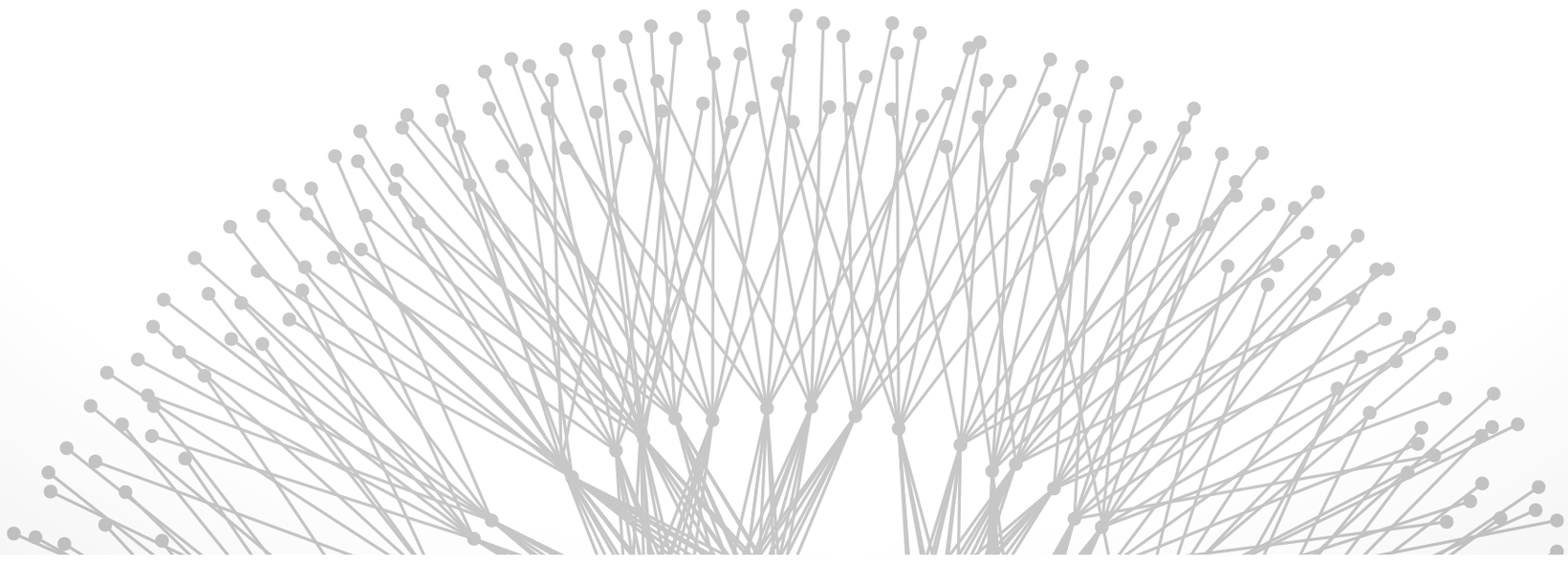
Interdomain Routing and Connectivity

Brighten Godfrey

CS 538 February 28 2018



Routing

Choosing paths along which messages will travel from source to destination.

Problems for intradomain routing

Distributed path finding

Optimize link utilization (traffic engineering)

React to dynamics

High reliability even with failures

Scale

Problems for interdomain routing

All of intradomain's problems

Bigger scale

Multiple parties

- No central control
- Conflicting interests
- Greater volume and diversity of attacks

Harder to change architecture

- Intradomain evolution: RIP, ISIS, OSPF, MPLS, OpenFlow, Segment Routing, …
- Interdomain: BGP.

Interdomain routing

BGP: Border Gateway Protocol

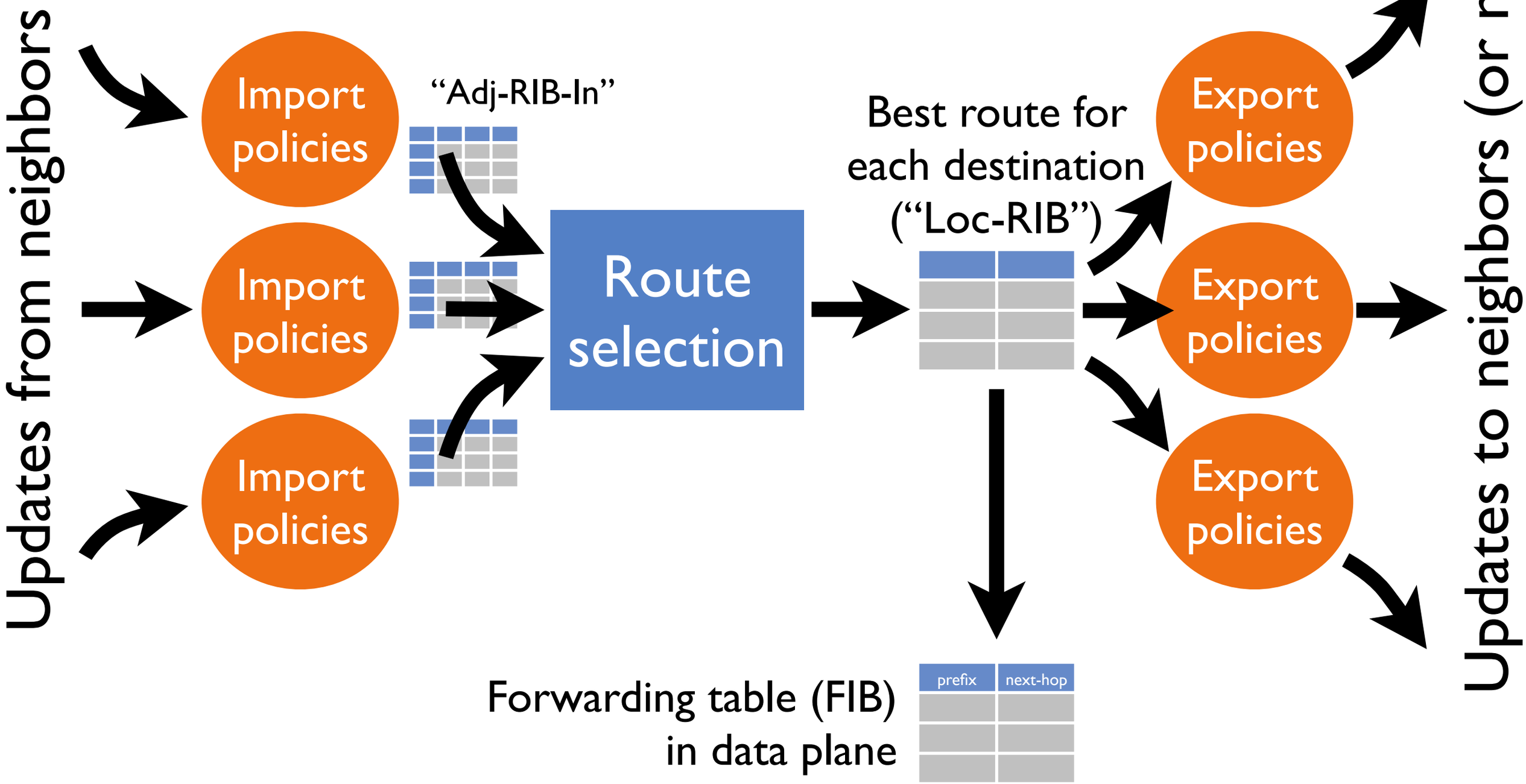
Distance vector variant

- Send incremental changes, not whole vector
- Path vector: Remember path instead of distance

Why path vector?

- Avoid DV's transient loops; but more importantly...
- Policy support: can pick any path offered by neighbors, not necessarily the shortest (Link State cannot)
- Privacy support: path choice policy is applied locally, not announced globally
  - Q: How much privacy is there?

BGP: The picture at one router



Route selection process

Import policies

Step	Attribute	Controlled by local or neighbor AS?
1.	Highest LocalPref	local
2.	Lowest AS path length	neighbor
3.	Lowest origin type	neither
4.	Lowest MED	neighbor
5.	eBGP-learned over iBGP-learned	neither
6.	Lowest IGP cost to border router	local
7.	Lowest router ID (to break ties)	neither

[Caesar, Rexford, IEEE Network Magazine, 2005]

This process is extended in many real implementations.

