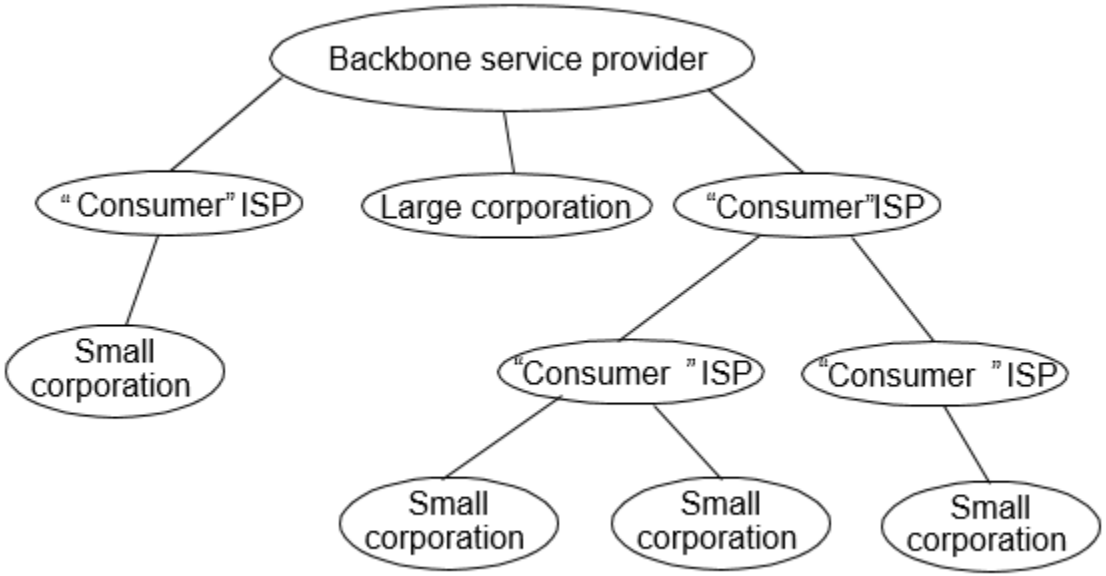
Inter-domain Routing

Outline

- Border Gateway Protocol

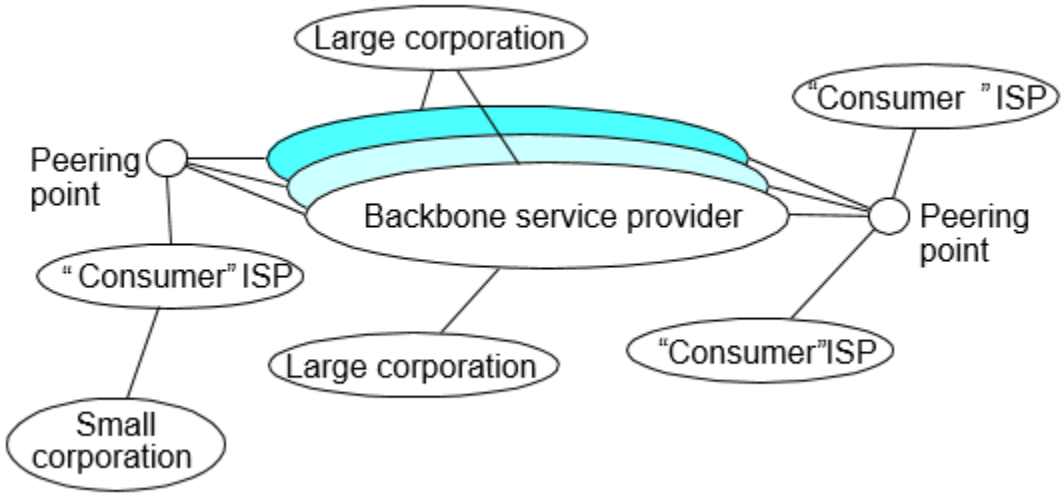
Internet Structure

Original idea



Internet Structure

Today



Route Propagation in the Internet

- Autonomous System (AS)
  - corresponds to an administrative domain
  - examples: University, company, backbone network
  - assign each AS a 16-bit number
- Two-level route propagation hierarchy
  - interior gateway protocol (each AS selects its own)
  - exterior gateway protocol (Internet-wide standard)
- Routes information is propagated at various levels
  - hosts know local router
  - local routers know site routers
  - site routers know core router
  - core routers know everything

