# BGP Route Flap Damping Algorithms

Wei Steve Shen
wshen@cs.sfu.ca

Communication Networks Laboratory
http://www.ensc.sfu.ca/cnl
Simon Fraser University

cnl communication networks laboratory

# Roadmap

- Introduction to Route Flap Damping (RFD)
- ns-2 implementation of RFD algorithms
- RFD simulation scenarios
- Performance analysis of RFD algorithms
- Improvements to RFD algorithms
- Conclusions and references

# BGP background

- Border Gateway Protocol:
  - inter-AS (Autonomous System) routing protocol
  - used to exchange network reachability information among BGP systems (routers)
  - BGP-4 is the current de facto inter-domain routing protocol
  - path vector protocol:
    - distribute route path information to peers
  - incremental:
    - send update messages as routing tables change

# Important terms

- Prefix:
  - specifies a network destination
  - represented by an IP address block, which consists of a 32-bit address and a mask length indicating the network size: 192.168.1.0/24
- Route:
  - defines a path to a particular destination
  - contains multiple attributes of the path: AS path, origin, next hop
- Route preference:
  - a metric (integer) indicating the degree of preference of a route in the BGP decision process

# Important terms

- Update:
  - advertisement (announces a feasible route to peers)
  - withdrawal (removes existing unfeasible routes from service)
- Convergence time:
  - BGP speakers (routers) may explore a number of transient routes before converging to a stable route
  - time difference between the instant when the origin router sent its update message and the instant when the last update message that resulted from the original update has been processed

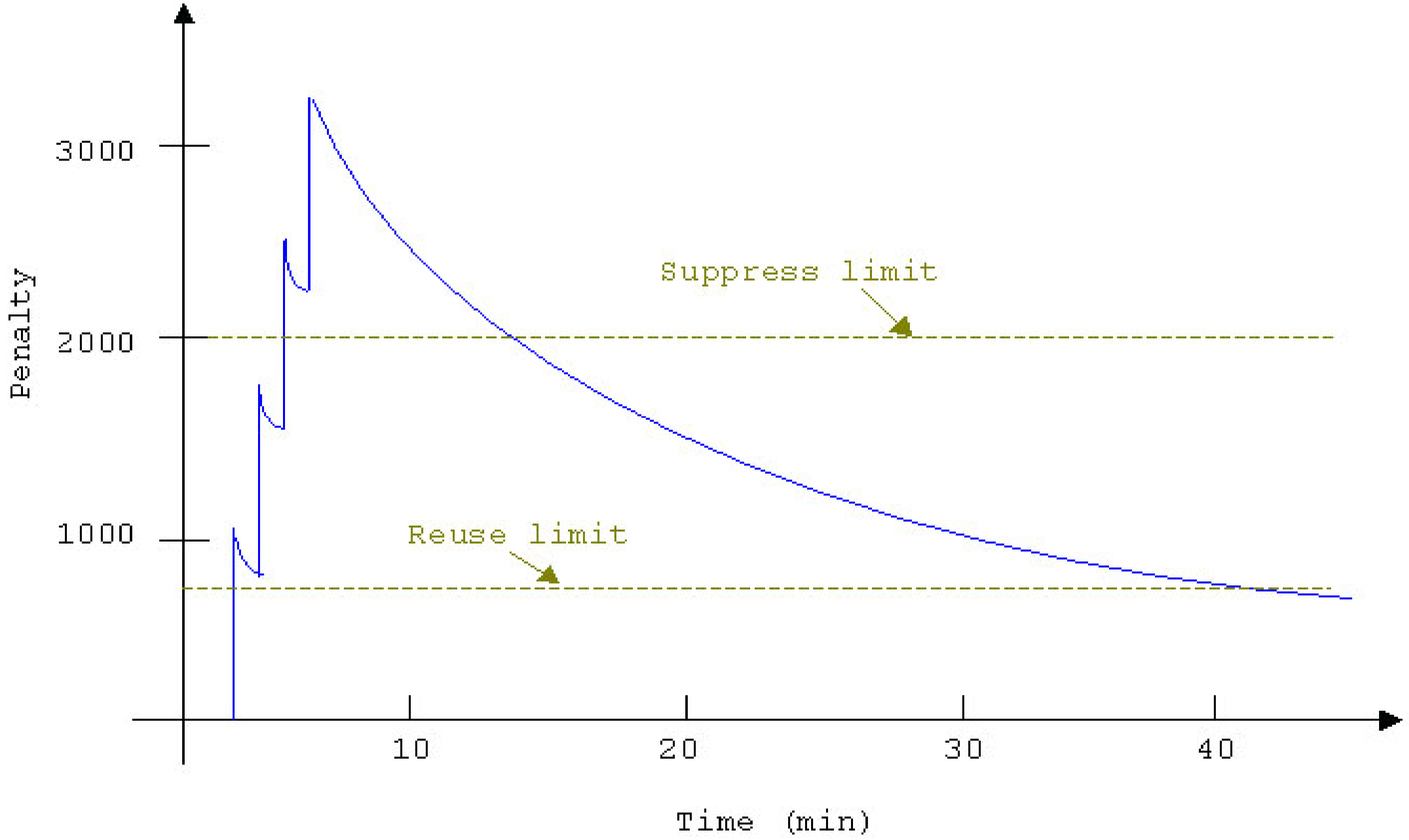
# Introduction to RFD

- A route flaps when it oscillates between being available and being unavailable
- Routing oscillations may be caused by:
  - router configuration errors
  - transient data link failures
  - software defects
- BGP employs RFD mechanisms to prevent persistent routing oscillations:
  - reduce the number of BGP update messages sent by BGP speakers
  - decrease the processing load imposed on BGP speakers

# Common approaches to RFD

- Assign a penalty to a route and increment the penalty value when the route flaps
- The route is suppressed and not advertised further when the penalty exceeds the suppress limit
- Penalty of a route decays exponentially based on the half life parameter
- If the penalty decreases below the reuse limit, the route is reused and may be advertised again

# Route penalty vs. time



# RFD algorithms

- Existing RFD algorithms, which identify and penalize route flaps:
  - Original RFD
  - Selective RFD
  - RFD+

Original RFD:

C. Villamizar, R. Chandra, and R. Govindan, “BGP route flap damping,” *IETF RFC 2439*, Nov. 1998.

Selective RFD:
Z. Mao, R. Govindan, G. Varghese, and R. Katz, “Route flap damping exacerbates Internet routing convergence,” in *Proc. SIGCOMM 2002*, Pittsburgh, PA, Aug. 2002, pp. 221–233.

RFD+:
Z. Duan, J. Chandrashekar, J. Krasky, K. Xu, and Z. Zhang, “Damping BGP route flaps,” in *Proc. IPCCC 2004*, Phoenix, AZ, Apr. 2004, pp. 131–138.

# Original RFD

- Defined in RFC 2439
- Each route withdrawal or route attribute change (route replacement) is considered to be a flap and is penalized accordingly
- Algorithm:

```
when receiving a route r with prefix d from peer j
 if r is a withdrawn route
      a flap is identified: route withdrawal
 else if (r is an advertised route and $r \neq p$)
      a flap is identified: route attribute change
 p = r
```

p: previous route with prefix d advertised from peer j

# Original RFD

- It may significantly delay the convergence of relatively well-behaved routes (routes that flap only occasionally):
  - BGP searches for alternatives if a route is withdrawn
  - this path exploration leads to increase of penalty due to interim updates
  - BGP may suppress a route due to a single route withdrawal

# Selective RFD

- Distinguishes path explorations from genuine route flaps:
  - routes are selected in order of non-increasing preference during path exploration after withdrawal
- How to identify flaps:
  - sender attaches route preference to each route advertisement
  - receiver compares the current route with previous route in terms of route preference
  - a flap is identified if a change of direction in route preference is detected (a decrease followed by an increase)

Z. Mao, R. Govindan, G. Varghese, and R. Katz, “Route flap damping exacerbates Internet routing convergence,” in *Proc. SIGCOMM 2002*, Pittsburgh, PA, Aug. 2002, pp. 221–233.

