IGP Deflections and BGP Divergence

Thesis Mid-Defense

PhD Student
Julian Martin Del Fiore

Committee
Chadi Barakat
Olivier Bonaventure
Pascal Merindol
Cristel Pelsser

msii ÉCOLE DOCTORALE

Université de Strasbourg

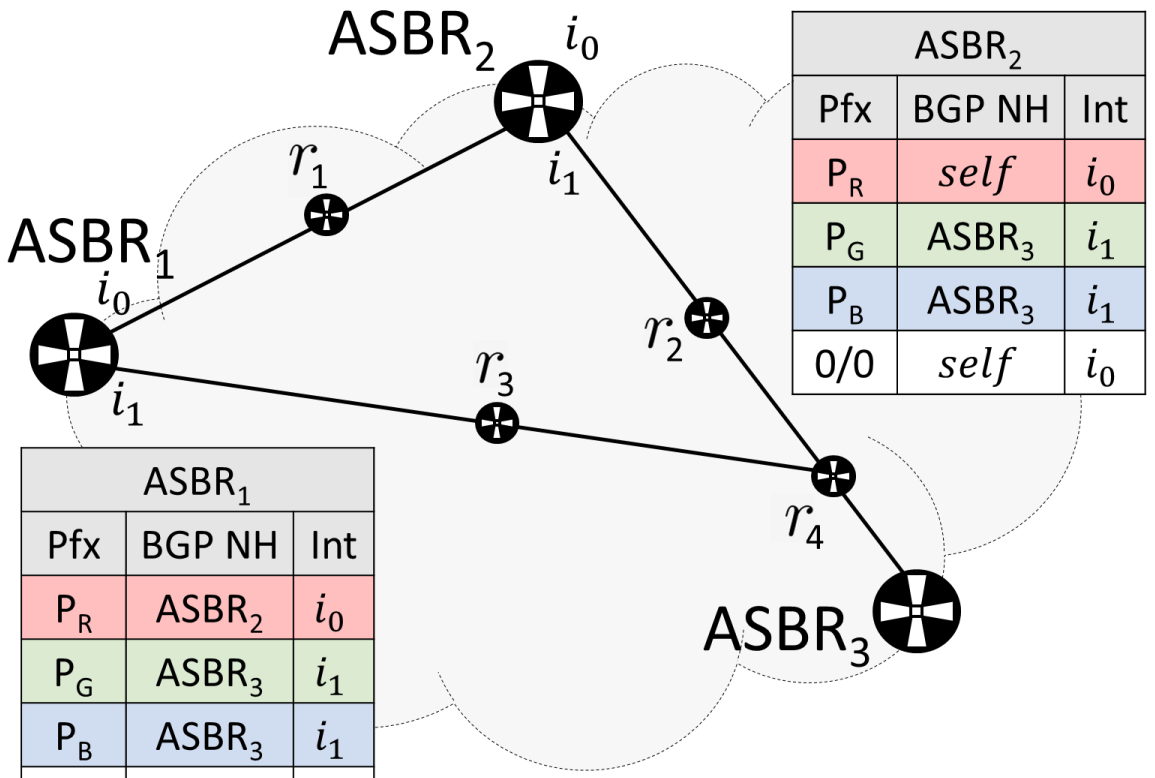
iCUBE

Submitted to IMC 2020

"Capturing Forwarding Deflections in the Wild Desired Load Balancing or Unwanted Detours?"

Julian M. Del Fiore, Valerio Persico, Pascal Merindol, Cristel Pelsser, Antonio Pescape

A case where all routers have **full-FIBs**

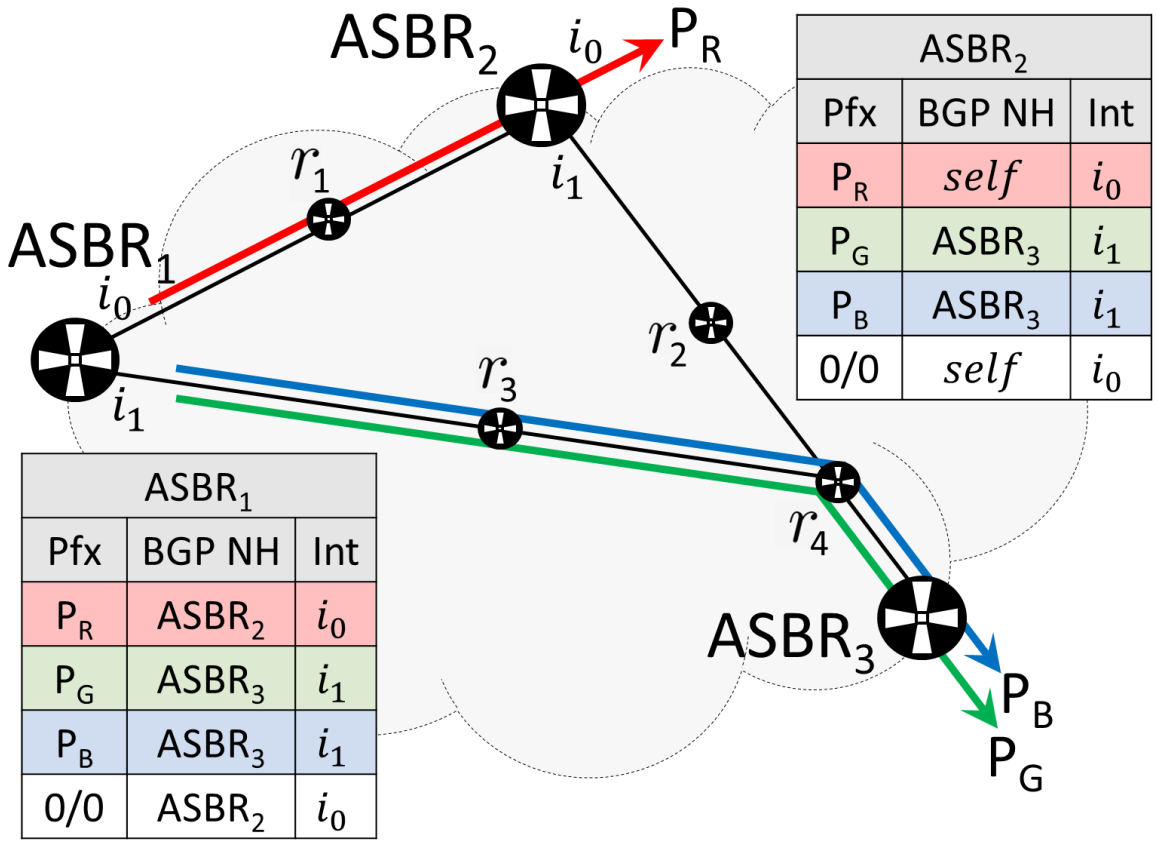


ASBR$_2$		
Pfx	BGP NH	Int
P$_R$	*self*	i_0
P$_G$	ASBR$_3$	i_1
P$_B$	ASBR$_3$	i_1
0/0	*self*	i_0

ASBR$_1$		
Pfx	BGP NH	Int
P$_R$	ASBR$_2$	i_0
P$_G$	ASBR$_3$	i_1
P$_B$	ASBR$_3$	i_1
0/0	ASBR$_2$	i_0

- The default routes are for **backup** and not used in “normal” behavior

A case where all routers have **full-FIBs**

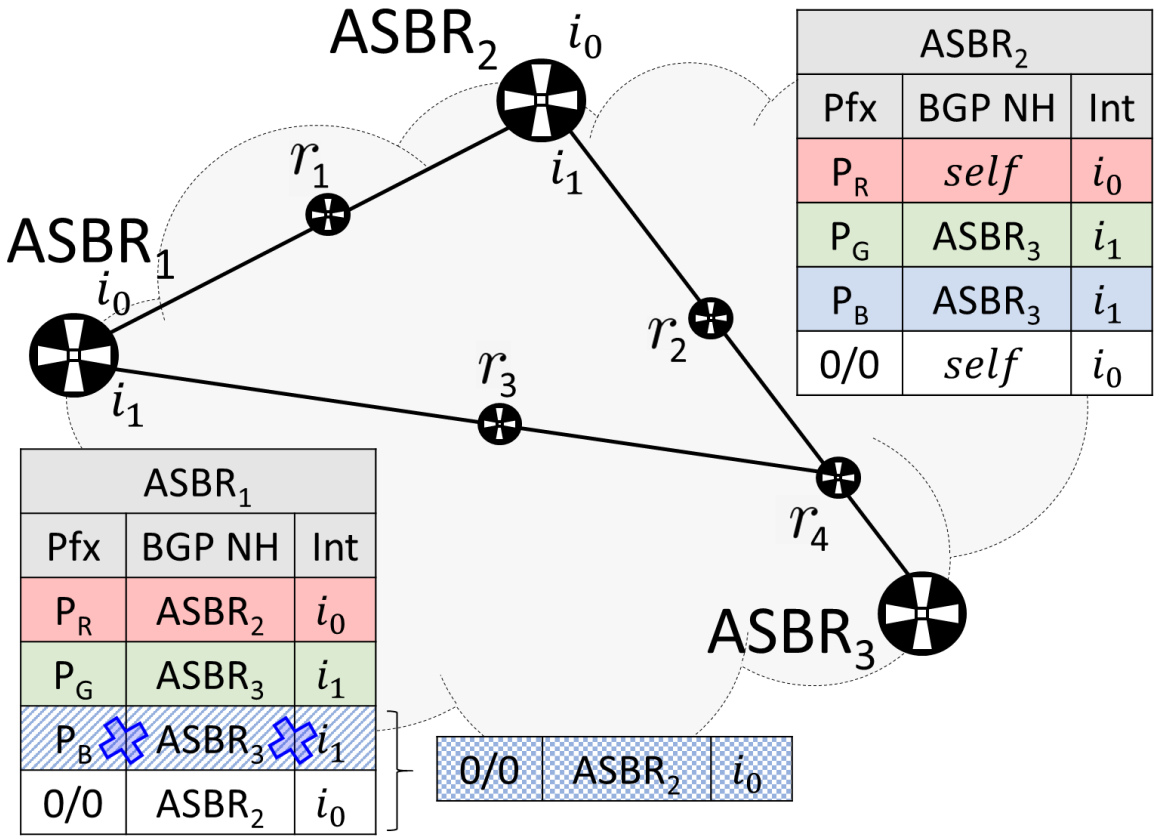


ASBR$_2$		
Pfx	BGP NH	Int
P$_R$	*self*	i_0
P$_G$	ASBR$_3$	i_1
P$_B$	ASBR$_3$	i_1
0/0	*self*	i_0

ASBR$_1$		
Pfx	BGP NH	Int
P$_R$	ASBR$_2$	i_0
P$_G$	ASBR$_3$	i_1
P$_B$	ASBR$_3$	i_1
0/0	ASBR$_2$	i_0

- The default routes are for **backup** and not used in “normal” behavior
- Between two ASBRs and any prefix, the **best IGP path** is followed

$ASBR_1$ has now has a **partial-FIB**



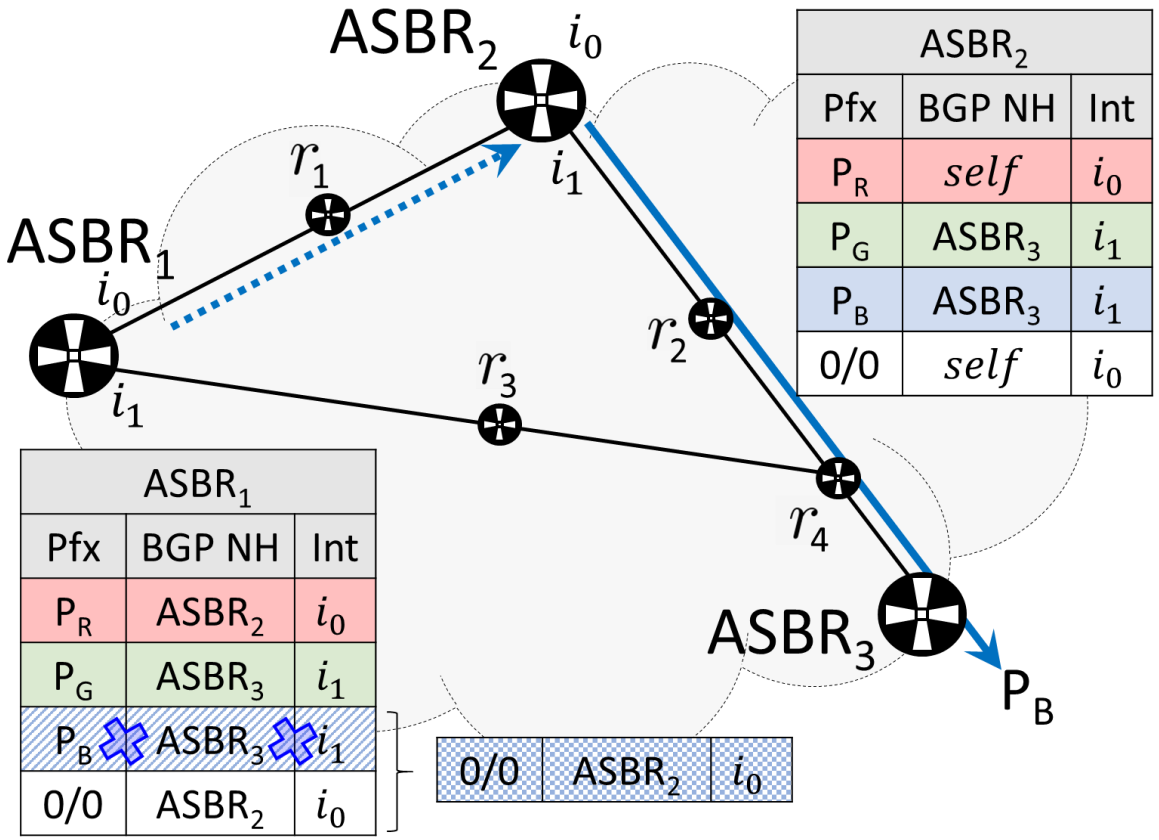
$ASBR_2$		
Pfx	BGP NH	Int
P_R	*self*	i_0
P_G	$ASBR_3$	i_1
P_B	$ASBR_3$	i_1
0/0	*self*	i_0

$ASBR_1$		
Pfx	BGP NH	Int
P_R	$ASBR_2$	i_0
P_G	$ASBR_3$	i_1
P_B	$ASBR_3$	i_1
0/0	$ASBR_2$	i_0

0/0	$ASBR_2$	i_0

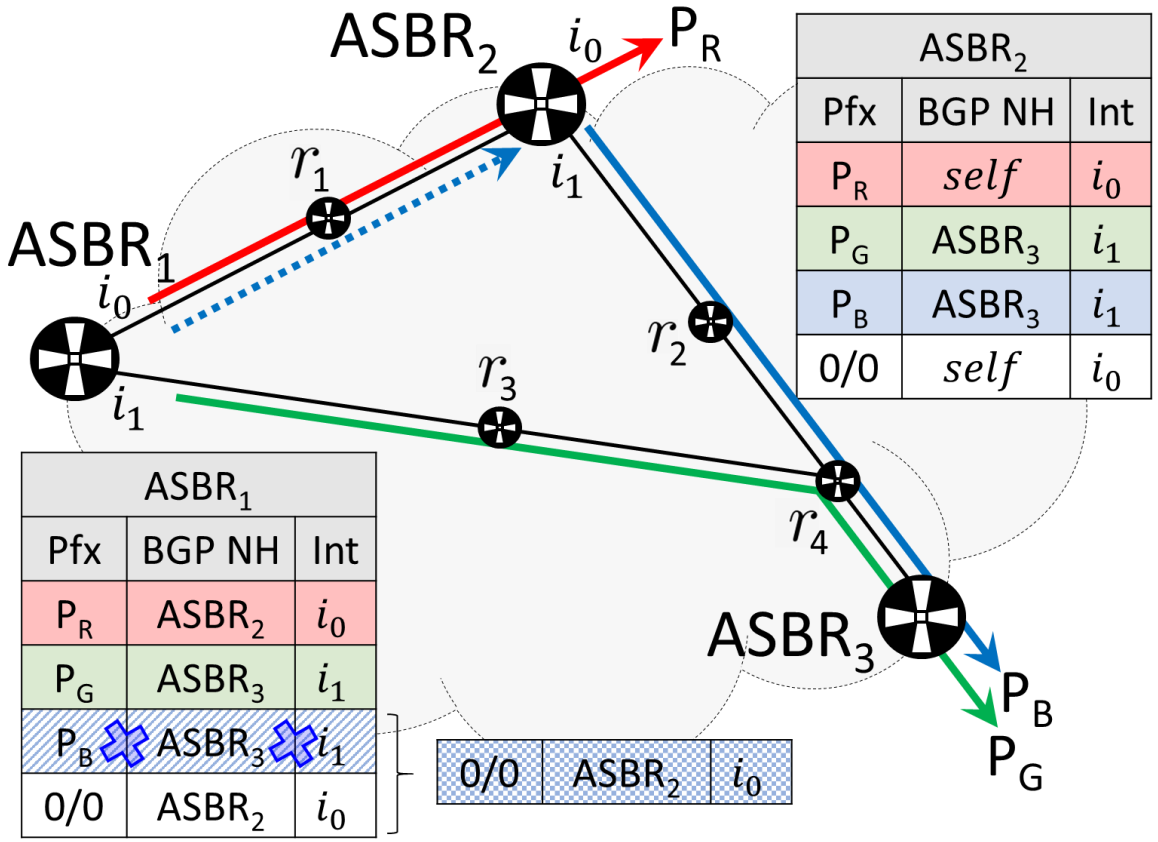
- The default route of $ASBR_1$ covers P_B

$ASBR_1$ has now has a **partial-FIB**



- The default route of $ASBR_1$ covers P_B
- The default routes is **actively** used, the route seen for P_B **detours**

$ASBR_1$ has now has a **partial-FIB**



ASBR2		
Pfx	BGP NH	Int
P_R	*self*	i_0
P_G	$ASBR_3$	i_1
P_B	$ASBR_3$	i_1
0/0	*self*	i_0

ASBR1		
Pfx	BGP NH	Int
P_R	$ASBR_2$	i_0
P_G	$ASBR_3$	i_1
P_B	$ASBR_3$	i_1
0/0	$ASBR_2$	i_0

0/0	$ASBR_2$	i_0

- The default route of $ASBR_1$ covers P_B
- The default routes is **actively** used, the route seen for P_B **detours**
- No detour for P_G between the same ASBRs, thus a **multipath** routing pattern

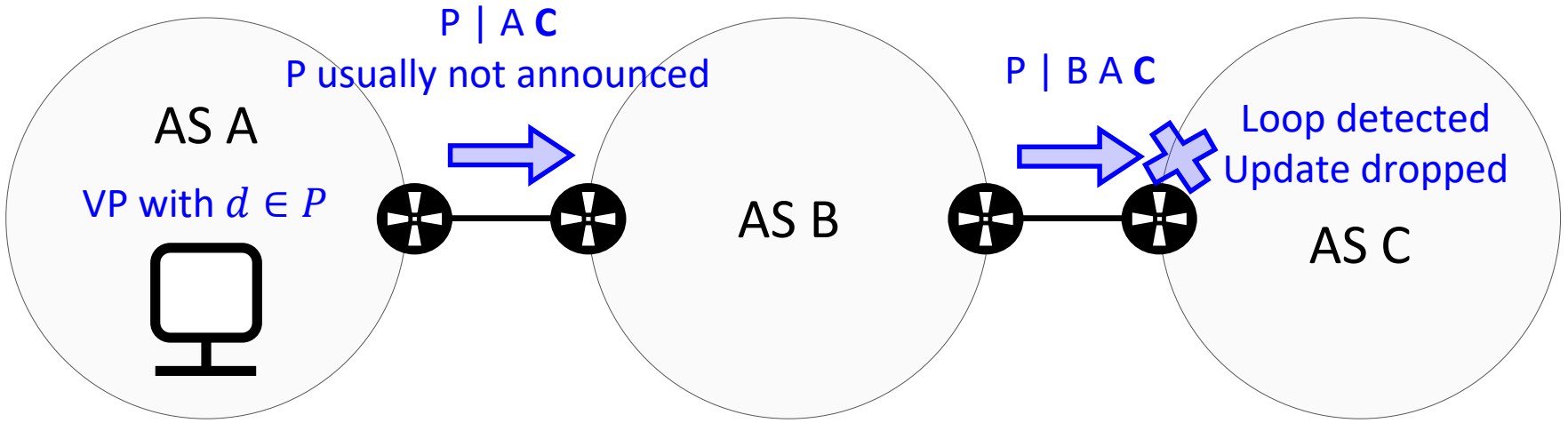
Previous work*: detect **backup** default routes with **AS poisoning**

*Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. 2009. **Internet optometry: assessing the broken glasses in internet reachability**. In IMC '09 Association for Computing Machinery, New York, NY, USA, 242–253.

