Detecting Hidden Broken Pieces of The Internet

PhD Defense by **Julian M. Del Fiore**

February 08, 2021

Outline

- Background, Research Goal and Questions
- Part I. Filtering the noise to reveal BGP lies
- Part II. Success and Failure of IXPs in Latin America
- Part III. The Art of Detecting Forwarding Detours
- Conclusions and Future Work

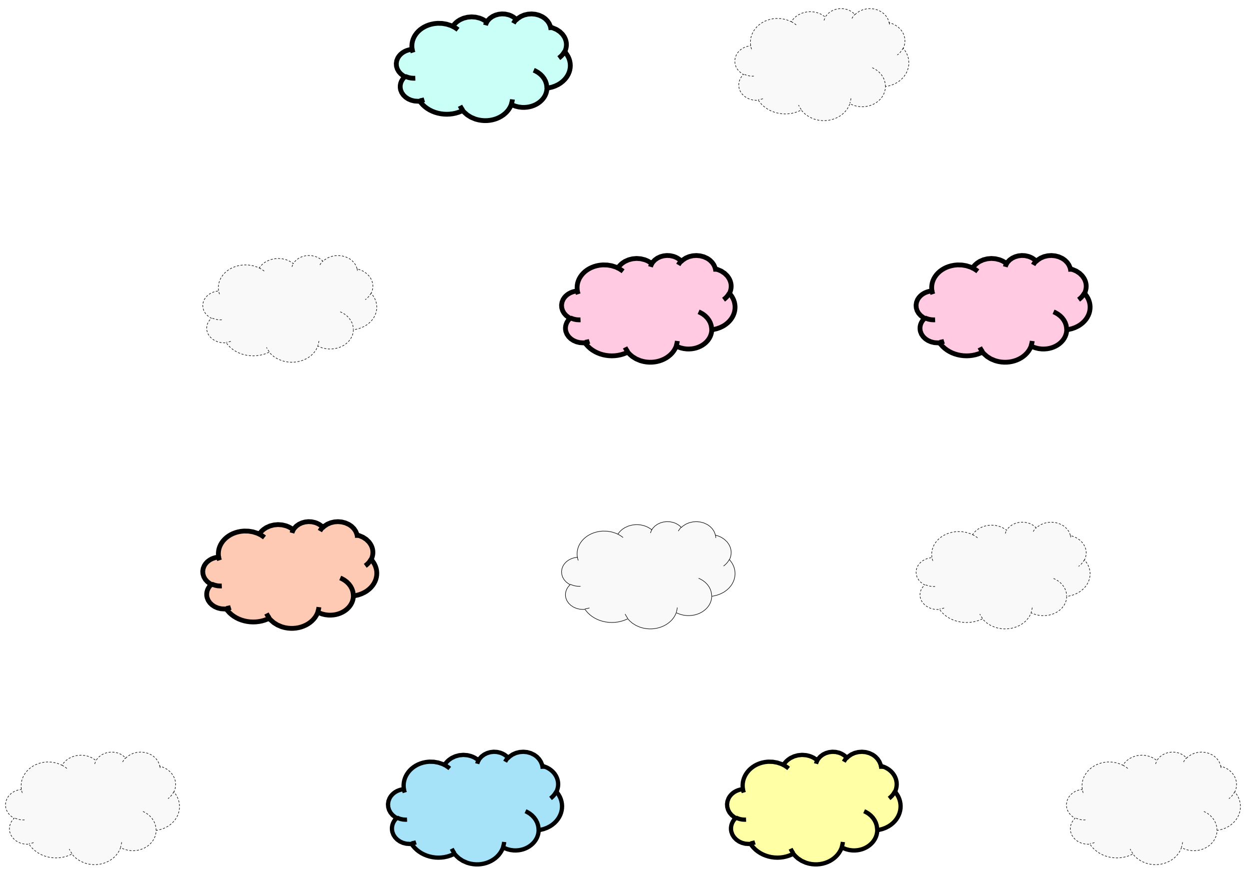
Outline

- Background, Research Goal and Questions
- Part I. Filtering the noise to reveal BGP lies
- Part II. Success and Failure of IXPs in Latin America
- Part III. The Art of Detecting Forwarding Detours
- Conclusions and Future Work

Background

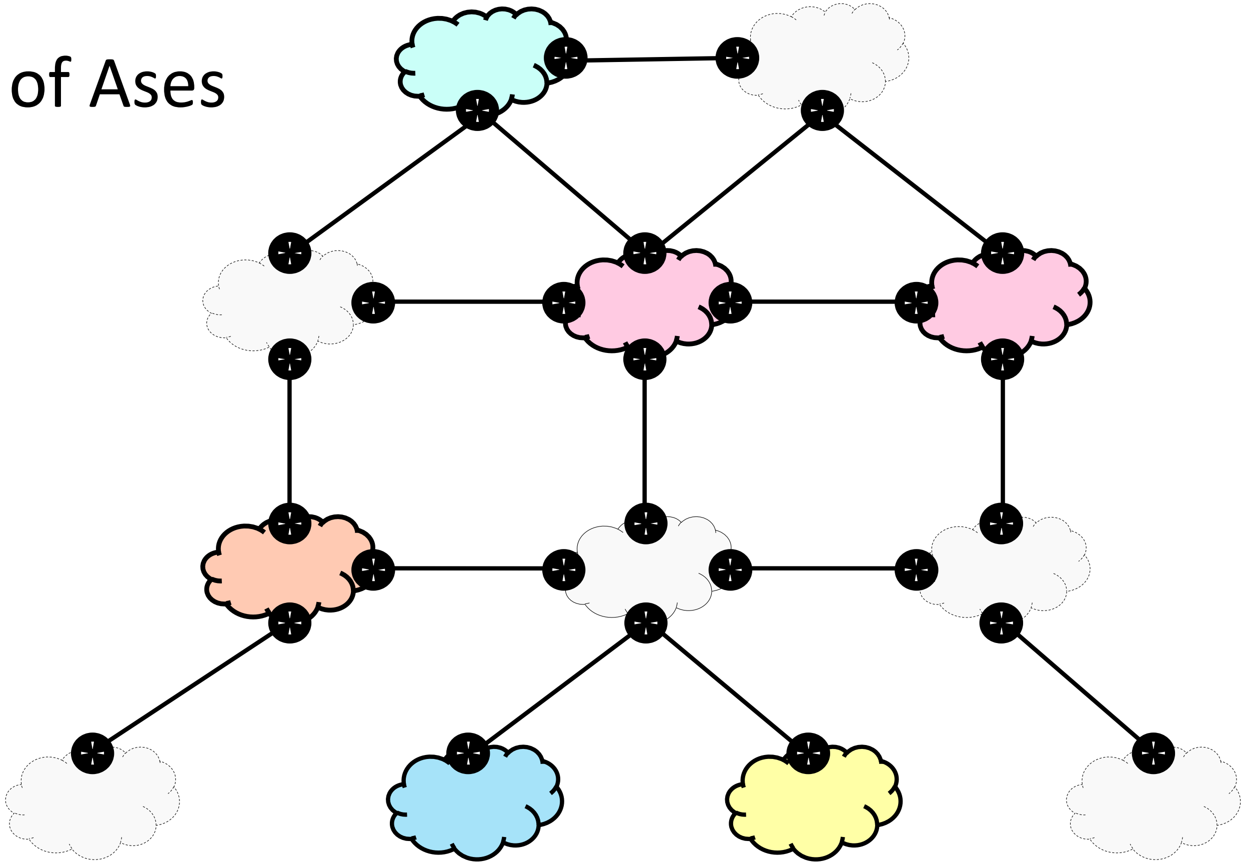
The Internet

- Autonomous Systems (ASes) are independent networks



The Internet

- Autonomous Systems (ASes) are independent networks
- The Internet is an Interconnection of Ases

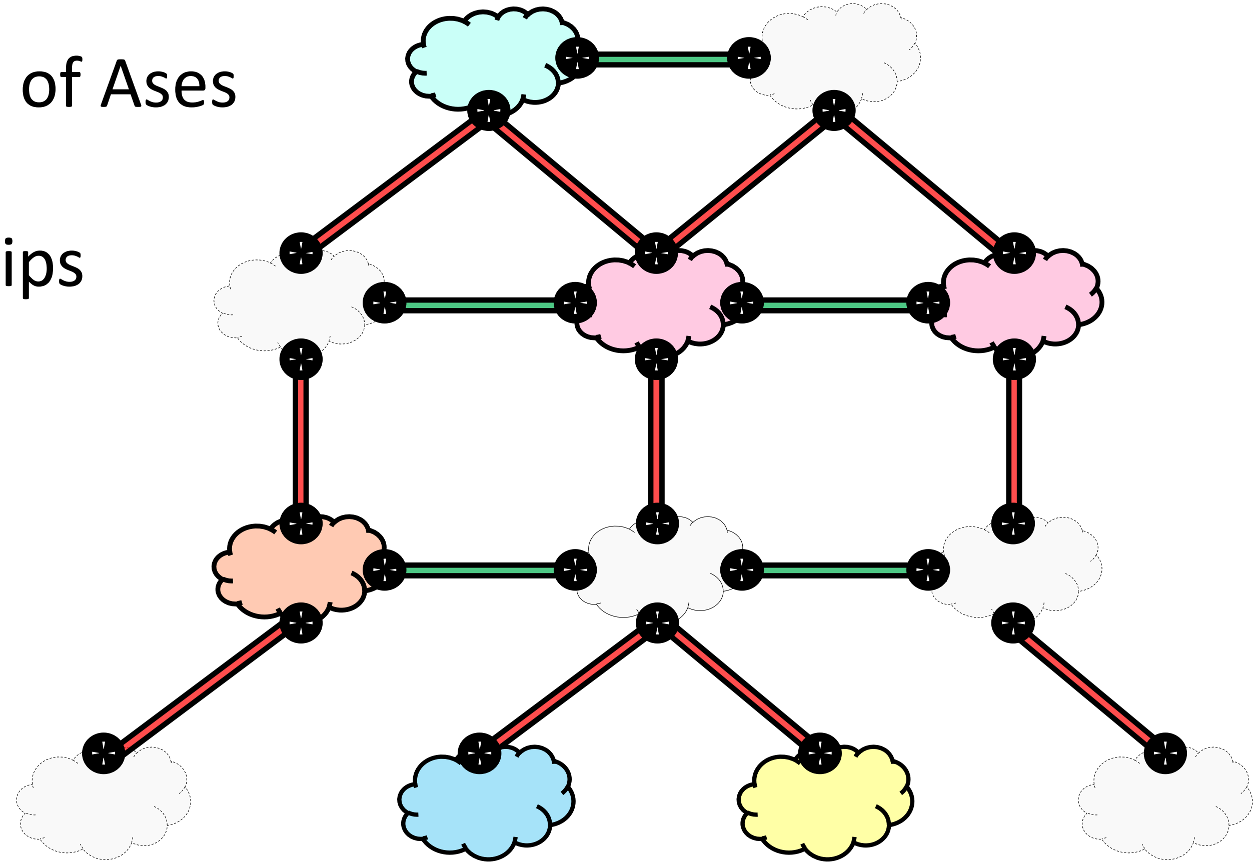


The Internet

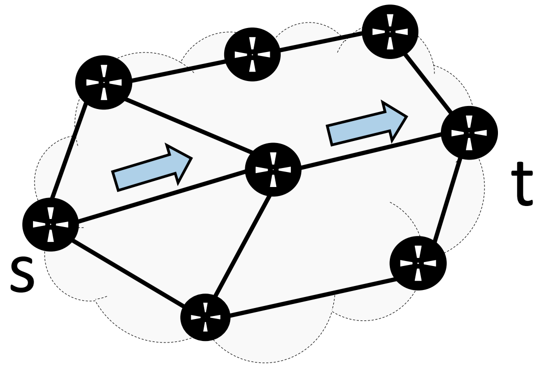
- Autonomous Systems (ASes) are independent networks

- The Internet is an Interconnection of Ases

- ASes establish business relationships
  - Customer-to-provider $$$
  - Peer-to-Peer Free



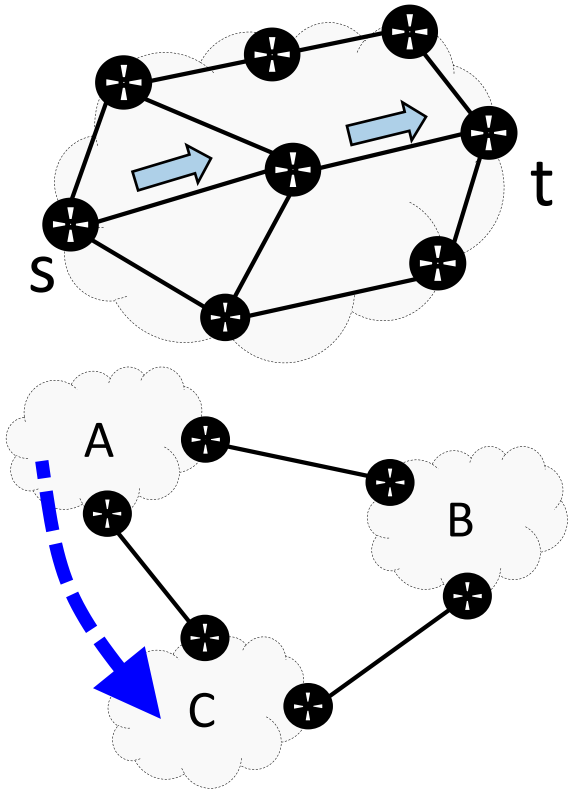
The Internet



- ASes run an Internal Gateway Protocol (IGP)
  - Deals with intra-domain routing

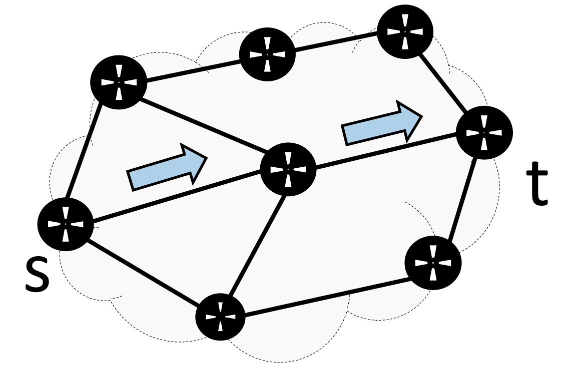
The Internet

- ASes run an Internal Gateway Protocol (IGP)
  - Deals with intra-domain routing
- ASes run the Border Gateway Protocol (BGP)
  - Deals with the inter-domain routing

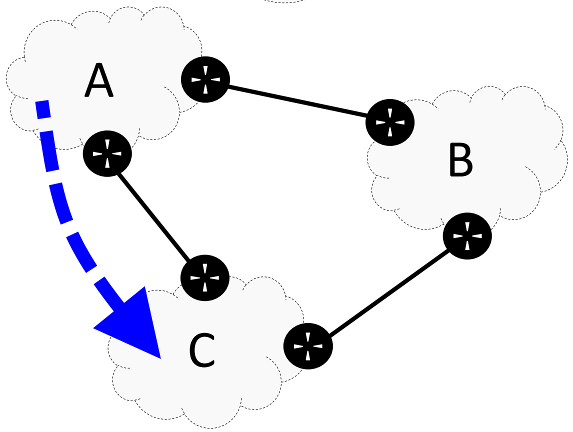


The Internet

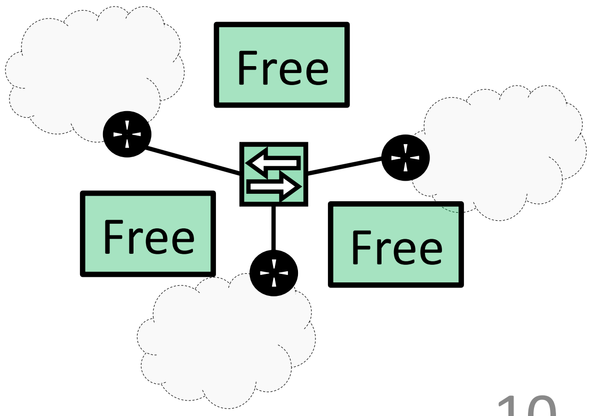
- ASes run an Internal Gateway Protocol (IGP)
  - Deals with intra-domain routing



- ASes run the Border Gateway Protocol (BGP)
  - Deals with the inter-domain routing



- ASes peer at Internet Exchange Points (IXPs)
  - Peer-to-peer relationships at a large scale



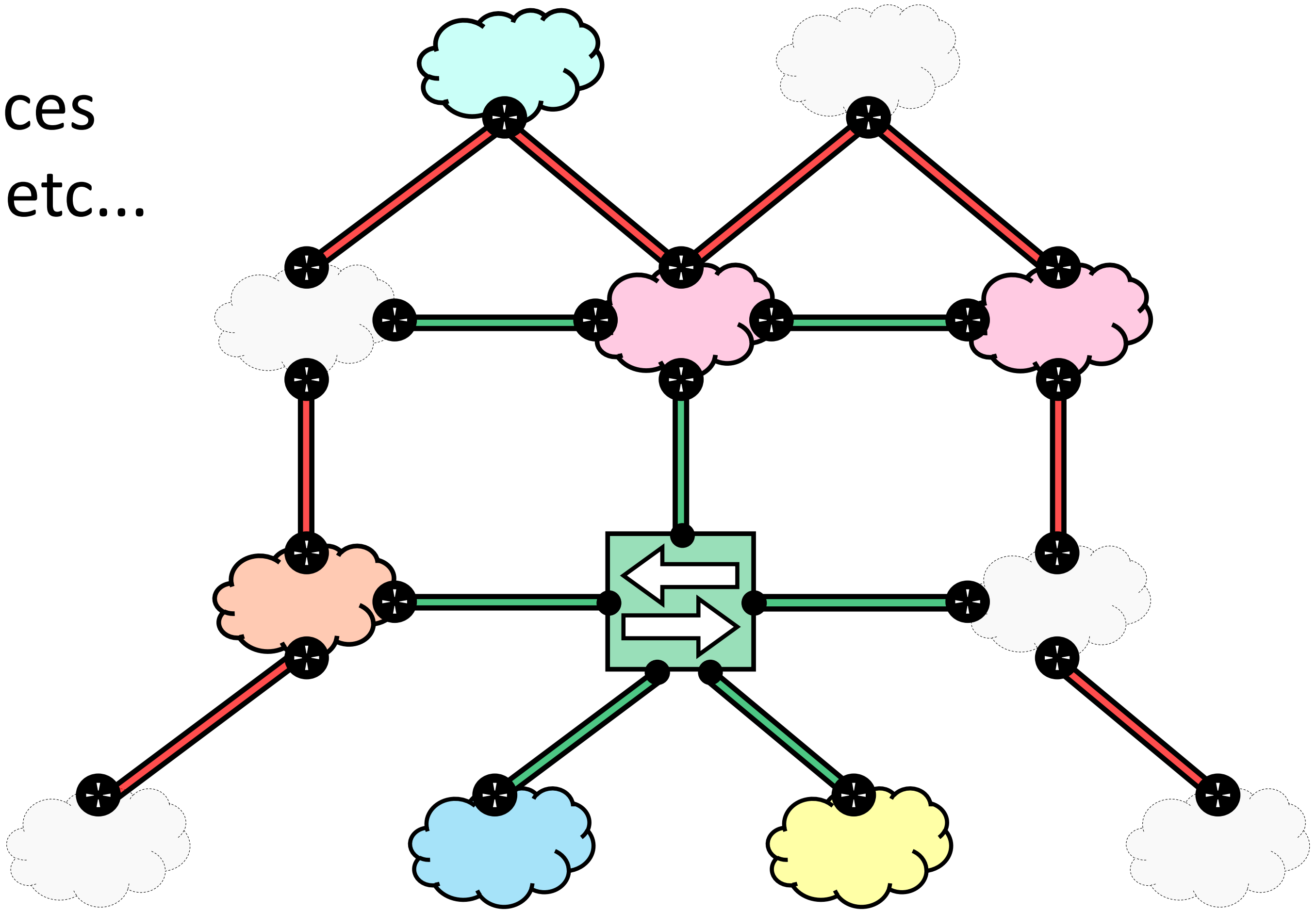
Research Goal

Research Goal

- Any system may have broken pieces
  - Problems, errors, limitations, etc...

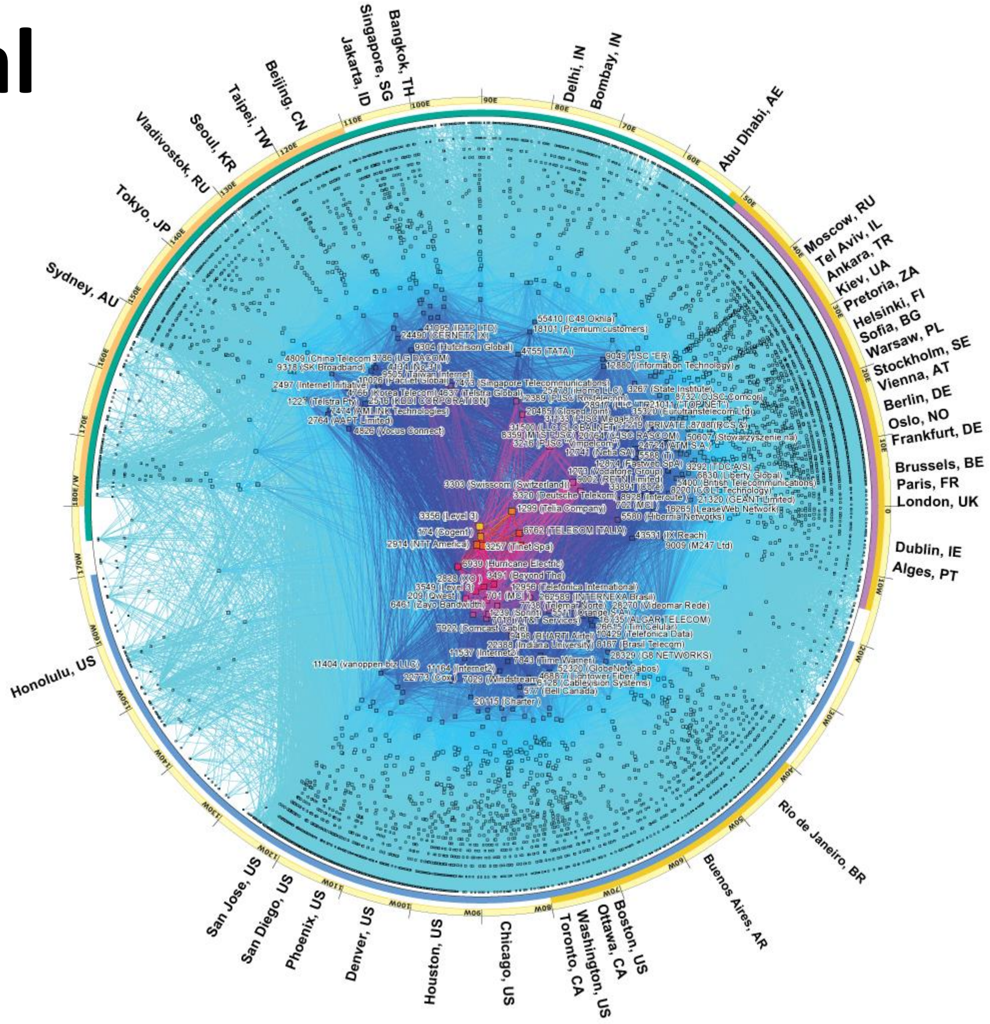
Research Goal

- Any system may have broken pieces
  - Problems, errors, limitations, etc...
- The Internet is a complex system
  - Protocols, facilities, networks
  - Hardware, software
  - Network operators



Research Goal

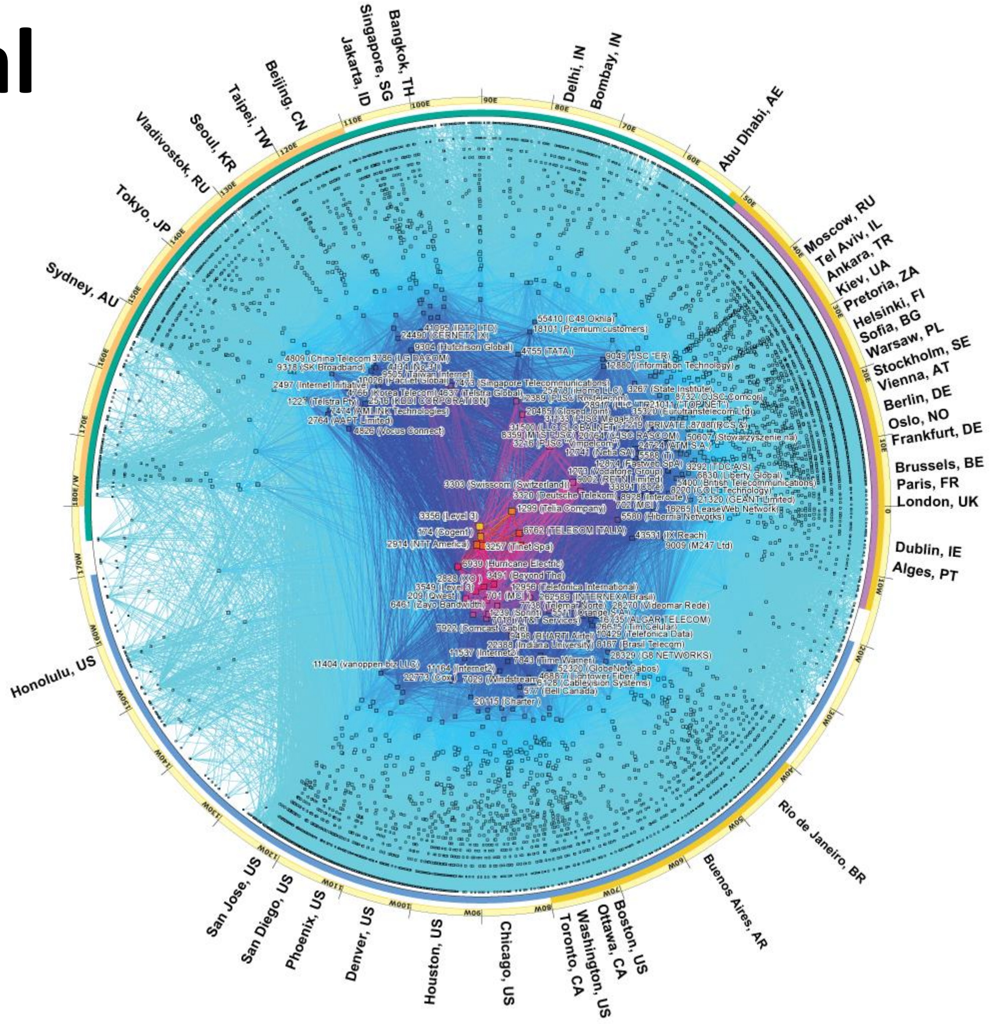
- Any system may have broken pieces
  - Problems, errors, limitations, etc...
- The Internet is a complex system
  - Protocols, facilities, networks
  - Hardware, software
  - Network operators
- The Internet is “big”...
  - Composed of 70K ASes
  - Point of observation matters



CAIDA’s IPv4 AS Core February 2017

Research Goal

- Any system may have broken pieces
  - Problems, errors, limitations, etc...
- The Internet is a complex system
  - Protocols, facilities, networks
  - Hardware, software
  - Network operators
- The Internet is “big”...
  - Composed of 70K ASes
  - Point of observation matters

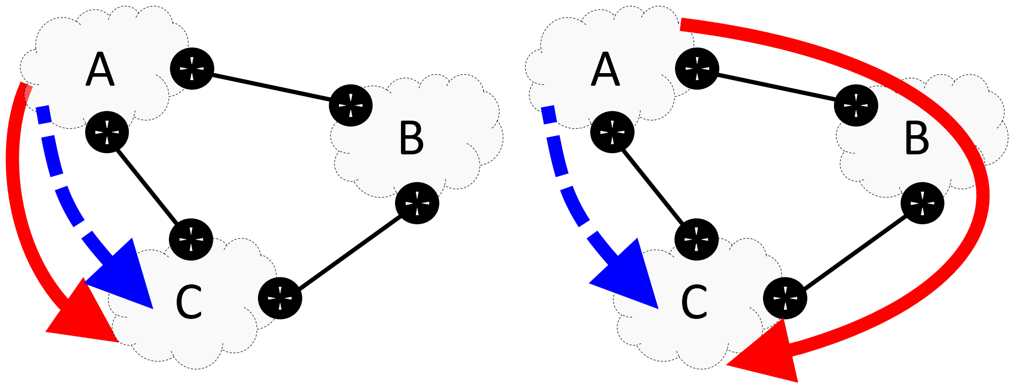


CAIDA’s IPv4 AS Core February 2017

Research Goal: Detecting Hidden Broken Pieces of The Internet

Research Questions

Research Questions



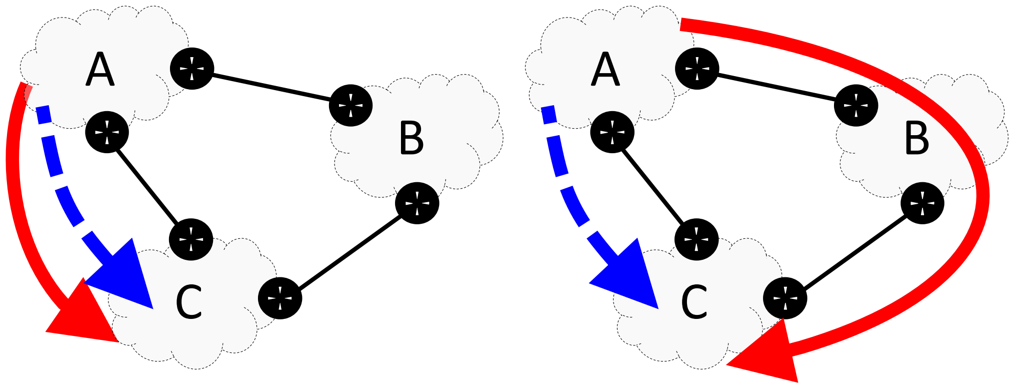
Q1: Can we detect BGP lies?

- Expected ≠ Practice

Research Questions

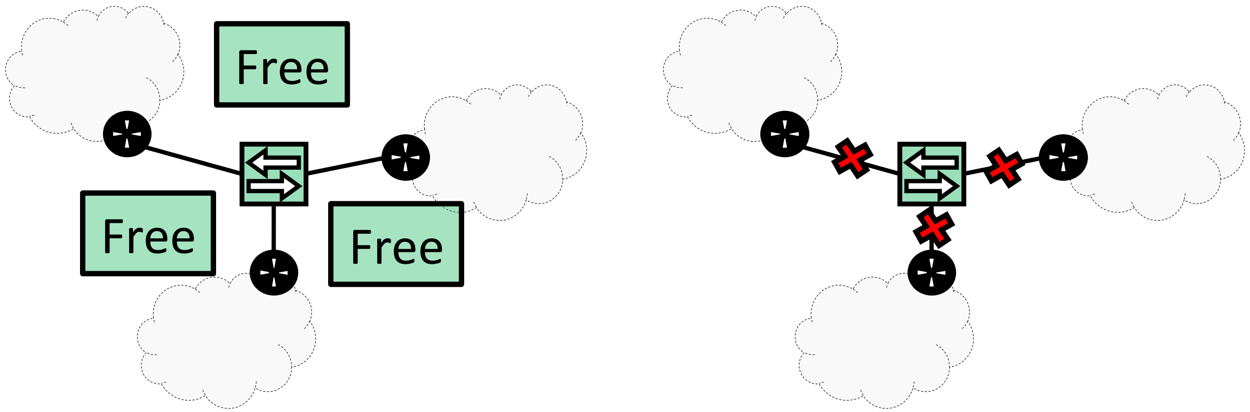
Q1: Can we detect BGP lies?

- Expected ≠ Practice



Q2: Are there failed IXPs? Why?

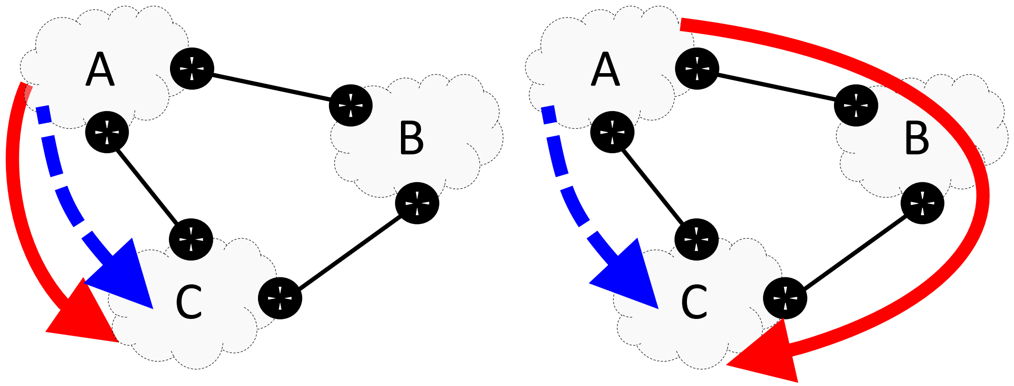
- IXPs with low coverage



Research Questions

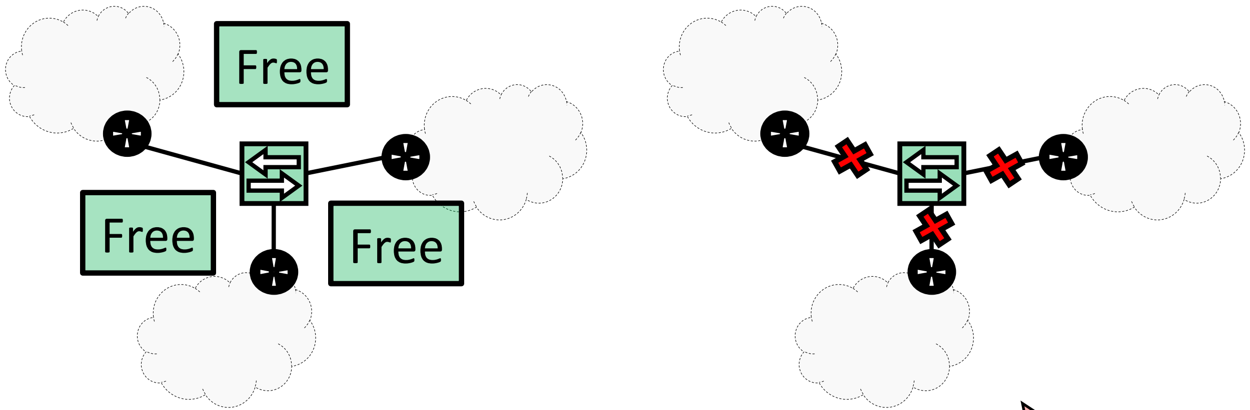
Q1: Can we detect BGP lies?

- Expected ≠ Practice



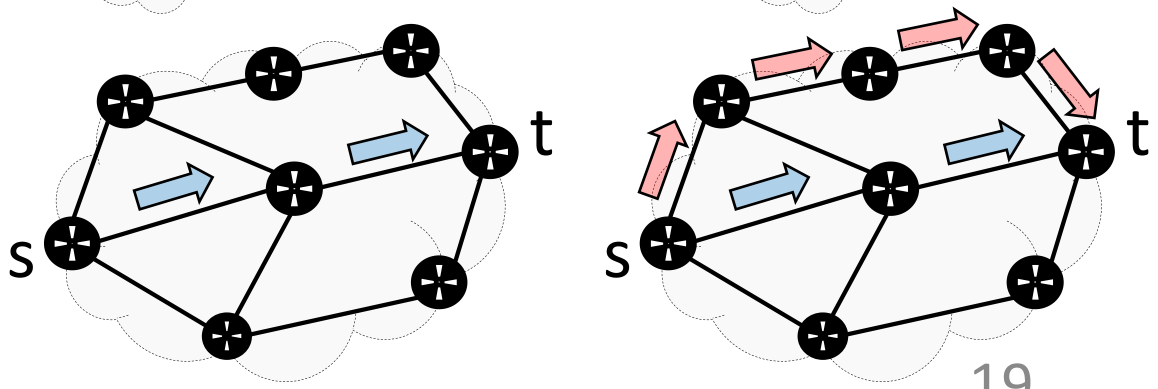
Q2: Are there failed IXPs? Why?

- IXPs with low coverage



Q3: Can we model and detect detours?

- Expected ≠ Practice



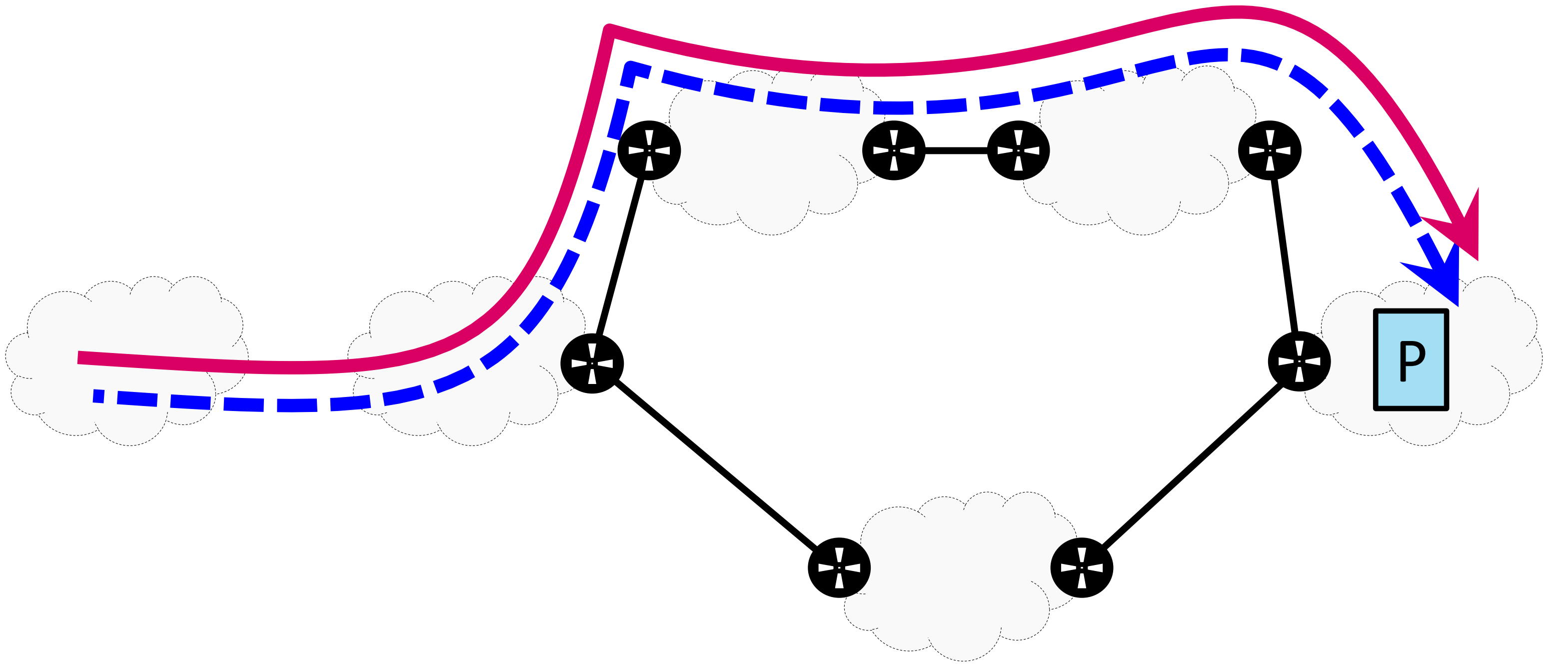
Outline

- Background, Research Goal and Questions
- Part I. Filtering the noise to reveal BGP lies
- Part II. Success and Failure of IXPs in Latin America
- Part III. The Art of Detecting Forwarding Detours
- Conclusions and Future Work

Background

Border Gateway Protocol (BGP)

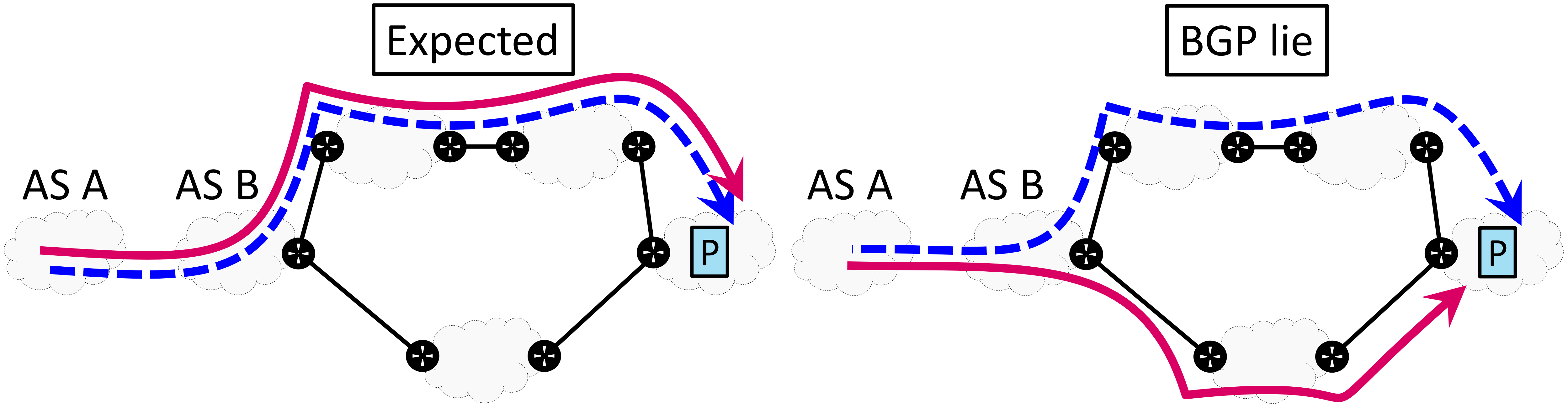
- For each external prefix P...
  - The control path (CP) that should **theoretically** be followed
  - The data path (DP) is the path used in **practice**



Problem Statement

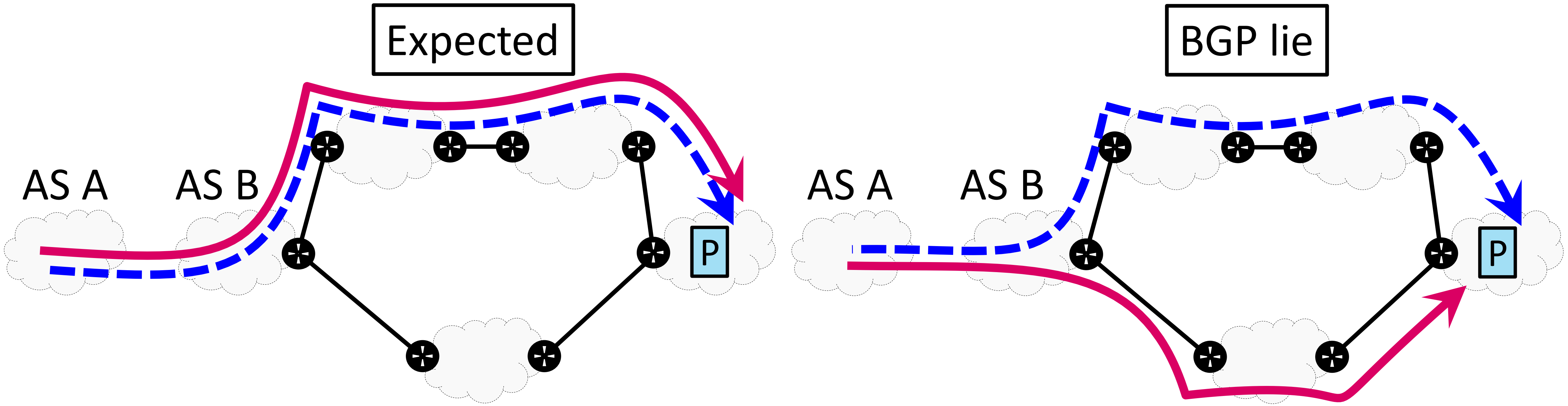
What are BGP lies?

When the control path (CP) and data path (DP) for a prefix P do not match



What are BGP lies?

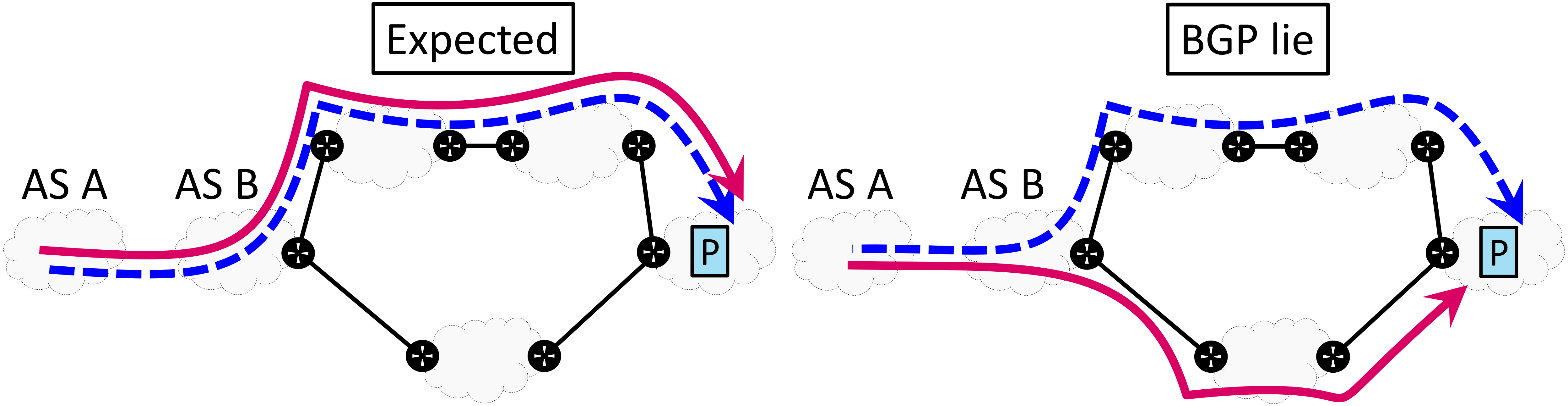
When the control path (CP) and data path (DP) for a prefix P do not match



AS B is lying to AS A

What are BGP lies?

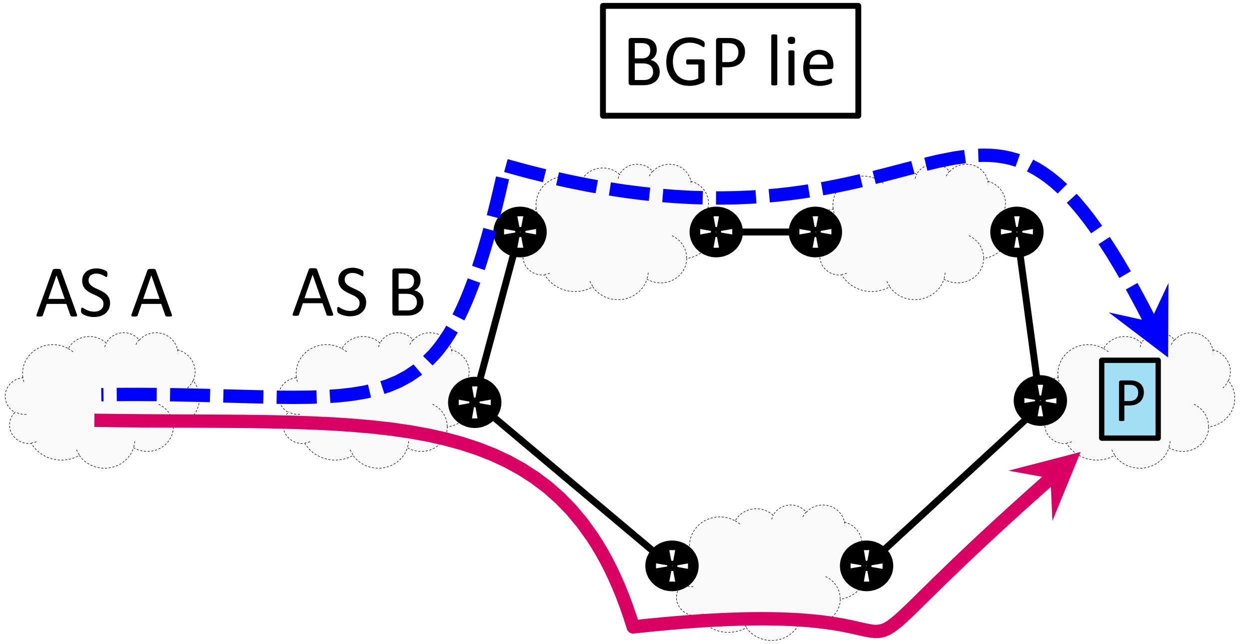
When the control path (CP) and data path (DP) for a prefix P do not match



AS B is lying to AS A

BGP lies may result from **malicious behavior** or **technical limitations**

Why detecting BGP lies (CP ≠ DP)?



