

Groupe de travail Réseau
Request pour Commentaires : 5043
 Catégorie : Sur la voie de la normalisation
 Traduction Claude Brière de L'Isle

C. Bestler, éd., Neterion
 R. Stewart, éd., Cisco Systems, Inc.
 octobre 2007

Adaptation du placement direct des données (DDP) sur le protocole de transmission de contrôle de flux (SCTP)

Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet sur la voie de la normalisation 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 "Protocoles officiels de l'Internet" (STD 1) pour voir l'état de normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Résumé

Le présent document spécifie une couche d'adaptation pour fournir un service de protocole de couche inférieure (LLP, *Lower Layer Protocol*) pour le placement direct de données (DDP, *Direct Data Placement*) utilisant le protocole de transmission de contrôle de flux (SCTP,).

Table des Matières

1. Introduction.....	2
1.1 Conventions.....	2
2. Définitions.....	2
3. Motivation.....	3
4. Vue d'ensemble.....	3
5. Formats de données.....	3
5.1 Indicateur de couche d'adaptation.....	3
5.2 Tronçons de données de charge utile.....	4
6. Sessions de flux DDP.....	5
6.1 Séquençage.....	5
6.2 Séquence légale : session active/passive acceptée.....	6
6.3 Séquence légale : session active/passive rejetée.....	6
6.4 Séquence légale : session active/passive non ULP rejetée.....	6
6.5 Séquençage spécifique d'ULP.....	6
6.6 Autres règles de séquençage.....	7
7. Points d'extrémité SCTP.....	7
7.1 Restriction d'indication de couche d'adaptation.....	7
7.2 Implications du multi rattachements.....	7
8. Nombre de flux.....	7
9. Fragmentation.....	8
10. Fonctionnement en séquence non ordonné.....	8
11. Procédures.....	8
11.1 Initialisation d'association.....	8
11.2 Regroupement de tronçons.....	9
11.3 Terminaison d'association.....	9
12. Considérations relatives à l'IANA.....	9
13. Considérations sur la sécurité.....	9
14. Contributeurs.....	10
15. Remerciements.....	10
16. Références.....	10
16.1 Références normatives.....	10
16.2 Références pour information.....	10
Adresse des auteurs.....	11
Déclaration complète de droits de reproduction.....	11

1. Introduction

Le présent document décrit une méthode pour adapter le placement direct de données [RFC5041] au protocole de transmission de contrôle de flux (SCTP) [RFC4960].

Certaines mises en œuvre peuvent inclure cette couche d'adaptation au sein de leurs mises en œuvre de SCTP pour obtenir un maximum de performances, mais le comportement de SCTP ne sera pas affecté. Une couche SCTP utilisée seulement par cette couche d'adaptation est capable de faire certaines optimisations sur la base du sous ensemble limité de capacités SCTP utilisé. Afin de permettre l'optimisation pour ces mises en œuvre, on spécifie l'indication de l'utilisation de la nouvelle couche d'adaptation comme défini dans la [RFC5061]

1.1 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].

2. Définitions

DDP : voir le protocole de placement direct de données.

point d'extrémité DDP : expéditeur/receveur logique de segments DDP. Une paire de flux SCTP n'est pas supposée avoir un point d'extrémité DDP. Les segments DDP peuvent seulement être envoyés une fois qu'un point d'extrémité DDP a été alloué à une paire de flux SCTP par une interface locale.

numéro de séquence de flux de source DDP (DDP-SSN, *DDP Source Stream Sequence Number*) : numéro de séquence spécifique de flux alloué par la couche d'adaptation à chaque tronçon de données SCTP envoyé. C'est l'ordre dans lequel les troncçons ont été soumis au SCTP, quel que soit l'ordre dans lequel ils sont réellement envoyés ou reçus.

segment DDP : plus petite unité de transfert de données pour le protocole DDP. Il inclut un en-tête DDP et une charge utile d'ULP (si elle est présente). Un segment DDP devrait être dimensionné pour tenir dans la Mulpdu (*Marker PDU Aligned (MPA) Upper Layer PDU*) du protocole de couche inférieure.

tronçon de segment DDP : tronçon de données de charge utile SCTP qui encapsule le DDP-SSN et un segment DDP.

flux DDP : séquence de segments DDP dont l'ordre est défini par le LLP. Pour SCTP, un flux DDP se transpose directement en une paire bidirectionnelle de flux SCTP avec les mêmes identifiants de flux. Noter que DDP n'a pas de garanties d'ordre entre les flux DDP.

