Groupe de travail Réseau
Request for Comments : 5474
Catégorie : Information

N. Duffield, éditeur, AT&T Labs - Research
D. Chiou, University of Texas
B. Claise, Cisco Systems, Inc.
A. Greenberg, Microsoft
M. Grossglauser, EPFL & Nokia
J. Rexford, Princeton University

Traduction Claude Brière de L'Isle
mars 2009

Cadre pour la sélection et le rapport de paquets

Statut de ce mémoire

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

Résumé

Le présent document spécifie un cadre pour le protocole d'échantillonnage de paquets (PSAMP, *Packet SAMPling*). Les fonctions de ce protocole sont de choisir des paquets dans un flux conformément à un ensemble de sélecteurs normalisés, pour former un flux de rapports sur les paquets choisis, et d'exporter les rapports à un collecteur. Ce cadre détaille les composants de cette architecture, puis décrit des exigences génériques, motivées par le double but de déploiement généralisé et d'utilité des rapports pour les applications. Les exigences détaillées pour la sélection, le rapport, et l'exportation sont décrits, avec les exigences de configuration des fonctions de PSAMP.

Table des matières

1. Introduction

Le présent document décrit le cadre PSAMP pour que des éléments de réseau choisissent des sous ensembles de paquets par des méthodes statistiques et autres, et pour exporter un flux de rapports sur les paquets choisis à un collecteur.

Le motif de la norme PSAMP vient du besoin d'un soutien fondé sur les mesures de la gestion et du contrôle de réseau à travers des domaines multi fabricants. Cela exige une cohérence des types de schéma de sélection disponibles au niveau du domaine, et de la manière dont les mesures résultantes sont présentées et interprétées.

Le motif d'opérations de sélection de paquet spécifiques vient des applications qu'elles permettent. Le développement de la norme PSAMP est ouvert aux influences par les exigences des normes des groupes de travail de l'IETF en rapport, par exemple, métriques des performances IP (IPPM, *IP Performance Metrics*) [RFC2330] et ingénierie du trafic Internet (TEWG, *Internet Traffic Engineering*).

Le nom PSAMP est la contraction des mots "Packet Sampling". Le mot "échantillonnage" rend l'idée que seul un sous ensemble de tous les paquets passant par un élément de réseau va être choisi pour le rapport. Mais les opérations de sélection de PSAMP incluent le choix aléatoire, le choix déterministe (filtrage) et des approximations déterministes de choix aléatoires (choix fondés sur le hachage).

2. Vue d'ensemble des documents PSAMP

Le présent document fait partie d'une série de documents du groupe PSAMP.

[RFC5474] (ce document) : "Cadre de travail pour la sélection et le rapport de paquet" décrit le cadre de PSAMP pour que les éléments de réseau choisissent des sous ensembles de paquets par des méthodes statistiques et autres, et pour exporter un flux de rapports sur les paquets choisis à un collecteur.

RFC 5475 : "Techniques d'échantillonnage et de filtrage pour la sélection de paquets IP" décrit l'ensemble de techniques de sélection de paquets prise en charge par PSAMP.

[RFC5476] : "Spécification du protocole d'échantillonnage de paquet (PSAMP)" spécifie l'exportation des informations de paquet d'un processus d'exportation PSAMP à un processus de collecte PSAMP.

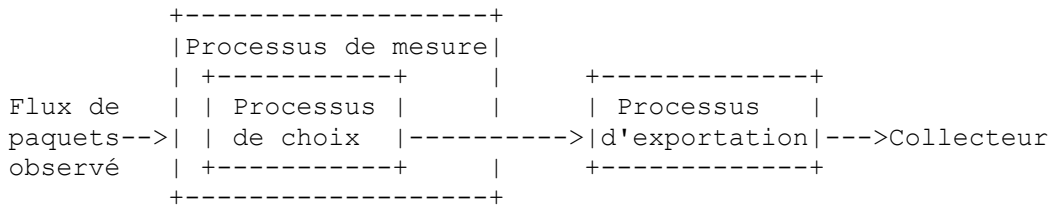
[RFC5477] : "Modèle d'information pour l'exportation d'échantillonnage de paquet" définit un modèle d'information et de données pour PSAMP.

3. Éléments, terminologie, et grandes lignes de l'architecture

3.1 Description générale de l'architecturePSAMP

Voici une description générale informelle du protocole PSAMP fonctionnant dans un appareil PSAMP (tous ces termes vont être définis). Un flux de paquets est observé à un point d'observation. Un processus de choix inspecte chaque paquet pour déterminer si il est ou non à choisir pour en faire rapport. Le processus de choix fait partie du processus de mesure, qui construit un rapport sur chaque paquet choisi, en utilisant le contenu du paquet, et éventuellement d'autres informations comme le traitement du paquet au point d'observation ou l'horodatage d'arrivée. Un processus d'exportation envoie les rapports de paquet à un collecteur, avec toutes les informations subsidiaires nécessaires pour leur interprétation.

La figure suivante indique la séquence des trois processus (choix, mesure, et exportation) au sein de l'appareil PSAMP.



Les paragraphes qui suivent donnent les définitions détaillées de chacun des objets désignés.

3.2 Points d'observation, flux de paquet, et contenu de paquet

Ce paragraphe contient la définition des termes pertinents pour obtenir l'entrée du paquet dans le processus de choix.

Point d'observation [RFC5101] : c'est une localisation dans le réseau où les paquets peuvent être observés. Des exemples incluent :
(i) une ligne à laquelle est attachée une sonde ;
(ii) un support partagé, comme un LAN fondé sur Ethernet ;
(iii) un seul accès d'un routeur, ou un ensemble d'interfaces (physiques ou logiques) d'un routeur ;
(iv) un sous système incorporé de mesures au sein d'une interface.
Noter que tout point d'observation est associé à un domaine d'observation et que ce point d'observation peut être un sur-ensemble de plusieurs autres points d'observation. Par exemple, un point d'observation peut être une carte de ligne entière. Cela serait le sur-ensemble des points d'observation individuels aux interfaces de la carte de ligne.

Flux de paquets observé : c'est l'ensemble de tous les paquets observés au point d'observation.

Flux de paquets : ensemble de paquets provenant du flux de paquets observé qui s'écoule après un point spécifié dans le processus de mesure. Un exemple de flux de paquets est le résultat du processus de sélection. Noter que les paquets choisis dans un flux, par exemple, par échantillonnage, ne possèdent pas nécessairement une propriété qui peut les distinguer des paquets qui n'ont pas été choisis. Pour cette raison, le terme de "flux" est préféré à celui de "courant", qui est défini comme un ensemble de paquets avec des propriétés communes [RFC3917].

Contenu de paquet : il note l'union de l'en-tête de paquet (qui inclut des champs d'en-tête de couche de liaison, de couche réseau, et autre couche d'encapsulation) et la charge utile du paquet.

3.3 Processus de choix

Ce paragraphe définit le processus de choix et les objets qui s'y rapportent.

Processus de choix : il prend en entrée le flux de paquets observé et choisit un sous ensemble de ce flux comme résultat.

État de sélection : un processus de choix peut maintenir des informations d'état à utiliser par le processus de choix. À un moment donné, l'état de sélection peut dépendre des paquets observés à ce moment ou avant, et d'autres variables. Des exemples incluent :
(i) les numéros de séquence des paquets à l'entrée des sélecteurs ;
(ii) un horodatage de l'observation du paquet au point d'observation ;
(iii) des itérateurs pour les générateurs de nombres pseudo aléatoires ;
(iv) des valeurs de hachage calculées durant la sélection ;
(v) des indicateurs de si le paquet a été choisi par un sélecteur donné.

Les processus de choix peuvent changer des portions de l'état de sélection par suite du traitement d'un paquet. L'état de sélection pour un paquet est de refléter l'état après le traitement du paquet.

Sélecteur : un sélecteur définit quelle action effectue un processus de choix sur un seul paquet de son entrée. Si il est choisi, le paquet devient un élément du flux de paquets de sortie.
Le sélecteur peut utiliser les informations suivantes pour déterminer si un paquet est choisi :
(i) le contenu du paquet ;
(ii) des informations déduites du traitement du paquet au point d'observation ;
(iii) tout état de sélection qui peut être maintenu par le processus de choix.

Sélecteur composite : c'est une composition ordonnée de sélecteurs, dans laquelle le flux de paquets de sortie provenant d'un sélecteur forme le flux de paquets d'entrée du sélecteur suivant.

Sélecteur primaire : un sélecteur est primaire si il n'est pas un sélecteur composite.

3.4 Rapports

Rapports de paquets : ils comprennent un sous ensemble configurable d'une entrée de paquets au processus de sélection, incluant le contenu de paquet, les informations relatives à son traitement (par exemple, l'interface de sortie) et son état de sélection associé (par exemple, un hachage du contenu du paquet).

Interprétation de rapport : elle comporte des informations subsidiaires, relatives à un ou plusieurs paquets, qui sont utilisées pour l'interprétation de leurs rapports de paquet. Des exemples incluent des paramètres de configuration du processus de choix.

Flux de rapports : c'est le résultat du processus de mesures, comprenant deux types distincts d'informations : les rapports de paquet et l'interprétation de rapport.

3.5 Processus de mesures

Un processus de mesure choisit les paquets à partir du flux de paquets observé en utilisant un processus de choix, et produit comme résultat un flux de rapport concernant les paquets choisis.

Le processus de mesure PSAMP peut être vu comme l'analogue du processus de mesure IPFIX [RFC5101], qui produit des enregistrements de flux comme résultat, avec la différence que le processus de mesure PSAMP contient toujours un processus de choix. La relation entre PSAMP et IPFIX est décrite dans les [RFC5477] et [RFC5474].