Common business relationships

Provider

- I pay for traffic we exchange

Peer

- Often “settlement-free”, i.e., neither party pays

Customer

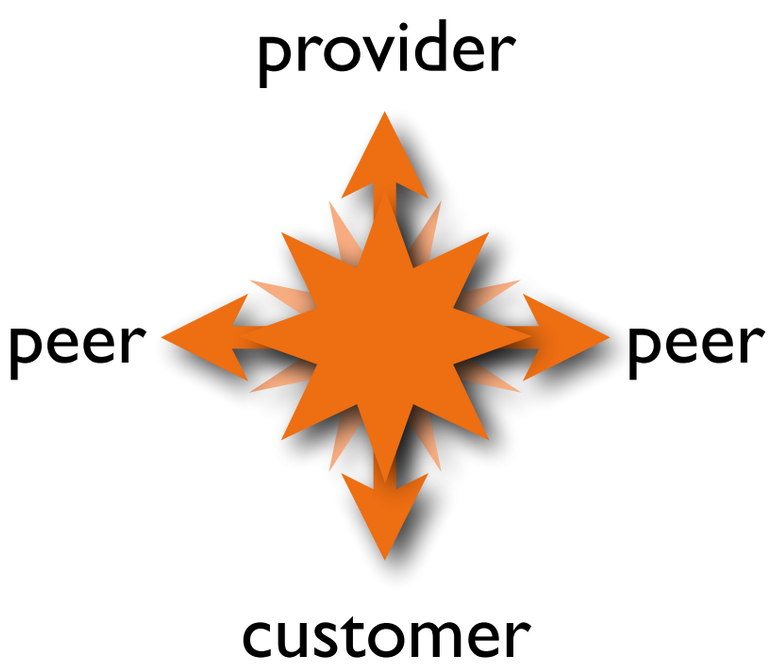
- They pay me for traffic we exchange

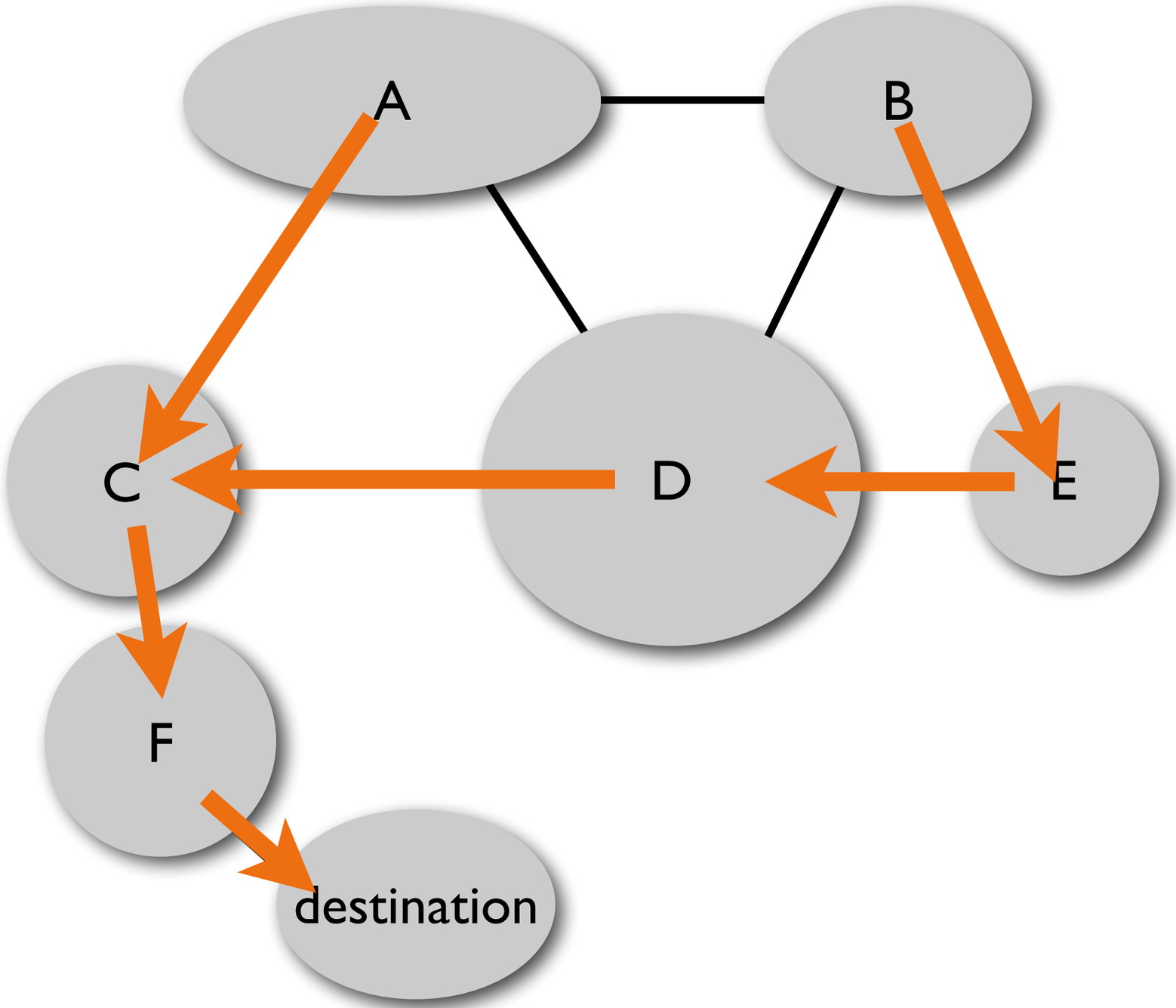
How might these classifications be used…

- in route selection?
- in route export?

Common policies

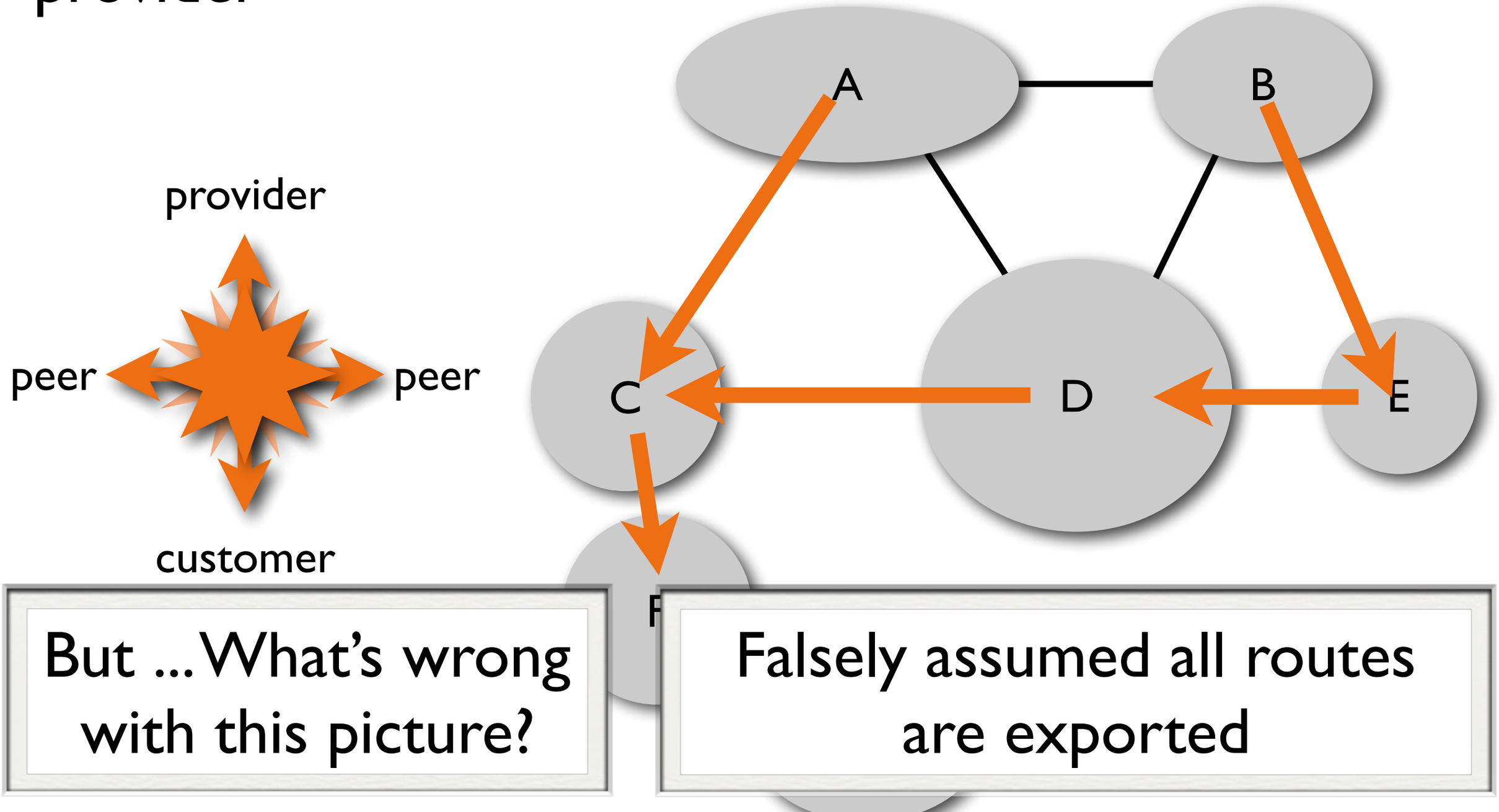
Route selection: prefer customer over peer over provider