- If not, what is the point of using BGP?
- Allows to detect possible malicious ASes
- Would allow to troubleshoot ASes

Detecting BGP lies

Required data

Control paths

P	CP
P_Y	BCD
P_R	D
P_V	E

Data paths

Vantage Point (VP)

Traceroute per destination

Technical Considerations

- ➢ Space-synchronization
  - Measurement platform
- ➢ Address space and time synchronization
  - Which DP should be compared with which CP
- ➢ IP-to-AS mapping
  - CPs come as AS-paths but DPs as IP-paths

Technical Considerations

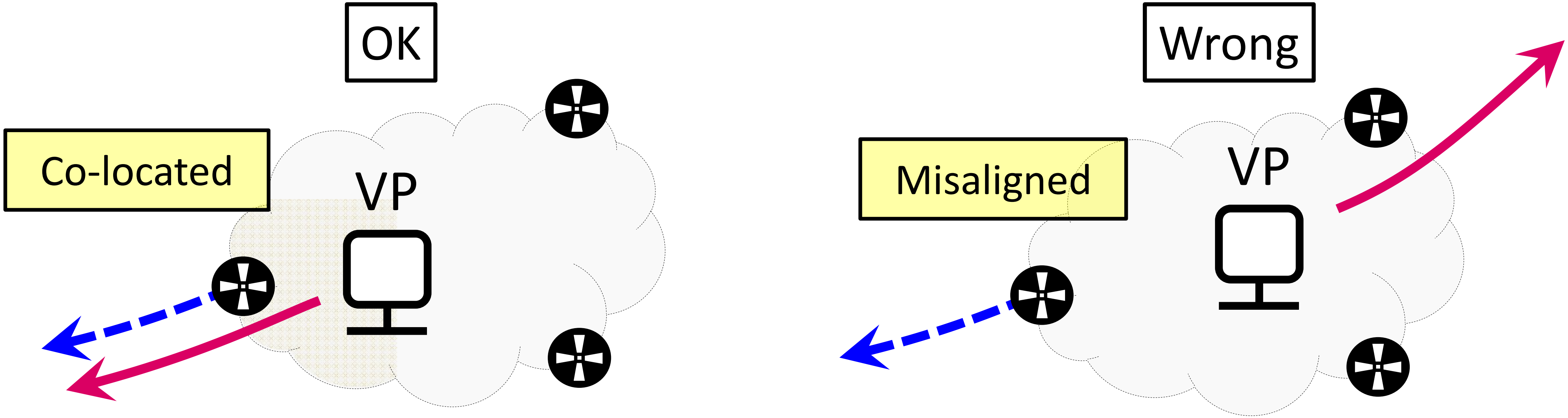
- Space-synchronization
  - Measurement platform

- Address space and time synchronization
  - Which DP should be compared with which CP

- IP-to-AS mapping
  - CPs come as AS-paths but DPs as IP-paths

Space-synchronization

- Control paths are obtained from a given router
- Data paths are gathered from a VP
- To be comparable, DPs need to go through the router that shared the CPs



IP-to-AS mapping

- While CPs are AS-paths, DPs are obtained as IP-paths

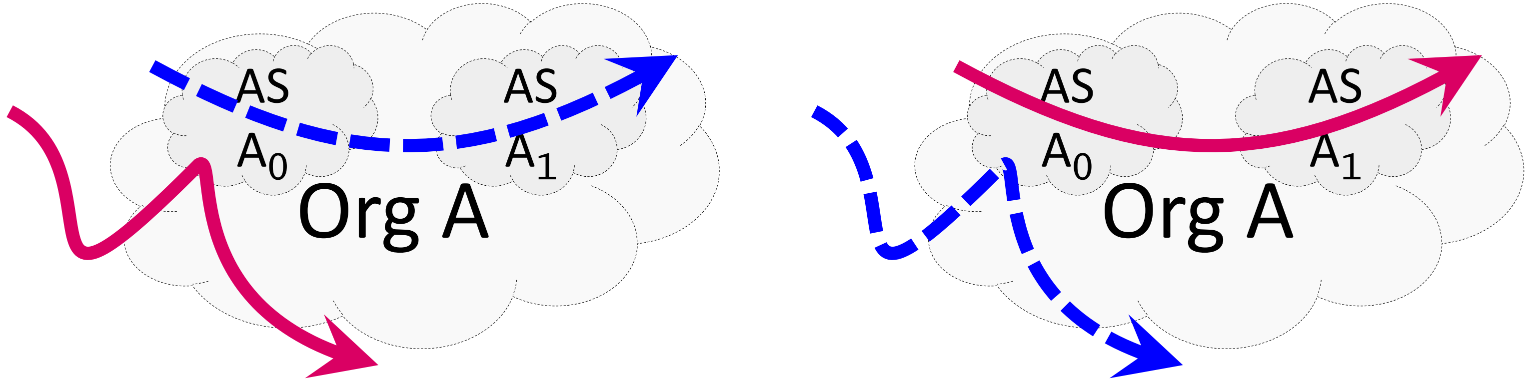
```
CP: AS A, AS B, AS C…
DP: IP1, IP2, IP3, IP4…
```

To compare them, an IP-to-AS mapping tool is needed !

The problem of IP-to-AS mapping

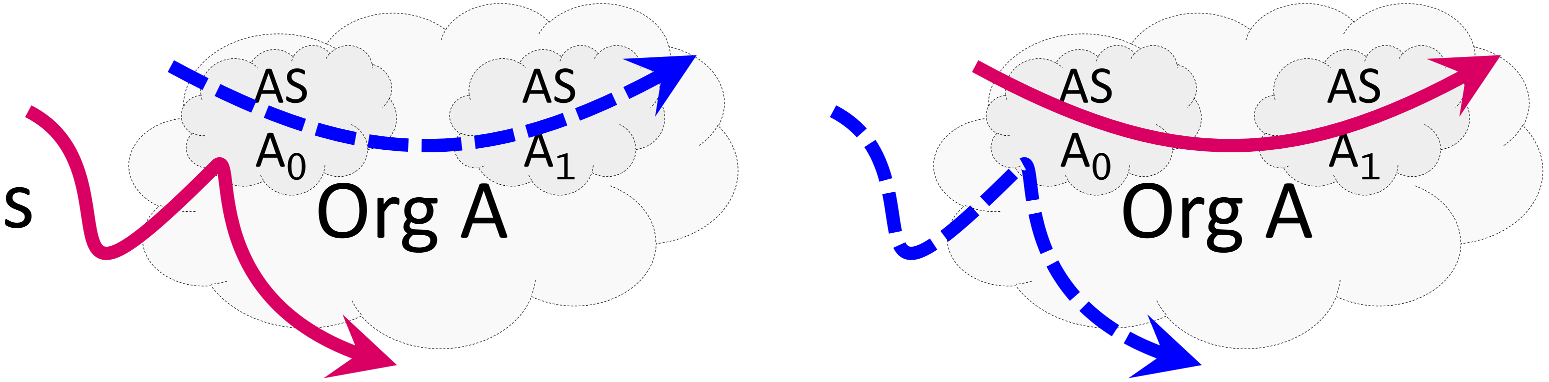
Noise or sources of errors

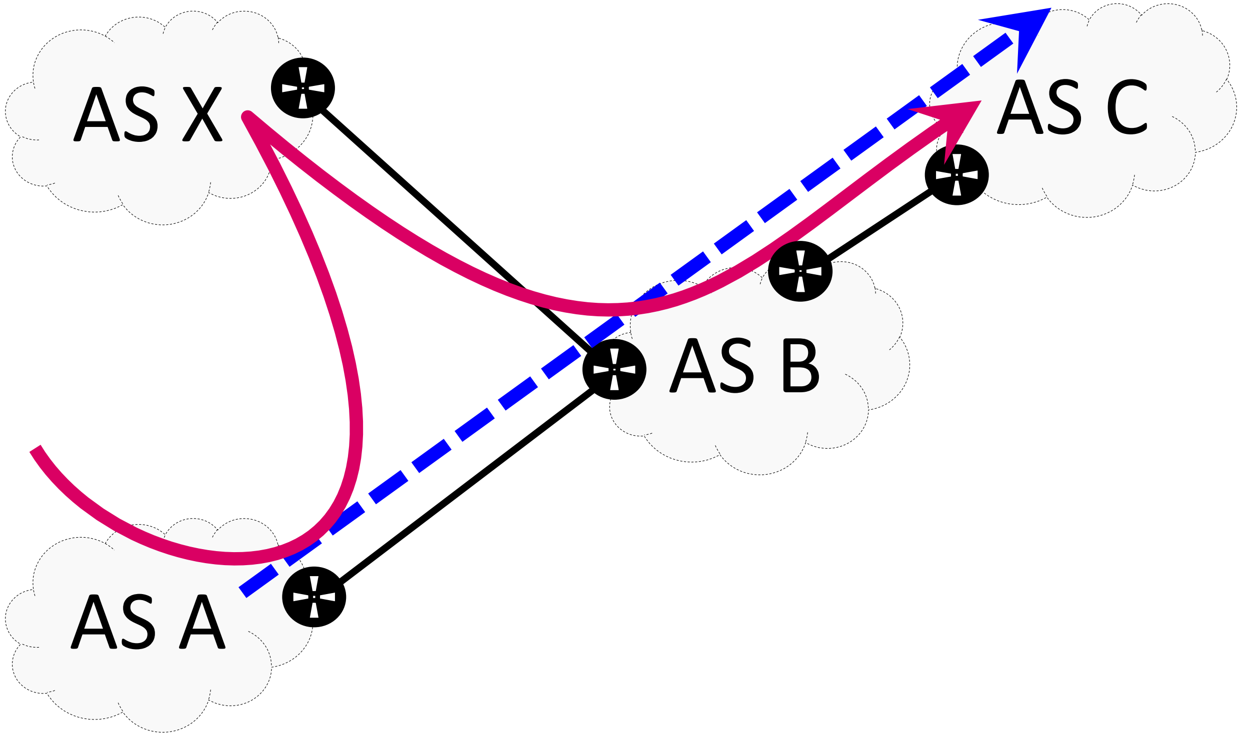
- AS siblings



Noise or sources of errors

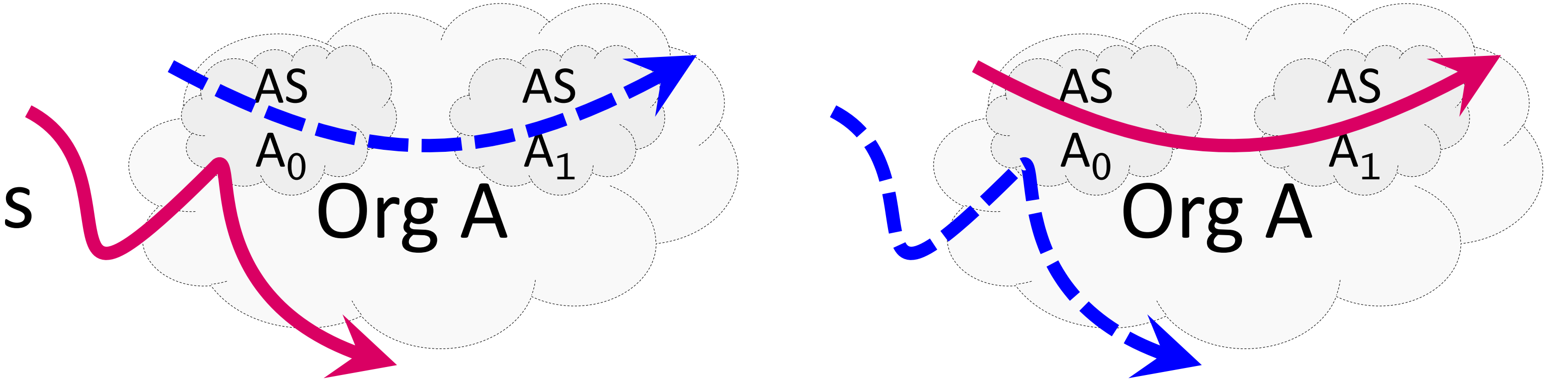
- AS siblings
- Third-party addresses

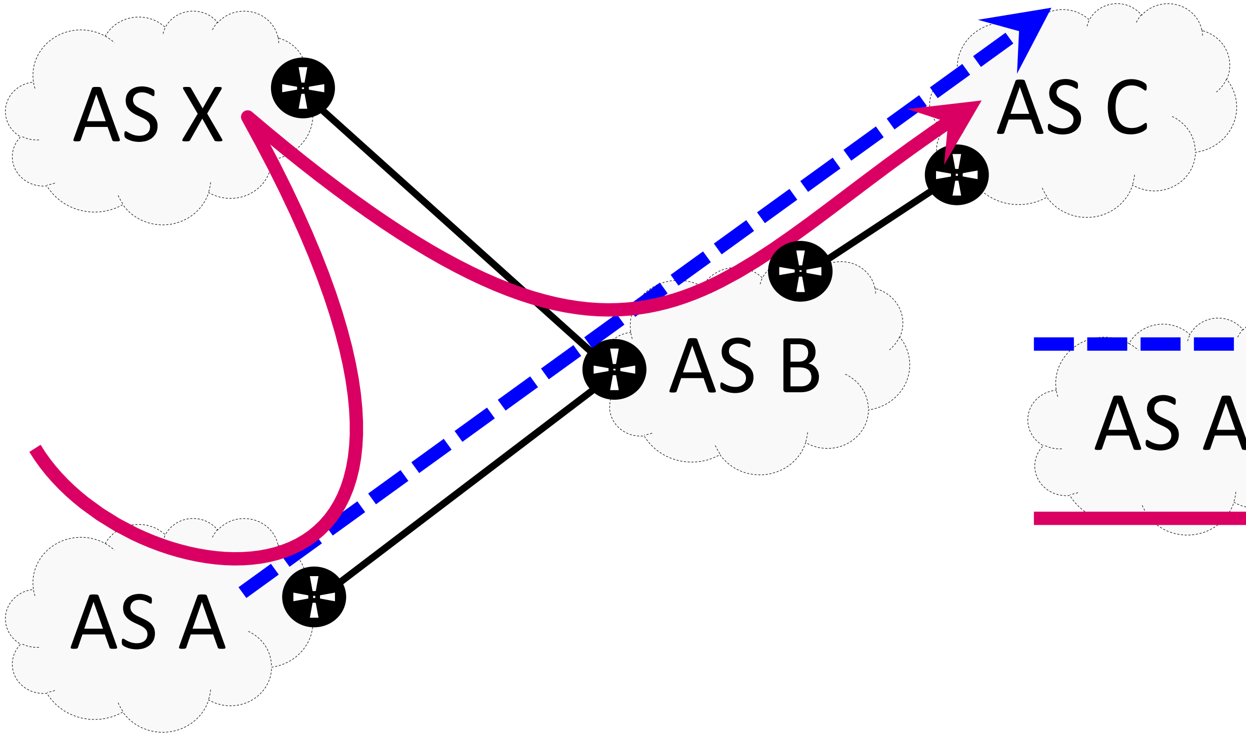


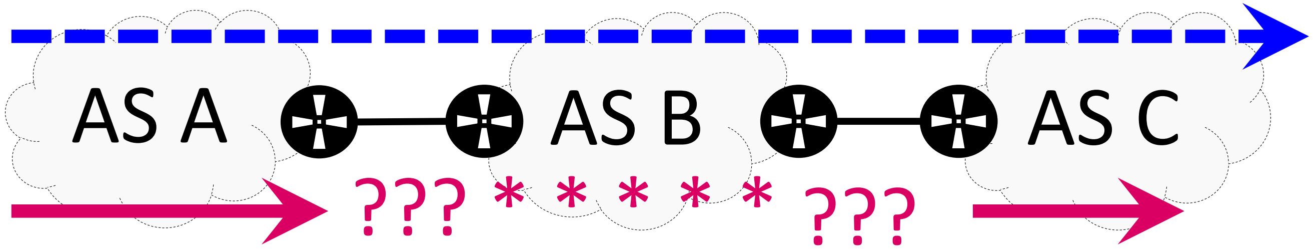


Noise or sources of errors

- AS siblings
- Third-party addresses
- Missing hops

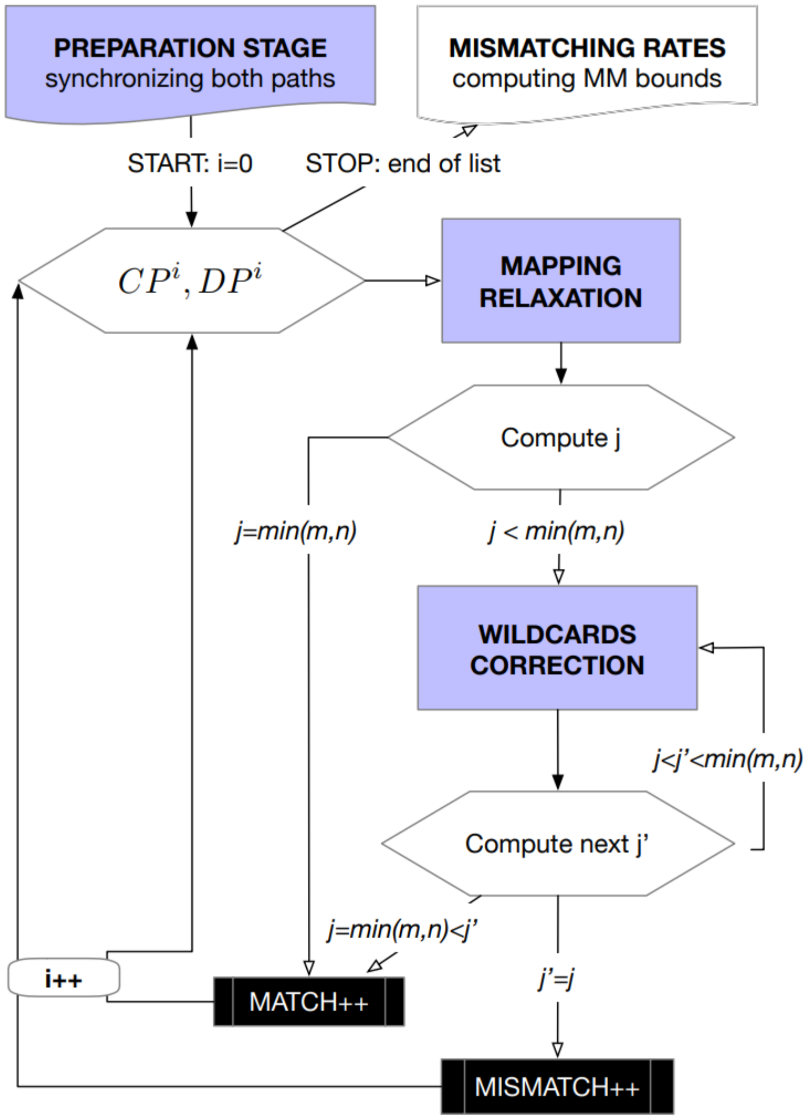






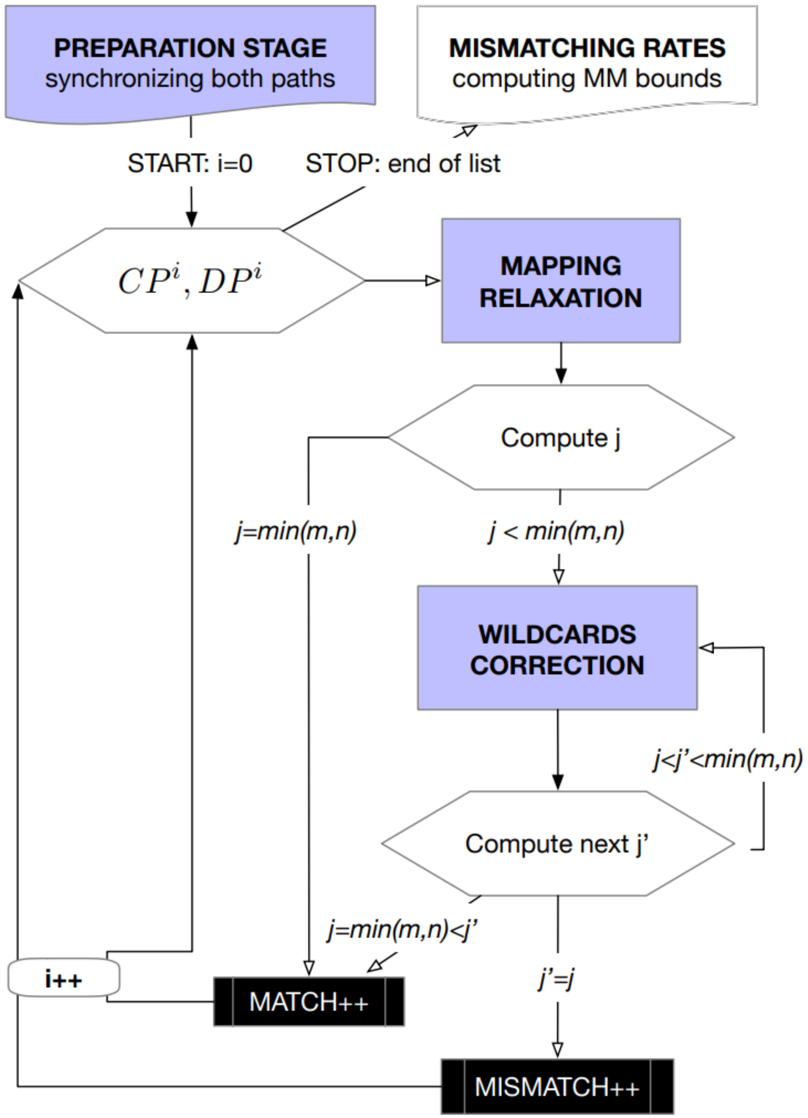
Our solution

A framework to detect BGP lies



- ✓ **Input:** CPs and DPs from a co-located VP
- ✓ **Output:** rate of BGP lies

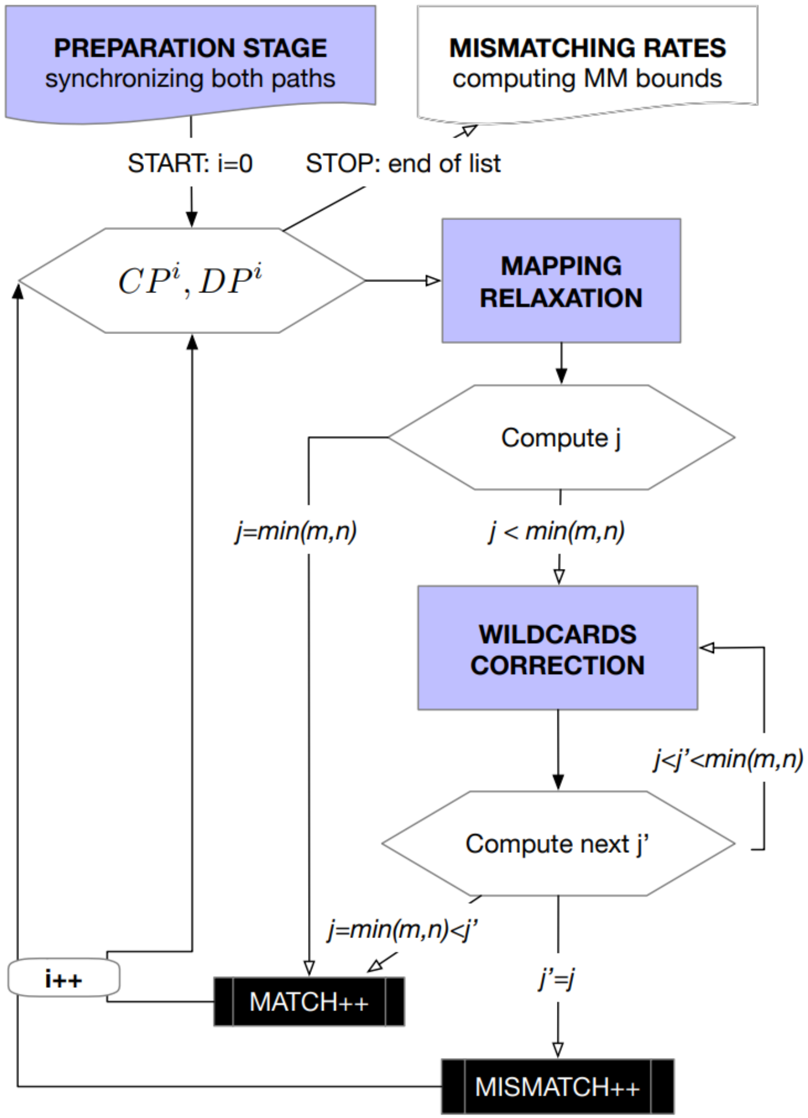
A framework to detect BGP lies



- ✓ **Input:** CPs and DPs from a co-located VP
- ✓ **Output:** rate of BGP lies

- ❑ **Preparation stage:**
  - Address space synchronization
  - Time synchronization
  - Basic IP-to-AS mapping
- ❑ **Mapping relaxation**
  - AS siblings
  - Third-party addresses
- ❑ **Wildcards correction stage**
  - Missing hops

A framework to detect BGP lies



- ✓ **Input:** CPs and DPs from a co-located VP
- ✓ **Output:** rate of BGP lies

- ❑ **Preparation stage:**
  - Address space synchronization
  - Time synchronization
  - Basic IP-to-AS mapping
- ❑ **Mapping relaxation**
  - AS siblings
  - Third-party addresses
- ❑ **Wildcards correction stage**
  - Missing hops

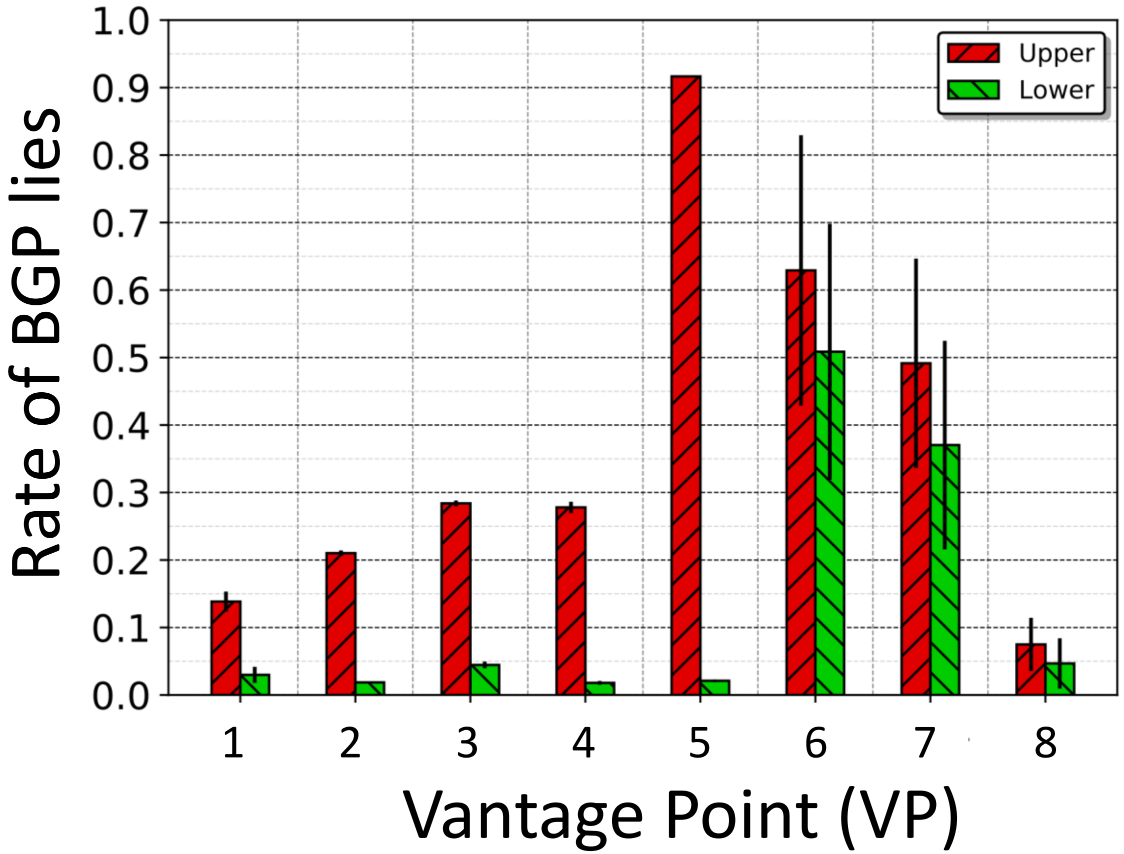
...we are conservative!

Results

Dataset

- Deployed 8 co-located VPs
- CPs collected every two hours
- DPs gathered targeting 80K destinations per day
- We run measurements multiple days (at least 13 days)

Filtering the noise with our framework



- VP 6,7: High rate, high variance
- VP 1-5. Quite effective, low variance
- Ground truth: BGP lies due to technical limitations in VP 7
- ...then in VP 6 too? ...and VP 1-5 malicious behaviour?

Conclusions

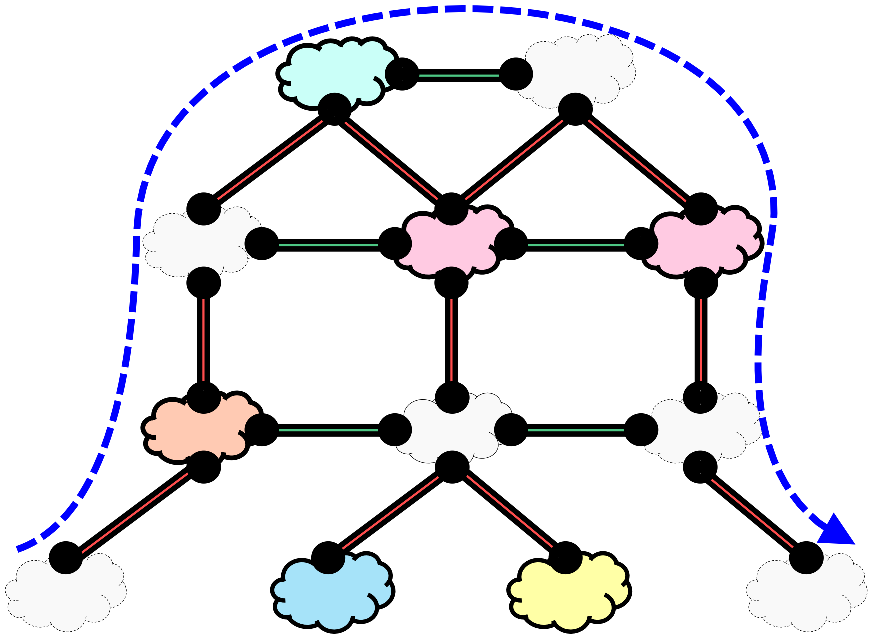
- ❖ A framework to detect BGP lies filtering the IP-to-AS mapping noise
- ❖ Deployed more co-located VPs than previous work
- ❖ Run the first time-analysis comparing CPs and DPs
- ❖ Patterns in results: technical limitations vs malicious Ases?

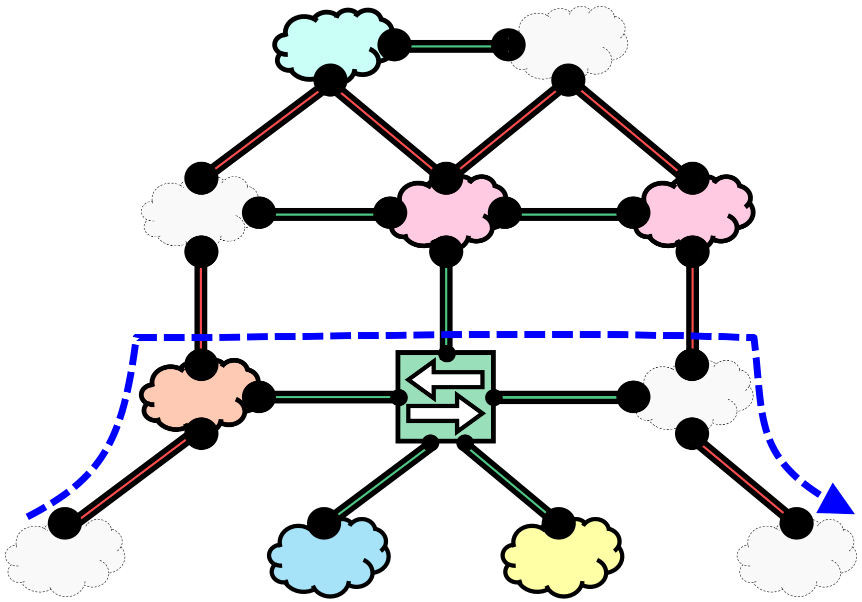
Outline

- Background, Research Goal and Questions
- Part I. Filtering the noise to reveal BGP lies
- Part II. Success and Failure of IXPs in Latin America
- Part III. The Art of Detecting Forwarding Detours
- Conclusions and Future Work

Why IXPs?

- Reshaped the structure of the Internet



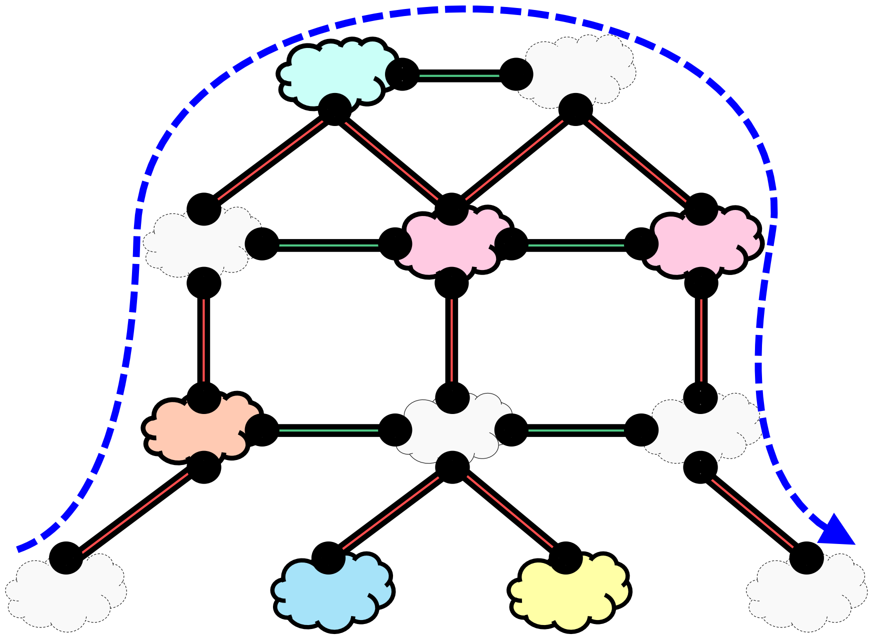


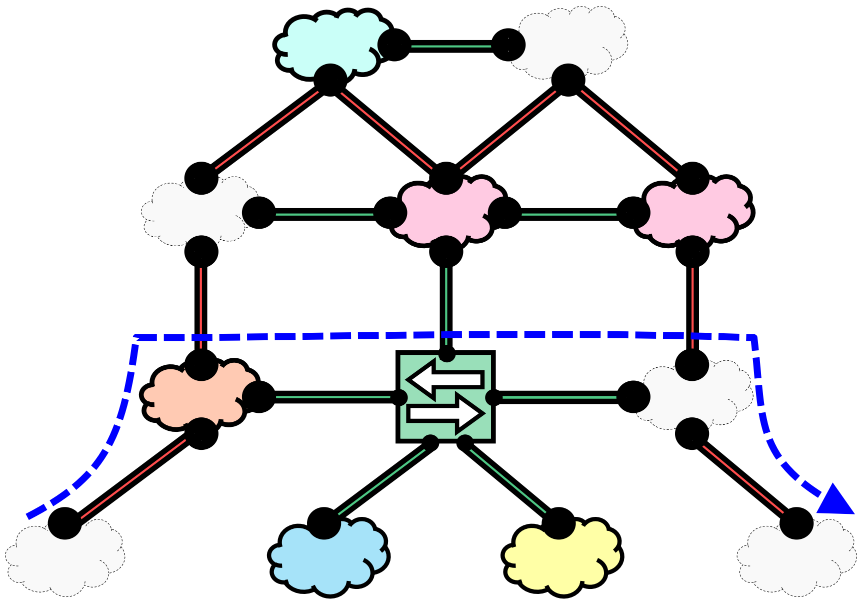
Why Latin America?

- Little previous work
- Discovered "new" datasets

Why IXPs?

- Reshaped the structure of the Internet





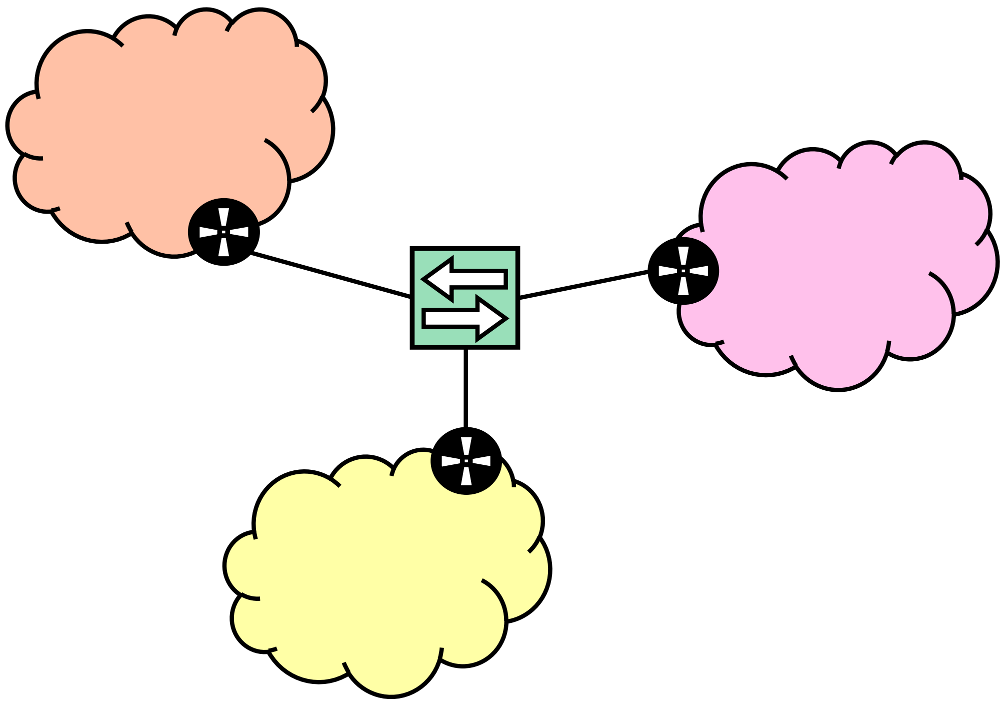
Why Latin America?

- Little previous work
- Discovered "new" datasets
- ...and I come from there

General Knowledge on IXPs

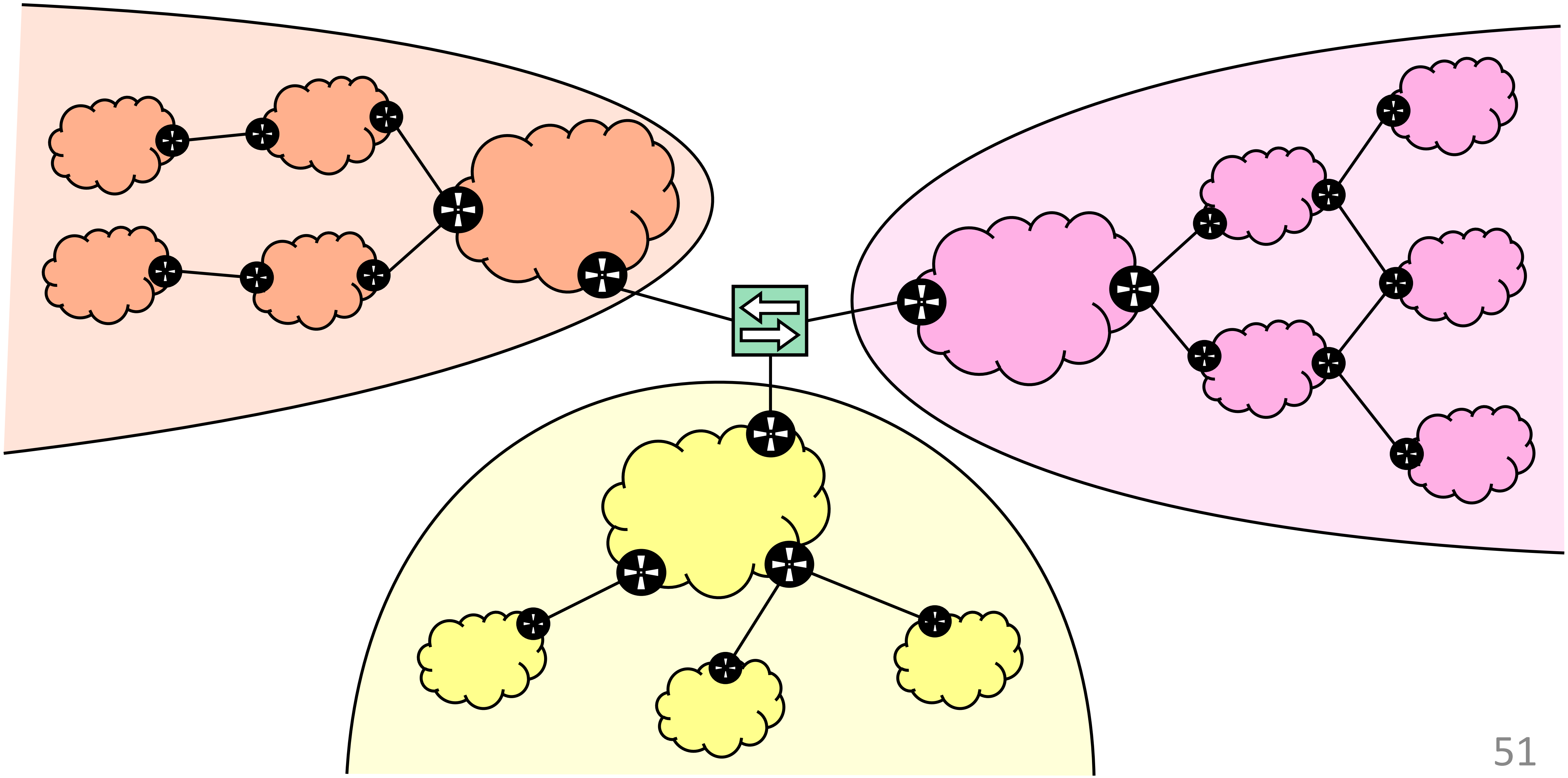
IXP Members

ASes that connect to the IXP and announce IP prefixes



Visible ASes of an IXP

IXP members + ASes seen in AS-paths announced by members



Preliminary Results

Dataset

- Control paths gathered in the IXPs
  - Members
  - Set of visible Ases
  - IP addresses announced

- Regional Internet registry files
  - Nationality of ASes

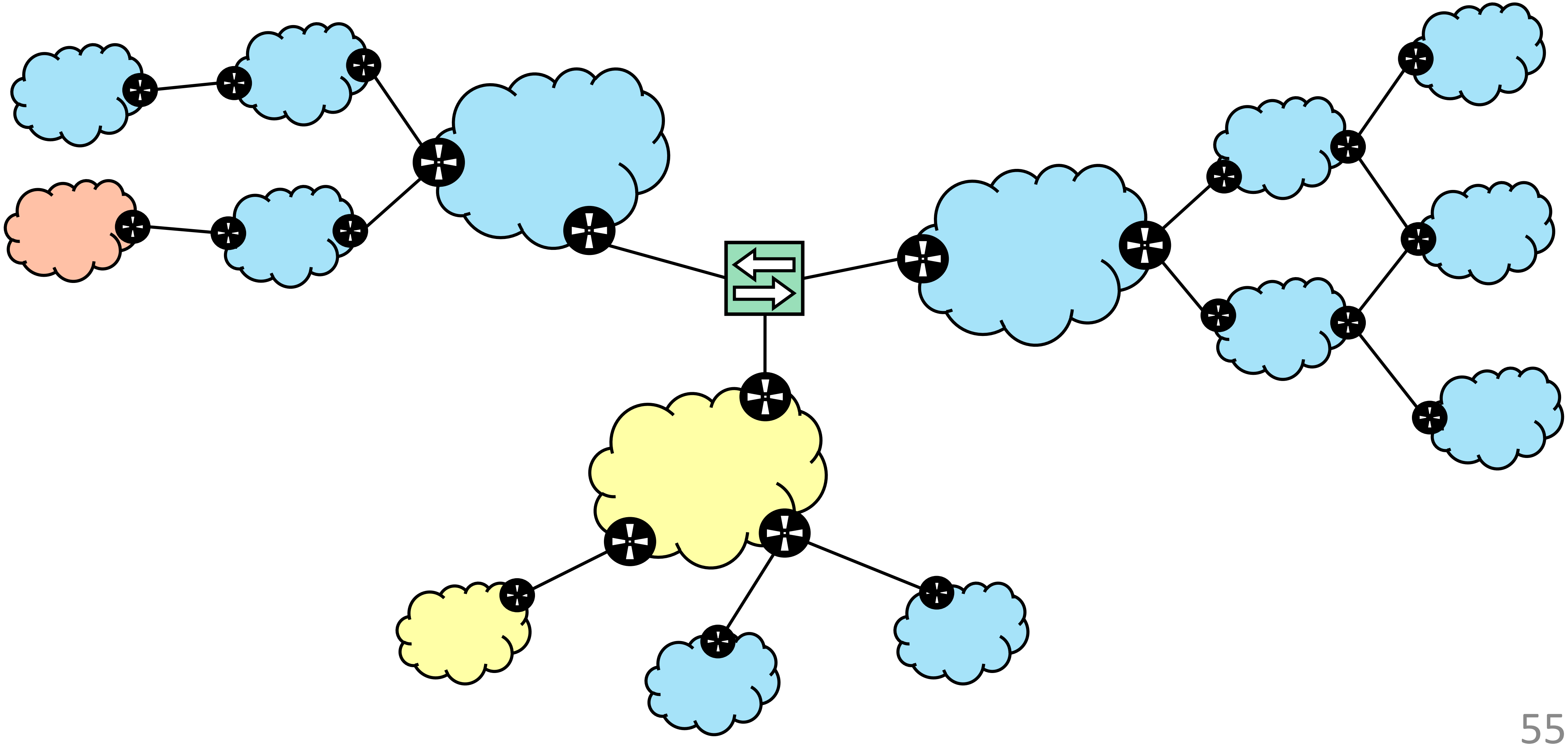
Success or Failure?

- Most IXPs in Latin America have low impact, or are failed IXPs
  - Less than 30 members
  - Less than 2M IP addresses announced

- The exception are Argentina, Brazil and Chile, the successful ones

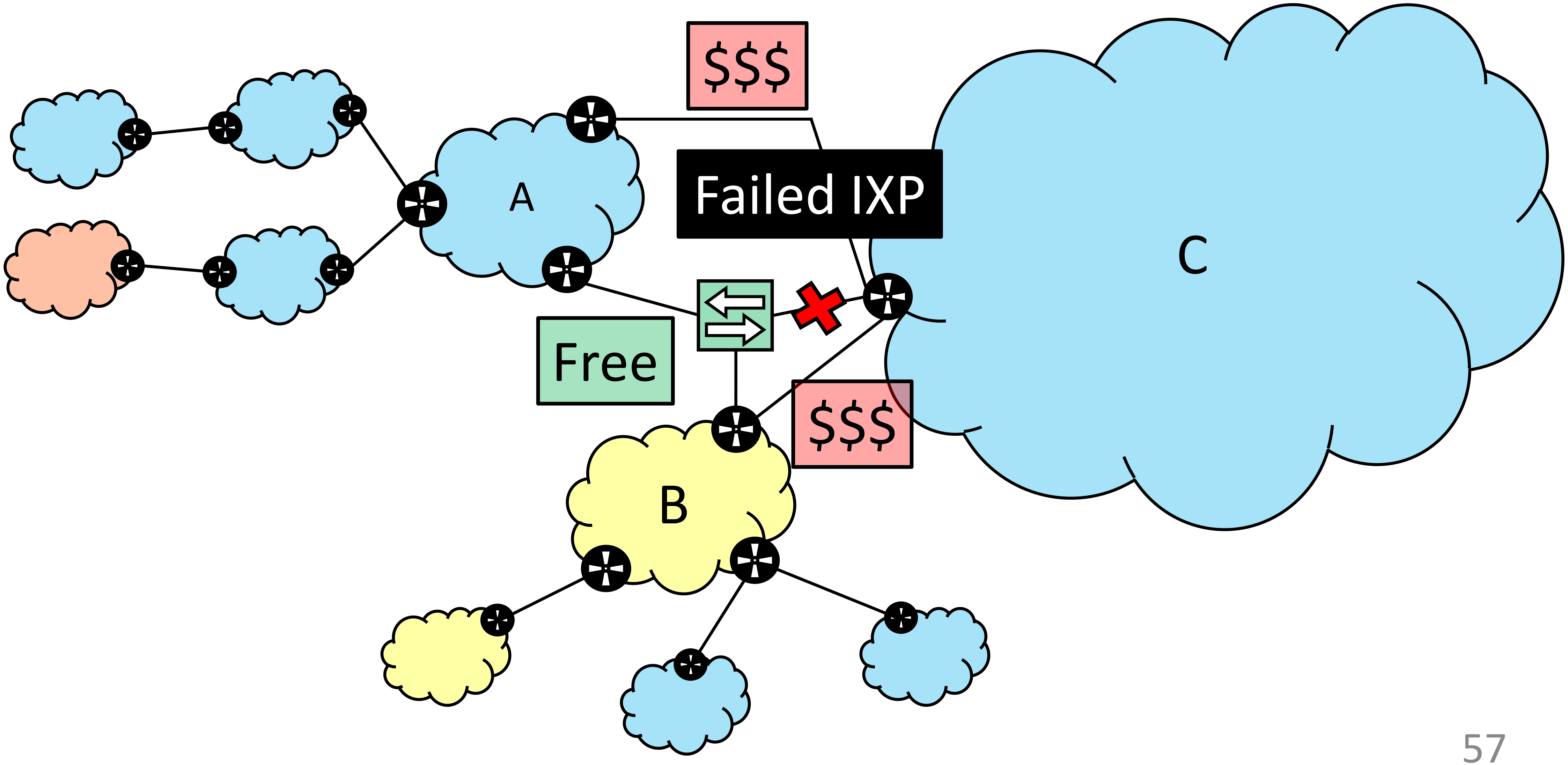
Most visible ASes in Latin American IXPs are local ASes

...consider color as nationality....



In the countries with Failed IXPs, are IP addresses fairly distributed among local Ases?

Maybe a monopolistic AS prefers not to peer in the IXP



How to measure whether the distribution is fair or not?

- We use the Herfindahl Hirschman Index (HHI)
  - Select a country
  - Choose 2 IPs of that country at random
  - Odds they belong to the same AS

How to measure whether the distribution is fair or not?

