Stabilizing BGP Routing without Harming Convergence

Xiaoqiang Wang[1], Olivier Bonaventure[2], Peidong Zhu[1]

[1] National University of Defense Technology, China

[2] University Catholique de Louvain, Belgium
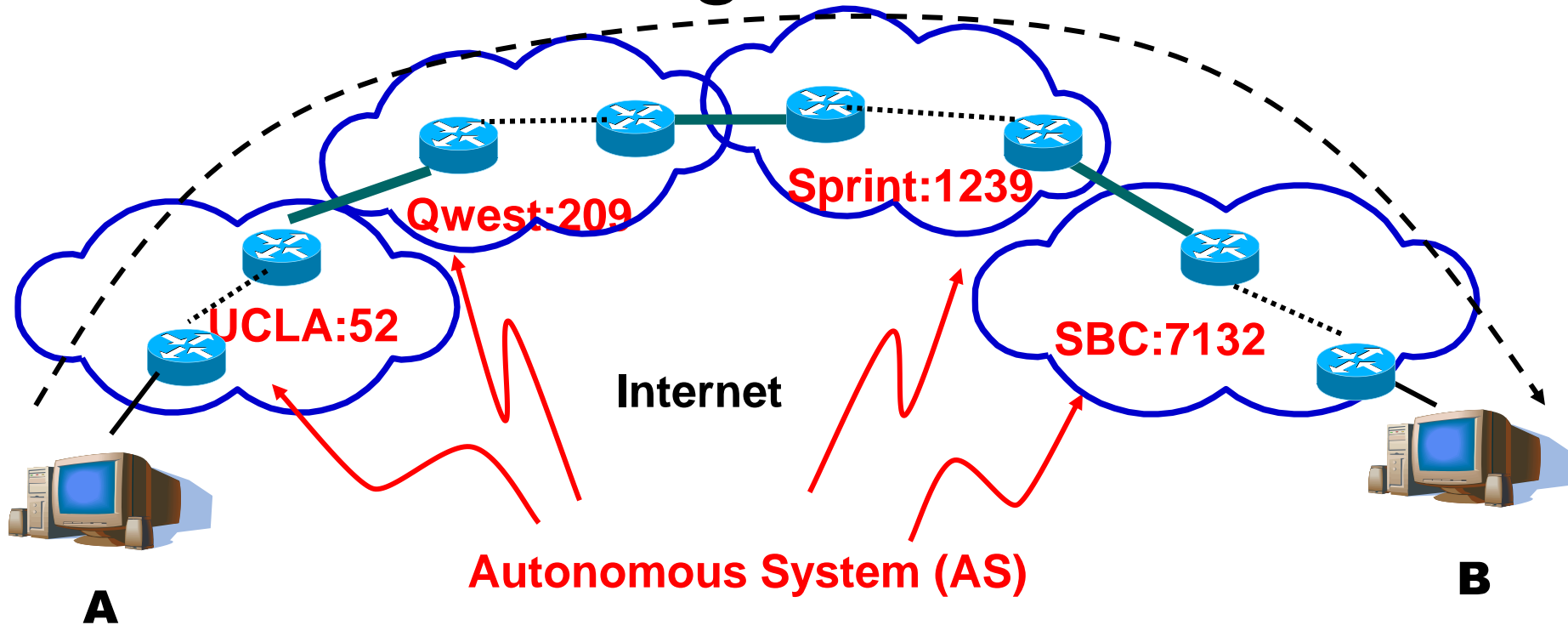
Outline