With this method, it is **impossible** to know if a detour occurred

Previous work*: detect **backup** default routes with **AS poisoning**

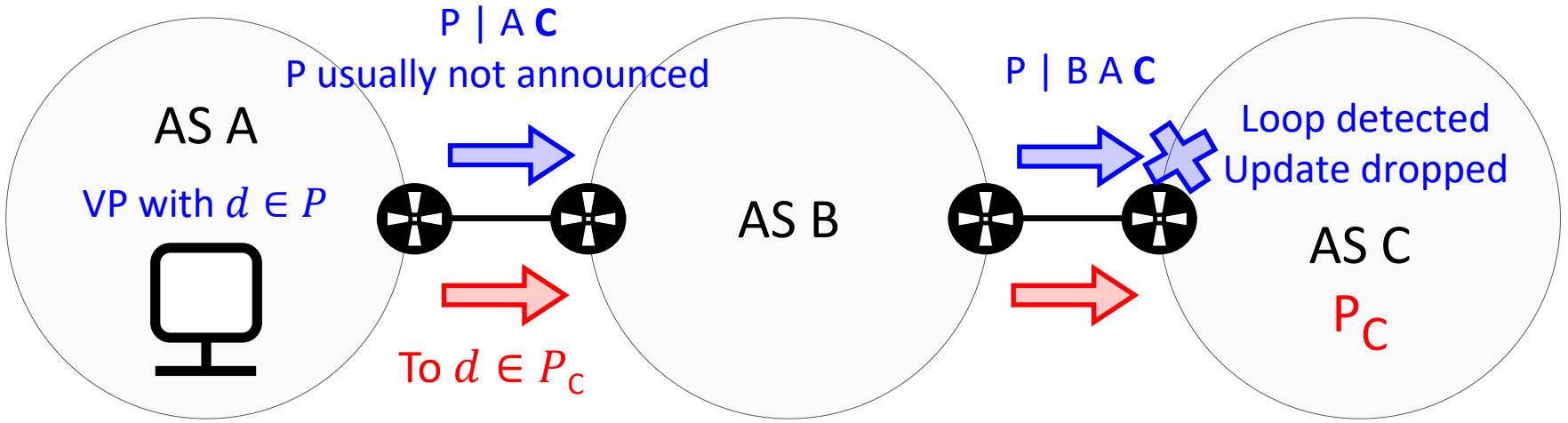
*Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. 2009. **Internet optometry: assessing the broken glasses in internet reachability**. In IMC '09 Association for Computing Machinery, New York, NY, USA, 242–253.



With this method, it is **impossible** to know if a detour occurred

Previous work*: detect **backup** default routes with **AS poisoning**

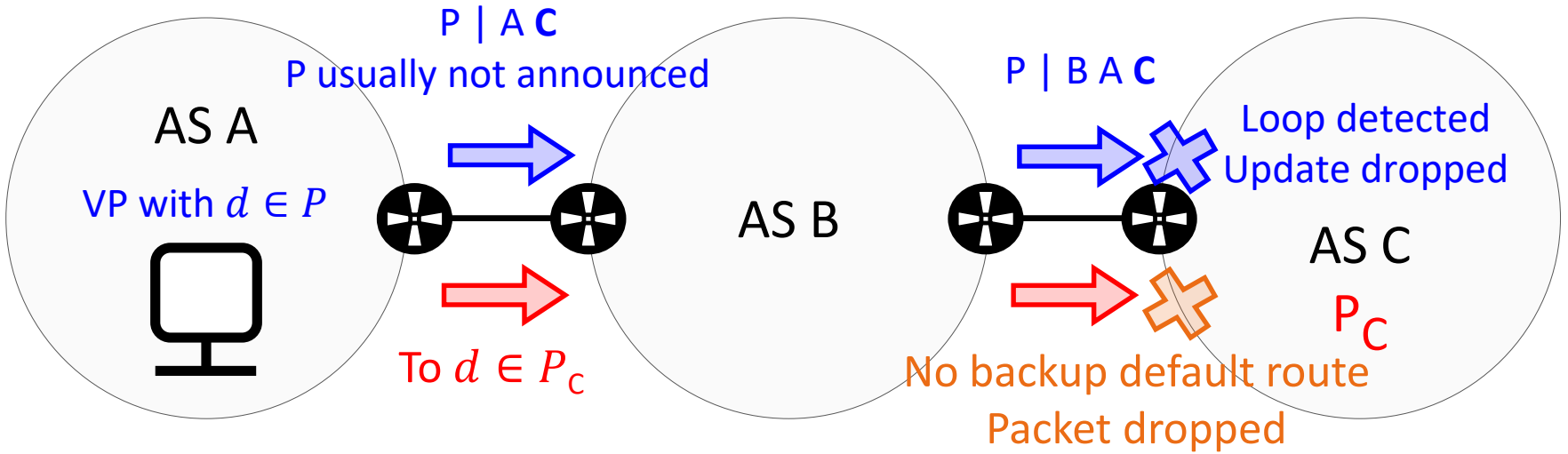
*Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. 2009. **Internet optometry: assessing the broken glasses in internet reachability**. In IMC '09 Association for Computing Machinery, New York, NY, USA, 242–253.



With this method, it is **impossible** to know if a detour occurred

Previous work*: detect **backup** default routes with **AS poisoning**

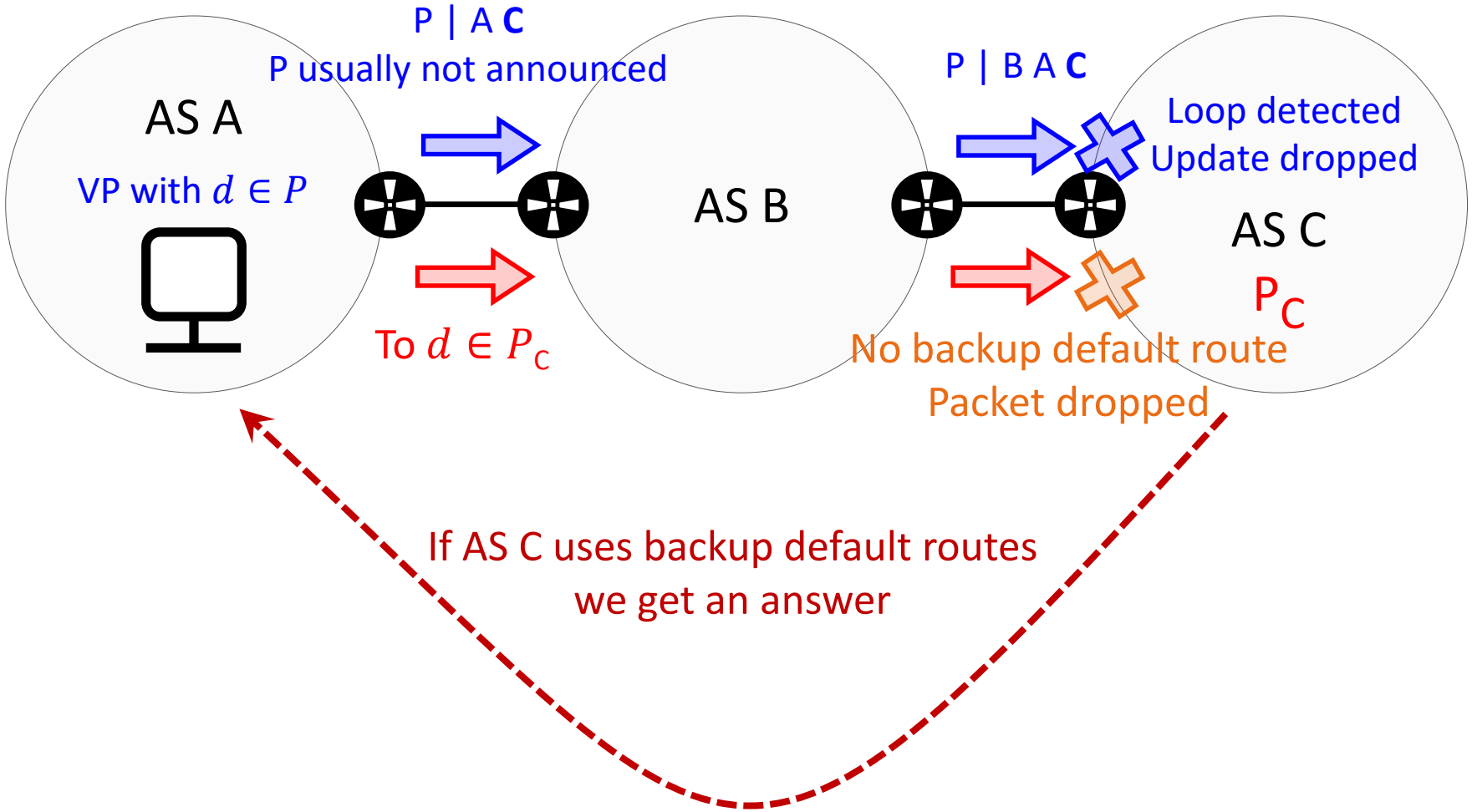
*Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. 2009. **Internet optometry: assessing the broken glasses in internet reachability**. In IMC '09 Association for Computing Machinery, New York, NY, USA, 242–253.



With this method, it is **impossible** to know if a detour occurred

Previous work*: detect **backup** default routes with **AS poisoning**

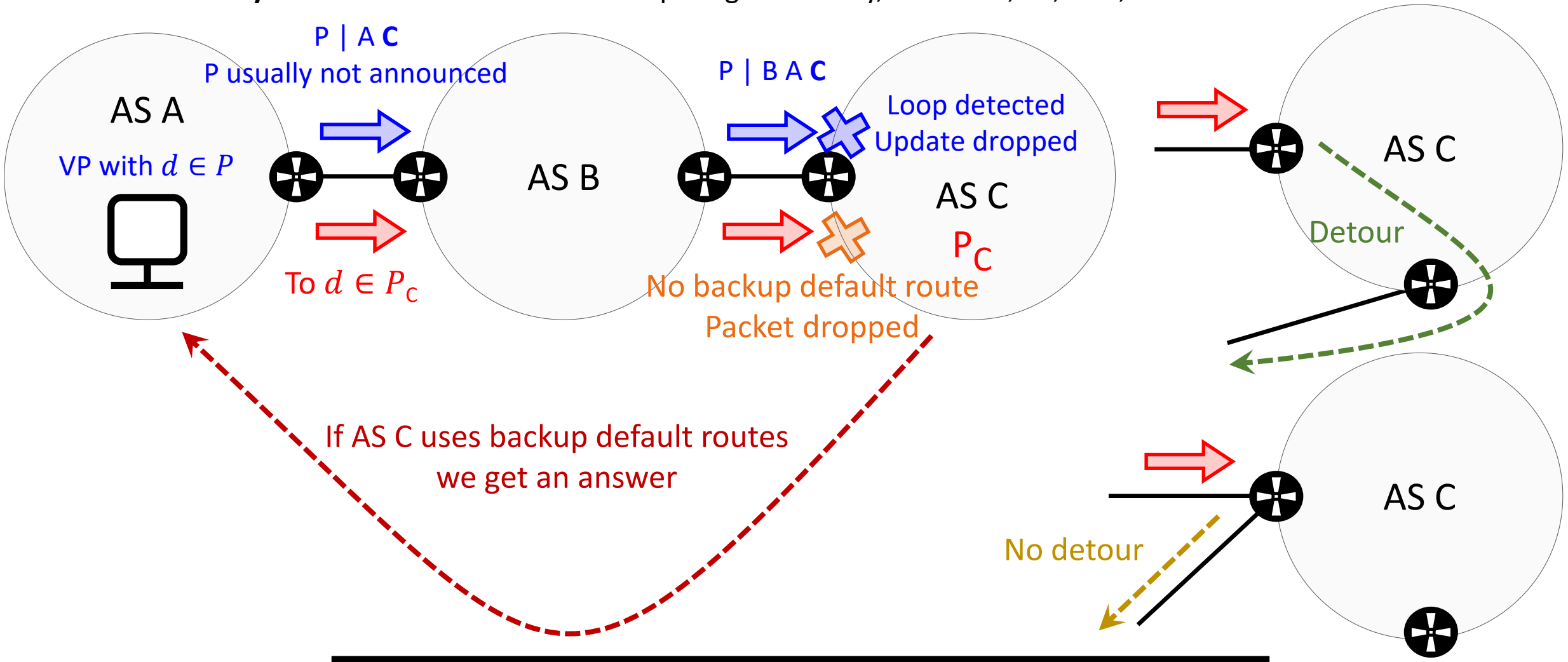
*Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. 2009. **Internet optometry: assessing the broken glasses in internet reachability**. In IMC '09 Association for Computing Machinery, New York, NY, USA, 242–253.



With this method, it is **impossible** to know if a detour occurred

Previous work*: detect **backup** default routes with **AS poisoning**

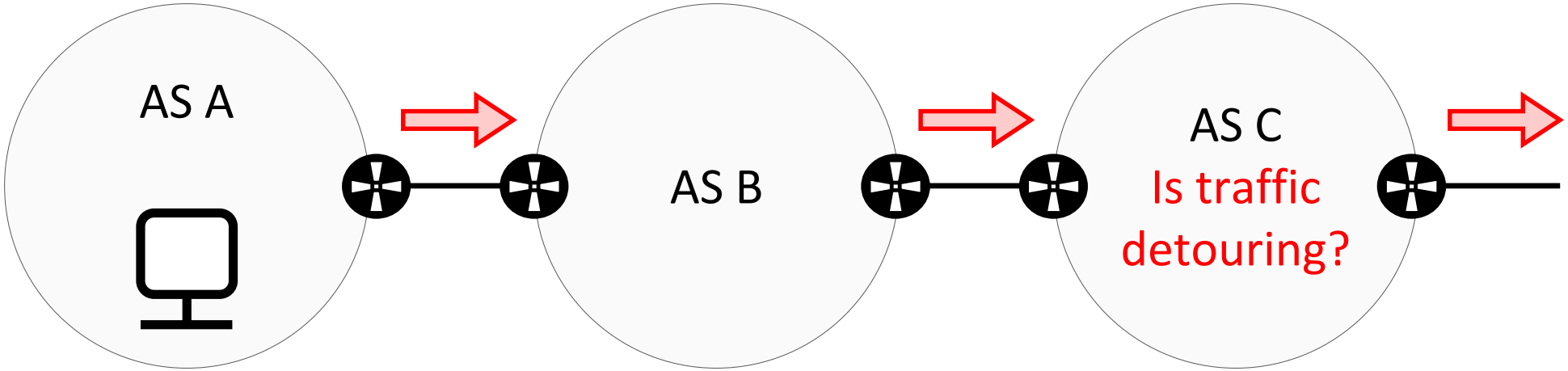
*Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. 2009. **Internet optometry: assessing the broken glasses in internet reachability**. In IMC '09 Association for Computing Machinery, New York, NY, USA, 242–253.



With this method, it is **impossible** to know if a detour occurred

Our Contribution

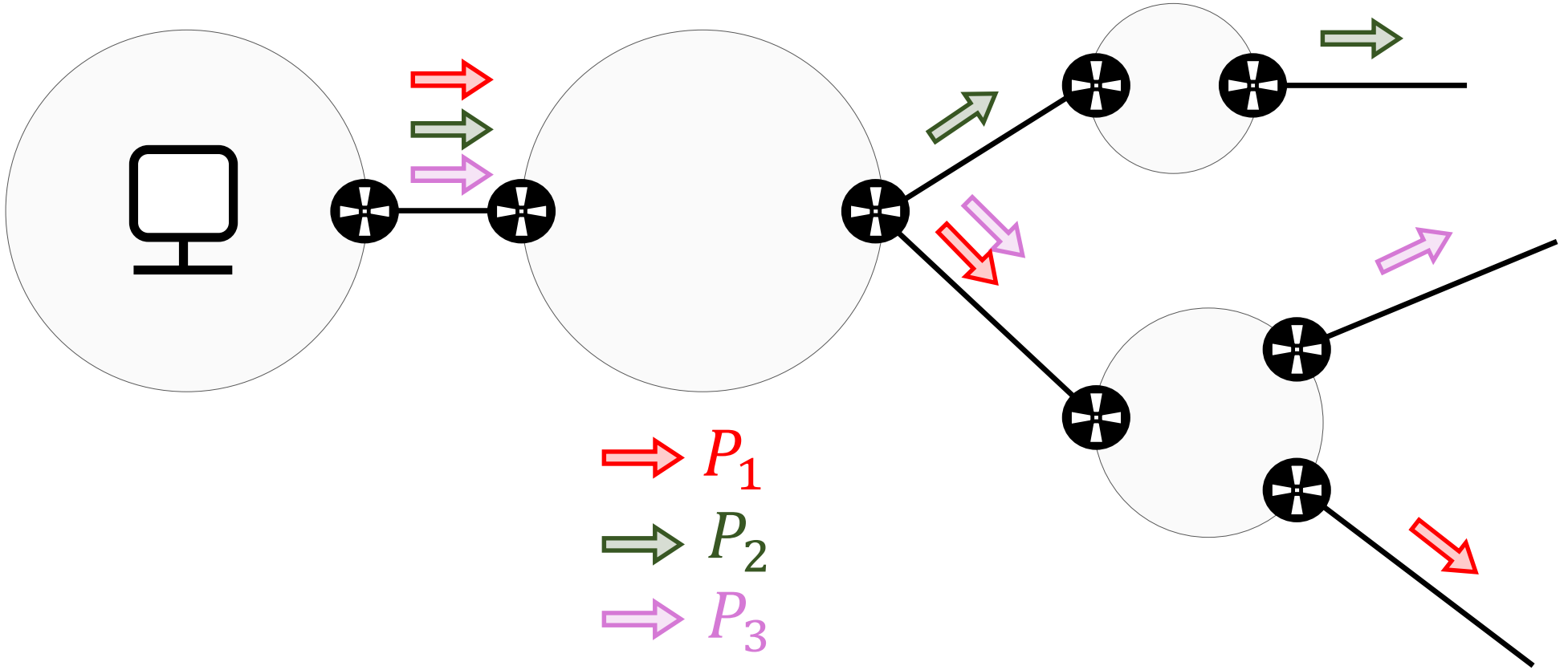
Use **traceroute** to detect **detours** likely resulting
from **actively** used default routes installed in **partial-FIB** routers



Detecting FDs in ASes that do **not deploy Load Balancing (LB)**

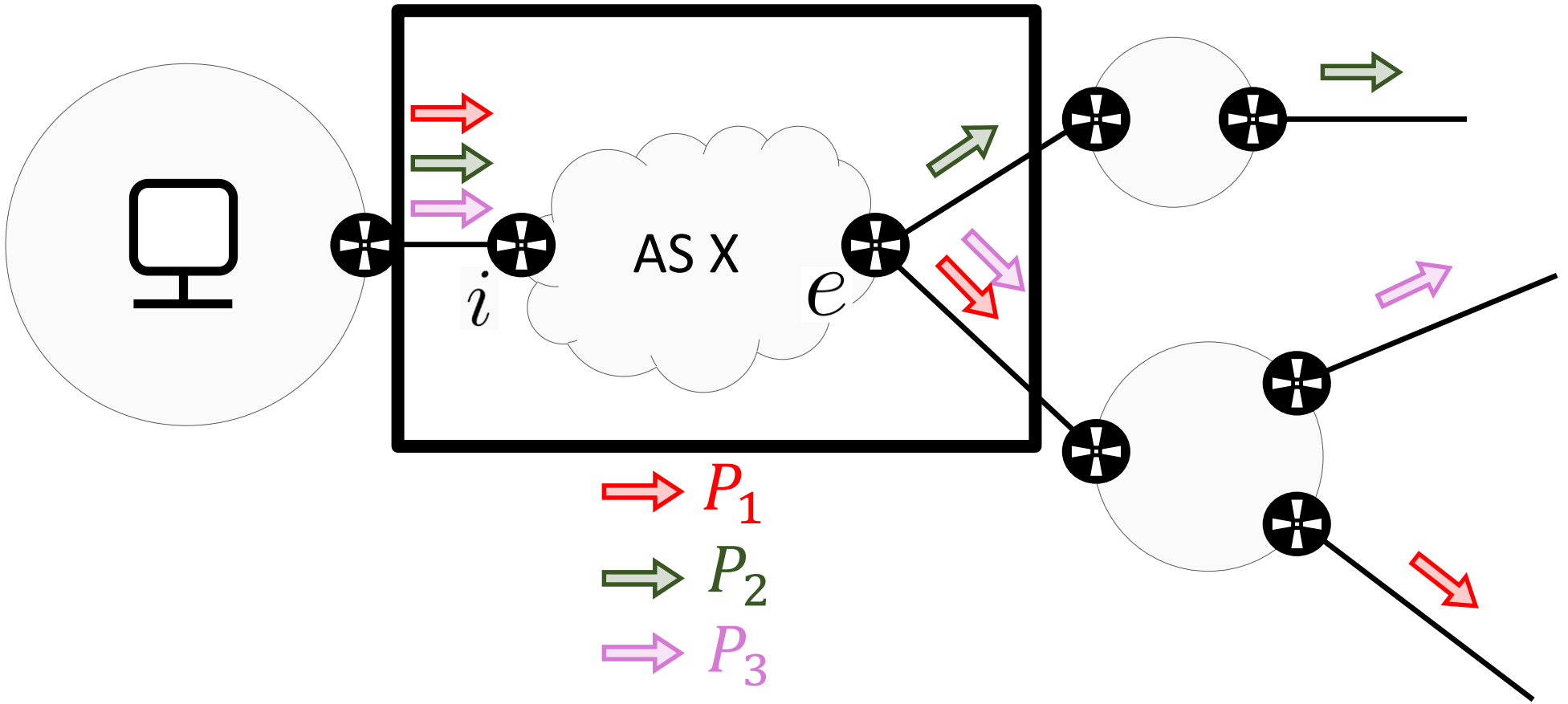
Detecting FDs in ASes that do **not deploy Load Balancing (LB)**

