Investigating occurrence of duplicate updates in BGP announcements

Jonathan Park, Dan Jen, Mohit Lab, Shane Amante, Danny McPherson, Lixia Zhang

GROW @ IETF75

July 27, 2009

Why This Work

- All BGP update messages should be unique
- We know that is not true in reality
- But exactly how bad is it?
  - Many papers mentioned existence of duplicate updates
  - No quantitative results
- Contributions of this work
  - Quantified the amount of duplicates
  - Looked impact of duplicates
  - First attempt to find the causes

Date Set

- From RouteViews and RIPE: data from all monitors
  - With full BGP table
  - Available for the whole month of March 2002-2009
  - The numbers of monitors from 2002 to 2009: 27, 37, 54, 67, 79, 100, 109, and 90 respectively

** The reason why we have less monitors in March 2009:

- RRC01 was down from March 20-31, 2009
- RRC13 was down from March 14-31, 2009
- RRC14 was down from March 24-31, 2009

Define BGP Duplicates

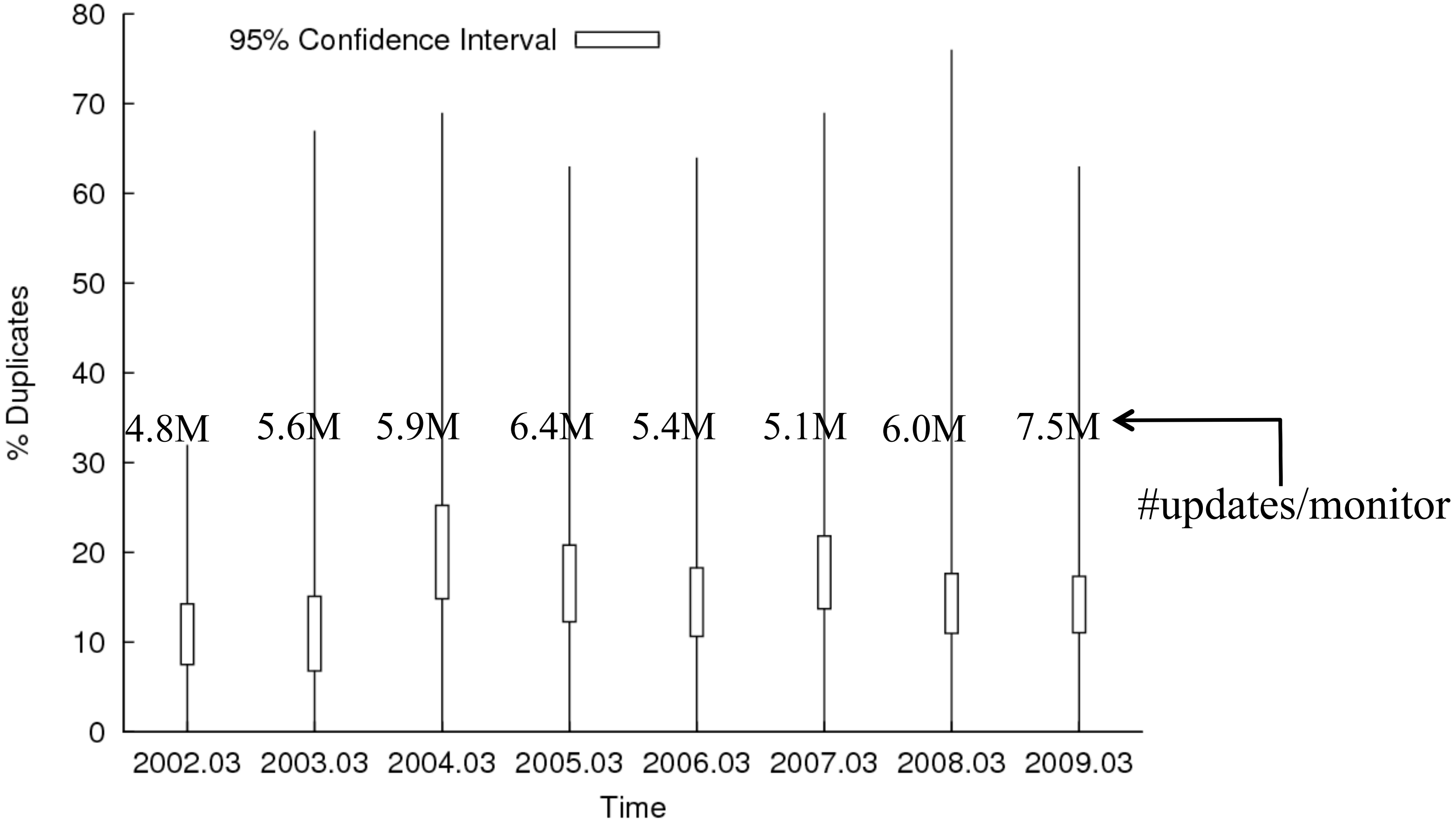
Pure adjacent identical updates

- Filtered out all updates due to session resets
- Did not count those with different attribute values

Now how many we saw for March 2009

- Total number of updates observed (90 monitors): 677 million
- Total number of duplicates: 91 million
  - About 13.5%

