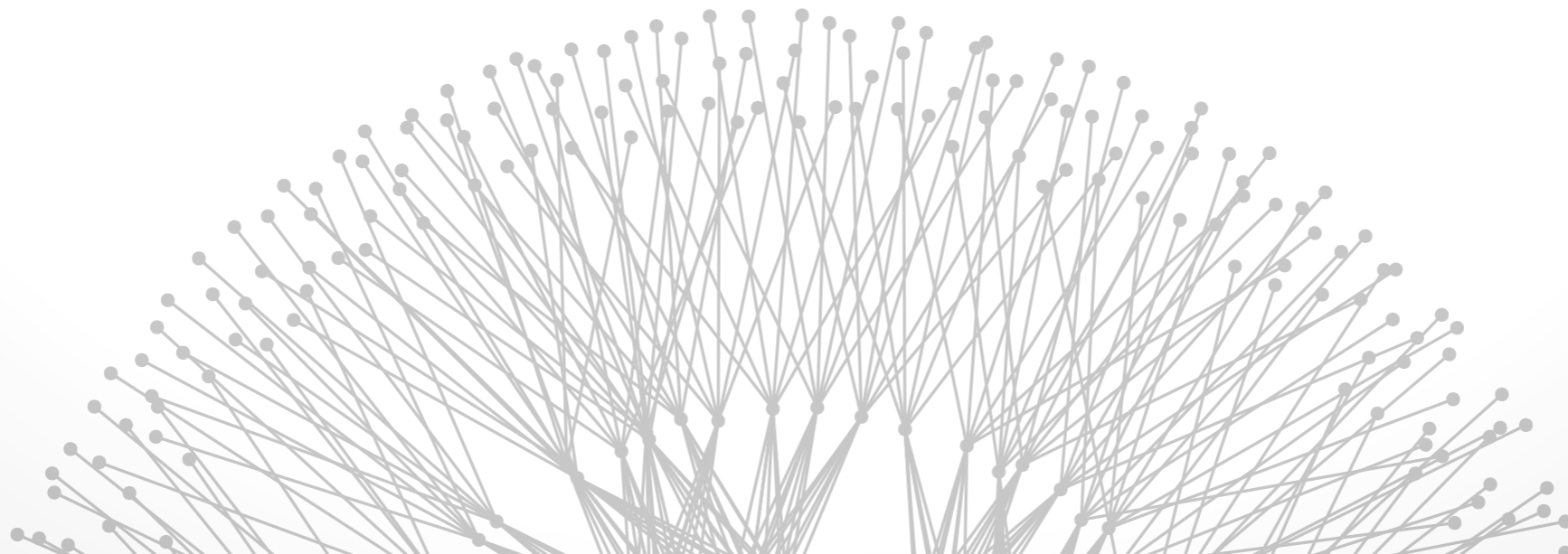# Interdomain Routing and Connectivity

Brighten Godfrey
CS 538 February 22 2017

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.

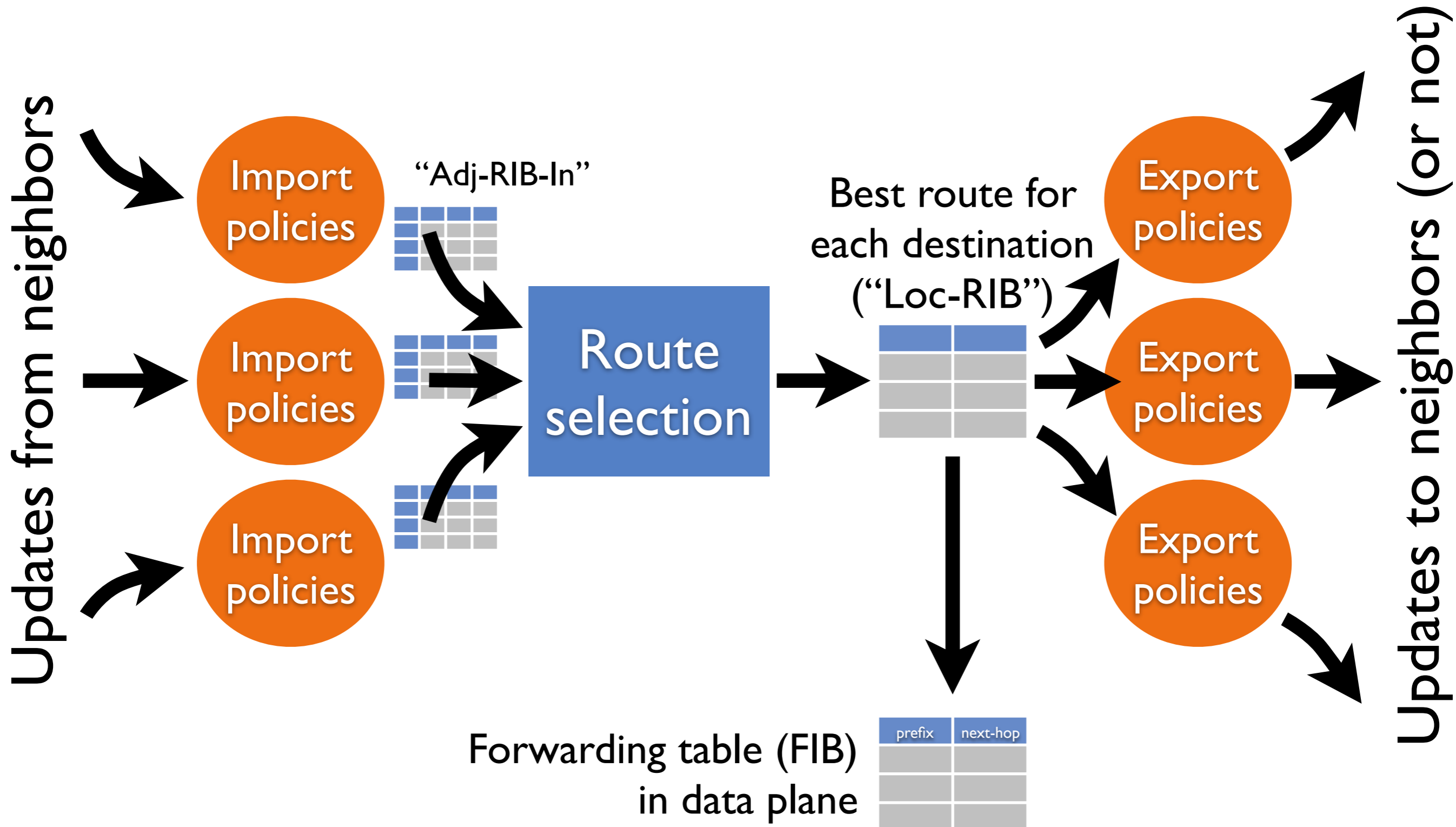# Interdomain routing

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

Updates from neighbors

Import policies

"Adj-RIB-In"

Route selection

Best route for each destination ("Loc-RIB")

Export policies

Updates to neighbors (or not)

Forwarding table (FIB) in data plane

prefix	next-hop

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]

