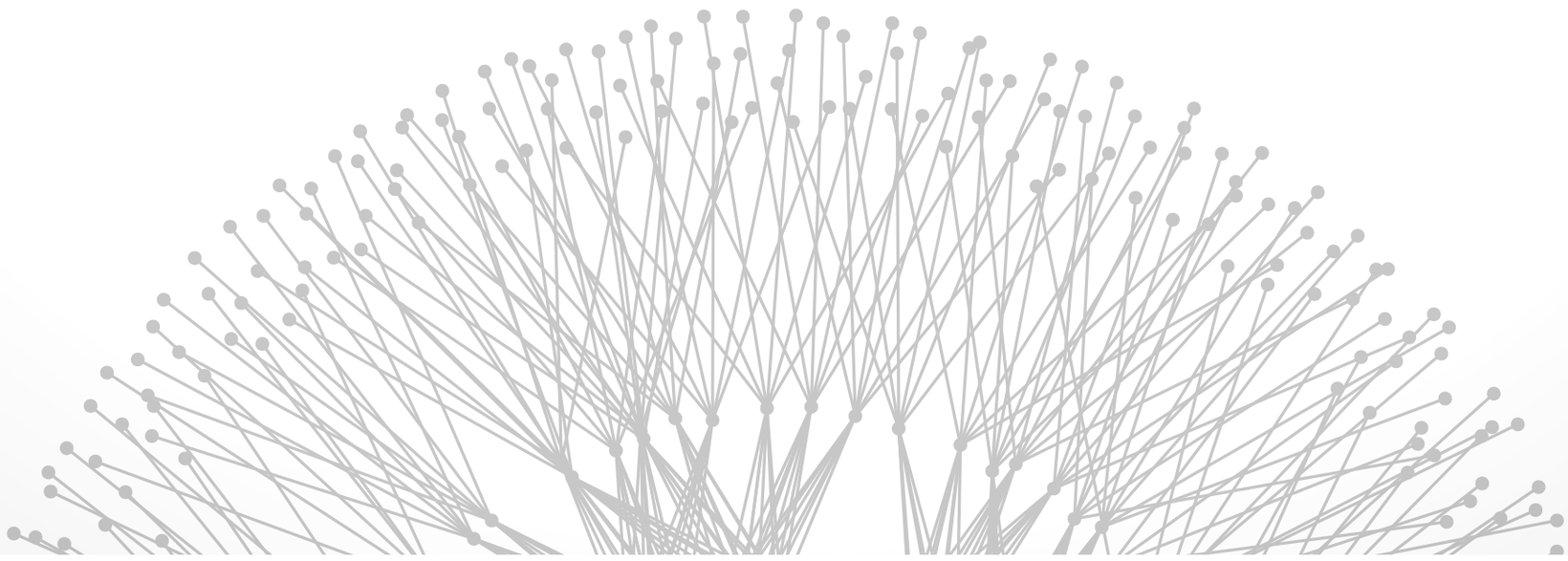
Interdomain Routing and Connectivity

Brighten Godfrey

CS 538 September 25 2012



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
- Attacks

Harder to change architecture

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

Interdomain routing

BGP: Border Gateway Protocol

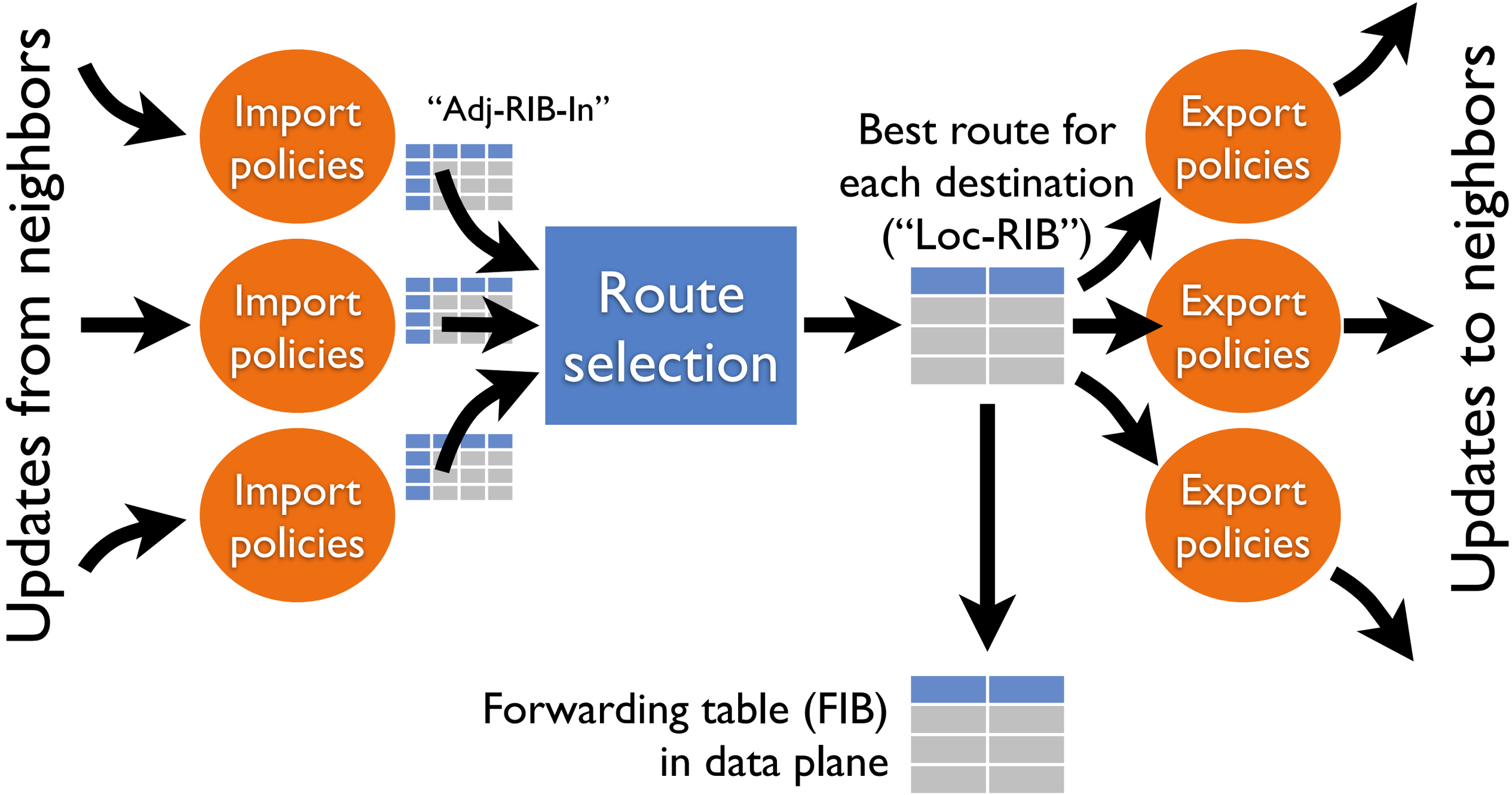
Distance vector variant

- Remember path instead of distance
- Hence, "path vector" instead of "distance vector"

Why path vector?

- Avoid DV's transient loops; but more importantly...
