bdrmap: Inference of Borders Between IP Networks

Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, David Clark, kc claffy

IMC 2016, November 15th 2016

KO TE TANGATA

CSAIL

caida

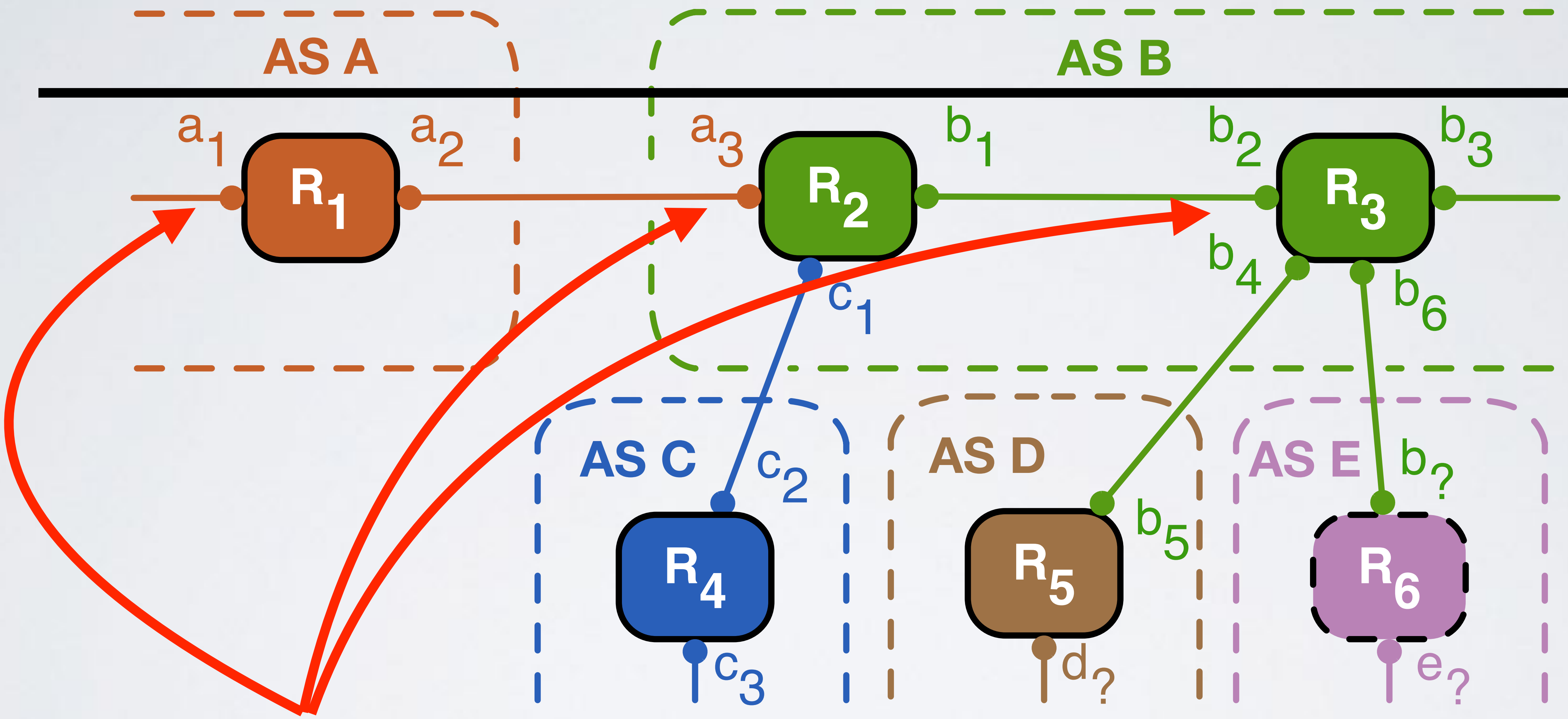
Who operates a router we observe in a traceroute path?

The Problem

1. The Internet architecture has no notion of interdomain boundaries at the network layer
2. Traceroute is a 30-year old hack with limitations
3. Using longest-matching prefix to infer ownership of routers is known to be error prone
4. Traceroute samples topology close to Vantage Point (VP), reducing topological constraints for inferring ownership for distant routers
5. Concerns about revealing topology information can align operator incentives away from transparency

Assumption

IP path: **a_1 a_3 b_2**

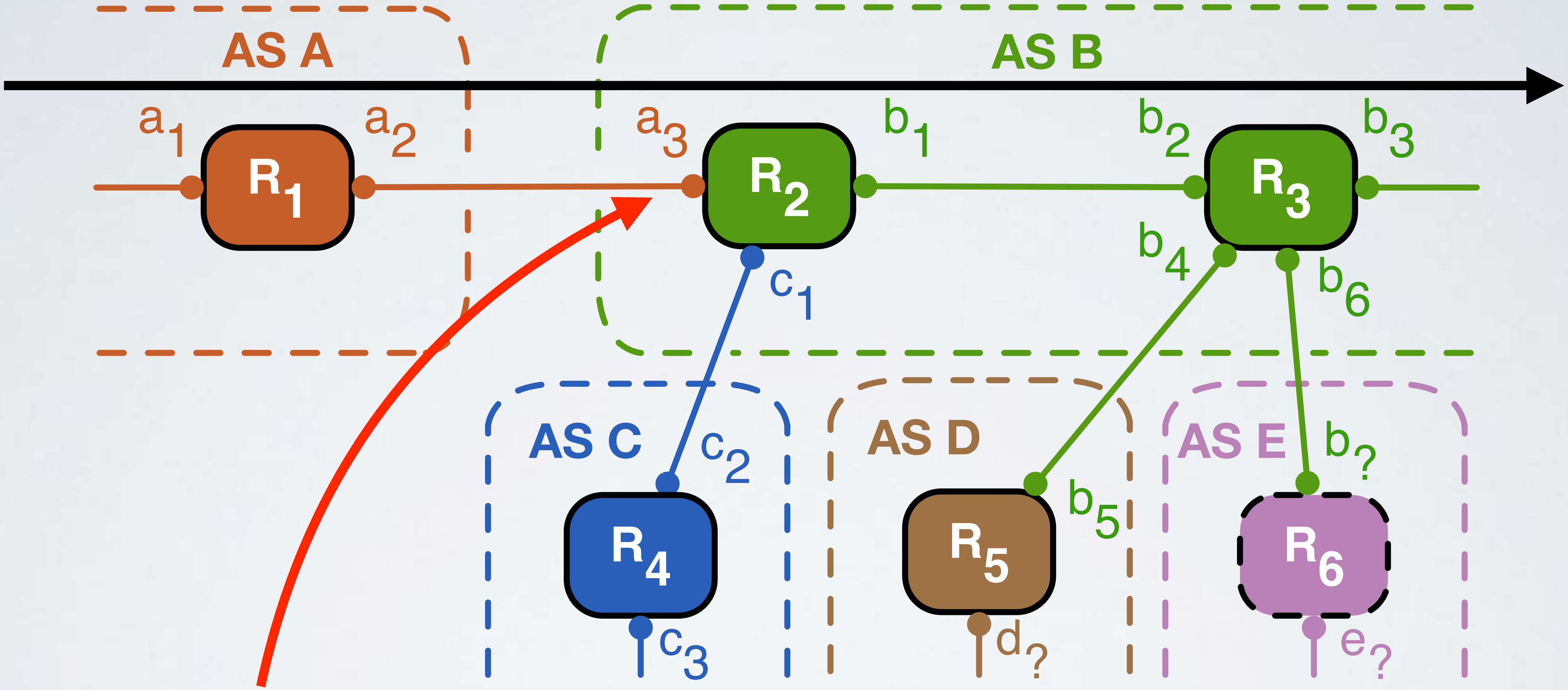


We assume that routers generally respond with an on-path interface facing the VP, and links between routers are point-to-point (IPv4 /30 or /31)

Challenges

IP path: a_1 a_3 b_2

Router path: R_1 R_2 R_3

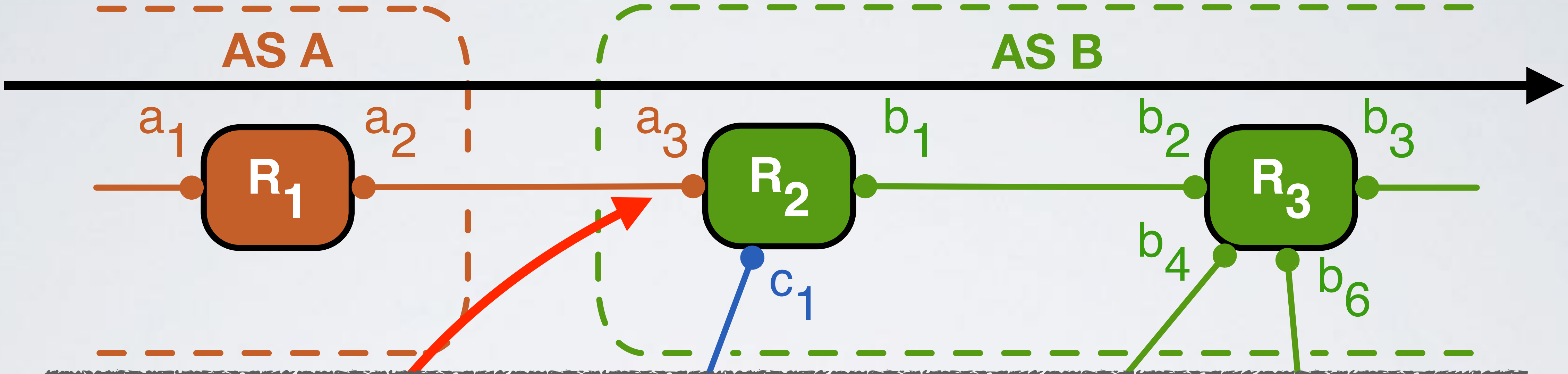


Neighbor router owned by **AS B** may respond with an IP address from **AS A** which the router uses to form the point to point link.

Challenges

IP path: a_1 a_3 b_2

Router path: R_1 R_2 R_3



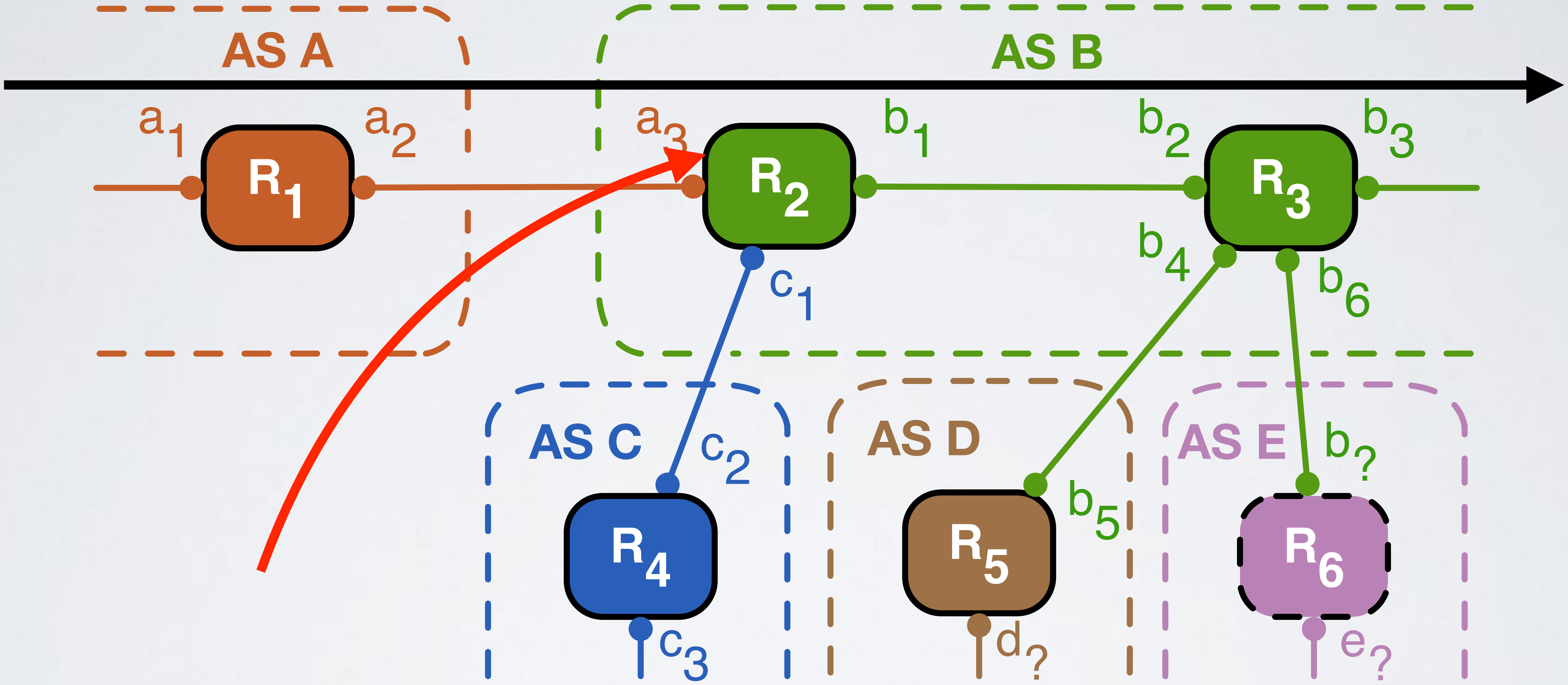
Industry convention for provider to assign interconnect IP address, but no convention for peering

Neighbor router owned by **AS B** may respond with an IP address from **AS A** which the router uses to form the point to point link.

Challenges

IP path: a_1 c_1 b_2

Router path: R_1 R_2 R_3

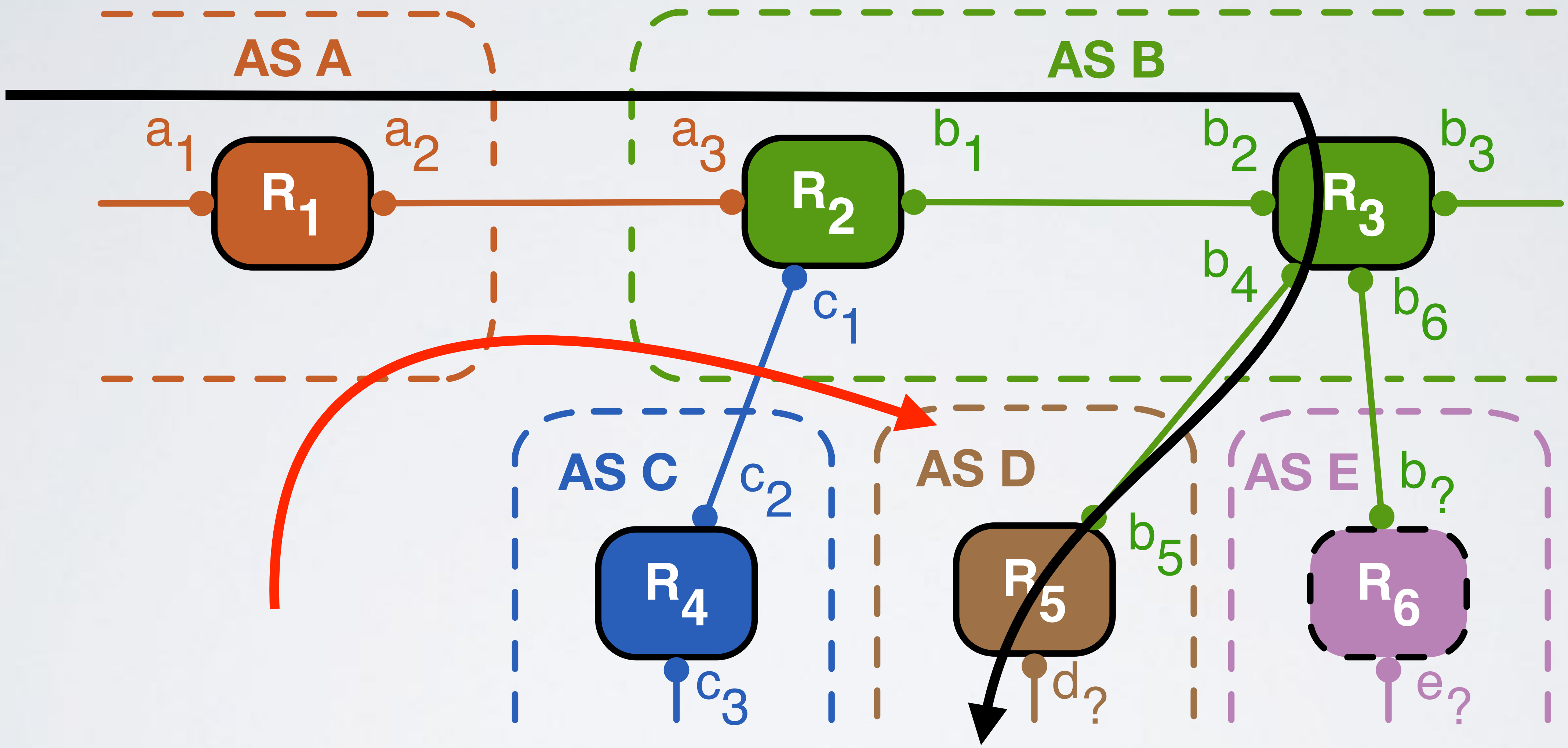


Neighbor router owned by **AS B** may respond with a third party IP address from **AS C**

Challenges

IP path: **a_1 a_3 b_2 b_5 ?**

Router path: R_1 R_2 R_3 R_5 ?

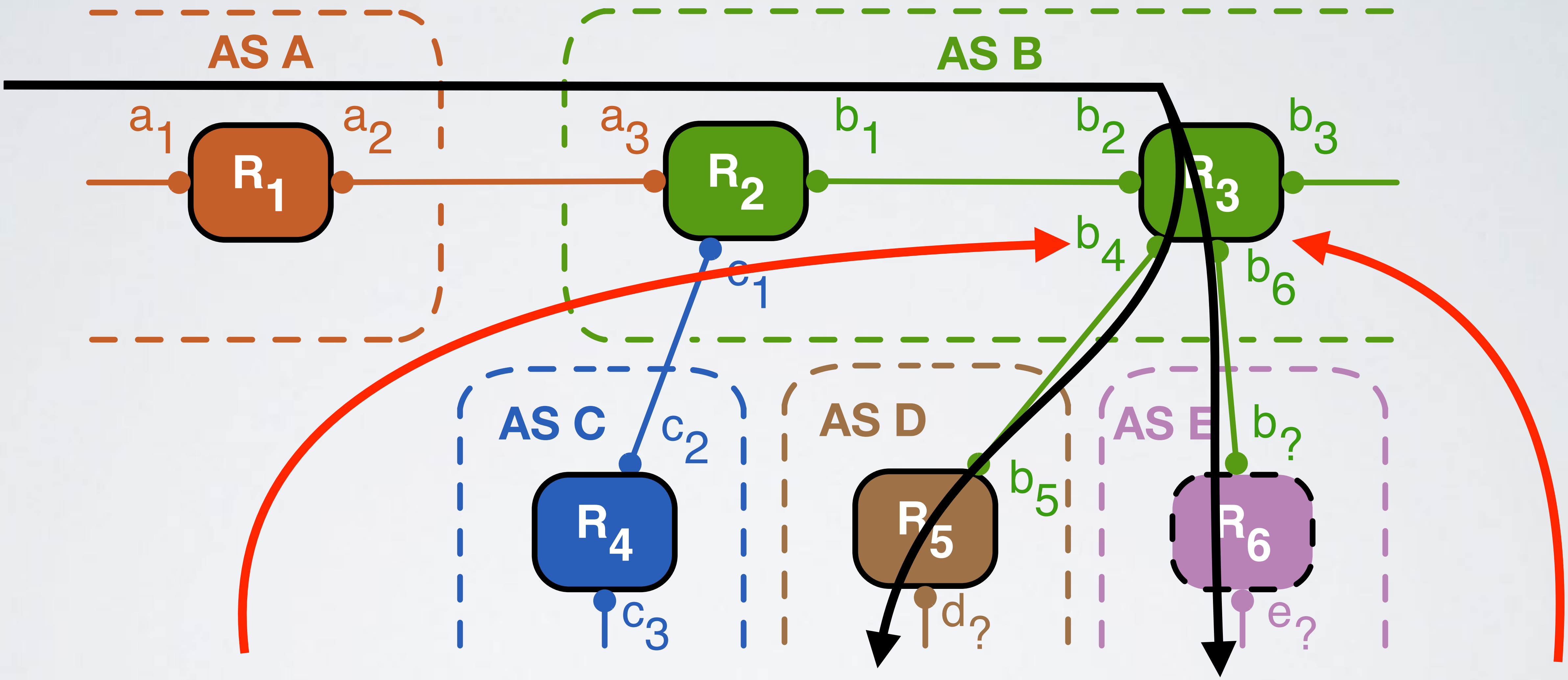


Border router operated by **AS D** may respond with an address from **AS B** used to form point-to-point link, but block probes from entering **AS D**

Challenges

IP paths: a_1 a_3 b_4 b_5 a_1 a_3 b_6 ?

Router paths: R_1 R_2 R_3 R_5 R_1 R_2 R_3 ?