- Background

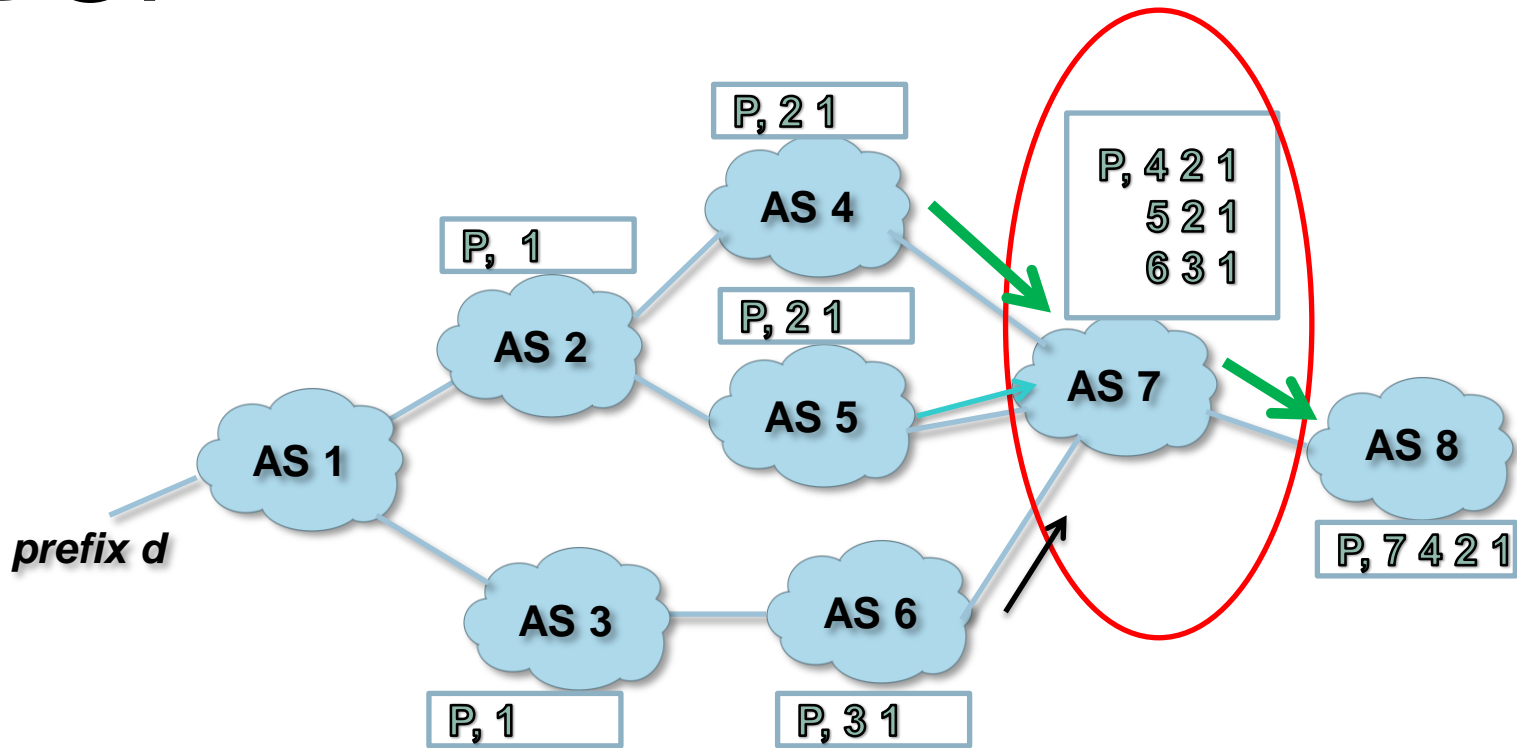- Motivation

- Methodology

- Evaluation

- Conclusion

Outline

- Background
- Motivation
- Methodology
- Evaluation
- Conclusion

Internet Routing

Qwest:209
Sprint:1239
UCLA:52
SBC:7132
Internet
A
B
Autonomous System (AS)