# Selective RFD

- Algorithm:

```
when receiving route r with prefix d from peer j
 if (r is a withdrawn route)
      remember it: a potential flap
 else
      if (rp(r) > rp(p) and rp(p) < rp(f))
          a flap is identified (add any potential flap)
      elseif (rp(r) < rp(p) and rp(p) > rp(f))
          a flap is identified (add any potential flap)
      f = p;  p = r; remove any potential flap
```

rp(x): route preference of a route x

p: advertised route previous to current route r, with prefix d received from peer j

f: advertised route previous to route p, with prefix d received from peer j

# Selective RFD

- Simulations in small networks indicated that selective RFD identifies genuine flaps better than original RFD
- Selective RFD assumes incorrectly that changes in route preference are monotonic during path exploration:
  - currently feasible paths at the router may change with time
  - a better path may become available afterwards during path exploration

Z. Duan, J. Chandrashekar, J. Krasky, K. Xu, and Z. Zhang, “Damping BGP route flaps,” in *Proc. IPCCC 2004*, Phoenix, AZ, Apr. 2004, pp. 131–138.

# RFD+

- Overcomes the problem of the selective RFD algorithm
- A flap is identified when:
  - current route has a higher degree of preference than the previous route
  - BGP speaker has received the current route more than once since its previous flap
- Simulations in small networks indicated that RFD+ could correctly identify route flaps in the case of a single flap

Z. Duan, J. Chandrashekar, J. Krasky, K. Xu, and Z. Zhang, “Damping BGP route flaps,” in *Proc. IPCCC 2004*, Phoenix, AZ, Apr. 2004, pp. 131–138.

# RFD+

- Algorithm:

```
when receiving a route r with prefix d from peer j
if ($r \notin R(d, j)$ )
    insert r into R(d, j)
else if ($r \in R(d, j)$ and $rp(r) > rp(p)^*$)
    a flap is identified
    clear R(d, j)
```

R(d, j): set of all routes with prefix d received from peer j
rp(x): route preference of a route x
p: route previous to current route r, with prefix d received from peer j
* Route preference for a withdrawal is considered to be the lowest

# Key difference: how to identify flaps?

- Original RFD:
  - any route withdrawal or route attribute change is considered a flap
- Selective RFD:
  - a flap is identified if a change of direction in route preference is detected (a decrease followed by an increase)
- RFD+:
  - a flap is identified if the current route has a higher preference than the previous one and the BGP speaker has received the current route more than once since the previous flap

# Example

Routes with a particular prefix from a particular peer:

| Type | AS path |
|---|---|
| A: | 1 3 5 |
| A: | 1 3 7 5 |
| A: | 1 3 7 9 5 |
| W | |
| A: | 1 3 5 |

Time

Number of identified flaps:

- Original RFD: 4
- Selective RFD: 2
- RFD+: 1

# Roadmap

- Introduction to Route Flap Damping (RFD)
- ns-2 implementation of RFD algorithms
- RFD simulation scenarios
- Performance analysis of RFD algorithms
- Improvements to RFD algorithms
- Conclusions and references

# ns-2 implementation of RFD

- Based on a BGP model developed for network simulator ns-2: ns-BGP 2.0
- Relevant source code ported from the SSFNet BGP-4 module for original RFD and selective RFD and made necessary modifications
- We implemented RFD+ in ns-2
- Two improvements to RFD have been added: modified RFD+ and combined RFD

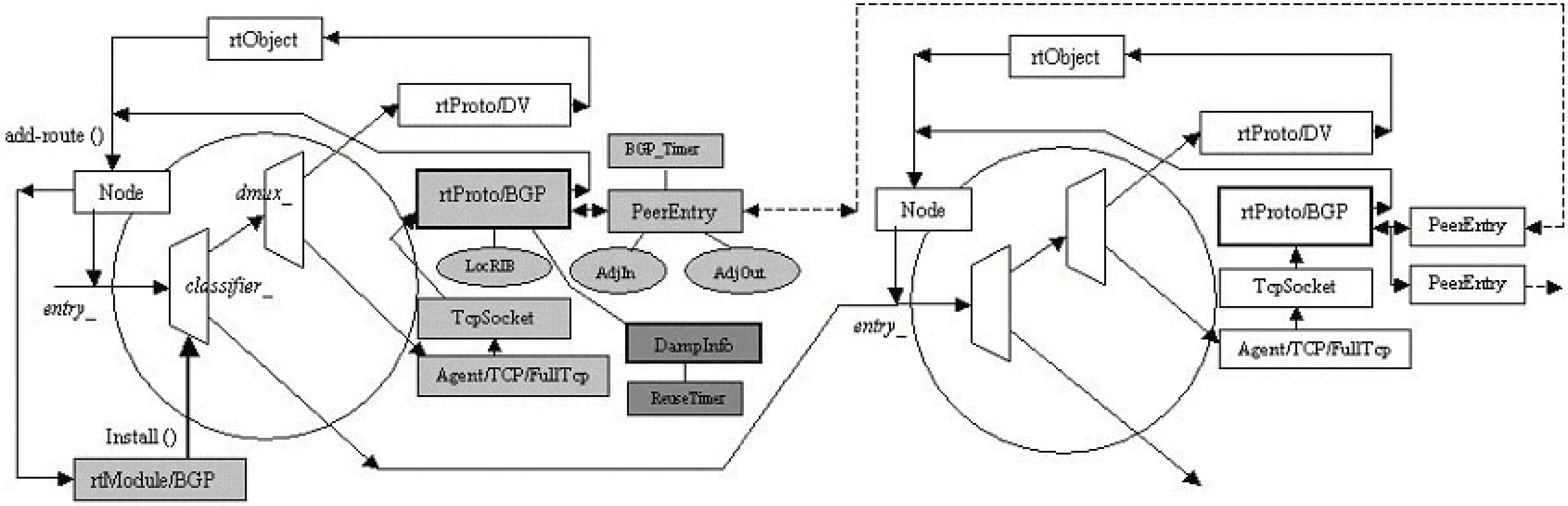
ns-BGP 2.0: http://www.ensc.sfu.ca/~ljilja/cnl/projects/BGP/

SSFNet: http://www.ssfnet.org/

RFD-AMRAI BGP: http://www.ensc.sfu.ca/~ljilja/cnl/projects/RFD-AMRAI/

# ns-2 Implementation of RFD

- Routing structure of modified ns-BGP with RFD:



# ns-2 implementation of RFD

- New C++ classes:
  - DampInfo: stores the damping structure for a prefix advertised from a peer of a BGP speaker and implements all five damping algorithms
  - ReuseTimer: keeps track of the reuse timer associated with a suppressed route
  - VecRoutes: maintains an array of interim routes
- Modified C++ files and tcl files:
  - implement route flap damping mechanisms when receiving updates and making routing decisions
  - set default global variables used in RFD algorithms

# Roadmap

- Introduction to Route Flap Damping (RFD)
- ns-2 implementation of RFD algorithms
- RFD simulation scenarios
- Performance analysis of RFD algorithms
- Improvements to RFD algorithms
- Conclusions and references

# ns-2 simulations of RFD

- Four elements of a simulation scenario:
  - network topology
  - inter-arrival time between updates
  - simulation time
  - nature of flaps

# Network topology

- Generated by network topology generator BRITE:
    - AS-level topologies
    - Generalized Linear Preference (GLP) model
    - network sizes: 100, 200, 300, 400, and 500 nodes
- Built from genuine BGP routing tables:
    - network sizes: 29 and 110 nodes
- Topologies built from routing tables are more densely connected than topologies generated by BRITE

BRITE: http://www.cs.bu.edu/brite

Multi-AS topologies from routing tables: http://www.ssfnet.org/Exchange/gallery/asgraph

GLP:

T. Bu and D. Towsley, “On distinguishing between Internet power law topology generators,” in *Proc. INFOCOM 2002*, New York, NY, June 2002, pp. 638–647.