- We use the Herfindahl Hirschman Index (HHI)
  - Select a country
  - Choose 2 IPs of that country at random
  - Odds they belong to the same AS

  - The closer to 0, the more fair
  - The closer to 1, the more concentrated

Results

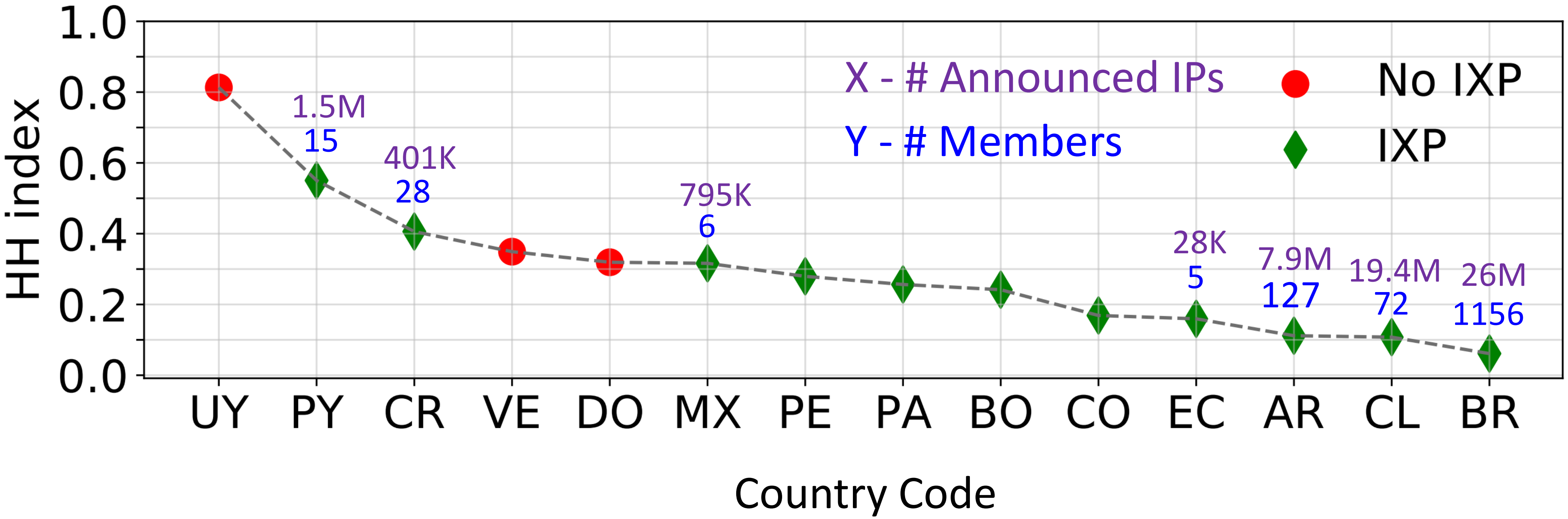
Dataset

- Control paths gathered in the IXPs
  - Members
  - Set of visible Ases
  - IP addresses announced

- Regional Internet registry files
  - Nationality of Ases

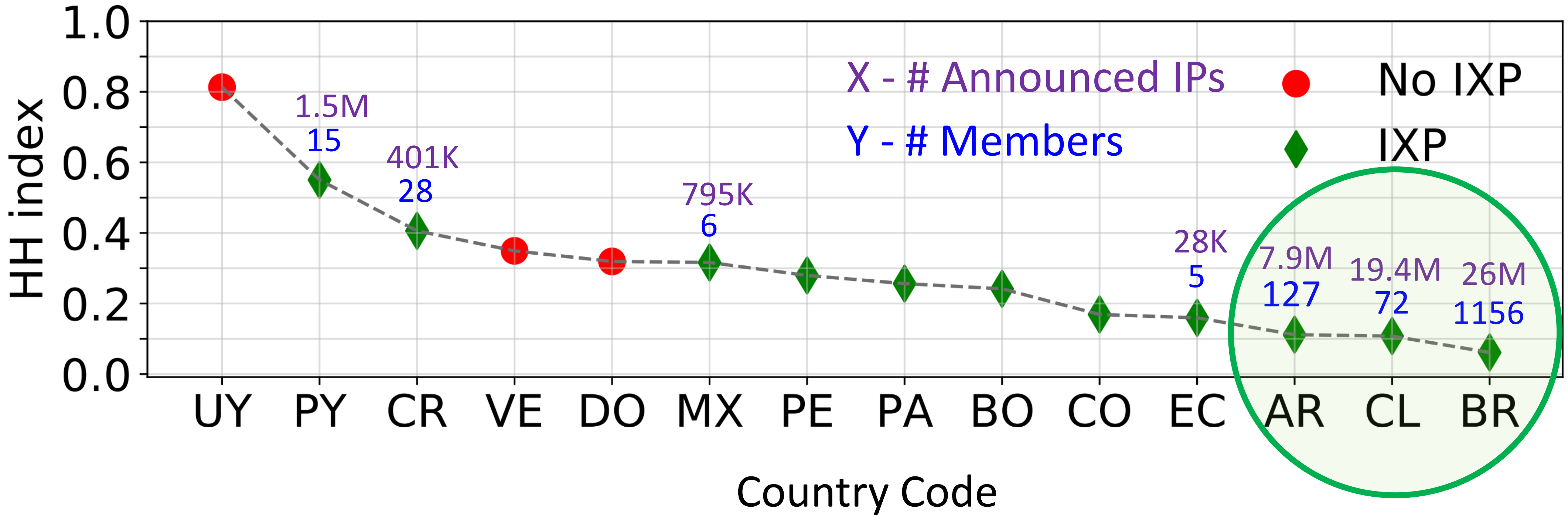
- Prefix-to-AS files
  - IP addresses that are actively used on the Internet

Concentration vs Success



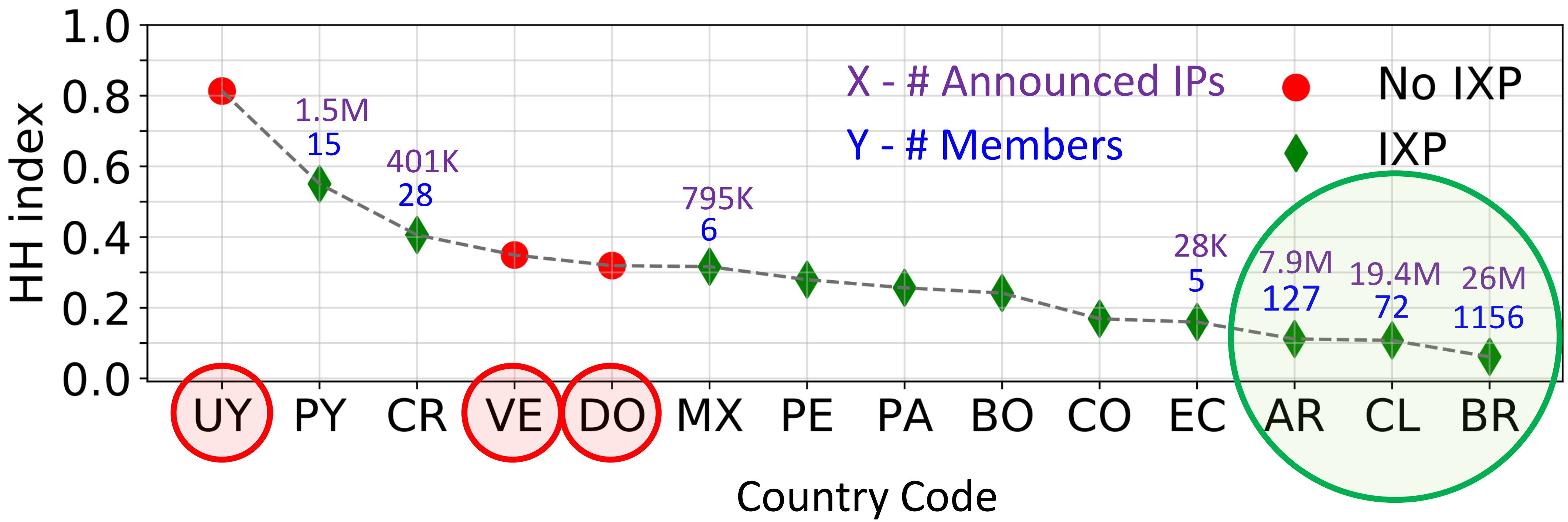
- Countries with more than 1M active IP addresses are displayed

Concentration vs Success



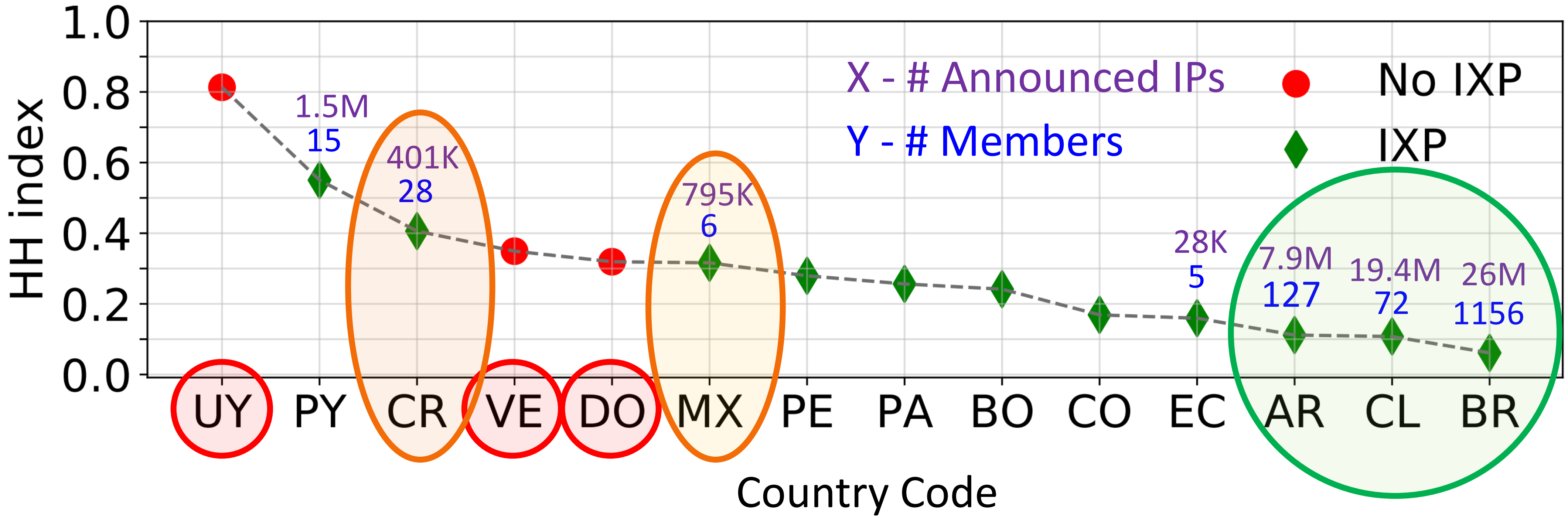
- Countries with more than 1M active IP addresses are displayed
- AR, CL, BR: largest IXPs, lowest HHI

Concentration vs Success



- Countries with more than 1M active IP addresses are displayed
- AR, CL, BR: largest IXPs, lowest HHI
- UY, VE, DO: no IXP at all

Concentration vs Success



- Countries with more than 1M active IP addresses are displayed
- AR, CL, BR: largest IXPs, lowest HHI
- UY, VE, DO: no IXP at all
- CR, MX: there is an IXP, but monopolistic local ASes not peering

Conclusions

- First to study Latin American IXPs in depth
- The region has many failed IXPs
- Visible ASes are mainly local ASes
- Possible correlation between failed IXPs and concentrated markets

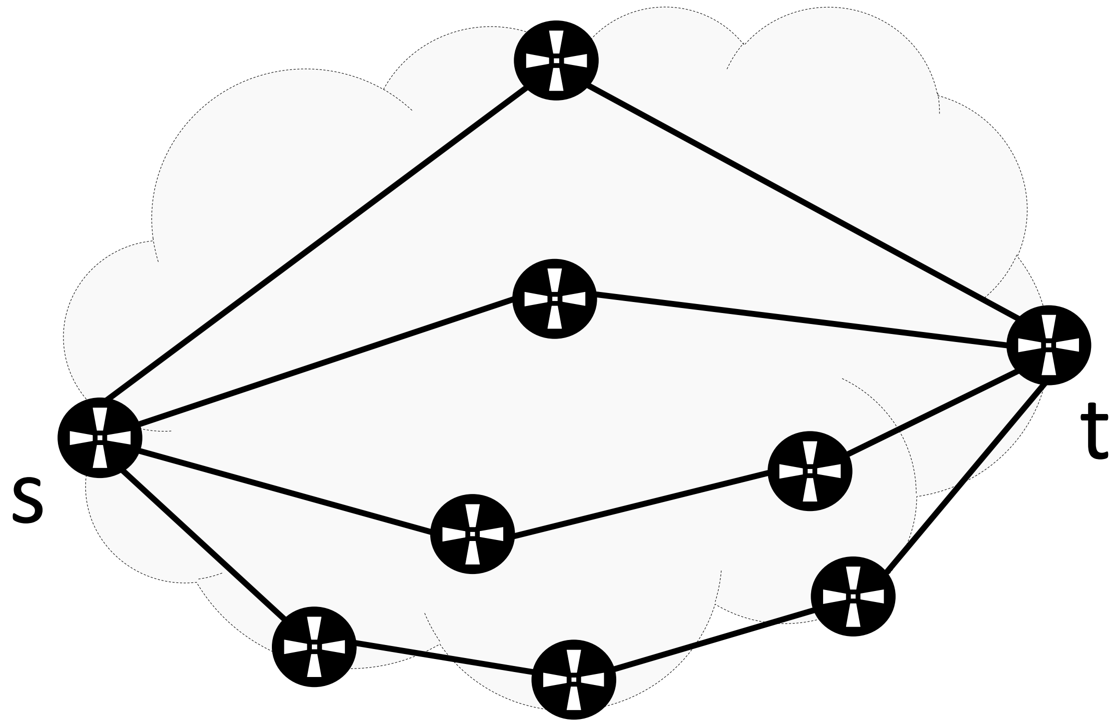
Outline

- Background, Research Goal and Questions
- Part I. Filtering the noise to reveal BGP lies
- Part II. Success and Failure of IXPs in Latin America
- Part III. The Art of Detecting Forwarding Detours
- Conclusions and Future Work

The basics

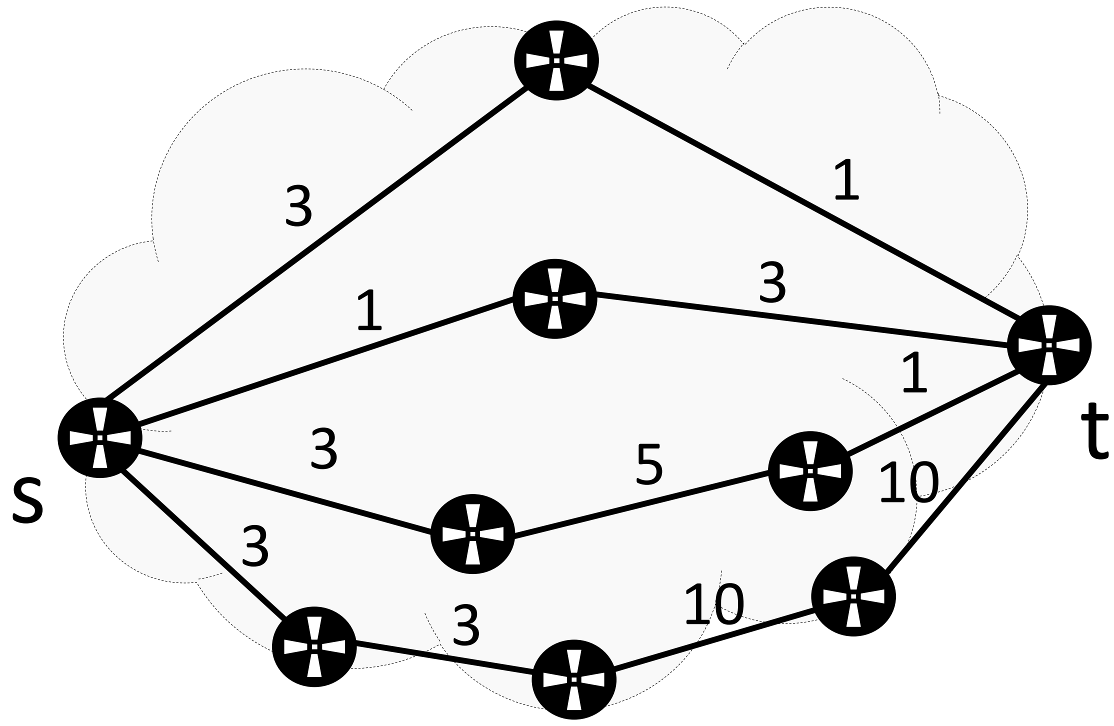
Internal Gateway Protocols (IGPs)

- Routing inside networks



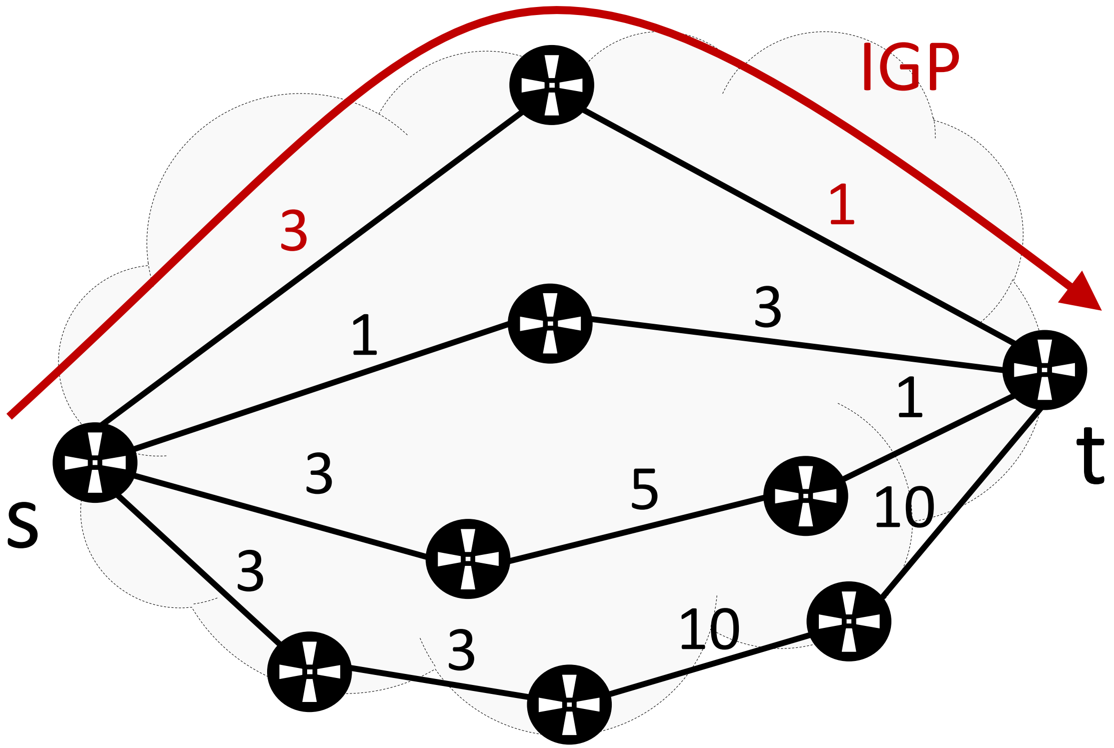
Internal Gateway Protocols (IGPs)

- Routing inside networks
- Links have a cost according to some metric



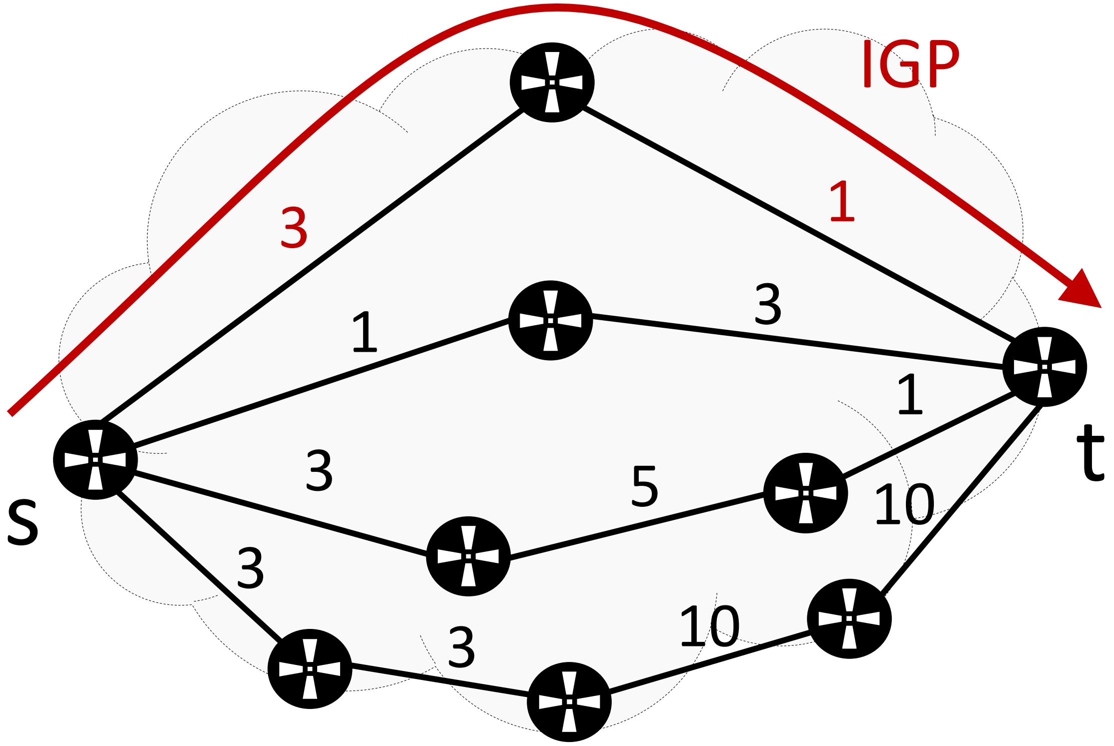
Internal Gateway Protocols (IGPs)

- Routing inside networks
- Links have a cost according to some metric
- The path with minimum cost is used



Internal Gateway Protocols (IGPs)

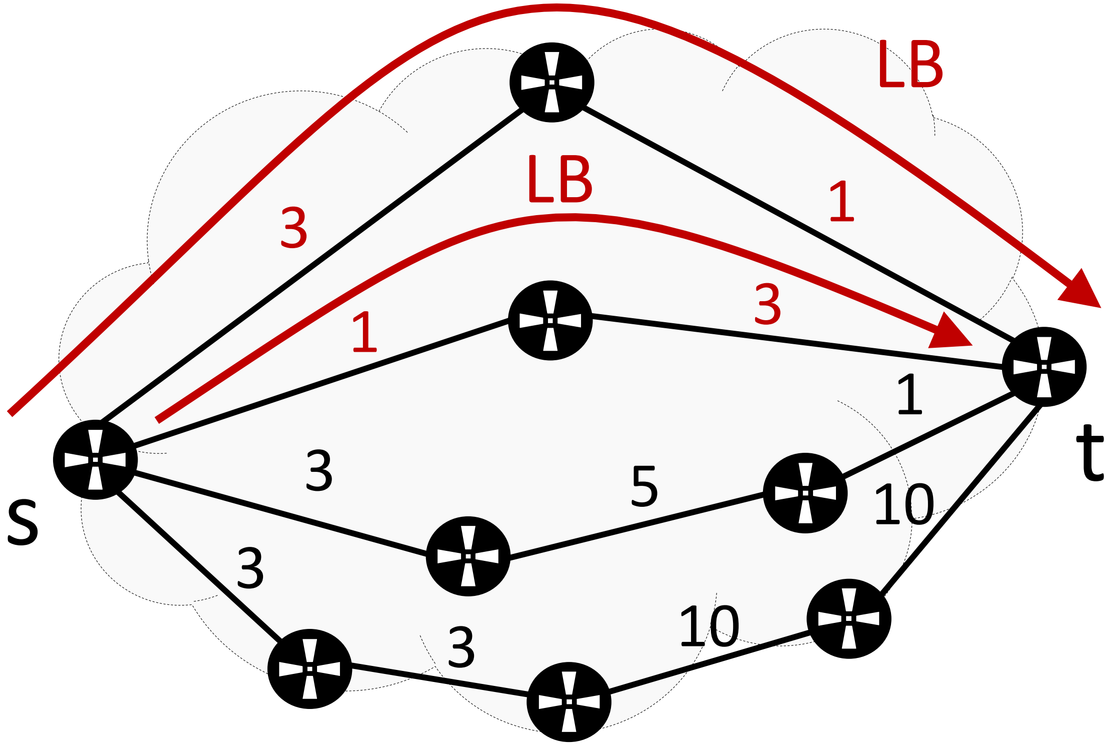
- Routing inside networks
- Links have a cost according to some metric
- The path with minimum cost is used



IGP		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1				
	P_2				
	P_3				
	P_4				
	P_5				
	P_6				
	P_7				
	P_8				

Load Balancing (LB)

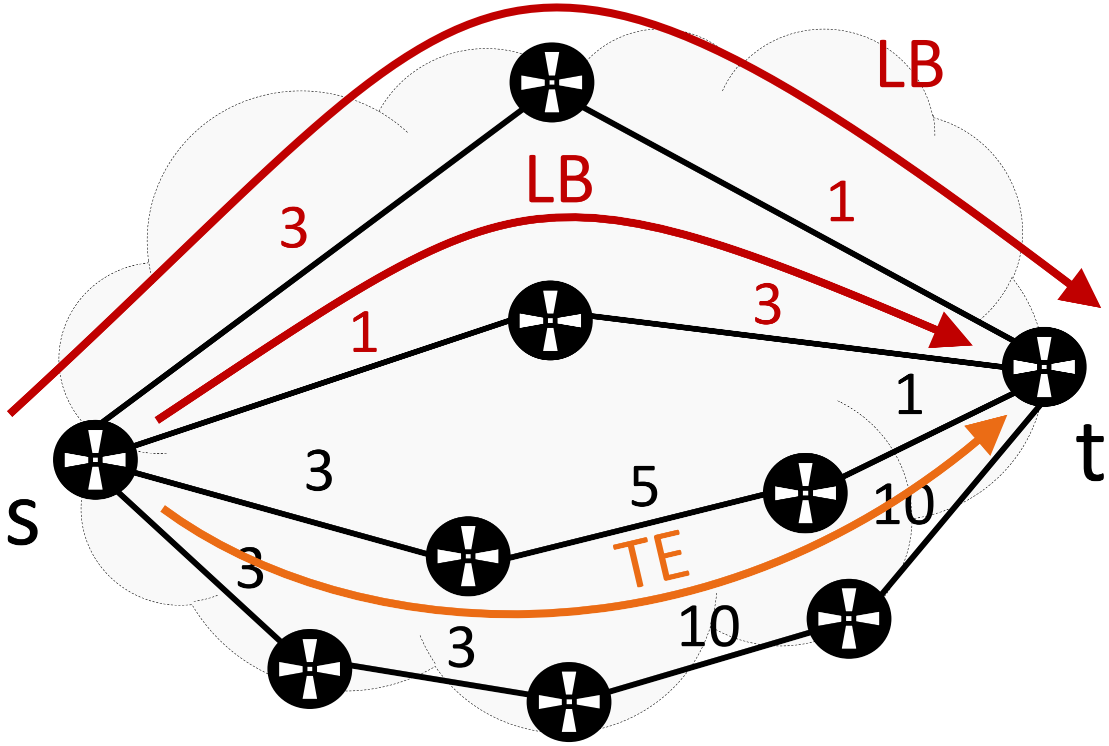
- From one to many best IGP paths
- Usually deployed with equal-cost multipath (ECMP)



LB		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1	⊙⊙	⊙⊙		
	P_2	⊙⊙	⊙⊙		
	P_3	⊙⊙	⊙⊙		
	P_4	⊙⊙	⊙⊙		
	P_5	⊙⊙	⊙⊙		
	P_6	⊙⊙	⊙⊙		
	P_7	⊙⊙	⊙⊙		
	P_8	⊙⊙	⊙⊙		

Traffic Engineering (TE)

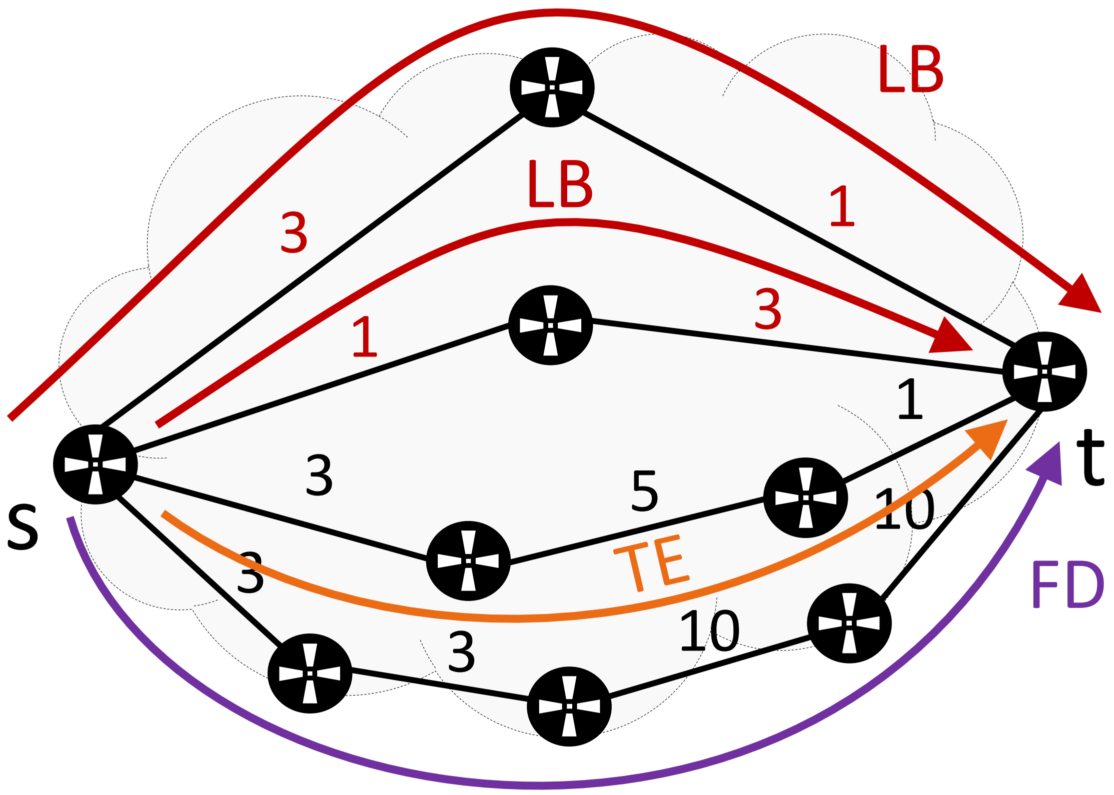
- Allows to craft paths “by hand”
- The crafted paths meet some requirements, e.g. low delay



LB TE		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1	⊙ ⊙	⊙ ⊙		
	P_2	⊙ ⊙	⊙ ⊙		
	P_3	⊙ ⊙	⊙ ⊙		
	P_4			⊙⊙ ⊙⊙	
	P_5	⊙ ⊙	⊙ ⊙		
	P_6	⊙ ⊙	⊙ ⊙		
	P_7	⊙ ⊙	⊙ ⊙		
	P_8	⊙ ⊙	⊙ ⊙		

Forwarding Detours (FDs)

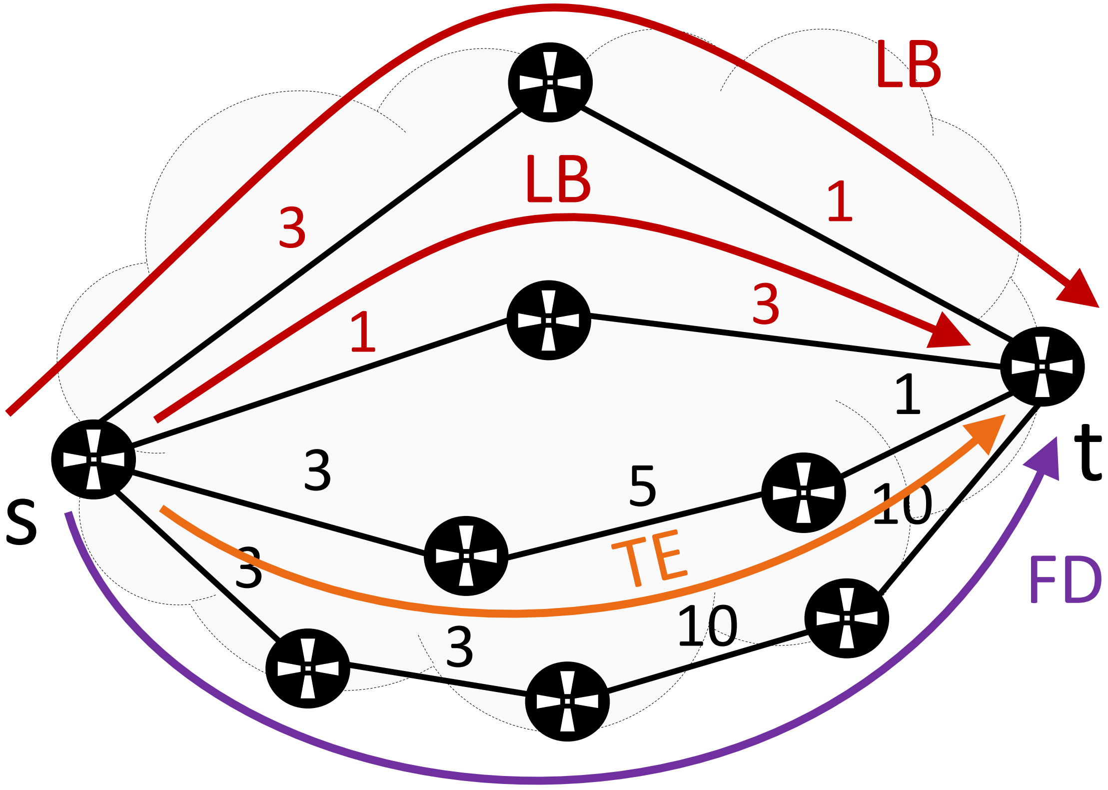
- When the forwarding route diverges from LB and TE paths



LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1				
	P_2				
	P_3				
	P_4				
	P_5				
	P_6				
	P_7				
	P_8				

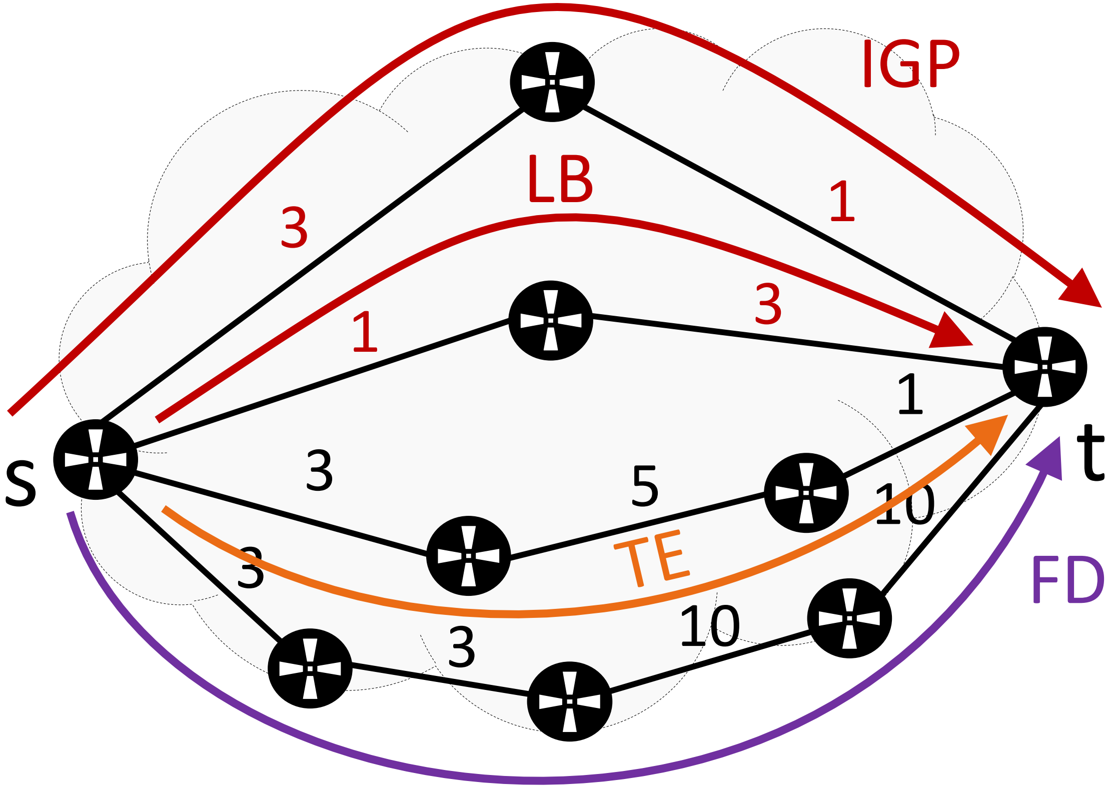
Why detecting FDs?

- FDs relate to unexpected paths being used
- Possible negative impact on performance



Methodology to detect FDs

Forwarding Pattern - Run measurements and find the matrix



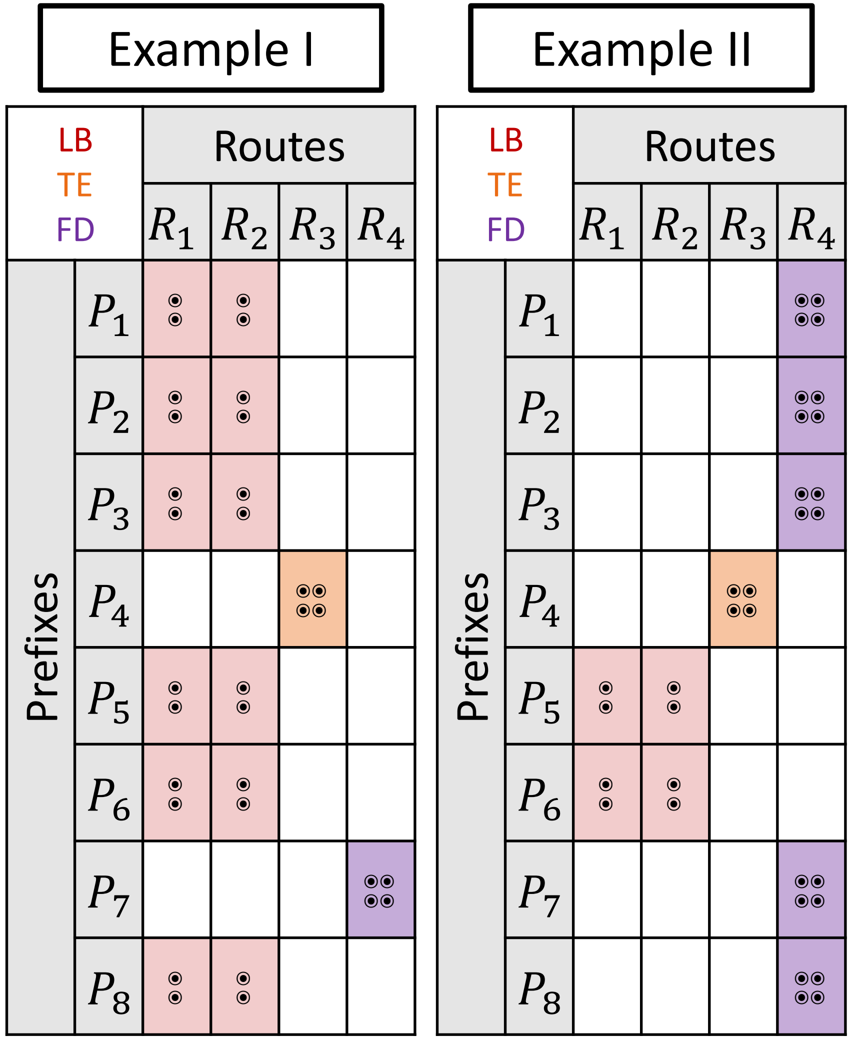
Example I

LB / TE / FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1	⦿ ⦿	⦿ ⦿		
	P_2	⦿ ⦿	⦿ ⦿		
	P_3	⦿ ⦿	⦿ ⦿		
	P_4			⦿⦿ ⦿⦿	
	P_5	⦿ ⦿	⦿ ⦿		
	P_6	⦿ ⦿	⦿ ⦿		
	P_7				⦿⦿ ⦿⦿
	P_8	⦿ ⦿	⦿ ⦿		

Example II

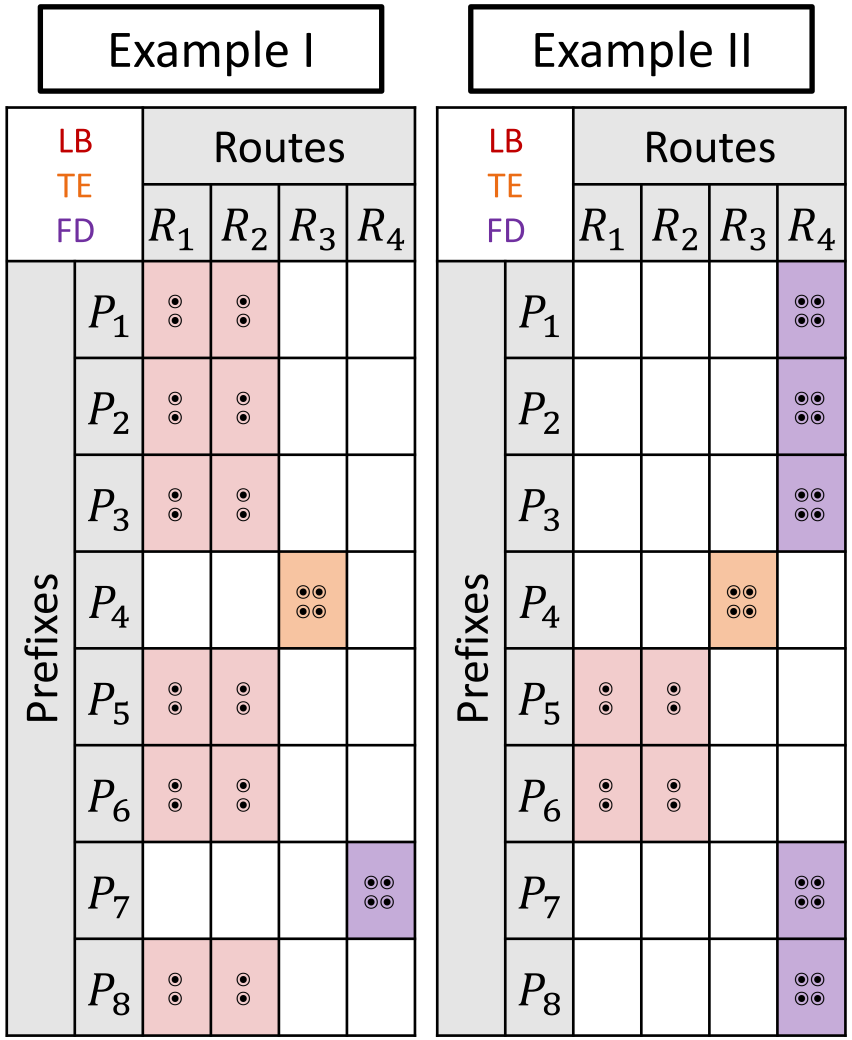
LB / TE / FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1				⦿⦿ ⦿⦿
	P_2				⦿⦿ ⦿⦿
	P_3				⦿⦿ ⦿⦿
	P_4			⦿⦿ ⦿⦿	
	P_5	⦿ ⦿	⦿ ⦿		
	P_6	⦿ ⦿	⦿ ⦿		
	P_7				⦿⦿ ⦿⦿
	P_8				⦿⦿ ⦿⦿

Concluding if FDs occur



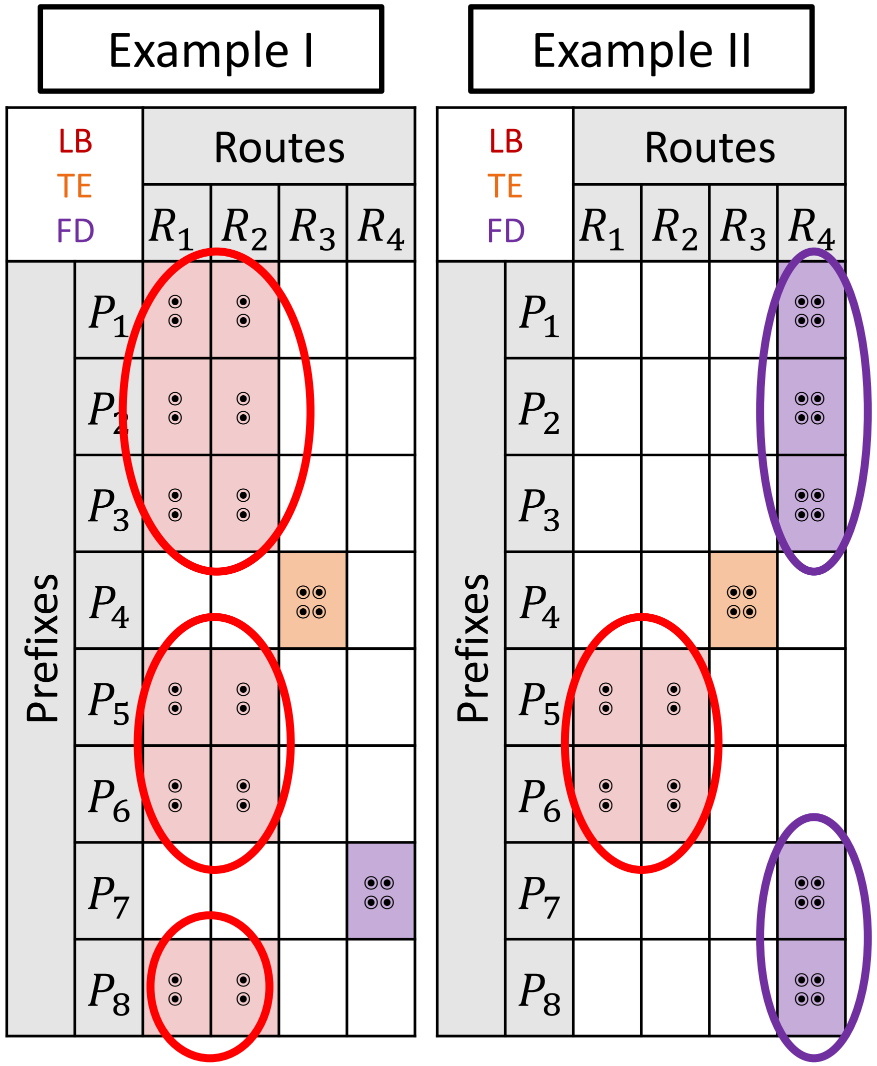
Concluding if FDs occur

1. Identify prefixes related to the same routes



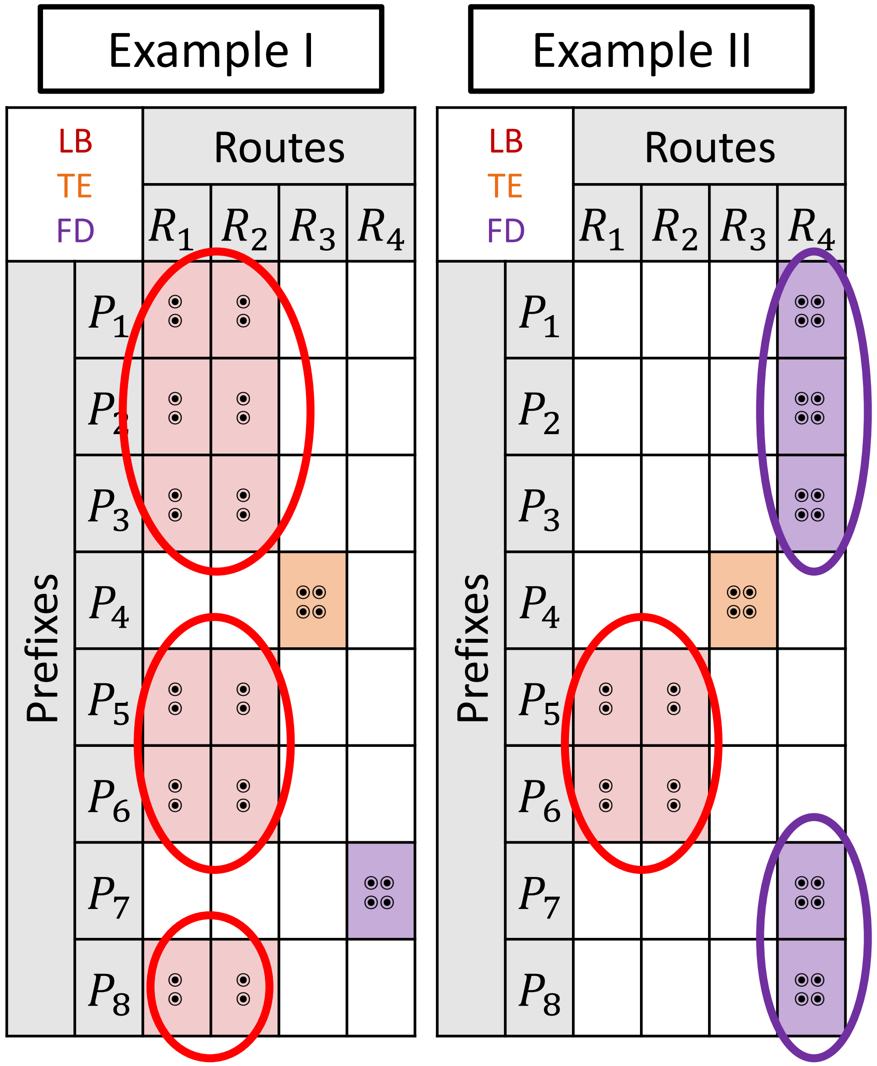
Concluding if FDs occur

1. Identify prefixes related to the same routes



Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets



Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets

Example I

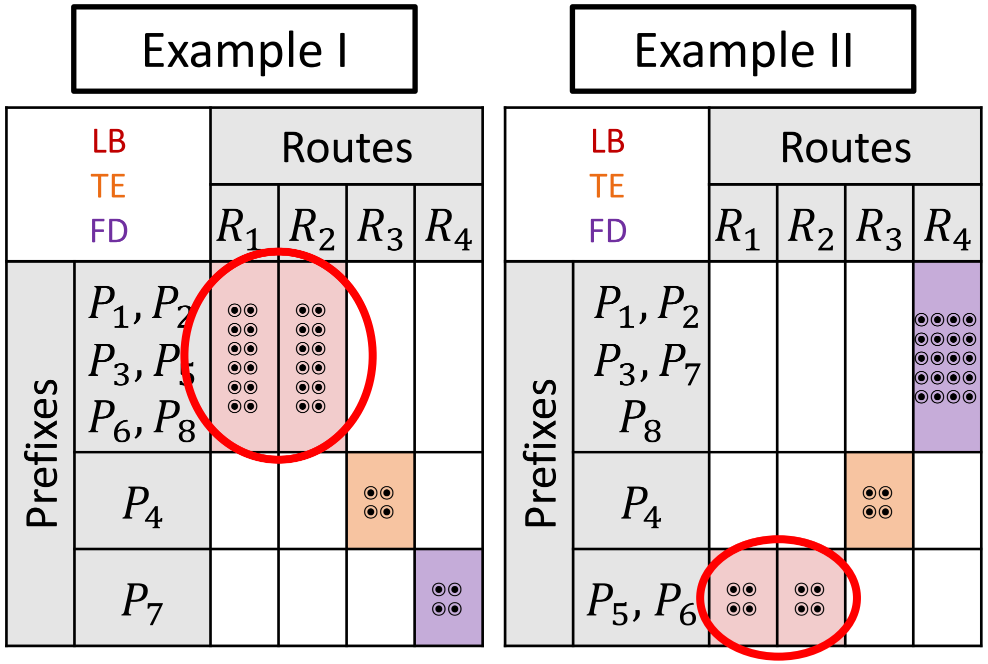
LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1, P_2 P_3, P_5 P_6, P_8				
	P_4				
	P_7				

Example II

LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1, P_2 P_3, P_7 P_8				
	P_4				
	P_5, P_6				