- Intra-AS: OSPF, IS-IS, RIP
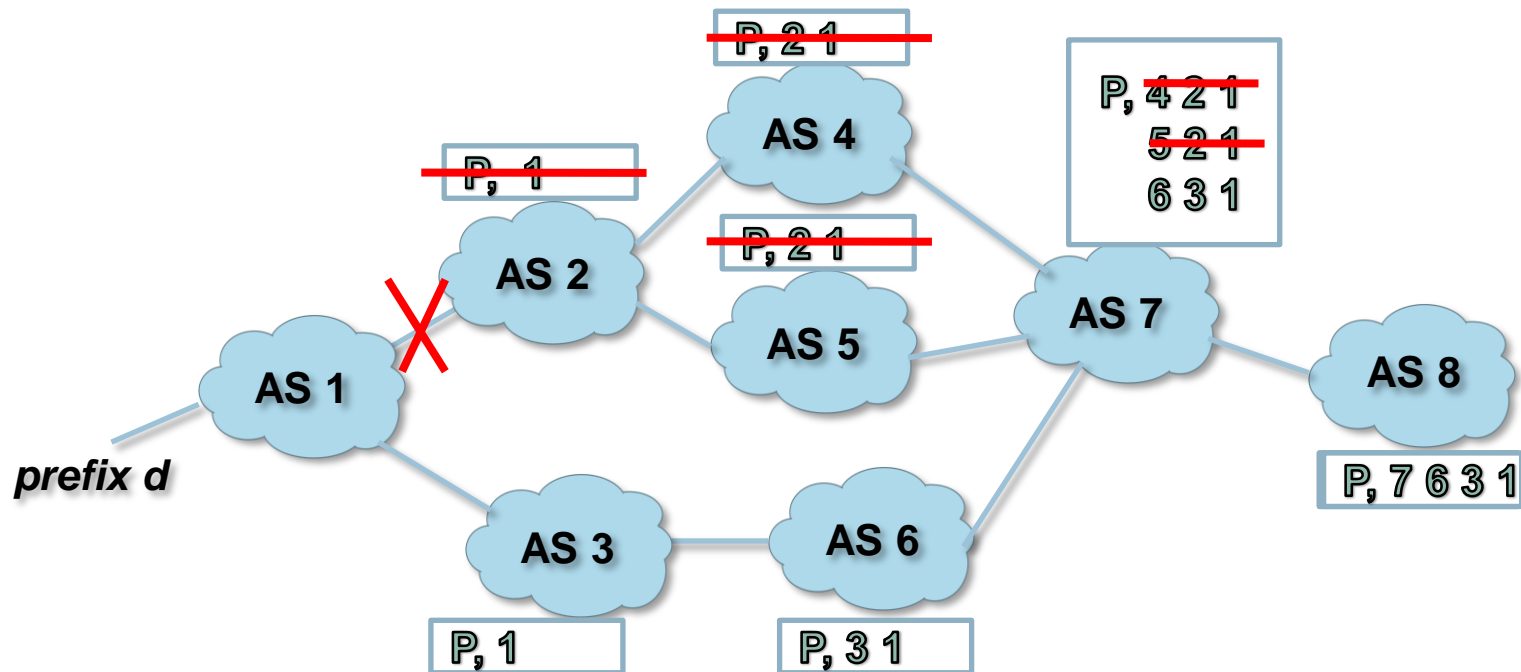- Inter-AS: BGP (Border Gateway Protocol)

BGP

- Prefix specific
- Path Vector Routing Protocol
- One-fits-all-model

BGP Churn

- Large volume of BGP updates
- Bad for routers
  - Overload CPU,memory, frequent FIB changes
- Major causes
  - BGP path exploration
  - Route flapping

BGP Path Exploration

- Single event triggers several updates

Route Flapping

- Routes periodically change

- Reasons are diverse

- Mice-elephant

  - a significant portion of churn is associated to a small set of highly active prefixes [Rexford 02, Oliveira05]

  - 3% prefixes → 36% updates [Pelsser PAM11]

Current countermeasures

- Path exploration acceleration
  - BGP-RCN, EPIC
  - Not deployed yet
- Suppress excessive BGP updates
  - Route flap Damping, MRAI
  - Only two built-in mechanisms in real router
  - Dying for breaking/delaying convergence
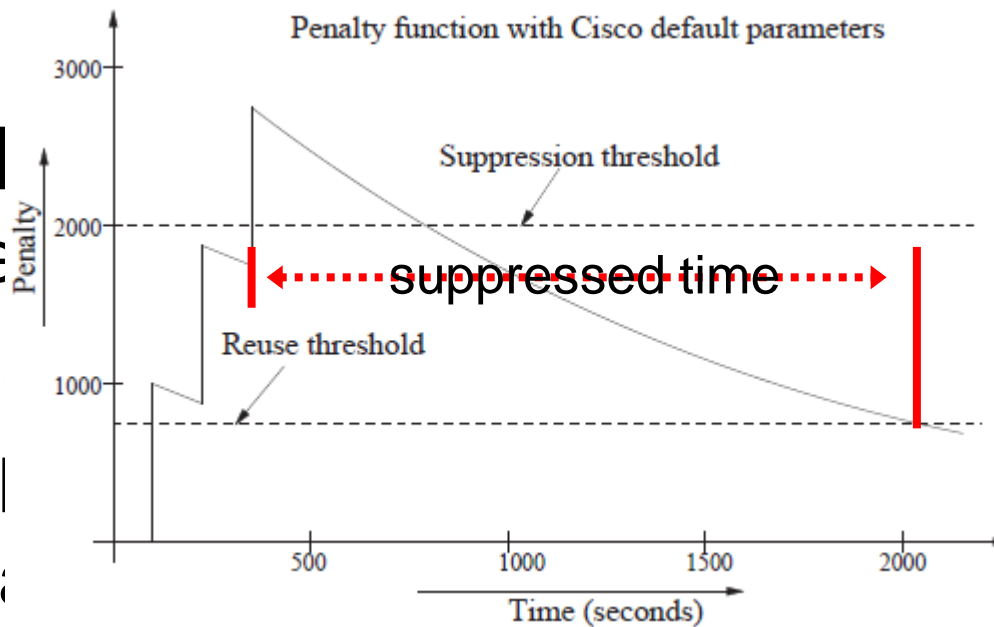
Route Flap Damping