- Run **traces** towards different /24s



Detecting FDs in ASes that do **not deploy Load Balancing (LB)**

- Run **traces** towards different /24s
- Identify those traversing AS X and the **ingress-egress** points (i, e)



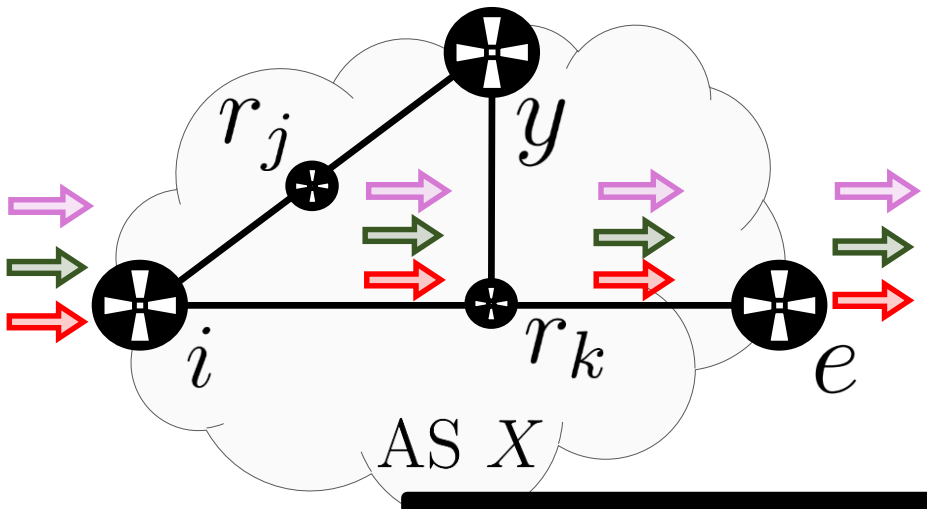
Detecting FDs in AS X that does **not deploy Load Balancing (LB)**

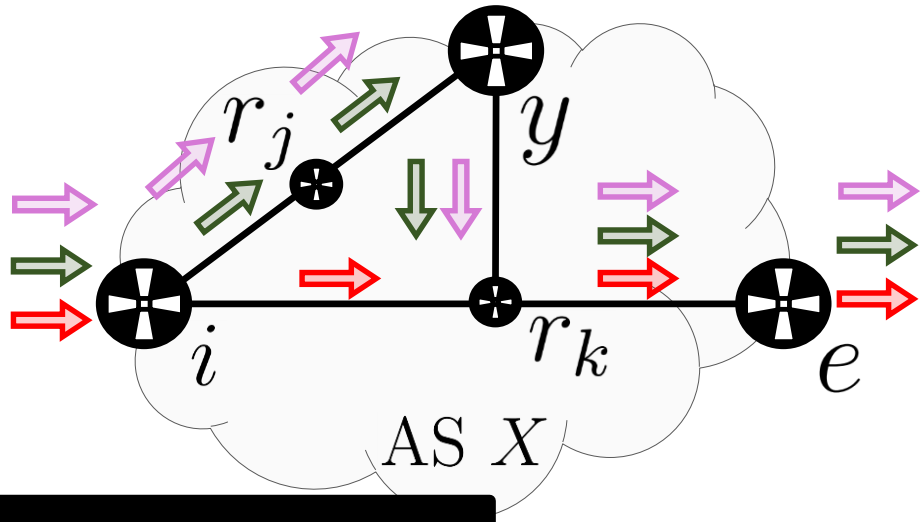
- Run **traces** towards different /24s
- Identify those traversing AS X and the **ingress-egress** points (i, e)
- **Compare** the internal routes per (i, e) and if two mismatch, FDs occur

OK

P_1
P_2
P_3

Forwarding Detours





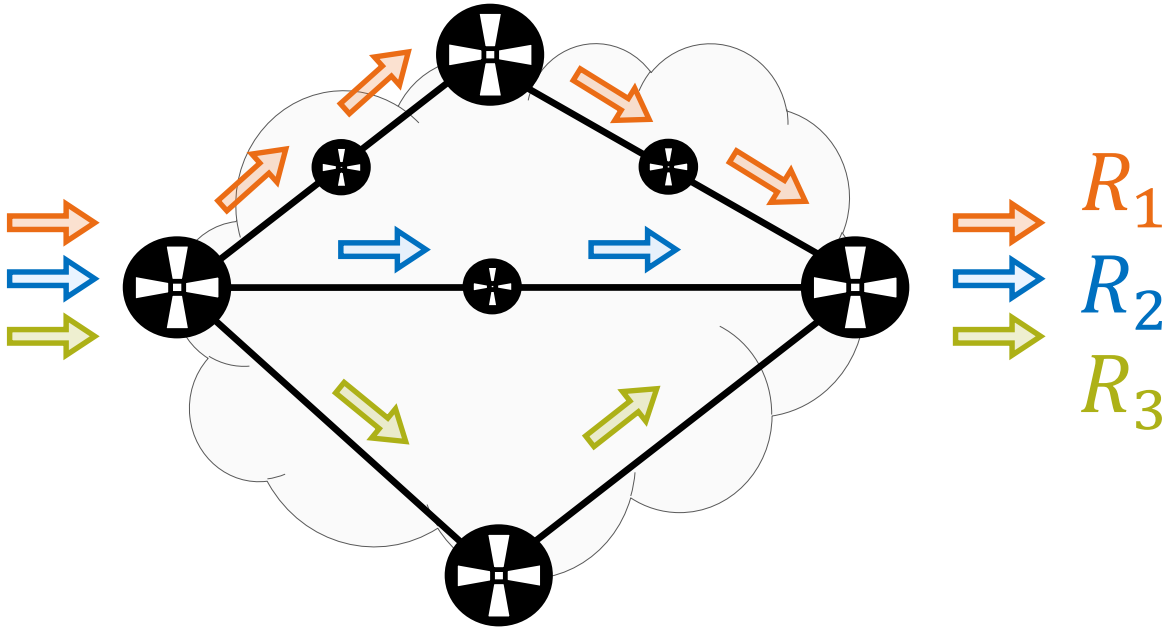
If no LB then: FDs ⇔ Multipath routing

When **LB is deployed** only FDs ⇒ multipath routing holds

This occurs since LB ⇒ multipath routing

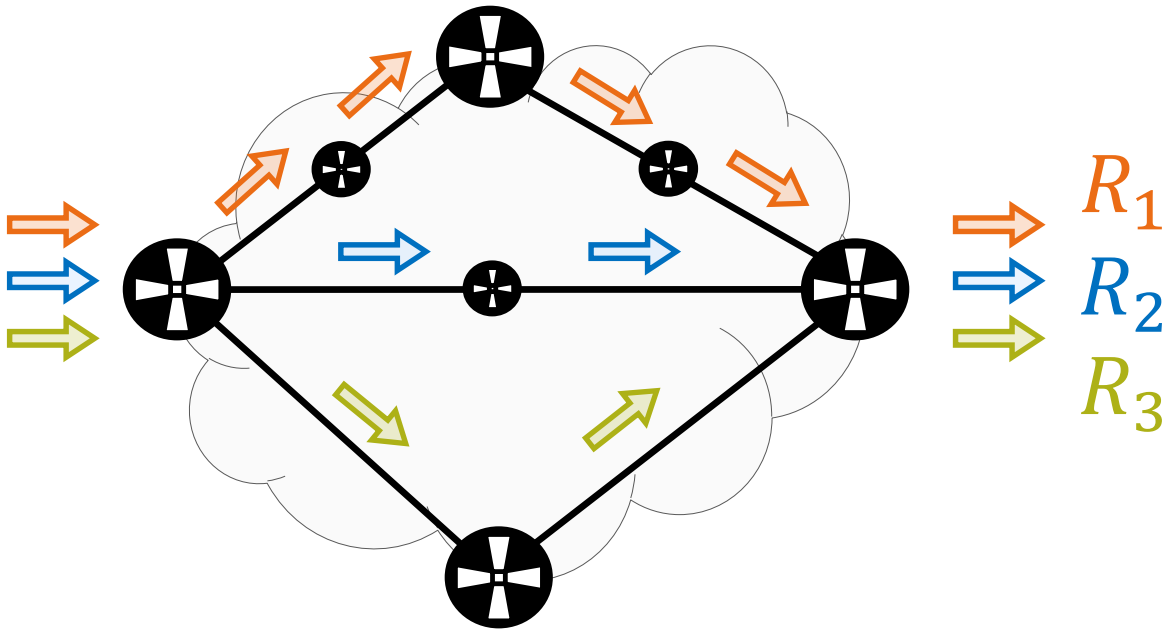
When **LB is deployed** only FDs ⇒ multipath routing holds
This occurs since LB ⇒ multipath routing

When LB is deployed, is the multipath observed LB or FDs?



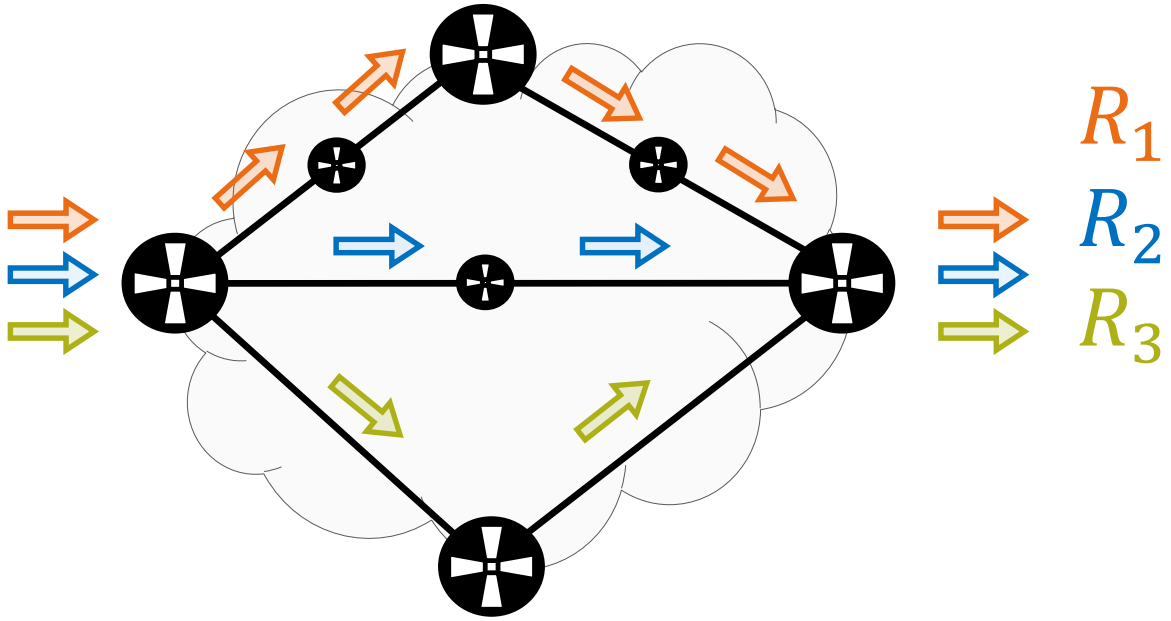
When **LB is deployed** only FDs ⇒ multipath routing holds
This occurs since LB ⇒ multipath routing

When LB is deployed, is the multipath observed LB or FDs?



We study the **forwarding pattern** across (i,e)

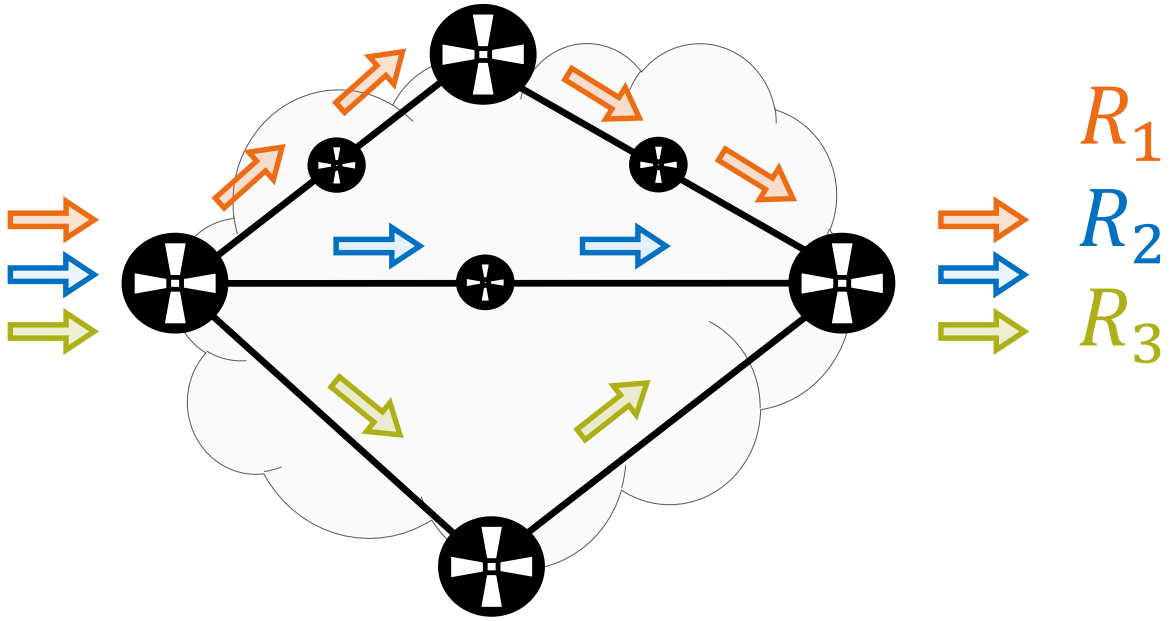
Forwarding pattern across (i,e)



$\mathbb{R}(i,e) \rightarrow$ routes between (i,e)

$\mathbb{R}(i,e) = \{R_1, R_2, R_3\}$

Forwarding pattern across (i,e)



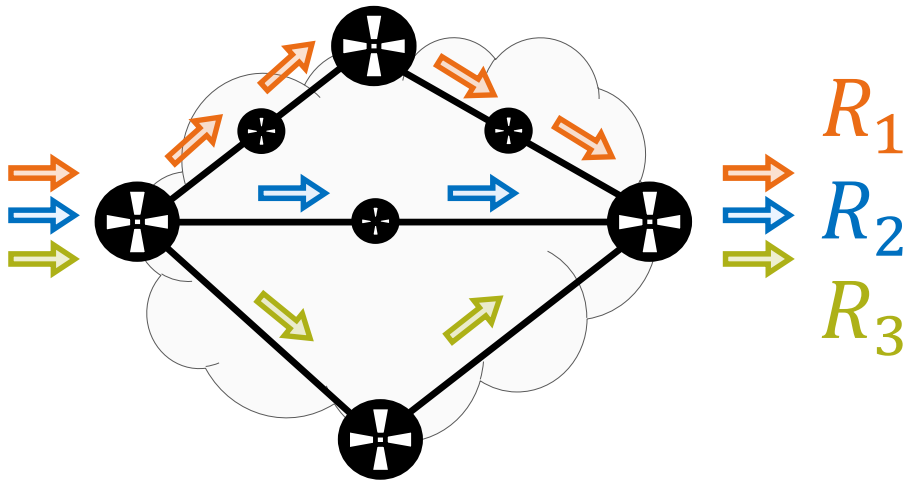
$\mathbb{R}(i, e) \rightarrow$ routes between (i, e)

$\mathbb{R}(i, e) = \{R_1, R_2, R_3\}$

We want to learn **which routes** are used for **which prefixes**

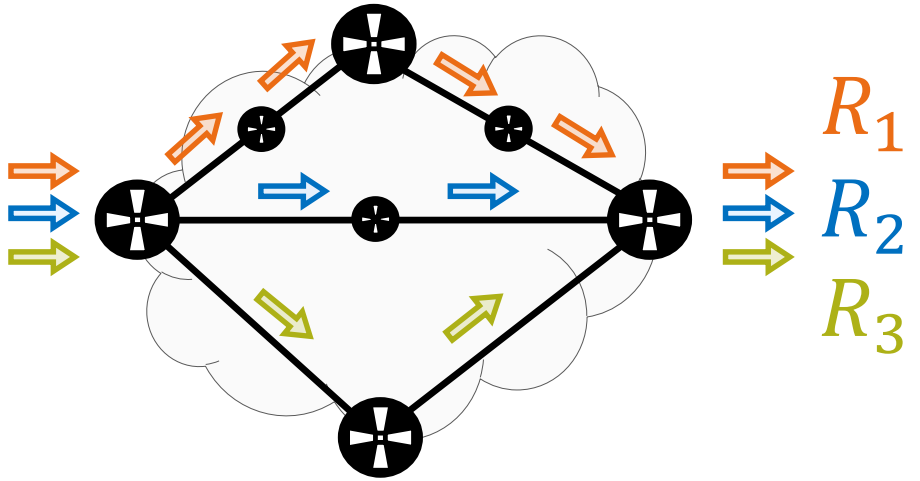
$\mathbb{R}(\mathbb{P}_i) \rightarrow$ set of routes seen for a set of prefixes $\mathbb{P}_i$

- Run one per /24 and identify prefixes $\mathbb{P}_i$ that use the same route

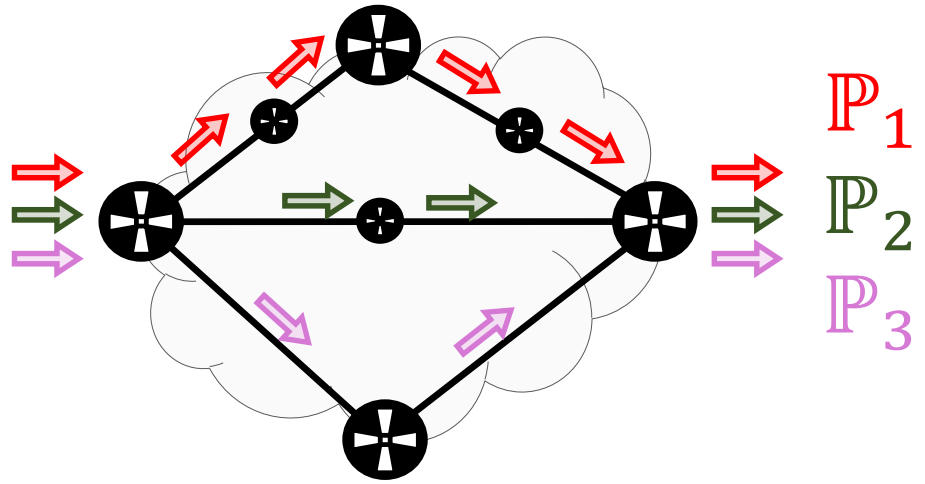


$$\mathbb{R}(i, e) = \{R_1, R_2, R_3\}$$

- Run one per /24 and identify prefixes $\mathbb{P}_i$ that use the same route
- For each $\mathbb{P}_i$ choose destinations and try to find more routes



