Detecting Thid-party Addresses in Traceroute Traces with IP Timestamp Option

P. Marchetta, W. de Donato, A. Pescapé

Dipartimento di Ingegneria Elettrica e delle Tecnologie dell’Informazione
University of Naples “Federico II”, Italy

Motivations

- An accurate knowledge of the Internet topology is essential for
  - network emulation and simulation
  - network management (e.g fault isolation, anomaly detection, etc.)
  - service and resource allocation
  - modeling the Internet cartography

- BGP derived AS-level topologies are incomplete
  - using Traceroute may help to overcome such incompleteness
  - Traceroute is known to be potentially inaccurate!

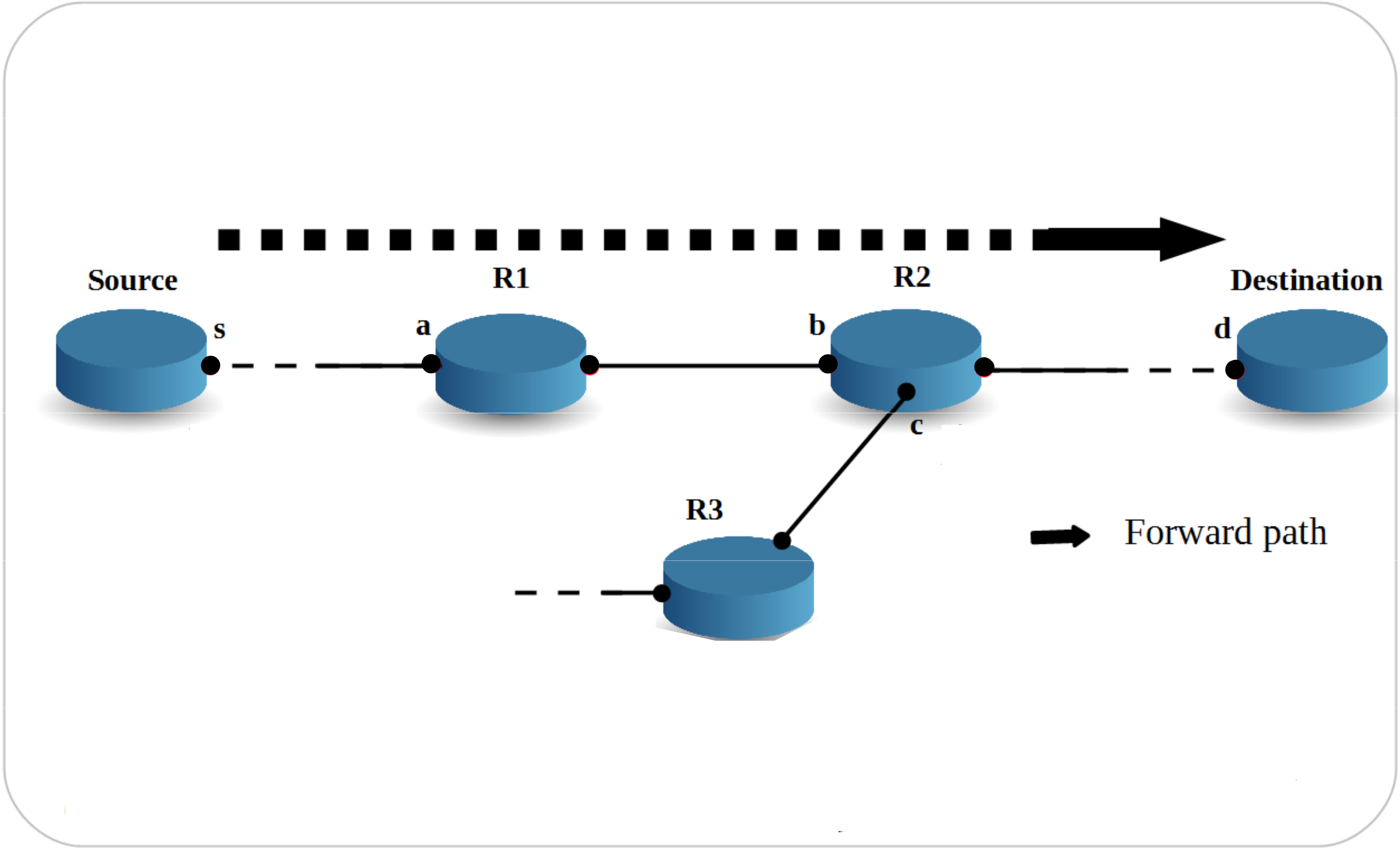
Motivations

- Potential sources of inaccuracy in Traceroute-derived AS-level topologies
  - anonymous routers
  - unmapped Traceroute hops
  - sibling ASes
  - multiple origin ASes prefixes
  - divergence between control and data paths
  - Internet Exchange Points
  - Third-party addresses
  - …

Motivations

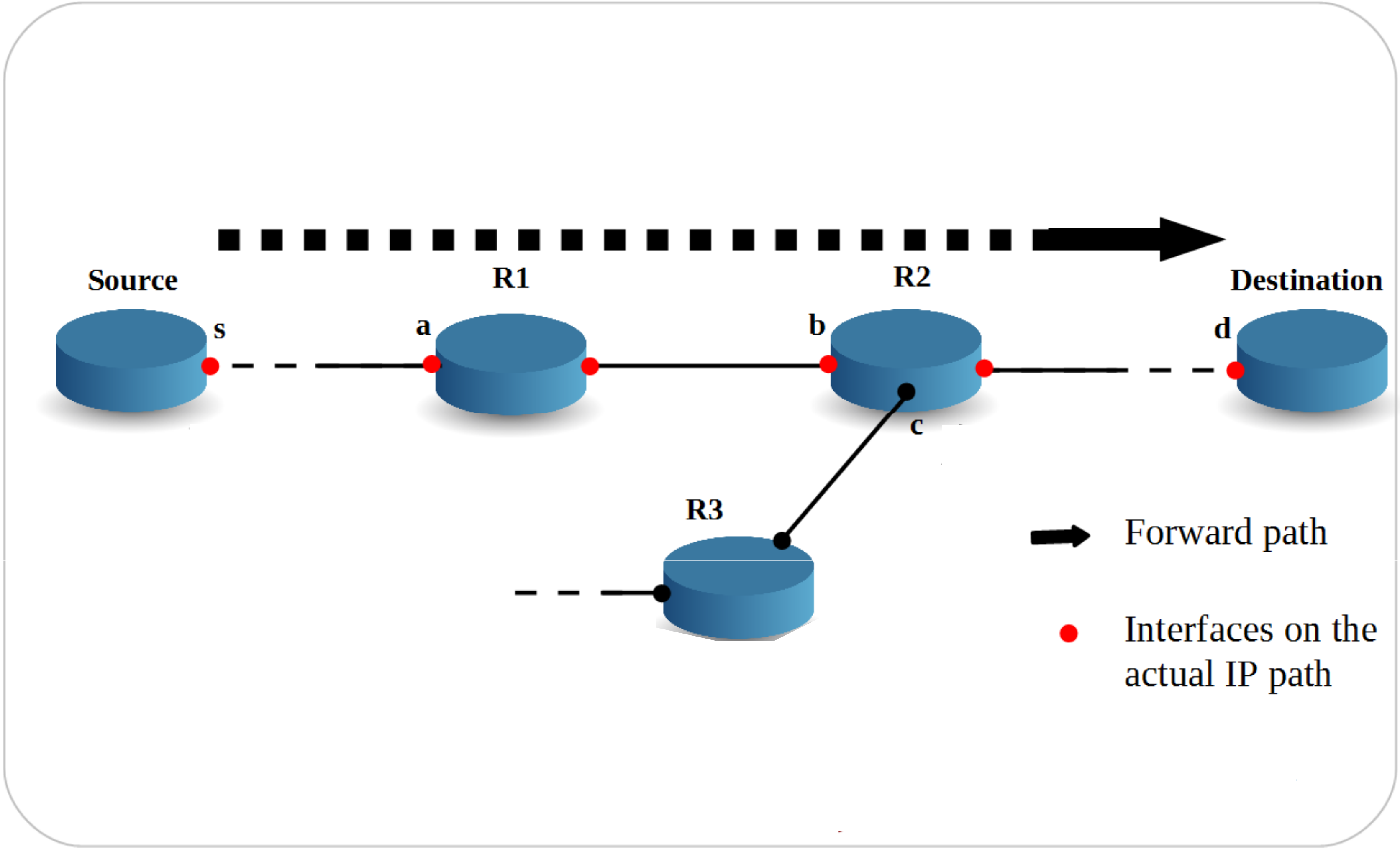
- Potential sources of inaccuracy in Traceroute-derived AS-level topologies
  - anonymous routers
  - unmapped Traceroute hops
  - sibling ASes
  - multiple origin ASes prefixes
  - divergence between control and data paths
  - Internet Exchange Points
  - Third-party addresses
  - …

Third-party Addresses



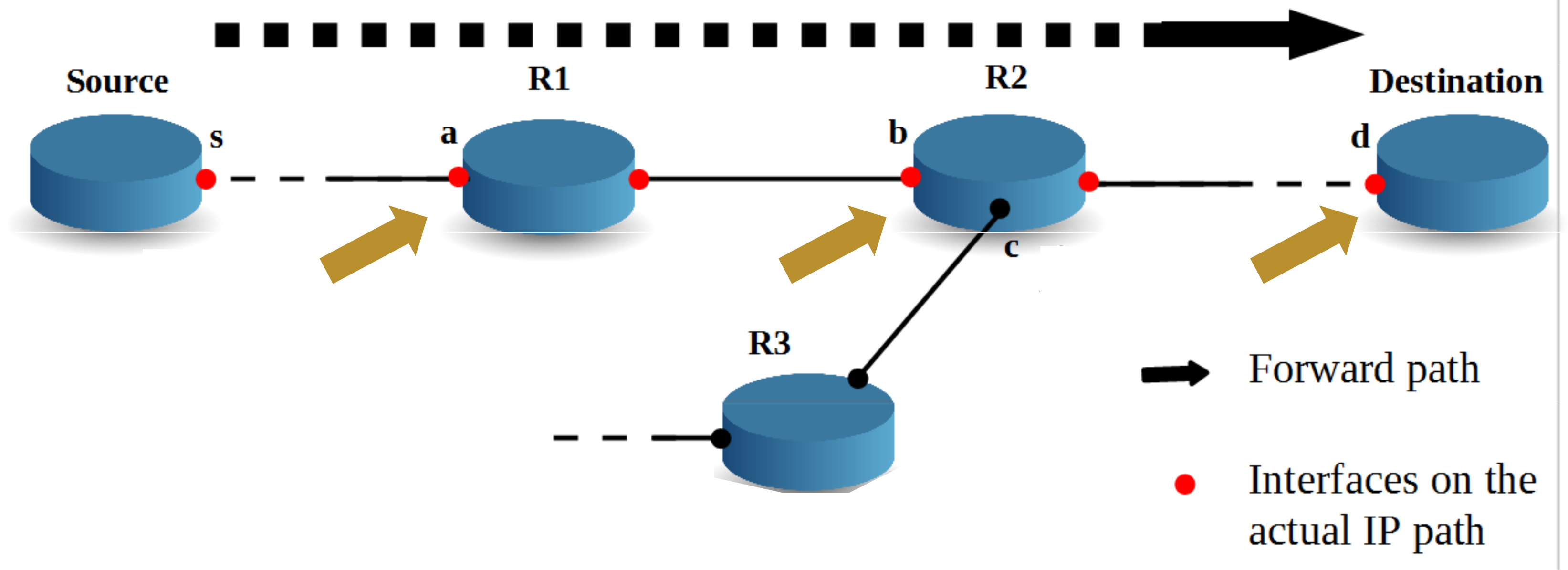
A Third-party (TP) address is an IP address discovered by Traceroute which does not belong to the actual IP path toward the destination

Third-party Addresses



A Third-party (TP) address is an IP address discovered by Traceroute which does not belong to the actual IP path toward the destination

