Hierarchical routing

Our routing study thus far - idealization
• all routers identical
• network “flat”
... *not* true in practice

**Scale:** with 600 million destinations:
• can’t store all dest’s in routing tables!
• routing table exchange would swamp links!

**Administrative autonomy**
• internet = network of networks
• each network admin may want to control routing in its own network
Hierarchical routing

• We aggregate routers into regions, “autonomous systems” (AS)

• Routers in same AS run same routing protocol
  ▪ “intra-AS” or “interior” routing protocol
  ▪ routers in different AS can run different intra-AS routing protocol

Gateway (or border) router:

• at “edge” of its own AS
• has link to router in another AS
Interconnected ASes

- Forwarding table configured by both intra- and inter-AS routing algs
  - intra-AS sets entries for internal dests
  - inter-AS & intra-AS sets entries for external dests
Inter-AS tasks

• Suppose router in AS1 receives a datagram destined outside of AS1:
  ▪ Router should forward packet to gateway router, but which one?

  **AS1 must:**
  1. Learn which dests are reachable through AS2, which through AS3
  2. Propagate this reachability info to all routers in AS1

  *Job of inter-AS routing!*
Example: setting forwarding table in router 1d

- Suppose AS1 learns (via inter-AS protocol) that AS with prefix $x$ is reachable via AS3 (gateway 1c), but not via AS2
  - inter-AS protocol propagates reachability to all internal routers
- Router 1d determines from intra-AS routing info that its interface $I$ is on the least cost path to 1c
  - Installs forwarding table entry $(x,I)$
If an external destination is reachable from multiple gateways, a router inside the AS should forward packets for that destination to

A. The closest gateway that can reach the destination.

B. The gateway that has the least-cost external path to the destination.

C. The gateway that has the least-cost path for both the internal and external path.

D. Somewhere else.
Routing Policy

• How should the ISP route the customer’s traffic to the destination?

Do what’s best for... who?
Hot Potato Routing

- Hot Potato: get rid of packets ASAP!
- Best path: get it as close to the destination as we can.
Hot Potato Routing

- **Hot Potato**: get rid of packets ASAP!
- **Best path**: get it as close to the destination as we can.

Not my problem!
(Don’t use resources)
Hot Potato Routing

• **Hot Potato**: get rid of packets ASAP!
• **Best path**: get it as close to the destination as we can.

Hold packet longer, use more resources....
Provide better service!
Route Selection

• Often dictated by non-technical factors

• When governed by protocols, two categories:
  – Intra-AS / Interior gateway protocols
  – Inter-AS / Exterior gateway protocols
Why different Intra-, Inter-AS routing?

Policy:
- inter-AS: admin wants control over how its traffic routed, who routes through its net.
- intra-AS: single admin, so no policy decisions needed

Scale:
- hierarchical routing saves table size, reduced update traffic

Performance:
- intra-AS: can focus on performance
- inter-AS: policy may dominate over performance
Intra-AS Routing

• Also known as *interior gateway protocols (IGP)*

• *Distance Vector:*
  - RIP: Routing Information Protocol
  - (E)IGRP: Interior Gateway Routing Protocol
    (Cisco proprietary)

• *Link State:*
  - OSPF: Open Shortest Path First
  - IS-IS: Intermediate system to Intermediate system

**OSPF and IS-IS are deployed most commonly today!**
Intra-AS Routing

• Also known as *interior gateway protocols (IGP)*

Goal:
Get traffic that is already in an AS to a destination inside that same AS.

OSPF and IS-IS are deployed most commonly today!
RIP (Routing Information Protocol)

• Included in BSD-UNIX distribution in 1982
  – distance metric: # hops (max = 15 hops), each link has cost 1
  – hops = number of subnets traversed
  – Distance vectors exchanged with neighbors every 30 sec
  – Each advertisement: list of up to 25 destination subnets

from router A to destination subnets:

<table>
<thead>
<tr>
<th>subnet</th>
<th>hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>u</td>
<td>1</td>
</tr>
<tr>
<td>v</td>
<td>2</td>
</tr>
<tr>
<td>w</td>
<td>2</td>
</tr>
<tr>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>y</td>
<td>3</td>
</tr>
<tr>
<td>z</td>
<td>2</td>
</tr>
</tbody>
</table>
OSPF (Open Shortest Path First)

• Link state protocol (reliable flooding of LSAs)

• “Open”: standardized, publicly available implementations

• Multiple equal-cost paths allowed (load balancing)

• Additional features:
  – OSPF messages authenticated (to prevent malicious intrusion)
  – Hierarchical OSPF for large autonomous systems.
Hierarchical OSPF

• **Two-level hierarchy:** local area, backbone.
  – link-state advertisements only in area
  – each node has detailed area topology; only know direction (shortest path) to nets in other areas.

