Route Flap Damping with Assured Reachability

Pei-chun Cheng, Jong Han Park, Keyur Patel, Lixia Zhang

UCLA
Contents

1. Introduction to RFD
2. Toward reviving RFD
3. Evaluation
4. Summary
BGP and Flap Damping

- Failures (flaps) happen from time to time
  - HW/SW defects
  - Mis-configurations
  - ...

- BGP runs a flat routing space
  - Event goes as far as possible
  - Ripple the entire Internet
  - $O(N \times N)$

- In 90's, RFD is invented to mitigate such persistent processing load [4]
How does RFD work?

For each route:
- Keep a penalty value
- Track the flapping history

Update the penalty using AIMD

Two thresholds:
- Suppress if above suppress-limit
- Re-use if below reuse-limit

Looks fine, but . . .
- What is a “flap”?  
- Difficult to identify

Problem: Reachability Loss

- In 2002, Mao et al. show that RFD *falsely* suppress occasional route changes, NANOG26 [2]
  - Unexpected interaction btw. RFD and “path exploration”
  - Lose reachability

- Over years, several enhancements are proposed
  - S-RFD [3], RFD+ [1], RFD-RCN [5], F-RFD[6]
  - All tried to correctly identify route flaps

- None is successful!
  - Introduce additional BGP complexity
  - Do *not* solve the reachability loss problem

- As the result, in 2006, RIPE suggested *NOT* to enable RFD in production networks
Why revisiting RFD now?

- Long term oscillation does exist
  - Responsible for a large fraction of updates
  - Last for days or even weeks

- Increasing popularity of real-time or VoIP applications
  - Suffers from persistent route flaps and slow convergence
  - Correlated with 50% of VoIP quality degradation

- Operators ask for faster convergence
  - Disable other rate limiting measures
  - Greatly increase the number of updates

► Our goal: fix the reachability loss and hopefully revive RFD
Let’s change the mindset!

Current approach
- *Flap detection* and *flap suppression* are tied together
  - False detection could hurt reachability
  - Sophisticated RFD enhancements are introduced
- Tradeoff between *reachability* and stability

Our response
- **Decouple** *flap detection* and *flap suppression*
  - This work only modify the flap suppression behavior
  - A simple, yet deployable patch
  - Compatible to the existing efforts
- Reachability comes first!
  - Only tradeoff between *optimality* and stability
Reachability in BGP

- **Protector routes**
  1. Alternative routes to the same prefix
     Ex: \((P.1,A)\) protects \((P.1,B)\)
  2. Routes to the covering prefixes
     Ex: \((P.1,A)\) protects \((P.1.2,A)\) and \((P.1.2,B)\)

only suppress when there exist **protector routes**
Illustrating the idea

Prefix Space

Instable Prefixes

* Not drawn to scale
Illustrating the idea

Prefix Space

RFD

Instable Prefixes

* Not drawn to scale

May lose reachability
Prefix Space

Illustrating the idea

- RFD
- Instable Prefixes
- May lose reachability
- S-RFD, RFD+, etc

* Not drawn to scale
Illustrating the idea

* Not drawn to scale

Prefix Space

- Instable Prefixes
- Prefixes have protector routes
- May lose reachability
- S-RFD, RFD+, etc

Only suppress these prefixes
Illustrating the idea

Prefix Space

- RFD
- Instable Prefixes
- Prefixes have protector routes
- May lose reachability
- S-RFD, RFD+, etc

* Not drawn to scale

Only suppress these prefixes
Q1: Do protectors exist?

- A quick peek of routing data (LINX exchange point)
  - 50% prefixes have protector prefixes
  - Num. paths $\approx$ num. providers

![Fraction of Prefixes (CDF)](image1)

![Number of Covering Prefixes](image2)

![Fraction of Prefixes (CDF)](image3)

![Number of Nexthop Routers](image4)
Q2: How to preserve reachability?

+RG (Reachability Guard): two checks

1. Reachability Check
2. Early Release
Q2: How to preserve reachability?

+RG (Reachability Guard): two checks

1. Reachability Check
2. Early Release
Q2: How to preserve reachability?

- **+RG (Reachability Guard): two checks**
  1. Reachability Check
  2. Early Release

![Diagram](image-url)
Q2: How to preserve reachability?

+RG (Reachability Guard): two checks
1. Reachability Check
2. Early Release

Diagram:
- **P**
  - **P.1**
    - **P.1.1**
      - A
    - **A, B**
  - **P.1.2**
    - A, [B]
  - **P.2**
    - [B]
Q2: How to preserve reachability?