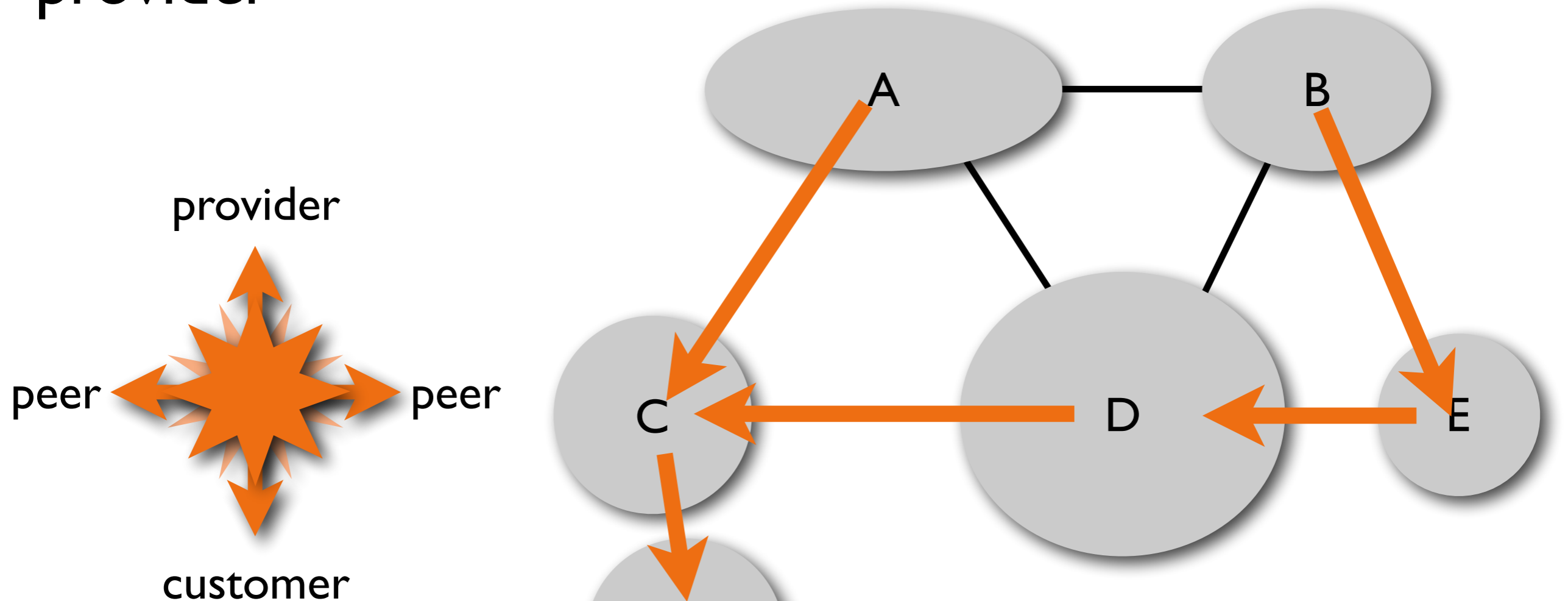This process is extended in many real implementations.

Common policies

Route selection: prefer customer over peer over provider
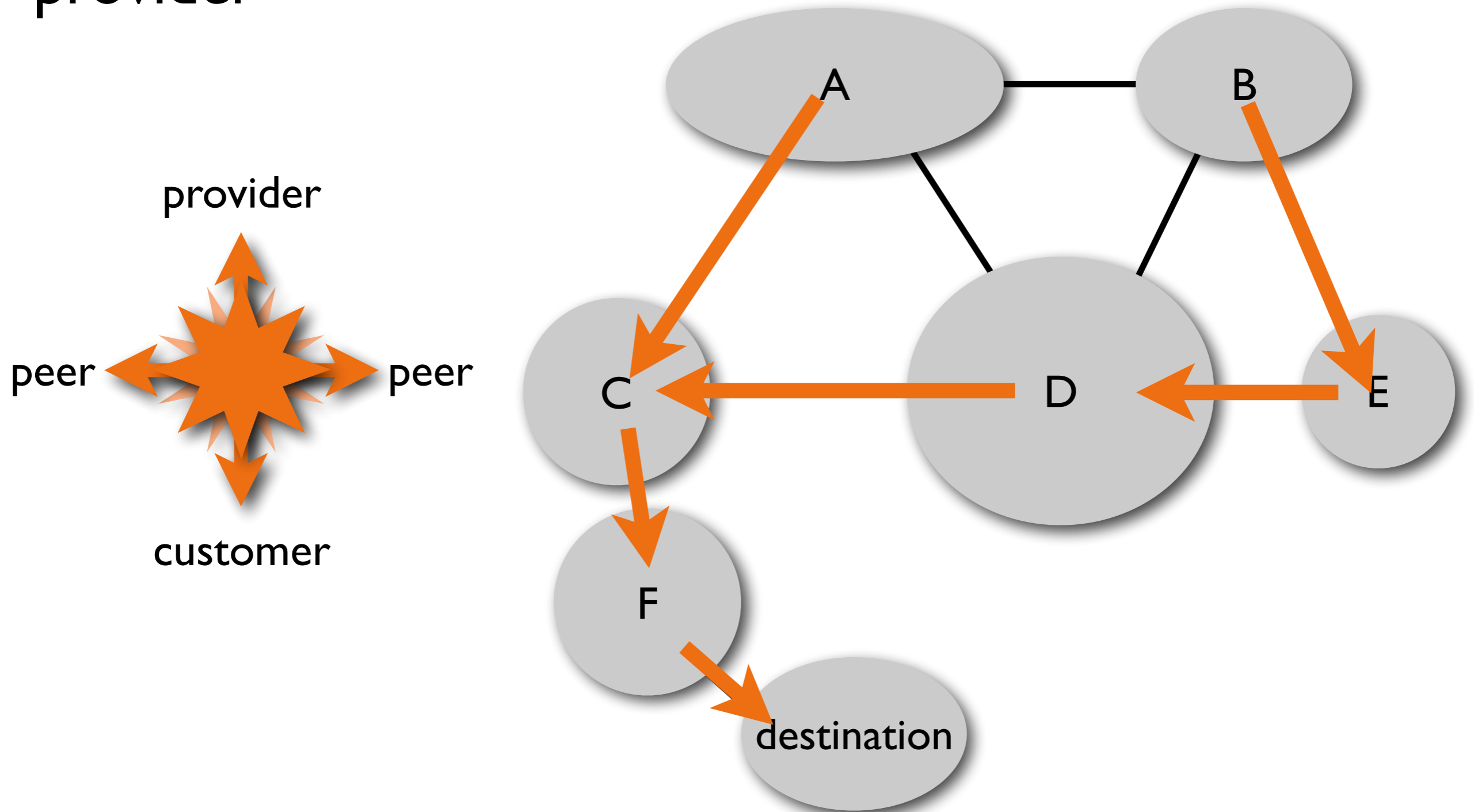
Common policies