Popular Interior Gateway Protocols

- RIP: Route Information Protocol
  - distributed with BSD Unix
  - distance-vector algorithm
  - based on hop-count (infinity set to 16)
- OSPF: Open Shortest Path First
  - recent Internet standard
  - uses link-state algorithm
  - supports load balancing
  - supports authentication

BGP-4: Border Gateway Protocol

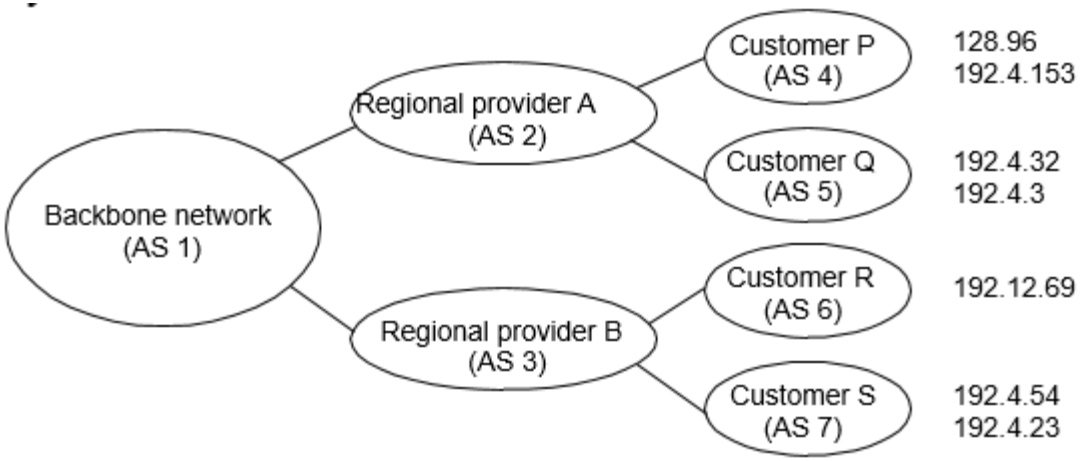
- BGP-1 developed in 1989 to address problems with EGP.
- Assumes Internet is an arbitrarily interconnected set of ASs
- AS traffic types
  - Local
    - starts or ends within an AS
  - Transit
    - passes through an AS
- AS Types
  - stub AS: has a single connection to one other AS
    - carries local traffic only
  - multihomed AS: has connections to more than one AS
    - refuses to carry transit traffic
  - transit AS: has connections to more than one AS
    - carries both transit and local traffic

BGP-4 contd.

- Each AS has:
  - one or more border routers
    - Handles inter-AS traffic
  - At least one BGP *speaker* for an AS that participates in routing
  - BGP speaker establishes BGP sessions with peers and advertises:
    - local network names
    - other reachable networks (transit AS only)
    - gives *path* information
    - withdrawn routes
- BGP goal: find loop free paths between ASs
  - Optimality is secondary goal
  - It's neither a distance-vector nor a link-state protocol
- Hard problem
  - Internet's size (~12K active ASs) means large tables in BGP routers
  - Autonomous domains mean different path metrics
  - Need for flexibility

BGP Example

- Speaker for AS2 advertises reachability to P and Q
  - network 128.96, 192.4.153, 192.4.32, and 192.4.3, can be reached directly from AS2



- Speaker for backbone advertises
  - networks 128.96, 192.4.153, 192.4.32, and 192.4.3 can be reached along the path (AS1, AS2).
- Speaker can cancel previously advertised paths

Some BGP details

- Path vectors are most important innovation in BGP
  - Enables loop prevention in complex topologies
  - If AS sees itself in the path, it will not use that path
- Routes can be aggregated
  - Based on CIDR (classless) addressing
- Routes can be filtered
- Runs over TCP
- ASes can apply a variety of policies

