Bulletin of Networking, Computing, Systems, and Software – www.bncss.org, ISSN 2186–5140
Volume 4, Number 1, pages 18–20, January 2015

Quantifying the Propagation Behavior of BGP Routing Update

Taihong Wang*‡, Xingang Shi†‡, Xia Yin*‡, Zhiliang Wang †‡
*Department of Computer Science and Technology, Tsinghua University
†Institute for Network Sciences and Cyberspace, Tsinghua University
‡Tsinghua National Laboratory for Information Science and Technology (TNLIST)
Beijing, P.R. China
Email: thwang@csnet1.cs.tsinghua.edu.cn, shixg@cernet.edu.cn, yxia@tsinghua.edu.cn, wzl@cernet.edu.cn

***Abstract*—Inter-domain routing convergence is the state of a set of routers that have the same topological information about the internetwork in which they operate. Previous measurement studies about inter-domain routing convergence are carried out through the analysis of BGP update traffic. However, this kind of studies only consider the time that inter-domain routers take to reach a consistent view of the network topology after internet path failure, failover and repair. There has been no study to quantify the time that inter-domain routers take to distribute a routing update across the network.**

In this paper we develop a system to measure the propagation behavior of BGP route updates, including both passive data collection and active data probe in the real world. This system has been continuously monitoring the Internet for more than three years. During this period, 143K route updates were detected, and 5172k detection were triggered. Our data shows that route oscillations that have only been theoretically studied before do exist in the Internet. Inter-domain routers take about 30 seconds to distribute a routing update across the network.

***Index Terms*—BGP; routing convergence; route oscillation; propagation time**

I. Introduction

The Internet is composed of tens of thousands of Autonomous Systems (ASes), defined as a collection of IP networks and routers under the control of one entity. BGP is the de facto inter-AS routing protocol used in the global Internet. It is designed to quickly adapt to network connectivity changes and converge on a new set of stable routes. However, a number of previous analytical and measurement studies ([1], [2], [3]) have shown that the existence of slow convergence and route oscillation in the inter-domain routing system. Labovitz et al. [2] found that the inter-domain routing convergence delay in multihomed fail-over now averages 3 minutes, and may trigger oscillations lasting as long as 15 minutes. Delayed inter-domain routing convergence is an important problem for the Internet today as the Internet contiues to grow in size.

A number of researchers focus on the problem of inter-domain routing convergence. Varadhan et al. [1] prove that there exist domain policies that cause BGP/IDRP to exhibit persistent oscillations through theoretical analysis. Labovitz et al. [2] investigate and quantitatively measure the time of inter-domain routing convergence through data collected from the RouteViews and fault injections into ISP backbone. Oliveira et al. [3] quantify the convergence times of different route change events through the analysis of BGP update traffic. Most of related works focus on the convergence delay caused by BGP path selection process. However, inter-domain routing convergence delays stem from not only the operation of BGP path selection process, but also the propagation time of BGP paths. This is a gap between the needs and reality of today's data networks.

In this paper we develop a system to investigate the time it takes to distribute a routing update across the network and some properties when inter-domain routers distribute routing updates. At a first step, we collect live BGP UPDATEs from the Internet. Then we select a AS-path segment from routing updates according to our filtering mechanism. Finally we randomly choose some monitors to detect whether they have received the AS-path. In this way, we can know the time it takes to distribute a routing update across the network. This system has been continuously monitoring the Internet for more than three years.

II. Methodology

Our system is composed of two main modules: the BGP Monitoring Module (BMM), the AS-path Detecting Module (ADM).

A. BGP Monitoring Module

If we measure the time that it takes to distribute every routing update, our system will be heavily loaded and not practical. So we build a filtering mechanism to decrease the number of BGP UPDATE to measure. We only measure two types of route updates: origin anomalies (OA), and AS-path segment anomalies (SA).

An origin anomaly (OA) happens when the origin AS in the AS-path for a prefix changes to a different AS. Our system maintains a database which keeps track of already announced origin ASes of each prefix to help detect origin anomalies.

Due to the enormous number of all the AS-paths, it is impractical to detect all the AS-path, so we only focus on the neighboring ASes pairs and triples, which we both call AS-path segment anomalies (SA).

