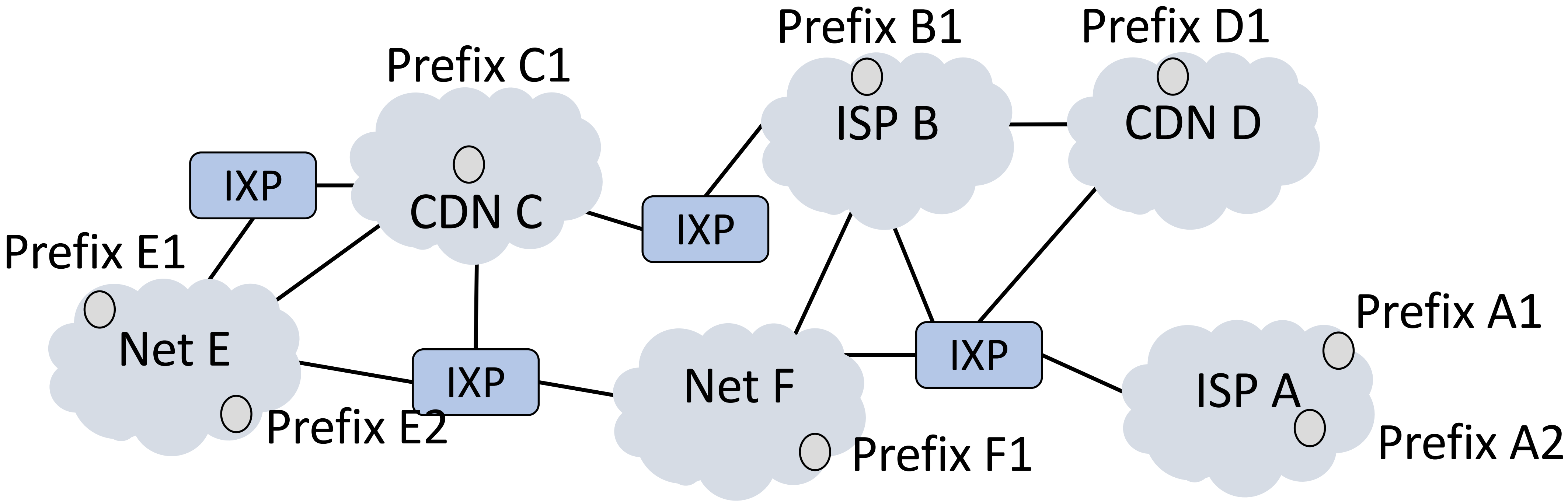
Border Gateway Protocol (BGP)

Structure of the Internet

- Networks (ISPs, CDNs, etc.) group with IP prefixes
- Networks are richly interconnected, often using IXPs



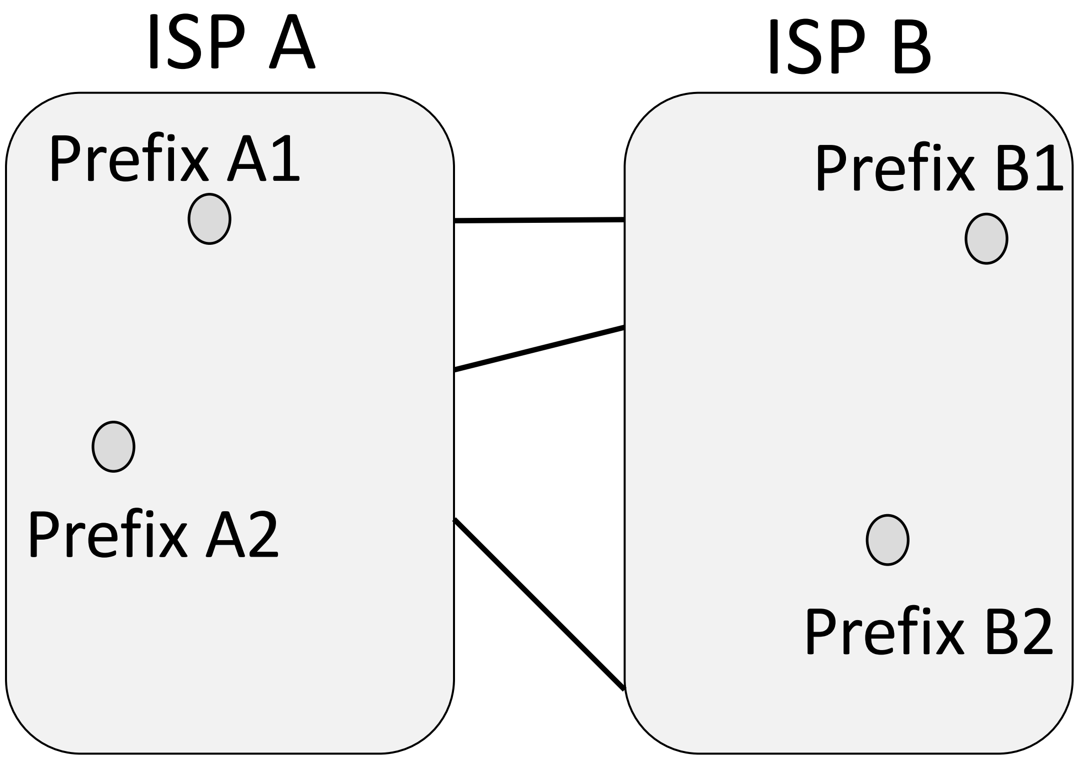
Internet-wide Routing

Two requirements not met by simple routing

1. Incorporating policy decisions
   - Letting parties choose routes to suit their own needs
2. Scaling to very large networks
   - Prefix aggregation

