Groupe de travail Réseau
Request for Comments : 2475
Catégorie : Information
décembre 1998

Traduction Claude Brière de L'Isle

S. Blake, Torrent Networking Technologies
D. Black, EMC Corporation
M. Carlson, Sun Microsystems
E. Davies, Nortel UK
Z. Wang, Bell Labs Lucent Technologies
W. Weiss, Lucent Technologies

Architecture pour les services différenciés

Statut de ce mémoire

Le présent mémoire apporte des informations pour la communauté de l'Internet. Il ne spécifie aucune sorte de norme de l'Internet. La distribution du présent mémoire n'est soumise à aucune restriction.

Résumé

Ce document définit une architecture pour la mise en œuvre d'une différenciation de services échelonnable dans l'Internet. Cette architecture réalise l'échelonnabilité en agrégeant des états de classification du trafic qui sont portés au moyen du marquage de paquets de couche IP en utilisant le champ DS [DSFIELD]. Les paquets sont classés et marqués pour recevoir un comportement de transmission par bond particulier sur les nœuds le long de leur chemin. Les opérations sophistiquées de classification, de marquage, de régulation et de mise en forme n'ont besoin d'être mises en œuvre qu'aux frontières de réseau ou aux hôtes. Les ressources réseau sont allouées aux flux de trafic par les politiques d'approvisionnement de service qui décident comment le trafic est marqué et conditionné à l'entrée dans un réseau capable de différenciation de services, et comment ce trafic est transmis au sein de ce réseau. Une large variété de services peut être mise en œuvre par dessus ces blocs de construction.

Table des Matières

1. Introduction

1.1 Généralités

Le présent document définit une architecture pour la mise en œuvre d'une différentiation de services échelonnable dans l'Internet. Un "service" définit des caractéristiques significatives de la transmission de paquet dans une direction à travers un ensemble de un ou plusieurs chemins au sein d'un réseau. Ces caractéristiques peuvent être spécifiées en termes quantitatifs ou statistiques de débit, de délai, de gigue, et/ou de perte, ou peuvent être autrement spécifiés en termes de priorité relative d'accès aux ressources du réseau. La différentiation de service est destinée à s'accommoder d'exigences d'applications et d'attentes des utilisateurs hétérogènes, et à permettre une appréciation différenciée du service de l'Internet.

Cette architecture se compose d'un certain nombre d'éléments fonctionnels mis en œuvre dans les nœuds de réseau, y compris un petit nombre de comportements de transmission par bonds, de fonctions de classement de paquets, et de fonctions de conditionnement du trafic, qui incluent la mesure, le marquage, le formatage, et la régulation. Cette architecture réalise l'échelonnabilité en mettant en œuvre des fonctions complexes de classification et de conditionnement dans les seuls nœuds frontière du réseau, et en appliquant des comportements par bond aux agrégats de trafic qui ont été marqués de façon appropriée en utilisant le champ DS dans les en-têtes IPv4 ou IPv6 [DSFIELD]. Les comportements par bond sont définis de façon à permettre une granularité raisonnable des moyens d'allocation des ressources de mémoire tampon et de bande passante à chaque nœud entre les flux de trafic en concurrence. L'état de transmission par flux d'application ou par utilisateur n'a pas besoin d'être entretenu au sein du cœur du réseau. Une distinction est faite entre :

- o le service fourni à un agrégat de trafic,
- o les fonctions de conditionnement et les comportements par bond utilisés pour réaliser les services,
- o la valeur du champ DS (codet DS) utilisé pour marquer les paquets pour choisir un comportement par bond, et
- o le mécanisme particulier de mise en œuvre du nœud qui réalise un comportement par bond.

Les politiques de provisionnement de service et de conditionnement du trafic sont suffisamment découplées des comportement de transmission au sein du réseau pour permettre de mettre en œuvre une large variété de comportements de service, avec de la place pour une future expansion.

Cette architecture ne traite de la différentiation de service que dans une seule direction du flux de trafic et elle est donc asymétrique. Le développement d'une architecture symétrique complémentaire fait l'objet de recherches actuelles mais sort du domaine d'application du présent document ; voir par exemple [EXPLICIT].

Le paragraphe 1.2 est un glossaire des termes utilisés au sein du présent document. Le paragraphe 1.3 fait la liste des exigences visées par cette architecture, et le paragraphe 1.4 donne une brève comparaison avec les autres approches de la différentiation de service. La Section 2 expose en détails les composants de l'architecture. La Section 3 propose des lignes directrices pour les spécifications de comportement par bond. La Section 4 expose les questions d'interopérabilité avec les nœud et réseaux qui ne mettent pas en œuvre les services différenciés tels que définis dans ce document et dans [DSFIELD]. La Section 5 expose les problèmes de la livraison de service de diffusion groupé. La Section 6 examine les considérations de sécurité et des tunnels.

1.2 Terminologie

Cette section fait un survol conceptuel général des termes utilisés dans ce document. Certains de ces termes sont définis plus précisément dans les sections suivantes.

Agrégat de comportement (BA, *Behavior Aggregate*)
C'est un agrégat de comportements DS.

Classeur BA
C'est un classeur qui choisit les paquets sur la seule base du contenu du champ DS.

Liaison frontière
C'est une liaison qui connecte les nœuds bordure de deux domaines.

Classeur
C'est une entité qui choisit les paquets sur la base du contenu des en-têtes de paquet selon des règles définies.

Agrégat de comportement DS
C"est une collection de paquets qui ont le même codet DS en traversant une liaison dans une direction particulière.

Nœud frontière DS
C'est un nœud DS qui connecte un domaine DS à un nœud dans un autre domaine DS ou dans un domaine sans capacité DS.

À capacité DS
Capable de mettre en œuvre des services différenciés tels que décrits dans cette architecture ; normalement utilisé en référence à un domaine consistant en nœuds conformes à DS.

Codet DS
C'est une valeur spécifique de la portion DSCP du champ DS, utilisé pour choisir un PHB.

Conforme à DS
Qui a la capacité de prendre en charge des fonctions et comportements de services différenciés tels que définis dans [DSFIELD], dans le présent document, et dans les autres documents de services différenciés ; utilisé normalement en référence à un nœud ou à un appareil.
Domaine DS
C'est un domaine à capacité DS ; un ensemble contigu de nœuds qui fonctionnent avec un ensemble commun de politiques d'approvisionnement de service et de définitions de PHB.

Nœud de sortie DS
C'est un nœud frontière DS dans son rôle de traitement du trafic qui quitte un domaine DS.

Nœud d'entrée DS
C'est un nœud frontière DS dans son rôle de traitement du trafic qui entre dans un domaine DS.

Nœud intérieur DS
C'est un nœud DS qui n'est pas un nœud frontière DS.

Champ DS
C'est l'octet TOS *(type de service)* de l'en-tête IPv4 ou l'octet Classe de trafic IPv6 lorsque il est interprété conformément à la définition donnée dans [DSFIELD]. Les bits du champ DSCP codent le codet DS, tandis que les bits restants sont actuellement non utilisés.

Nœud DS
C'est un nœud conforme à DS.

Région DS
C'est un ensemble de domaines DS contigus qui peuvent offrir des services différenciés sur des chemins qui traversent ces domaines DS.

Domaine DS aval
C'est le domaine DS vers l'aval du flux de trafic sur une liaison frontière.

Élimineur
Appareil qui effectue l'élimination.

Élimination
C'est le processus d'élimination des paquets sur la base de règles spécifiées ; régulation.

Nœud traditionnel
C'est un nœud qui met en œuvre la préséance IPv4 telle que définie dans les [RFC791], [RFC1812] mais qui par ailleurs n'est pas conforme à DS.

Marqueur
C'est un appareil qui effectue le marquage.

Marquage
C'est le processus d'établissement du codet DS dans un paquet sur la base de règles définies ; prémarquage, re-marquage.

Mécanisme
C'est un algorithme ou opération spécifique (par exemple, discipline de mise en file d'attente) qui est mise en œuvre dans un nœud pour réaliser un ensemble d'un ou plusieurs comportements par bond.

Mesureur
Appareil qui effectue des mesures.

Mesurage
C'est le processus de mesure des propriétés temporelles (par exemple, le débit) d'un flux de trafic choisi par un classeur. L'état instantané de ce processus peut être utilisé pour affecter le fonctionnement d'un marqueur, d'un formateur, ou d'un élimineur, et/ou peut être utilisé pour des besoins de comptabilité et de mesure.

Microflux
C'est une seule instance d'un flux de paquets d'application à application qui est identifié par une adresse de source, un accès de source, une adresse de destination, un accès de destination et un identifiant de protocole.

Classeur MF
C'est un classeur multi-champ (MF, *Multi-Field*) qui choisit des paquets sur la base du contenu d'un certain nombre arbitraire de champs d'en-tête, normalement une combinaison d'adresse de source, d'adresse de destination, de champ DS, d'identifiant de protocole, d'accès de source et de destination.

Comportement par bond (PHB, *Per-Hop-Behavior*)
C'est le comportement de transmission observable de l'extérieur appliqué sur un nœud conforme à DS à un agrégat de comportement DS.

Groupe de PHB
C'est un ensemble de un ou plusieurs PHB qui ne peuvent être significativement spécifiés et mis en œuvre que simultanément, du fait d'une contrainte commune qui s'applique à tous les PHB de l'ensemble comme une politique de traitement de mise en file d'attente ou de gestion de file d'attente. Un groupe de PHB fournit un bloc de construction de service qui permet à un ensemble de comportements de transmission qui s'y rapportent d'être spécifiés ensemble (par exemple, les quatre priorités d'abandon). Un seul PHB est un cas particulier de groupe de PHB.

Régulation
C'est le processus d'élimination de paquets (par un élimineur) au sein d'un flux de trafic conformément à l'état d'un mesureur correspondant qui met en application un profil de trafic.

Prémarque
Elle établit le codet DS d'un paquet avant qu'il entre dans un domaine DS situé en aval.

Fournisseur de domaine DS
C'est le fournisseur de services à capacité DS à un domaine source.

Re-marquage
Il consiste à changer le codet DS d'un paquet, normalement effectué par un marqueur conformément à un TCA.

Service
C'est le traitement global pour un sous ensemble défini du trafic d'un consommateur au sein d'un domaine DS ou de bout en bout.

Accord de niveau de service (SLA, *Service Level Agreement*)
C'est un contrat de service entre un consommateur et un fournisseur de service qui spécifie le service de transmission d'un consommateur devrait recevoir. Un consommateur peut être une organisation utilisatrice (domaine de source) ou un autre domaine DS (domaine en amont). Un SLA peut inclure des règles de conditionnement de trafic qui constituent en tout ou partie un TCA.