session de flux DDP : un seul appariement de points d'extrémité DDP sur un flux DDP qui dure d'un message d'initiation jusqu'au ou aux messages de terminaison.

message de contrôle de session de flux DDP : message utilisé pour contrôler l'association du point d'extrémité DDP avec le flux DDP.

protocole de placement direct de données (DDP) : protocole réseau qui prend en charge le placement direct de données en associant des informations explicites de placement de mémoire tampon avec les unités de charge utile de LLP.

protocole de couche inférieure (LLP) : dans le contexte de DDP, la couche de protocole en dessous de RDMA qui fournit un service de transport fiable. L'adaptation DDP de SCTP est un des LLP initialement défini pour DDP.

domaine de protection : convention commune d'interface locale pour contrôler quelles étiquettes de pilotage (STag, *Steering Tag*) sont valides avec quels points d'extrémité DDP. Sous cette convention, l'étiquette de pilotage et le point d'extrémité DDP sont créés au sein du contexte d'un domaine de protection, et l'étiquette de pilotage peut seulement être activée pour les points d'extrémité DDP créés sous le même domaine de protection.

RDMA (*Remote Direct Memory Access*) : accès direct à la mémoire distante

RNIC (*RDMA Network Interface Card*) : carte d'interface réseau RDMA.

association SCTP : relation de protocole entre deux points d'extrémité SCTP. Une association SCTP prend en charge plusieurs flux SCTP.

tronçon de données SCTP : tronçon SCTP utilisé pour porter des données de charge utile. Il peut y avoir plusieurs tronçons au sein de chaque paquet SCTP. D'autres tronçons sont utilisés pour contrôler l'association SCTP.

point d'extrémité SCTP : envoyeur/receveur logique des paquets SCTP. Sur un hôte multi rattachements, un point d'extrémité SCTP est représenté à ses homologues comme une combinaison d'un numéro d'accès SCTP et un ensemble d'adresses de transport de destination éligibles auxquelles les paquets SCTP peuvent être envoyés.

flux SCTP : canal unidirectionnel logique établi d'un point d'extrémité SCTP associé à un autre. Il peut y avoir plusieurs flux SCTP au sein de chaque association SCTP. Un flux SCTP est utilisé pour former une direction d'un flux DDP.

numéro de séquence de transmission (TSN, *Transmission Sequence Number*) : numéro de séquence de 32 bits utilisé en interne par SCTP. Un TSN est attaché à chaque tronçon contenant des données d'utilisateur pour permettre au point d'extrémité SCTP receveur d'en accuser réception et détecter les livraisons dupliquées.

protocole de couche supérieure (ULP, *Upper Layer Protocol*) : dans le contexte des spécifications du protocole RDMA, c'est la couche qui utilise les services RDMA. Normalement, c'est une application ou un logiciel médiateur. Un des buts principaux des protocoles RDMA est de permettre un transfert direct de charge utile de/vers les mémoires tampon d'ULP.

3. Motivation

Le présent document spécifie une couche d'adaptation qui satisfait aux exigences d'un protocole de couche inférieure (LLP) pour DDP en utilisant un sous ensemble spécifique des capacités de SCTP.

Le protocole défini est destiné à être mis en œuvre sur les piles SCTP existantes, mais en définissant clairement quelles portions de SCTP sont exigées pour permettre à une mise en œuvre d'être optimisée spécifiquement pour prendre en charge DDP.

4. Vue d'ensemble

La couche d'adaptation utilise une paire de flux SCTP pseudo-numérotés au sein d'une association SCTP pour fournir un flux DDP fiable entre deux points d'extrémité DDP. Sauf comme spécifiquement noté, chaque segment DDP soumis par la couche DDP est codé comme un seul tronçon de données SCTP non ordonné. En plus du segment DDP, le tronçon de données contient aussi un numéro de séquence (DDP-SSN) qui reflète l'ordre dans lequel DDP a soumis les segments pour ce flux.

Une session de flux DDP est définie par les tronçons de contrôle de session de flux DDP qui gèrent l'état de la session de flux DDP. Ces tronçons lient dynamiquement les points d'extrémité DDP à la session de flux DDP, et les tronçons de segment DDP sont utilisés pour livrer de façon fiable les segments DDP avec la session.