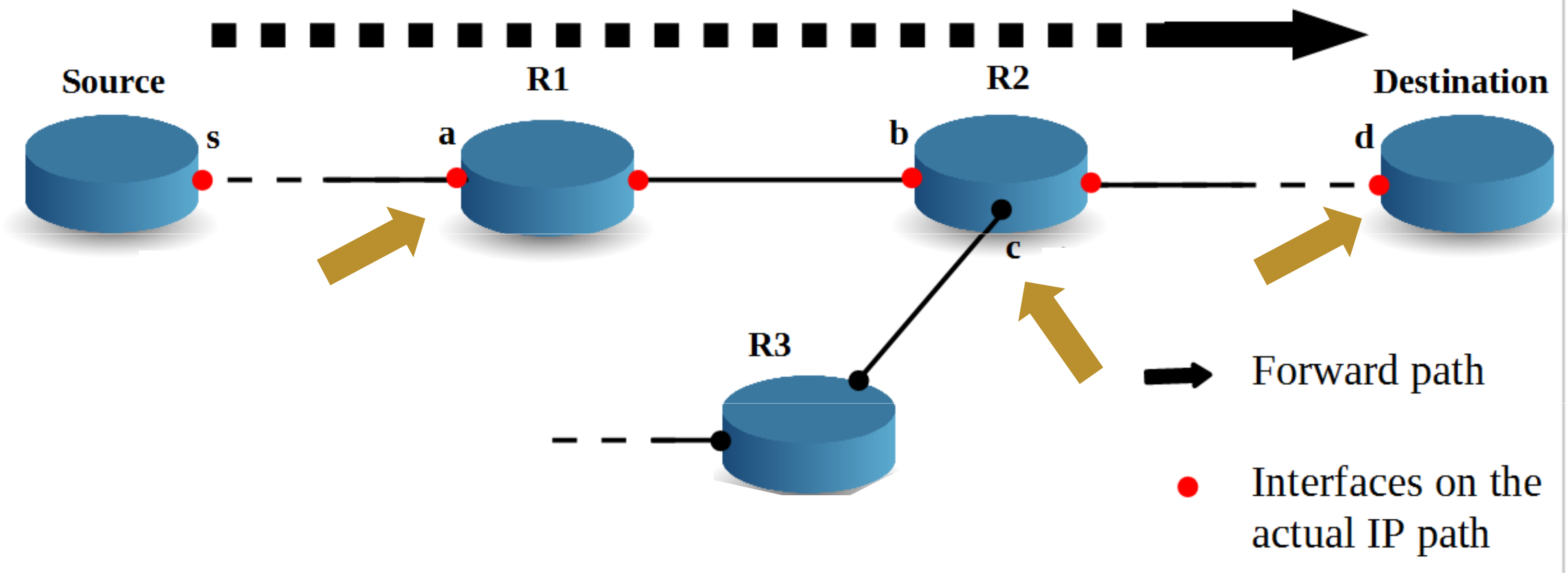
Third-party Addresses



Expected Traceroute: a – b – d

A Third-party (TP) address is an IP address discovered by Traceroute which does not belong to the actual IP path toward the destination

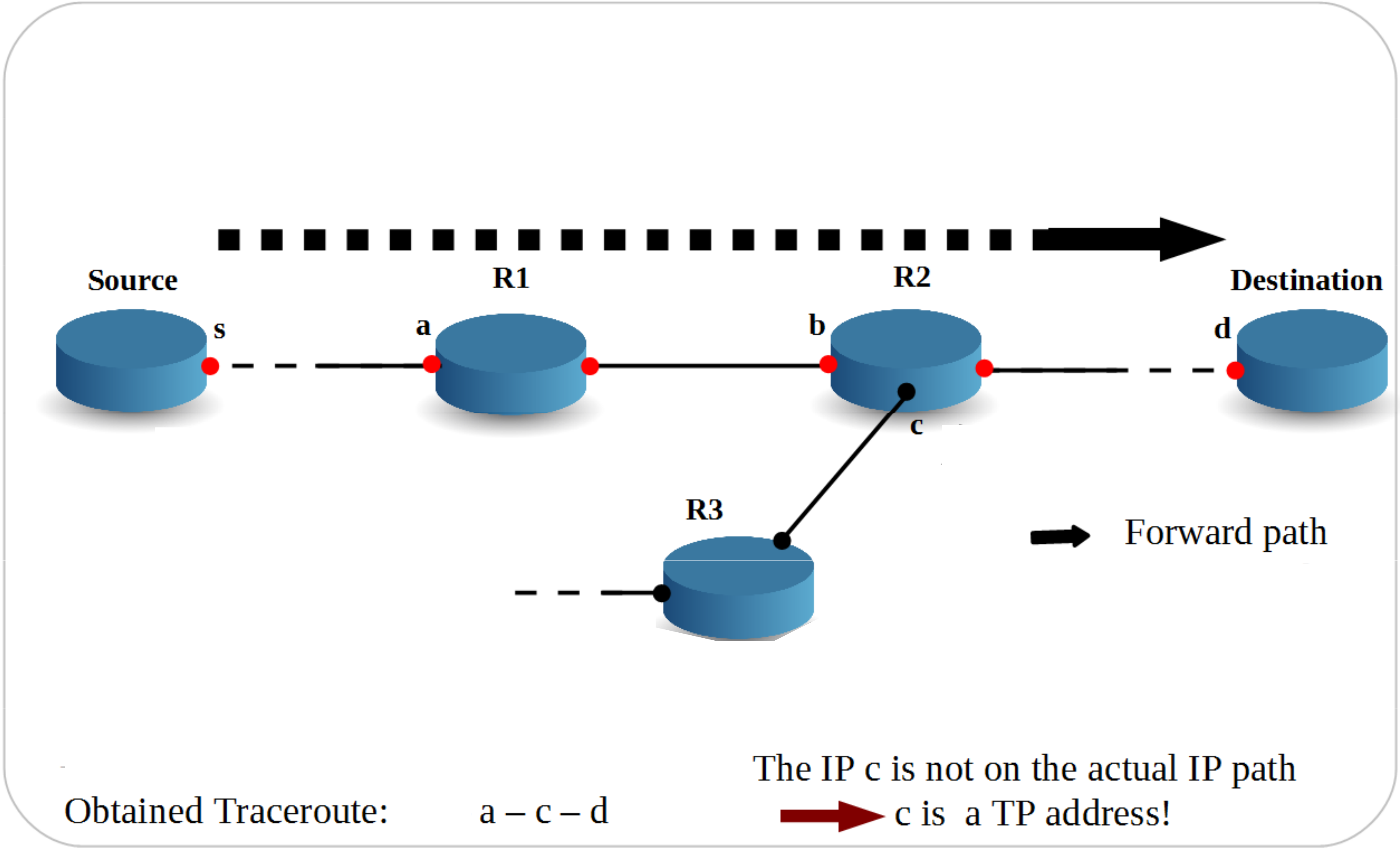
Third-party Addresses



Obtained Traceroute: a – c – d

A Third-party (TP) address is an IP address discovered by Traceroute which does not belong to the actual IP path toward the destination

Third-party Addresses



Obtained Traceroute: a – c – d

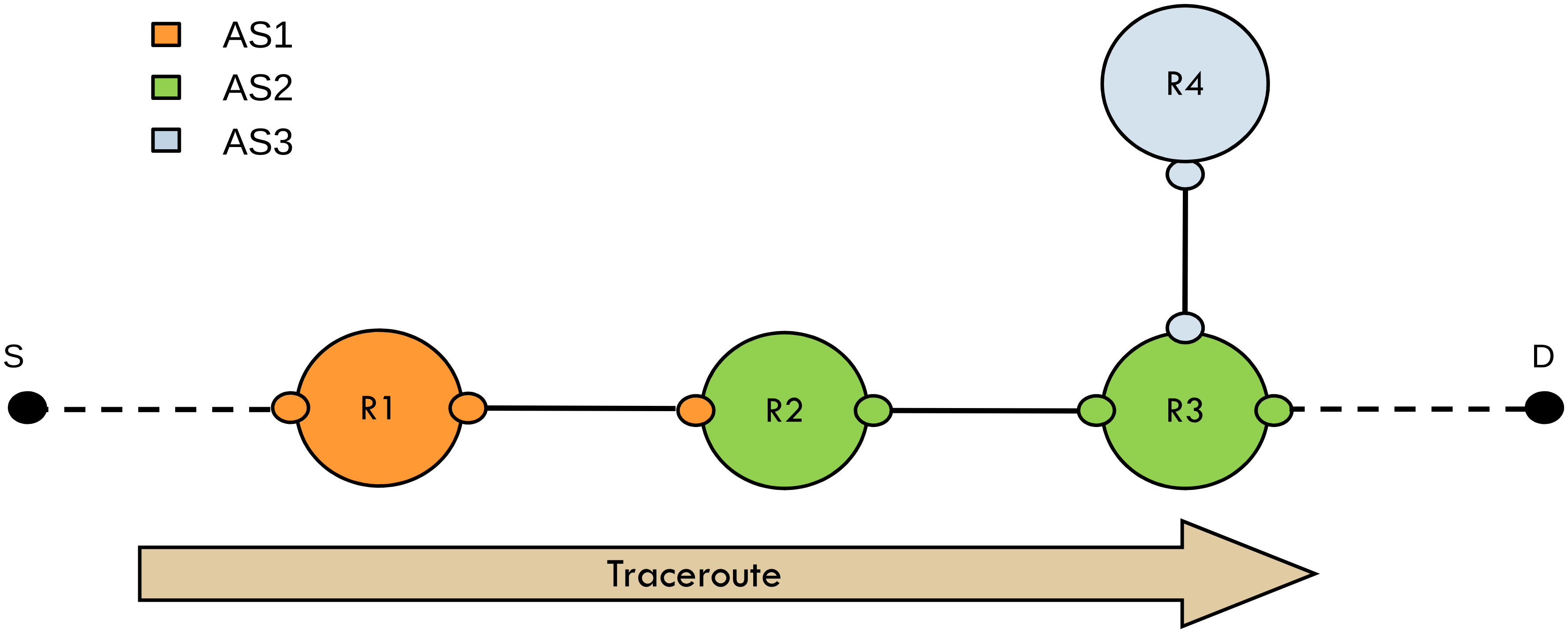
The IP c is not on the actual IP path

c is a TP address!

A Third-party (TP) address is an IP address discovered by Traceroute which does not belong to the actual IP path toward the destination

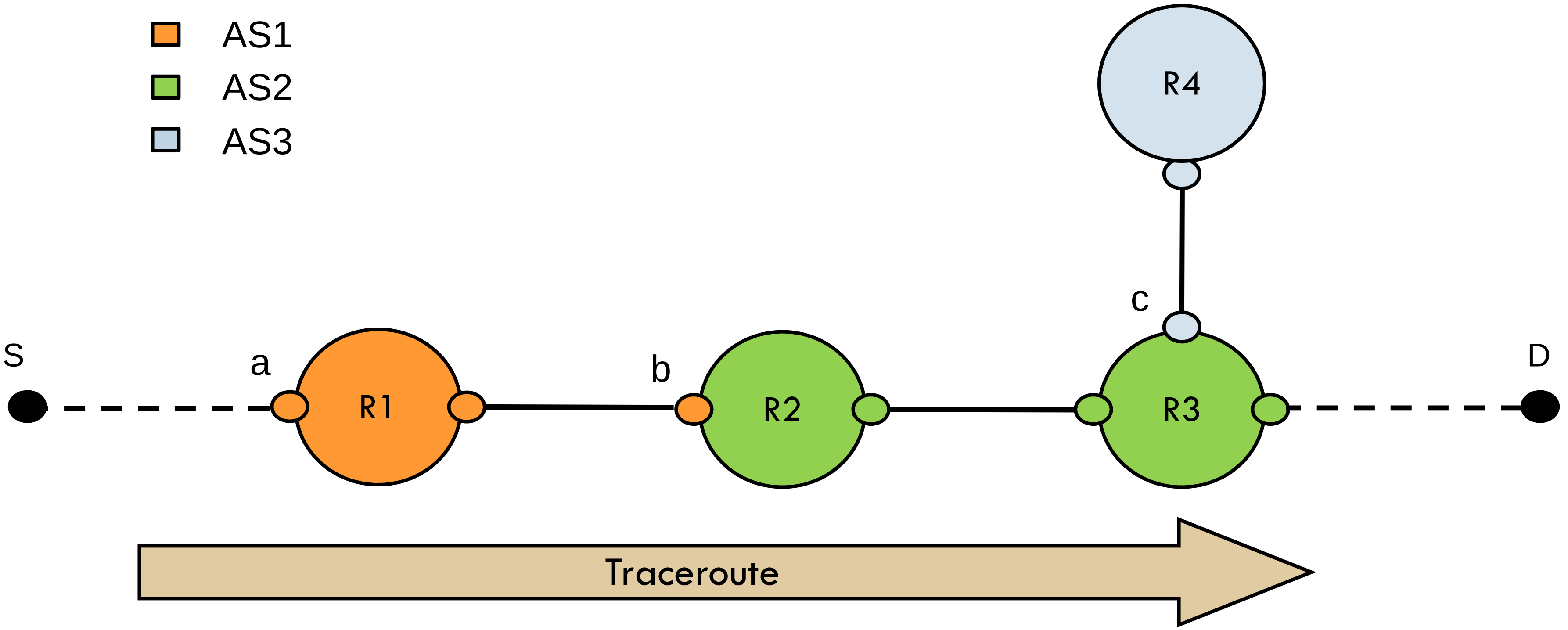
Third-party Addresses impact

TP addresses may cause the inference of false AS-level links



Third-party Addresses impact

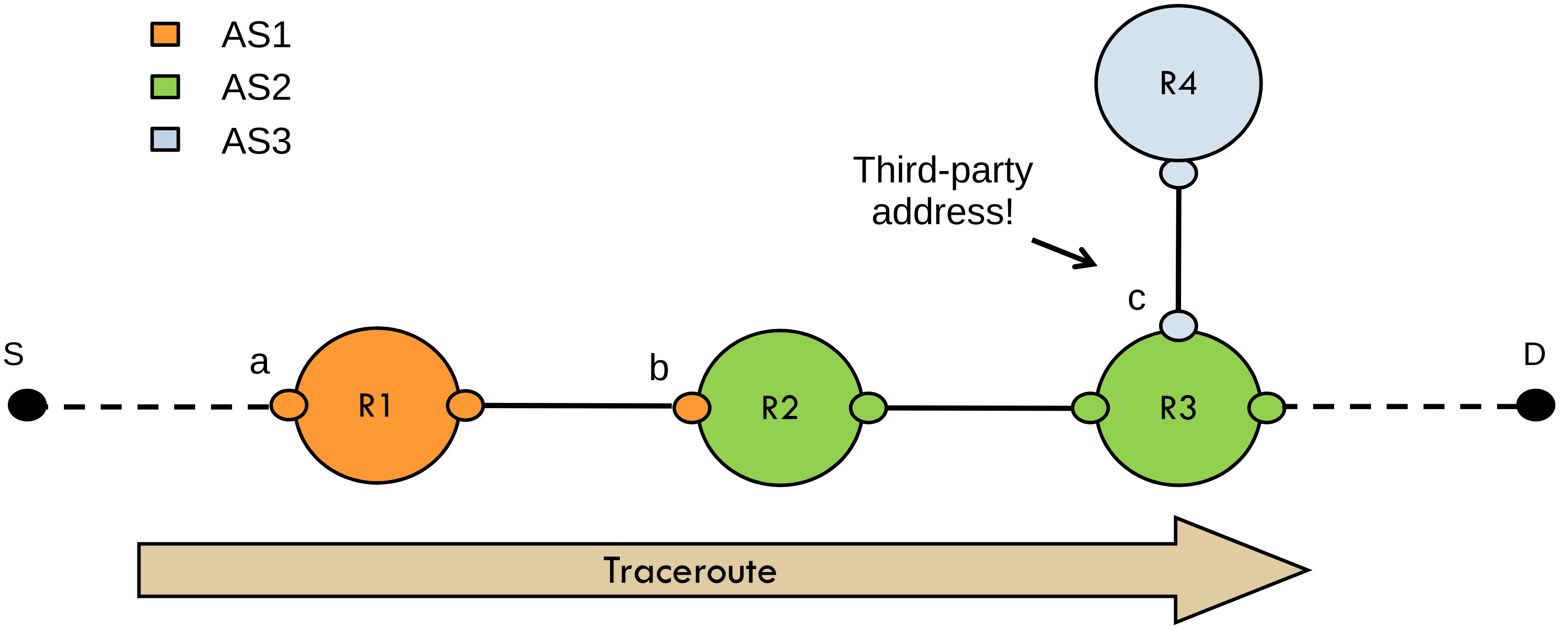
TP addresses may cause the inference of false AS-level links



