How Do You Run an ISP?

(a.k.a. Intra-AS Routing)

Paul Gauthier Jed Liu

14 October 2002

What does an ISP's customer care about?

- Cost
- Global connectivity
- Reliability/availability
- Quality of connection
- Privacy
- Added value (e.g., e-mail accounts, web hosting space)

Global connectivity

- ISP must either be a Tier 1 or buy transit from a Tier 1 provider
- Peering can improve connectivity to other non-Tier-1 ISPs

Forwarding options (within a single ISP)

- Using global addresses (IP)
  - Each router manages a table of (IP address prefix → next hop) mappings
  - CIDR − Classless Inter-Domain Routing
    - ability to use variable-length IP address prefixes for routing
- Virtual circuits
  - MPLS is a technique for establishing VCs through a network of routers
  - an MPLS header is added to each packet, identifying the virtual circuit being traversed
  - *s*wizzling lets routers avoid maintaining a table of $O(N^2)$ possible virtual circuits (and each packet having a large MPLS header)
    - each router only knows how to route the VCs that pass through it
    - router maintains a table mapping (incoming tag → (next hop, outgoing tag)). This is used to rewrite the incoming MPLS header with a new outgoing header, specific to the table in the next-hop router.

Maintaining routing tables

- Static tables
  - maintain routing tables by hand
  - handling failures is hard
- Routing protocols
  - Distance vector
    - routers exchange their (IP prefix → (next hop, distance)) routing tables with their neighbours
    - No way to determine the origin of these routing entries.

    - If links change rapidly, can create a persistent routing loop.
    - Count-to-infinity problem:
        1. Routers $A \leftrightarrow B \leftrightarrow C$.
        2. B announces "$(C \rightarrow (C, 1))$"
        3. A announces "$(C \rightarrow (B, 2))$"
        4. The $B \leftrightarrow C$ link fails: $A \leftrightarrow B - \times - C$
        5. B receives A's announcement and announces "$(C \rightarrow (A, 3))$", which is clearly wrong
        6. A receives B's announcement and concludes "$(C \rightarrow (B, 4))$"
        7. count to infinity...
    - Solution to count to infinity: exchange paths instead of links
- Link state
    - Every router broadcasts its neighbour info to every other router
      Everyone has a complete topology
    - All announcements have "authority" since the actual owner of the link spreads the information.
    - Broadcast latency can still cause problems

Traffic engineering

- If using a shortest-path routing protocol, can vary link cost
    - Usually, operators define hop costs using some magic formula to engineer traffic
    - Cisco defaults to: $\frac{1}{\text{capacity}}$ + transmission latency.
- MPLS provides more flexibility
    - By setting up virtual circuits, can control traffic more easily
    - Failure tolerance not as automatic − one router failure can take down many VCs