5. Formats de données

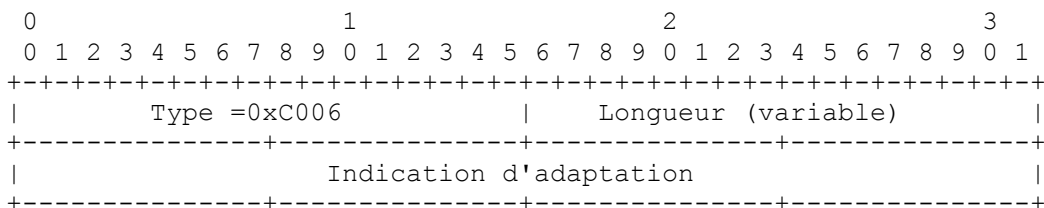
5.1 Indicateur de couche d'adaptation

La couche d'adaptation DDP/SCTP utilise tous les flux au sein d'une association SCTP. Une association SCTP dont l'indication d'adaptation DDP a été négociée va porter seulement des tronçons de données SCTP comme défini dans le présent document.

On suppose que le traitement des tronçons de données entrants pour les associations à capacité DDP est suffisamment différent de celui des associations SCTP habituelles pour qu'il ne soit pas souhaitable d'exiger de prendre en charge le mélange de flux DDP et non DDP dans une seule association. Plus d'une seule association est nécessaire si une application

désire utiliser du trafic à la fois DDP et non DDP avec le même hôte distant.

On définit une indication d'adaptation qui DOIT apparaître dans le INIT ou INIT-ACK avec le format suivant comme défini dans la [RFC5061].



Indication d'adaptation : la valeur 0x00000001 a été allouée pour DDP.

5.2 Tronçons de données de charge utile

L'adaptation DDP SCTP utilise deux types de tronçons de données de charge utile SCTP, différenciés par l'identifiant de protocole de charge utile:

Les tronçons de segment DDP sont utilisés pour livrer fiablement les segments DDP envoyés entre points d'extrémité DDP.

Les messages de contrôle de session de flux DDP sont utilisés pour établir et supprimer les sessions de flux DDP, en particulier en contrôlant le lien des points d'extrémité DDP avec le flux SCTP.

Identifiant de protocole de charge utile :

Les valeurs suivantes définies pour DDP dans ce document ont été allouées par l'IANA :

- Tronçon de segment 16
- Contrôle de session de flux DDP 17

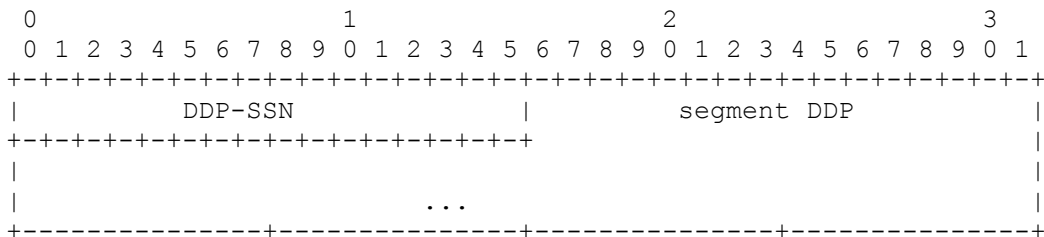
5.2.1 Numéro de séquence de source DDP (DDP-SSN)

Tous les tronçons de données de charge utile SCTP utilisés par cette couche d'adaptation incluent un numéro de séquence de source DDP (DDP-SSN, *DDP Source Sequence Number*). Le DDP-SSN retrace la séquence dans laquelle les messages ont été soumis à la couche SCTP pour le flux SCTP utilisé. Le DDP-SSN DOIT avoir la même valeur que le numéro de séquence du flux SCTP (SSN, *SCTP Sequence Number*) qui aurait été alloué si des tronçons de données de charge utile SCTP ordonnés avaient été utilisés plutôt que des tronçons non ordonnés.

La raison pour spécifier le DDP-SSN est la suivante :

- o Le numéro de séquence de flux SCTP (SSN) ne convient pas dans ce cas parce que tous les messages définis par le présent document utilisent des tronçons de données de charge utile non ordonnés pour assurer une livraison rapide à partir de la couche SCTP receveuse.
- o Le numéro de séquence de transmission SCTP ne convient pas pour déterminer l'ordre original des tronçons de données au sein d'un flux. La couche SCTP envoyeuse peut optimiser la séquence de transmission des tronçons de données non ordonnés pour encourager au groupement des tronçons , ou pour d'autres raisons.