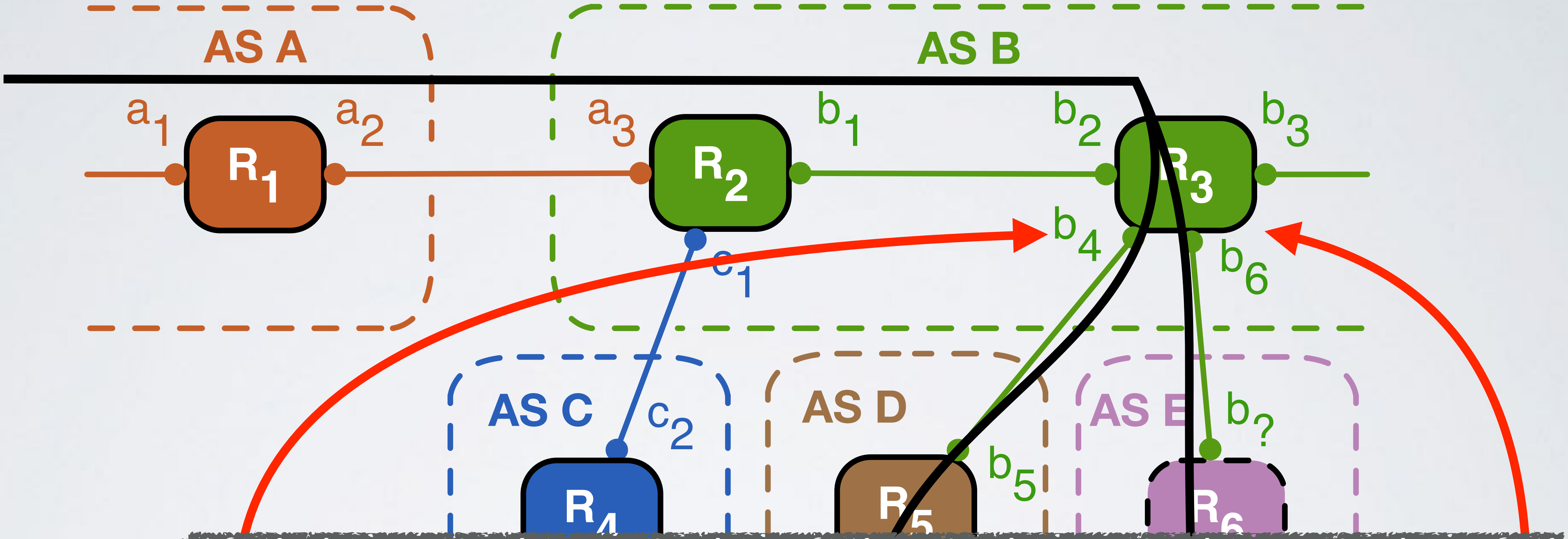


Border router owned by **AS B** may use virtual routing features; the router will respond with different IP addresses that form the point-to-point link with **AS D** and **AS E**

Challenges

IP paths: a_1 a_3 b_4 b_5 a_1 a_3 b_6 ?

Router paths: R_1 R_2 R_3 R_5 R_1 R_2 R_3 ?



If b_4 and b_6 are not resolved for aliases, and E's router is silent, b_6 might be incorrectly inferred to be neighbor E's router

Border r

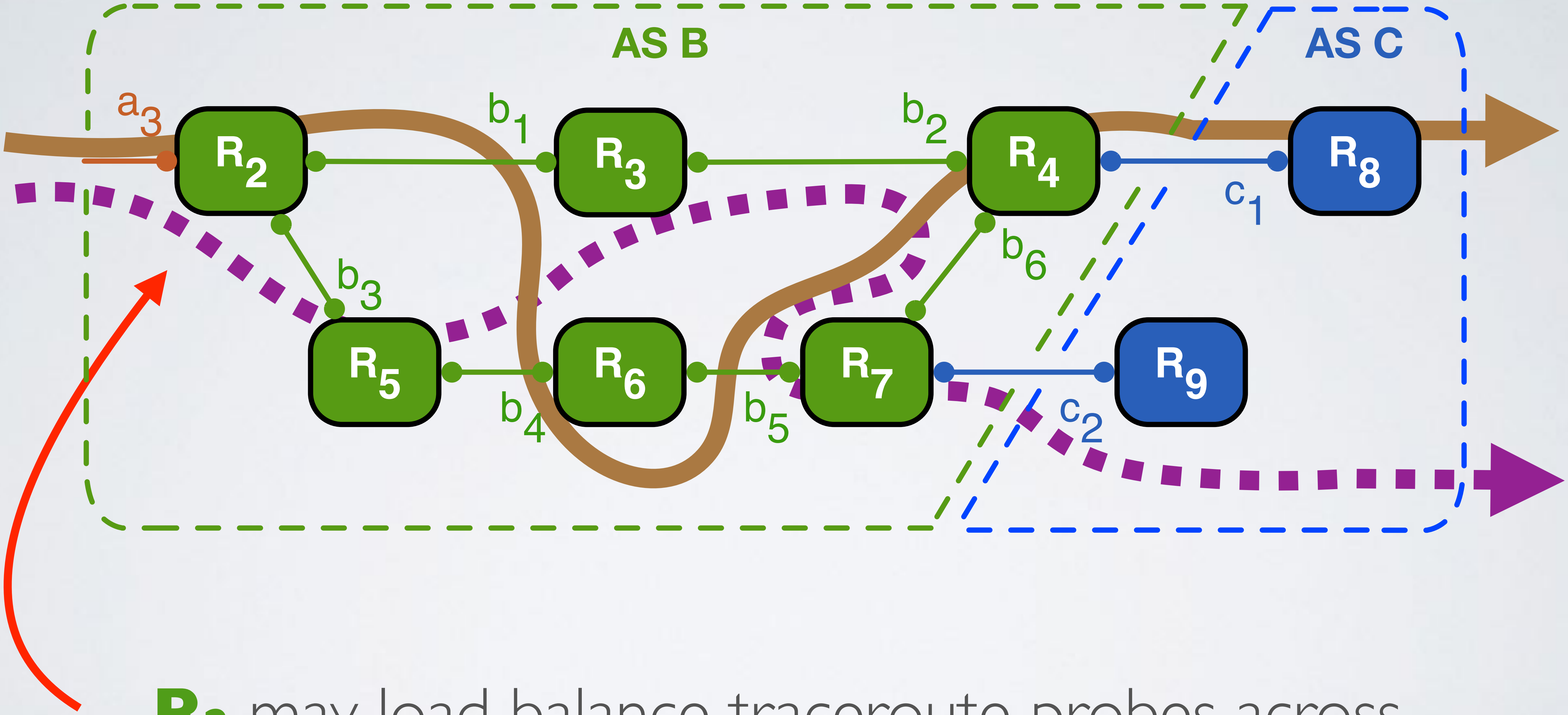
features; th

that forn

Challenges

IP paths: a_3 b_1 b_4 b_2 a_3 b_3 b_2 b_5

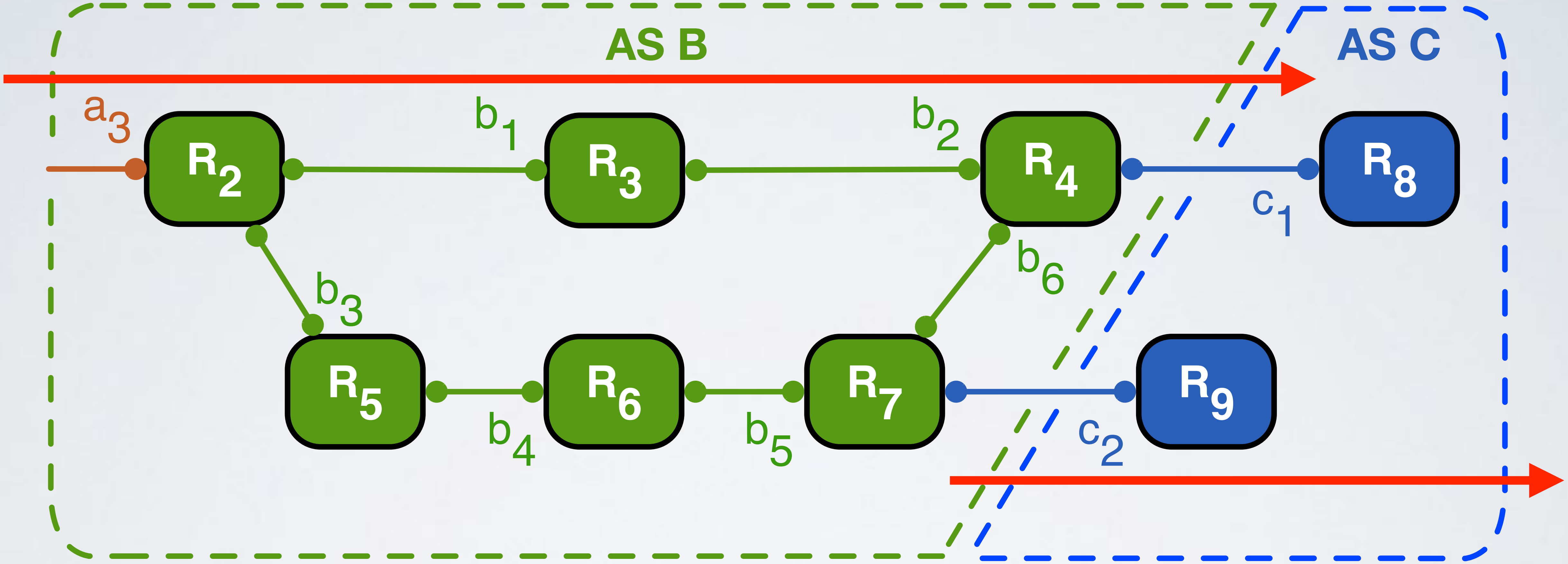
Router paths: R_2 R_3 R_6 R_7 R_2 R_5 R_4 R_8



R_2 may load balance traceroute probes across topologically diverse paths, resulting in false links.

Challenges

IP paths:	a_3	b_1	b_2	c_1	c_2
Router paths:	R_2	R_3	R_4	R_8	R_9



R_2 may choose a different next hop as a traceroute measurement proceeds

Additional Challenges

1. **Sibling ASes** may confuse attempts to infer connectivity between organizations
   - sibling information has known false and missing inferences
2. **IXP addresses** may appear inconsistently in paths
   - an IXP and/or member(s) may originate prefix into BGP, or it might not be originated at all
3. **Multiple ASes** may originate a prefix into BGP
   - The more ASes, the more challenging to infer ownership

Motivation for Border Router Ownership Inference

- **Network Modeling and Resilience**
  - Early Internet models considered topology at AS-level, with a single link between pairs of ASes
  - Our work enables the construction of a router-level Interdomain connectivity map
- **Interdomain Congestion**
  - Public policy community has growing interest in identifying persistent congestion on interdomain links
  - Greatest measurement challenge is identifying interdomain links to probe, and associating observed evidence of congestion to specific interdomain links

Related Work

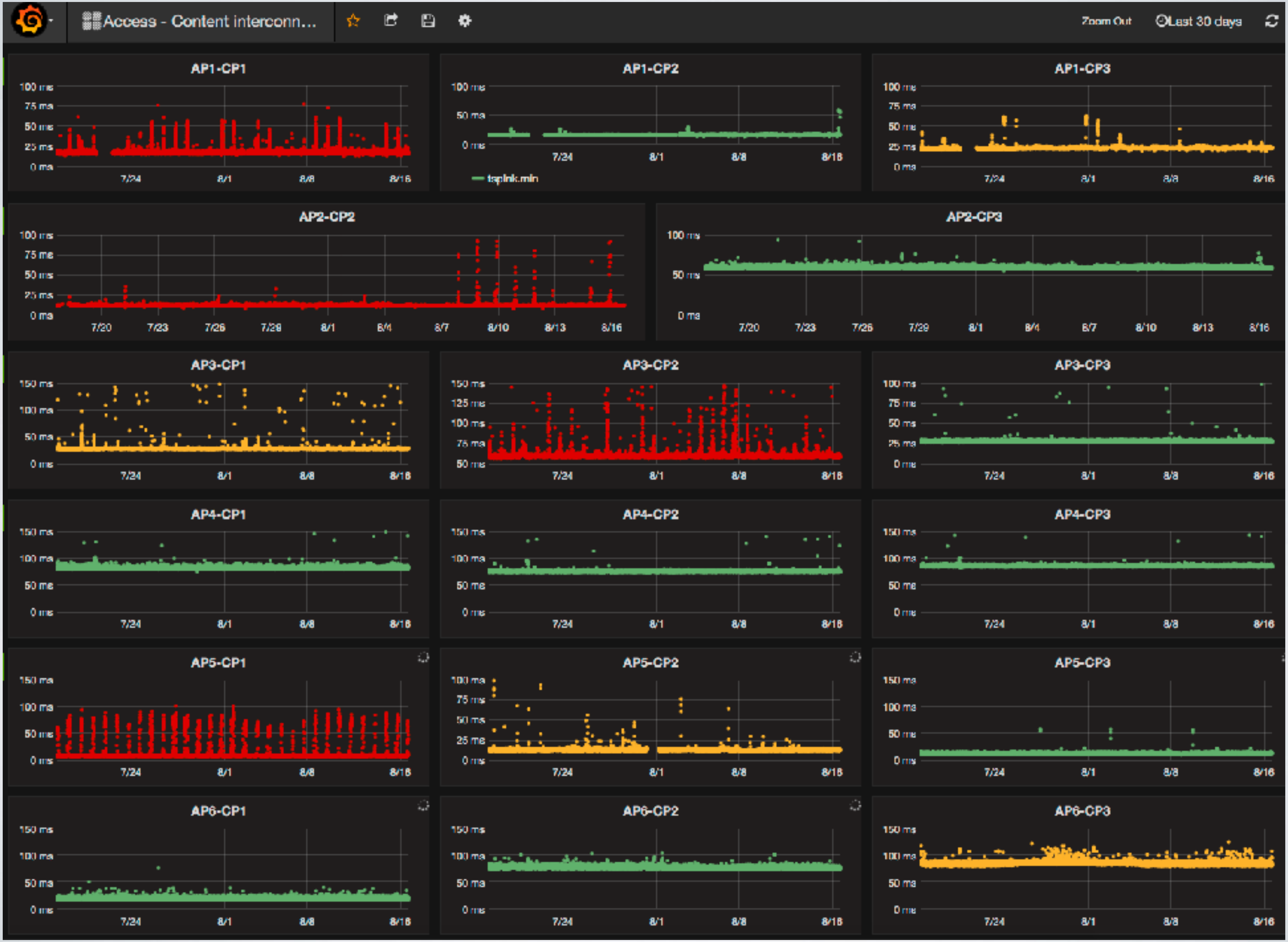
- **Significant work on inferring router aliases**: e.g., Mercator, Ally, Pre-specified timestamps, mrinfo, Discarte, Radargun, MIDAR, APAR + kapar,
- **Significant work inferring AS-level connectivity**
  - AS traceroute (SIGCOMM 2002 and SIGMETRICS 2003); adjust IP-AS mappings with colocated traceroute and BGP
  - Where the Sidewalk ends (CoNEXT 2009): goal of accurately inferring AS-level connectivity
  - Topology dualism (PAM 2010): evaluation of heuristics

Our Contributions

1. **Scalable, accurate method** for inferring interdomain boundaries for the network hosting the VP
2. **Efficient system** to allow deployment on resource-limited devices (e.g. SamKnows)
3. **Validation using ground truth** from four network operators and IXP databases
4. **Analysis of interdomain connectivity** of a large access ISP (Comcast): 45 links w/ Level3, Jan 2016
5. **We release our data collection and analysis system** as part of scamper, with man pages

https://www.caida.org/tools/measurement/scamper/

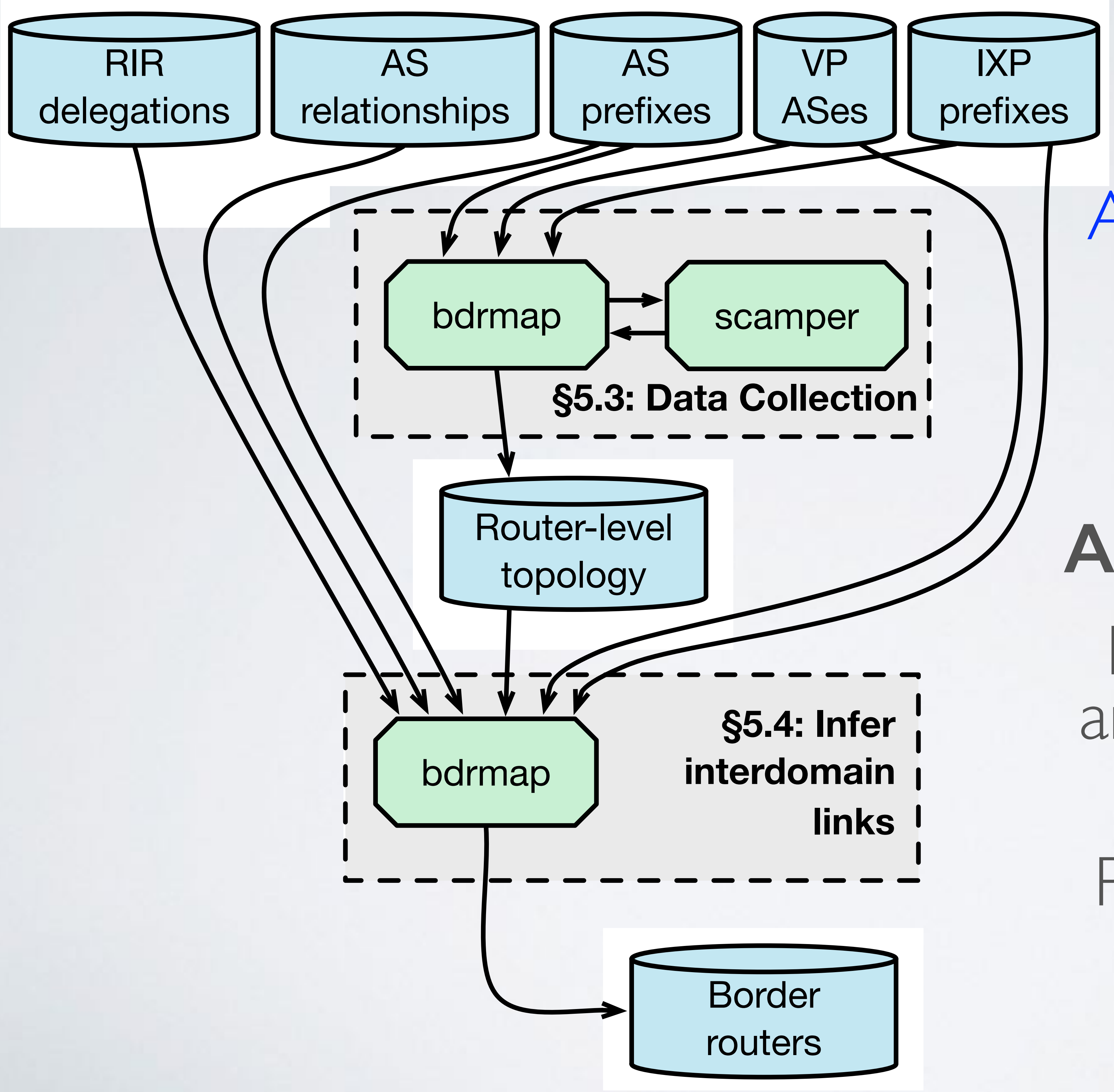
Select Interconnections from Top 3 Content to Top 6 Access



Roadmap

- bdrmap
  - Input data
  - Data collection
  - Analysis: overview of heuristics
- Validation, coverage of BGP-observed links
- Systems challenges and solutions
- Interconnection Insights

Approach to Border Mapping (I)



Assemble Input Data

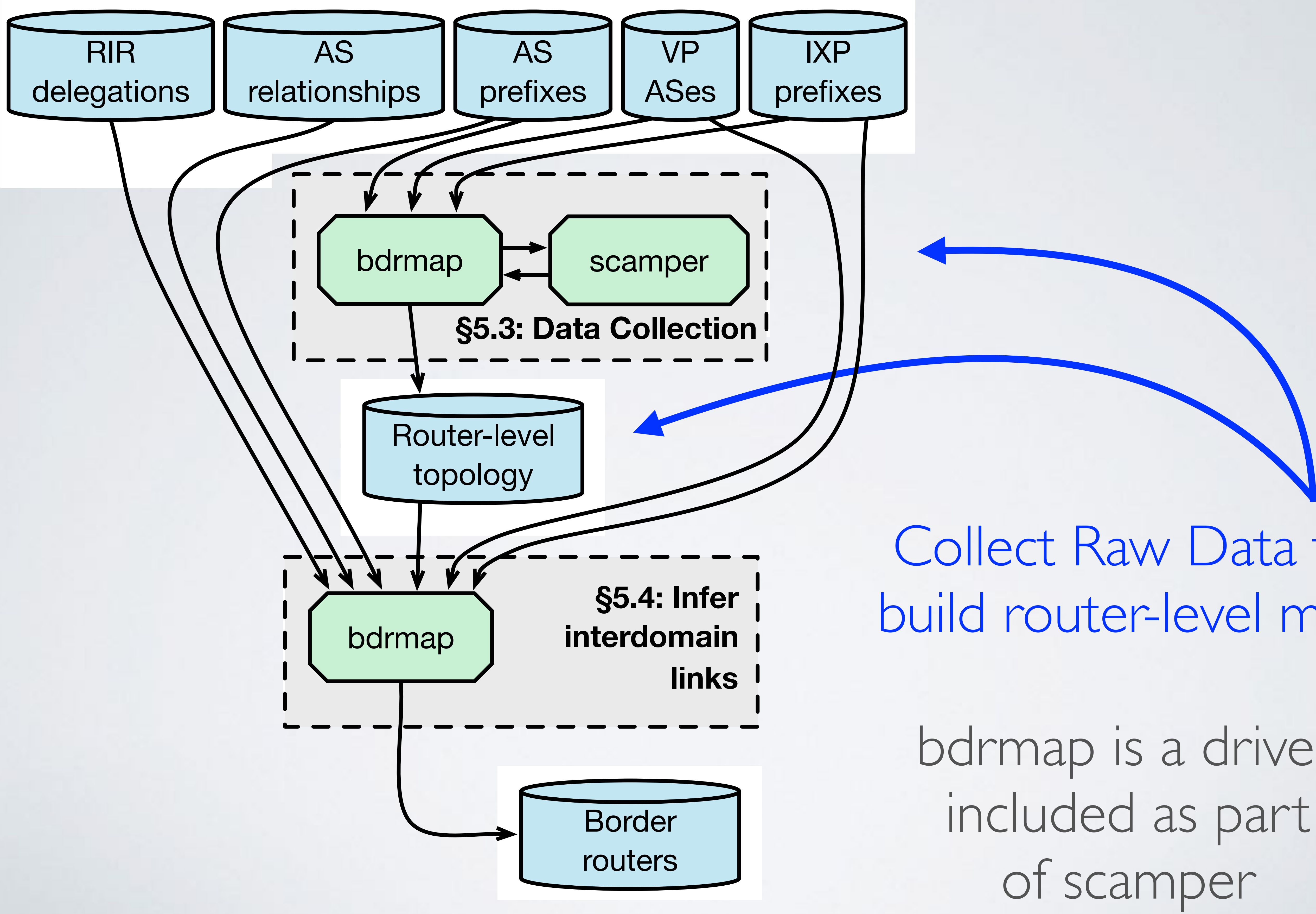
RIR delegations: RIR statistics files

AS relationships and prefixes: BGP data and as-rank algorithm

IXP prefixes: PeeringDB and PCH

VP ASes: manual oversight

Approach to Border Mapping (2)