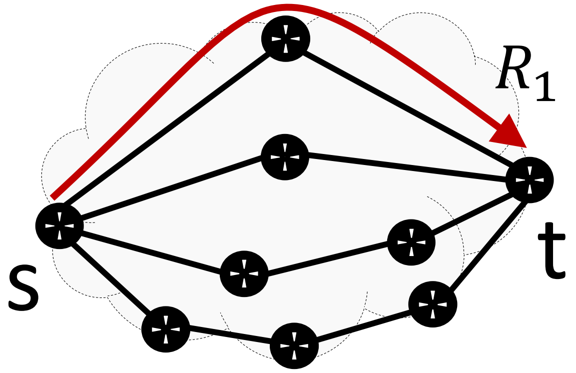
Concluding if FDs occur

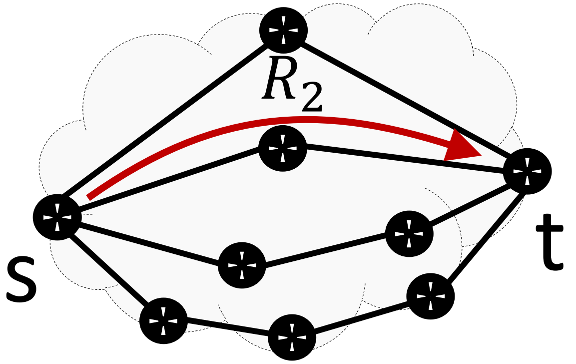
1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t



Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t





Example I

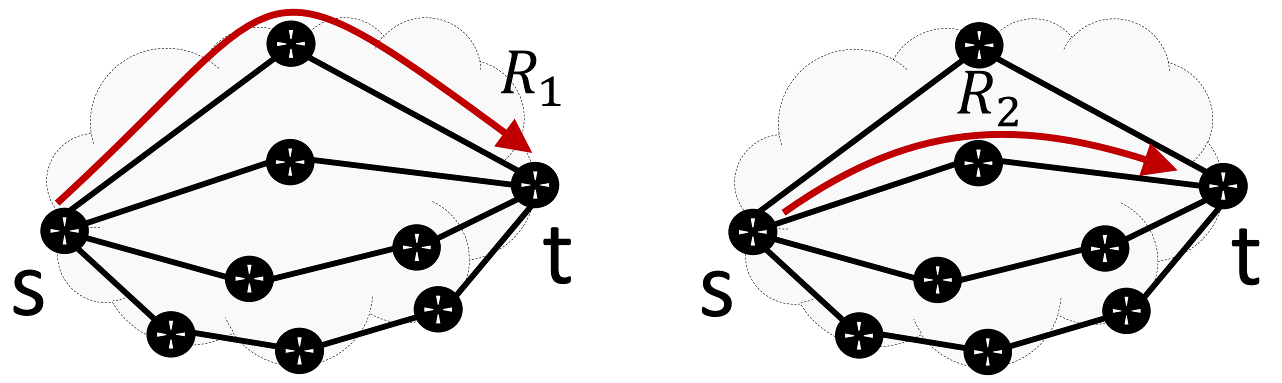
LB / TE / FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	$P_1, P_2, P_3, P_5, P_6, P_8$				
	P_4				
	P_7				

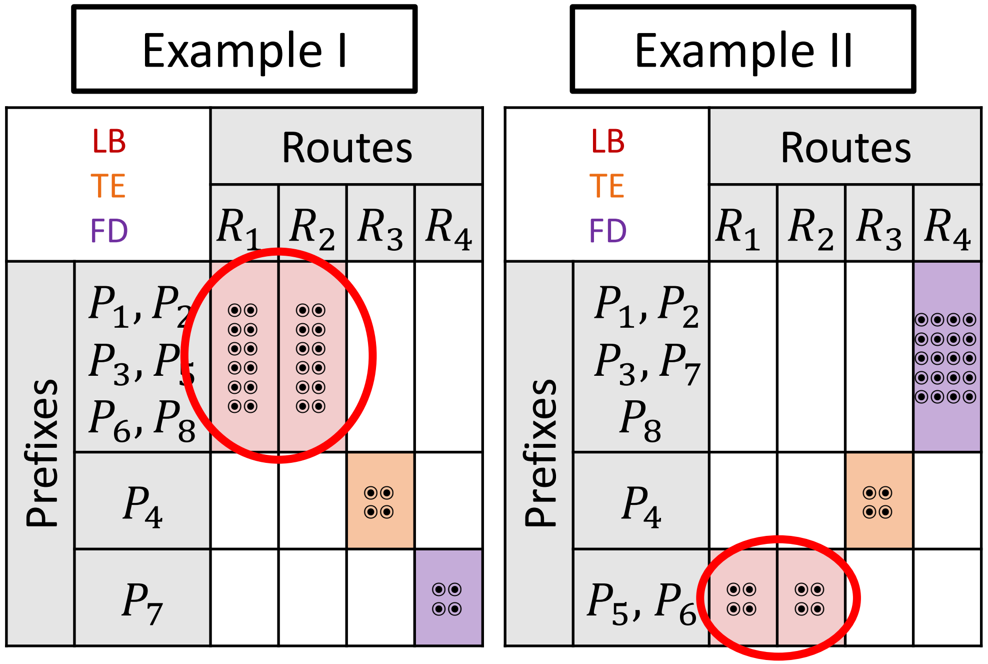
Example II

LB / TE / FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1, P_2, P_3, P_7, P_8				
	P_4				
	P_5, P_6				

Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t

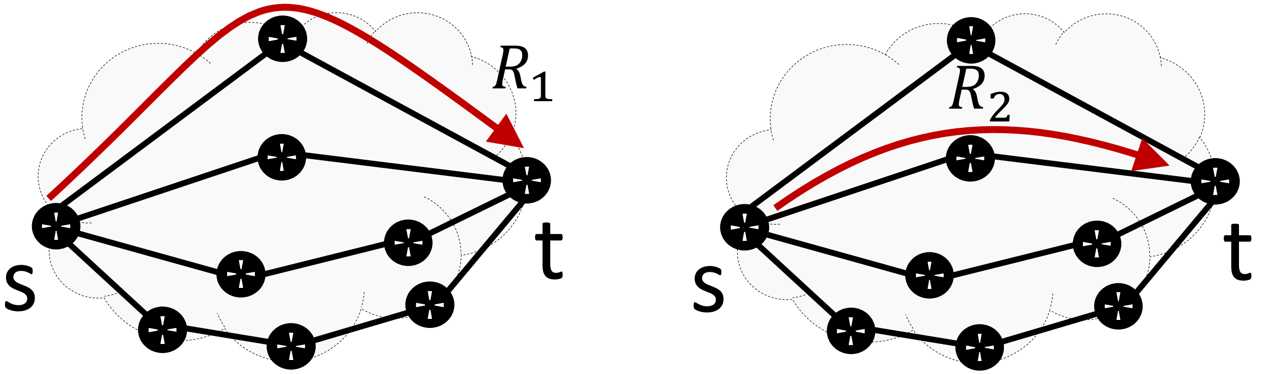


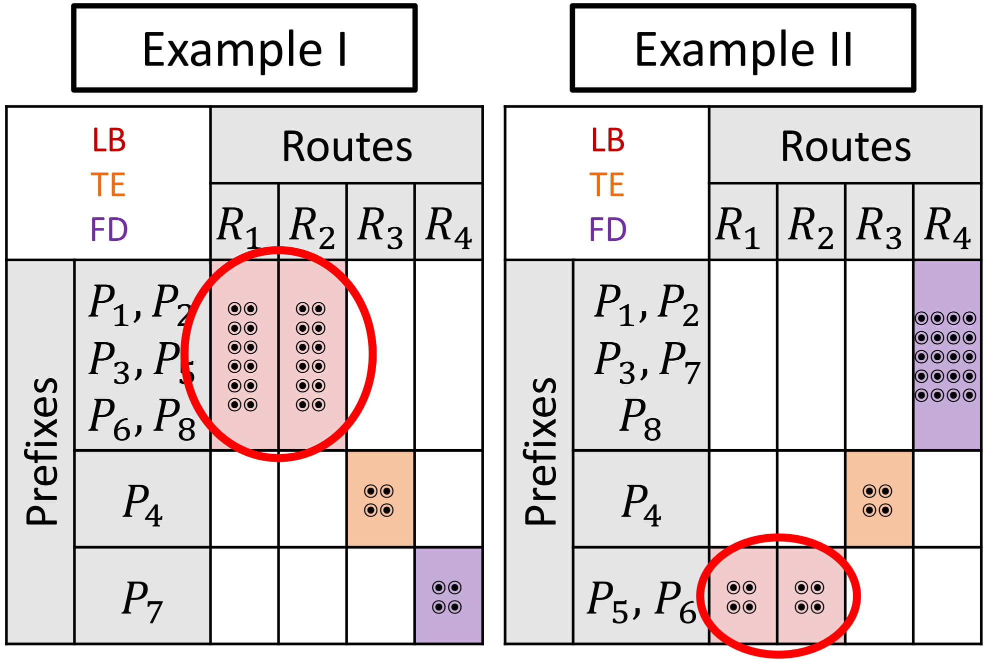


4. Compute #pfxs in each set: (6, 1, 1) and (5, 1, 2)

Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t

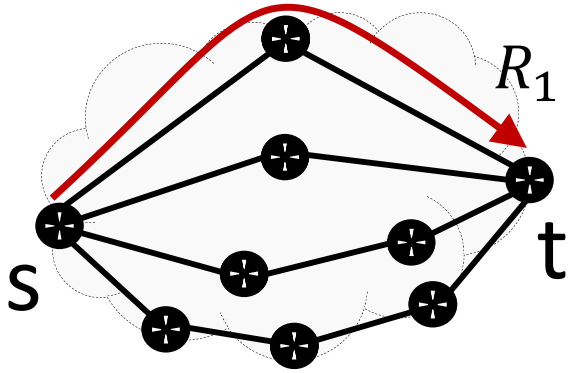


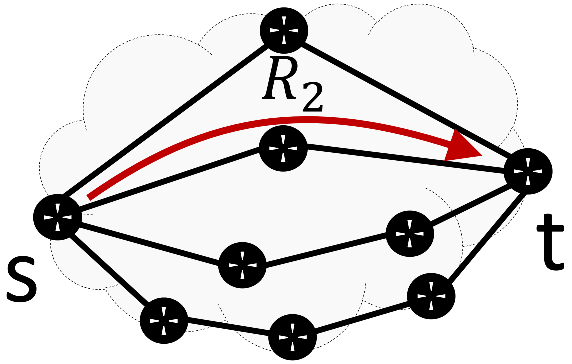


4. Compute #pfxs in each set: (6, 1, 1) and (5, 1, 2)
5. Turn it into proportions: (0.75, 0.125, 0.125) and (0.625, 0.125, 0.25)

Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t





Example I

LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	$P_1, P_2, P_3, P_5, P_6, P_8$				
	P_4				
	P_7				

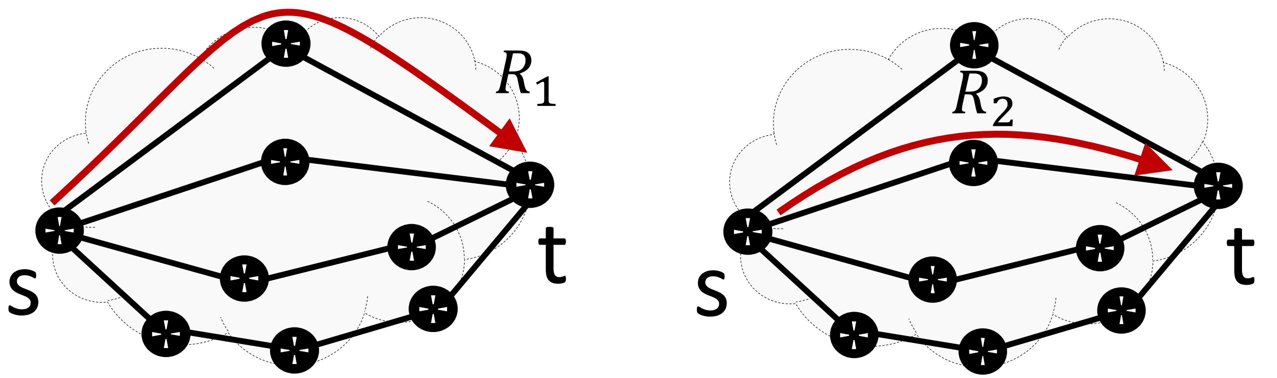
Example II

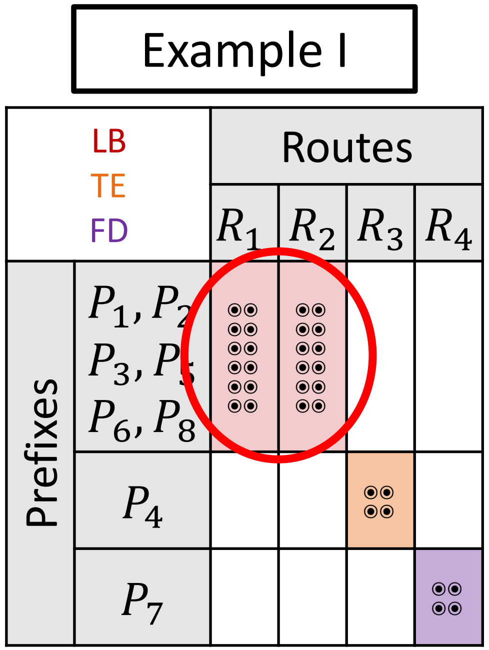
LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1, P_2, P_3, P_7, P_8				
	P_4				
	P_5, P_6				

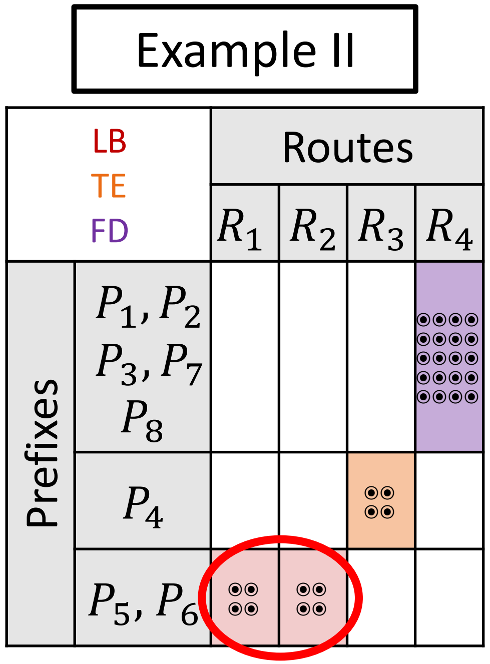
4. Compute #pfxs in each set: (6, 1, 1) and (5, 1, 2)
5. Turn it into proportions: (0.75, 0.125, 0.125) and (0.625, 0.125, 0.25)
6. Compute the n number of sets … in this case $n = 3$ for both examples…

Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t



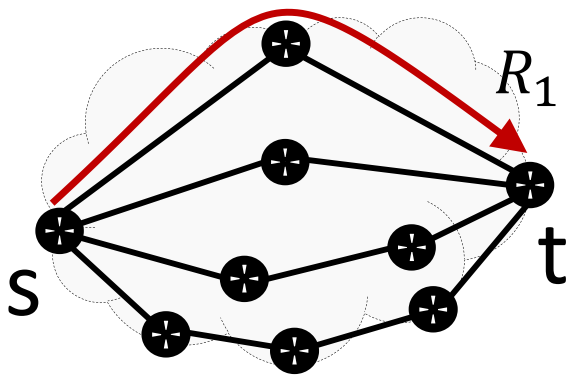


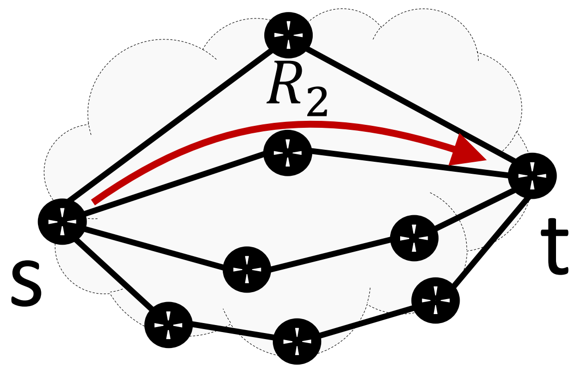


4. Compute #pfxs in each set: (6, 1, 1) and (5, 1, 2)
5. Turn it into proportions: (0.75, 0.125, 0.125) and (0.625, 0.125, 0.25)
6. Compute the n number of sets ... in this case $n = 3$ for both examples...
7. Conclude that FDs occur if LB is associated to less than $\frac{1}{n} = 0.33$ pfxs...

Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t





Example I

LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	$P_1, P_2,$ $P_3, P_5,$ P_6, P_8				
	P_4				
	P_7				

Example II

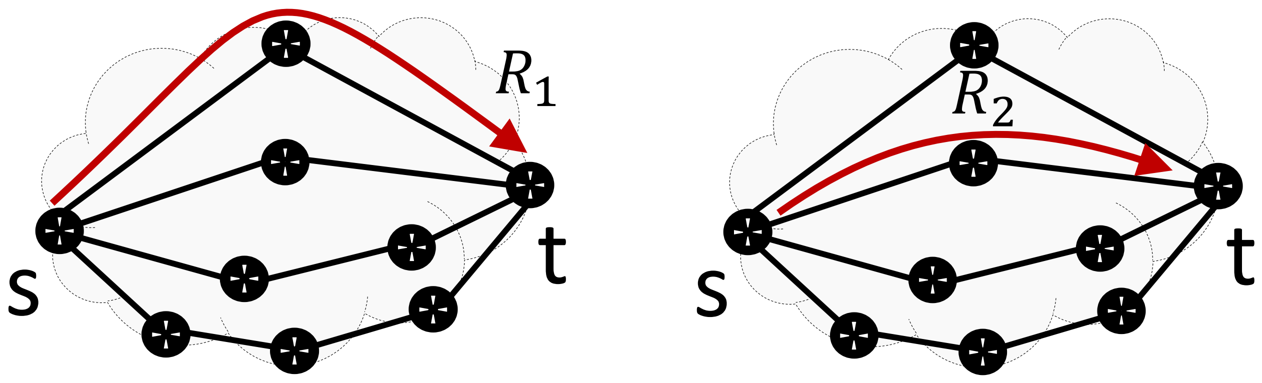
LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	$P_1, P_2,$ P_3, P_7 P_8				
	P_4				
	P_5, P_6				

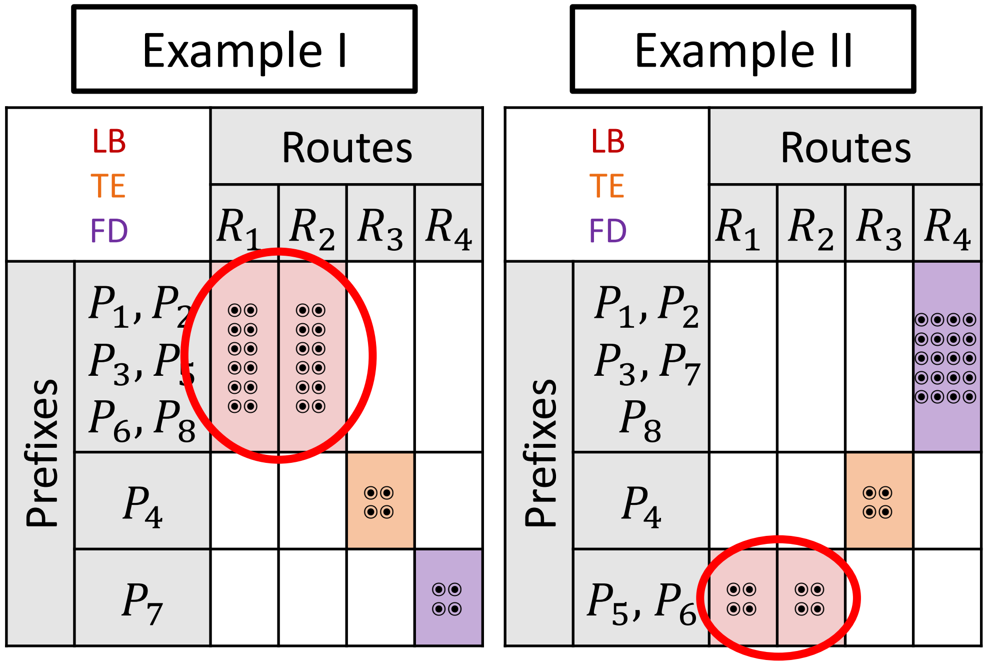
4. Compute #pfxs in each set: (6, 1, 1) and (5, 1, 2)
5. Turn it into proportions: (0.75, 0.125, 0.125) and (0.625, 0.125, 0.25)
6. Compute the n number of sets ... in this case $n = 3$ for both examples...
7. Conclude that FDs occur if LB is associated to less than $\frac{1}{n} = 0.33$ pfxs...

0.33 < 0.75 ... no FDs and 0.33 > 0.25 ... there are FDs

Concluding if FDs occur

1. Identify prefixes related to the same routes
2. Group the related prefixes in sets
3. Identify the LB set targeting router t





...we are conservative!

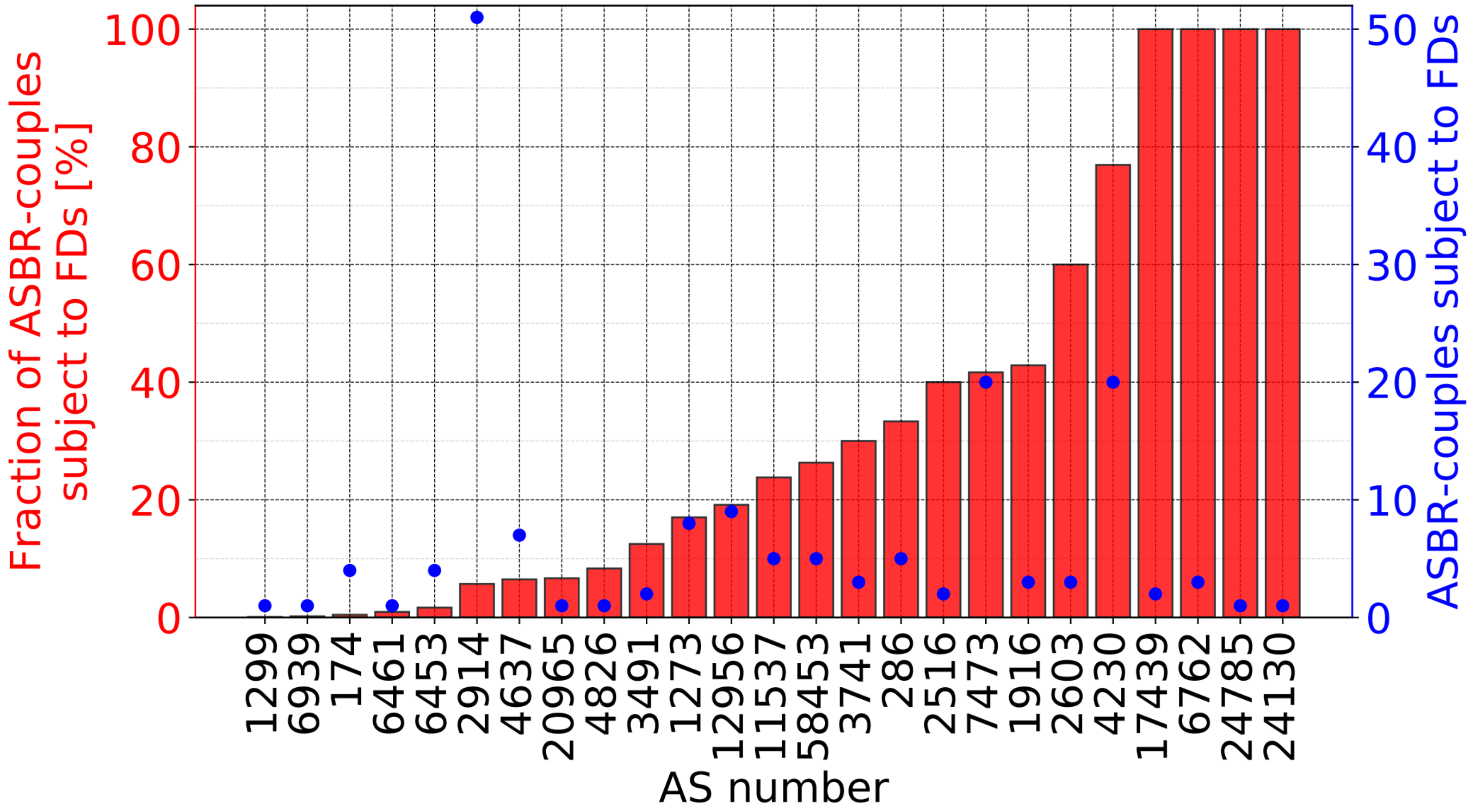
4. Compute #pfxs in each set: (6, 1, 1) and (5, 1, 2)
5. Turn it into proportions: (0.75, 0.125, 0.125) and (0.625, 0.125, 0.25)
6. Compute the n number of sets ... in this case $n = 3$ for both examples...
7. Conclude that FDs occur if LB is associated to less than $\frac{1}{n} = 0.33$ pfxs...

0.33 < 0.75 ... no FDs and 0.33 > 0.25 ... there are FDs

Results

In the wild, FDs are a thing!

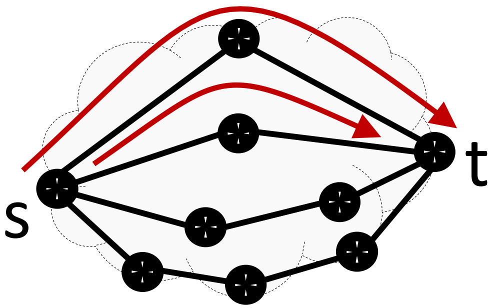
- We measure from 100 VPs
- We look for FDs between AS border routers (ASBRs) and request #pfxs > 100
- We find FDs in 25/54 Ases, with an heterogeneous distribution



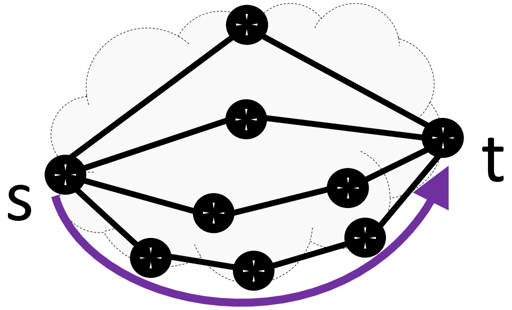
Digging into the results: a binary pattern

- According to the FDs we found, all traffic detours or none does

LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1, P_2 P_3, P_4 P_5, P_6 P_7, P_8				



LB TE FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1, P_2 P_3, P_4 P_5, P_6 P_7, P_8				



...in other words...

No cases like this!

LB TE, FD		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1, P_2 P_3, P_5 P_6, P_8				
	P_4				
	P_7				

Conclusions

- ❖Routing inconsistencies produce FDs

- ❖First methodology to systematically detect FDs

- ❖We built the first FD-detector and run measurements

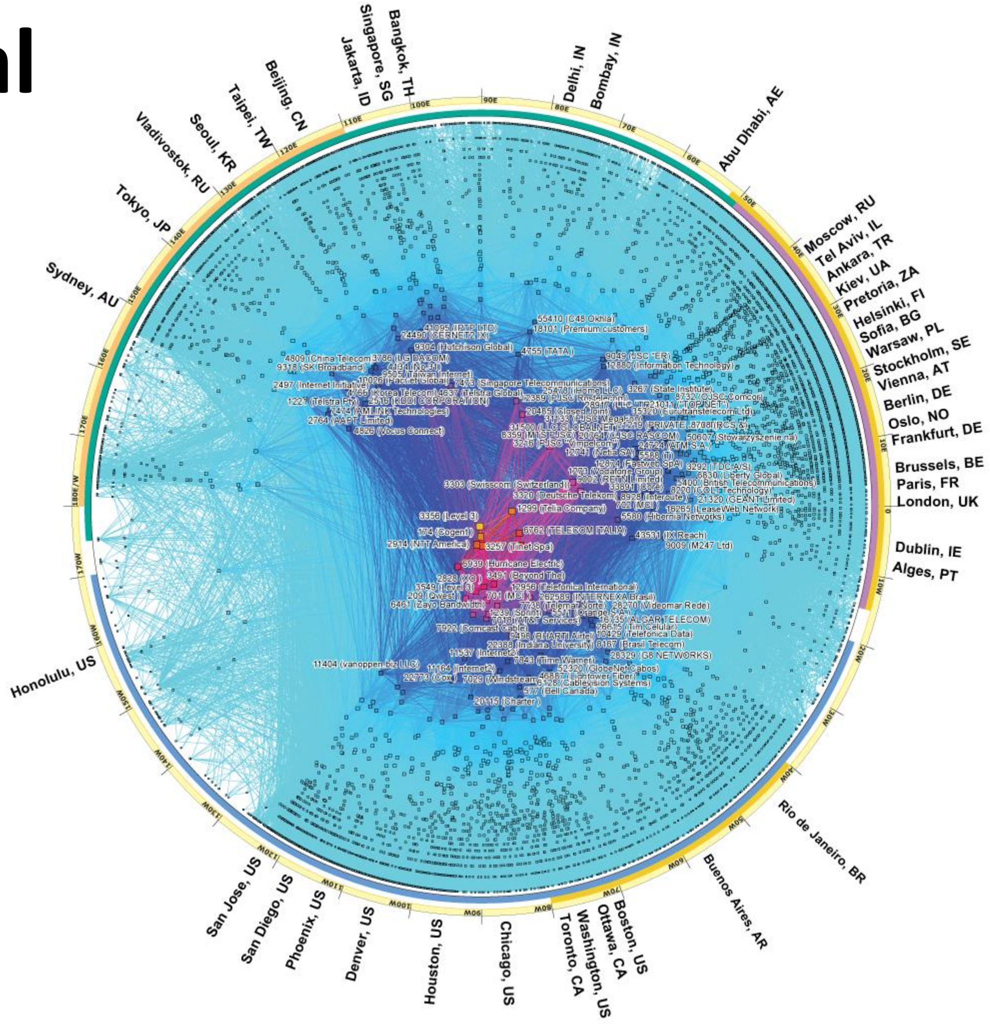
- ❖FDs exist, distribute heterogeneously and have a binary pattern

Outline

- Background, Research Goal and Questions
- Part I. Filtering the noise to reveal BGP lies
- Part II. Success and Failure of IXPs in Latin America
- Part III. The Art of Detecting Forwarding Detours
- Conclusions and Future Work

Research Goal

- Any system may have broken pieces
  - Problems, errors, limitations, etc...
- The Internet is a complex system
  - Protocols, facilities, networks
  - Hardware, software
  - Network operators, people
- The Internet is “big”...
  - Composed of 70K ASes
  - Point of observation matters



Research Goal: Detecting Hidden Broken Pieces of The Internet

Research Questions...and answers!

Q1: Can we detect BGP lies?

- Expected != Practice

Yes, filtering the noise with our framework

Q2: Are there failed IXPs? Why?

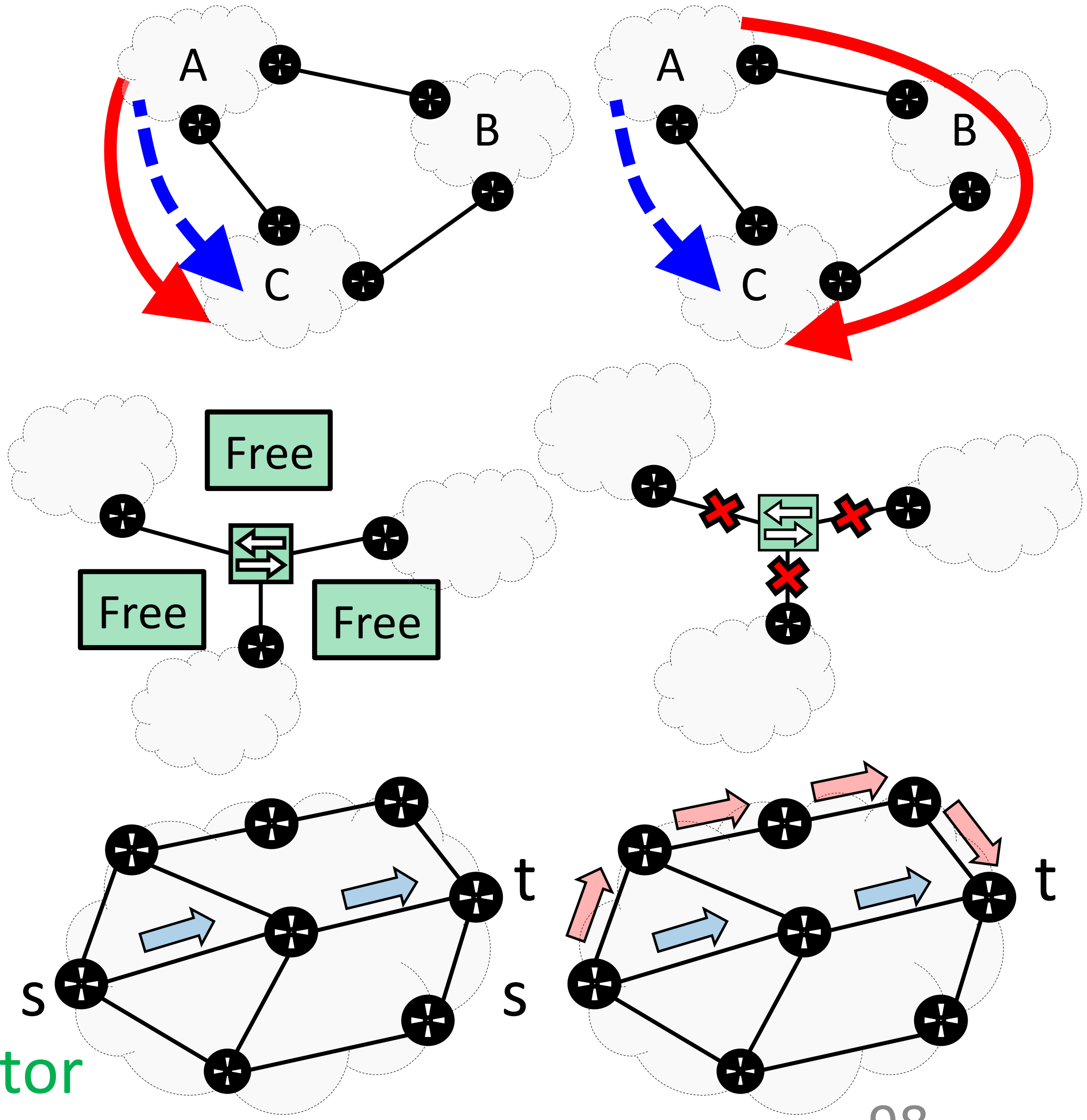
- IXPs with low impact

In Latin America, yes. Possibly due to the presence of monopolistic local Ases

Q3: Can we model and detect detours?

- Expected != Practice

Yes^2: RIes produce them; use our FD-detector



Publications

Contribution 1
Filtering the Noise to Reveal Inter-Domain Lies
In Network Traffic Measurement and Analysis Conference (TMA) 2019
Julián M. Del Fiore, Pascal Merindol, Valerio Persico, Cristel Pelsser and Antonio Pescapè.

Contribution 2
A first look at the Latin American IXPs
In SIGCOMM Computer Communications Review (CCR), January 2020
Esteban Carisimo, **Julián M. Del Fiore**, D. Dujovne, Cristel Pelsser, and J. I. Alvarez-Hamelin

Contribution 3
The Art of Detecting Forwarding Detours
Minor revision in IEEE Transactions on Network and Service Management (IEEE TNSM) 2021
Julián M. Del Fiore, Valerio Persico, Pascal Merindol, Cristel Pelsser and Antonio Pescapè.

Future Work

Short term: enlarging the measurements

- We used 8 co-located VPs to detect BGP lies
- Our study of IXPs relied on BGP data

- New contributions:
  1. Use co-located VPs placed in IXPs
  2. Run active measurements for the IXPs work

Medium term: digging more into FDs

- Currently, we focus on the detection of FDs
- New contributions:
  1. Detect the router introducing the FA leading to a FD
  2. Measuring impact of FDs on performance
  3. Building an FD-detector-lite leveraging (2)

Long term: topology discovery and LB studies

- The multipath discovery algorithm (MDA):
    - Discovers multi-path routing patterns
    - Probing cost updated following a mathematical model
    - Measurements on a per-prefix basis
    - Campaigns usually comprise multiple destinations

- New contributions:
    1. Two step measurement process (Topology Feedback, TF-MDA)
    2. Add network knowledge to probing model (Bayesian-MDA)
    3. Combine the ideas of (1) and (2) (Ultimate, U-MDA)

Thank you for your attention

Questions ?

Complementary Slides

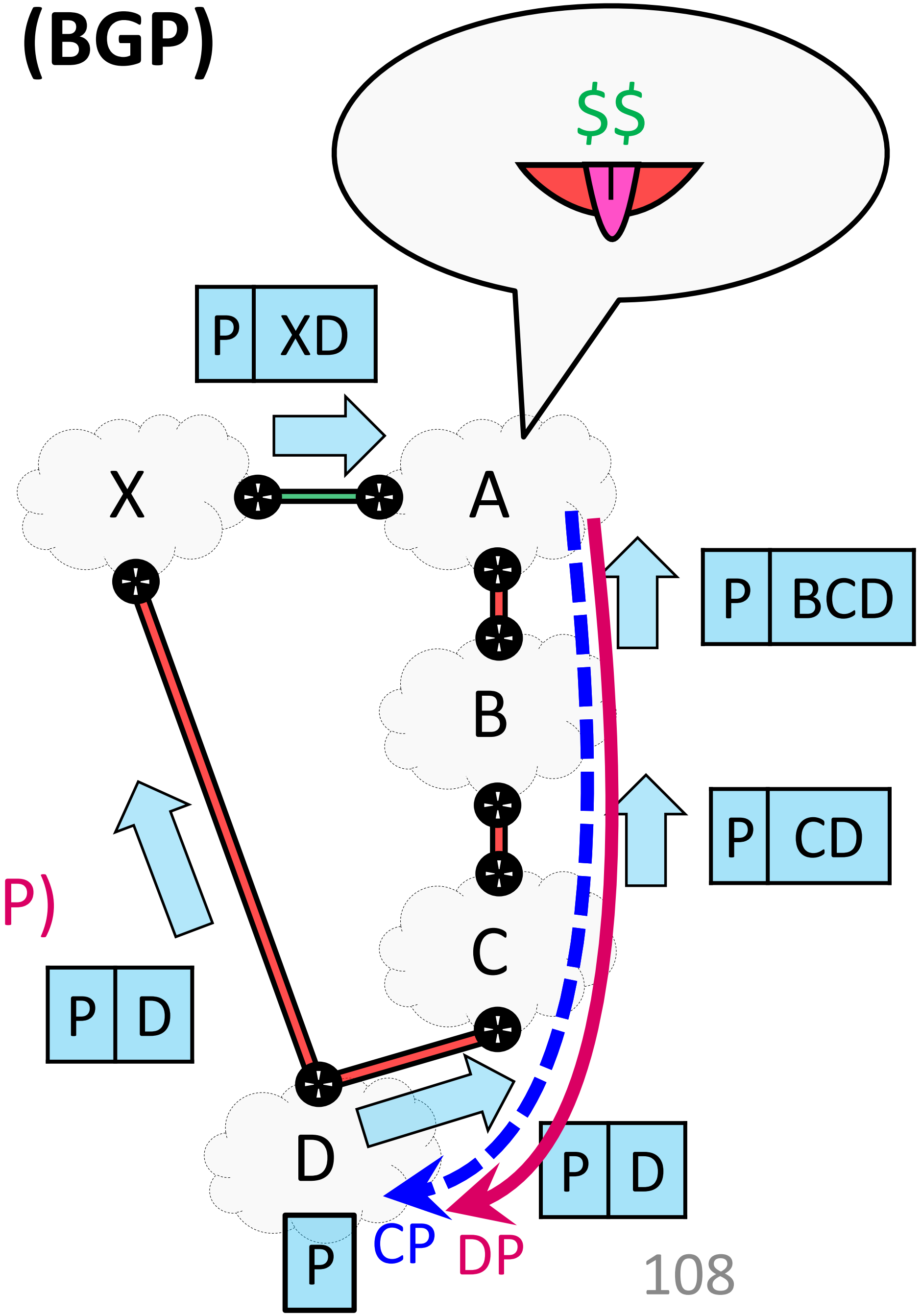
You told me the Internet was perfect!

Yeah, in my dreams

BGP: Extended Background

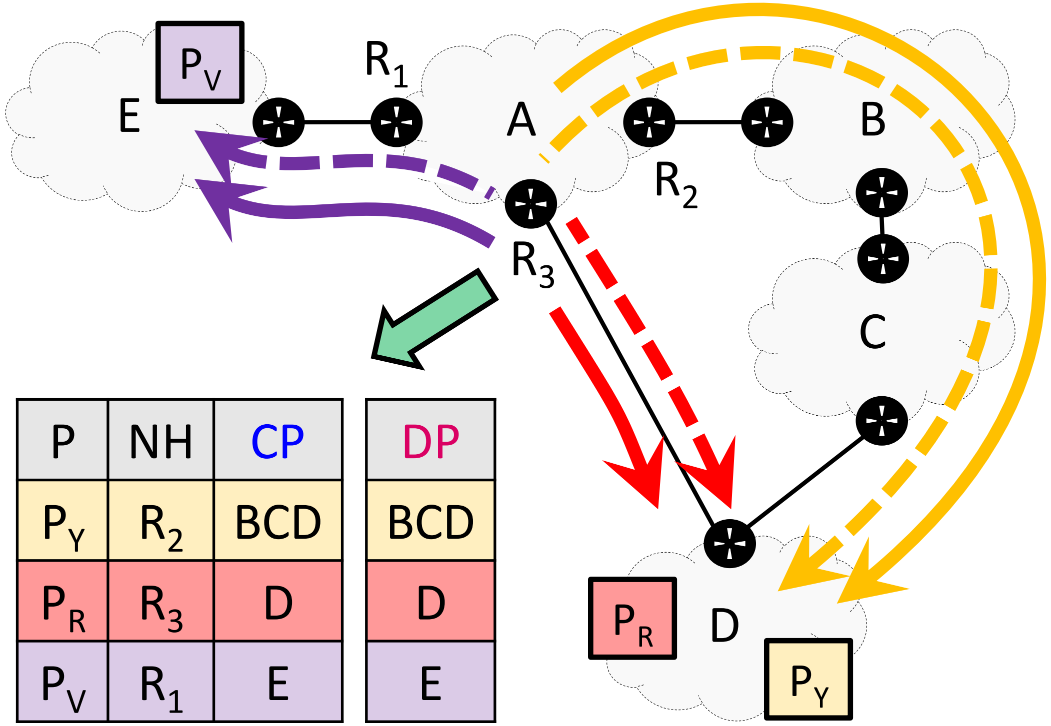
Border Gateway Protocol (BGP)

- Announce the IP prefixes they own
- Relay announcements updating the messages
- Decision process to choose the best path
- Resulting AS-path as the control path (CP)
- Packets flow towards P through a data path (DP)



Border Gateway Protocol (BGP)

- BGP is run by routers called BGP speakers
- For each external IP prefix (P):
  - the next-hop (NH) to be reached
  - the control path (CP) that should **theoretically** be followed
- The data path (DP) is the path used in **practice**



P	NH	CP	DP
P_Y	R_2	BCD	BCD
P_R	R_3	D	D
P_V	R_1	E	E

Detecting BGP lies
Technical considerations

Address space synchronization

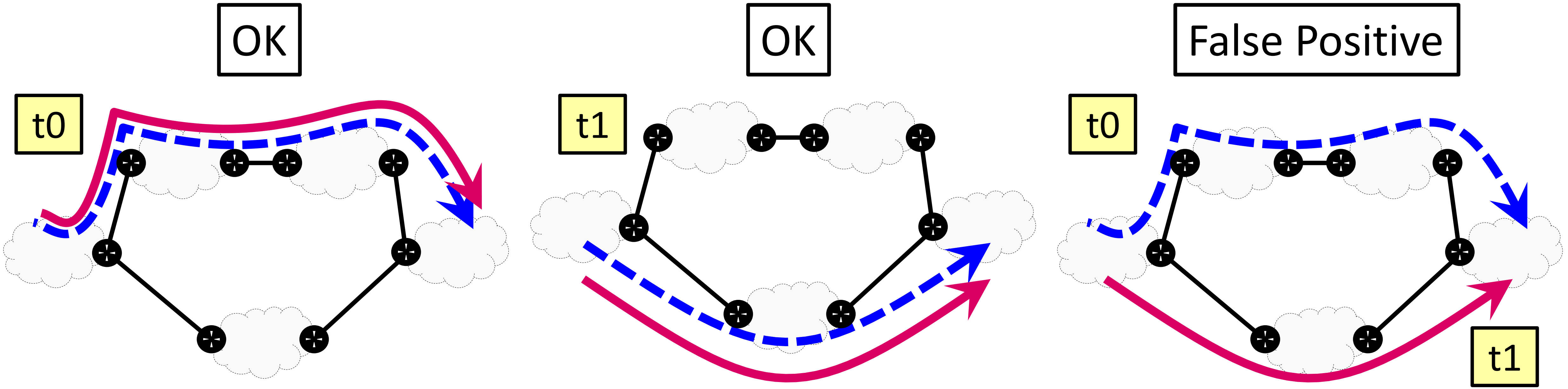
- After the measurements, we have a "bag" of CPs and DPs
- Question...which DP should be compared with which CP?
- Each DP is associated with a given destination d
- Compare DP with the CP of the longest matching prefix

P	NH	CP
P_Y	R_2	BCD
P_R	R_3	D
P_V	R_1	E

DP
BCD
D
E

Time-synchronization

- The CP is not static, at t0 and t1 it may be different
- Imagine no BGP lies occur...then the DP also changes over time!
- To avoid false positives, then CPs and DPs need to be collected “close” in time



A basic IP-to-AS mapping method

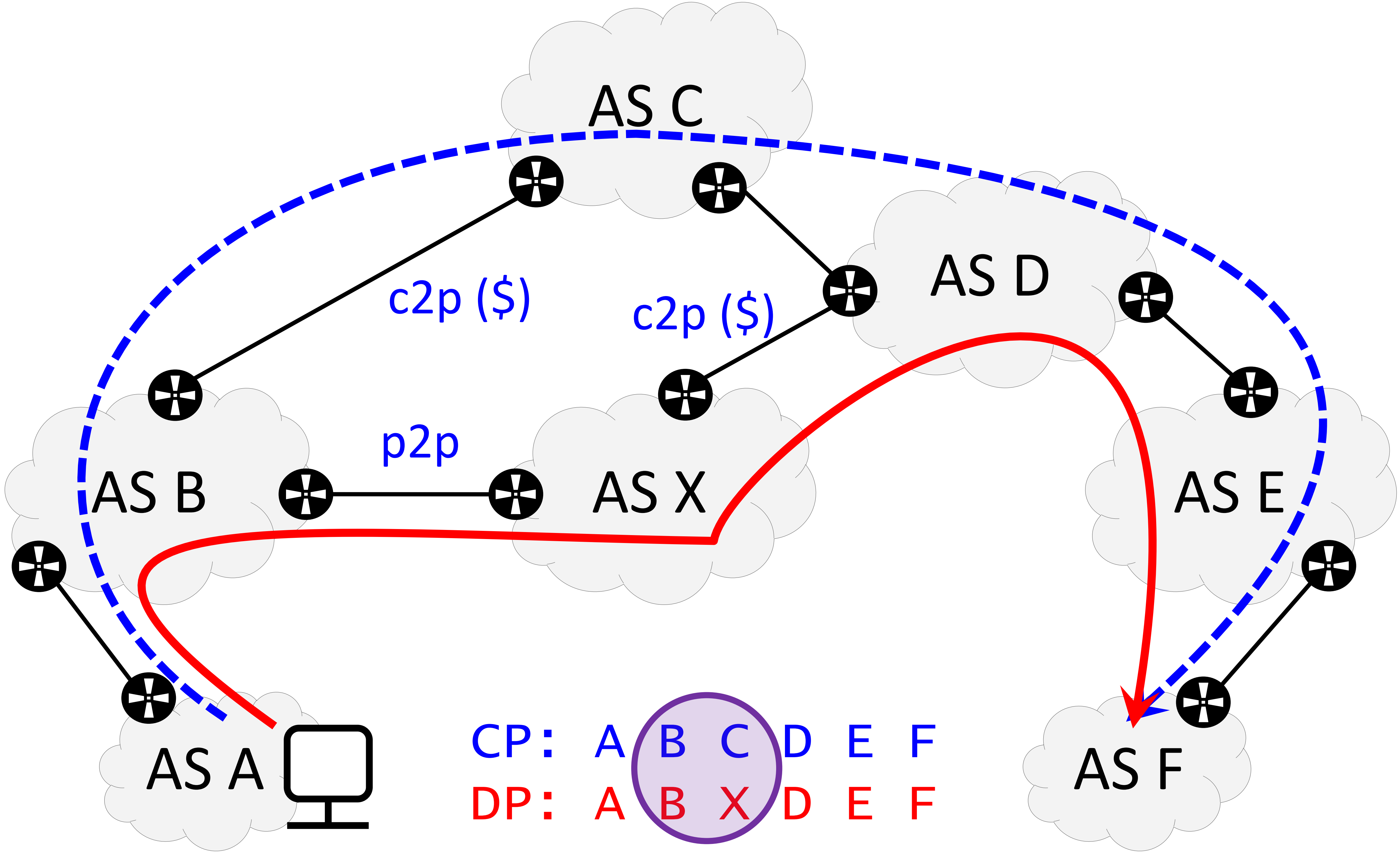
- For each IP address...
  - Look for longest matching prefix
  - Map to the first AS in the AS-path associated to that entry
  - Collapse the repetitions