# Inter-arrival time between updates

- Use at least three values for the inter-arrival time between updates:
  - a value smaller than the default MRAI value of 30 s: 10 s
  - an intermediate value: 100 s
  - a value large enough for BGP to converge: 1,000 s

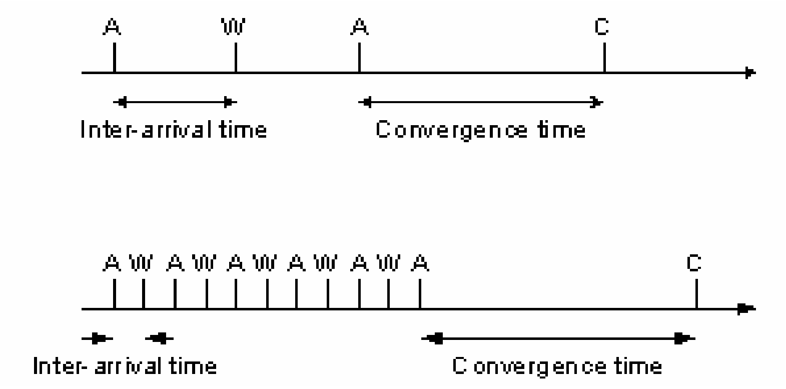
MRAI: Minimum Route Advertisement Interval

# Simulation time

- Depends on:
  - with route suppression period
  - without route suppression period
- Comparisons illustrate the impact of route suppression on individual BGP speakers and on the overall network

# Nature of flaps

- Occasional flaps:
  - one flap
  - inter-arrival time between updates: 1,000 s
- Persistent flaps:
  - five flaps
  - inter-arrival time between updates: 300 s



A: Advertise

W: Withdraw

C: Converge

# Nature of flaps: inter-arrival time between updates

- 1,000 s:
  - sufficiently large for BGP to converge in simulated networks
  - enables adequate comparison of RFD algorithms in identifying route flaps and examining the impact of route suppression in the case of occasional flaps
- 300 s:
  - suggests that route flaps occur at more frequent intervals
  - a route suppressed due to persistent flaps remains suppressed after the origin router sends its last advertisement
  - this route suppression indicates that the RFD algorithm can effectively suppress persistent flaps

# Roadmap

- Introduction to Route Flap Damping (RFD)
- ns-2 implementations of RFD algorithms
- RFD simulation scenarios
- Performance analysis of RFD algorithms
- Improvements to RFD algorithms
- Conclusions and references

# RFD performance analysis

- Default MRAI value: 30 s
- Default Cisco settings:

| | |
|---|---|
| Suppress limit | 2000 |
| Reuse limit | 750 |
| Half life (s) | 900 |
| Withdrawal penalty | 1000 |
| Attribute change penalty | 500 |
| Re-advertisement penalty | 0 |
| Maximum suppression time (s) | 3600 |

# RFD performance analysis

- Compare:
  - RFD disabled, original RFD, selective RFD, and RFD+

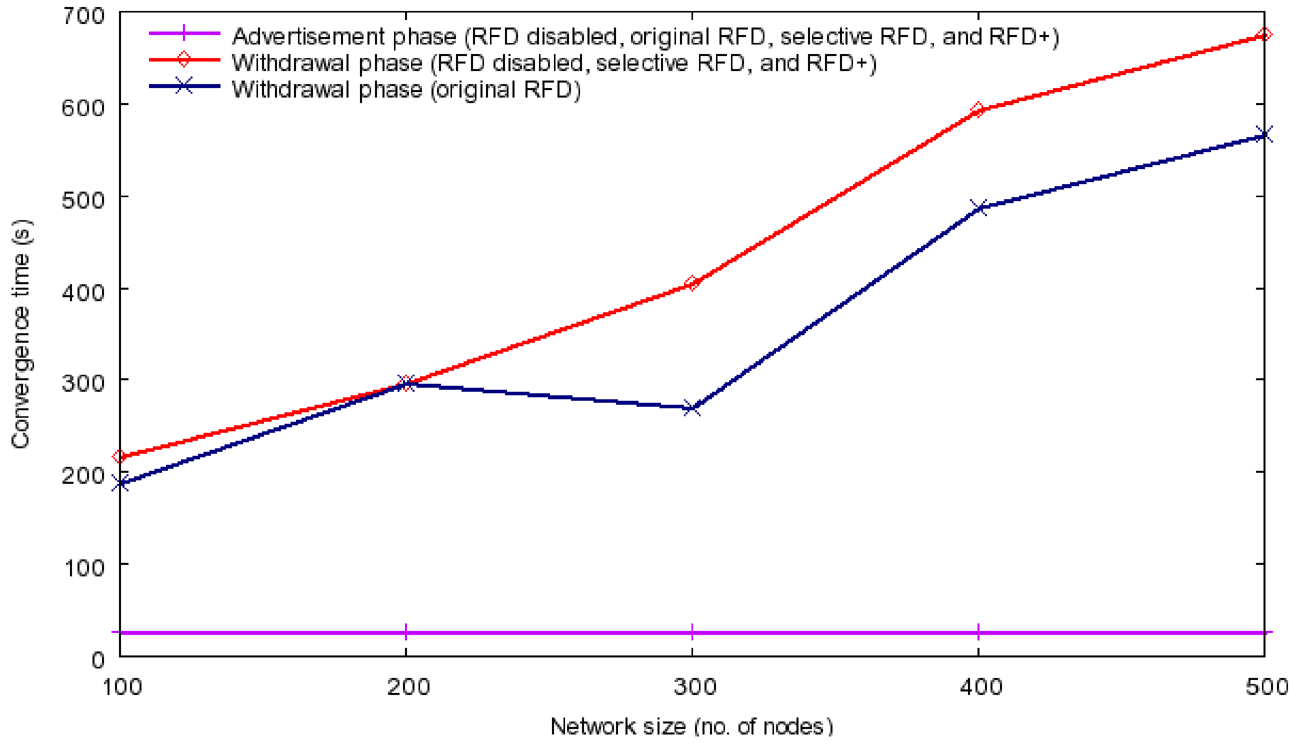
    in cases of occasional and persistent flaps in various networks
- Examine:
  - advertisement and withdrawal phases
  - effect of inter-arrival time between updates
  - effect of location of the origin router

# RFD performance analysis

- Important variables to compare:
  - overall number of updates
  - overall number of reported flaps
  - number of flaps reported by each BGP speaker
  - maximum number of flaps associated with a single peer of each BGP speaker
  - overall number of suppressions caused by all the flaps
  - convergence time

# Advertisement vs. withdrawal: convergence time

- A withdrawal message causes BGP to converge significantly slower than in the case of an advertisement message:

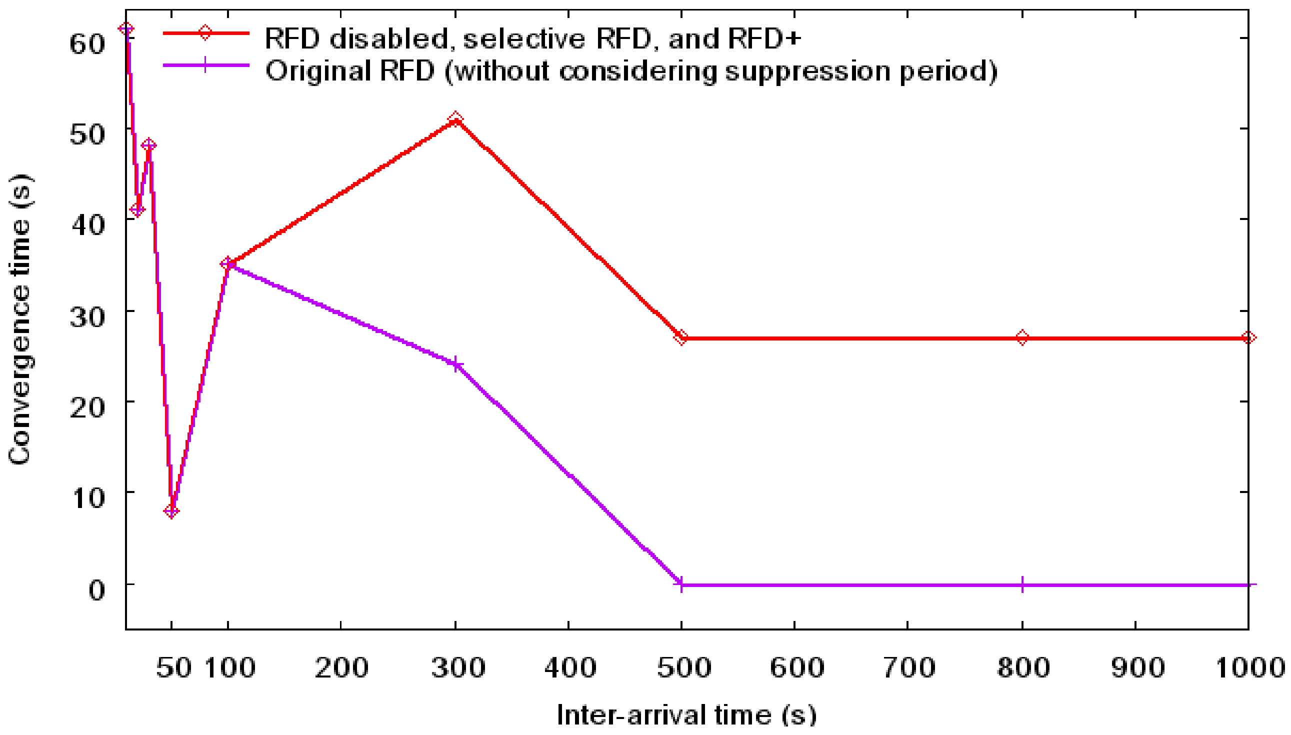


# Advertisement vs. withdrawal

- Withdrawal phase: original RFD has the fastest convergence
- Withdrawal phase depends heavily on network size and network topology (dense or sparse)
- Damping algorithms have little effect on advertisement phase, but play an important role during the withdrawal phase