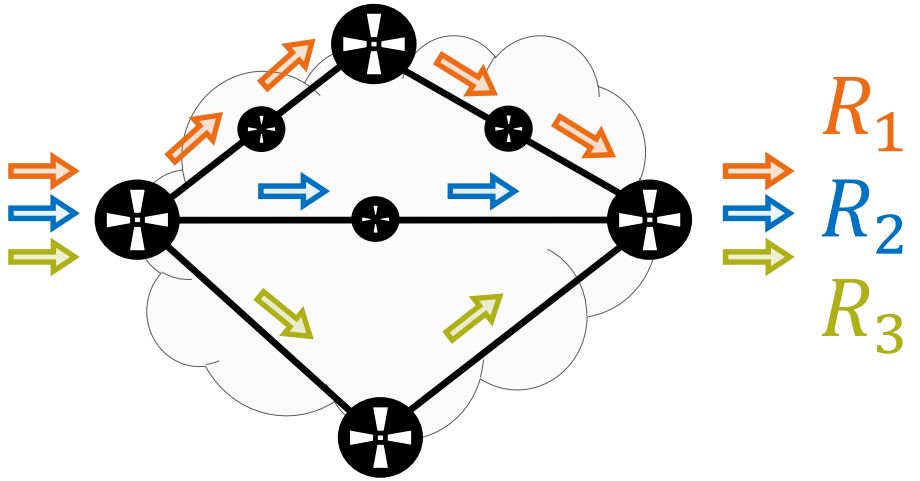
$$\mathbb{R}(i, e) = \{R_1, R_2, R_3\}$$



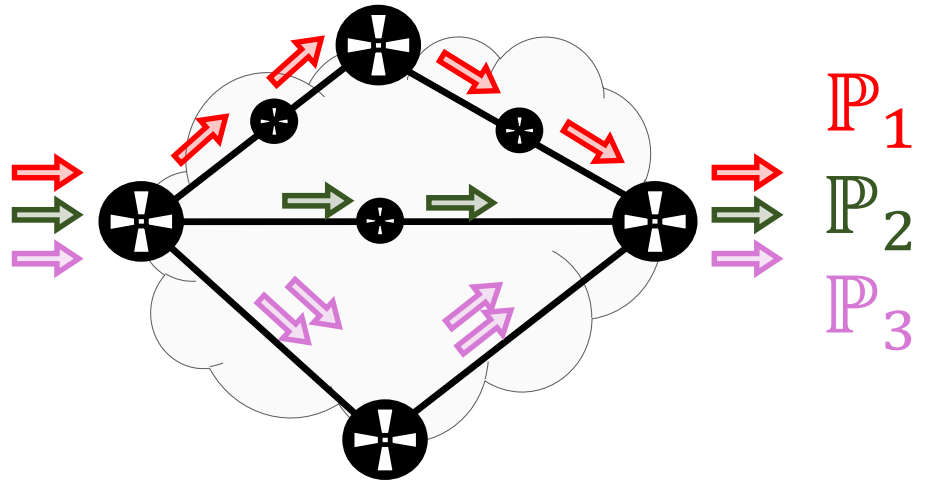
$$\widehat{\mathbb{R}(i, e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\mathbb{P}_2) \cup \mathbb{R}(\mathbb{P}_3)$$

$$\mathbb{R}(\mathbb{P}_1) = R_1$$
$$\mathbb{R}(\mathbb{P}_2) = R_2$$
$$\mathbb{R}(\mathbb{P}_3) = R_3$$

- Run one per /24 and identify prefixes $\mathbb{P}_i$ that use the same route
- For each $\mathbb{P}_i$ choose destinations and try to find more routes



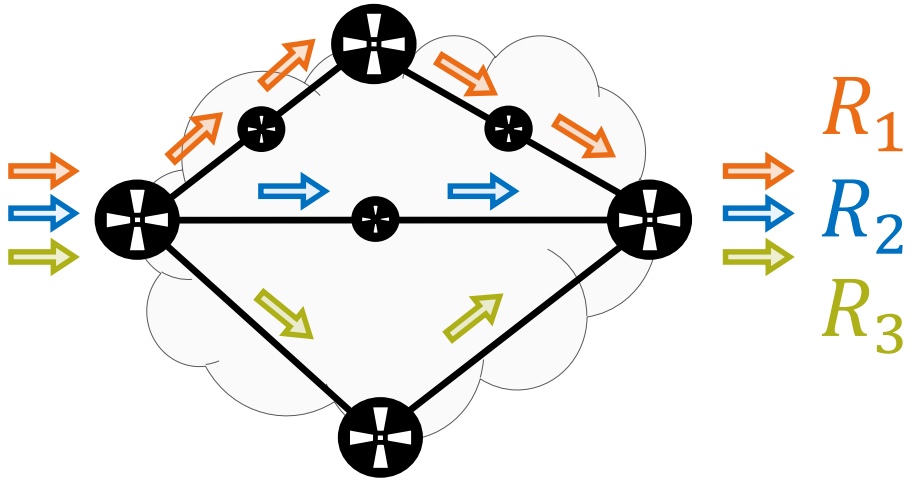
$$\mathbb{R}(i, e) = \{R_1, R_2, R_3\}$$



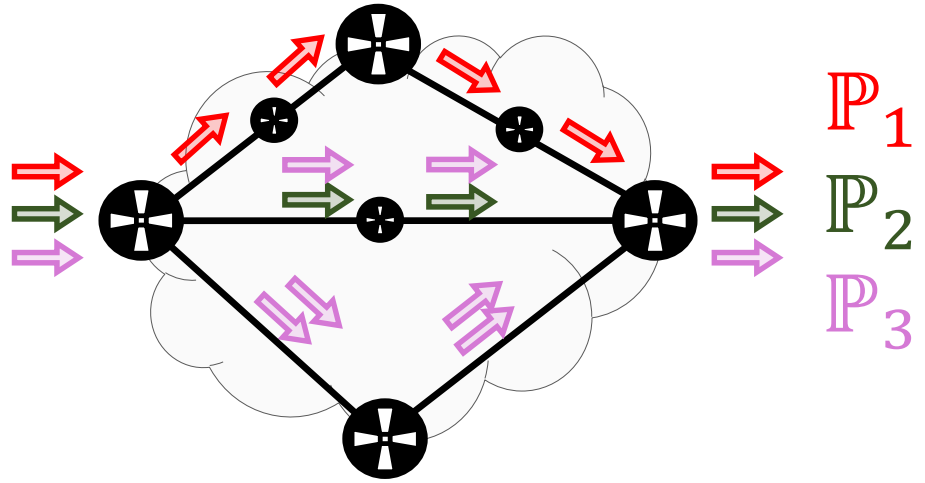
$$\widehat{\mathbb{R}(i, e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\mathbb{P}_2) \cup \mathbb{R}(\mathbb{P}_3)$$

$$\mathbb{R}(\mathbb{P}_1) = R_1$$
$$\mathbb{R}(\mathbb{P}_2) = R_2$$
$$\mathbb{R}(\mathbb{P}_3) = R_3$$

- Run one per /24 and identify prefixes $\mathbb{P}_i$ that use the same route
- For each $\mathbb{P}_i$ choose destinations and try to find more routes



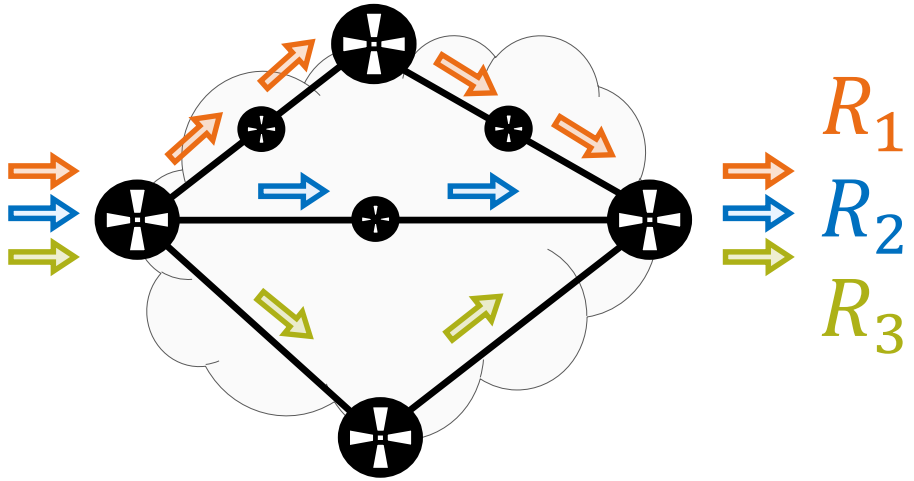
$$\mathbb{R}(i,e) = \{R_1, R_2, R_3\}$$



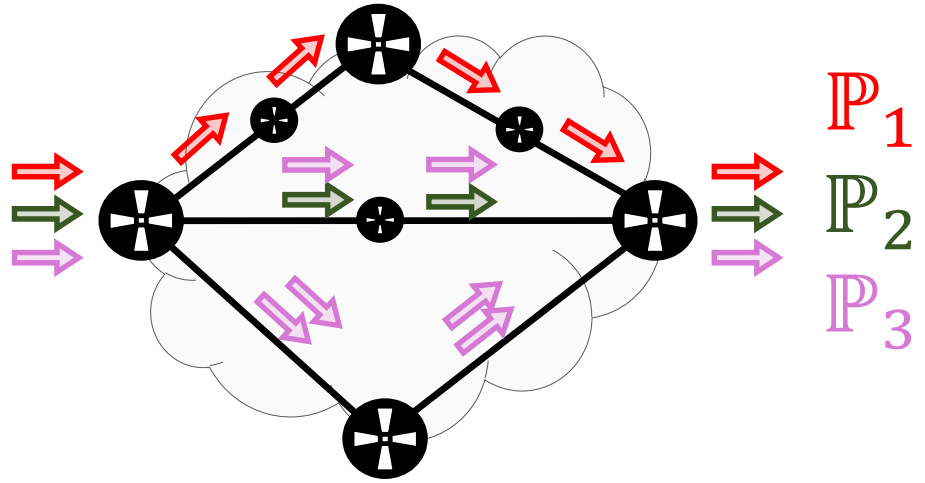
$$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\mathbb{P}_2) \cup \mathbb{R}(\mathbb{P}_3)$$

$$\mathbb{R}(\mathbb{P}_1) = R_1$$
$$\mathbb{R}(\mathbb{P}_2) = R_2$$
$$\mathbb{R}(\mathbb{P}_3) = \{R_3, R_2\}$$

- Run one per /24 and identify prefixes $\mathbb{P}_i$ that use the same route
- For each $\mathbb{P}_i$ choose destinations and try to find more routes
- When sets of routes for different $\mathbb{P}_i$ intersect, we **merge** both routes and prefixes



$$\mathbb{R}(i, e) = \{R_1, R_2, R_3\}$$

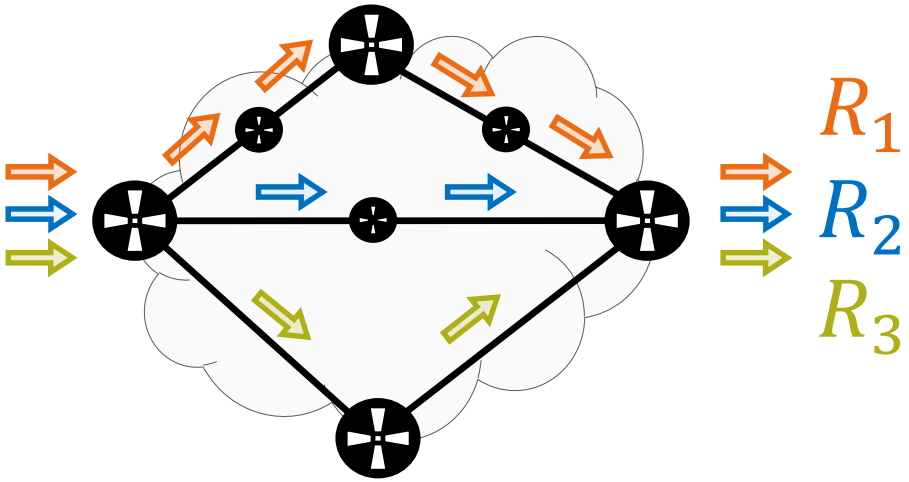


$$\widehat{\mathbb{R}(i, e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\{\mathbb{P}_2 \cup \mathbb{P}_3\})$$

$$\mathbb{R}(\mathbb{P}_1) = R_1$$
$$\mathbb{R}(\{\mathbb{P}_2 \cup \mathbb{P}_3\}) = \{R_2, R_3\}$$

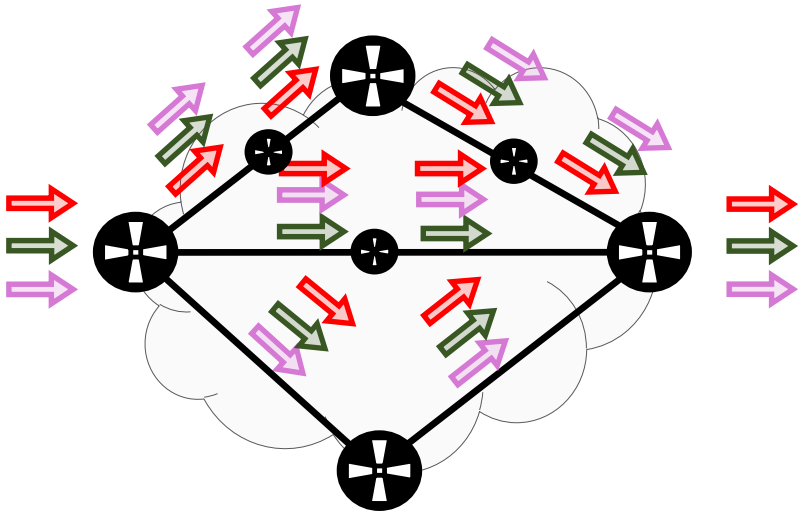
Results of the route discovery and merging phases

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$



R_1
R_2
R_3

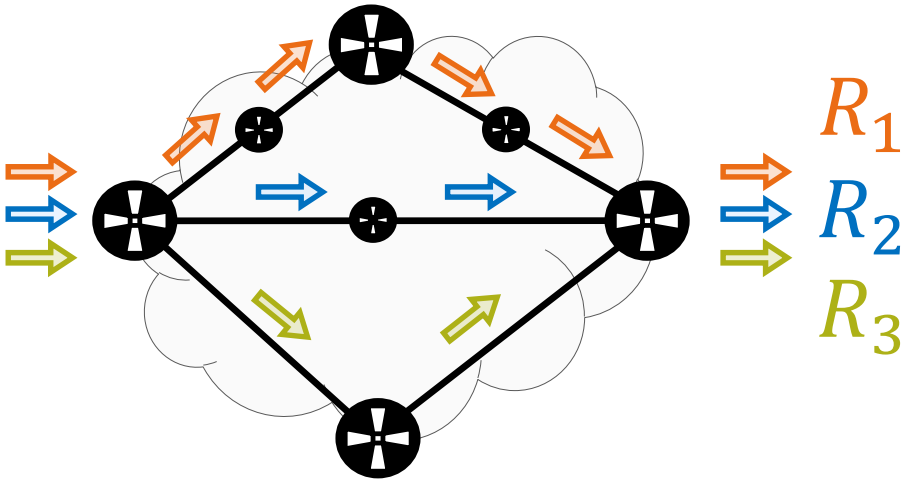
LB Case



$\mathbb{P}_1$
$\mathbb{P}_2$
$\mathbb{P}_3$

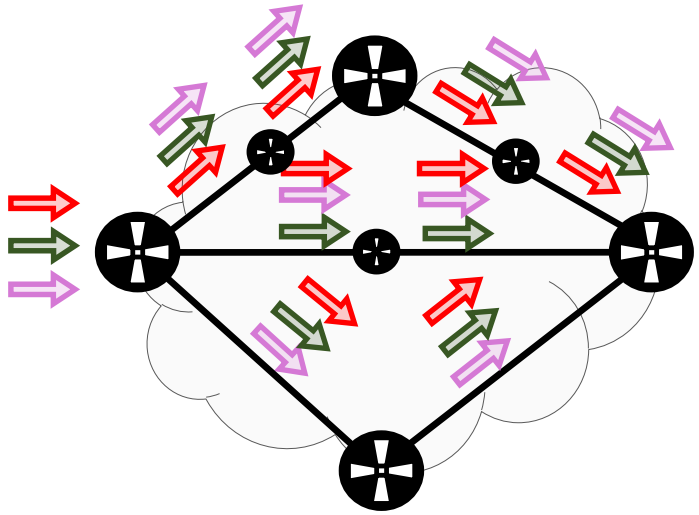
Results of the route discovery and merging phases

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$



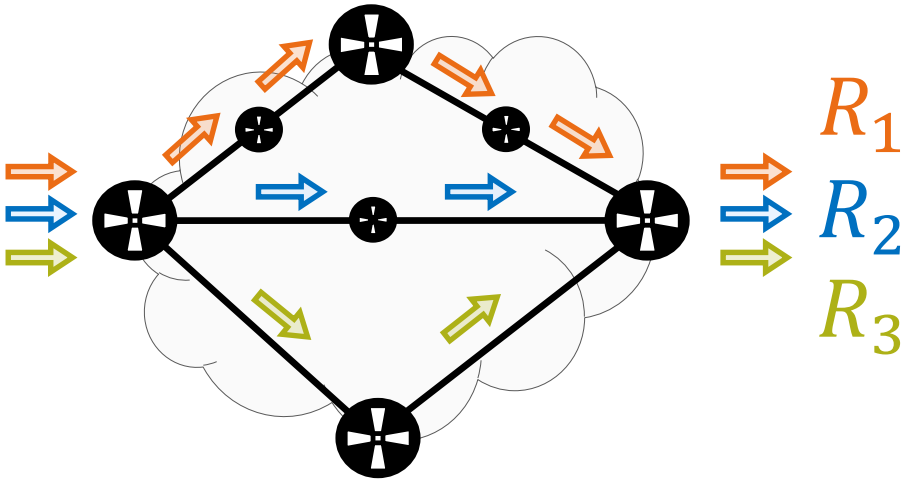
R_1
R_2
R_3

LB Case



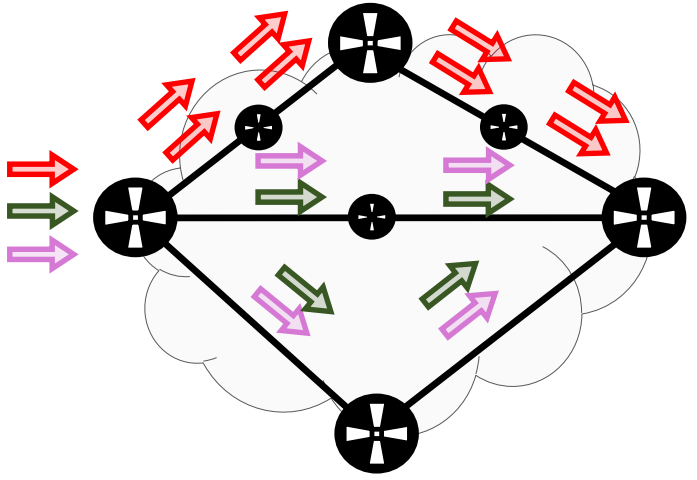
$\mathbb{P}_1$
$\mathbb{P}_2$
$\mathbb{P}_3$

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\{\mathbb{P}_2 \cup \mathbb{P}_3\})$