Politique d'approvisionnement de service
C'est une politique qui définit comment sont configurés les conditionneurs de trafic aux nœuds frontière DS et comment les flux de trafic sont transposés en agrégats de comportement DS pour réaliser une gamme de services.

Formateur
C'est un appareil qui effectue le formatage.

Formatage
C'est le processus qui consiste à retarder des paquets au sein d'un flux de trafic pour qu'il se conforme à un profil de trafic défini.

Domaine source

C'est un domaine qui contient le ou les nœuds qui génèrent le trafic qui reçoit un service particulier.

Conditionneur de trafic
C'est une entité qui effectue des fonctions de conditionnement de trafic et qui peut contenir des mesureurs, des marqueurs, des élimineurs et des formateurs. Les conditionneurs de trafic sont normalement déployés seulement sur une frontière DS. Un conditionneur de trafic peut remarquer un flux de trafic ou peut éliminer ou formater des paquets pour altérer les caractéristiques temporelles du flux et l'amener à être conforme à un profil de trafic.

Conditionnement de trafic
Ce sont les fonctions de contrôle effectuées pour mettre en application les règles spécifiées dans un TCA, qui incluent le mesurage, le marquage, le formatage et la régulation.

Accord de conditionnement de trafic (TCA, *Traffic Conditioning Agreement*)
C'est un accord qui spécifie les règles de classement et tous les profils et les règles correspondantes de mesurage, marquage, élimination et/ou formatage qui sont à appliquer aux flux de trafic choisis par le classeur. Un TCA renferme toutes les règles de conditionnement du trafic explicitement spécifiées dans un SLA ainsi que toutes les règles implicites découlant des exigences de service pertinentes et/ou d'une politique d'approvisionnement de service d'un domaine DS.

Profil de trafic
C'est une description des propriétés temporelles d'un flux de trafic telles que son débit et sa taille de salve.

Flux de trafic
C'est un ensemble administrativement significatif d'un ou plusieurs microflux qui traversent un segment de chemin. Un flux de trafic peut consister en un ensemble de microflux actifs qui sont choisis par un classeur particulier.

Domaine DS amont
C'est le domaine DS en amont du flux de trafic sur une liaison frontière.

1.3 Exigences

L'histoire de l'Internet est celle d'une croissance continue du nombre des hôtes, du nombre et de la variété des applications, et des capacités de l'infrastructure du réseau, et il est prévu que cette croissance continuera dans l'avenir prévisible. Une architecture échelonnable pour la différenciation de service doit être capable de s'accommoder de la poursuite de cette croissance.

Les exigences suivantes ont été identifiées et sont traitées dans cette architecture :

- elle devrait s'accommoder d'une grande diversité de services et de politiques d'approvisionnement, s'étendant de bout en bout ou au sein de réseaux (ou ensemble de réseaux) particuliers,

- elle devrait permettre de découpler le service de l'application particulière utilisée,

- elle devrait fonctionner avec les applications existantes sans qu'il soit besoin de changer l'interface de programmation d'application ou de modifier des logiciels d'hôte (en supposant un déploiement convenable des classeurs, marqueurs, et autres fonctions de conditionnement du trafic),

- elle devrait découpler les fonctions de conditionnement du trafic et d'approvisionnement de service des comportements de transmission mis en œuvre au sein des nœuds du cœur de réseau,

- elle ne devrait pas dépendre de la signalisation d'application bond par bond,

- elle ne devrait exiger qu'un petit ensemble de comportements de transmission dont la complexité de mise en œuvre ne surcharge pas le coût des appareils du réseau, et n'introduise pas de goulets d'étranglement pour les futures mises en œuvre de systèmes à grande vitesse,

- elle devrait éviter les états par microflux ou par usager au sein des nœuds de cœur de réseau,

- elle devrait utiliser seulement des états de classement agrégés au sein du cœur de réseau,

- elle devrait permettre une mise en œuvre simple de classement de paquet dans les nœuds de cœur de réseau (classeur BA),

- elle devrait permettre une interopérabilité raisonnable avec les nœuds de réseau non conformes à DS,

- elle devrait s'accommoder d'un développement incrémentaire.

1.4 Comparaisons avec les autres approches

L'architecture de services différenciés spécifiée dans le présent document peut être comparée avec d'autres modèles existants de différenciation de service. Nous classons ces autres modèles dans les catégories suivantes : marquage de priorité relative, marquage de service, commutation d'étiquettes, intégration de services/RSVP, et statique par bond.

Les exemples de marquage de priorité relative sont le marquage de préséance IPv4 défini dans la [RFC791], la priorité d'anneau à jetons 802.5 [TR], et l'interprétation par défaut des classes de trafic de 802.1p [802.1p]. Dans ce modèle, l'application, l'hôte, ou le nœud mandataire, choisit une priorité relative ou "préséance" pour un paquet (par exemple, la priorité de délai ou d'élimination) et les nœuds du réseau le long du chemin de transit appliquent le comportement de priorité de transmission approprié correspondant à la valeur de priorité contenue dans l'en-tête du paquet. Notre architecture peut être considérée comme une amélioration de ce modèle, dans la mesure où nous spécifions plus clairement le rôle et l'importance des nœuds frontière et des conditionneurs de trafic, et où notre modèle de comportement par bond permet des comportements de transmission plus généraux que la priorité de délai ou d'élimination relative.

Un exemple du modèle de marquage de service est le TOS IPv4 tel que défini dans la [RFC1349]. Dans ce modèle, chaque paquet est marqué avec une demande d'un "type de service", qui peut inclure de "minimiser le délai", "maximiser le débit", "maximiser la fiabilité", ou de "minimiser le coût". Les nœuds du réseau peuvent choisir les chemins ou les comportements de transmission qui sont constitués de façon convenable pour satisfaire à la demande de service. Ce modèle est légèrement différent de notre architecture. Noter que nous ne décrivons pas l'utilisation du champ DS comme élément du choix du chemin. Les marquages de TOS définis dans la [RFC1349] sont très génériques et ne s'étendent pas sur toute la gamme possible de la sémantique des services. De plus, la demande de service est associée à chaque paquet individuel, tandis que certaines sémantiques de service peuvent dépendre du comportement de transmission agrégé d'une séquence de paquets. Le modèle de marquage de service ne s'accommode pas facilement de la croissance du nombre et de la portée des futurs services (car l'espace des codets est restreint) et implique la configuration d'associations "TOS – comportement de transmission" dans chaque nœud de cœur de réseau. La normalisation des marquages de service implique de normaliser l'offre de service, ce qui ne relève pas des compétences de l'IETF. Noter que des dispositions sont prises, dans l'allocation de l'espace des codets DS, pour permettre que des codets à signification locale puissent être utilisés par un fournisseur pour prendre en charge la sémantique de marquage de service [DSFIELD].

Les exemples du modèle de commutation d'étiquette (ou de circuit virtuel) incluent le relais de trame, l'ATM, et MPLS [FRELAY], [ATM]. Dans ce modèle l'état de transmission du chemin et la gestion du trafic ou l'état de la QS est établi pour les flux de trafic st chaque bond le long d'un chemin du réseau. Les agrégats de trafic de diverses granularité sont associés à un chemin à commutation d'étiquette à un nœud d'entrée, et les paquets/cellules au sein de chaque chemin à commutation d'étiquette sont marqués avec une étiquette de transmission qui est utilisée pour rechercher le nœud du prochain bond, le comportement de transmission par bond, et l'étiquette de remplacement à chaque bond. Ce modèle permet une granularité plus fine de l'allocation des ressources aux flux de trafic, car les valeurs d'étiquettes n'ont pas de signification globale, mais ne sont significatives que sur une seule liaison ; donc, les ressources peuvent être réservées pour les agrégats de paquets/cellules reçus sur une liaison avec une étiquette particulière, et la sémantique de la commutation d'étiquettes gouverne le choix du prochain bond, ce qui permet à un flux de trafic de suivre un chemin spécialement aménagé à travers le réseau. Cette granularité améliorée permet, au prix d'exigences supplémentaires de gestion et de configuration d'établir et de maintenir les chemins de commutation d'étiquettes. De plus, la quantité d'état de transmission entretenue à chaque nœud s'échelonne proportionnellement au nombre de nœuds bordure du réseau dans le meilleur cas (en supposant des chemins de commutation d'étiquette de multipoint à point) et elle s'échelonne proportionnellement au carré du nombre de nœuds bordure dans le pire des cas, lorsque les chemins de commutation d'étiquette de bord à bord sont employés avec les ressources provisionnées.

Le modèle de service intégré/RSVP s'appuie sur la transmission traditionnelle des datagrammes dans le cas par défaut, mais permet aux sources et aux receveurs d'échanger des messages de signalisation qui établissent un classement supplémentaire des paquets et un état de transmission sur chaque nœud le long du chemin entre eux [RFC1633], [RSVP]. En l'absence d'agrégation d'état, la quantité d'état sur chaque nœud s'échelonne proportionnellement au nombre de réservations concurrentes, qui peut éventuellement être grand sur les liaisons à grande vitesse. Ce modèle exige aussi la prise en charge par les applications du protocole de signalisation RSVP. Les mécanismes de services différenciés peuvent être utilisé pour agréger l'état de services intégrés/RSVP dans le cœur du réseau [Bernet].

Une variante du modèle de services intégrés/RSVP élimine l'exigence de la signalisation bond par bond en utilisant

seulement le classement et les politiques de transmission "statiques" qui sont mises en œuvre dans chaque nœud le long d'un chemin du réseau. Ces politiques sont mises à jour selon des modalités administratives et non pas en réponse au mélange instantané des microflux actifs dans le réseau. Les exigences d'état pour cette variante sont potentiellement pires que celles qui se rencontrent avec l'utilisation de RSVP, en particulier dans les nœuds de cœur de réseau, car le nombre de politiques statiques qui peuvent être applicables à un nœud au fil du temps peut être supérieur au nombre de sessions d'envoyeurs/receveurs actives qui peuvent avoir installé des états de réservation sur un nœud. Bien que la prise en charge de grands nombres de règles de classeurs et de politiques de transmission puisse être calculable, la charge de gestion associée à l'installation et à la maintenance de ces règles sur chaque nœud au sein d'un cœur de réseau qui va être traversé par un flux de trafic est substantiel.