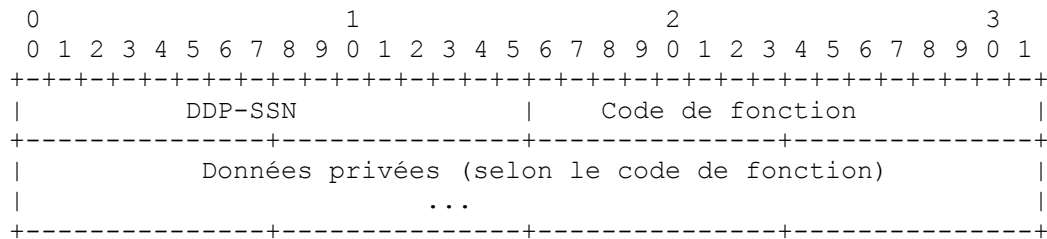
5.2.2 Tronçon de segment DDP



Les segments DDP sont comme défini dans la [RFC5041]. Le tronçon de segment DDP sert au même objet que la PDU de couche supérieure de MPA (MULPDU, *MPA Upper Layer PDU*) [RFC5044] en ce qu'elle porte des segments DDP sur un

protocole fiable avec des informations de séquence en plus.

5.2.3 Contrôle de session de flux DDP



Les valeurs de code de fonction suivantes sont définies pour DDP :

Initier la session de flux	0x001
Accepter la session de flux	0x002
Rejeter la session de flux	0x003
Terminer la session de flux	0x004

Les données privées fournies par l'ULP DOIVENT être incluses pour les messages Initier la session de flux, Accepter la session de flux, et Rejeter la session de flux. Cependant, les données privées fournies par l'ULP PEUVENT être de longueur zéro.

Dans tout message, la longueur des données privées NE DOIT PAS excéder 512 octets.

Les données privées NE DOIVENT PAS être incluses dans le messages Terminer la session de flux.

Les messages de contrôle de session de flux DDP reçus DEVRAIENT être rapportés à l'ULP. Si ils le sont, toutes les données privées fournies DOIVENT être disponible pour que l'ULP les examine.

La couche d'adaptation DDP/SCTP PEUT limiter le nombre de demandes d'initiation de session qu'elle soumet à l'ULP. Quand un Initier la session de flux DDP ne peut pas être transmis à l'ULP à cause d'une telle limite, la couche d'adaptation DOIT répondre avec un message Terminer la session de flux.

6. Sessions de flux DDP

Un point d'extrémité DDP est l'envoyeur/receveur logique des segments DDP. Un flux DDP connecte deux points d'extrémité DDP en utilisant une paire correspondante de flux SCTP ayant les mêmes identifiants de flux SCTP.

Une session de flux DDP définit la séquence de tronçons de données échangés entre deux points d'extrémité DDP sur un flux DDP qui a un commencement et une fin distincts comme défini au paragraphe suivant. Les tronçons de données provenant d'une session de flux DDP ne sont jamais portés sur la prochaine session. Chaque tronçon de données appartient sans ambiguïté à exactement une session. Les DDP-SSN alloués aux tronçons de données pour une session NE DOIVENT PAS avoir de trous.

L'interface locale PEUT associer dynamiquement un point d'extrémité DDP au flux DDP sur la base des échanges initiaux d'une session DDP, et terminer dynamiquement cette association à la fin de la session. Autrement, une interface locale spécialisée pourrait simplement transposer statiquement les points d'extrémité DDP en flux DDP.

Conventionnellement, les interfaces locales pour RDMA ont différé le choix du point d'extrémité DDP jusqu'après la décision de l'ULP d'accepter une demande de connexion RDMA. Mais c'est un choix de l'interface locale et non un exigence du protocole réseau.

Un flux DDP est associé au plus à un domaine de protection durant une seule session de flux DDP. Sur le côté passif, l'association est normalement différée jusqu'au message Accepter la session de flux DDP.

6.1 Séquençage

Le DDP-SSN est remis à zéro au début de chaque session de flux DDP.

La séquence normative pour examiner les tronçons de données de charge utile au sein d'une session se fonde sur le DDP-SSN de chaque tronçon de données. Lorsque elles sont examinées dans cette séquence normative, toutes les sessions DOIVENT se conformer à un des schémas définis dans cette section.

Si la couche d'adaptation reçoit un tronçon de données de charge utile qui ne se conforme à aucun des schémas légaux énumérés, la session de flux DDP DOIT être terminée.

6.2 Séquence légale : session active/passive acceptée

Dans cette séquence de session de flux DDP, un point d'extrémité DDP assume le rôle actif en demandant une session de flux DDP, que l'autre côté accepte.

Le côté actif envoie un message Initier la session de flux DDP.

Le côté passif envoie un message Accepter la session de flux DDP.

Chaque côté peut alors envoyer zéro, un ou plusieurs segments DDP avec des DDP-SSN croissants, sous réserve des contrôles de flux imposés par les autres couches de protocole.