3.6 Processus d'exportation

Processus d'exportation : il envoie, sous la forme de paquets d'exportation, le résultat d'un ou plusieurs processus de mesure à un ou plusieurs collecteurs.

Paquet d'exportation : combinaison d'interprétations de rapport et/ou d'un ou plusieurs rapports de paquet qui sont groupés par le processus d'exportation dans un paquet d'export à exporter à un collecteur.

3.7 Appareil PSAMP

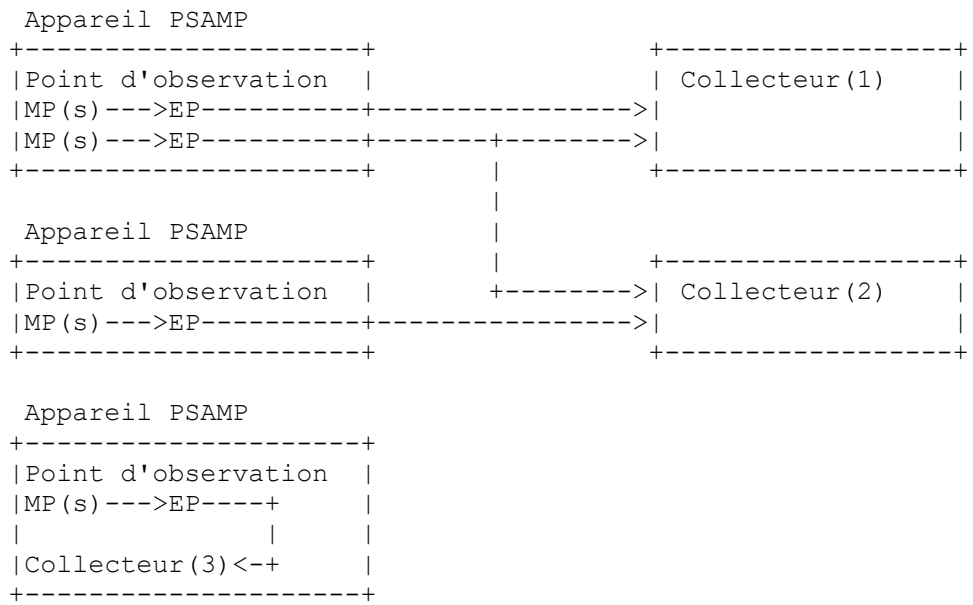
Appareil PSAMP : appareil qui héberge au moins un point d'observation, un processus de mesure (qui inclut un processus de choix) et un processus d'exportation. Normalement, le ou les points d'observation, le ou les processus de mesure, et le ou les processus d'exportation correspondants sont co-localisés dans cet appareil, par exemple, dans un routeur.

3.8 Collecteur

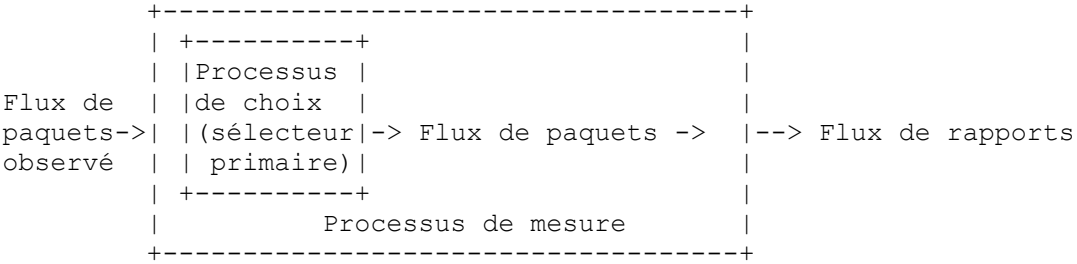
Collecteur : un collecteur reçoit un flux de rapport exporté par un ou plusieurs processus d'exportation. Dans certains cas, l'hôte des processus de mesure et/ou d'exportation peut aussi servir de collecteur.

3.9 Configurations possibles

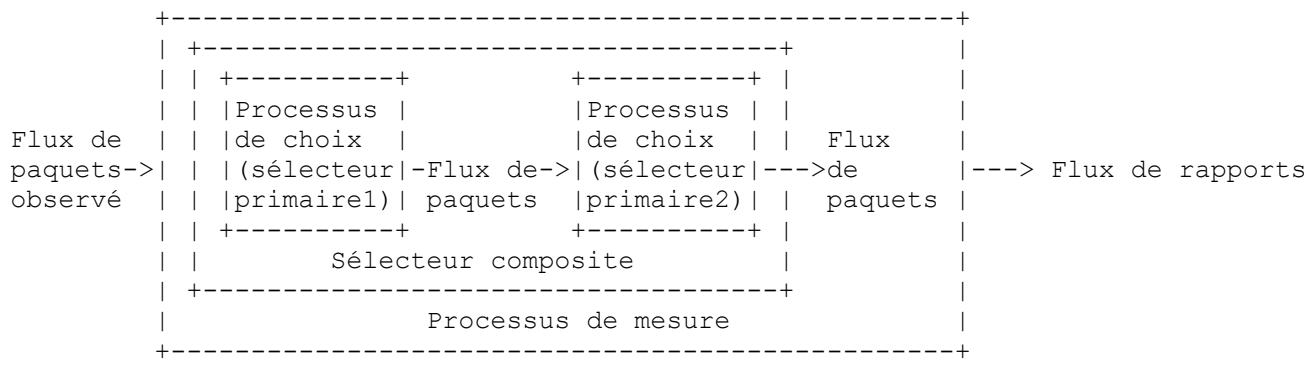
Les diverses possibilités pour l'architecture générale de ces éléments sont les suivantes.
MP *(Metering Process)* = processus de mesure ; EP *(Exporting process)* = processus d'exportation



La plus simple configuration de processus de mesure est composé de :



Un processus de mesure avec un sélecteur composite est constitué de :



4. Exigences générales pour PSAMP

Cette section décrit les exigences générales pour le protocole PSAMP, dont certaines sont réalisées comme des exigences spécifiques dans les sections suivantes.

4.1 Exigences générales pour le processus de choix

(a) Ubiquité : les sélecteurs doivent être assez simples pour être mis en œuvre partout au taux de ligne maximal.

(b) Applicabilité : l'ensemble de sélecteurs doit être assez riche pour prendre en charge une gamme d'applications et protocoles existants et émergents fondés sur la mesure. Cela exige un compromis acceptable entre la gamme d'applications d'ingénierie du trafic et des tâches opérationnelles qu'elles permettent, et la complexité de l'ensemble de capacités.

(c) Extensibilité : le protocole doit être capable de s'accommoder de sélecteurs de paquet supplémentaires non définis actuellement.

(d) Souplesse : le protocole doit supporter le choix de paquets utilisant divers protocoles réseau ou couches d'encapsulation, incluant la version 4 du protocole Internet (IPv4) [RFC0791], la version 6 du protocole Internet (IPv6) [RFC2460], et la commutation d'étiquettes multi protocoles (MPLS, *Multiprotocol Label Switching*) [RFC3031].

(e) Sélection robuste : le choix de paquets doit être robuste contre les tentatives de fabriquer un flux de paquets observé à partir duquel les paquets seraient choisis de façon disproportionnée (par exemple, pour éviter la sélection ou pour surcharger les systèmes de mesure).

(f) Processus de mesure en parallèle : le protocole doit supporter le fonctionnement simultané de plusieurs processus de mesure indépendants sur le même hôte.

(g) Causalité : la décision de choix pour chaque paquet devrait dépendre seulement faiblement, sinon pas du tout, de l'arrivée des paquets futurs. Cela favorise l'ubiquité en limitant la complexité de la logique de sélection.

(h) Paquets chiffrés : les sélecteurs qui interprètent les champs de paquets doivent être configurables à ignorer (c'est-à-dire, ne pas choisir) les paquets chiffrés, quand ils sont détectés.

Des sélecteurs spécifiques sont mentionnés à la Section 5, et décrits plus en détail dans la [RFC5475].

4.2 Exigences générales pour les rapports

(i) Auto défini : le flux de rapports doit être complet en ce sens qu'aucune information supplémentaire n'a besoin d'être restituée du point d'observation afin d'interpréter et analyser les rapports.

(j) Indication de perte d'informations : le flux de rapports doit inclure des informations suffisantes pour indiquer ou

permettre la détection de pertes survenant dans les processus de choix, de mesure, et/ou d'exportation, ou dans le transport. Cela peut être réalisé par l'utilisation de numéros de séquence.

(k) Précision : le flux de rapports doit inclure des informations qui permettent que soit déterminée la précision des mesures.

(l) Fidélité : toutes les quantités rapportées qui se rapportent au traitement de paquet doivent refléter l'état du routeur et de la configuration rencontrés par le paquet au moment où il est reçu par le processus de mesure.

(m) Confidentialité : bien que le choix du contenu des rapports de paquet doive répondre aux besoins des applications de mesure, il doit aussi se conformer à la [RFC2804]. En particular, la capture complète de paquet de flux de paquets arbitraires sort explicitement du domaine d'application.

Voir à la Section 6 plus d'explications sur les rapports.

4.3. Exigences générales pour le processus d'exportation

(n) Opportunité : la configuration doit permettre de limiter les délais de mise en mémoire tampon pour la formation et la transmission des paquets d'exportation. Voir les détails au paragraphe 8.5.

(o) Évitement d'encombrement : l'exportation d'un flux de rapports à travers un réseau doit éviter de créer de l'encombrement, en accord avec la [RFC2914]. Ceci est développé au paragraphe 8.3.

(p) Exportation sûre :
(i) confidentialité : l'option de chiffrer les données exportées doit être fournie.
(ii) intégrité : les altérations dans le transit aux données exportées doivent être détectables au collecteur.
(iii) authenticité : l'authenticité des données exportées doit être vérifiable par le collecteur afin de détecter les données falsifiées.