• **Area border routers:** “summarize” distances to nets in own area, advertise to other Area Border routers.

• **Backbone routers:** route between local areas

• **Boundary routers:** connect to other AS’s.
Hierarchical OSPF

- Backbone
- Border routers
- Area 1
- Area 2
- Area 3
- Internal routers
- Boundary router
- Backbone router
Route Selection

• Often dictated by non-technical factors

• When governed by protocols, two categories:
  – Intra-AS / Interior gateway protocols
  – Inter-AS / Exterior gateway protocols
Internet inter-AS routing: BGP

Goal: Get traffic from one AS to another.
Internet inter-AS routing: BGP

- **BGP (Border Gateway Protocol):** The de facto inter-domain routing protocol
- BGP provides each AS a means to:
  - **external BGP:** obtain subnet reachability information from neighboring ASs.
  - **internal BGP:** propagate reachability information to all AS-internal routers.
  - determine “good” routes to other networks based on reachability information and policy.
- Allows a subnet to advertise its prefix to the rest of the Internet
BGP

- **BGP session**: two BGP routers ("peers") exchange BGP messages:
  - Advertising *paths* to different destination network prefixes ("path vector")
  - Exchanged over long-term TCP connections

- When AS3 advertises a prefix to AS1:
  - AS3 *promises* it will forward datagrams towards that prefix
  - AS3 can aggregate prefixes in its advertisement
BGP: Distributing Path Information

- Using eBGP session between 3a and 1c, AS3 sends prefix reachability info to AS1.
  - 1c can then use iBGP do distribute new prefix info to all routers in AS1
  - 1b can then re-advertise new reachability info to AS2 over 1b-to-2a eBGP session
- When a router learns of a new prefix, it creates an entry for the prefix in its forwarding table.
Path attributes and BGP routes

• An advertised prefix includes BGP attributes
  – prefix + attributes = “route”

• Two important attributes:
  – **AS-PATH**: contains list of ASs through which prefix advertisement has passed:
    • If AS2 advertises a prefix to AS1, AS1 will advertise path: AS1 AS2...
    • Ignore routes that include yourself in them!
  – **NEXT-HOP**: indicates specific internal-AS router to next-hop AS.
    (may be multiple links from current AS to next-hop-AS)

• Gateway router receiving route advertisement uses import policy to accept/decline
  – e.g., never route through AS x
  – *policy-based* routing
BGP Route Selection

• Router may learn about more than one route to destination AS, selects route based on:
  – local preference value attribute: administrative policy
  – shortest AS-PATH
  – closest NEXT-HOP router: hot potato routing
  – additional criteria
Which routes a BGP router advertises will depend on...

A. which ISPs have contractual agreements.

B. the shortest path to a subnet/prefix.

C. which subnets are customers of an ISP.

D. More than one of the above. (which?)
BGP routing policy

- A, B, C are provider networks
- X, W, Y are customer (of provider networks)
- X is dual-homed: attached to two networks
  - X does not want to route from B via X to C
  - .. so X will not advertise to B a route to C
BGP routing policy (2)

- A advertises path AW to B
- B advertises path BAW to X
- Should B advertise path BAW to C?
  - B gets no “revenue” for routing CBAW since neither W nor C are B’s customers
  - B wants to force C to route to w via A
  - B wants to route *only* to/from its customers!
Summary

• As we’ve seen before (DNS), a hierarchy can help manage state storage constraints.
  – intra-AS routing: lots of info about local routes
  – inter-AS routing: less info about far away routes

• BGP: the inter-AS routing protocol for the Internet
  – Decisions often contractual

• BGP advertises AS prefixes, including:
  – entire path of ASes along the way
  – which border router heard the advertisement (Next Hop)