R_1
R_2
R_3

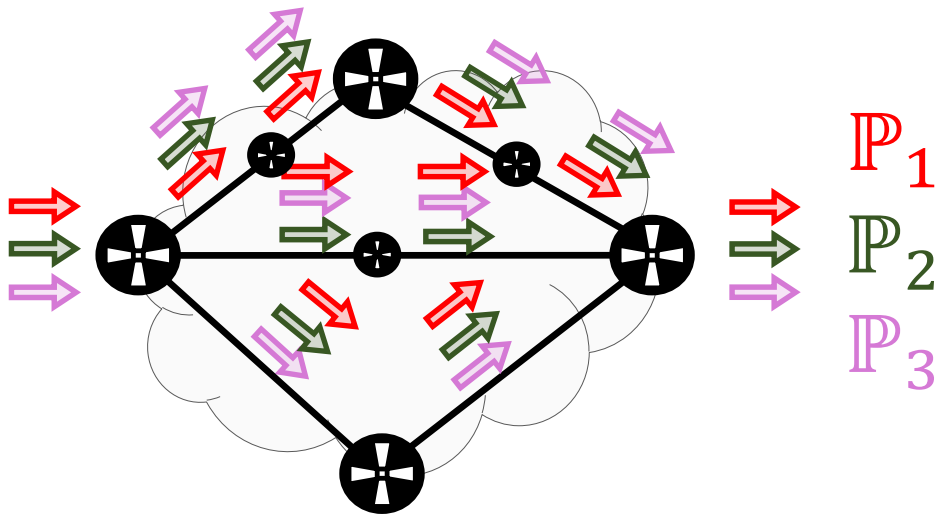
FDs Case



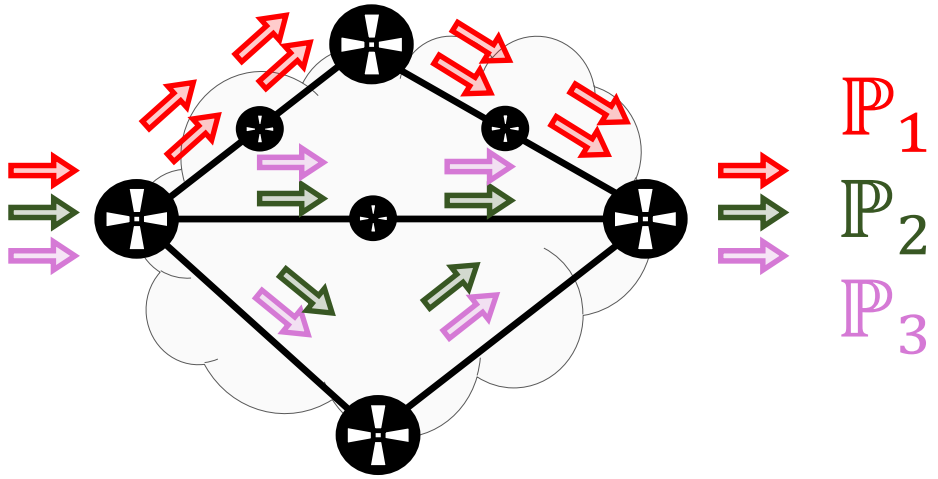
$\mathbb{P}_1$
$\mathbb{P}_2$
$\mathbb{P}_3$

Results of the route discovery and merging phases

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$

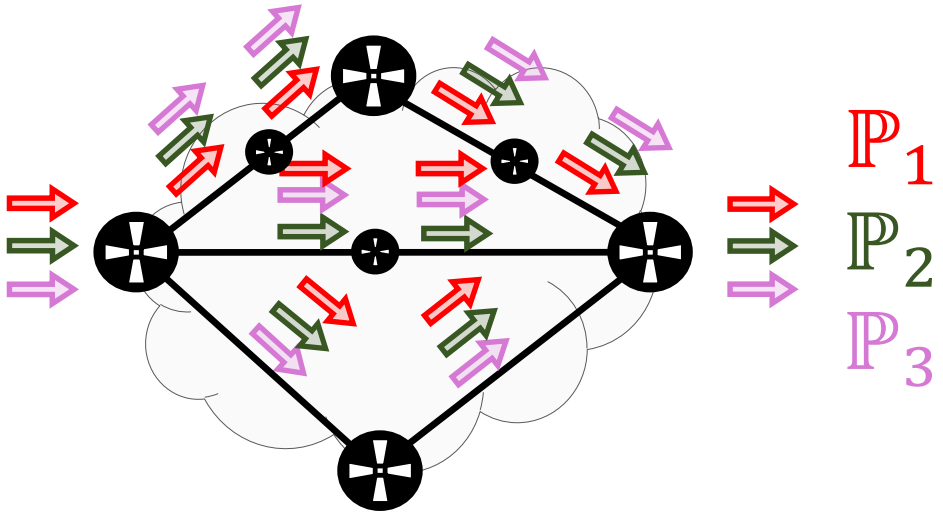


$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\{\mathbb{P}_2 \cup \mathbb{P}_3\})$

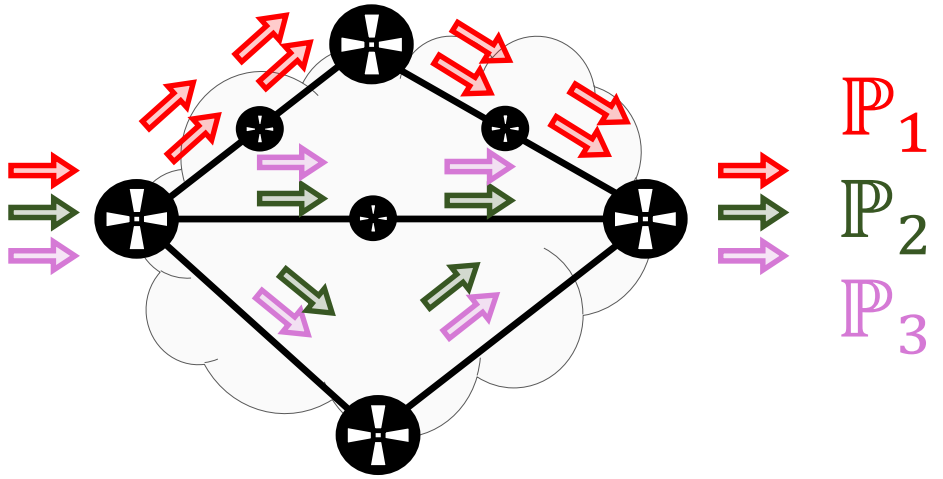


Results of the route discovery and merging phases

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$



$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\{\mathbb{P}_2 \cup \mathbb{P}_3\})$



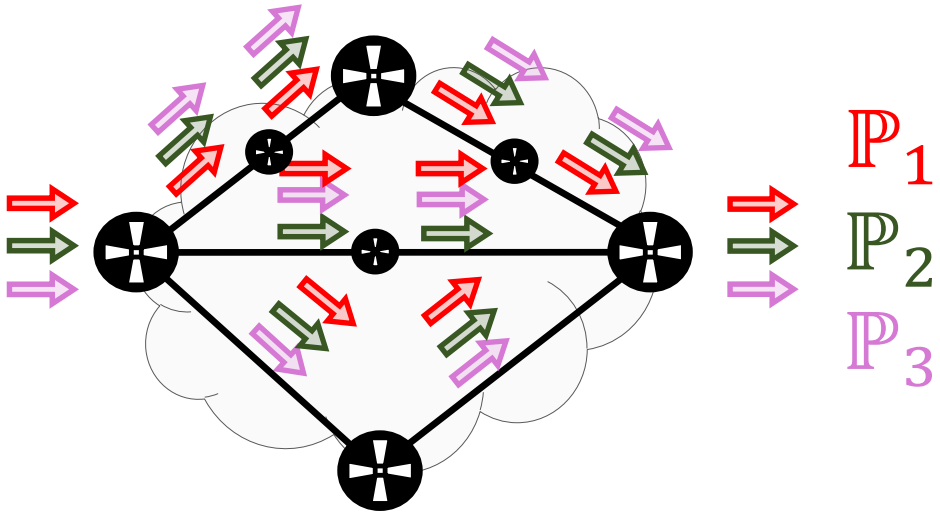
s → # sets of prefixes / blocks conforming $\widehat{\mathbb{R}(i,e)}$

s = 1 ⇔ LB

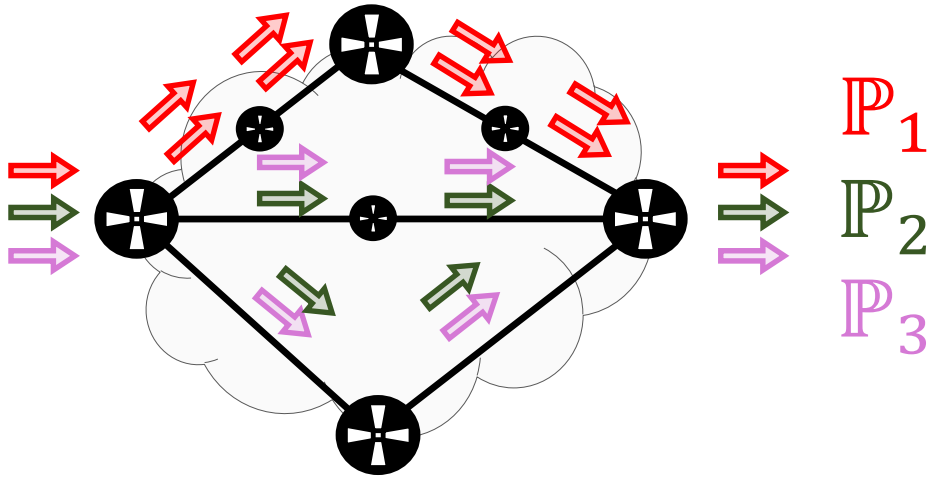
s ≥ 2 ⇔ FDs

Results of the route discovery and merging phases

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$



$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\{\mathbb{P}_2 \cup \mathbb{P}_3\})$



s → # sets of prefixes / blocks conforming $\widehat{\mathbb{R}(i,e)}$

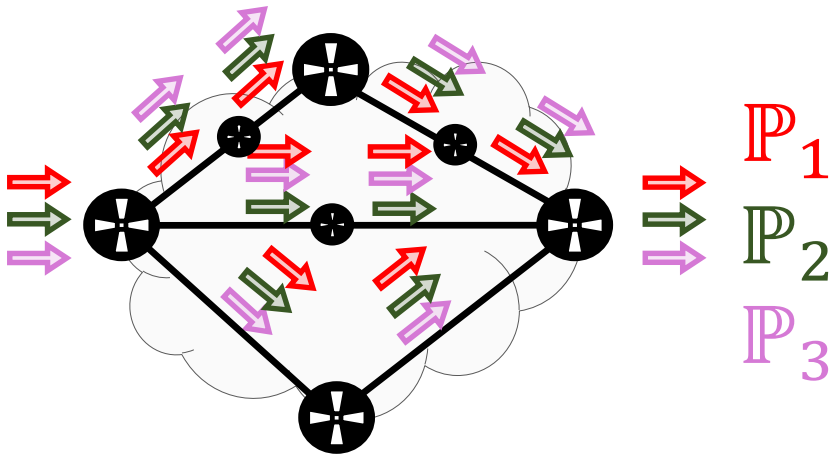
s = 1 ⇔ LB

s ≥ 2 ⇔ FDs

However… **not all LB flavors** produce the **s = 1** pattern

Fine Grained LB (F-LB)
e.g.: per-dest, per-flow, per-app
s = 1

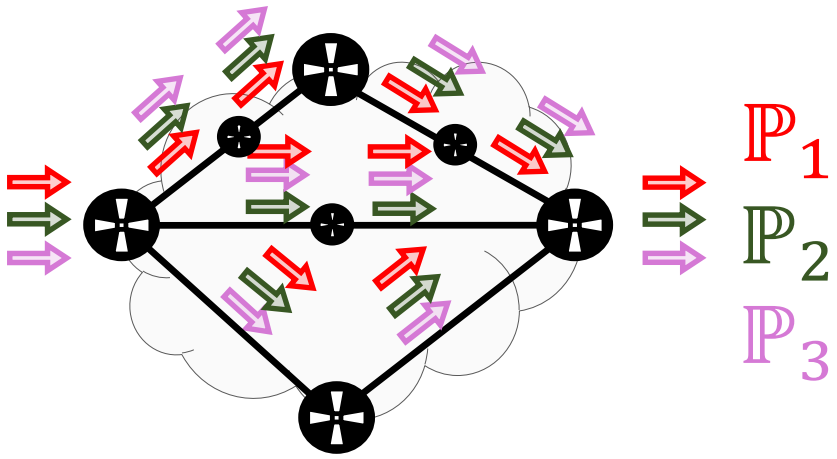
$$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$$



Some LB flavor **reserve routes** for certain **prefixes**

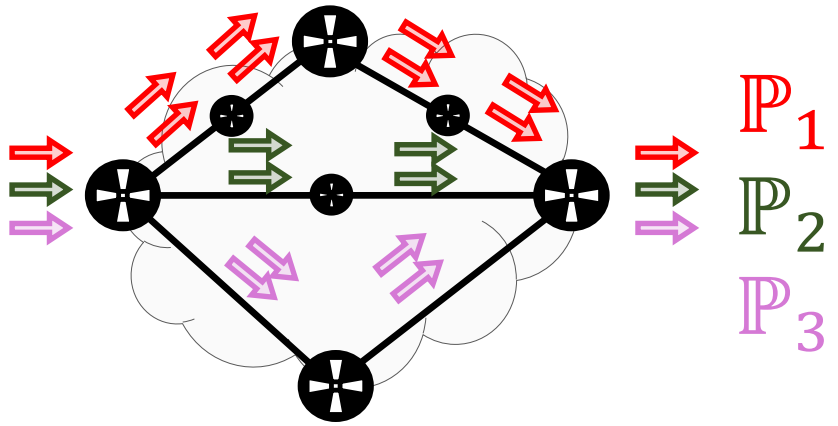
Fine Grained LB (F-LB)
e.g.: per-dest, per-flow, per-app
s = 1

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$



Coarse Grained LB (C-LB)
e.g.: per-prefix
s = $|\mathbb{R}(i,e)|$ > 1

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\mathbb{P}_2) \cup \mathbb{R}(\mathbb{P}_3)$

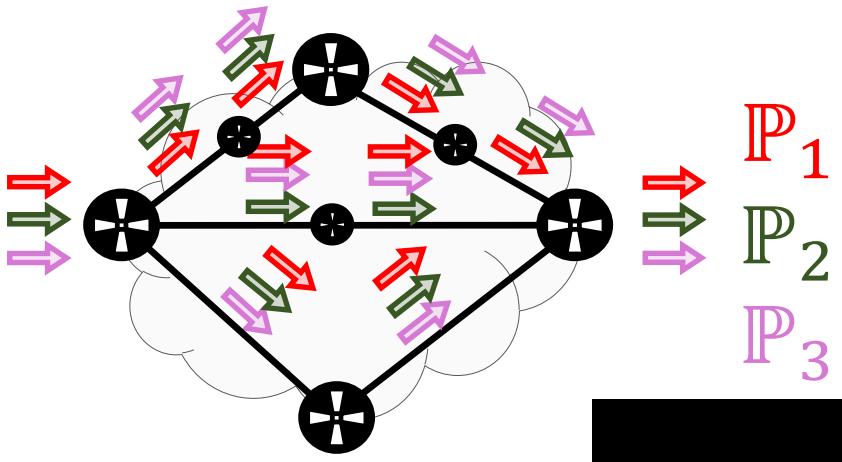


Some LB flavors **reserve routes** for certain **prefixes**

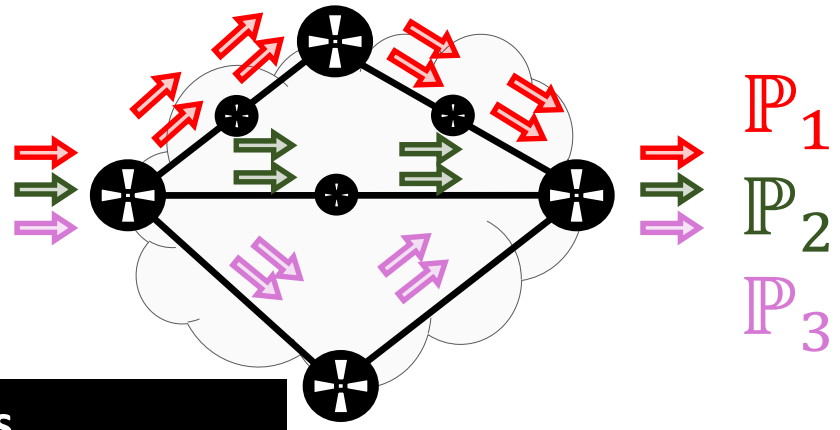
Fine Grained LB (F-LB)
e.g.: per-dest, per-flow, per-app
s = 1

Coarse Grained LB (C-LB)
e.g.: per-prefix
s = $|\mathbb{R}(i,e)|$ > 1

$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\{\mathbb{P}_1 \cup \mathbb{P}_2 \cup \mathbb{P}_3\})$



$\widehat{\mathbb{R}(i,e)} = \mathbb{R}(\mathbb{P}_1) \cup \mathbb{R}(\mathbb{P}_2) \cup \mathbb{R}(\mathbb{P}_3)$



s = 1 ⇔ No FDs
No LB or F-LB
s ≥ 2 ⇒ ???
Either C-LB, FDs or C-LB + FDs
We need further analysis

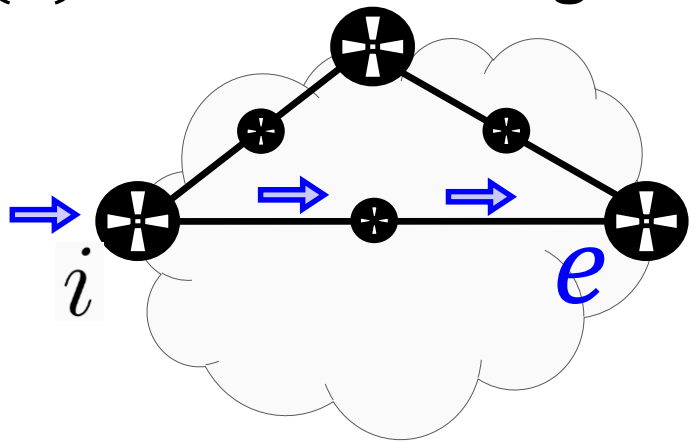
- Assume **IGP prefixes are not subject to RIes and FDs**
  - The IPs of the egress-ASBR belong to the IGP

$ASBR_2$		
Pfx	BGP NH	Int
P_R	$self$	i_0
P_G	$ASBR_3$	i_1
P_B	$ASBR_3$	i_1
0/0	$self$	i_0

ASBR$_2$		
Pfx	BGP NH	Int
P_R	$self$	i_0
P_G	$ASBR_3$	i_1
P_B	$ASBR_3$	i_1
0/0	$self$	i_0

- Assume **IGP prefixes are not subject to RIes and FDs**
  - The IPs of the egress-ASBR belong to the IGP

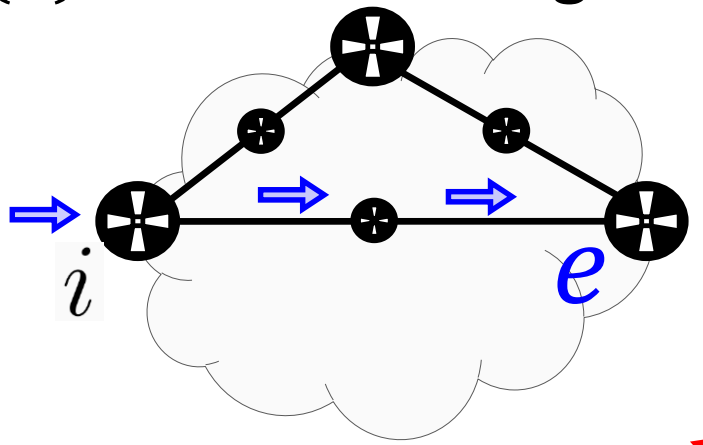
- Best IGP path == Direct trace $R_X(e)$ towards the egress-ASBR
  - $\boldsymbol{R_X(e)}$ **does not detour**
  - Note that $R_X(e)$ must cross i



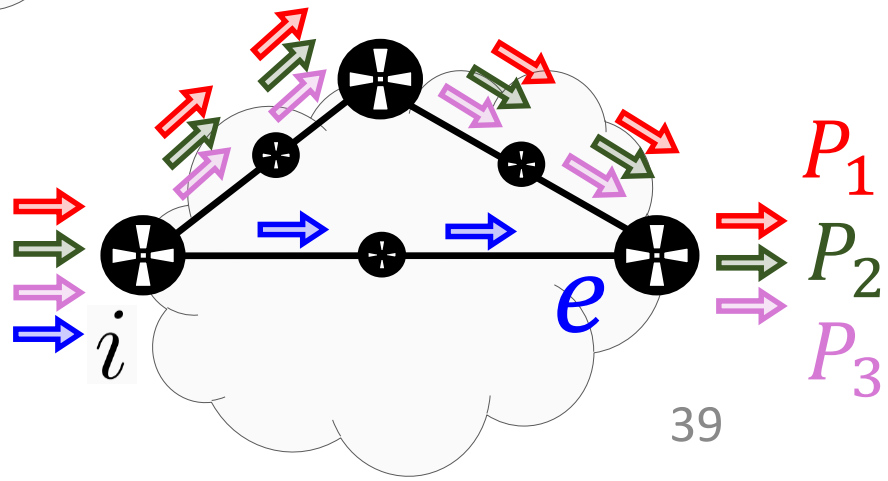
- Assume **IGP prefixes are not subject to RIes and FDs**
  - The IPs of the egress-ASBR belong to the IGP

$ASBR_2$		
Pfx	BGP NH	Int
P_R	$self$	i_0
P_G	$ASBR_3$	i_1
P_B	$ASBR_3$	i_1
0/0	$self$	i_0

- Best IGP path == Direct trace $R_X(e)$ towards the egress-ASBR
  - $\boldsymbol{R_X(e)}$ **does not detour**
  - Note that $R_X(e)$ must cross i



- Since $R_X(e)$ does not detour…
  - **Neither does** a block of routes including it!