IP level trace: a - b - c

Third-party Addresses impact

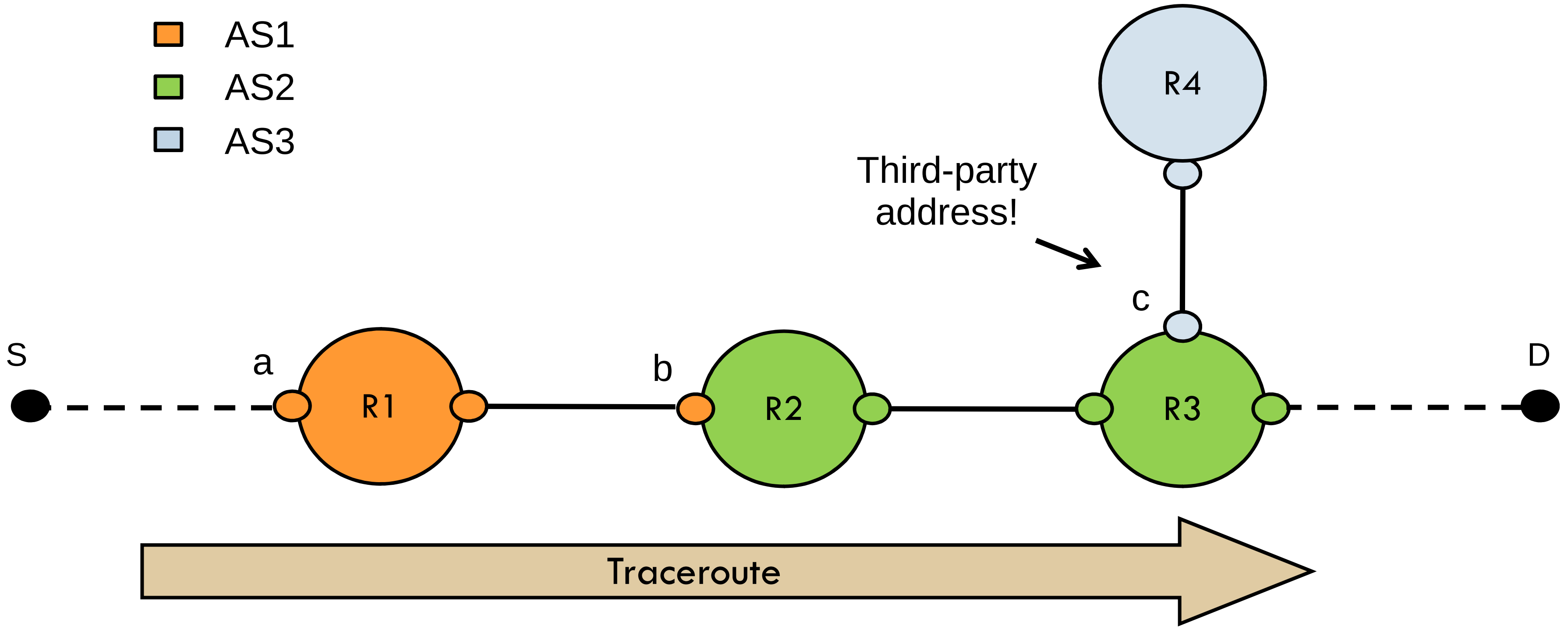
TP addresses may cause the inference of false AS-level links



IP level trace: a - b - c

Third-party Addresses impact

TP addresses may cause the inference of false AS–level links

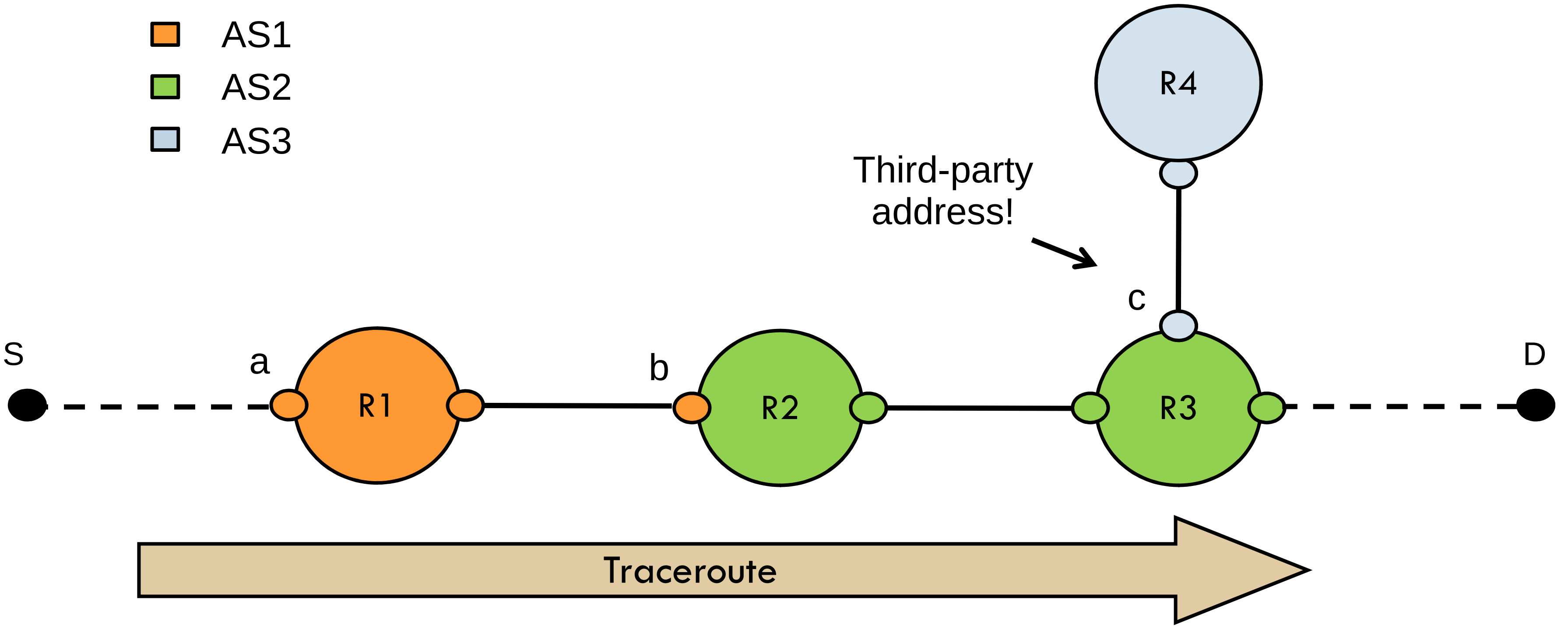


IP level trace: a - b - c

AS mapping: AS1 - AS1 - AS3

Third-party Addresses impact

TP addresses may cause the inference of false AS-level links



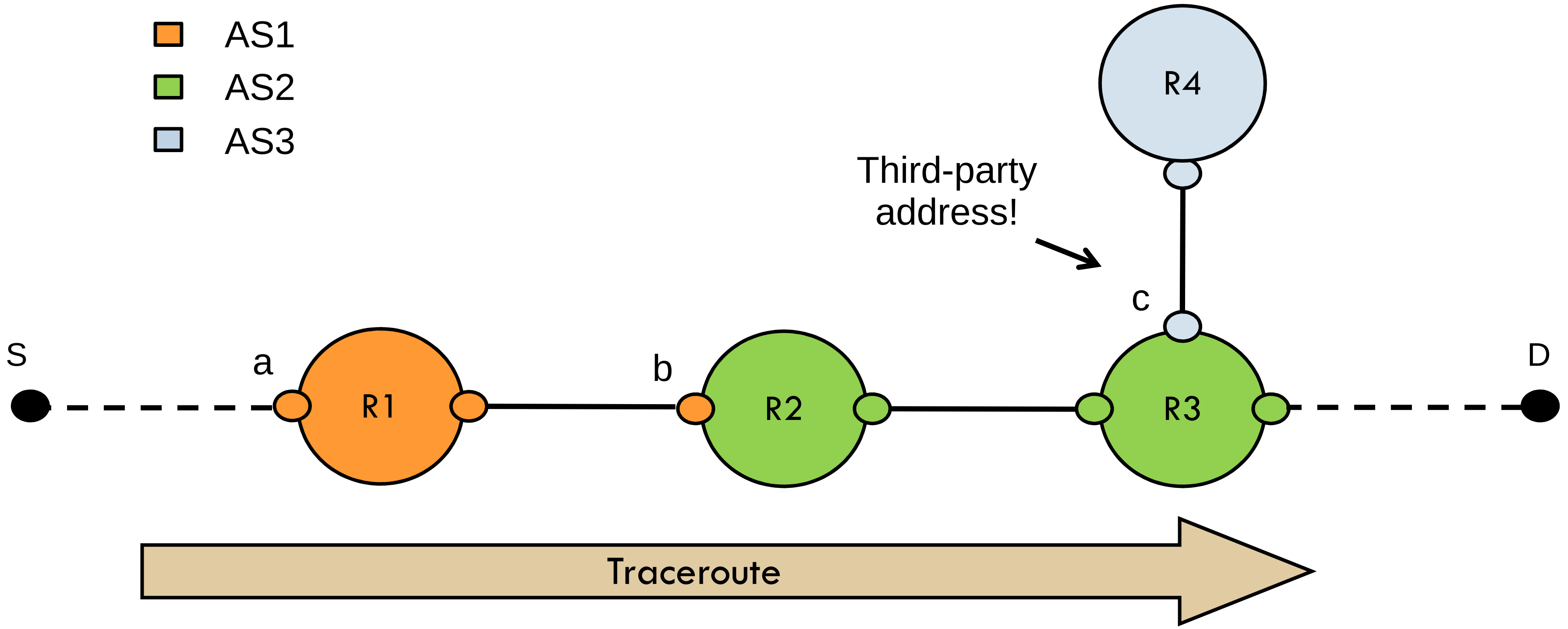
IP level trace: a - b - c

AS mapping: AS1 - AS1 - AS3

False AS-level link AS1 – AS3

Third-party Addresses impact

TP addresses may cause the inference of false AS-level links



IP level trace: a - b - c

AS mapping: AS1 - AS1 - AS3

False AS-level link AS1 – AS3

AS2 is totally invisible

Literature

Is an IP address discovered by Traceroute a Third-party address or is it part of the actual traversed path?

Hyun *et al*. (PAM' 03)

- Assessing candidate TP addresses with heuristic methods based on IP to AS mapping
- TP addresses cannot be considered as a significant source of AS map distortion

Zhang *et al.* (JSAC' 11)

- Pre-collected information about the topology
- TP addresses represent a *huge obstruction towards the accuracy of Traceroute measurements* and the last and most difficult cause of inaccuracy to be inferred

Identifying Third-Party addresses

- Exploiting the IP Pre-specified Timestamp (TS) Option (RFC791)
  - allows to pre-specify in a single packet probe up to four IP addresses from which a timestamp is requested

- Common router behaviors in Internet [1]
  - Routers not managing the TS option
  - Any-interface stamping routers
    - insert *all* the requested Timestamps when the pre-specified IPs are associated to *any* owned interface
  - Per-interface stamping routers
    - insert a *single* Timestamp every time the packet passes through the interface associated to the pre-specified IP address

[1] W. de Donato, P. Marchetta, and A. Pescapé. "A Hands-on Look at Active Probing using the IP Prespecified Timestamp option". In PAM'12, Vienna, Austria, 2012.

