

Groupe de travail Réseau
Request for Comments : 3758
 Catégorie : En cours de normalisation

R. Stewart & M. Ramalho, Cisco Systems, Inc.
 Q. Xie, Motorola, Inc.
 M. Tuexen, Univ. of Applied Sciences Muenster
 P. Conrad, University of Delaware
 mai 2004

Traduction Claude Brière de L'Isle

Extension de fiabilité partielle du protocole de transmission de contrôle de flux (SCTP)

Statut de ce mémoire

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

Notice de copyright

Copyright (C) The Internet Society (2004). Tous droits réservés.

Résumé

Le présent mémoire décrit une extension au protocole de transmission de contrôle de flux (SCTP, *Stream Control Transmission Protocol*) qui permet à un point d'extrémité SCTP de signaler à son homologue qu'il devrait déplacer vers l'avant le point d'accusé de réception cumulatif. Lorsque les deux côtés d'une association SCTP prennent en charge cette extension, elle peut être utilisée par une mise en œuvre SCTP pour fournir un service de transmission de données partiellement fiable à un protocole de couche supérieure. Le présent mémoire décrit les extensions du protocole, qui consistent en un nouveau paramètre pour INIT et INIT ACK, et en un nouveau type de tronçon TSN FORWARD, et donne un exemple de service partiellement fiable qui peut être fourni aux couches supérieures via ce mécanisme.

Table des Matières

1. Introduction.....	2
1.1 Vue d'ensemble des extensions du protocole.....	2
1.2 Vue d'ensemble des nouveaux services fournis à la couche supérieure.....	2
1.3 Avantages de PR-SCTP.....	2
2. Conventions.....	3
3. Changements du protocole pour la prise en charge de PR-SCTP.....	3
3.1 Paramètre Forward-TSN-Supported pour INIT et INIT ACK.....	3
3.2 Définition de tronçon TSN cumulatif vers l'avant (FORWARD TSN).....	3
3.3 Négociation du paramètre Forward-TSN-Supported.....	4
3.4 Définition de "abandonné" dans le contexte de PR-SCTP.....	5
3.5 Mise en œuvre côté expéditeur de PR-SCTP.....	5
3.6 Mise en œuvre côté receveur de PR-SCTP.....	7
4. Services fournis par PR-SCTP à la couche supérieure.....	8
4.1 Définition de service PR-SCTP pour "fiabilité en temps".....	8
4.2 Établissement d'association PR-SCTP.....	9
4.3 Lignes directrices pour définir d'autres services PR-SCTP.....	10
4.4 Notes d'utilisation.....	11
5. Variables.....	11
6. Remerciements.....	11
7. Considérations sur la sécurité.....	11
8. Considérations relatives à l'IANA.....	11
9. Références.....	12
9.1 Références normatives.....	12
9.2 Références pour information.....	12
10. Adresse des auteurs.....	12
11. Déclaration complète de droits de reproduction.....	12

1. Introduction

Le présent mémoire décrit une extension au protocole de transmission de contrôle de flux (SCTP, *Stream Control Transmission Protocol*) [RFC2960] qui permet à un expéditeur SCTP de signaler à son homologue qu'il ne devrait plus espérer recevoir un ou plusieurs tronçons de données.

1.1 Vue d'ensemble des extensions du protocole

L'extension de protocole décrite dans le présent document consiste en deux nouveaux éléments :

1. un seul nouveau paramètre dans l'échange INIT/INIT-ACK qui indique si le point d'extrémité prend en charge l'extension,
2. un seul nouveau type de tronçon, FORWARD TSN, qui indique que le receveur DEVRAIT déplacer vers l'avant son point d'accusé de réception cumulatif (en sautant éventuellement un ou plusieurs tronçons DATA passés qui PEUVENT n'avoir pas encore été reçus et/ou acquittés.)

1.2 Vue d'ensemble des nouveaux services fournis à la couche supérieure

Lorsque cette extension est prise en charge par les deux côtés d'une association SCTP, elle peut être utilisée pour fournir un service de transport partiellement fiable sur une association SCTP. On définit un service de transport partiellement fiable comme un service qui permet à l'utilisateur de spécifier, message par message, les règles qui gouvernent quelle DEVRAIT être la persistance du service pour tenter d'envoyer le message au receveur.

Un exemple de service partiellement fiable est spécifié dans le présent document, à savoir le service "fiabilité en temps" (*timed reliability*). Ce service permet à l'utilisateur du service d'indiquer une limite à la durée pendant laquelle l'expéditeur DEVRAIT essayer de transmettre/retransmettre le message (c'est une extension naturelle du paramètre "durée de vie" qui est déjà dans le protocole de base).

En plus de cet exemple, on montrera aussi que définir la sémantique d'un service partiellement fiable particulier implique deux éléments, à savoir :

1. comment l'utilisateur du service indique le niveau de fiabilité requis pour un message particulier, et
2. comment la mise en œuvre du côté de l'expéditeur utilise ce niveau de fiabilité pour déterminer quand abandonner les tentatives de retransmission de ce message.

Noter qu'à part le fait que le tronçon FORWARD-TSN est requis, aucun de ces deux éléments n'impacte le protocole "sur le fil" ; seules l'API et la mise en œuvre du côté de l'expéditeur sont affectées par la façon dont le service est défini à la couche supérieure. Donc, en principe, il est faisable de mettre en œuvre de nombreuses variétés de services partiellement fiables dans une mise en œuvre SCTP particulière dans changer le protocole réseau. Aussi, le receveur SCTP n'a pas nécessairement besoin de savoir quelle sémantique de service partiellement fiable est utilisée par l'expéditeur, car le seul rôle du receveur est d'interpréter correctement les tronçons FORWARD TSN, sautant les messages passés que l'expéditeur a décidé de ne plus transmettre (ou retransmettre).

