

Groupe de travail Réseau
Request for Comments : 3081
 Catégorie : En cours de normalisation
 Traduction Claude Brière de L'Isle

M. Rose
 Invisible Worlds, Inc.
 mars 2001

Transposition du cœur BEEP en TCP

Statut de ce mémoire

Le présent document spécifie un protocole Internet en cours de normalisation pour la communauté de l'Internet, et appelle à des discussions et des suggestions pour son amélioration. Prière de se reporter à l'édition actuelle du STD 1 "Normes des protocoles officiels de l'Internet" pour connaître l'état de 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 (2001). Tous droits réservés

Résumé

Le présent mémoire décrit comment une session du protocole d'échange de blocs extensibles (BEEP, *Blocks Extensible Exchange Protocol*) est transposée en une seule connexion du protocole de contrôle de transmission (TCP, *Transmission Control Protocol*).

Table des Matières

1. Introduction.....	1
2. Gestion de session.....	1
3. Échange de message.....	2
3.1 Contrôle de flux.....	2
3.1.1 Création de canal.....	2
3.1.2 Envoi des messages.....	2
3.1.3 Traitement des trames SEQ.....	2
3.1.4 Utilisation du contrôle de flux.....	3
4. Considérations sur la sécurité.....	3
Références.....	4
Adresses de l'auteur.....	4
Appendice A Remerciements.....	4
Déclaration complète de droits de reproduction.....	4

1. Introduction

Le présent mémoire décrit comment une session BEEP [RFC3080] est transposée en une seule connexion TCP [RFC0793]. Se reporter au paragraphe 2.5 de la [RFC3080] pour une explication des exigences de la transposition.

2. Gestion de session

La transposition de la gestion de session BEEP sur le service TCP est directe.

Une session BEEP est établie lorsque une session TCP est établie entre deux homologues BEEP :

- o l'homologue BEEP qui produit un appel passif TCP OPEN est appelé l'écouteur ;
- o l'homologue BEEP qui produit un appel active TCP OPEN est appelé l'initiateur.

Des appels TCP OPEN simultanés résulteraient en ce que les deux homologues BEEP croiraient qu'ils sont l'initiateur et aucun des homologues ne serait capable de démarrer un canal. À cause de cela, les services fondés sur BEEP doivent être conçus de telle façon que des TCP OPEN simultanés ne puissent pas se produire.

Si les deux homologues s'accordent pour libérer une session BEEP (c.f. le paragraphe 2.4 de la [RFC3080]) l'homologue qui envoie la réponse "ok" produit immédiatement l'appel TCP CLOSE. À réception de la réponse, l'autre homologue produit immédiatement l'appel TCP CLOSE.

Une session BEEP se termine lorsque l'un ou l'autre homologue produit l'appel TCP ABORT, et la connexion TCP est ensuite interrompue.

3. Échange de message

La transposition des échanges BEEP sur le service TCP est moins directe.

Les messages sont envoyés de façon fiable et reçus en utilisant les appels TCP SEND et RECEIVE. (Cela assure aussi une livraison ordonnée des messages sur le même canal.)

Bien que TCP impose le contrôle de flux sur la base de la connexion, si plusieurs canaux sont utilisés simultanément dans une session BEEP, BEEP doit fournir un mécanisme pour éviter la famine et le blocage. Pour ce faire, BEEP réintroduit un mécanisme utilisé par TCP : le contrôle de flux fondé sur la fenêtre – chaque canal a une fenêtre glissante qui indique le nombre d'octets de charge utile qu'un homologue peut transmettre avant de recevoir une autre permission.

3.1 Contrôle de flux

On rappelle d'après le paragraphe 2.2.1.2 de la [RFC3080] que chaque octet de charge utile envoyé dans chaque direction sur un canal est associé à un numéro de séquence. La numérotation des octets de charge utile au sein d'une trame de données est telle que le premier octet de charge utile a le numéro le plus faible, et que les octets de charge utile suivants sont numérotés en ordre croissant.

L'espace réel des numéros de séquence est fini, bien que très grand, allant de 0 à 4 294 967 295 ($2^{32} - 1$). Comme l'espace est fini, toute arithmétique qui traite les numéros de séquence est effectuée modulo 2^{32} . Cette arithmétique non signée préserve les relations des numéros de séquence car ils effectuent un cycle de $2^{32} - 1$ à 0 à nouveau. Consulter les sections 2 à 5 de la [RFC1982] pour un exposé sur les propriétés arithmétiques des numéros de séquence.