Identifying Third-Party addresses

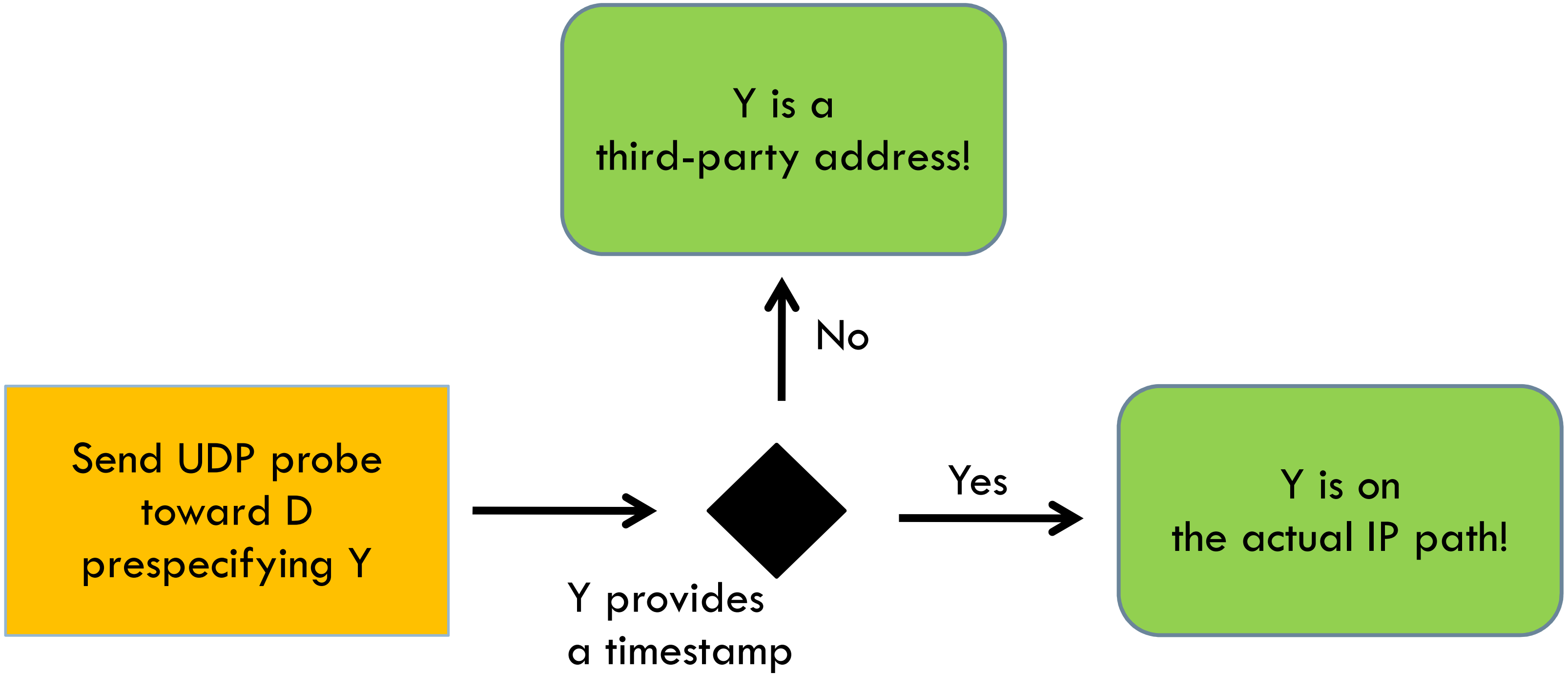
- Exploiting the IP Pre-specified Timestamp (TS) Option (RFC791)
  - allows to pre-specify in a single packet probe up to four IP addresses from which a timestamp is requested

- Common router behaviors in Internet [1]
  - Routers not managing the TS option
  - Any-interface stamping routers
    - insert *all* the requested Timestamps when the pre-specified IPs are associated to *any* owned interface
  - Per-interface stamping routers
    - insert a *single* Timestamp every time the packet passes through the interface associated to the pre-specified IP address

[1] W. de Donato, P. Marchetta, and A. Pescapé. "A Hands-on Look at Active Probing using the IP Prespecified Timestamp option". In PAM'12, Vienna, Austria, 2012.

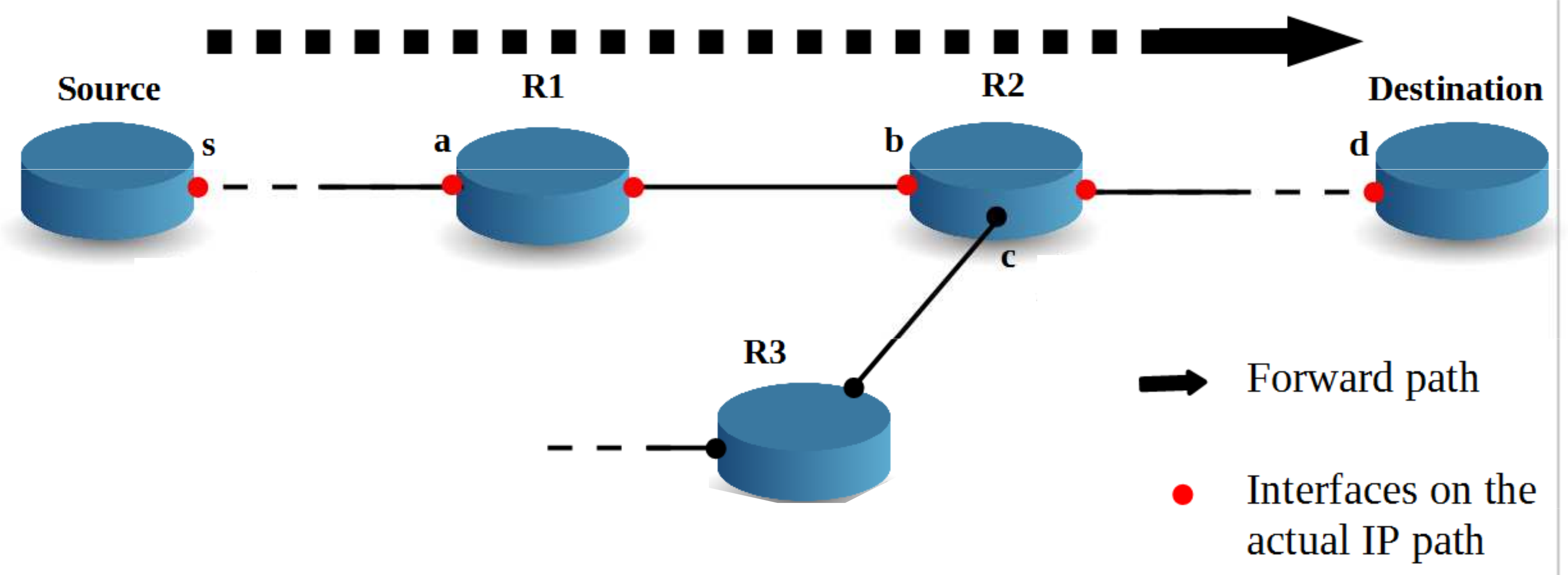
Identifying Third-Party addresses

- Let Y be
  - an IP address discovered by Traceroute toward a destination D
  - owned by a per-interface stamping router
- Is Y a Third-party address or not?