Néanmoins, il est recommandé qu'un nombre limité de définitions standard de services partiellement fiables soient normalisées par l'IETF afin que les concepteurs de protocole de couche application de l'IETF puissent satisfaire aux exigences de leurs protocoles de couche supérieure pour les définitions de service standard fournies par une mise en œuvre SCTP particulière. Une telle définition, "fiabilité en temps", est incluse dans le présent document. Étant données les extensions proposées dans le présent document, d'autres définitions PEUVENT être normalisées lorsque le besoin s'en fera sentir sans autres changements du protocole réseau.

1.3 Avantages de PR-SCTP

Ci-après, on utilise la notation "protocole de transmission de contrôle de flux partiellement fiable" (PR-SCTP, *Partial Reliable Stream Control Transmission Protocol*) pour se référer au protocole SCTP, étendu comme défini dans le présent document.

Voici quelques uns des avantages de l'intégration du service de données partiellement fiable dans SCTP, c'est-à-dire, les avantages de PR-SCTP :

1. Certains protocoles de couche application PEUVENT tirer parti d'être capables d'utiliser une seule association SCTP pour porter à la fois du contenu fiable, comme des pages de texte, des informations de facturation et de comptabilité, la signalisation d'établissement, et du contenu non fiable, par exemple, l'état qui est très sensible à l'opportunité, où la

génération d'un nouveau paquet est plus avantageuse que la transmission d'un vieux [NGP].

2. Le trafic de données partiellement fiable porté par PR-SCTP bénéficiera de la même détection de défaillance de communication et des capacités de protection que le trafic normal de données SCTP fiable. Cela inclut la capacité de détecter rapidement un échec d'adresse de destination, la reprise sur défaillance d'une autre adresse de destination, et la notification de ce que le receveur des données devient injoignable.
3. En plus de fournir un transfert de données non ordonné, non fiable comme le fait UDP, PR-SCTP peut fournir un service de transfert de données ordonné, non fiable.
4. PR-SCTP emploie le même contrôle d'encombrement et évitement d'encombrement pour tout le trafic de données, qu'il soit fiable ou partiellement fiable – ceci est très utile car SCTP met en application la compatibilité TCP (à la différence de UDP.)
5. À cause de la fonction d'empaquetage de tronçon de SCTP, des messages fiables et non fiables peuvent être multiplexés sur une seule association PR-SCTP. Donc, le nombre de datagrammes IP (et donc la redondance du réseau) peut être réduit au lieu d'avoir à envoyer ces différents types de données en utilisant des protocoles séparés. De plus, ce multiplexage permet une économie d'accès plutôt que d'utiliser des accès différents pour les connexions fiables et non fiables.

2. Conventions

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

Les comparaisons et l'arithmétique de numéros de séquence de transport (TSN, *Transport Sequence Number*) sont gouvernées par les règles du paragraphe 1.6 de la [RFC2960].

3. Changements du protocole pour la prise en charge de PR-SCTP

3.1 Paramètre Forward-TSN-Supported pour INIT et INIT ACK

Le nouveau paramètre FACULTATIF est ajouté aux tronçons INIT et INIT ACK.

Nom du paramètre	Statut	Valeur du type
Forward-TSN-Supported	FACULTATIF	49152 (0xC000)