If **all prefixes** flow through a block of routes **not including $R_X(e)$**, then **FDs occur**

Setup

100 nodes from NLNOG RING
heterogeneously distributed

100K IP Internet Address Hitlist

Xun Fan and John Heidemann. 2010.
Selecting representative IP addresses for Internet topology studies.
In Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM, 411–423

Filter every couple for which less than 100 prefixes are collected

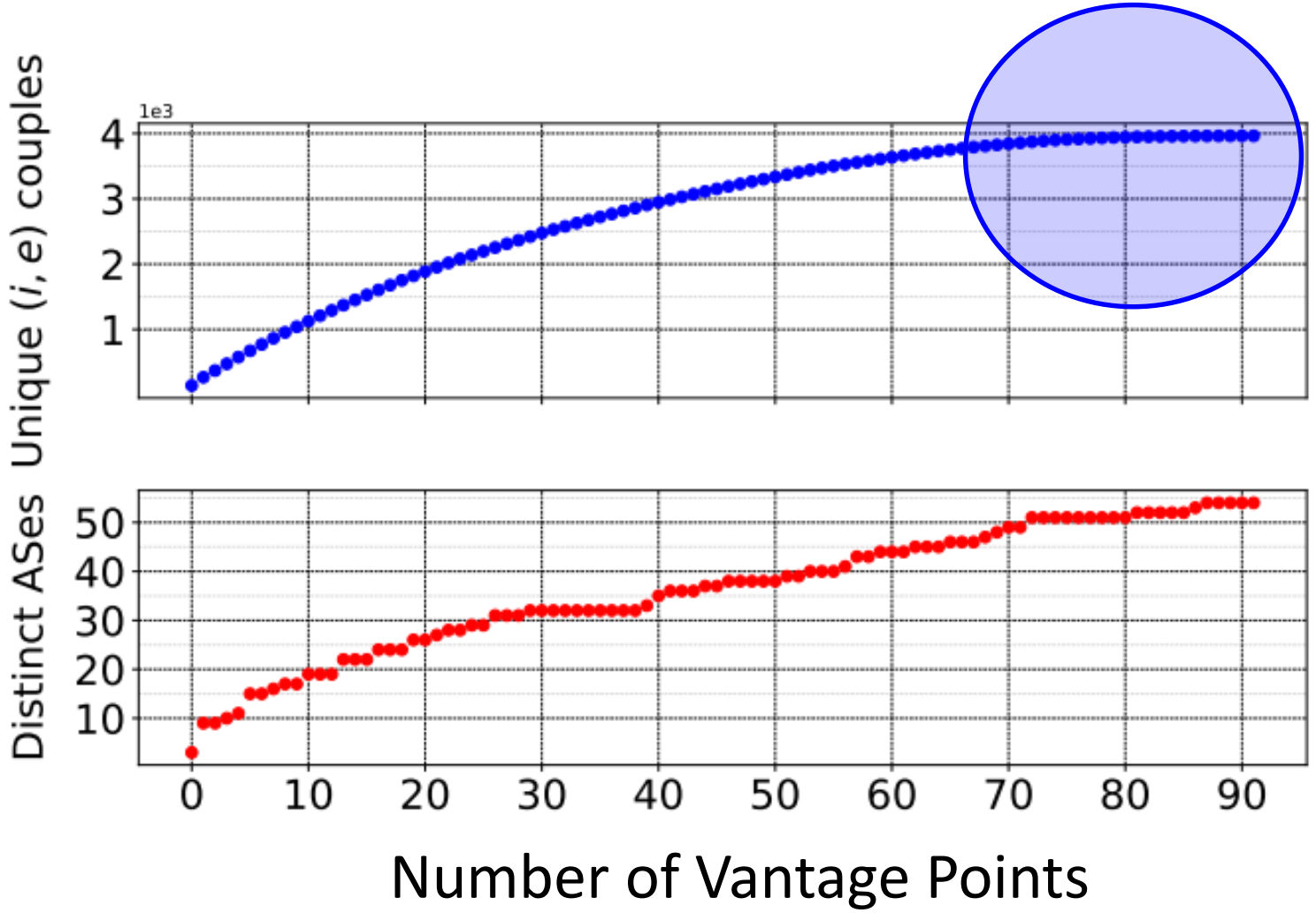
Setup

100 nodes from NLNOG RING **heterogeneously** distributed

100K IP Internet Address Hitlist

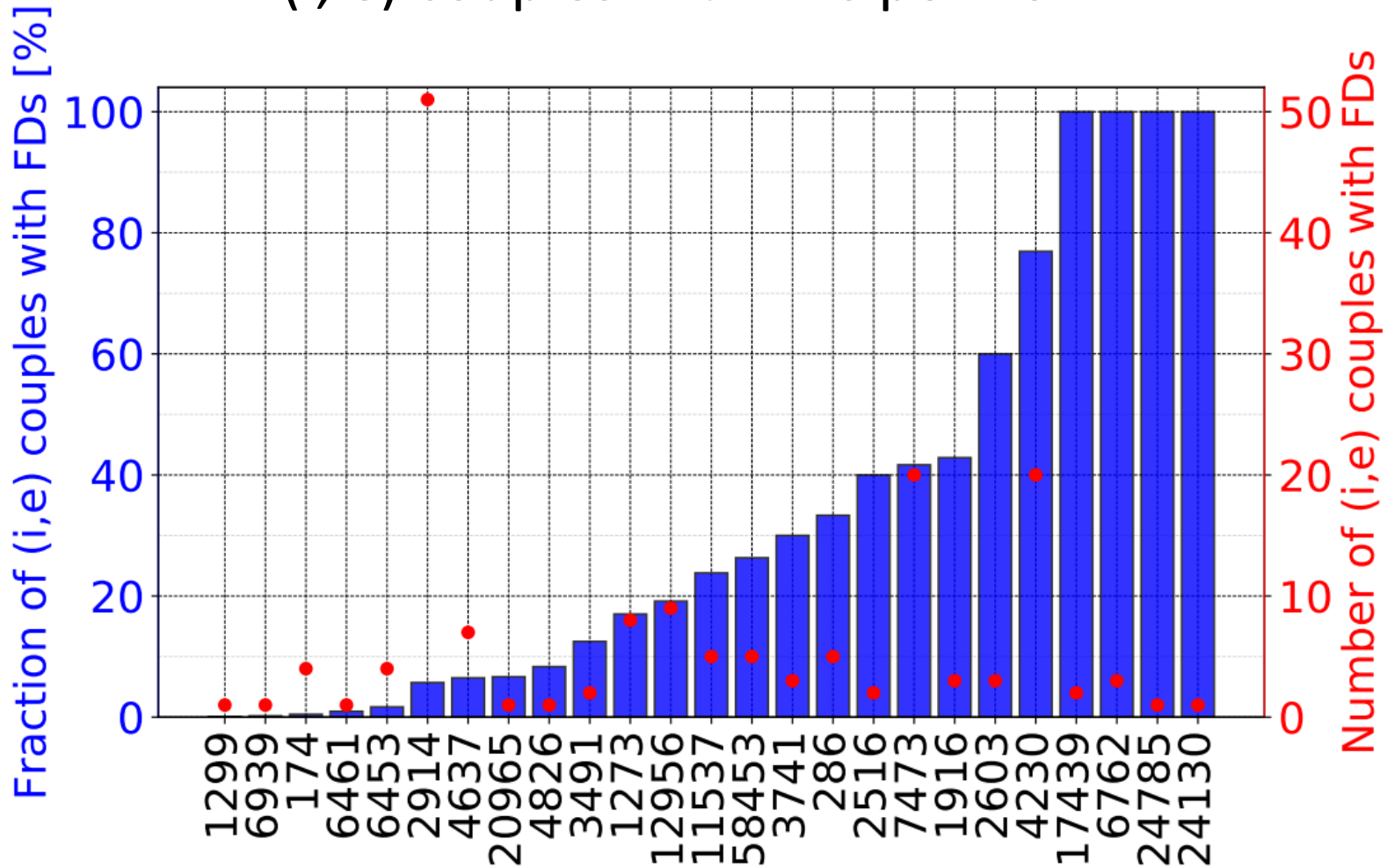
Xun Fan and John Heidemann. 2010.
Selecting representative IP addresses for Internet topology studies.
In Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM, 411–423

Filter every couple for which less than 100 prefixes are collected



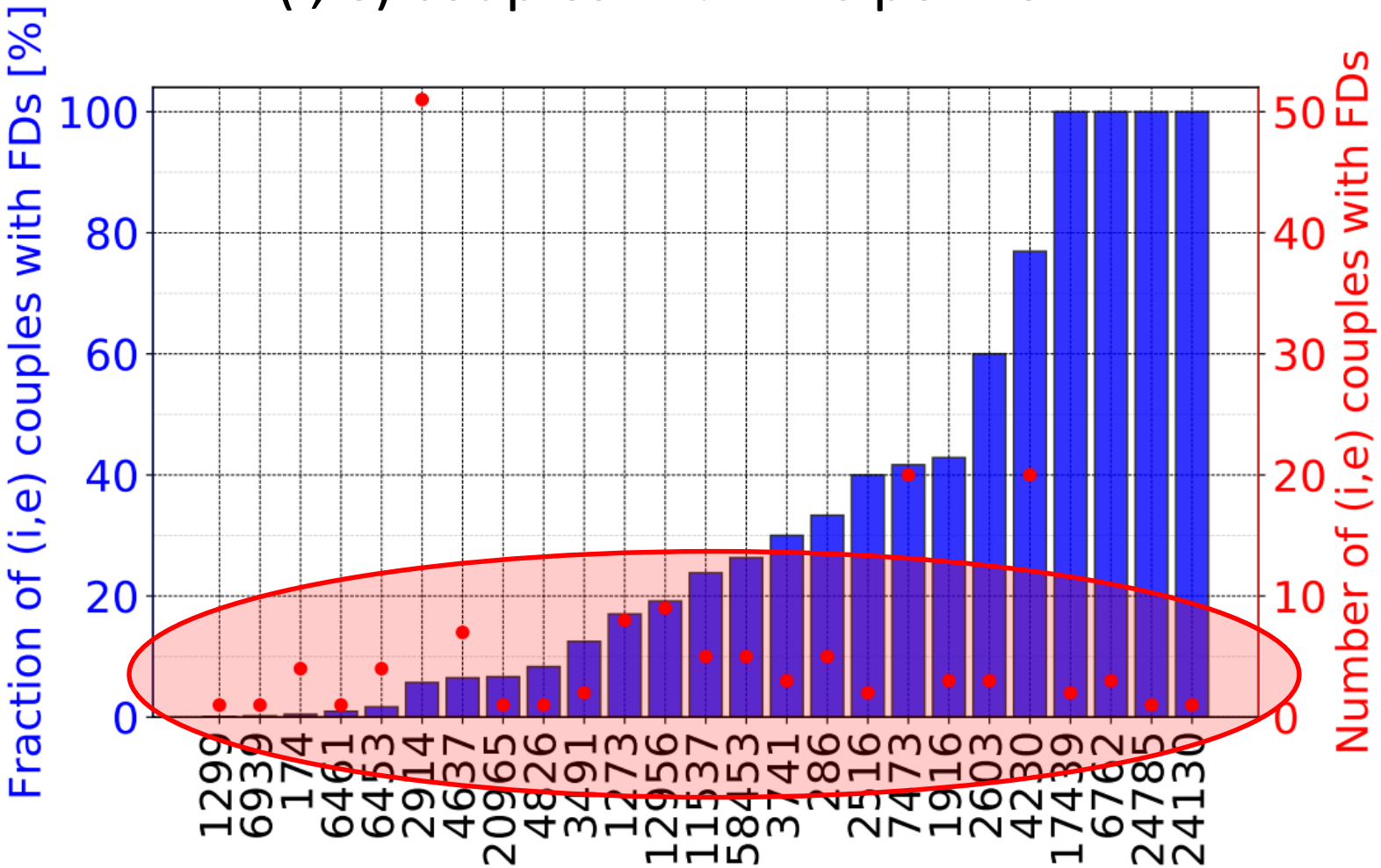
- General view
  - 54 ASes
  - ~ 4000 (i, e) couples
  - Marginal utility at the (i, e) level **negligible after 70 VPs**

(i, e) couples with FDs per AS



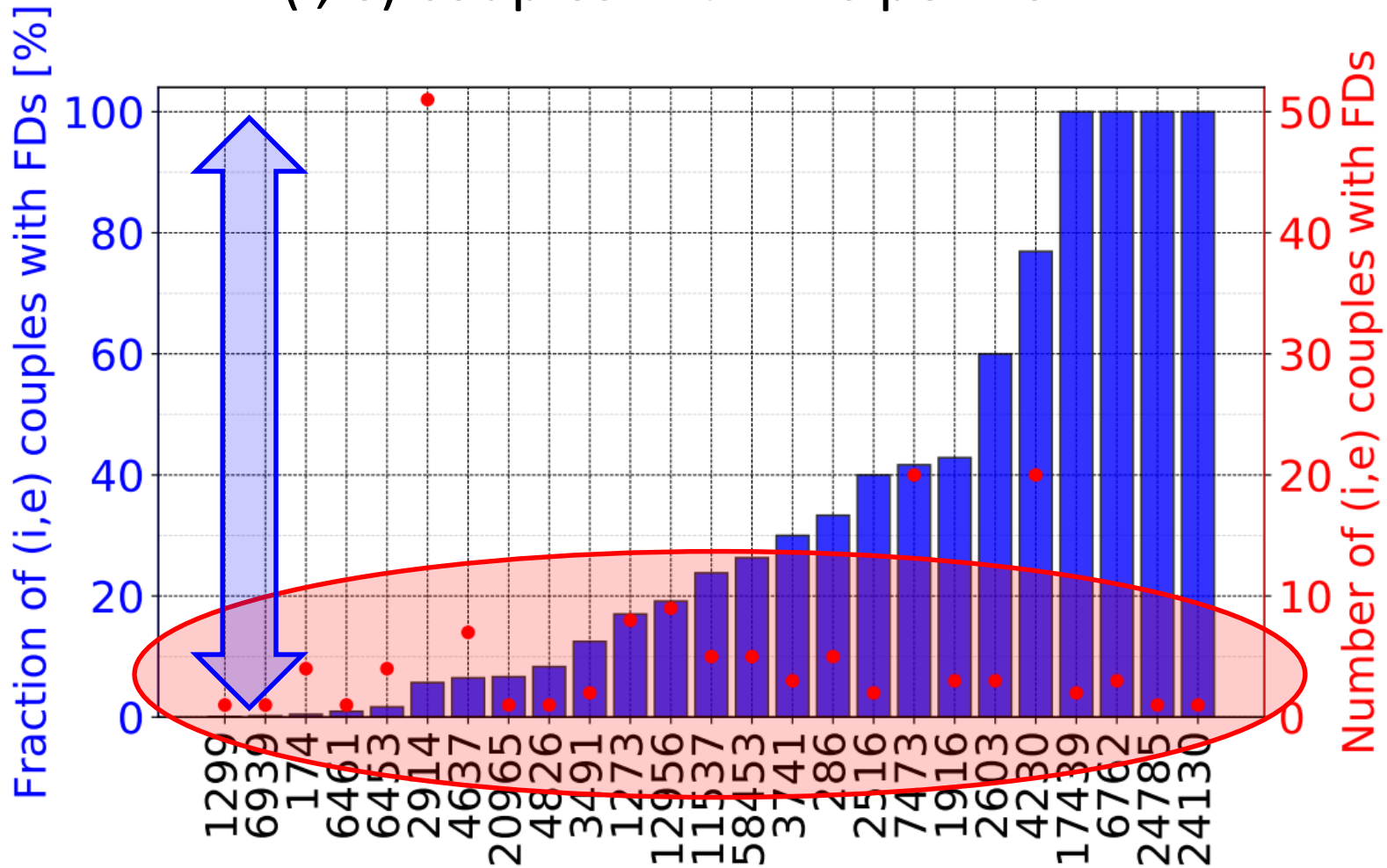
- 168 couples distributed in 25 (45%) ASes show FDs

(i, e) couples with FDs per AS



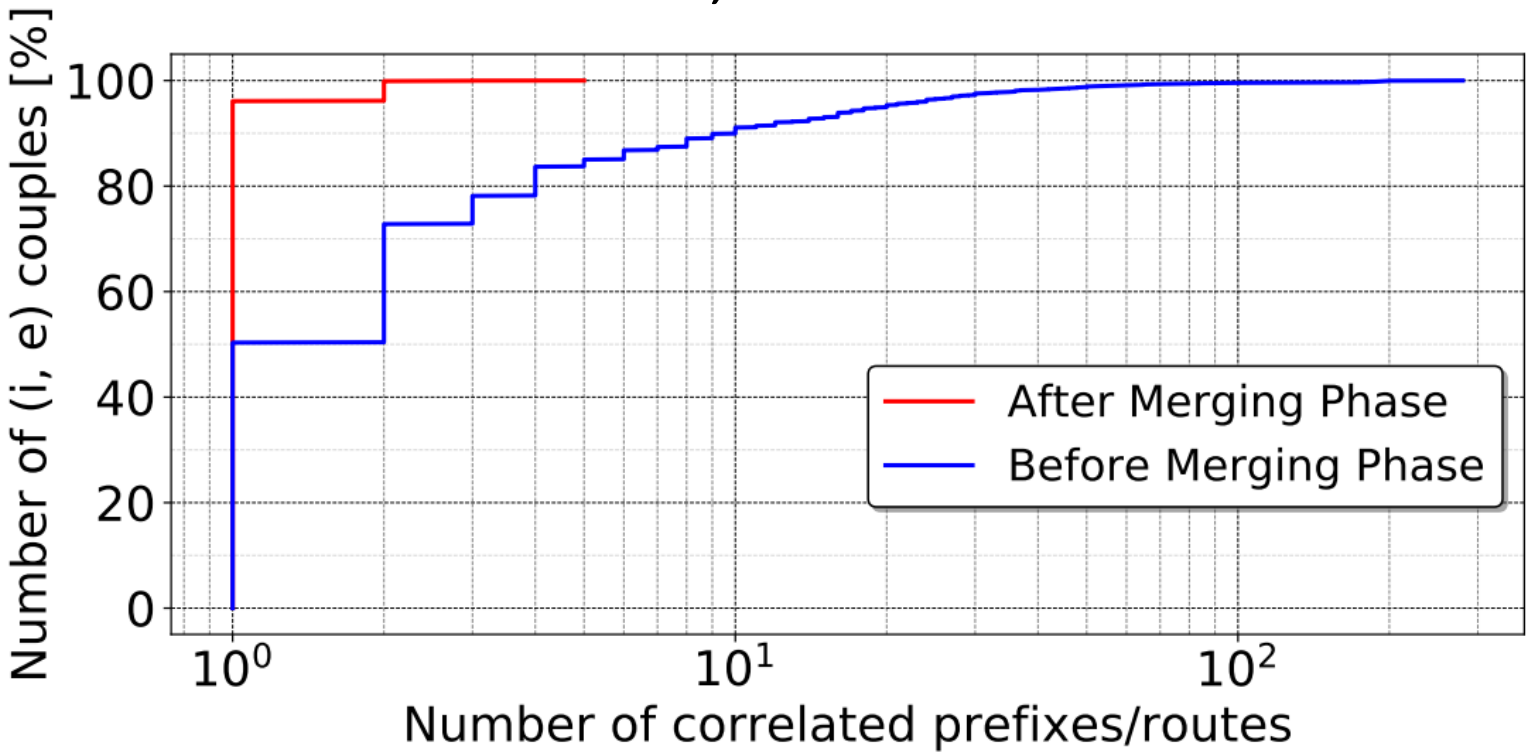
- 168 couples distributed in 25 (45%) ASes show FDs
- In general, ASes have **few couples** with FDs