Collect Raw Data to build router-level map

bdrmap is a driver, included as part of scamper

Data Collection

- Parts of our data collection process are similar to Rocketfuel
    - targeted traceroutes, informed by public BGP data
    - alias resolution to infer router-level topology
- **Rocketfuel** maps topologies of networks from the outside
- **bdrmap** maps interdomain topologies from the inside
- bdrmap data collection time depends on diameter and complexity of hosting network;
    - typically 12-48 hours at 100pps

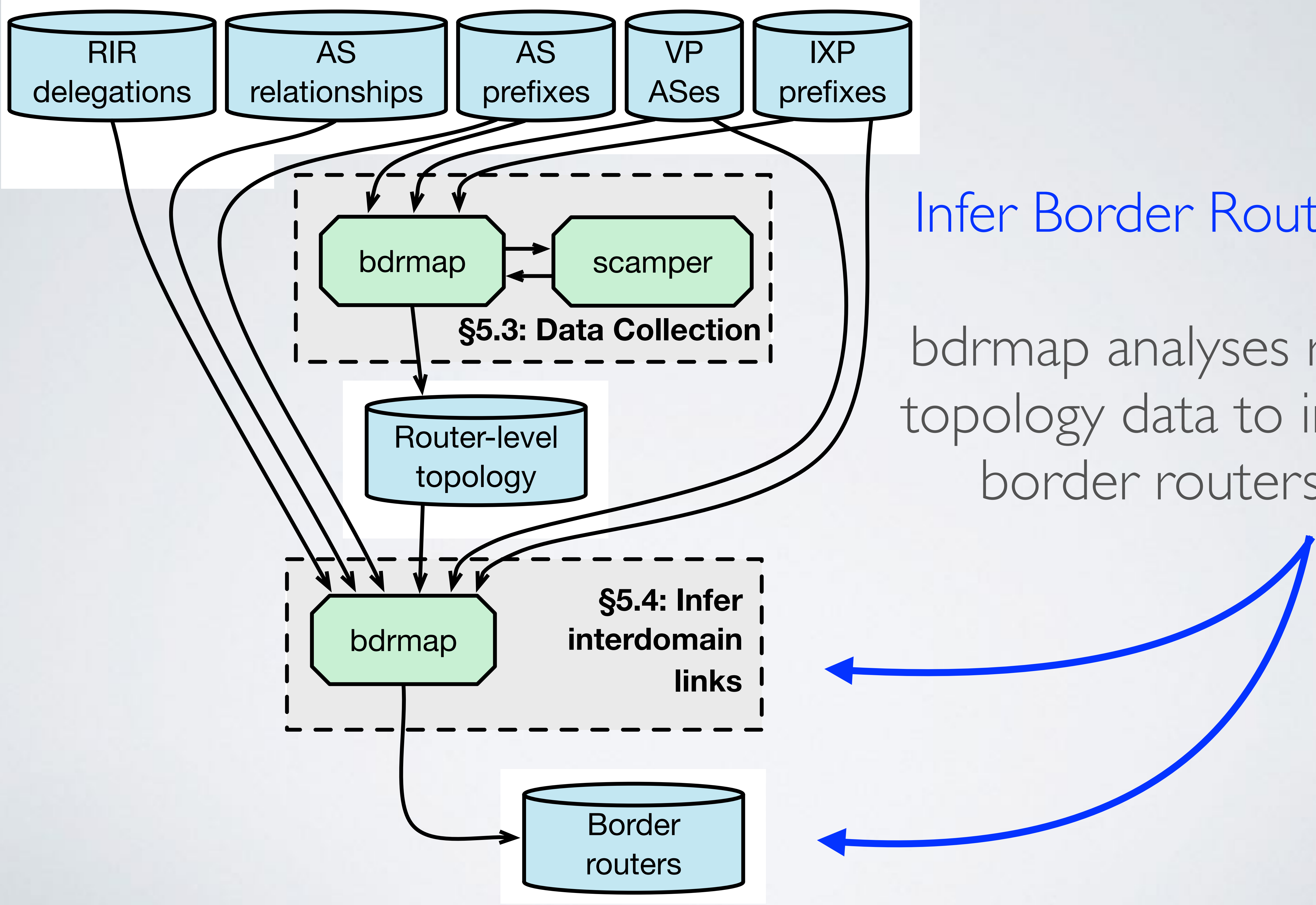
Data Collection

- **Generate list of address blocks from BGP data**
- **Gather traceroutes**
  - we focus on first-hop interdomain links, so use a **stop-set** (DoubleTree) to halt traceroutes from probing beyond hops in a neighbor network we have seen before
  - bdrmap tries up to five different addresses per block, to avoid interpreting third-party addresses as neighbors

Data Collection

- **Perform Alias Resolution**
  - We use the Ally, Mercator, and Prefixscan alias resolution techniques as we collect traceroutes to collect raw data to support building a router-level graph
  - We use MIDAR's Monotonic Bounds Test where we use IP-ID based techniques, as well as repeated tests, to reduce the chance we infer false aliases
- **Build Router Level Graph**
  - Focus on interfaces observed in ICMP TTL-expired messages; the source address on an ICMP echo response could be on any of the interfaces on the router

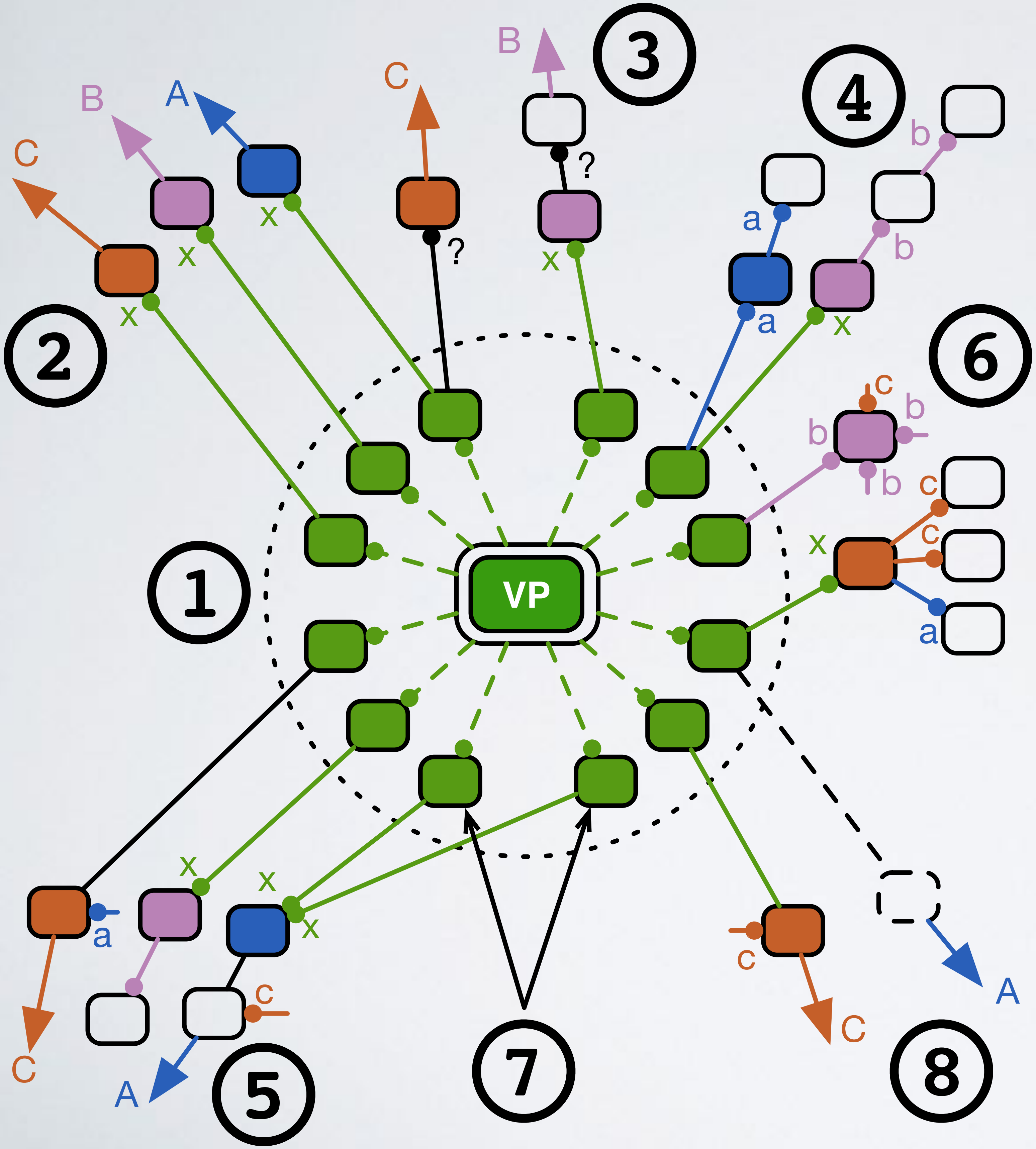
Approach to Border Mapping (3)



Infer Border Routers

bdrmap analyses raw topology data to infer border routers

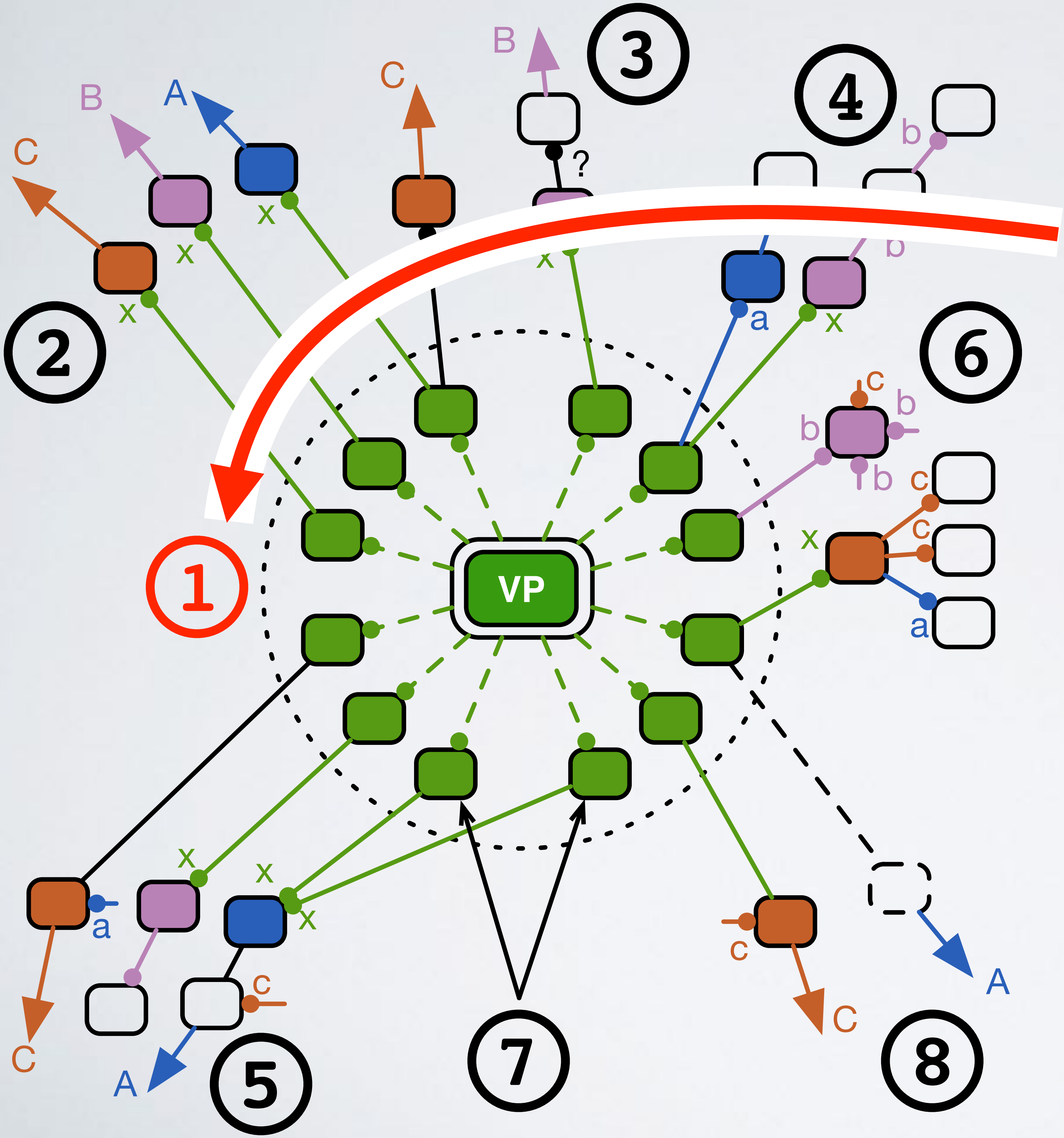
Heuristic Overview



Set of heuristics that reverse-engineer operator and router behaviors

Applied in the order presented based on constraints available

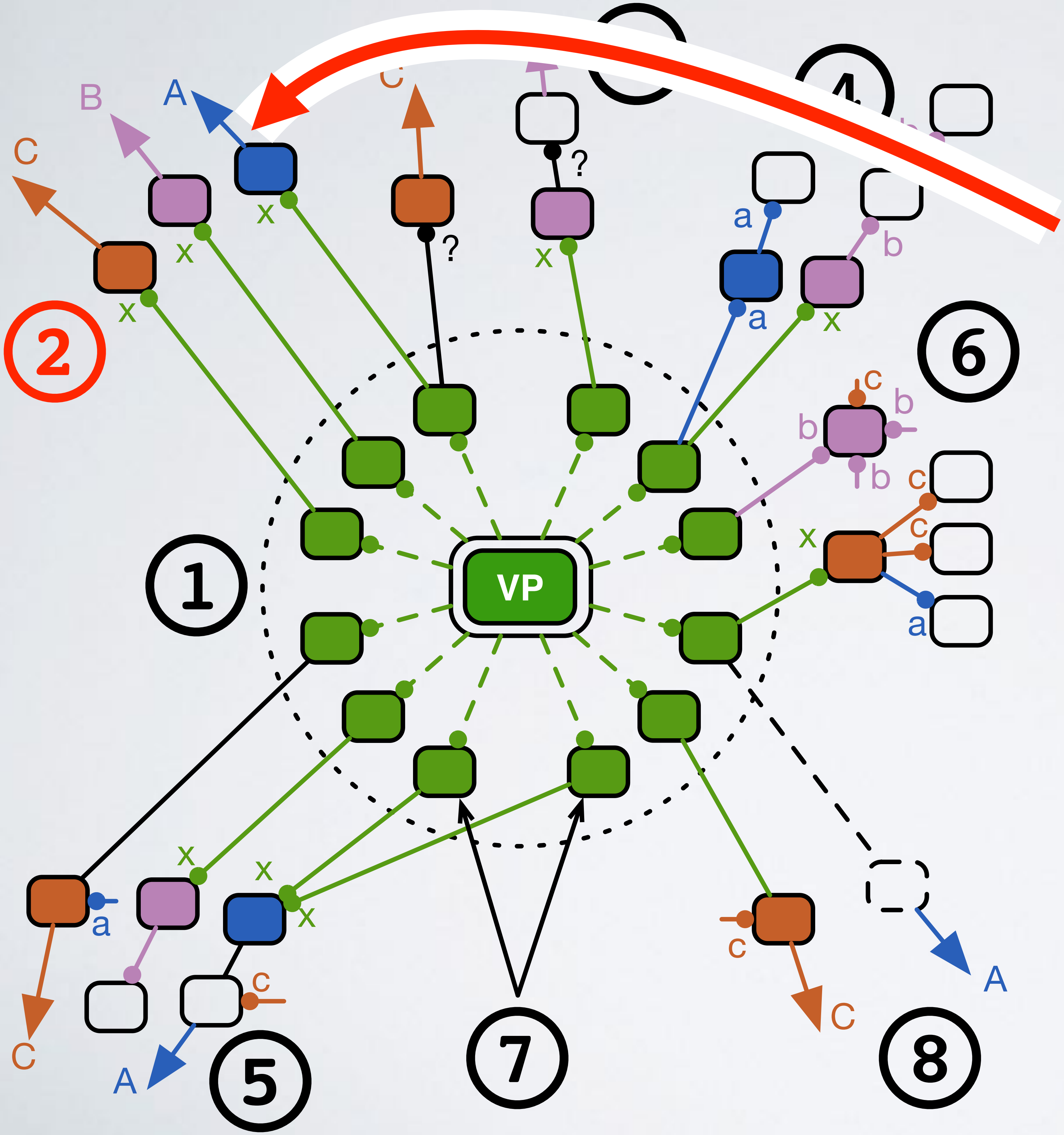
Heuristic Overview



Infer if the router is operated by the **network hosting the VP.**

Other routers must be operated by a neighbor AS

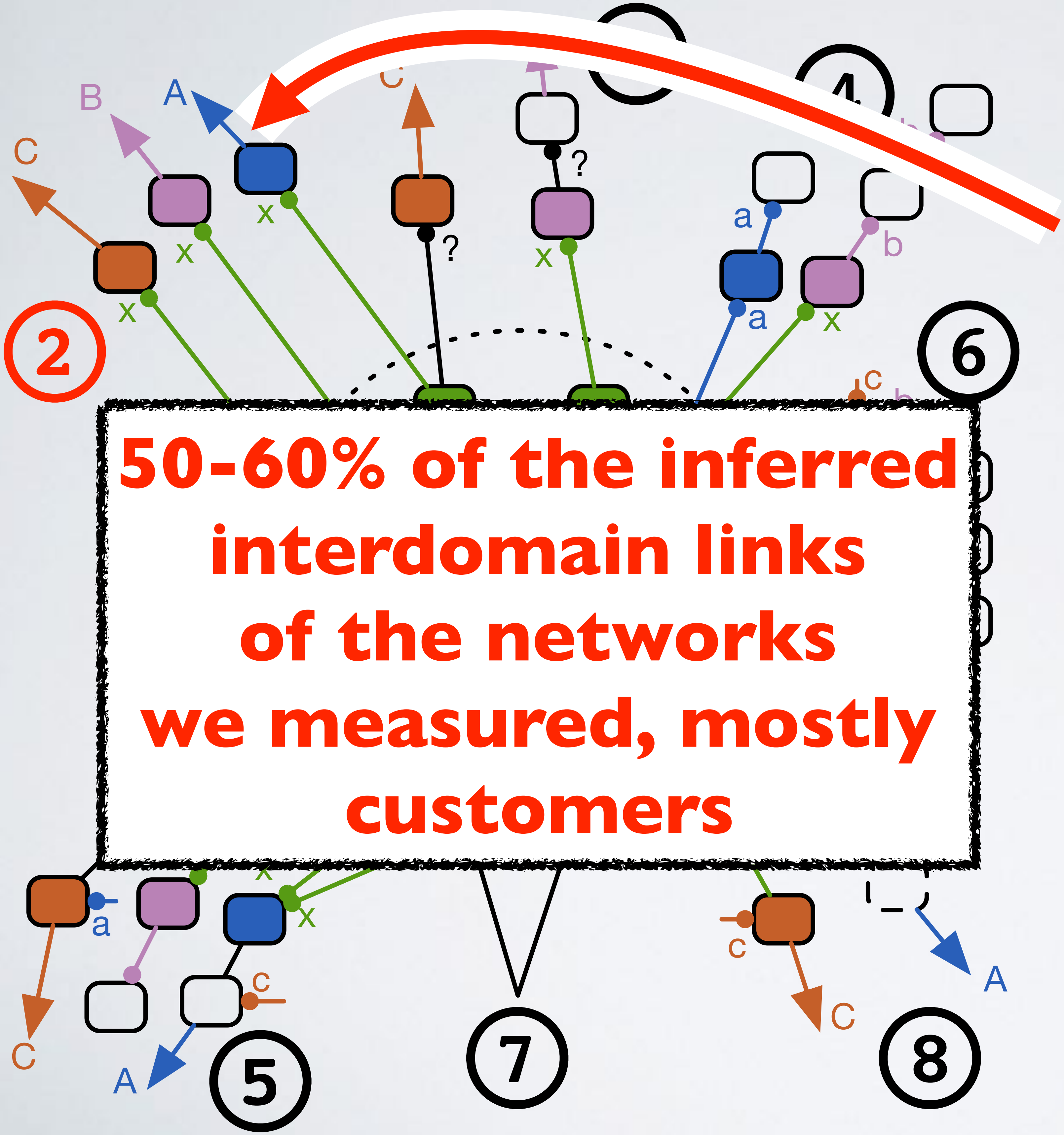
Heuristic Overview



Sometimes we do not observe other topology after a neighbor router.

