Universidad
Carlos III de Madrid

Power Prefixes Prioritization for Smarter BGP Reconvergence

Juan Brenes

Marcelo Bagnulo
Alberto García-Martínez
Andra Lutu
Cristel Pelsser

Wait... What!?

- We propose a simple change over the convergence algorithm to reduce the impact of BGP failure events.

- Our proposal does NOT require any change in the protocol

- The traffic losses are reduced by one order of magnitude.

OK...But why?

- MONEY:
  - CDNs & other Internet services
  - ISPs
- Service Level Agreements.
- Delay sensitive services in Internet.

Lot of previous research… but

- Most of them are based in:
  - Decrease the time it takes to converge.
  - Decrease the amount of information exchanged.
- Usually require changes in the protocol
- Not universally aplicable:
  - PIC and ADD-PATH with Next-hop self

We follow a completely different approach based on 3 observations.

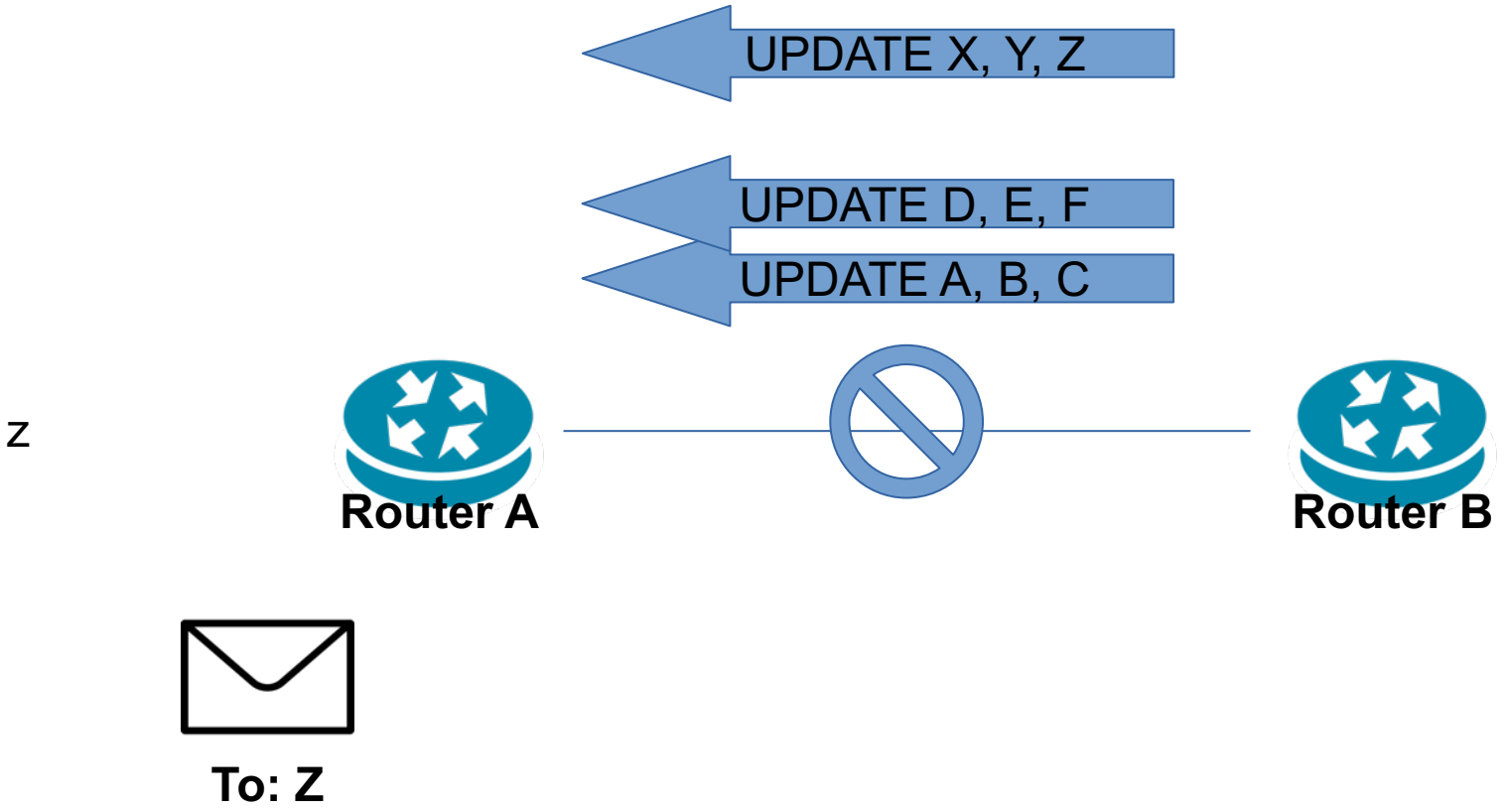
1) A single BGP event may affect a large number of prefixes.