Common policies

Route selection: prefer customer over peer over provider

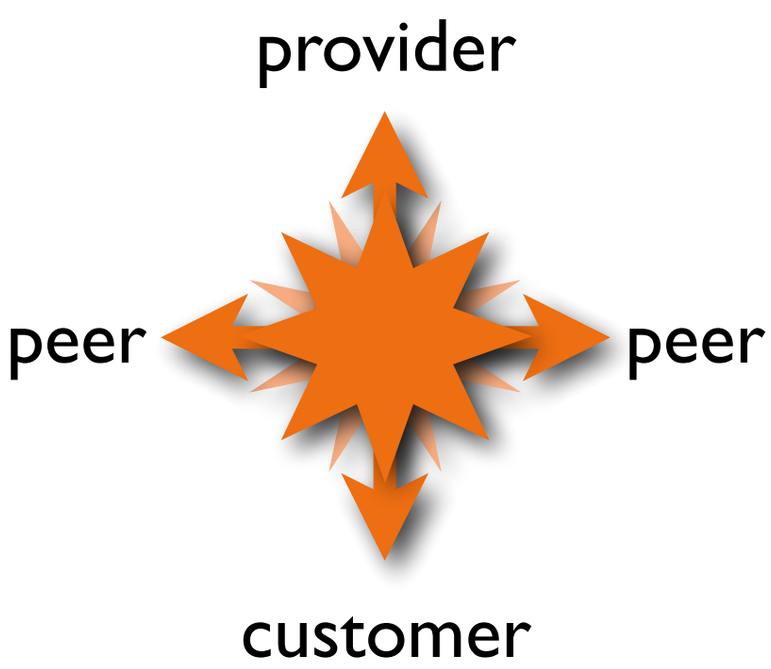


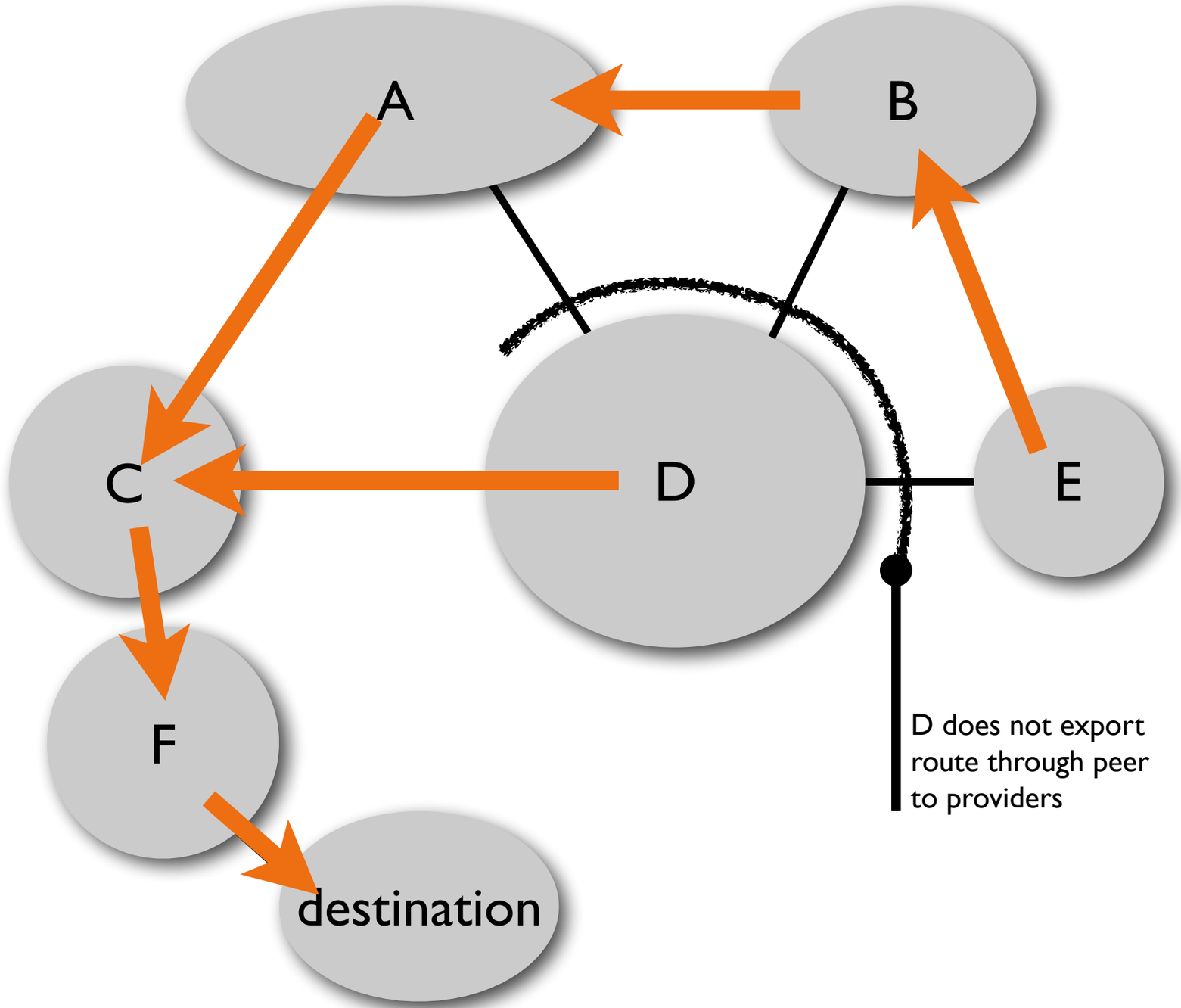
But ... What's wrong with this picture?

Falsely assumed all routes are exported

Common policies

Route export (most common): to/from customer only ("valley-free")

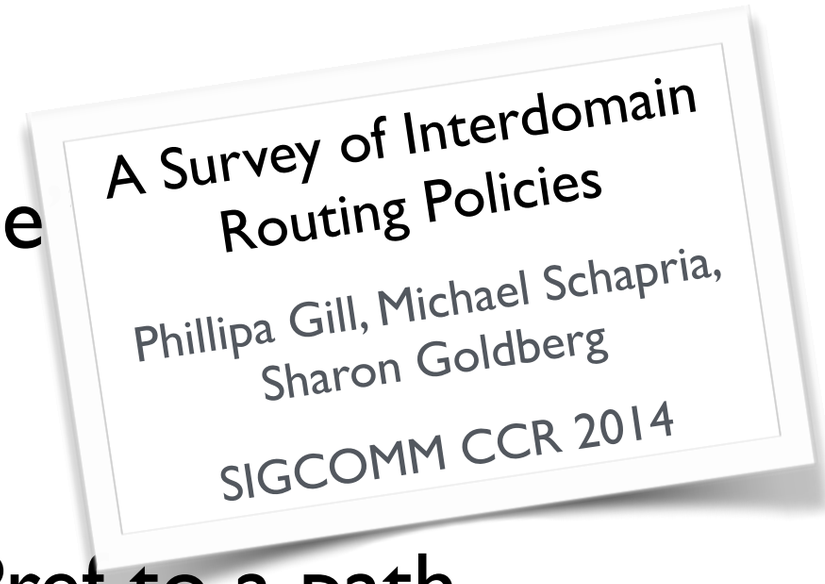




Common policies: summary

"Gao-Rexford" policies:

- Prefer customer > peer > provider
- Export all routes to customers
- Export customer routes to everyone
- (…and export nothing else: "valley-free