We can only reason about ownership using the **destination ASes probed** where we observed the router

Heuristic Overview

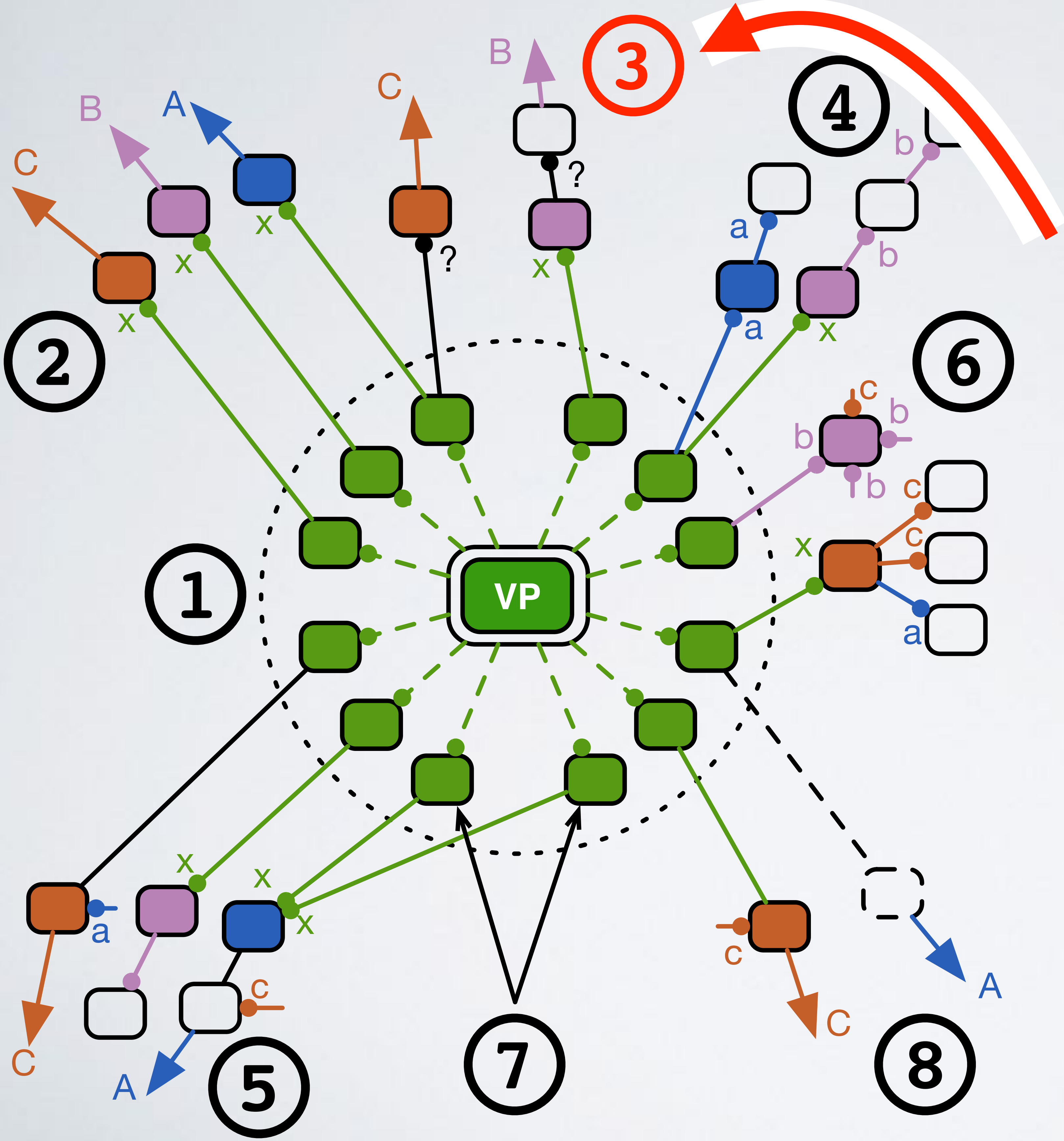


50-60% of the inferred interdomain links of the networks we measured, mostly customers

Sometimes we do not observe other topology after a neighbor router.

We can only reason about ownership using the **destination ASes probed** where we observed the router

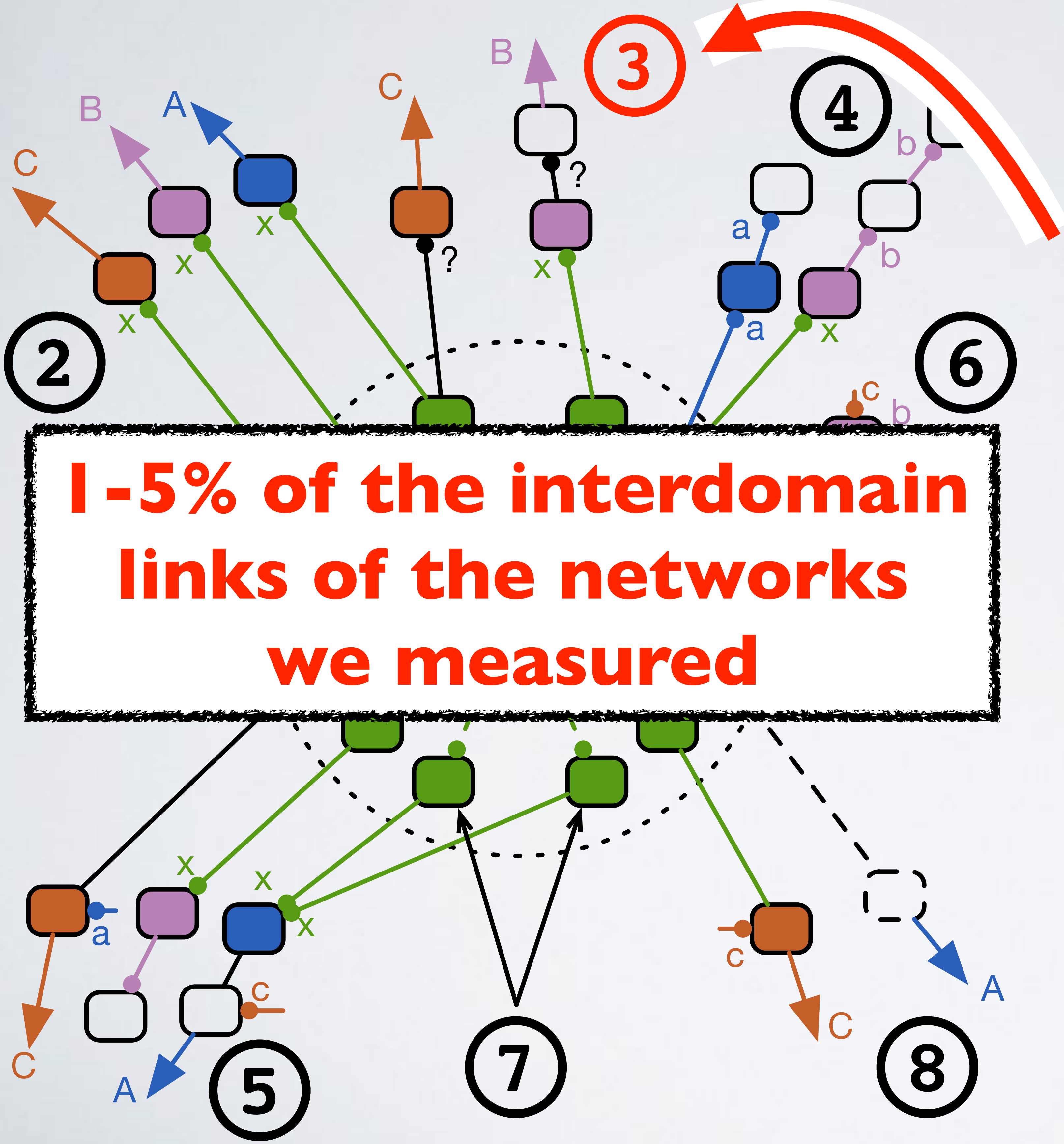
Heuristic Overview



Some routers only respond with an IP address **not routed in BGP**.

We can only reason about ownership using subsequent IPs in the path that are routed, or the destination ASes probed for paths where we observed the router

Heuristic Overview

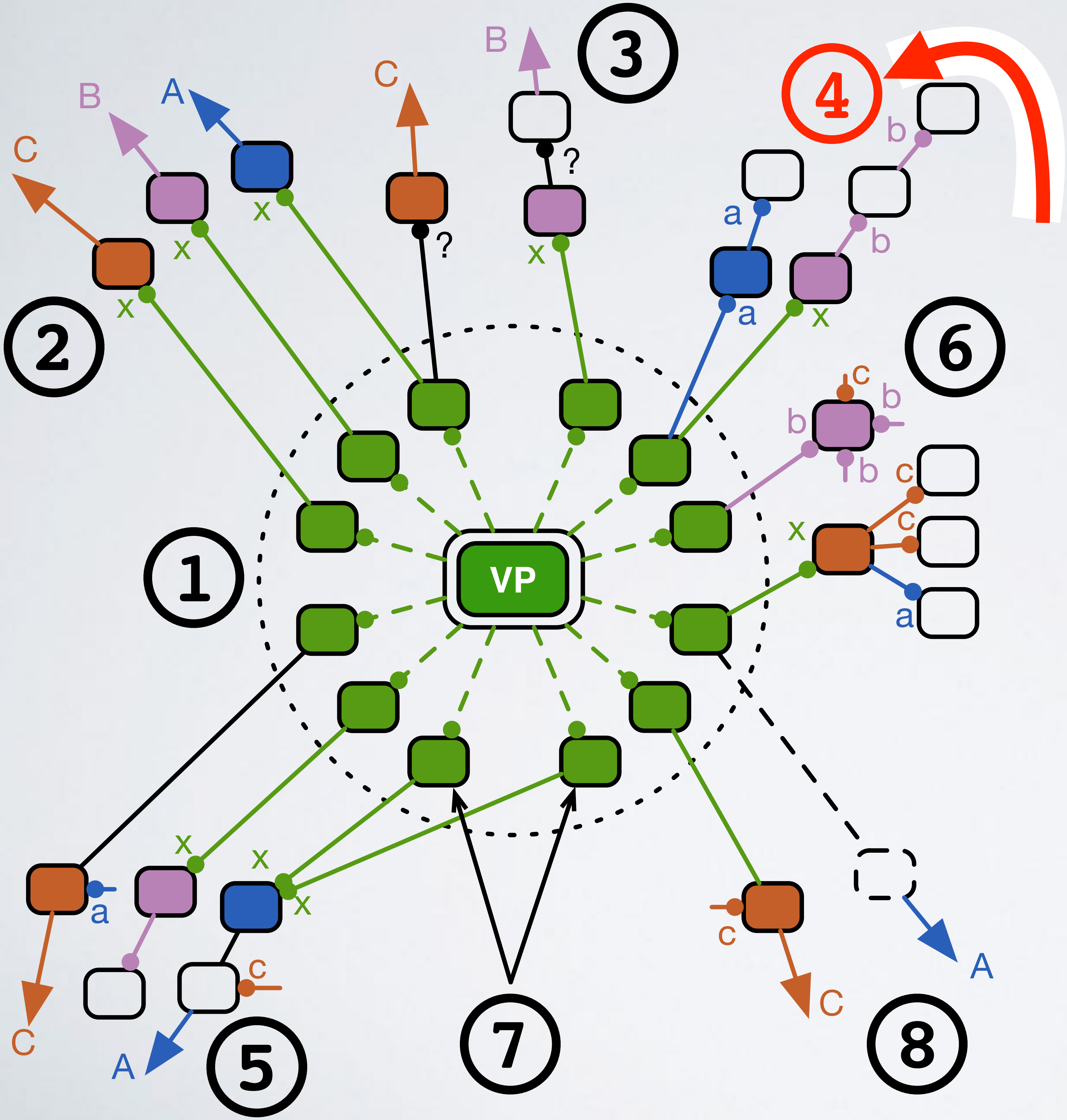


1-5% of the interdomain links of the networks we measured

Some routers only respond with an IP address **not routed in BGP**.

We can only reason about ownership using subsequent IPs in the path that are routed, or the destination ASes probed for paths where we observed the router

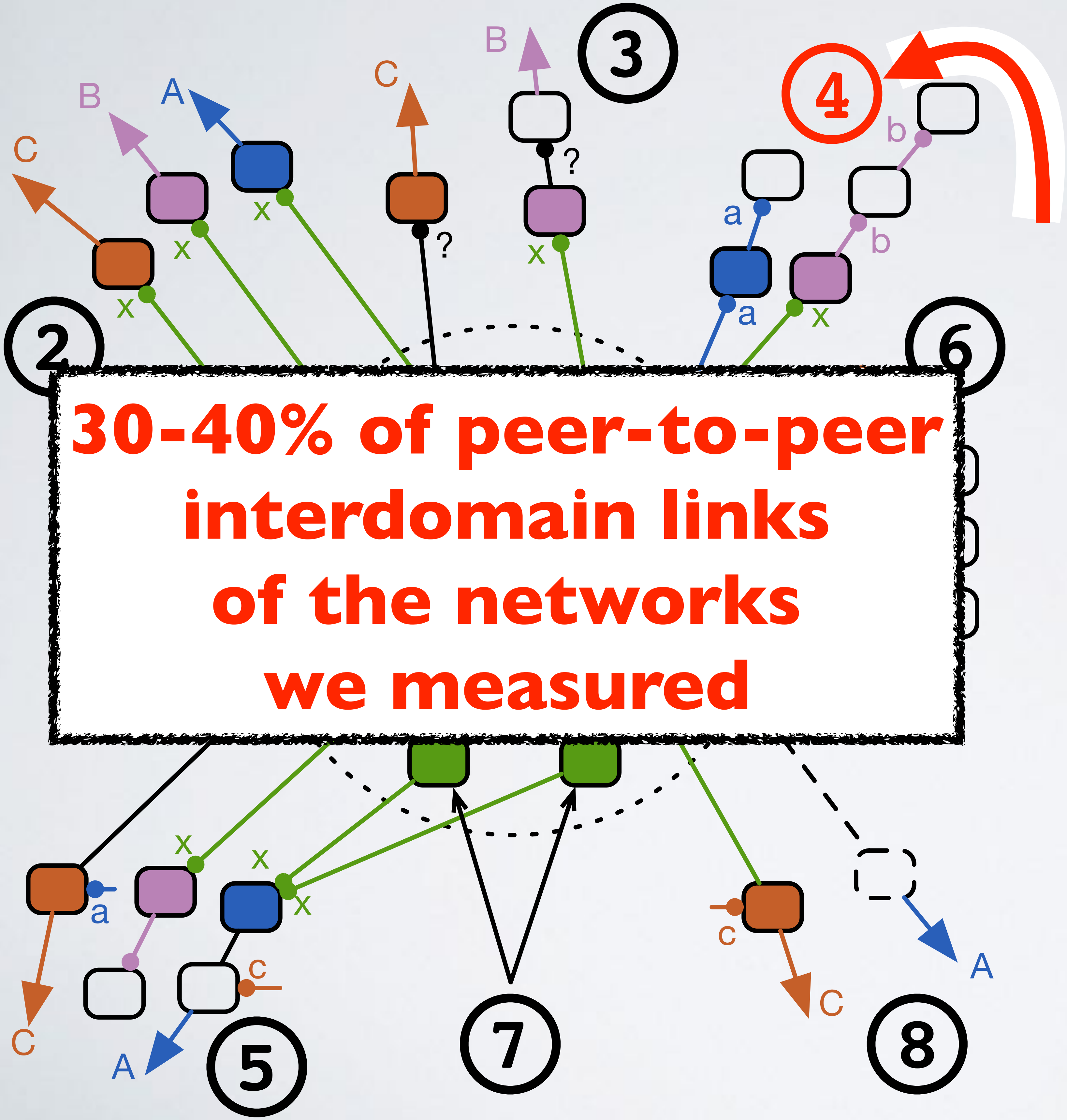
Heuristic Overview



Reason about ownership using IP-AS mappings where the **same AS on two consecutive routers**

We are unlikely to observe two consecutive third-party IP addresses in a path

Heuristic Overview

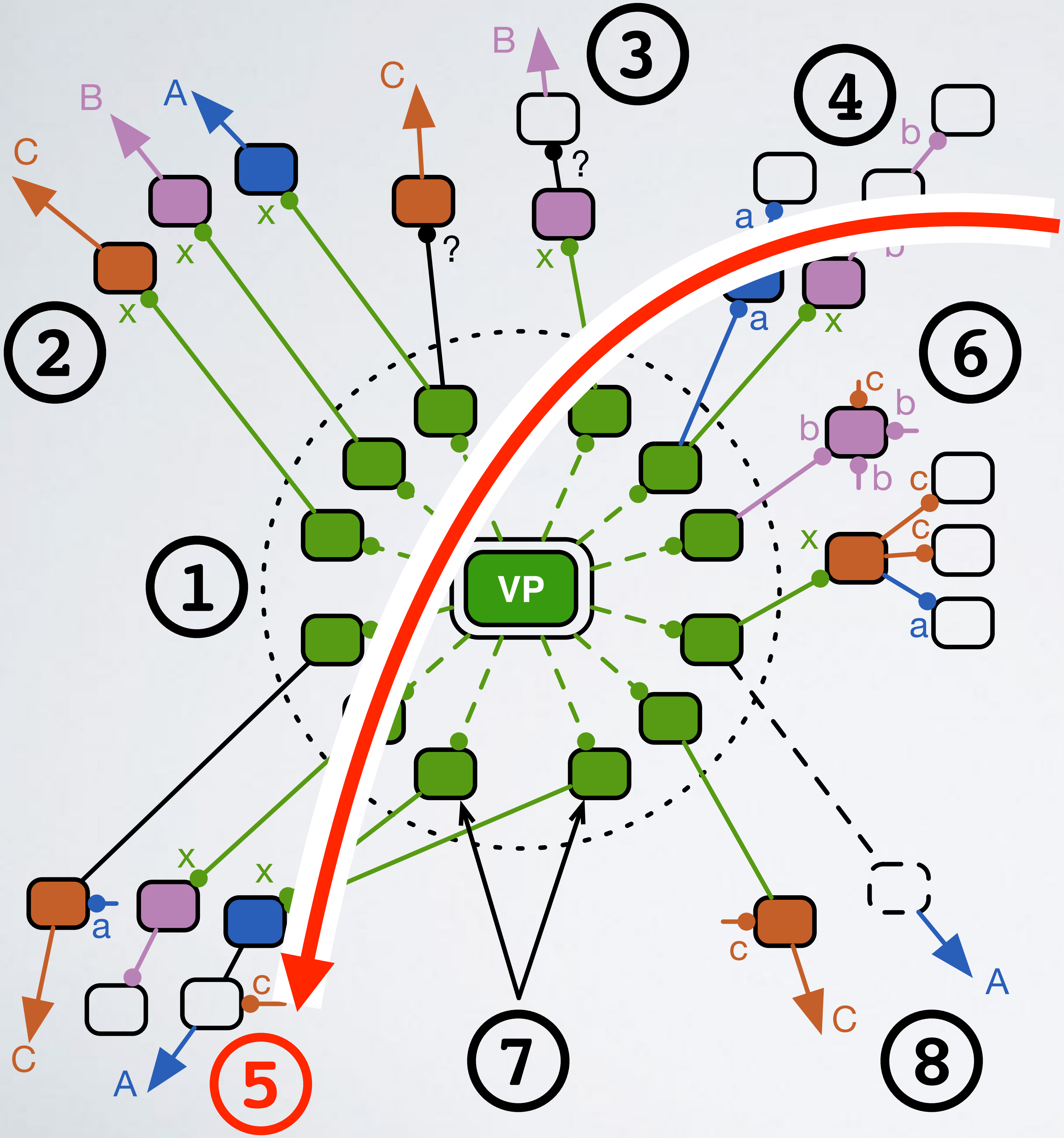


30-40% of peer-to-peer interdomain links of the networks we measured

Reason about ownership using IP-AS mappings where the **same AS on two consecutive routers**

We are unlikely to observe two consecutive third-party IP addresses in a path

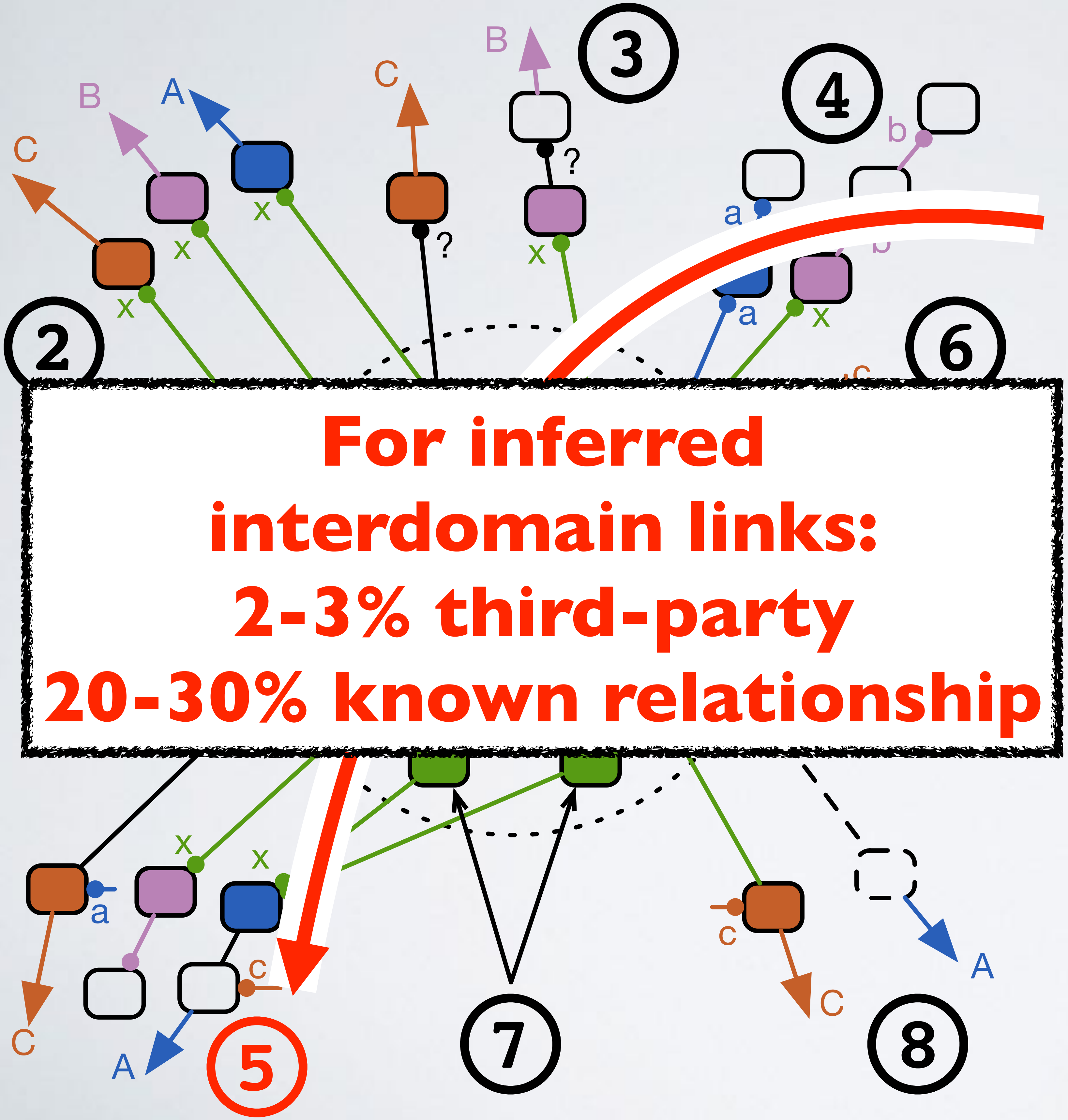
Heuristic Overview



Reason about ownership **using AS relationships** and IP-AS mappings.