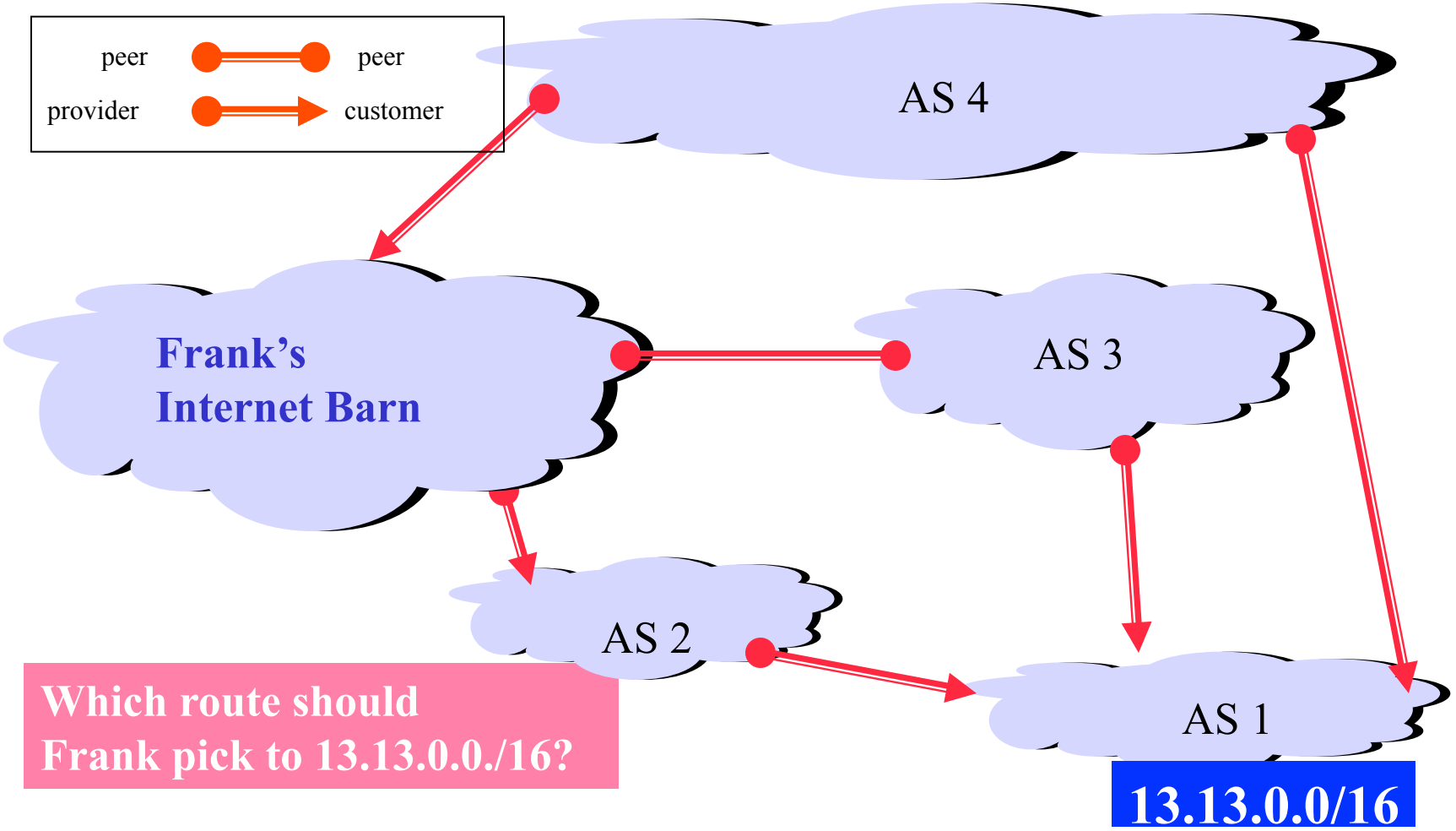
BGP in practice

- 10-20 “backbone” ASs which are fairly richly connected to each other
  - Peers
- Other “lower tier” ASs hang off the backbone networks -> Customers
  - Some of them may also connect with each other at peering points → Peers
- Corporations connect as Customers to lower tier ASs or to backbone ASs depending on their need/ willingness to pay