Les motivations sont ici les mêmes que pour la sécurité dans l'exportation IPFIX ; voir le paragraphe 6.3 et la Section 10 de la [RFC3917].

4.4 Exigences générales pour la configuration

(q) Facilité de configuration : cela s'applique à la facilité de configuration des paramètres d'échantillonnage et d'exportation, par exemple, pour la reconfiguration automatique à distance en réponse aux rapports collectés.

(r) Configuration sûre : l'option de configurer via des protocoles qui empêchent la reconfiguration non autorisée ou l'espionnage des communications de configuration doit être disponible. L'espionnage d'une configuration pourrait permettre à un attaquant d'obtenir la connaissance de ce qui lui serait utile pour fabriquer un flux de paquets pour éviter la sélection ou surcharger l'infrastructure de mesure.

La configuration est discutée à la Section 9.

5. Choix de paquet

Cette section détaille les exigences spécifiques pour le processus de choix, motivées par les exigences génériques du paragraphe 3.3.

5.1 Deux types de sélecteurs

PSAMP classe les sélecteurs en deux types :

Filtrage : un filtre est un sélecteur qui choisi un paquet de façon déterminée sur la base du contenu de paquet, ou de son traitement, ou de fonctions de cela qui se produisent dans l'état de sélection. Deux exemples sont :
(i) Filtrage sur la correspondance de propriété : un paquet est choisi si un champ spécifique dans le paquet est égal à une valeur prédéfinie.
(ii) Choix fondé sur le hachage : une fonction de hachage est appliquée au contenu de paquet, et le paquet est choisi si le

résultat tombe dans une gamme spécifiée.

Échantillonnage : un sélecteur qui n'est pas un filtre est appelé une opération d'échantillonnage. Cela reflète la notion intuitive que si le choix d'un paquet ne peut pas être déterminé à partir de son seul contenu, il doit y avoir lieu à un certain type d'échantillonnage.

Les opérations d'échantillonnage peuvent être divisées en deux sous types :
(i) Échantillonnage indépendant du contenu, qui n'utilise pas le contenu de paquet pour prendre les décisions d'échantillonnage. Des exemples incluent l'échantillonnage systématique, et l'échantillonnage pseudo aléatoire uniforme fondé sur un nombre pseudo aléatoire dont la génération est indépendante du contenu de paquet. Noter que dans l'échantillonnage indépendant du contenu, il n'est pas nécessaire d'accéder au contenu de paquet afin de prendre la décision de choix.
(ii) Échantillonnage dépendant du contenu, dans lequel le contenu du paquet est utilisé pour prendre les décisions de choix. Une application est une sélection pseudo aléatoire avec une probabilité qui dépend du contenu d'un champ du paquet, par exemple, l'échantillonnage de paquets avec une probabilité dépendant de leur numéro d'accès TCP/UDP. Noter que ceci n'est pas un filtre.

5.2 Sélecteurs de paquet PSAMP

Un ensemble de sélecteurs de paquets est décrit en détails dans la [RFC5475]. Ici, on résume seulement brièvement leur signification pour être complet.

Un processus de sélection PSAMP doit prendre en charge au moins un des sélecteurs suivants.

* Échantillonnage systématique fondé sur le compte : le choix de paquets est déclenché périodiquement par le compte de paquets, un certain nombre de paquets successifs étant sélectionné à la suite de chaque déclenchement.

* Échantillonnage systématique fondé sur le temps : c'est similaire à l'échantillonnage systématique fondé sur le compte sauf que le choix est effectué par rapport au temps plutôt que sur le compte. Le choix des paquets est déclenché périodiquement à des instants séparés par un intervalle appelé l'espacement. Tous les paquets qui arrivent dans un certain intervalle de déclenchement (appelé la longueur d'intervalle) sont choisis.
'
* Échantillonnage probabiliste de n parmi N : de chaque bloc successif de N paquets fondé sur le compte, n sont choisis au hasard.

* Échantillonnage probabiliste uniforme : les paquets sont choisis indépendamment avec une probabilité d'échantillonnage fixe de p.

* Échantillonnage probabiliste non uniforme : les paquets sont choisis indépendamment avec une probabilité p qui dépend du contenu du paquet.

* Filtrage par correspondance de propriéte : avec cette méthode de filtrage, un paquet est choisi si un champ spécifque au sein du paquet et/ou des propriétés de l'état de routeur sont égales à une valeur prédéfinie. Les champs de filtre possibles sont tous les attributs de flux IPFIX spécifiés dans la [RFC5102]. D'autres champs peuvent être définis par des extensions spécifiques de fabricant.
  Un paquet est choisi si Champ:=Valeur. Les gabarits et les gammes ne sont pris en charge que dans la mesure où la [RFC5102] les permet, par exemple, en fournissant des champs explicites comme les gabarits de réseau pour les adresses de source et de destination.

  Les opérations ET logiques sont possibles en enchaînant les filtres, produisant donc une opération de choix composite. Dans ce cas, l'ordre dans lequel le filtrage se produit est implicitement défini (les filtres externes viennent après les filtres internes). Cependant, tant que l'enchaînement est seulement sur les filres, le résultat de la cascade de filtres est indépendant de l'ordre, mais l'ordre peut être important pour des besoins de mise en œuvre, car le premier filtre va devoir travailler à un taux plus élevé. Dans tous les cas, une mise en œuvre n'est pas obligée de respecter l'ordre des filtres, pour autant que le résultat soit le même, et elle peut même appliquer le filtrage composite en une seule étape.

  Les opérations OU logiques ne sont pas prises en charge dans le modèle de base. Des filtres plus sophistiqués (par exemple, qui prennent en charge des gabarits binaires, des gammes, ou des opérations OU) peuvent être réalisés par des schémas spécifiques de fabricant.

Les opérations de correspondance à une propriété devraient être disponibles pour différentes portions de protocole de l'en-tête de paquet :
(i) en-tête IP (à l'exclusion des options dans IPv4, des en-têtes en pile dans IPv6)
(ii) en-tête de transport
(iii) en-têtes d'encapsulation (par exemple, la pile d'étiquettes MPLS, si elle est présente).

Quand l'appareil PSAMP offre le filrage par correspondance de propriété, et, dans sa capacité ordinaire autre que d'effectuer des fonctions PSAMP, identifie ou traite des informations provenant des protocoles IP, de transport, ou d'encapsulation, alors les informations devraient être rendues disponibles au filtrage. Par exemple, quand un appareil PSAMP est un routeur qui achemine sur la base de d'adresse IP de destination, ce champ devrait être rendu disponible au filtrage. À l'inverse, un appareil PSAMP qui n'achemine pas n'est pas supposé être capable de localiser une adresse IP au sein d'un paquet, ou de la rendre disponible pour le filtrage, bien qu'il puisse le faire.

Comme le chiffrement de paquet altère la signification des champs chiffrés, le filtrage par correspondance à une propriété doit être configurable à ignorer les paquets chiffrés quand ils sont détectés.

Le processus de choix peut prendre en charge le filtrage sur la base des propriétés de l'état de routeur :
(i) L'interface d'entrée à laquelle le paquet arrive est égale à une valeur spécifiée,
(ii) L'interface de sortie à laquelle le paquet est acheminé est égale à une valeur spécifiée,
(iii) Le paquet a violé une liste de contrôle d"accès (ACL, *Access Control List*) sur le routeur,
(iv) Échec de transmission sur le chemin inverse (RPF, *Reverse Path Forwarding*). Les paquets qui correspondent à la condition d'échec de transmission sur le chemin inverse sont des paquets pour lesquels le filtrage d'entrée a échoué, comme défini dans la [RFC3704].
(v) Échec de protocole de réservation de ressources (RSVP, *Resource Reservation Protocol*). Les paquets qui correspondent à la condition d'échec RSVP sont des paquets qui ne respectent pas la spécification RSVP comme définie dans la [RFC2205].
(vi) Pas de chemin trouvé pour le paquet.
(vii) Le système autonome (AS, *Autonomous System*) du protocole de passerelle frontière (BGP, *Border Gateway Protocol*) [RFC4271] d'origine est égal à une valeur spécifiée, ou se tient dans une gamme donnée.
(viii) L'AS BGP de destination est égal à une valeur spécifiée ou se tient dans une gamme donnée.

Des considérations architecturales de routeur peuvent empêcher que soient disponibles certaines informations concernant le traitement de paquet au débit de ligne pour le choix de paquets. Par exemple, le processus de choix peut n'être pas mis en œuvre dans le chemin rapide qui est capable d'accéder à l'état de routeur au débit de ligne. Cependant, quand le filtrage suit l'échantillonnage (ou une autre opération de sélection) dans un sélecteur composite, le taux de sortie du flux de paquets de l'échantillonneur et le taux d'entrée au filtre peuvent être suffisamment faibles pour que le filtre puisse choisir sur la base de l'état de routeur.

* Choix fondé sur le hachage : le choix fondé sur le hachage va employer une ou plusieurs fonctions de hachage pour être normalisé. Une fonction de hachage est appliquée à un sous ensemble du contenu de paquet, et le paquet est choisi si le hachage résultant tombe dans la gamme spécifiée. Plus la fonction de hachage est forte, plus étroitement le choix fondé sur le hachage s'approche de l'échantillonnage aléatoire uniforme. La confidentialité de la gamme de choix de hachage et des paramètres de la fonction de hachage empêchent la subversion du sélecteur par des paquets qui sont fabriqués soit pour éviter la sélection, soit pour être choisis. La confidentialité de la fonction de hachage n'est pas exigée. Les considérations de robustesse et de sécurité du choix fondé sur le hachage sont discutées plus en détails dans la [RFC5475]. Les applications d'échantillonnage fondé sur le hachage sont décrites à la Section 11.