We infer owners of routers that responded using a third-party address, as well as known peers and customers

Heuristic Overview

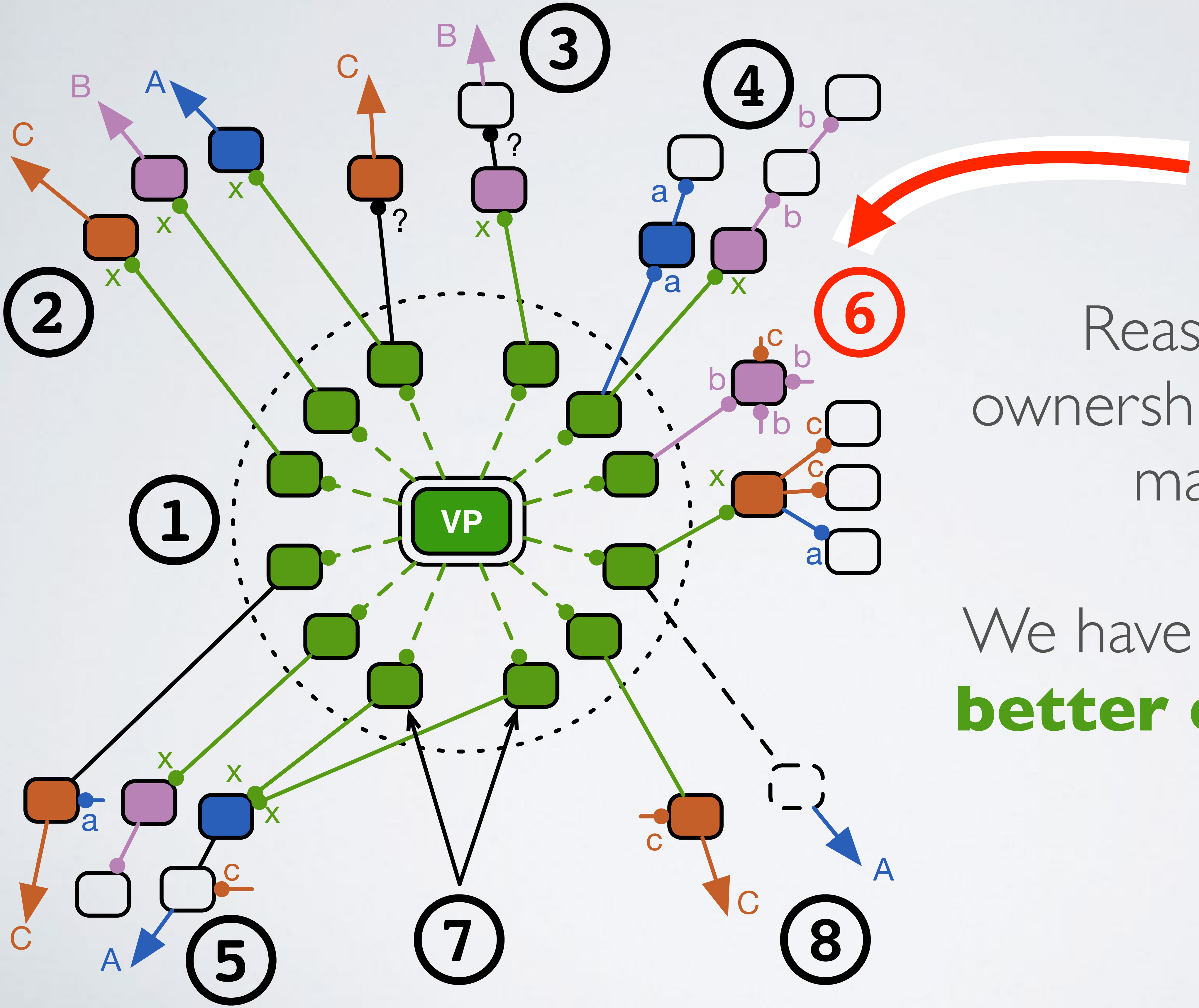


For inferred interdomain links:
2-3% third-party
20-30% known relationship

Reason about ownership **using AS relationships** and IP-AS mappings.

We infer owners of routers that responded using a third-party address, as well as known peers and customers

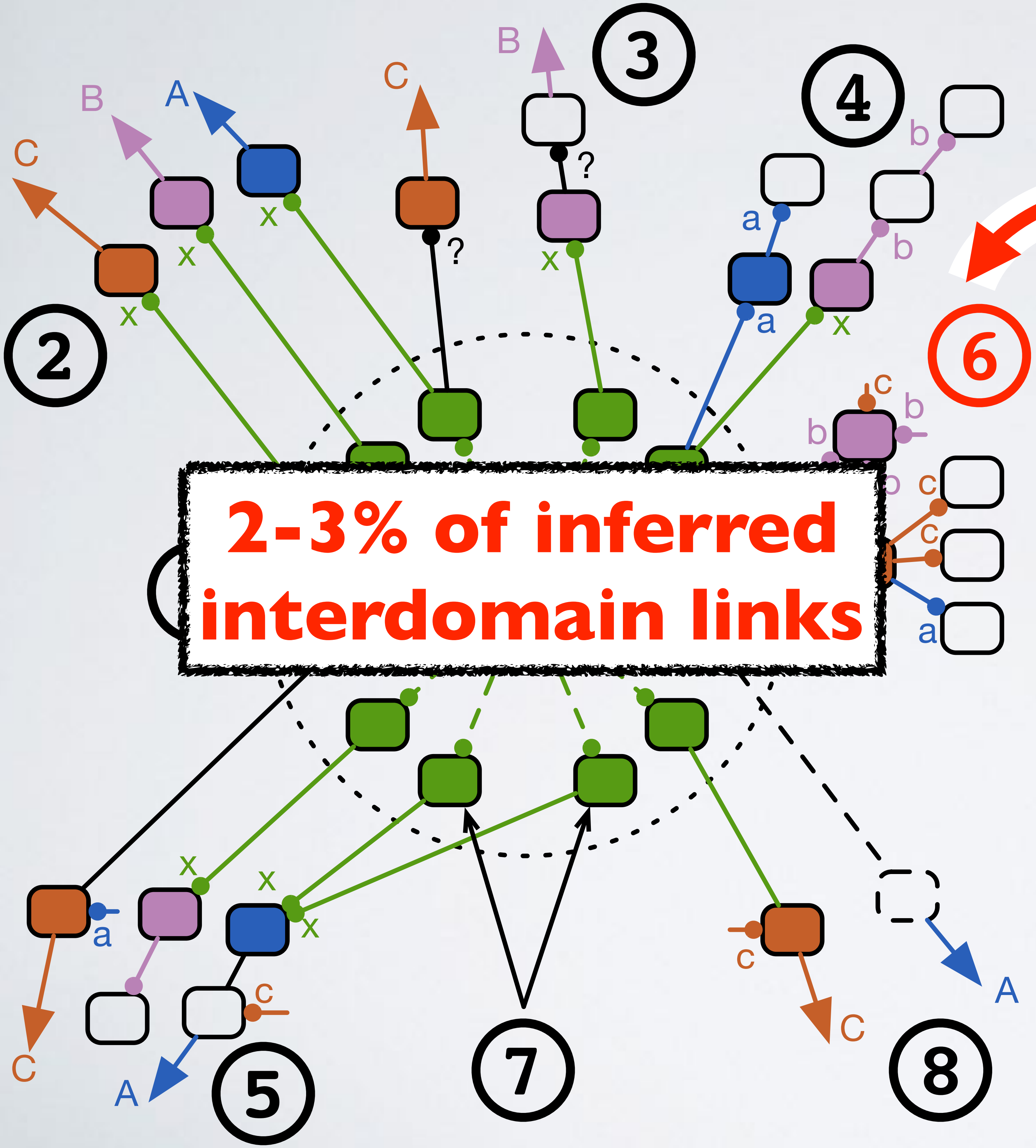
Heuristic Overview



Reason about ownership using IP-AS mappings.

We have **exhausted better constraints**

Heuristic Overview

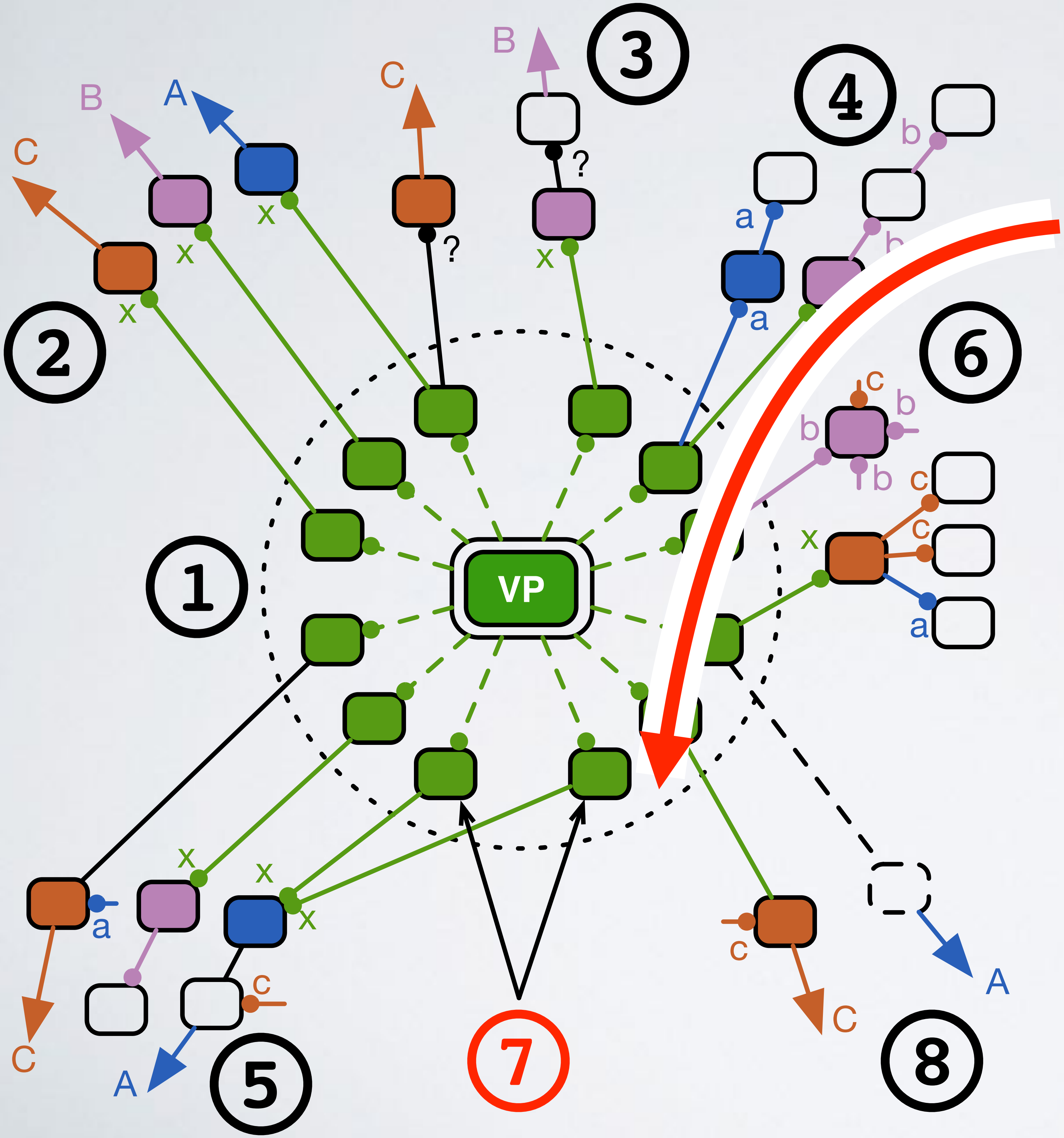


2-3% of inferred interdomain links

Reason about ownership using IP-AS mappings.

We have **exhausted better constraints**

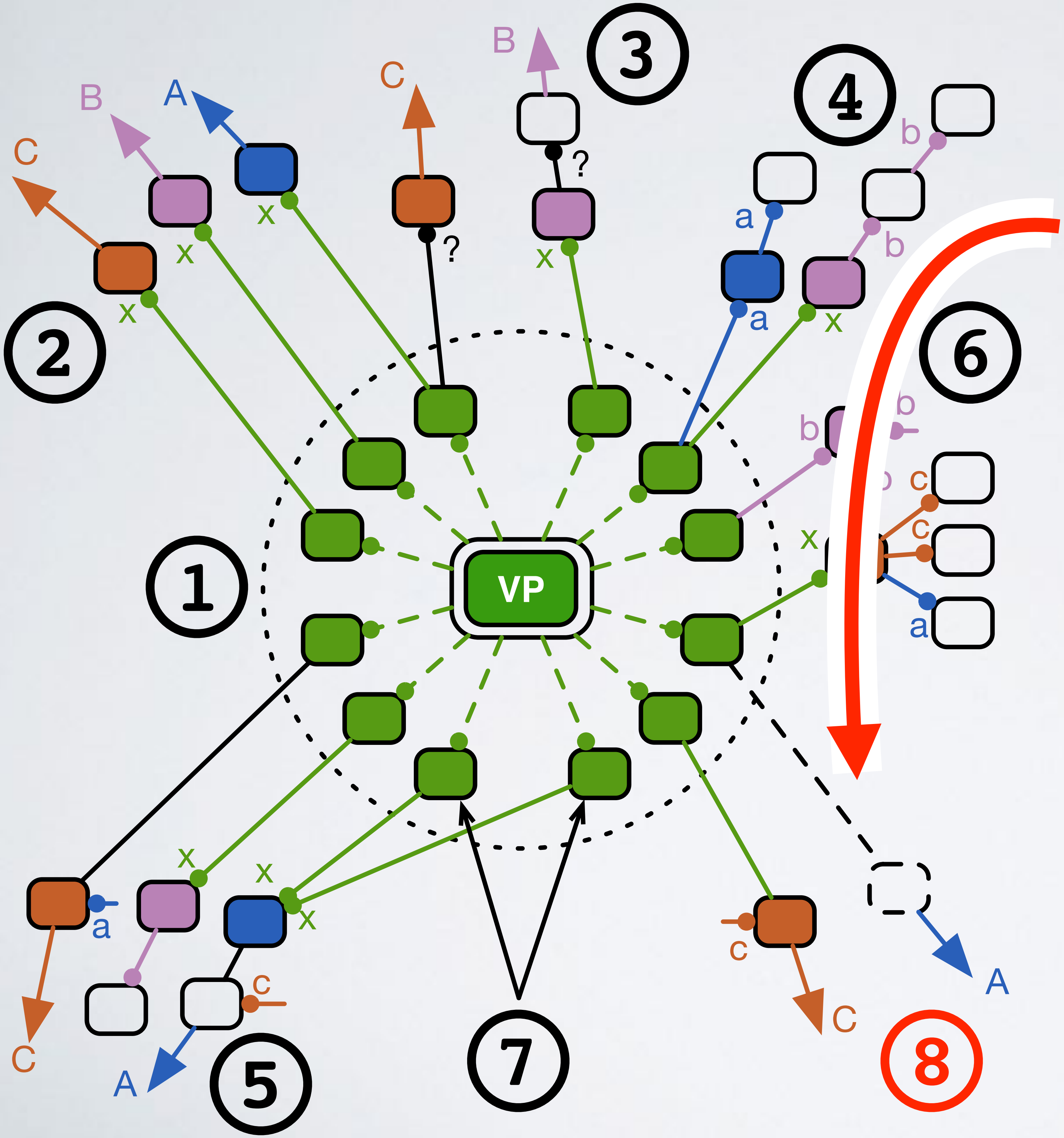
Heuristic Overview



Infer additional aliases for routers operated by the network hosting the VP

A single neighbor router is likely connected to a single VP router with a point-to-point link

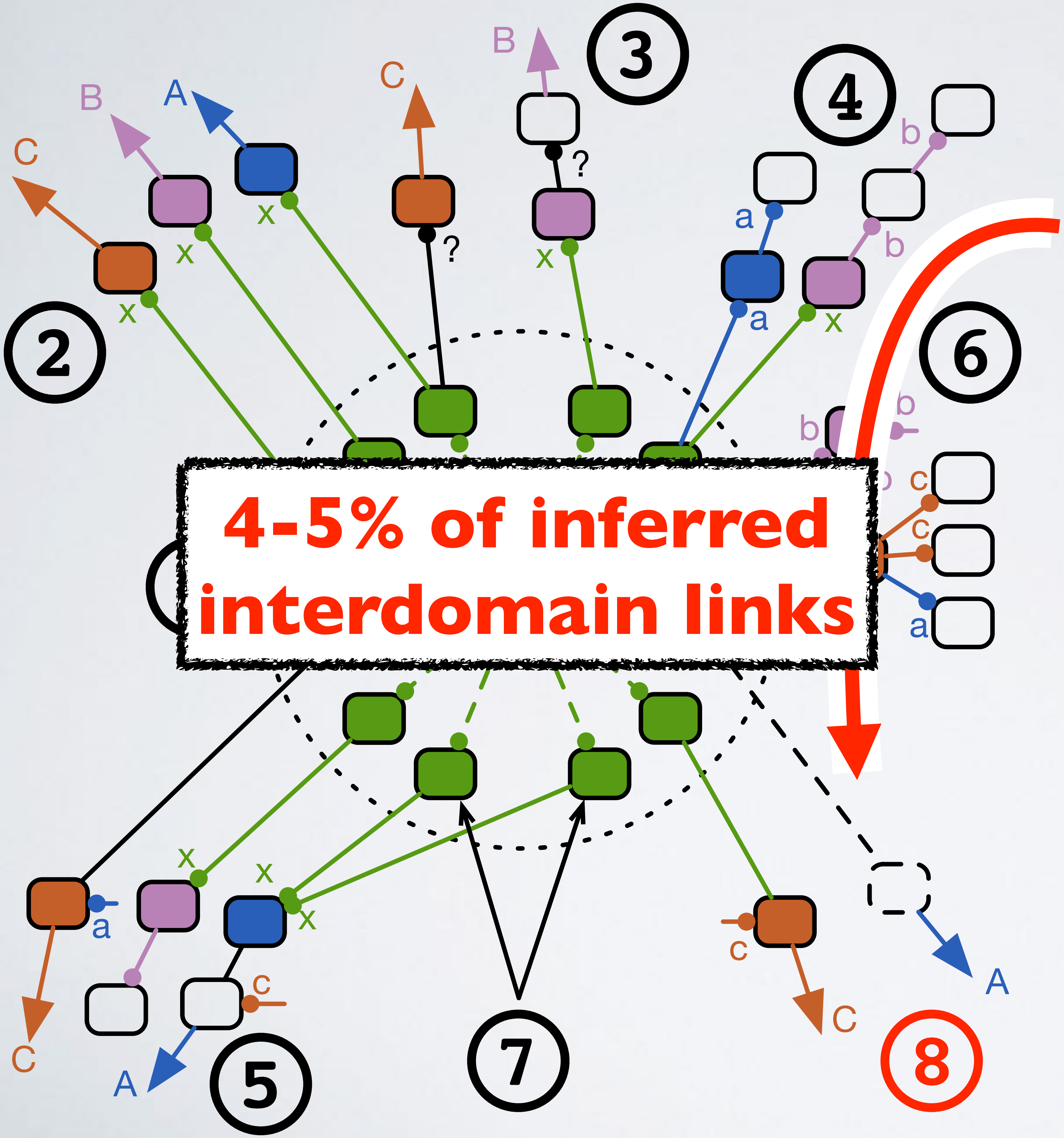
Heuristic Overview



Infer presence of **silent neighbor routers**, and owners of routers that responded without ICMP time exceeded messages.

To avoid false link inferences, we only infer presence of a neighbor when we know link exists through BGP

Heuristic Overview



4-5% of inferred interdomain links

Infer presence of **silent neighbor routers**, and owners of routers that responded without ICMP time exceeded messages.

To avoid false link inferences, we only infer presence of a neighbor when we know link exists through BGP

Coverage + Validation

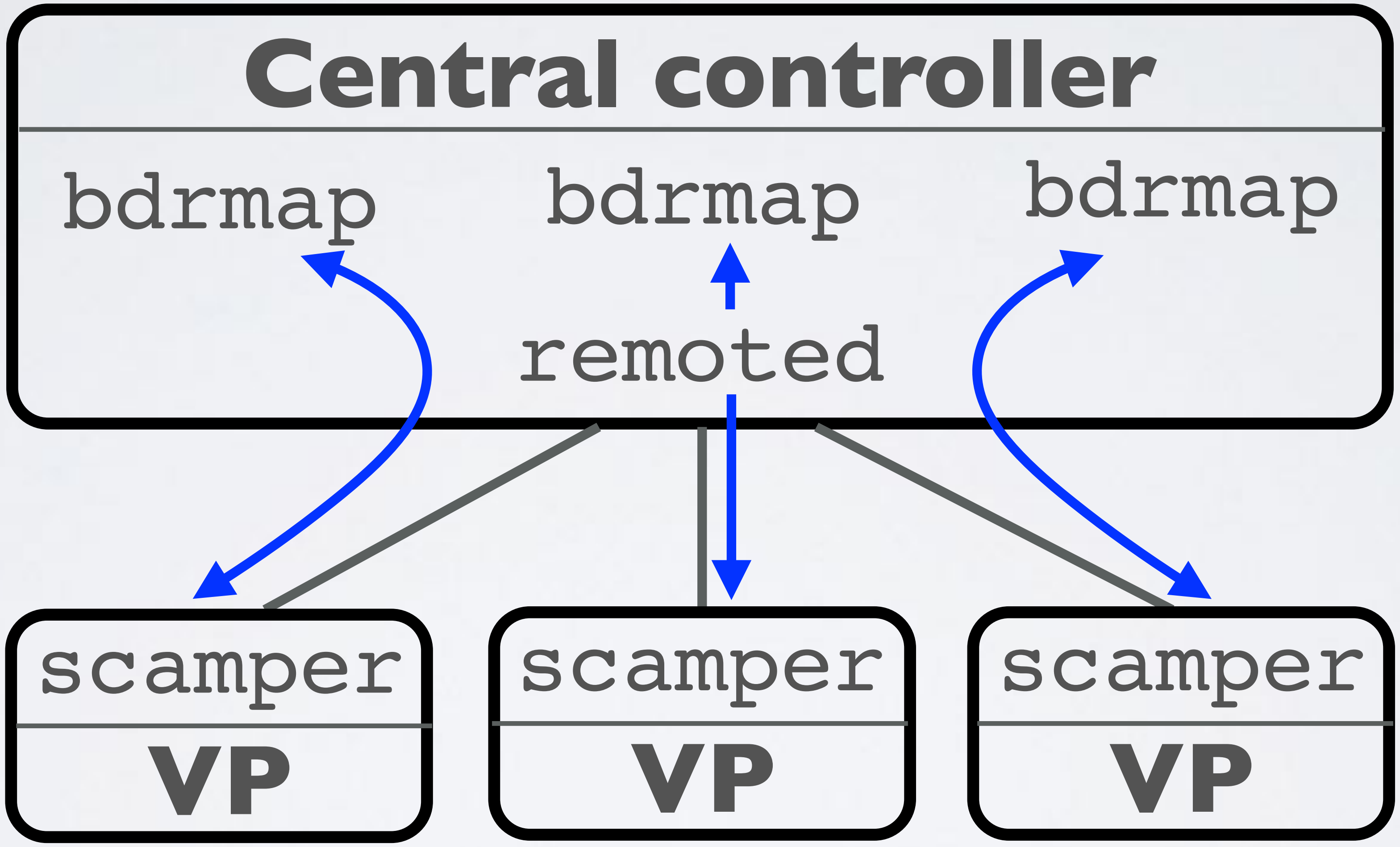
- Overall, 92-97% coverage of VP-network links in BGP
- We accurately find additional VP-network links not in BGP
- **Validation**: contacted 10 networks, received validation for 4
  - R&E network: 131 of 136 links correct (96.3%)
  - Large access network: 97.0% - 98.9% correct, depend on VP
  - Tier-1 network: 2584 of 2650 links correct (97.5%)
  - Small access network: 283 of 293 links (96.6%)

Limitations

- bdrmap **depends on observing topology at interconnection** points to assemble useful constraints
  - not always possible as traceroute may observe other paths
- Still restricted by limitations of what is possible with traceroute
- Alias resolution techniques are not always able to map IP addresses to the same underlying router

Using low-resource VPs

We extended scamper to be remote controlled.
Algorithm state and collected data can be kept off the device.