2) The time it takes for BGP to restore reachability after a BGP event that affects a large number of routes is different for each of the prefixes affected.

3) In practice only a small number of prefixes are relevant for the AS operation.

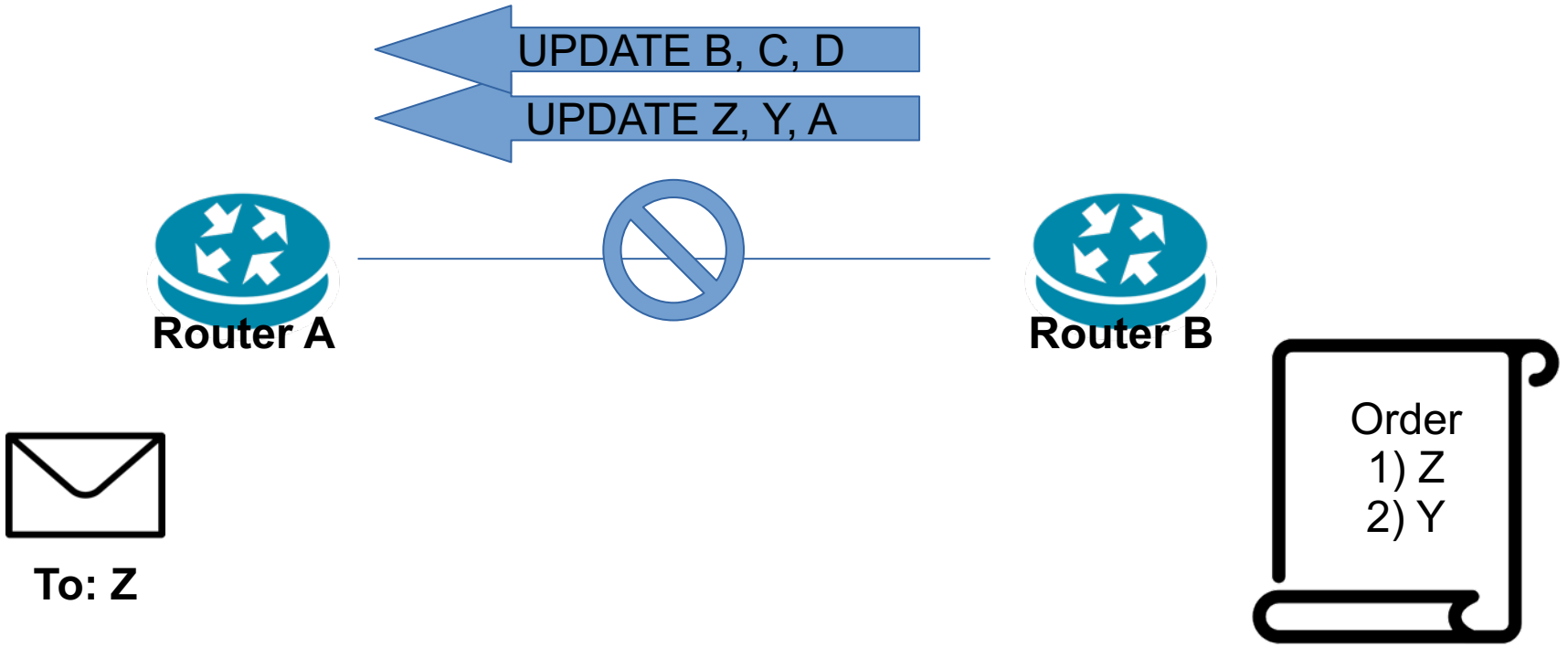
PPP proposes to significantly reduce the impact of a failure event by ensuring that the most relevant routes converge before routes to less relevant prefixes.

