SABIS : Simulation des Instabilités du protocole BGP

Houssame Yahiaoui

Laboratoire PRiSM
Université de Versailles St-Quentin de Yvelines
Houssame.Yahiaoui@prism.uvsq.fr

1. Introduction

La dernière décennie a vu l'utilisation d'un unique protocole de routage inter-domaine s'étendre au point de devenir le protocole officiel pour cette catégorie d'échanges de données. Ce protocole, nommé BGP (*Border Routing Protocol*), permet l'interconnexion de milliers de domaines de routage, appelés *Autonomous Systems* (AS). Ces domaines s'échangent leurs routes respectives par l'intermédiaire des *speakers* que possède chaque AS. Les speakers (routeurs BGP) s'envoient des annonces de routes, ainsi que des messages de perte de ces routes.
Afin de limiter la consommation des ressources sur Internet, les annonces s'effectuent uniquement lors de changements des routes connues par les AS. Malgré cette restriction, la taille des AS (parfois des centaines de routeurs), ainsi que le nombre grandissant de plages d'adresse IP à propager entre AS, font que le volume des échanges BGP reste important. Et il est d'autant plus important, que l'observation du comportement de BGP durant les dernières années a exposé un fait incontestable : le protocole n'est pas stable au niveau mondial. En examinant l'évolution du volume de messages de mise à jour au niveau du moindre AS -grâce aux *logs* du projet *BGP Instability Report* [3]-, nous constatons un même schéma : une évolution en dent de scie et, fait alarmant, une corrélation forte entre l'évolution de la masse de traffic utilisateur circulant sur le réseau et le degré de fluctuation du protocole BGP. Seul un diagnostic peut expliquer tous ces symptômes : l'instabilité de BGP. Cette instabilité est d'autant plus inquiétante, qu'elle s'accentue à la suite d'augmentations ponctuelles du volume de traffic TCP : l'illustration de ce fait réside dans l'augmentation de l'instabilité durant les périodes connues de propagation de vers (*worms*). Ainsi, l'étude des vers *RedCode I*, et *II*, ainsi que *Nimda*, a montré l'influence importante du volume du traffic utilisateur sur les échanges de routage inter-domaine. Nous avons proposé une explication possible du phénomène et modélisé cette correlation [5].

En plus du fait que ces instabilités peuvent accroitre la possibilité de perte des paquets utilisateurs (augmentant les délais de bout-en-bout), et qu'elles peuvent créer des circuits dans le graphe, il est important de noter qu'elles provoquent une sur-consommation, non négligeable, des ressources du réseau : les routeurs allouent d'avantage de ressources à traiter les échanges BGP. Cette surcharge (des routeurs et des canaux de communication) induit un stress sur l'infrastructure, et peut, à terme, provoquer l'arrêt de certains de ses composants.
Cet état de fait a été constaté depuis plusieurses années [1] sans qu'une véritable solution n'y soit apportée. A cela plusieurs raisons :

- Il est difficile de trouver les véritables sources de ces symptomes : aucune étude n'a été capable de préciser toutes les raisons derrière l'instabilité ;
- L'inertie des FAI qui estiment, pour beaucoup, que le fonctionnement du réseau est satisfaisant et qui sont rétissant à l'introduction de modifications au protocole BGP dans leurs routeurs ;
- Les preuves incomplètes du fonctionnement des méthode proposées. Elles sont validées par des simulations réelement éloignés des conditions du réseau : typiquement, la preuve est apportée qu'une certaine méthode réduit le temps de convergence et le volume de paquets échangé. Le problème réside dans notre ignorance du comportement de la même méthode dans un environnement totalement instable, et physiquement différent de celui de la simulation.

Pour toutes ces raisons, nous pensons qu'un début de solution viendrait d'un simulateur effectivement proche du comportement du protocole réel : Il est indispensable de simuler l'instabilité du protocole dans une topologie imitant le réseau inter-domaine actuel, en taille et en configuration
Une telle simulation permettera de générer le comportement de BGP, si ce n'est quantativement, par manque de données réelles et de ressources de simulation suffisantes, au moins qualitativement afin de déduire des schémas de comportement du protocole. Cette simulation permettera de tester dans des conditions réalistes, les améliorations à apporter au protocole : Pour chaque méthode nous obtenons l'impact sur l'instabilité, sur la consommation de ressources, une indication sur le taux de pénétration de la méthode sur le réseau nécessaire à l'obtention d'un réel gain, etc ...