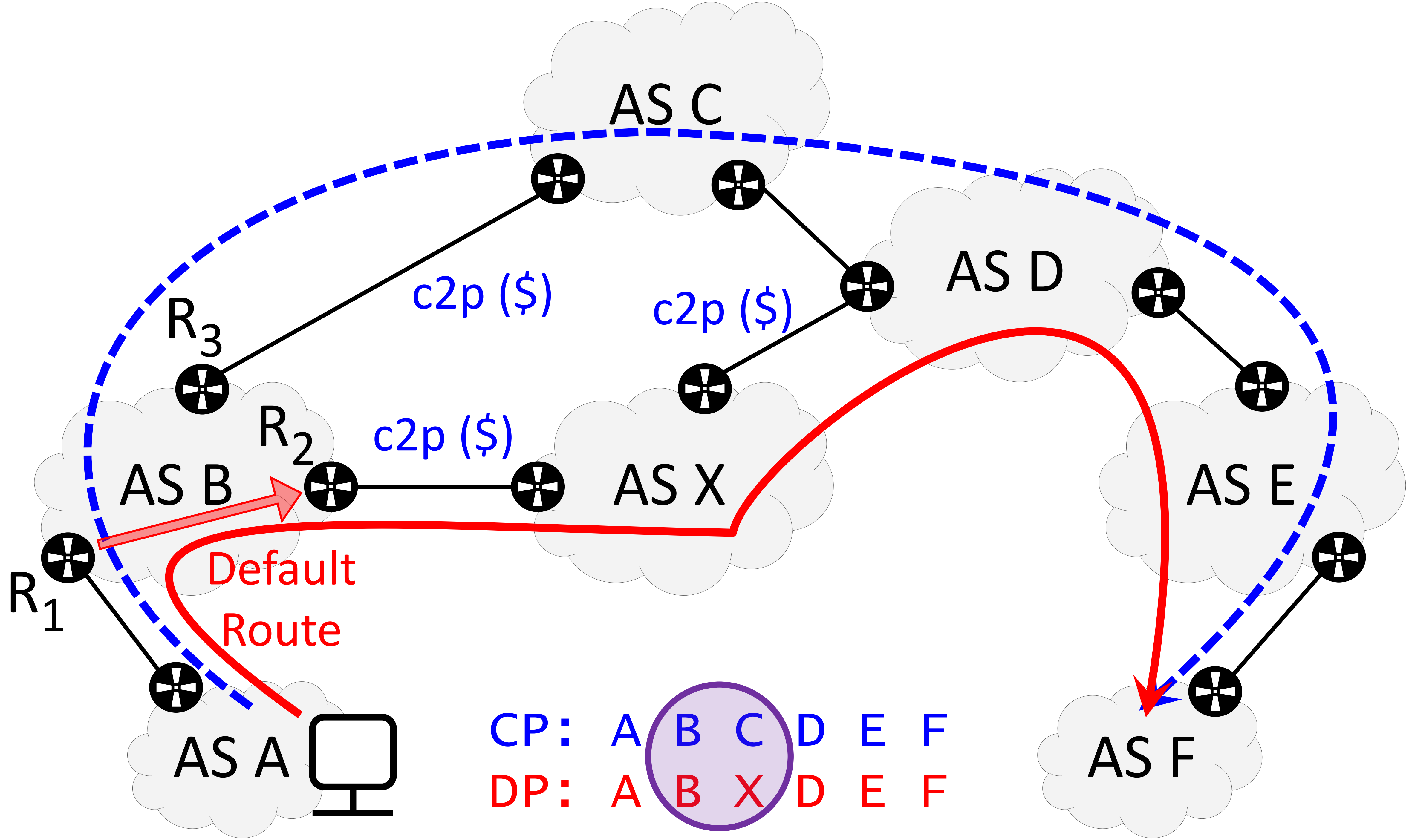
Prefix	AS-path
P1	X Y A
P2	X Z A
P3	X W B

CP: A B C vs DP: $\underbrace{\underbrace{\text{I1}}_{\text{P1}}}_{\text{A}}\ \underbrace{\underbrace{\text{I2}}_{\text{P1}}}_{\text{A}}\ \underbrace{\underbrace{\text{I3}}_{\text{P2}}}_{\text{B}}$ I4 I5 I6 I7 I8

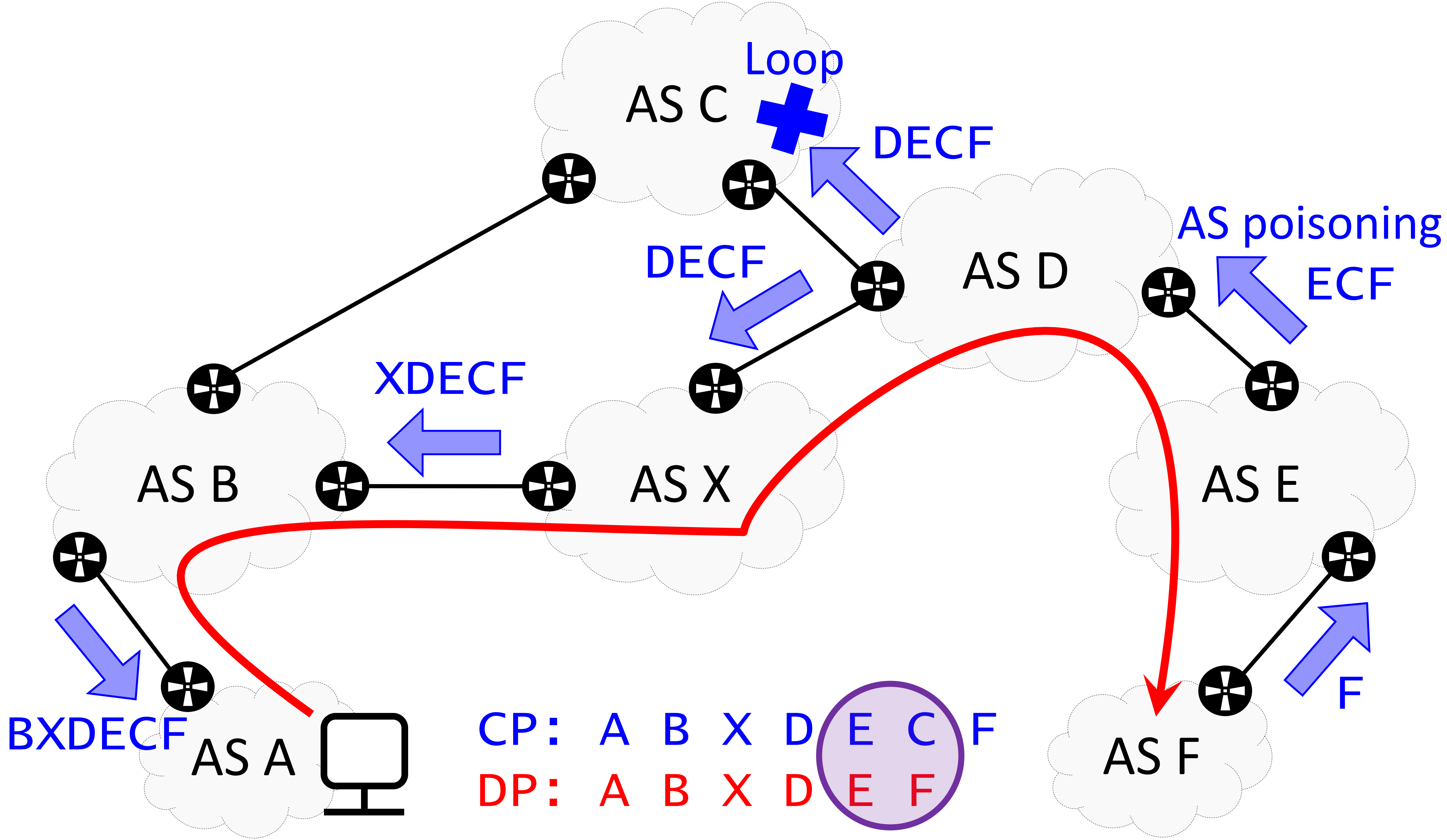
DP: $\underbrace{\text{I1 I2}}_{\text{A}}\ \underbrace{\text{I3 I4 I5}}_{\text{B}}\ \underbrace{\text{I6 I7 I8}}_{\text{C}}$

BGP lies: examples

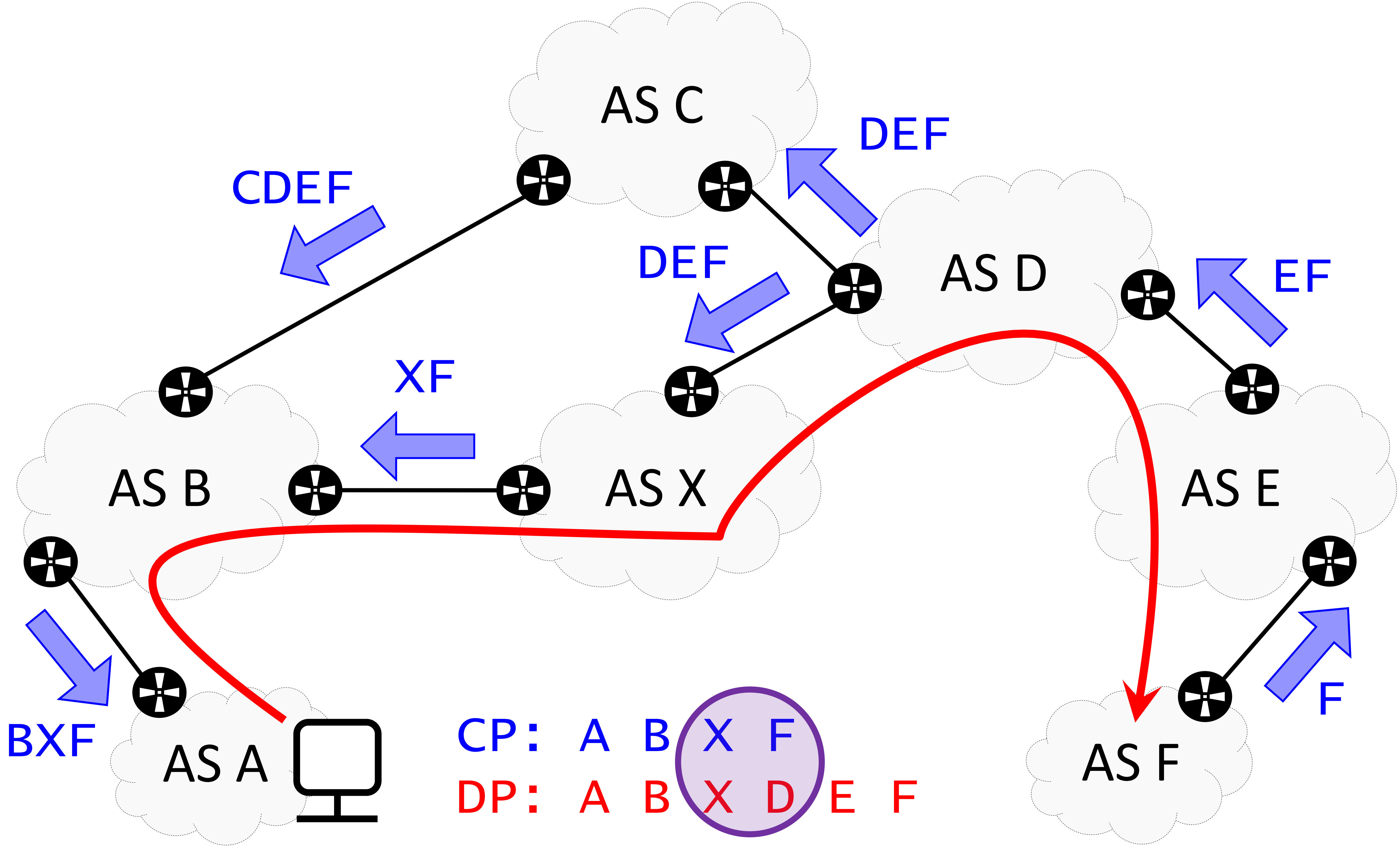

AS C
AS D
AS B
AS X
AS E
AS A
AS F
c2p ($)
c2p ($)
p2p
CP: A B C D E F
DP: A B X D E F



AS C
AS D
AS B
AS X
AS E
AS A
AS F
R3
R2
R1
c2p ($)
c2p ($)
c2p ($)
Default Route
CP: A B C D E F
DP: A B X D E F



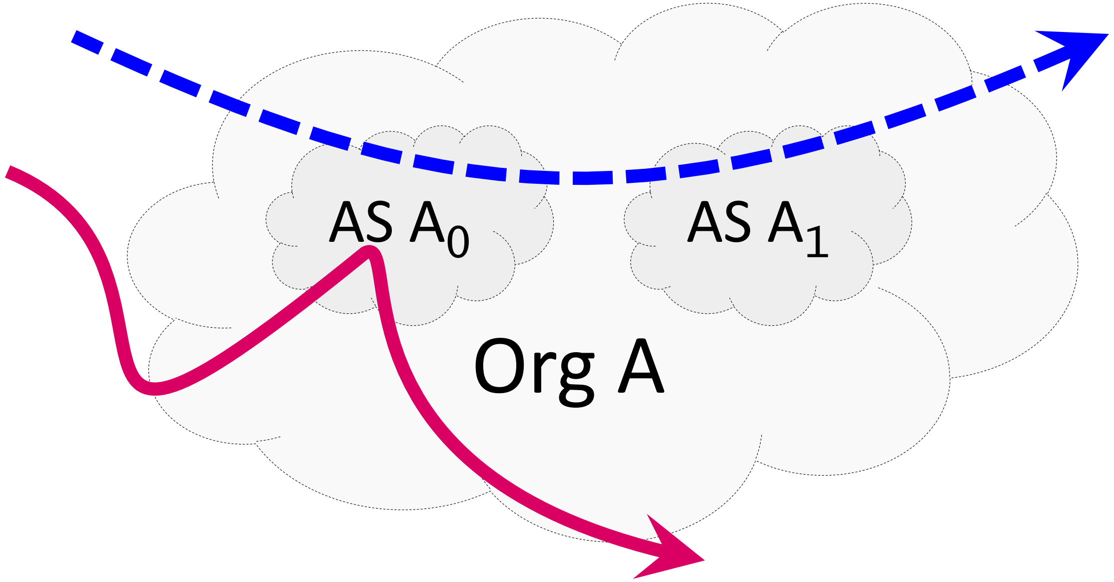
Loop
AS C
DECF
DECF
AS D
AS poisoning
ECF
XDECF
AS B
AS X
AS E
BXDECF
AS A
F
AS F
CP: A B X D E C F
DP: A B X D E F


AS C
DEF
CDEF
DEF
AS D
EF
XF
AS B
AS X
AS E
BXF
AS A
CP: A B X F
DP: A B X D E F
F
AS F

Framework: Our filters

Mapping relaxation - SIB Rule

- SIB rule: Apply an AS-to-organization mapping
- We construct the mapping with CAIDA's AS Organizations Dataset



CP: Z A_0 A_1 B
DP: Z A_0 B

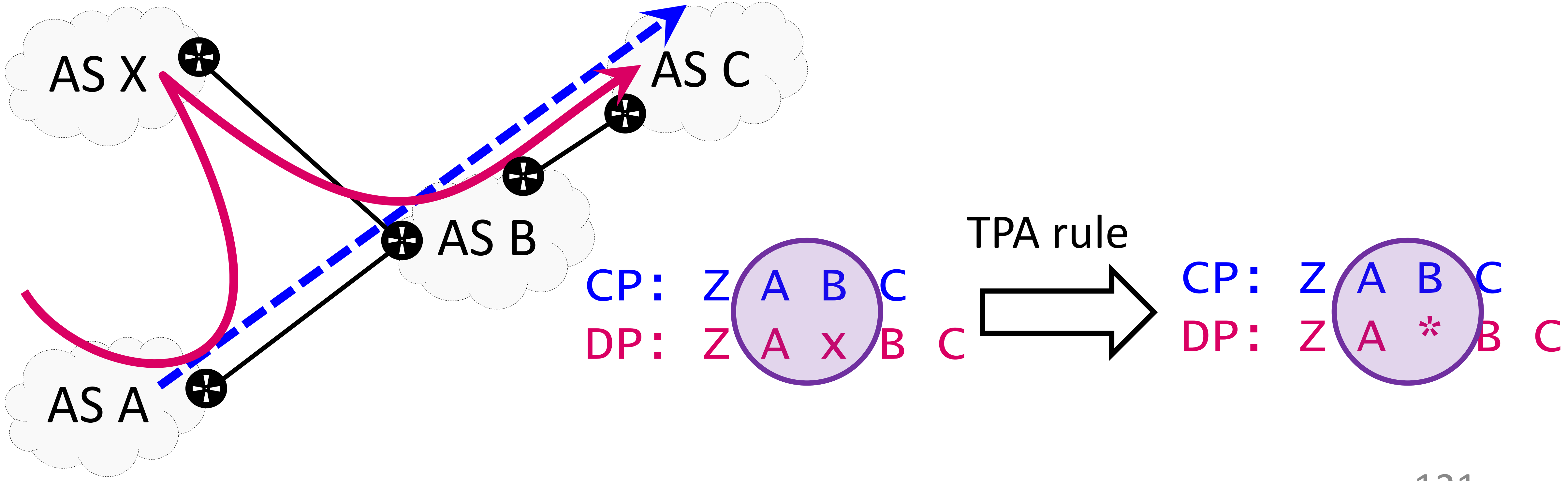
SIB rule $\Rightarrow$

CP: Z A B
DP: Z A B

OK

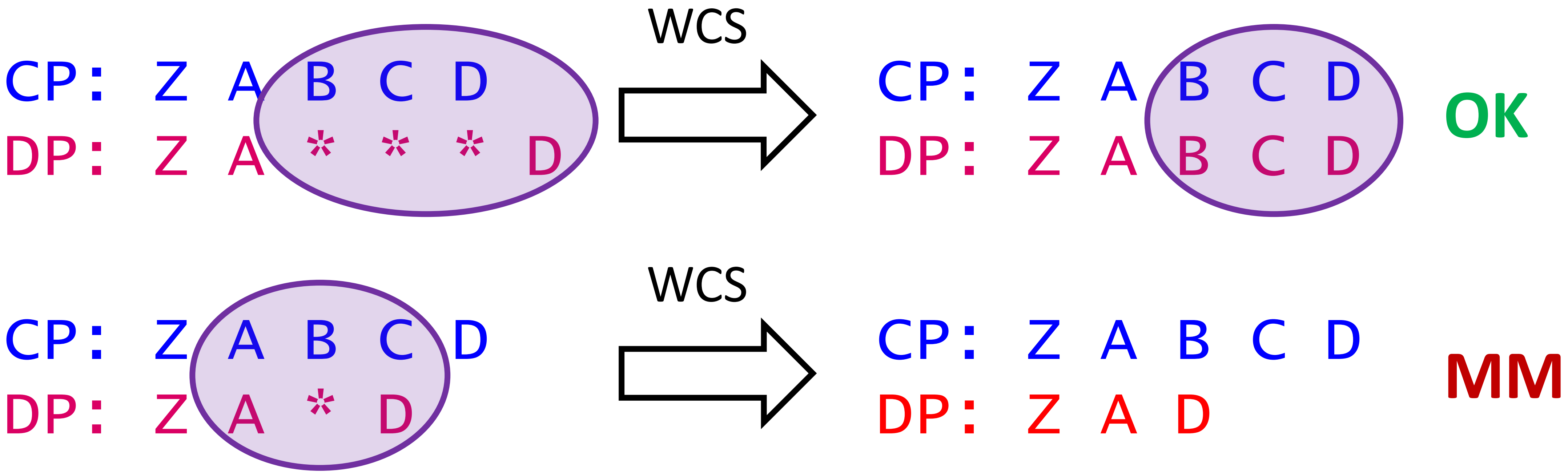
Mapping relaxation – TPAs Rule

- TPA rules: replace TPAs with wildcards.
- When only one IP maps to an AS, we label it as candidate TPAs (cTAPs)
    - looseTPA: all cTPAs are inferred to be TPAs
    - strictTPA: exclude cTPAs surrounded by cTPAs or missing hops

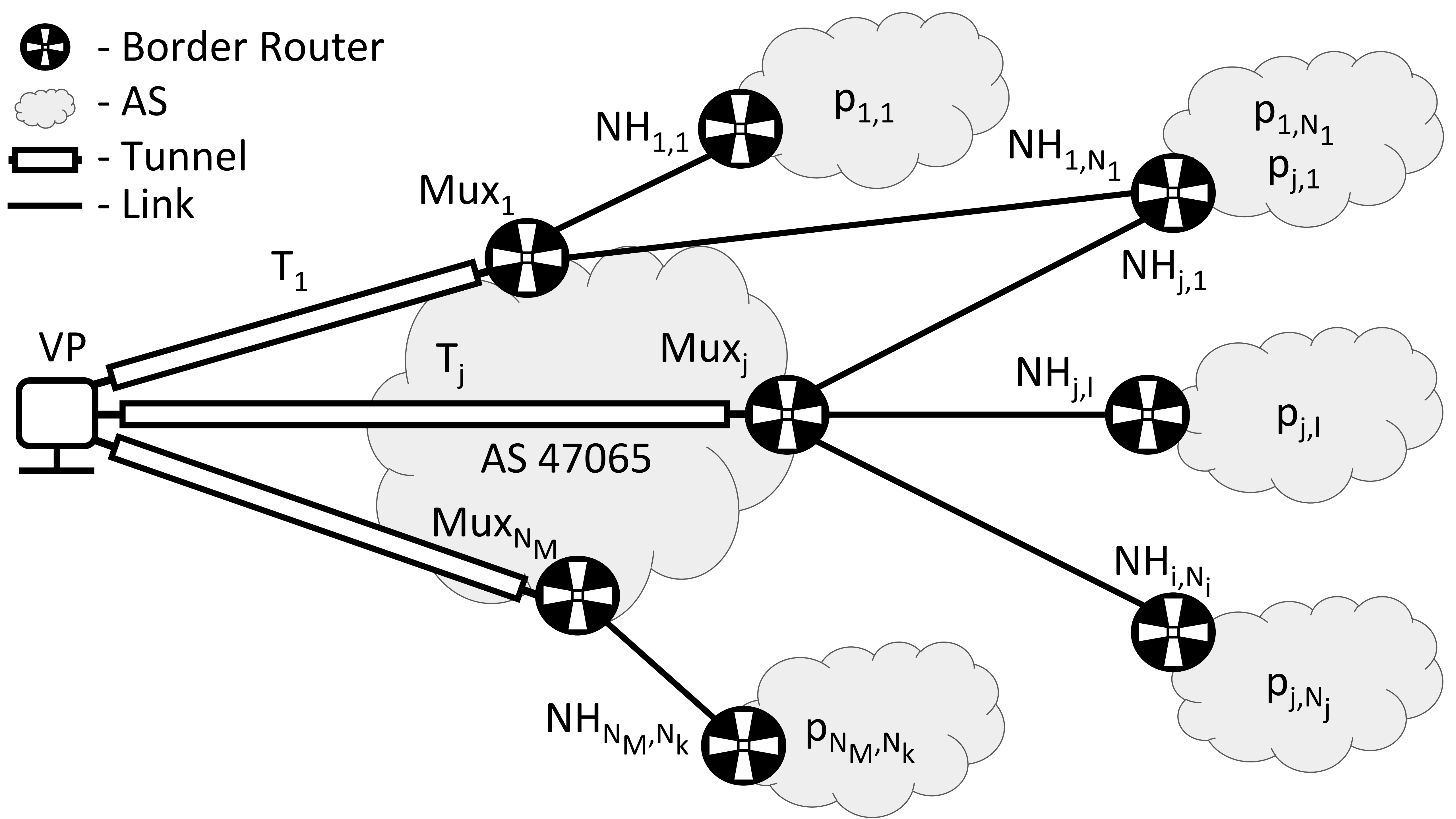


Wildcards Correction Stage (WCS)

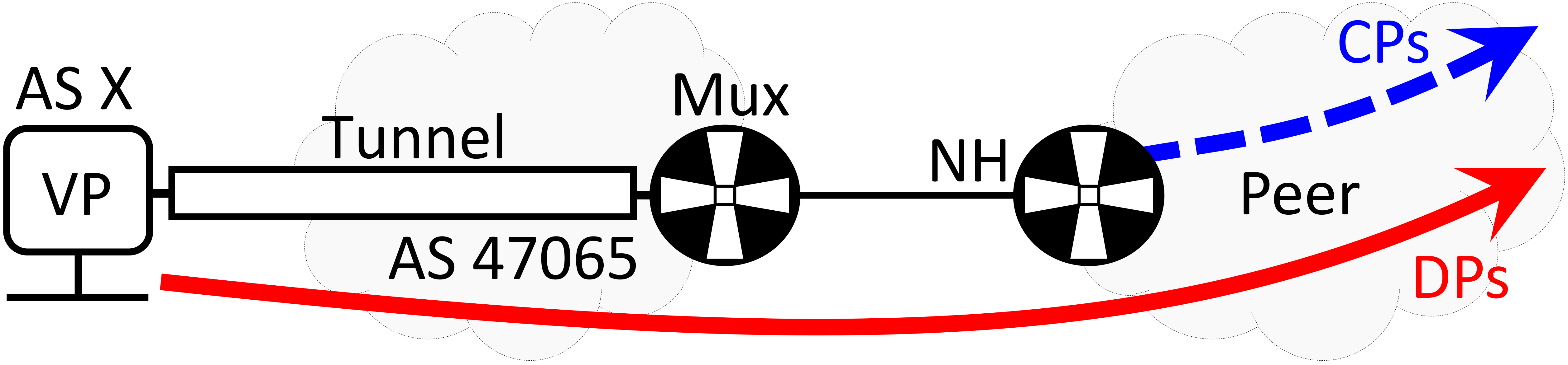
- Try to infer a value for the wildcards and see if paths mismatch (MM)
- Note that wildcards are either missing hops or inferred TPAs.



Measuring Platform



- Border Router
- AS
- Tunnel
- Link
VP
T_1
T_j
Mux_1
Mux_j
Mux_{N_M}
AS 47065
$NH_{1,1}$
$p_{1,1}$
NH_{1,N_1}
p_{1,N_1}
$p_{j,1}$
$NH_{j,1}$
$NH_{j,l}$
$p_{j,l}$
NH_{i,N_i}
p_{j,N_j}
NH_{N_M,N_k}
p_{N_M,N_k}



AS X
VP
Tunnel
AS 47065
Mux
NH
Peer
CPs
DPs

Peer	**Organization**	**ASN**	**CP-DP match [%]**
isi	Los Nettos	226	77.92
uw	University of Washington	101	77.93
neu	Northeastern University	156	76.84
uth	University of Utah	210	69.51
grt	GRNet	5408	77.93
cle	Clemson University	12148	77.93
hm1	University of Strasbourg	2259	77.94
hm2	RGnet, LLC	3130	77.90

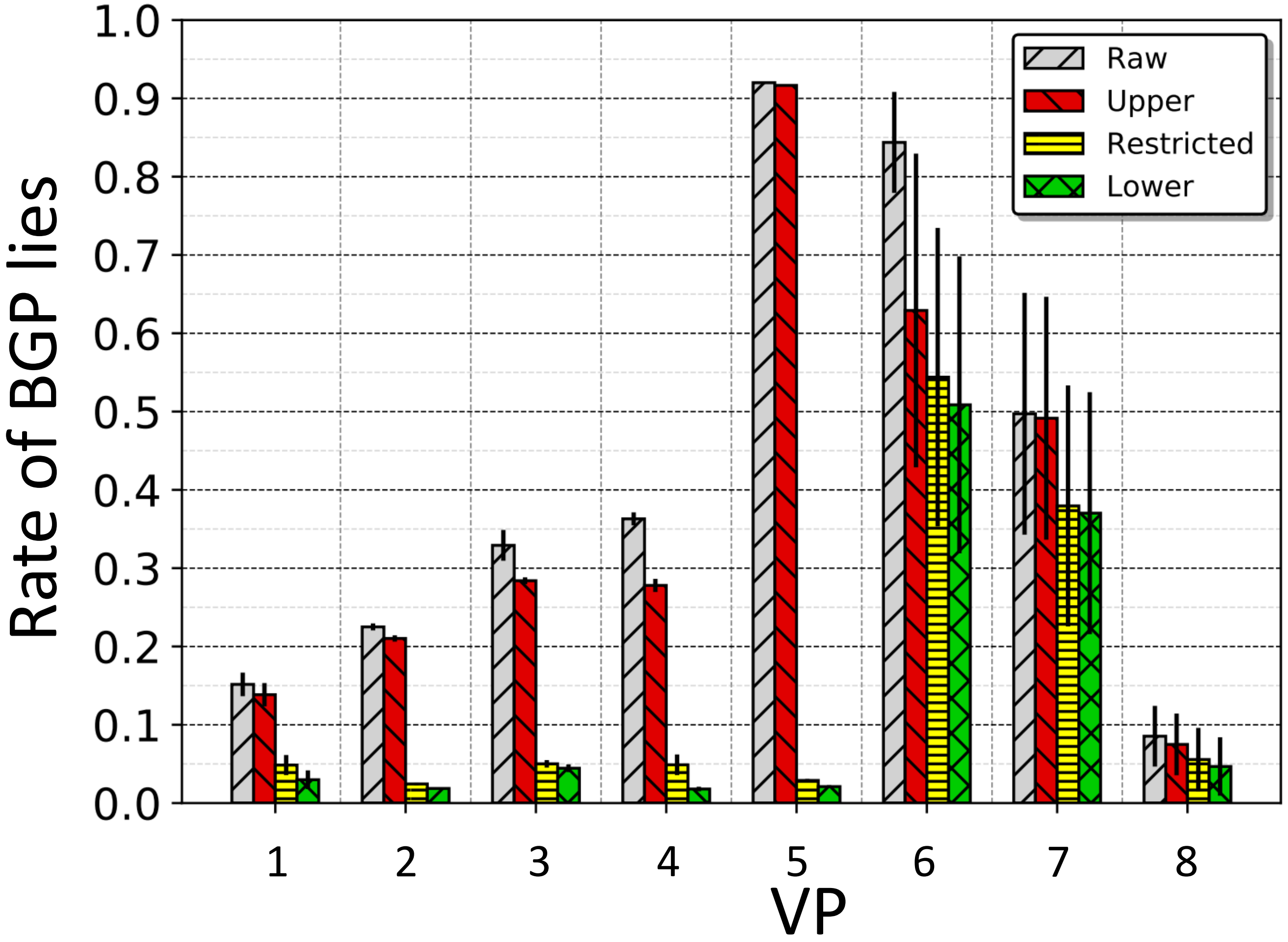
Modular Framework
Different models, different results

Modularity

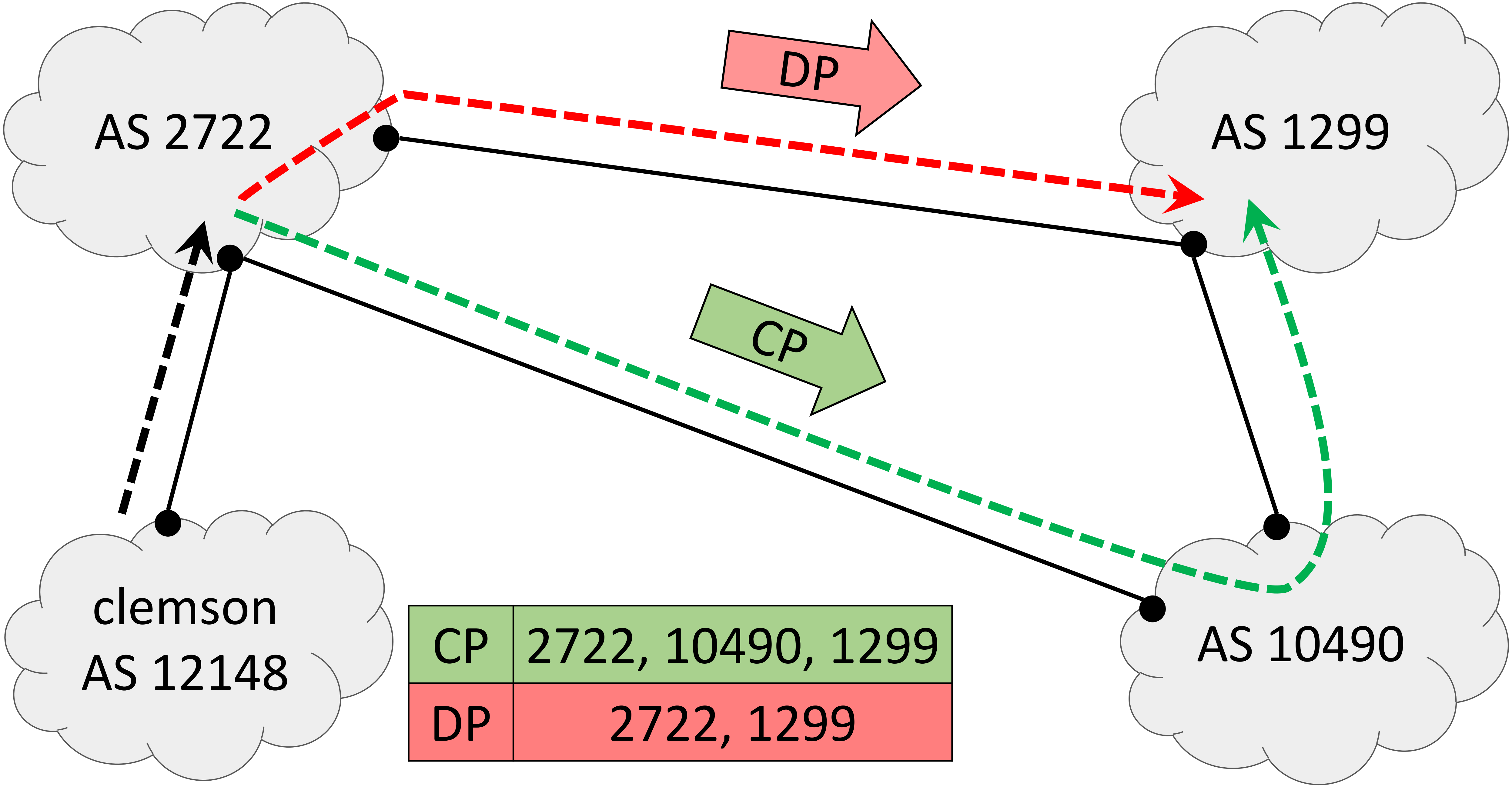
- Our framework allows to implement different noise-filtering models

	Mapping Relaxation			Wildcards Correction	
Model/Rules	SIB	looseTPA	strictTPA	match*	nomatch*
Raw	✗	✗	✗	✗	i
Upper	✗	✗	✗	i	ii
Restricted	i	✗	ii	iii	iv
Lower	ii	i	✗	iii	iv

Mismatch (MM) rate in the wild

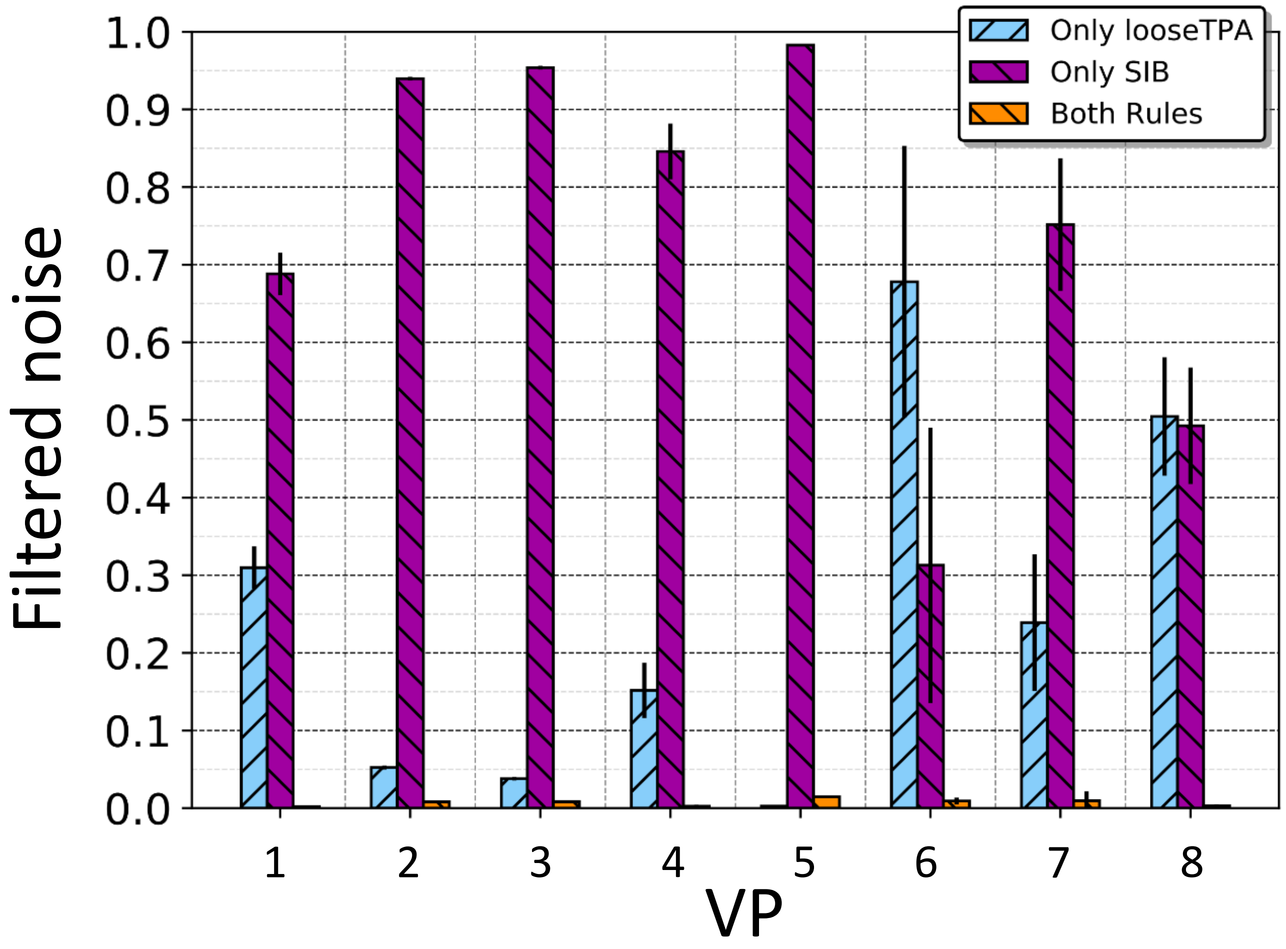


- The models implementing the mapping relaxation outperform the others
- The looseTPA does not outperform the strictTPA for much


AS 2722
DP
AS 1299
CP
clemson
AS 12148
AS 10490
CP 2722, 10490, 1299
DP 2722, 1299

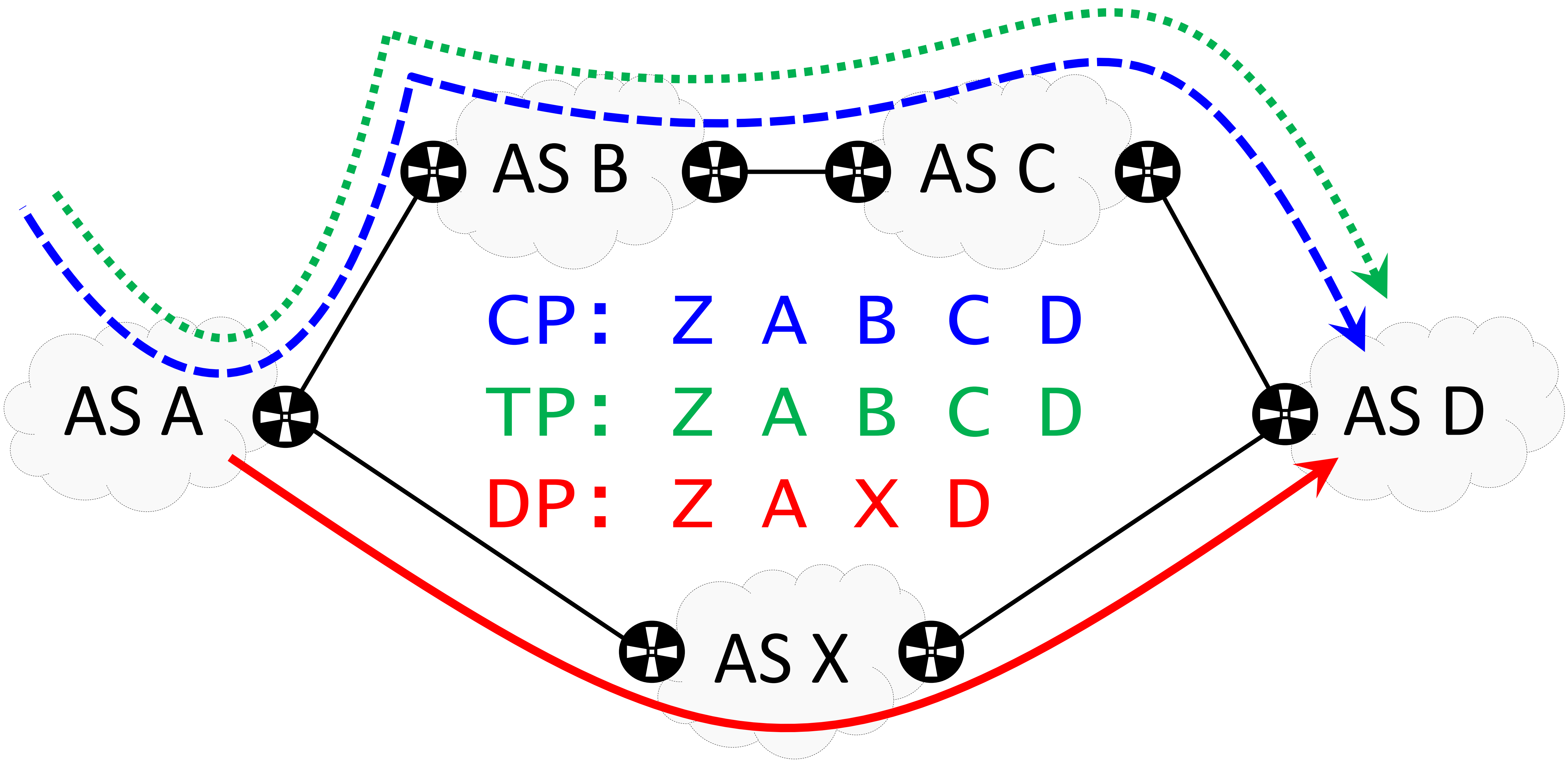
Characterizing the mapping noise

Looking at the filtered noise



- In general, AS siblings and third-party addresses not combine
- The worse source of noises varies depending on the VP

Future Work BGP lies


AS A
AS B
AS C
AS D
AS X
CP: Z A B C D
TP: Z A B C D
DP: Z A X D

Future Work BGP lies

All about Latin America
And IXPs

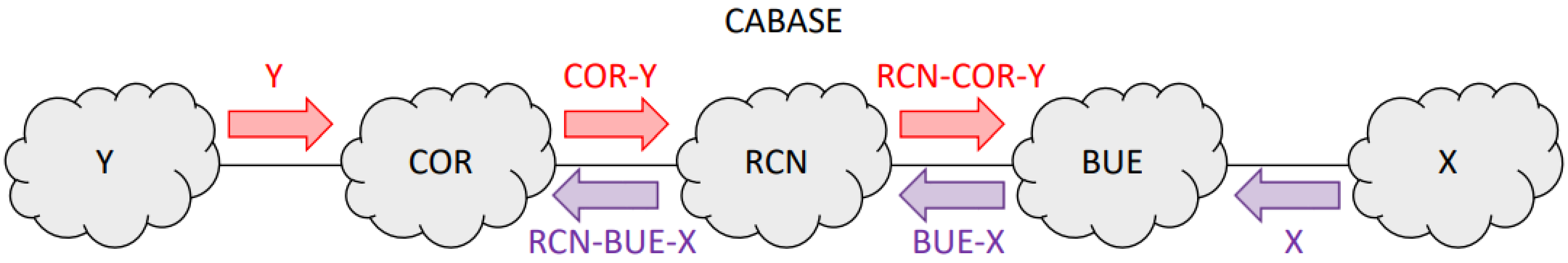
Public Policies

	Country	AR	BO	BR	BZ	CL	CO	CR	CU	EC	HT	HN	MX	PA	PY	PE	TT
Sponsored by		CABASE	Law	CGI	PUC	PIT CL	CCIT	Ex.Ord.	State	IXP.EC	AHTIC	CONATEL	IFT	SENACYT	SENATICS	NAP.PE	TTIX
Operated by		CABASE	State	NIC.br	UoBZ	PIT CL	CCIT	NIC.cr	NAP.CU	IXP.EC	AHTIC	UNAH	CITI	InteRED	NIC.py	NAP.PE	TTIX
BGP TDs	Monitor	PCH		RVs/LGs	PCH	PCH		PCH		PCH	PCH	PCH	PCH		PCH		PCH
	#Memb	127	×	1156	6	72	×	28	×	5	4	4	6	×	15	×	5
	#AggIPs	7.9M		26M	67K	19.4M		401K		28K	102K	131K	795K		1.5M		196K

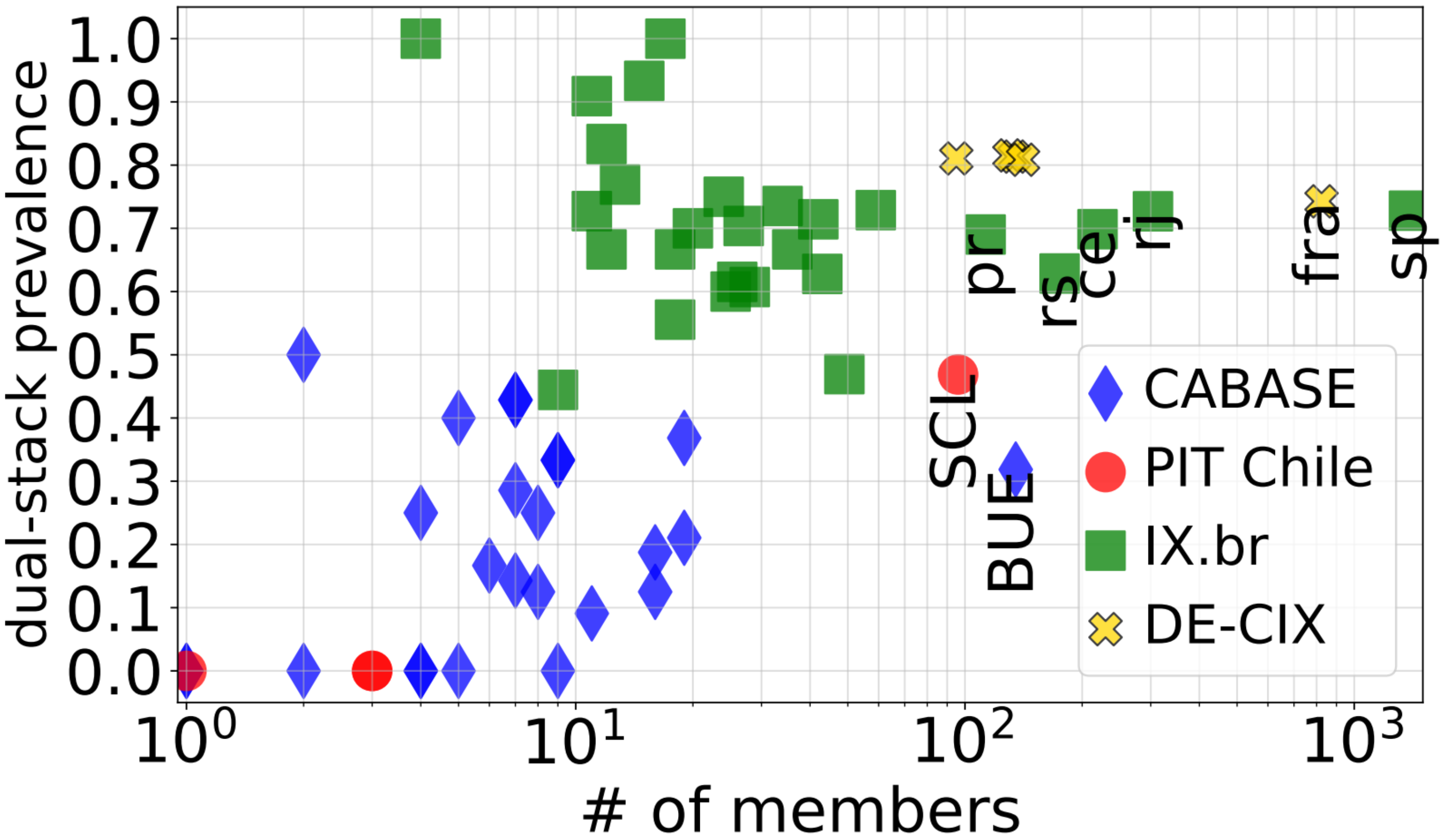
- B, Y and V represent state agencies, non-profit organizations and universities, respectively
- Governments involved in the creation of their national IXPs in more than 55% of the cases
- Similar to the European IXP model, in LatAm many non-profit orgs created and run IXPs

IXP networks topology

	CABASE	PIT-CL	IX.br	DE-CIX
CC	AR	CL	BR	DE
#IXPs in CC	28	5	31	5
ASN per IXP	✓	✓	✗	✓
IXP facilities	1/IXP	1/IXP	PIXes	Sites
IXPs Linked	✓	✓	✗	✓
Enforced Policy	MMPP	✗	✗	✗

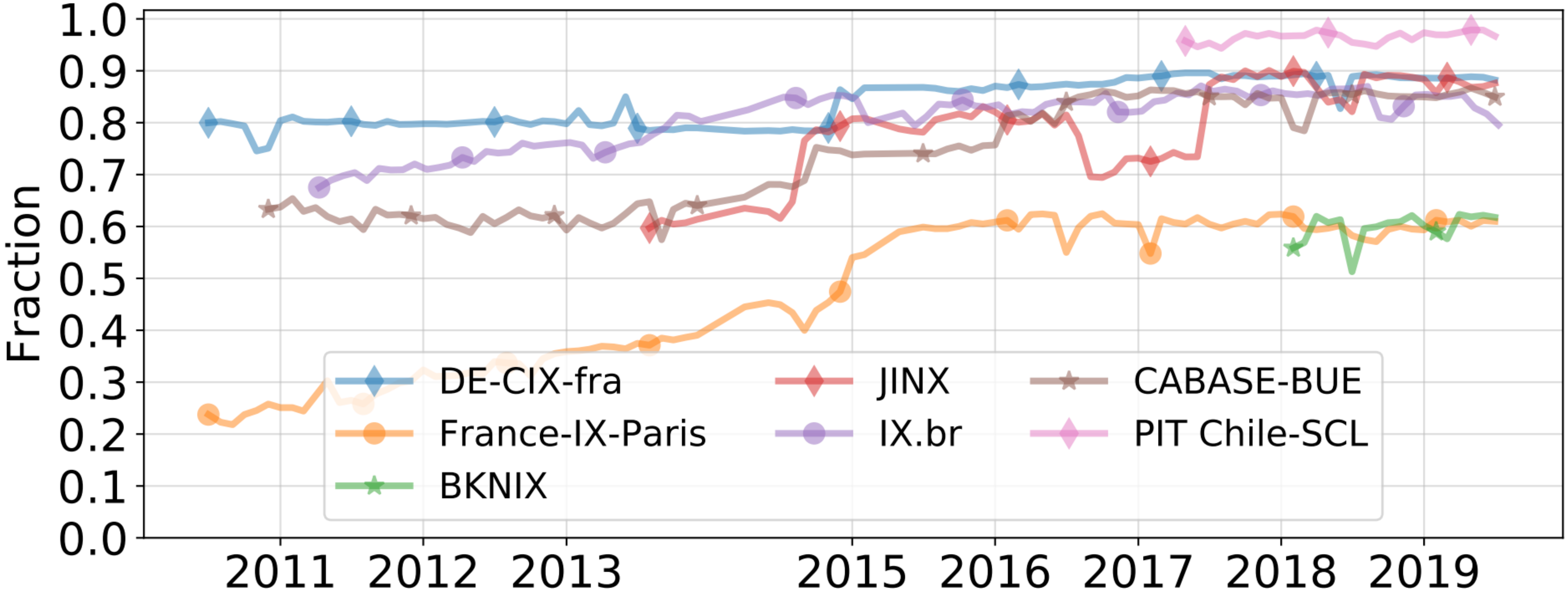


DS-prevalence vs #members



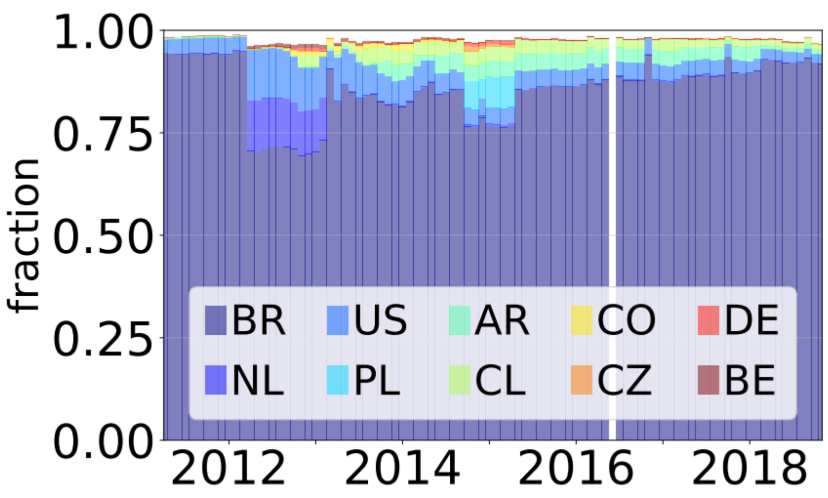
- IX.br-SP is the largest and the remaining in the TOP5 too
- IX.br is much larger than CABASE and PIT Chile
- Largest regional IXPs in cities that are economically central
- DS-prevalence if BR similar to DE, but AR and CL lower