Bien qu'on différencie notre architecture de ces autres modèles de différenciation de service, on notera que les liaisons et les nœuds qui emploient ces techniques peuvent être utilisés pour étendre les comportements et sémantiques de services différenciés à travers une infrastructure commutée de couche 2 (par exemple, de LAN 802.1p, de cœurs de réseau de relais de trame/ATM) qui interconnecte des nœuds DS, et dans le cas de MPLS, elle peut être utilisée comme technologie de mise en œuvre intra domaine de remplacement. Les contraintes imposées par l'utilisation d'une technologie spécifique de couche liaison dans des régions particulières d'un domaine DS (ou dans un réseau qui donne accès à des domaines DS) peut impliquer la différenciation du trafic sur une granularité plus grossière. Selon la transposition des PHB sur différents services de couche liaison et de la façon dont les paquets sont programmés sur un ensemble restreint de classes de priorités (ou de circuits virtuels de différentes catégories et capacités) tous les PHB (ou sous ensemble des PHB) utilisés peuvent être pris en charge (ou être indistinguables).

2. Modèle architectural des services différenciés

L'architecture de services différenciés se fonde sur un modèle simple dans lequel le trafic entrant dans un réseau est classé et éventuellement conditionné à la frontière du réseau, et alloué à différents agrégats de comportement. Chaque agrégat de comportement est identifié par un seul codet DS. Au sein du cœur de réseau, les paquets sont transmis conformément au comportement par bond associé au codet DS. Dans cette section, on exposera les composants clés au sein d'une région de services différenciés, les fonctions de classement et de conditionnement du trafic, et comment les services différenciés sont réalisés à travers la combinaison du conditionnement du trafic et de la transmission fondée sur le PHB.

2.1 Domaine des services différenciés

Un domaine DS est un ensemble contigu de nœuds DS qui fonctionnent avec une politique d'approvisionnement de service commune et un ensemble de groupes de PHB communs qui sont mis en œuvre sur chaque nœud. Un domaine DS a une frontière bien définie qui consiste en nœuds frontière DS qui classent et éventuellement conditionnent le trafic d'entrée pour s'assurer que les paquets qui transitent par le domaine sont marqués de façon appropriée pour choisir un PHB parmi les groupes de PHB pris en charge au sein du domaine. Les nœuds à l'intérieur du domaine DS choisissent le comportement de transmission pour les paquets sur la base de leur codet DS, en transposant cette valeur en un des PHB pris en charge en utilisant la transposition recommandée de codet en PHB ou une transposition personnalisée localement [DSFIELD]. L'inclusion de nœuds non conformes à DS au sein d'un domaine DS peut résulter en des performance imprévisibles et peut entraver la capacité à satisfaire aux accords de niveau de service (SLA, *service level agreement*).

Un domaine DS consiste normalement en un ou plusieurs réseaux sous la même administration ; par exemple, l'intranet d'une organisation ou d'un FAI. L'administration du domaine est chargée de s'assurer que des ressources adéquates sont provisionnées et/ou réservée pour prendre en charge les SLA offerts par le domaine.

2.1.1 Nœuds frontière et nœuds intérieurs de services différenciés

Un domaine DS comporte des nœuds DS frontières et des nœuds DS intérieurs. Les nœuds DS frontières interconnectent le domaine DS avec les autres domaines DS ou les domaines sans capacité DS, tandis que les nœuds DS intérieurs connectent seulement avec les autres nœuds DS intérieurs ou les nœuds frontières au sein du même domaine DS.

Les nœuds DS frontières et les nœuds intérieurs doivent tous deux être capables d'appliquer le PHB approprié aux paquets sur la base du codet DS ; un comportement imprévisible peut autrement en résulter. De plus, il peut être demandé aux nœuds DS frontières d'effectuer des fonctions de conditionnement du trafic comme défini par l'accord de conditionnement de trafic (TCA, *conditionnement de trafic agreement*) entre leur domaine DS et le domaine homologue auquel ils se connectent (voir au paragraphe 2.3.3).

Les nœuds intérieurs peuvent être capables d'effectuer des fonctions limitées de conditionnement de trafic telles que le re-marquage des codets DS. Les nœuds intérieurs qui mettent en œuvre des fonctions plus complexes de classement et conditionnement de trafic sont analogues aux nœuds DS frontières (voir au paragraphe 2.3.4.4).

Un hôte qui se trouve dans un réseau qui contient un domaine DS peut agir comme un nœud DS frontière pour du trafic provenant d'applications qui fonctionnent sur cet hôte ; on dit donc que l'hôte est dans le domaine DS. Si un hôte n'agit pas comme un nœud frontière, le nœud DS topologiquement le plus proche de cet hôte agit alors comme nœud DS frontière pour le trafic de cet hôte.

2.1.2 Nœuds d'entrée et nœuds de sortie de services différenciés

Les nœuds DS frontières agissent à la fois comme nœud d'entrée DS et comme nœud de sortie DS pour les différentes directions de trafic. Le trafic entre dans un domaine DS à un nœud d'entrée DS et quitte un domaine DS à un nœud de sortie DS. Un nœud d'entrée DS est chargé de s'assurer que le trafic qui entre dans le domaine DS se conforme à tout TCA entre lui et l'autre domaine auquel le nœud d'entrée est connecté. Un nœud de sortie DS peut effectuer des fonctions de conditionnement de trafic sur le trafic transmis à un domaine homologue directement connecté, selon les détails du TCA entre les deux domaines. Noter qu'un nœud DS frontière peut agir comme nœud DS intérieur pour un certain ensemble d'interfaces.

2.2 Région de services différenciés

Une région de services différenciés (région DS) est un ensemble d'un ou plusieurs domaines DS contigus. Les régions DS sont capables de prendre en charge les services différenciés le long des chemins qui parcourent le domaine au sein de la région.

Les domaines DS dans une région DS peuvent prendre en charge différents groupes de PHB en interne et différentes transpositions de codet en PHB. Cependant, pour permettre des services qui s'étendent à travers les domaines, les domaines DS homologues doivent chacun établir un SLA d'homologues qui définit (explicitement ou implicitement) un TCA qui spécifie comment est conditionné le trafic de transit d'un domaine DS à l'autre à la frontière entre les deux domaines DS.

Il est possible que plusieurs domaines DS au sein d'une région DS adoptent une politique d'approvisionnement de services commune et qu'ils puissent prendre en charge un ensemble commun de groupes de PHB et de transposition de codets, éliminant ainsi le besoin de conditionnement de trafic entre ces domaines DS.

2.3 Classification et conditionnement du trafic

Les services différenciés sont étendus à travers une frontière de domaine DS en établissant un SLA entre un réseau amont et un domaine DS aval. Le SLA peut spécifier une classification de paquets et des règles de re-marquage, et peut aussi spécifier des profils de trafic et des actions pour les flux de trafic qui sont dans ou hors du profil (voir au paragraphe 2.3.2). Le TCA entre les domaines est déduit (explicitement ou implicitement) de ce SLA.

La politique de classification des paquets identifie le sous ensemble de trafic qui peut recevoir un service différencié en étant conditionné et/ou transposé sur un ou plusieurs agrégats de comportement (par un re-marquage du codet DS) au sein du domaine DS.

Le conditionnement du trafic effectue le mesurage, le formatage, la régulation et/ou le re-marquage pour assurer que le trafic entrant dans le domaine DS se conforme aux règles spécifiées dans le TCA, conformément à la politique d'approvisionnement de services du domaine. L'étendue du conditionnement de trafic requis dépend des spécificités de l'offre de service, et peut aller du simple re-marquage de codet à des opérations complexes de régulation et de formatage. Les détails des politiques de conditionnement de trafic qui sont négociées entre les réseaux sortent du domaine d'application du présent document.

2.3.1 Classeur

Les classeurs de paquets choisissent les paquets dans le flux de trafic sur la base du contenu de certaines portions de l'en-tête du paquet. On définit deux types de classeurs. Le classeur BA (Behavior Aggregate, *agrégat de comportement*) classe les paquets sur la seule base du codet DS. Le classeur MF (Multi-Field, *multi champs*) choisit les paquets sur la base de la valeur d'une combinaison d'un ou plusieurs champs d'en-tête, tels que l'adresse de source, l'adresse de destination, le champ

DS, l'identifiant de protocole, les numéros d'accès de source et de destination, et d'autres informations telles que l'interface entrante.

Les classeurs sont utilisés pour "piloter" les paquets qui correspondent à une règle spécifiée sur un élément d'un conditionneur de trafic pour un traitement ultérieur. Les classeurs doivent être configurés par une procédure de gestion conforme au TCA approprié.

Le classeur devrait authentifier les informations qu'il utilise pour classer le paquet (voir la Section 6).

Noter que dans le cas d'une fragmentation de paquet en amont, les classeurs MF qui examinent le contenu des champs d'en-tête de couche transport peuvent classer incorrectement les fragments de paquet qui suivent le premier fragment. Une solution possible à ce problème est de conserver l'état de fragmentation ; cependant, ceci n'est pas une solution générale à cause de la possibilité d'un réarrangement de fragments en amont ou de l'acheminement par des chemins divergents. La politique à appliquer aux fragments de paquets sort du domaine d'application du présent document.

2.3.2 Profils de trafic

Un profil de trafic spécifie les propriétés temporelles d'un flux de trafic choisi par un classeur. Il fournit des règles pour déterminer si un paquet particulier est dans le profil ou hors du profil. Par exemple, un profil fondé sur un baquet de jetons peut ressembler à :

codet=X, utilise le baquet de jetons r, b

Le profil ci-dessus indique que tous les paquets marqués avec le codet DS X devraient être mesurés par rapport à un mesureur de baquet de jetons au débit r et la taille de salve b. Dans cet exemple, les paquets hors profil sont ceux qui dans le flux de trafic arrivent quand un nombre insuffisant de jetons est disponible dans le baquet. Le concept de "dans" et "hors profil" peut être étendu à plus de deux niveaux, par exemple, plusieurs niveaux de conformité à un profil peuvent être définis et mis en application.

Différentes actions de conditionnement peuvent être appliquées aux paquets dans le profil et aux paquets hors profil, ou différentes actions de comptabilité peuvent être déclanchées. Les paquets dans le profil peuvent être autorisés à entrer dans le domaine DS sans autre conditionnement, ou autrement, leur codet DS peut être changé. Ce dernier cas survient lorsque le codet DS est réglé à une valeur non par défaut pour la première fois [DSFIELD], ou quand les paquets entrent dans un domaine DS qui utilise un groupe de PHB différent ou une politique de transposition de codet en PHB différente pour ce flux de trafic. Les paquets hors profil peuvent être mis en file d'attente jusqu'à ce qu'ils soient dans le profil (formatés), éliminés (régulés), marqués avec un nouveau codet (re-marqués), ou transmis inchangés tout en déclanchant une procédure comptable. Les paquets hors profil peuvent être transposés en un ou plusieurs agrégats de comportement qui sont "inférieurs" dans certaines dimensions des performances de transmission à l'agrégat de comportements dans lequel les paquets dans le profil sont transposés.

