

RFC 7786 : TCP modifications for Congestion Exposure

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 15 mai 2016

Date de publication du RFC : Mai 2016

<https://www.bortzmeyer.org/7786.html>

C'est après une très longue genèse (ce document avait été adopté par le groupe de travail il y a plus de quatre ans) que voici les protocoles concrets pour la signalisation de la congestion aux routeurs situés en **aval**. Ici, l'utilisation de ce système avec TCP.

ConEx (pour "*Congestion Exposure*", signalement de la congestion) est décrit dans le RFC 6789¹. L'idée de base est d'informer le réseau qu'un flot de données rencontre de la congestion, afin que les éléments actifs de ce réseau puissent prendre des décisions intelligentes. ConEx est divisé en une spécification abstraite (RFC 7713) et une ou plusieurs spécifications de protocoles concrets. Ce RFC ne spécifie pas un encodage sur le câble, juste le changement de comportement des mises en œuvre de TCP. (Un autre RFC concret spécifie un encodage pour IPv6, le RFC 7837.)

Il n'y a pas besoin de négociation au début (section 1 du RFC), comme avec les options TCP. Les émetteurs qui connaissent ConEx utilisent les informations existantes (pertes de paquets et ECN, stockées dans deux compteurs différents, cf. la section 3 du RFC).

Pour bien utiliser ConEx, le TCP de l'émetteur a besoin d'informations. Bien qu'il ne soit pas indispensable, c'est mieux si SACK (RFC 2018) est disponible.

La section 2 du RFC résume les modifications chez l'émetteur TCP (le récepteur n'a pas besoin d'être changé pour ConEx). Le comportement de l'émetteur dépendra du fait que le récepteur met en œuvre (ou pas) SACK et ECN. L'émetteur est responsable du marquage des paquets (suivant l'encodage du RFC 7837) avec des signaux ConEx. Il doit tenir compte du nombre d'octets, pas du nombre de paquets (RFC 7141).

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc6789.txt>

En cas de perte de paquets, l'émetteur va mettre le signal ConEx L ("*loss experienced*"). Dans certains cas (retransmission inutile parce que le paquet était bien arrivé, c'est l'accusé de réception qui s'était perdu), l'émetteur ne connaît pas le nombre exact d'octets perdus. Il peut donc surestimer la congestion dans les signaux ConEx qu'il envoie. Pour avoir de meilleures mesures, l'émetteur peut toujours mettre en œuvre les RFC 5682, RFC 3708 ou RFC 3522.

Si c'est par des marques ECN, et non pas par des pertes de paquets, que l'émetteur a appris qu'il y avait congestion, il va utiliser les marques ConEx E ("*ECN experienced*").

Dans tous les cas (perte de paquets ou bien ECN), les paquets marqués ConEx devront également porter une marque X, qui indique que l'émetteur sait faire du ConEx (section 4 du RFC).

Un émetteur ConEx est aussi censé envoyer des **crédits** lorsqu'il n'y a **pas** de congestion. Ces crédits seront ensuite « consommés » lors des épisodes de congestion. Cela se fait avec le signal C ("*Credit*").

Les paquets contenant les signaux ConEx peuvent se perdre, comme tous les paquets IP (section 5 du RFC). Cela peut mener à des pénalités injustes (un émetteur détecte qu'il y a congestion en aval, le signale avec un L ou un E, le paquet portant le signal est perdu, un routeur qui fait de l'audit ConEx se dit alors « ah, ah, il essaie de tricher, il n'a pas signalé la congestion »). Le problème étant du second ordre (si la probabilité de perdre un paquet est P, la probabilité qu'il y ait perte du paquet **et** perte du signal est de P au carré), on peut choisir de l'ignorer.

Plus délicat, le problème de la fraîcheur des informations ConEx (section 6 du RFC). Ces informations ne sont utiles que si elles sont très récentes (typiquement moins d'un RTT depuis que la congestion est apparue). L'émetteur doit donc faire attention à ne pas retarder les signaux ConEx. Parfois, il n'a pas trop le choix, par exemple si l'application arrive d'envoyer des données, il n'y aura pas de paquets à marquer.