Visible ASes: domestic impact and foreign attraction

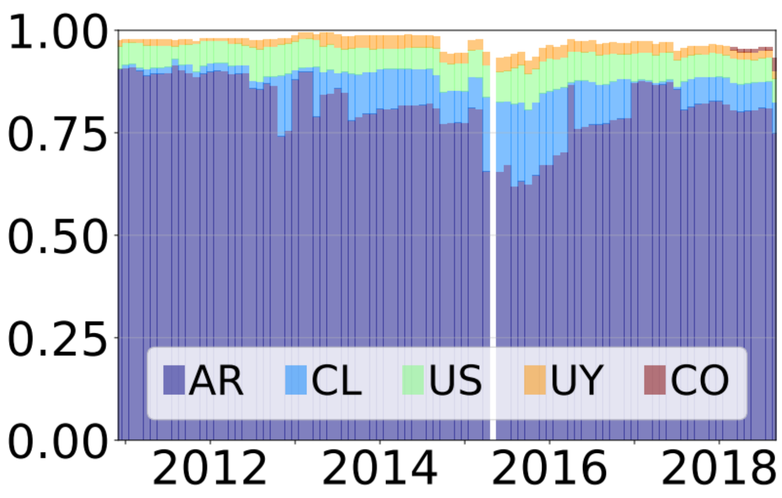


- Ratio of local visible ASes to all active ASes (with AS rels) in each country
- Lately, values in LatAm similar to those in Europe. Similar for Africa.
- PIT Chile is surprising given it’s a “young” IXP, as BKNIX is also

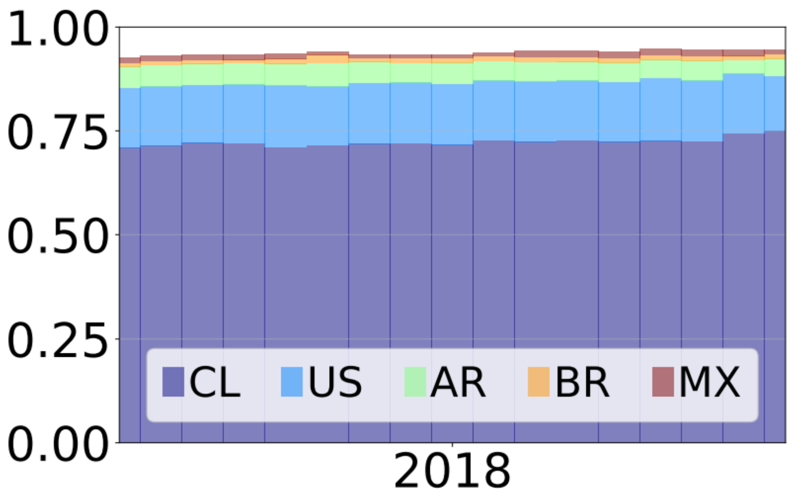
Visible ASes: foreign attraction



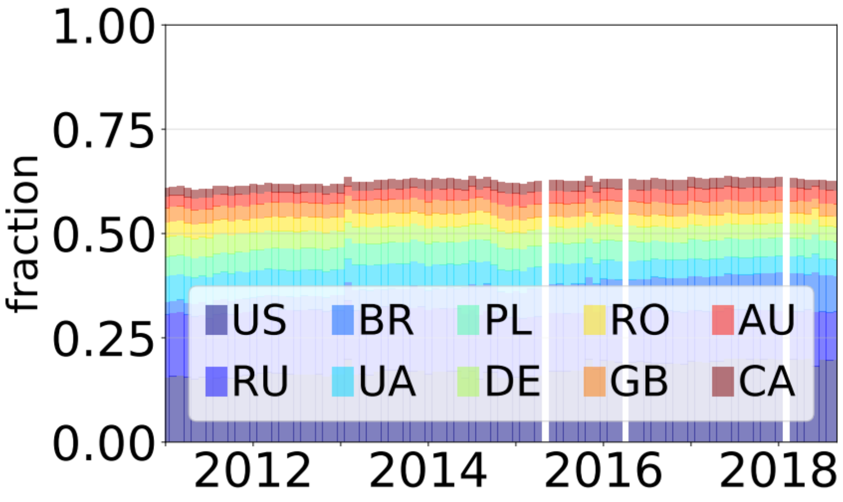
(a) IX.br-SP



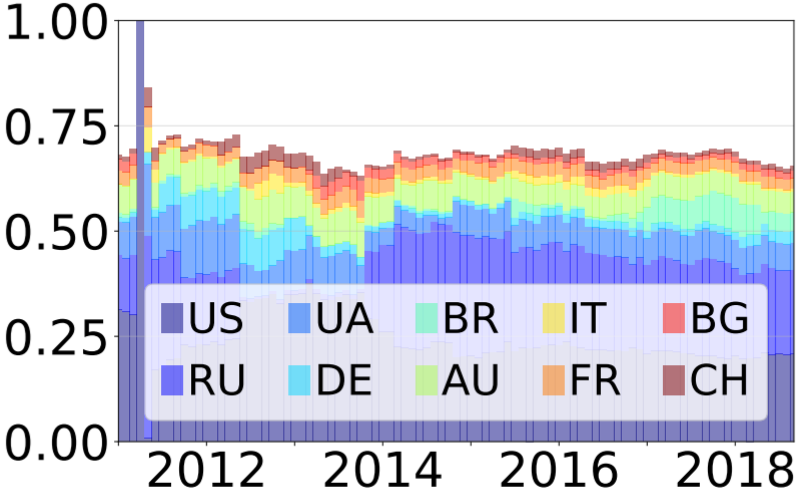
(b) CABASE-BUE



(c) PIT Chile-SCL



(f) DE-CIX-fra



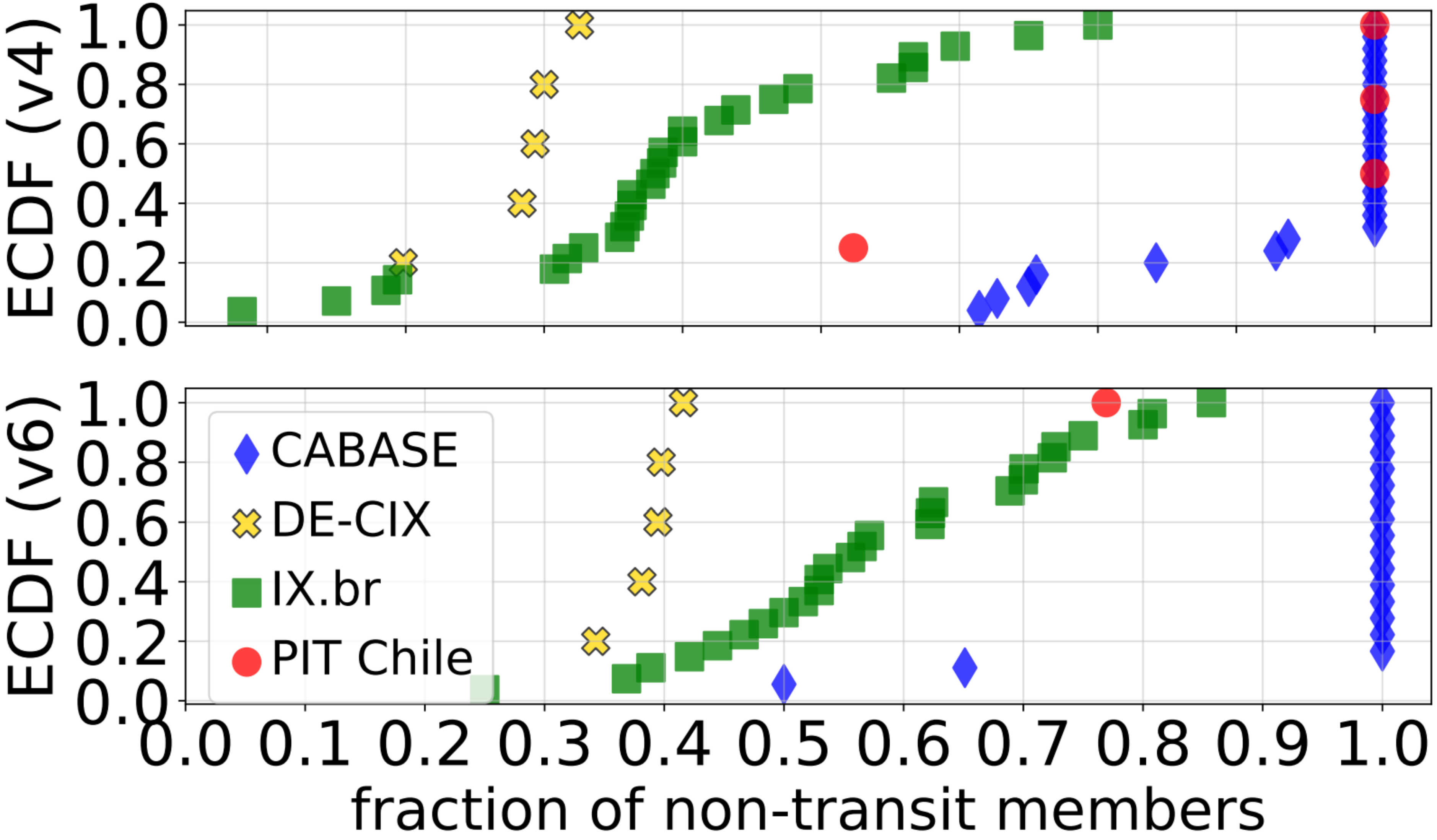
(g) France-IX-Paris

Reaching IXPs: transit members

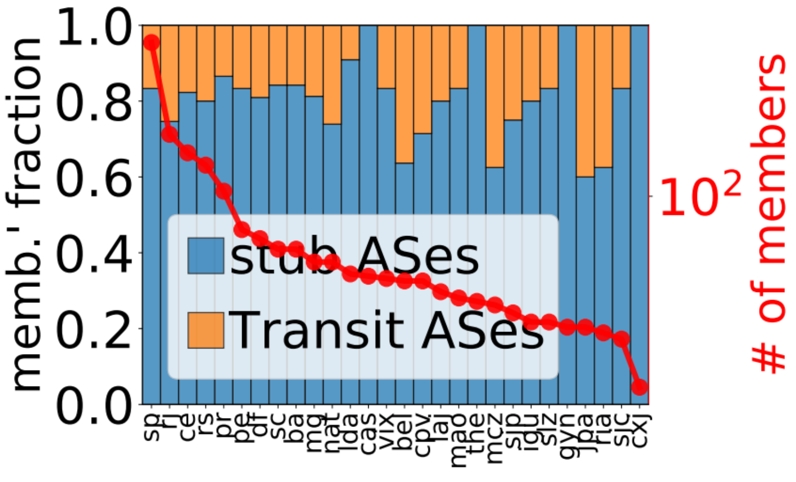
	CABASE	IX.br	PIT-CL	DE-CIX	FR-IX	BKNIX	JINX
1k-10k	16	404	12	299	118	11	25
100-1k	12	95	7	216	42	6	12
10-100	1	16	1	39	13	0	1
1-10	1	1	0	4	1	0	1

IX.br-SP	ASN	16735	262589	7049	61832	28329
	#	903	381	218	209	207
CABASE-BUE	ASN	3549	52361	7049	19037	11664
	#	219	113	100	82	81
PIT Chile-SCL	ASN	7004	22661	52280	19228	14259
	#	88	87	70	57	57

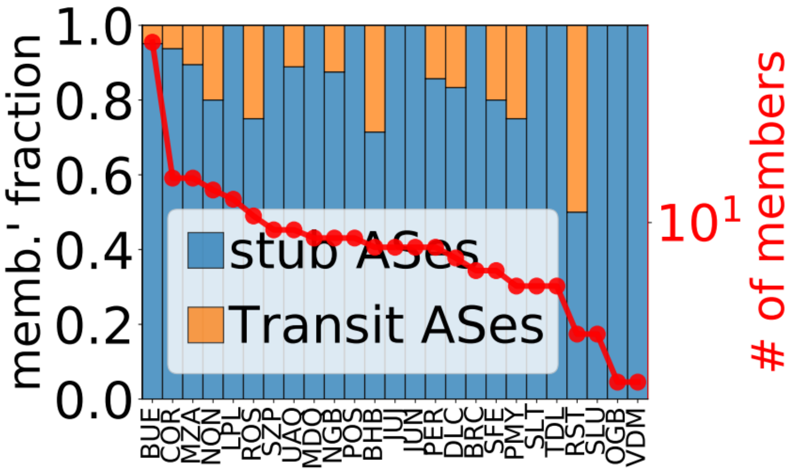
Reaching IXPs: non-transit members



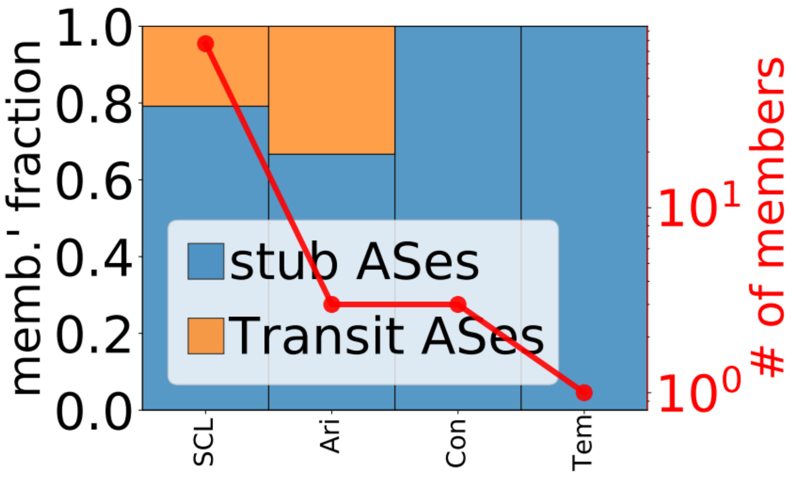
Non-transit members: transit vs stub ASes



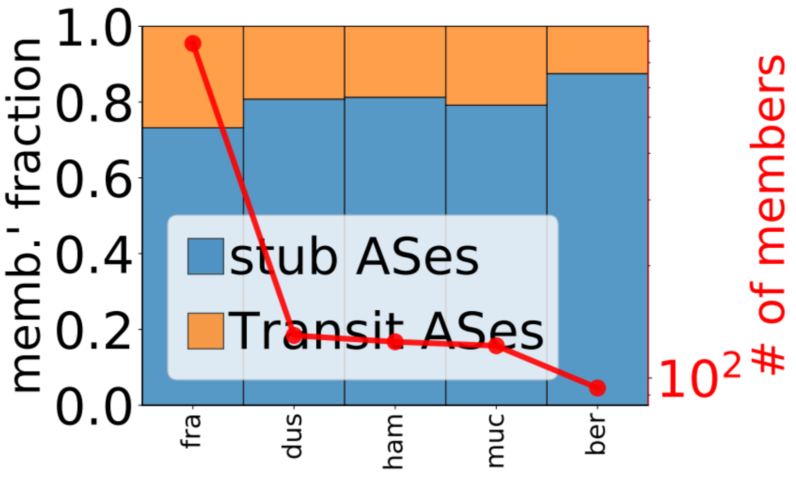
(a) IX.br.



(b) CABASE

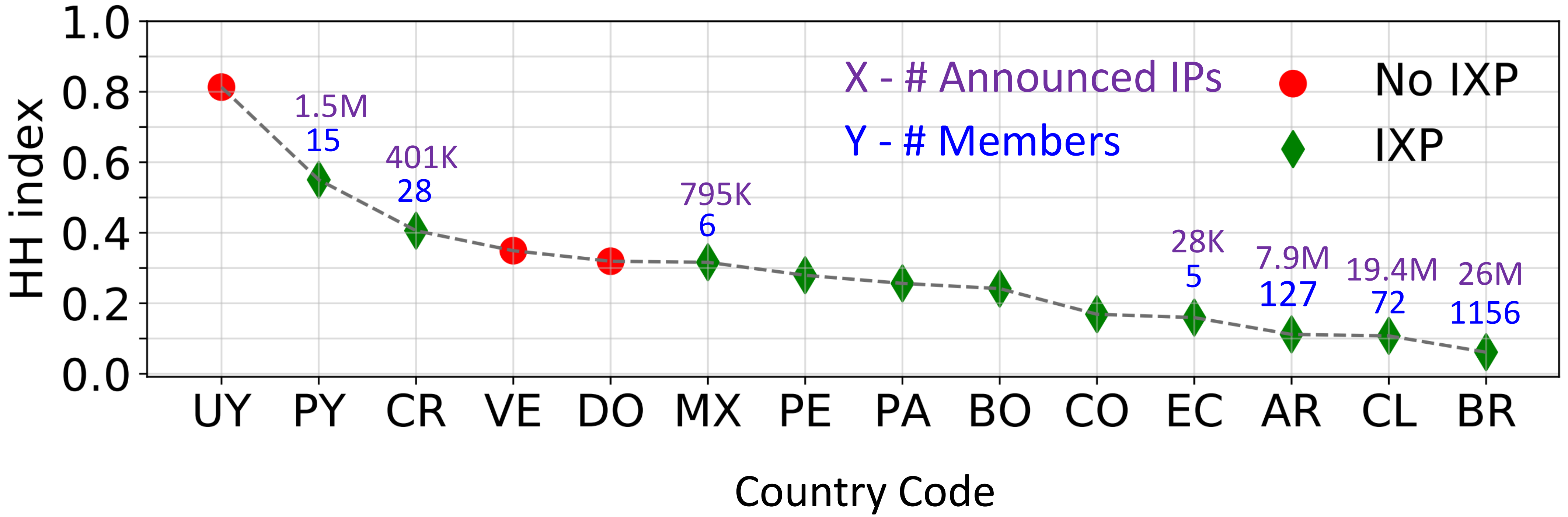


(c) PIT-CL



(d) DE-CIX

IXPs and concentration



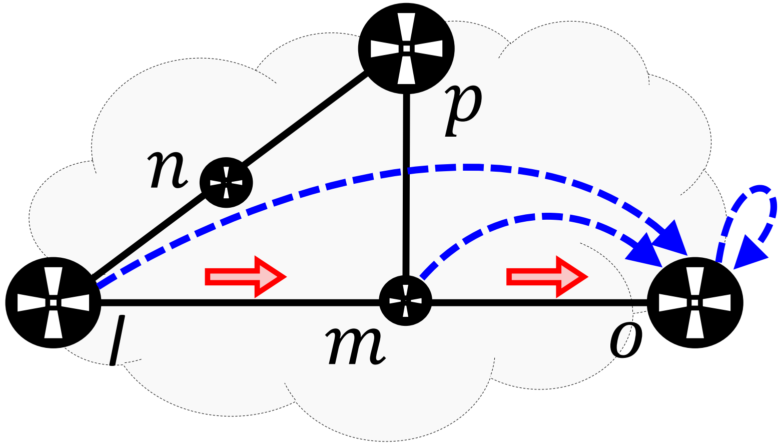
	UY		VE		CR		MX	
ASN	6057*	19422	8048*	6306	11830*	52228	8151	13999
ip-cnt$_{cc}$	2.38M		5.15M		2.42M		24.9M	
ip-cnt	2.15M	90.1k	2.84M	629k	1.52M	197k	13.7M	2.05M
ip-frac	0.90	0.04	0.55	0.14	0.63	0.08	0.55	0.08

Routing Inconsistencies, Forwarding Alterations, Forwarding Detours

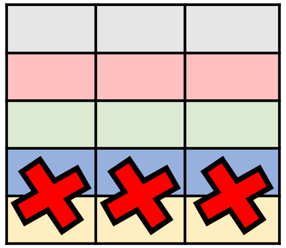
What produces FDs?

- **BGP(d):** the exit point to use to reach d
- **IGP o BGP(d):** the next-hop towards that exit point

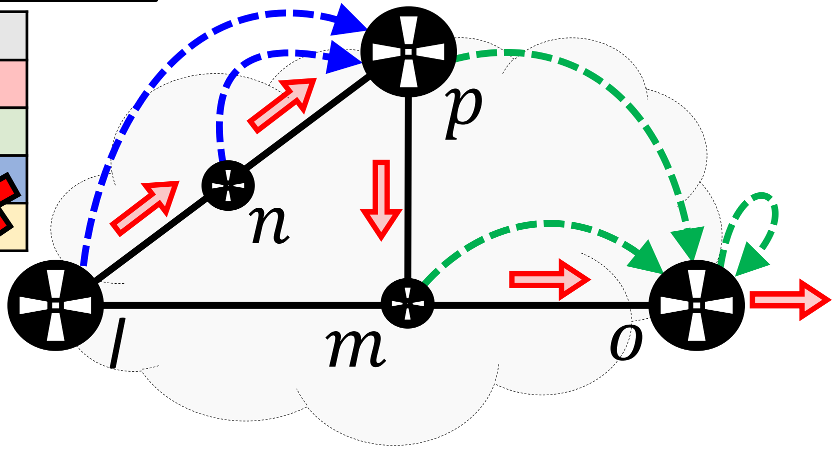
Ideal: best IGP path



Memory limitation



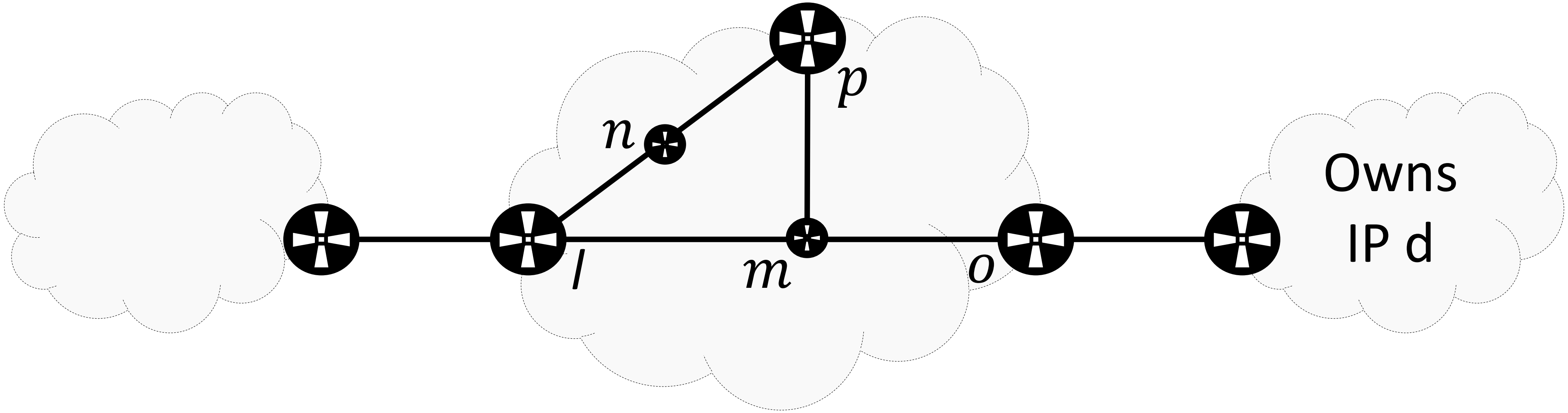
RIes and FD



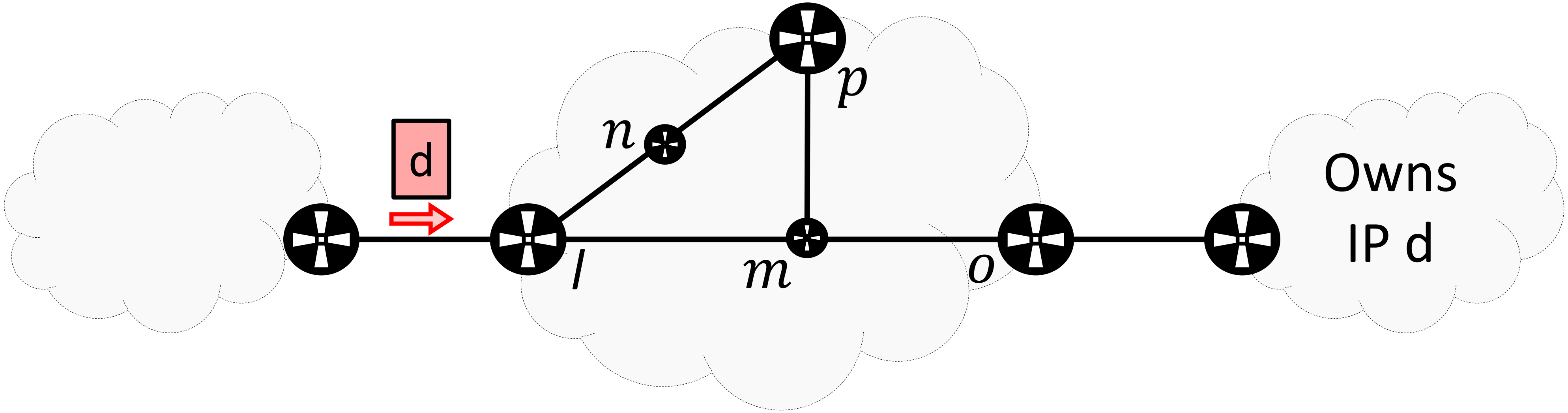
- Routing consistency
  - Agreement on BGP(d)

- **Routing inconsistency (RI)**
  - Disagreement on BGP(d)
  - May lead to a FD
  - Due to scalability workarounds

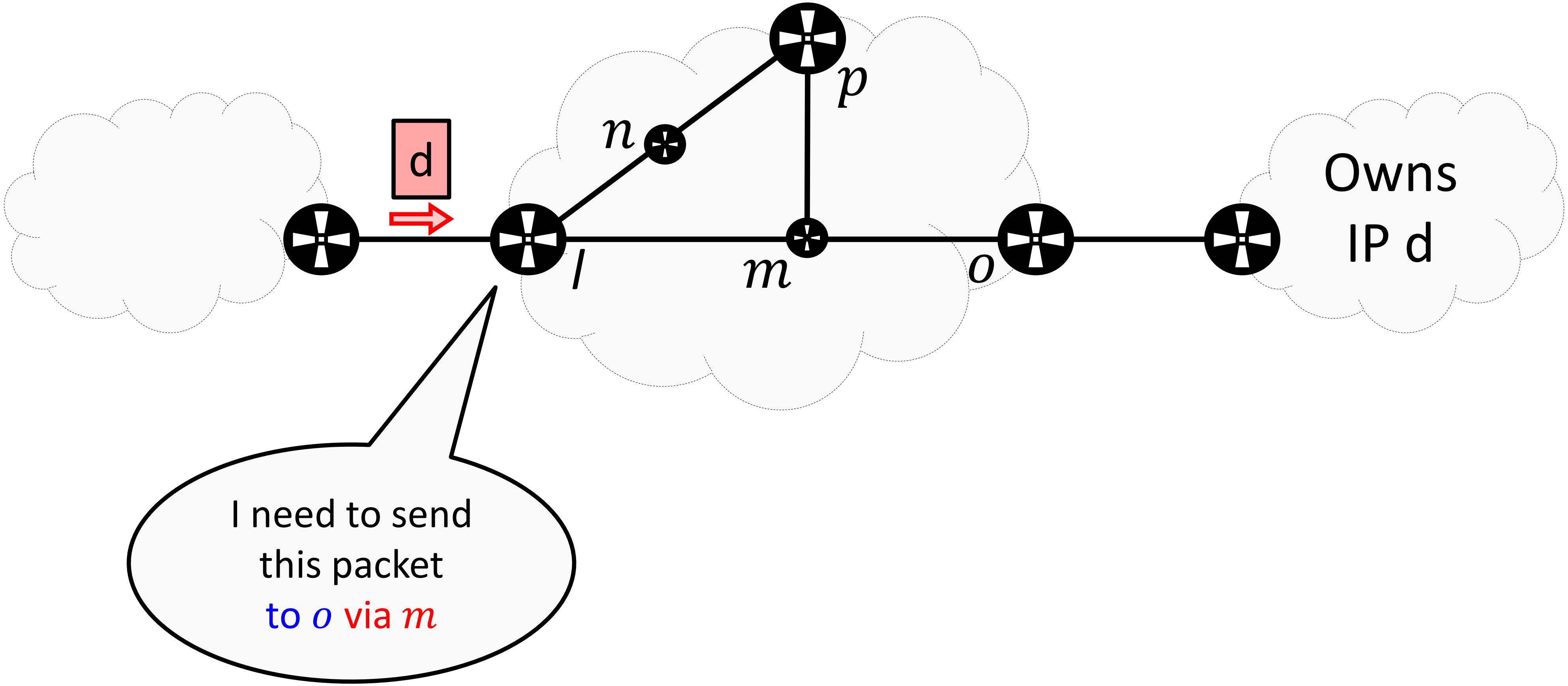
How does the forwarding work?



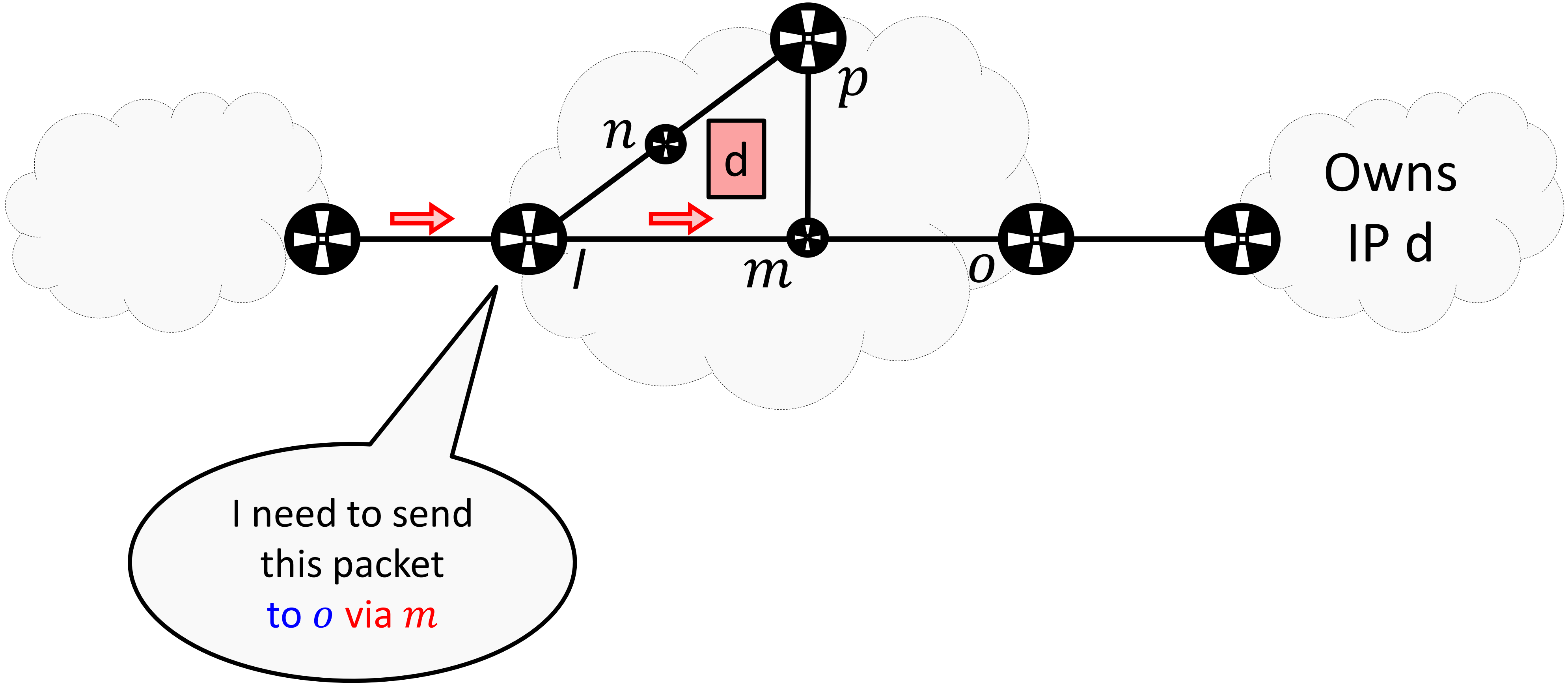
How does the forwarding work?



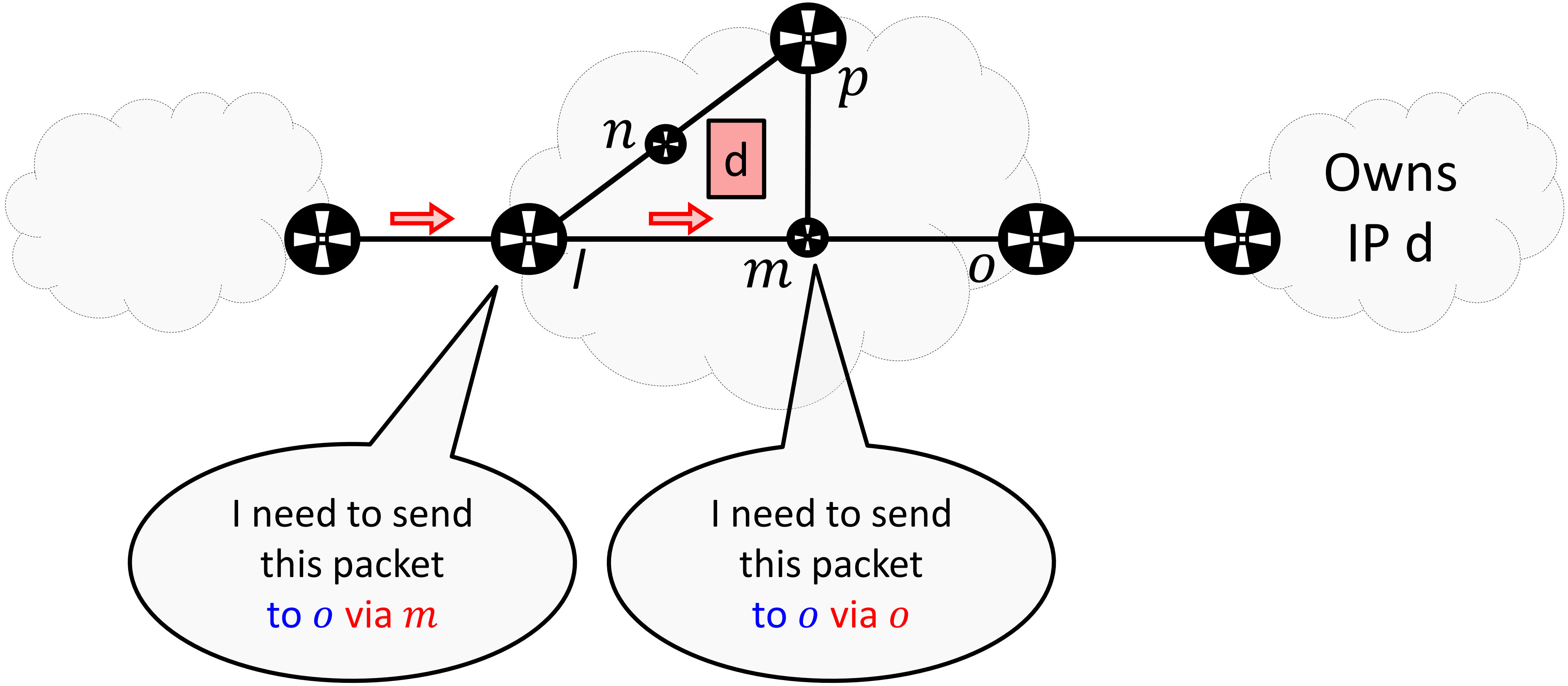
How does the forwarding work?



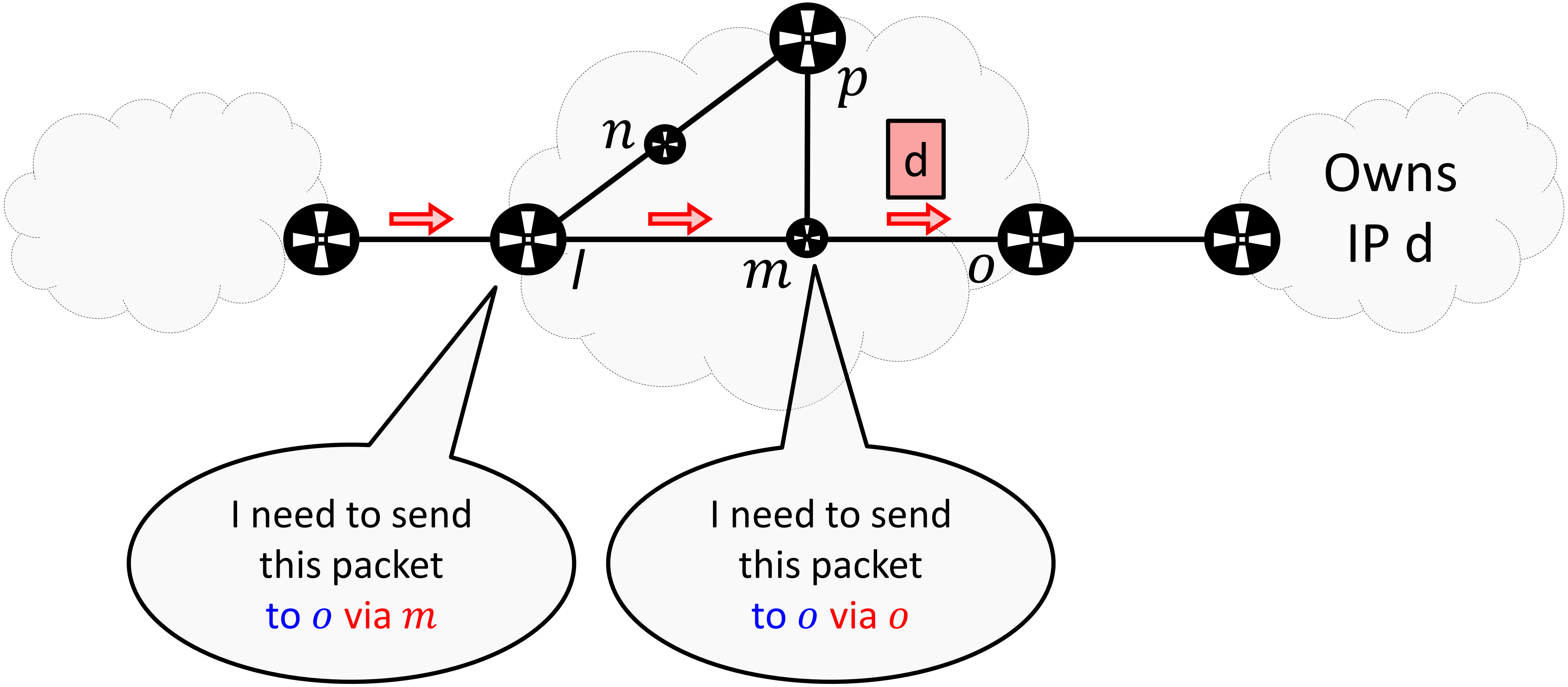
How does the forwarding work?



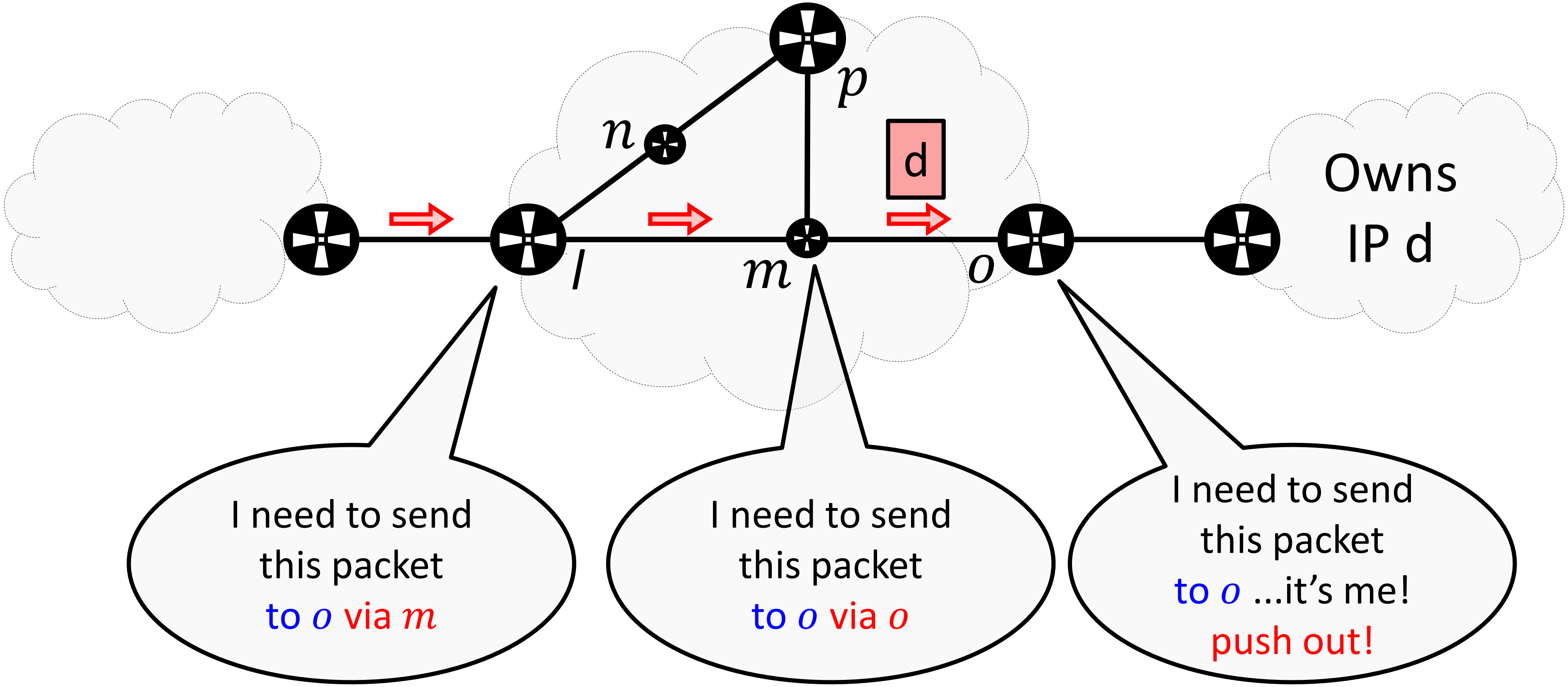
How does the forwarding work?



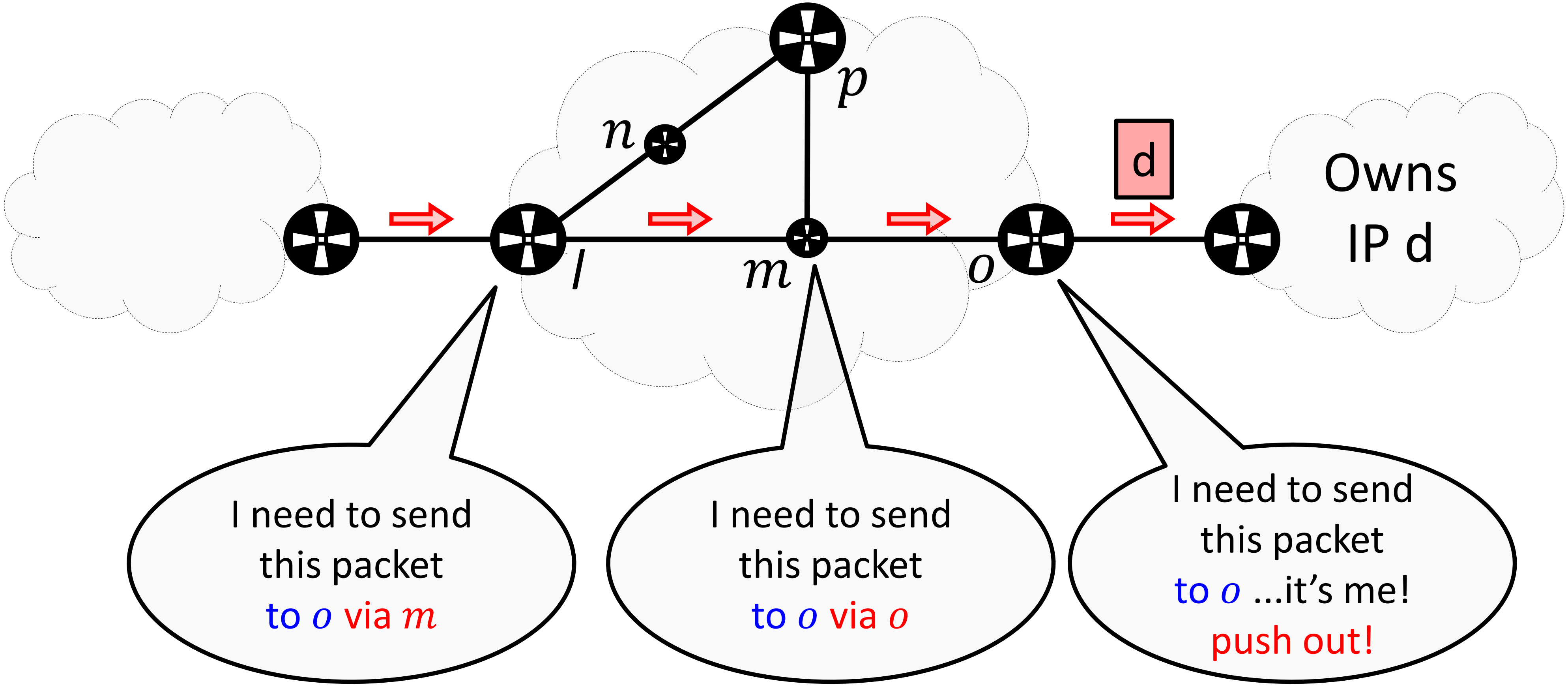
How does the forwarding work?



How does the forwarding work?

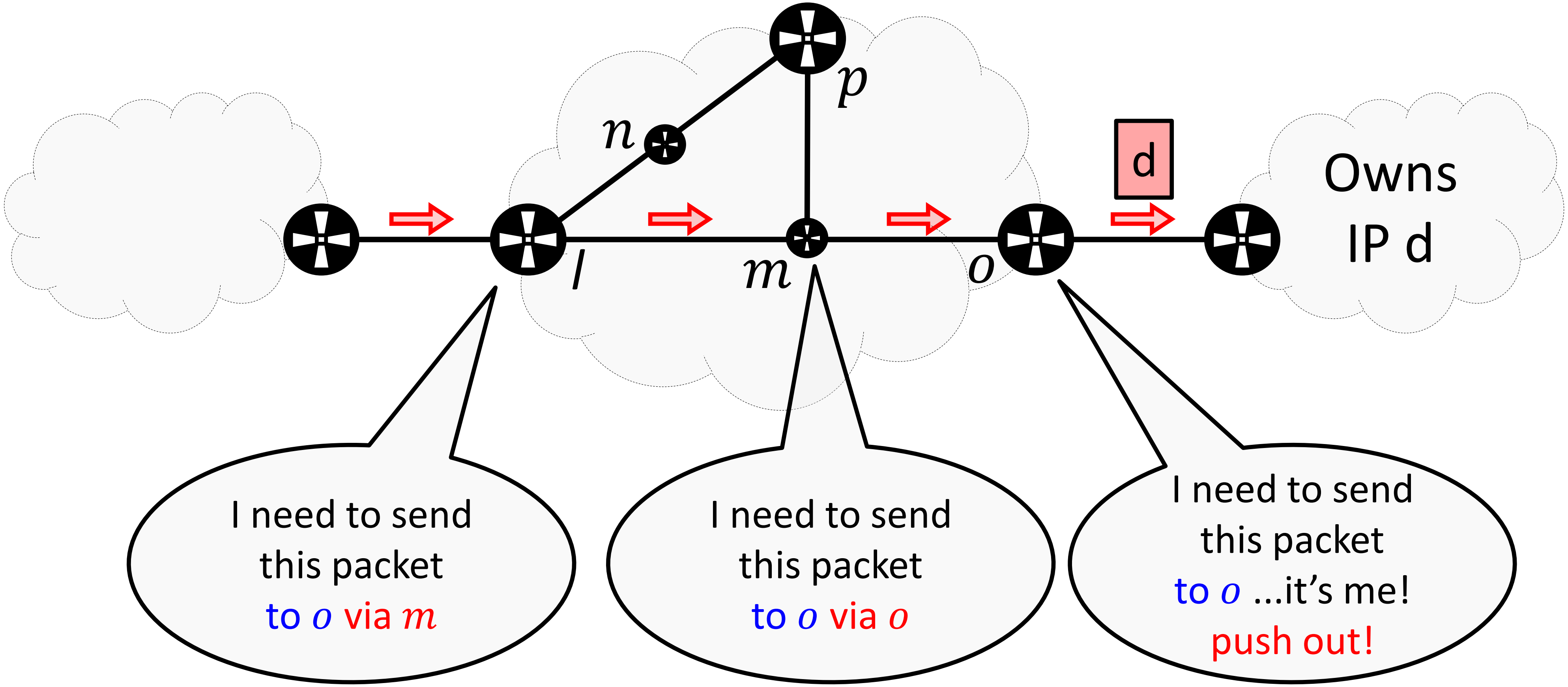


How does the forwarding work?



