Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 24, 2012

C. Pelsser
R. Bush
Internet Initiative Japan
K. Patel
P. Mohapatra
Cisco Systems
O. Maenel
Loughborough University

December 22, 2011

# Making Route Flap Damping Usable
draft-ymbk-rfd-usable-02

## Abstract

Route Flap Damping (RFD) was first proposed to reduce BGP churn in routers. Unfortunately, RFD was found to severely penalize sites for being well-connected because topological richness amplifies the number of update messages exchanged. Many operators have turned RFD off. Based on experimental measurement, this document recommends adjusting a few RFD algorithmic constants and limits, to reduce the high risks with RFD, with the result being damping a non-trivial amount of long term churn without penalizing well-behaved prefixes' normal convergence process.

## Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

## Status of this Memo

This Internet-Draft will expire on June 24, 2012.

Copyright Notice

Table of Contents

## 1. Suggested Reading

It is assumed that the reader understands BGP, [RFC4271] and Route Flap Damping, [RFC2439]. This work is based on the measurements in the paper [pelsser2011]. A survey of Japanese operators' use of RFD and their desires is reported in [I-D.shishio-grow-isp-rfd-implement-survey].

## 2. Introduction

Route Flap Damping (RFD) was first proposed (see [ripe178] and [RFC2439]) and subsequently implemented to reduce BGP churn in routers. Unfortunately, RFD was found to severely penalize sites for being well-connected because topological richness amplifies the number of update messages exchanged, see [mao2002]. Subsequently, many operators turned RFD off, see [ripe378]. Based on experimental measurements, this document recommends adjusting a few RFD algorithmic constants and limits, with the result being damping of a non-trivial amount of long term churn without penalizing well-behaved prefixes' normal convergence process.

Very few prefixes are responsible for a large amount of the BGP messages received by a router, see [huston2006] and [pelsser2011]. For example, the measurements in [pelsser2011] showed that only 3% of the prefixes were responsible for 36% percent of the BGP messages at a router with real feeds from a Tier-1 and an Internet Exchange Point during a one week experiment. Only these very frequently flapping prefixes should be damped. The values recommended in Section 6

achieve this. Thus, RFD can be enabled, and some churn reduced.

The goal is to, with absolutely minimal change, ameliorate the danger of current RFD implementations and use. It is not a panacea, nor is it a deep and thorough approach to flap reduction.

## 3. RFD Parameters

The following RFD parameters are common to all implementations. Some may be tuned by the operator, some not.





| Parameter | Tunable? | Cisco | Juniper |
|---|---|---|---|
| Withdrawal | No | 1000 | 1000 |
| Re-Advertisement | No | 0 | 1000 |
| Attribute Change | No | 500 | 500 |
| Suppress Threshold | Yes | 2000 | 3000 |
| Half-Life (min) | Yes | 15 | 15 |
| Reuse Threshold | Yes | 750 | 750 |
| Max Suppress Time (min) | Yes | 60 | 60 |

Default RFD Paramaters of Juniper and Cisco

Table 1

## 4. Suppress Threshold Versus Churn

By turning RFD back on with the values recommended in Section 6 churn is reduced. Moreover, with these values, prefixes going through normal convergence are generally not damped.

[pelsser2011] estimates that, with a suppress threshold of 6,000, the BGP update rate is reduced by 19% compared to a situation without RFD enabled. With this 6,000 suppress threshold, 90% fewer prefixes are damped compared to use of a 2,000 threshold. I.e. far fewer well-behaved prefixes are damped.

Setting the suppress threshold to 12,000 leads to very few damped prefixes (1.7% of the prefixes damped with a threshold of 2,000, in the experiments in [pelsser2011] yielding an average hourly update reduction of 11% compared to not using RFD.





| Suppress Threshold | Damped Instances | % of Table Damped | Update Rate (one hour bins) |
|---|---|---|---|
| 2,000 | 43342 | 13.16% | 53.11% |
| 4,000 | 11253 | 3.42% | 74.16% |
| 6,000 | 4352 | 1.32% | 81.03% |
| 8,000 | 2104 | 0.64% | 84.85% |
| 10,000 | 1286 | 0.39% | 87.12% |
| 12,000 | 720 | 0.22% | 88.74% |
| 14,000 | 504 | 0.15% | 89.97% |
| 16,000 | 353 | 0.11% | 91.01% |
| 18,000 | 311 | 0.09% | 91.88% |
| 20,000 | 261 | 0.08% | 92.69% |

Damped Prefixes vs. Churn, from [pelsser2011]

Note overly-aggressive current default Suppress Threshold

Table 2

## 5. Maximum Penalty

It is important to understand that the parameters shown in Table 1, and the implementation's sampling rate, impose an upper bound on the penalty value, which we can call the 'computed maximum penalty'.