Route selection: prefer customer over peer over provider

provider

peer   peer
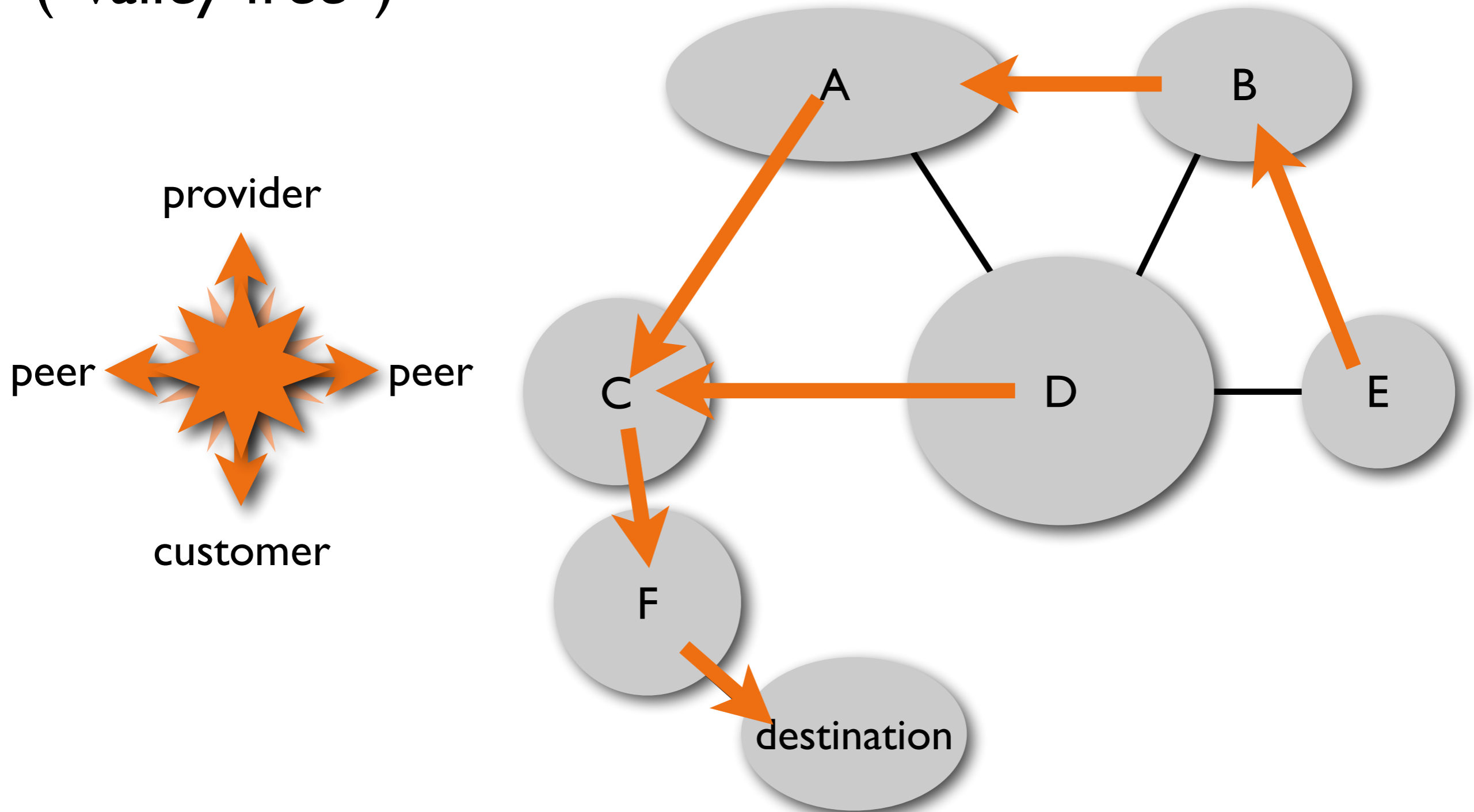
customer

A — B

C   D   E

F

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")

Paper discussion

How does BGP inbound 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

Tier 1's

Mid-tier

Stub / Leaf
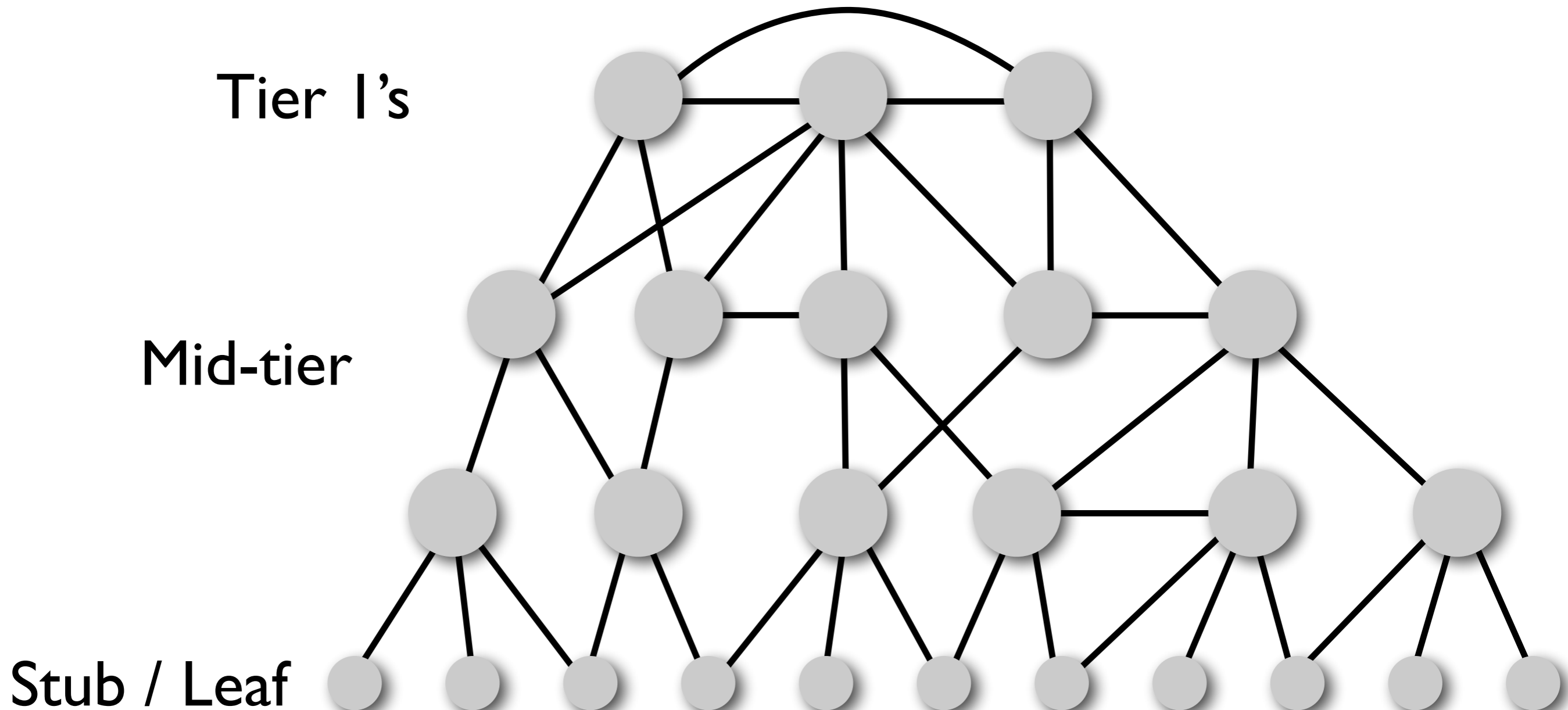
Hierarchical, limited peering at lower tiers