- Support policies: can pick any path offered by neighbors, not necessarily the shortest (Link State cannot)
- Support privacy: path choice policy is applied locally, not announced globally

BGP: The picture at one router



Route selection process

Import policies

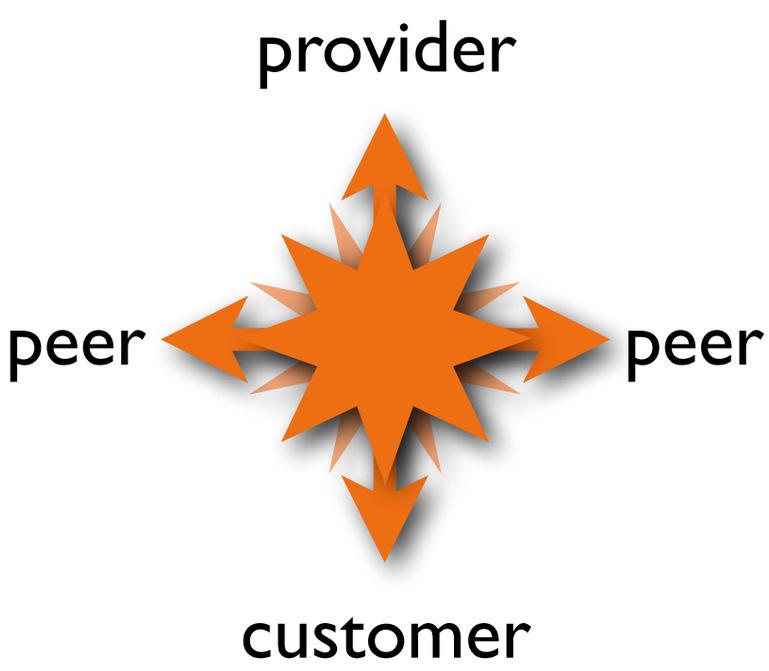
Step	Attribute	Controlled by local or neighbor AS?
1.	Highest LocalPref	local
2.	Lowest AS path length	neighbor
3.	Lowest origin type	neither
4.	Lowest MED	neighbor
5.	eBGP-learned over iBGP-learned	neither
6.	Lowest IGP cost to border router	local
7.	Lowest router ID (to break ties)	neither