À l'initialisation de l'association, l'envoyeur du tronçon INIT ou INIT ACK PEUT inclure ce paramètre FACULTATIF pour informer son homologue qu'il est capable de prendre en charge le tronçon Forward TSN (voir au paragraphe 3.3 pour les détails). Le format de ce paramètre est défini comme suit :

```

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Type de paramètre = 49152 | Longueur de paramètre = 4 |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Type : 16 bit u_int
49152, indiquant le paramètre Forward-TSN-Supported

Longueur : 16 bit u_int
Indique la taille du paramètre, c'est-à-dire, 4.

3.2 Définition de tronçon TSN cumulatif vers l'avant (FORWARD TSN)

Le nouveau type de tronçon suivant est défini :

Type de tronçon	Nom du tronçon
192 (0xC0)	Forward Cumulative TSN (FORWARD TSN)

Ce tronçon devra être utilisé par l'envoyeur des données pour informer le receveur des données qu'il faut ajuster son point de TSN cumulatif reçu vers l'avant parce que des TSN manquants sont associés à des tronçons de données qui NE DEVRAIENT PAS être transmis ou retransmis par l'envoyeur.

Le tronçon TSN cumulatif vers l'avant a le format suivant :

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Type = 192  |Fanions = 0x00 |      Longueur = variable      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Nouveau TSN cumulatif                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Flux-1          |          Séquence de flux-1          |
+-----+-----+-----+-----+-----+-----+-----+-----+
\                                                                    /
/                                                                    \
+-----+-----+-----+-----+-----+-----+-----+-----+
|          Flux-N          |          Séquence de flux-N          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Fanions de tronçons : réglés tous à zéro à l'émission et ignorés en réception.

Nouveau TSN cumulatif : 32 bit `u_int`Cela indique le nouveau TSN cumulatif au receveur des données. À réception de cette valeur, le receveur des données DOIT considérer tout TSN manquant plus tôt que ou égal à cette valeur comme reçu, et arrête de le rapporter comme trou dans les SACK suivants.

FluxN : 16 bit `u_int`

Ce champ contient un numéro de flux qui a été sauté par ce FWD-TSN.

Séquence de flux-N : 16 bit `u_int`

Ce champ contient le numéro de séquence associé au flux qui a été sauté. Le champ séquence de flux contient le plus grand numéro de séquence de flux dans ce flux sauté. Le receveur des FWD-TSN peut utiliser les champs Flux-N et Séquence de flux-N pour permettre la livraison de tout TSN échoué qui reste dans les files d'attente de remise en ordre de flux. Ce champ NE DOIT PAS rapporter de TSN correspondant aux tronçons DATA qui sont marqués comme déclassés. Pour les tronçons de données classés, ce champ DOIT être rempli.

3.3 Négociation du paramètre Forward-TSN-Supported

3.3.1 Envoi du paramètre Forward-TSN-Supported dans INIT

Si un point d'extrémité SCTP prend en charge le tronçon FORWARD TSN, chaque fois qu'il envoie un INIT durant l'établissement d'association, il PEUT alors inclure le paramètre Forward-TSN-supported dans le tronçon INIT pour indiquer ce fait à son homologue.

Noter que si le point d'extrémité choisit de NE PAS inclure le paramètre, alors à aucun moment de la vie de l'association il ne peut envoyer ou traiter un FORWARD TSN. Il DOIT à la place agir comme si il ne prenait PAS en charge le tronçon FORWARD TSN, et retourner une ERREUR à l'homologue à réception de tout FORWARD TSN.

3.3.2 Réception du paramètre Forward-TSN-Supported dans INIT ou INIT-ACK

Lorsque un receveur d'un INIT détecte un paramètre Forward-TSN-Supported et qu'il ne prend pas en charge le type de tronçon Forward-TSN, le receveur DOIT suivre les règles définies au paragraphe 3.3.3 de la [RFC2960].

Lorsque un receveur d'un INIT-ACK détecte un paramètre Forward-TSN-Supported et qu'il ne prend pas en charge le type de tronçon Forward-TSN, le receveur DOIT suivre les règles définies au paragraphe 3.3.3 de la [RFC2960].

Lorsque un receveur d'un INIT détecte un paramètre Forward-TSN-Supported et qu'il ne prend pas en charge le type de tronçon Forward-TSN, le receveur PEUT répondre avec un paramètre Forward-TSN-supported dans le tronçon INIT-ACK.

Noter que si le point d'extrémité choisit de NE PAS inclure le paramètre, il ne peut alors à aucun moment de la vie de l'association envoyer ou traiter un FORWARD TSN. Il DOIT à la place agir comme si il ne prenait PAS en charge le

tronçon FORWARD TSN, et retourner une ERREUR à l'homologue à réception de tout FORWARD TSN.

Lorsque un point d'extrémité qui prend en charge le tronçon FORWARD TSN reçoit un INIT qui ne contient pas le paramètre Forward-TSN-Supported, ce point d'extrémité :

- o PEUT inclure le paramètre Forward-TSN-Supported dans le INIT-ACK,
- o DEVRAIT enregistrer le fait que l'homologue ne prend pas en charge le tronçon FORWARD TSN,
- o NE DOIT PAS envoyer un tronçon FORWARD TSN pendant la durée de vie de l'association,
- o DEVRAIT informer la couche supérieure si celle-ci a demandé une telle notification.

3.3.3 Réception d'une erreur de fonctionnement pour le paramètre Forward-TSN-Supported

Lorsque un point d'extrémité SCTP qui désire utiliser la caractéristique tronçon FORWARD TSN pour un transfert de données partiellement fiable reçoit une erreur de fonctionnement du point d'extrémité distant (groupée avec le COOKIE ou comme paramètre non reconnu dans le INIT-ACK) qui indique que le point d'extrémité distant ne reconnaît pas le paramètre Forward-TSN-Supported, le point d'extrémité local DEVRAIT informer sa couche supérieure de l'incapacité du point d'extrémité distant à prendre en charge le transfert de données partiellement fiable.

Le point d'extrémité local PEUT alors choisir :

- 1) de terminer le procès d'initiation (si celui-ci s'est déjà terminé, le point d'extrémité PEUT avoir besoin d'envoyer un ABORT) en considération de l'incapacité de l'homologue à fournir les caractéristiques demandées pour la nouvelle association,