Looking Over Time

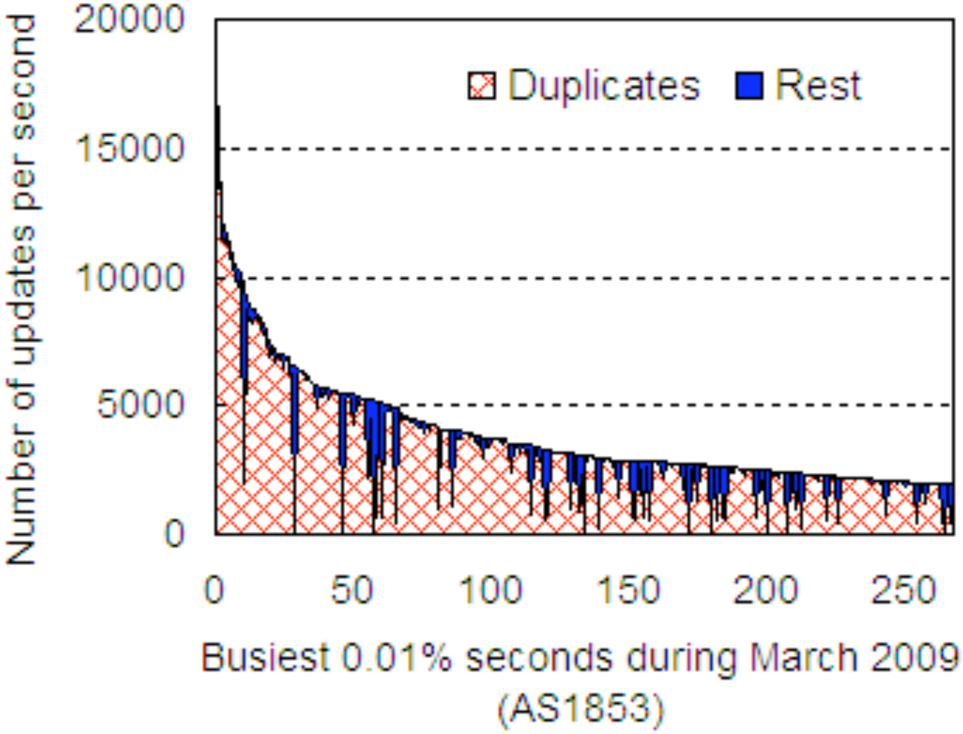


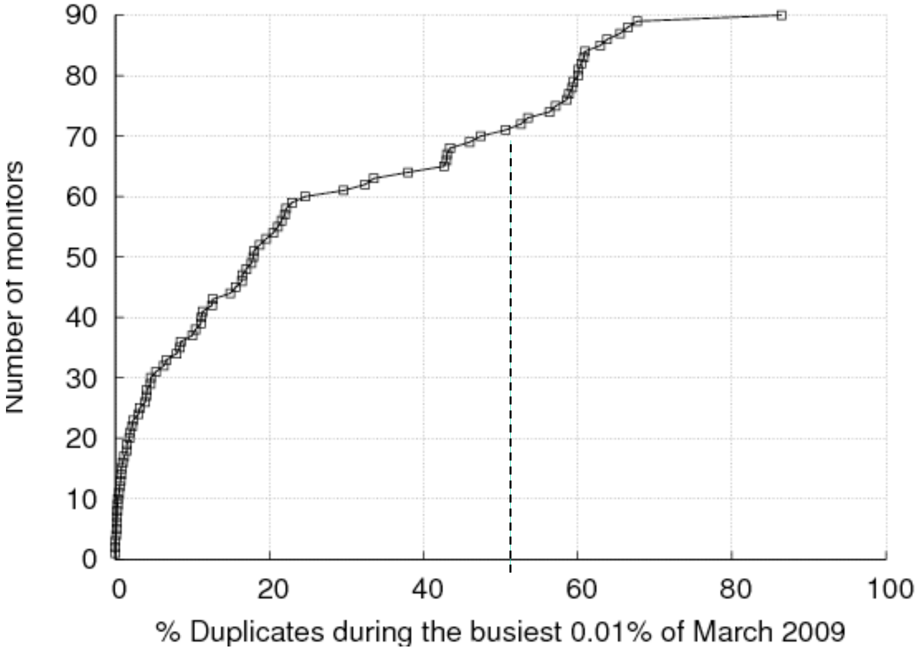
4.8M 5.6M 5.9M 6.4M 5.4M 5.1M 6.0M 7.5M

#updates/monitor

For the last 8 years, the percentage of dup. updates has not changed much

Are duplicate updates bad?