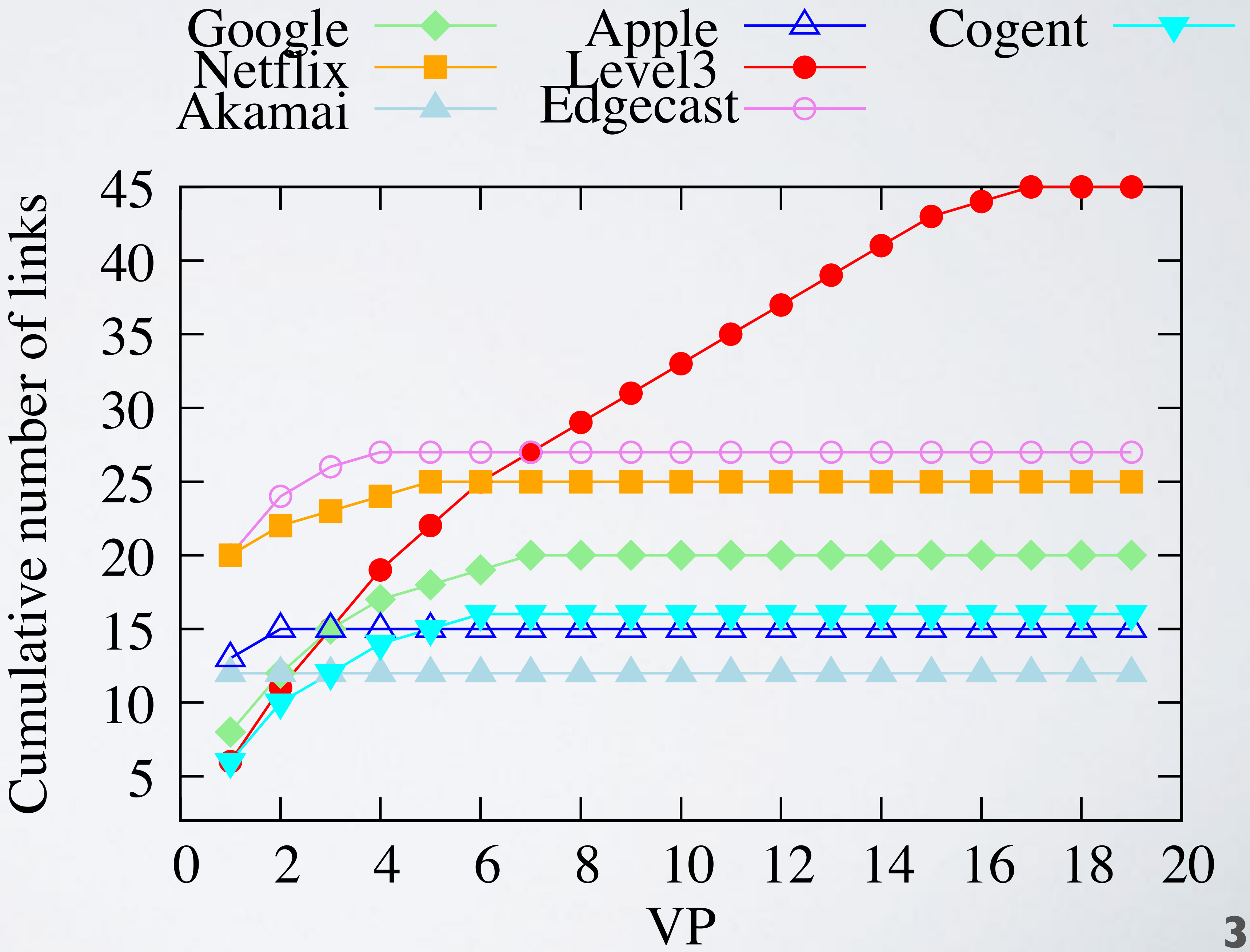
SamKnows / BISmark: 450Mhz MIPS CPU,
64-128MB RAM, 16MB flash storage

Interconnection Insights

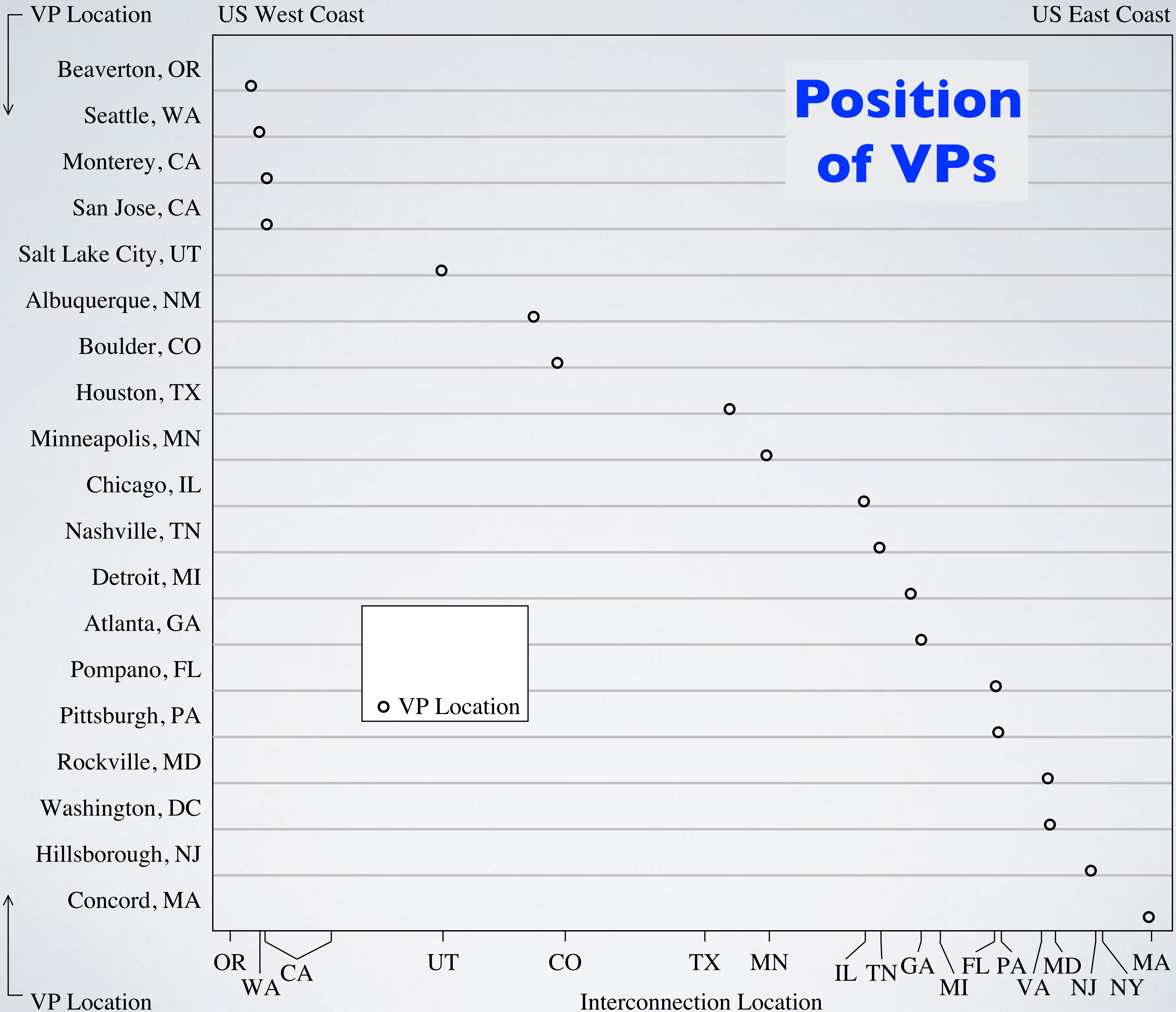
- We used 19 geographically distributed VPs inside Comcast to map router-level interdomain connectivity of Comcast in January 2016

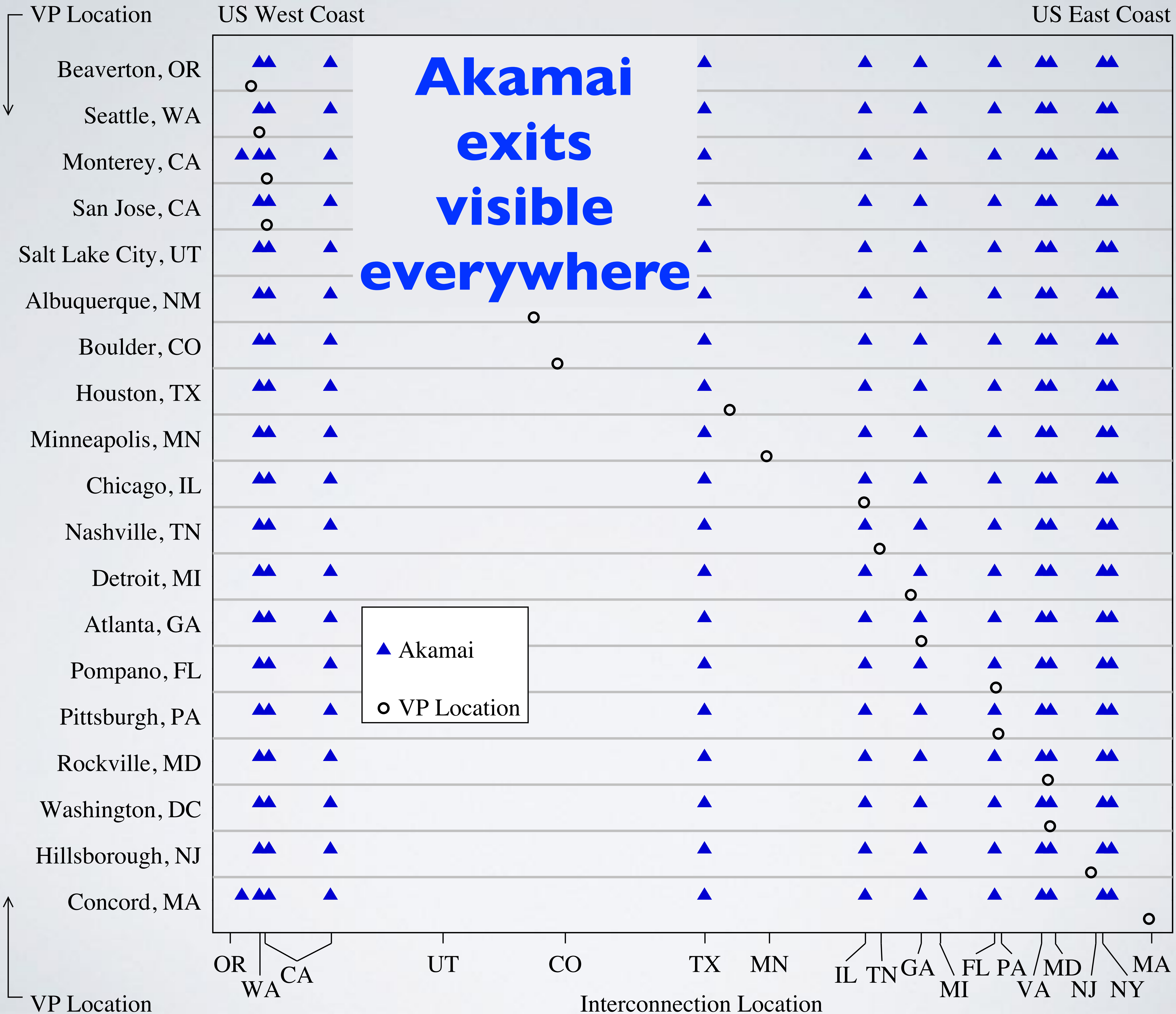
Max interconnection density: 45 router links with Level3.

Required 17 VPs to observe them all

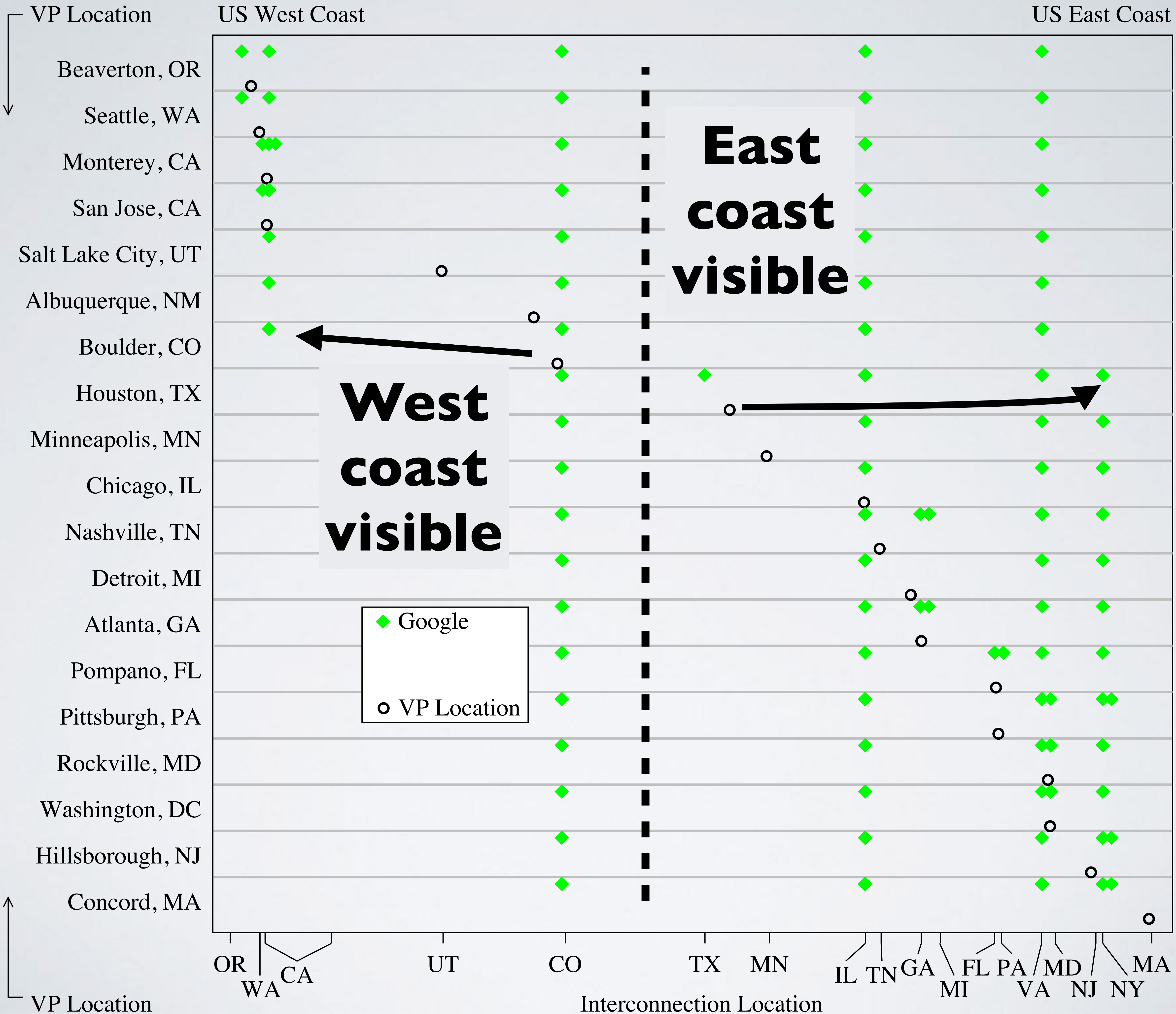


Position of VPs



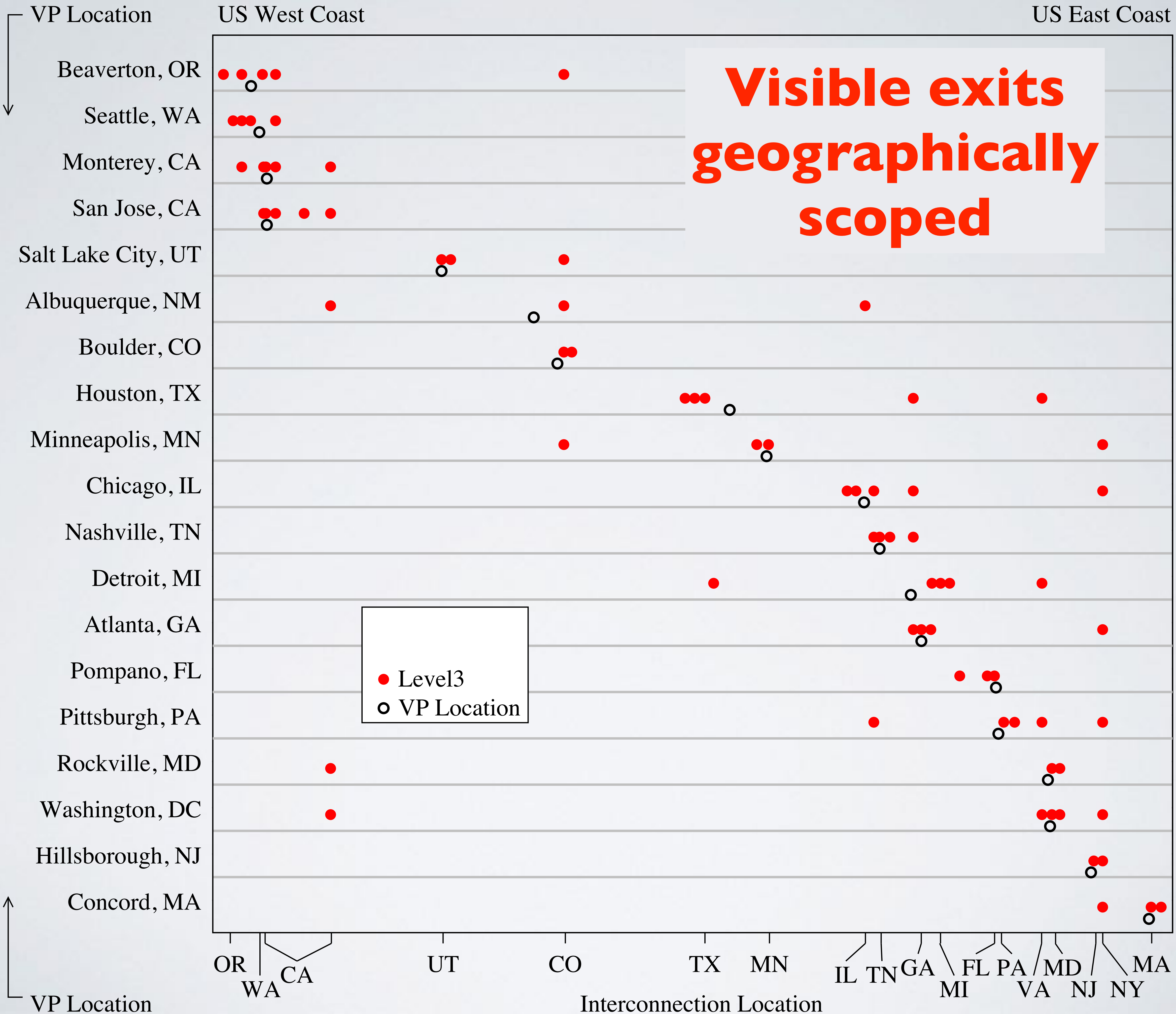


Akamai exits visible everywhere



East coast visible

West coast visible



Visible exits geographically scoped

Summary

- We used **active measurement** techniques to build a router-level map focused on **router ownership inference** for interdomain links for a network hosting a VP
- We **developed and validated heuristics** to distinguish VP-routers from neighbor routers, and to infer the operator of neighbor routers
- We used our system to **investigate modern interconnection arrangements**
- We **publicly release** our source code implementation