Penalty function with Cisco default parameters

- Princi[...]
  - □ A pe[...]
  - □ Upd[...]update
  - □ Sup[...]alty > *T*
  - □ Pen[...]time
- Triggered by 3 flaps under cisco parameter[Mao02, Randy02]
- Interactions between RFD and BGP path exploration
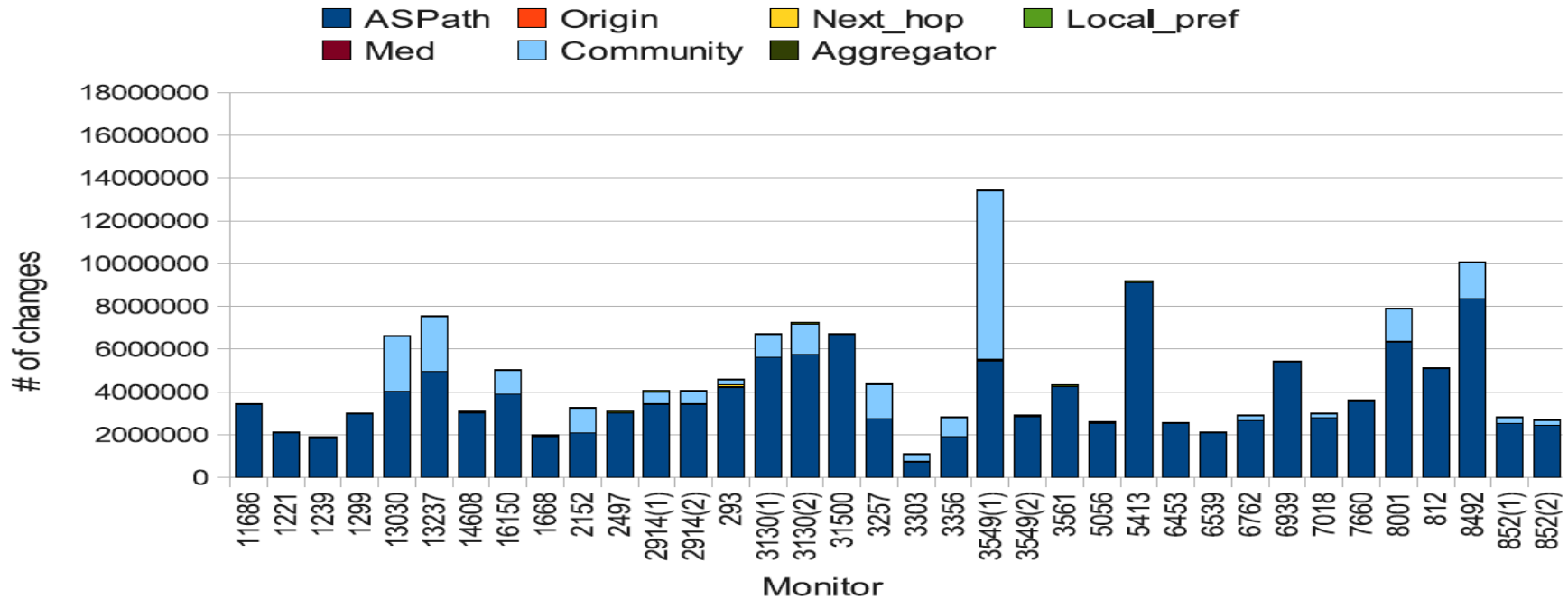
MRAI

- Minimum Route Advertisement Interval
- Supposed to be per peer and per prefix
- Rate-limit BGP updates
  - Two consecutive announcements are spaced at least a MRAI interval*jitter[0.75,1]
  - Typical setting: 30s for eBGP, 5s for iBGP
- BGP updates are heavily delayed