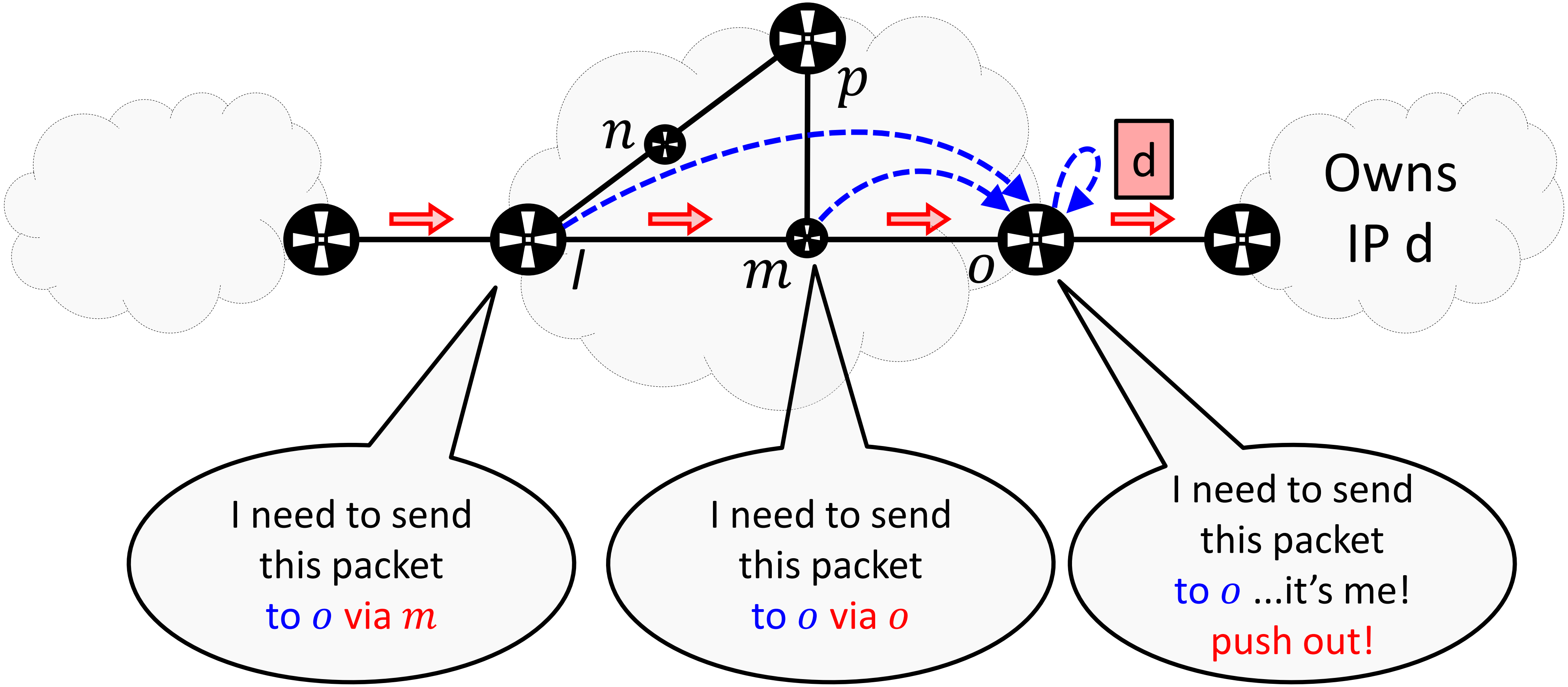
Forwarding Model

- **BGP(d):** the exit point to use to reach d
- **IGP o BGP(d):** the next-hop towards that exit point



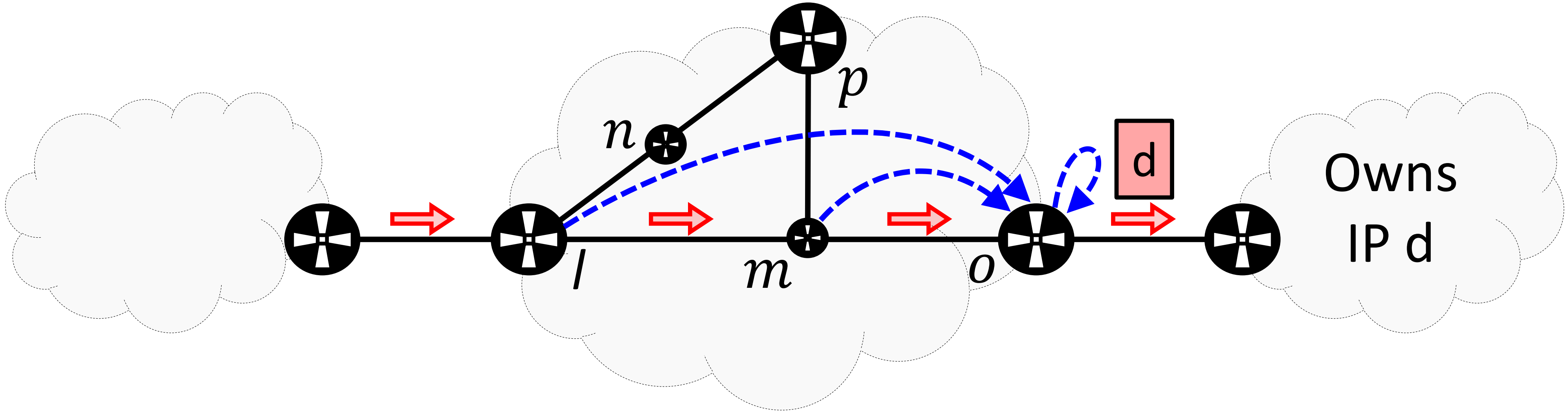
Forwarding Model

- **BGP(d):** the exit point to use to reach d
- **IGP o BGP(d):** the next-hop towards that exit point



Forwarding Model

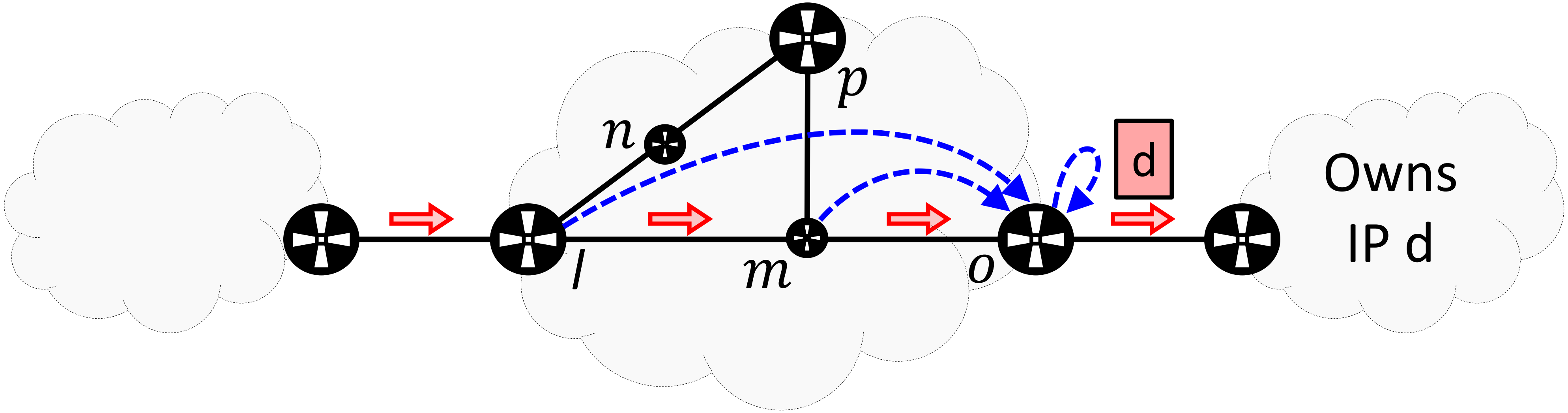
- **BGP(d):** the exit point to use to reach d
- **IGP o BGP(d):** the next-hop towards that exit point



- Routing consistency – BGP(d) is the same for all routers

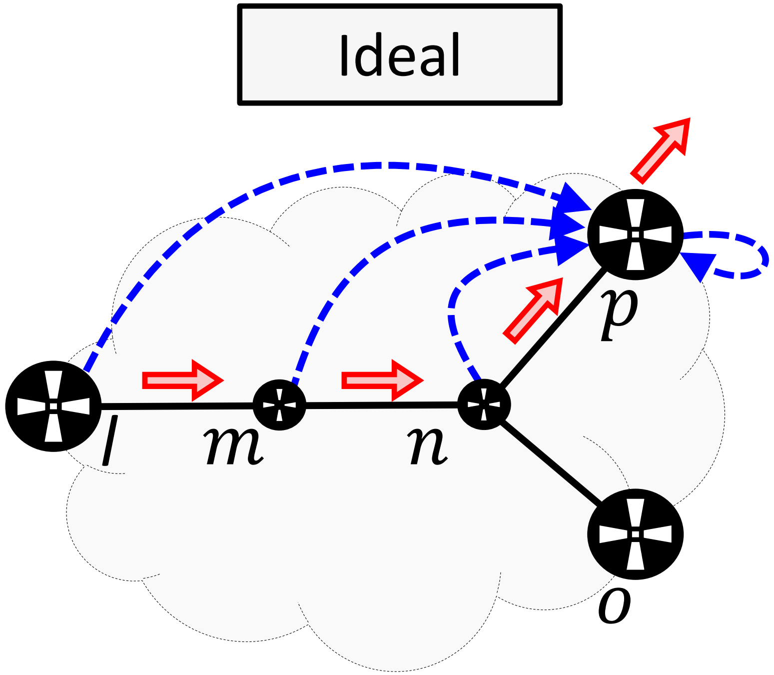
Forwarding Model

- **BGP(d):** the exit point to use to reach d
- **IGP o BGP(d):** the next-hop towards that exit point

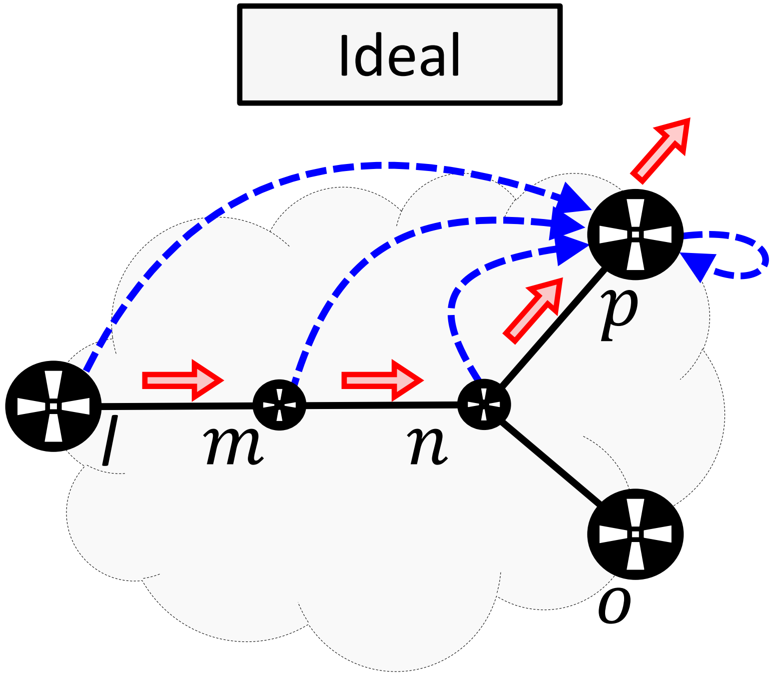


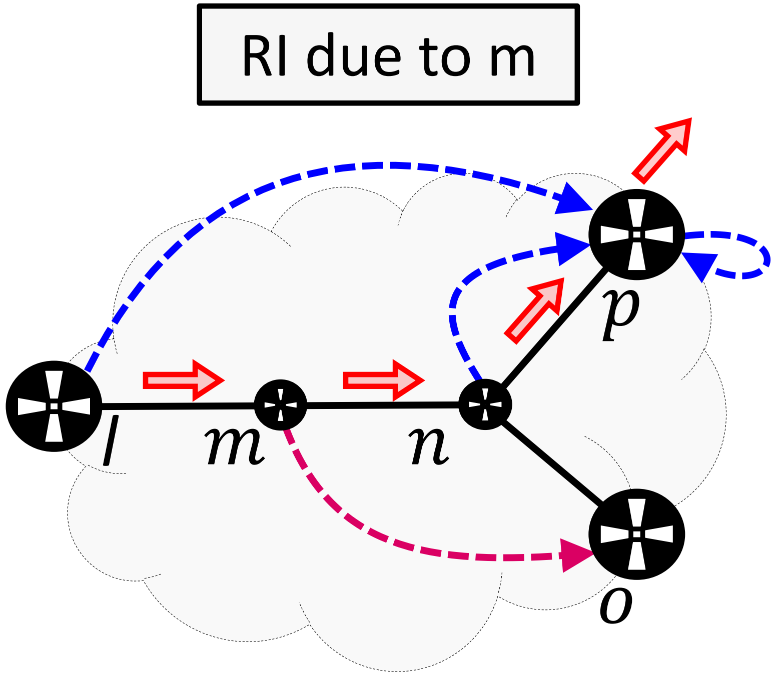
- Routing consistency – BGP(d) is the same for all routers
- **Routing inconsistency (RI) – routers disagree on BGP(d)**

What happens when RIes occur?

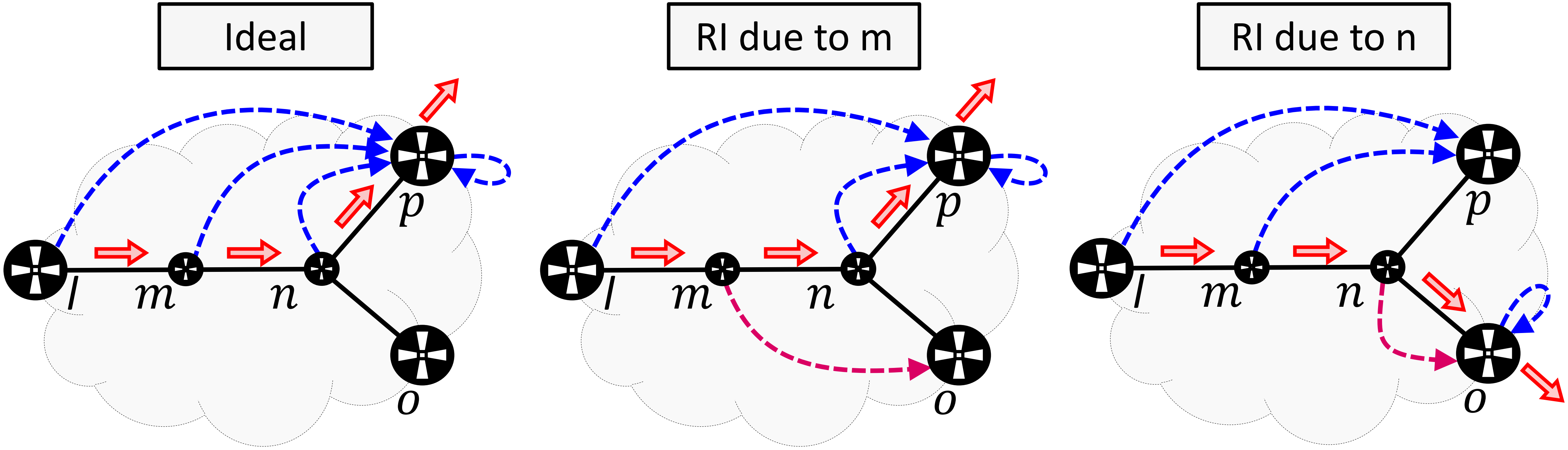


What happens when RIes occur?

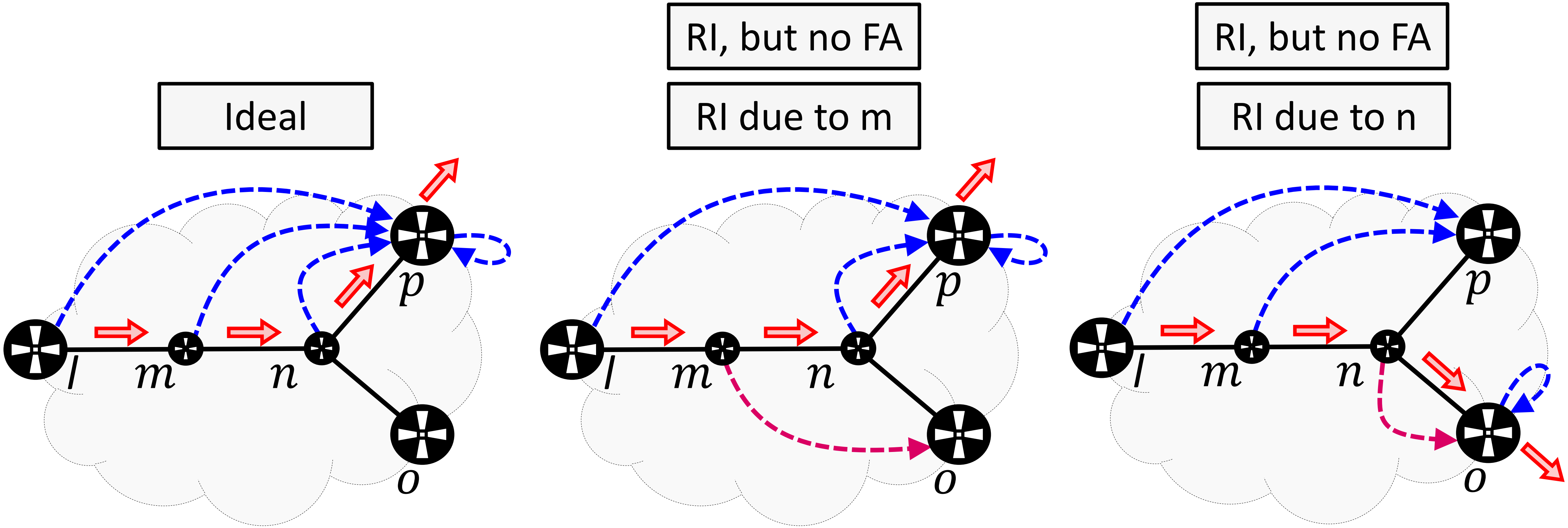




What happens when RIes occur?



What happens when RIes occur?

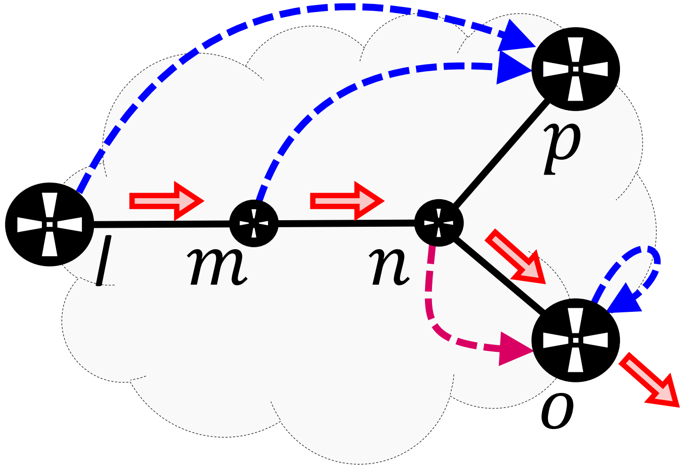


- **Forwarding alteration (FA) –** RI leading to a new route

What happens when FAs occur?

Example I

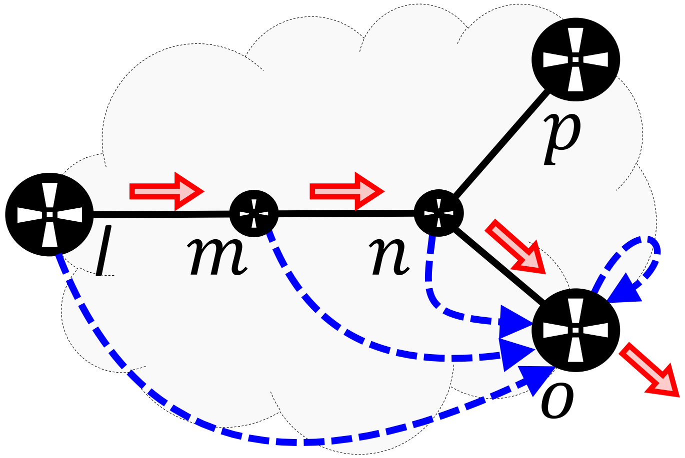
RI and FA



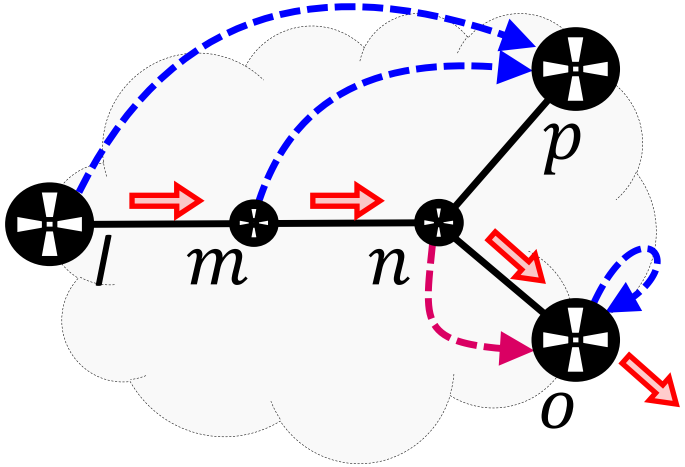
What happens when FAs occur?

Example I

Ideal



RI and FA

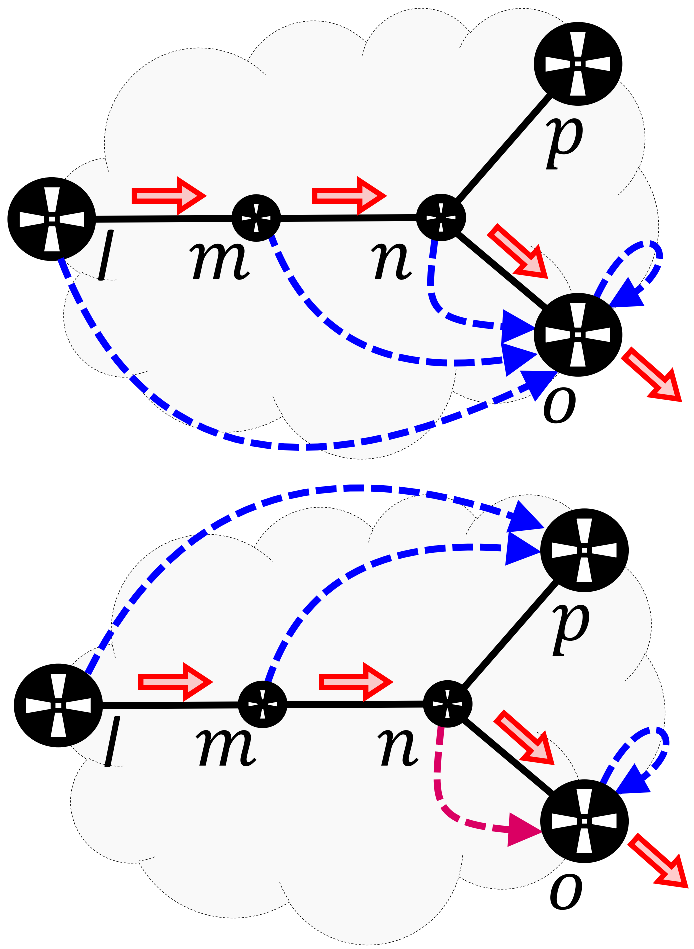


What happens when FAs occur?

Example I

Ideal

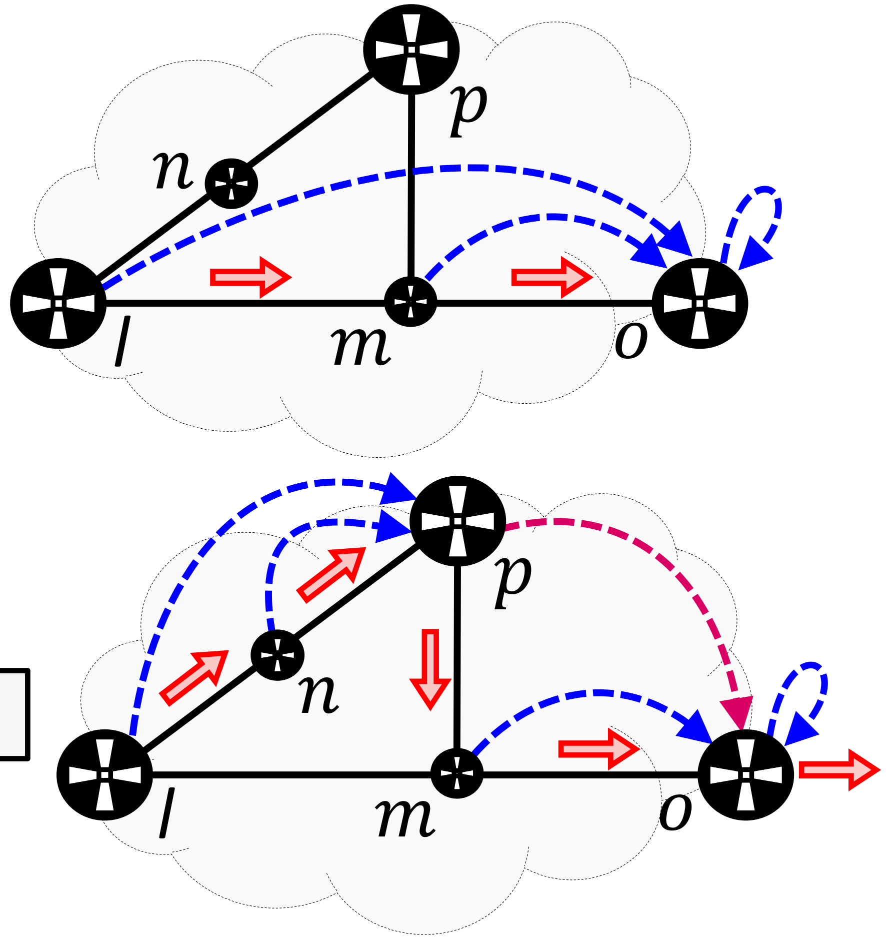
RI and FA



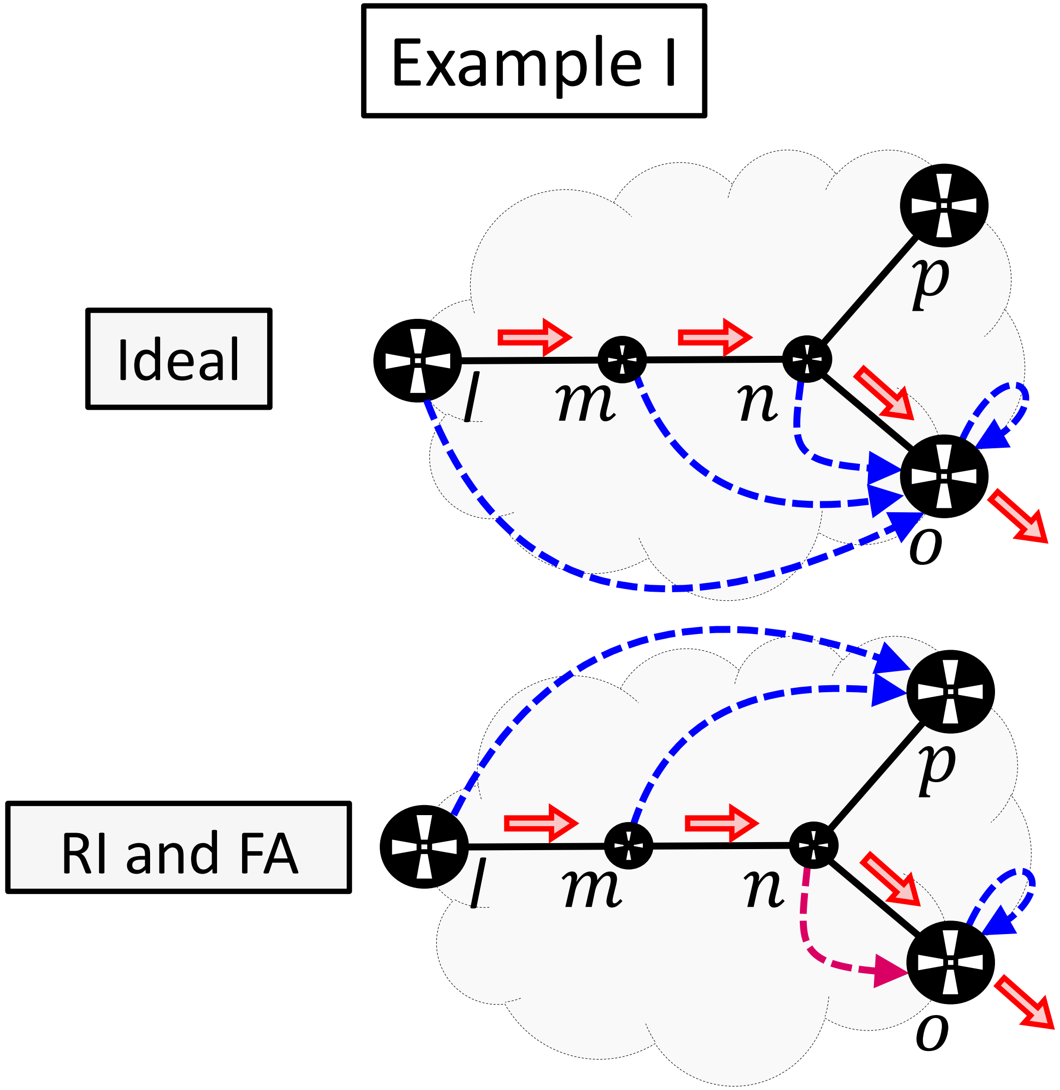
Example II

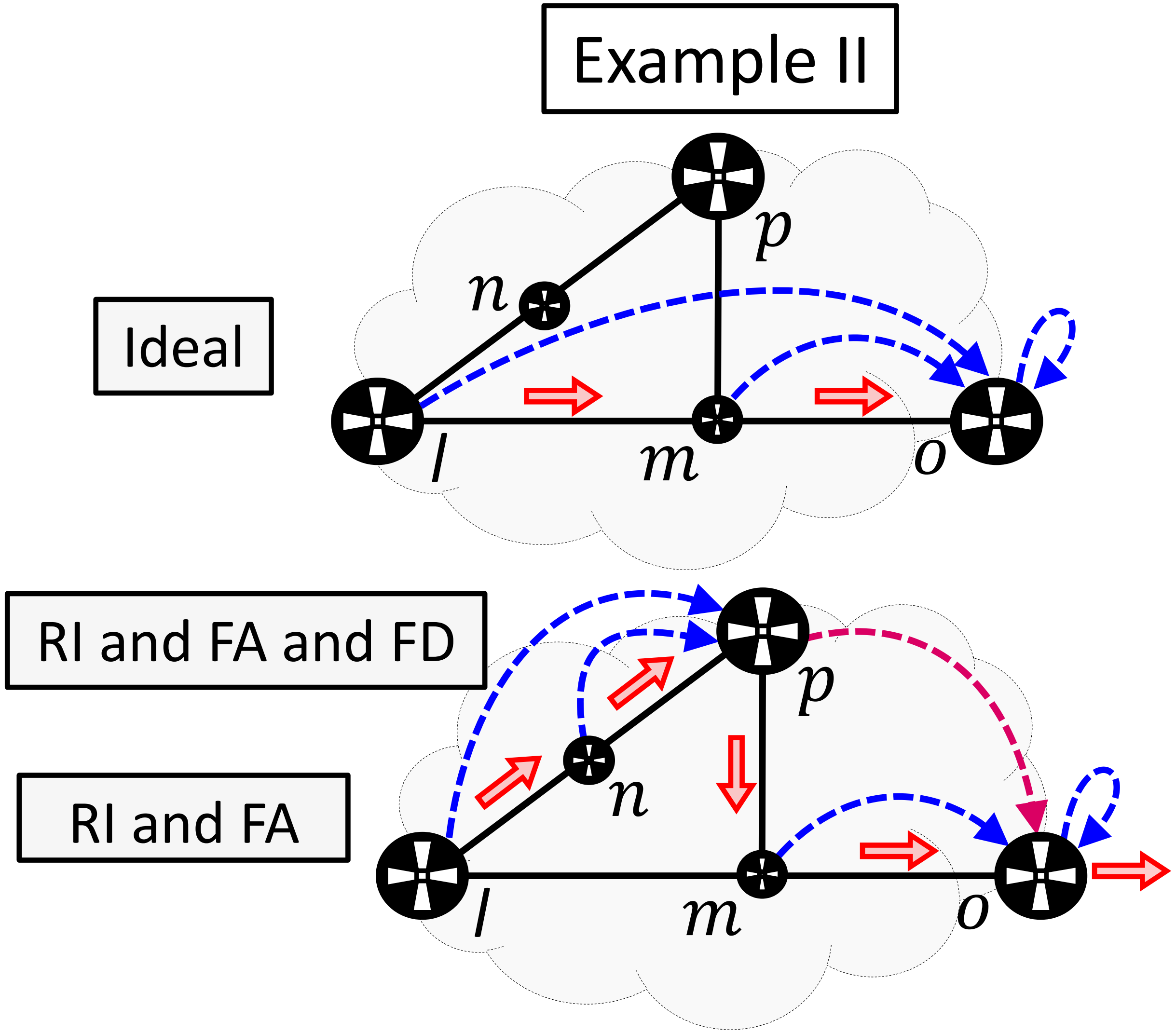
Ideal

RI and FA



What happens when FAs occur?



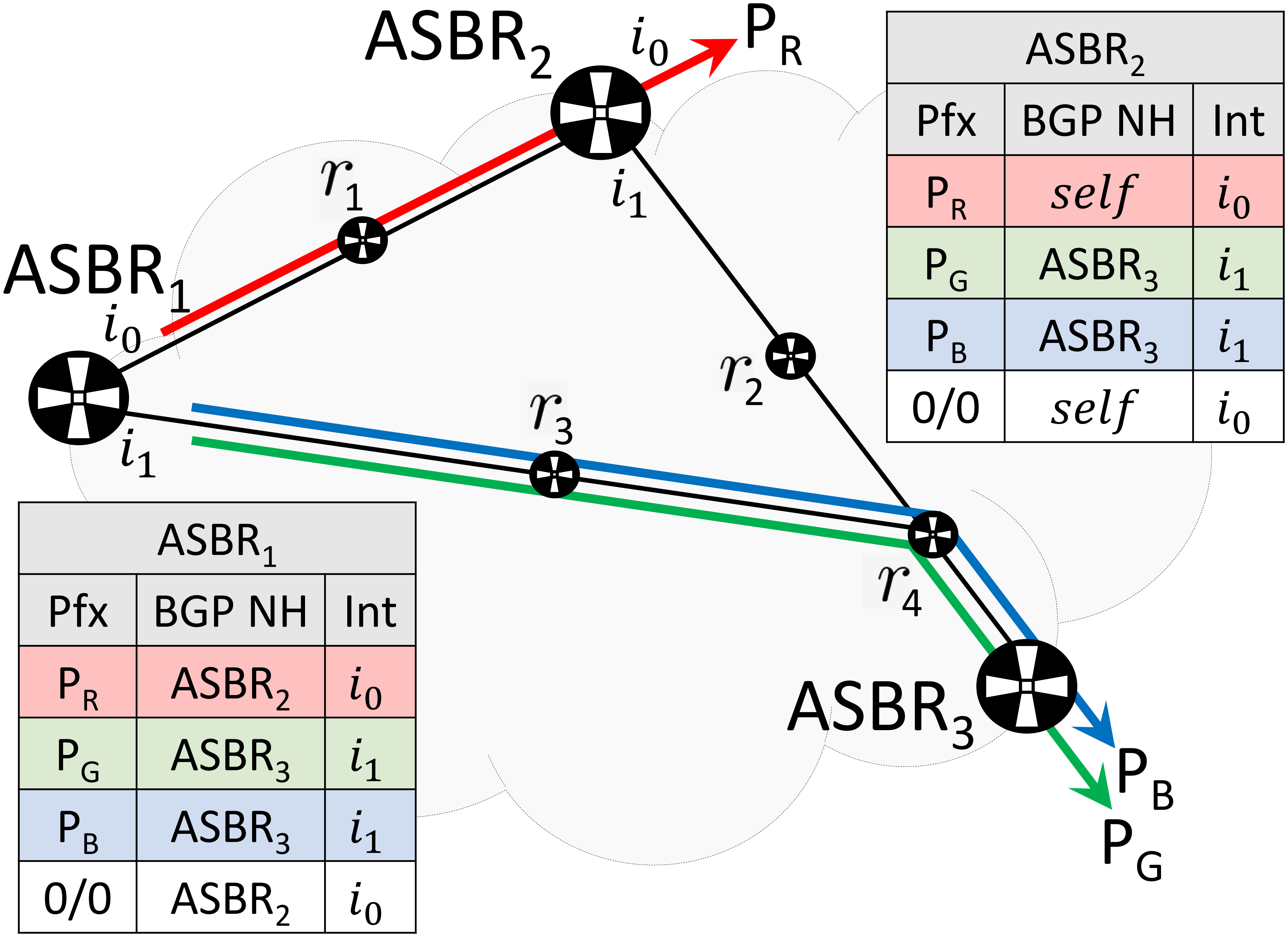


- **Forwarding Detour (FD) –** FA leading to a sub-optimal route

Conclusions

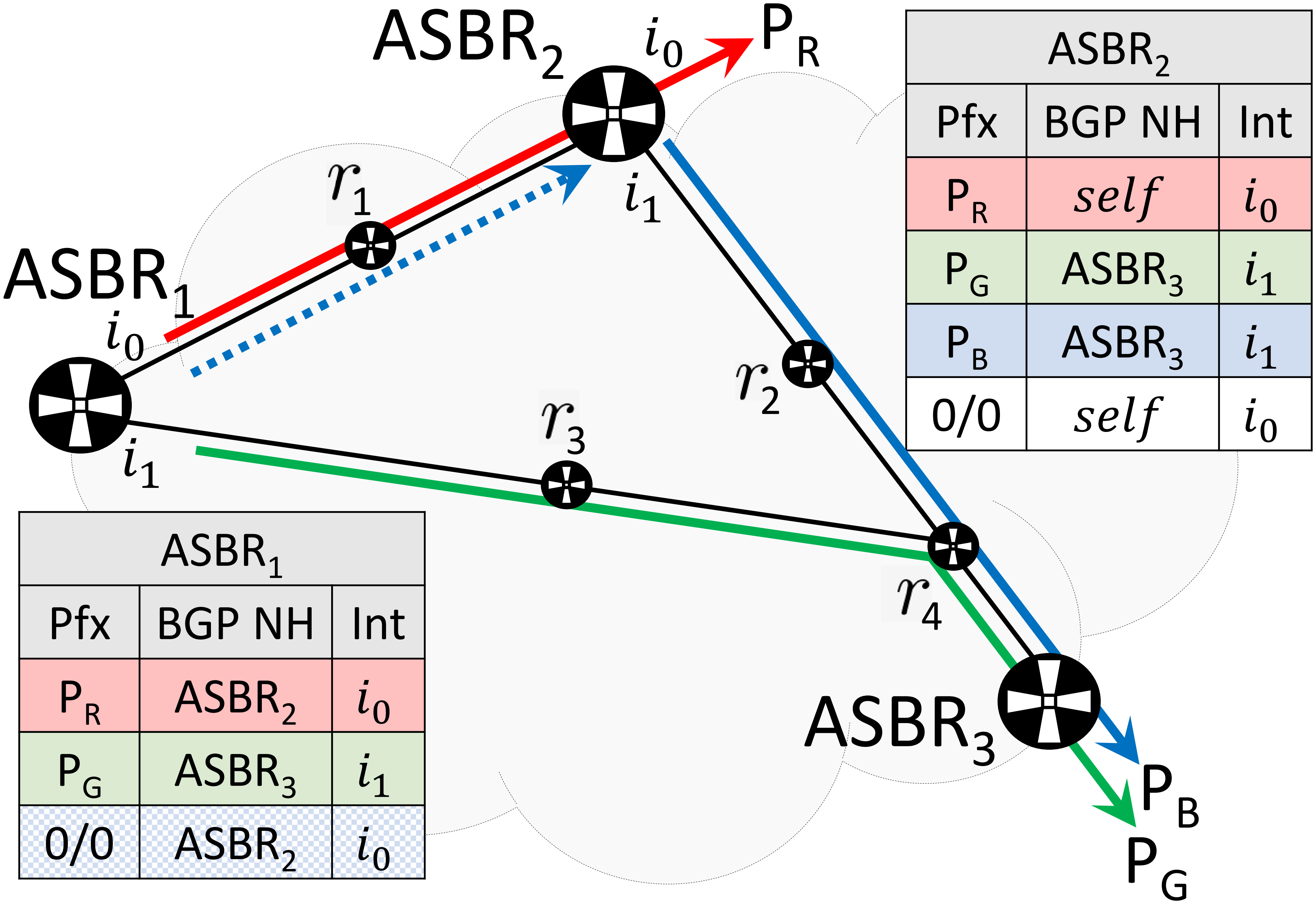
- A forwarding model
- Two new concepts: RIes and FAs
- Two theorems: FDs $\Rightarrow$ FAs $\Rightarrow$ RIes
- Observable FDs are a lower bound of RIes

Full-FIB vs Partial-FIB



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

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

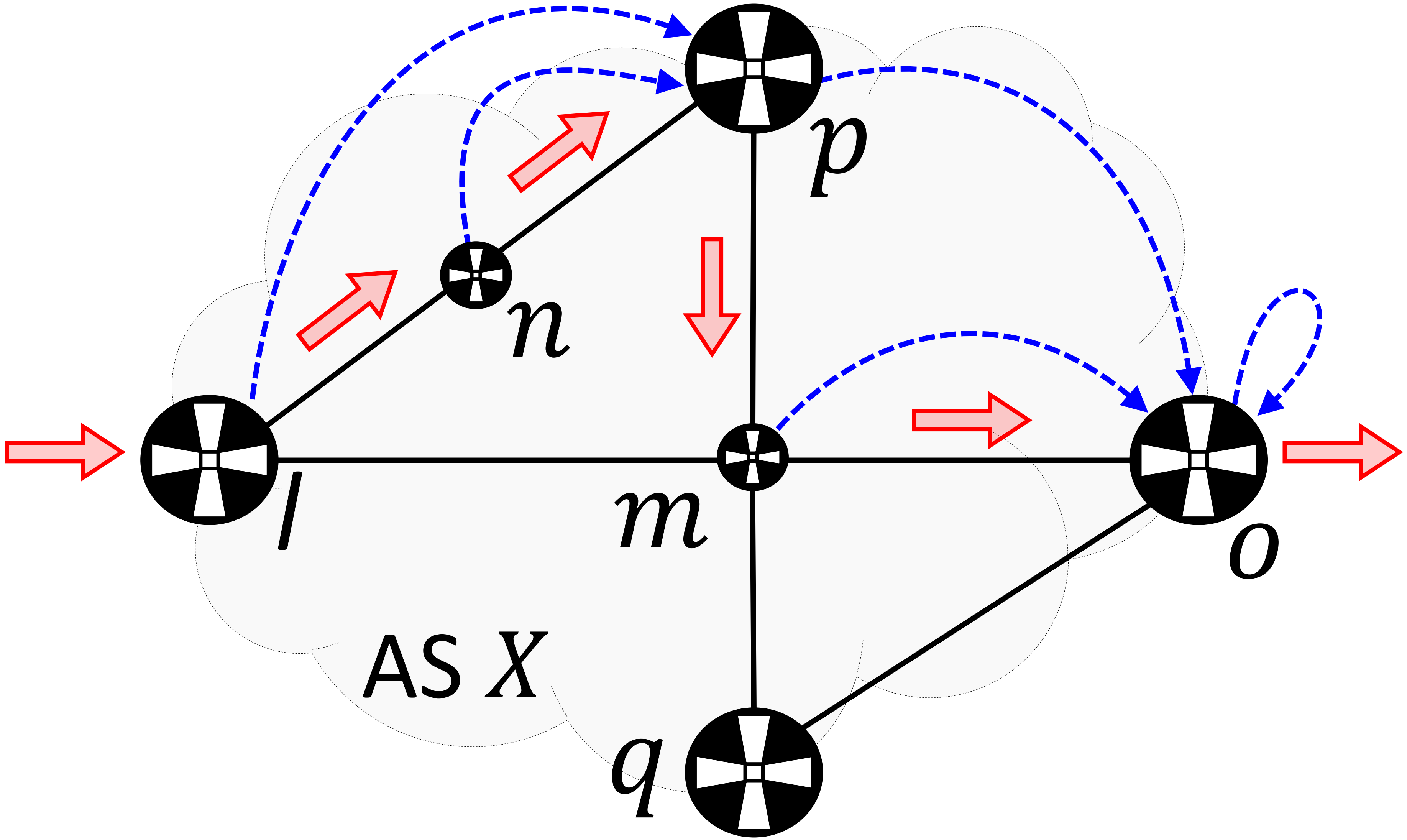


ASBR$_2$		
Pfx	BGP NH	Int
P$_R$	*self*	i_0
P$_G$	ASBR$_3$	i_1
P$_B$	ASBR$_3$	i_1
0/0	*self*	i_0

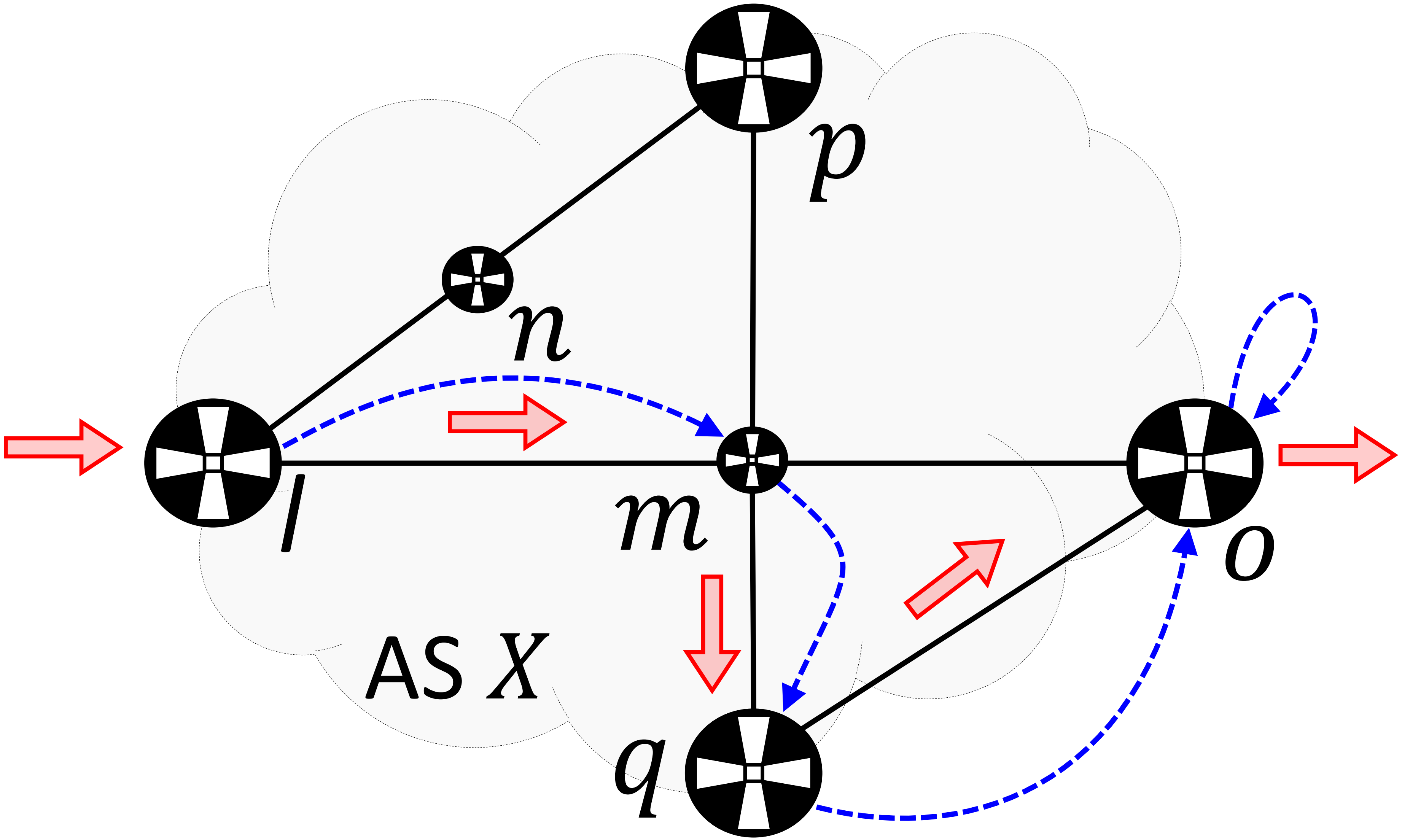
ASBR$_1$		
Pfx	BGP NH	Int
P$_R$	ASBR$_2$	i_0
P$_G$	ASBR$_3$	i_1
0/0	ASBR$_2$	i_0