5.3 Terminologie de fraction de choix

Population : c'est un flux de paquets ou un sous ensemble d'un flux de paquets. Une population peut être considérée comme un ensemble de base à partir duquel les paquets sont choisis. Un exemple est "tous les paquets dans le flux de paquets observés qui sont observés dans un intervalle de temps spécifié".

Taille de population : c'est le nombre total de paquets dans la population.

Taille d'échantillon : c'est le nombre de paquets choisis dans la population par un sélecteur.

Fraction configurée de choix : c'est le ratio attendu de la taille d'échantillon sur la taille de population, telle que fondée sur les paramètres de choix configurés.

Fraction de choix atteinte : c'est le ratio de la taille d'échantillon réelle sur la taille de population. Pour certaines méthodes d'échantillonnage, la fraction de choix atteinte peut différer de la fraction de choix configurée du fait, par exemple, de la variabilité statistique inhérente des décisions d'échantillonnage d'échantillonnage probabiliste et de choix fondé sur le hachage. Néanmoins, pour les grandes tailles de population et des sélecteurs configurés de façon appropriée, la fraction de choix atteinte approche généralement la fraction configurée de choix.

Les notions de fraction de choix configurée/atteinte s'étendent au delà des sélecteurs. Un exemple pour l'illustrer est la fraction de choix configurée de la composition du processus de mesure avec le processus d'exportation. Ici, la population est le flux de paquets observé ou un de ses sous ensembles. La fraction de choix configurée est la fraction de la population pour laquelle les rapports de paquets sont supposés atteindre le collecteur. Cette quantité peut refléter des paramètres supplémentaires, non nécessairement décrits dans le protocole PSAMP, qui déterminent le degré de pertes subi par les rapports de paquet en route vers le collecteur, par exemple, la bande passante de transmission disponible au processus d'exportation. Dans cet exemple, la fraction de choix atteinte est la fraction de la population de paquets pour laquelle les rapports ont réellement atteint le collecteur, et donc incorpore l'effet de toute pertes de rapports de paquet dues, par exemple, à une limitation de ressources au point d'observation ou durant la transmission.

5.4 Numéros de séquence d'entrée pour les sélecteurs primaires

Chaque instance d'un sélecteur primaire doit tenir un compte des paquets présentés à son entrée. La valeur du compteur est à inclure dans un numéro de séquence pour les paquets choisis. Les numéros de séquence sont considérés comme faisant partie de l'état de sélection du paquet.

L'utilisation de l'entrée de numéros de séquence permet aux applications de déterminer la fraction de choix atteinte, et donc de normaliser correctement les estimations d'usage du réseau sans considération des pertes d'informations, ni de si ces pertes se sont produites à cause de l'élimination de rapports de paquets dans le processus de mesure (par exemple, du fait de limitations de ressources chez l'hôte de ces processus) ou de la perte de paquets d'exportation dans la transmission ou la collecte. Voir les détails dans la [RFC3176].

Par exemple, considérons un ensemble de n rapports de paquet consécutifs r1, r2,... , rn, choisis par une opération d'échantillonnage et reçus à un collecteur. Soient s1, s2,..., sn les numéros de séquence d'entrée rapportés par les paquets. La fraction de choix atteinte pour le composé des processus de mesure et d'exportation, en prenant en compte l'échantillonnage de paquet au point d'observation et les pertes lors de la transmission, est calculée comme R = (n-1)/(sn-s1). (Noter que R serait 1 si tous les paquets étaient sélectionnés et si il n'y avait pas de perte de transmission.)

La fraction de choix atteinte peut être utilisée pour estimer le nombre d'octets présents dans une portion du flux de paquets observé. Soit b1, b2,..., bn le nombre d'octets rapportés dans chaque paquet qui atteint le collecteur, et soit B = b1+b2+... +bn. Alors le total des octets présents dans les paquets du flux de paquets observé dont les numéros de séquence d'entrée se tiennent entre s1 et sn est estimé par B/R, c'est-à-dire, réduisant les octets mesurés de la division par la fraction de choix atteinte.

Avec les sélecteurs composites, un numéro de séquence d'entrée doit être rapporté pour chaque sélecteur dans la composition.

5.5 Sélecteurs composites

La capacité de composer des sélecteurs dans un processus de choix devrait être fournie. Les combinaisons suivantes paraissent être les plus utilisées pour les applications :

* Enchaînement de filtres de correspondance de propriété. C'est utile pour la construction du ET des filtres composants.
* Filtrage suivi d'échantillonnage.
* Échantillonnage suivi de filtrage.

Les sélecteurs composites sont utiles pour les applications d'examen. Le premier composant d'un sélecteur composite peut être utilisé pour réduire la charge sur le second composant. Dans cette disposition, l'avantage gagné d'un certain ordre peut dépendre de la composition du flux de paquets.

5.6 Contraintes sur la fraction choisie

L'échantillonnage au plein taux de ligne, c'est-à-dire, avec une probabilité de 1, n'est pas exclue par principe, bien que des contraintes de ressources puissent ne pas le permettre en pratique.

6. Rapport

Cette Section détaille les exigences spécifiques des rapports, motivées par les exigences génériques du paragraphe 3.4.

6.1 Contenu obligatoire des rapports de paquet : rapports de base

Les rapports de paquet doivent inclure ce qui suit :
(i) le ou les numéros de séquence d'entrée de tous les sélecteurs qui ont agi sur le paquet dans l'instance d'un processus de mesure qui a produit le rapport.
(ii) l'identifiant du processus de mesure qui a produit le paquet choisi.

Le processus de mesure doit prendre en charge l'inclusion de ce qui suit dans chaque rapport de paquet, comme option configurable :
(iii) un rapport de base sur le paquet, c'est-à-dire, un certain nombre d'octets contigus depuis le début du paquet, incluant l'en-tête de paquet (qui inclut les en-têtes de couche réseau et tout en-tête d'encapsulation) et des octets suivants de la charge utile de paquet.

Certains appareils peuvent n'avoir pas les capacités de ressources ou de fonctionnalités pour fournir des rapports de paquet plus détaillés que ceux des points (i), (ii), et (iii) ci-dessus. En utilisant cette fonctionnalité minimum de rapport requise, le processus de mesure place la charge de l'interprétation sur le collecteur ou sur les applications qu'il fournit. Certains appareils peuvent avoir la capacité de fournir des rapports de paquet étendus, décrits dans le paragraphe suivant.

6.2 Rapports de paquet étendus

Le processus de mesure peut prendre en charge l'inclusion dans les rapports de paquet des informations suivantes, l'inclusion d'une ou de toutes étant configurable comme une option.

(iv) les champs qui se rapportent aux protocoles suivants utilisés dans le paquet : IPv4, IPV6, protocoles de transport, et protocoles d'encapsulation incluant MPLS.

(v) traitement de paquet, incluant :
- des identifiants pour toutes les interfaces d'entrée et de sortie du point d'observation qui a été traversé par le paquet
- l'AS BGP de source et de destination.

(vi) l'état de sélection associé au paquet, incluant :
- L'horodatage de l'observation du paquet au point d'observation. L'horodatage devrait être rapporté à une résolution de microseconde.
- Les valeurs de hachage, lorsque elles sont calculées.

Il est envisagé que la sélection des champs pour le rapport de paquet étendu puisse être utilisé pour réduire la bande passante de rapport, et dans ce cas, l'option de rapporter des informations en (iii) ne peut pas être effectuée.

6.3 Rapports de paquet étendus en présence de IPFIX

Si un processus de mesure IPFIX est pris en charge au point d'observation, alors pour être conforme à PSAMP, les rapports de paquet étendus doivent être capables d'inclure tous les champs exigés dans le modèle d'information IPFIX [RFC5102], avec les modifications appropriées pour le rapport sur des paquets seuls plutôt que sur des flux.

6.4 Interprétation de rapport

L'interprétation de rapport doit inclure :
(i) les paramètres de configuration des sélecteurs des paquets rapportés ;
(ii) le format du rapport de paquet ;
(iii) l'indication de la précision inhérente des quantités rapportées, par exemple, de l'horodatage du paquet.

La mesure de la précision en (iii) est d'une importance fondamentale pour estimer l'erreur probable attachée aux estimations

formées à partir des rapports de paquet par les applications.

Les exigences de robustesse et de transparence sont le motif de l'inclusion de l'interprétation de rapport dans le flux de rapports : elle rendent le flux de rapports auto-défini. Le cadre PSAMP exclut de s'appuyer sur un modèle de remplacement dans lequel l'interprétation serait récupérée hors bande. Cette dernière approche n'est pas robuste à l'égard de changements non documentés de la configuration de sélecteur, et peut donner lieu à de futurs problèmes architecturaux pour que les systèmes de gestion de réseau gèrent de façon cohérente la configuration et la collecte des données.

Il n'est pas envisagé que toute l'interprétation de rapport soit incluse dans chaque rapport de paquet. Beaucoup des quantités mentionnées ci-dessus sont supposées être relativement statiques ; elles pourraient être communiquées périodiquement et lorsque elles changent.

7. Processus de mesure en parallèle

À cause du nombre croissant d'applications de mesure distinctes avec des exigences variables, il est souhaitable d'établir un processus de mesure parallèle sur un flux de paquets observé donné. Un appareil capable d'héberger un processus de mesure devrait être capable de prendre en charge simultanément plus d'un processus de mesure configurable indépendamment. Chacun de ces processus de mesure devrait avoir l'option d'être équipé avec son propre processus d'exportation ; autrement, le processus de mesure parallèle peut partager le même processus d'exportation.

