Hierarchical Routing

• Introduce a larger routing unit
  – IP prefix (hosts) ← from one host
  – Region, e.g., ISP network

• Route first to the region, then to the IP prefix within the region
  – Hide details within a region from outside of the region
Hierarchical Routing (2)
Hierarchical Routing (3)

Full table for 1A

<table>
<thead>
<tr>
<th>Dest.</th>
<th>Line</th>
<th>Hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1B</td>
<td>1B</td>
<td>1</td>
</tr>
<tr>
<td>1C</td>
<td>1C</td>
<td>1</td>
</tr>
<tr>
<td>2A</td>
<td>1B</td>
<td>2</td>
</tr>
<tr>
<td>2B</td>
<td>1B</td>
<td>3</td>
</tr>
<tr>
<td>2C</td>
<td>1B</td>
<td>3</td>
</tr>
<tr>
<td>2D</td>
<td>1B</td>
<td>4</td>
</tr>
<tr>
<td>3A</td>
<td>1C</td>
<td>3</td>
</tr>
<tr>
<td>3B</td>
<td>1C</td>
<td>2</td>
</tr>
<tr>
<td>4A</td>
<td>1C</td>
<td>3</td>
</tr>
<tr>
<td>4B</td>
<td>1C</td>
<td>4</td>
</tr>
<tr>
<td>4C</td>
<td>1C</td>
<td>4</td>
</tr>
<tr>
<td>5A</td>
<td>1C</td>
<td>4</td>
</tr>
<tr>
<td>5B</td>
<td>1C</td>
<td>5</td>
</tr>
<tr>
<td>5C</td>
<td>1B</td>
<td>5</td>
</tr>
<tr>
<td>5D</td>
<td>1C</td>
<td>6</td>
</tr>
<tr>
<td>5E</td>
<td>1C</td>
<td>5</td>
</tr>
</tbody>
</table>

Hierarchical table for 1A

<table>
<thead>
<tr>
<th>Dest.</th>
<th>Line</th>
<th>Hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1B</td>
<td>1B</td>
<td>1</td>
</tr>
<tr>
<td>1C</td>
<td>1C</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1B</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1C</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1C</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>1C</td>
<td>4</td>
</tr>
</tbody>
</table>
Hierarchical Routing (4)

- Penalty is longer paths

1C is best route to region 5, except for destination 5C.
Observations

• Outside a region, nodes have one route to all hosts within the region
  – This gives savings in table size, messages and computation

• However, each node may have a different route to an outside region
  – Routing decisions are still made by individual nodes; there is no single decision made by a region
Topic

• How to route with multiple parties, each with their own routing policies
  – This is Internet-wide BGP routing
Structure of the Internet

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

• Two problems beyond routing within an individual 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
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 symmetric too!
- This is a consequence of independent goals and decisions, not hierarchy
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 (Border Gateway Protocol)

- BGP is the **interdomain** routing protocol used in the Internet
  - Path vector, a kind of distance vector
Routing with BGP (2)

• Different parties like ISPs are called AS (Autonomous Systems)
• Border routers of ASes announce BGP routes to each other
• Route announcements contain an 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 (3)
Routing with BGP (4)

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 (5)

- **TRANSIT**: AS1 says \([B, (AS1, AS3)], [C, (AS1, AS4)]\) to AS2
Routing with BGP (6)

- CUSTOMER (other side of TRANSIT): AS2 says [A, (AS2)] to AS1
Routing with BGP (7)

- PEER: AS2 says [A, (AS2)] to AS3, AS3 says [B, (AS3)] to AS2
Routing with BGP (8)

- AS2 hears two routes to B (via AS1, AS3) and chooses AS3 (Free!)
BGP Thoughts

• Much more beyond basics to explore!

• Policy is a substantial factor
  – Can we even be independent decisions will be sensible overall?

• Other important factors:
  – Convergence effects
  – How well it scales
  – Integration with intradomain routing
  – And more ...