3.1.1 Création de canal

Lorsque un canal est créé, le numéro de séquence associé au premier octet de charge utile de la première trame de données est 0, et la taille de fenêtre initiale pour ce canal est de 4096 octets. Après la création du canal, un homologue BEEP peut mettre à jour la taille de fenêtre en envoyant une trame SEQ (paragraphe 3.1.3).

Si il est demandé à un homologue BEEP de créer un canal et si il n'est pas capable d'allouer au moins 4096 octets à ce canal, il doit refuser la création du canal, comme spécifié au paragraphe 2.3.1.2 de la [RFC3080]. De même, durant l'établissement de la session BEEP, si l'homologue BEEP qui agit dans le rôle d'écouteur n'est pas capable d'allouer au moins 4096 octets au canal 0, il doit alors retourner une réponse négative, comme spécifié au paragraphe 2.4 de la [RFC3080], au lieu d'un accord.

3.1.2 Envoi des messages

Avant l'envoi d'un message, l'homologue BEEP envoyeur doit s'assurer que la taille de la charge utile est dans la fenêtre annoncée par l'homologue BEEP receveur. Si non, il a le choix suivant :

- o si la fenêtre permettrait qu'au moins un octet de charge utile soit envoyé, l'homologue BEEP peut segmenter le message et commencer par envoyer une plus petite trame de données (jusqu'à la taille de fenêtre restante) ;
- o l'homologue BEEP peut différer l'envoi du message jusqu'à ce que la fenêtre soit assez grande; ou,
- o l'homologue BEEP peut signaler à son application qu'il est dans l'incapacité d'envoyer le message, permettant à l'application d'essayer ultérieurement (ou peut-être de signaler à son application quand une fenêtre plus grande est disponible).

Le choix appartient à la mise en œuvre, bien qu'il soit recommandé que l'application qui utilise BEEP ait un mécanisme pour influencer la décision.

3.1.3 Traitement des trames SEQ

Lorsque une application accepte la responsabilité des trames de données entrantes, son homologue BEEP devrait envoyer des trames SEQ pour annoncer une nouvelle fenêtre.

L'ABNF [RFC2234] pour une trame SEQ est :

```
seq    = "SEQ" SP channel SP ackno SP window CR LF
ackno  = seqno
window = size                ; channel, seqno, et size sont définis au paragraphe 2.2.1 de la [RFC3080].
```

La trame SEQ a trois paramètres :

- o un numéro de canal ;
- o un numéro d'accusé de réception, qui indique la valeur du prochain numéro de séquence que l'expéditeur s'attend à recevoir sur ce canal ;
- o une taille de fenêtre, qui indique le nombre d'octets de charge utile commençant par celui indiqué par le numéro d'accusé de réception que l'expéditeur s'attend à recevoir sur ce canal.

Un seul caractère espace (code décimal 32, " ") sépare chaque composant. La trame SEQ se termine par une paire CRLF.

Lorsque une trame SEQ est reçue, si un des numéros de canal, numéro d'accusé de réception, ou taille de fenêtre ne peut pas être déterminé ou est invalide, la session BEEP se termine alors sans générer de réponse, et il est recommandé qu'une entrée de diagnostic soit enregistrée.

3.1.4 Utilisation du contrôle de flux

La clé de la réussite de l'utilisation du contrôle de flux au sein de BEEP est d'équilibrer performances et équité :

- o les gros messages devraient être segmentés en trames pas plus grandes que les deux tiers de la taille maximum de segment TCP négociée ;
- o les trames pour des canaux différents avec du trafic prêt à envoyer devraient être envoyées en round-robin ;
- o chaque fois d'une trame est reçue, une trame SEQ devrait être envoyée chaque fois que la taille de fenêtre qui sera envoyée est au moins de la moitié de l'espace de mémoire tampon disponible pour ce canal ;
- o si le service de transport présente plusieurs trames simultanément à un homologue BEEP, une seule trame SEQ de consolidation peut être envoyée.

Pour éviter des interactions inattendues avec le service de transport, il est important qu'un homologue BEEP annonce les fenêtres sur la base de l'espace de mémoire tampon disponible, pour permettre que les données soient lues du service de transport aussitôt qu'elles sont disponibles. De plus, les trames SEQ pour un canal doivent avoir une priorité supérieure à celles des messages pour ce canal.