# Effect of inter-arrival time: convergence time

- No visible trend (monotonic increase or decrease) in the relationship between inter-arrival time and convergence time:

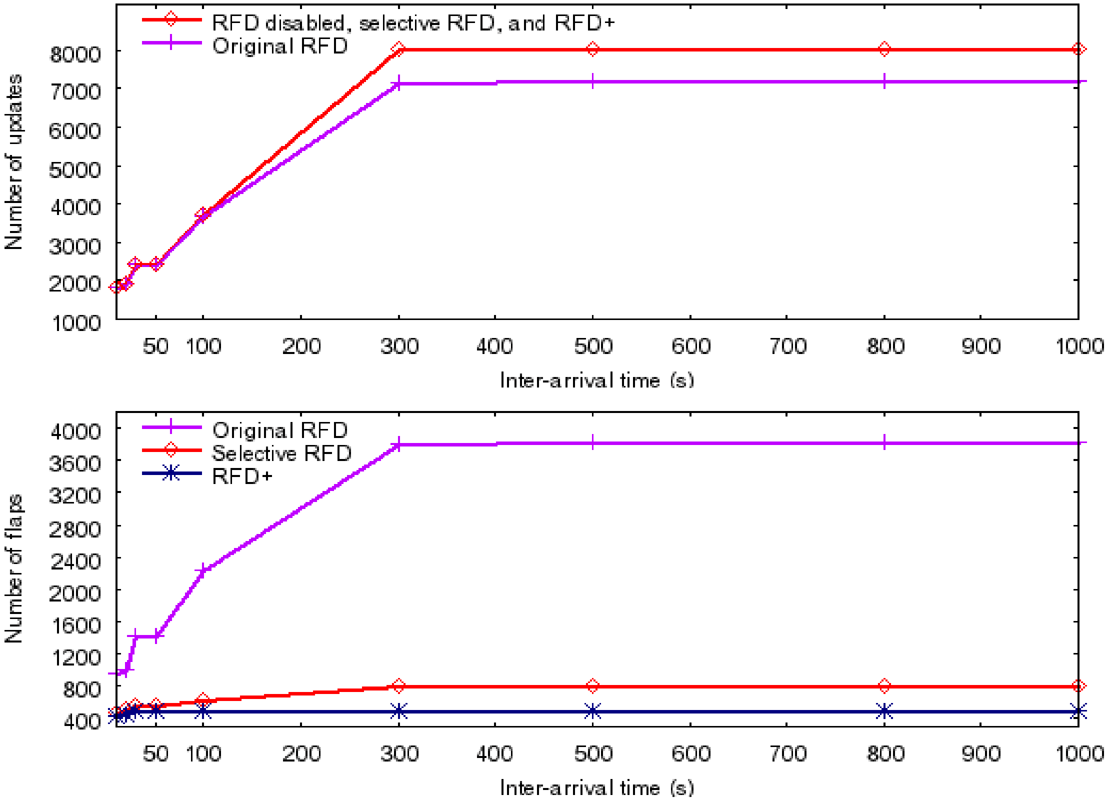


200-node network

# Effect of inter-arrival time: number of updates/flaps

- Number of updates and number of flaps tend to grow as inter-arrival time increases:

200-node network



# Effect of inter-arrival time: summary

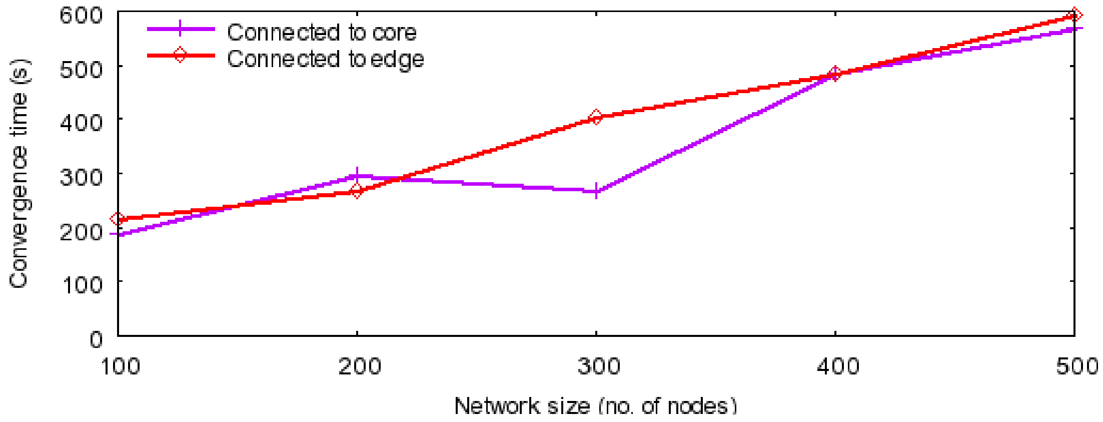
- Convergence time and number of updates: not affected by increase of inter-arrival time beyond a certain threshold
- Convergence time: affected by the difference between the instances when an update is ready to be sent and when the MRAI timer expires
- Convergence time and number of updates: not affected by damping algorithms when inter-arrival time is very short
- Number of updates and number of flaps: no decrease as inter-arrival time increases
- Number of flaps and number of route suppressions depend on variations in inter-arrival time :
  - RFD+: least sensitive
  - original RFD: most sensitive

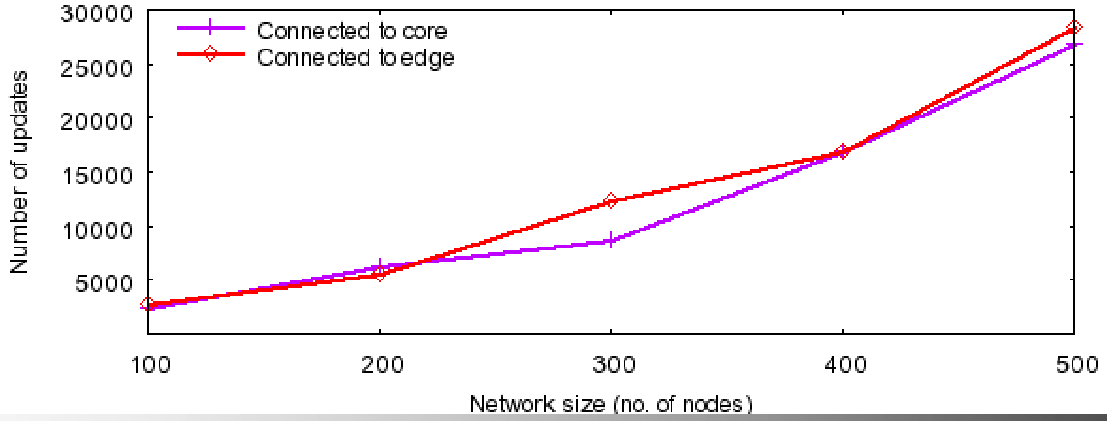
# Location of the origin router

- Location:
  - core of the network
  - edge of the network:
    - often takes up to ~ 20% longer for BGP to converge
    - usually increases the number of updates by up to ~ 25%
- Effect on BGP performance depends on:
  - network topology
  - phase: advertisement or withdrawal
  - damping algorithm