The current situation (lexicographical order)

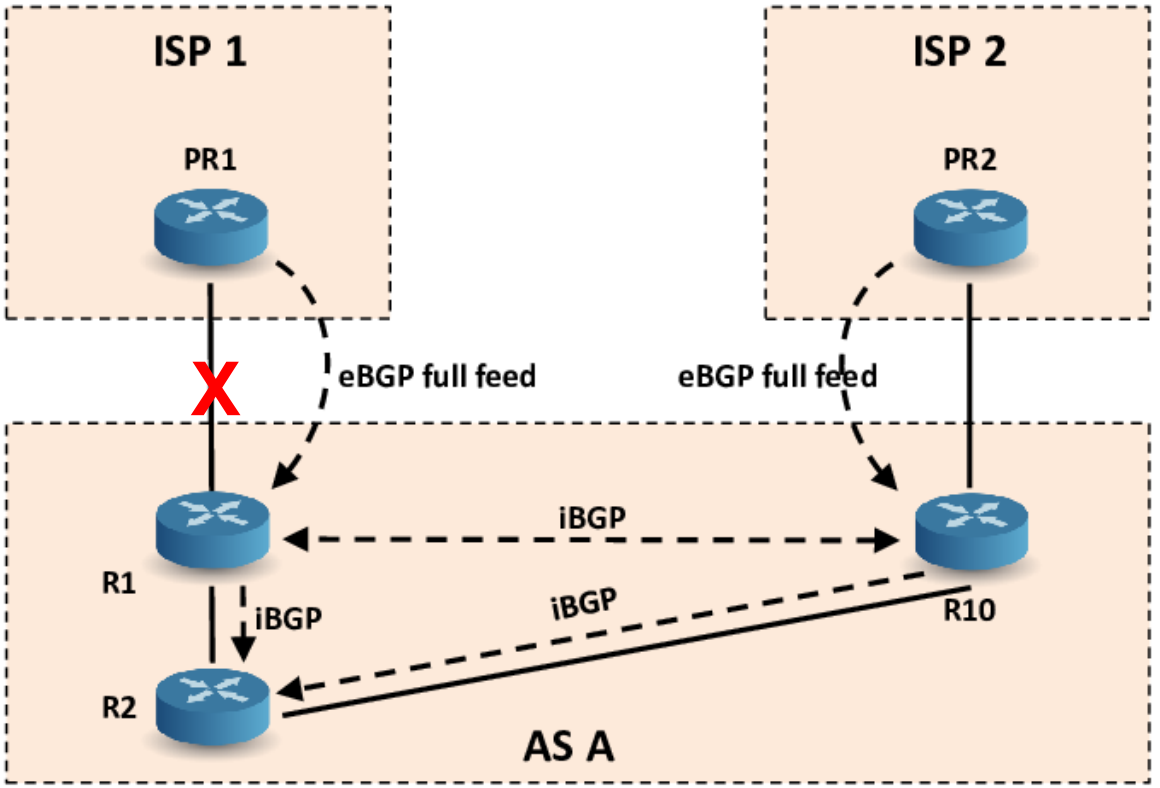


z

Our proposal



BGP toy example



The implementation is just a list!!!

