

RFC 6789 : ConEx Concepts and Use Cases

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 6 décembre 2012

Date de publication du RFC : Décembre 2012

<https://www.bortzmeyer.org/6789.html>

Premier RFC du groupe de travail CONEX <<http://tools.ietf.org/wg/conex>> de l'IETF, ce document expose les principes de bases du mécanisme de « publication de la congestion » ("*CONgestion EXposure*"). L'idée est d'indiquer dans les paquets la congestion prévue, pour que les équipements réseaux puissent prendre des décisions appropriées. Une fois normalisée (ce RFC 6789¹ n'est qu'une toute première étape), ce sera une des nombreuses techniques permettant de gérer la congestion dans les réseaux TCP/IP et notamment l'Internet.

Il n'y a pas encore de protocole normalisé pour cette « publication de la congestion » (les deux premiers utiliseront une option IPv6 et une option de TCP). Pour l'instant, notre RFC s'en tient à des usages : à quoi cela pourrait servir. La congestion est clairement une des plaies du réseau. La force de TCP/IP est de multiplexer la capacité des liens au niveau des paquets et pas des circuits. Cela permet une utilisation bien plus efficace du réseau. Mais cela ne crée pas magiquement de la capacité : quand trop de paquets rencontrent trop peu de capacité, la congestion survient. Elle se manifeste par des pertes de paquets (le routeur, n'arrivant pas à suivre, abandonne certains paquets), mais aussi par des retards (si les tampons du routeur amortissent la congestion, ils augmentent les délais d'acheminement), et par le marquage ECN (RFC 3168) de certains paquets.

Le récepteur des données est censé détecter ces signaux (un trou dans la séquence TCP, des RTT qui augmentent, les marques ECN) et il doit alors dire à l'émetteur de ralentir. Ce fonctionnement où seules les machines terminales agissent (les routeurs intermédiaires, à part éventuellement des marques ECN, n'ont pas grand'chose à faire, ils se contentent de router) est typique de l'Internet mais, dans certains cas, exposés dans ce RFC, il est insuffisant.

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

L'idée de base de ConEx est que l'émetteur mette dans les paquets IP qu'il envoie des indications sur la congestion qu'on lui a signalé. Ainsi, des équipements qui sont purement de niveau 3 sur le trajet peuvent être informés de la congestion (ECN les informe aussi mais ne concerne que les équipements en **aval** du point où est détectée la congestion, cf. figure 1 du RFC). Ainsi, le réseau pourra avoir une meilleure vision de la congestion, comptant les paquets qui vont rencontrer de la congestion aussi facilement qu'il compte aujourd'hui les paquets ou les octets.

Pourquoi est-ce important de connaître les paquets qui vont circuler dans des parties du réseau où il y a congestion? Déjà, cela a une importance économique : un réseau ne coûte pas plus cher qu'il soit utilisé ou pas. Un trafic qui emprunte le réseau à un moment où celui-ci est peu chargé ne coûte donc rien à l'opérateur du réseau. Par contre, la congestion coûte cher puisqu'elle est le signe qu'il faut sortir son carnet de chèques pour mettre à jour le réseau, vers des capacités plus grandes. Ainsi, ConEx pourrait être une brique de base pour un système qui pénaliserait uniquement les flots de données qui contribuent à la congestion.

La section 2 présente en détail les concepts et la terminologie ConEx. D'abord, évidemment, la **congestion**. Tout le monde en parle mais il n'y a pas de définition rigoureuse et unique pour ce concept. Pour une analyse de ce terme, voir « *"The Evolution of Internet Congestion"* <http://mitas.csail.mit.edu/papers/Bauer_Clark_Lehr_2009.pdf> » de S. Bauer, S., D. Clark et W. Lehr. Dans ConEx, la congestion est définie comme la probabilité de perte de paquet (ou de marquage par ECN). Un autre effet de la congestion, le retard, n'est pas pris en compte (donc, ConEx ne mesurera pas l'effet du "bufferbloat").

Deuxième concept, le **volume de congestion**. Il s'agit du nombre d'octets perdus suite à la congestion. L'idée est de rendre les émetteurs responsables : envoyer 10 Go sur un lien congestionné n'est pas la même chose que d'y envoyer quelques octets. Par exemple, si Alice envoie un fichier d'un Go sur un lien où la probabilité de perte est de 0,2 %, son volume de congestion est de 2 Mo. Si, plus tard dans la journée, elle envoie un fichier de 30 Go alors que la ligne, moins encombrée, ne perd plus que 0,1 % des paquets, elle ajoutera 3 Mo à son volume de congestion (donc, 5 Mo en tout). L'un des intérêts de cette métrique est qu'elle est très facilement mesurable par un routeur : il lui suffit de compter le volume total des paquets qu'il a dû jeter (ou marquer avec ECN, s'il utilise cette technique). C'est donc quasiment la même chose que les compteurs de volume actuels.