# Location of the origin router: convergence time/number of updates

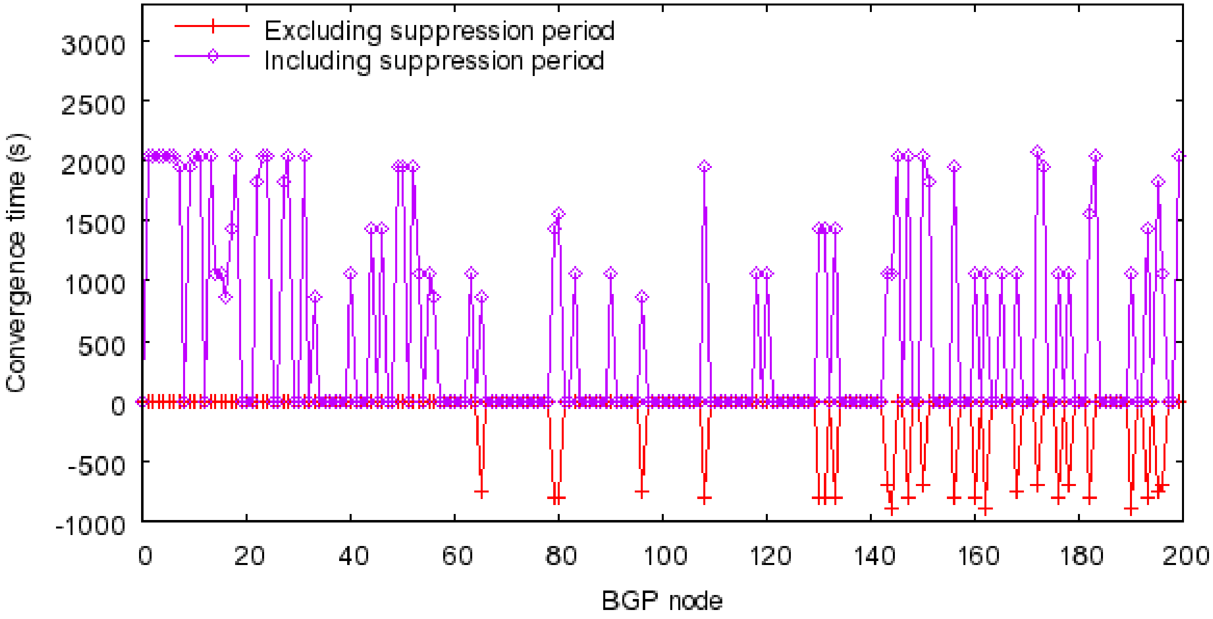
- Original RFD: withdrawal phase





# Impact of route suppression: occasional flaps

- Original RFD: one flap may cause many nodes to suffer from a significant delay in convergence (up to ~ 1 h)

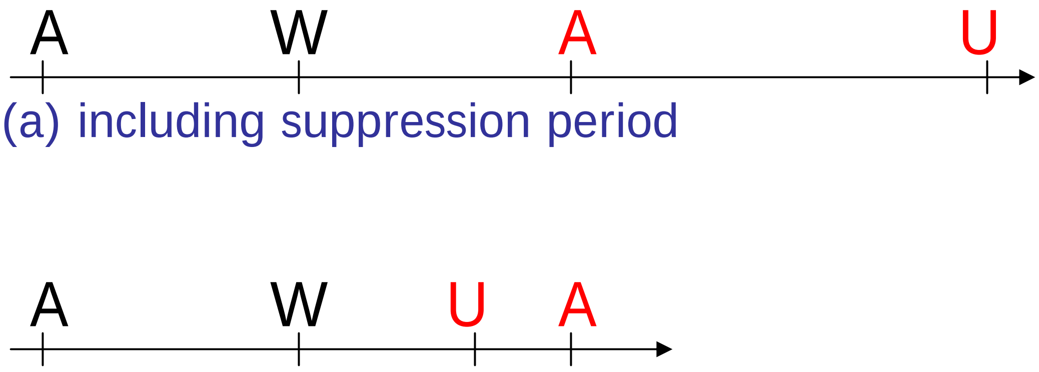


200-node network

Negative values imply that the nodes do not receive the route re-advertisement after withdrawal and will wait until other nodes become reused and start to advertise

# Negative convergence time?

- Convergence time: time difference between the re-advertisement (second A) and the last update (U)
  - U occurs after A when including suppression period (a)
  - U occurs before A when excluding suppression period (b)



(a) including suppression period

(b) excluding suppression period

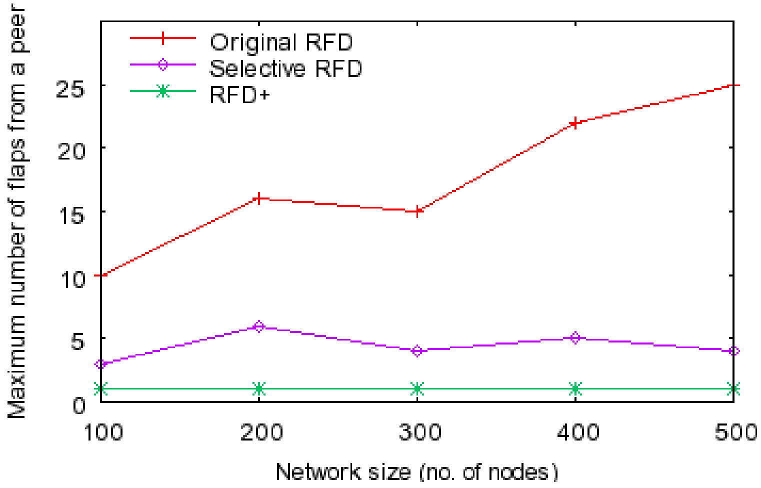
A: Advertisement

W: Withdrawal

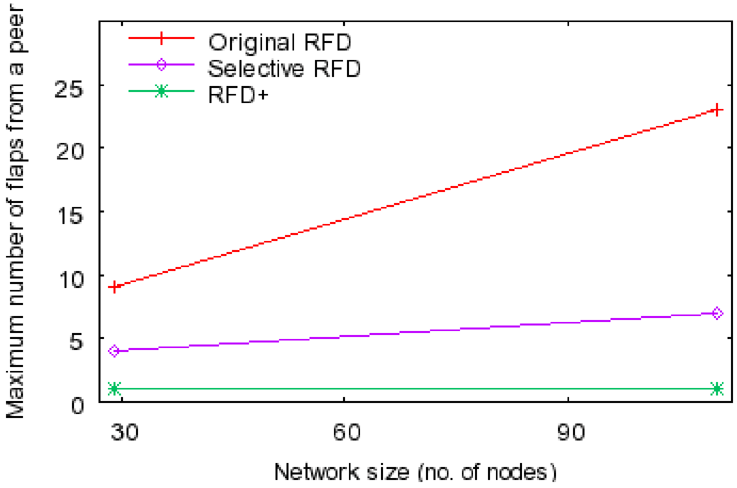
U: Last update

# Identified flaps: occasional flaps

- Selective RFD performs better than original RFD in terms of the number of flaps and suppressions
- RFD+ has the best behavior because it does not misinterpret path explorations as route flaps



BRITE-generated topologies



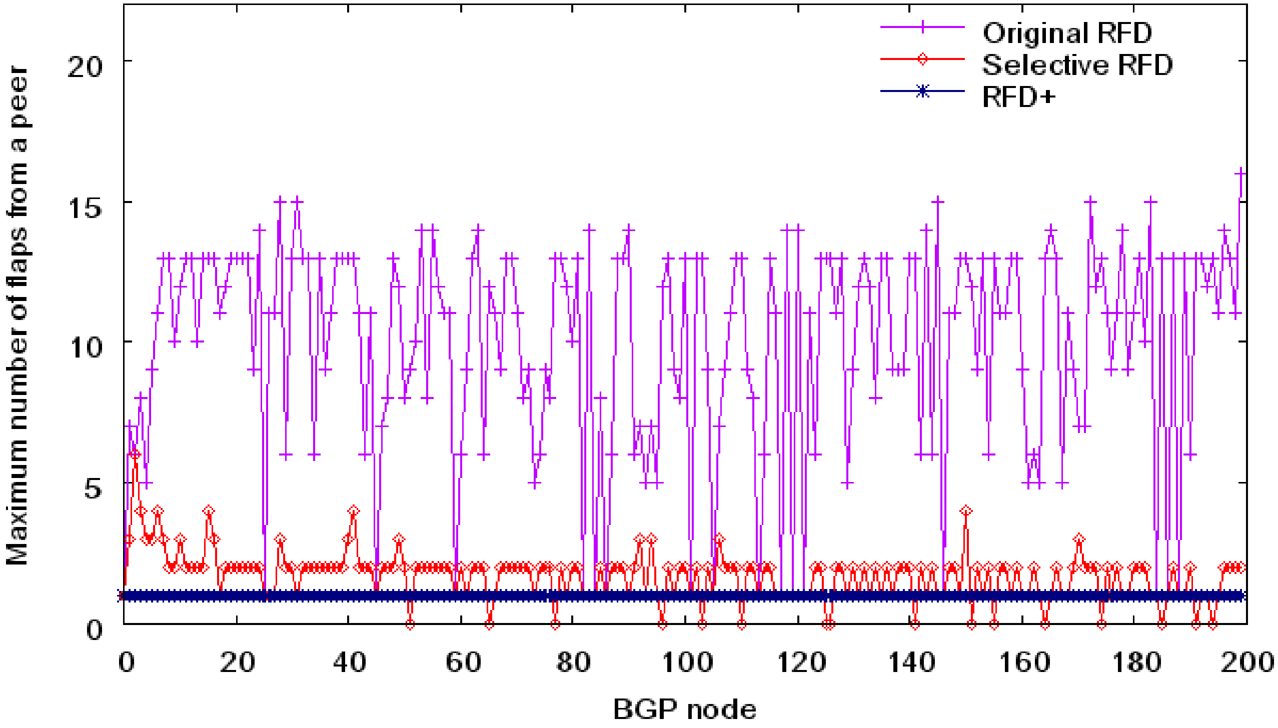
Topologies built from routing tables

# Identified flaps: occasional flaps

Original RFD: 16 (max)