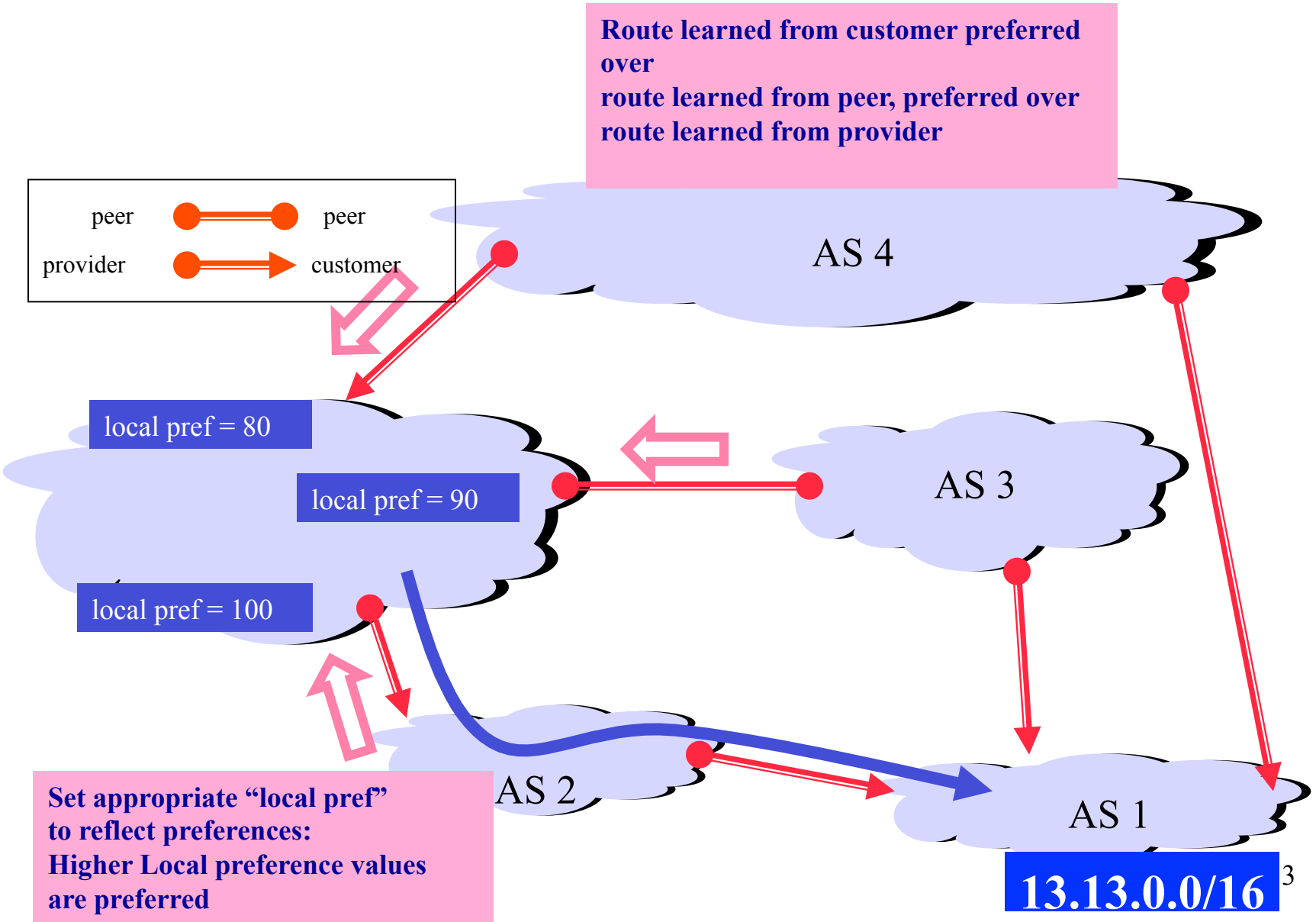
Policy with BGP

- BGP provides capability for enforcing various policies
- Policies are **not** part of BGP: they are provided to BGP as configuration information
- **Enforces** policies by
  - *Choosing appropriate paths* from multiple alternatives
  - *Controlling advertisement* to other AS's