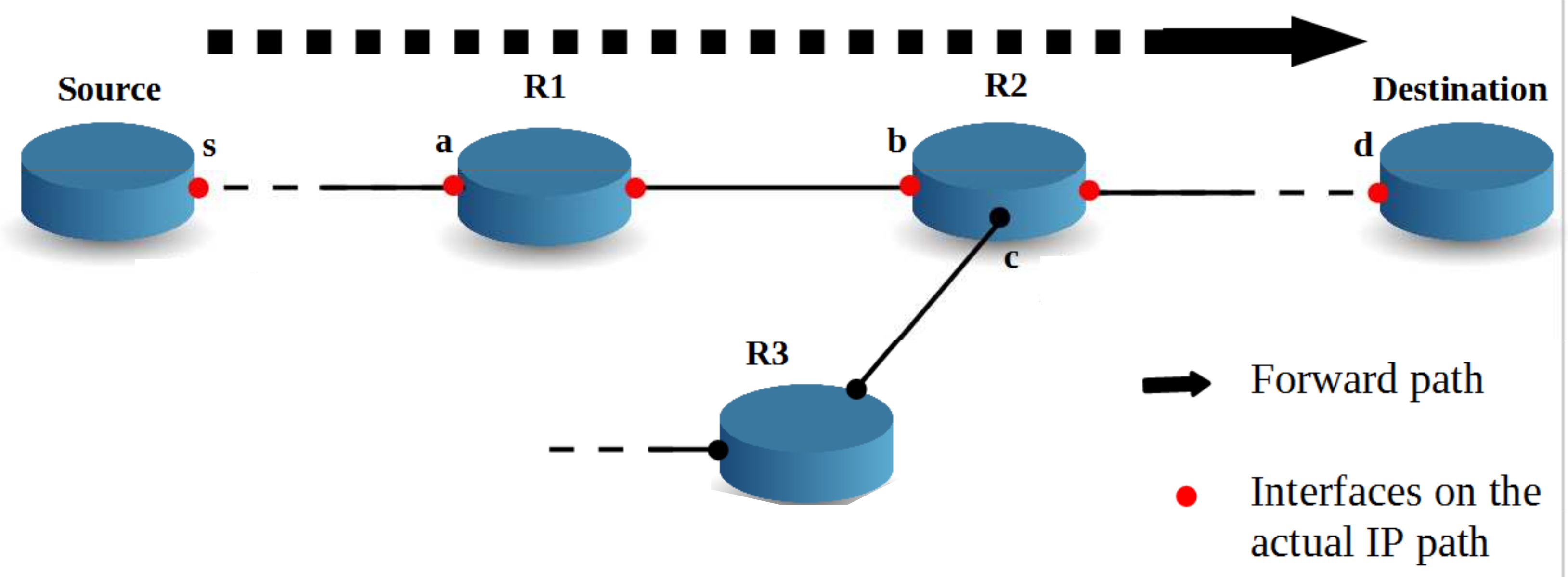
Identifying Third-Party addresses

IP		
TSopt	b	
UDP		



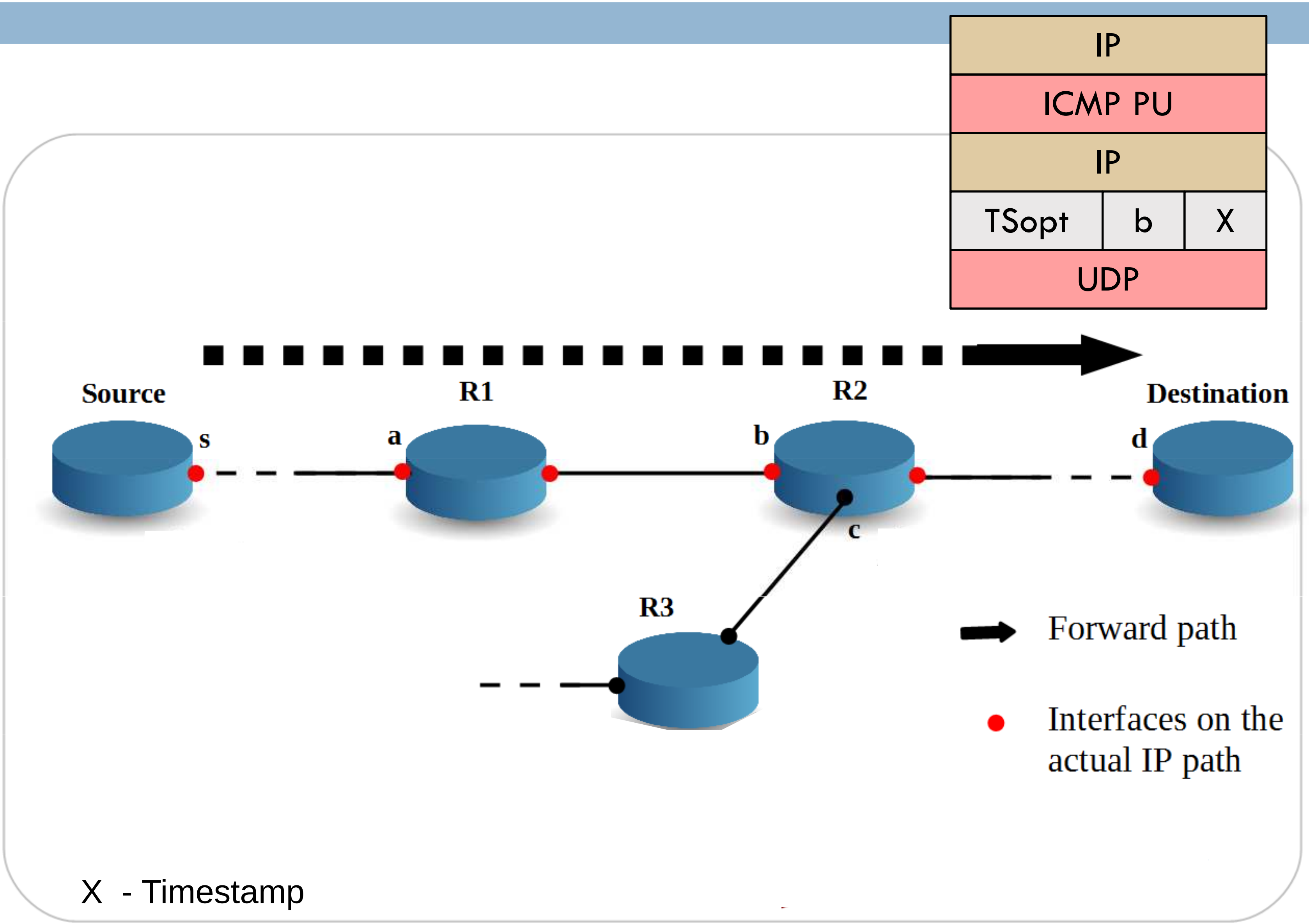
Identifying Third-Party addresses

IP		
TSopt	b	X
UDP		



X - Timestamp

Identifying Third-Party addresses

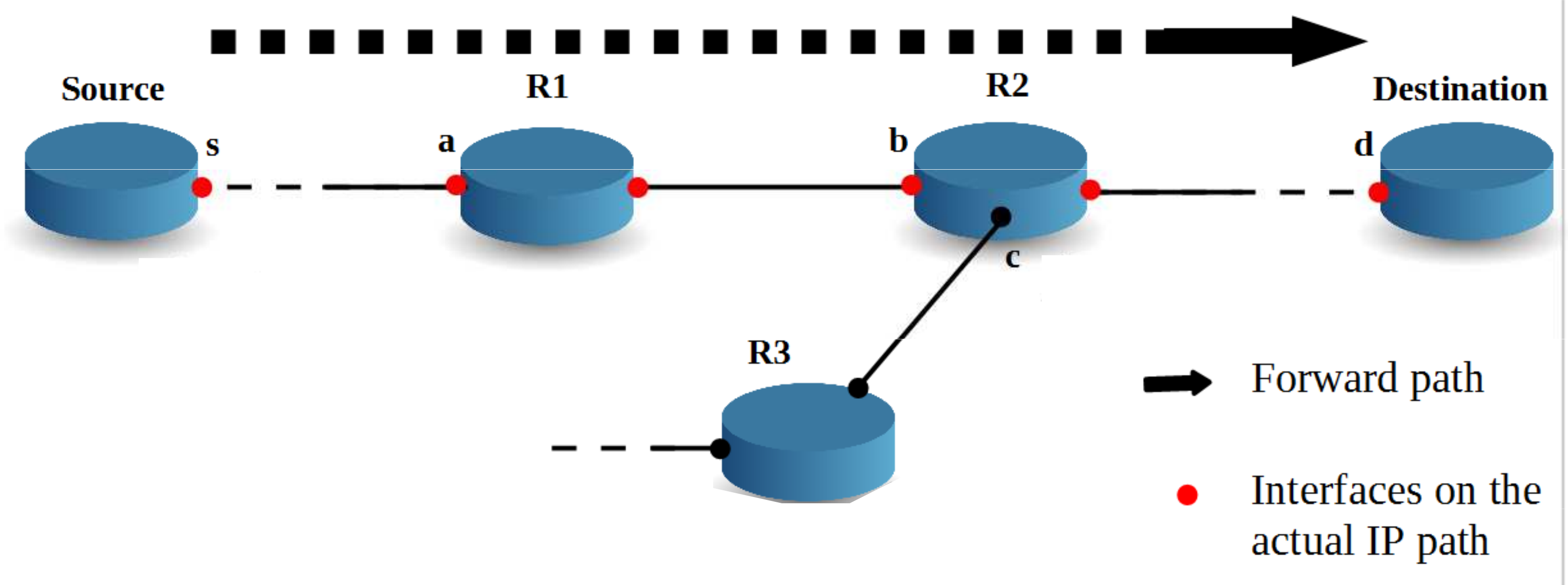


IP
ICMP PU
IP
TSopt | b | X
UDP

X - Timestamp

Identifying Third-Party addresses

IP		
ICMP PU		
IP		
TSopt	b	X
UDP		

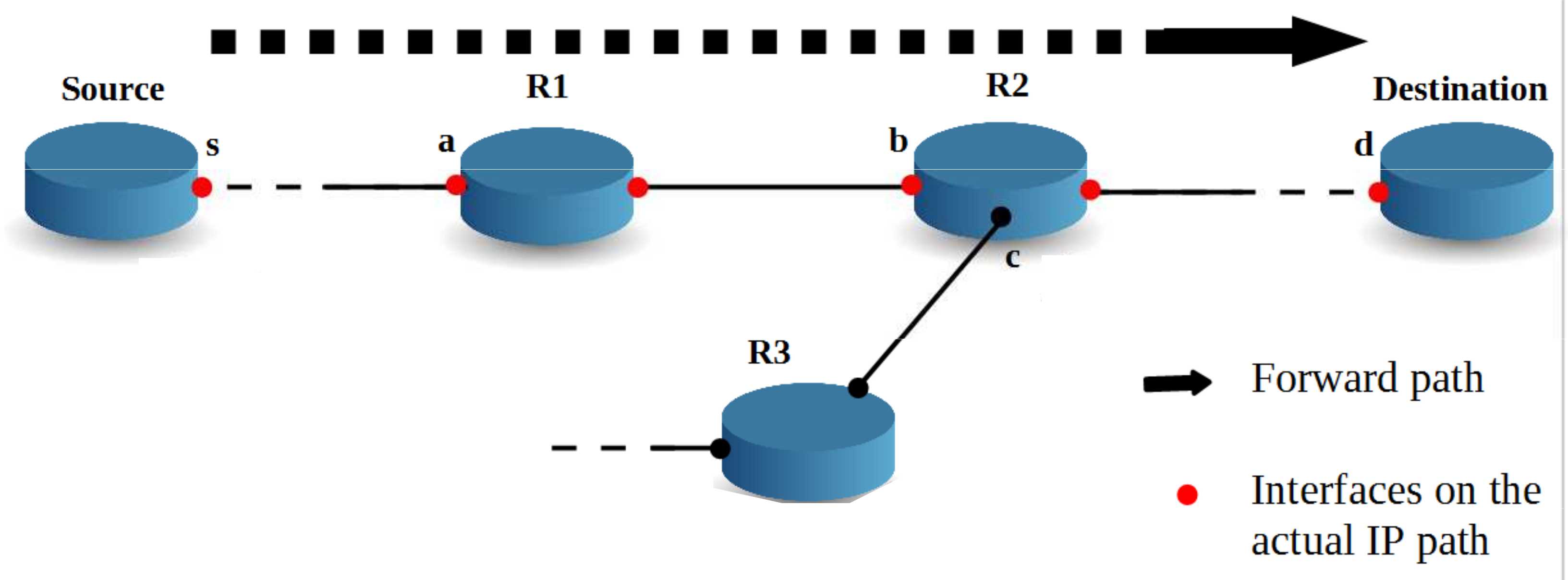


X - Timestamp

Identifying Third-Party addresses

<table>
<tr><td colspan="3">IP</td></tr>
<tr><td colspan="3">ICMP PU</td></tr>
<tr><td colspan="3">IP</td></tr>
<tr><td>TSopt</td><td>b</td><td>X</td></tr>
<tr><td colspan="3">UDP</td></tr>
</table>

b is on the IP path!