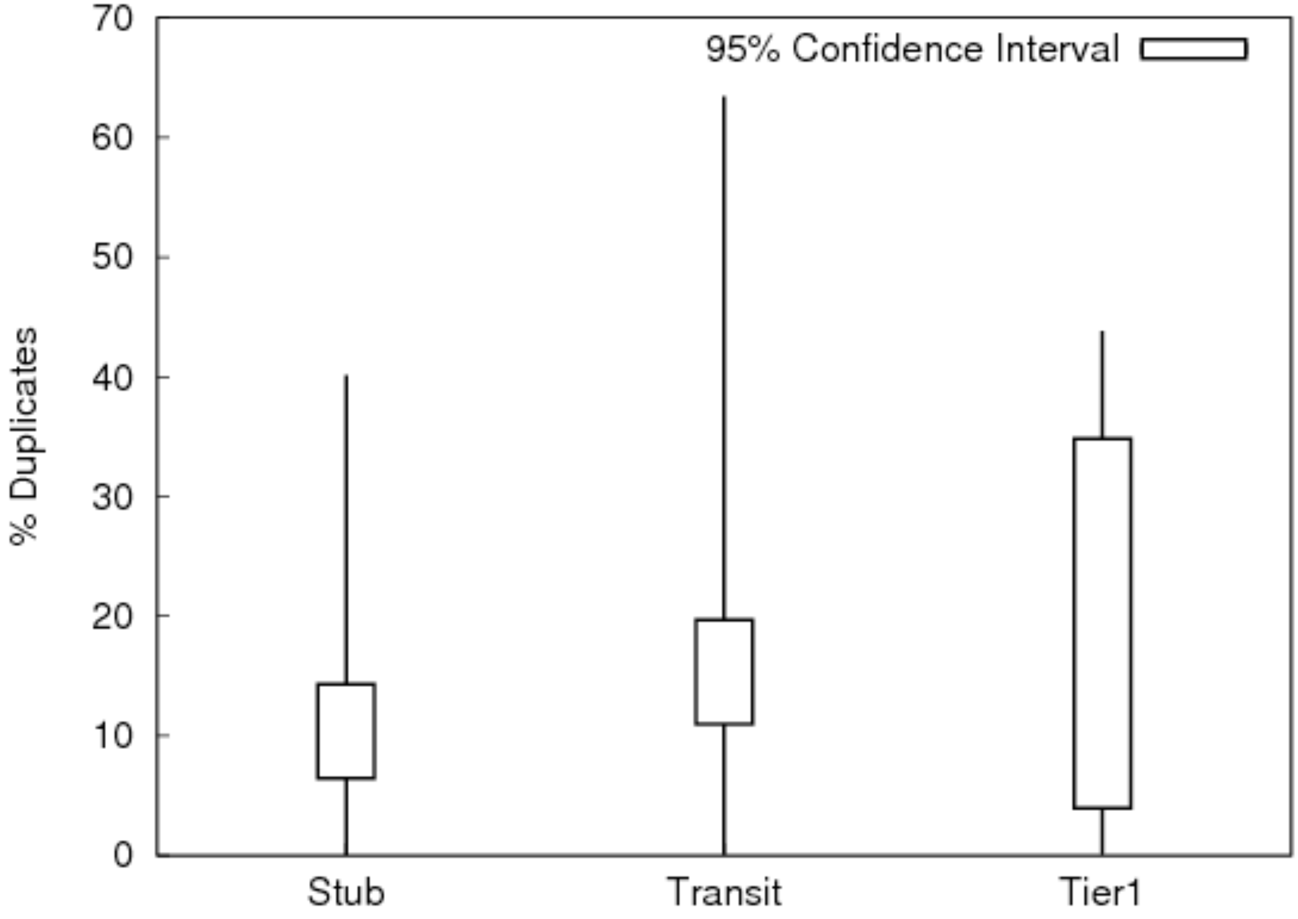


- For AS1853, 86.42% of the total updates during the busiest 0.01% sec in March 2009 are duplicates
- 20% of the monitors have more than 52.6% of total updates as duplicates during busiest 0.01% sec

Sorting Data Sources

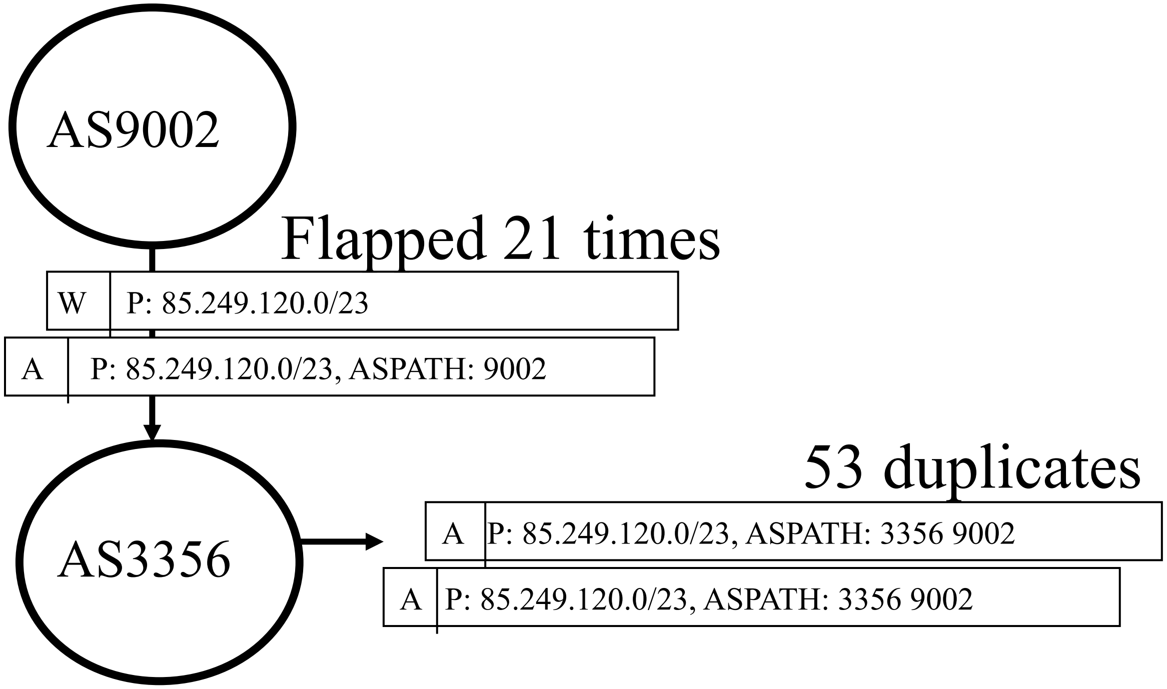
- The numbers of monitors used 90
- Classified monitors into 3 types
  - Tier-1: AS with no providers
  - Transit: Neither Tier-1s or Stubs
  - Stub: AS with less than 5 down stream ASes
- Number of monitors in Tier-1s, transit, and stub are 8, 55, 27 respectively (March 2009)
  - Tier-1s: AS7018, AS3549, AS2914, AS209, AS6453, AS701, AS3561, AS1299