Le tronçon de données d'utilisateur final pour chaque côté PEUT être un flux DDP Terminer. Au moins un côté DOIT envoyer un flux DDP Terminer. Noter que cela va suivre tout message RDMAP Terminer, qui pour la couche d'adaptation est simplement un autre segment DDP.

6.3 Séquence légale : session active/passive rejetée

Les sessions de flux DDP permettent à chaque partie d'envoyer un seul message non de charge utile avant que l'autre extrémité engage des ressources spécifiques pour la session. Cela permet à chaque extrémité de déterminer quelles ressources sont à utiliser, et comment elles sont à configurer, ou même si la session devrait être accordée.

Ces décisions PEUVENT être influencées par le besoin d'allouer un domaine de protection spécifique, pour déterminer combien de crédits RDMA Read sont requis, ou de déterminer combien d'opérations de réception l'ULP devrait permettre.

À cause de ces facteurs ou d'autres, le côté passif PEUT choisir de rejeter une demande de session de flux DDP. Il en résulte la séquence légale suivante :

Le côté actif envoie un message Initier la session de flux DDP.

Le côté passif envoie un message Rejeter la session de flux DDP.

Un message Rejeter la session de flux DDP NE DOIT PAS être envoyé si le rejet n'est pas en direction de l'ULP.

6.4 Séquence légale : session active/passive non ULP rejetée

L'acceptation ou le rejet des messages Initier la session de flux DDP DEVRAIT être sous le contrôle de l'ULP. Cela PEUT exiger de passer un événement à l'ULP. Il DOIT y avoir une limite finie au nombre de ces demandes qui sont en attente d'une décision de l'ULP. Quand plus de demandes de session sont reçues, le côté passif DOIT répondre au message Initier avec un message Terminer le flux DDP.

6.5 Séquençage spécifique d'ULP

Une mise en œuvre PEUT choisir de prendre en charge des séquences supplémentaires spécifiques de l'ULP, mais NE DOIT PAS le faire si l'ULP ne le lui demande pas.

Un ULP défini DOIT être capable de fonctionner en utilisant seulement les séquences de session obligatoires définies. Toutes les séquences supplémentaires doivent être utilisées seulement pour des optimisations facultatives.

6.6 Autres règles de séquençage

Un message de contrôle de session de flux DDP NE DOIT PAS être envoyé si il peut être reçu avant un message de contrôle de session de flux DDP antérieur au sein de la même session de flux DDP.

Un côté actif d'une session de flux DDP NE DOIT PAS envoyer de segment DDP qui pourrait être reçu avant le message Initier la session de flux DDP.

Ceci PEUT être déterminé par SCTP lorsque il accuse réception du tronçon de données utilisé pour porter le message Initier la session de flux DDP, ou par la réception d'un message de réponse de contrôle de session de flux DDP.

Un identifiant de flux DDP NE DOIT PAS être réutilisé pour une autre session de flux DDP alors qu'un tronçon de données provenant d'une session antérieure pourrait être en instance.

7. Points d'extrémité SCTP

7.1 Restriction d'indication de couche d'adaptation

L'interface locale DOIT permettre à l'ULP de spécifier un point d'extrémité SCTP pour utiliser une indication d'adaptation spécifique. Elle PEUT exiger que l'ULP le fasse.

Une fois qu'un point d'extrémité décide de ses indications d'adaptation acceptables, il DEVRAIT terminer toutes les demandes d'établir une association avec toute indication d'adaptation différente.

Une mise en œuvre SCTP PEUT choisir de n'accepter les demandes d'association pour un point d'extrémité SCTP que quand une association pour le point d'extrémité a été établie. À ce point, elle PEUT choisir d'interdire toute autre association pour le même point d'extrémité pour utiliser la même indication d'adaptation.

7.2 Implications du multi rattachements

SCTP permet à un point d'extrémité SCTP d'être associé à plusieurs adresses IP, représentant potentiellement des appareils d'interface différents. La distribution de la logique pour un seul flux DDP à travers plusieurs appareils d'entrée peut être très indésirable, résultant en des défis de cohérence de mémoire tampon complexes. Donc, l'interface locale PEUT restreindre les points d'extrémité SCTP à capacité DDP à une seule adresse IP, ou à un ensemble d'adresses IP qui sont toutes allouées au même appareil d'entrée ("RNIC").

Le lien par défaut d'un point d'extrémité SCTP à capacité DDP NE DEVRAIT PAS couvrir plus d'une seule adresse IP sauf si le faire ne résulte ni en trafic de bus supplémentaire ni en duplication de ressources d'enregistrement de mémoire. Cela va fréquemment résulter en un lien par défaut différent de celui pour les points d'extrémité SCTP qui ne sont pas à capacité DDP.

