

Route Flap Damping with Assured Reachability

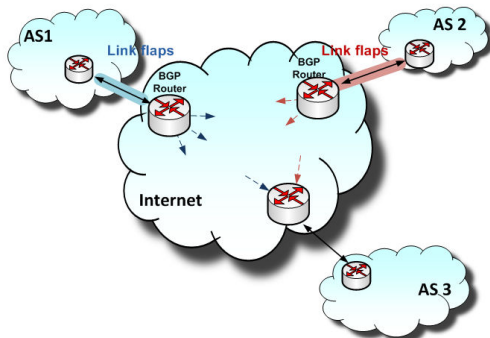
Pei-chun Cheng, Jong Han Park, Keyur Patel, Lixia Zhang

UCLA

Contents

- 1 Introduction to RFD
- 2 Toward reviving RFD
- 3 Evaluation
- 4 Summary

BGP and Flap Damping

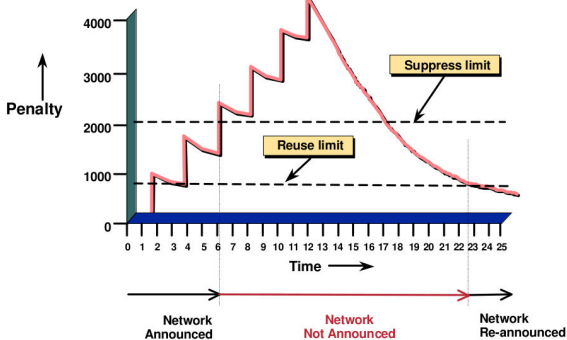


- Failures (flaps) happen from time to time
 - HW/SW defects
 - Mis-configurations
 - ...
- BGP runs a flat routing space
 - Event goes as far as possible
 - Ripple the **entire Internet**
 $\sim O(N \times N)$

► 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 ¹



- 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

¹ source: <ftp://ftp-eng.cisco.com/cons/workshops/isp-workshop/StudentCD-Rev-E/Routing/b5-1up.pdf>

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

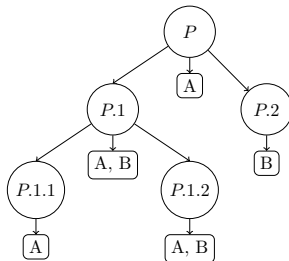
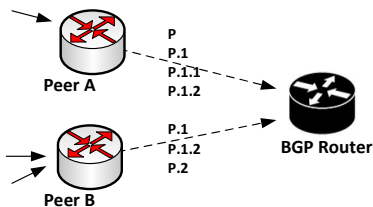
Let's change the mindset!