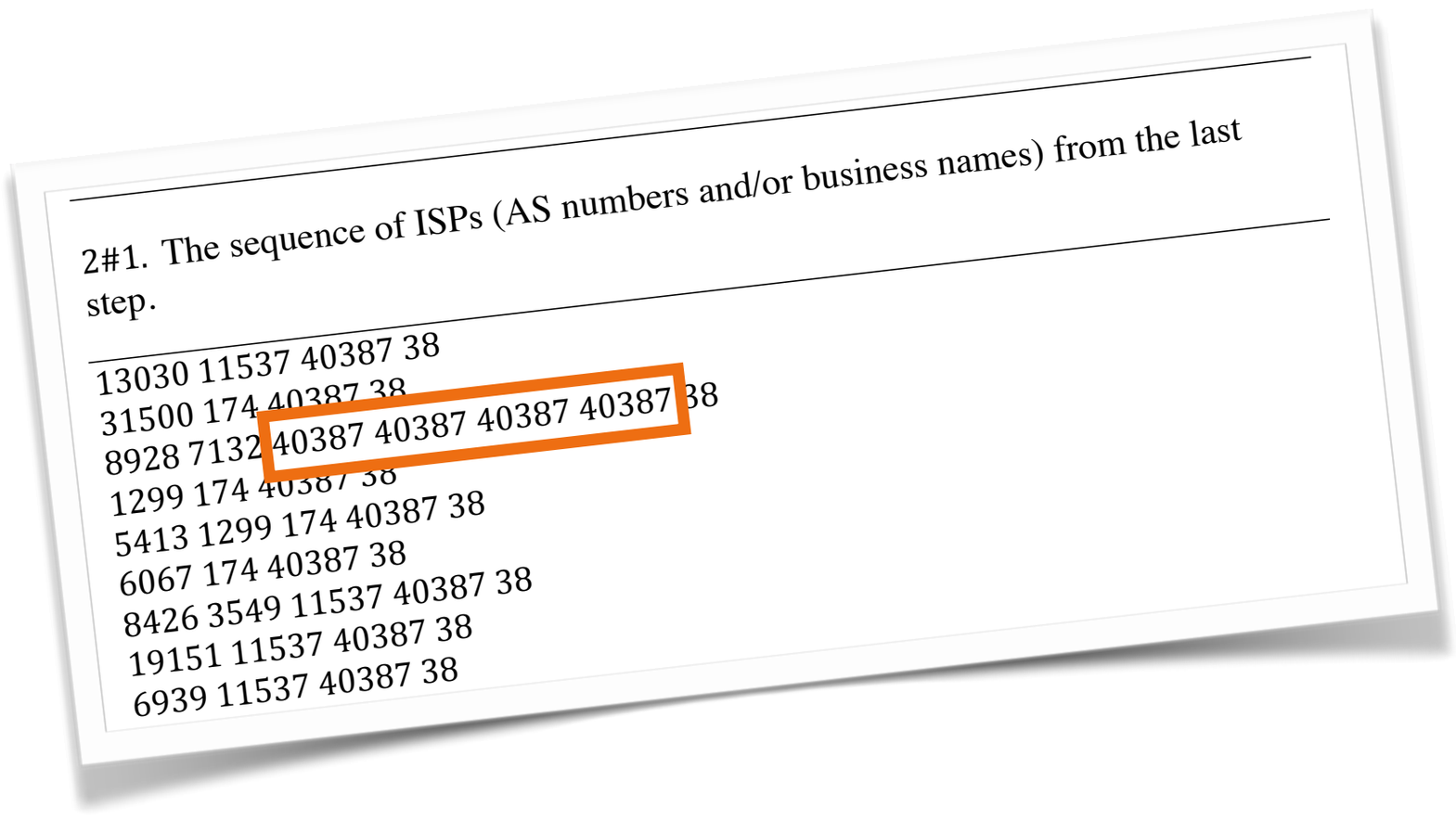
Are they used in the real world?

- "Do you always assign a higher LocalPref to a path through your customer than to a path through your peer or transit provider? (Note: exclude cases where routes through customers are tagged as backup.)" 79% yes
- Does your LocalPref configuration depend only on the next-hop AS? 56% yes

Paper discussion

How does BGP traffic engineering fit with TeXCP? Are they solving the same problem?

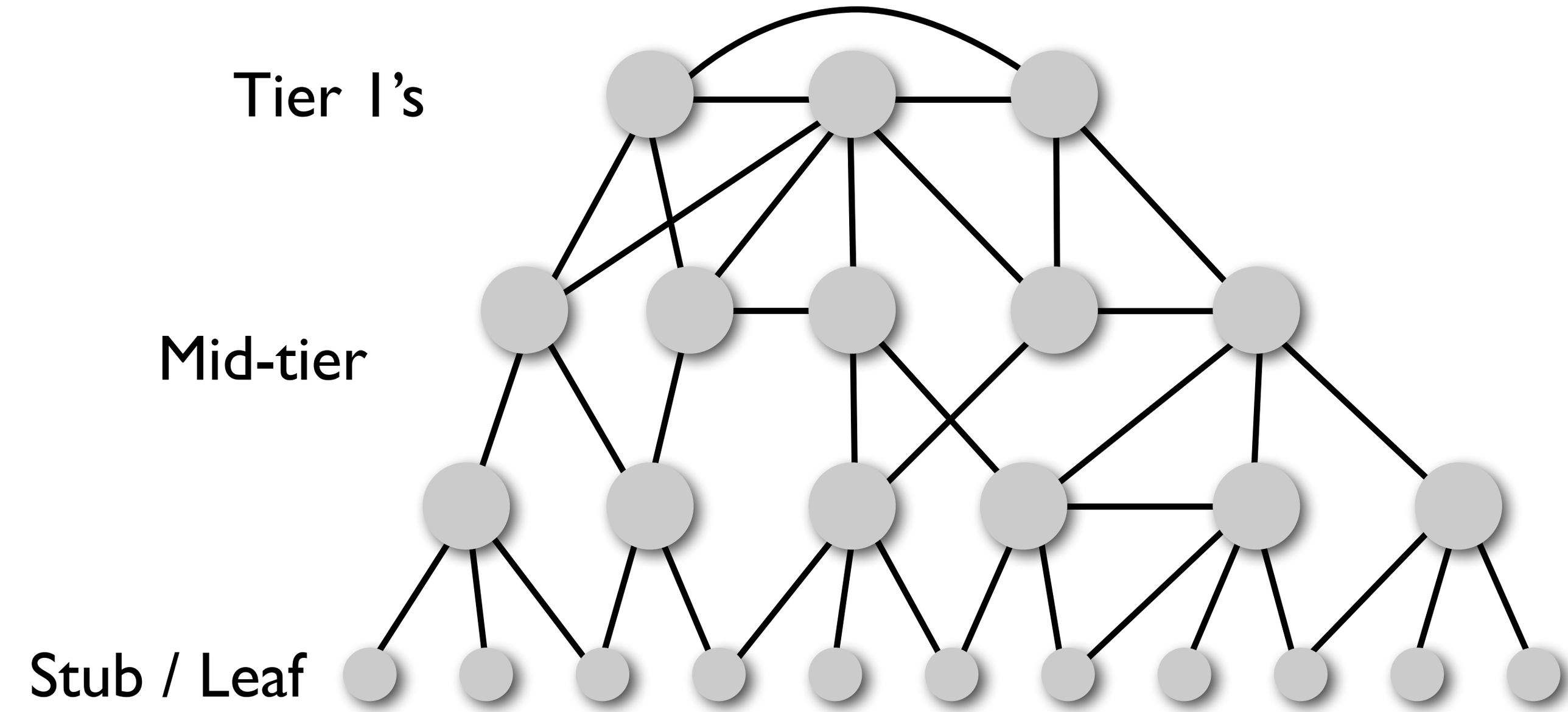
How can ISPs perform interdomain outbound TE?



2#1. The sequence of ISPs (AS numbers and/or business names) from the last step.

13030 11537 40387 38
31500 174 40387 38
8928 7132 40387 40387 40387 40387 38
1299 174 40387 38
5413 1299 174 40387 38
6067 174 40387 38
8426 3549 11537 40387 38
19151 11537 40387 38
6939 11537 40387 38