Les applications PEUVENT choisir d'éviter d'utiliser des méthodes hors bande pour communiquer l'ensemble d'adresses IP utilisé par un point d'extrémité SCTP quand il y a une confusion potentielle quand à la portée prévue du point d'extrémité SCTP. Par exemple, en supposant que le point d'extrémité SCTP consiste en ce que toutes les adresses IP annoncées par le DNS peuvent fonctionner pour un point d'extrémité SCTP d'utilisation générale mais pas pour un à capacité DDP.

Même quand le multi rattachements est pris en charge, les ULP sont prévenus qu'ils NE DEVRAIENT PAS utiliser le contrôle d'ULP de l'adresse de source pour tenter un équilibrage de charge d'un flux sur plusieurs chemins. Une mise en œuvre DDP/SCTP receveuse qui choisit de prendre en charge le multi rattachements DEVRAIT optimiser sa conception dans l'hypothèse où le multi rattachements va être utilisé pour la tolérance aux fautes du réseau, et non pour un équilibrage de charge entre les chemins. Ceci est cohérent avec les pratiques recommandées de SCTP.

8. Nombre de flux

Les flux DDP sont bidirectionnels. Ils sont toujours composés en appariant le flux SCTP entrant et sortant avec le même identifiant de flux SCTP.

La couche d'adaptation devrait demander le nombre maximum de flux SCTP qu'elle souhaite utiliser sur la durée de vie de l'association. Le flux SCTP doit quand même être lié aux points d'extrémité DDP, et une association SCTP à capacité DDP ne doit pas prendre en charge les tronçons de données ordonnés. Donc, la simple existence d'un flux SCTP ne va probablement pas exiger des ressources de support significatives.

Cette transposition utilise une association SCTP pour porter un ou plusieurs flux DDP. Chaque flux DDP va être transposé en une paire de flux SCTP avec le même nombre de flux SCTP. L'adaptation DOIT initialiser toutes ses associations SCTP avec le même nombre de flux entrants et sortants.

9. Fragmentation

Un receveur DDP/SCTP traite déjà la fragmentation aux deux couches IP et DDP. Donc, l'utilisation de la segmentation de couche SCTP va être évitée dans la plupart des cas.

Comme un protocole de couche inférieure (LLP) pour DDP, la couche d'adaptation SCTP DOIT informer la couche DDP de la taille maximum de segment DDP qui va être prise en charge. Ce devrait être la plus grande valeur qui peut être prise en charge sans utilisation de la fragmentation IP ou SCTP, ou 516 octets, selon ce qui est le plus grand.

Un minimum de 516 octets est exigé pour permettre un message de contrôle de session de flux DDP avec 512 octets de données privées.

La fragmentation de tronçon de données SCTP NE DOIT PAS être utilisée sauf si la solution de remplacement est la fragmentation IP.

La couche d'adaptation SCTP DEVRAIT régler la taille maximum de segment DDP en dessous du maximum théorique afin de permettre le groupement de tronçons de contrôle dans le même paquet SCTP.

La couche d'adaptation SCTP DOIT rejeter les segments DDP qui font plus que la taille maximum spécifiée.

10. Fonctionnement en séquence non ordonné

La couche d'adaptation DOIT utiliser l'option Non ordonnée sur tous les tronçons de données (fanion U réglé à un). La couche SCTP est supposée livrer les tronçons de données non ordonnés sans délai.

Parce que DDP emploie la livraison SCTP non ordonnée, le receveur NE DOIT PAS s'appuyer sur le numéro de séquence de transmission SCTP (TSN, *Transmission Sequence Number*) pour en déduire un ordre des segments DDP. Le fait que le tronçon de données SCTP pour un segment DDP soit avant le point d'accusé de réception cumulatif ne garantit pas que tous les segments DDP antérieurs ont été placés. L'envoyeur SCTP n'est pas obligé de transmettre des tronçons de données non ordonnés dans l'ordre présenté.

Le DDP-SSN peut être utilisé sans logique particulière pour déterminer la séquence de soumission quand le nombre maximum de messages en cours est inférieur à 32 768. Cela s'applique aussi si l'envoyeur SCTP n'accepte pas plus de 32 767 tronçons de données pour un seul flux sans allouer de TSN.