- 2) ou de continuer le procès d'initiation (si celui-ci s'est déjà terminé, le point d'extrémité DOIT juste marquer l'association comme ne prenant pas en charge la fiabilité partielle) mais en comprenant que la transmission de données partiellement fiable n'est pas prise en charge. Dans ce cas, le point d'extrémité qui reçoit l'erreur de fonctionnement DEVRAIT noter que le tronçon FORWARD TSN n'est pas pris en charge, et NE DOIT PAS transmettre un tronçon FORWARD TSN pendant toute la durée de vie de l'association.

3.4 Définition de "abandonné" dans le contexte de PR-SCTP

À certains moments, une mise en œuvre PR-SCTP envoyeuse PEUT déterminer qu'un tronçon de données particulier NE DEVRAIT PAS être transmis ou retransmis, conformément aux règles qui gouvernent une certaine définition de service PR-SCTP particulière (comme la définition de "fiabilité en temps" au paragraphe 4.1). Pour les besoins du présent document, on définit le terme "abandonné" comme se référant à tout tronçon de données sur lequel l'envoyeur SCTP a fait cette détermination.

Chaque service PR-SCTP définit les règles pour déterminer quand un TSN est "abandonné", et en conséquence, les règles qui gouvernent comment, si, et quand "abandonner" un TSN PEUVENT varier d'une définition de service à une autre. Cependant, les règles qui gouvernent les actions effectuées lorsque un TSN est "abandonné" NE varient PAS entre les définitions de service ; ces règles figurent au paragraphe 3.5.

3.5 Mise en œuvre côté envoyeur de PR-SCTP

La mise en œuvre côté envoyeur de PR-SCTP est identique à celle du protocole SCTP de base, sauf :

- o les actions qu'une mise en œuvre PR-SCTP côté envoyeur DOIT effectuer lorsque un TSN est "abandonné" (selon les règles de la définition des services PR-SCTP en effet)
- o les actions particulières qu'une mise en œuvre PR-SCTP DOIT effectuer à réception d'un SACK
- o les règles qui gouvernent la génération des tronçons FORWARD TSN.

En détail, ces exceptions sont les suivantes :

- A1) L'envoyeur maintient un "Advanced.Peer.Ack.Point" pour chaque homologue pour retracer un point de TSN cumulatif théorique de l'homologue. (Noter que c'est une variable de protocole `_new_` et sa valeur N'EST PAS nécessairement la même que le "point d'accusé de réception TSN cumulatif" SCTP défini au paragraphe 1.4 de la [RFC2960], et comme exposé dans le présent document.)
- A2) De temps en temps, selon les règles d'une définition de service PR-SCTP particulière (voir la Section 4) l'envoyeur de données SCTP PEUT déterminer qu'un tronçon de données particulier à qui a déjà été alloué un TSN DEVRAIT être "abandonné". Lorsque un tronçon de données est "abandonné", l'envoyeur DOIT traiter le tronçon de données comme ayant été finalement acquitté et comme n'étant plus en instance. L'envoyeur NE DOIT PAS créditer un tronçon de données "abandonné" des `partial_bytes_acked` comme défini au paragraphe 7.2.2 de la [RFC2960], et NE DOIT PAS avancer la `cwnd` sur la base de ce tronçon de données "abandonné".

- A3) Lorsque un TSN est "abandonné", si il fait partie d'un message fragmenté, tous les autres TSN au sein de ce message fragmenté DOIVENT être abandonnés en même temps.
- A4) Chaque fois que l'expéditeur des données reçoit un SACK du receveur de données, il DOIT d'abord traiter le SACK en utilisant les procédures normales définie au paragraphe 6.2.1 de la [RFC2960].

L'expéditeur des données DOIT alors effectuer les étapes supplémentaires suivantes :

- C1) Soit SackCumAck l'ACK cumulatif de TSN porté dans le SACK reçu.
Si (Advanced.Peer.Ack.Point < SackCumAck), alors mettre à jour Advanced.Peer.Ack.Point pour qu'il soit égal à SackCumAck.
- C2) Essayer d'avancer encore "Advanced.Peer.Ack.Point" en local, c'est-à-dire, augmenter "Advanced.Peer.Ack.Point" jusqu'à ce que le prochain tronçon dans l'espace de file d'attente de sortie soit marqué "abandonné", comme le montre l'exemple suivant :

En supposant qu'un SACK est arrivé avec l'ACK TSN cumulatif = 102 et que le Advanced.Peer.Ack.Point est mis à jour à cette valeur :

file d'attente de sortie à la fin du traitement ==> file d'attente de sortie après l'avance locale de Adv.Ack.Point normal de SACK

...	...
Adv.Ack.Pt-> 102 acquitté	102 acquitté
103 abandonné	103 abandonné
104 abandonné	Adv.Ack.P-> 104 abandonné
105	105
106 acquitté	106 acquitté
...	...

Dans cet exemple, l'expéditeur des données a bien avancé le "Advanced.Peer.Ack.Point" de 102 à 104 en local.

- C3) Si, après les étapes C1 et C2, le "Advanced.Peer.Ack.Point" est supérieur à l'ACK TSN cumulatif porté dans le SACK reçu, l'expéditeur des données DOIT envoyer au receveur des données un tronçon FORWARD TSN contenant la dernière valeur de "Advanced.Peer.Ack.Point". Noter que l'expéditeur PEUT retarder l'envoi d'un FORWARD TSN comme défini dans la règle F2 ci-dessous.

Note de mise en œuvre : L'adresse de destination à laquelle l'expéditeur est une décision de mise en œuvre, la seule restriction étant que l'adresse DOIT être CONFIRMÉE.

- C4) Pour chaque TSN "abandonné", l'expéditeur du FORWARD TSN DOIT déterminer si le tronçon a un numéro valide de flux et de séquence (c'est-à-dire, si il est ordonné). Si le tronçon a un numéro valide de flux et de séquence, l'expéditeur DOIT l'inclure dans le FORWARD TSN. Cette information permettra au receveur de trouver facilement tout TSN en attente sur les files d'attente de remise en ordre de flux. Chaque flux DEVRAIT n'être rapporté qu'une seule fois ; cela signifie que si plusieurs messages abandonnés surviennent dans le même flux, seul le plus fort numéro de séquence de flux abandonné est rapporté. Si la taille totale du FORWARD TSN ne tient pas dans une seule MTU, l'expéditeur du FORWARD TSN DEVRAIT alors diminuer le Advanced.Peer.Ack.Point au dernier TSN qui tient dans une seule MTU.
- C5) Si un FORWARD TSN est envoyé, l'expéditeur DOIT s'assurer qu'au moins un temporisateur T3-rtx fonctionne.