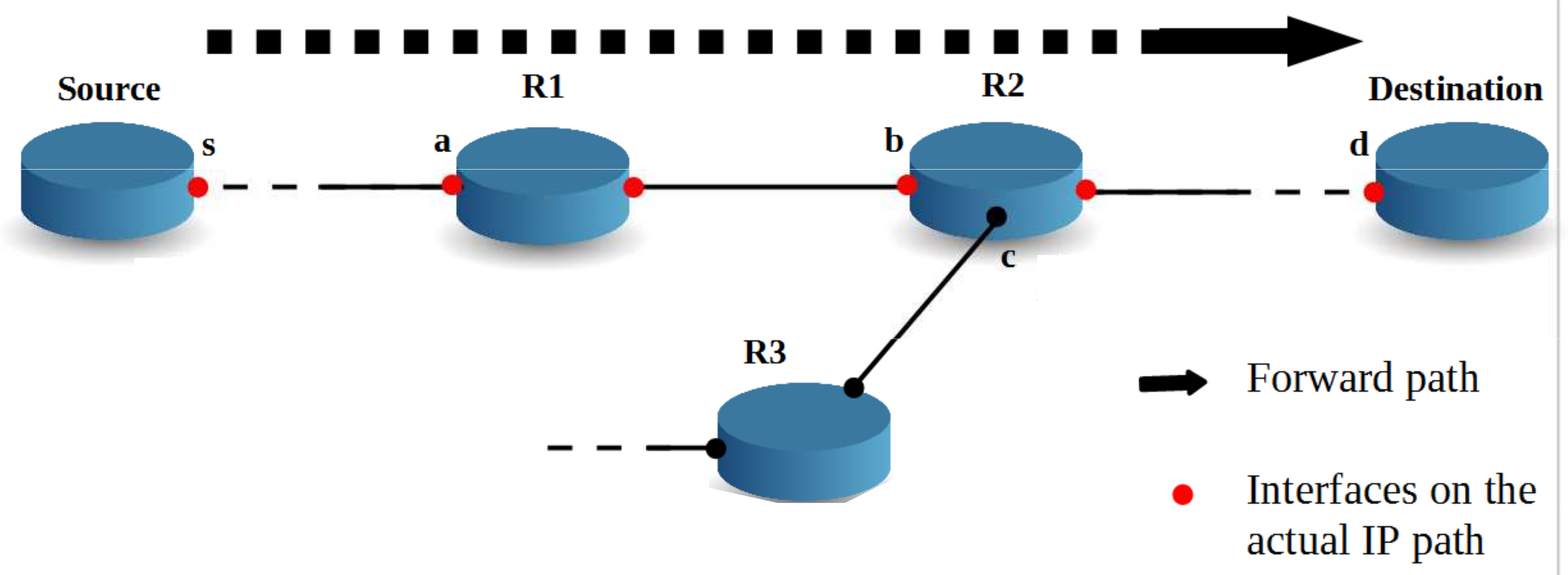


X - Timestamp

UDP packet probes allow to avoid ambiguities caused by the reverse path

Identifying Third-Party addresses

IP		
TSopt	c	
UDP		

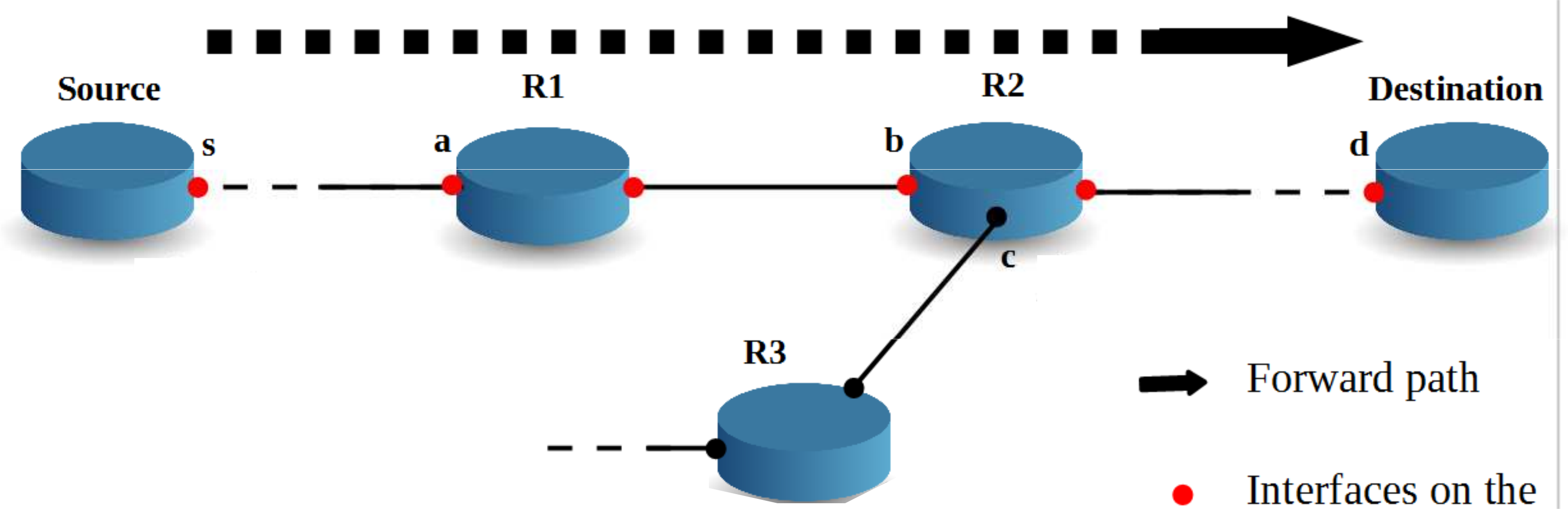


X - Timestamp

UDP packet probes allow to avoid ambiguities caused by the reverse path

Identifying Third-Party addresses

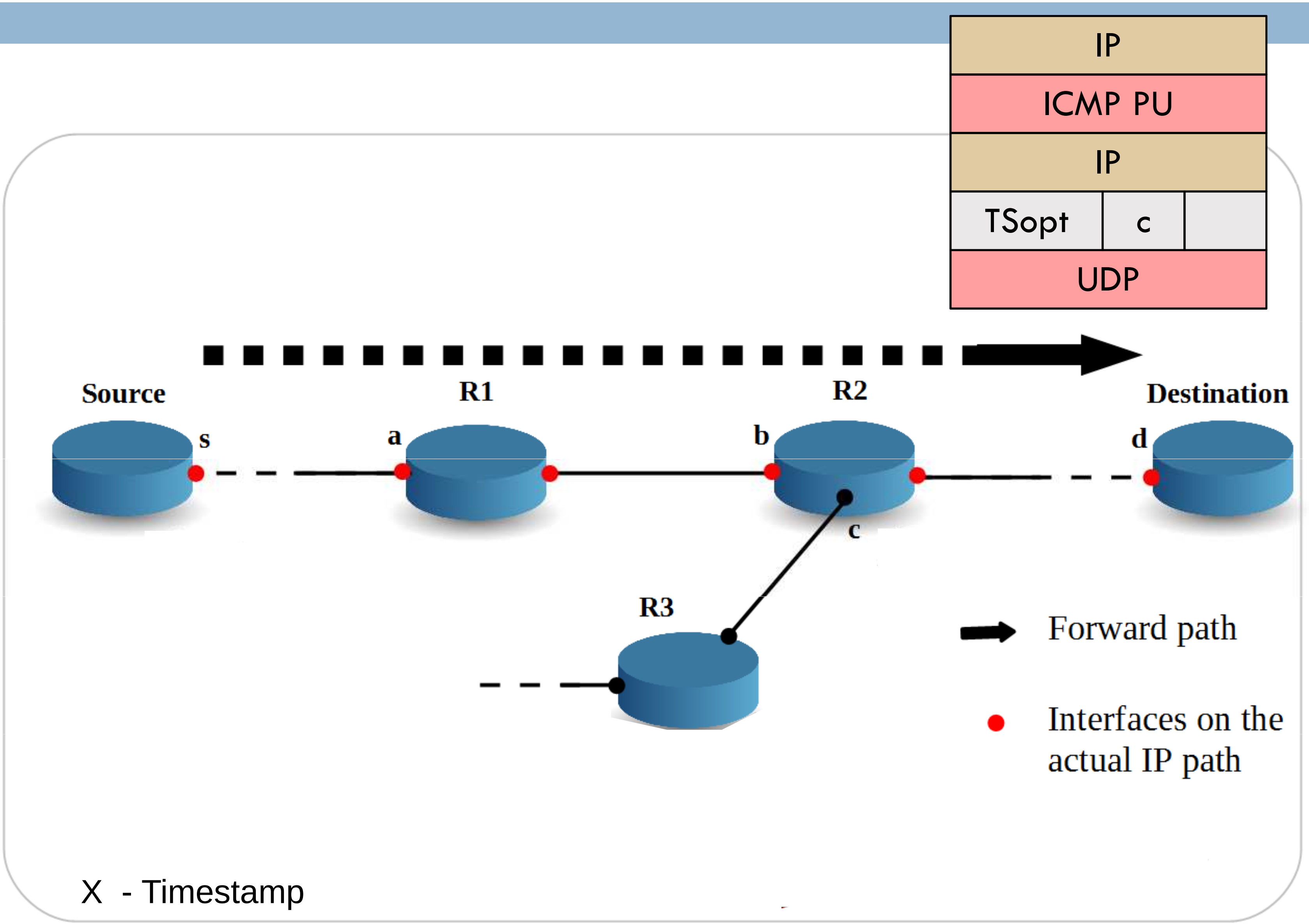
IP		
TSopt	c	
UDP		



X - Timestamp

UDP packet probes allow to avoid ambiguities caused by the reverse path

Identifying Third-Party addresses



UDP packet probes allow to avoid ambiguities caused by the reverse path

Identifying Third-Party addresses

IP

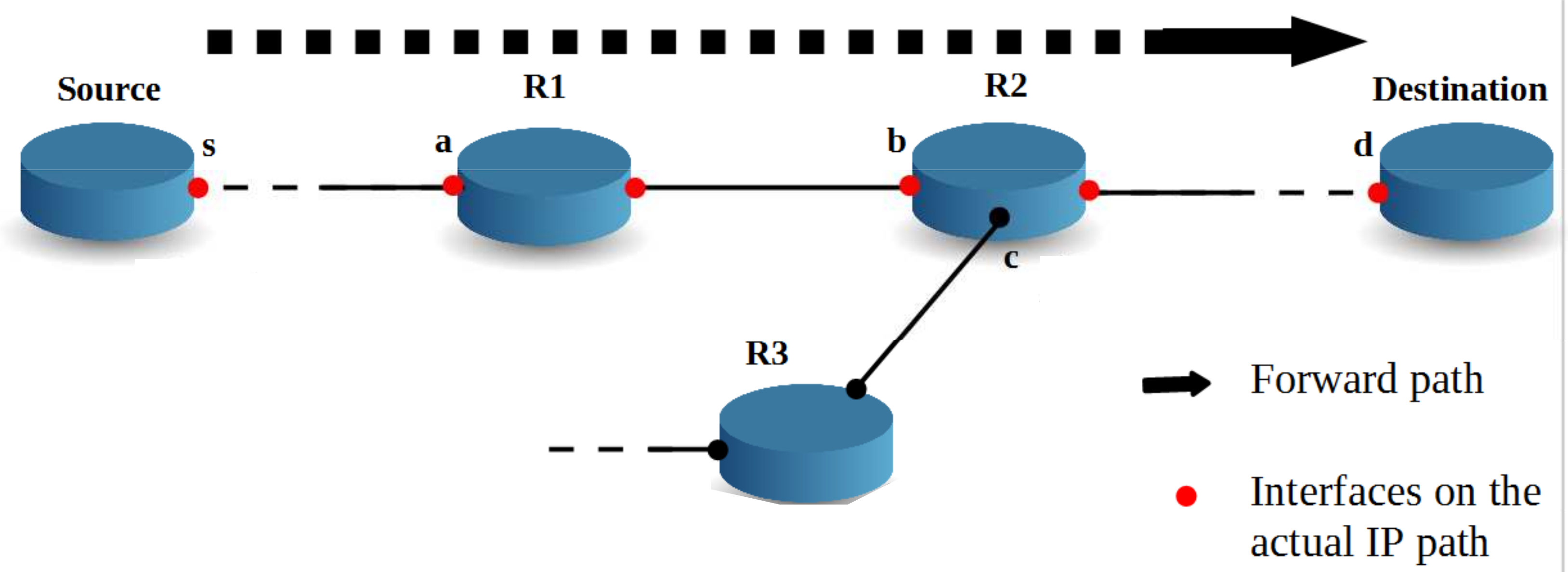
ICMP PU

IP

TSopt | c |

UDP

c is a Third-party address!



X - Timestamp

UDP packet probes allow to avoid ambiguities caused by the reverse path

Identifying Third-Party addresses

- A preliminary step is necessary to state if Y is owned by a per-interface stamping router
  - Ping Y pre-specifying Y four times
  - Typically the TS option is replicated in the Ping Reply

- Y is considered non-classifiable when
  - It is not clear if the owning router manages the TS option
    - Y is a private address
    - Y does not reply to Ping
    - Y does not provide timestamps in the Ping Reply
    - The TS option is removed from Ping Reply
  - It is owned by an any-interface stamping router [1]
    - Y provides 4 timestamps in the Ping Reply

- Y is considered classifiable only when it provides at least 1 Timestamp but less than 4 Timestamps in the Ping Reply

[1] W. de Donato, P. Marchetta, and A. Pescapé. "A Hands-on Look at Active Probing using the IP Prespecified Timestamp option". In PAM'12, Vienna, Austria, 2012.

Evaluation Methodology

Targeting 327K IPs in 14K distinct ASes showing stable responsiveness to

- Ping according to the PREDICT project
- UDP packet probes carrying the TS option

53 Planetlab nodes as vantage points located in distinct ASes performing the following steps for each destination

1. trace the IP path with UDP paris-traceroute
2. classify each IP address discovered along the path

Final dataset

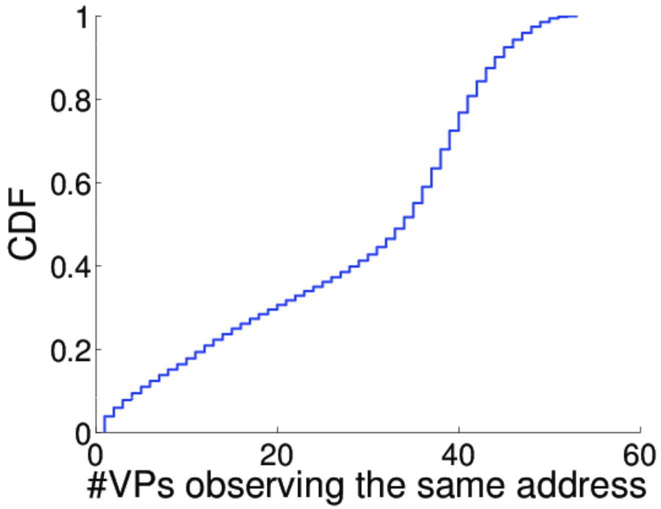
- removed traces containing loops or unable to reach the destination
- about 12M Traceroute traces and 443K Ips

Experimental results

- hop classifiability
- classification results
- impact on AS-level links and paths
- Comparison with the Hyun’s method

Hop classifiability

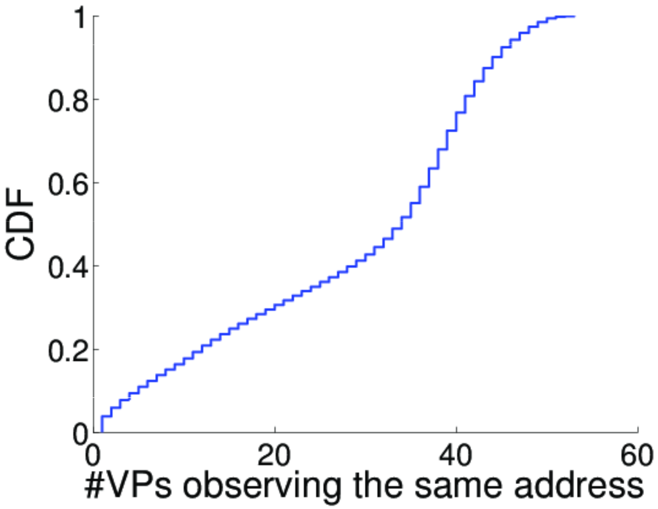
- The same IP address is captured by multiple VPs



+50% of IPs is observed
by more than 35 VPs

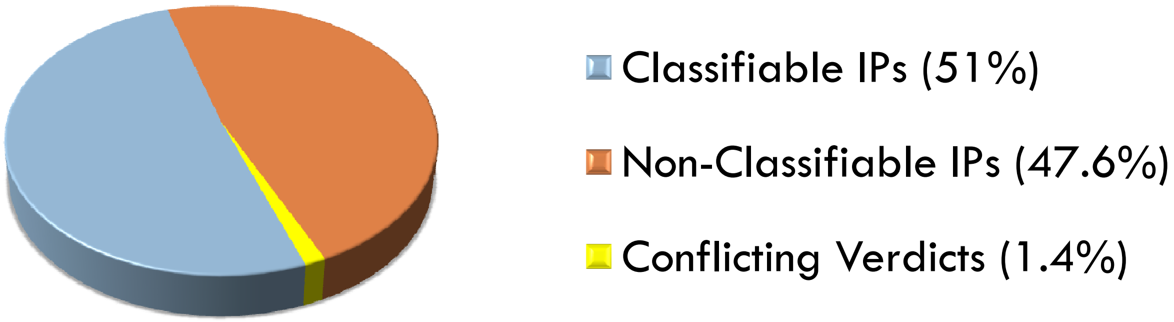
Hop classifiability

- □ The same IP address is captured by multiple VPs



+50% of IPs is observed
by more than 35 VPs

- □ The same IP address has been judged as classifiable or not by each vantage point



Hop classifiability

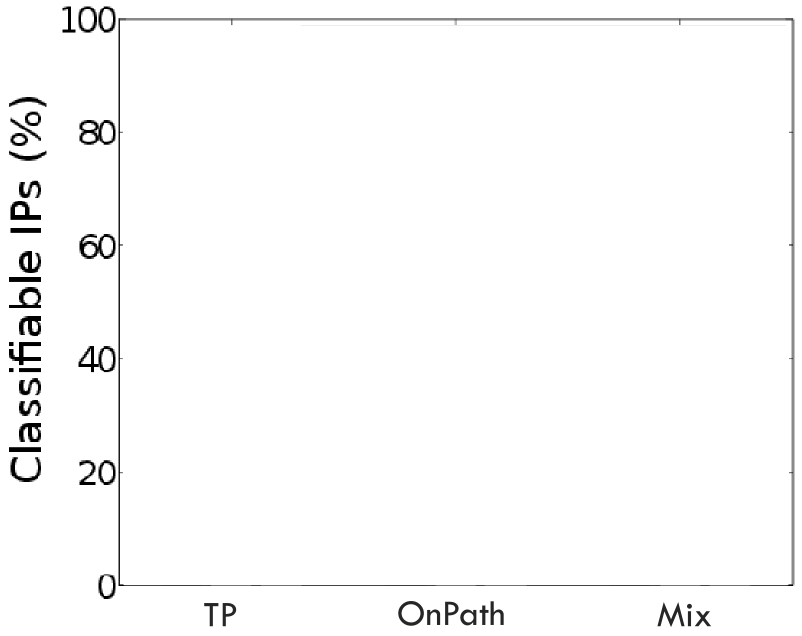
- Conflicting verdicts are mainly caused by filtering events
  - in–transit filtering of the Ping reply
  - removal of the TS option from the Ping reply