Si SCTP accepte plus de 32 768 tronçons de données pour un seul flux sans allouer de TSN, l'envoyeur DDP doit simplement s'abstenir d'envoyer plus de 32 767 tronçons de données pour un seul flux sans accusé de réception. Noter qu'il NE DOIT PAS s'appuyer sur le contrôle de flux d'ULP pour cela. Le contrôle de flux d'ULP normal va traiter exclusivement des messages non étiquetés, non des segments DDP.

Le receveur PEUT effectuer une vérification de validité sur les DDP-SSN reçus pour s'assurer qu'aucun trou ne pourrait être pris en compte pour des tronçons de données non reçus. Les mises en œuvre NE DEVRAIENT PAS allouer de ressources avec l'hypothèse que les DDP-SSN sont valides sans d'abord effectuer une telle vérification de validité. Un DDP-SSN invalide PEUT résulter en la terminaison du flux DDP.

11. Procédures

11.1 Initialisation d'association

Au début d'une association, une couche d'adaptation DDP/SCTP DOIT inclure une indication de couche d'adaptation dans son INIT ou INIT-ACK (comme défini au paragraphe 5.1). Après l'échange des deux premiers tronçons SCTP initiaux (INIT et INIT-ACK) un point d'extrémité DOIT vérifier et inspecter l'indication d'adaptation et la comparer au tableau suivant pour déterminer l'action appropriée.

Type d'indication	Action
NONE	Cela indique que l'homologue NE PREND PAS en charge d'adaptation DDP ou RDMA, et donc les procédures de RDMA et DDP NE DOIVENT PAS être effectuées sur cette association.
DDP	Cela indique que l'homologue prend bien en charge la couche d'adaptation DDP/SCTP définie ici.
ANY-OTHER	Cela indique que l'homologue NE PREND PAS en charge l'adaptation DDP, et donc les procédures d'indication DDP NE DOIVENT PAS être effectuées sur cette association.

Une mise en œuvre PEUT exiger que toutes les associations pour un certain point d'extrémité SCTP soient placées dans le même mode.

L'interface locale PEUT permettre à l'ULP d'accepter seulement les demandes d'établissement d'une association dans un mode spécifié.

11.2 Regroupement de tronçons

SCTP permet que plusieurs tronçons de données soient groupés dans un seul paquet SCTP. Les tronçons de données contenant des segments DDP avec des messages non étiquetés NE DEVRAIENT PAS être retardés pour faciliter le groupement. Les tronçons de données contenant des segments DDP avec des messages étiquetés vont généralement être de pleine taille, et donc non susceptibles de groupement. Cependant, des messages étiquetés de taille partielle PEUVENT être retardés, car ils sont fréquemment suivis par un court message non étiqueté.

11.3 Terminaison d'association

La terminaison d'une association SCTP due à des erreurs devrait être traitée à la couche SCTP. Le message RDMAP Terminé défini par RDMAP NE DEVRAIT PAS être envoyé sur chaque flux DDP quand il a été décidé de terminer une association SCTP. L'envoi de ce message sur chaque flux SCTP pourrait sévèrement retarder la terminaison de l'association.

L'interface locale DEVRAIT notifier à tous les consommateurs de flux DDP quand le flux SCTP sous-jacent a été terminé.

Les autres messages Terminé définis par RDMAP DOIVENT être générés comme spécifié quand un flux DDP est terminé. Noter qu'avec la transposition SCTP, la terminaison d'un flux DDP ne rend pas obligatoire la terminaison de l'association.

12. Considérations relatives à l'IANA

Le présent document définit un nouveau codet d'indication de couche d'adaptation SCTP pour DDP (0x00000001). La [RFC5061] crée le registre dans lequel ce codet a été alloué.

Le présent document définit aussi deux nouveaux identifiants de protocole de charge utile SCTP (PPID). La [RFC4960] crée le registre dans lequel ces identifiants ont été alloués. Les valeurs suivantes sont allouées :

Tronçon de segment DDP - 16

Contrôle de session de flux DDP - 17

13. Considérations sur la sécurité

Tout placement direct en mémoire pourrait poser un risque de sécurité significatif si des contrôles locaux adéquats ne sont pas fournis. Ces menaces sont traitées dans les documents DDP [RFC5041], RDMA [RFC5040], ou de sécurité [RFC5042]

appropriés. Le présent document n'ajoute aucun risque supplémentaire pour la sécurité à ceux de la [RFC4960].