Chacun des processus de mesure parallèle devrait être indépendant. Cependant, des contraintes de ressources peuvent empêcher un rapport complet sur un paquet choisi par plusieurs processus de choix. Dans ce cas, le rapport pour le paquet doit être complet pour au moins un processus de mesure ; les autres processus de mesure ont seulement besoin d'enregistrer qu'ils ont choisi le paquet, par exemple, en incrémentant un compteur. La priorité parmi les processus de mesure sous des contraintes de ressources devrait être configurable.

Il n'est pas proposé de normaliser le nombre de processus de mesure parallèles.

8. Processus d'exportation

Cette section détaille les exigences spécifiques pour le processus d'exportation, motivées par les exigences génériques du paragraphe 3.6.

8.1 Utilisation de IPFIX

PSAMP va utiliser le protocole d'exportation d'informations de flux IP (IPFIX, *IP Flow Information Export*) pour exporter les flux de rapports. Le protocole IPFIX convient bien pour cela, parce que l'architecture IPFIX correspond très bien à l'architecture de PSAMP et que les moyens fournis par le protocole IPFIX sont suffisants pour les besoins de PSAMP. Par ailleurs, toutes les caractéristiques du protocole IPFIX n'ont pas besoin d'être mises en œuvre par certains appareils PSAMP. Par exemple, un appareil qui offre seulement l'échantillonnage indépendant du contenu et le rapport de base PSAMP n'a pas besoin de prendre en charge les capacités IPFIX fondées sur les champs de paquet.

8.2 Paquets d'exportation

Les paquets d'exportation peuvent contenir un ou plusieurs rapports de paquet, et/ou interprétations de rapport. Les paquets d'exportation doivent aussi contenir :
(i) un identifiant pour le processus d'exportation,
(ii) un numéro de séquence de paquet d'exportation.

Un numéro de séquence de paquet d'exportation permet au collecteur d'identifier les pertes de paquets d'exportation en transit. Noter que certains protocoles de transport, par exemple, UDP, ne fournissent pas de numéro de séquence. De plus, avoir des numéros de séquence disponibles au niveau de l'application permet au collecteur de calculer le taux de perte de paquets pour s'en servir, par exemple, à l'estimation des volumes de trafic original provenant des paquets d'exportation qui atteignent le collecteur.

8.3 Transport non fiable sensible à l'encombrement

L'exportation des flux de rapports n'exige pas une exportation fiable. Le paragraphe 5.4 montre que l'utilisation de numéros de séquence d'entrée dans les sélecteurs de paquet signifie que la capacité d'estimer les taux de trafic n'est pas entravée par les pertes d'exportation. Les pertes de paquet d'exportation deviennent une autre forme d'échantillonnage, bien que moins désirable, et moins contrôlée.

De plus, la retransmission des paquets d'exportation perdus consomme des ressources réseau supplémentaires. L'exigence de mémoriser des données non acquittées est un inconvénient pour avoir une prise en charge généralisée de PSAMP.

Afin de satisfaire à la fois les exigences d'opportunité et d'évitement d'encombrement du paragraphe 4.3, un protocole de transport non fiable sensible à l'encombrement peut être utilisé. IPFIX est compatible avec cette exigence, car il rend obligatoire la prise en charge du protocole de transmission de commandes de flux (SCTP, *Stream Control Transmission Protocol*) [RFC4960] et l'extension SCTP de fiabilité partielle [RFC3758].

IPFIX permet aussi l'utilisation du protocole de datagramme d'utilisateur (UDP, *User Datagram Protocol*) [RFC0768], bien que ce ne soit pas un protocole sensible à l'encombrement. Cependant, dans ce cas, les paquets d'exportation doivent rester entièrement dans le domaine administratif des opérateurs [RFC5101]. Le processus d'exportation PSAMP est équipé d'une limite de taux d'exportation configurable (voir le paragraphe 8.4) qui peut être utilisée pour limiter le taux d'exportation quand un protocole de transport sensible à l'encombrement n'est pas utilisé. Le collecteur, quand il détecte la perte d'un paquet d'exportation du fait de numéros de séquence d'export manquants, peut reconfigurer la limite de taux d'exportation afin d'éviter l'encombrement.

8.4 Limite de taux d'exportation configurable

Le processus d'exportation doit avoir une limite de taux d'exportation, configurable par processus d'exportation. C'est utile pour deux raisons :

(i) Même sans encombrement du réseau, le taux de choix de paquet peut excéder la capacité du collecteur de traiter les rapports, en particulier quand de nombreux processus d'exportation alimentent un collecteur commun. L'utilisation d'une limite de taux d'exportation permet le contrôle du taux d'entrée global au collecteur.

(ii) IPFIX fournit l'exportation en utilisant UDP comme protocole de transport dans certaines circonstances. Une limite de taux d'exportation permet de contrôler le taux d'exportation pour correspondre à la fois à la vitesse des liaisons du chemin et aux capacités du collecteur.

8.5 Limitation de délai pour les paquets d'exportation

Une faible latence de mesure permet au système de surveillance du trafic de mieux répondre aux événements de réseau en temps réel, par exemple, pour identifier rapidement les sources d'encombrement. L'opportunité est généralement une bonne chose pour les appareils qui effectuent l'échantillonnage car elle minimise la quantité de mémoire nécessaire pour mettre les échantillons en mémoire tampon.

Maintenir faible le délai d'expédition de paquet a d'autres avantages que de limiter les exigences de mémoire tampon. Pour de nombreuses applications, une résolution de 1 seconde est suffisante. Les applications de cette catégorie pourraient inclure d'identifier les sources associées à l'encombrement, de retracer les attaques de déni de service (DoS) à travers le réseau, et de construire des matrices de trafic. De plus, maintenir le délai d'expédition dans la résolution requise par les applications élimine le besoin d'horodater par des horloges synchronisées aux points d'observation, ou pour les points d'observation et collecteurs de maintenir une communication bidirectionnelle afin de saisir les décalages d'horloges. Le collecteur peut simplement traiter les rapports de paquet dans l'ordre où ils sont reçus, utilisant sa propre horloge comme base de temps "globale". Cela évite la complexité de la mise en mémoire tampon et du réarrangement des échantillons. Voir un exemple dans [DuGeGr02].

Le délai entre l'observation d'un paquet et la transmission d'un paquet d'exportation contenant un rapport sur ce paquet a plusieurs composants. Il est difficile de normaliser une exigence de délai numérique donnée, car en pratique le délai peut être sensible à la charge du processeur au point d'observation. Donc, PSAMP vise à contrôler la portion du délai au sein du point d'observation qui est due à la mise en mémoire tampon dans la formation et la transmission des paquets d'exportation.

Afin de limiter le délai dans la formation des paquets d'exportation, le processus d'exportation doit fournir la capacité de clore et mettre en file d'attente de transmission tout paquet d'exportation durant la formation aussitôt qu'il inclut un rapport

de paquet.

Afin de limiter le délai de transmission des paquets d'exportation, une limite supérieure configurable du délai d'un paquet d'exportation avant transmission doit être fourni. Si la limite est franchie, le paquet d'exportation est éliminé. Cette fonction peut être fournie par le service de fiabilité temporelle de l'extension SCTP de fiabilité partielle [RFC3758].

Le processus d'exportation peut mettre en file d'attente le flux de rapports afin d'exporter plusieurs rapports de paquet dans un seul paquet d'exportation. Tout délai conséquent doit quand même permettre la disponibilité en temps utile des rapports de paquet qu'on vient de décrire. Le service de fiabilité temporelle de l'extension SCTP de fiabilité partielle [RFC3758] permet d'éliminer des paquets de la mémoire tampon d'exportation quand leur âge dans la mémoire tampon excède une limite configurable. Une valeur par défaut convenable pour la limite devrait être utilisée afin d'éviter un faible taux de transmission dû à une mauvaise configuration.

8.6 Compression de paquet d'exportation

Pour conserver la bande passante du réseau et les ressources chez le collecteur, les paquets d'exportation peuvent être compressés avant l'exportation. La compression est supposée être assez efficace car les paquets choisis peuvent avoir de nombreux champs en commun, par exemple, si un filtre se concentre sur des paquets avec certaines valeurs dans des champs d'en-tête particuliers. En utilisant la compression, on pourrait cependant impacter l'opportunité des rapports de paquet. Un retard conséquent ne doit pas violer l'exigence d'opportunité pour la disponibilité des rapports de paquet chez le collecteur.

8.7 Collecteur de destination

Quand on exporte à un collecteur distant, celui-ci est identifié par l'adresse IP, le protocole de transport, et le numéro d'accès du transport.

8.8 Exportation locale

Le flux de rapports peut être directement exporté à des applications embarquées fondées sur la mesure, par exemple, celles qui forment des statistiques composites pour plus d'un paquet. L'exportation locale peut être présentée à travers une interface directement à des applications de niveau supérieur, c'est-à-dire, à travers une API, plutôt que d'employer le transport utilisé pour l'exportation au loin. La spécification d'une telle API sort du domaine du cadre PSAMP.

Un exemple possible d'exportation locale pourrait être des paquets choisis par le processus de mesure PSAMP qui servent d'entrée au protocole IPFIX, qui forme alors des enregistrements de flux à partir du flux des paquets choisis.

9. Configuration et gestion

Une exigence clé pour PSAMP est la facilité de reconfiguration des paramètres du processus de mesure, incluant ceux pour le choix et les rapports de paquet, et du processus d'exportation. Un important exemple est de prendre en charge les applications fondées sur la mesure qui veulent examiner de façon adaptative les détails du trafic en temps réel.

Pour faciliter la restitution et la surveillance des paramètres, ils doivent résider dans une base de données d'informations de gestion (MIB, *Management Information Base*). Les objets de surveillance obligatoires vont couvrir toutes les fonctions obligatoires de PSAMP. L'alarme sur des paramètres spécifiques pourrait être déclenchée avec des mécanismes de seuil comme les événements et alarmes de la surveillance de réseau à distance (RMON, R*emote Network Monitoring*) [RFC2819] ou la MIB d'événements [RFC2981].