Noter qu'un profil de trafic est un composant facultatif d'un TCA et son utilisation dépend des spécificités de l'offre de services et de la politique d'approvisionnement de services du domaine.

2.3.3 Conditionneurs de trafic

Un conditionneur de trafic peut contenir les éléments suivants : mesureur, marqueur, formateur, et élimineur. Un flux de trafic est choisi par un classeur, qui dirige les paquets vers une instance logique d'un conditionneur de trafic. Un mesureur est utilisé (lorsque c'est approprié) pour mesurer le flux de trafic par rapport à un profil de trafic. L'état du mesureur par rapport à un paquet particulier (par exemple, si il est dans ou hors profil) peut être utilisé pour affecter une action de marquage, d'élimination ou de formatage.

Lorsque les paquets sortent du conditionneur de trafic d'un nœud DS frontière, le codet DS de chaque paquet doit être réglé à une valeur appropriée.

La Figure 1 montre le diagramme d'état d'un classeur et d'un conditionneur de trafic. Noter qu'un conditionneur de trafic peut ne pas nécessairement contenir les quatre éléments. Par exemple, dans le cas où aucun profil de trafic n'est en œuvre, les paquets peuvent seulement passer à travers un classeur et un marqueur.

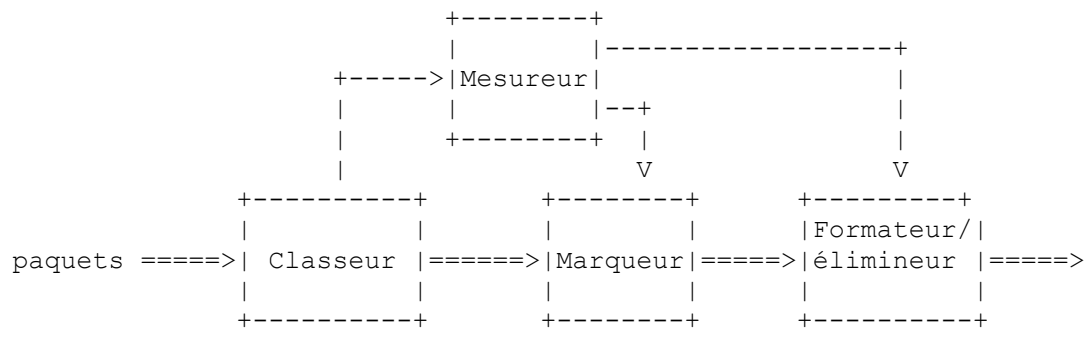

Figure 1 : Vue logique d'un classeur de paquet et d'un conditionneur de trafic

2.3.3.1 Mesureurs

Les mesureurs de trafic mesurent les propriétés temporelles du flux de paquets choisi par un classeur par rapport à un profil de trafic spécifié dans un TCA. Un mesureur passe les informations d'état aux autres fonctions de conditionnement pour déclancher une action particulière pour chaque paquet qui est soit dans, soit hors profil (dans une certaine mesure).

2.3.3.2 Marqueurs

Les marqueurs de paquet règlent le champ DS d'un paquet à un codet particulier, en ajoutant le paquet marqué à un agrégat de comportement DS particulier. Le marqueur peut être configuré pour marquer tous les paquets qui sont dirigés sur lui avec un seul codet, ou il peut être configuré pour marquer un paquet avec un ensemble de codets utilisés pour choisir un PHB dans un groupe de PHB, conformément à l'état d'un mesureur. Lorsque le marqueur change le codet dans un paquet, il est dit avoir "re-marqué" le paquet.

2.3.3.3 Formateurs

Les formateurs retardent certains paquets, ou tous, dans un flux de trafic afin d'amener le flux à se conformer à un profil de trafic. Un formateur a normalement une mémoire tampon de dimension finie, et les paquets peuvent être éliminés si il n'y a plus d'espace de mémoire tampon suffisant pour contenir les paquets retardés.

2.3.3.4 Élimineurs

Les élimineurs éliminent certains paquets, ou tous, dans un flux de trafic afin d'amener le flux à se conformer à un profil de trafic. Ce processus est appelé "réguler" le flux. Noter qu'un élimineur peut être mis en œuvre comme cas particulier de formateur en réglant la taille de la mémoire tampon du formateur à zéro (ou quelques) paquets.

2.3.4 Localisation des conditionneurs de trafic et des classeurs MF

Les conditionneurs de trafic sont normalement situés au sein des nœuds d'entrée et de sortie de frontière DS, mais ils peuvent aussi être situés dans des nœuds à l'intérieur d'un domaine DS, ou au sein d'un domaine sans capacité DS.

2.3.4.1 Au sein du domaine source

On définit le domaine source comme le domaine qui contient le ou les nœuds qui génèrent le trafic qui reçoit un service particulier. Les nœuds de source de trafic et les nœuds intermédiaires au sein d'un domaine source peuvent effectuer les fonctions de classification et de conditionnement. Le trafic généré du domaine source à travers une frontière peut être marqué par les sources de trafic directement ou par des nœuds intermédiaires avant de quitter le domaine source. C'est ce qu'on appelle le marquage initial ou "pré-marquage".

Considérons l'exemple d'une compagnie qui a comme politique que les paquets de son dirigeant devraient avoir une priorité supérieure. L'hôte du dirigeant peut marquer le champ DS de tous les paquets sortants avec un codet DS qui indique "priorité supérieure". Autrement, le routeur du premier bond directement connecté à l'hôte du dirigeant peut classer le trafic et marquer les paquets du dirigeant avec le codet DS correct. Un tel trafic de priorité élevée peut aussi être conditionné près de la source afin qu'il y ait une limite à la quantité de trafic à priorité élevée transmis à partir d'une source particulière.

Il y a certains avantages à marquer les paquets près de la source du trafic. D'abord, une source de trafic peut plus facilement prendre en compte les préférences d'une application lorsque elle doit décider quels paquets devraient recevoir un meilleur traitement de transmission. Ensuite, la classification des paquets est beaucoup plus simple avant que le trafic ait été agrégé avec les paquets provenant d'autres sources, car le nombre de règles de classement à appliquer au sein d'un seul nœud est réduit.

Comme le marquage de paquet peut être réparti entre plusieurs nœuds, le domaine DS de source est chargé de s'assurer que le trafic agrégé vers son domaine DS fournisseur se conforme au TCA approprié. Des mécanismes d'allocation

supplémentaires tels que des courtiers en bande passante ou RSVP peuvent être utilisés pour allouer de façon dynamique des ressources pour un agrégat de comportement DS particulier au sein du réseau du fournisseur [2BIT], [Bernet]. Le nœud frontière du domaine source devrait aussi surveiller la conformité au TCA, et peut réguler, formater, ou re-marquer les paquets en tant que de besoin.

2.3.4.2 À la frontière d'un domaine DS

Les flux de trafic peuvent être classés, marqués, et autrement conditionnés à l'une ou l'autre extrémité d'une liaison frontière (le nœud de sortie DS du domaine amont ou le nœud d'entrée DS du domaine aval). Le SLA entre les domaines devrait spécifier quel domaine a la responsabilité de la transposition des flux de trafic en agrégats de comportement DS et de conditionner ces agrégats en conformité avec le TCA approprié. Cependant, un nœud d'entrée DS doit supposer que le trafic entrant peut n'être pas conforme au TCA et doit être prêt à mettre en application le TCA conformément à la politique locale.

Lorsque les paquets sont pré-marqués et conditionnés dans le domaine amont, moins de règles de classification et de conditionnement de trafic ont potentiellement besoin d'être prises en charge dans le domaine DS aval. Dans ces circonstances, le domaine DS aval peut avoir seulement besoin de re-marquer ou réguler les agrégats de comportement entrants pour mettre en application le TCA. Cependant, des services plus sophistiqués qui dépendent du chemin ou de la source peuvent exiger la classification MF dans les nœuds d'entrée du domaine DS aval.

Si un nœud d'entrée DS est connecté à un domaine DS amont sans capacité DS, le nœud d'entrée DS doit être capable d'effectuer toutes les fonctions de conditionnement de trafic nécessaires sur le trafic entrant.

2.3.4.3 Dans les domaines sans capacité DS

Les nœuds de source de trafic ou les nœuds intermédiaires dans un domaine sans capacité DS peuvent employer des conditionneurs de trafic pour pré-marquer le trafic avant qu'il atteigne l'entrée d'un domaine DS aval. De cette façon, les politiques locales de classification et de marquage peuvent être masquées.

2.3.4.4 À l'intérieur des nœuds DS

Bien que l'architecture de base suppose que les fonctions complexes de classification et de conditionnement de trafic ne soient localisées que dans les nœuds frontière d'entrée et de sortie d'un réseau, le déploiement de ces fonctions dans l'intérieur du réseau n'est pas interdit. Par exemple, des politiques d'accès plus restrictives peuvent être mises en application sur une liaison transocéanique, exigeant une fonctionnalité de classification et conditionnement de trafic MF dans le nœud amont sur la liaison. Cette approche peut avoir des limites d'échelle, dues au nombre potentiellement grand de règles de classification et de conditionnement qui devraient être entretenues.

2.4 Comportements par bond

Un comportement par bond (PHB) est une description du comportement de transmission observable de l'extérieur d'un nœud DS appliqué à un agrégat de comportements DS particulier. "Comportement de transmission" est un concept général dans ce contexte. Par exemple, dans le cas où un seul agrégat de comportement occupe une liaison, le comportement de transmission observable (c'est-à-dire, de perte, de délai, de gigue) va souvent dépendre seulement de la charge relative de la liaison (c'est-à-dire, dans le cas où le comportement suppose une discipline de programmation qui conserve le travail). Des distinctions comportementales utiles sont principalement observées lorsque plusieurs agrégats de comportement sont en concurrence pour les ressources de mémoire tampon et de bande passante sur un nœud. Le PHB est le moyen par lequel un nœud alloue des ressources aux agrégats de comportement, et il est sur le dessus de ce mécanisme d'allocation de ressource bond par bond de base sur lequel des services différenciés utiles peuvent être construits.

Le plus simple exemple de PHB est celui qui garantit une allocation minimale de X % d'une liaison sur un intervalle de temps raisonnable) à un agrégat de comportements. Ce PHB peut être très facilement mesuré dans diverses conditions de trafic concurrentes. Un PHB légèrement plus complexe serait la garantie d'une allocation de bande passante minimale de X % d'une liaison, avec un partage équitable proportionnel de tout excédent de capacité de la liaison. En général, le comportement observable d'un PHB peut dépendre de certaines contraintes des caractéristiques de trafic de l'agrégat de comportement associé, ou des caractéristiques des autres agrégats de comportement.

