Making Route Flap Damping Usable

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 in order to reduce the high risks with RFD. The result is damping a non-trivial amount of long-term churn without penalizing well-behaved prefixes’ normal convergence process.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7196.
1. 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 the measurements of [PELSSER2011], [RIPE580] now recommends that RFD is usable with some changes to the parameters. Based on the same measurements, this document recommends adjusting a few RFD algorithmic constants and limits. The result is 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 provider 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.

1.1. 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 [RFD-SURVEY].

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to
be interpreted as described in RFC 2119 [RFC2119] only when they
appear in all upper case. They may also appear in lower or mixed
case as English words, without normative meaning.

3. RFD Parameters

The following RFD parameters are common to all implementations. Some
may be tuned by the operator, some not. There is currently no
consensus on a single set of default values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Tunable?</th>
<th>Cisco</th>
<th>Juniper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Withdrawal</td>
<td>No</td>
<td>1,000</td>
<td>1,000</td>
</tr>
<tr>
<td>Re-Advertisement</td>
<td>No</td>
<td>0</td>
<td>1,000</td>
</tr>
<tr>
<td>Attribute Change</td>
<td>No</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Suppress Threshold</td>
<td>Yes</td>
<td>2,000</td>
<td>3,000</td>
</tr>
<tr>
<td>Half-Life (min.)</td>
<td>Yes</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Reuse Threshold</td>
<td>Yes</td>
<td>750</td>
<td>750</td>
</tr>
<tr>
<td>Max Suppress Time (min.)</td>
<td>Yes</td>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>

Note: Values without units specified are dimensionless constants.

Table 1: Default RFD Parameters of Juniper and Cisco
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. [PELSSER2011] studies the number of prefixes damped over a week between September 29, 2010 and October 6, 2010. With this 6,000 suppress threshold, 90% fewer prefixes are damped compared to use of a 2,000 threshold. That is, far fewer well-behaved prefixes are damped.

Setting the suppress threshold to 12,000 leads to very few damped prefixes (0.22% of the prefixes were damped with a threshold of 12,000 in the experiments in [PELSSER2011], yielding an average hourly update reduction of 11% compared to not using RFD).

<table>
<thead>
<tr>
<th>Suppress Threshold</th>
<th>Damped Prefixes</th>
<th>% of Table Damped</th>
<th>Update Rate (one-hour bins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,000</td>
<td>43,342</td>
<td>13.16%</td>
<td>53.11%</td>
</tr>
<tr>
<td>4,000</td>
<td>11,253</td>
<td>3.42%</td>
<td>74.16%</td>
</tr>
<tr>
<td>6,000</td>
<td>4,352</td>
<td>1.32%</td>
<td>81.03%</td>
</tr>
<tr>
<td>8,000</td>
<td>2,104</td>
<td>0.64%</td>
<td>84.85%</td>
</tr>
<tr>
<td>10,000</td>
<td>1,286</td>
<td>0.39%</td>
<td>87.12%</td>
</tr>
<tr>
<td>12,000</td>
<td>720</td>
<td>0.22%</td>
<td>88.74%</td>
</tr>
<tr>
<td>14,000</td>
<td>504</td>
<td>0.15%</td>
<td>89.97%</td>
</tr>
<tr>
<td>16,000</td>
<td>353</td>
<td>0.11%</td>
<td>91.01%</td>
</tr>
<tr>
<td>18,000</td>
<td>311</td>
<td>0.09%</td>
<td>91.88%</td>
</tr>
<tr>
<td>20,000</td>
<td>261</td>
<td>0.08%</td>
<td>92.69%</td>
</tr>
</tbody>
</table>

Note: the current default Suppress Threshold (2,000) is overly aggressive.

Table 2: Damped Prefixes vs. Churn, from [PELSSER2011]

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’, and the current computed penalty may not exceed it.

6. Recommendations

Use of the following values is 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, existing BGP implementations, including the examples in Table 1, SHOULD NOT change their default values.

Minimum Suppress Threshold: Operators that want damping that is much less destructive than the current damping, 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 can 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. Acknowledgments

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

9.1. Normative References


9.2. Informative References


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  
EMail: randy@psg.com

Keyur Patel  
Cisco Systems  
170 W. Tasman Drive  
San Jose, CA 95134  
US  
EMail: keyupate@cisco.com

Pradosh Mohapatra  
Sproute Networks  
41529 Higgins Way  
Fremont, CA 94539  
US  
EMail: mpradosh@yahoo.com

Olaf Maennel  
Loughborough University  
Department of Computer Science - N.2.03  
Loughborough  
UK  
Phone: +44 115 714 0042  
EMail: o@maennel.net