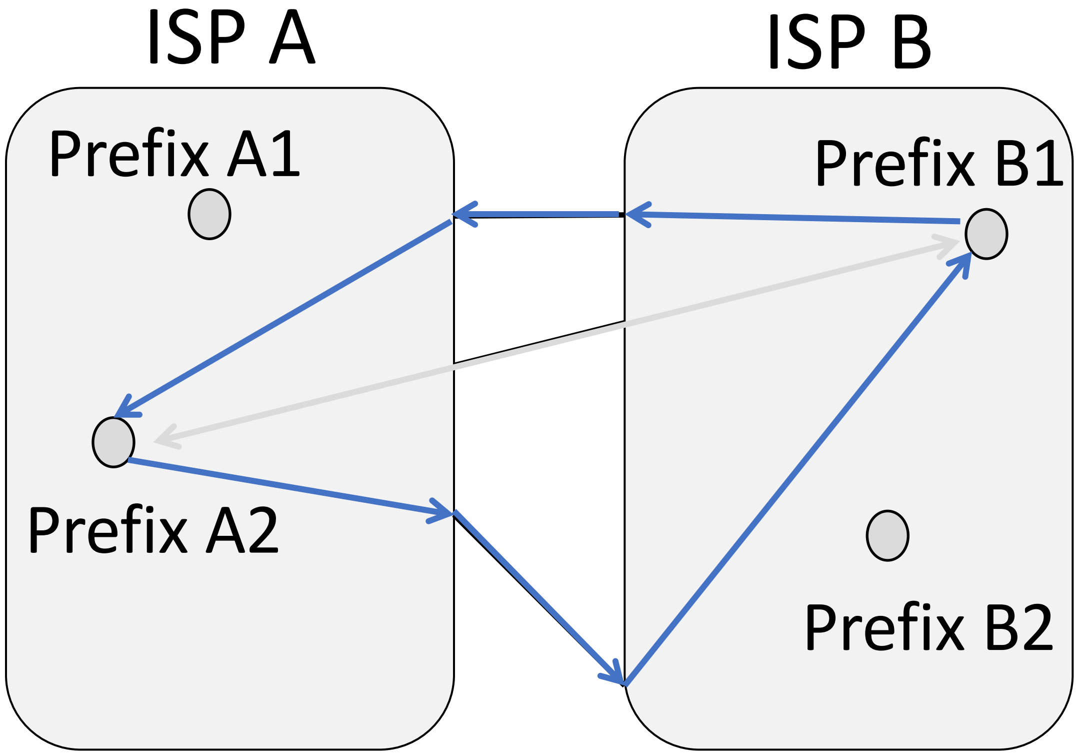
Effects of Independent Parties

- Each party selects routes to suit its own interests
  - e.g, shortest path in ISP
- What path will be chosen for A2→B1 and B1→A2?
  - What is the best path?



Effects of Independent Parties (2)

- Selected paths are longer than overall shortest path
  - And asymmetric too!
- Consequence of independent decisions

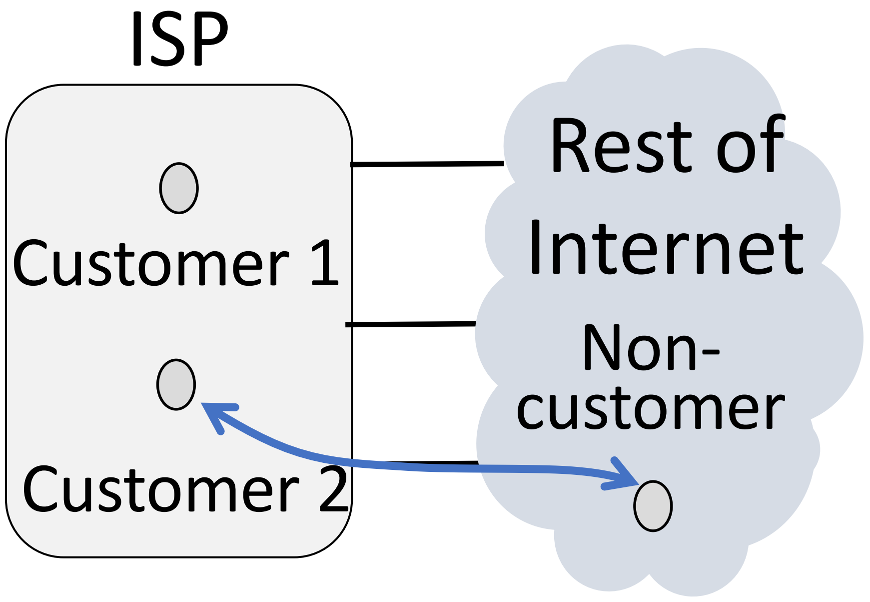


Routing Policies

- Capture the goals of different parties
  - Could be anything
  - E.g., Internet2 only carries non-commercial traffic
- Common policies we’ll look at:
  - ISPs give TRANSIT service to customers
  - ISPs give PEER service to each other

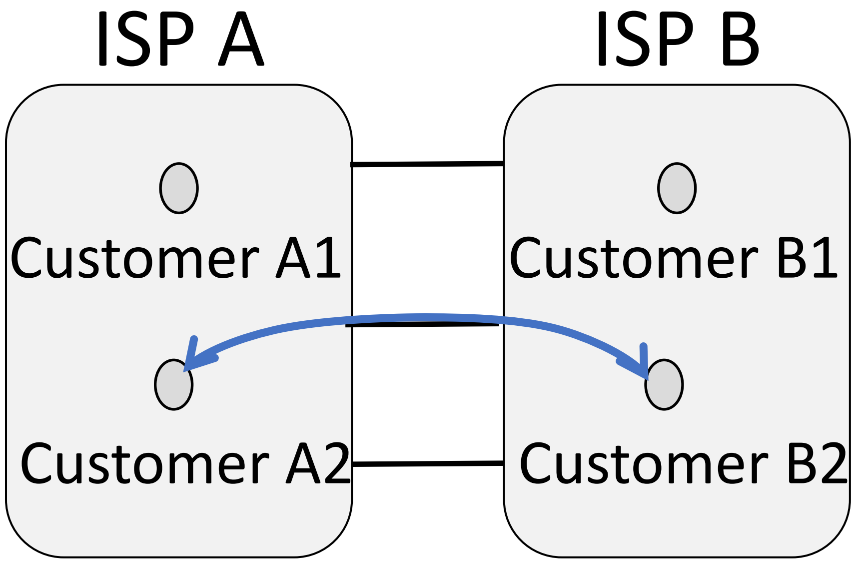
Routing Policies – Transit

- One party (customer) gets TRANSIT service from another party (ISP)
  - ISP accepts traffic for customer from the rest of Internet
  - ISP sends traffic from customer to the rest of Internet
  - Customer pays ISP for the privilege