Les PHB peuvent être spécifiés en termes de priorité de ressources (par exemple, mémoire tampon, bande passante) par rapport aux autres PHB, ou en termes de caractéristiques de trafic observables relatives (par exemple, délai, perte). Ces PHB peuvent être utilisés comme les blocs de construction pour allouer des ressources et devraient être spécifiés comme un groupe (groupe de PHB) pour rester cohérent. Les groupes de PHB vont normalement partager une contrainte commune qui s'applique à chaque PHB du groupe, comme une politique de programmation de paquets ou de gestion de mémoire tampon. Les relations entre les PHB dans un groupe peuvent être en termes de priorité absolue ou relative (par exemple, la priorité à l'élimination au moyen de seuils déterministes ou stochastiques) mais ce n'est pas exigé (par exemple, N parts égales de la

liaison). Un seul PHB défini isolément est un cas particulier de groupe de PHB.

Les PHB sont mis en œuvre dans des nœuds au moyen de mécanismes de gestion de mémoire tampon et de programmation des paquets. Les PHB sont définis en termes de caractéristiques de comportement pertinentes pour les politiques d'approvisionnement de service, et non en termes de mécanismes de mise en œuvre particuliers. En général, divers mécanismes de mise en œuvre peuvent convenir pour un groupe de PHB particulier. De plus, il est vraisemblable que plus d'un groupe de PHB peut être mis en œuvre sur un nœud et utilisé au sein d'un domaine. Les groupes de PHB devraient être définis de telle sorte que l'allocation des ressources appropriées entre les groupes puisse en être déduite, et des mécanismes intégrés peuvent être mis en œuvre pour la prise en charge simultanée de deux groupes ou plus. Une définition de groupe de PHB devrait indiquer les conflits possibles avec les groupes de PHB déjà documentés qui pourraient empêcher un fonctionnement simultané.

Comme décrit dans [DSFIELD], un PHB est choisi à un nœud par une transposition du codet DS dans un paquet reçu. Les PHB normalisés ont un codet recommandé. Cependant, l'espace total des codets est plus grand que l'espace disponible pour les codets recommandés pour les PHB normalisés, et [DSFIELD] donne des dispositions pour les transpositions configurables localement. Un tableau de transposition de codet en PHB peut contenir aussi bien des transpositions injectives que surjectives. Tous les codets doivent être transposés en un PHB ; en l'absence de politique locale, les codets qui ne sont pas transposés en un PHB normalisé conformément à la spécification de ce PHB devraient être transposés dans le PHB par défaut.

2.5 Allocation des ressources du réseau

La mise en œuvre, la configuration, le fonctionnement et l'administration des groupes de PHB pris en charge dans les nœuds d'un domaine DS devraient effectivement partager les ressources de ces nœuds et des liaisons inter nœuds entre les agrégats de comportement, conformément à la politique d'approvisionnement de services du domaine. Les conditionneurs de trafic peuvent pousser plus loin le contrôle de l'utilisation de ces ressources à travers la mise en application des TCA et éventuellement des retours opérationnels de la part des nœuds et des conditionneurs de trafic dans le domaine. Bien qu'une gamme de services puisse être déployée en l'absence de fonctions complexes de conditionnement de trafic (par exemple, en utilisant seulement des politiques de marquage statique) les fonctions telles que de régulation, de formatage, et de re-marquage dynamique permettent le déploiement de services qui fournissent des métriques de performances quantitatives.

La configuration et l'interaction entre conditionneurs de trafic et nœuds intérieurs devrait être gérée par le contrôle administratif du domaine et peut exiger un contrôle du fonctionnement à travers les protocoles et une entité de contrôle. Il y a une large gamme de modèles de contrôle possibles.

La nature précise et la mise en œuvre de l'interaction entre ces composants sort du domaine d'application de la présente architecture. Cependant, les besoins de modularité exigent que le contrôle du domaine ne demande pas la gestion détaillée des ressources du réseau. Le modèle de contrôle le plus modulable fonctionnerait sur des nœuds en boucle ouverte dans le cadre temporel de fonctionnement, et n'exigerait qu'une gestion des échelles temporelles administratives avec les divers SLA. Ce modèle simple peut ne pas convenir dans certaines circonstances, et certains contrôles de fonctionnement automatisés à variation lente (en minutes plutôt qu'en secondes) peuvent être souhaitables pour équilibrer l'utilisation du réseau par rapport au dernier profil de charge.

3. Lignes directrices de la spécification du comportement par bond

Les exigences de base de la normalisation du comportement bond par bond sont données dans [DSFIELD]. La présente section développe à partir de ce texte en décrivant des lignes directrices supplémentaires pour les spécifications de (groupes de) PHB. Elles sont destinées à aider à garder la cohérence des mises en œuvre. Avant que soit proposée la normalisation d'un groupe de PHB, il devrait satisfaire à ces lignes directrices, en tant que de besoin, pour préserver l'intégrité de cette architecture.

G.1 : Un PHB standard doit spécifier un codet DS recommandé choisi dans l'espace des codets réservé pour les transpositions standard [DSFIELD]. Les codets recommandés seront alloués par l'IANA. Une proposition de PHB peut recommander un codet temporaire à partir de l'espace EXP/LU pour faciliter l'expérimentation inter domaine. La détermination du PHB d'un paquet ne doit pas exiger l'inspection de champs d'en-tête de paquet supplémentaires au delà du champ DS.

G.2 : La spécification de chaque nouvelle proposition de groupe de PHB devrait inclure une vue générale du comportement et de l'objet du comportement proposé. Cette vue générale devrait inclure une déclaration du ou des problèmes que vise ce groupe de PHB. Elle devrait inclure les concepts de base qui sous-tendent le groupe de PHB.

Ces concepts devraient inclure, sans s'y limiter, le comportement de mise en file d'attente, le comportement d'élimination, et le comportement de choix de la liaison de sortie. Enfin, elle devrait spécifier la méthode par laquelle le groupe de PHB résout le ou les problèmes spécifiés dans la déclaration des problèmes.

G.3 : Une spécification de groupe de PHB devrait indiquer le nombre de PHB individuels spécifiés. Dans le cas où plusieurs PHB sont spécifiés, les interactions entre ces PHB et les contraintes qui doivent être globalement respectées par tous les PHB au sein du groupe devraient être clairement spécifiées. Par exemple, la spécification doit indiquer si la probabilité de réarrangement de paquets au sein d'un microflux est augmentée si différents paquets de ce microflux sont marqués pour différents PHB au sein du groupe.

G.4 : Lorsque le bon fonctionnement d'un groupe de PHB dépend de contraintes telles que de restrictions d'approvisionnement, la définition du PHB devrait alors décrire le comportement lorsque ces contraintes ne sont pas respectées. De plus, si des actions telles que l'élimination ou le re-marquage de paquet sont requises en cas de violation de ces contraintes, ces actions devraient alors être spécifiquement stipulées.

G.5 : Un groupe de PHB peut être spécifié pour une utilisation locale au sein d'un domaine afin de fournir une fonctionnalité ou des services spécifiques de ce domaine. Dans ce cas, la spécification du PHB est utile pour donner aux fabricants une définition cohérente du groupe de PHB. Cependant, un groupe de PHB qui est défini pour une utilisation locale ne devrait pas être pris en considération pour la normalisation, mais il peut être publié dans une RFC d'information. À l'opposé, un groupe de PHB qui est destiné à l'usage général va suivre un processus de normalisation plus strict. Donc, toutes les propositions de PHB devraient spécifiquement établir si elles sont pour utilisation générale ou locale.

Il est vrai que les groupes de PHB peuvent être conçus dans l'intention de fournir des services d'hôte à hôte, de bordure de WAN à bordure de WAN, et/ou de bordure de domaine à bordure de domaine. L'utilisation du terme "de bout en bout" dans une définition de PHB devrait être interprétée comme signifiant "d'hôte à hôte" pour être cohérent.

D'autres groupes de PHB peuvent être définis et déployés localement au sein de domaines, pour des besoins expérimentaux ou opérationnels. Il n'est pas exigé que ces groupes de PHB fassent l'objet d'une documentation publique, mais ils devraient utiliser des codets DS pris dans un des pôles de EXP/LU comme défini dans [DSFIELD].

G.6 : Il est éventuellement possible ou approprié pour un paquet marqué pour un PHB au sein d'un groupe de PHB d'être re-marqué pour choisir un autre PHB au sein du groupe ; au sein d'un domaine ou à travers une frontière de domaine. Il y a normalement trois raisons à une telle modification de PHB :

a. les codets associés au groupe de PHB sont collectivement destinés à porter l'état du réseau,

b. il existe des conditions qui exigent la promotion ou la réduction du PHB d'un paquet (cela suppose que les PHB au sein du groupe puissent être rangés dans un certain ordre),

c. la frontière entre deux domaines n'est pas couverte par un SLA. Dans ce cas, le codet/PHB à choisir lors du franchissement de la liaison frontière sera déterminé par la politique locale du domaine amont.

Une spécification de PHB devrait déclarer clairement les circonstances dans lesquelles des paquets marqués pour un PHB au sein d'un groupe de PHB peuvent, ou devraient être modifiés (par exemple, promus ou réduits) en un autre PHB au sein du groupe. Si il n'est pas souhaitable que le PHB d'un paquet soit modifié, la spécification devrait clairement déclarer les risques encourus lorsque le PHB est modifié. Un risque possible du changement du PHB d'un paquet, aussi bien au sein d'un groupe qu'au dehors, est une plus forte probabilité de réarrangement des paquets au sein d'un microflux. Les PHB au sein d'un groupe peuvent porter une sémantique d'hôte à hôte, de bordure de WAN à bordure de WAN, et/ou de bordure de domaine à bordure de domaine qu'il peut être difficile de répliquer si les paquets sont re-marqués pour choisir un autre PHB à partir du groupe (ou autrement).

Pour certains groupes de PHB, il peut être approprié de refléter un changement d'état dans le nœud en re-marquant les paquets pour spécifier un autre PHB provenant du même groupe. Si un groupe de PHB est conçu pour refléter l'état d'un réseau, la définition du PHB doit décrire de façon adéquate la relation entre les PHB et les états qu'ils reflètent. De plus, si ces PHB limitent de quelque façon les actions de transmission que peut effectuer un nœud, ces contraintes peuvent être spécifiées comme des actions que le nœud devrait, ou doit effectuer.

G.7 : Une spécification de groupe de PHB devrait aussi comporter une section définissant les implications du tunnelage sur l'utilité du groupe de PHB. Cette section devrait spécifier les implications sur l'utilité du groupe de PHB d'un en-tête externe nouvellement créé lorsque le champ DS d'origine de l'en-tête interne est encapsulé dans un tunnel. Cette section devrait aussi exposer les changements possibles qui devraient s'appliquer à l'en-tête interne à la sortie du

tunnel, lorsque les codets provenant de l'en-tête interne et de l'en-tête externe sont tous deux accessibles (voir au paragraphe 6.2).