Pour configurer les paramètres du processus de mesure, plusieurs solutions de remplacement sont disponibles, incluant un module de MIB avec des objets inscriptibles, ainsi que d'autres protocoles de configuration. Pour configurer les paramètres du processus d'exportation, le rapport de paquet, et l'interprétation de rapport, qui sont des tâches de IFPIX, la ou les méthodes de configuration IPFIX devraient être utilisées.

Bien que la gestion et la configuration des collecteurs sortent de notre domaine d'application, un appareil PSAMP, dans la mesure où il emploie IPFIX comme protocole d'exportation, hérite de IPFIX la capacité de détecter et récupérer d'une défaillance du collecteur ; voir le paragraphe 8.2 de la [RFC5470].

10. Faisabilité et complexité

Afin que PSAMP soit pris en charge sur le spectre entier des équipements de réseautage, il doit être simple et peut coûteux à mettre en œuvre. On peut envisages des instances faciles à mettre en œuvre des mécanismes décrits dans le présent document. Donc, pour ce sous ensemble d'instances, il devrait être facile pour virtuellement tous les fabricants de système de les inclure dans leurs produits. Bien sûr, les opérations d'échantillonnage et de filtrage sont déjà réalisées dans les équipements disponibles.

On donne ici des arguments spécifiques pour démontrer la faisabilité et commenter sur la complexité des mises en œuvre de matériels. On souligne ici que l'objet de ces arguments n'est pas de favoriser ou recommander une mise en œuvre particulière, ni de suggérer une voie pour la normalisation, mais plutôt de montrer que l'ensemble des mises en œuvre possibles n'est pas vide.

10.1 Faisabilité

10.1.1 Filtrage

Le filtrage consiste en un petit nombre d'opérations de gabarits (logiques au bit près) de comparaison, et de gammes (plus grand que). La mise en œuvre d'au moins un petit nombre de ces opérations est directe. Par exemple, les filtres pour les listes de contrôle d'accès de sécurité (ACL, *Access Control List*) sont largement mis en œuvre. Cela pourrait être aussi simple qu'une correspondance exacte sur certains champs, ou impliquer de plus complexes comparaisons et gammes.

10.1.2 Échantillonnage

L'échantillonnage fondé sur des compteurs (réglage de compteur, décrément, essai d'égalité à zéro) ou correspondance de gamme sur le hachage d'un paquet (plus grand que) est possible avec un petit nombre de sélecteurs, bien qu'il puisse y avoir des différences dans la facilité de mise en œuvre pour les plate-formes de matériels par rapport à celles de logiciel.

10.1.3 Hachage

La complexité des fonctions de hachage est très varaible. L'exécution d'un petit nombre de fonctions de hachage suffisamment simples peut être mise en œuvre au taux de ligne. Concernant l'entrée à la fonction de hachage, les champs d'en-tête IP invariants selon le bond (adresse IP, identification IP) et champs d'en-tête TCP/UDP (numéros d'accès, numéros de séquence TCP) tirés des 40 premiers octets du paquet se sont trouvés posséder une variabilité considérable ; voir [DuGr01].

10.1.4 Rapport

Le plus simple rapport de paquet va dupliquer les n premiers octets du paquet. Cependant, un tel format non compressé peut taxer la bande passante disponible au processus d'exportation pour les forts taux d'échantillonnage ; rapporter des champs choisis économiserait cette bande passante. Donc, il y a un compromis à faire entre simplicité et limitations de bande passante.

10.1.5 Exportation

La facilité d'exportation des paquets d'exportation dépend de l'architecture du système. La plupart des systèmes devraient être capables de prendre en charge l'exportation en insérant les paquets d'exportation, même à travers le chemin logiciel.

10.2 Complexité potentielle du matériel

Réaliser de faibles constantes pour les performances tout en minimisant les ressources matérielles est, bien sûr, un défi, en particulier à de très hautes fréquences d'horloge. La plupart des sélecteurs sont cependant très basiques et leurs mises en œuvre très bien comprises ; en fait, le concepteur moyen de circuit intégré spécifique d'application (ASIC, *Application-Specific Integrated Circuit*) utilise simplement des instances de bibliothèque incorporées de ces opérations plutôt que de les

concevoir à partir de zéro. De plus, l'équipement de réseautage n'a généralement pas besoin de fonctionner aux plus hauts débits d'horloge, réduisant encore l'effort requis pour obtenir des mises en œuvre raisonnablement efficaces.

Les opérations au bit près simples sont faciles à mettre en œuvre dans le matériel. De telles opérations (NET/NOU/XNOU) se traduisent directement en portes à quatre transistors. Chaque bit d'une opération logique à plusieurs bits est complètement indépendante et peut donc être effectuée en parallèle, ne causant aucun coût additionnel de performances au delà d'une opération à un seul bit.

Les comparaisons (EQ/NEQ) prennent O(log(M)) étapes de logique, où M est le nombre de bits impliqués dans la comparaison. Le log(M) est obligé d'accumuler le résultat dans un seul bit.

Les opérations "plus grand que", comme utilisées pour déterminer si un hachage tombe dans une gamme de sélection, sont une détermination du bit de plus fort poids non équivalent dans les deux opérandes. L'opérande avec le bit de plus fort poids non égal établi à un est plus grand que l'autre.

Donc, une opération "plus grand que" est aussi une opération à O(log(M)) étapes de logique. Les mises en œuvre optimisées d'opérations arithmétiques sont aussi O(log(M)) du fait de la propagation du bit de report.

Le réglage d'un compteur est simplement le chargement d'un registre avec un état. Une telle opération est un O(1) simple et rapide. Incrémenter ou décrémenter un compteur est une lecture, suivie par une opération arithmétique, suivie d'une mémorisation. Rendre le registre à double accès prend de l'espace supplémentaire, mais c'est une technique bien comprise. Donc, l'incrément/décrément est aussi une opération O(log(M)).

Les fonctions de hachage prennent diverses formes. Le calcul impliqué, par exemple, dans une vérification de redondance cyclique (CRC, *Cyclic Redundancy Check*) standard est essentiellement un ensemble d'opérations OUX, où le résultat intermédiaire est mémorisé et OUXé avec le prochain tronçon de données. Il y a seulement O(1) opérations et aucune opération de complexité logarithmique. Donc, une simple fonction de hachage, comme un CRC ou ses généralisations, peut être mise en œuvre très efficacement dans le matériel.

À l'autre extrémité de la gamme de complexité, la fonction MD5 utilise un grand nombre d'opérations conditionnelles au bit près et d'opérations arithmétiques. Les premières sont des opérations O(1) et les dernière des opérations O(log(M)). MD5 spécifie 256 opérations ADD de 32 bits par 16 octets d'entrée traités. Si on considère le traitement de 10 Gbit/s à 100 MHz (ce taux de traitement est actuellement disponible). Cela exige de traiter 12,5 octets/cycle, et donc au moins 200 additionneurs, un nombre plutôt grand. À cause de la dépendance des données au sein de l'algorithme MD5, les additionneurs ne peuvent pas être simplement effectués en parallèle, exigeant donc soit des débits d'horloge plus rapides, soit des architectures plus évoluées. Donc, le choix de fonctions de hachage aussi complexes que MD5 peut être empêché pour une utilisation universelle au plein débit de ligne. Cela motive d'explorer l'utilisation de fonctions de hachage de sélection d'une complexité intermédiaire entre celle de MD5 et du CRC. Dans certaines applications (voir la Section 11) un second hachage peut être calculé sur seulement les paquets choisis ; MD5 est faisable pour cela si le taux de production des paquets choisis est suffisamment faible.

11. Applications

On décrit d'abord plusieurs applications opérationnelles représentatives qui exigent des mesures de trafic à divers niveaux de granularité temporelle et spatiale. Certains des buts apparaissent ici similaires à ceux de IPFIX, au moins dans les grandes classes d'applications prises en charge. L'avantage majeur de PSAMP est la prise en charge de nouvelles applications de gestion de réseau, spécifiquement, celles permises par les sélecteurs de paquet qu'il supporte.

11.1 Mesure de base et prélèvement

L'échantillonnage de paquet convient idéalement pour déterminer la composition du trafic à travers un réseau. L'approche est de permettre la mesure sur un ensemble limité des liaisons du réseau afin que chaque paquet entrant dans le réseau soit vu au moins une fois, par exemple, sur toutes les liaisons d'entrée. L'échantillonnage non filtré avec le choix d'une fraction relativement faible établit les mesures de base du trafic du réseau. Les rapports de paquet incluent des attributs de paquet d'intérêt commun : adresse de source et de destination et numéros d'accès, préfixe, numéro de protocole, type de service, etc. Les matrices de trafic sont indiquées en rapportant les matrices d'AS de source et destination. Les volumes absolus de trafic sont estimés en renormalisant les volumes de trafic échantillonné en le divisant par la fraction de choix configuré ou par la fraction de choix atteinte (comme déduit des compteurs d'entrée de paquets inclus dans le flux de rapports).

Supposons qu'un opérateur ou une application fondée sur la mesure détecte un sous ensemble intéressant d'un flux de paquets, identifié par un attribut de paquet particulier. L'examen en temps réel de ce sous ensemble est réalisé par l'instanciation d'un nouveau processus de mesure sur le même flux de paquets observé d'où le sous ensemble a été rapporté. Le processus de choix du nouveau processus de mesure filtre en accord avec l'attribut intéressant, et compose avec l'échantillonnage si nécessaire pour gérer la fraction de paquets choisis atteinte.