Note de mise en œuvre : tout temporisateur de destination PEUT être utilisé pour les besoins de la règle C5.

- A5) Chaque fois que le temporisateur T3-rtx arrive à expiration, sur toute destination, l'expéditeur DEVRAIT essayer d'avancer le "Advanced.Peer.Ack.Point" en suivant les procédures mentionnées en C2 - C5.

Les règles supplémentaires suivantes gouvernent la génération des tronçons FORWARD TSN :

- F1) Un point d'extrémité NE DOIT PAS utiliser le FORWARD TSN pour un autre objet que les circonstances décrites dans le présent document.
- F2) L'expéditeur des données DEVRAIT toujours tenter de lier un FORWARD TSN sortant avec des tronçons de DATA sortants pour une meilleure efficacité. Un expéditeur PEUT même choisir de retarder l'envoi du FORWARD TSN dans l'espoir de le lier avec un tronçon DATA sortant.

Note de mise en œuvre : Une mise en œuvre PEUT souhaiter limiter le nombre de tronçons FORWARD TSN dupliqués qu'elle envoie soit en envoyant seulement un FORWARD TSN dupliqué tous les deux SACK, soit en attendant un RTT complet avant d'envoyer un FORWARD TSN dupliqué.

Note de mise en œuvre : Une mise en œuvre PEUT permettre que le délai maximum de génération d'un FORWARD TSN soit configuré soit statiquement soit dynamiquement afin de satisfaire les exigences spécifiques de temps du protocole porté, mais voir la règle suivante:

F3) Un retard appliqué à l'envoi d'un tronçon FORWARD TSN NE DEVRAIT PAS excéder 200 ms et NE DOIT PAS excéder 500 ms. En d'autres termes, une mise en œuvre PEUT diminuer cette valeur en dessous de 500 ms mais NE DOIT PAS l'augmenter au delà de 500 ms.

Note : Retarder l'envoi des tronçons FORWARD TSN PEUT causer des retards de la capacité du receveur à livrer d'autres données qui sont détenues chez le receveur pour remise en ordre. Les valeurs de 200 ms et 500 ms correspondent aux valeurs exigées pour l'accusé de réception retardé dans la [RFC2960] car retarder un FORWARD TSN a les mêmes conséquences mais dans la direction inverse.

F4) Le critère de détection des SACK déclassés DOIT rester le même que ce qui est déclaré dans la RFC 2960, c'est-à-dire, un SACK n'est considéré déclassé que si le ACK TSN cumulatif porté dans le SACK est plus tôt que le SACK reçu précédemment (c'est-à-dire, la comparaison NE DOIT PAS être faite avec "Advanced.Peer.Ack.Point").

F5) Si la décision d'abandon d'un tronçon est prise, quelle que soit la façon dont elle est prise, l'ajustement d'encombrement approprié DOIT être fait comme spécifié dans la RFC 2960 si le tronçon avait été marqué pour retransmission ultérieure (par exemple, soit par la fin du temporisateur T3, soit par retransmission rapide).

3.6 Mise en œuvre côté receveur de PR-SCTP

La mise en œuvre côté receveur de PR-SCTP à un point d'extrémité SCTP A est capable de prendre en charge toute définition de service PR-SCTP utilisée par l'expéditeur au point d'extrémité B, même si cette définition de service n'est pas prise en charge par la fonction de côté d'envoi de l'hôte A. Tout ce qui est nécessaire est que le côté receveur traite correctement le paramètre Forward-TSN-Supported comme spécifié au paragraphe 3.3, et traite correctement la réception des tronçons FORWARD TSN comme spécifié ci-dessous.

L'arrivée d'un tronçon DATA à un receveur PR-SCTP se fait exactement comme pour celle d'un tronçon DATA chez un receveur SCTP du protocole de base – c'est-à-dire que le receveur DOIT effectuer le même traitement de TSN, incluant la détection de dupliqués, la détection de trous, la génération de SACK, l'avancement de TSN cumulatifs, etc. comme défini dans la [RFC2960] avec les exceptions et ajouts suivants.

Lorsque un tronçon FORWARD TSN arrive, le receveur des données DOIT d'abord mettre à jour son point de TSN cumulatif à la valeur portée dans le tronçon FORWARD TSN, et DOIT ensuite avancer son point de TSN cumulatif en local si possible, comme le montre l'exemple suivant :

En supposant que le nouveau TSN cumulatif porté par le FORWARD TSN qui arrive est 103 :

File d'attente entrante avant traitement du FORWARD TSN	==>	File d'attente après traitement du FORWARD TSN et son avancement
cum.TSN.Pt-> 102 reçu		102 --
103 manquant		103 --
104 reçu		104 --
105 reçu		cum.TSN.Pt-> 105 reçu
106 manquant		106 manquant
107 reçu		107 reçu
...		...

Dans cet exemple, le point de TSN cumulatif du receveur est d'abord mis à jour à 103 et ensuite avancé à 105.

Après le traitement ci-dessus, le receveur des données DOIT arrêter de rapporter des TSN manquants inférieurs ou égaux au nouveau point de TSN cumulatif.

Note : si la valeur du "nouveau TSN cumulatif" portée dans le tronçon FORWARD TSN arrivé se trouve être derrière ou au point actuel de TSN cumulatif, le receveur des données DOIT traiter ce FORWARD TSN comme périmé et NE DOIT PAS mettre à jour le TSN cumulatif. Le receveur DEVRAIT envoyer un SACK à son homologue (l'expéditeur du FORWARD TSN) car un tel dupliqué PEUT indiquer que le SACK précédent s'est perdu dans le réseau.

Chaque fois qu'un tronçon FORWARD TSN arrive, pour les besoins de l'envoi d'un SACK, le receveur DOIT suivre les mêmes règles que si un tronçon DATA avait été reçu (c'est-à-dire, suivre les règles de SACK retardé spécifiées au paragraphe 6.2 de la [RFC2960]).

Chaque fois qu'un tronçon DATA arrive avec le bit 'U' réglé à '0' (ce qui indique une livraison dans l'ordre) et est décalé, le receveur DOIT conserver le tronçon pour le remettre dans l'ordre. Comme il est possible avec PR-SCTP qu'un tronçon DATA qu'on attend ne soit pas retransmis, des actions particulières devront être prises à l'arrivée d'un FORWARD TSN.

En particulier, durant le traitement d'un FORWARD TSN, le receveur DOIT utiliser les informations de séquence de flux pour examiner toutes les files d'attente de remise en ordre de flux, et rendre immédiatement disponibles à la livraison les numéros de séquence de flux antérieurs ou égaux au numéro de séquence de flux mentionné dans le FORWARD TSN. Toutes ces données en attente DEVRAIENT être rendues immédiatement disponibles à l'application de couche supérieure.