https://www.caida.org/tools/measurement/scamper/

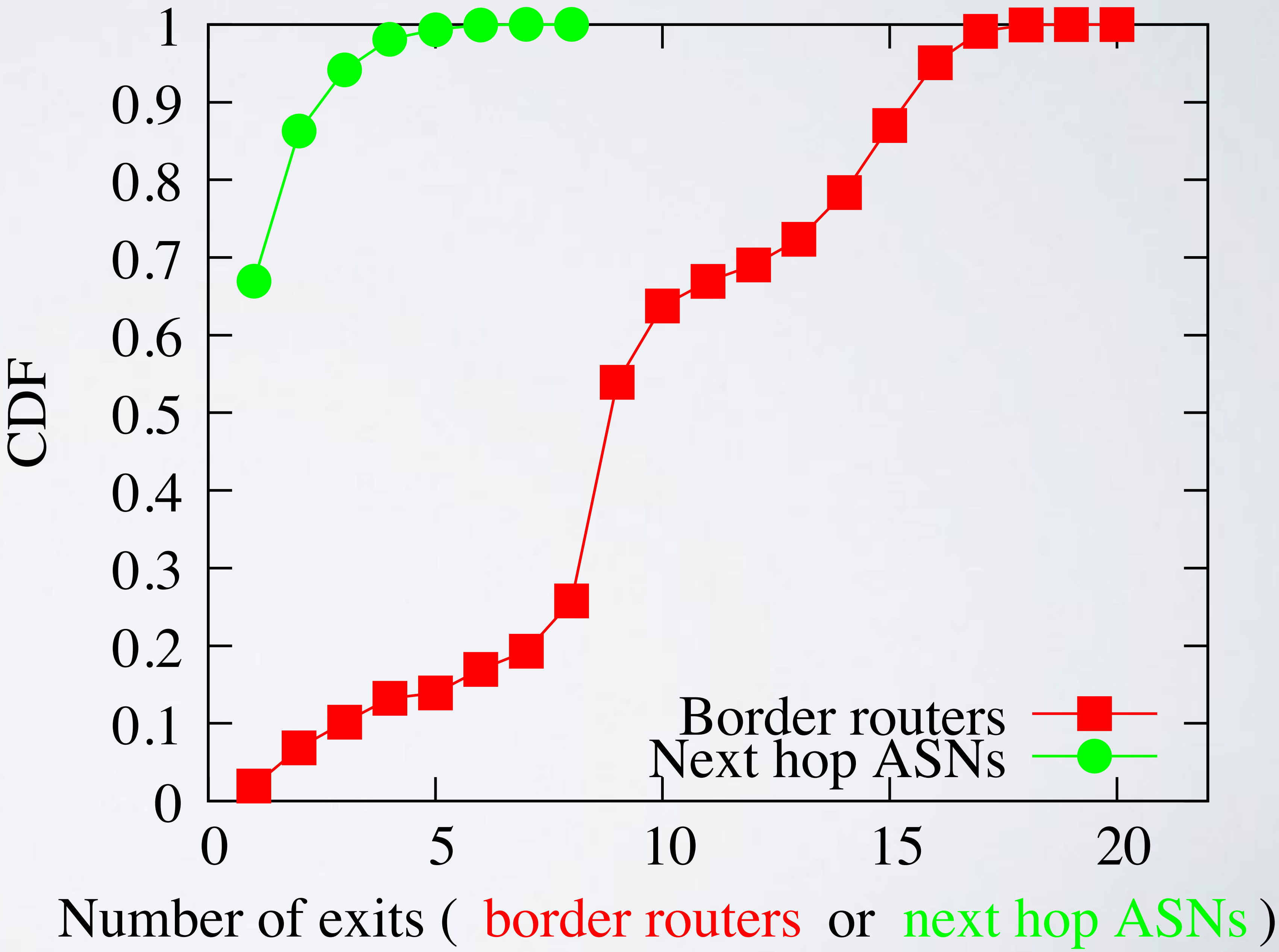
BACKUP SLIDES

Interconnection Insights

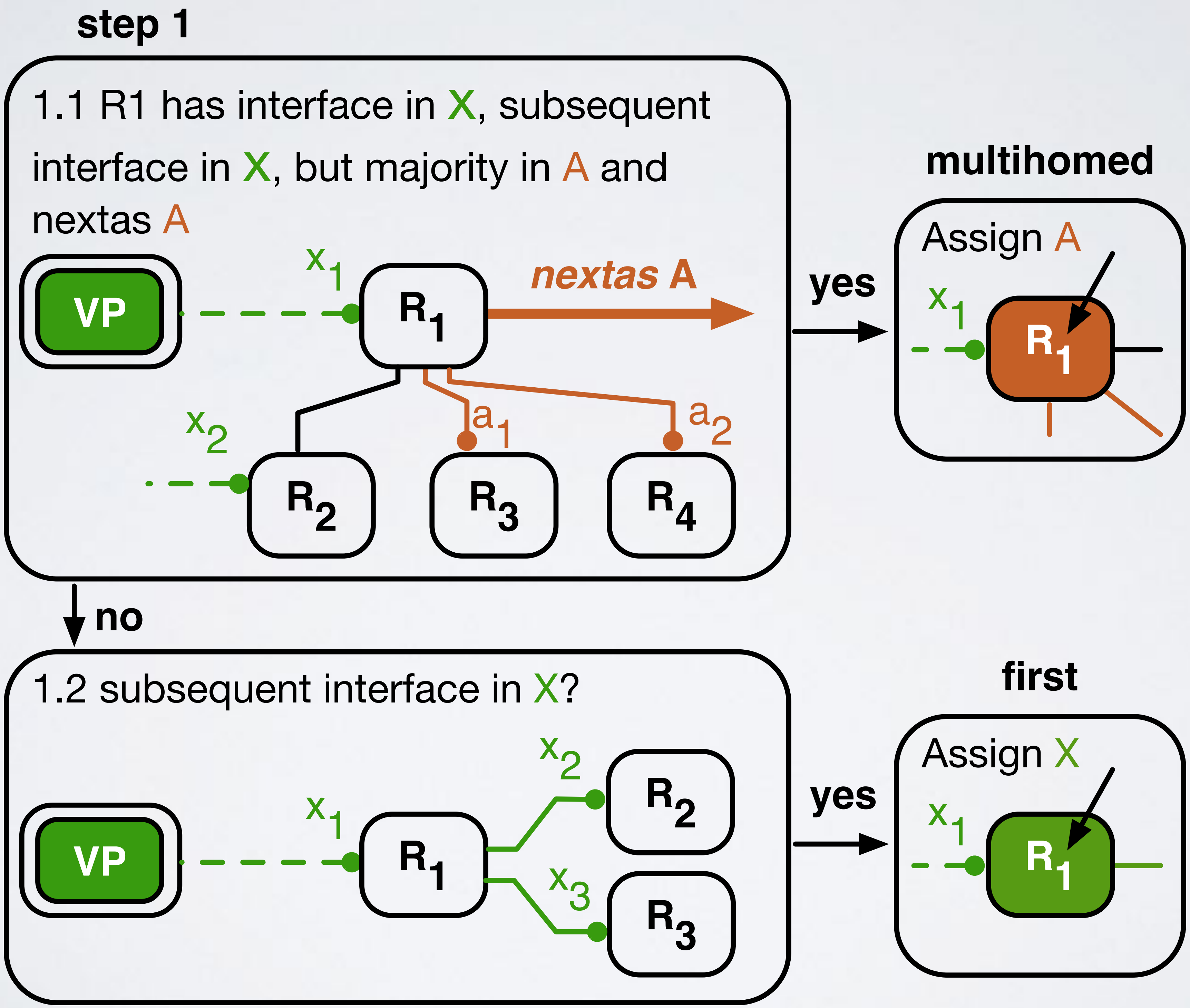
- We used 19 geographically distributed VPs inside Comcast to map router-level interdomain connectivity of Comcast in January 2016

Fewer than 2% of prefixes left network via the same border for each VP.

For 73% of prefixes, we observed 5-15 distinct border routers, and 13% of prefixes had more than 15 exits.



Infer routers operated by the network hosting the VP



step 1

1.1 R1 has interface in X, subsequent interface in X, but majority in A and nextas A

no

1.2 subsequent interface in X?

yes

multihomed

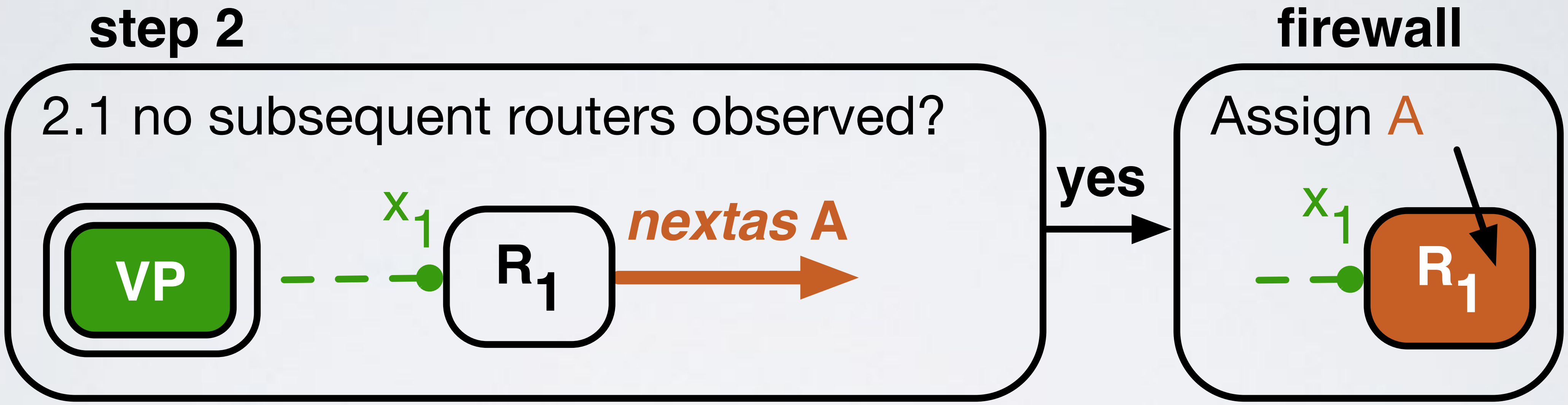
Assign A

yes

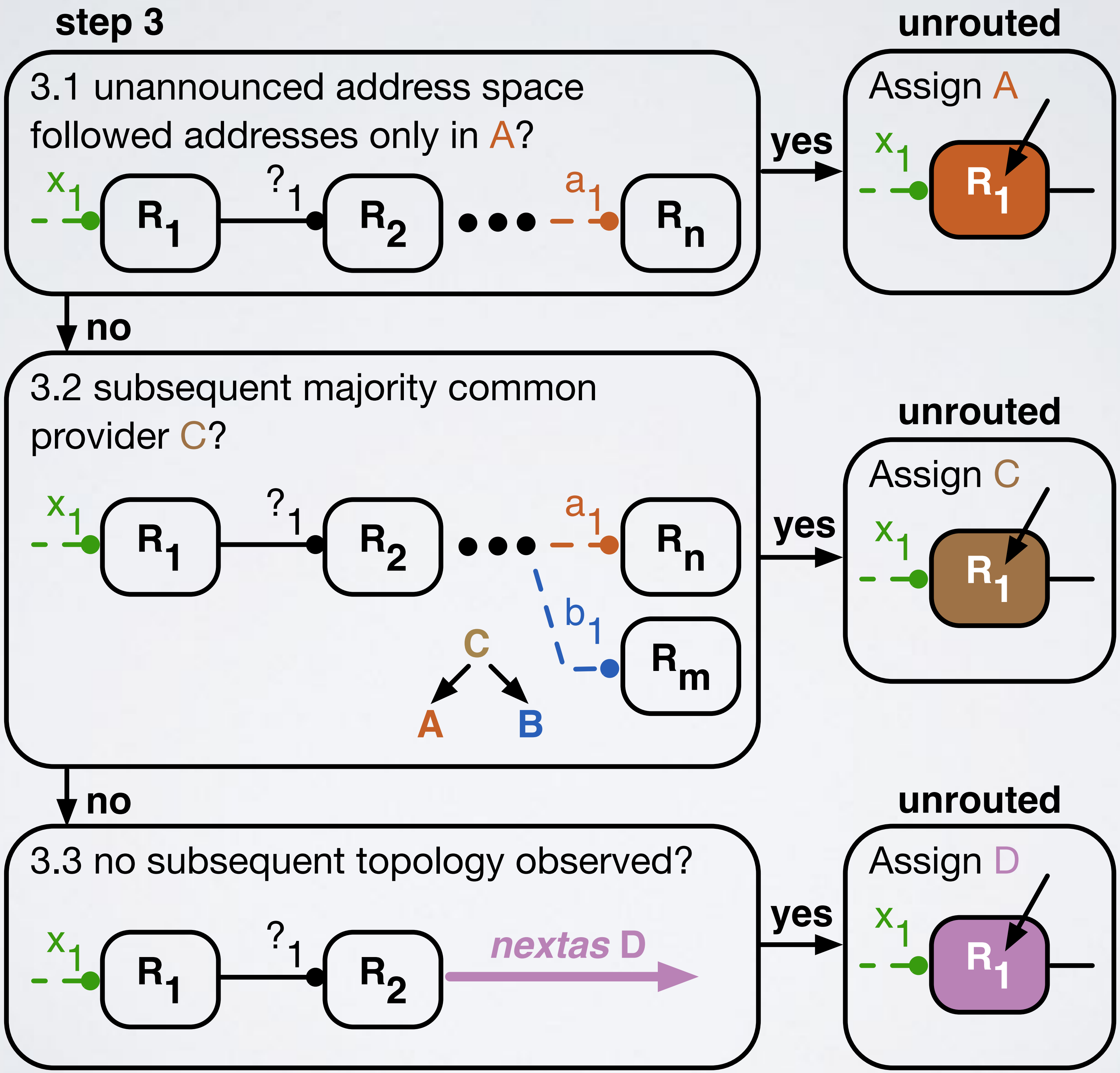
first

Assign X

Inferring owner of neighbor routers with firewalls

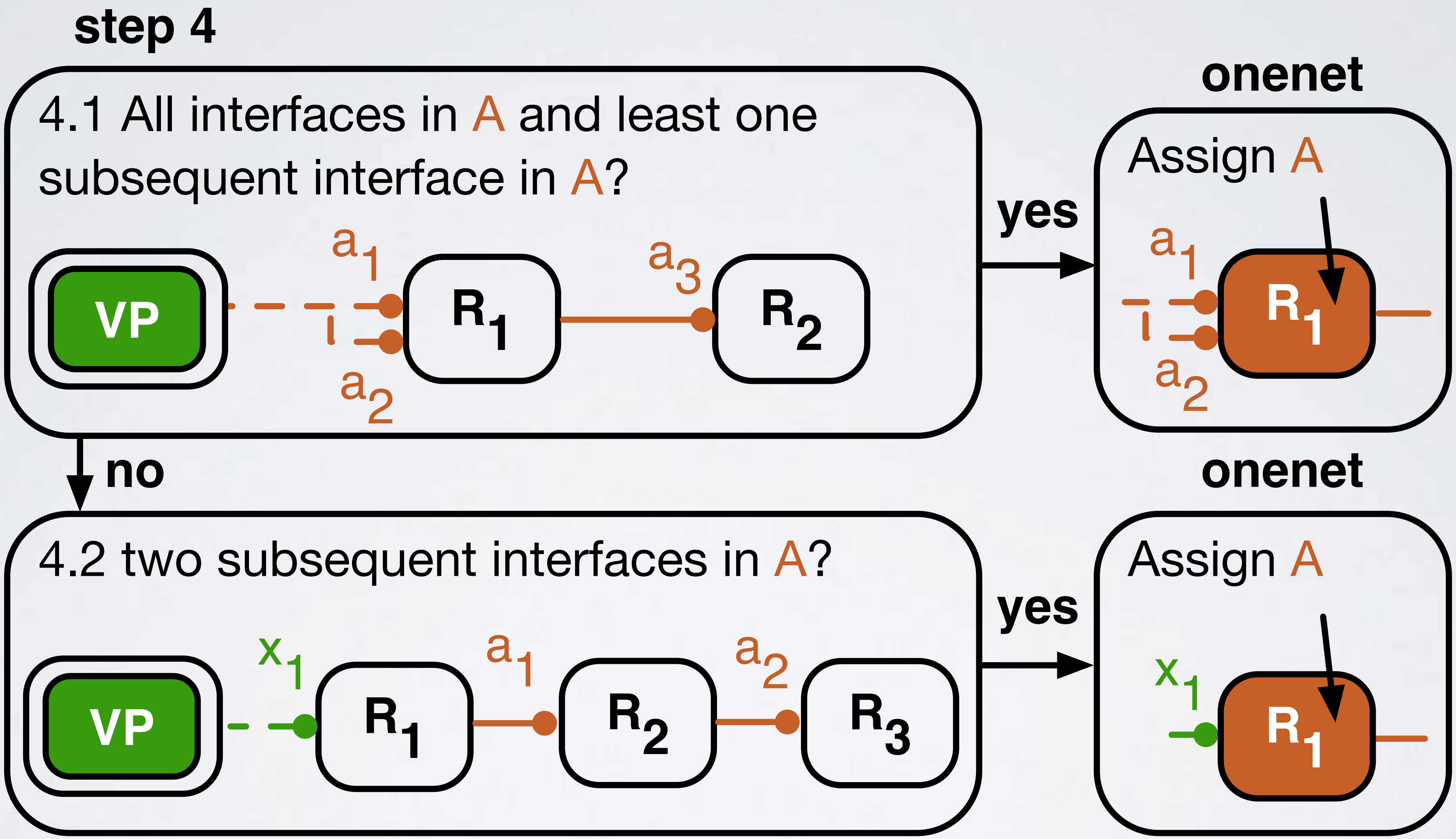


Infer operator of neighbor routers that use unrouted IP



step 3
3.1 unannounced address space followed addresses only in A?
x1
R1
?1
R2
a1
Rn
yes
unrouted
Assign A
x1
R1
no
3.2 subsequent majority common provider C?
x1
R1
?1
R2
a1
Rn
b1
Rm
C
A
B
yes
unrouted
Assign C
x1
R1
no
3.3 no subsequent topology observed?
x1
R1
?1
R2
nextas D
yes
unrouted
Assign D
x1
R1

Use IP-AS mappings to infer operator of neighbor routers



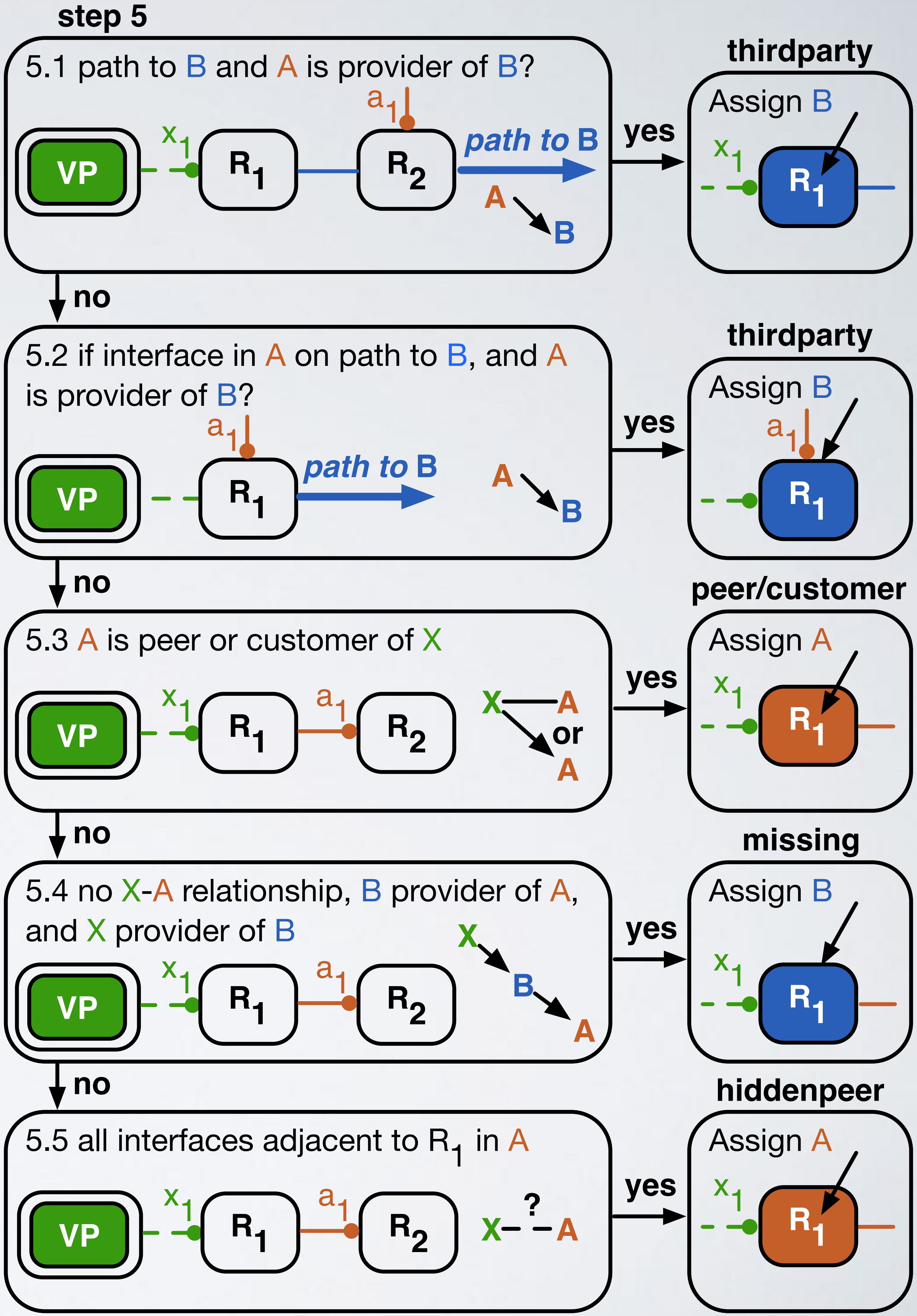
step 4

4.1 All interfaces in A and least one subsequent interface in A?

4.2 two subsequent interfaces in A?