11.2 Échantillonnage de trajectoire

Le but de l'échantillonnage de trajectoire est de choisir un sous ensemble de paquets à tous les points d'observation activés auxquels ces paquets sont observés dans un domaine de réseau. Donc, les décisions de choix sont cohérentes dans le sens où chaque paquet est choisi soit à tous les points d'observation activés, soit à aucun. L'échantillonnage de trajectoire est réalisé par le choix fondé sur le hachage si tous les points d'observation activés appliquent une fonction de hachage commune à une portion du contenu de paquet invariante le long du chemin du paquet. (Donc, les champs comme le TTL et le CRC sont exclus.)

La trajectoire suivie par un paquet est reconstruite à partir des rapports de paquet sur lui qui atteignent le collecteur. Les rapports sur un paquet donné sont associés en correspondant à une étiquette comprenant le contenu de paquet invariant rapporté ou éventuellement un résumé de celui-ci. La reconstruction des trajectoires et les méthodes pour traiter les possibles ambiguïtés dues aux collisions d'étiquettes (étiquettes identiques rapportées par des paquets différents) et aux pertes potentielles de rapports dans la transmission sont traitées dans [DuGr01], [DuGeGr02], et [DuGr04].

11.3 Mesure passive des performances

L'échantillonnage de trajectoire permet de retracer les performances du trafic d'abonnés, abonnés identifiés par une liste de préfixes de source ou de destination, ou par des interfaces d'entrée ou de sortie. Les utilisations opérationnelles incluent la vérification des accords de niveau de service (SLA, *Service Level Agreement*) et les réparations, suite à la plainte d'un abonné.

Dans cette application, l'échantillonnage de trajectoire est activé à toutes les interfaces d'entrée et de sortie du réseau. Les taux de pertes dans le transit entre l'entrée et la sortie sont estimées à partir de la proportion de trajectoires pour lesquelles aucun rapport de sortie n'est reçu. Noter que la perte de paquets d'abonnés se distingue de la perte des rapports de paquet par l'utilisation des numéros de séquence des rapports. En supposant la synchronisation des horloges entre les différentes entités, le délai du trafic d'abonné à travers le réseau peut aussi être mesuré ; voir [Zs02].

Étendre la sélection de hachage à toutes les interfaces dans le réseau permettrait l'attribution des mauvaises performances aux liaisons individuelles de réseau.

11.4 Réparation

Les rapports de paquet PSAMP peuvent aussi être utilisés pour diagnostiquer des problèmes dont l'occurrence est évidente d'après des statistiques agrégées, par utilisation d'interface et des statistiques de perte de paquet. Ces statistiques sont généralement des moyennes mobiles sur des périodes de temps relativement longues, par exemple, 5 minutes, et servent d'indication grossière de la santé opérationnelle du réseau. La méthode la plus courante pour obtenir de telles mesures est par les MIB SNMP appropriées (MIB-II [RFC1213] et MIB spécifiques de fabricant).

Supposons qu'un opérateur détecte une liaison qui est constamment surchargée et subit des taux d'élimination de paquets significatifs. Il y a une large gamme de causes potentielles : paramètres d'acheminement (par exemple, pondération de liaison OSPF) mal adaptés à la matrice de trafic, par exemple, à cause d'un glissement dans cette matrice ; une attaque de DoS, une poussée soudaine de trafic, ou un problème d'acheminement (oscillation de liaison). Dans la plupart des cas, des statistiques de liaison agrégées ne sont pas suffisantes pour distinguer entre de telles causes et pour décider d'une action corrective appropriée. Par exemple, si l'acheminement sur deux liaisons est instable, et si les liaisons oscillent entre surcharge et inactivité, cela pourrait être moyenné sur une fenêtre de 5 minutes, indiquant une charge modérée sur les deux liaisons.

La mesure de base de PSAMP de l'encombrement de liaison décrite au paragraphe 11.1, permet des mesures fines dans l'espace et le temps. L'opérateur doit être capable de déterminer combien d'octets/paquets sont générés pour chaque adresse de source/destination, numéro, d'accès et préfixe, ou autres attributs, comme un numéro de protocole, une classe d'équivalence de transmission (FEC, *Forwarding Equivalence Class*) MPLS, type de service, etc. Cela permet la

détermination précise de la nature du trafic en cause. Par exemple, dans le cas d'une attaque de déni de service répartie (DDoS, *Distributed Denial of Service*) l'opérateur va voir une fraction significative du trafic avec une adresse de destination identique.

Dans certaines circonstances, des informations précises sur le flux spatial de trafic à travers le domaine de réseau sont requises pour détecter et diagnostiquer les problèmes et vérifier le comportement correct du réseau. Dans le cas d'une liaison surchargée, il serait très utile de connaître l'ensemble précis de chemins que suivent les paquets qui traversent cette liaison. Cela révélerait directement un problème d'acheminement tel qu'une boucle, ou une liaison avec une pondération mal configurée. Plus généralement, les scénarios de diagnostic complexes peuvent bénéficier de mesures des intensités de trafic (et autres attributs) sur un ensemble de chemins qui est contraint d'une certaine façon. Par exemple, si un consommateur multi rattachements se plaint de problèmes de performances sur une des liaisons d'accès provenant d'un préfixe d'adresses de source particulier, l'opérateur devrait être capable d'examiner en détails le trafic provenant de ce préfixe de source qui traverse aussi la liaison d'accès spécifiée vers le consommateur.

Bien qu'il soit en principe possible d'obtenir le flux spatial de trafic par des informations d'état de réseau auxiliaire, par exemple, en téléchargeant les tableaux d'acheminement et de transmission provenant des routeurs, ces informations sont souvent peu fiables, périmées, volumineuses, et dépendantes d'un modèle de réseau. Pour les besoins du fonctionnement, une observation directe des flux de trafic fournie par l'échantillonnage de trajectoire est plus fiable, car elle ne dépend pas de ces informations auxiliaires. Par exemple, si il y a une faute dans le logiciel d'un routeur, l'observation directe va permettre de diagnostiquer l'effet de cette faute, alors qu'une méthode indirecte ne le permettrait pas.

12. Considérations sur la sécurité

12.1 Relation de la sécurité de PSAMP et de IPFIX pour le processus d'exportation

Comme précisé au paragraphe 4.3, PSAMP partage avec IPFIX les exigences de sécurité pour l'exportation, à savoir la confidentialité, l'intégrité, et l'authenticité des données exportées ; voir aussi le paragraphe 6.3 et la Section 10 de la [RFC3917]. Comme PSAMP va utiliser IPFIX pour l'exportation, il peut employer le protocole IPFIX [RFC5101] pour satisfaire ses exigences.

12.2 Considérations de confidentialité spécifiques de PSAMP

À la différence de IPFIX, un appareil PSAMP peut, dans certaines configurations, rapporter un certain nombre d'octets initiaux du paquet, qui peuvent inclure une partie d'une charge utile de paquet. Cette option est conforme aux exigences de la [RFC2804] car elle ne rend pas obligatoire des configurations qui permettraient de capturer un flux de paquets entier d'un flux : ni un taux d'échantillonnage unitaire (échantillonnage de un sur un) ni le rapport d'un nombre spécifique d'octets initiaux n'est exigé du protocole PSAMP.

Pour préserver la confidentialité de tout utilisateur agissant comme envoyeur ou receveur du trafic observé, le contenu des rapports de paquet doit être capable de rester confidentiel dans le transit entre l'appareil PSAMP exportateur et le collecteur. PSAMP va utiliser IPFIX comme protocole d'exportation, et le protocole IPFIX doit fournir des mécanismes pour assurer la confidentialité du processus d'exportation, par exemple, le chiffrement des paquets d'exportation [RFC5101].

12.3 Considérations de confidentialité pour le choix fondé sur le hachage

12.3.1 Modes et impact des vulnérabilités

Un problème du choix fondé sur le hachage est si un grand ensemble de paquets en relation pourrait être échantillonné de façon disproportionnée, soit
(i) par un comportement imprévu dans la fonction de hachage, ou
(ii) parce que les paquets ont été délibérément fabriqués pour avoir cette propriété.

Comme précisé ci-dessous, seules les fonctions de hachage cryptographiques (par exemple, celles fondées sur MD5) qui emploient un paramètre privé sont suffisamment fortes pour résister à la gamme des attaques concevables. Cependant, des considérations de mise en œuvre peuvent empêcher de mettre en œuvre les plus fortes fonctions de hachage au taux de ligne. Pour cette raison, PSAMP n'est pas supposé normaliser une fonction de hachage cryptographique pour le moment. L'objet de cette section est d'informer des vulnérabilités et des compromis associés aux différents choix de fonction de hachage. Le paragraphe 6.2.2 de la [RFC5475] le fait plus en détails.

Un attaquant capable de prédire les résultats d'un échantillonnage de paquet pourrait fabriquer un flux de paquets qui éviterait la sélection, ou un autre qui pourrait surcharger l'infrastructure de mesure en faisant que tous ses paquets soient choisis. Un attaquant peut tenter de faire cela sur la base de la connaissance de la fonction de hachage. Un attaquant pourrait employer la connaissance des résultats de la sélection d'un flux de paquets connu pour inverser les paramètres de la fonction de hachage. Cette connaissance pourrait être obtenue, par exemple, des informations de facturation, des réactions des systèmes de détection d'intrusion, ou de l'observation d'un flux de rapports.

Comme le choix fondé sur le hachage est déterministe, il est vulnérable aux attaques en répétition. La répétition d'un seul paquet peut être remarquée par d'autres méthodes de mesure si il en est d'employées (par exemple, la collecte des statistiques de flux) tandis qu'un ensemble de paquets distincts qui paraissent statistiquement similaires à du trafic régulier peut être moins remarquable. L'impact des attaques en répétition sur le choix fondé sur le hachage peut être atténué en changeant de façon répétée les paramètres de la fonction de hachage.