G.8 : Le processus de spécification des groupes de PHB sera vraisemblablement de nature incrémentaire. Lorsque de nouveaux groupes de PHB sont proposés, leurs interactions connues avec des groupes de PHB spécifiés précédemment devrait être documentée. Lorsque un nouveau groupe de PHB est créé, il peut être entièrement nouveau dans sa portée ou il peut être une extension d'un groupe de PHB existant. Si le groupe de PHB est entièrement indépendant de certaines ou de toutes les spécifications de PHB existantes, une section devrait être incluse dans la spécification du PHB qui précise comment le nouveau groupe de PHB peut coexister avec les groupes de PHB déjà normalisés. Par exemple, cette section pourrait indiquer la possibilité du réarrangement des paquets au sein d'un microflux pour les paquets marqués avec des codets associés à deux groupes de PHB séparés. Si le fonctionnement concurrent de deux groupes (ou plus) de PHB différents dans le même nœud est impossible ou nuisible, cela devrait être déclaré. Si le fonctionnement concurrent de deux groupes (ou plus) de PHB différents exige des comportements spécifiques de la part du nœud lorsque des paquets marqués pour les PHB provenant de ces groupes de PHB différents sont traités en même temps par le nœud, ces comportements devraient être déclarés. Il faut faire attention à éviter des définitions en boucle *("qui se mordent la queue")* de groupes de PHB.

Si le groupe de PHB proposé est une extension d'un groupe de PHB existant, on devrait inclure dans la spécification du groupe de PHB une section qui précise comment cette extension interopère avec le comportement étendu. De plus, si l'extension altère ou définit de quelque façon plus restrictive le comportement existant, cela devrait aussi être clairement indiqué.

G.9 : Chaque spécification de PHB devrait inclure une section spécifiant les exigences minimales de conformité pour les mises en œuvre du groupe de PHB. Cette section de conformité est destinée à donner le moyen de spécifier les détails d'un comportement tout en permettant aux mises en œuvre des variations dans la mesure permise par la spécification de PHB. Cette section de conformité peut prendre la forme de règles, de tableaux, de pseudo-code, ou d'essais.

G.10 : Une spécification de PHB devrait comporter une section détaillant les implications de ce comportement sur la sécurité. Cette section devrait comporter une discussion sur le re-marquage du codet de l'en-tête interne à la sortie d'un tunnel et son effet sur le comportement de transmission désiré. De plus, cette section devrait aussi exposer comment le groupe de PHB proposé pourrait être utilisé dans des attaques de déni de service, des attaques en réduction du contrat de service, et des attaques en violation du contrat de service. Enfin, elle devrait exposer les moyens possibles de détection de telles attaques lorsque elles sont pertinentes pour le comportement proposé.

G.11 : Une spécification de PHB devrait comporter une section qui précise les questions de configuration et de gestion qui peuvent affecter le fonctionnement du PHB et qui peuvent impacter les services candidats à l'utilisation de ce PHB.

G.12 : Il est fortement recommandé que soit fourni un appendice à chaque spécification de PHB, considérant les implications du comportement proposé sur les services actuels et potentiels. Ces services pourraient inclure, sans s'y limiter, les services spécifiques de l'utilisateur, de l'appareil, du domaine, ou de bout en bout. Il est aussi fortement recommandé que l'appendice comporte une section qui décrirait comment les services sont vérifiés par les utilisateurs, les appareils, et/ou les domaines.

G.13 : Il est recommandé que soit fourni un appendice à chaque spécification de PHB qui vise une utilisation locale au sein d'un domaine, fournissant des lignes directrices pour le choix par un PHB des paquets qui sont transmis dans un domaine homologue qui ne prend pas en charge le groupe de PHB.

G.14 : Il est recommandé que soit fourni un appendice à chaque spécification de PHB qui examine l'impact du groupe de PHB proposé sur les protocoles existants de couche supérieure. Dans certaines circonstances, les PHB peuvent permettre des changements éventuels en faveur de protocoles de couche supérieure qui peuvent augmenter ou diminuer l'utilité du groupe de PHB proposé.

G.15 : Il est recommandé que soit fourni un appendice à chaque spécification de PHB qui recommande l'abandon des mécanismes de qualité de service de couche Liaison pour prendre en charge le comportement prévu du PHB à travers un support partagé ou une couche Liaison commutée. La détermination de la transposition la plus appropriée entre un PHB et un mécanisme de qualité de service de couche Liaison dépend de nombreux facteurs et sort du domaine d'application du présent document ; cependant, la spécification devrait essayer d'offrir des lignes directrices.

4. Interopérabilité avec les nœuds non conformes aux services différenciés

On définit un nœud non conforme aux services différenciés (nœud non conforme DS) comme tout nœud qui n'interprète pas le champ DS comme spécifié dans [DSFIELD] et/ou ne met pas en œuvre certains ou tous les PHB normalisés (ou ceux utilisés dans un domaine DS particulier). Cela peut être dû aux capacités ou à la configuration du nœud. On définit un nœud traditionnel comme un cas particulier de nœud non conforme DS qui met en œuvre la classification et la transmission de préséance IPv4 comme défini dans les [RFC791] [RFC1812], mais qui est par ailleurs non conforme à DS. Les valeurs de préséance dans l'octet TOS IPv4 sont compatibles en intention avec les codets de sélecteur de classe définis dans [DSFIELD], et les comportements de transmission de préséance définis dans les [RFC791] [RFC1812] se conforment aux exigences du PHB de sélecteur de classe définis aussi dans [DSFIELD]. Une distinction clé entre un nœud traditionnel et un nœud conforme à DS est que le nœud traditionnel peut ou non interpréter les bits 3 à 6 de l'octet de TOS comme défini dans la [RFC1349] (les bits "DTRC") ; en pratique, il ne va pas interpréter ce bit comme spécifié dans [DSFIELD]. On suppose que l'utilisation des marquages de TOS définis dans la [RFC1349] est déconseillée. Les nœuds qui sont non conformes à DS et qui ne sont pas des nœuds traditionnels affichent des comportements de transmission imprévisibles pour les paquets qui ont des codets DS qui ne sont pas à zéro.

Les services différenciés dépendent des mécanismes d'allocation de ressources fournis par la mise en œuvre du comportement par bond dans les nœuds. La qualité ou le niveau d'assurance statistique d'un service peut être rompu lorsque le trafic transite par un nœud non conforme à DS, ou un domaine sans capacité DS.

On examinera deux cas distincts. Le premier cas concerne l'utilisation des nœuds non conformes à DS au sein d'un domaine DS. Noter que la transmission du PHB est surtout utile pour allouer des ressources rares en nœud et en liaison d'une façon contrôlée. Sur des liaisons à grande vitesse et faible charge, le plus mauvais cas de délai de paquet, de gigue et de perte peut être négligeable, et l'utilisation d'un nœud non conforme à DS sur l'extrémité amont d'une telle liaison peut ne pas entraîner de dégradation du service. Dans des circonstances plus réalistes, l'absence de transmission du PHB dans un nœud peut rendre impossible d'offrir des services de faible délai, de faible perte, ou de bande passante garantie sur des chemins qui traversent le nœud. Cependant, l'utilisation d'un nœud traditionnel peut être une solution de remplacement acceptable, en supposant que le domaine DS se restreint lui-même à la seule utilisation des codets de sélecteur de classe définis dans [DSFIELD], et en supposant que la mise en œuvre particulière de la préséance dans le nœud traditionnel fournit des comportements de transmission qui sont compatibles avec les services offerts le long de chemins qui traversent ce nœud. Noter qu'il est important de restreindre les codets utilisés à ceux de sélecteur de classe, car le nœud traditionnel peut ou non interpréter les bits 3-5 conformément à la [RFC1349], d'où peut résulter des solutions de transmission imprévisibles.

Le second cas concerne le comportement de services qui traversent des domaines sans capacité DS. On suppose pour les besoins de l'exposé qu'un domaine sans capacité DS ne déploie pas de fonction de conditionnement de trafic sur les nœuds frontière du domaine ; donc, même dans le cas où le domaine comporte des nœuds intérieurs traditionnels ou conformes à DS, l'absence de régulation du trafic aux frontières va limiter la capacité à fournir de façon cohérente certains types de services sur le domaine. Un domaine DS et un domaine sans capacité DS peuvent négocier un accord qui règle la façon dont le trafic sortant du domaine DS devrait être marqué avant d'entrer dans le domaine sans capacité DS. Cet accord peut prévoir la surveillance de la conformité par l'échantillonnage du trafic au lieu d'un conditionnement de trafic rigoureux. Autrement, lorsque on sait que le domaine sans capacité DS consiste en nœuds traditionnels, le domaine DS amont peut re-marquer le trafic de services différenciés de façon opportuniste sur un ou plusieurs des codets de sélecteur de classe. Si on ne connaît pas les capacités de gestion de trafic du domaine aval, et si aucun accord n'est établi, un nœud de sortie de domaine DS peut choisir de re-marquer les codets DS à zéro, faisant l'hypothèse que le domaine sans capacité DS va traiter uniformément le trafic comme service au mieux.

Dans le cas où un domaine sans capacité DS échange du trafic avec un domaine DS, le trafic qui s'écoule du domaine sans capacité DS devrait être conditionné au nœud d'entrée DS du domaine DS conformément au SLA ou politique approprié.

5. Considérations sur la diffusion groupée

L'utilisation des services différenciés par le trafic en diffusion groupée introduit un certain nombre de problèmes pour l'approvisionnement de service. Tout d'abord, les paquets de diffusion groupée qui entrent dans un domaine DS à un nœud d'entrée peuvent prendre simultanément plusieurs chemins à travers des segments du domaine du fait de la réplication du paquet de diffusion groupée. De cette façon, ils consomment plus de ressources du réseau que les paquets en envoi individuel. Lorsque l'adhésion au groupe de diffusion groupée est dynamique, il est difficile de prédire à l'avance la quantité des ressources du réseau qui peut être consommée par le trafic de diffusion groupée généré à partir d'un réseau amont pour un groupe particulier. Une conséquence de cette incertitude est qu'il peut être difficile de fournir des garanties de service quantitatives aux envoyeurs de diffusions groupées. De plus, il peut être nécessaire de réserver des codets et des PHB pour l'usage exclusif du trafic en envoi individuel, pour préserver des ressource de la voracité du trafic de diffusion groupée.

