Filter-Based RFD: Can We Stabilize Network Without Sacrificing Reachability Too Much?

Ke Zhang and S. Felix Wu

Computer Science Department, The University of California, Davis
{kezhang, sfwu}@ucdavis.edu

1 Introduction

Internet instability, also referred to as route flaps, can propagate to the whole Internet and consume remarkable computational resource of the routers. Route Flap Damping (RFD) [1] is designed to stabilize the Internet by suppressing persistent route flaps. RFD is a penalty based mechanism.The magnitude of the penalty value indicates the degree of instability of an inter-domain route. Once the penalty reaches a certain threshold, the route will be suppressed. This simple mechanism does not work perfectly. First, the way it identifies route flaps or accumulates penalty is too aggressive and may suppress a fairly stable route with a few occasional flaps. Second, a route may be suppressed even after it has converged.

Extensive researches on the side-effects of Route Flap Damping have been done recently [2–4]. Mao *et al.* [2] proposed an intriguing approach, selective route flap damping. Their approach tries to solve the first problem by distinguishing persistent route flap from occasional route flap. However, the ISP industry seems to lack interest in the new RFD implementations [5]. A major reason is that as a penalty system, RFD cannot stabilize the network without sacrificing reachability. Although stabilization and high reachability are highly desired, RFD can not optimize the two aspects at the same time. In order to persuade ISP industry to adopt the new RFD ideas, we need to demonstrate that the new mechanisms can really achieve optimal trade-off between the two aspects of RFD.

This paper tries to fill this gap. We propose an empirical RFD tune-up to improve route flap damping based on two heuristics. First, an occasional flap can trigger excessive route updates due to path exploration. These updates are usually observed as a burst of updates. If we accumulate RFD penalty only based on sampled updates, the occasional flap will not be punished and we can still capture the long term persistent route flaps. Second, by examining Internet BGP updates, people observe that when failure is recovered, most of the time the route converges to the previous primary AS path [6]. The primary path can be viewed as a signal indicating that the route is converged into a stable state. Thus, the suppressed route can be reused when the primary path appears, which will significantly reduce the suppression time.

We conduct extensive experiments to evaluate our optimized RFD and SRFD mechanisms and try to answer the question: can we stabilize the network without sacrificing reachability too much? We examined the following factors in the experiments: 3 typical BGP update burst sequences, 2 typical MRAI timer, single homing or mutli-homing network. We evaluate these RFD mechanisms from the following perspectives:

- How many BGP update messages triggered by route flaps are suppressed?
- How much does RFD impact routing convergence?
- How much is network stability improved?
- How well the network reachability is maintained?

The results show that both SRFD and FRFD can significantly reduce the side-effects of RFD and stabilize the network as expected.

2 Filter-based RFD

Burst of BGP updates caused by path exploration usually lasts for a shorter time period compared to persistent route flaps. Current RFD is oversensitive to a short time burst of BGP updates. To solve this problem, we design a dynamic slide-window based filter. The principle of the design is simple and lightweight, because RFD is performed for every route received from a single peer. Sophisticated algorithm requires large memory and intensive CPU computing, which cannot be a scalable solution.

We apply a window-based filter to sample incoming updates. Within this time window, only the first update is penalized among all the incoming updates. If an incoming update is beyond the window, the window correspondingly slides and the window size is dynamically adjusted. Initially, the size is set to a predefined maximum value. If there is an incoming update beyond the window, we reduce the window size to half each time until the size is reduced to the minimum size. When the route is stable for a long period of time or the penalty drops below the reuse threshold, the sampling window is set to maximum again. The maximum size should be defined carefully to cover the path exploration period, which depends on the topology and peers. The minimum size should be at least equal to Minimum Route Advertisement Interval(MRAI). In the experiments, we choose $16 * MRAI$ as the maximum size.

The heuristic of this design is based on two observations. First, a burst of BGP updates caused by a single route flap should only be penalized once. Second, the purpose of RFD is to prevent persistent route oscillations caused by link/router failure or mis-configuration. Although the sampling may miss some route flaps, it is capable of detecting and quickly penalizing long-term route flaps through the decreased sampling window.

Current RFD suppresses the route till the penalty drops the below reuse threshold. Even when the route is not flapping, it still cannot be selected as the best route. This can significantly delay the fail-over efforts. In the early work, we observed that many prefixes have a primary AS path. The primary AS path is the route that has been used for most of the time. When route flaps happen, the primary route is very likely to be selected as the best path again after convergence. Thus, the primary path could indicate that route convergence is completed and a stable route has been selected. Based on this heuristic, we propose **early reuse** – reduce the penalty value to one half whenever the primary path is received.

3 Evaluation and Comparison

We performed a set of experiments on SSFNet. Our goal is to reveal that given the same network topology and route flap events, how three RFD mechanisms impact network convergence, reachability and stability.

