellement égal au) le taux alloué à la classe de trafic PCN. Aussi, le taux de PCN en excès est supérieur à, ou éventuellement égal, au taux de seuil de PCN.

Comme au paragraphe B.5, "fonctionnellement équivalent" permet une certaine souplesse de mise en œuvre, par exemple, l'algorithme exact quand le baquet de jetons est presque vide ou presque plein.

Le paragraphe 2.4 spécifie, "Un paquet NE DEVRAIT PAS être mesuré (par cette fonction de mesure de trafic en excès) ... si le paquet est déjà marqué comme trafic en excès à l'arrivée au nœud PCN". Cela évite la sur-termination (avec certains comportements bordures) dans le cas où le trafic PCN passe par plusieurs goulets d'étranglement dans le domaine PCN [Charny07]. Noter qu'une mise en œuvre pourrait déterminer si le paquet est déjà marqué comme trafic en excès comme partie intégrante de sa fonction de classification de BA. Le comportement est défini comme un "NE DEVRAIT PAS", plutôt que comme un "NE DOIT PAS", parce que il peut être un peu plus difficile à mettre en œuvre qu'une fonction de mesure aveugle aux marquages de paquets précédents.

Le paragraphe 2.4 spécifie, "Un paquet NE DEVRAIT PAS être mesuré (par cette fonction de mesure de trafic en excès) ... si ce nœud PCN élimine le paquet". Cela évite la sur-termination [Menth10]. (Une déclaration similaire pourrait aussi être faite pour la fonction de mesure de seuil mais n'est pas pertinente, car une liaison qui est surchargée va déjà être substantiellement pré-encombrée et donc va marquer comme seuil tous les paquets.) Il semble naturel d'effectuer la fonction d'élimination avant les fonctions de mesure, bien que pour certains équipements cela puisse être plus difficile à mettre en œuvre ; donc, le comportement est défini comme un "NE DEVRAIT PAS", plutôt que comme un "NE DOIT PAS".

"Marquage indépendant de la taille de paquet" -- marquage de trafic en excès qui est indépendant de la taille de paquet -- est spécifié comme un "DEVRAIT" plutôt que comme un "DOIT" au paragraphe 2.4 parce que il peut être un peu plus difficile à certains équipements de le mettre en œuvre, et l'impact de ne pas le faire est indésirable mais modéré (le trafic suffisant est terminé, mais des flux avec de grands paquets ont plus de probabilité d'être terminés). Avec le comportement "classique" de mesure de trafic en excès, les grands paquets ont plus de probabilité d'être marqués comme trafic en excès que les petits paquets (parce que les paquets sont marqués si le nombre de jetons dans le baquet de jetons est plus petit que la taille de paquet). Cela signifie que, avec certains comportements bordures, les flux avec de grands paquets sont plus susceptibles d'être terminés que les flux avec de petits paquets ([RFC7141], [Menth10]). Le "marquage indépendant de la taille de paquet" peut être réalisé par une petite modification de la mesure de trafic en excès "classique". Le nombre de jetons dans le baquet peut devenir négatif ; si ce nombre est négatif à l'arrivée d'un paquet, le paquet est marqué ; autrement, la quantité de jetons égale à la taille du paquet est retirée du baquet. Noter qu'avec le "marquage indépendant de la taille de paquet", soit le paquet est marqué soit les jetons sont retirés -- jamais les deux. Donc, le baquet de jetons ne peut pas devenir plus négatif que la taille maximum de paquet sur la liaison. L'algorithme décrit à l'Appendice A met en œuvre ce comportement.

Noter que BS_ets est indépendant de BS_tm, F_ets est indépendant de F_tm (sauf qu'un paquet peut changer les deux) et les deux taux configurés (taux de PCN en excès et taux de seuil de PCN) sont indépendants (sauf que taux de PCN en excès $\geq$ taux de seuil de PCN).

B.7 Marquage

Le paragraphe 2.5 définit, "Un nœud PCN NE DOIT PAS ...changer un paquet PCN en un paquet non PCN". Cela signifie qu'il n'est pas permis à un nœud PCN de dégrader un paquet PCN en un BA Diffserv de priorité inférieure (donc, la dégradation n'est pas permise comme solution de remplacement de l'élimination).

Le paragraphe 2.5 définit, "Un nœud PCN NE DOIT PAS ... marquer comme PCN un paquet qui n'est pas un paquet PCN". Cela signifie que dans le scénario où des paquets non PCN en compétition sont traités comme des paquets mesurés, une mesure peut indiquer qu'un paquet est à marquer comme PCN, mais la fonction de marquage sait qu'il ne peut pas être marqué. Ce que faire exactement dans ce cas est laissé à la mise en œuvre ; une possibilité simple est de marquer le paquet PCN suivant. Noter que sauf si les paquets PCN sont une large fraction de tous les paquets mesurés, les mécanismes de PCN ne peuvent pas bien fonctionner.

Bien que les fonctions de mesure soient décrites séparément de la fonction de marquage, elles peuvent être mises en œuvre de façon intégrée

Adresse de l'éditeur

Philip Eardley
BT
Adastral Park, Martlesham Heath
Ipswich IP5 3RE

UK

mél : philip.eardley@bt.com