FDs: may be a set of routes

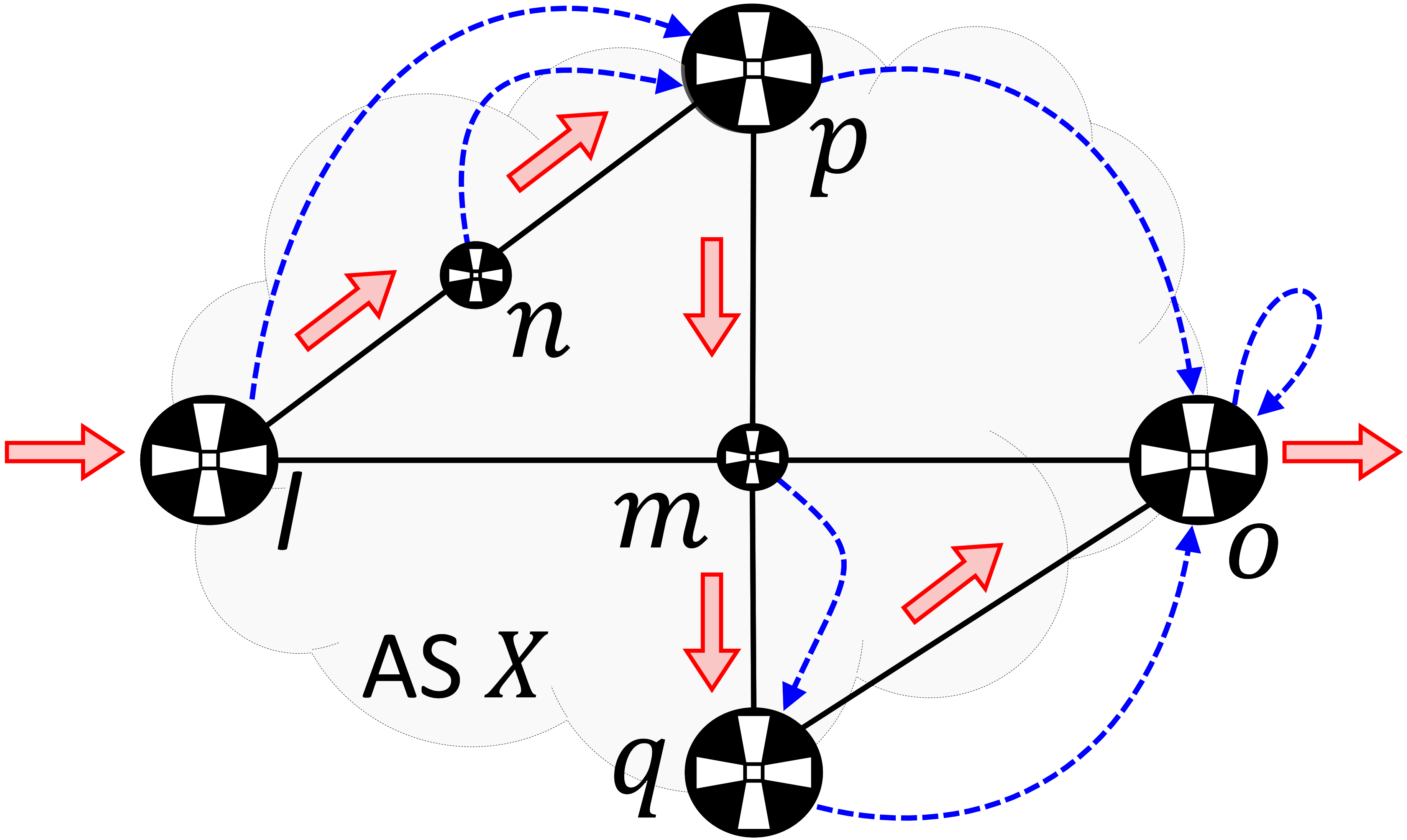
Forwarding Detour I



Forwarding Detour II



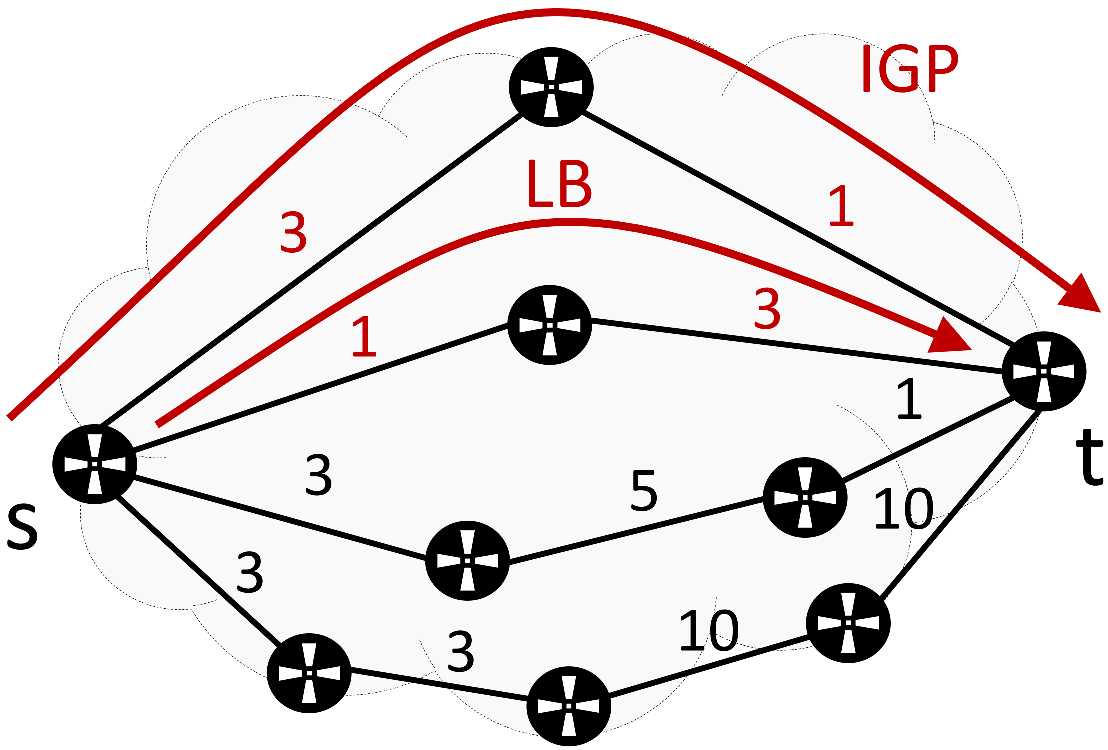
Forwarding Detour III



Load Balancing
F-LB and C-LB

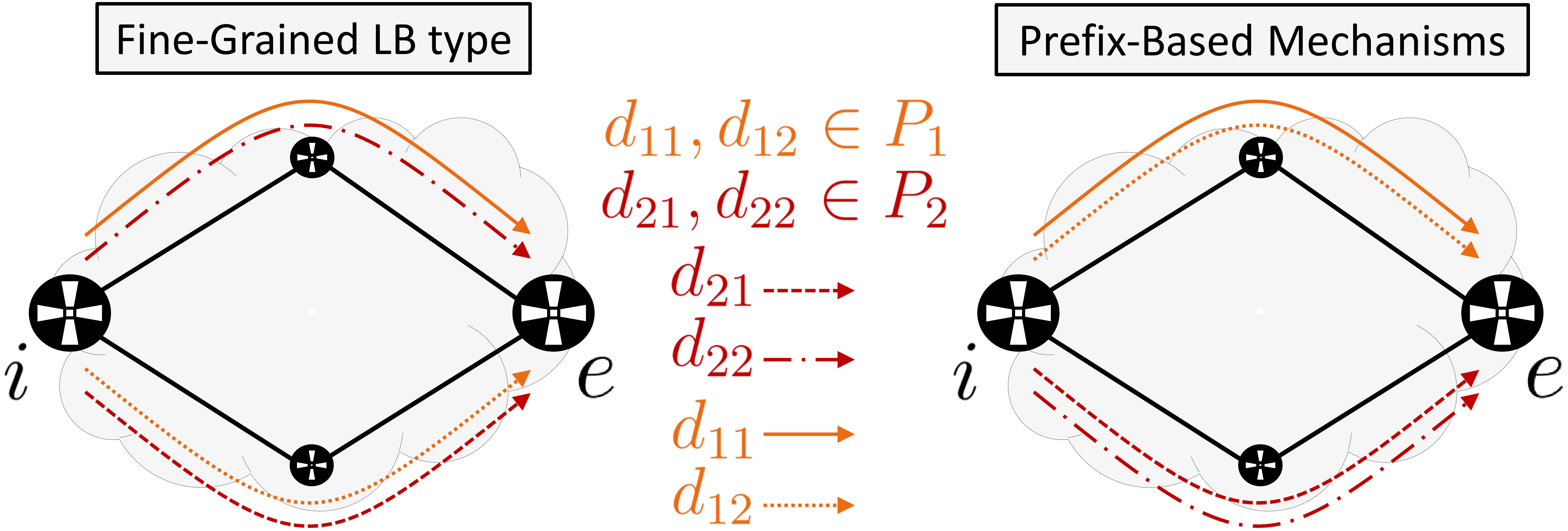
Load Balancing (LB)

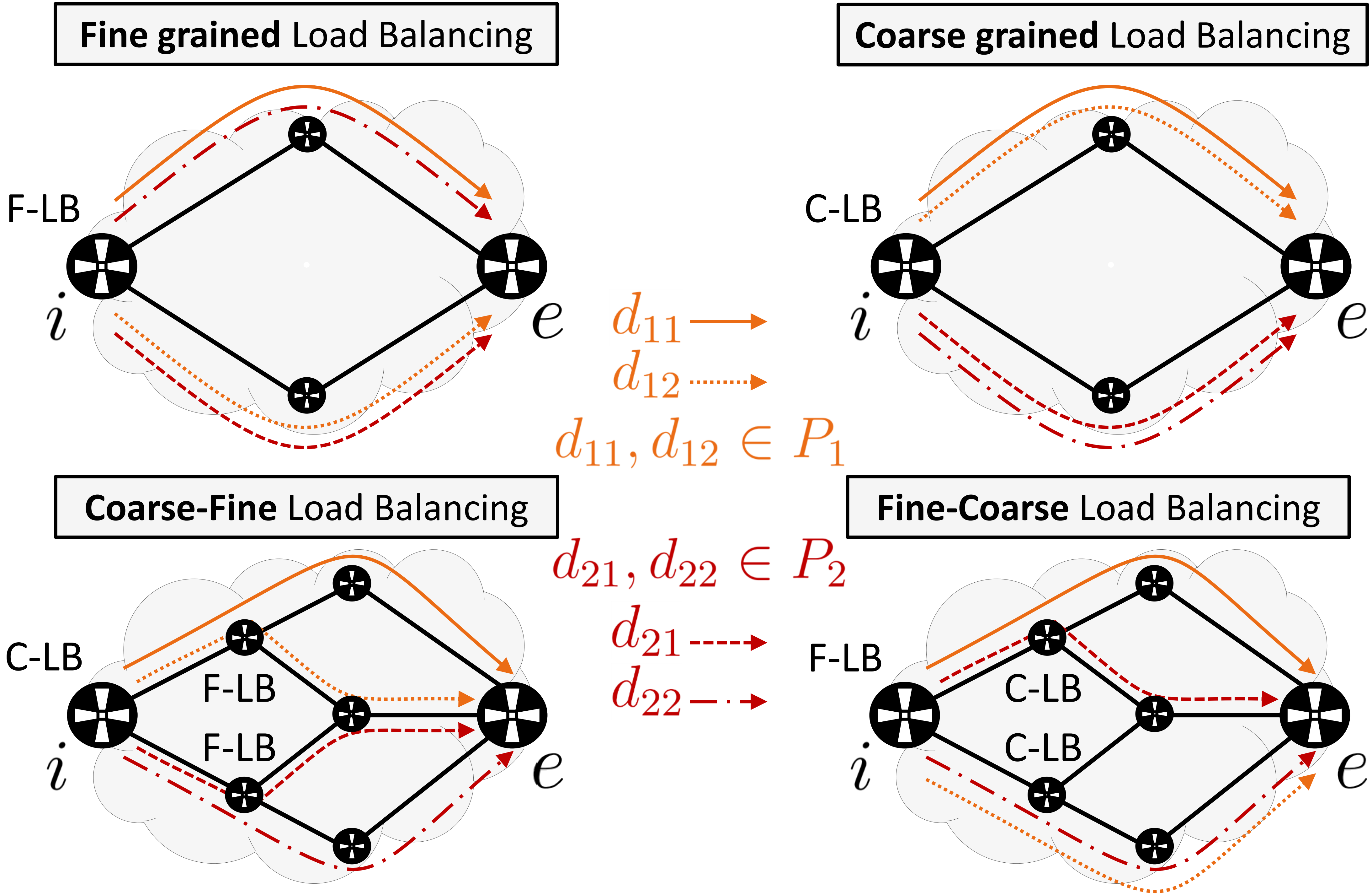
- There exist different LB flavors:
  - F-LB: different destination, then route may change
  - C-LB: same prefix, same route



F-LB		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1	⦿ ⦿	⦿ ⦿		
	P_2	⦿ ⦿	⦿ ⦿		
	P_3	⦿ ⦿	⦿ ⦿		
	P_4	⦿ ⦿	⦿ ⦿		
	P_5	⦿ ⦿	⦿ ⦿		
	P_6	⦿ ⦿	⦿ ⦿		
	P_7	⦿ ⦿	⦿ ⦿		
	P_8	⦿ ⦿	⦿ ⦿		

C-LB		Routes			
		R_1	R_2	R_3	R_4
Prefixes	P_1	⦿⦿ ⦿⦿			
	P_2	⦿⦿ ⦿⦿			
	P_3	⦿⦿ ⦿⦿			
	P_4	⦿⦿ ⦿⦿			
	P_5		⦿⦿ ⦿⦿		
	P_6		⦿⦿ ⦿⦿		
	P_7		⦿⦿ ⦿⦿		
	P_8		⦿⦿ ⦿⦿		


Fine-Grained LB type
Prefix-Based Mechanisms
$d_{11}, d_{12} \in P_1$
$d_{21}, d_{22} \in P_2$
d_{21}
d_{22}
d_{11}
d_{12}
i
e
i
e

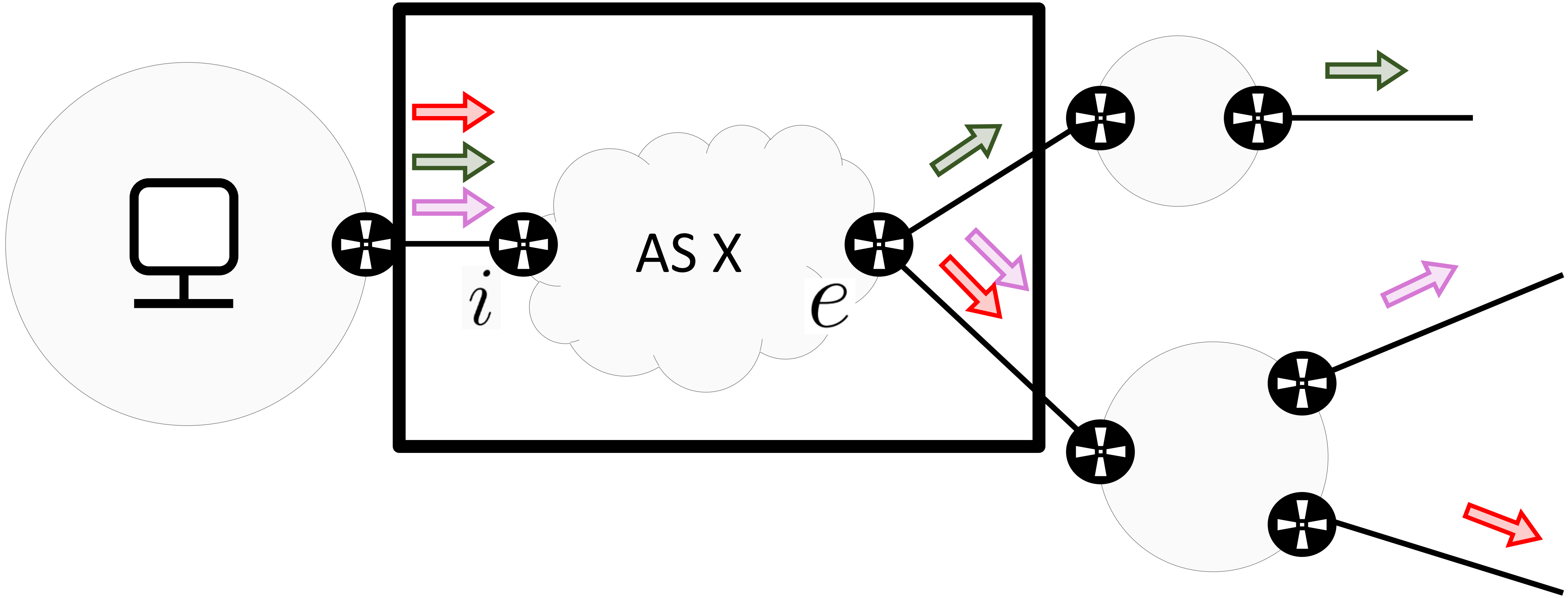


Fine grained Load Balancing
F-LB
i
e
Coarse grained Load Balancing
C-LB
i
e
d_{11}
d_{12}
$d_{11}, d_{12} \in P_1$
Coarse-Fine Load Balancing
C-LB
F-LB
F-LB
i
e
$d_{21}, d_{22} \in P_2$
d_{21}
d_{22}
Fine-Coarse Load Balancing
F-LB
C-LB
C-LB
i
e

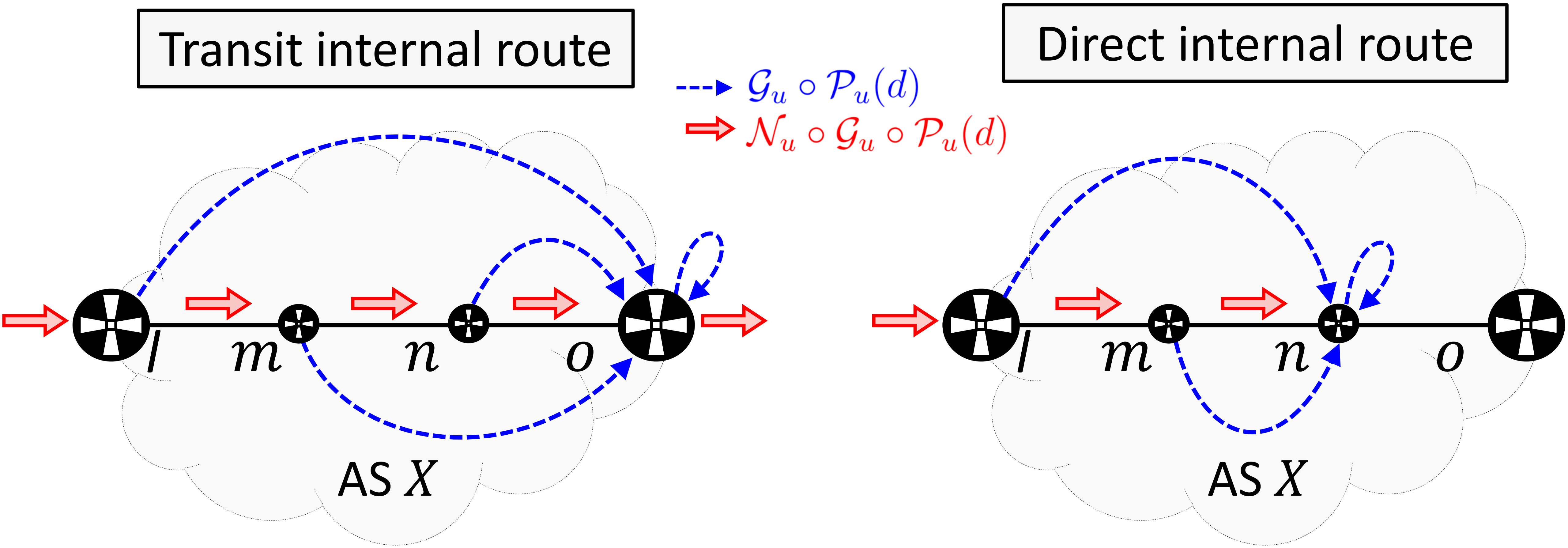
FD-detector

Exploration phase

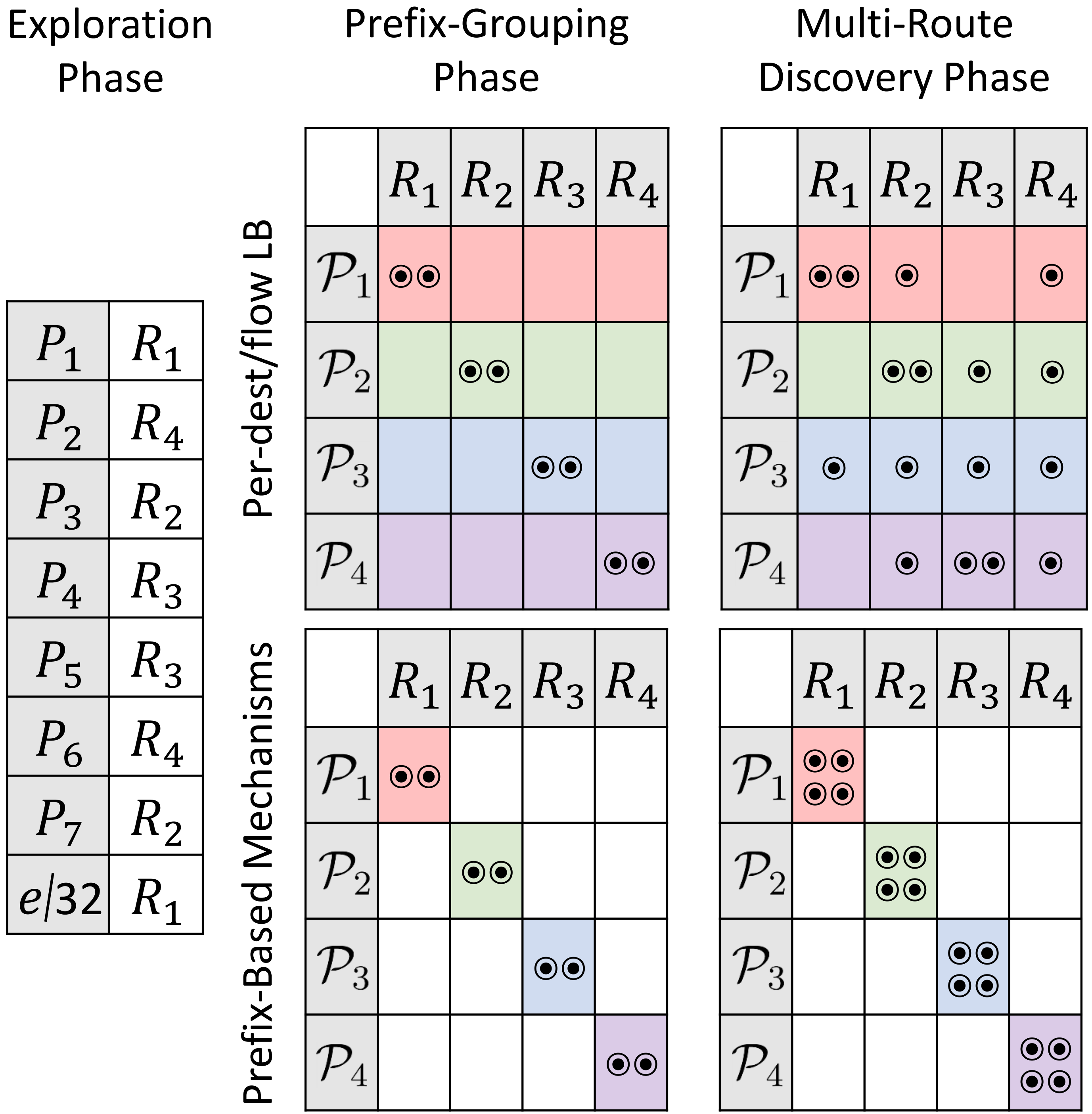
- Run traces to randomly chosen destinations
- Identify ASBR couples (i, e) in each traversed AS X
- Trace router e and annotate routes traversed for each prefix



P_1	R_1
P_2	R_4
P_3	R_2
P_4	R_3
P_5	R_3
P_6	R_4
P_7	R_2
$e/32$	R_1



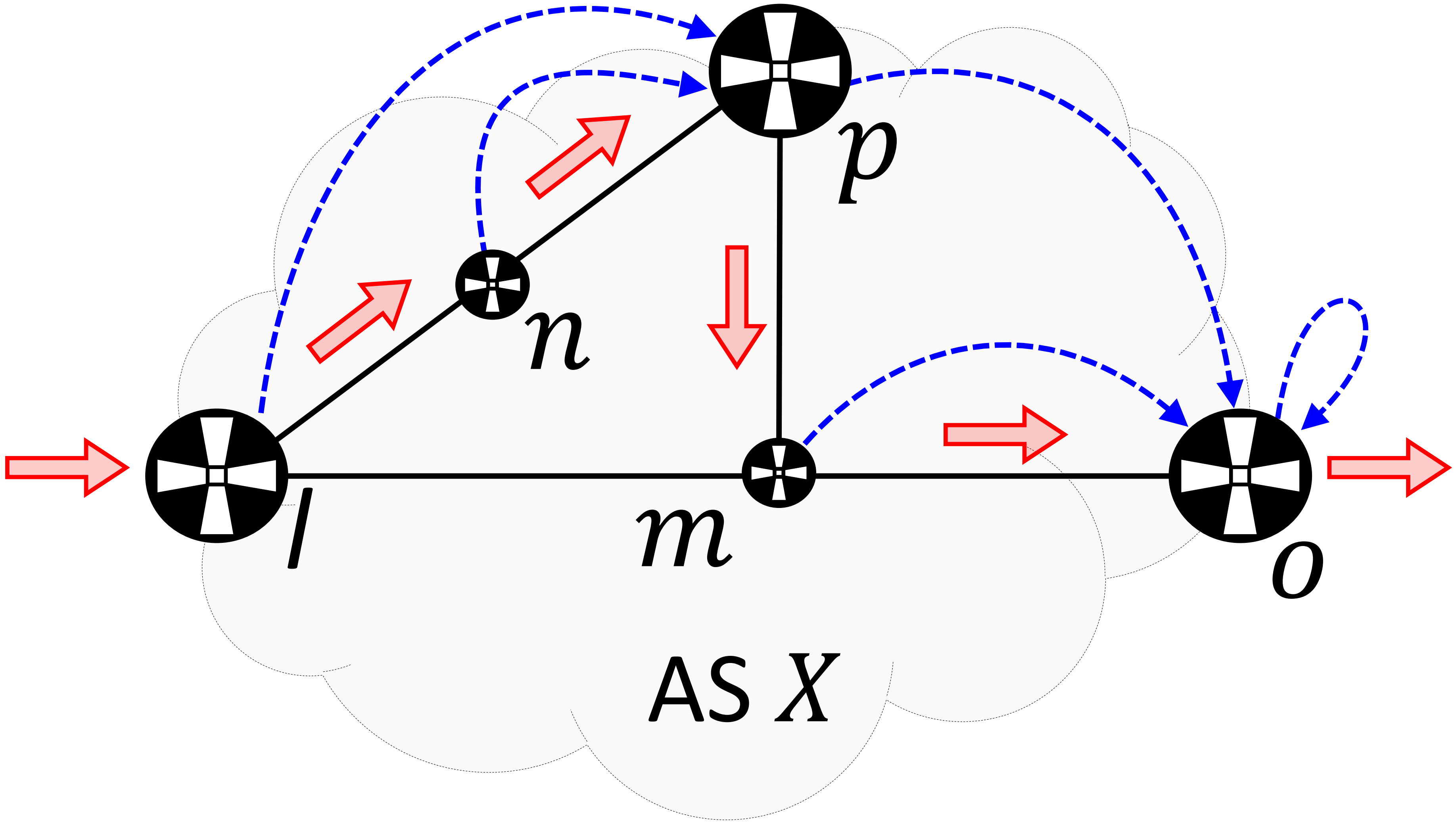
Transit internal route
Direct internal route
$\mathcal{G}_u \circ \mathcal{P}_u(d)$
$\mathcal{N}_u \circ \mathcal{G}_u \circ \mathcal{P}_u(d)$
l
m
n
o
AS X
l
m
n
o
AS X



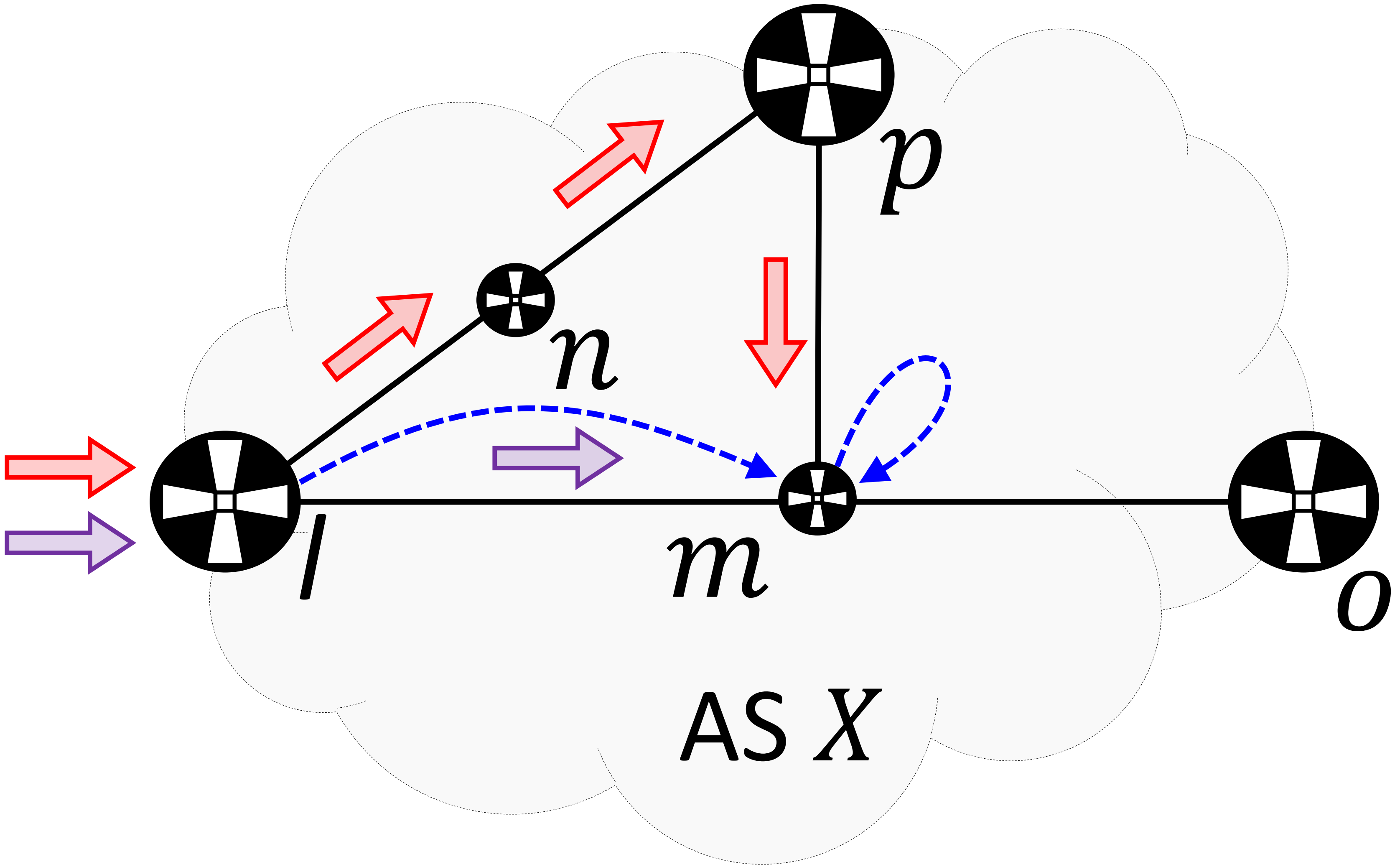
Exploration Phase
Prefix-Grouping Phase
Multi-Route Discovery Phase
Per-dest/flow LB
Prefix-Based Mechanisms
P_1 R_1
P_2 R_4
P_3 R_2
P_4 R_3
P_5 R_3
P_6 R_4
P_7 R_2
$e/32$ R_1
R_1 R_2 R_3 R_4
$\mathcal{P}_1$
$\mathcal{P}_2$
$\mathcal{P}_3$
$\mathcal{P}_4$

Detecting Forwarding Alterations

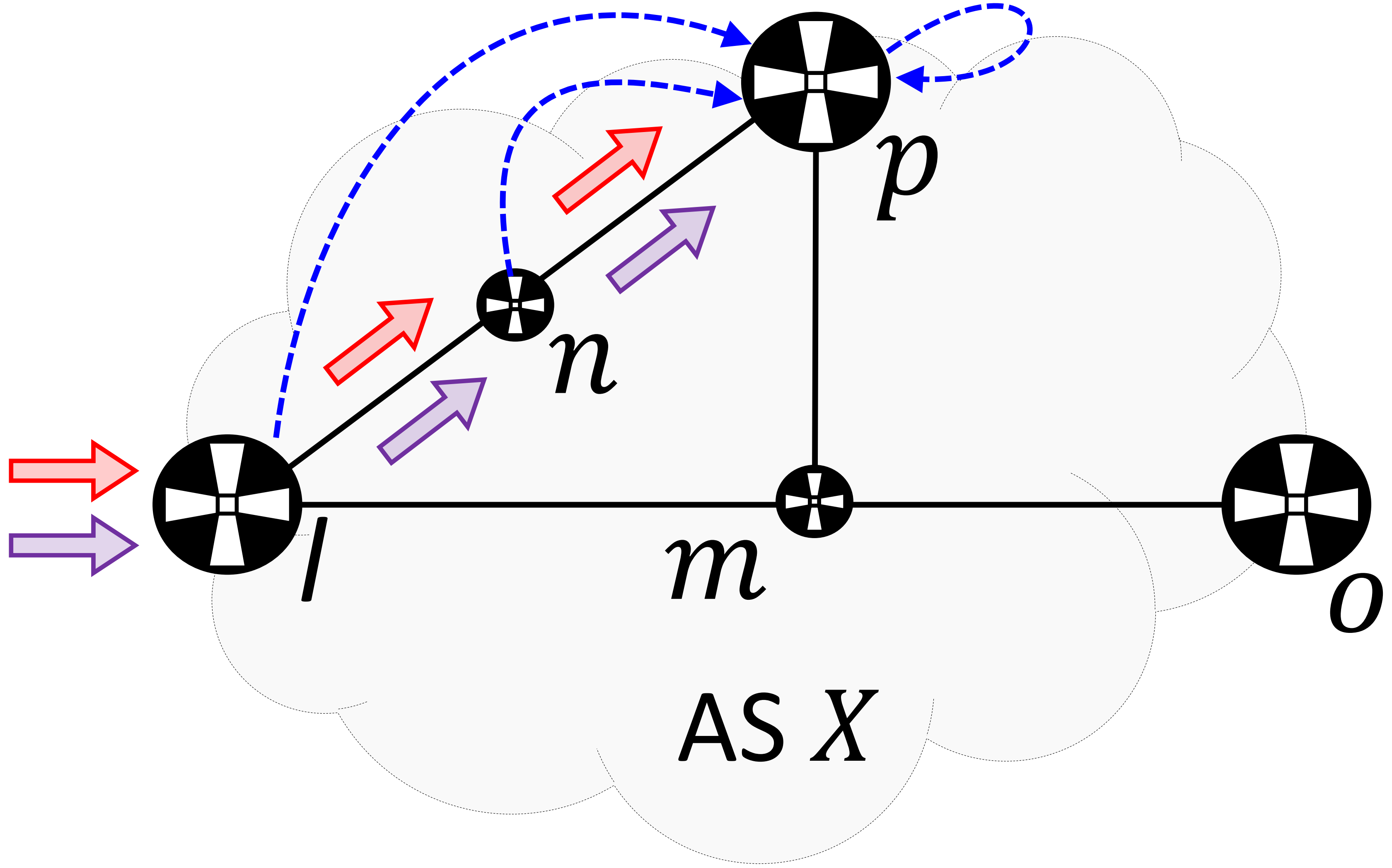
Forwarding Detour



Step I - target m

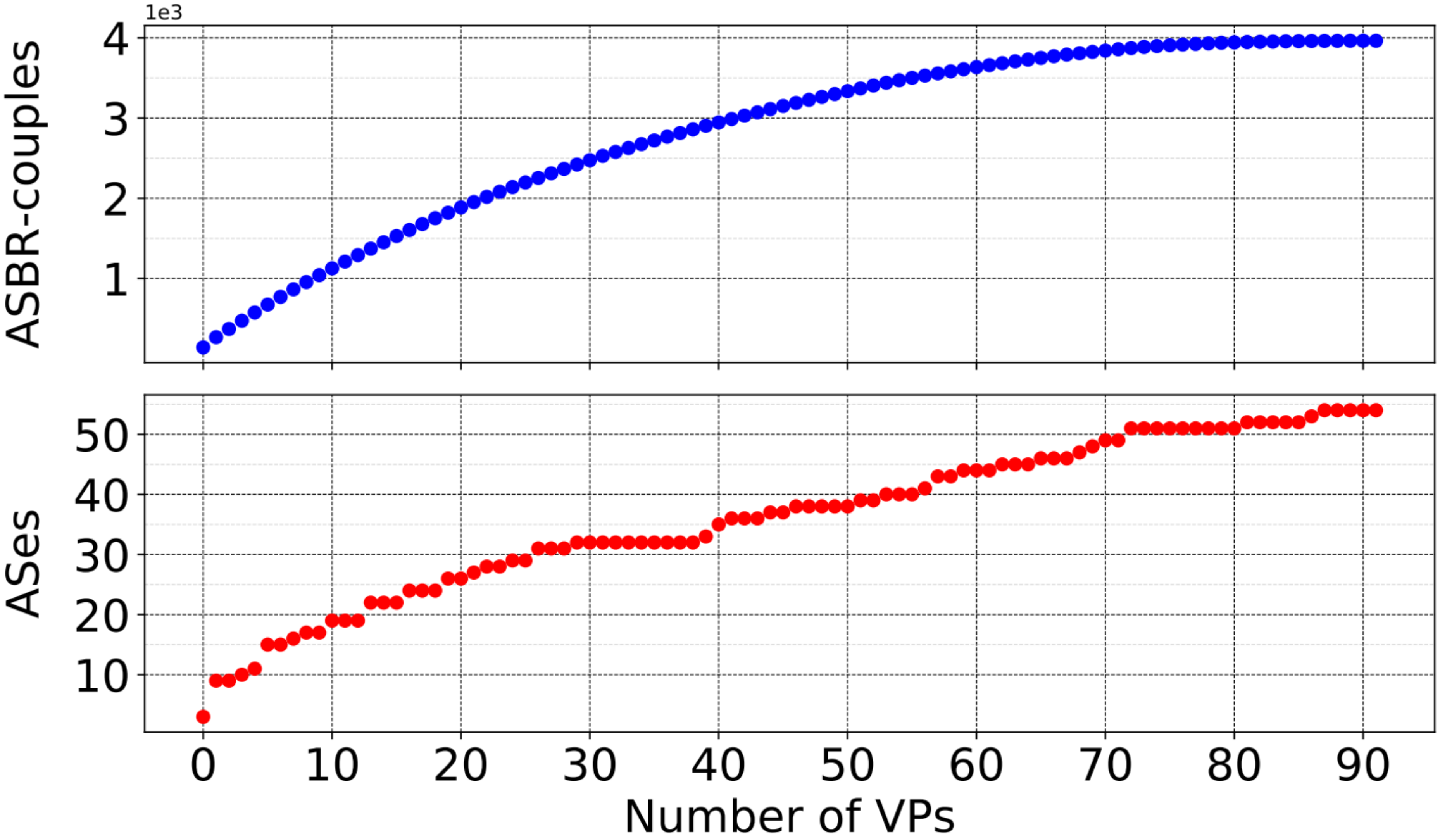


Step II - target p

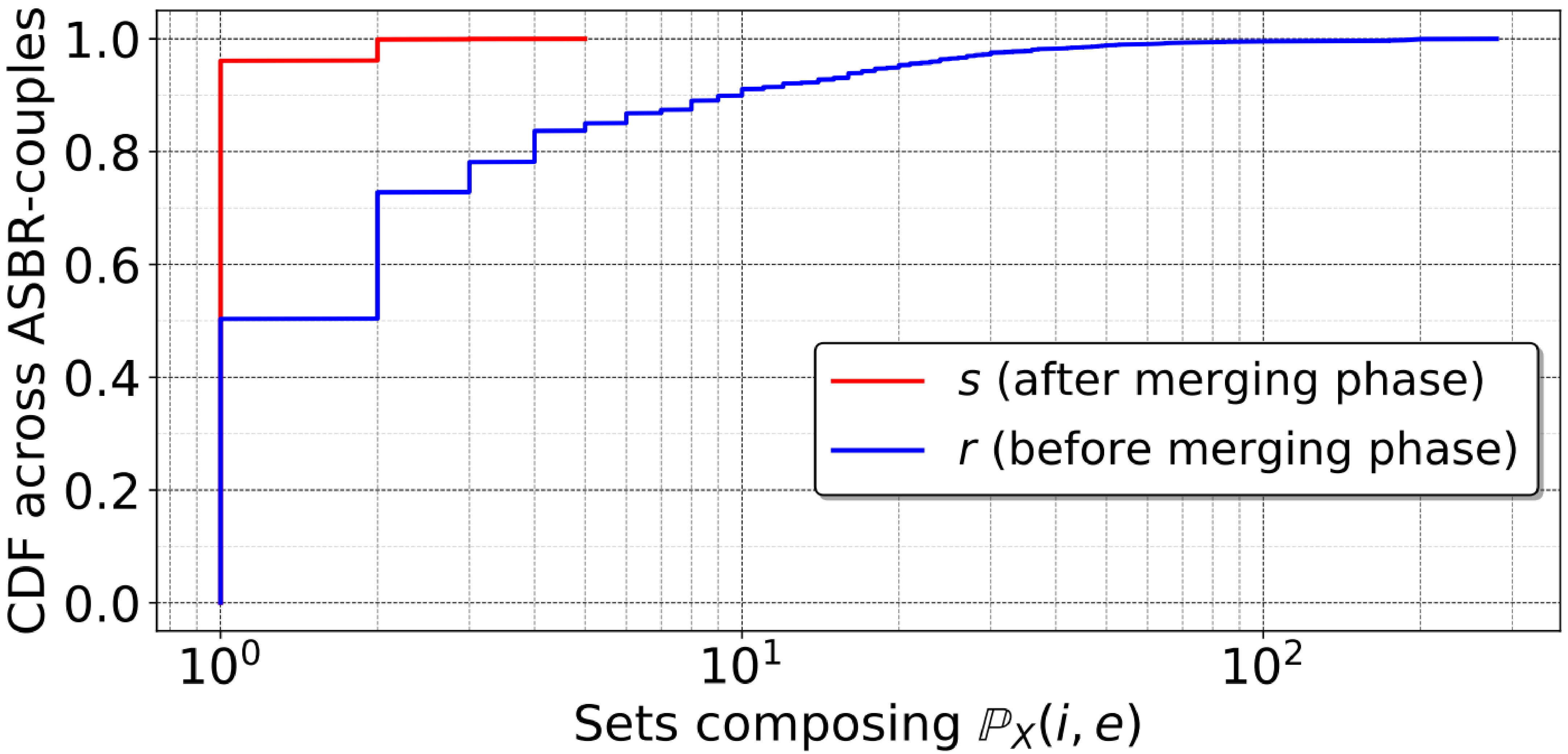


Results detection of FD

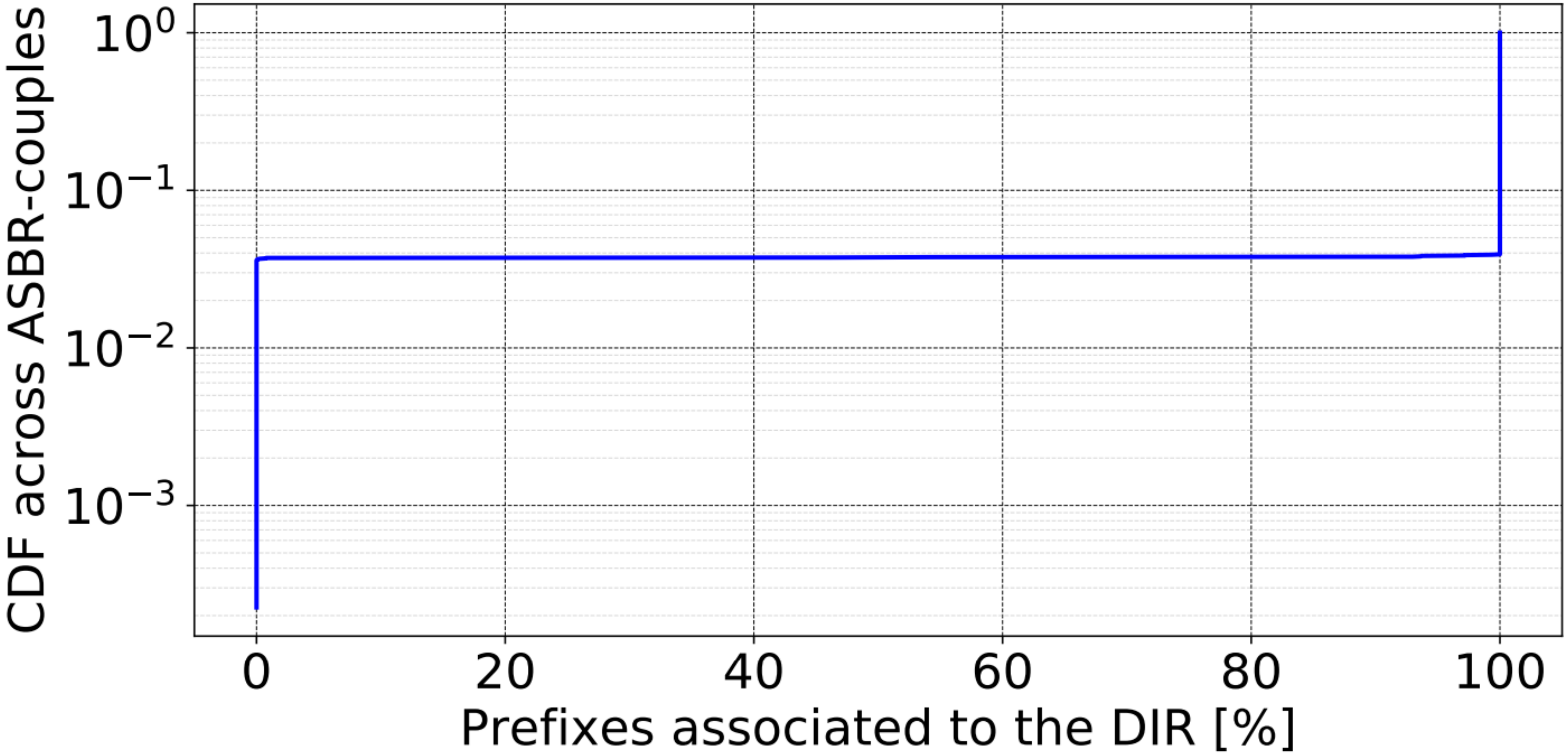
Marginal utility



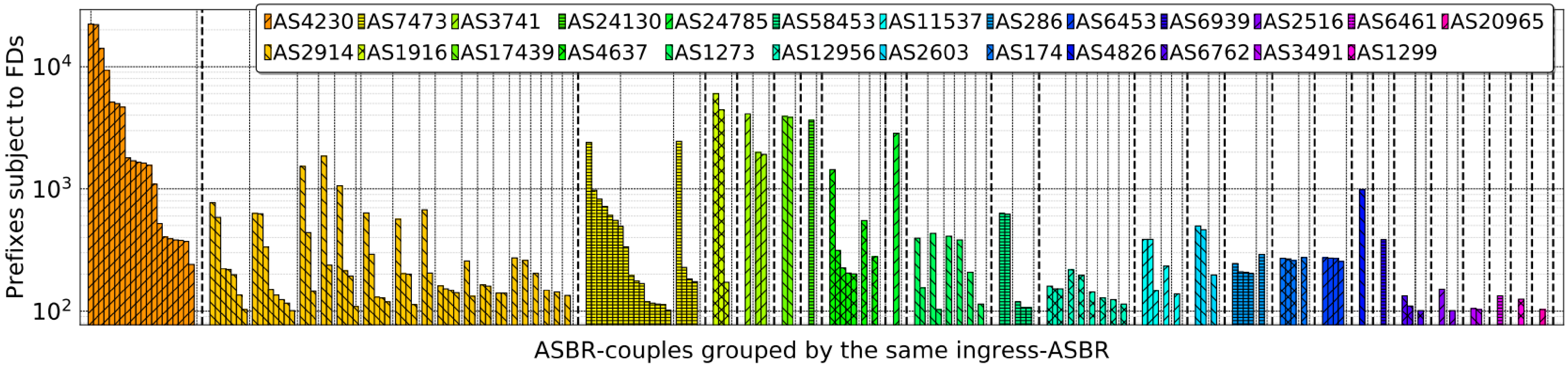
Merging-phase



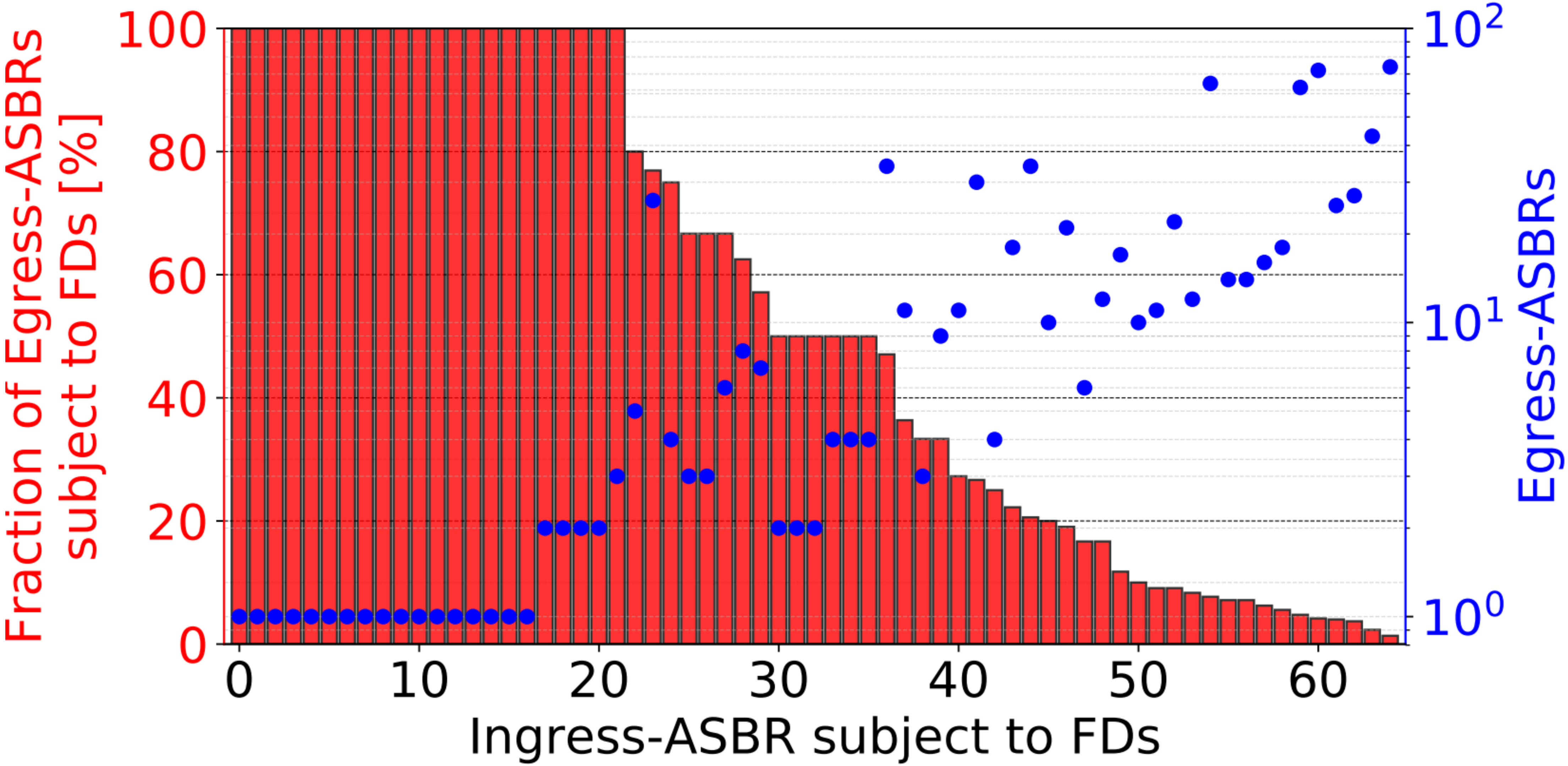
Binary pattern



FDs per AS, ASBR couple and ingress-ASBR



Analysis per ingress-ASBR



BGP lies and FDs

BGP lies and FDs may be correlated