Outline

- Background
- Motivation
- Methodology
- Evaluation
- Conclusion

Motivation(1/2)

BGP churn mostly caused by AS_PATH and COMMUNITY changes
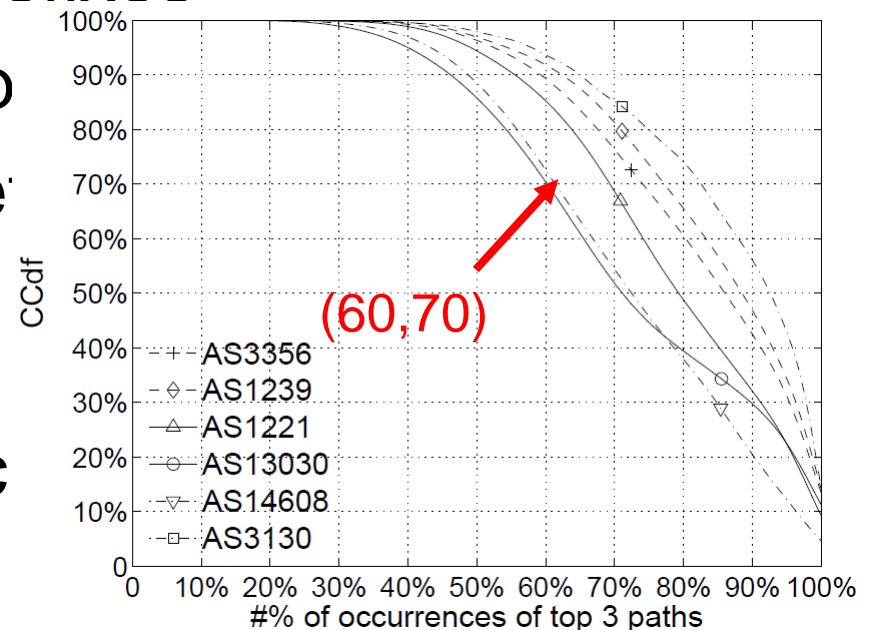
Motivation(2/2)

- ## Path Locality

  - ☐ An AS explores limit number of AS_PATHs to reach highly active prefixes

  - ☐ Same data set as in p

  - ☐ for each prefix, we de

    $$likelyhood =$$

  - ☐ Results are similar ac

(60,70)

Legend:
- AS3356
- AS1239
- AS1221
- AS13030
- AS14608
- AS3130

Y-axis: CCdf (0 to 100%)
X-axis: #% of occurrences of top 3 paths (0 to 100%)
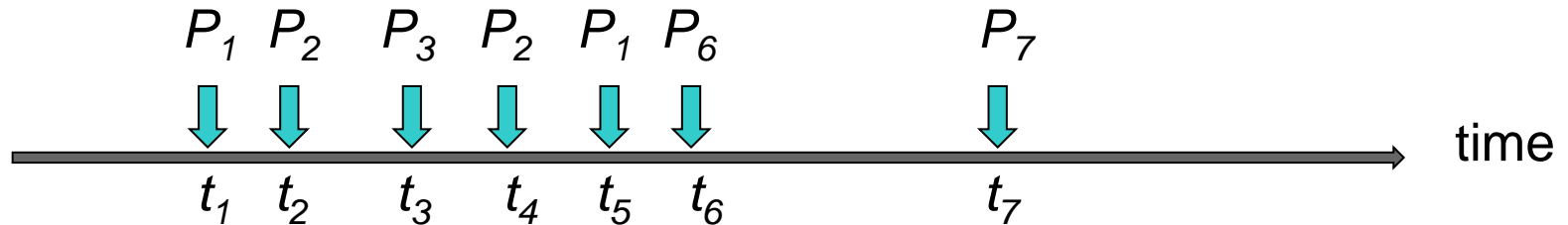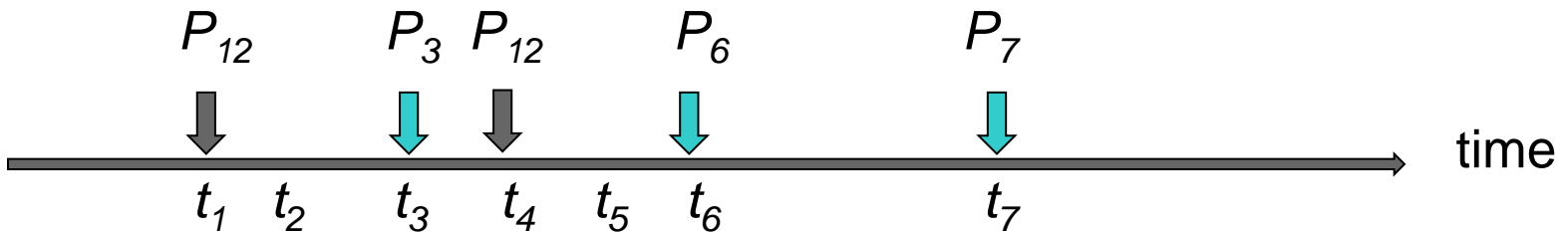
Outline