Selective RFD: 6 (max)

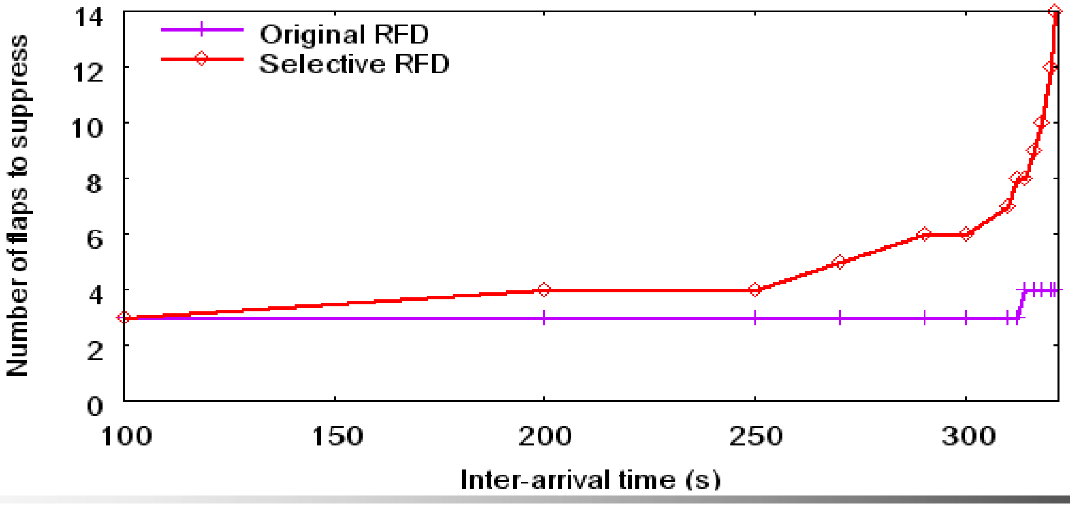
RFD+: 1 (max)



200-node network

# Persistent flaps: original RFD and selective RFD

- Original RFD prevents the spread of routing oscillations as early as possible:
  - 4 flaps are usually sufficient to suppress the flapping route
- Selective RFD may require additional flaps in order to suppress a flapping route:
  - the number depends on inter-arrival time between updates



Based on default Cisco settings

# Calculation of flaps required to suppress a route

1. The initial value of penalty is set to 0
2. When a flap is identified, it incurs a penalty of 1,000 (for withdrawals) or 500 (for route attribute changes) *
3. The penalty decays exponentially over time:
   $penalty(t2) = penalty(t1) * e^{(- \ln 2 / half\_life * (t2-t1))}$ $(t2 > t1)$
4. When penalty is larger than 2,000, the route is suppressed

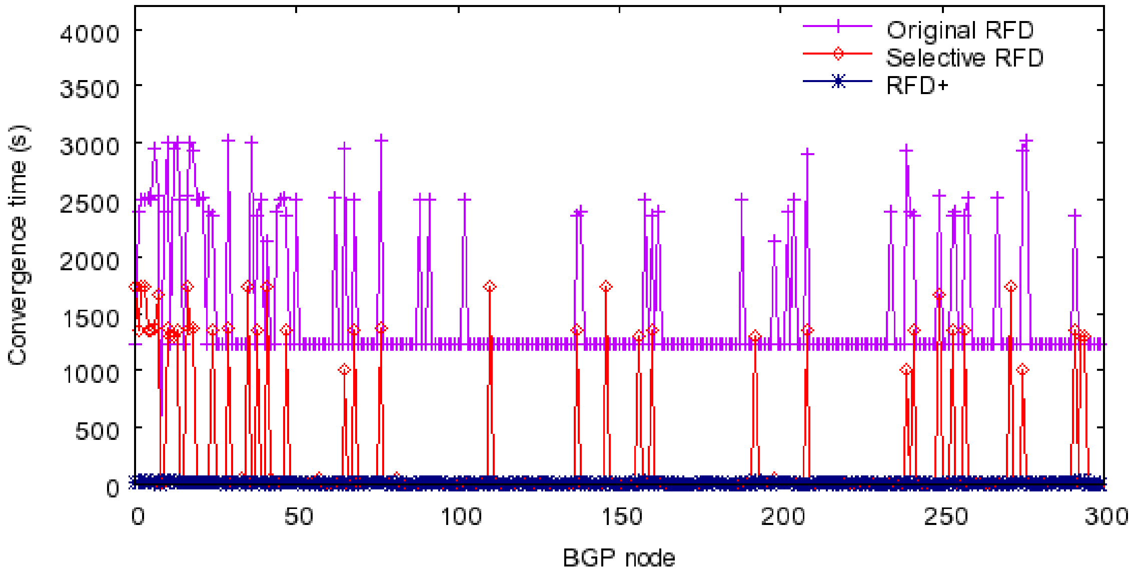
* Selective RFD postpones treating a withdrawal as a flap and only remembers the temporary withdrawal penalty. Nevertheless, the temporary withdrawal penalty decays exponentially over time. If the next update is identified as a flap, the decayed withdrawal penalty will be added to the overall penalty of the route.

# Persistent flaps: RFD+

- RFD+ underestimates the number of genuine flaps, causing a delay in route suppression:
  - reports floor((N+1)/2) flaps if the origin router experiences failure and then recovery for N times
  - treats a series of 5 updates (A, W, A, W, A) as 1 flap, rather than 2 flaps:
    - the last 2 updates are not sufficient to cause an additional flap
- RFD+ may have a large memory consumption:
  - remembers all interim routes for each prefix from each peer

# Persistent flaps

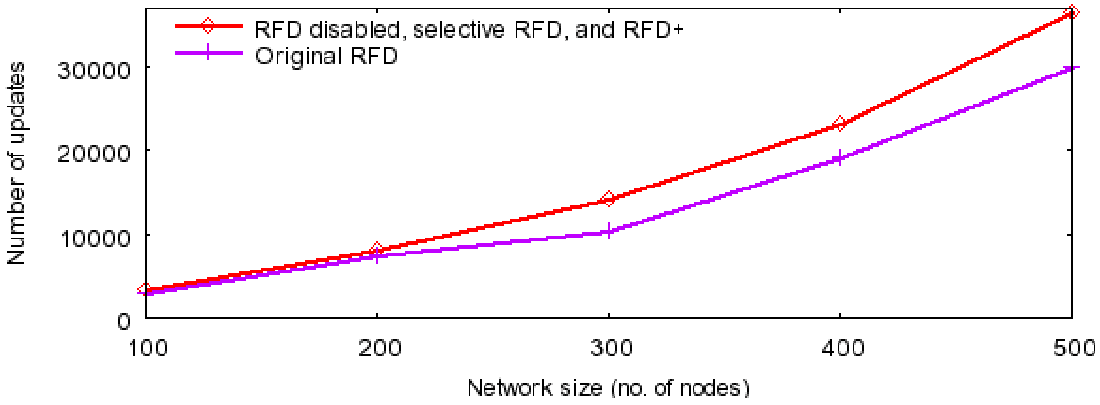
- Selective RFD and RFD+ are less aggressive than original RFD in suppressing persistently flapping routes:
  - may cause a higher number of updates and a heavier processing load on BGP speakers (undesirable)



