

La grande panne DNS de Chine de mai 2009

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 6 Novembre 2009

<http://www.bortzmeyer.org/panne-dns-chine.html>

Le 19 mai 2009, la Chine a connu sa plus grande panne de l'Internet. Sur le moment, de nombreux articles ont été publiés, sans détails pratiques la plupart du temps, à part le fait qu'il s'agirait d'un problème lié au DNS. Le 5 novembre, à la réunion OARC (<http://www.dns-oarc.net>) à Pékin, Ziqian Lu, de China Telecom, a fait un remarquable exposé (https://www.dns-oarc.net/files/workshop-200911/Ziqian_Liu.pdf) détaillant les causes de la panne.

C'est un bel exercice de transparence, avec plein de détails techniques. Je ne suis pas sûr que les opérateurs Internet français en fassent autant, si une telle panne frappait la France.

Donc, le 19 mai vers 21 h, heure locale, les téléphones se mettent à sonner : « l'Internet est en panne ». China Telecom, mais également d'autres FAI, constate à la fois que les utilisateurs se plaignent mais que le trafic a chuté considérablement. Les « tuyaux » ne sont donc pas surchargés, bien au contraire. En outre, le service n'est pas complètement interrompu : parfois, cela marche encore un peu. En raison de la baisse du trafic, on soupçonne un problème dans le routage. Il faut un certain temps pour que quelqu'un remarque ces lignes dans le journal des serveurs de noms récursifs du FAI :

```
19-May-2009 22:21:13.186 client: warning: client 218.77.186.180#51939: recursive-clients soft limit exceeded, abor
19-May-2009 22:21:13.213 client: warning: client 59.50.182.161#1151: recursive-clients soft limit exceeded, abor
```

Et la vérité se fait jour : le problème est dans le service DNS récursif. La majorité des requêtes DNS échouent. Quelques unes passent, ce qui explique que l'Internet ne semble pas complètement en panne à certains utilisateurs. Comme presque toutes les activités Internet dépendent du DNS, le trafic réseau chute.

Pour résumer la panne, il y avait bien une attaque DoS mais, comme au billard, l'attaque n'a pas frappé directement les serveurs DNS. L'attaque a touché un service très populaire, Baofeng, qui distribue de la vidéo et de la musique. Les attaquants frappent un serveur de jeux en ligne et l'arrêtent. Rien d'extraordinaire, ce genre de choses arrive tous les jours sur l'Internet. Sauf que l'attaque stoppe également certains des serveurs du domaine `baofeng.com`, qui partagent la même infrastructure. Et que les logiciels clients de Baofeng, devant la panne, réagissent en faisant encore plus de requêtes DNS, qui restent sans réponse. Le logiciel résolveur utilisé par tous les FAI chinois, BIND, a une limite sur le nombre de requêtes DNS récursives en attente, 1000 par défaut. En temps normal, c'est largement suffisant mais, ici elle était vite remplie par les innombrables requêtes en attente des serveurs de `baofeng.com`.

Si vous gérez un récurseur BIND, vous pouvez voir l'état des requêtes en cours avec `rndc` :

```
% rndc status
...
recursive clients: 13/1900/2000
```

Le dernier chiffre est la limite dure au nombre de requêtes en attente (il se règle avec l'option `recursive-clients` dans `named.conf`). L'avant-dernier est la limite « douce » à partir de laquelle BIND commencera à laisser tomber des requêtes, provoquant le message ci-dessus. Le premier chiffre est le nombre de requêtes actuellement en attente.

En raison du nombre de clients attendant `baofeng.com`, cette limite a été vite dépassée, supprimant toute résolution DNS, même pour les domaines n'ayant rien à voir avec Baofeng. Pour votre récurseur, faites le calcul : prenez la différence entre la limite douce et le nombre de clients en temps normal (ici, c'est 1900 - 13, mettons 1900) et divisez là par le taux de requêtes : cela vous donnera une idée du nombre de secondes que vous pourrez tenir en cas de panne. Ici, si le taux de requêtes est de 100 par seconde (ce qui est une valeur pour un petit FAI), vous avez droit à seulement dix-neuf secondes de marge en cas de panne d'un gros domaine très populaire... La plupart des récurseurs ont probablement une valeur de `recursive-clients` trop basse.

Conclusion : si quelqu'un réussit à planter tous les serveurs DNS de `google.com` ou `ebay.com`, il peut théoriquement planter tout le DNS et donc tout l'Internet.

Dans le cas chinois, tous les résolveurs étaient des BIND (comme c'est probablement le cas dans la plupart des pays). Il n'a pas été possible de tester avec d'autres résolveurs comme Unbound mais rien n'indique qu'ils auraient fait mieux. Le choix des développeurs de BIND était d'avoir un tableau de taille limitée pour les requêtes en attente. Si ce tableau était par contre dynamique, le récurseur aurait, à la place, avalé toute la mémoire du serveur.

Quelques-uns des articles les moins mal informés qui ont été publiés sur cette panne :

- « *"DNS Attack Downs Internet in Parts of China"* (http://www.pcworld.com/businesscenter/article/165319/dns_attack_downs_interne) »
- « *"A month after web chaos, Baofeng issues new media player"* (<http://www.telecomasia.net/content/month-after-web-chaos-baofeng-issues-new-media-player>) »
- « *"Internet attack 'organized' says Ministry"* (http://www.shanghaidaily.com/sp/article/2009/200905/20090521/article_401635.htm) »

Merci à Ziqian Lu pour ses explications détaillées.