We use a two dimensional $10X10$ grid topology to simulate a flat transit network. Each node represents a single EBGP router. Route selection is solely based on the length of AS path. Node 0 announces a prefix p and generates different update sequence for p. To simulate multi-homing environment, we attach another node (node 101) to announce the same prefix p.

We simulate three types of BGP update sequence as input to the network.

1. *Route Flapping (RF)*: A sequence of route UP and DOWN.
2. *Route Oscillation (RO)*: A sequence of oscillating routes. It simulates the persistent route attribute changes.
3. *Slow Convergence (SC)*: A sequence of path exploration followed by a route withdrawal and an announcement. This represents BGP path exploration process corresponding to a failure and fail-over event.

We apply four metrics to measure the behavior of RFD from different perspectives.

1. total number of BGP update messages.
2. delayed convergence time. It is defined as the interval between the time when node 0 re-advertises the initial updates and the time when the network stops generating updates.
3. total number of nodes that lose routes to the prefix advertised by node 0.
4. α-instability. α is the time limit for a router to use a nexthop for switching. Any fast change of nexthop in the forwarding table (in our experiment, nexthop change is equal to FIB interface change) will be counted as an unstable change of the forwarding plane if the time of using this nexthop is less than α (seconds). Thus, α-unstable nexthop is defined as a nexthop which is installed in FIB shorter than α. α-instability is defined as the total time that α-unstable nexthop is used for forwarding. It is a score to measure the forwarding instability of the whole network.

We present the experiment results and compare three RFD mechanisms. These results are based on experiments where MRAI is set to 5 seconds. We also analyze results based on experiments where MRAI is set to 30 seconds. There is no significant difference when different MRAI timer is used.

Our findings include the following:

- All three RFD mechanisms reduce the number of updates as expected. For the first event, persistent route flaps, RFD reduce the number of updates by 87% in single-homing network and 90% in multi-homing network. SRFD cuts off 70% and 83% of the updates. FRFD suppresses 65% and 80% of the updates. However, in the second and third events, the reduction is not as drastic. Especially in the multi-homing network, the RFD mechanism even increase the BGP updates!!

- Although regular RFD performs better in terms of BGP updates deduction, it indeed sacrifices convergence time and route availability. Convergence is delayed for approximately 4000-7000 seconds. SRFD improves the convergence by avoiding the reuse-triggered suppression. Compared to regular RFD, FRFD reduces the delayed convergence by half.
- RFD hurts reachability. Due to the reuse-triggered suppression in RFD, some nodes lose route for more than 4000 seconds. In the event of persistent route flaps, although the reuse-triggered suppression is unavoidable, SRFD achieves a shorter suppression. FRFD keeps the route reachable on approximately half of all the nodes. In the second event, path oscillation, FRFD also keeps half of the nodes reachable due to the early-reuse. In the third event, slow convergence, both SRFD and FRFD do not suppress the route and the number of updates is not reduced. On the contrary, although RFD reduces 19% of the updates, reachability has been cut off for 2000 - 4000 seconds.
- In terms of forwarding instability, RFD achieves the most stable forwarding in all three events. FRFD is more preferred than SRFD in the first two events and has the same score as SRFD in the third event. In addition, if we only consider nexthop changes within 4 seconds to be unstable, the stability achieved by FRFD is about same as RFD.

4 Conclusion

In this paper, we proposed a Fliter-based RFD, applying two simple heuristics to current RFD. We performed extensive experiments to evaluated the Filter-based RFD. Our experiments measure the number of BGP updates suppressed, network convergence, network reachability and stability. We took into account of various factors that may influence the performance of RFD, including MRAI timer, different BGP events, single-homing, and multi-homing. We demonstrated that, by optimizing current RFD, we can stabilize the network without sacrificing reachability too much.

References

1. R. Chandra C. Villamizar and R. Govindan. BGP Route Damping. RFC 2439, May 1998.
2. Z. Mao, R. Govindan, G. Varghese, and R. Katz. Route Flap Damping Exacerbates Internet Routing Convergence. August 2002.
3. Zhenhai Duan, Jaideep Chandrashekar, Jeffrey Krasky, Kuai Xu, and Zhi-Li Zhang. Damping bgp route flaps. In *23rd IEEE International Performance Computing and Communications Conference*, 2004.
4. Beichuan Zhang, Dan Pei, Daniel Massey, and Lixia Zhang. Timer Interaction in Route Flap Damping. In *The 25th International Conference on Distributed Computing Systems (ICDCS)*, 2005.
5. P. Smith and C. Panigl. RIPE routing-gw recommendations on route-flap damping. Technical Report 378, RIPE, May 2006.
6. Olaf Maennel and Anja Feldmann. Realistic BGP Traffic for Test Labs. In *Proceedings of the ACM SIGCOMM '02*, August 2002.