By considering the above two kinds of anomalies, our system can detect different kinds of BGP UPDATEs.

The BGP Monitoring Module receives live BGP UPDATEs from BGPmon [4], a real-time BGP feed in the Route Views project [5] that collects UPDATEs from routers all over the world. When the AMM receives an UPDATE, it will check whether the embedded AS-path has an anomalous origin AS or an anomalous AS-path segment according to its local routing information database.

Once the BMM detects an anomaly, it will notify the ADM to detect whether this BGP update reach to other routers.

B. AS-path Detecting Module

The AS-path Detecting Module (ADM) employs a number of public route-servers and looking-glasses, called as the *eyes* of our system, to detect where the BGP update reach to this eye. When anomaly is detected for a prefix f, each eye gathers the control-plane route status of the prefix f.

Specifically, our system uses *show ip bgp* to extract the best BGP route $p_{t,j}(f)$ that its j-th *eye* chooses for the prefix f, in the t-th second. Then it checks whether $p_{t,j}(f)$ is affected by the anomalous origin AS (or AS pair/triple) reported by the *BMM*.

After an anomaly is reported, we continuously do that for W seconds on N *eyes*. At the t-th second ($1 \leq t \leq W$), we compose the control-plane results of all *eyes* into a binary vector $C_t = \{C_{t,j} | 1 \leq j \leq N\}$, where

$$C_{t,j} = \begin{cases} 0, & \text{if } p_{t,j}(f) \text{ is affected by the anomaly} \\ 1, & \text{if } p_{t,j}(f) \text{ is not affected by the anomaly} \end{cases}$$

the AS-path Detecting Module (ADM) has a number of 376 *eyes*, which distribute in 38 ASes. Through these eyes, our system can make a close and pervasive monitoring over the Internet.

III. Analysis

Our system has been continuously monitoring the Internet for three years, starting from May 2, 2011. During this period, 143K anomalous route events were reported, and 5172k detection were triggered. Some detections show that route updates don't reach the eye. These detections don't contribute to the study of routing convergence, so we filter them. After filtering, there are 1375k useful detections.

A. Propagation Time

Since the Internet can often converge in less than two minutes except for some unusual cases [3], We set W = 120 seconds. Due to the delay of the network data plane, our system may receive the response of eyes at different time. To remove this mistake made by the delay of response, we remove detections that firstly receive the route update later than the first. After this process, there are 217k useful detections.