Il existe bon nombre de simulateur, plus ou moins complexes et réalistes. Mais aucun ne répond aux exigences liés à la reproduction de l'instabilité observée, en terme de taille ou de diversité des paramètres de BGP. Notre approche est, donc, nouvelle en ce sens qu'elle ne tente non seulement d'implémenter fidèlement le protocole, mais qu'elle exige la restitution d'assez de faits pour la synthèse des instabilités de BGP.

2. SABIS : Simulateur d'Instabilités de BGP

La mise en oeuvre de SABIS a exigé la spécification de trois élements, qui, conjointement, produisent les instabilités : (1) Une architecture élaborée de *speaker BGP* ; (2) Des topologies de simulation répliquant la structure externe et interne des AS du réseau ; (3) Une implémentation fidèle des aspects temporels influanant les échanges.

2.1. Les Speakers dans SABIS

Le speaker constitue l'entité communiquante de base dans SABIS, et est très proches de la spécification de BGP-4. Il dispose de deux filtres (en entrée et en sortie) implémentant les politiques de routage de l'AS du speaker (elles reflètent les accords de transit entre AS sous forme de régles d'acceptation/modification/refus des routes importées/exportées). Le speaker choisit ses propres routes à partir des routes importées suivant un processus de décision complexe. Ce processus allié aux politiques de routage permettent un choix bien plus riche qu'un algorithme de *Plus Court Chemin*. Ces choix de routes sont, ensuite, annonés aux voisins. en utilisant des messages d'annonce (et de rejet) de route. Ce type de message, avec les messages KEEPALIVE constituent les seuls moyen de communication des speakers. Les message KEEPALIVE servent à informer les voisins du maintient des connexions avec eux. Lorsqu'un speaker ne reçoit pas de messages de ce type pendant un certain temps, il considère la connexion perdu et supprime toute route qu'il aurait apprise grâce à elle.

2.2. Topologies de Simulation

Les topologies de simulation de SABIS contiennent, contrairement à la majorité des simulateurs de BGP, des AS contenant de multiples speakers. Conserver la complexité de l'AS, est à notre sens un élement déterminant dans l'obtention d'une part des instabilités de BGP (ceci a été prouvé par des travaux d'observation du traffic BGP). Pour cause d'insuffisance d'information de l'intérieur des AS, nous avons conçu une méthode de génération de topologies d'AS [4], qui, si elle ne produit pas des topologies exactes (nous ne disposons pas de données de comparaison complètes), procure des objets aux caractéristiques désirées : des topologies aux AS multi-connectés.

2.3. Composantes temporelles dans SABIS

Nous tentons de combler un des manques les plus marquants des simulateurs de BGP par la combinaison d'une simulation exacte des timers du protocole BGP au sein de notre speaker, et d'une estimation de l'impact temporel du traffic utilisateur sur les traitements BGP. Ce dernier point a été obtenu par l'estimation pour chaque message BGP échangé d'un délai induit par la charge du speaker. En effet, les speakers étant implémentés sur des routeurs, ils doivent effectuer nombre d'autres tâches dépendant du traffic utilisateur. Ces tâches peuvent retarder le traitement BGP, et ce retard combiné au mécanisme de détection de fin de session de BGP peut provoquer des avalanches de déconnexions, qui participent pour beaucoup à l'instabilité de BGP.

L'implémentation de ces élements sur l'environnement de simulation OMNeT++ a produit SABIS. Afin de maximiser les possibilités du simulateur, nous veillons à la flexibilité et à l'extensibilité du simulateur. Divers mécanismes permettent de tester toute modification sur le protocole, et d'en simuler les besoins, l'impact et les effets indésirables face à divers scénario.
Un résultat obtenu sur SABIS, fut la reproduction d'oscillations de routes par les politiques de routage [2], sur des topologies de grande ampleur. Nous avons utilisé un algorithme de génération de politiques oscillantes sur des graphes d'AS de différentes ordres de grandeur et avons vérifié l'apparition des oscillations. Nous avons implémenté une méthode de détection et d'élimination de ces oscillations, et prouvé leur disparition.

Bibliographie

1. C. Labovitz, et al. , – Origins of Internet Routing Instability, – IEEE INFOCOM, 1999.
2. T. Griffin et G. Wilfong, – An analysis of BGP convergent properties, – ACM Sigcomm, 1999.
3. Geoff Huston, – The BGP Instability Report, – `http://bgpupdates.potaroo.net`.
4. J-M. Fourneau et H. Yahiaoui, – Internet Topology Generation for Large Scale BGP Simulation,– IPS-MoMe, 2005.
5. J-M. Fourneau et H. Yahiaoui, – Modelling The effects of a Worm propagation on a BGP Router, – ASMTA, 2007.