In addition, BGP implementations have an internal constant which we will call the 'maximum penalty' which the current computed penalty may not exceed.

## 6. Recommendations

The following changes are recommended:

Router Maximum Penalty: The internal constant for the maximum penalty value MUST be raised to at least 50,000.

Default Configurable Parameters: In order not to break existing operational configurations, BGP implementations SHOULD NOT change the default values in Table 1.





Minimum Suppress Threshold: Operators wishing damping which is much less destructive than current, but still somewhat aggressive SHOULD configure the Suppress Threshold to no less than 6,000.

Conservative Suppress Threshold: Conservative operators SHOULD configure the Suppress Threshold to no less than 12,000.

Calculate But Do Not Damp: Implementations MAY have a test mode

where the operator could see the results of a particular configuration without actually damping any prefixes. This will allow for fine tuning of parameters without losing reachability.

## 7. Security Considerations

It is well known that an attacker can generate false flapping to cause a victim's prefix(es) to be damped.

As the recommendations merely change parameters to more conservative values, there should be no increase in risk.

In fact, the parameter change to more conservative values should slightly mitigate the false flap attack.

## 8. IANA Considerations

This document has no IANA Considerations.

## 9. Acknowledgments

Nate Kushman initiated this work some years ago. Ron Bonica, Seiichi Kawamura, and Erik Muller contributed useful suggestions.

## 10. References

### 10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2439] Villamizar, C., Chandra, R., and R. Govindan, "BGP Route Flap Damping", RFC 2439, November 1998.

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.





[mao2002] Mao, Z. M., Govidan, R., Varghese, G., and Katz, R.,

"Route Flap Damping Excacerbates Internet Routing Convergence", In Proceedings of SIGCOMM , August 2002, <http://www.acm.org/sigcomm/sigcomm2002/papers/routedampening.pdf>.

[pelsser2011]
Pelsser, C., Maennel, O., Mohapatra, P., Bush, R., and Patel, K., "Route Flap Damping Made Usable", Passive and Active Measurement (PAM), March 2011, <http://archive.psg.com/110103.pam-rfd.pdf>.

[ripe378] Panigl, P. and Smith, P., "RIPE Routing Working Group Recommendations On Route-flap Damping", 2006, <http://www.ripe.net/ripe/docs/ripe-378>.

## 10.2. Informative References

[I-D.shishio-grow-isp-rfd-implement-survey]
Tsuchiya, S., Kawamura, S., Bush, R., and C. Pelsser, "Route Flap Damping Deployment Status Survey", draft-shishio-grow-isp-rfd-implement-survey-02 (work in progress), June 2011.

[huston2006]
Huston, G., "BGP Extreme Routing Noise", RIPE 52 , 2006, <http://meetings.ripe.net/ripe-52/presentations/ripe52-plenary-bgp-review.pdf>.

[ripe178] Barber, T., Doran, S., Karrenberg, D., Panigl, C., and Schmitz, J., "RIPE Routing-WG Recommendation for Coordinated Route-flap Damping Parameters", 2001, <http://www.ripe.net/ripe/docs/ripe-178>.

## Authors' Addresses

Cristel Pelsser
Internet Initiative Japan
Jinbocho Mitsui Buiding, 1-105
Kanda-Jinbocho, Chiyoda-ku, Tokyo 101-0051
JP

Phone: +81 3 5205 6464
Email: cristel@iij.ad.jp

Randy Bush
Internet Initiative Japan
5147 Crystal Springs
Bainbridge Island, Washington  98110
US

Phone: +1 206 780 0431 x1
Email: randy@psg.com

Keyur Patel
Cisco Systems
170 W. Tasman Drive
San Jose, CA  95134
US

Email: keyupate@cisco.com

Pradosh Mohapatra
Cisco Systems
170 W. Tasman Drive
San Jose, CA  95134
US

Email: pmohapat@cisco.com

Olaf Maennel
Loughborough University
Department of Computer Science - N.2.03
Loughborough
UK

Phone: +44 115 714 0042
Email: o@maennel.net