Duplicates by monitor classes



Tier-1s have wider confidence Intervals due to fewer data points (8 only)

One example of duplicate update occurrence

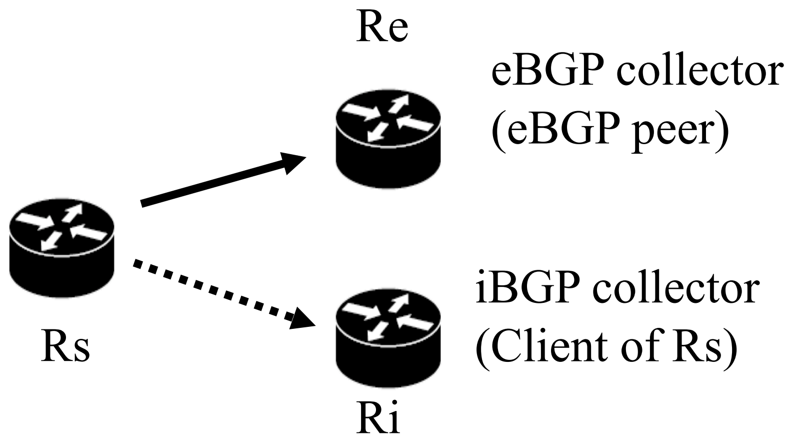

AS9002
Flapped 21 times
W
P: 85.249.120.0/23
A
P: 85.249.120.0/23, ASPATH: 9002
53 duplicates
AS3356
A
P: 85.249.120.0/23, ASPATH: 3356 9002
A
P: 85.249.120.0/23, ASPATH: 3356 9002

Observed from monitor in AS9002: A/W/A/W/… on 85.249.120.0/23

Observed from monitor in AS335: never withdraws the prefix; sent generates duplicates.

Why duplicate updates: investigation

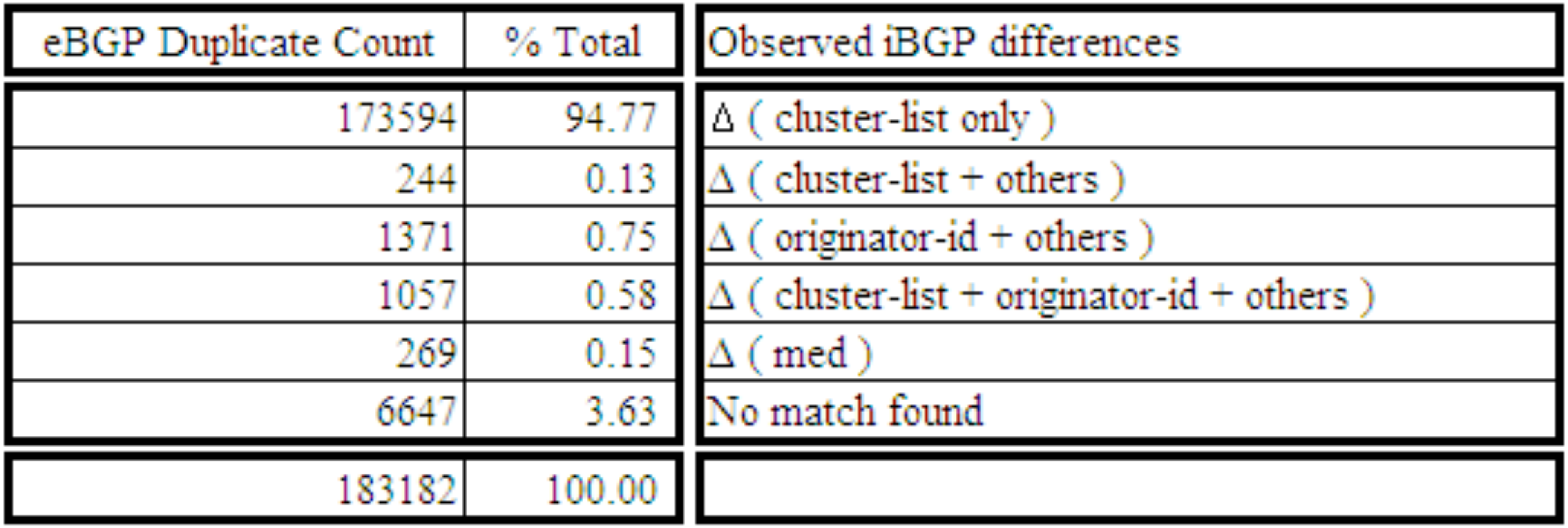
- Suspect that the duplicates are due to eBGP-iBGP interactions
- Measurement setting: one router providing both eBGP and iBGP data; MRAI timer set to 0