- 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

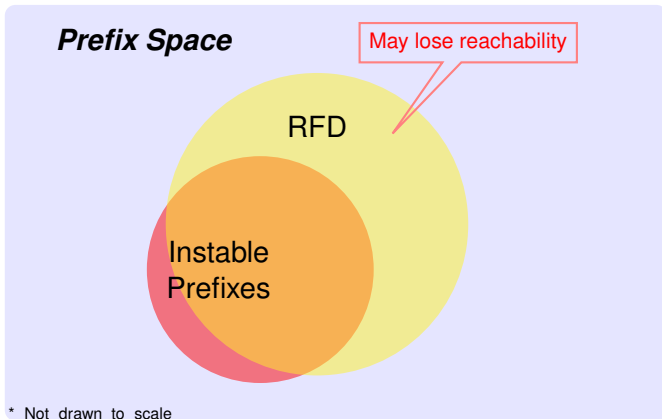
Prefix Space



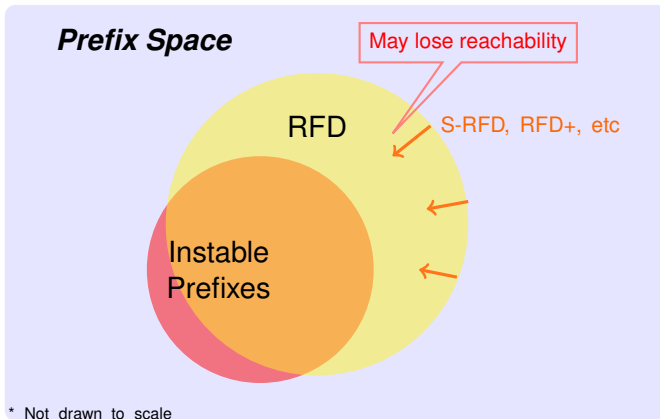
Instable
Prefixes

* Not drawn to scale

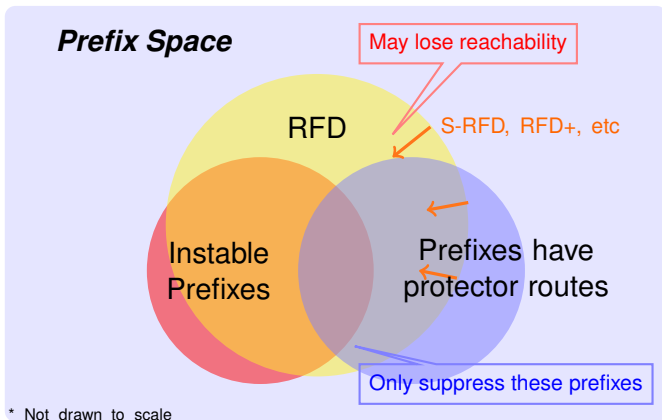
Illustrating the idea



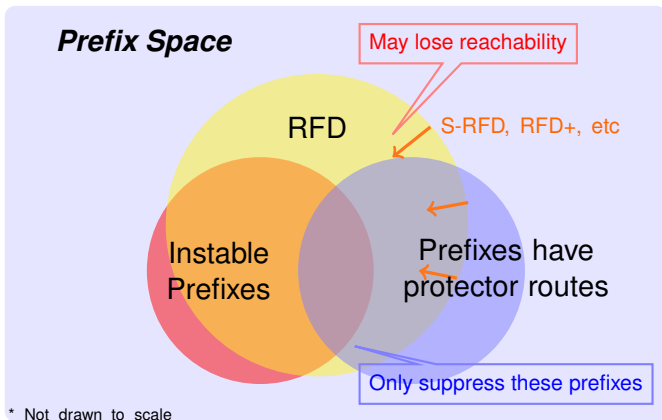
Illustrating the idea



Illustrating the idea

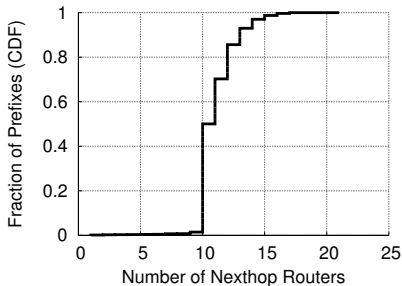
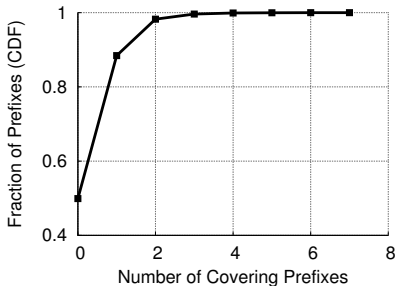


Illustrating the idea



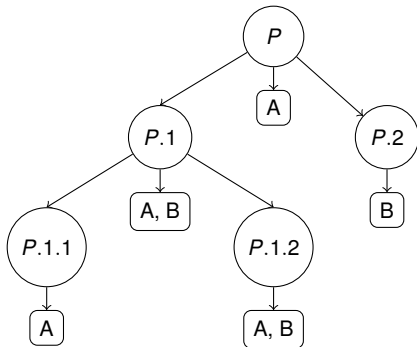
Q1: Do protectors exist?

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



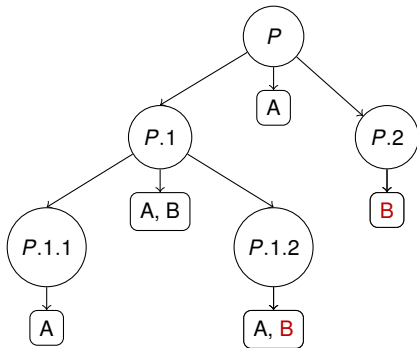
Q2: How to preserve reachability?

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



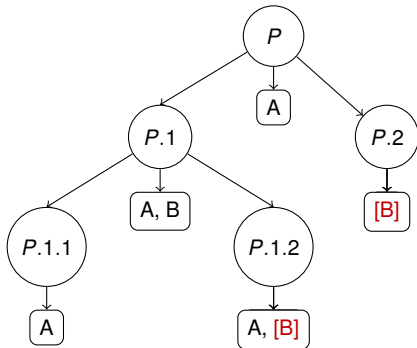
Q2: How to preserve reachability?