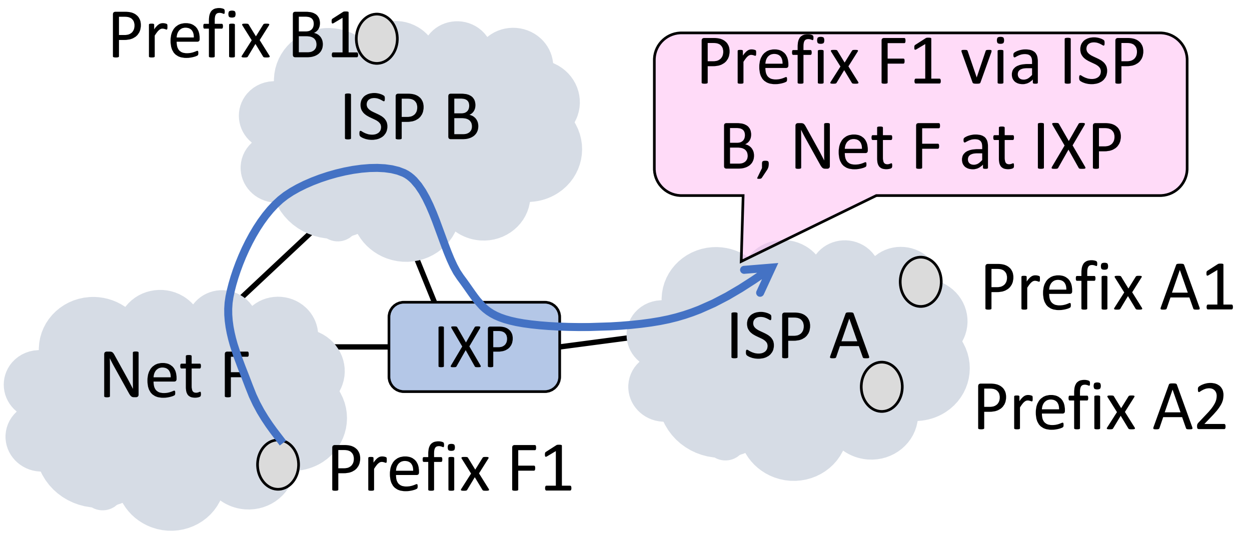
Routing Policies – Peer

- Both party (ISPs in example) get PEER service from each other
  - Each ISP accepts traffic from the other ISP only for their customers
  - ISPs do not carry traffic to the rest of the Internet for each other
  - ISPs don't pay each other



Routing with BGP

- iBGP is used for "internal" routing
- eBGP is <u>interdomain</u> routing for the Internet
  - Path vector, a kind of distance vector



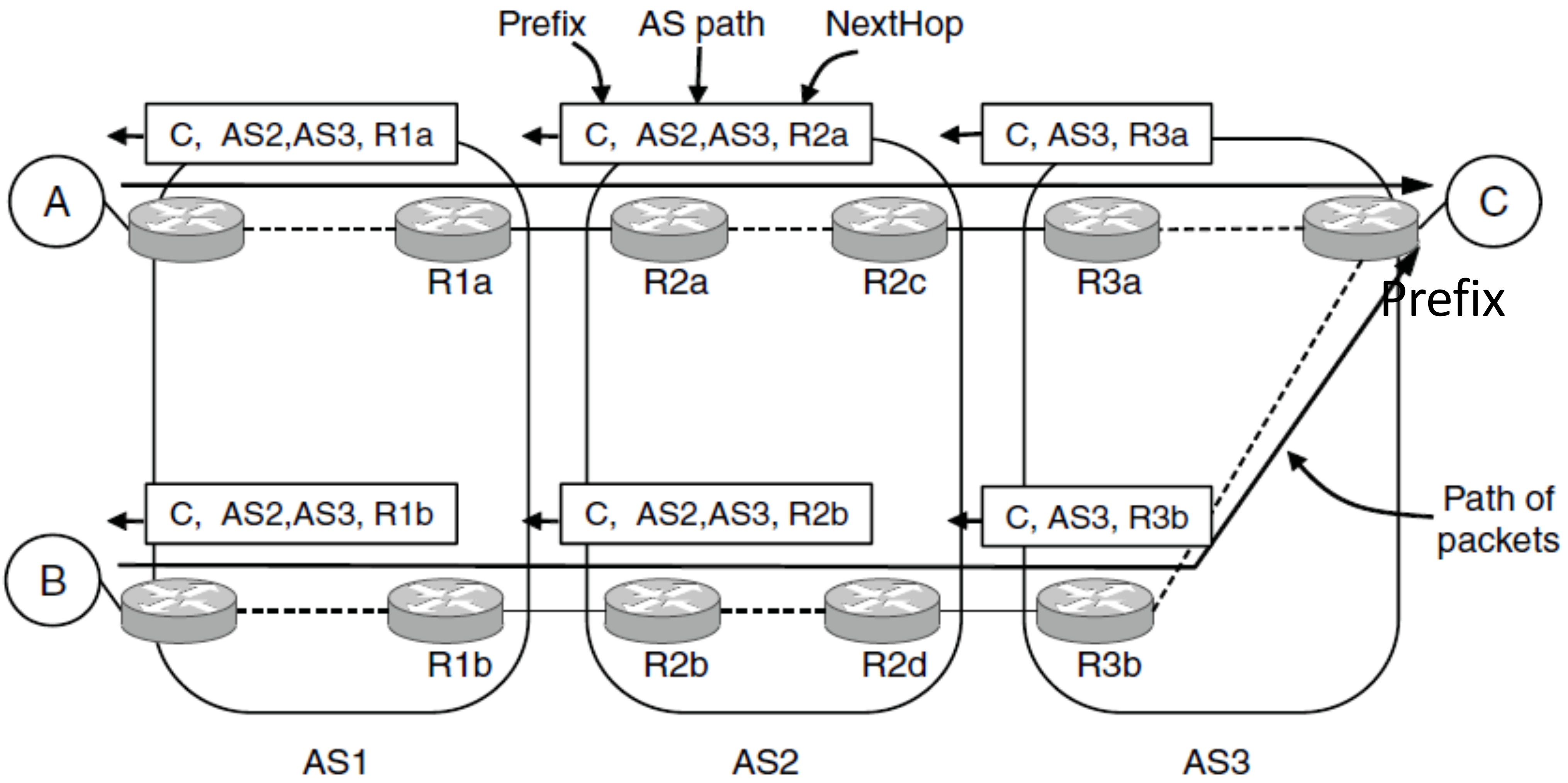
Routing with BGP (2)