Interconnection: Modern view

Tier 1's

Mid-tier

IXP

IXP

Stub / Leaf
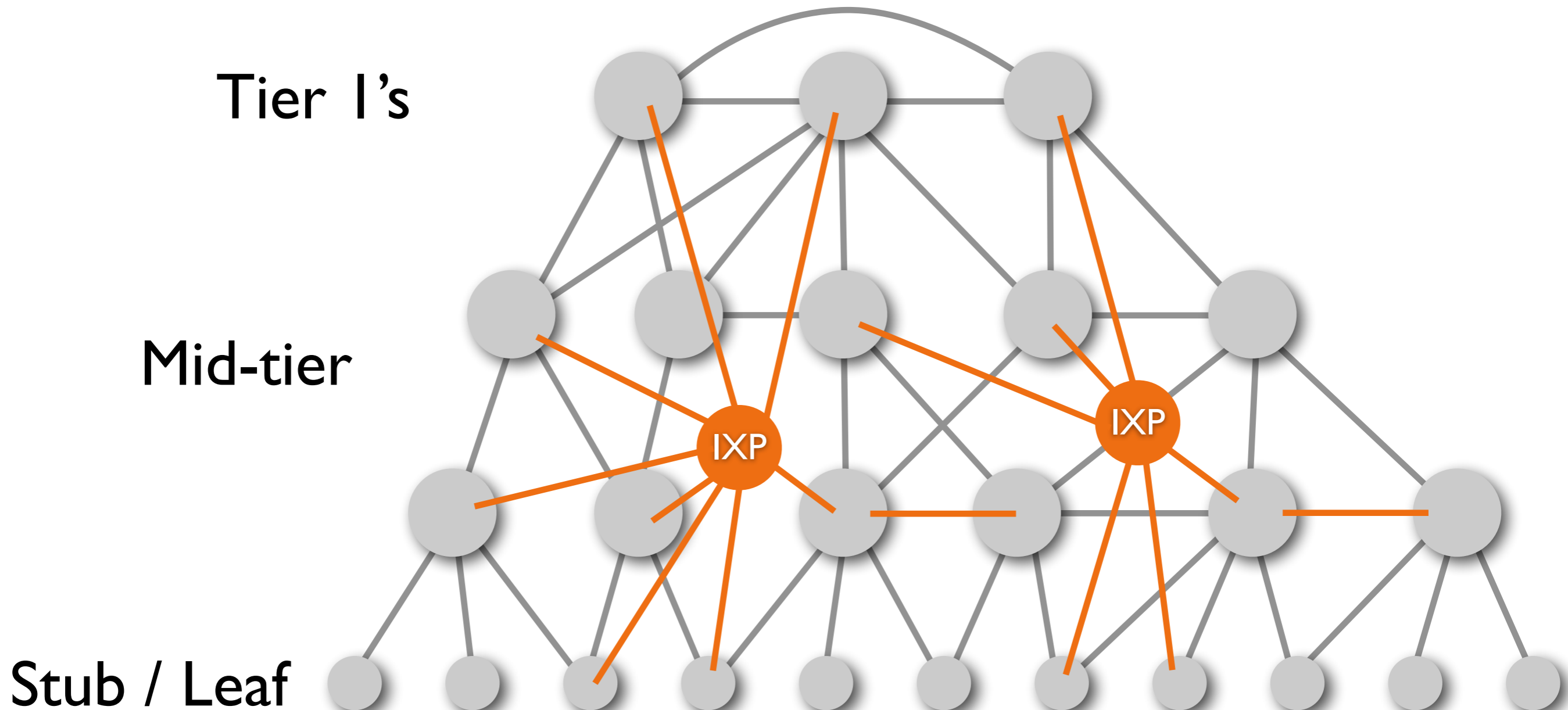
Significant and increasing peering at lower tiers