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



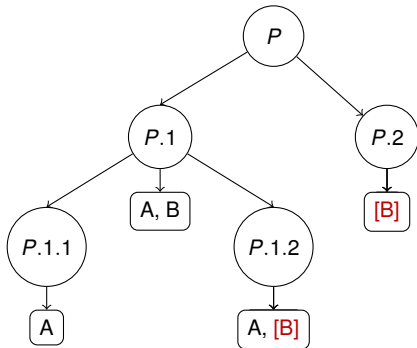
Q2: How to preserve reachability?

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



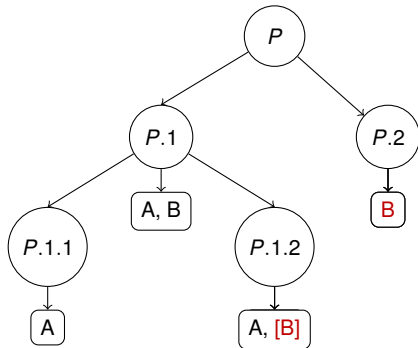
Q2: How to preserve reachability?

- +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



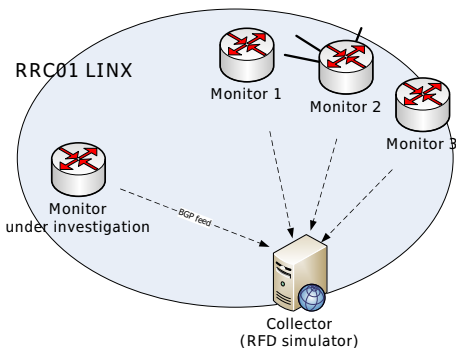
Evaluation methodology

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 ○

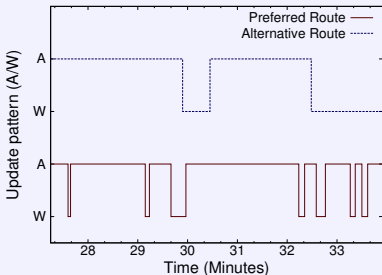
2 Tradeoff
(Quality of Reachability)