- Parties like ISPs are called AS (Autonomous Systems)
  - AS numbers are unique identifiers
- AS's configure their internal BGP routes
- External routes go through complicated filters
- Intra-AS BGP routers communicate (via iBGP) to keep consistent routing information

Routing with BGP (3)

- Border routers of ASes announce BGP routes
- Route announcements have IP prefix, path vector, next hop
  - Path vector is list of ASes on the way to the prefix
  - List is to find loops
- Route announcements move in the opposite direction to traffic

Routing with BGP (4)



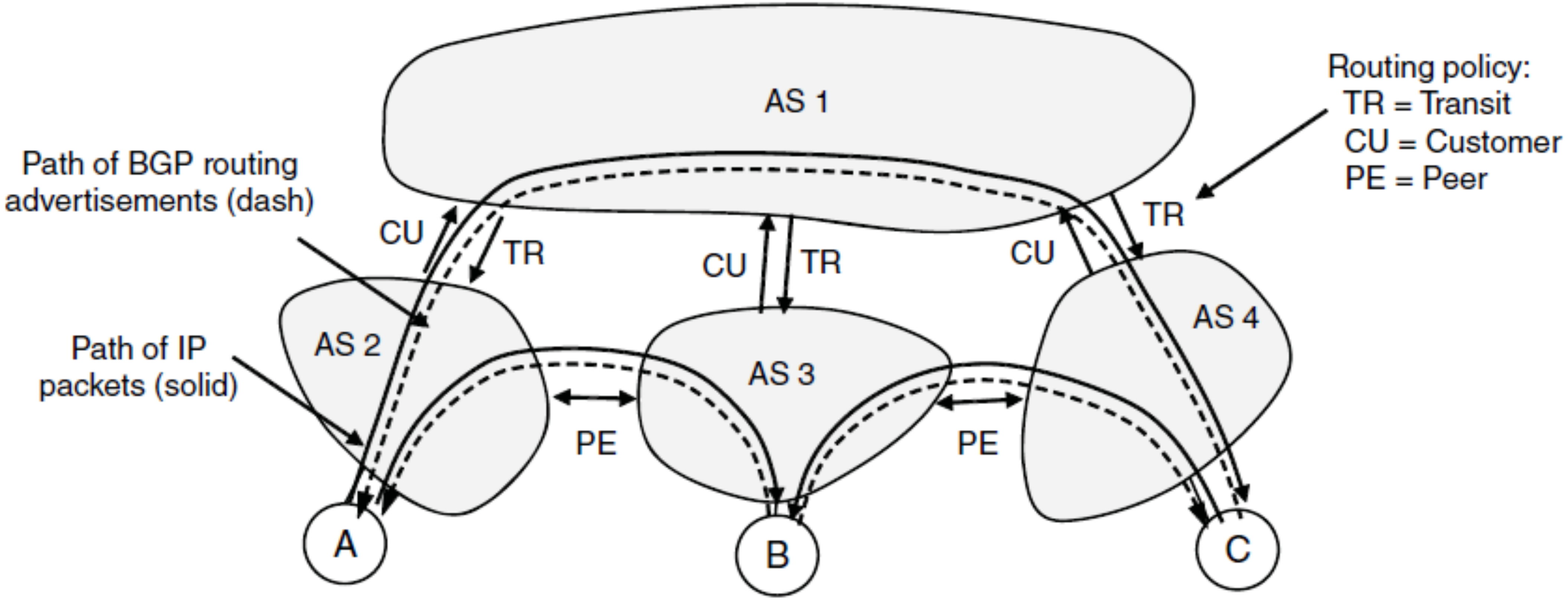
Routing with BGP (5)

Policy is implemented in two ways:

1. Border routers of ISP announce paths only to other parties who may use those paths
    - Filter out paths others can't use
2. Border routers select the best path of the ones they hear in any way (not necessarily shortest)

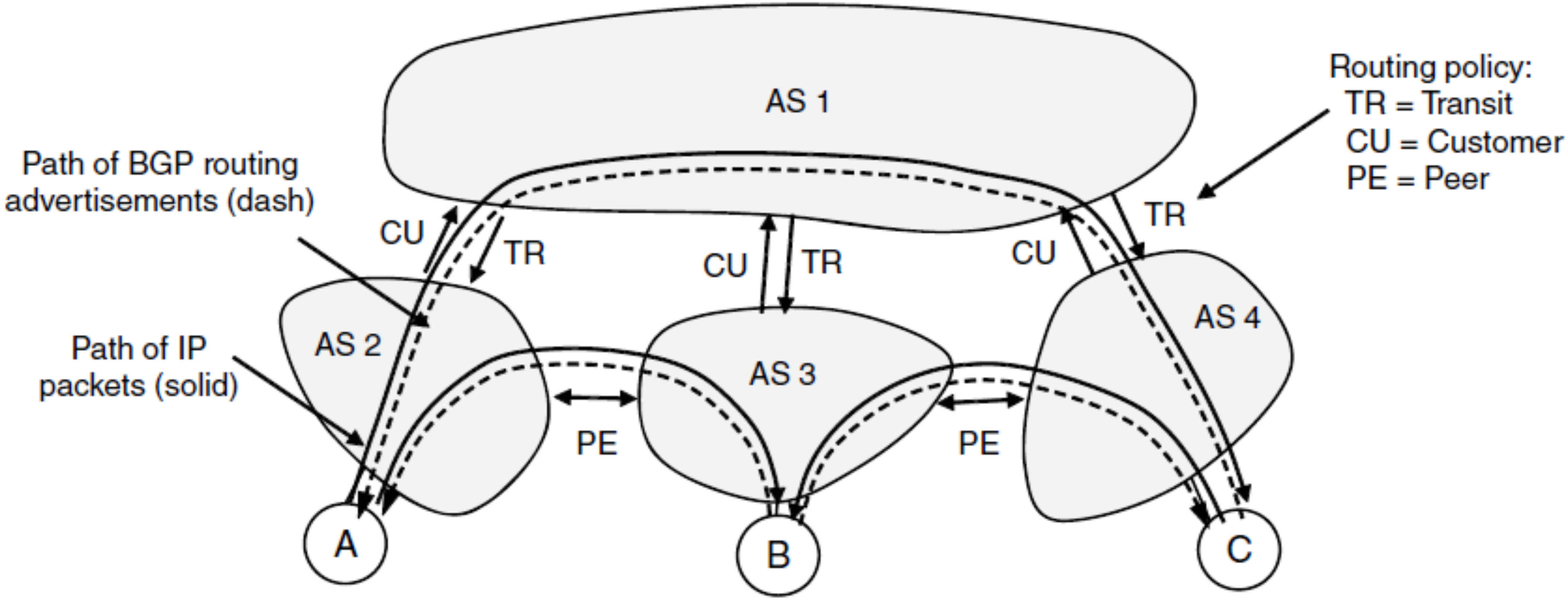
Routing with BGP (6)