(i, e) couples with FDs per AS

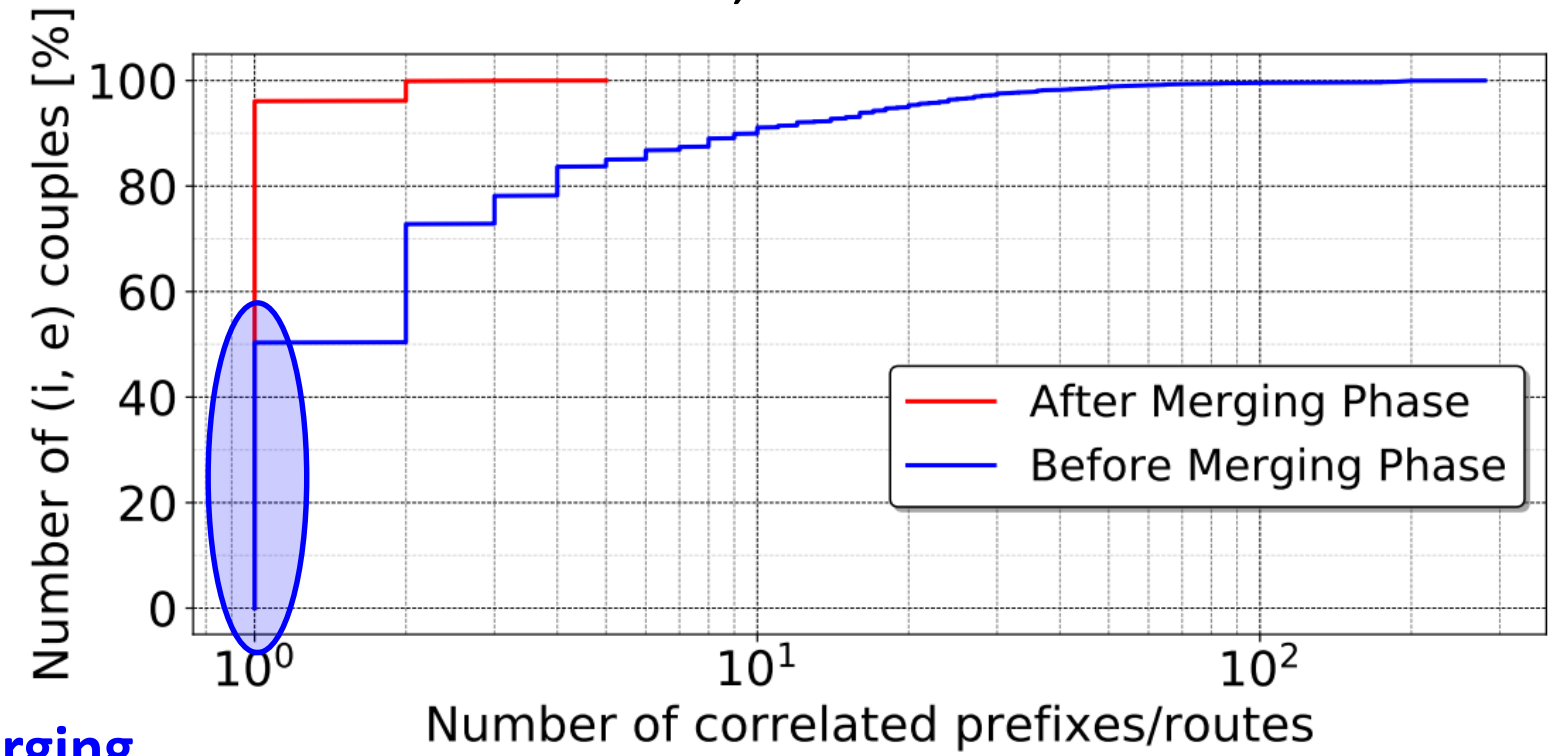


- 168 couples distributed in 25 (45%) ASes show FDs
- In general, ASes have **few couples** with FDs
- **FDs are AS-specific**

Do we see F-LB, C-LB besides FDs?

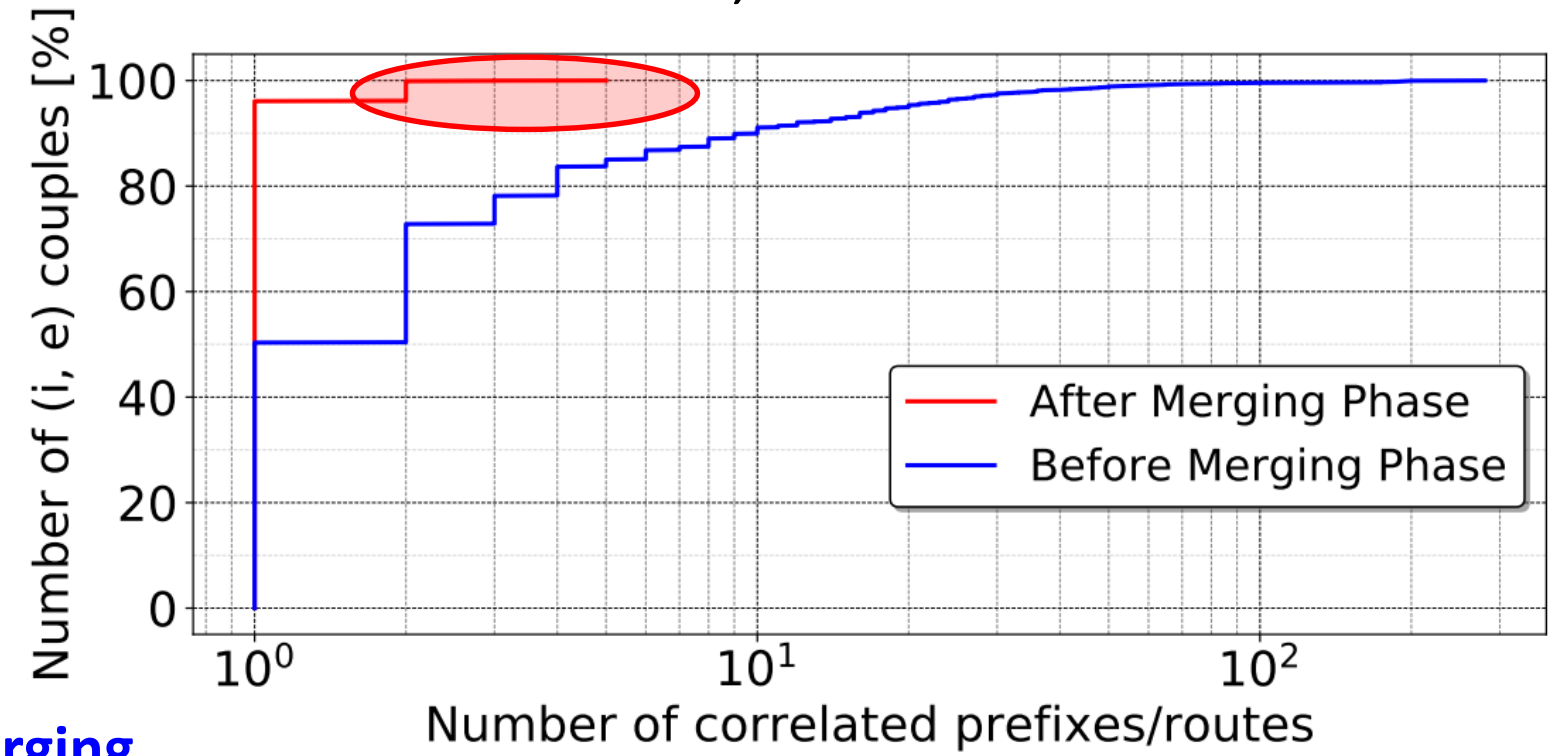


Do we see F-LB, C-LB besides FDs?



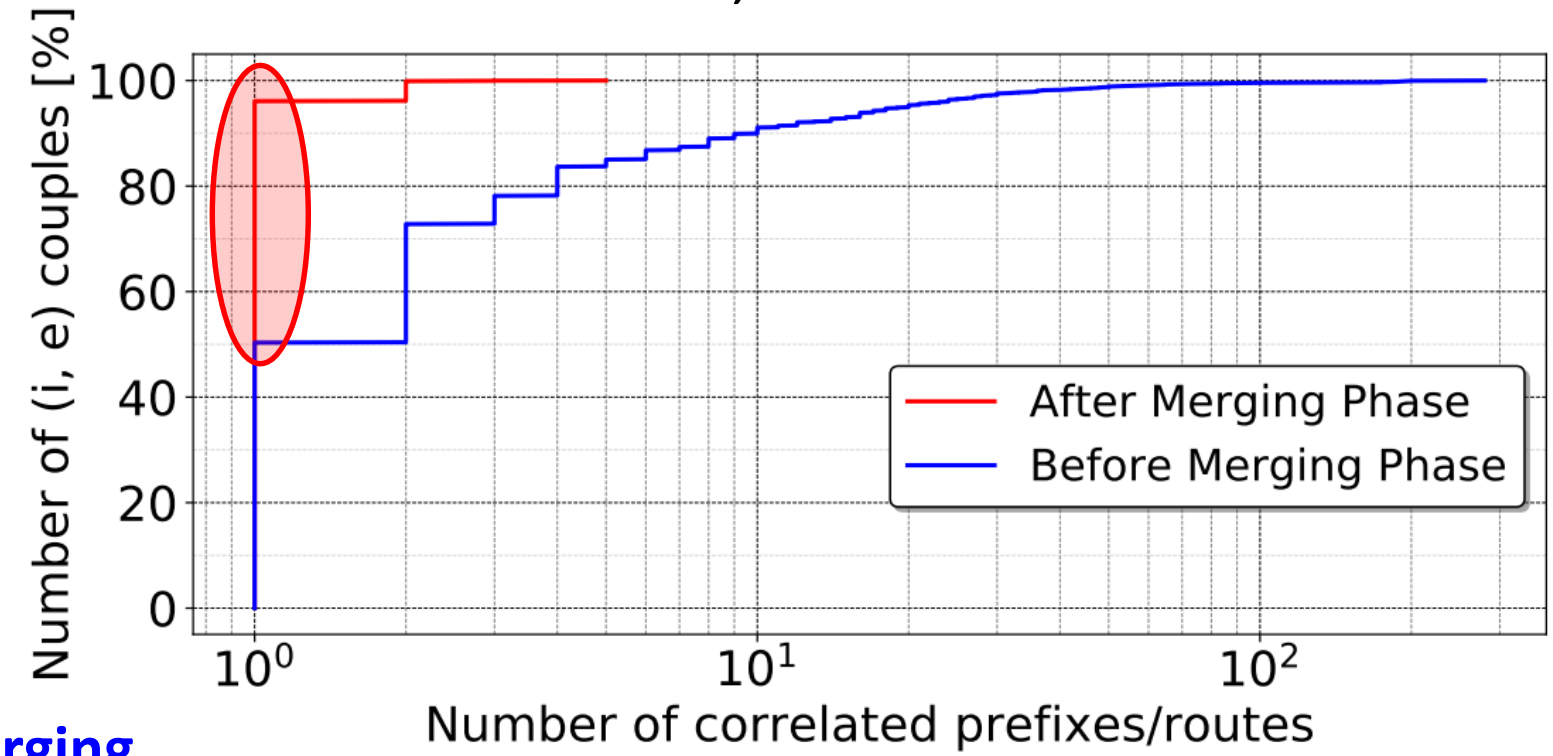
- **Before merging**
  - s = 1 for 50% cases, so **no LB sign**

Do we see F-LB, C-LB besides FDs?



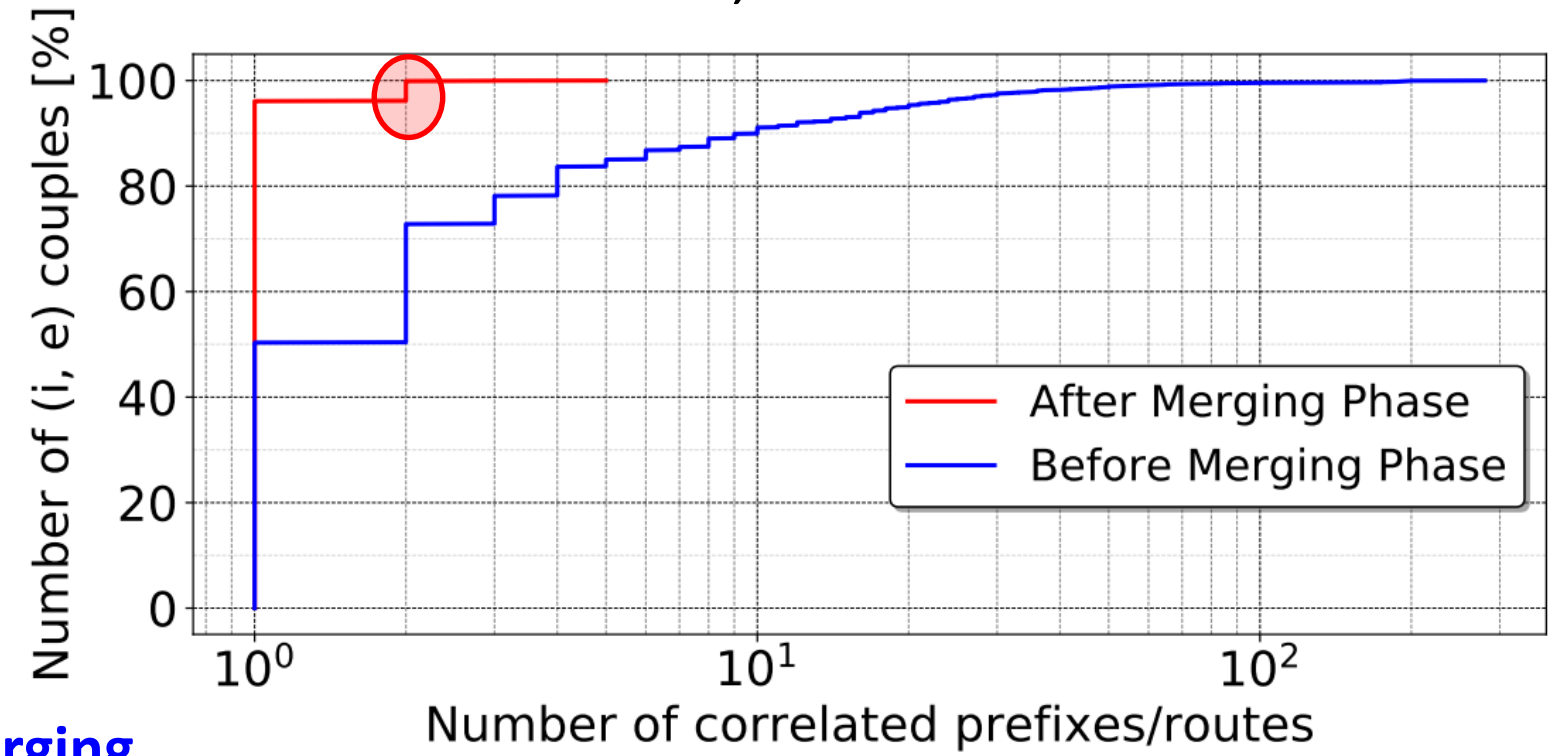
- **Before merging**
  - s = 1 for 50% cases, so **no LB sign**
- **After merging**
  - **C-LB flavors** do not seem to be popular

Do we see F-LB, C-LB besides FDs?



- **Before merging**
  - s = 1 for 50% cases, so **no LB sign**
- **After merging**
  - **C-LB flavors** do not seem to be popular
  - Additional s = 1 cases are **F-LB cases**

Do we see F-LB, C-LB besides FDs?



- **Before merging**
  - s = 1 for 50% cases, so **no LB sign**
- **After merging**
  - **C-LB flavors** do not seem to be popular
  - Additional s = 1 cases are **F-LB cases**
  - s = 2 is around 5% and **all are FDs**

Conclusions

- We propose a method to **detect** FDs
- We analyze LB with a new perspective, in terms of **forwarding patterns**
- FDs **can be found** in the wild, and they are AS-specific.
- FDs are **extreme**: when they occur all traffic of an couple detours.
- C-LB flavors are **not popular** nowadays

Complementary Work

- Presentations
  - **AIMS-KISMET 2020 Workshop, CAIDA, San Diego, Feb 2020.**
- Teaching
  - Mission d'enseignement (64 hs) + vacataire (23hs)
    - TCP/IP + C programming
  - Collaborated guiding a master-II intern
- Internships/Research Visits
  - 3-month internship in Telefonica Research
  - 20 day visit in University of Napoli Federico II
- Scientific Communication
  - Organized a seminar on crytocurrencies
- Participation in PhD Summer Schools
  - TMA PhD School 2018 (presented posters) and 2019

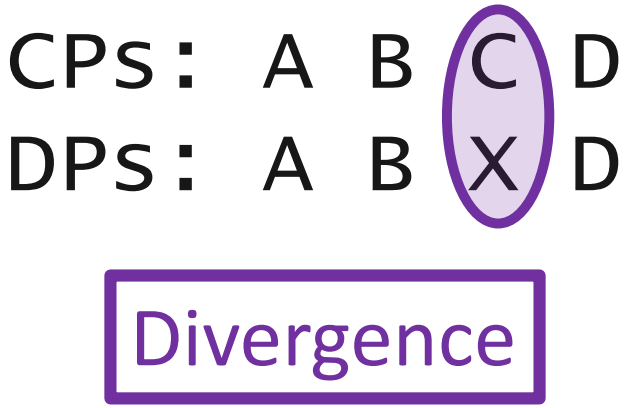
Future work

Leveraging TMA 2019

"Filtering the Noise to Reveal Inter-Domain Lies"

Julian M. Del Fiore, Pascal Merindol, Valerio Persico
Cristel Pelsser, Antonio Pescape