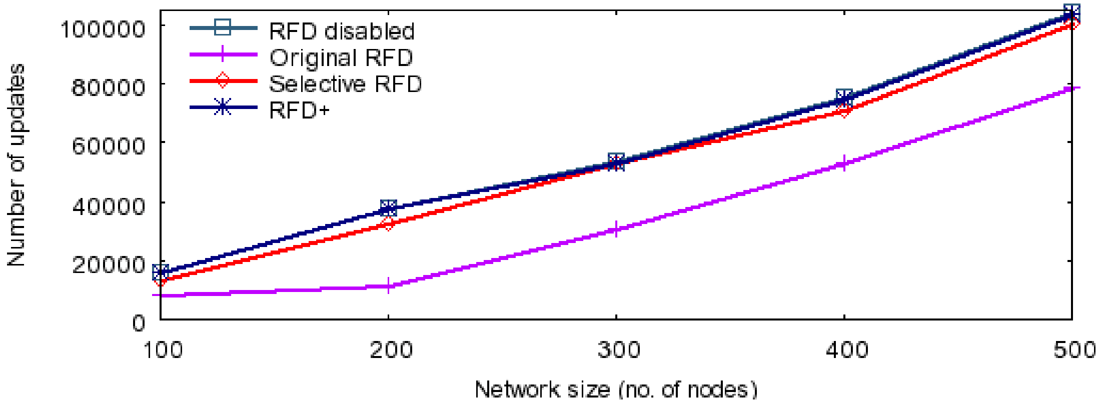
300-node network

# Number of updates

- Original RFD: optimal in reducing the number of updates



Occasional flaps



Persistent flaps

# Roadmap

- Introduction to Route Flap Damping (RFD)
- ns-2 implementation of RFD algorithms
- RFD simulation scenarios
- Performance analysis of RFD algorithms
- Improvements to RFD algorithms
- Conclusions and references

# Improving RFD+

- Aims to remedy RFD+'s problem of underestimating genuine route flaps
- Simple modification to RFD+ (modified RFD+):
  - keeps track of the "up-down-up" state of a route: advertise, withdraw, and re-advertise
  - reports a flap either when identified by RFD+ or when a route is advertised, withdrawn, and advertised again:
    - identifies all N flaps if origin router fails and then recovers for N consecutive times
    - in rare cases, may report additional flaps

# Modified RFD+

- Algorithm:

```
when receiving an advertised route r with prefix d from peer j
if (p is a withdrawn route and r == f)
    a flap is identified and clear R(d, j)      //up-down-up
elseif (r $\notin$ R(d, j) )
    insert r into R(d, j)
elseif (r $\in$ R(d, j) and rp(r) > rp(p)*)
    a flap is identified and clear R(d, j)
f = p;   p = r
```

R(d, j): set of all routes with prefix d advertised from peer j
rp(x):  route preference of a route x
p: route previous to current route r, with prefix d advertised from peer j
f: route previous to route p, with prefix d advertised from peer j

* Route preference for a withdrawal is considered to be the lowest

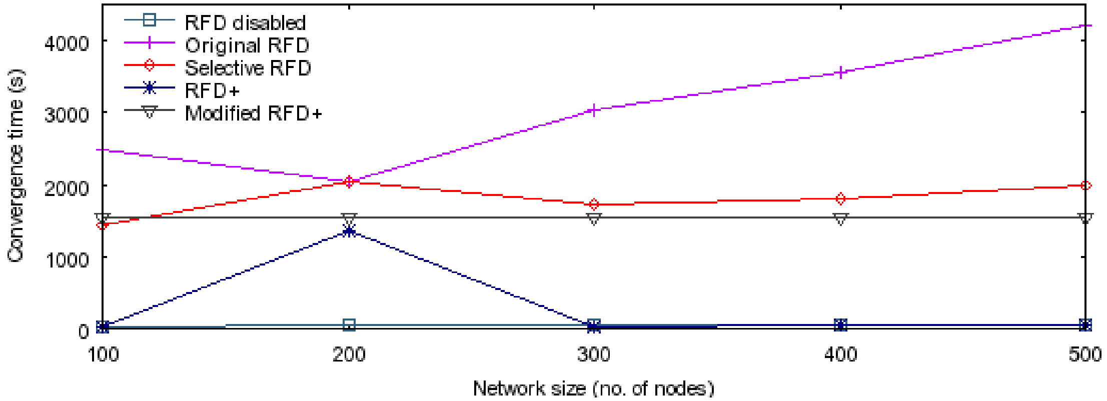
# Modified RFD+

- Advantage:
  - identifies genuine flaps better than other RFD algorithms in cases of occasional and persistent flaps:
    - does not significantly increase the BGP convergence time in the case of occasional flaps
    - identifies and suppresses persistent flaps
- Disadvantage:
  - not ideal in reducing the number of updates
- Modified RFD+ should be used when the main concern is to properly identify route flaps

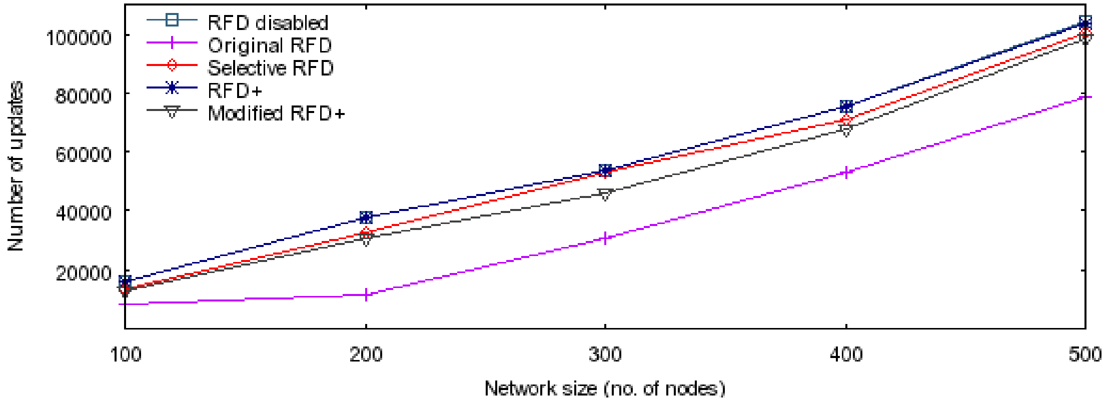
# Modified RFD+: persistent flaps

- Modified RFD+ suppresses persistent flaps and reduces updates:

Convergence time



Number of updates



# Combined RFD: adaptive approach

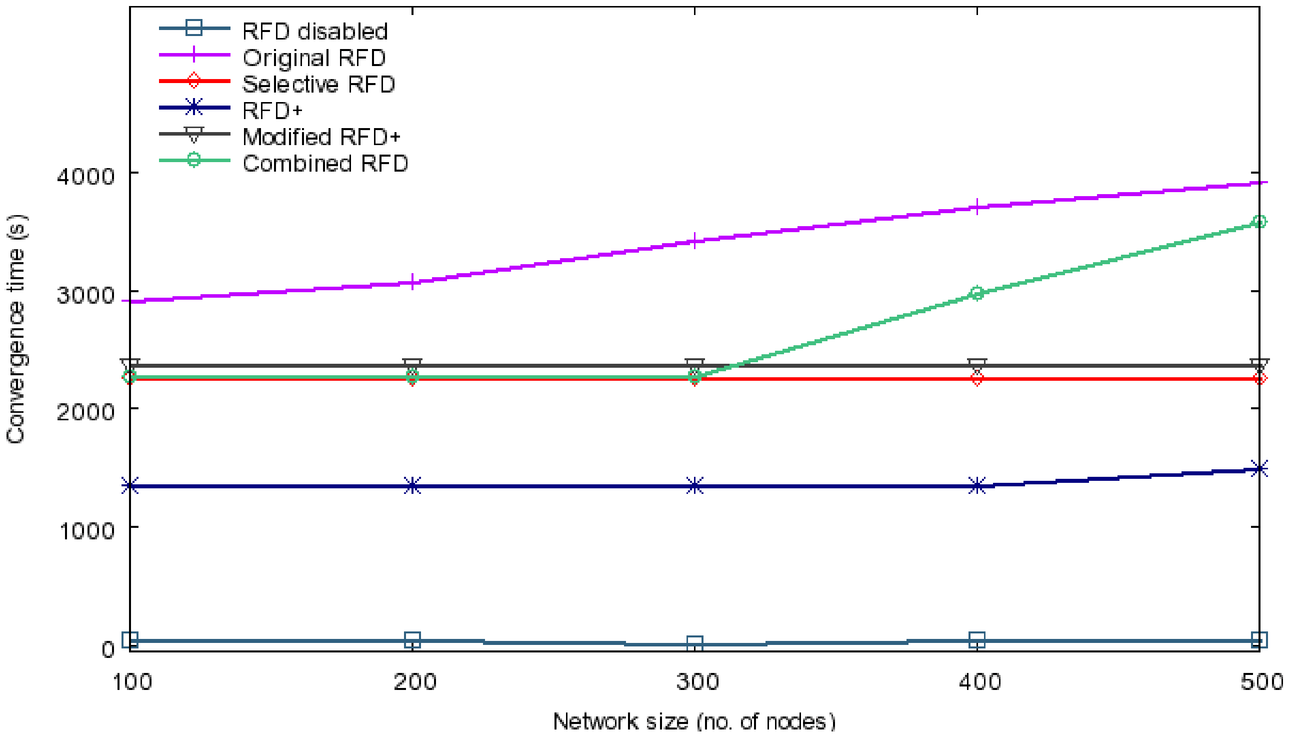
- Combined RFD:
  - integrates original RFD and modified RFD+
  - uses modified RFD+ for first two flaps because modified RFD+ behaves well in the case of occasional flaps
  - uses original RFD starting from the third flap because original RFD is efficient in suppressing persistent flaps and in reducing the number of updates