- TRANSIT: AS1 says [B, (AS1, AS3)], [C, (AS1, AS4)] to AS2



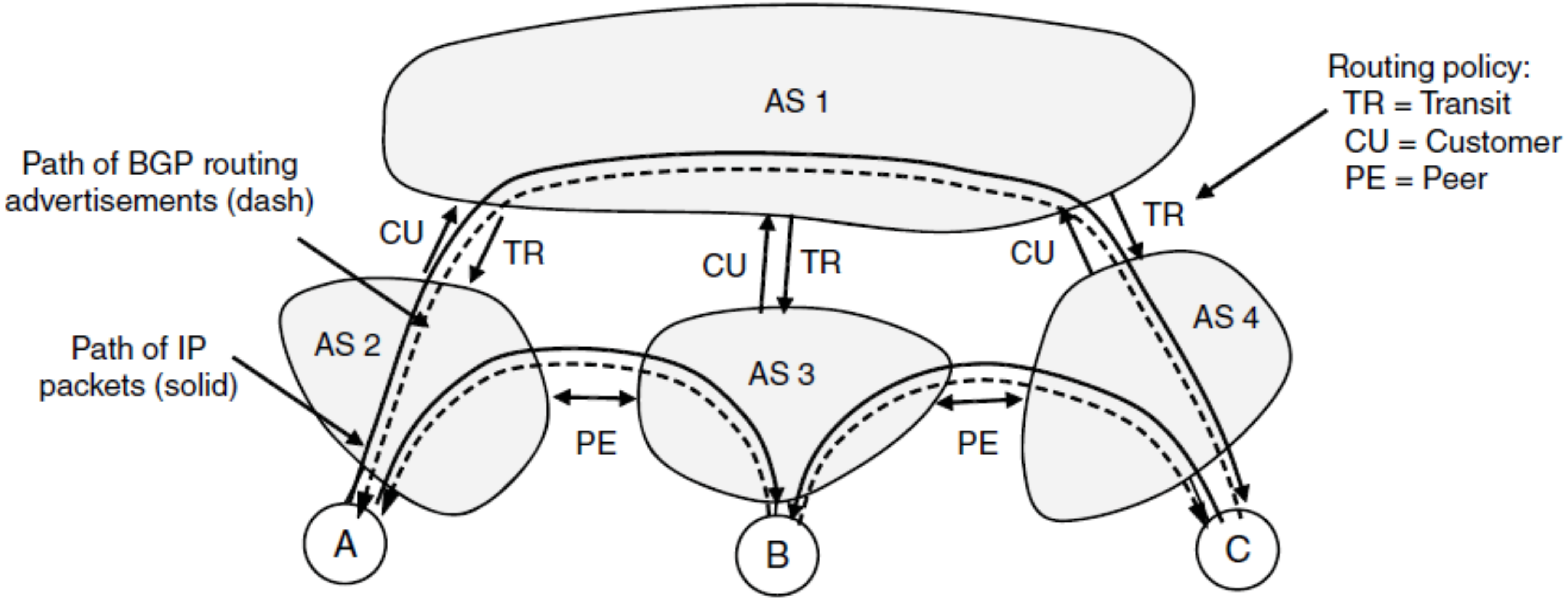
Routing with BGP (7)