Une application utilisant PR-SCTP qui reçoit des données DEVRAIT être informée de messages éventuellement manquants. Le numéro de séquence de flux peut être utilisé, dans un tel cas, pour déterminer qu'un message intermédiaire a été sauté. Lorsque des messages intermédiaires manquent, c'est à une décision de l'application que revient le choix de traiter les messages ou de prendre une autre action corrective.

Après la réception et le traitement d'un FORWARD TSN, le receveur des données DOIT faire attention lors de la mise à jour de sa file d'attente de réassemblage. Le receveur DOIT retirer tout message partiellement réassemblé, qui manque encore à un ou plusieurs TSN antérieurs ou égaux au nouveau point de TSN cumulatif. Dans le cas où le receveur a invoqué l'API de livraison partielle, une notification DEVRAIT aussi être générée pour informer l'API de couche supérieure que le message partiellement livré NE SERA PAS complété.

Noter qu'après la réception d'un FORWARD TSN et la mise à jour du point d'accusé de réception cumulatif, si un TSN qui a été sauté arrive (c'est-à-dire, à cause d'une remise en ordre par le réseau) le receveur va suivre les règles normales définies dans la [RFC2960] pour le traitement des données dupliquées. Cela implique que le receveur va abandonner le tronçon et le rapporter comme dupliqué dans le prochain SACK sortant.

4. Services fournis par PR-SCTP à la couche supérieure

Comme décrit au paragraphe 1.2, on peut mettre en œuvre divers services de transport partiellement fiables en utilisant les nouveaux mécanismes de protocole présentés à la Section 3 ; introduire ces nouveaux services exige seulement de faire des changements à l'API côté envoyeur, et à la mise en œuvre de protocole côté receveur. Donc, il PEUT être tentant de ne normaliser que le protocole, et de laisser la définition de service comme "spécifique de la mise en œuvre" ou de la laisser être définie dans des documents "pour information".

Cependant, pour ceux qui PEUVENT souhaiter écrire des normes de l'IETF pour des protocoles de couche supérieure mis en œuvre sur PR-SCTP, il est important d'être capable de se référer à une définition standard des services fournis. Donc, cette section donne des exemples de définition d'un tel service, tout en donnant des lignes directrices pour la définition de services supplémentaires en tant que de besoin. Chacun de ces services PEUT être proposé comme une nouvelle RFC séparée.

La Section 4 est organisée comme suit :

- o le paragraphe 4.1 donne la définition d'un service PR-SCTP spécifique : fiabilité en temps,
- o le paragraphe 4.2 décrit comment une définition particulière de service PR-SCTP est demandée par la couche supérieure durant l'établissement d'association, et comment la couche supérieure est notifiée si cette demande ne peut être satisfaite,
- o le paragraphe 4.3 donne des lignes directrices pour la spécification de services PR-SCTP autres que celui défini ici,
- o le paragraphe 4.4 décrit des notes d'usage supplémentaires qui peuvent être utiles aux concepteurs et mises en œuvre de protocole.

4.1 Définition de service PR-SCTP pour "fiabilité en temps"

Le service "fiabilité en temps" est une extension naturelle du concept de "durée de vie" déjà présent dans le protocole SCTP de base.

Lorsque ce service est demandé pour une association SCTP, il change la signification du paramètre Durée de vie spécifié dans la primitive SEND (voir au paragraphe 10.1, partie (E) de la [RFC2960] ; noter que le paramètre est appelé "durée de vie dans le présent document.)

Dans le protocole SCTP de base, le paramètre Durée de vie est utilisé pour éviter d'envoyer des données périmées. Lorsque une valeur de durée de vie est indiquée pour un message particulier et que cette durée de vie arrive à expiration, SCTP annule l'envoi de ce message, et notifie le protocole de couche supérieure (ULP, *upper layer protocol*) si la première transmission des données n'a pas eu lieu (à cause des limitations de *rwnd* ou *cwnd*, ou pour toute autre raison). Cependant, dans le protocole de base, si SCTP a envoyé la première transmission avant l'expiration de la durée de vie, le message DOIT alors être envoyé comme un message fiable normal. Durant les épisodes d'encombrement, ceci est particulièrement malencontreux, car la retransmission gaspille la bande passante qui aurait pu être utilisée pour d'autres messages (dont la durée de vie n'est pas expirée).

Lorsque le service "fiabilité en temps" est invoqué, cette dernière restriction est supprimée. En particulier, lorsque le service "fiabilité en temps" est activé, les règles suivantes gouvernent tous les messages qui sont envoyés avec un paramètre Durée de vie :

- TR1) Si le paramètre Durée de vie d'un message est `SCTP_LIFETIME_RELIABLE` (ou non spécifié, voir la Section 5) ce message est traité comme un message normal SCTP fiable, tout comme dans le protocole SCTP de base.
- TR2) Si le paramètre Durée de vie n'est pas `SCTP_LIFETIME_RELIABLE` (voir la Section 5) l'expéditeur SCTP DOIT alors traiter le message tout comme si il était un message normal SCTP fiable, tant que la durée de vie n'a pas encore expiré.
- TR3) Avant d'allouer un TSN à un message, l'expéditeur SCTP DOIT évaluer la durée de vie de ce message. Si elle est expirée, l'expéditeur SCTP NE DOIT PAS allouer de TSN à ce message, mais plutôt DEVRAIT produire une notification à la couche supérieure et abandonner le message.
- TR4) Avant de transmettre ou retransmettre un message pour lequel un TSN est déjà alloué, l'expéditeur SCTP DOIT évaluer la durée de vie du message. Si la durée de vie du message est expirée, l'expéditeur SCTP DOIT "abandonner" le message, selon les règles spécifiées au paragraphe 3.5, en marquant ce TSN comme éligible pour l'avancée de TSN. Noter que ceci satisfait à l'exigence G1 définie au paragraphe 4.3.

Note de mise en œuvre : Une mise en œuvre DEVRAIT retarder l'allocation de TSN comme mentionné au paragraphe 10.1 de la [RFC2960]. Dans ce cas, le paramètre Durée de vie DEVRAIT être vérifié AVANT l'allocation d'un TSN, permettant donc que le message soit abandonné sans qu'il soit besoin d'envoyer un FORWARD TSN.