- Background

- Motivation

- Methodology

- Evaluation

- Conclusion

Basic idea

aggregate P_1 and P_2 into P_{12}

- Conclusion:
  - 2 fewer changes
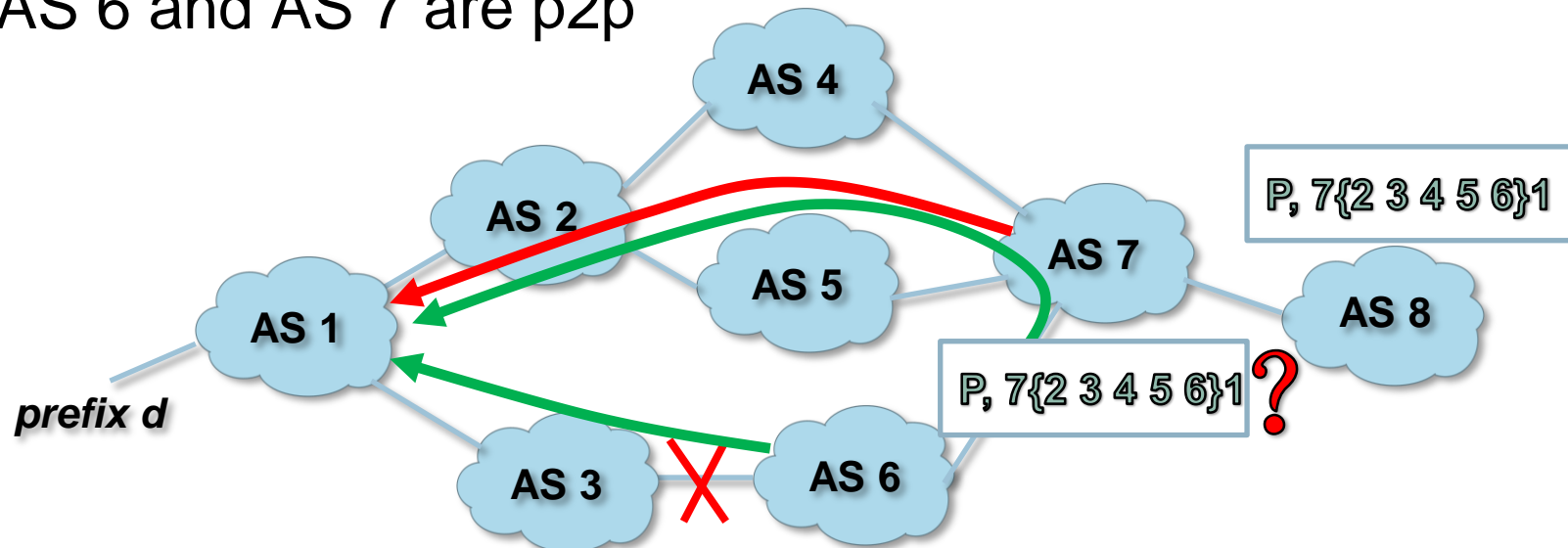 - *4* fewer changes *if P_3 is further involved*
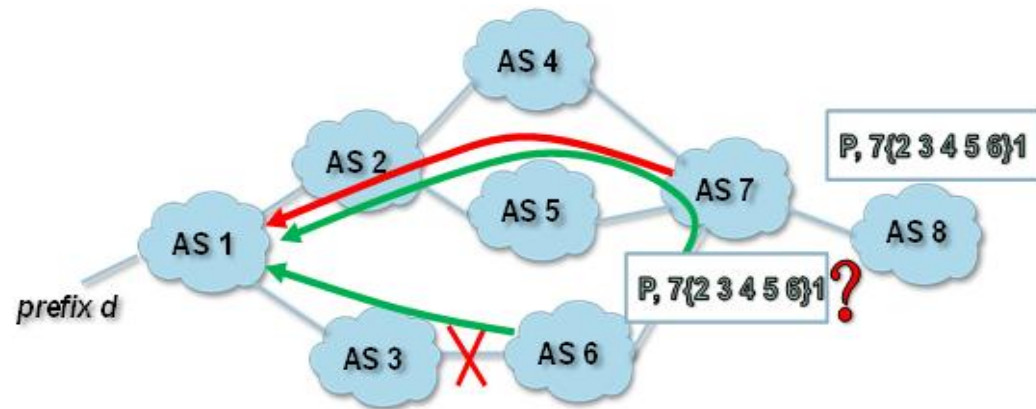
Routing issues