+RG (Reachability Guard): two checks

1. Reachability Check
2. Early Release
Evaluation methodology

Our approach ~ simulation with BGP feeds

- A partial BGP simulator
  - Flap damping
    - Vanilla RFD (default cisco parameters)
    - +RG, +RG (without early release)
  - Simple shortest path selection

- Actual BGP data source
  - From exchange points (collected by RIPE/RIS)
    - LINX
    - Others
  - One week in 2009 December (168 hours)
  - Full IPv4 routing table (~ provider-customer)
Evaluation metrics

- Measure the performance and trade-off

1. Reachability
2. Tradeoff
   (Quality of Reachability)
   - Router processing load
   - Stability
   - Route preference
An example case

- An example prefix 137.119.0.0/20, December 1st
  - No covering prefix
  - Reachable via two different peers
    - Preferred path flaps persistently
    - Less preferred path is stable (failed occasionally)

### Input feed

- Update pattern (A/W)
- Preferred Route
- Alternative Route

### Damping result

- Reachable pattern (R/U)
- No Damping
- RFD
- RFD+RG
- Route Change
An example case

- An example prefix 137.119.0.0/20, December 1st
  - No covering prefix
  - Reachable via two different peers
    - Preferred path flaps persistently
    - Less preferred path is stable (failed occasionally)

Input feed

- Update pattern (A/W)

Damping result

- Reachable pattern (R/U)
  - No Damping
  - RFD
  - RFD+RG
  - Route Change
  - Reach. loss

Cheng et al. (UCLA)

Nov 2010 13 / 22
An example case

- An example prefix 137.119.0.0/20, December 1st
  - No covering prefix
  - Reachable via two different peers
    - Preferred path flaps persistently
    - Less preferred path is stable (failed occasionally)

Input feed

Damping result

Cheng et al. (UCLA)
Preserve reachability

- All damped prefixes (41,086 prefixes)

- For RFD, 20% of damped prefixes lose reachability
  - 3% (1.3K) prefixes lose more than 10 minutes reachability
- RFD+RG guarantees reachability!
✓ **Gain - Fewer updates**

- Reduce number of updates
  - \( \sim \) best path selection
  - RFD (26.0%)
  - RFD+RG (24.2%)
- +RG saves a little less than RFD
  - 2.0% of updates
  - Essential updates for maintaining reachability

### Evaluation Summary

<table>
<thead>
<tr>
<th></th>
<th>Saved</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>RG</td>
<td>1,468,421</td>
<td>24.21</td>
</tr>
<tr>
<td>RG (w/o ER)</td>
<td>1,539,993</td>
<td>25.39</td>
</tr>
<tr>
<td>RFD</td>
<td>1,576,991</td>
<td>26.00</td>
</tr>
</tbody>
</table>

Number of total updates: 6,065,353
Gain - Fewer route changes

- Reduce next hop changes (~ FIB changes)
  - RFD (23.5%)
  - RFD+RG (21.7%)

- +RG saves a little less than RFD
  - 1.8% number of next hop changes

<table>
<thead>
<tr>
<th></th>
<th>Saved</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>RG</td>
<td>205,461</td>
<td>21.70</td>
</tr>
<tr>
<td>RG(w/o ER)</td>
<td>212,657</td>
<td>22.46</td>
</tr>
<tr>
<td>RFD</td>
<td>222,978</td>
<td>23.55</td>
</tr>
</tbody>
</table>

Number of total next hop changes: 946,829
Tradeoff - Stability and Preference

All damped prefixes (41,086 prefixes)

+RG makes BGP choose the less preferred route, if...
- Preferred path is unstable
- Less preferred path is stable

Deviate from the preference settings
- Might be an acceptable trade-off?
Result of other exchange points

- Exchange points across different topological locations
  - Guarantee reachability
  - Overhead reduction could be different

<table>
<thead>
<tr>
<th>Location</th>
<th>Reach. loss</th>
<th>Damped prefixes</th>
<th>Reduced updates (%)</th>
<th>Reduced NH changes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINX</td>
<td>0</td>
<td>46,488</td>
<td>24.21</td>
<td>21.70</td>
</tr>
<tr>
<td>AMS-IX</td>
<td>0</td>
<td>13,464</td>
<td>13.28</td>
<td>19.09</td>
</tr>
<tr>
<td>CIXP</td>
<td>0</td>
<td>9,125</td>
<td>5.82</td>
<td>16.40</td>
</tr>
<tr>
<td>NETNOD</td>
<td>0</td>
<td>45,496</td>
<td>27.16</td>
<td>20.30</td>
</tr>
<tr>
<td>MIX</td>
<td>0</td>
<td>31,543</td>
<td>11.10</td>
<td>14.33</td>
</tr>
<tr>
<td>NYIIX</td>
<td>0</td>
<td>15,907</td>
<td>8.99</td>
<td>15.20</td>
</tr>
<tr>
<td>DE-CIX</td>
<td>0</td>
<td>26,708</td>
<td>17.89</td>
<td>27.07</td>
</tr>
<tr>
<td>MSK-IX</td>
<td>0</td>
<td>29,314</td>
<td>12.97</td>
<td>19.77</td>
</tr>
</tbody>
</table>
This work . . .

- a simple addition to guarantee reachability
  - Address the long overdue reachability loss problem
  - Offer a better trade-off

- compatible with all existing damping schemes
  - Previous efforts are not wasted
  - Incrementally deployable

- NOT yet another damping scheme

- NOT a routing enhancement, but a defensive measure

Future questions: implementation issues, overhead, system wide impact . . .
Thank you!
References


Z. M. Mao, R. Govindan, G. Varghese, and R. H. Katz. Route Flap Damping: Harmful?


References (cont.)