Troisième concept important, la **congestion aval** ("*Rest-of-path congestion*" ou "*downstream congestion*" dans la langue de Van Jacobson). C'est celle que le flot de données va supporter entre le point de mesure et sa destination (la congestion amont étant évidemment celle entre la source et le point de mesure : aujourd'hui, elle est bien plus difficile à mesurer si on n'utilise pas ECN).

La section 3 décrit ensuite le principal scénario d'usage envisagé pour ConEx. Si vous vous demandez « mais à quoi sert ce truc? » depuis tout à l'heure, vous allez bientôt avoir une réponse. Soit un opérateur réseau qui a des clients. La plupart consomment peu mais certains sont des acharnés et font passer en permanence des gigaoctets à travers le réseau. L'opérateur voudrait que les mesures désagréables qu'il va prendre pour limiter la congestion frappent en priorité ces gros consommateurs.

Il va pour cela placer un équipement qui regarde les indications de congestion ConEx et va ensuite limiter le trafic en fonction de ces indications (en fonction du volume de congestion). (Les limitations de trafic habituelles visent le débit, pas la participation à la congestion. Or comme on l'a vu, débit 100 Mb/s sur un réseau vide n'est pas du tout la même chose que de le débit sur un réseau chargé. Il est donc dommage de limiter le débit dans tous les cas.) L'article « *"Policing Freedom to Use the Internet Resource Pool"* <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.159.7151&rep=rep1&type=pdf>> » de B. Briscoe, A. Jacquet et T. Moncaster, détaille ce mécanisme. La force de cette méthode est qu'aucune limite, aucune mesure contraignante et désagréable, n'est prise lorsqu'il

n'y a pas de congestion (car, dans ce cas, l'usage du réseau ne coûte rien). Avec ConEx, un bit n'est plus égal à un bit : celui envoyé pendant la congestion coûte plus cher au réseau et doit donc, d'une certaine façon, être « payé » par l'utilisateur.

Si l'opérateur est un FAI ADSL, un bon endroit pour mettre cette supervision de la congestion et cette limitation du trafic est sans doute le BRAS. (Voir le rapport du DSL Forum « *Technical Report TR-059 : Requirements for the Support of QoS-Enabled IP Services* » mais il ne semble pas en ligne.)

Ainsi, les utilisateurs seront encouragés à utiliser des protocoles que notre RFC appelle « charognards » dans la mesure où ils n'utilisent que la capacité dont personne ne veut (le charognard a une mauvaise image de marque mais c'est bien à tort : il consomme les ressources négligées par tous les autres). Ces protocoles n'envoient du trafic que lorsque le réseau est inutilisé. Leur coût est donc à peu près nul pour l'opérateur. Le RFC 6297 décrit ces protocoles. Pour encourager leur usage, on ne peut pas seulement compter sur des exhortations citoyennes, dont l'écologie montre le peu d'efficacité. Il faut aussi des motivations plus pratiques.

Mais il existe d'autres approches, déjà, pour ce genre de gestion (ralentir les gros consommateurs, ceux qui sortent franchement du rang), non ? La section 3.3 les examine :

- On l'a vu, il est courant de limiter le débit, ou bien d'imposer un total de données transféré maximum (cette dernière technique est courante en 3G). Cette approche est très grossière car elle mène à sous-utiliser ou sur-utiliser le réseau. S'il n'y a pas de congestion, on va sous-utiliser le réseau, alors qu'il est libre (pire, on décourage l'utilisation de protocoles « charognards » puisque leur utilisateur serait pénalisé pour le volume envoyé, au lieu d'être récompensé pour le soin qu'il met à ne l'envoyer qu'en l'absence de congestion). S'il y a congestion, on risque de sur-utiliser le réseau car, tant que le quota n'est pas atteint, l'utilisateur n'a pas de raison de se gêner.
- Les techniques de "*fair queuing*" visent à égaliser le trafic entre les différents utilisateurs. Agissant en temps réel et n'ayant pas de mémoire, elles favorisent, sur le long terme, les plus gros consommateurs alors qu'on voudrait le contraire.
- Plus subtile, la technique décrite dans le RFC 6057 ne mesure le débit que lorsqu'il y a congestion et donc effectivement gêne pour le réseau. Mais, lors de la congestion, tout le trafic est compté comme y contribuant et cette technique, comme les deux précédentes, n'encourage absolument pas à utiliser les protocoles « charognards ».
- Dernière méthode possible, un examen plus ou moins poussé du paquet (allant jusqu'à la DPI) pour décider quelle est l'application utilisée, suivi d'une dégradation du service pour cette application particulière (par exemple BitTorrent). Elle est évidemment incompatible avec des techniques de sécurité comme IPsec et elle soulève de sérieuses questions politiques <<https://www.bortzmeyer.org/neutralite.html>>. ConEx, lui, est neutre par rapport à l'application utilisée.

