15. Inter-Domain Routing

- General Concepts
- BGP

Jon Turner – slides adapted from Kurose and Ross
Hierarchical Routing

- The Internet is divided among many distinct networks
  - owned and operated by different organizations
  - networks called *Autonomous Systems* (aka routing domains)
- Leads to a two level routing structure
  - intra-domain routing: finding most efficient paths within an AS
  - inter-domain routing: finding paths among ASes
  - makes Internet routing more scalable
  - allows ASes to operate independently and to keep their internal network structure private
- Drawbacks of hierarchical routing
  - lack of global knowledge of network topology prevents selection of best routes
  - motivates AS-owners to focus on reducing their own costs, not providing best service to users
Intra-Domain & Inter-Domain Routing

- Forwarding table configured by both intra-domain and inter-domain routing
  - intra-domain sets entries for internal destinations
  - both collaborate to set entries for external destinations

Diagram shows network topologies and routers connecting different autonomous systems (AS) with labels 1a, 1b, 2a, 2b, 3a, 3b, 3c.
Inter-AS Tasks

- Suppose router in AS1 receives datagram destined outside of AS1
  - router should forward packet to gateway router, but which one?

  AS1 must:
  1. learn which destinations are reachable through AS2, which through AS3
  2. propagate this reachability info to all routers in AS1
Setting Forwarding Table in Router 1d

- Suppose AS1 learns (via inter-domain protocol) that X is reachable via AS3 (gateway 1c) but not via AS2
  - inter-domain protocol propagates reachability info to all internal routers
- Router 1d determines from intra-domain routing info that its interface I is on the least cost path to 1c
  - installs forwarding table entry (X,I)
Choosing Among Multiple Paths

- Now suppose AS1 learns from inter-domain protocol that subnet x is reachable from AS3 and from AS2.
- To configure forwarding table, router 1d must select from among multiple choices:
  - common approach, choose path with fewest AS-hops
    - break ties by picking closest gateway
    - may not be the true least-cost path
Internet Inter-Domain Routing: BGP

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

- BGP session – two BGP routers (“peers”) exchange BGP messages
  - advertising paths to different destination network prefixes (“path vector” protocol)
  - exchanged over semi-permanent TCP connections
- Example: if AS3 advertises a prefix 1.2.0.0/16 to AS1
  - AS3 promises it will forward matching datagrams from AS1
  - AS3 can aggregate prefixes in its advertisements
Distributing Path Information

- Using eBGP session between 3a and 1c, AS3 sends prefix reachability info to AS1
  - 1c can then use iBGP to 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 router learns of new prefix, it creates entry for prefix in its forwarding table
Path Attributes and BGP Routes

- Advertised prefix includes BGP attributes
  - prefix + attributes = “route”
- Two important attributes:
  - AS-PATH: contains ASs through which prefix advertisement has passed: e.g., AS 67, AS 17
  - NEXT-HOP: indicates specific internal-AS router to next-hop AS (may be multiple links from current router to next-hop-AS)
- Gateway router receiving route advertisement uses import policy to accept/decline
  - for example, an AS may choose to not route transit traffic to domains that are not its “customers”
  - policy-based routing
BGP Route Selection

- Router may learn about more than 1 route to destination AS, selects route based on:
  1. local preference value attribute: policy decision
  2. shortest AS-PATH
     - no essential relationship to user-perceived performance
  3. closest NEXT-HOP router: hot potato routing
  4. additional criteria
BGP Messages

- BGP messages exchanged between peers over TCP connection
- BGP messages
  - OPEN – opens TCP connection to peer and authenticates sender
  - UPDATE – advertises new path (or withdraws old)
  - KEEPALIVE – keeps connection alive in absence of UPDATES; also ACKs OPEN request
  - NOTIFICATION – reports errors in previous message; also used to close connection
BGP Routing Policy

- A, B, C are provider networks
- X, W, Y are customer (of provider networks)
- X is dual-homed: attached to two provider 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

- A advertises path $AW$ to $B$
- B advertises path $BAW$ to $X$
- Should B advertise path $BAW$ to $C$?
  - No way! 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!
  - unless B and C agree to forward traffic for each other (peering)
Why Use Different Routing Protocols?

- **Policy**
  - Inter-domain: administrator wants control over how its traffic is routed and who routes through its net
  - Intra-domain: single admin, so no policy decisions needed

- **Scale**
  - Hierarchical routing reduces table size and update traffic
    - Arguably, modern processors can handle Internet scale

- **Performance**
  - Intra-domain: can focus on performance
  - Inter-domain: policy may dominate over performance

- **Drawbacks**
  - Independent policy choices can lead to routing instability
  - Inter-domain routing largely ignores user-perceived performance
BGP Example
Exercises

Assume BGP when answering the following questions.

1. List five distinct inter-AS paths leading to AS4 that router C might learn of using BGP. For each path, give the path and the “next-hop-address” for that path. For each of these inter-AS paths, what is the intra-AS path that would be used with it? Which path would you expect it to actually select?

2. What path would router B use to reach AS8? What path would it use to reach AS9?

3. How could AS1 avoid carrying packets between AS2 and AS7? Might this have some unintended consequences?
Exercises

1. Give an example illustrating how the routes computed by BGP can lead to packets traveling distances that are much longer than the shortest path distance between the sender and the receiver. How common do you think such sub-optimal paths are? What are some of the negative consequences of packets taking sub-optimal paths?

2. One justification for BGP’s AS-hop-based metric is that it allows ISPs to conceal the topologies of their networks. Why do you think ISPs consider it important to keep this information secret? Do you think that these reasons are sufficient justification for the negative impacts of sub-optimal routing?