- ## AS_PATH functions
  - Prevent routing loops, influent BGP decision process
- ## Backup path
  - AS 6 and AS 7 are p2p

Solution

P, 7{2 3 4 5 6}1

prefix d

P, 7{2 3 4 5 6}1 ?

- Per peer and per prefix
- SSLD(Sender sider loop detection) [Labovitz02]
- Example
  - To AS 8: [7 4 2 1], [7 5 2 1], [7 6 3 1]
    - Aggregated path is 7{2 3 4 5 }1
  - To AS 6: [7 4 2 1], [7 5 2 1]
   - Aggregated path is 7{4 5}2 1

Workflow

- Upon receiving a route *r* regarding *p*
  - ☐ Update the prefix penalty associated to *p*
  - ☐ Update the path penalty associated to *r.path*
  - ☐ Update the path penalty in *p*'s history cache
  - ☐ If prefix penalty regarding *p* > *threshold*
    - AS_PATH aggregation is triggered

- Clean process is scheduled every *T* hours
  - ☐ Remove those paths whose path penalties are small enough

Outline

- Background
- Motivation
- Methodology
- Evaluation
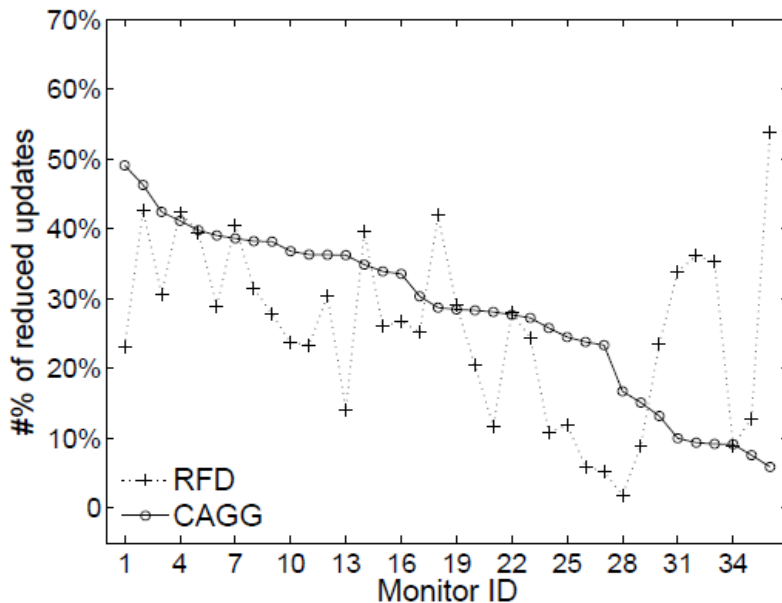- Conclusion