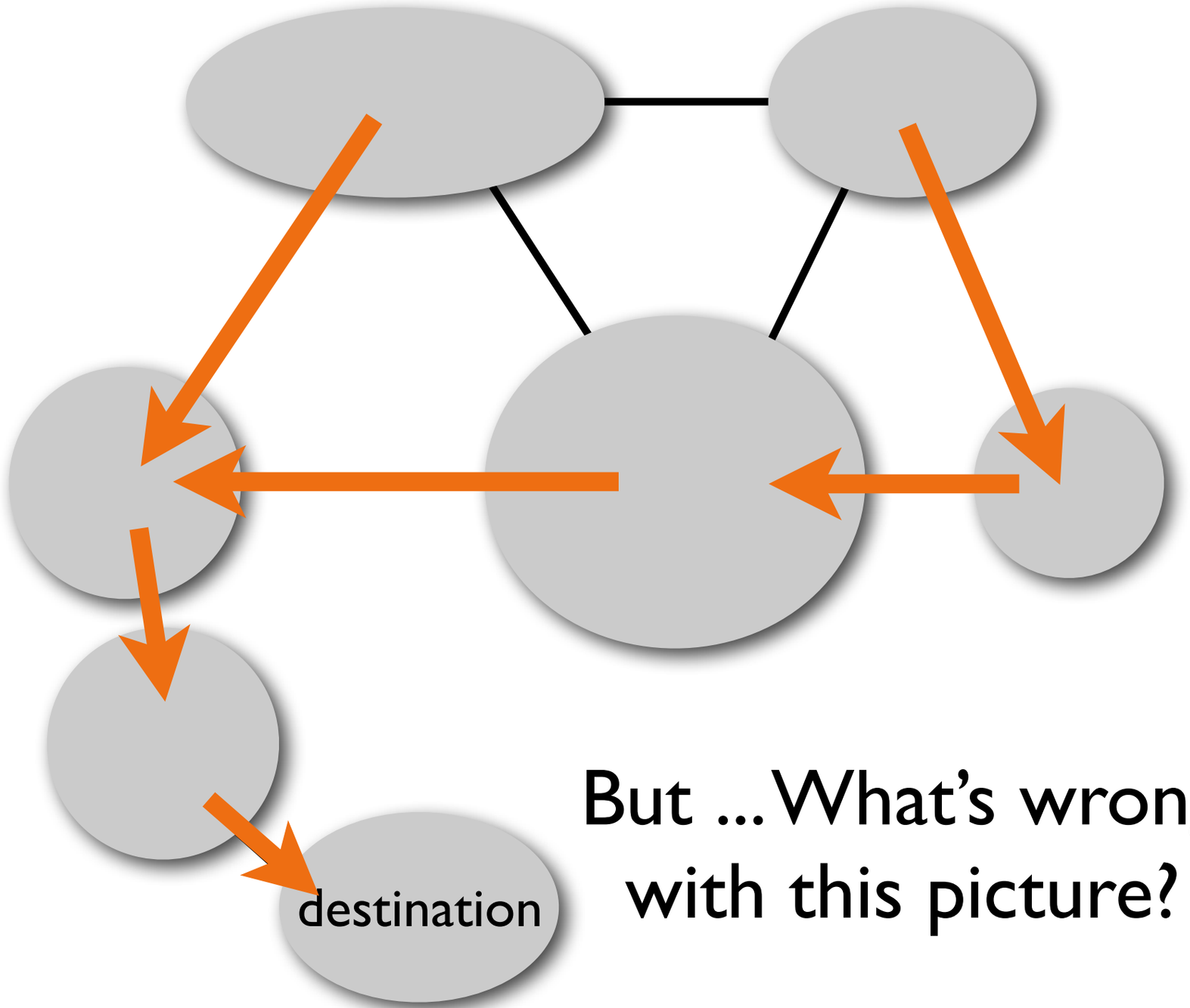
[Caesar, Rexford, IEEE Network Magazine, 2005]

This process is extended in many real implementations.