Assign A

Use AS relationship inferences to infer operator of neighbor routers



step 5

5.1 path to B and A is provider of B?

yes → thirdparty: Assign B

no

5.2 if interface in A on path to B, and A is provider of B?

yes → thirdparty: Assign B

no

5.3 A is peer or customer of X

yes → peer/customer: Assign A

no

5.4 no X-A relationship, B provider of A, and X provider of B

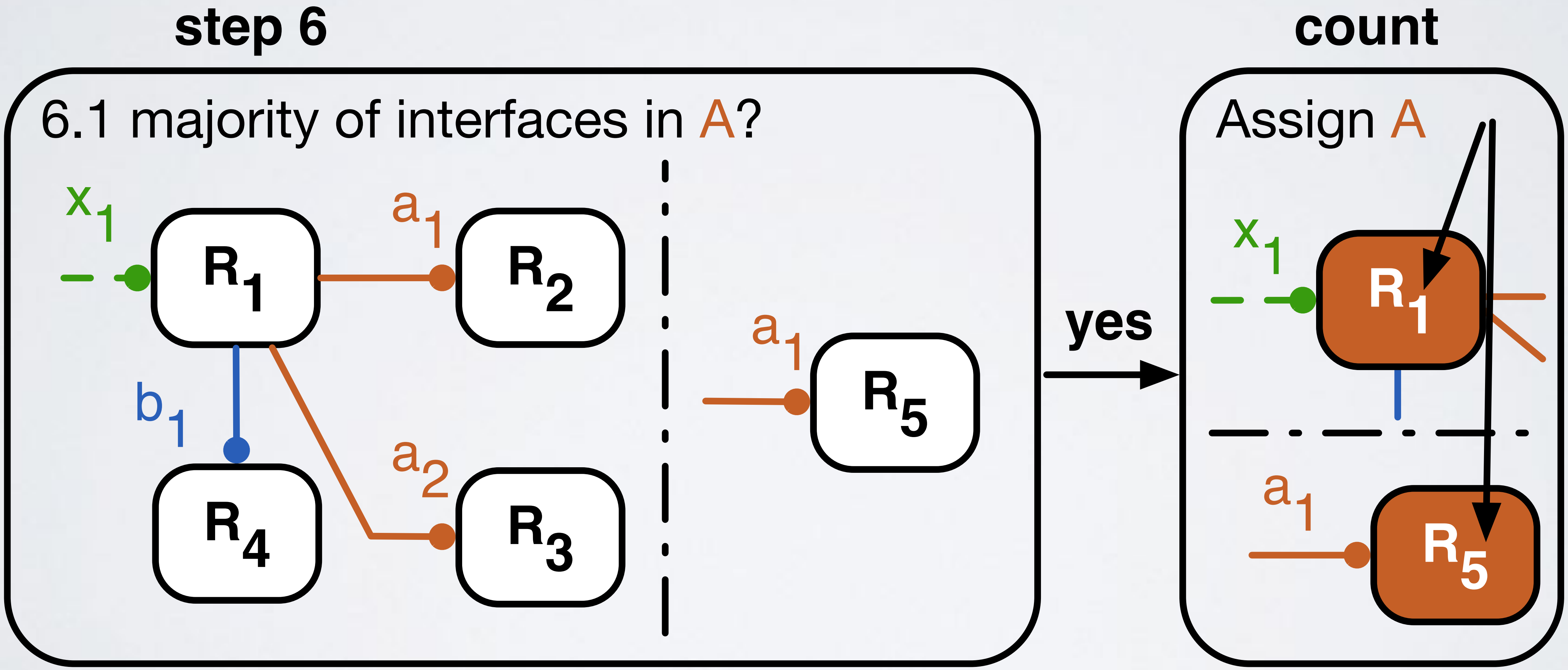
yes → missing: Assign B

no

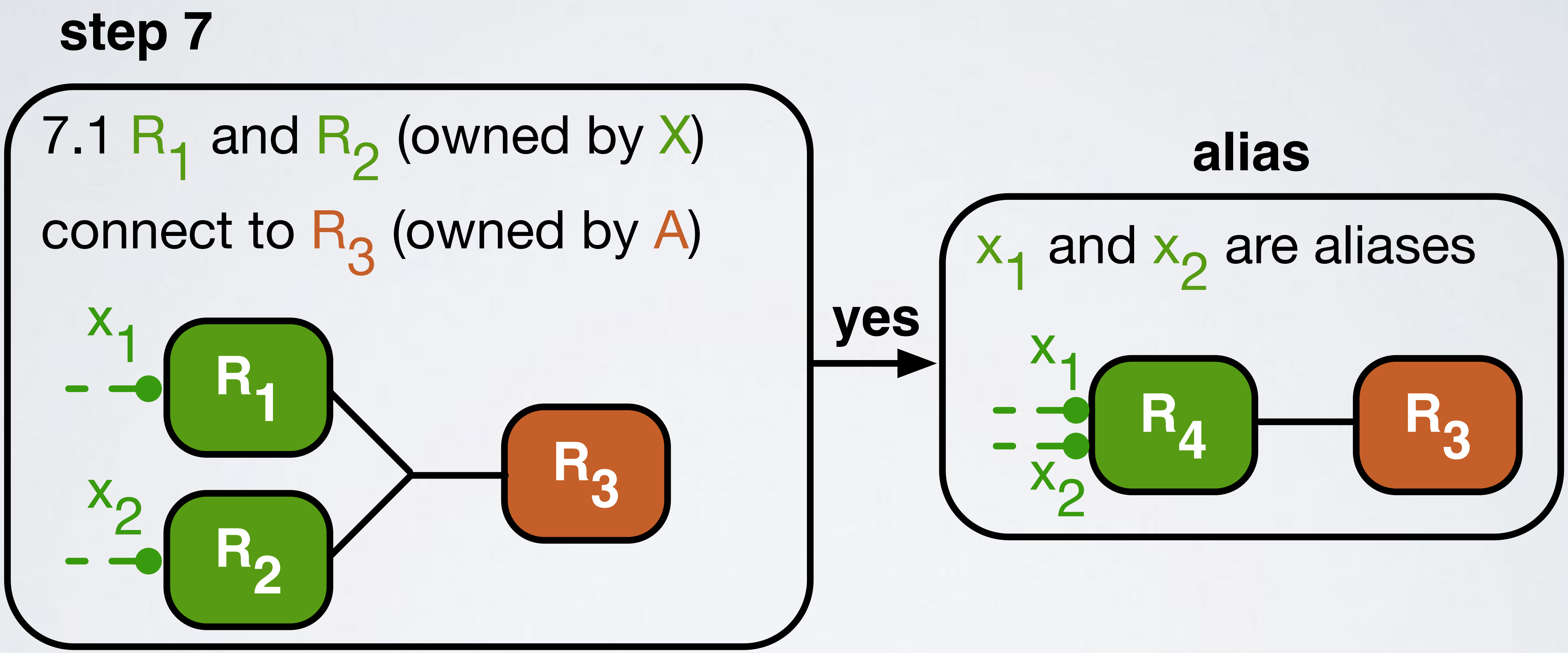
5.5 all interfaces adjacent to R_1 in A

yes → hiddenpeer: Assign A

Use IP-AS mappings to infer operator of neighbor routers in ambiguous scenarios



Infer additional aliases for border routers



step 7

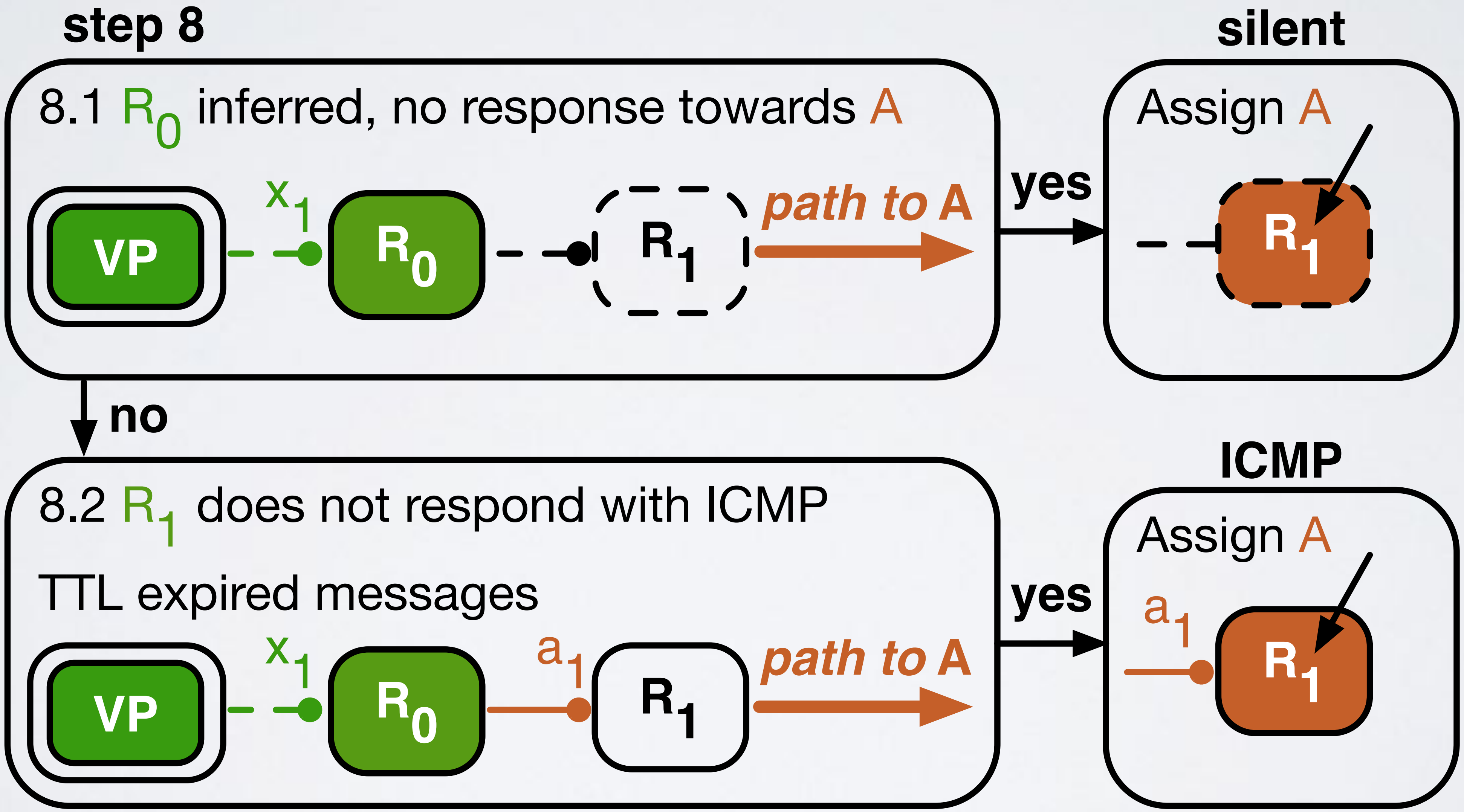
7.1 R_1 and R_2 (owned by X) connect to R_3 (owned by A)

yes

alias

x_1 and x_2 are aliases

Infer operator of neighbor routers without TTL expired messages



step 8

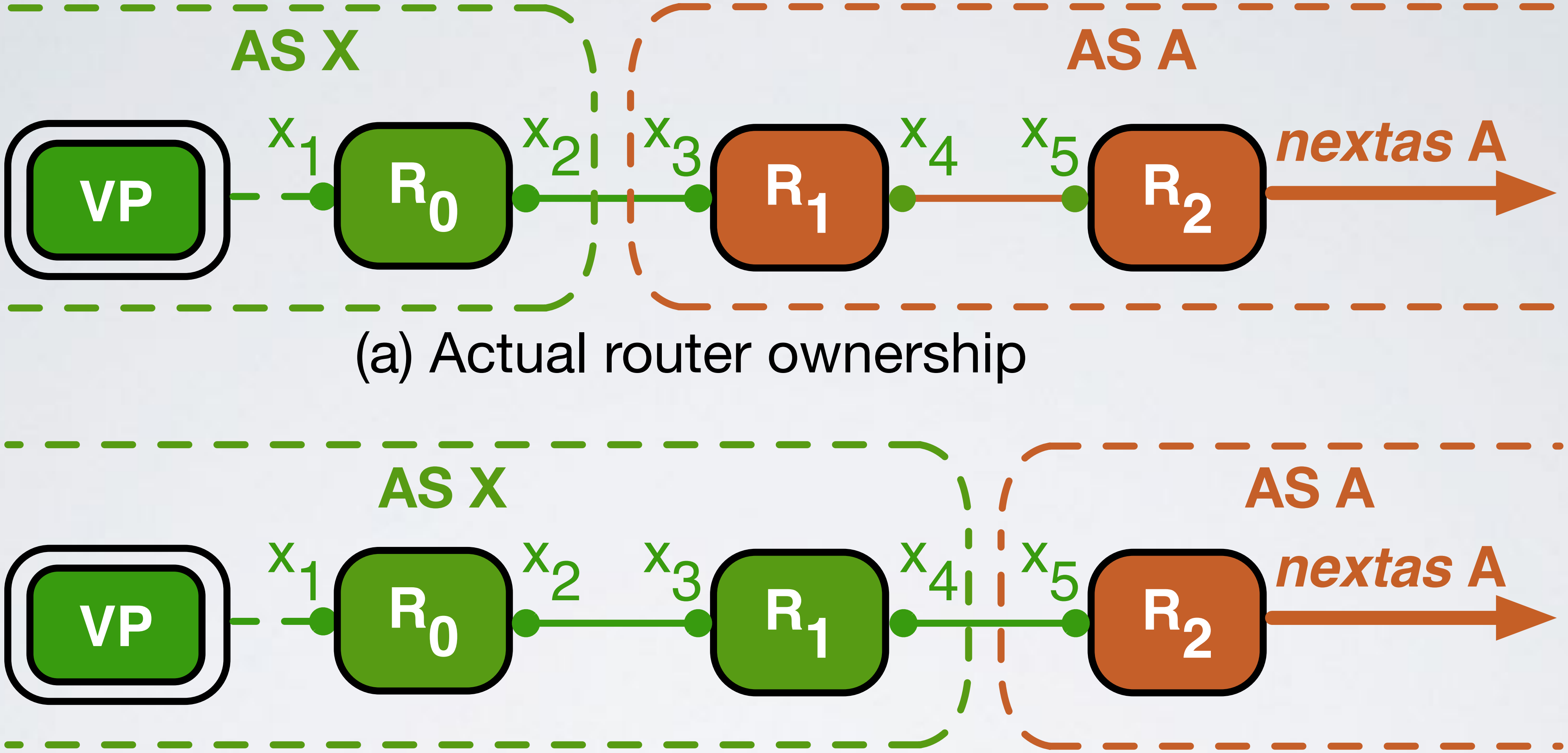
8.1 R_0 inferred, no response towards A

Assign A

8.2 R_1 does not respond with ICMP TTL expired messages

Assign A

Limitations

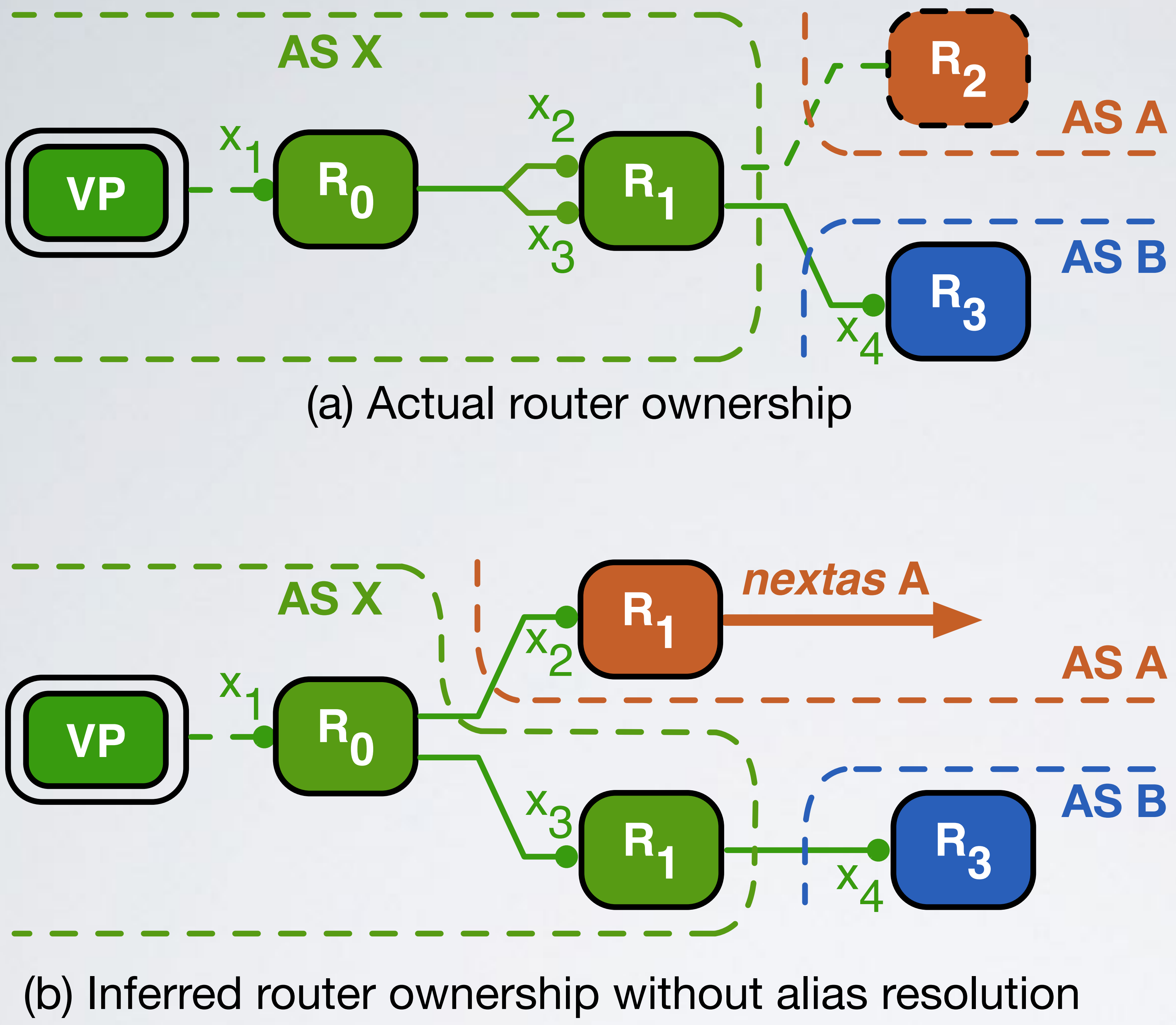


(a) Actual router ownership

(b) Inferred router ownership

If an AS uses provider-aggregatable address space from their provider on interfaces on their internal routers, bdrmap may incorrectly infer the position of interdomain link.

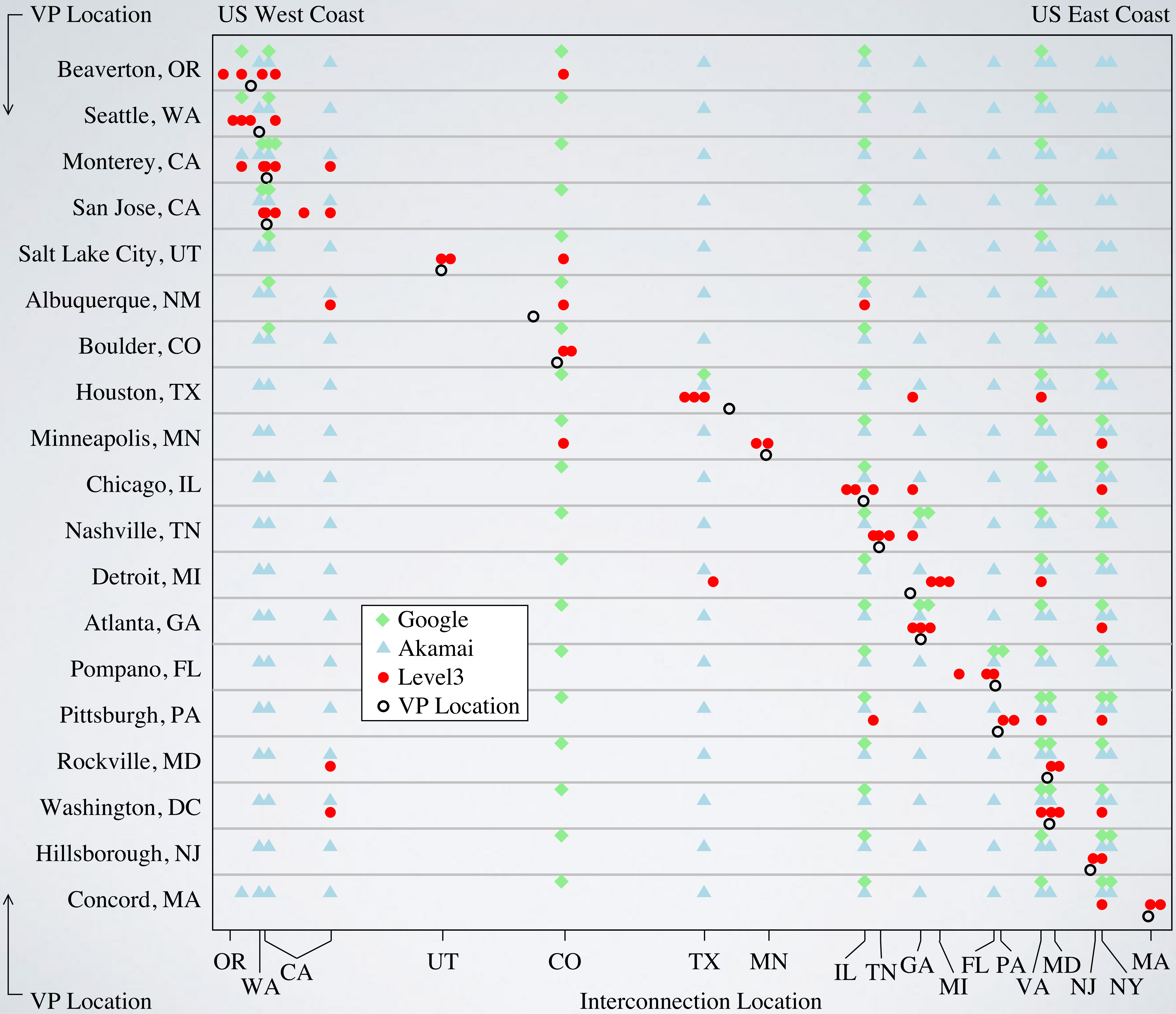
Limitations



(a) Actual router ownership

(b) Inferred router ownership without alias resolution

If router R1 responds with different IP addresses depending on the destination probed, and those addresses are not inferred to be aliases, bdrmap may incorrectly infer the position of an interdomain link.


VP Location
US West Coast
US East Coast
Beaverton, OR
Seattle, WA
Monterey, CA
San Jose, CA
Salt Lake City, UT
Albuquerque, NM
Boulder, CO
Houston, TX
Minneapolis, MN
Chicago, IL
Nashville, TN
Detroit, MI
Atlanta, GA
Pompano, FL
Pittsburgh, PA
Rockville, MD
Washington, DC
Hillsborough, NJ
Concord, MA
Google
Akamai
Level3
VP Location
OR
WA
CA
UT
CO
TX
MN
IL
TN
GA
MI
FL
PA
VA
MD
NJ
NY
MA
VP Location
Interconnection Location

Development Approach

- We designed and implemented our algorithm over the course of a year, without validation data.

- We used DNS-naming, where available, to infer if our methods appeared to yield correct inferences

- Border routers with high out-degree usually implied an incorrect inference

- We did not use DNS-naming for validation as we found mislabeled interfaces, as well as names containing organization names, rather than AS numbers