Interconnection: Traditional view



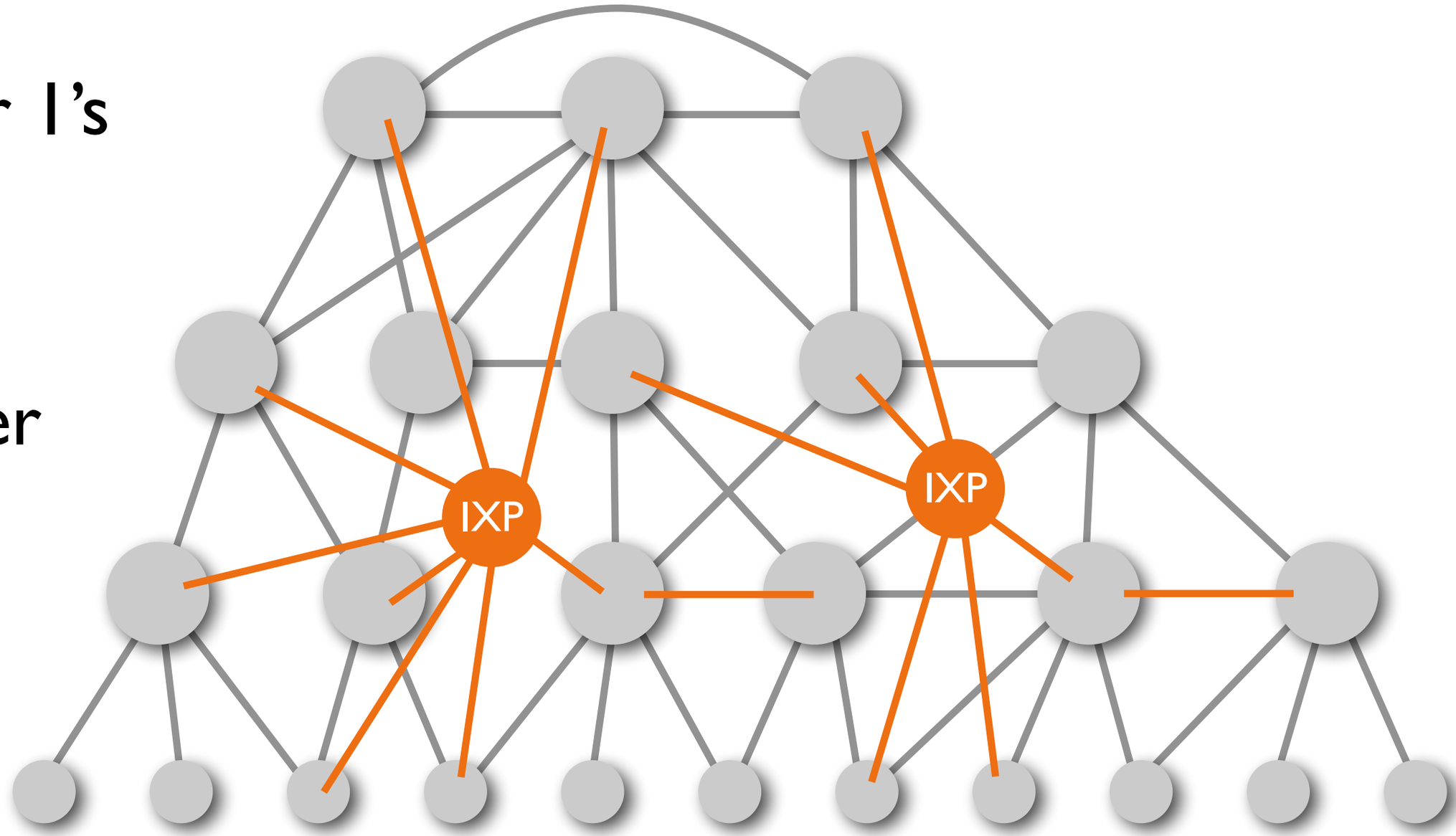
Hierarchical, limited peering at lower tiers

Interconnection: Modern view

Tier 1's

Mid-tier

Stub / Leaf



Significant and increasing peering at lower tiers

Data from Ager et al, SIGCOMM'12