Significant peering

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

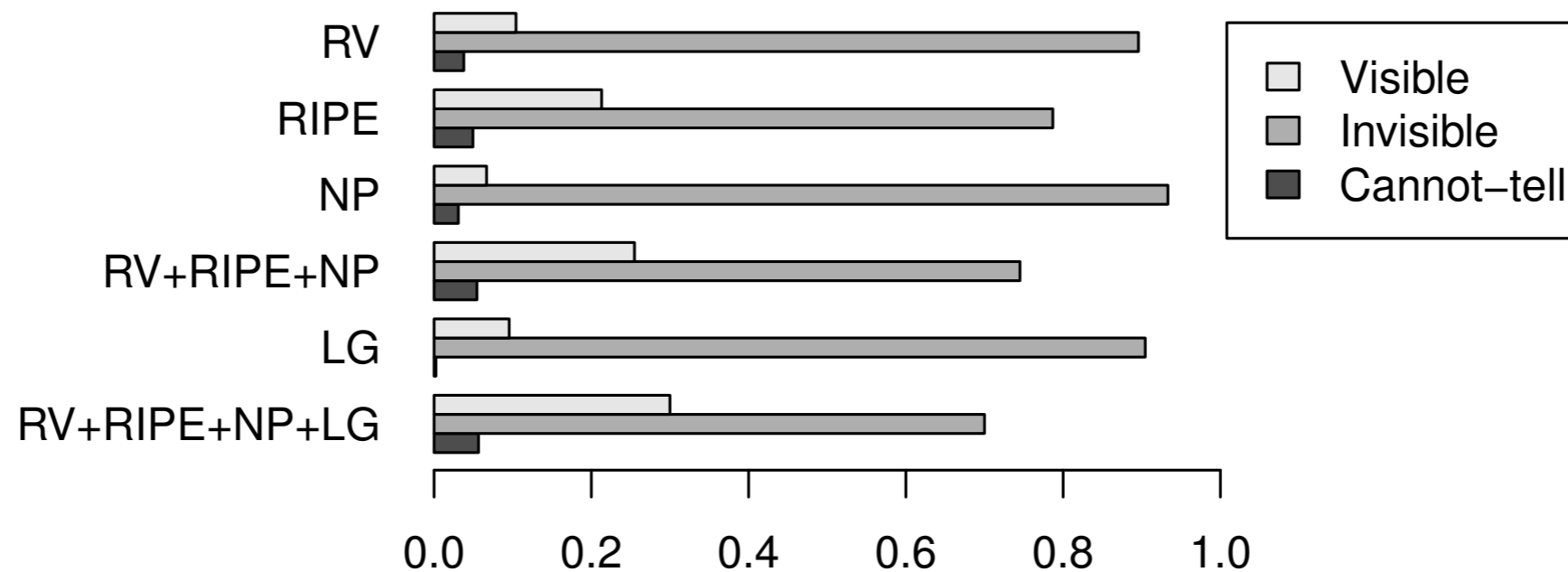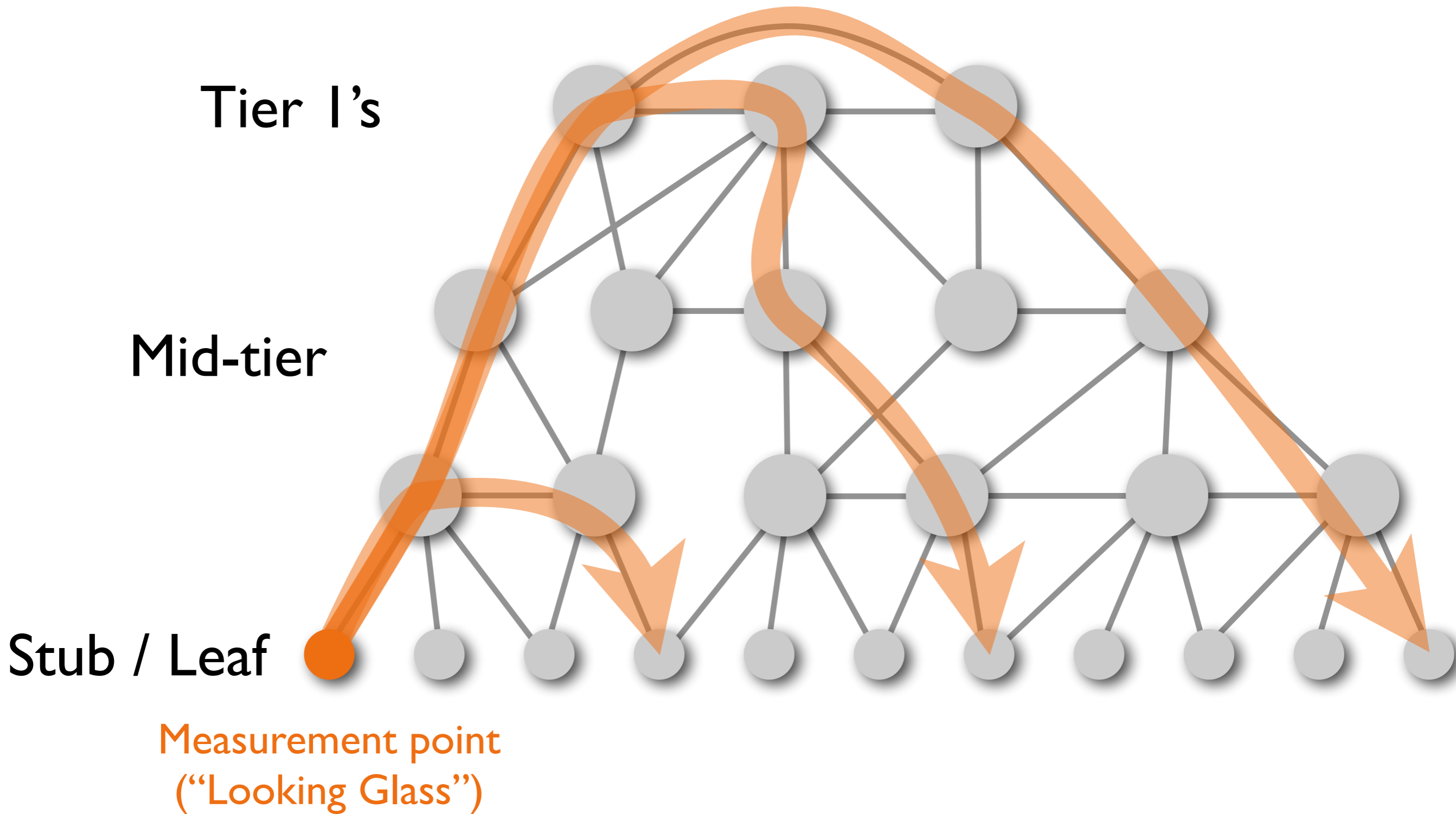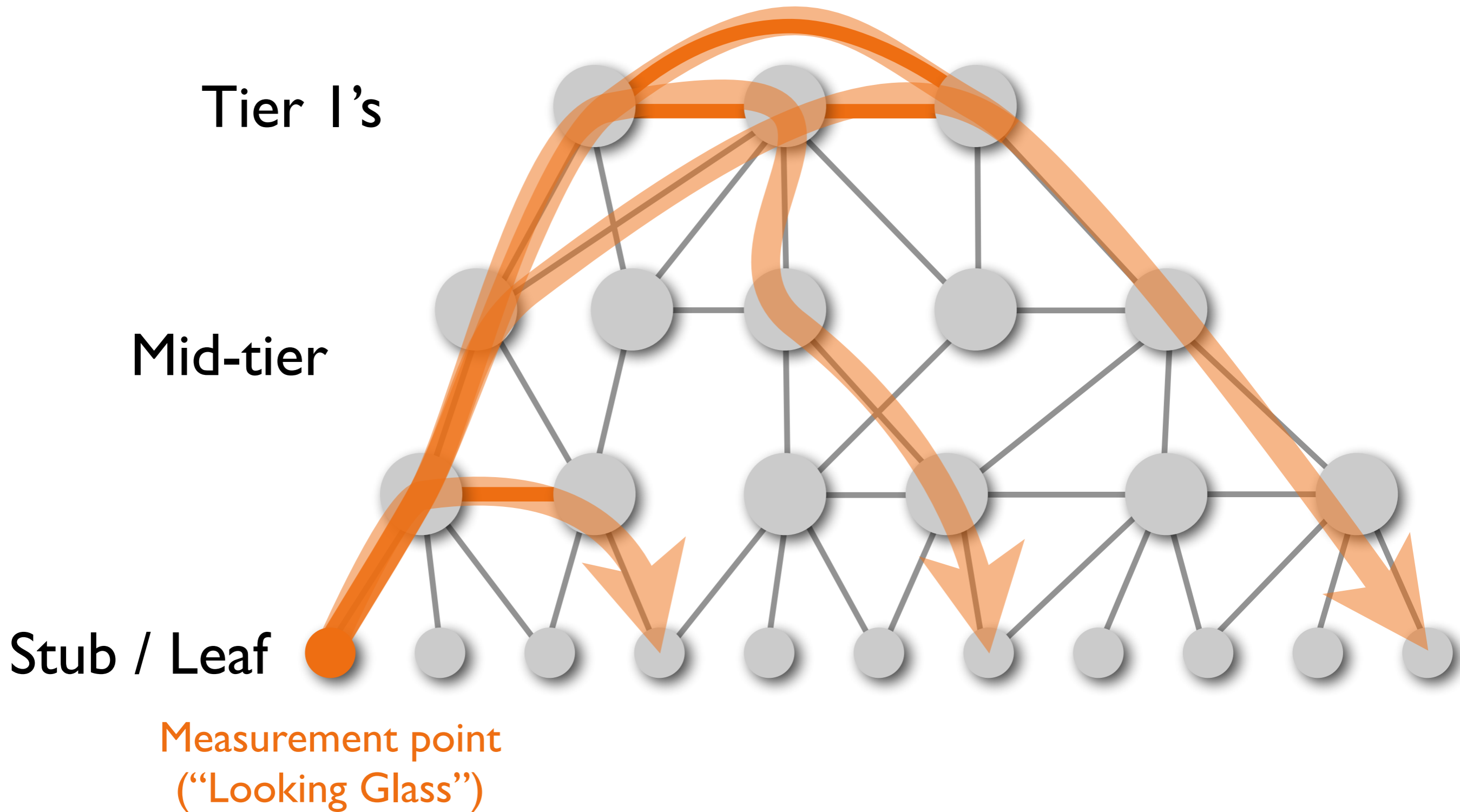## Past measurements missed these peerings