- CUSTOMER (other side of TRANSIT): AS2 says [A, (AS2)] to AS1



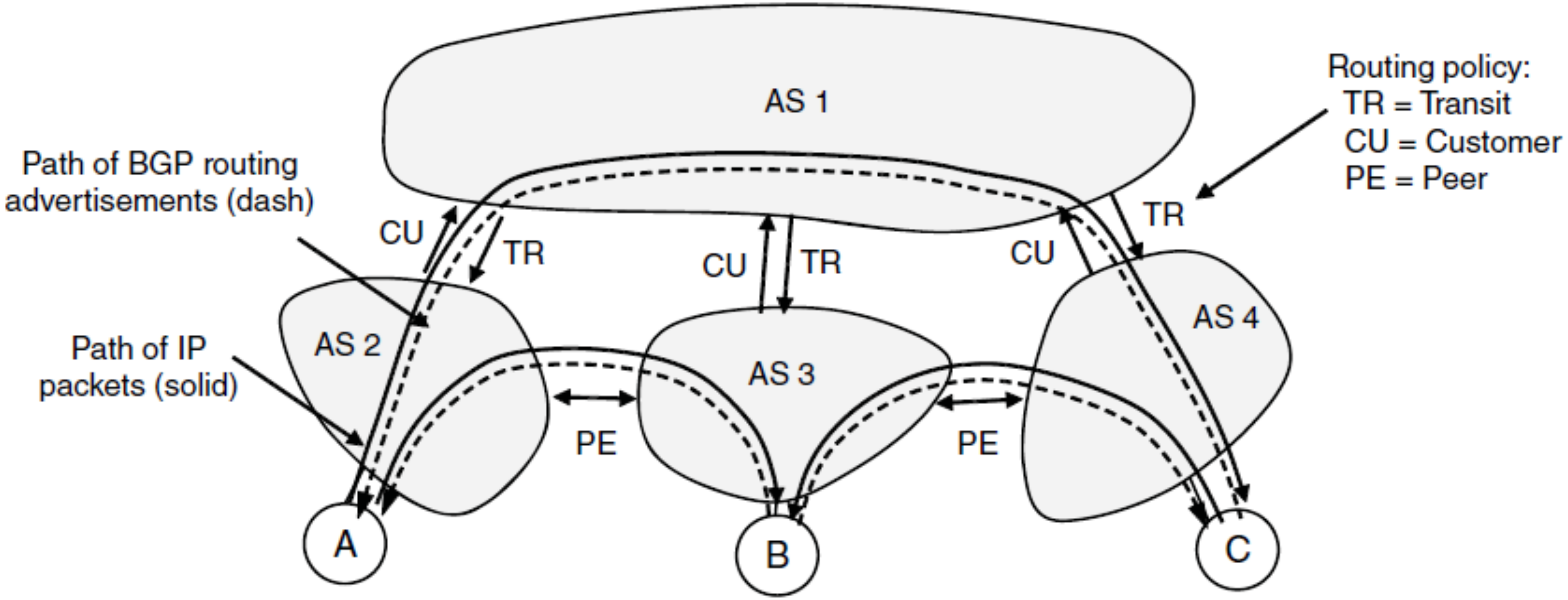
Routing with BGP (8)

- PEER: AS2 says [A, (AS2)] to AS3, AS3 says [B, (AS3)] to AS2



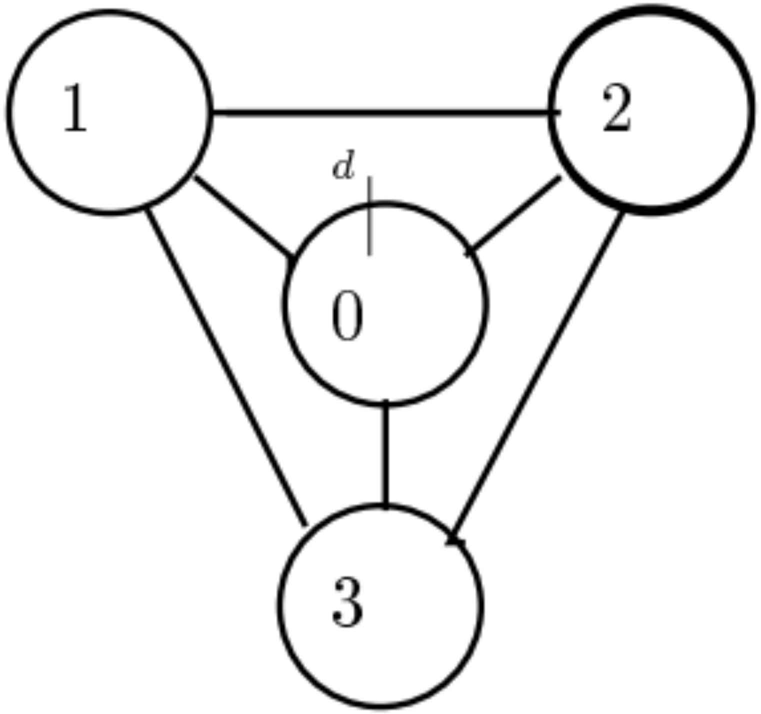
Routing with BGP (9)

- AS2 has two routes to B (AS1, AS3) and chooses AS3 (Free!)



BGP “bad gadget”: Non-convergence

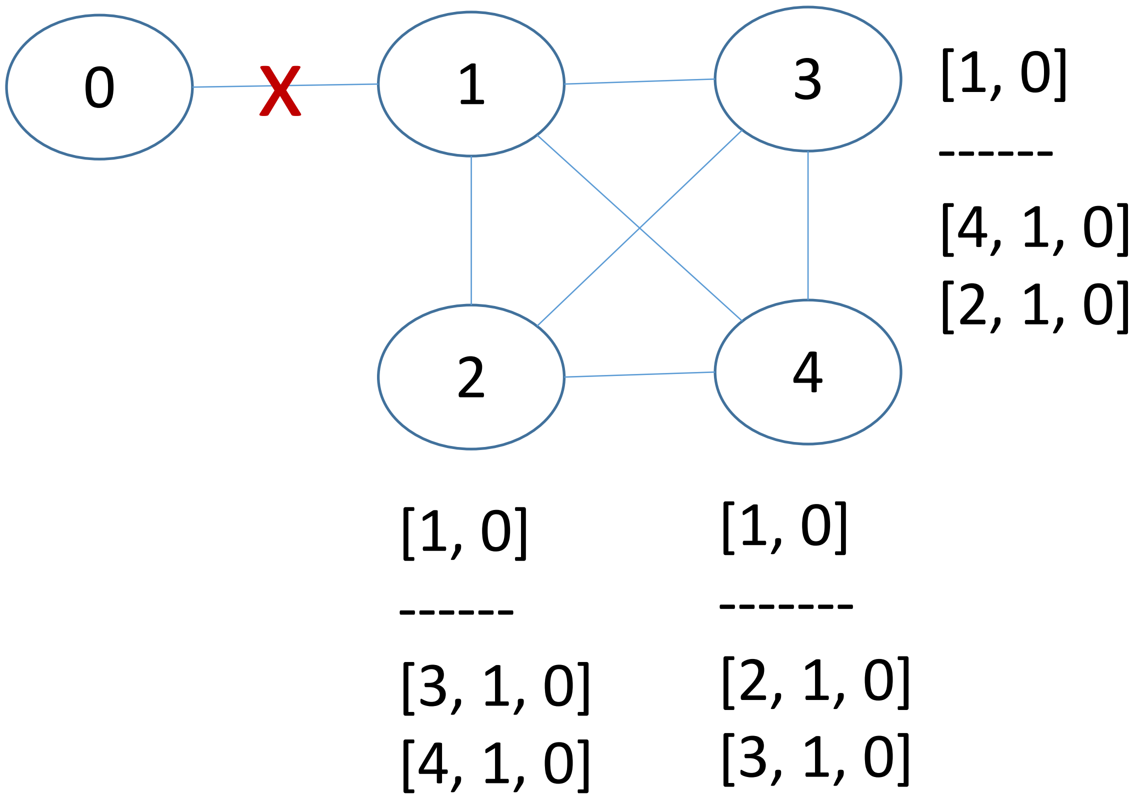
[2, 0] > [0] > [2, 3, 0]



