Network Layer (Routing)
Border Gateway Protocol
Structure of the Internet

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

• Two problems beyond routing within a network

1. Scaling to very large networks
   • Techniques of IP prefixes, hierarchy, prefix aggregation

2. Incorporating policy decisions
   • Letting different parties choose their routes to suit their own needs
   Yikes!
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!
- This is a consequence of independent goals and decisions, not hierarchy
Routing Policies

• Capture the goals of different parties
  • Could be anything

• 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 (Border Gateway Protocol)

• iBGP is for internal routing
• eBGP is **interdomain** routing for the Internet
  • Path vector, a kind of distance vector

![Diagram showing BGP routing](attachment:diagram.png)
Routing with BGP

• Parties like ISPs are called AS (Autonomous Systems)
• AS’s **MANUALLY** configure their internal BGP routes/advertisements
• External routes go through complicated filters for forwarding/filtering
• AS BGP routers communicate with each other to keep consistent routing rules
Routing with BGP

• 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
Routing with BGP

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 of ISP select the best path of the ones they hear in any, non-shortest way
Routing with BGP

- **TRANSIT**: AS1 says \([B, (AS1, AS3)], [C, (AS1, AS4)]\) to AS2
Routing with BGP

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

Path of BGP routing advertisements (dash)

Routing policy:
- \(TR =\) Transit
- \(CU =\) Customer
- \(PE =\) Peer

Path of IP packets (solid)
Routing with BGP

- **PEER**: AS2 says \([A, (AS2)]\) to AS3, AS3 says \([B, (AS3)]\) to AS2
Routing with BGP

• AS2 has two routes to B (AS1, AS3) and chooses AS3 (Free!)
BGP Thoughts

• Much more beyond basics to explore!
• Policy is a substantial factor
  • Can independent decisions be sensible overall?
• Other important factors:
  • Convergence effects
  • How well it scales
  • Integration with intradomain routing
  • And more …