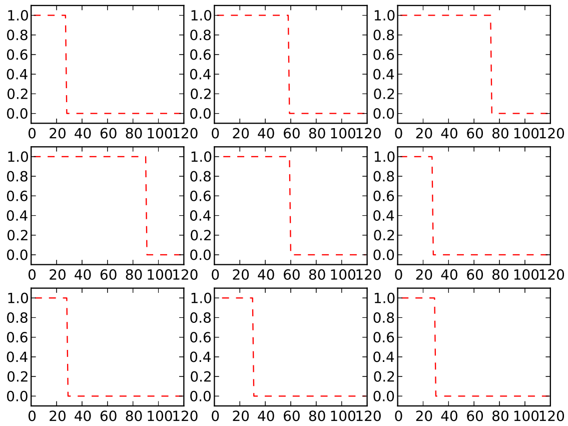


Fig. 1. Routing Propagation

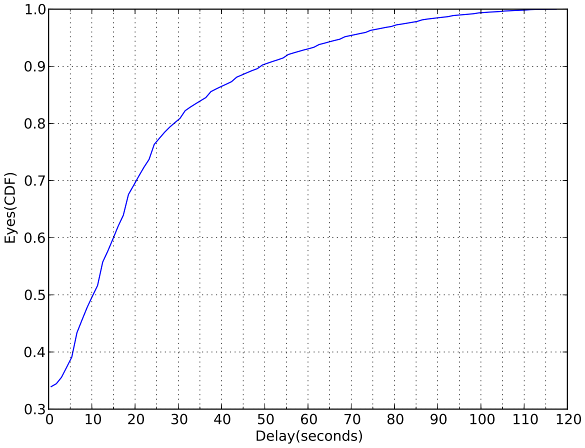


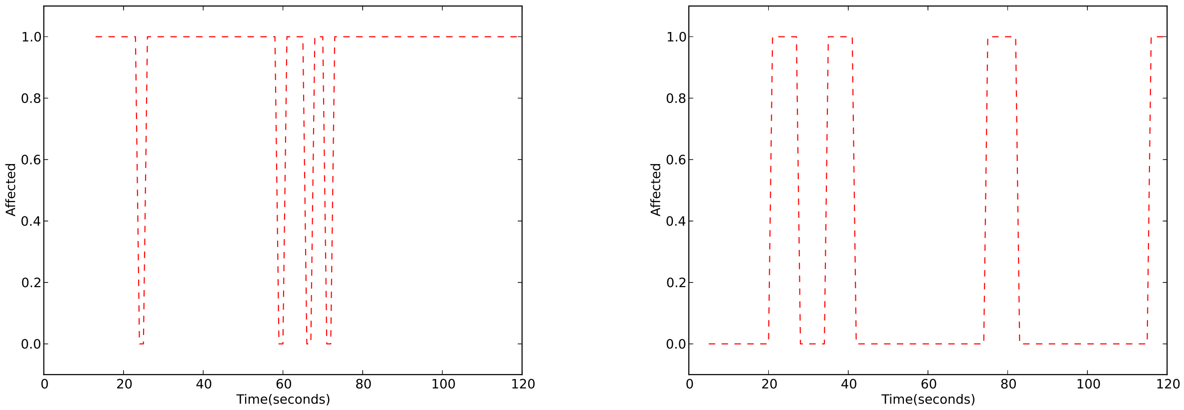
Fig. 2. Propagation Time(CDF)

Fig. 1 plots a example of the propagation behavior of a route update. X-axis indicates the time, the unit of which is *seconds*. Y-axis indicates whether the eye is affected by the anomalous origin AS (or AS pair/triple) for the prefix f, *zero* means affected, *one* means unaffected. So does Fig. 3. During the detection, about N = 40 eyes are selected to detect, but we only draw nine eyes' behavior due to the limit of capacity. In the figure, a BGP update announced that the origin AS for prefix 74.84.68.0/24 changed from AS6478 to AS30036. We can know from the figure that different eyes receive the BGP update at different time, some eyes received the BGP update around 24th seconds, other eyes received the BGP update after a while.

Fig. 2 shows the cumulative distribution function (CDF) of the propagation time of all route updates. About 30 percent of the eyes received the route update immediately, and about 80 percent of the eyes received the route update after 30 seconds. We can draw a conclusion that inter-domain routers take about 30 seconds to distribute a routing update across the network.

B. Route Oscillation

Among all the detections, there are two kinds of route behavior. One kind of route behavior is that the AS-path for a prefix change and recover within two seconds, called as a



(a) spike

(b) instability

Fig. 3. route oscillation

spike. Another kind of behavior is that the AS-path for a prefix change and recover several times within a short time, called as a *instability*. We consider both of them as route oscillation.

Fig. 3(a) shows that the AS-path for a prefix change and recover within two seconds. Fig. 3(b) shows that the AS-path for a prefix change and recover several times within two minutes.

Among all the detections, there are 211 detections that have more than one *s*pikes, and 36K detections (2.6%) that the AS-path for a prefix changed more than two times.

IV. Conclusion

In this paper, we develop a system to measure the propagation behavior of BGP route updates. We show that route oscillations that have only been theoretically studied before do exist in the Internet, and inter-domain routers take about 30 seconds to distribute a routing update across the network.

References

[1] K. Varadhan, R. Govindan, and D. Estrin, "Persistent route oscillations in inter-domain routing," Computer Networks, Tech. Rep., 1996.

[2] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian, "Delayed internet routing convergence," *SIGCOMM Comput. Commun. Rev.*, vol. 30, no. 4, pp. 175–187, Aug. 2000.

[3] R. Oliveira, B. Zhang, D. Pei, R. Izhak-Ratzin, and L. Zhang, "Quantifying path exploration in the internet," in *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*, ser. IMC '06. ACM, 2006, pp. 269–282.

[4] "The BGPmon project," http://bgpmon.netsec.colostate.edu.

[5] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the internet as-level topology," *SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 1, Jan. 2005.