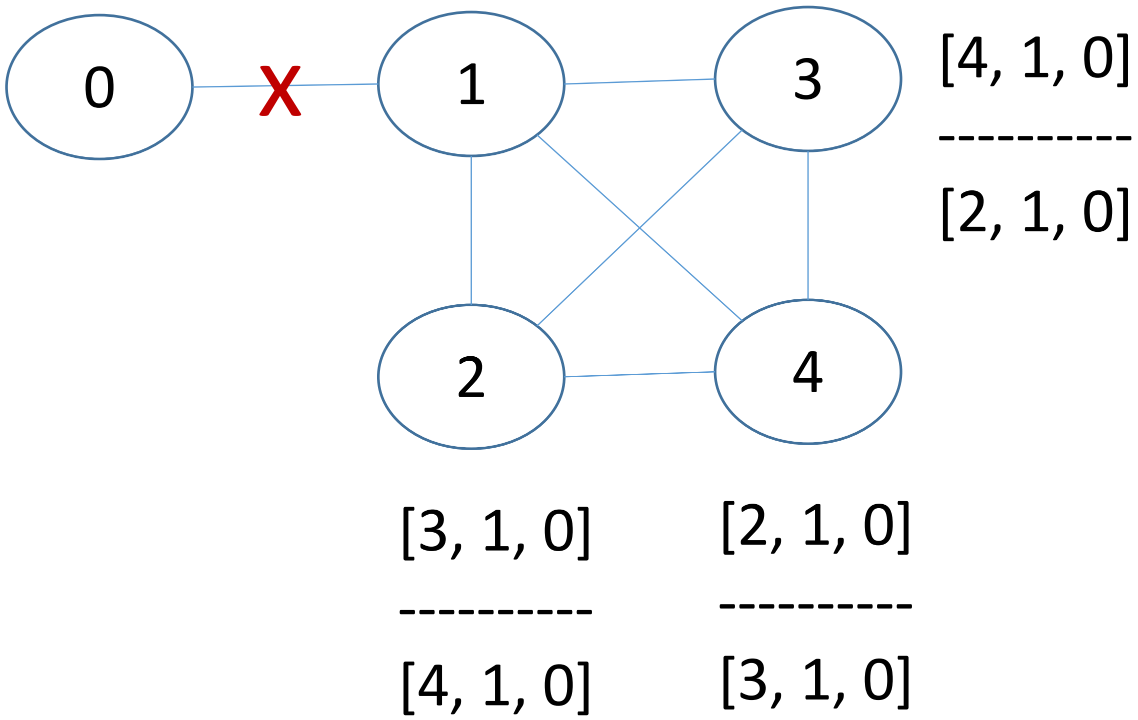
[3, 0] > [0] > [3, 1, 0]

[1, 0] > [0] > [1, 2, 0]

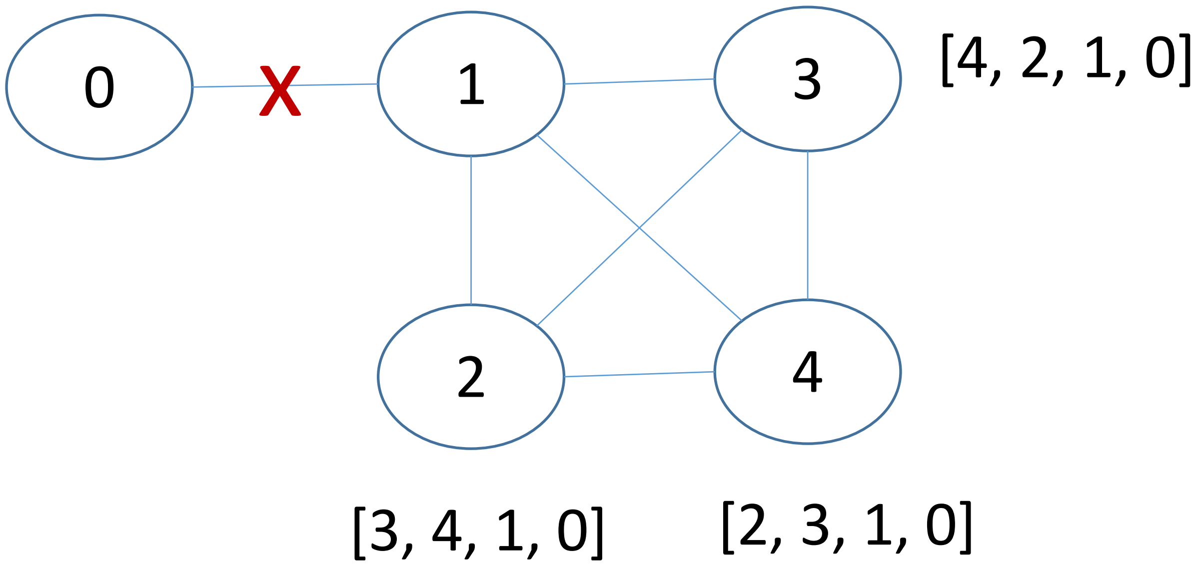
BGP slow convergence



BGP slow convergence



BGP slow convergence



Implementing policy in BGP

1. Export policy
   - Determines what to announce to whom
2. Import policy
   - Determine how to modify (or drop) incoming announcements
3. Decision process
   - Determine “best” path among all those available