Significant peering

- Estimated 200,000 peerings just in Europe
- More than 2x as many as non-peering links!

Past measurements missed these peerings

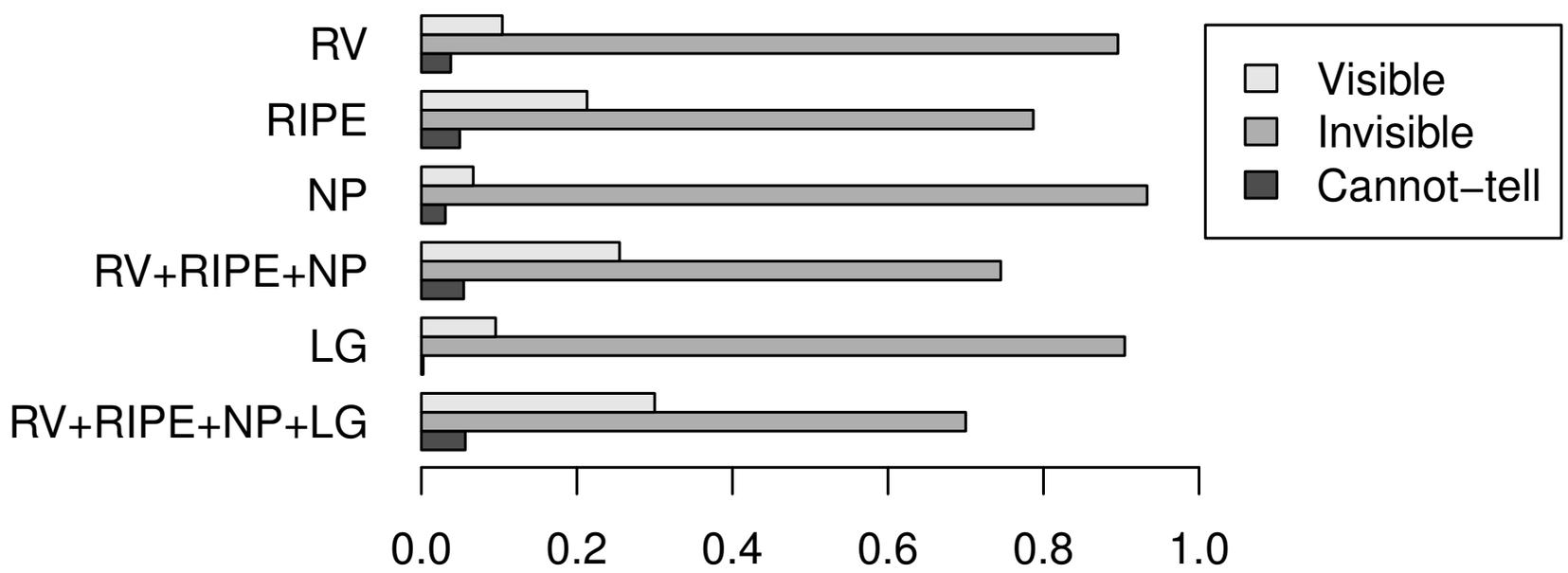
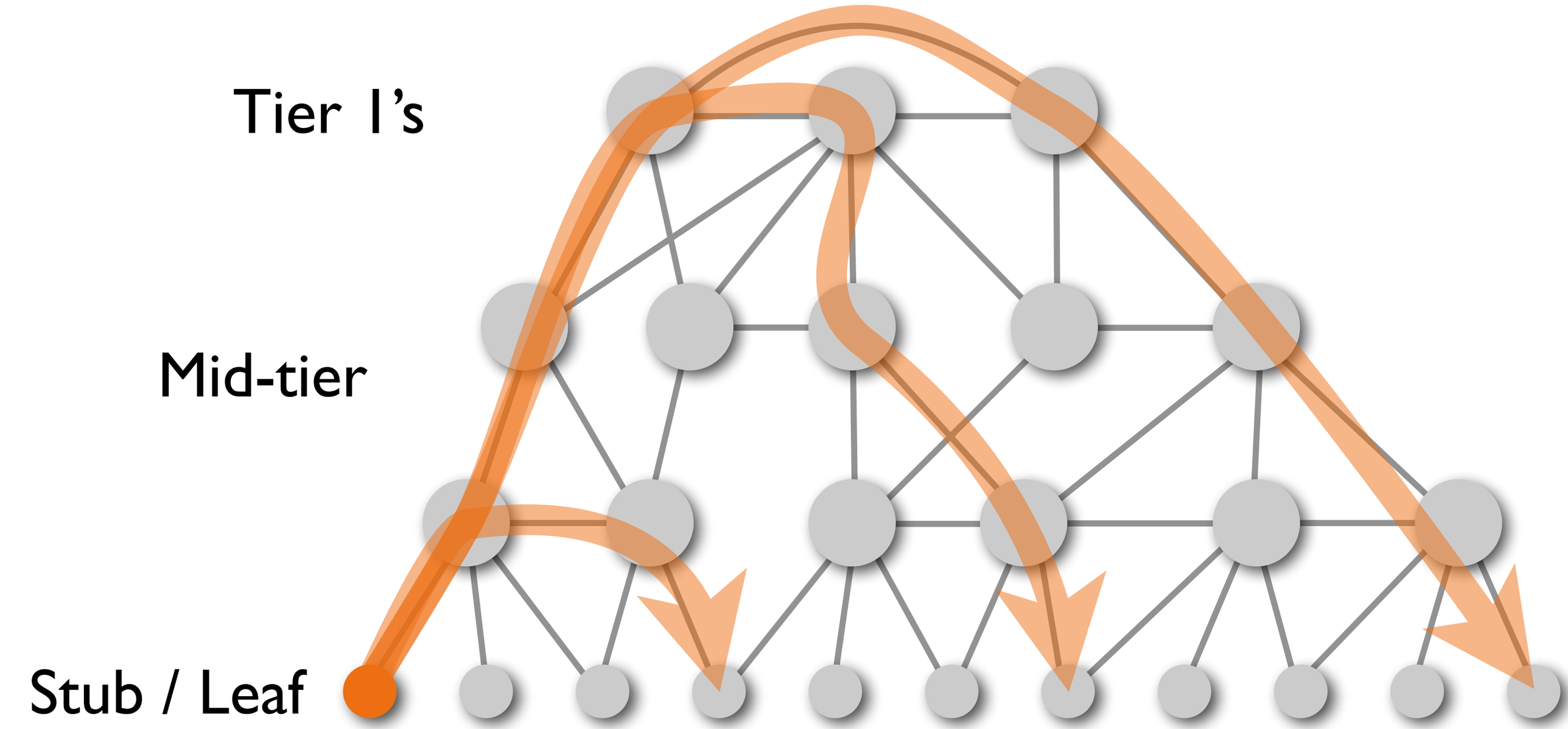
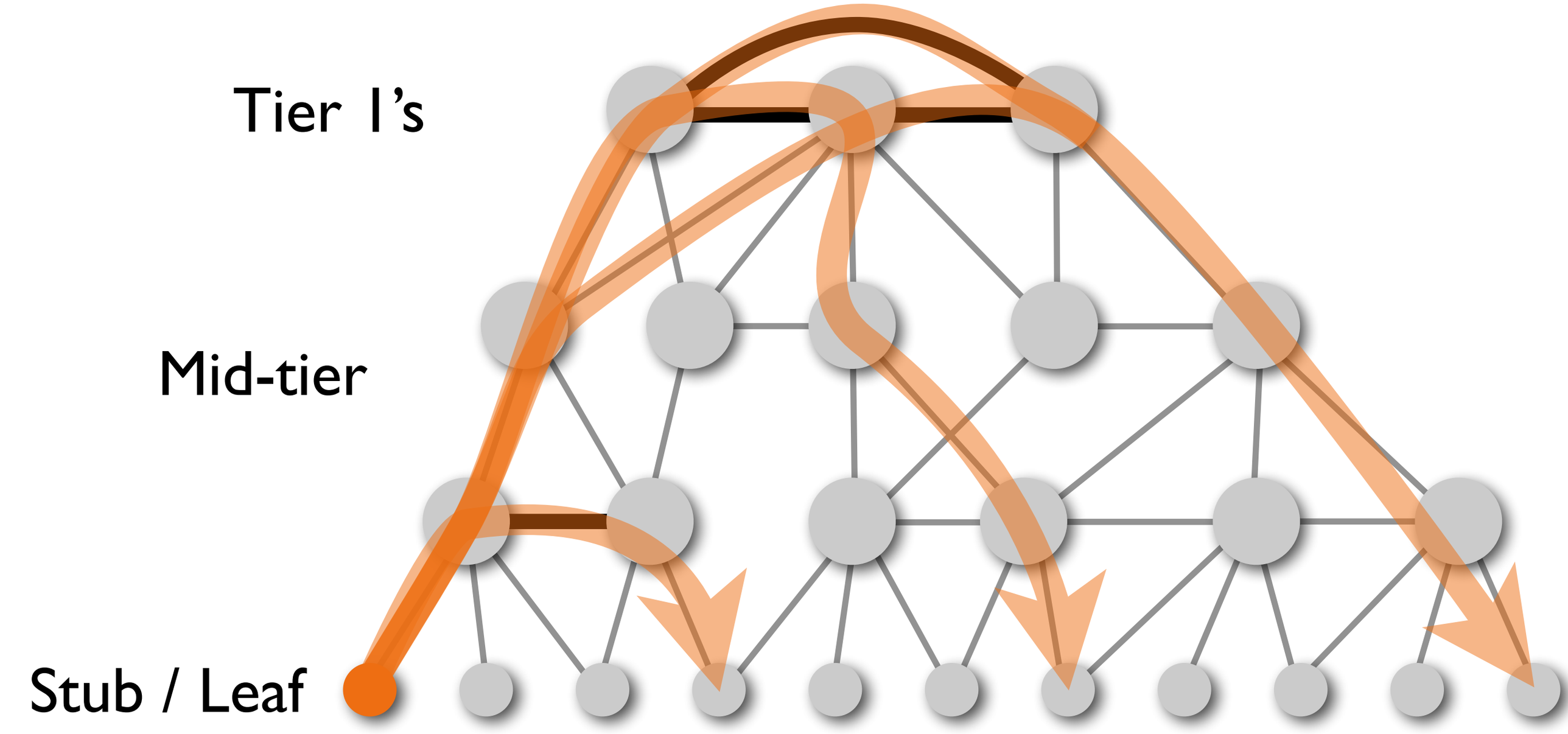