- TR5) Le SCTP expéditeur PEUT évaluer la durée de vie des messages à tout moment. Les messages expirés qui n'ont pas reçu de TSN PEUVENT être traités selon la règle TR3. Les messages expirés qui ont reçu un TSN PEUVENT être traités selon la règle TR4.
- TR6) L'application envoyeuse NE DOIT PAS changer le paramètre Durée de vie une fois que le message a été passé au SCTP expéditeur.

Note de mise en œuvre : Les règles TR1 à TR4 sont conçues de façon à éviter d'exiger de la mise en œuvre le maintien d'un temporisateur séparé pour chaque message ; la durée de vie a seulement besoin d'être évaluée à des instants de la vie du message où des actions sont déjà entreprises, comme l'allocation du TSN, la transmission, ou l'expiration d'un temporisateur de retransmission. La règle TR5 est destinée à donner à la mise en œuvre de SCTP la souplesse d'évaluer la durée de vie à toute autre opportunité convenable, SANS exiger que la durée de vie soit évaluée immédiatement à l'instant où elle expire.

4.2 Établissement d'association PR-SCTP

Un protocole de couche supérieure (ULP) qui utilise PR-SCTP PEUT avoir besoin de savoir si PR-SCTP peut être pris en charge sur une certaine association. Donc, l'ULP doit avoir certaines indications de si le tronçon FORWARD-TSN est pris en charge par son homologue.

Le paragraphe 10.1 de la [RFC2960] décrit les primitives abstraites pour l'interface d'ULP à SCTP, et note que "les mises en œuvre individuelles DOIVENT définir leur propre format exact, et PEUVENT fournir des combinaisons ou des sous ensembles de fonctions de base dans des appels individuels."

Dans ce paragraphe, on décrit une valeur de retour supplémentaire qui PEUT être ajoutée à la primitive ASSOCIATE pour permettre à un utilisateur de service SCTP d'indiquer si le tronçon FORWARD-TSN est pris en charge par son homologue.

La RFC 2960 indique que la primitive ASSOCIATE "permet à la couche supérieure d'initier une association à un point

d'extrémité spécifique de l'homologue". Il est structuré comme suit :

Format : ASSOCIATE(nom d'instance locale SCTP, adresse de transport de destination, compte de flux sortant) -> identifiant d'association [,liste d'adresses de transport de destination] [,compte de flux sortants]

Cette extension ajoute une nouvelle valeur de retour FACULTATIVE, de sorte que la nouvelle primitive se lit comme suit :

Format : ASSOCIATE(nom d'instance locale SCTP, adresse de transport de destination, compte de flux sortant) -> identifiant d'association [,liste d'adresses de transport de destination] [,compte de flux sortants] [,forward tsn pris en charge]

Note : selon la RFC 2960, si la primitive ASSOCIATE est mise en œuvre comme un appel non bloquant, la nouvelle valeur FACULTATIVE de retour devra être passée avec les paramètres d'association en utilisant la notification COMMUNICATION UP.

Le nouveau paramètre FACULTATIF "forward tsn pris en charge" est un fanion booléen :
 (0) faux [par défaut] indique que FORWARD TSN n'est pas activé par les deux points d'extrémité.
 (1) vrai indique que FORWARD TSN est activé par les deux points d'extrémité.

On ajoute aussi une nouvelle primitive pour permettre que l'application d'utilisateur active/désactive le service PR-SCTP sur son point d'extrémité avant l'établissement d'une association.

Format : ENABLE_PRSCTP(nom d'instance locale SCTP, active booléen)

Le paramètre booléen, s'il est réglé à vrai, va activer PR-SCTP sur les futures associations de points d'extrémité. Si le paramètre booléen est réglé à faux, le point d'extrémité local ne va pas annoncer la prise en charge de PR-SCTP et donc désactiver la caractéristique sur les futures associations. Il est recommandé que cette option soit désactivée par défaut, c'est-à-dire, afin d'activer PR-SCTP, l'utilisateur devra invoquer cette option d'API avec le fanion d'activation réglé à "vrai".

4.3 Lignes directrices pour définir d'autres services PR-SCTP

D'autres services PR-SCTP PEUVENT être définis et mis en œuvre comme dicté par les besoins des protocoles de couche supérieure. Si de tels protocoles de couche supérieure devaient être normalisés et exiger un service PR-SCTP particulier autre que celui défini dans le présent document (c'est-à-dire, "fiabilité en temps") ces services PR-SCTP supplémentaires DEVRAIENT alors aussi être spécifiés et normalisés dans une nouvelle RFC.

Il est suggéré que toute définition de service supplémentaire soit modélisée d'après le contenu du paragraphe 4.1. En particulier, la définition de service DEVRAIT fournir :

1. Une description de comment l'utilisateur du service spécifie les paramètres qui doivent être associés à un message particulier (et/ou toute autre communication qui a lieu entre l'application et l'expéditeur de transport SCTP) qui fournit à l'expéditeur de transport SCTP les informations nécessaires pour déterminer quand abandonner la transmission d'un message particulier.
 De préférence, cette description DEVRAIT faire référence aux primitives dans l'API abstraite fournie à la Section 10 de la [RFC2960], indiquant tout :
 - * changement de l'interprétation des paramètres existants des primitives existantes,
 - * les paramètres supplémentaires à ajouter aux primitives existantes (ils DEVRAIENT être FACULTATIFS, et les valeurs par défaut DEVRAIENT être indiquées),
 - * primitives supplémentaire qui PEUT être nécessaire.
2. Une description des règles utilisées par le côté expéditeur de la mise en œuvre pour déterminer si elle abandonne les messages auxquels un TSN n'a pas encore été alloué. Cette description DEVRAIT aussi indiquer quels événements du protocole déclenchent l'évaluation, et quelles actions entreprendre (par exemple, des notifications.)
3. Une description des règles utilisées par le côté expéditeur de la mise en œuvre pour déterminer quand abandonner la transmission ou retransmission des messages auxquels un TSN a déjà été alloué, et PEUVENT avoir été transmis et éventuellement retransmis zéro, une ou plusieurs fois.

