RIPE Routing Working Group Recommendations on Route-flap Damping

Philip Smith
Christian Panigl

Document ID: ripe-378
Date: 11 May 2006
Obsoletes: ripe-229, ripe-210, ripe-178

Abstract

This document discusses Route-flap Damping and recommends acceptable practices for ISPs who are considering deploying Route-flap Damping.

1.0 Introduction

Route-flap Damping (RFD) [1] is a mechanism for BGP speaking routers intended to improve the overall stability of the Internet routing table and reduce the load on the CPUs of the core routers. Unfortunately, due to the dynamics of the protocol, common simple configurations can do more harm than good, see [3,4].

1.1 Background

In the early 1990s the accelerating growth in the number of prefixes being announced to the Internet (often due to inadequate prefix-aggregation), the denser meshing through multiple inter-provider paths, and increased instabilities started to cause significant impact on the performance and efficiency of the Internet backbone routers. Every time a routing prefix became unreachable because of a single line-flap, the withdrawal was advertised to the whole core Internet and handled by every single router that carried the full Internet routing table.

It was soon realized that the increasing routing churn created significant processing load on routing engines, sometimes sufficiently high load to cause router crashes.

To overcome this situation RFD was developed in 1993 and has since been integrated into most router BGP software implementations. RFD is described in detail in RFC 2439. RFD is now used in many service provider networks in the Internet.

1.2 Coordination of flap damping parameters

When RFD was first implemented in commercial routers, vendor implementations had different default values and different characteristics. As inconsistency would result in different rates of flap damping, and therefor introduce inconsistent path selection and thus behavior that was very hard to diagnose, the ISP community introduced a consistent set of recommendations for flap damping parameters, so that ISPs deploying RFD would treat flapping prefixes in the same way.

This call for consistency resulted in the RIPE Routing Working Group producing first ripe-178, then ripe-210, and finally the ripe-229 documents [2], following consensus of the Routing Working Group. The parameters documented in ripe-229 were considered, at time of publication in 2001, the best current practice.

2.0 Current Status of Route-flap Damping

Research in the years following the introduction of RFD into BGP implementations, and the publication of the RIPE Routing Working Group recommendations, has demonstrated that there are real and signficant problems with RFD as deployed on the Internet today.

2.1 Impeded Convergence

Perhaps the best known work highlighting major problems with RFD is that by Zhuoquing Mao and colleagues, presented at Sigcomm in 2002. Following presentations by Randy Bush and colleagues explain the research work more accessibly.

The major issue is that if one path is withdrawn, all BGP speakers will use best path selection to pick the next best path, and advertise this best path to all their neighbours. These neighbours will see a change in path; a change in path is a change in attribute, so the prefix as seen on a neighbouring router will attract a flap penalty - even though that path is perfectly valid and there has been no disappearance of the prefix from the routing table [5].

And this path "hunting" goes on throughout the Internet - a simple prefix withdrawal can result in the appearance of a major flap event a few AS hops away in the Internet, with the result that vendor default and even the RIPE-229 recommended flap damping parameters will mark the prefix to be suppressed. While the operator can see this is an error, the routers are simply reacting to the circumstances presented to them.

2.2 Updates transiting the network

Problems are not just caused by path "hunting". Each implementation of BGP either has differing values of the Minimum Route Advertisement Interval (MRAI) Timer (the amount of time a router waits before passing on a route update) or does not implement MRAI at all in favour of the vendor's own throttling algorithm.

Some implementations pass on the update without waiting at all, others wait for 30 seconds, etc. These differences mean that update messages transiting different ASNs using different vendor equipment will arrive at the target router at different times. This router will see these different messages, and will consider each one for best path options. This will more than likely result in a different best path offered to its neighbours for each message update arriving.

The result of this is that a simple update message from one ASN would be seen as a multiple route flap event a few ASN hops away - when in fact there was no instability whatsoever. There have been actual measurements where this resulted in a single prefix withdrawal producing 41 BGP events a few hops away!

Not only is the MRAI timer a potential source of problems, but also differences in CPU loadings and CPU speed will result in different update times for prefixes announcements passing from router to router. These differences will also contribute to the effects described above.

3.0 Solutions

Possible solutions to the problems summarised above have been proposed and analysed in the work by Zhouqing Mao and colleagues.

However, despite publication in 2002, there has since then been no desire expressed from the ISP industry for these modifications to be made to the BGP implementations. Nor has there been any activity by the BGP implementors to enhance their flap damping implementations to follow those recommendations.

As the power of routers has increased, the original needs for BGP Flap Damping is no longer a major concern for operators or router equipment vendors as it was in the mid-1990s when route flapping consumed a signficant percentage of the CPU of early routers. In fact, the negative effects of RFD, as described above, have become the major concern, the cure has become worse than the disease!

4.0 Recommendation

This Routing Working Group document proposes that with the current implementations of BGP flap damping, the application of flap damping in ISP networks is NOT recommended. The recommendations given in ripe-229 and previous documents [2] are considered obsolete henceforth.

If flap damping is implemented, the ISP operating that network will cause side-effects to their customers and the Internet users of their customers' content and services as described in the previous sections. These side-effects would quite likely be worse than the impact caused by simply not running flap damping at all.

5.0 Conclusion

With current vendor implementations, BGP flap damping is harmful to the reachability of prefixes across the Internet. We would like to encourage more work to correct some of the issues highlighted by the work of Mao et al [3], to allow the viewing of prefix flap statistics without applying flap damping, and permit more flexible per eBGP neighbour damping configuration features for network operators.

6.0 Acknowledgements

We would like to acknowledge valuable contributions and feedback from Randy Bush.

7.0 References

[1] Curtis Villamizar, Ravi Chandra, Ramesh Govindan
RFC2439: BGP Route-flap Damping (Proposed Standard)
ftp://ftp.ietf.org/rfc/rfc2439.txt

[2] RIPE Documents

ftp://ftp.ripe.net/ripe/docs/ripe-178.txt
ftp://ftp.ripe.net/ripe/docs/ripe-210.txt
ftp://ftp.ripe.net/ripe/docs/ripe-229.txt

[3] Zhouqing Mao, Ramesh Govindan, George Varghese, Randy Katz
Route-flap Damping Exacerbates Internet Routing Congerence Sigcomm 2002
http://www.eecs.umich.edu/~zmao/Papers/sig02.pdf

[4] Randy Bush, Tim Griffin, Zhouqing Mao
Route-flap Damping: Harmful?
NANOG 26
http://www.nanog.org/mtg-0210/ppt/flap.pdf

[5] Craig Labovitz, Abha Ahuja, Abhijit Bose, Farnam Jihanian
Delayed Internet Routing Convergence
Sigcomm 2000
http://www.acm.org/sigs/sigcomm/sigcomm2000/conf/paper/sigcomm2000-5-2.pdf