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

Why measurements miss so much
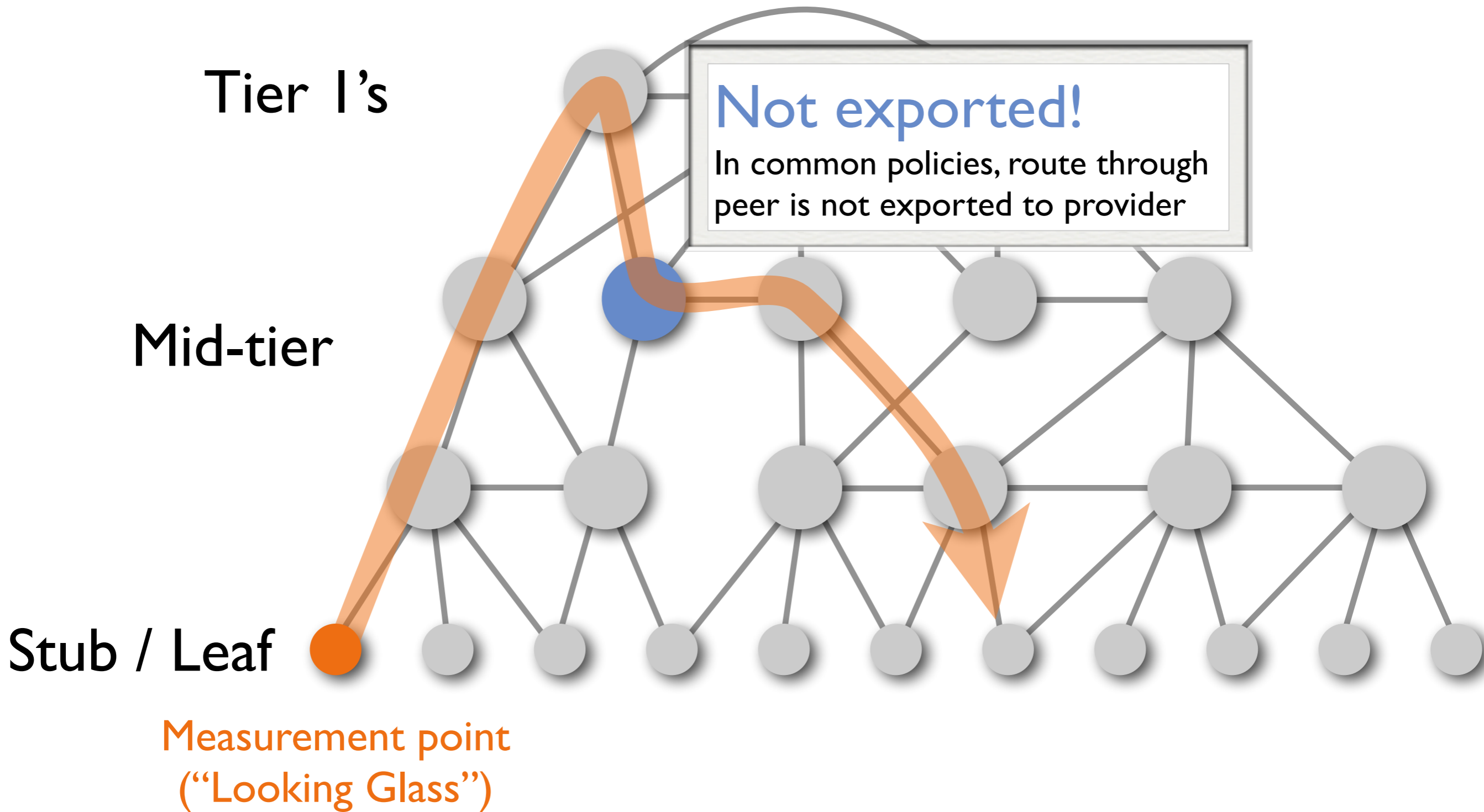
Tier 1's

Mid-tier

Stub / Leaf

Measurement point
("Looking Glass")

Tier 1's

Mid-tier

Stub / Leaf

Measurement point
("Looking Glass")

Tier 1's

Mid-tier

Stub / Leaf

Measurement point
("Looking Glass")

Tier 1's

Mid-tier

Stub / Leaf

Not exported!

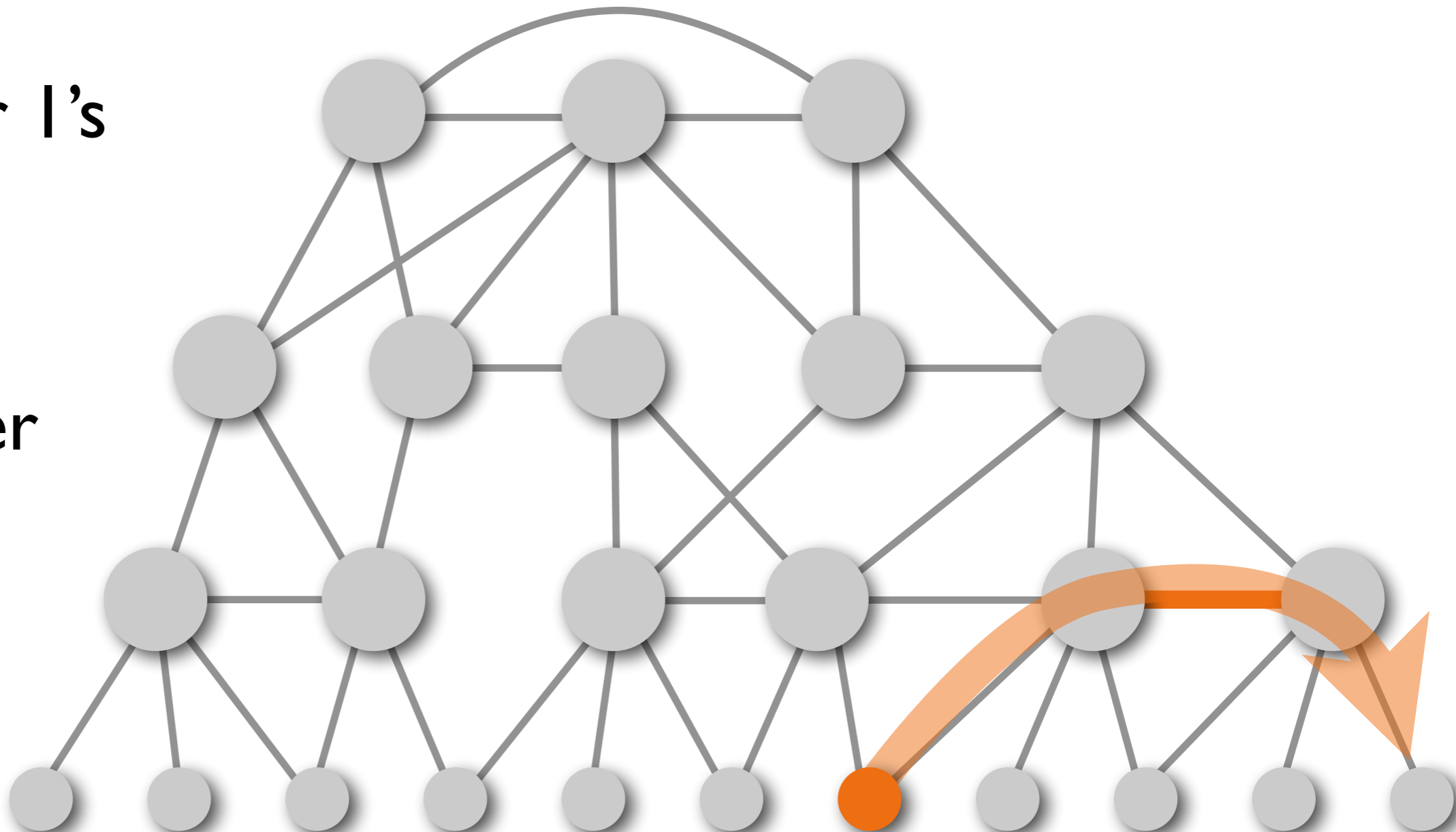In common policies, route through peer is not exported to provider