Le second problème est le choix du codet DS pour un paquet en diffusion groupée qui arrive à un nœud d'entrée DS. Comme ce paquet peut sortir du domaine DS à plusieurs nœuds de sortie DS qui échangent du trafic avec plusieurs domaines vers l'aval, le codet DS utilisé ne devrait pas résulter en une demande d'un service qui serait en violation d'un SLA d'homologue de la part d'un domaine DS. Lorsque on établit l'état d'un classeur et conditionneur de trafic à un nœud d'entrée DS pour un agrégat de trafic qui reçoit un service différencié qui s'étend à travers la frontière de sortie du domaine, l'identité du domaine de transit adjacent vers l'aval et les spécificités du SLA homologue correspondant peuvent être prises en compte dans la décision de configuration (sous réserve de la politique d'acheminement et de la stabilité de l'infrastructure d'acheminement). De cette façon, les SLA pour l'échange de trafic avec les domaines DS vers l'aval peuvent être partiellement mis en application à la sortie du domaine amont, ce qui réduit la charge de classification et de conditionnement de trafic au nœud de sortie du domaine amont. Cela n'est pas aussi facile dans le cas de trafic en diffusion groupée, à cause de la possibilité d'adhésion dynamique au groupe. Il en résulte que la garantie de service pour le trafic en envoi individuel peut être impactée. Un moyen de régler ce problème est d'établir un SLA d'échange de trafic distinct pour le trafic en diffusion groupée, et soit d'utiliser un ensemble particulier de codets pour les paquets de diffusion groupée, soit de mettre en œuvre les mécanismes nécessaires de classification et de conditionnement de trafic dans les nœuds de sortie DS pour fournir l'isolement préférentiel du trafic en envoi individuel en conformité avec le SLA d'échange de trafic avec le domaine aval.

6. Considérations sur la sécurité et le tunnelage

Cette section traite des problèmes de sécurité soulevés par l'introduction de services différenciés, et principalement des attaques potentielles de déni de service, et du potentiel de vol de service par du trafic non autorisé (paragraphe 6.1). De plus, le fonctionnement des services différenciés en présence de IPsec et leur interaction avec IPsec est aussi exposé (paragraphe 6.2), ainsi que les exigences d'audit (paragraphe 6.3). Cette section examine les problèmes introduits par l'utilisation des tunnels IPsec et non IPsec.

6.1 Vol et déni de service

Le principal objectif des services différenciés est de permettre que soient fournis différents niveaux de service pour les flux de trafic sur une infrastructure de réseau commune. Diverses techniques de gestion des ressources peuvent être utilisées pour réaliser cela, mais le résultat final sera que certains paquets reçoivent un service différent (par exemple, meilleur) que d'autres. La transposition du trafic réseau dans les comportements spécifiques qui résultent en différents services (par exemple, meilleur ou moins bon) est indiquée principalement par le champ DS, et donc un adversaire peut être capable d'obtenir un meilleur service en modifiant le champ DS par des codets qui indiquent des comportements utilisés pour des services améliorés ou en injectant des paquets avec le champ DS réglé avec de tels codets. À la limite, ce vol de service devient une attaque de déni de service lorsque le trafic modifié ou injecté épuise les ressources disponibles pour le transmettre lui et les autres flux de trafic. La défense contre un tel vol et les attaques de déni de service consiste en la combinaison du conditionnement de trafic aux nœuds DS frontière avec la sécurité et l'intégrité de l'infrastructure du réseau au sein d'un domaine DS.

Comme décrit à la Section 2, les nœuds d'entrée DS doivent conditionner tout le trafic entrant dans un domaine DS pour s'assurer qu'il a des codets DS acceptables. Cela signifie que les codets doivent se conformer au ou aux TCA applicables et à la politique d'approvisionnement de service du domaine. Donc, les nœuds d'entrée sont la principale ligne de défense contre les attaques de vol et de déni de service fondées sur des codets DS modifiés (par exemple, des codets auxquels le trafic n'a pas droit) car le succès d'une telle attaque constitue une violation du ou des TCA applicables et/ou des politiques d'approvisionnement de service. Une importante instance d'un nœud d'entrée est que tout nœud générant du trafic dans un domaine DS est le nœud d'entrée pour ce trafic, et doit s'assurer que tout le trafic généré porte des codets DS acceptables.

Une politique d'approvisionnement de service et les TCA d'un domaine peuvent tous deux exiger des nœuds d'entrée qu'ils changent les codets sur certains paquets entrants (par exemple, un routeur d'entrée peut régler le codet DS du trafic d'un consommateur conformément au SLA approprié). Les nœuds d'entrée doivent conditionner tout le reste du trafic entrant pour s'assurer que les codets DS sont acceptables : les paquets qui se trouvent avoir des codets non acceptables doivent être soit éliminés, soit avoir leurs codets DS modifiés à des valeurs acceptables avant d'être transmis. Par exemple, un nœud d'entrée qui reçoit du trafic provenant d'un domaine avec lequel n'existe aucun accord de service amélioré peut rétablir le codet DS au codet de PHB par défaut [DSFIELD]. L'authentification du trafic peut être exigée pour valider l'utilisation de certains codets DS (par exemple, ceux correspondant aux services améliorés) et une telle authentification peut être effectuée par des moyens techniques (par exemple, IPsec) et/ou non techniques (par exemple, la liaison entrante est connue pour être connectée à exactement un site d'usager).

Un accord inter domaine peut réduire ou éliminer l'obligation du conditionnement de trafic par le nœud d'entrée en rendant le domaine amont partiellement ou complètement responsable de s'assurer que le trafic a des codets DS acceptables pour le

domaine aval. Dans ce cas, le nœud d'entrée peut encore effectuer des vérifications redondantes du conditionnement de trafic pour réduire la dépendance au domaine amont (par exemple, de telles vérifications peuvent empêcher les attaques de vol de service de se propager à travers les frontières du domaine). Si une telle vérification échoue à cause du non respect de ses responsabilités par le domaine amont, cet échec est un événement d'audit ; l'enregistrement de journal d'audit généré devrait comporter la date/heure de réception du paquet, les adresses de source et de destination IP, et le codet DS qui a causé la défaillance. En pratique, les gains limités de telles vérifications doivent être mis en balance avec l'impact potentiel sur les performances pour déterminer quelles vérifications effectuer dans ces circonstances.

Les nœuds intérieurs dans un domaine DS peuvent s'appuyer sur le champ DS pour associer le trafic de services différenciés avec les comportements utilisés pour mettre en œuvre les services améliorés. Tout nœud qui fait ainsi dépend du fonctionnement correct du domaine DS pour empêcher l'arrivée de trafic avec des codets DS inacceptables. Le souci de robustesse impose que l'arrivée de paquets avec des codets DS inacceptables ne doive pas causer de défaillance (par exemple, un blocage) des nœuds du réseau. Les nœuds intérieurs ne sont pas responsables de l'application de la politique d'approvisionnement de service (ou des SLA individuels) et ne sont donc pas obligés de vérifier les codets DS avant de les utiliser. Les nœuds intérieurs peuvent effectuer des vérifications de conditionnement de trafic sur les codets DS (par exemple, vérifier des codets DS qui ne sont jamais utilisés pour le trafic sur une liaison spécifique) pour améliorer la sécurité et la robustesse (par exemple, la résistance aux attaques de vol de service sur la base des modifications de codets DS). Toute défaillance détectée lors d'une telle vérification est un événement d'audit et l'enregistrement d'événement d'audit généré devrait comporter la date/heure à laquelle le paquet a été reçu, les adresses IP de source et de destination, et le codet DS qui a causé la défaillance. En pratique, les gains limités tirés de telles vérifications devraient être mis en balance avec l'impact potentiel sur les performances pour déterminer quelles vérifications effectuer aux nœuds intérieurs.

Une liaison qui ne peut pas être correctement sécurisée contre la modification des codets DS ou l'injection de trafic par des adversaires devrait être traitée comme une liaison frontière (et donc, tout trafic qui arrive sur cette liaison est traité comme si il entrait dans le domaine à un nœud d'entrée). Les politiques locales de sécurité fournissent la définition de "correctement sécurisé," et une telle définition peut comporter la détermination que les risques et les conséquences de la modification d'un codet DS et/ou de l'injection de trafic ne justifient aucune mesure supplémentaire de sécurité pour une liaison. La sécurité d'une liaison peut être améliorée via des contrôles d'accès physiques et/ou des moyens logiciels tels que des tunnels qui assurent l'intégrité du paquet.

6.2 Interactions entre IPsec et tunnelage

Le protocole IPsec, tel que défini dans [ESP], [AH], ne comporte pas le champ DS de l'en-tête IP dans ses calculs de chiffrement (dans le cas d'un mode tunnel, c'est le champ DS de l'en-tête IP externe qui n'est pas inclus). Donc la modification du champ DS par un nœud du réseau n'a pas d'effet sur la sécurité IPsec de bout en bout, parce qu'elle ne peut causer l'échec d'aucune vérification d'intégrité IPsec. Par conséquent, IPsec ne donne aucune défense contre une modification hostile du champ DS (c'est à dire, une attaque par interposition) car la modification hostile n'aura aucun effet sur la sécurité IPsec de bout en bout. Dans certains environnements, la capacité à modifier le champ DS sans affecter les vérifications d'intégrité d'IPsec peut constituer un chemin détourné ; si il est nécessaire d'éliminer un tel chemin ou de réduire sa bande passante, le domaine DS devrait être configuré de telle sorte que le traitement nécessaire (par exemple, régler sur le trafic sensible tous les champs DS à une seule valeur) puisse être effectué aux nœuds de sortie DS où le trafic sort vers des domaines à plus forte sécurité.

Le mode tunnel d'IPsec apporte la sécurité pour le champ DS de l'en-tête IP encapsulé. Un paquet IPsec en mode tunnel contient deux en-têtes IP : un en-tête externe fourni par le nœud d'entrée du tunnel et un en-tête interne encapsulé fourni par la source originale du paquet. Lorsque un tunnel IPsec est hébergé (en tout ou en partie) sur un réseau à services différenciés, les nœuds intermédiaires de réseau fonctionnent sur le champ DS dans l'en-tête externe. Au nœud de sortie du tunnel, le traitement IPsec comporte l'élimination de l'en-tête externe et la transmission du paquet (si nécessaire) en utilisant l'en-tête interne. Si l'en-tête IP interne n'a pas été traité par un nœud d'entrée DS pour le domaine DS du nœud de sortie du tunnel, celui-ci est le nœud d'entrée DS pour le trafic qui sort du tunnel, et doit donc porter la responsabilité du conditionnement de trafic correspondante (voir au paragraphe 6.1). Si le traitement IPsec comporte une vérification d'intégrité suffisamment forte du paquet encapsulé (où "suffisamment" est déterminé par la politique locale de sécurité) le nœud de sortie du tunnel peut supposer en toute sécurité que le champ DS dans l'en-tête a la même valeur que celle qu'il avait au nœud d'entrée du tunnel. Cela permet à un nœud de sortie de tunnel dans le même domaine DS que le nœud d'entrée du tunnel de traiter en toute sécurité un paquet qui réussit une telle vérification d'intégrité comme si il était arrivé d'un autre nœud au sein du même domaine DS, en omettant le conditionnement de trafic de nœud d'entrée DS qui aurait été exigé autrement. Une importante conséquence est que des liaisons par ailleurs non sûres internes à un domaine DS peuvent être sécurisée par un tunnel IPsec suffisamment fort.