- Non–classifiable IPs breakdown

Category	IPs (K)	IPs (%)
IPs not replying to Ping	72.7	16.4
IPs not providing Timestamps	64.6	14.6
Any-interface stamping router	45.9	10.4
IPs providing a Ping Reply not containing the TS option	18.0	4.0
Private addresses	9.5	2.2
Total	**210**	**47.6**

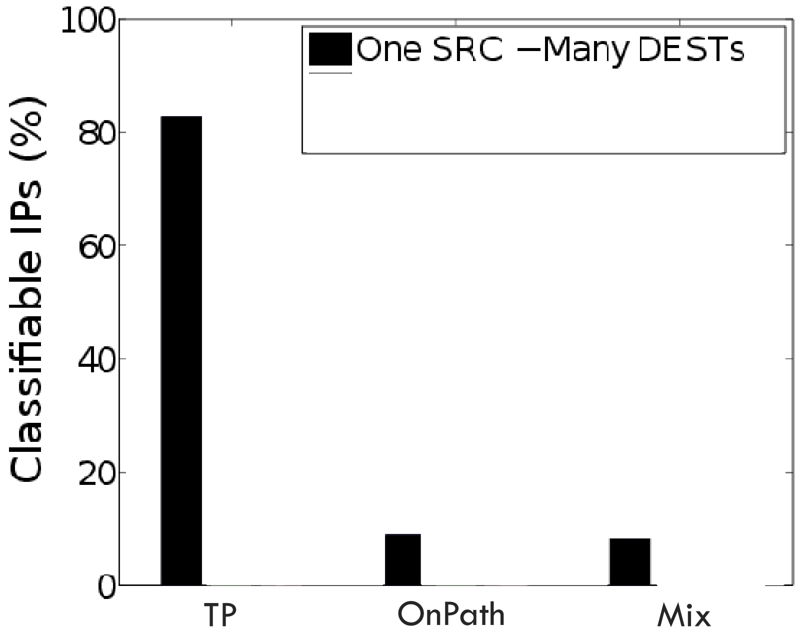
Classification Results

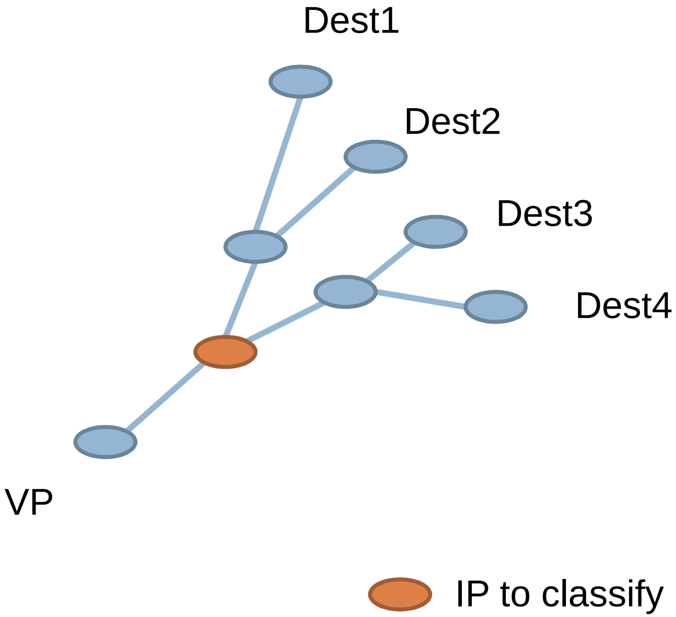
- An IP address appears in several paths and each time it has been classified
  - TP IPs always classified as Third-party addresses
  - OnPath IPs always classified as on the IP path
  - Mix IPs classified sometimes as Third-party addresses sometimes on the path



Classification Results

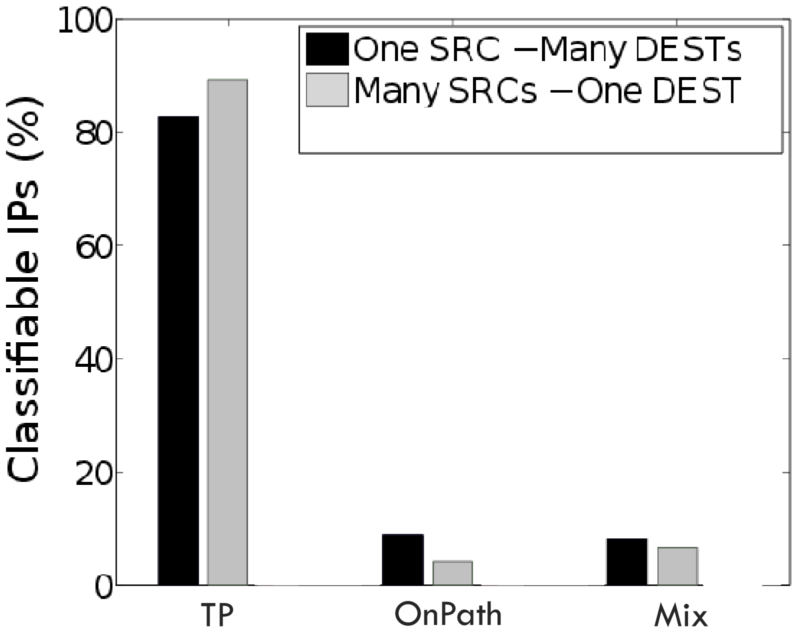
- An IP address appears in several paths and each time it has been classified
  - TP IPs always classified as Third-party addresses
  - OnPath IPs always classified as on the IP path
  - Mix IPs classified sometimes as Third-party addresses sometimes on the path

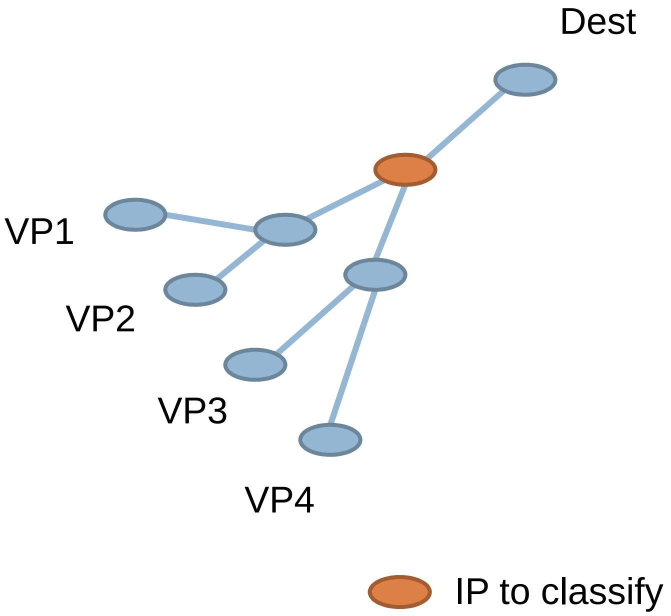




Classification Results

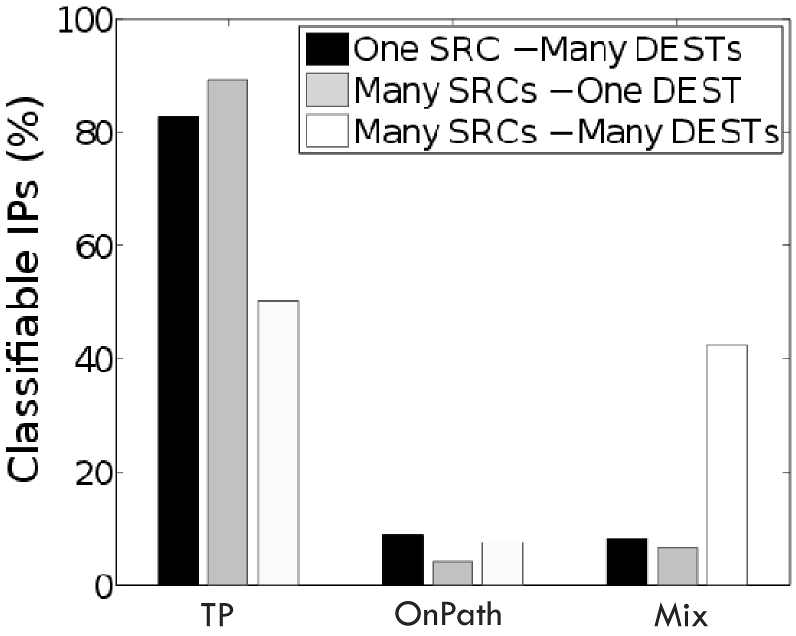
- An IP address appears in several paths and each time it has been classified
  - TP IPs always classified as Third-party addresses
  - OnPath IPs always classified as on the IP path
  - Mix IPs classified sometimes as Third-party addresses sometimes on the path

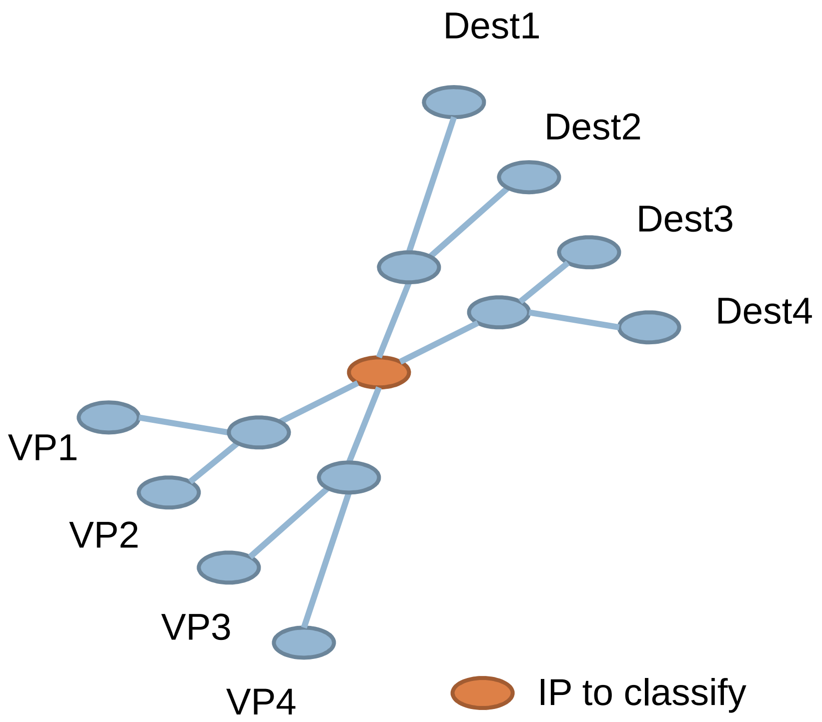




Classification Results

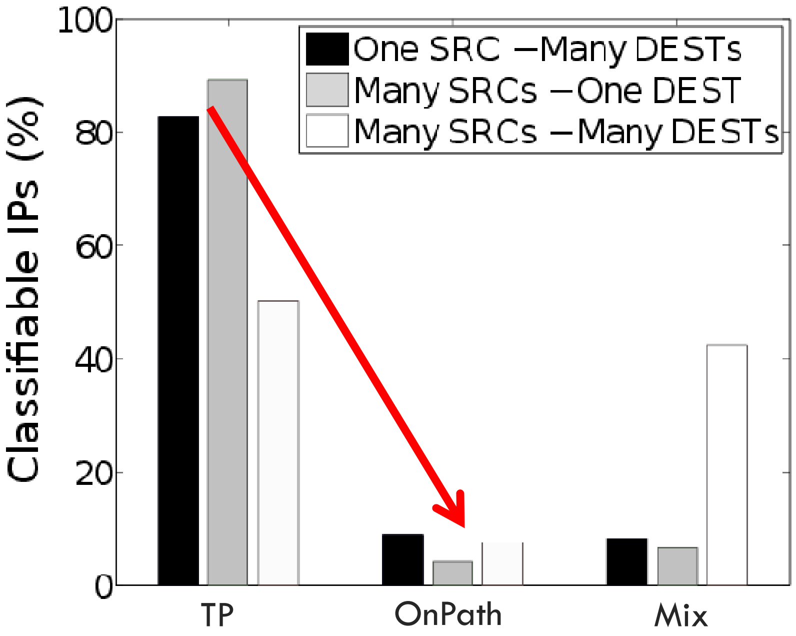
- An IP address appears in several paths and each time it has been classified
  - TP IPs always classified as Third-party addresses
  - OnPath IPs always classified as on the IP path
  - Mix IPs classified sometimes as Third-party addresses sometimes on the path





Classification Results

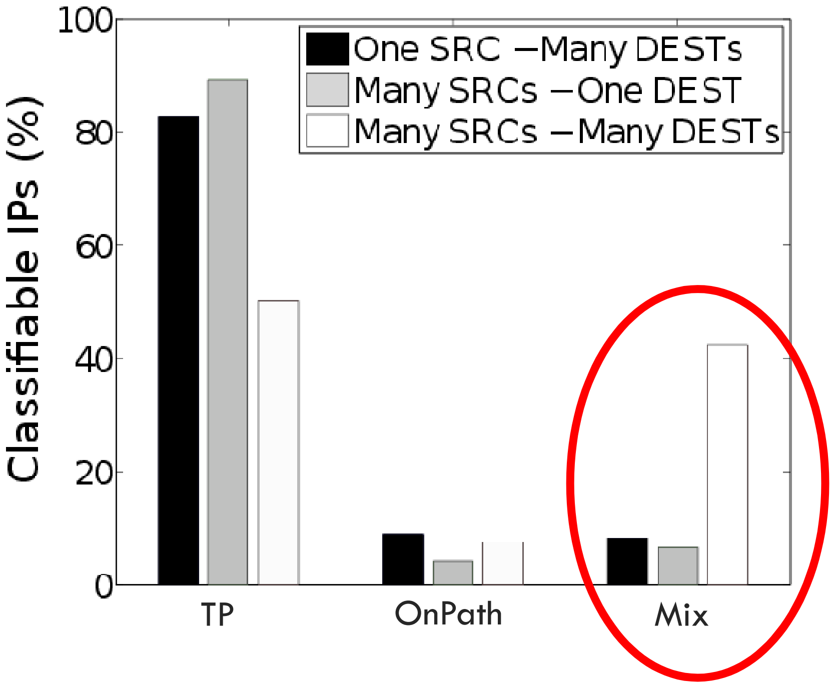
- An IP address appears in several paths and each time it has been classified
  - TP IPs always classified as Third-party addresses
  - OnPath IPs always classified as on the IP path
  - Mix IPs classified sometimes as Third-party addresses sometimes on the path



Routers often replies by exploiting an interface different from those actually traversed toward the Traceroute destination!