More Details on the Measurement

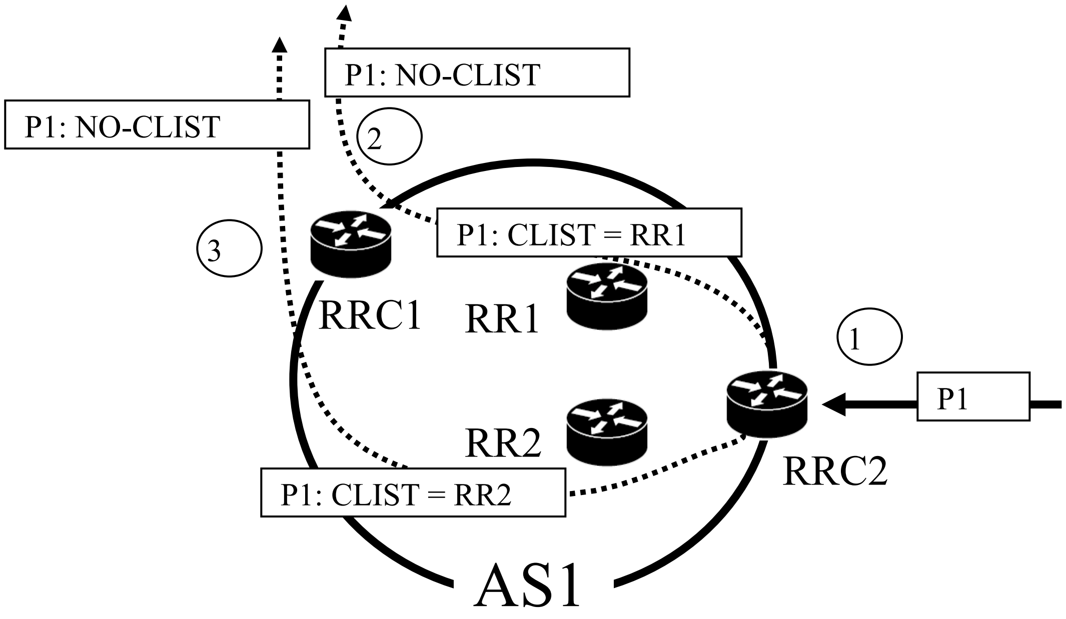
- one day of iBGP and eBGP data
- For every eBGP duplicate we find,
  - look for the same sequence of signatures within a time window of T to find the matching signatures in iBGP
    - $sig(u) = \mathrm{peer} \,||\, \mathrm{asn} \,||\, \mathrm{prefix} \,||\, \mathrm{aspath} \,||\, \mathrm{origin} \,||\, \mathrm{comm} \,||\, \mathrm{agg}$
    - T = 5 min
  - For the matching iBGP update found, compare it with the previous update for this prefix to find the difference
- Total eBGP duplicates examined: 183182

The Results



eBGP Duplicate Count	% Total	Observed iBGP differences
173594	94.77	Δ (cluster-list only)
244	0.13	Δ (cluster-list + others)
1371	0.75	Δ (originator-id + others)
1057	0.58	Δ (cluster-list + originator-id + others)
269	0.15	Δ (med)
6647	3.63	No match found
183182	100.00	