Measurement point
("Looking Glass")

Tier 1's

Mid-tier

Stub / Leaf

Measurement point
("Looking Glass")

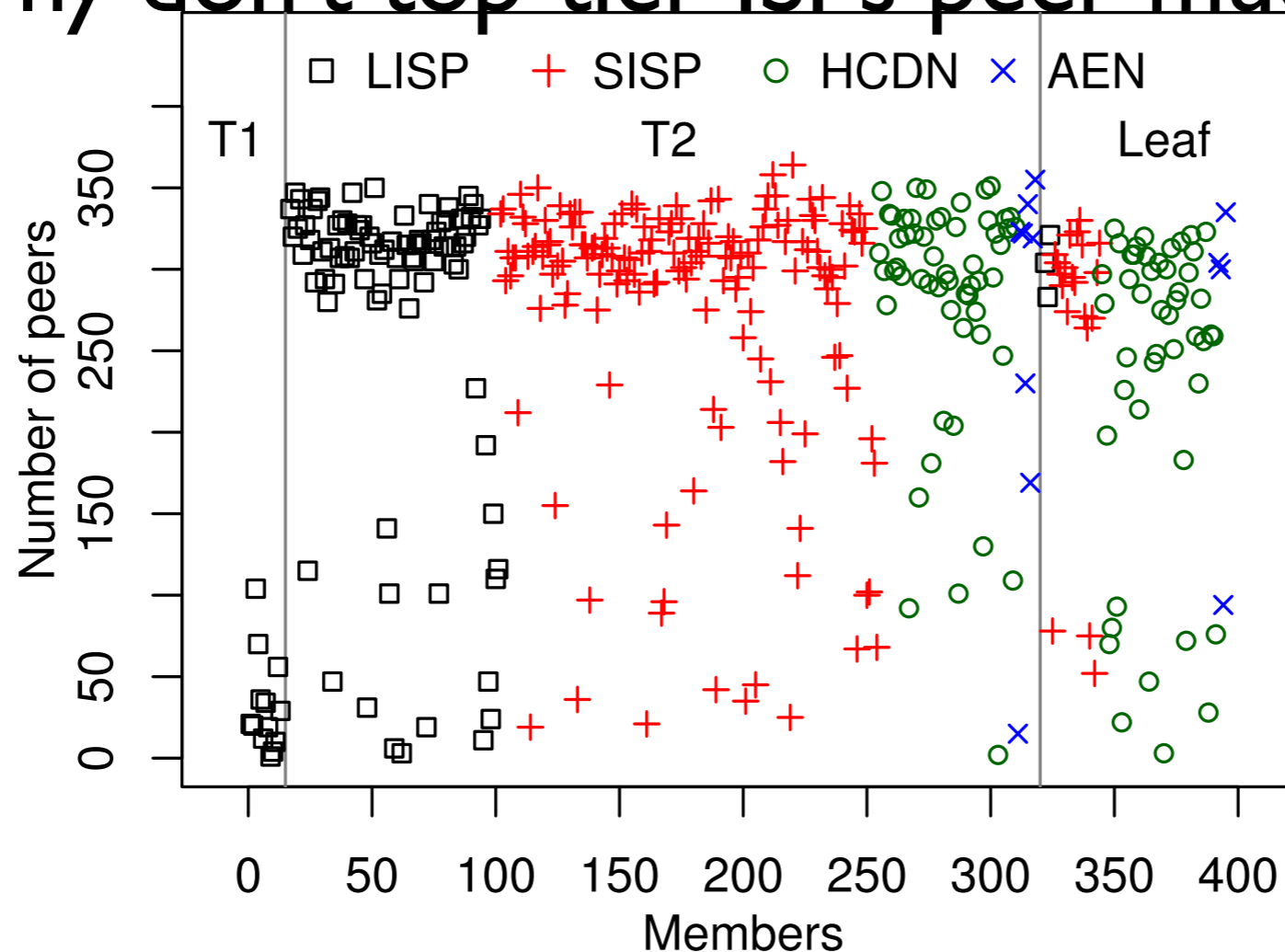To see peer-peer link, *both* probe source & dest. must be in localized area

What's the purpose of an IXP?
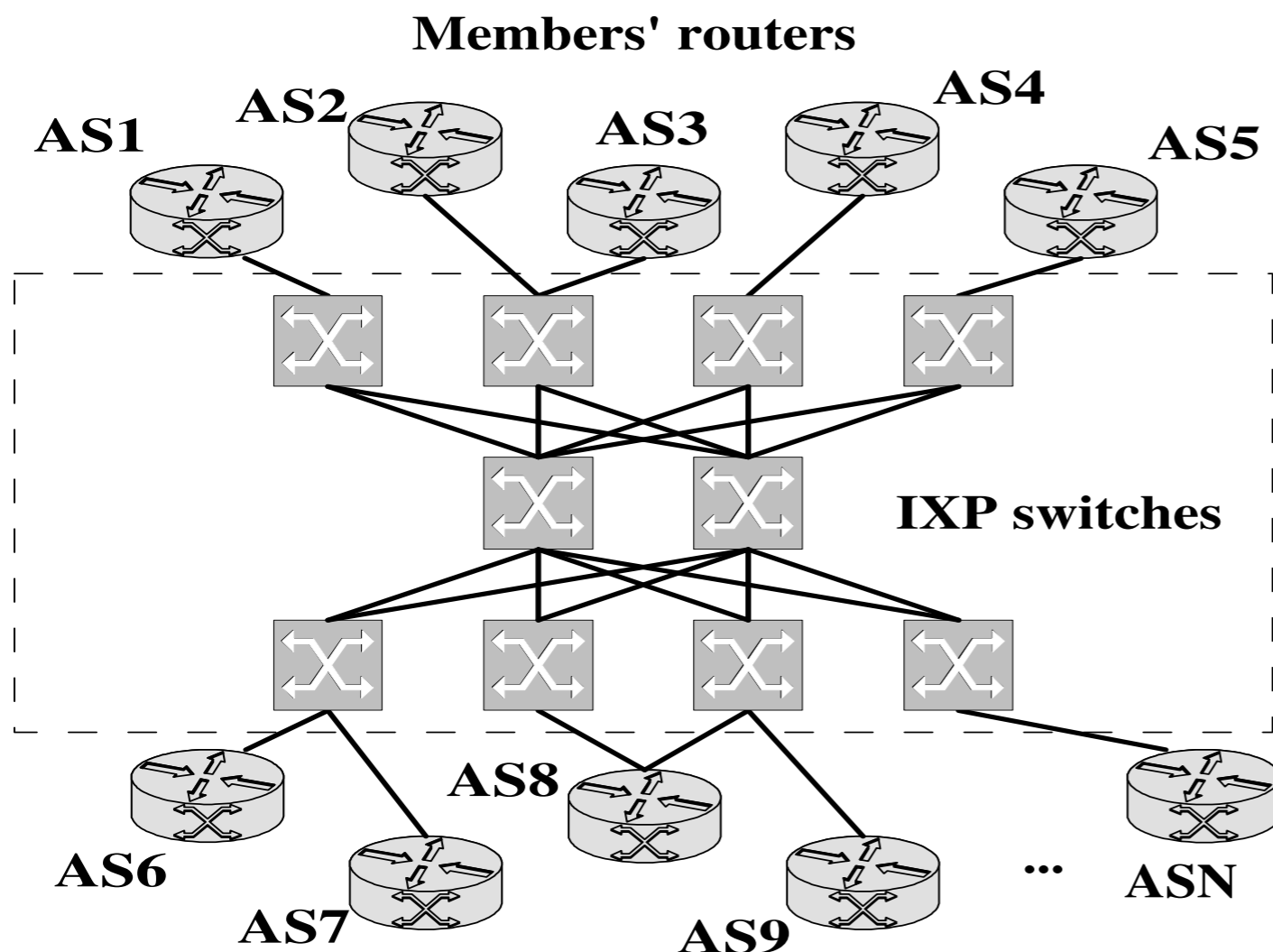
- "Metcalf's law": value of net is $O(n^2)$ when n participants