Figure 2: Peering links and visibility in control/data plane (normalized by number of detected P-P links).

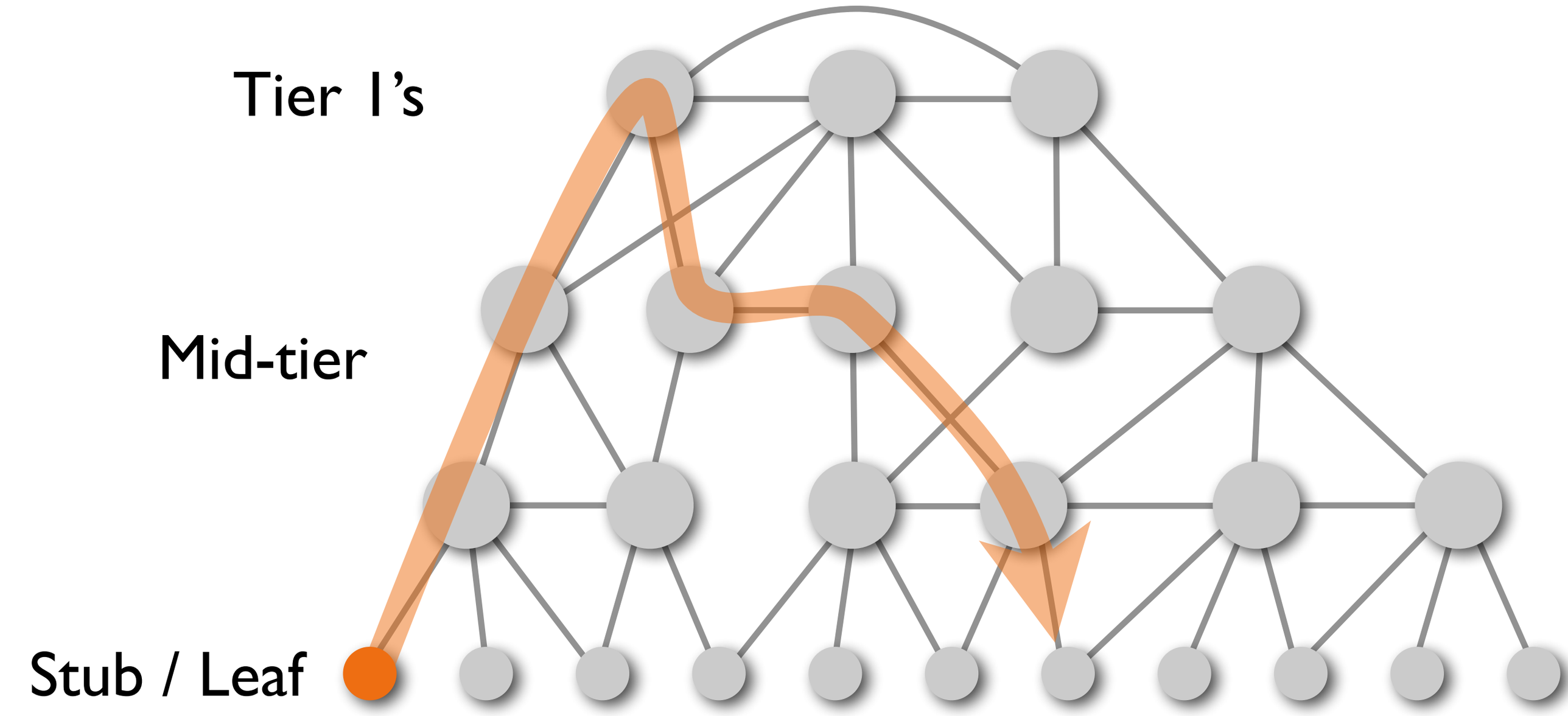
Why measurements miss so much



Why measurements miss so much



Why measurements miss so much

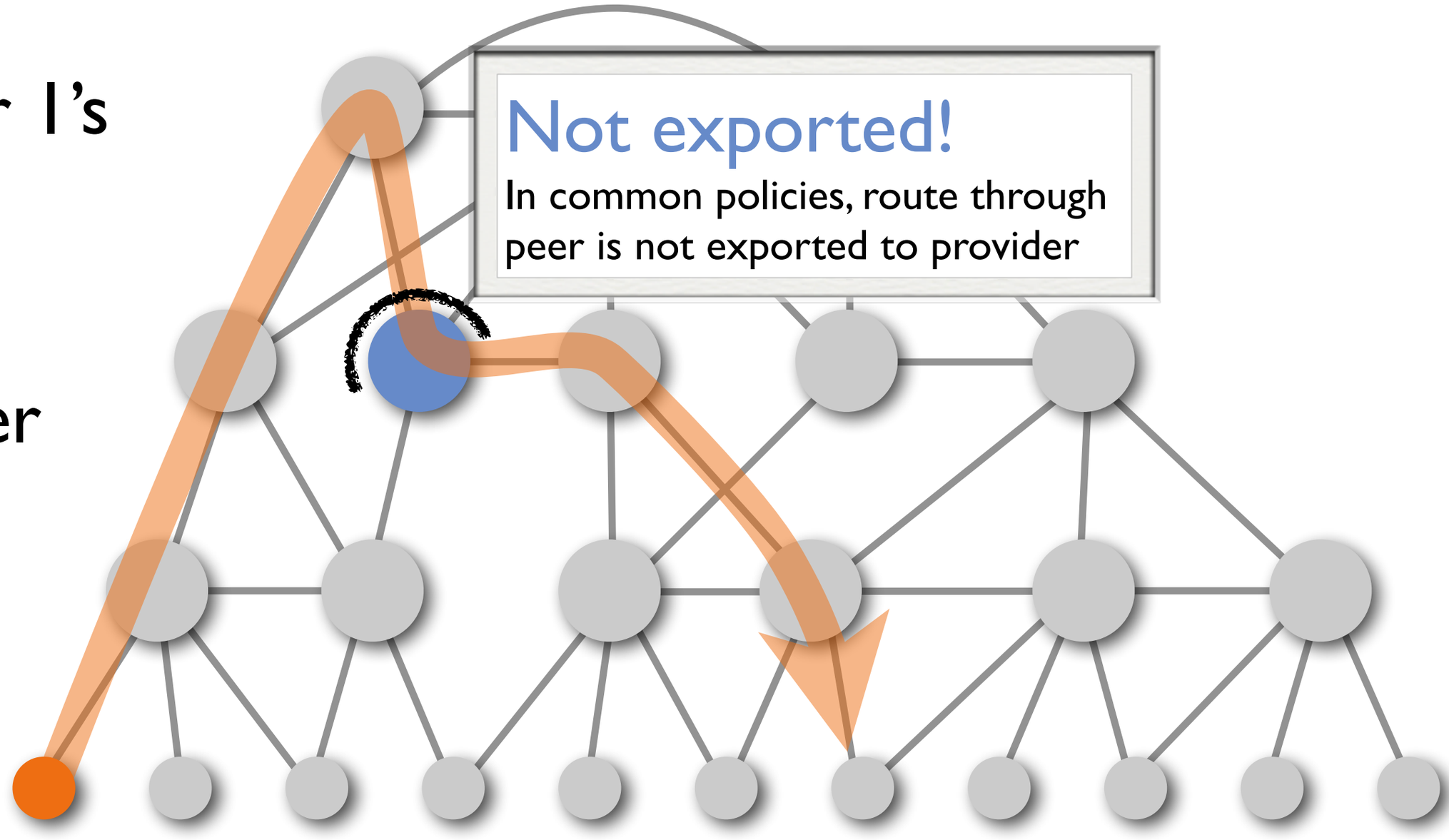


Why measurements miss so much

Tier 1's

Mid-tier

Stub / Leaf



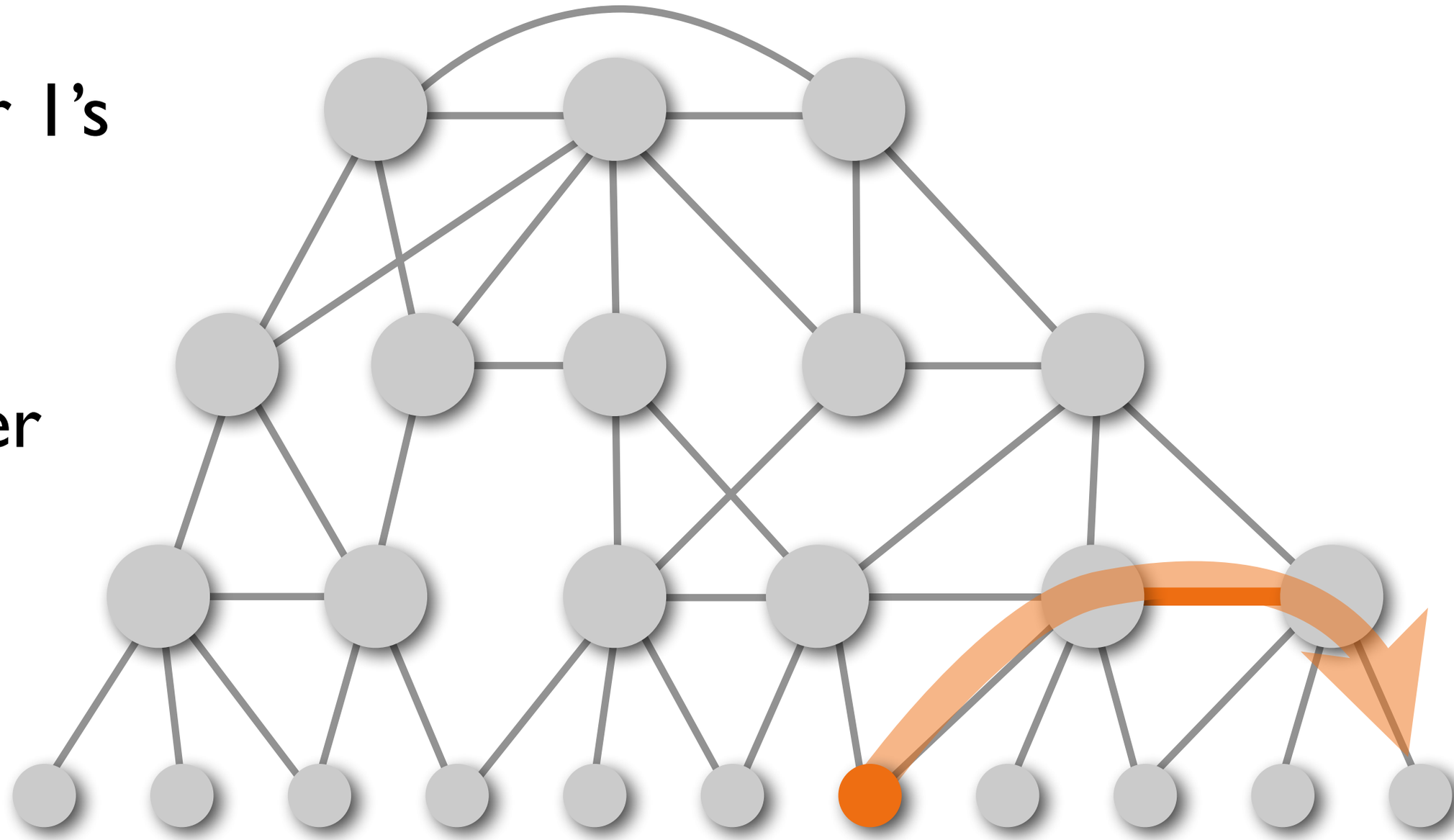
Measurement point
("Looking Glass")

Why measurements miss so much

Tier 1's

Mid-tier

Stub / Leaf



Measurement point
("Looking Glass")