...it is giving control in the order

…we wanted results

We verified the results of our algorithm using real ISP traces

- 2 datasets (2014 and 2015) containing traces from a ISP to its transit provider.
- We used the amount of traffic as the ranking parameter.
- BGP dumps to match the destinations.

Variables

- For how long can we use the same ranking? (**validity period**)
- Does the time of taking the samples affect? (**measuring interval**)
- Is it efficient to sample the traffic? (**sampling rate**)

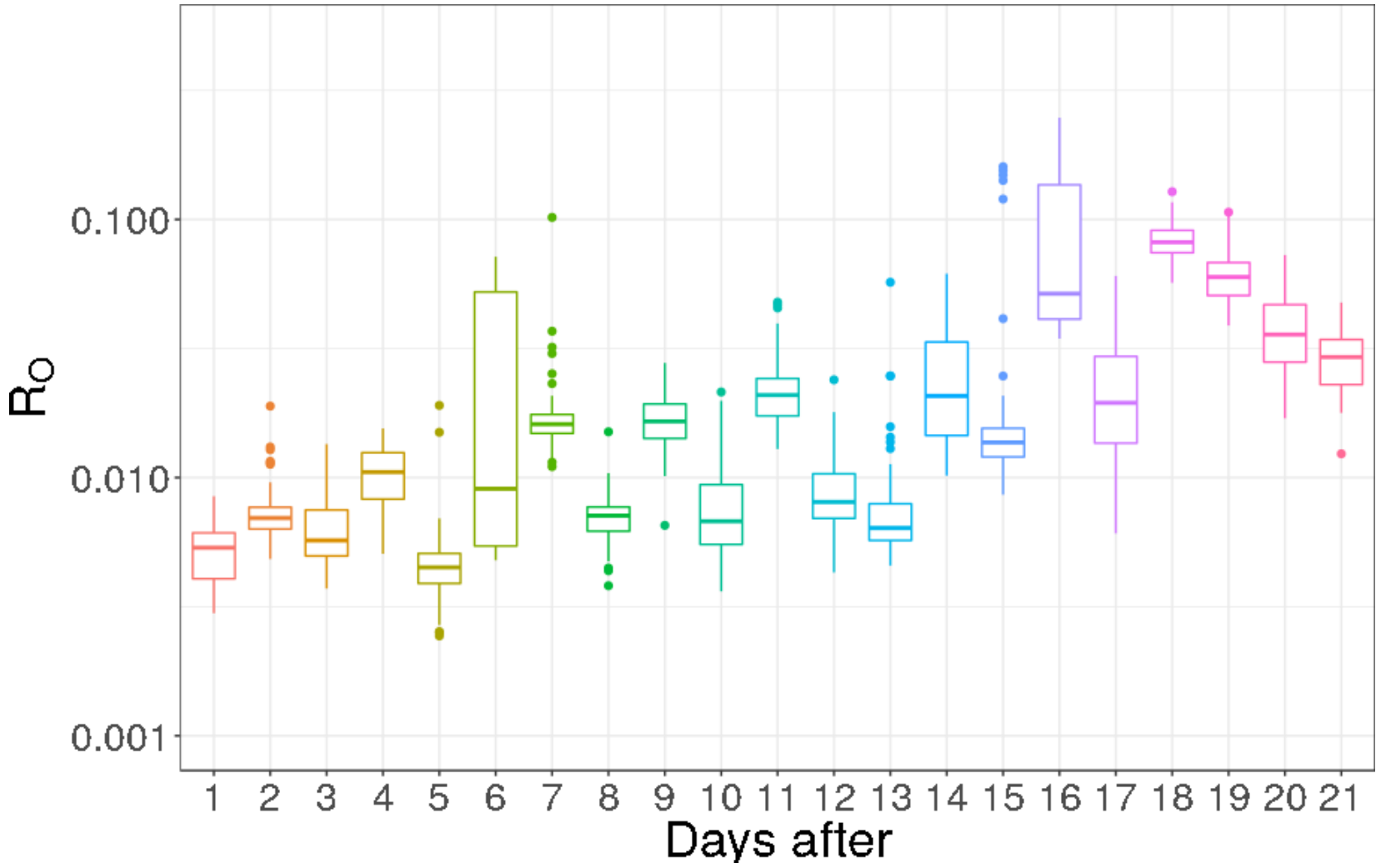
We use the ratio of traffic loss

- We calculate the traffic losses at a given moment using lexicographical order.