# Combined RFD

- Advantages:
  - does not suppress a route that flaps only once or twice
  - efficiently suppresses persistently flapping routes
  - tends to generate fewer updates than selective RFD, RFD+, and modified RFD+
- Disadvantage:
  - does not always accurately identify genuine route flaps
- Combined RFD should be used when the main concern is to keep the number of update messages close to optimal (original RFD)

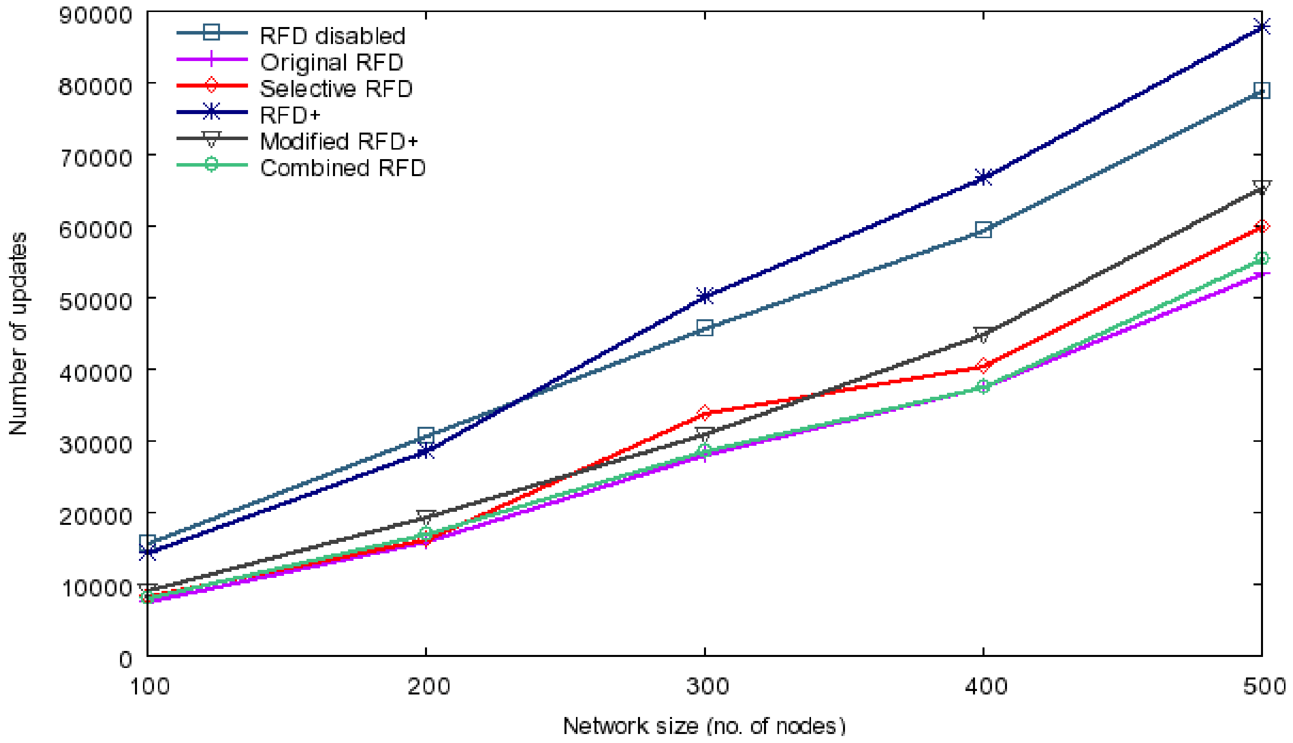
# Comparison of convergence time: 8 flaps

- Convergence time of Combined RFD lies between selective RFD and original RFD and depends on network topology



# Comparison of number of updates: 8 flaps

- Combined RFD is the second best in reducing update messages (close to original RFD) in the case of persistent flaps



# Security and RFD

- Route flap damping may be exploited by malicious attackers to cause long AS-to-AS or AS-to-prefix isolation:
  - malicious user may attack the underlying TCP connection and reset the BGP connection
  - withdrawals followed by advertisements of entire routing tables are sent between peers
  - route flap damping amplifies the adverse effects caused by BGP session attacks, causing a potential denial of service (DOS) attack

K. Sriram, D. Montgomery, O. Borchert, O. Kim, and R. Kuhn, “Autonomous System (AS) Isolation under Randomized BGP Session Attacks with RFD Exploitation,” private communication.

# Conclusions

- Implemented five RFD algorithms in ns-BGP: original RFD, selective RFD, RFD+, modified RFD+, and combined RFD
- Compared three existing RFD algorithms using reasonably large and realistic network topologies:
  - no algorithm performs optimally in all circumstances:
    - Original RFD behaves favorably in the case of persistent flaps
    - RFD+ performs well in the case of occasional flaps

# Conclusions

- Proposed two improvements:
  - modified RFD+ (modification to RFD+):
    - identifies genuine flaps better than other RFD algorithms in cases of occasional and persistent flaps
  - combined RFD (adaptive approach):
    - efficiently suppresses routes that flap persistently and reduces update messages
    - does not suppress routes that flap only once or twice
- Future RFD implementations may incorporate algorithms that deal with unfair suppression, slow convergence, and low level of security

# References

- C. Villamizar, R. Chandra, and R. Govindan, “BGP route flap damping,” *IETF RFC 2439*, Nov. 1998.
- Z. Mao, R. Govindan, G. Varghese, and R. Katz, “Route flap damping exacerbates Internet routing convergence,” in *Proc. SIGCOMM 2002*, Pittsburgh, PA, Aug. 2002, pp. 221–233.
- Z. Duan, J. Chandrashekar, J. Krasky, K. Xu, and Z. Zhang, “Damping BGP route flaps,” in *Proc. IPCCC 2004*, Phoenix, AZ, Apr. 2004, pp. 131–138.
- ns-2 [Online]. Available: http://www.isi.edu/nsnam/ns.
- ns-BGP [Online]. Available: http://www.ensc.sfu.ca/~ljilja/cnl/projects/BGP.
- SSFNet [Online]. Available: http://www.ssfnet.org/.
- RFD-AMRAI BGP [Online]. Available: http://www.ensc.sfu.ca/~ljilja/cnl/projects/RFD-AMRAI.
- BRITE [Online]. Available: http://www.cs.bu.edu/brite.
- Multi-AS topologies from routing tables [Online]. Available: http://www.ssfnet.org/ Exchange/gallery/asgraph.
- Y. Rekhter and T. Li, “A border gateway protocol 4 (BGP-4),” *RFC 1771*, Mar. 1995.
- S. Halabi and D. McPherson, *Internet Routing Architectures*. Indianapolis, IN: Cisco Press, 2000.
- C. Huitema, *Routing in the Internet*. Upper Saddle River, NJ: Prentice Hall, 2000.

# References

- T. Bu and D. Towsley, “On distinguishing between Internet power law topology generators,” in *Proc. INFOCOM 2002*, New York, NY, June 2002, pp. 638–647.
- T. G. Griffin and B. J. Premore, “An experimental analysis of BGP convergence time,” in *Proc. ICNP 2001*, Riverside, CA, Nov. 2001, pp. 53–61.
- T. D. Feng, R. Ballantyne, and Lj. Trajkovic, “Implementation of BGP in a network simulator,” in *Proc. ATS ’04*, Arlington, VA, Apr. 2004, pp. 149–154.
- W. Shen and Lj. Trajkovic, “BGP route flap damping algorithms,” in *Proc. SPECTS ’05*, Philadelphia, PA, July 2005, pp. 488–495.