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

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]

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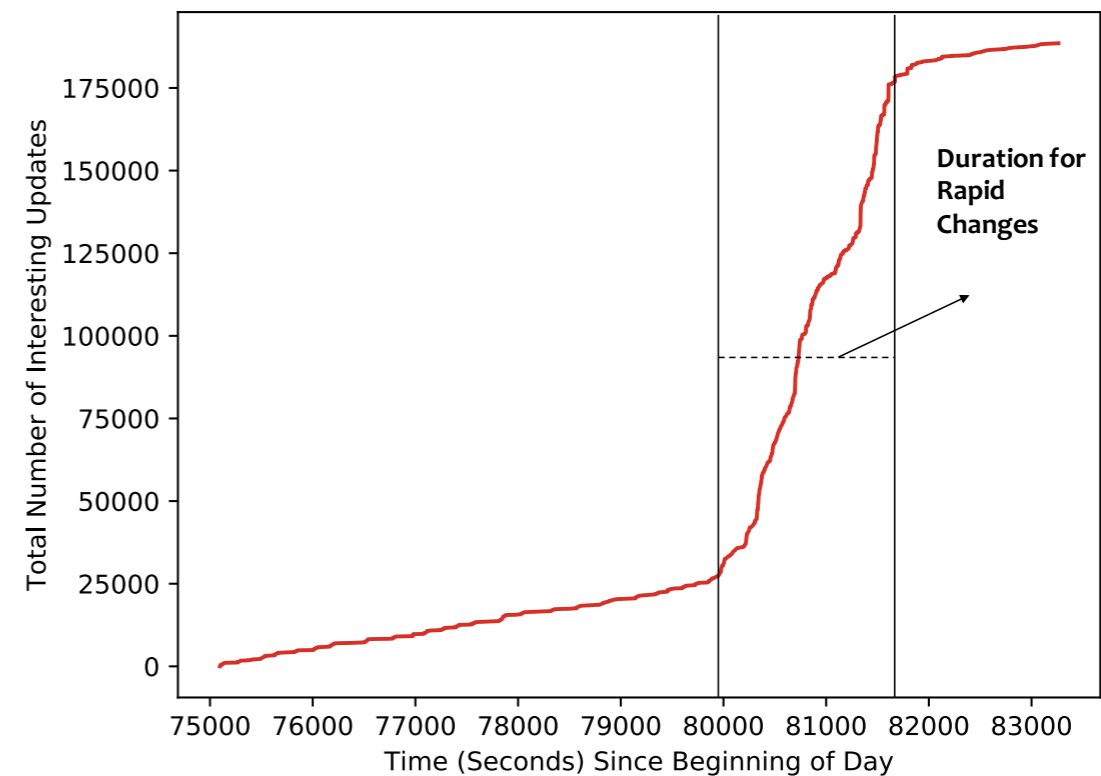- If you treat the IXP as an AS "in the middle", what fraction of the links are observable?

Announcements etc.

Assignment 1 and project comments will be returned today

Anthony

Faria

What's to come
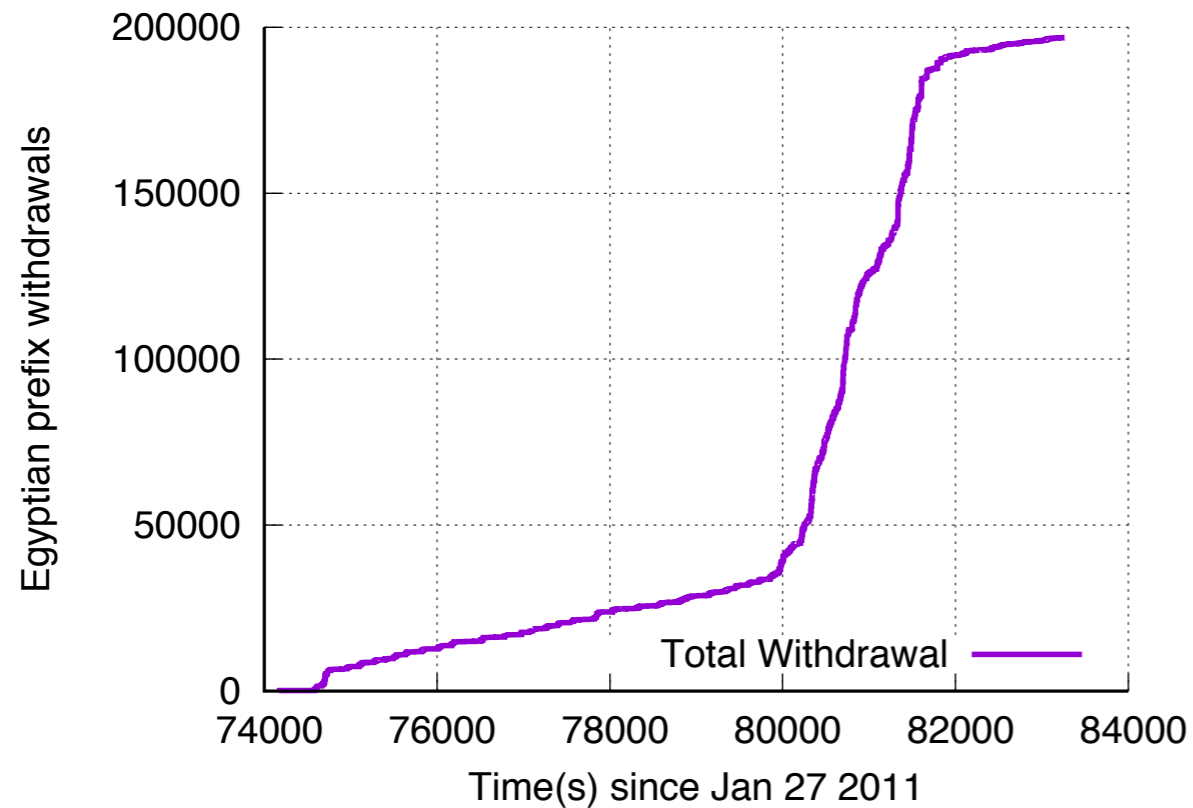
Next: Part Two of the course: Grand Challenges

- complexity & programmability: SDN
- reliability
- selfishness
- security & privacy

Project midterm presentations March 27, 29

Schedule updated through end of March