12.3.2 Utilisation de paramètres privés dans les fonctions de hachage

Parce que les fonctions de hachage pour le choix fondé sur le hachage vont être normalisées et donc publiques, la décision de choix de paquet doit être contrôlée par une quantité privée associée au sélecteur de choix fondé sur le hachage. Rendre confidentielle la gamme de valeurs de hachage pour laquelle les paquets sont choisis n'est pas par soi-même suffisant pour empêcher un attaquant de construire un flux de paquets distincts qui vont être choisis de façon disproportionnée. Un paramètre confidentiel doit être utilisé au sein de la fonction de hachage, par exemple, un module confidentiel dans une fonction de hachage, ou en enchaînant l'entrée de hachage avec une chaîne secrète avant le hachage.

12.3.3 Force des fonctions de hachage

Le choix spécifique de la fonction de hachage et son usage déterminent les types de vulnérabilités potentielles :

* Fonctions de hachage cryptographiques : quand un paramètre privé est utilisé, les futurs résultats de sélection ne peuvent pas être prédits même par un attaquant qui connaît les résultats des sélections passées.

* Fonctions de hachage non cryptographiques : utilisation de la connaissance des résultats passés de sélection : des fonctions de hachage bien connues, par exemple, CRC-32, sont vulnérables aux attaques, parce que leur paramètre privé peut être déterminé en connaissant suffisamment de sélections passées, même quand un paramètre privé est utilisé ; voir [GoRe07].

Aucune connaissance des résultats des sélections passées : l'utilisation d'un paramètre privé rend plus résistante la fonction de hachage aux classes d'attaques qui fonctionnent quand le paramètre est public, bien que la vulnérabilité à de futures attaques ne soit pas exclue.

12.4 Lignes directrices pour la sécurité de la configuration de PSAMP

Les paramètres de fonction de hachage configurés dans un appareil PSAMP sont des informations sensibles, qui doivent être gardées secrètes. Tout comme l'utilisation de techniques de sondage pour découvrir les paramètres des fonctions de hachage non cryptographiques décrites ci-dessus, les faiblesses de mise en œuvre et de procédure peuvent conduire des attaquants à découvrir les paramètres, quelle que soit la classe de fonction de hachage utilisée. Les mesures suivantes peuvent empêcher cela de se produire :

Les paramètres de fonction de hachage ne doivent pas être affichables en clair sur les appareils PSAMP. Cela réduit les chances que les paramètres soient découverts par un accès non autorisé à l'appareil PSAMP.

Les paramètres de fonction de hachage ne doivent pas être établis à distance en clair sur un canal qui peut être espionné.

Les paramètres de fonction de hachage doivent être changés régulièrement. Noter que ces changements doivent être synchronisés sur tous les appareils PSAMP dans un domaine dans lequel l'échantillonnage de trajectoire est employé afin de conserver un échantillonnage de paquets cohérent sur le domaine.

Des valeurs de paramètre de fonction de hachage par défaut devraient être initialisées au hasard, afin d'éviter des valeurs prévisibles que des attaquants pourraient exploiter.

13. Contributeurs

Sharon Goldberg a contribué au paragraphe 12.3 sur les considérations sur la sécurité du choix fondé sur le hachage.

Sharon Goldberg
Department of Electrical Engineering
Princeton University
USA
mél : goldbe@princeton.edu

14. Remerciements

Les auteurs tiennent à remercier Peram Marimuthu et Ganesh Sadasivan de leurs apports dans les premiers projets du présent document.

15. Références

15.1 Références normatives

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

[RFC3758] R. Stewart et autres, "Extension de fiabilité partielle du protocole de transmission de contrôle de flux (SCTP)", mai 2004. *(P.S.)*

[RFC4960] R. Stewart, éd., "Protocole de transmission de commandes de flux (SCTP)", septembre 2007. (*Remplace* RFC2960, RFC3309 ; *P.S. ; Remplacée par* RFC9260)

[RFC5101] B. Claise, éd., "Spécification du protocole d'exportation d'informations de flux IP (IPFIX) pour l'échange d'informations de flux de trafic IP", janvier 2008. *(P.S.) (Obsolète, voir* RFC7011, STD77)

[RFC5102] J. Quittek et autres, "Modèle d'informations pour l'exportation d'informations de flux IP", janvier 2008. *(P.S.) (Remplacée par* RFC7012)

[RFC5475] T. Zseby et autres, "Techniques d'échantillonnage et de filtrage pour sélection de paquet IP", mars 2009. *(P.S.)*

[RFC5476] B. Claise et autres "Spécification du protocole d'échantillonnage de paquet (PSAMP)", mars 2009. *(P.S.)*

[RFC5477] T. Dietz et autres, "Modèle d'information pour l'exportation d'échantillonnage de paquet", mars 2009. *(P.S.)*

15.2 Références pour.information

[DuGeGr02] N. G. Duffield, A. Gerber, M. Grossglauser, "Trajectory Engine: A Backend for Trajectory Sampling", IEEE Network Operations and Management Symposium 2002, Florence, Italie, 15-19 avril 2002.

[DuGr04] N. G. Duffield and M. Grossglauser, "Trajectory Sampling with Unreliable Reporting", Proc IEEE Infocom 2004, Hong Kong, mars 2004.

[DuGr08] N. G. Duffield and M. Grossglauser, "Trajectory Sampling with Unreliable Reporting", IEEE/ACM Trans. on Networking, 16(1), février 2008.

[GoRe07] S. Goldberg, J. Rexford, "Security Vulnerabilities and Solutions for Packet Sampling", IEEE Sarnoff Symposium, Princeton, NJ, mai 2007.

[RFC0768] J. Postel, "Protocole de datagramme d'utilisateur (UDP)", (STD 6), 28 août 1980.

[RFC1213] K. McCloghrie et M. Rose, "Base de données d'informations de gestion pour la gestion de réseau des internets fondés sur TCP/IP : MIB-II", STD 17, mars 1991.

[RFC2205] R. Braden, éd., L. Zhang, S. Berson, S. Herzog, S. Jamin, "Protocole de réservation de ressource (RSVP) -- version 1, spécification fonctionnelle", septembre 1997. (*MàJ par* RFC2750, 3936, 4495, 6780)) *(P.S.)*

[RFC2330] V. Paxson, G. Almes, J. Mahdavi, M. Mathis, "Cadre pour la mesure des performances d'IP", mai 1998. *(Information ; MàJ par* RFC8468, RFC9198)

[RFC2460] S. Deering et R. Hinden, "Spécification du protocole Internet, version 6 (IPv6)", décembre 1998. *(MàJ par* 5095, 6564 *; D.S ; Remplacée par* RFC8200, STD 86*)*

[RFC2804] IAB, IESG, "Politique de l'IETF en matière d'écoutes", mai 2000. *(Information)*

[RFC2819] S. Waldbusser, "Base de données d'informations de gestion de surveillance à distance de réseau", mai 2000. (STD0059)

[RFC2914] S. Floyd, "Principes du contrôle d'encombrement", BCP 41, DOI 10.17487/RFC2914, septembre 2000.

[RFC2981] R. Kavasseri, éd., "MIB d'événement", octobre 2000. *(P.S.)*

[RFC3031] E. Rosen, A. Viswanathan, R. Callon, "Architecture de commutation d'étiquettes multi protocoles", janvier 2001. *(P.S.)* (*MàJ par la* RFC6790)

[RFC3176] P. Phaal, S. Panchen, N. McKee, "sFlow de InMon Corporation : méthode de surveillance du trafic dans les réseaux commutés et routés", septembre 2001. *(Information)*

[RFC3704] F. Baker, P. Savola, "Filtrage d'entrée pour réseaux à rattachement multiples", mars 2004. (BCP0084) (*MàJ par* RFC8704)

[RFC3917] J. Quittek, T. Zseby, B. Claise, S. Zander, "Exigences pour l'exportation d'informations de flux IP (IPFIX)", octobre 2004. *(Information)*

[RFC4271] Y. Rekhter, T. Li et S. Hares, "Protocole de routeur frontière version 4 (BGP-4)", janvier 2006. *(D.S.) (MàJ par* RFC6608, RFC8212, RFC9072)

[RFC5470] G. Sadasivan et autres, "Architecture pour l'exportation d'informations de flux IP", mars 2009. *(Information)*

[Zs02] T. Zseby, "Deployment of Sampling Methods for SLA Validation with Non-Intrusive Measurements", Proceedings of Passive et Active Measurement Workshop (PAM 2002), Fort Collins, CO, USA, 25-26 mars 2002.

Adresse des auteurs

Derek Chiou
Department of Electrical and Computer Engineering
University of Texas at Austin
1 University Station, Stop C0803, ENS Building room 135,
Austin TX, 78712
USA
téléphone : +1 512 232 7722
mél : Derek@ece.utexas.edu

Jennifer Rexford
Department of Computer Science
Princeton University
35 Olden Street
Princeton, NJ 08540-5233
USA
téléphone : +1 609-258-5182
mél : jrex@cs.princeton.edu

Benoit Claise
Cisco Systems
De Kleetlaan 6a b1
1831 Diegem
Belgium
téléphone : +32 2 704 5622
mél : bclaise@cisco.com

Nick Duffield, Editor
AT&T Labs - Research
Room B139
180 Park Ave
Florham Park NJ 07932
USA
mél : duffield@research.att.com

Albert Greenberg
One Microsoft Way
Redmond, WA 98052-6399
USA
téléphone : +1 425-722-8870
mél : albert@microsoft.com

Matthias Grossglauser
School of Computer and Communication Sciences
EPFL
1015 Lausanne
Switzerland
mél : matthias.grossglauser@epfl.ch