Common policies

Route selection: prefer customer over peer over provider

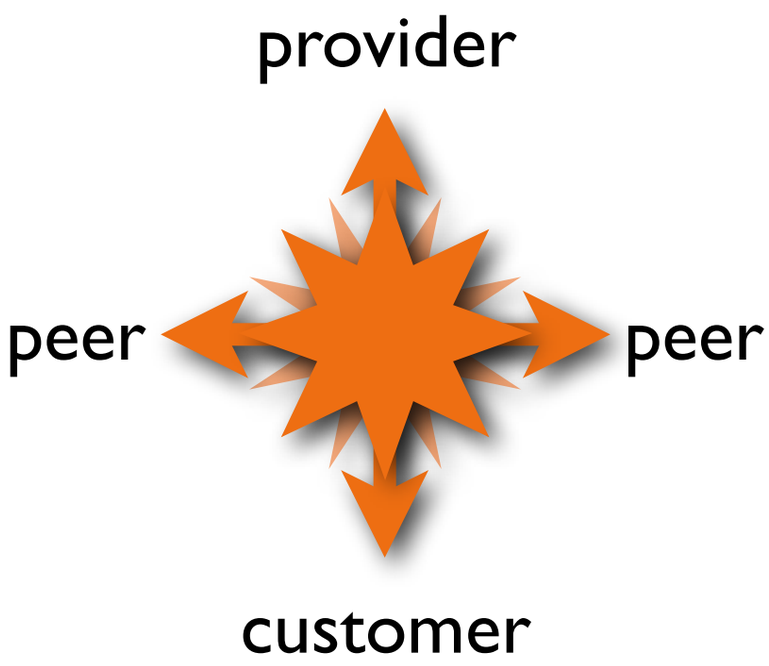


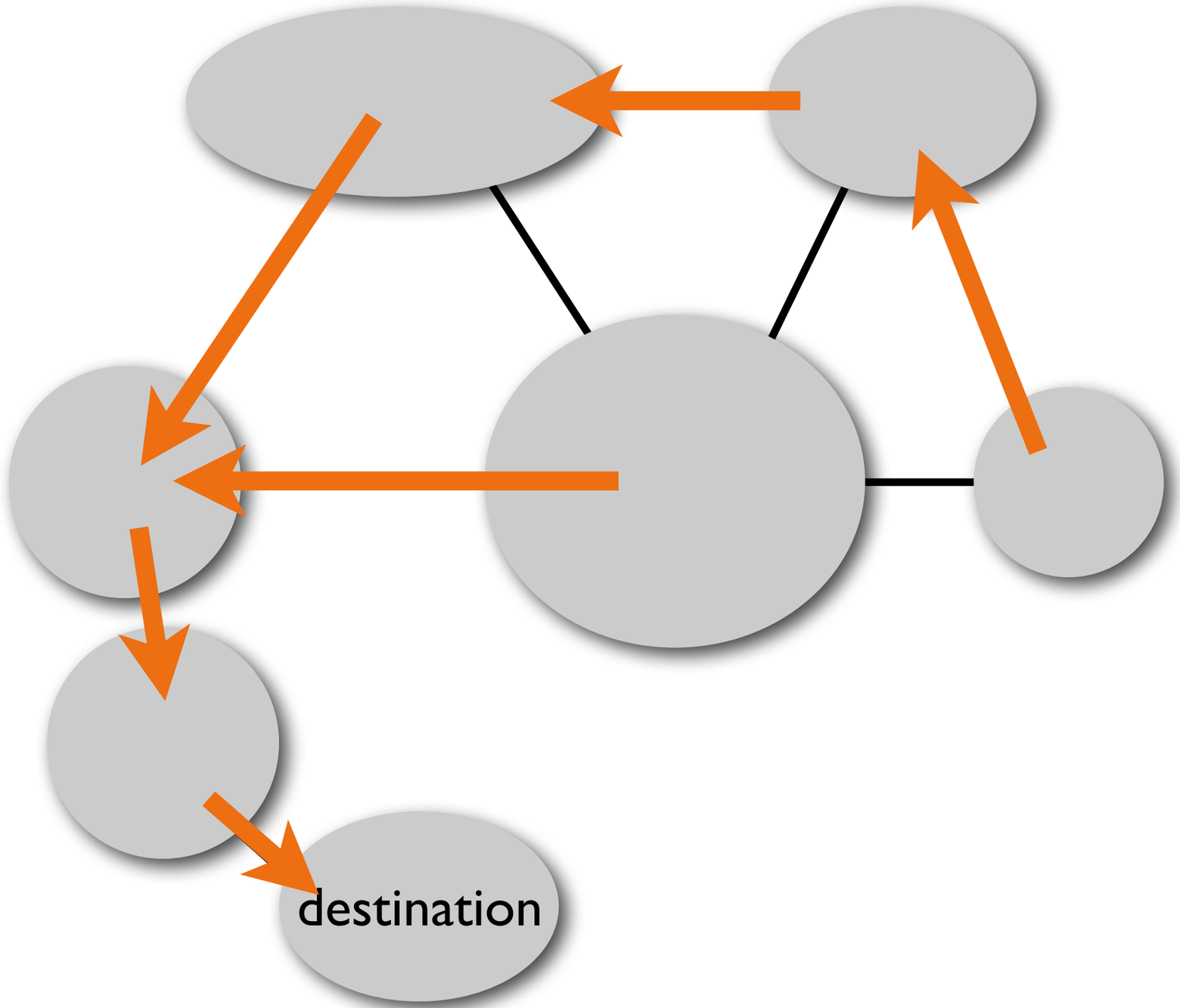


But ... What's wrong with this picture?