Les éléments (2) et (3) de la liste ci-dessus DEVRAIENT aussi indiquer quels événements du protocole déclenchent l'évaluation, et quelles actions entreprendre si la détermination est faite que l'expéditeur DEVRAIT abandonner la transmission du message (par exemple, des notifications à l'ULP.)

Noter que dans tout service PR-SCTP, les règles suivantes DOIVENT être spécifiées pour éviter une impasse du protocole :

(G1) Lorsque le côté expéditeur de la mise en œuvre abandonne la transmission d'un message qui a reçu un TSN (c'est-à-

dire, lorsque ce message est "abandonné", comme défini au paragraphe 3.4) le côté envoyeur DOIT marquer ce TSN comme éligible pour l'avancée de TSN, et les règles du paragraphe 3.4 concernant l'envoi de tronçons FORWARD TSN DOIVENT être suivies.

Finalement, une définition de service PR-SCTP DEVRAIT spécifier un "nom canonique de service" pour identifier de façon univoque le service, et le distinguer des autres services PR-SCTP. Ce nom peut alors être utilisé dans les normes de protocole de couche supérieure pour indiquer quelle définition de service PR-SCTP est exigée par ce protocole de couche supérieure. Il peut aussi être utilisé dans la documentation des API de mise en œuvre de PR-SCTP pour indiquer comment une couche supérieure indique quelle définition de service PR-SCTP DEVRAIT s'appliquer. Le nom canonique de service pour le service PR-SCTP défini au paragraphe 4.1 est "fiabilité en temps".

4.4 Notes d'utilisation

Détecter des données manquantes dans un flux PR-SCTP est utile pour certaines applications (par exemple, canal fibre ou SCSI sur IP). Avec PR-SCTP, cela devient possible – la couche supérieure a simplement besoin d'examiner le numéro de séquence de flux des messages d'utilisateurs arrivés dans ce flux pour détecter toutes données manquantes. Noter que cette détection ne fonctionne que quand tous les messages de ce flux sont envoyés dans l'ordre, c'est-à-dire que le bit "U" n'est pas établi.

5. Variables

Cette Section définit les variables utilisées dans le présent document:

SCTP_LIFETIME_RELIABLE - indication d'interface d'utilisateur définie par une mise en œuvre et utilisée pour indiquer quand un message est à considérer comme pleinement fiable.

6. Remerciements

Les auteurs tiennent à remercier Brian Bidulock, Scott Bradner, Jon Berger, Armando L. Caro Jr., John Loughney, Jon Peterson, Ivan Arias Rodriguez, Ian Rytina, Chip Sharp, et les autres, de leurs commentaires.

7. Considérations sur la sécurité

Le présent document n'introduit pas de nouveaux problèmes de sécurité dans SCTP autres que ceux déjà documentés dans la [RFC2960]. En particulier, le présent document partage les problèmes de sécurité de données en désordre avec la [RFC2960] identifiés par la [RFC3436]. Une application qui utilise l'extension PR-SCTP NE DEVRAIT PAS utiliser la sécurité de couche transport ; on trouvera plus de détails sur ce sujet dans la [RFC3436].

Noter que la capacité de faire sauter un message (c'est-à-dire le tronçon FORWARD TSN) ne fournit aucune nouvelle opportunité d'attaque par interposition (MIM), car l'interposé est déjà capable de changer et/ou retirer des données, sautant donc effectivement des messages. Cependant, le tronçon FORWARD TSN fournit un mécanisme qui facilite à un interposé de choisir les messages à sauter lorsque l'application a activé cette caractéristique car l'interposé aura moins d'état à conserver.

8. Considérations relatives à l'IANA

L'IANA a alloué 192 comme nouveau type de tronçon à SCTP.

L'IANA a alloué 49152 comme nouveau code de type de paramètre à SCTP.

9. Références

9.1 Références normatives

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

[RFC2960] R. Stewart et autres, "Protocole de transmission de commandes de flux", octobre 2000. (*Obsolète, voir RFC4960*) (P.S.)

9.2 Références pour information

[NGP] Clark, D. and D. Tennenhouse, "Architectural Considerations for a New Generation of Protocols", SIGCOMM 1990 pp. 200-208, septembre 1990.

[RFC3436] A. Jungmaier, E. Rescorla, M. Tuexen, "[Sécurité de la couche Transport sur le protocole de transmission](#) de contrôle de flux", décembre 2002. (P.S.)

10. Adresse des auteurs

Randall R. Stewart
Cisco Systems, Inc.
8725 West Higgins Road
Suite 300
Chicago, IL 60631
USA
téléphone : +1-815-477-2127
mél : rrs@cisco.com

Michael A. Ramalho
Cisco Systems, Inc.
1802 Rue de la Porte
Wall Township, NJ 07719-3784
USA
téléphone : +1.732.449.5762
mél : mramalho@cisco.com

Phillip T. Conrad
University of Delaware
Department of Computer & Information Sciences
Newark, DE 19716
USA
téléphone : +1 302 831 8622
mél : conrad@acm.org
URI : <http://www.cis.udel.edu/~pconrad>

Qiaobing Xie
Motorola, Inc.
1501 W. Shure Drive, #2309
Arlington Heights, IL 60004
USA
téléphone : +1-847-632-3028
mél : qxie1@email.mot.com

Michael Tuexen
Univ. of Applied Sciences Muenster
Stegerwaldstr. 39
48565 Steinfurt
Germany
mél : tuexen@fh-muenster.de

11. Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2004).

Le présent document est soumis aux droits, licences et restrictions contenues dans le BCP 78 et sauf comme mentionné ci-dessous, les auteurs conservent tous leurs droits.

Le présent document et les informations contenues sont fournis sur une base "EN L'ÉTAT" et le contributeur, l'organisation qu'il ou elle représente ou qui le/la finance (s'il en est) la INTERNET SOCIETY et la INTERNET ENGINEERING TASK FORCE déclinent toutes garanties, exprimées ou implicites, y compris mais non limitées à toute garantie que l'utilisation des informations incluses ne viole aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Propriété intellectuelle

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

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat

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

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

Remerciement

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