Evaluation(1/3)
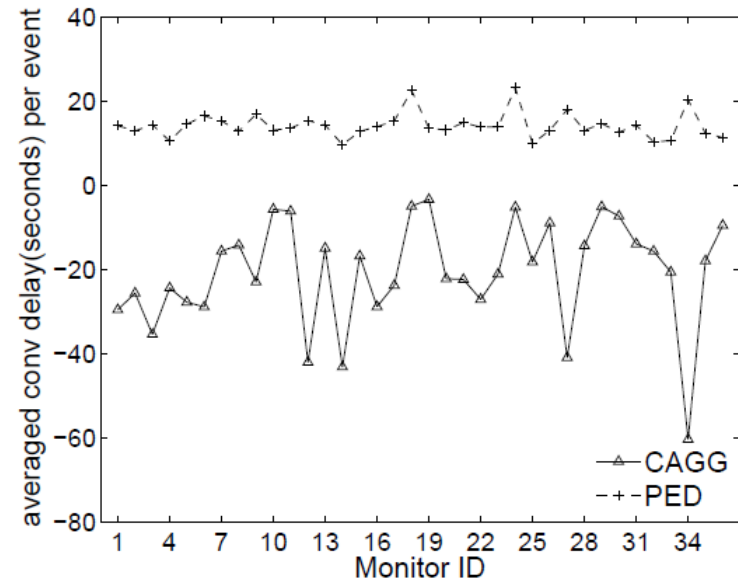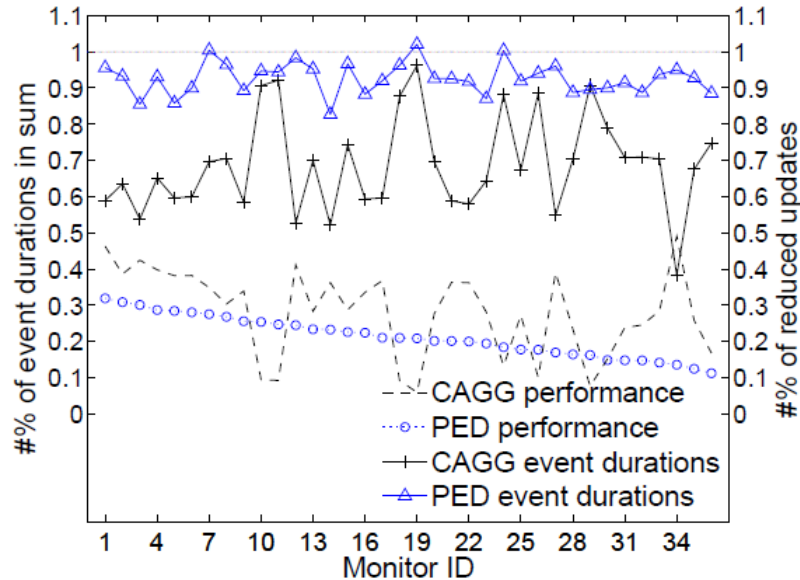
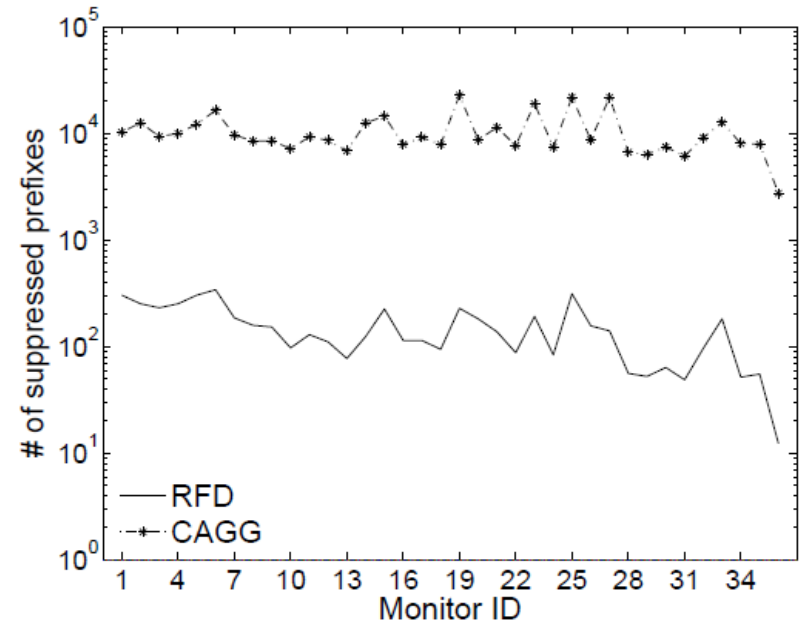(a) Event durations and performance

(b) Relative convergence delay

- Performance: better in 29/36 monitors
- Convergence: better in all monitors

Evaluation(2/3)

(a) Reduced updates

(b) # of suppressed prefixes

- Perform better in 21/36 monitors
- Suppress more prefixes

Evaluation(3/3)

- Memory cost
- AS_PATH sharing
- Only upper bound is evaluated
  - At most 5,000 more paths per router
  - Higher ASes buffer fewer AS_PATHs

Outline

- Background
- Motivation
- Methodology
- Evaluation
- Conclusion

Conclusion

- BGP churn is a problem, especially for those highly active prefixes

- To utilize path locality is a potential choice

- Next step is to extend our approach to iBGP so that AS itself can benefit from this technology as well

Thanks!
Q&A