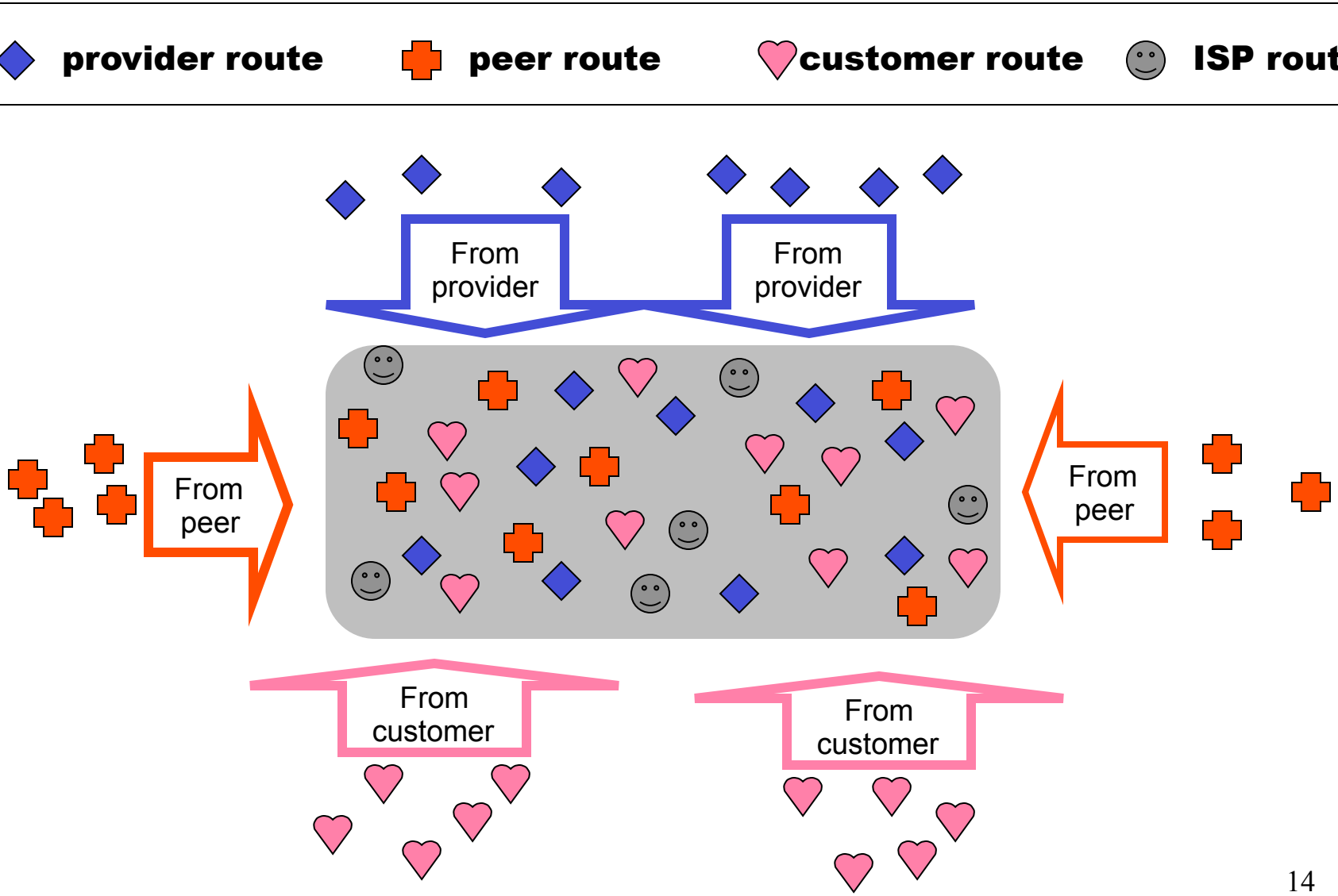
Illustrating BGP Policies



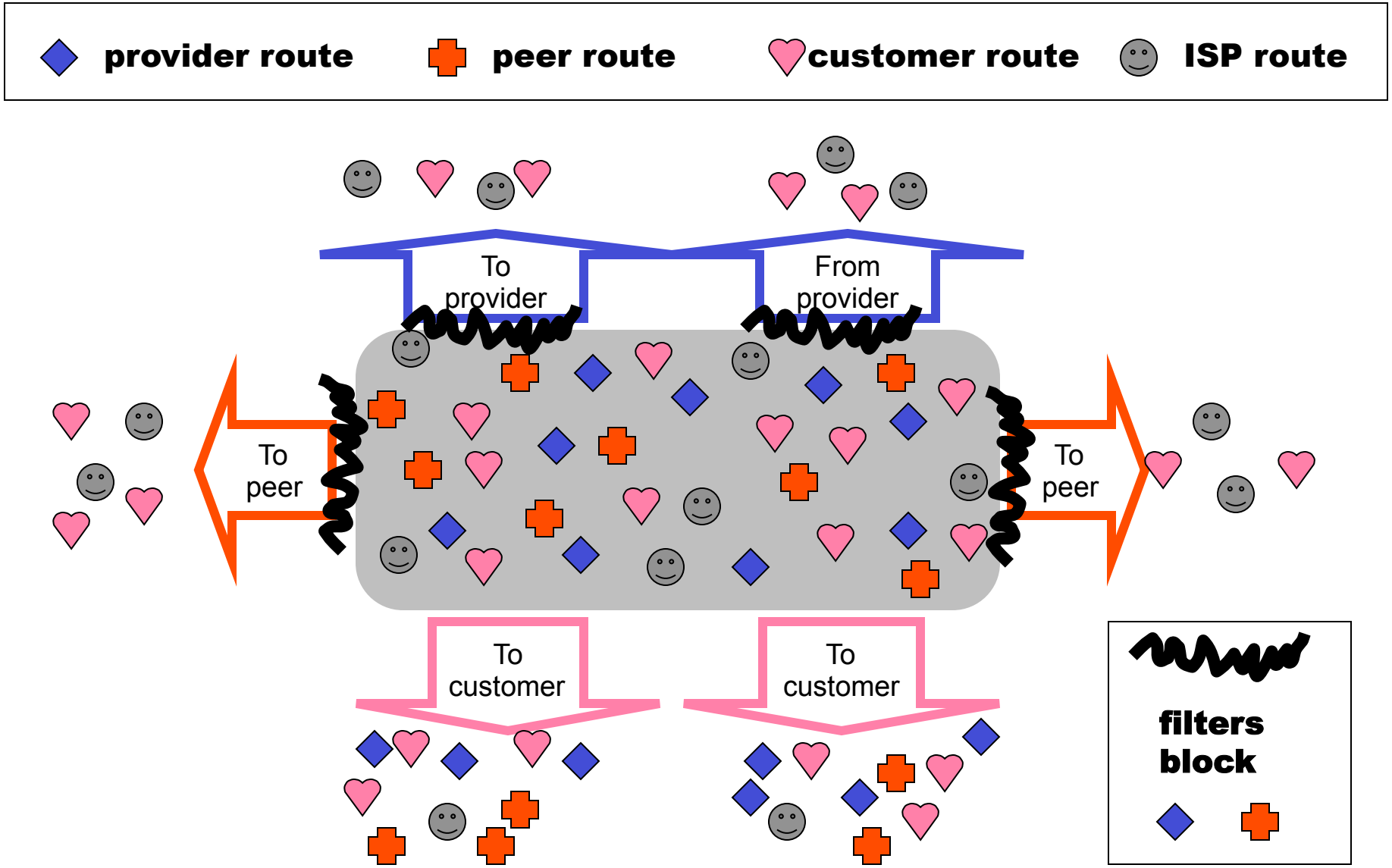
Policy I: Prefer Customer routing



Policy II: Import Routes



Policy II: Export Routes



BGP Policies Summarized

Route learned From ↓	Advertise to → Customer	Provider	Peer
Customer	✓	✓	✓
Provider	✓	✗	✗
Peer	✓	✗	✗

Export Policies: Valley-Free Routes

- "Valley-free" routing
  - Label links as (+1, 0, -1) for provider, peer and customer
  - In any *valid* path should only see sequence of +1, followed by at most one 0, followed by sequence of -1
  - Why?
    Export incentives

- How to make these choices?
  - Prefer-customer routing: LOCAL_PREF
  - Valley-free routes: control route advertisements (see previous slides)