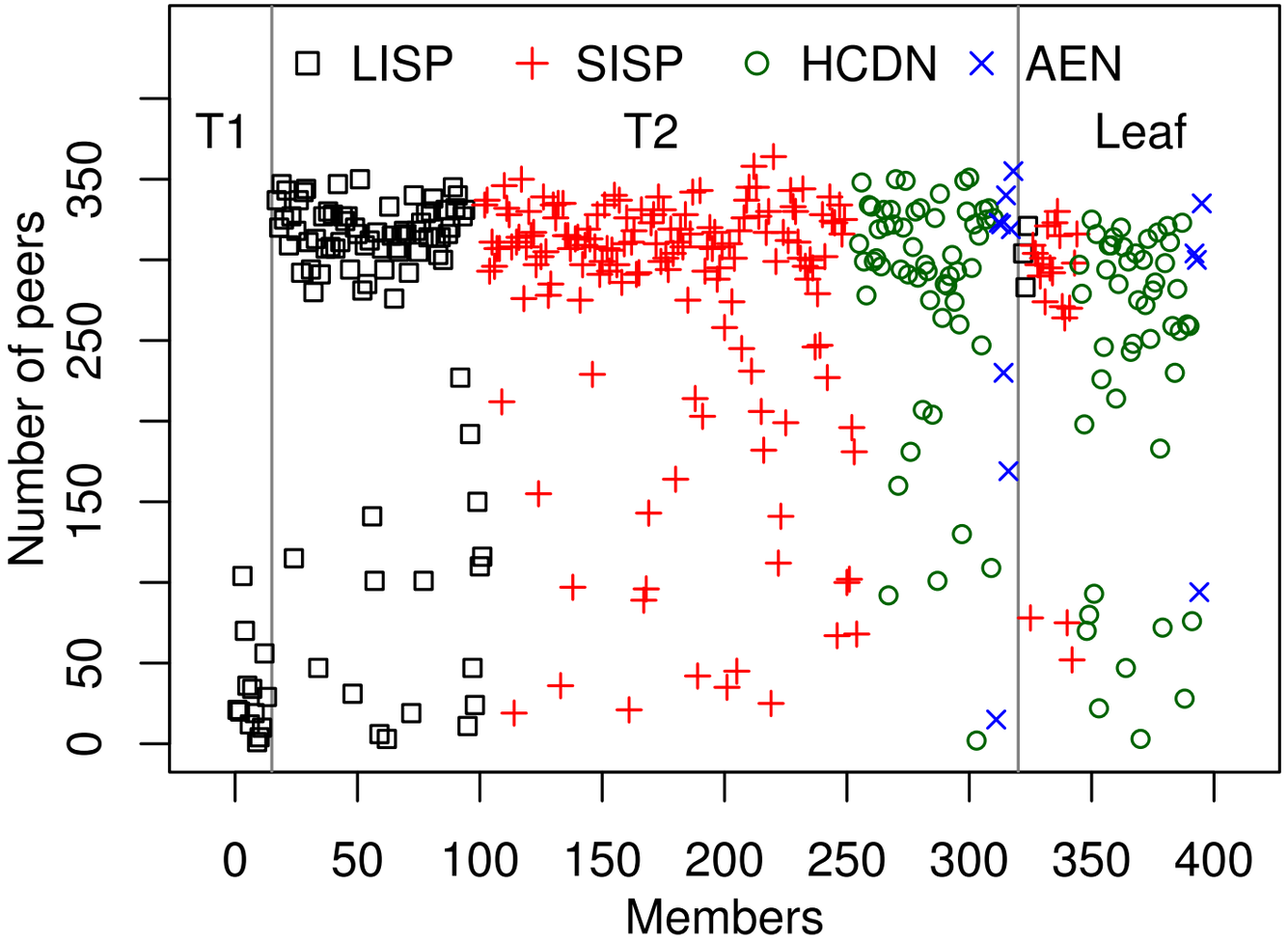
To see peer-peer link, *both* source & dest. of the probe must be in localized area

Paper discussion

What's the purpose of an IXP?

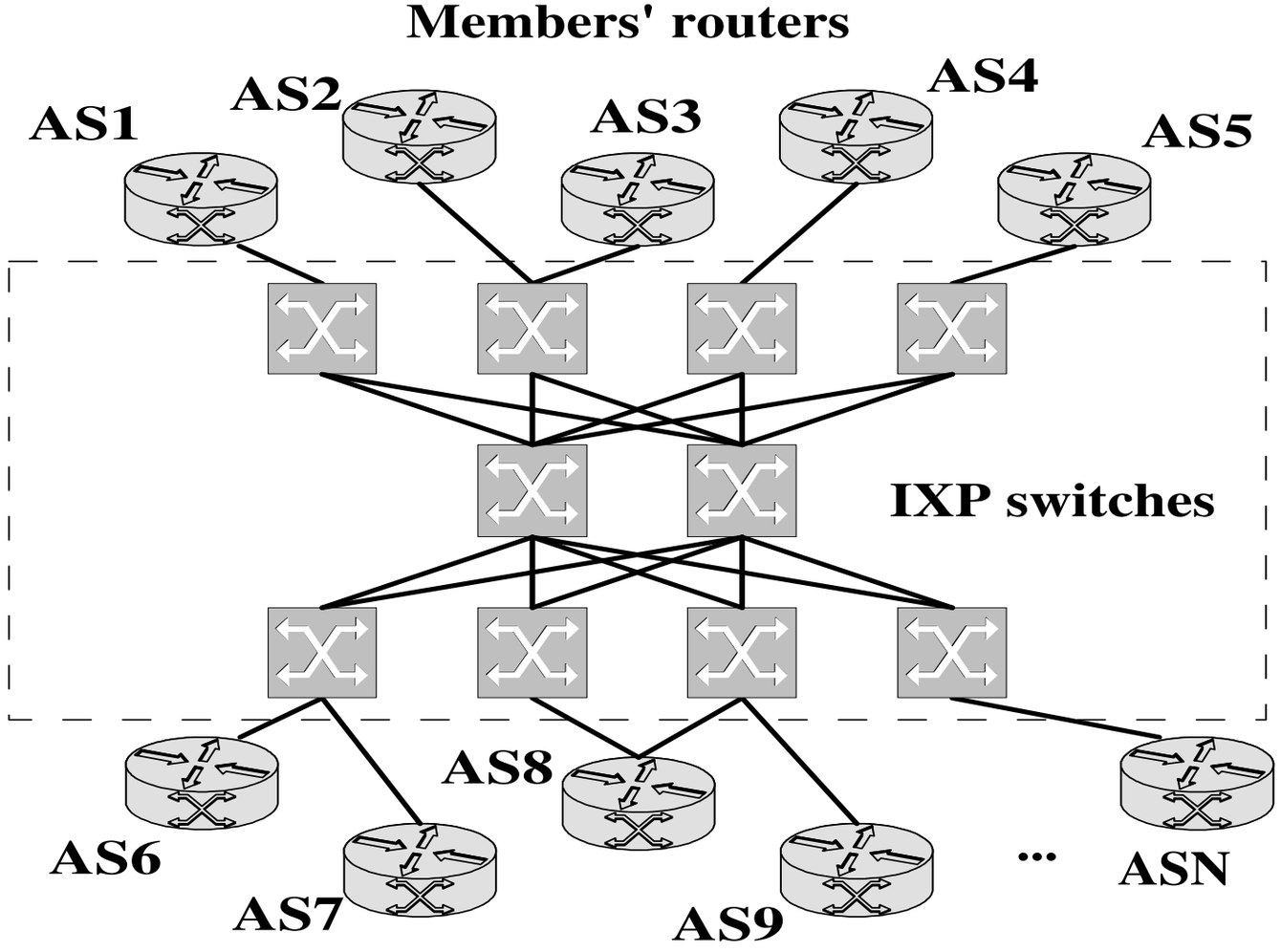
- "Metcalf's law": value of net is $O(n^2)$ when n participants

Why don't top-tier ISPs peer much at the IXP?



Paper discussion

How might router-level interconnection differ from AS-level peering? Would this paper's conclusions be the same for router-level?



physical links = 27

potential peerings = 45

[Ager, Chatzis, Feldmann, Sarrar, Uhlig, Willinger, SIGCOMM 2012]

Paper discussion

Similarly ... suppose we treat the IXP as an AS "in the middle" of each member AS-to-AS connection

Now how many links are there?

- 396 total members of this IXP, so 396 links
- vs. 50,000 reported in the paper!
- $O(n^2)$ peering relationships among n member ASes

This suggests interesting measurement projects:

- If you care about only the router level, what fraction of the links are observable?
- If you treat the IXP as an AS "in the middle", what fraction of the links are observable?

What's to come

Next: Part Two of the course: Grand Challenges

- programmability: capturing intent
- reliability
- selfishness
- security & privacy

March 12: Project midterm presentations

- Be ready to present on Monday March 12
- Some groups will present on Wednesday March 14

Project Midterm Presentations

Two key goals

- Benchmark: Demonstrate concrete progress
- Feedback & discussion with your peers

Content

- What problem are you solving?
- Why has past work not addressed the problem?
- What is your approach for solving it?
- What are your preliminary results & progress?

Logistics

- 10 minutes total: 6:40 min presentation + 4 min discuss
- PechaKucha format: 20 slides x 20 seconds, auto-advance