BGP vs Traceroute-AS paths or CPs vs DPs

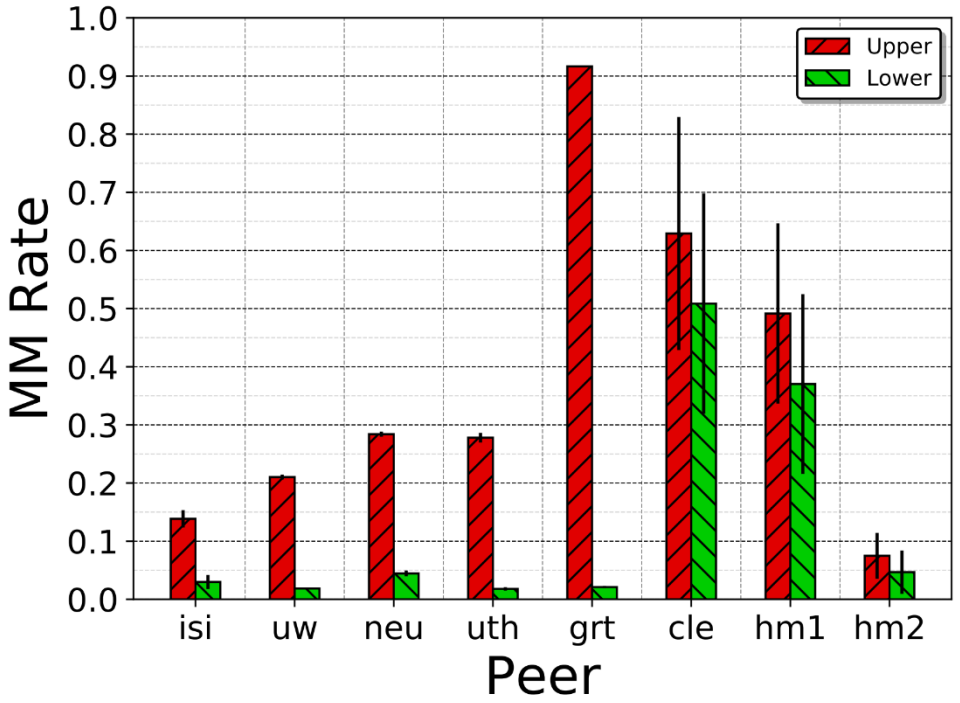


BGP vs Traceroute-AS paths or CPs vs DPs

CPs: A B C D

DPs: A B X D

Divergence

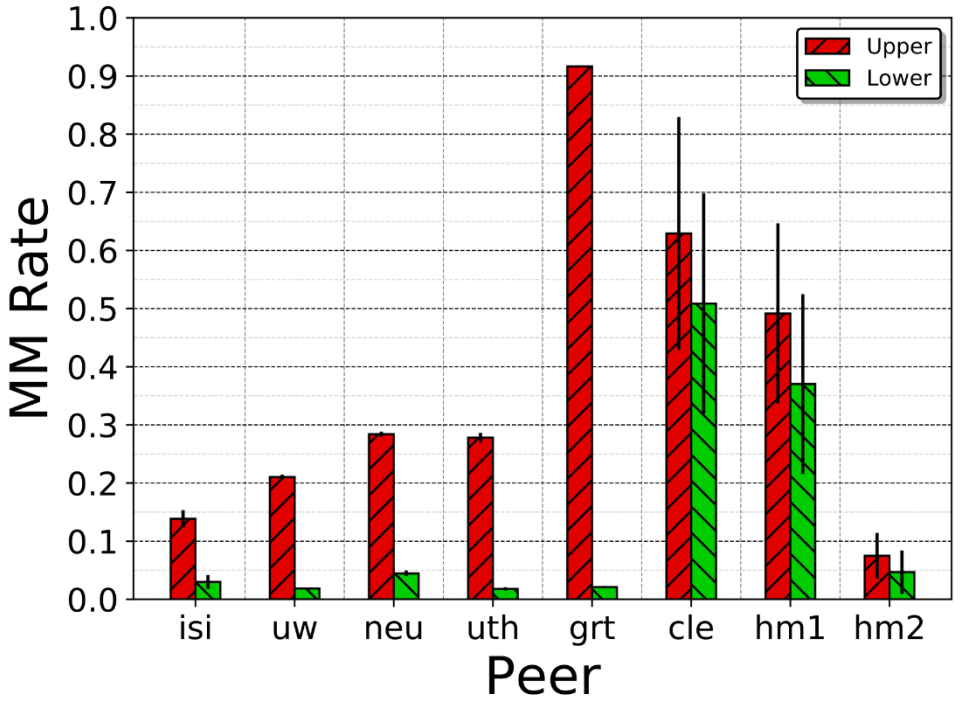


BGP vs Traceroute-AS paths or CPs vs DPs

CPs: A B C D

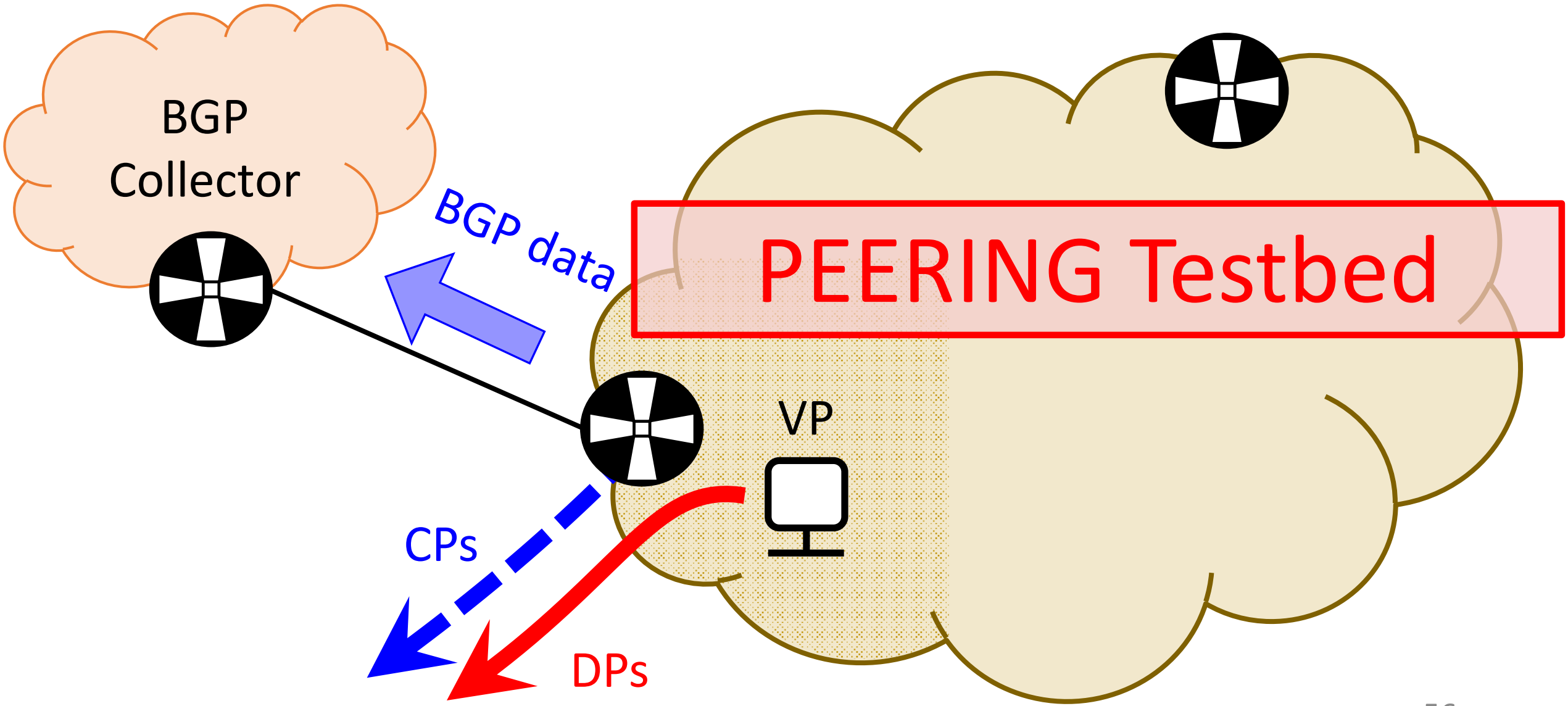
DPs: A B X D

Divergence



hm1: had technical limitations, e.g., router with a **partial-FIB**

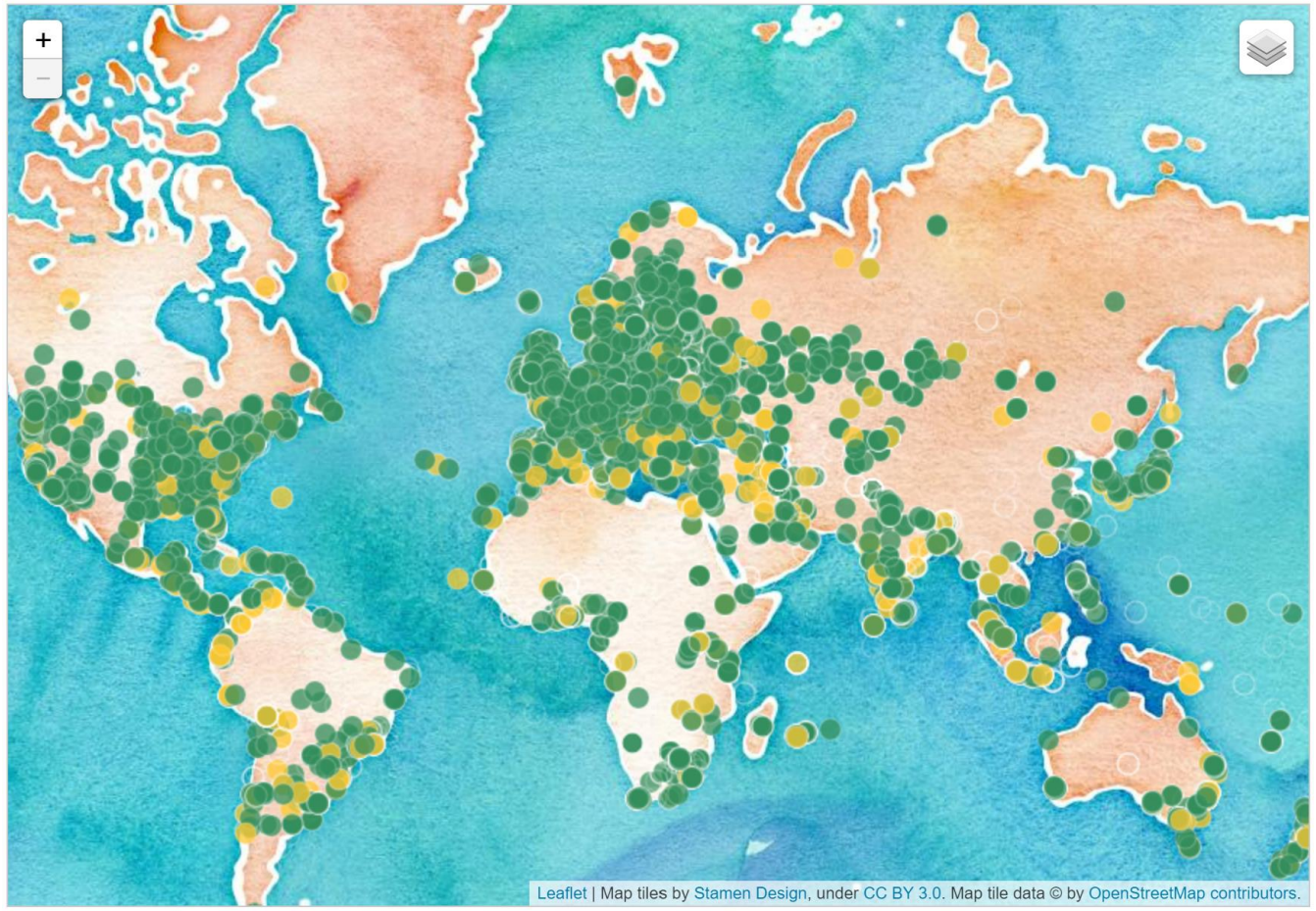
Co-located VPs



Infrastructure?...That's easy!

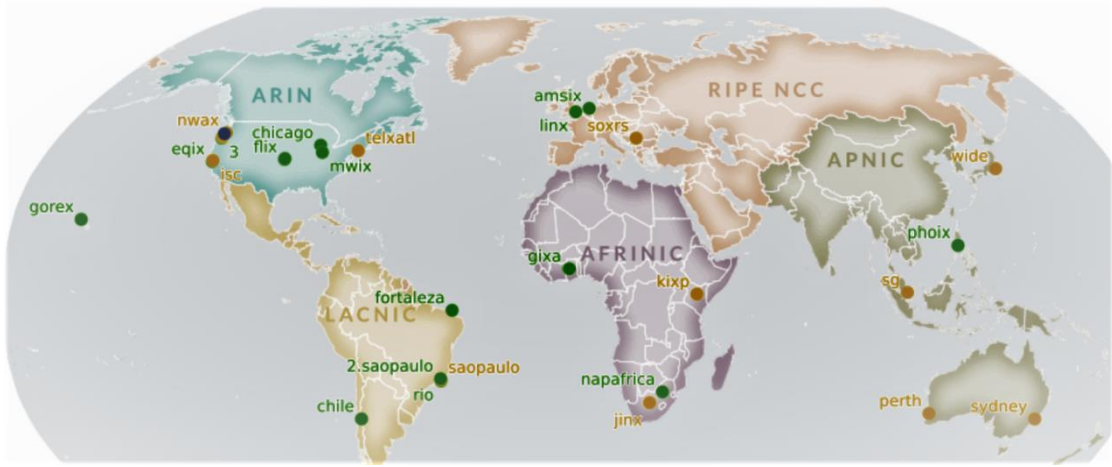
RIPE ATLAS: 11k++ probes (02/2020)

https://atlas.ripe.net/results/maps/network-coverage/

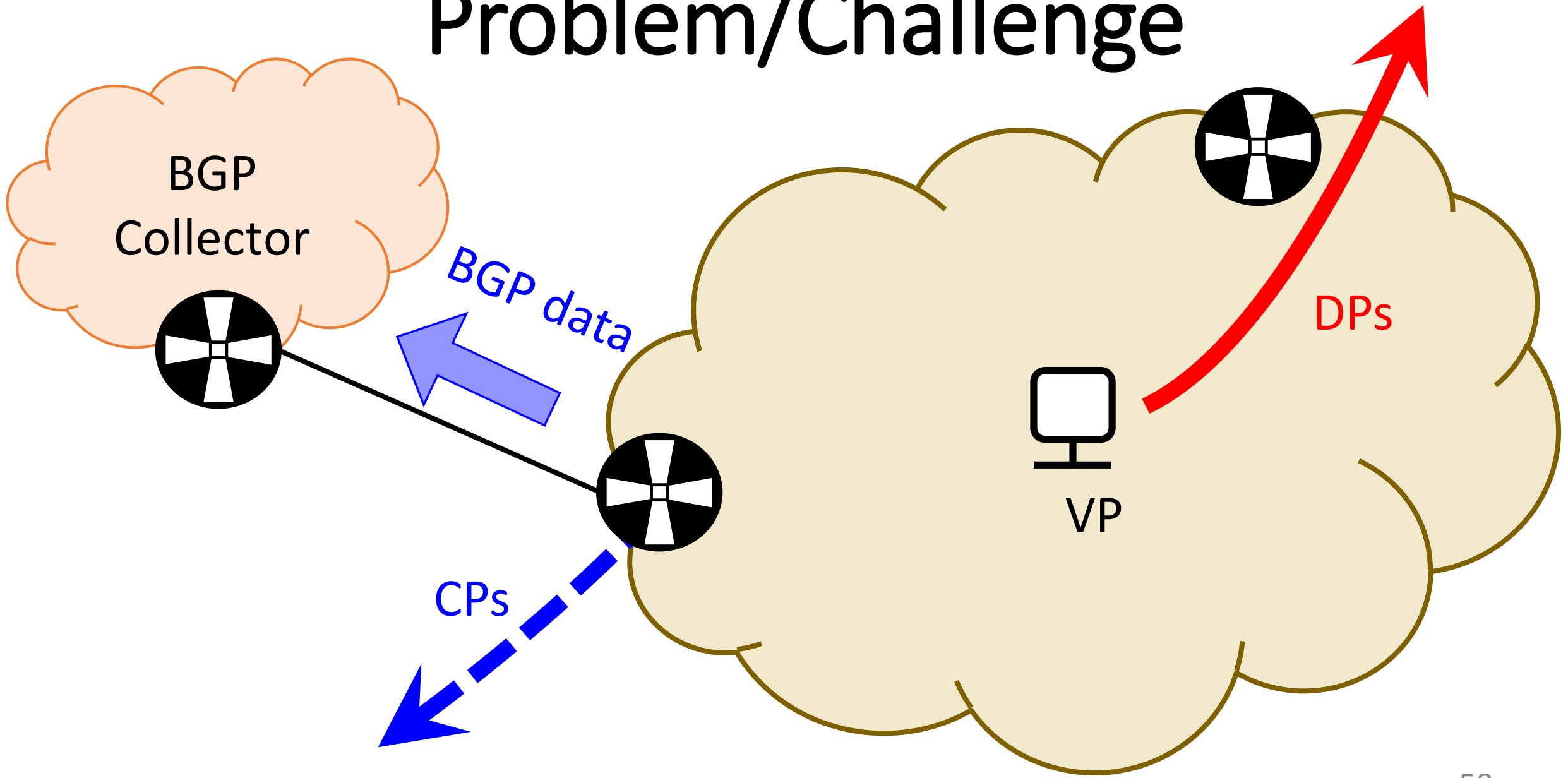


Routeviews: 31 collectors

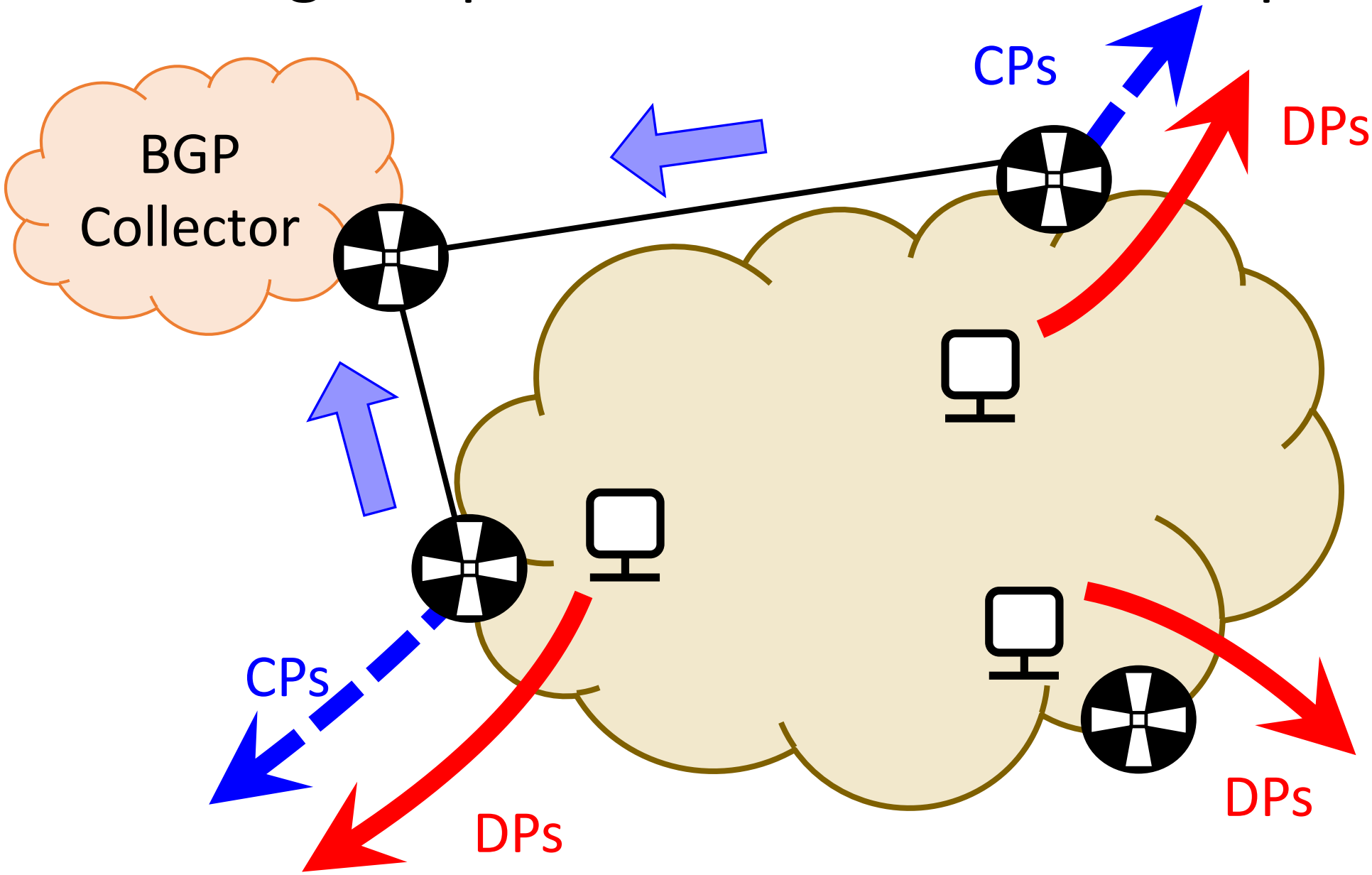
http://www.routeviews.org/routeviews/index.php/map/



Problem/Challenge



Extending the problem to consider multiple VPs



My Obligations concerning l'ED

- Formations Scientifiques: 83/54
- Formations Transversales: 77/54
- Formations Europe: 18/18
- Seminaires: 12/18 (due to COVID…)

My Contributions

“Capturing Forwarding Deflections in the Wild: Desired Load balancing or Unwanted Detours?”
Under submission in IMC 2020, University of Strasbourg/ICube, University of Napoli Federico II

“Filtering the Noise to Reveal Inter-Domain Lies”
In TMA 2019, University of Strasbourg/ICube, University of Napoli Federico II

“Understanding LatAm's IXP ecosystem in an International Context”, Under submission in CoNEXT 2020, University of Buenos Aires/CONICET, University of Strasbourg/ICube

“A first Look at The Latin American IXPs”, in CCR 2020, January Issue
University of Buenos Aires/CONICET, University of Strasbourg/ICube, University Diego Portales

"Country-level influence of IXPs in Latin America“
In LANCOMM 2019, University of Buenos Aires/CONICET, University of Strasbourg/ICube

---

"From Best-Effort to Deterministic Packet Delivery for Wireless Industrial IoT Networks“
In IEEE Transactions on Industrial Informatics 2018, IMT Atlantique, University of Bristol

"Toward Deterministic Industrial Networks“
In AlgoTel-CoRes 2017, IMT Atlantique

The End

Questions ?