Cette analyse et ses implications s'appliquent à tout protocole de tunnelage qui effectue des vérifications d'intégrité, mais le niveau d'assurance du champ DS de l'en-tête interne dépend de la force de la vérification d'intégrité effectuée par le protocole de tunnelage. En l'absence d'assurance suffisante sur un tunnel qui peut transiter par des nœuds en-dehors du

domaine DS actuel (ou qui sont par ailleurs vulnérables) le paquet encapsulé doit être traité comme si il était arrivé à un nœud d'entrée DS de l'extérieur du domaine.

Le protocole IPsec exige actuellement que le champ DS de l'en-tête interne ne soit pas changé par le traitement IPsec de désencapsulation au nœud de sortie d'un tunnel. Cela assure que des modifications hostiles au champ DS ne peuvent pas être utilisées pour lancer des attaques de vol ou de déni de service à travers un point d'extrémité de tunnel IPsec, car de telles modifications seraient éliminées par le point d'extrémité du tunnel. Le présent document n'apporte aucun changement à cette exigence d'IPsec.

Si les spécifications d'IPsec sont modifiées à l'avenir pour permettre qu'un nœud de sortie de tunnel modifie le champ DS dans l'en-tête IP interne sur la base de la valeur du champ DS dans l'en-tête externe (par exemple, en copiant tout ou partie du champ DS externe dans le champ DS interne) des considérations supplémentaires devraient alors être prises en compte. Pour un tunnel contenu entièrement au sein d'un seul domaine DS et pour lequel les liaisons sont adéquatement sécurisées contre les modifications du champ DS externe, les seules limites aux modifications du champ DS interne seraient celles imposées par la politique d'approvisionnement de service du domaine. Autrement, le nœud de sortie de tunnel effectuant de telles modifications agirait comme un nœud d'entrée DS pour le trafic sortant du tunnel et devrait porter la responsabilité du conditionnement de trafic d'un nœud d'entrée, y compris la défense contre les attaques de vol et de déni de service (voir au paragraphe 6.1). Si le tunnel entre dans le domaine DS à un nœud différent du nœud de sortie du tunnel, celui-ci peut dépendre de ce que le nœud d'entrée DS amont s'est assuré que les valeurs du champ DS externe sont acceptables. Même dans ce cas, il y a certaines vérifications qui ne peuvent être effectuées que par le nœud de sortie du tunnel (par exemple, une vérification de cohérence entre les codets DS internes et externes pour un tunnel chiffré). La détection de toute défaillance à une telle vérification est un événement d'audit et l'entrée de journal d'audit générée devrait comporter la date/heure de réception du paquet, les adresses IP de source et de destination, et le codet DS qui n'était pas acceptable.

Un tunnel IPsec peut être vu d'au moins deux façons différentes d'un point de vue architectural. Si le tunnel est vu comme un "lien virtuel" logique d'un seul bond, les actions des nœuds intermédiaires pour la transmission du trafic tunnelé ne devraient pas être visibles au delà des extrémités du tunnel et donc le champ DS ne devrait pas être modifié au titre du processus de désencapsulation. À l'opposé, si le tunnel est vu comme un participant multi bonds dans la transmission du trafic, la modification du champ DS au titre du processus de désencapsulation du tunnel peut être souhaitable. Un exemple spécifique de cette dernière situation survient lorsque un tunnel se termine à un nœud intérieur d'un domaine DS auquel l'administrateur du domaine ne souhaite pas déployer de logique de conditionnement de trafic (par exemple, pour simplifier la gestion du trafic). Cela pourrait être pris en charge en utilisant le codet DS dans l'en-tête IP externe (qui a été soumis au conditionnement de trafic au nœud d'entrée DS) pour rétablir le codet DS dans l'en-tête IP interne, déplaçant effectivement les responsabilités du conditionnement du trafic d'entrée DS du nœud de sortie du tunnel IPsec au nœud d'entrée DS amont approprié (qui doit déjà assurer cette fonction pour le trafic non encapsulé).

6.3 Audit

Tous les systèmes qui prennent en charge les services différenciés ne vont pas mettre en œuvre l'audit. Cependant, si la prise en charge de services différenciés est incorporée dans un système qui accepte l'audit, la mise en œuvre de services différenciés devrait aussi accepter l'audit. Si une telle prise en charge est présente, la mise en œuvre doit permettre à un administrateur de système d'activer ou désactiver les services différenciés comme un tout, et peut permettre qu'un tel audit soit partiellement activé ou désactivé.

La granularité de l'audit est pour la plus grande part une affaire locale. Cependant, plusieurs événements auditables sont identifiés dans le présent document et pour chacun de ces événements est défini un ensemble minimum des informations qui devraient être incluses dans un journal d'audit. Des informations supplémentaires (par exemple, les paquets en rapport avec celui qui a déclanché l'événement auditable) peuvent aussi être inclus dans le journal d'audit pour chacun de ces événements, et des événements supplémentaires, non explicitement invoqués dans la présente spécification, peuvent aussi résulter en entrées de journal d'audit. Il n'est pas exigé que le receveur transmette de message à l'envoyeur supposé en réponse à la détection d'un événement auditable, à cause du potentiel induit de déni de service via une telle action.

7. Remerciements

Le présent document a bénéficié de projets antérieurs de Steven Blake, David Clark, Ed Ellesson, Paul Ferguson, Juha Heinanen, Van Jacobson, Kalevi Kilkki, Kathleen Nichols, Walter Weiss, John Wroclawski, et Lixia Zhang.

Les auteurs tiennent à remercier les personnes suivantes de leurs commentaires et suggestions utiles : Kathleen Nichols, Brian Carpenter, Konstantinos Dovrolis, Shivkumar Kalyana, Wu-chang Feng, Marty Borden, Yoram Bernet, Ronald Bonica, James Binder, Borje Ohlman, Alessio Casati, Scott Brim, Curtis Villamizar, Hamid Ould- Brahi, Andrew Smith, John Renwick, Werner Almesberger, Alan O'Neill, James Fu, et Bob Braden.

8. Références

[802.1p] ISO/IEC Final CD 15802-3 "Information technology - Telecommunications and information exchange entre systems - Local and metropolitan area networks - Common specifications - Part 3: Media Access Control (MAC) bridges", (projet actuel disponible sous la référence IEEE P802.1D/D15).

[AH] S. Kent et R. Atkinson, "En-tête d'authentification IP", RFC2402, novembre 1998. (*Obsolète, voir* RFC4302, 4305)

[ATM] "ATM Traffic Management Specification Version 4.0" <af-tm-0056.000>, ATM Forum, avril 1996.

[Bernet] Y. Bernet et autres, "Cadre de fonctionnement de services intégrés sur réseaux Diffserv", RFC2998, novembre 2000. *(Information)*

[DSFIELD] K. Nichols, S. Blake, F. Baker et D. Black, "Définition du champ Services différenciés (DS Field) dans les en-têtes IPv4 et IPv6", RFC2474, décembre 1998. (*MàJ par* RFC3168, RFC3260) *(P.S.)*

[EXPLICIT] D. Clark and W. Fang, "Explicit Allocation of Best Effort Packet Delivery Service", IEEE/ACM Trans. on Networking, vol. 6, n° 4, août 1998, pp. 362-373.

[ESP] S. Kent et R. Atkinson, "Encapsulation de charge utile de sécurité IP (ESP)", RFC2406, novembre 1998. *(Obsolète, voir RFC4303)*

[FRELAY] ANSI T1S1, "DSSI Core Aspects of Frame Rely", mars 1990.

[RFC0791] J. Postel, éd., "Protocole Internet - Spécification du protocole du programme Internet", STD 5, septembre 1981.

[RFC1349] P. Almquist, "Type de service dans la suite de protocole Internet", juillet 1992. *(Obsolète, voir RFC2474)*

[RFC1633] R. Braden, D. Clark et S. Shenker, "Intégration de services dans l'architecture de l'Internet : généralités", juin 1994. *(Info.)*

[RFC1812] F. Baker, "Exigences pour les routeurs IP version 4", juin 1995. *(Mise à jour par la RFC 2644)*

[RSVP] R. Braden, éd., L. Zhang, S. Berson, S. Herzog, S. Jamin, "Protocole de réservation de ressource (RSVP) -- version 1, spécification fonctionnelle", RFC2205, septembre 1997. (*MàJ par* RFC2750, RFC3936, RFC4495) *(P.S.)*

[2BIT] K. Nichols, V. Jacobson, and L. Zhang, "A Two-bit Differentiated Services Architecture for the Internet", ftp://ftp.ee.lbl.gov/papers/dsarch.pdf, novembre 1997.

[TR] ISO/CEI 8802-5 "Technologies de l'information – Échanges d'informations entre systèmes de télécommunications – Réseaux de zone locale et métropolitaine - Spécifications communes - Partie 5 : Spécifications de la couche physique et de la méthode d'accès par anneau à jetons", (aussi norme ANSI/IEEE 802.5- 1995).

Adresse des auteurs

Steven Blake
Torrent Networking Technologies
3000 Aerial Center, Suite 140
Morrisville, NC 27560
téléphone : +1-919-468-8466 x232
mél : slblake@torrentnet.com

David L. Black
EMC Corporation
35 Parkwood Drive
Hopkinton, MA 01748
téléphone : +1-508-435-1000 x76140
mél : black_david@emc.com

Mark A. Carlson
Sun Microsystems, Inc.
2990 Center Green Court South
Boulder, CO 80301
téléphone : +1-303-448-0048 x115
mél : mark.carlson@sun.com

Elwyn Davies
Nortel UK
London Road
Harlow, Essex CM17 9NA, UK
téléphone : +44-1279-405498
mél : elwynd@nortel.co.uk

Zheng Wang
Bell Labs Lucent Technologies
101 Crawfords Corner Road
Holmdel, NJ 07733
mél : zhwang@bell-labs.com

Walter Weiss
Lucent Technologies
300 Baker Avenue, Suite 100
Concord, MA 01742-2168
mél : wweiss@lucent.com

Déclaration complète de droits de reproduction

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