Toutes ces solutions ont un autre défaut en commun : elles laissent l'utilisateur dans l'incertitude des performances qu'il obtiendra. Le principal intérêt des tarifs forfaitaires pour l'accès Internet (je parle des vrais forfaits, ceux où le prix est connu d'avance, pas de ce que les commerciaux menteurs de la téléphonie mobile appellent « forfait ») est la certitude : pas de mauvaises surprises au moment de l'arrivée de la facture. Mais cette certitude ne porte que sur les prix. Les performances, elles, peuvent varier, par exemple si le trafic émis était tel qu'on est désormais limité. Au contraire, avec ConEx, l'utilisateur est en mesure de voir sa propre contribution à la congestion, et peut ajuster son comportement en fonction de cela. (À mon avis, c'est une vision un peu idéale, car elle suppose que le réseau n'agit sur les performances qu'en réponse aux indicateurs ConEx.)

Ce scénario d'usage de la section 3, informer l'opérateur sur la contribution des utilisateurs à la congestion, est le principal but de ConEx à l'heure actuelle. Les versions initiales de ce RFC traitaient sur un pied d'égalité plusieurs scénarios mais le groupe de travail, après bien des discussions, a décidé de prioriser celui qui semblait le plus prometteur. Ceci dit, la section 4 présente quelques autres cas d'usage :

- Mesurer la contribution à la congestion non pas par utilisateur mais par opérateur. Comme les indicateurs ConEx sont vus par tous les routeurs intermédiaires (y compris les routeurs de bordure entre opérateurs), chaque opérateur peut savoir ce que chacun de ses voisins va lui apporter comme congestion (et réciproquement). De beaux débats sur la répartition des investissements à prévoir!
- Avoir de meilleures indications sur l'utilisation globale du réseau, pour mieux planifier l'avitaillement ("*provisioning*") en capacité.

La section 5 se penche ensuite sur le déploiement de ConEx. Ce système ne nécessite pas un accord simultané de tous les acteurs (heureusement). ConEx est donc utile même si certains routeurs du trajet ne marquent pas et même si certains ne lisent pas l'information ConEx.

Par contre, il faut qu'au moins deux acteurs bougent : au moins un routeur doit marquer les paquets et au moins une machine doit réagir et signaler la congestion dans les paquets qu'elle émet. Les développeurs d'une application Ledbat (voir RFC 6297 et RFC 6817), par exemple, peuvent se lancer en espérant que cela motivera les routeurs pour s'y mettre.

Si certains paquets sont marqués par ConEx et d'autres non, que va faire le routeur ? Pour le déploiement de ConEx, il serait préférable que les paquets marqués soient favorisés d'une manière ou d'une autre. Mais comme la raison du marquage est leur contribution à la congestion, ce n'est pas forcément idéal... En outre, cela soulève un problème de sécurité (le marquage n'est pas authentifié et une machine peut mentir) sur lequel ce RFC 6789 est complètement muet (le problème sera traité dans les futurs RFC, plus concrets).

Comme ConEx va tripoter dans la couche 3, cœur de l'Internet, il risque de perturber le fonctionnement d'IP. Il faut donc considérer ConEx comme plutôt expérimental et la section 6 fait la liste des questions encore ouvertes, par exemple :

- ConEx est actuellement limité à IPv6, l'idée est que le délai de déploiement sera de toute façon tel qu'IPv6 sera largement dominant lorsque ConEx sera généralisé. Était-ce une bonne idée ?
- ConEx ne fait pas de différence entre l'auto-congestion (un utilisateur qui se gêne lui-même) et la congestion infligée aux autres (lorsqu'on encombre des ressources partagées). Moralement, il serait intéressant de séparer les deux mais n'est-ce pas trop complexe ?
- ConEx est plus ambitieux qu'ECN et s'appuie sur lui. Mais le déploiement d'ECN est très limité. Est-ce que ConEx sera quand même utile si le déploiement d'ECN en reste où il est maintenant ?