

Impact of Routing Lags on Internet Routing Failures ^{*}

Feng Wang and Lixin Gao

Department of Electrical and Computer Engineering
University of Massachusetts, Amherst
MA 01002, USA
{fewang, lgao}@ecs.umass.edu

Abstract. Studies have shown that the Internet has experienced the widespread failures such as router crash, fiber cut, or scheduled maintenance. Ideally, routing protocols should be able to quickly find alternate paths to reroute around failures. In this paper, we consider one important factor that prevents routing protocols from achieving this goal: the delay of finding and obtaining alternate paths, defined as *routing lag*. We show that a significant number of routing failures in the Internet are caused by routing lags, and they can last for a significant period of time.

1 Introduction

One of the goals of routing protocols is to be able to reroute around failures such as router crash, fiber cut, or scheduled maintenance. When a failure occurs, routing protocols should be able to quickly find alternate paths to reroute around failures. Routing protocols failing to achieve this goal is called as *routing failure*. Widespread routing failures in the Internet have been observed in experimental studies [1, 2]. Furthermore, Border Gateway Protocol (BGP), a path vector protocol for interdomain routing in the Internet, experiences considerable delay in reaching convergence [3–5]. BGP can experience transient loss of reachability during route convergence. For example, BGP applies “poison reverse” to avoid routing loops. If router *A* uses router *B* to reach a destination, router *A* does not announce its best route to router *B*. This can limit the route visibility of router *B* so that there is latency of obtaining an alternate route from *A* during the path exploration. We define the delay of finding and obtaining alternate routes as *routing lag*.

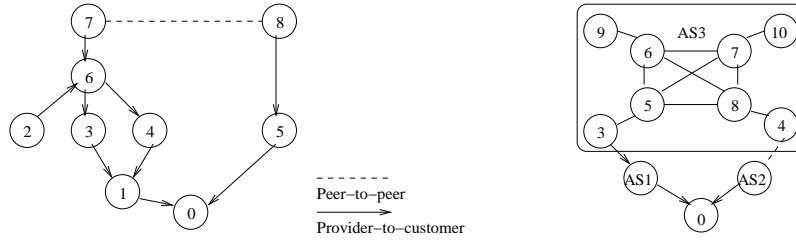
In this paper, we study routing failures due to routing lags in the Internet. First, we show that routing lags are prevalent in the sense that any router can experience routing lags. Second, we derive upper bound for routing failure durations due to routing lags, and consider the impact of rate limiting timers on the duration. We find that if the rate limiting timers are applied on the granularity of a peer (current implementation) instead of a prefix, routing lags can last for a significant period of time. Our results imply that there is a clear need to reconsider the practice of rate limiting timers.

^{*} This work was supported by the NSF under NSF CNS-0208116 and NSF CNS 0325868. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

2 Routing Lags in the Internet

We study routing lags in a typical BGP system, which means that every router in the system applies *common routing policies*. Routing policies are guided by commercial relationships between ASes: *customer-to-provider* and *peer-to-peer*. In a typical BGP system, the import routing policies are guided by the *prefer-customer* routing policies, while export routing policies by the *no-valley* routing policies.

Routing lags can be prevalent in a typical BGP system. Figure 1(a) shows a typical BGP system, in which we assume that each AS contains only one router and the destination is situated in AS 0. Suppose that the link between nodes 1 and 0 breaks, node 1, 3, 4 and 6 will definitely, and node 2 will possibly experience routing lags. On the other hand, large ASes typically have a hierarchical iBGP structure. For example, in Figure 1(b), AS 3 has the customer AS 1 and the peer AS 2 so that all routers inside AS 3 will use the path via router 3 to reach the destination in AS 0. Once the link between router 3 and AS 1 fails, all routers inside the AS except router 4 may experience routing lag.



(a) Routing lags involving multiple ASes

(b) Routing lags inside an AS

Fig. 1. Routing lags take place in a typical BGP system

3 Routing Failure Duration Induced by Routing Lags

Routing failure duration induced by a routing lag is the time interval between the time when there is no route to a destination and the time when the first route appears after the failure. The major factor determining the failure duration is the rate limiting timer, *MinRouteAdver (MRAI)*, which is used to determine the minimum amount of time that must elapse between routing updates. There are two ways to apply rate limiting timers: on the granularity of a prefix and on a peer. For each way, rate limiting timers can be applied to only announcements or both announcements and withdrawals. When the timer is applied to withdrawals (announcements), it is started again after finishing sending the withdrawal (announcement) messages. Next, we will derive the upper bounds of routing failure duration, which is caused by routing lags under different MRAI implementations.

We use an AS graph $G = (V, E, d)$ to represent a network connectivity to a destination d . The node set V consists of routers, and the edge set E represents the connectivity

to the destination d . Router u has a set of path $P_{(u,l)} = \{P_1, P_2, \dots, P_r\}$ to the destination. Suppose a link fails, it will lead router u to reroute the failure. It will request other routers provide alternate paths if they have. We denote these nodes where node u can obtain a path with a set $B = \{\beta_1, \beta_2, \dots, \beta_m\}$. So failure duration due to routing lags consists of (1) the delay of propagating the request (withdrawal message) from u to any router in B , denoted as $d_{(u,\beta)}$, and (2) the delay of obtaining an alternate path (announcement message) from the router, denoted as $d_{(\beta,u)}$.