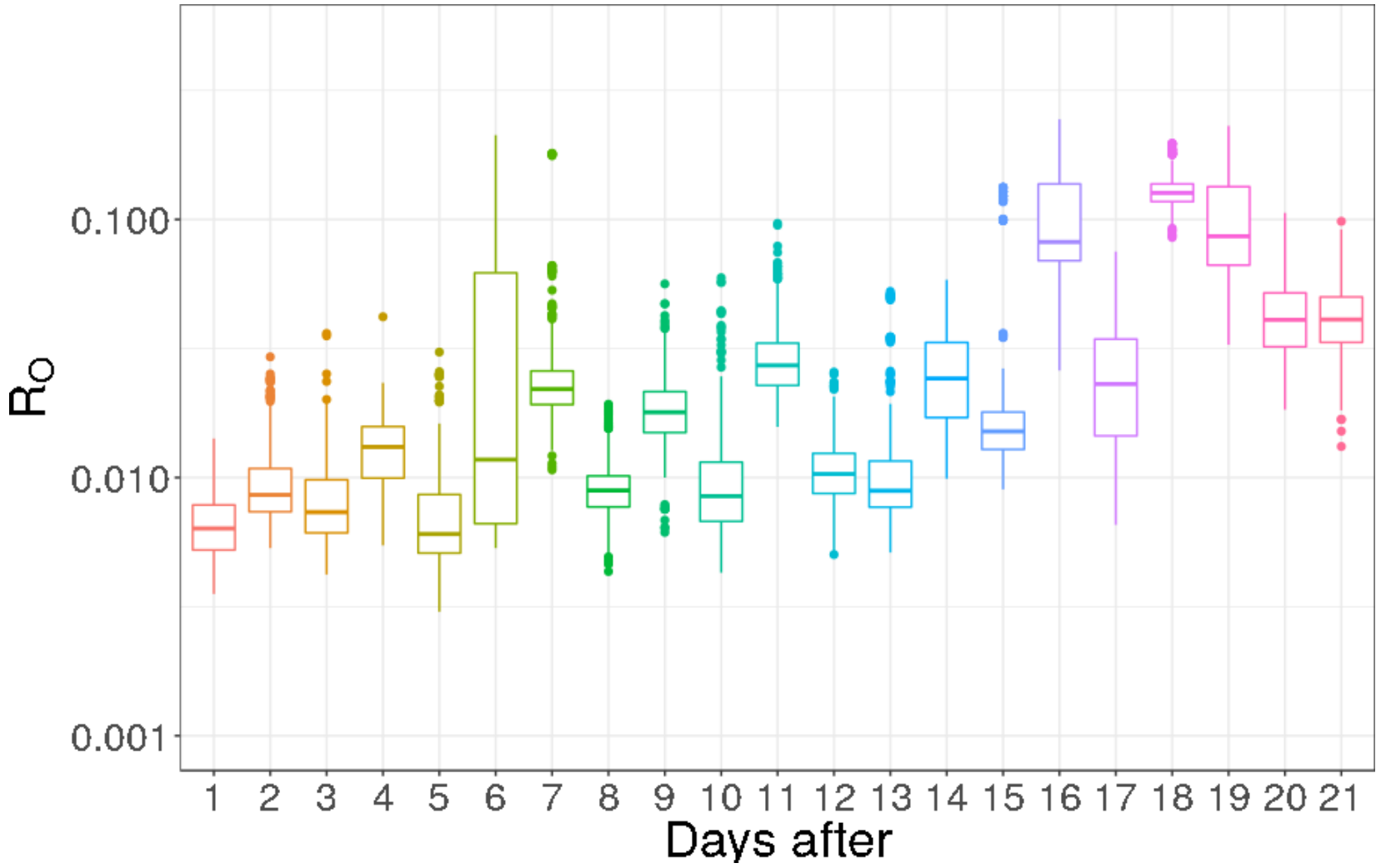
- We calculate the traffic losses at a given moment using a PPP rank.