Export policy

Arbitrary transformation and filtering of route attributes

- Legal transformations for (10.10.10.0/24, [12, 42], [com1, com2])
  - (10.10.10.0/24, [12, 24, 93], [com1, com2])
  - (10.10.10.0/24, [12, 24, 93, 93, 93, 93], [com1, com2])
  - (10.10.10.0/24, [93], [com1, com2])
  - (10.10.10.0/24, [12, 24, 93], [])
  - (10.10.10.0/24, [12, 24, 93], [com1, com32])

In addition, one may aggregate

- Only announce, 10.10.0.0/16
- Typically done only if a sub-prefix is present

Import policy

Arbitrary transformation and filtering of route attributes

And attach a numerical “local preference” attribute

Decision process

Standard process (all the mauling happens beforehand)

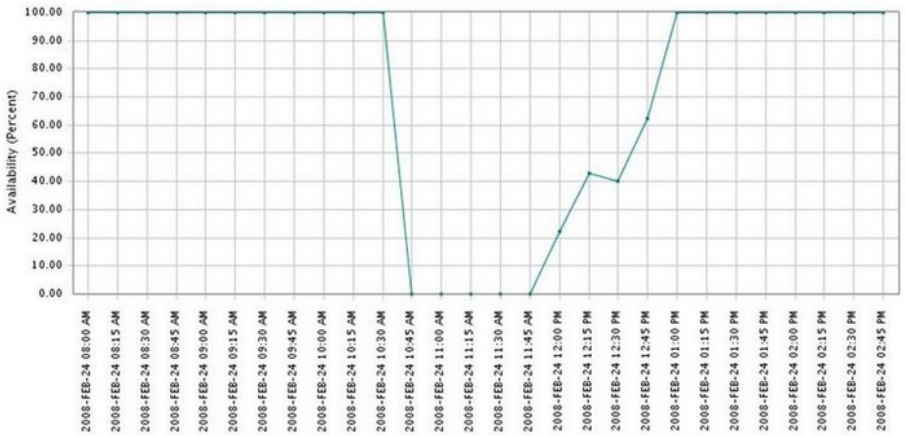
1. Prefer higher local preference
2. Prefer locally originated routes
3. Prefer shorter AS path
4. Prefer eBGP over iBGP
5. Prefer lower internal cost
6. Prefer lower router ID

CULTURE

How Pakistan knocked YouTube offline (and how to make sure it never happens again)

YouTube becoming unreachable isn't the first time that Internet addresses were hijacked. But if it spurs interest in better security, it may be the last.

BY DECLAN MCCULLAGH | FEBRUARY 25, 2008 4:28 PM PST



This graph that network-monitoring firm Keynote Systems provided to us shows the worldwide availability of YouTube.com dropping dramatically from 100 percent to 0 percent for over an hour. It didn't recover completely until two hours had elapsed.

Keynote Systems

A high-profile incident this weekend in which Pakistan's state-owned telecommunications company managed to cut YouTube off the global Web highlights a long-standing security weakness in the way the Internet is managed.

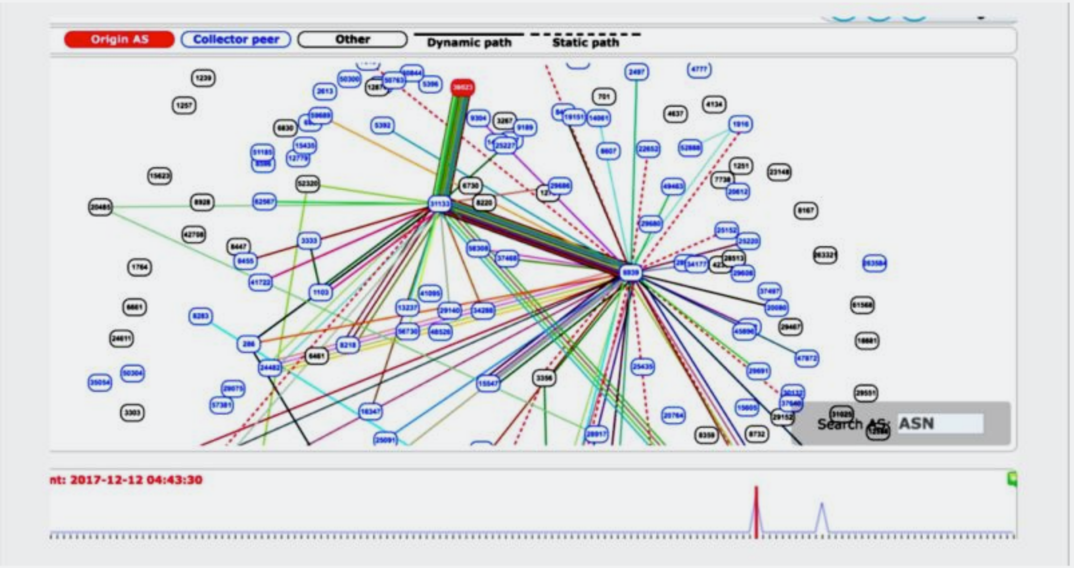
After receiving a censorship order from the telecommunications ministry directing that YouTube.com be blocked, Pakistan Telecom went even further. By accident or

BIZ & IT —

"Suspicious" event routes traffic for big-name sites through Russia

Google, Facebook, Apple, and Microsoft all affected by "intentional" BGP mishap.

DAN GOODIN - 12/13/2017, 2:43 PM



BGPMon

Enlarge

Traffic sent to and from Google, Facebook, Apple, and Microsoft was briefly routed through a previously unknown Russian Internet provider Wednesday under circumstances researchers said was suspicious and intentional.

FURTHER READING
Russian-controlled telecom hijacks financial services' Internet traffic

The unexplained incident involving the Internet's Border Gateway Protocol is the latest to raise troubling questions about the trust and reliability of communications sent over the global network. BGP routes large-scale amounts of traffic among Internet backbones, ISPs, and other large networks. But despite the sensitivity and amount of data it controls, BGP's security is often based on trust and word of mouth. Wednesday's event comes eight months after large chunks of network traffic belonging to MasterCard, Visa, and more than two dozen other financial services were briefly routed through a Russian government-