Les mises en œuvre peuvent souhaiter fournir des facilités de gestion de file d'attente pour les applications qui utilisent BEEP, par exemple, des priorités de canal, des allocations (relatives) de mémoire tampon, et ainsi de suite. En particulier, les mises en œuvre ne devraient pas permettre qu'un certain canal monopolise la fenêtre de transport sous-jacente (par exemple, les lecteurs lents devraient obtenir de petites fenêtres).

De plus, lorsque possible, les mises en œuvre devraient prendre en charge les API de couche transport qui convoient les informations d'encombrement. Ces API permettent à une mise en œuvre de déterminer sa part de la bande passante disponible, et aussi d'avoir notification des changements d'estimation de la bande passante du chemin. Noter que lorsque une session BEEP a plusieurs canaux qui échangent simultanément de grands messages, les mises en œuvre sans accès à ces informations peuvent avoir des propriétés incertaines d'équité et de progrès durant les moments d'encombrement du réseau.

Finalement, les mises en œuvre devraient suivre les lignes directrices données dans les portions pertinentes de la [RFC1122] qui traitent du contrôle de flux (et garder présent à l'esprit que des questions comme la retransmission, lors d'interactions avec le contrôle de flux dans TCP, ne sont pas applicables à ce mémoire). Par exemple, le paragraphe 4.2.2.16 de la RFC1122 [5] indique qu'un "receveur NE DEVRAIT PAS restreindre la fenêtre, c'est-à-dire, déplacer le bord droit de la fenêtre vers la gauche" et expose ensuite l'impact de cette règle sur les données non acquittées. Dans le contexte de la transposition de BEEP sur une seule connexion TCP, seule les portions concernant le contrôle de flux devraient être mises en œuvre.

4. Considérations sur la sécurité

Consulter la Section 9 de la [RFC3080] pour l'exposé sur les questions de sécurité.

Références

- [RFC0793] J. Postel (éd.), "Protocole de [commande de transmission](#) – Spécification du protocole du programme Internet DARPA", STD 7, septembre 1981.
- [RFC1122] R. Braden, "[Exigences pour les hôtes Internet](#) – couches de communication", STD 3, octobre 1989. (MàJ par la RFC6633)
- [RFC1982] R. Elz, R. Bush, "[Arithmétique des numéros de série](#)", août 1996. (MàJ RFC1034, RFC1035) (P.S.)
- [RFC2234] D. Crocker et P. Overell, "[BNF augmenté pour les spécifications de syntaxe](#) : ABNF", novembre 1997. (Obsolète, voir RFC5234)
- [RFC3080] M. Rose, "Cœur du [protocole extensible d'échange de blocs](#) (BEEP)", mars 2001. (P.S.)

Adresses de l'auteur

Marshall T. Rose
Invisible Worlds, Inc.
1179 North McDowell Boulevard
Petaluma, CA 94954-6559
USA
téléphone : +1 707 789 3700
mél : mrose@invisible.net
URI : <http://invisible.net/>

Appendice A Remerciements

L'auteur remercie de leurs contributions Dave Crocker, Steve Harris, Eliot Lear, Keith McCloghrie, Craig Partridge, Vernon Schryver, et Joe Touch. En particulier, Dave Crocker a fourni d'utiles suggestions sur la nature du contrôle de flux dans la transposition.

Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2001). Tous droits réservés.

Le présent document et ses traductions peuvent être copiés et fournis aux tiers, et les travaux dérivés qui les commentent ou les expliquent ou aident à leur mise en œuvre peuvent être préparés, copiés, publiés et distribués, en tout ou partie, sans restriction d'aucune sorte, pourvu que la déclaration de droits de reproduction ci-dessus et le présent paragraphe soient inclus dans toutes telles copies et travaux dérivés. Cependant, le présent document lui-même ne peut être modifié d'aucune façon, en particulier en retirant la notice de droits de reproduction ou les références à la Internet Society ou aux autres organisations Internet, excepté autant qu'il est nécessaire pour le besoin du développement des normes Internet, auquel cas les procédures de droits de reproduction définies dans les procédures des normes Internet doivent être suivies, ou pour les besoins de la traduction dans d'autres langues que l'anglais.

Les permissions limitées accordées ci-dessus sont perpétuelles et ne seront pas révoquées par la Internet Society ou ses successeurs ou ayant 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 encloses ne violent aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Remerciement

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