Common policies

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



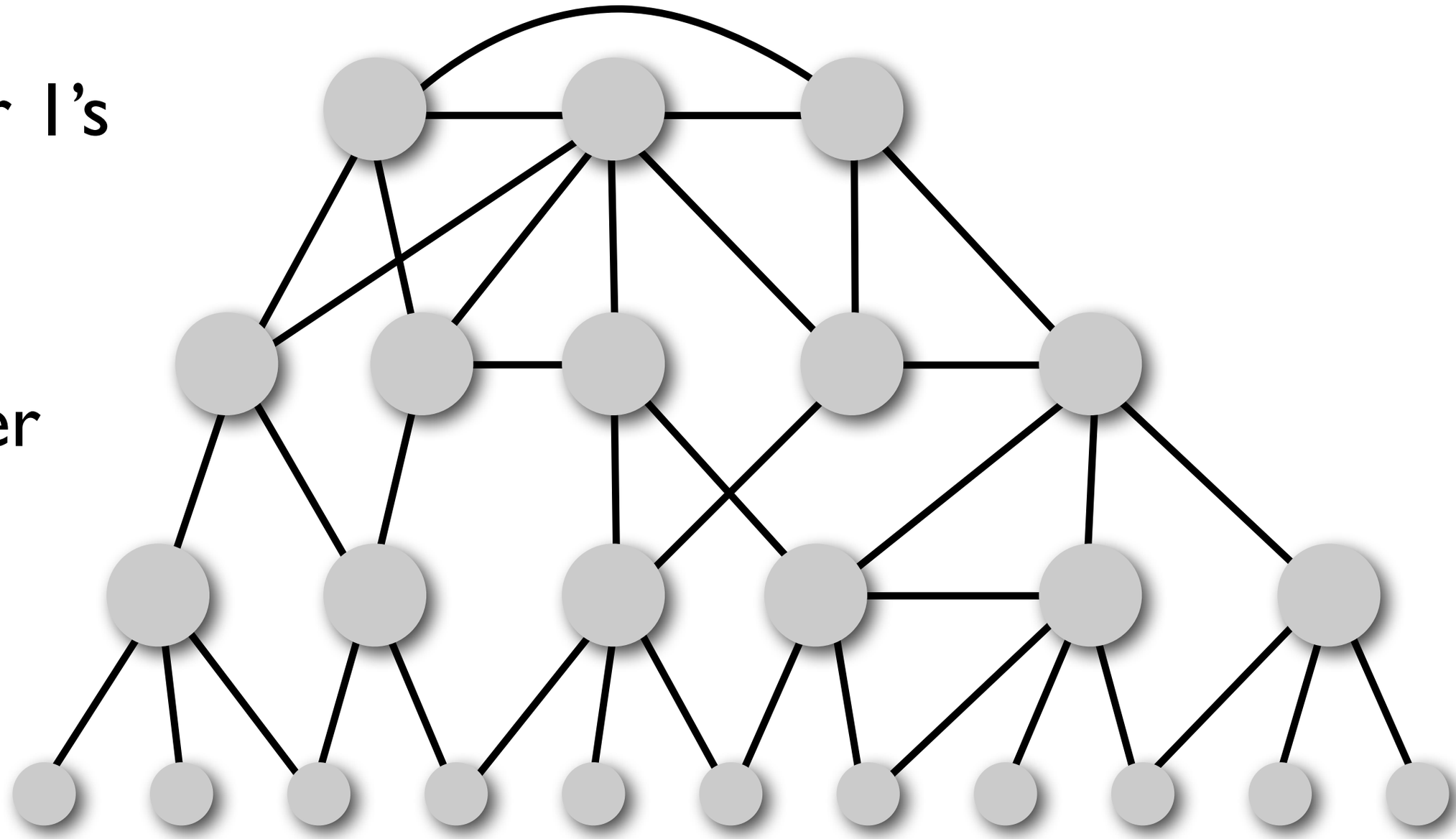


Interconnection: Traditional view

Tier 1's

Mid-tier

Stub / Leaf



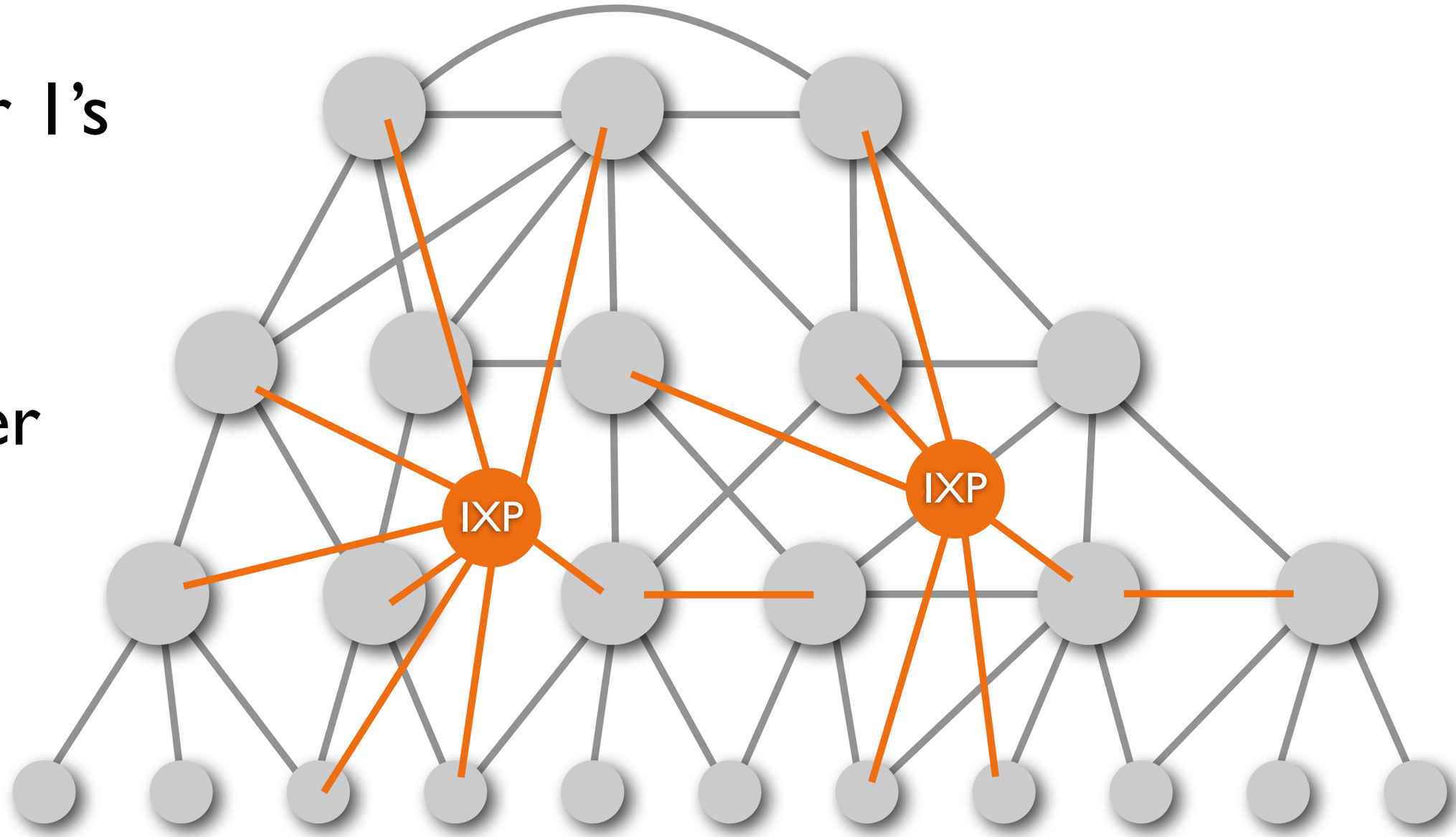
Hierarchical, limited peering at lower tiers

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!