Classification Results

- An IP address appears in several paths and each time it has been classified
  - TP IPs always classified as Third-party addresses
  - OnPath IPs always classified as on the IP path
  - Mix IPs classified sometimes as Third-party addresses sometimes on the path



An address is a TP address or not depending on

1. the originating host
2. the targeted destination
3. both of them!

Impact on AS-level links

- IP-to-AS mapping to extract AS-level links
  - not considering IXPs hops and links involving Sibling ASes, MOAS and unmapped hops

- An AS-level link may appear in several traces being determined by several distinct pairs of IP addresses
  - Not Affected AS link
    - at least once, both the involved IPs have been classified as on path
  - Potentially Not Affected AS link
    - at least once, both IPs have been labeled as non-classifiable
  - Affected AS link
    - always, at least one of the involved IPs has been classified as TP

Impact on AS-level links

- IP-to-AS mapping to extract AS-level links
  - not considering IXPs hops and links involving Sibling ASes, MOAS and unmapped hops

- An AS-level link may appear in several traces being determined by several distinct pairs of IP addresses
  - Not Affected AS link
    - at least once, both the involved IPs have been classified as on path
  - Potentially Not Affected AS link
    - at least once, both IPs have been labeled as non-classifiable
  - Affected AS link
    - always, at least one of the involved IPs has been classified as TP

Total AS links: 34,414

AS-level link classification	AS-links (%)
Not Affected AS-links	**6.2**
Potentially Not Affected AS-links	**76**
Affected AS-links	**17.8**

Impact on AS-level links

- IP-to-AS mapping to extract AS-level links
  - not considering IXPs hops and links involving Sibling ASes, MOAS and unmapped hops

- An AS-level link may appear in several traces being determined by several distinct pairs of IP addresses
  - Not Affected AS link
    - at least once, both the involved IPs have been classified as on path
  - Potentially Not Affected AS link
    - at least once, both IPs have been labeled as non-classifiable
  - Affected AS link
    - always, at least one of the involved IPs has been classified as TP

Total AS links: 34,414

Third-party addresses affect a significant percentage of AS-level links!

AS-level link classification	AS-links (%)
Not Affected AS-links	**6.2**
Potentially Not Affected AS-links	**76**
Affected AS-links	**17.8**

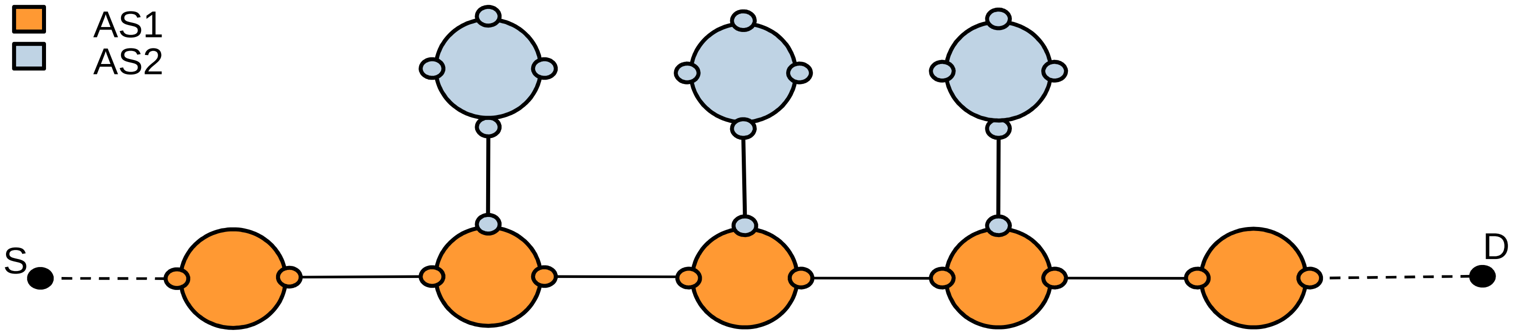
!

TP addresses and AS-level loops

- About 587K traces normally reaching the destinations contain AS-level loops
  - about 37% of the loops involves TP addresses
    - 105K loops start with a TP address
    - 149K loops end with a TP address
    - 6K loops contain consecutive TP addresses

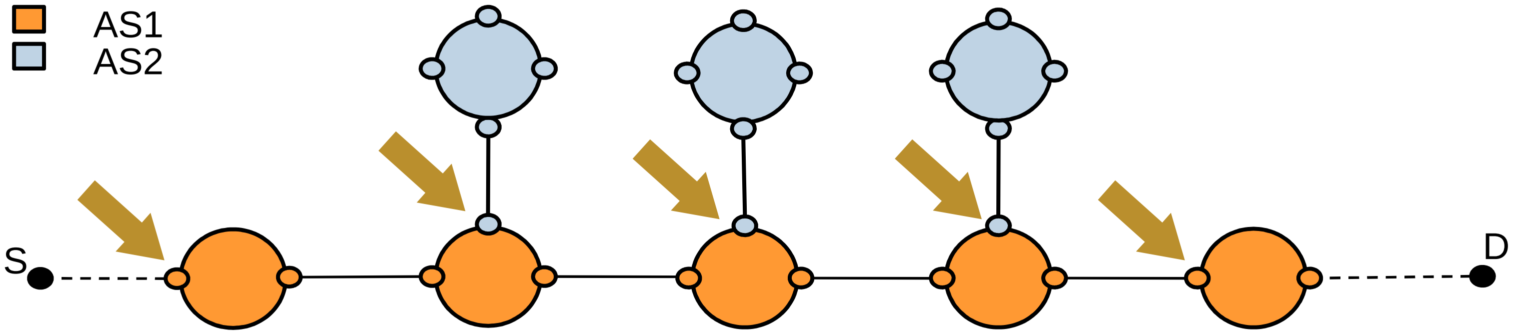
TP addresses and AS-level loops

- About 587K traces normally reaching the destinations contain AS-level loops
  - about 37% of the loops involves TP addresses
    - 105K loops start with a TP address
    - 149K loops end with a TP address
    - 6K loops contain consecutive TP addresses



TP addresses and AS-level loops

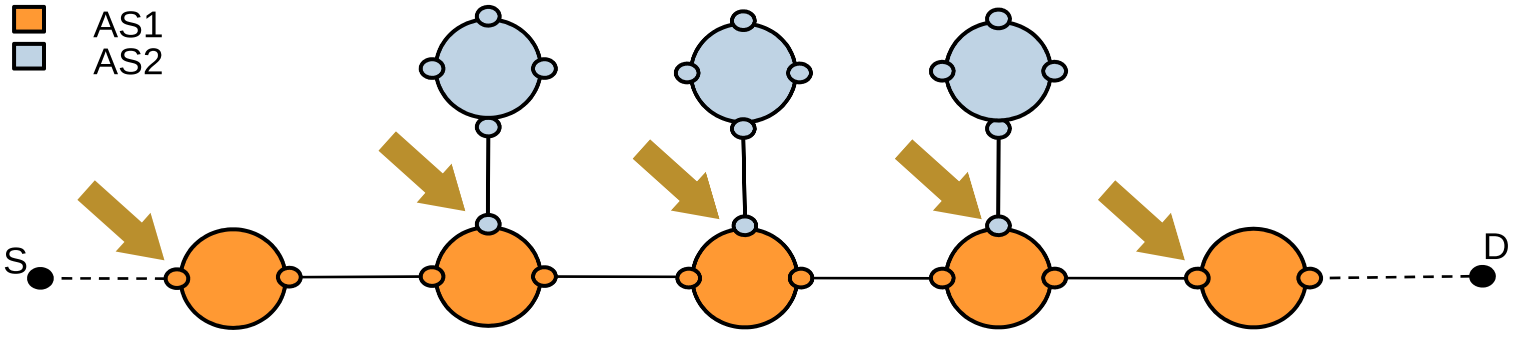
- About 587K traces normally reaching the destinations contain AS-level loops
  - about 37% of the loops involves TP addresses
    - 105K loops start with a TP address
    - 149K loops end with a TP address
    - 6K loops contain consecutive TP addresses



AS-level path: AS1 AS2 AS2 AS2 AS1

TP addresses and AS-level loops

- About 587K traces normally reaching the destinations contain AS-level loops
  - about 37% of the loops involves TP addresses
    - 105K loops start with a TP address
    - 149K loops end with a TP address
    - 6K loops contain consecutive TP addresses

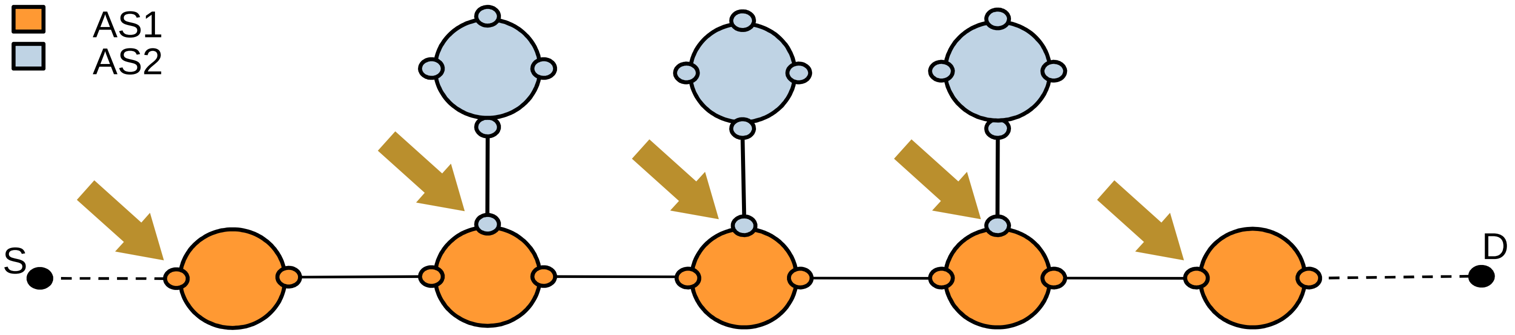


AS-level path: AS1 AS2 AS2 AS2 AS1

We never left AS1 !

TP addresses and AS-level loops

- About 587K traces normally reaching the destinations contain AS-level loops
  - about 37% of the loops involves TP addresses
    - 105K loops start with a TP address
    - 149K loops end with a TP address
    - 6K loops contain consecutive TP addresses



AS-level path: AS1 AS2 AS2 AS2 AS1

We never left AS1 !

Third-party addresses may be responsible for bogus AS-level loops

Comparison with the Hyun's method

> *"An intermediate address that resolves to an AS that differs from the ASes of both adjacent addresses in the same path is a candidate Third-party address"*
>
> Hyun *et al*. (PAM'03)

- Just 1.5% of the TP addresses identified by our technique is also detected by the Hyun's method
  - a TP address is such independently from the AS point of view
  - an address is a TP address depending on the source and the destination
  - also a single AS-level transition may be affected by TP addresses
  - a Traceroute trace may contain multiple consecutive TP addresses

Summary

An active probing technique able to identify TP addresses in Traceroute traces

- exploiting the IP Pre-specified Timestamp Option
- no BGP information
- no pre-collected information about the topology

Main findings

- the same IP address is a Third-party address or not depending on the Traceroute originating host, the targeted destination, both of them
- routers often reply to Traceroute by exploiting an interface different from those actually traversed toward the Traceroute destination
- a significant percentage of Traceroute-derived AS-level links are affected by Third-party addresses
- Third-party addresses may be responsible for bogus AS-level loops

When Traceroute is used to infer the AS-level topology,
TP addresses may represent a strong source of AS map distortion!

IP options based measurements

- Many researchers believe that *IP options are not an option* (Fonseca et al. 2005)
  - IP options expose packet probes to filtering policies
  - IP options are not widely supported
  - IP options are poorly implemented

- Large scale experiments demonstrate how filtering actually depends on
  - the type of packet probes (UDP, ICMP, TCP, …)
  - the type of IP options (TS, RR, SSRR, …)
  - the type of routers on the path

- More and more active probing techniques based on IP options have been recently proposed
  - Reverse Traceroute (best paper, NSDI'10)
  - Alias resolution (IMC'10)
  - Quantifying violations of destination-based forwarding (IMC'12)
  - Detecting and locating Hidden routers (GI'13)
  - Detecting Third-party addresses (best poster, SIGCOMM'12)
  - Inferring router statistics (CONEXT'10)
  - and more …

IP options represent an amazing though limited tool
for Internet measurements!