Les exigences de IPsec pour le placement direct de données à distance (RDDP) se fondent sur la version de IPsec spécifiée dans la [RFC2401] et les RFC en rapport, et le profil de la [RFC3723], en dépit de l'existence d'une version plus récente de IPsec spécifiée dans la [RFC4301] et les RFC en rapport. Une des applications précoces importantes des protocoles RDDP est leur utilisation avec iSCSI iSER [RFC5046] ; les exigences de IPsec de RDDP suivent celles de IPsec afin de faciliter cet usage en permettant qu'un profil commun de IPsec soit utilisé avec iSCSI et les protocoles RDDP. À l'avenir, la RFC3723 pourrait être mise à jour avec la nouvelle version de IPsec ; les exigences de sécurité de IPsec d'une telle mise à jour devraient s'appliquer également à iSCSI et aux protocoles RDDP.

Des exigences supplémentaires s'appliquent à la sécurité pour RDDP sur SCTP, à cause de l'utilisation de SCTP comme protocole de transport. Une mise en œuvre de IPsec pour RDDP sur SCTP :

- 1) DOIT prendre en charge une fonctionnalité IPsec pour SCTP équivalente à la fonctionnalité IPsec pour TCP qui est exigée par la RFC 3723,
- 2) DEVRAIT prendre en charge le même niveau de fonctionnalité IPsec pour SCTP et TCP sauf si il n'y a pas de prise en charge pour TCP, et
- 3) DOIT prendre en charge au moins le niveau de fonctionnalité de sélecteur de protocole et d'accès pour SCTP que celui pris en charge pour TCP.

14. Contributeurs

Merci à nos contributeurs qui ont passé de nombreuses heures à lire et réviser ce document pour nous garder dans le droit chemin : Sukanta Ganguly, consultant indépendant, Vivek Kashyap de IBM, Jim Pinkerton de Microsoft, et Hemal Shah de Broadcom. Merci pour tout ce dur travail.

15. Remerciements

Les auteurs tiennent à remercier les personnes suivantes qui ont fourni des commentaires et des remarques : Stephen Bailey, David Black, Douglas Otis, Allyn Romanow, et Jim Williams.

16. Références

16.1 Références normatives

- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))
- [RFC3723] B. Aboba et autres, "Protocoles de [sécurisation de mémorisation de blocs](#) sur IP", avril 2004. (P.S.)
- [RFC4960] R. Stewart, éd., "[Protocole de transmission de contrôle de flux](#) (SCTP)", septembre 2007. (Remplace [RFC2960](#), [RFC3309](#) ; P.S. ; Remplacée par [RFC9260](#))
- [RFC5040] R. Recio et autres, "[Spécification d'un protocole d'accès direct](#) à une mémoire distante", octobre 2007. (P.S. ; MàJ par [RFC7146](#))
- [RFC5041] H. Shah et autres, "[Placement direct des données](#) sur transports fiables", octobre 2007. (P.S. ; MàJ par [RFC 7146](#))
- [RFC5042] J. Pinkerton, E. Deleganes, "[Sécurité du protocole de placement direct](#) des données (DDP) / protocole d'accès direct à une mémoire distante (RDMAP)", octobre 2007. (P.S. ; MàJ par [RFC7146](#))
- [RFC5061] R. Stewart et autres, "[Reconfiguration dynamique d'adresse](#) pour le protocole de transmission de contrôle de flux (SCTP)", septembre 2007. (P.S.)

16.2 Références pour information

- [RFC2401] S. Kent et R. Atkinson, "[Architecture de sécurité](#) pour le protocole Internet", novembre 1998. (*Obsolète, voir RFC4301*)
- [RFC4301] S. Kent et K. Seo, "[Architecture de sécurité](#) pour le protocole Internet", décembre 2005. (*P.S.*) (*Remplace la RFC2401*)
- [RFC5044] P. Culley et autres, "[Tramage verrouillé sur la PDU](#) de marqueur pour la spécification de TCP", octobre 2007. (*P.S.* ; *MàJ par RFC6581, RFC7146*)
- [RFC5046] M. Ko et autres, "[Extensions pour l'accès direct](#) à une mémoire distante (RDMA) à l'interface système de petit ordinateur à l'Internet (iSCSI)", octobre 2007. (*P.S.* ; *Remplacée par RFC7145*)

Adresse des auteurs

Caitlin Bestler (éditeur)
Neterion
20230 Stevens Creek Blvd.
Suite C
Cupertino, CA 95014
USA
téléphone : 408-366-4639
mél : caitlin.bestler@neterion.com

Randall R. Stewart (éditeur)
Cisco Systems, Inc.
Forest Drive
Columbia, SC 29036
USA
téléphone : +1-815-342-5222
mél : rrs@cisco.com

Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2007)

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et sauf pour ce qui est mentionné ci-après, 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, le IETF TRUST 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 encloses 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.