These peerings missed in past measurements

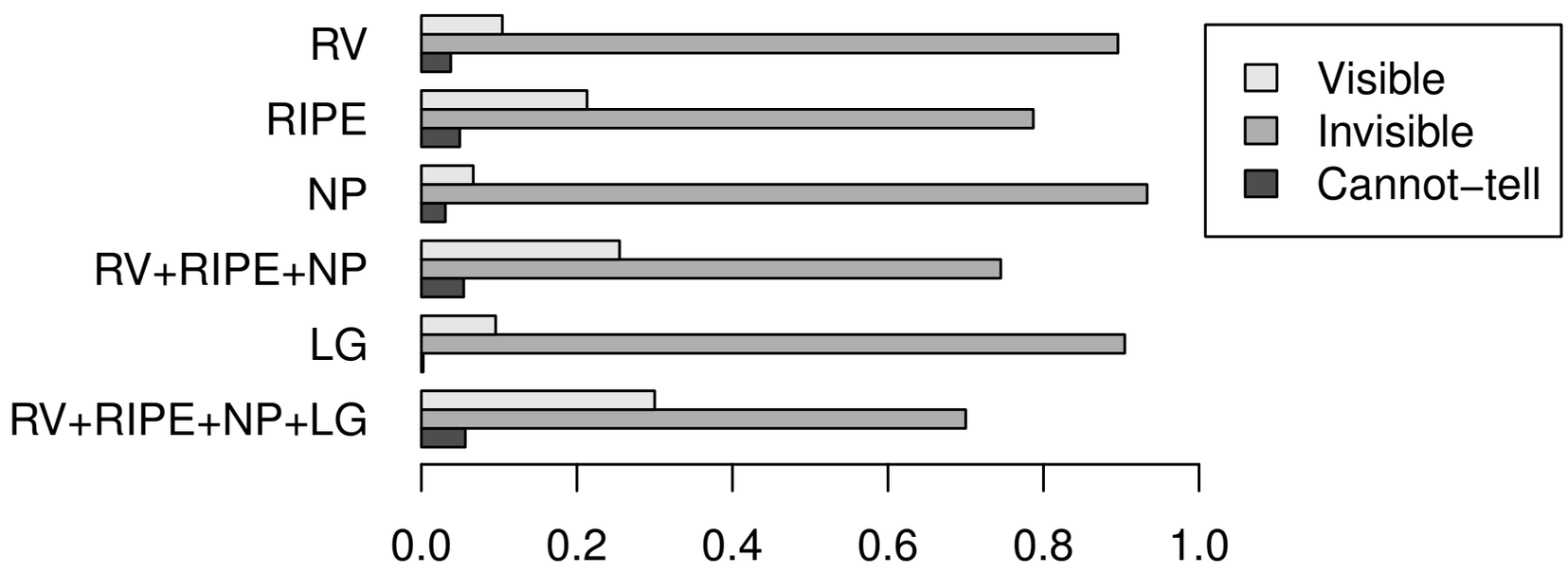


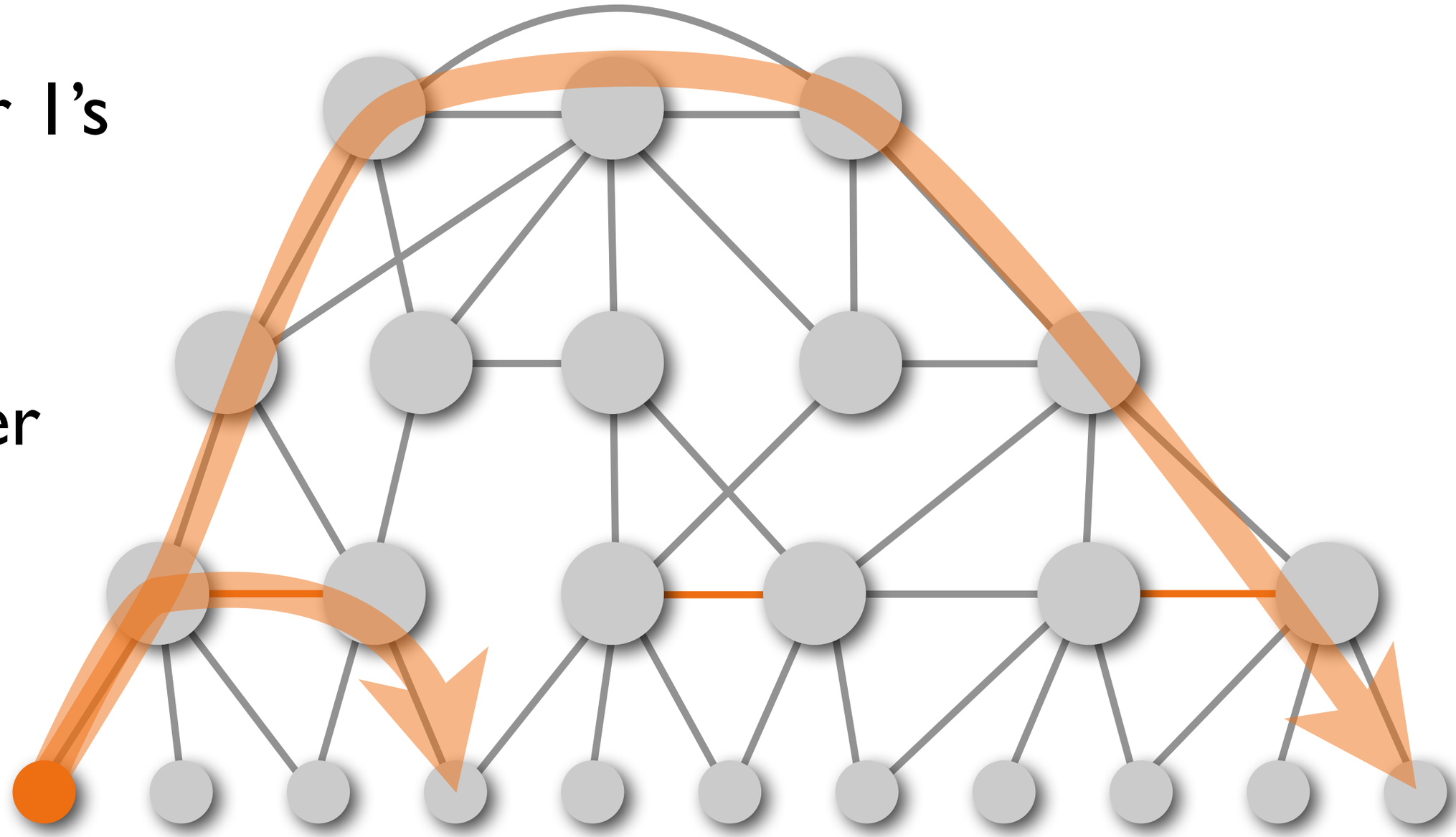
Figure 2: Peering links and visibility in control/data plane (normalized by number of detected P-P links).

Why measurements miss so much

Tier 1's

Mid-tier

Stub / Leaf



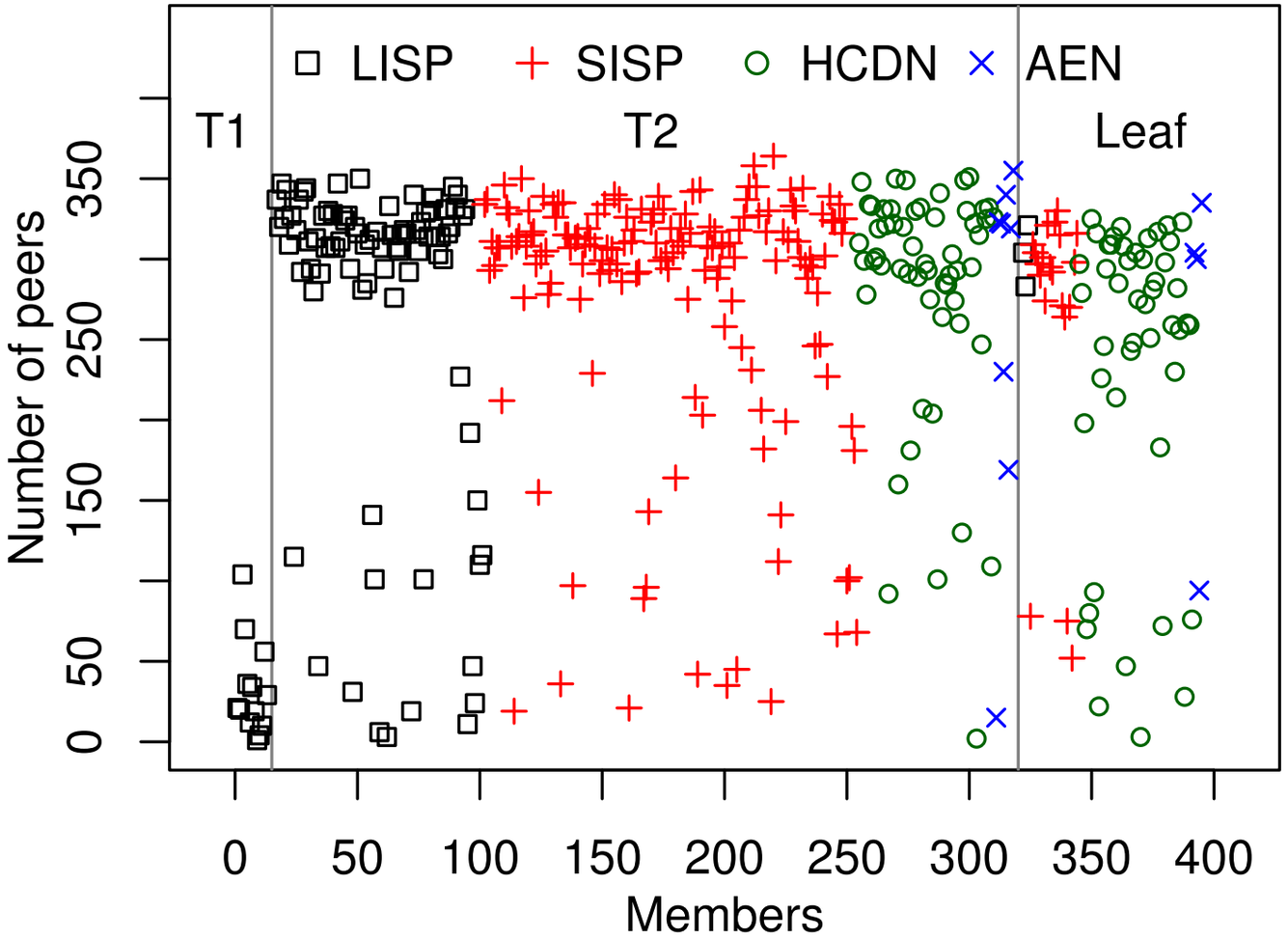
Measurement point
("Looking Glass")