- Router processing load ✓
- Stability ✓
- Route preference ✗

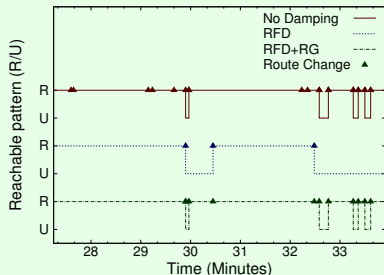
An example case

- 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



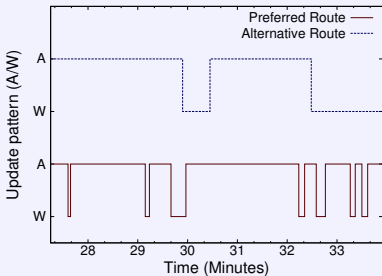
Damping result



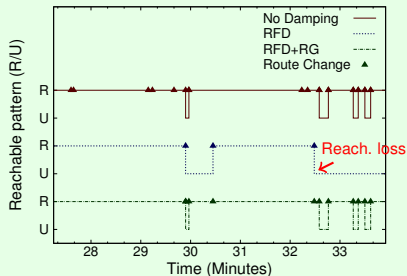
An example case

- 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



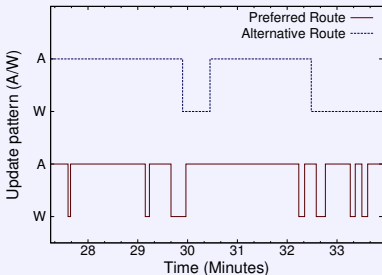
Damping result



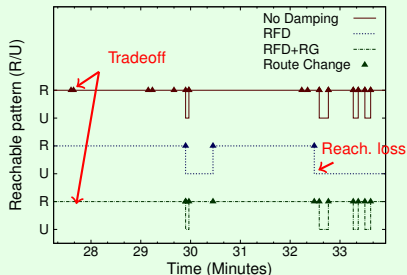
An example case

- 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

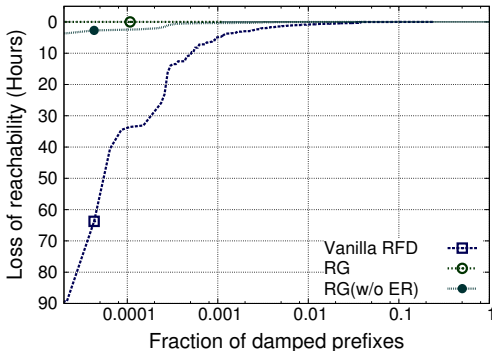


Damping result



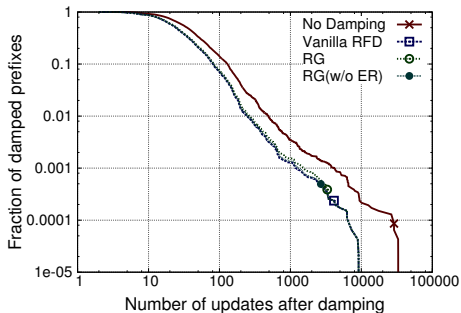
Preserve reachability

- ▶ 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

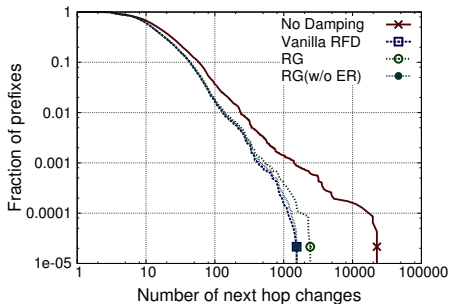


	Saved	Ratio
RG	1,468,421	24.21
RG(w/o ER)	1,539,993	25.39
RFD	1,576,991	26.00

Number of total updates: 6,065,353

- Reduce number of updates (~ 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

✓ Gain - Fewer route changes



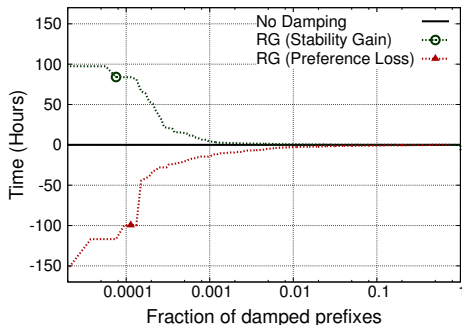
- Reduce next hop changes (~ FIB changes)
 - RFD (23.5%)
 - RFD+RG (21.7%)
- +RG saves a little less than RFD
 - 1.8 % number of next hop changes

	Saved	Ratio
RG	205,461	21.70
RG(w/o ER)	212,657	22.46
RFD	222,978	23.55

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

	Location	Reach. loss	Damped prefixes	Reduced updates (%)	Reduced NH changes (%)
LINX	London	0	46,488	24.21	21.70
AMS-IX	Amsterdam	0	13,464	13.28	19.09
CIXP	Geneva	0	9,125	5.82	16.40
NETNOD	Stockholm	0	45,496	27.16	20.30
MIX	Milan	0	31,543	11.10	14.33
NYIIX	New York	0	15,907	8.99	15.20
DE-CIX	Frankfurt	0	26,708	17.89	27.07
MSK-IX	Moscow	0	29,314	12.97	19.77

Summary

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. Duan, J. Chandrashekar, J. Krasky, K. Xu, and Z.-L. Zhang.
Damping BGP route flaps.

In Proc. IEEE International Conference on Performance, Computing, and Communications, pages 131–138, 2004.



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



Z. M. Mao, R. Govindan, G. Varghese, and R. H. Katz.
Route flap damping exacerbates internet routing convergence.

In SIGCOMM '02: Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications, pages 221–233, New York, NY, USA, 2002. ACM.



C. Villamizar, R. Chandra, and R. Govindan.
BGP Route Flap Damping.

RFC 2439 (Proposed Standard), Nov. 1998.

References (cont.)



B. Zhang, D. Pei, D. Massey, and L. Zhang.
Timer Interaction in Route Flap Damping.

*In Proc. 25th IEEE International Conference on Distributed Computing Systems
ICDCS 2005*, pages 393–403, 2005.



K. Zhang and S. Wu.

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

LECTURE NOTES IN COMPUTER SCIENCE, 4479:1200, 2007.