Example of a duplicate update occurrence

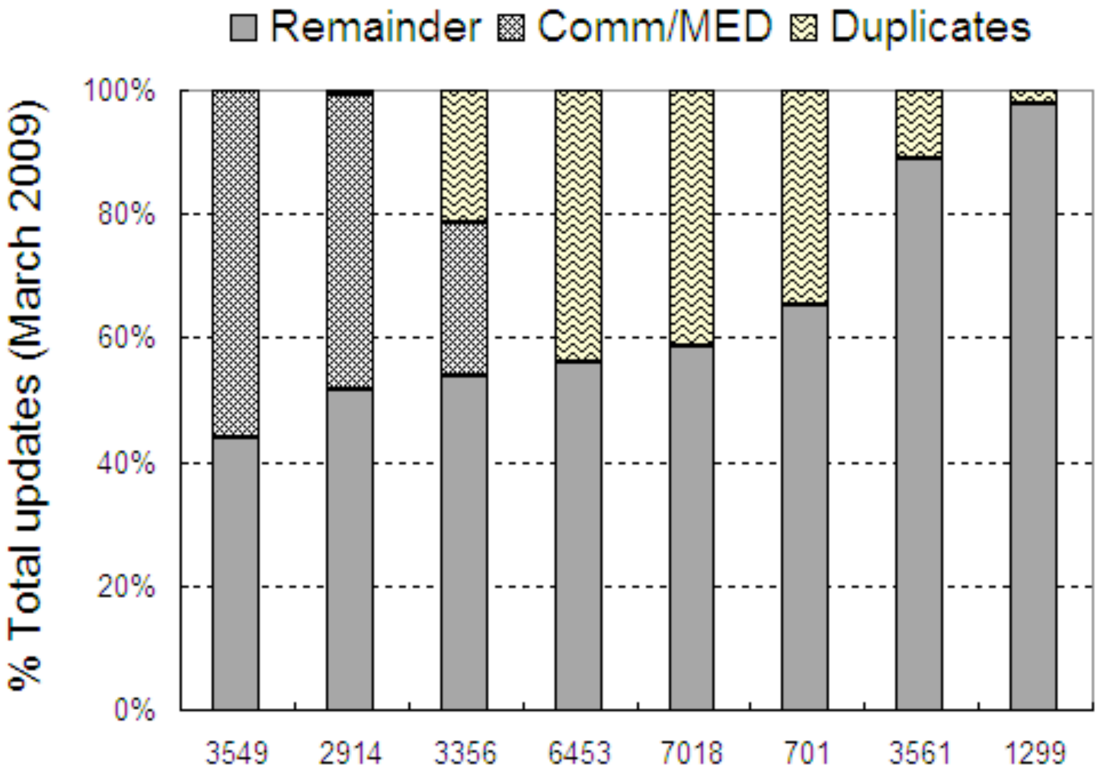


- Path from which the announcement is delivered flaps between RRC2-RR1-RRC1 and RRC2-RR2-RRC1
- When sending the update to eBGP peers, CLIST field is striped off by RRC1
- More alternative paths within AS → more internal path exploration → more duplicates

Discussion

- For this particular Tier1 ISP,
  - Duplicates are due to router software
    - Internal routing dynamics → external duplicates
  - More internal path exploration → more duplicate updates
  - Prefixes can be dampened if there are internal route flaps within the provider network regardless of the stability of the originator
- We conjecture that the same phenomenon happens in other ASes, and we need to verify if this is true

duplicates may exist in other forms



- We also saw from the example tier-1's iBGP data that internal non-transitive attribute's oscillation (cluster-list) is coupled with transitive attributes (community) values changes
- we conjecture that this is the case for AS2914 and AS3549, where MED and MED+comm are coupled with internal flapping

Summary

- We observe non-trivial amount of eBGP duplicate updates
- Duplicate updates can affect reachability (if caused dampening), add to router load (during peak load time)
- Our measurement suggests one cause of duplicates that is responsible for most, if not all, duplicates
  - the internal dynamics leak to the outside in the form of duplicate updates
- There exist other forms of noise in BGP, and this work is a first step in reducing such noise