Given two neighboring nodes u and v , the MRAI timer of u configured for v is denoted by $M_{(u,v)}$. Obviously, the delay for a routing update from u to v is $d_{(u,v)} \leq M_{(u,v)}$. Note that the value of M can be either for an iBGP or for an eBGP session.

With MRAI timer based on per peer implementation, the upper bound of failure duration is shown as follows:

- If MRAI timer is applied to both announcement and withdrawal, the failure duration due to routing lag is bounded by

$$\min_{\beta \in B} (d_{(u,\beta)} + d_{(\beta,u)}) \leq \min_{\beta \in B} \left(\sum_{(i,j) \in P_{(u,\beta)}} M_{(i,j)} + \sum_{(i,j) \in P_{(\beta,u)}} M_{(i,j)} \right) \quad (1)$$

- If MRAI timer is applied to announcement only, the failure duration due to routing lag is bounded by

$$\min_{\beta \in B} (d_{(u,\beta)} + d_{(\beta,u)}) \leq \min_{\beta \in B} \left(\sum_{(i,j) \in P_{(\beta,u)}} M_{(i,j)} \right) \quad (2)$$

With MRAI timer based on per prefix implementation, the upper bound of failure duration is shown as follows:

- If MRAI timer is applied to both announcement and withdrawal, the failure duration due to routing lag is bounded by

$$\min_{\beta \in B} (d_{(u,\beta)} + d_{(\beta,u)}) \leq M \quad (3)$$

In this case, we assume that the first update can be propagated *without delay* imposed by those timers.

- If MRAI timer is applied only to announcement, the failure duration due to routing lag is bounded by

$$\min_{\beta \in B} (d_{(u,\beta)} + d_{(\beta,u)}) \leq \delta, \quad (4)$$

where δ represents the transmission delay of routing updates at each link and process delay at each router, without any delay imposed by MRAI timer.

We find that applying rate limiting timer on the granularity of a peer can indeed prolong the failure duration compared with on the granularity of a prefix. We also find that when MRAI timer is applied to just announcement but based on per prefix, the duration is the smallest. Further details on the proof of failure duration are presented in [6].

4 Measurement Results

We first measure the prevalence of routing failures due to routing lags at 5 tier-1 ASes by using BGP updates in August 2003 from Oregon RouteView server. The percentage of routing failures at those ASes is shown as following: AS1239 (4.4%), AS2914 (18%), AS3356 (3%), 3561 (2.8%), and AS7018 (8.7%). Figure 2 shows the cumulative distribution of failure duration due to routing lags at 5 tier-1 ASE, which obtain alternate paths from neighboring ASes, in August, 2003. We find that more than 95% of those failures have less than 90 seconds of failure duration, and more than 90% of those failures last more than one MRAI time (30 seconds). Details on the method of identifying routing lags are presented in [6].

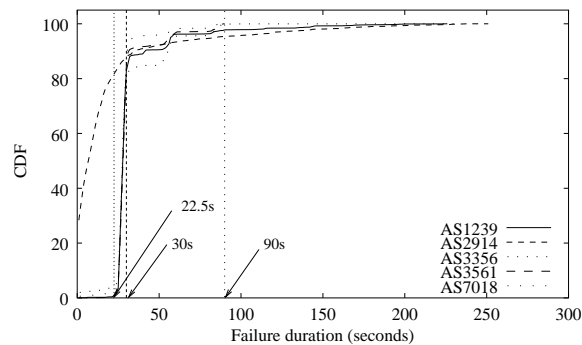


Fig. 2. Distribution of the duration of routing failures due to routing lags.

5 Conclusions

In this paper, we analyze the impact of routing lags on routing failures. Our results show that routing lags are prevalent in the Internet, and can indeed lead to longer failures. We also consider the impact of the implementation of rate limiting timer on routing lags. We find that if rate limiting timers are applied on the granularity of a peer instead of a prefix, the failure duration is the longest. Our results imply that there is a clear need to reconsider the practice of rate limiting timers.

References

1. C. Labovitz, G. Malan, F.Jahanian. Internet Routing Instability, In *Proc. of SIGCOMM*, September 1997.
2. C. Labovitz, A. Ahuja, F. Jahanian, Experimental Study of Internet Stability and Backbone Failures. *FTCS 1999*: 278-285.
3. C. Labovitz, A. Ahuja, and et al, Delayed Internet Routing Convergence. *IEEE/ACM trans. On Networking*, Vol.9, No.3, June 2001.
4. C. Labovitz, A. Ahuja, The Impact of Internet Policy and Topology on Delayed Routing Convergence. In *Proc. of INFOCOM 2001*, Anchorage, Alaska, April 2001.
5. T. Griffin and B. J. Premore, An experimental analysis of BGP convergence time. In *Proc. ICNP 2001*, Nov 2001.
6. Feng Wang, Lixin Gao, Impact of Routing Lags on Internet Routing Failures, Technical report, <http://rio.ecs.umass.edu/~wangwind/lag.pdf>.