- We use the ratio between these two numbers as a comparing mechanism.

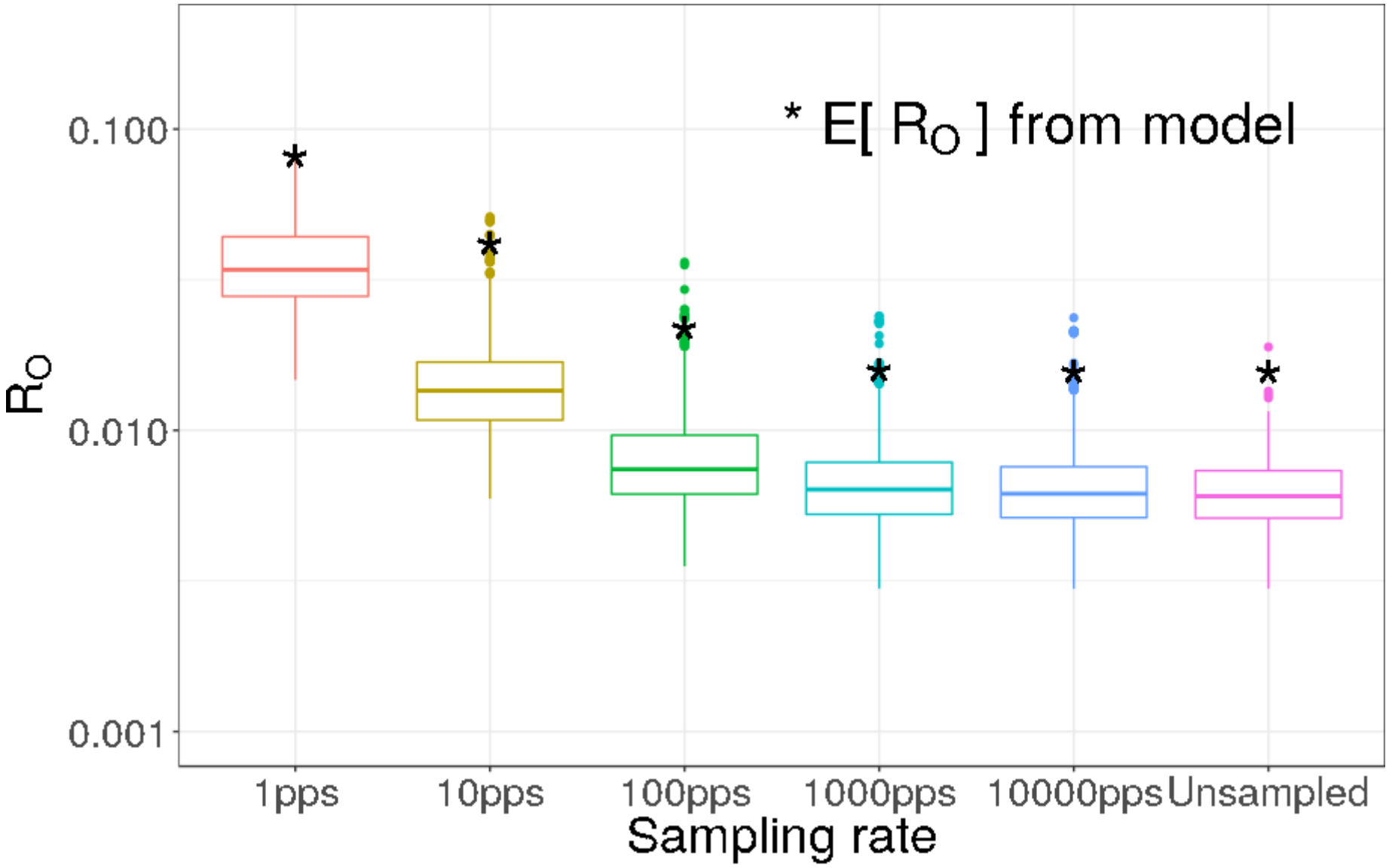
Validity period



Validity period 100pps



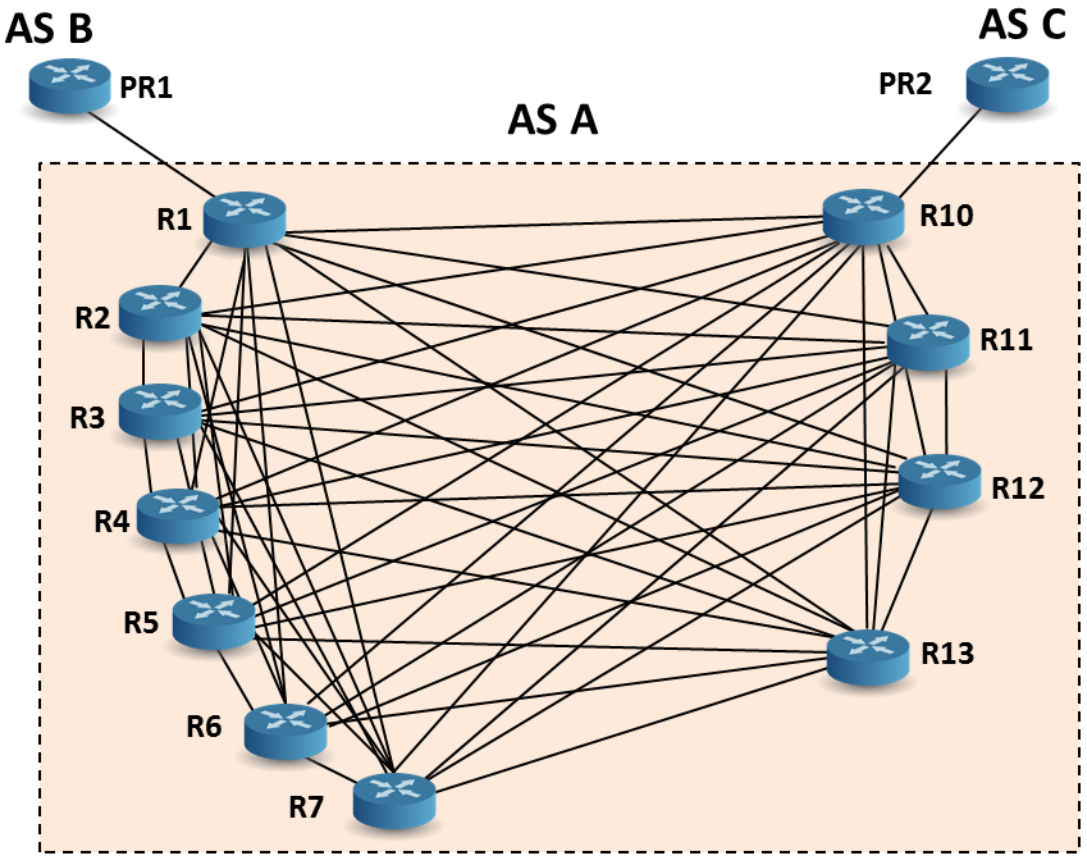
Sampling rate analysis



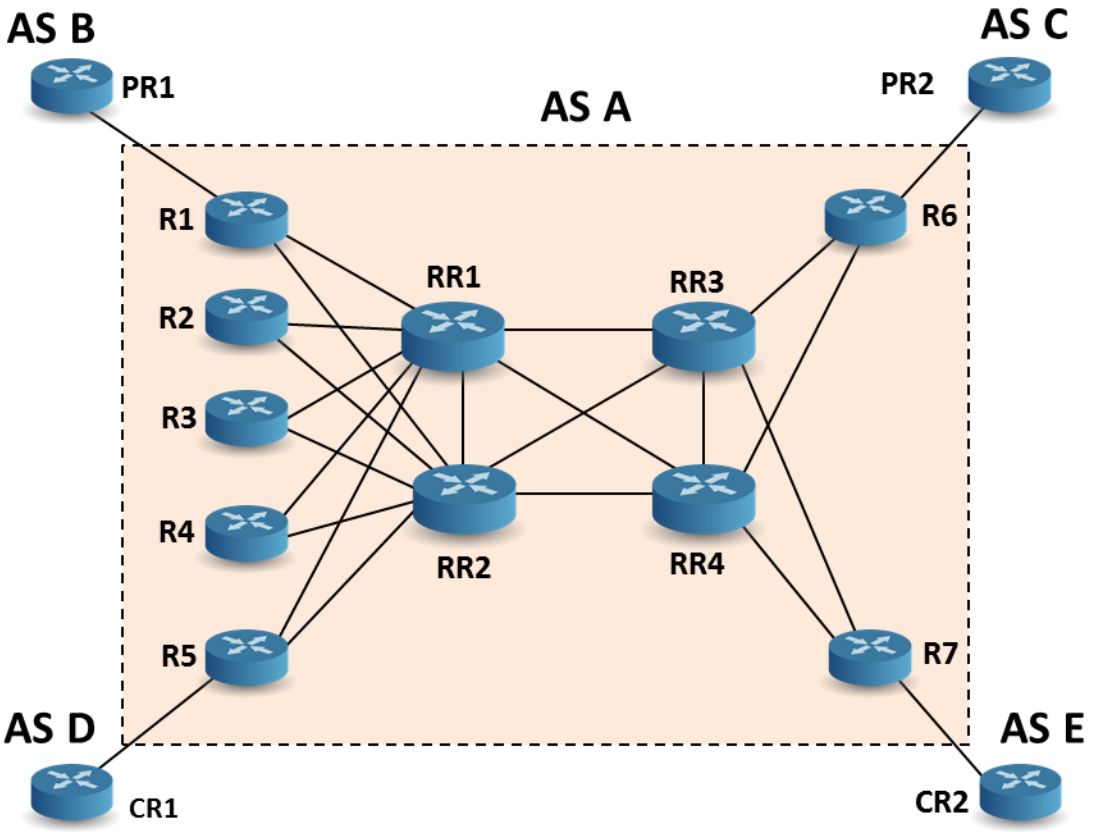
We modified bgp quagga

- BGP-PPP uses a list of prefixes to establish the orders of the updates.
- Prefix ranks obtained from the 24-hour measuring interval, using different sampling rates.
- 15-second bins from the 3 days after the predictor dataset.

Topologies – Full-mesh



Topologies – Route-reflector



Results

TABLE III: $E[R_O]$ and R_O results for full-mesh and route reflector topologies, 2014

Type	**Sampling Rate**					
	1	10	100	1000	10000	Unsampled
$E[R_O]$	.0863	.0477	.0299	.0239	.0236	.0236
R2, full-mesh	.0298	.0108	.0062	.0055	.0053	.0051
RR1, route reflector	.0321	.0126	.0080	.0074	.0074	.0074

Conclusions

- We can do better without changing the protocol itself!
- PPP is universally applicable.
- It is feasible to automatically generate the ranks using the amount of traffic.


KEEP
CALM
AND
ASK
QUESTIONS

It is my turn now!

1. Suggestions in how to follow? IETF?

2. Shall we extend this?