Paper discussion

What's the purpose of an IXP?

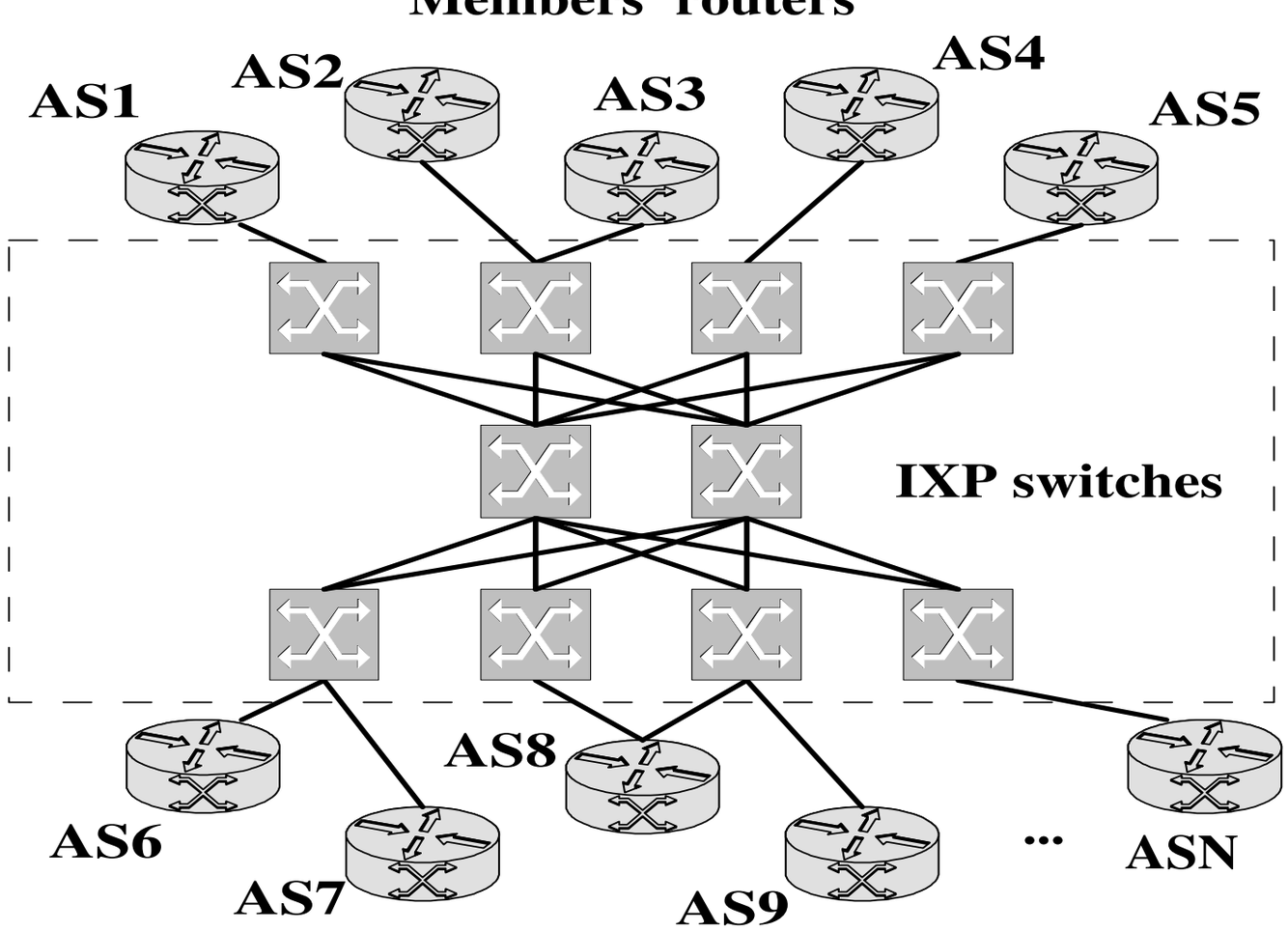
- "Metcalf's law": value of net is $O(n^2)$ when n participants

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]

What's to come

Project proposals

- Due tonight, 11:59 pm, plain text email to Brighten
- Be sure to read spec in Syllabus and include related work
- Comments back to you next week

Part Two of the course: Grand Challenges

- scalability
- reliability
- selfishness
- security & privacy
- complexity