University of Massachusetts Amherst

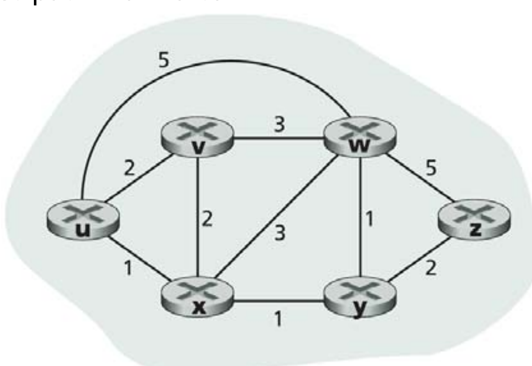
ECE697AA – Lecture 9

Routing: Distributed Routing Algorithms

Tilman Wolf
Department of Electrical and Computer Engineering
10/02/08

Routing algorithms

- Representation of network as graph
  - Routers as nodes
  - Links as edges
    - » Link weights determine cost
- Routing problem as graph problem
  - Find the least cost path from u to z

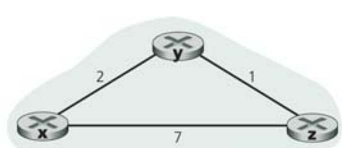


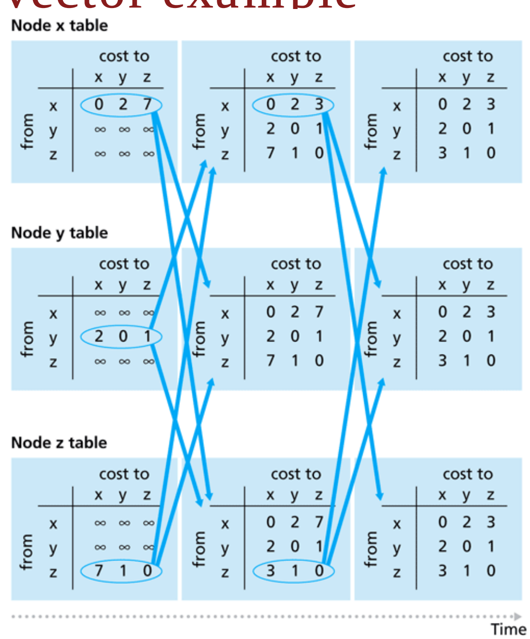
Distance vector algorithm

- Features
  - Distributed
  - Iterative
  - Asynchronous
- Each node reports local view
  - Cost to neighbors
  - Routes to others via neighbors
- Each node picks the best option
  - Bellman-Ford equation: $d_x(y) = \min_v\{c(x,v)+d_v(y)\}$
- Information is exchanged as "distance vector"
  - Shortest distance to all nodes as seen locally
- With enough exchanges, routing converges

Distance vector example

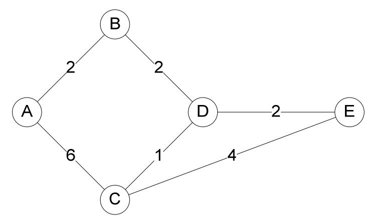
- Example:





Worksheet

- Try yourself:



D^A	B	C
B		
C		
D		
E		

D^B	A	D
A		
C		
D		
E		

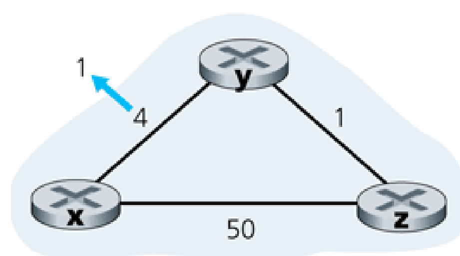
D^C	A	D	E
A			
B			
D			
E			

D^D	B	C	E
A			
B			
C			
E			

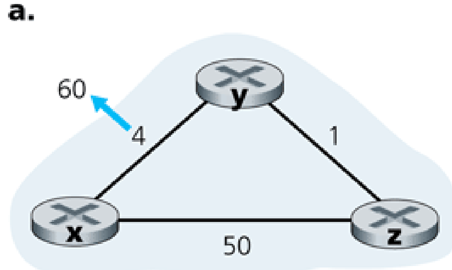
D^E	C	D
A		
B		
C		
D		

Distance vector problem

- Good news travels fast
  - y can reach x in 1
  - z can reach x in 2
- Bad news travels slowly
  - y can reach x in 6 (via z)
  - z can reach x in 7 (via y)
  - ...
- "Count-to-infinity" problem
  - Fix: "poisoned reverse"



a.



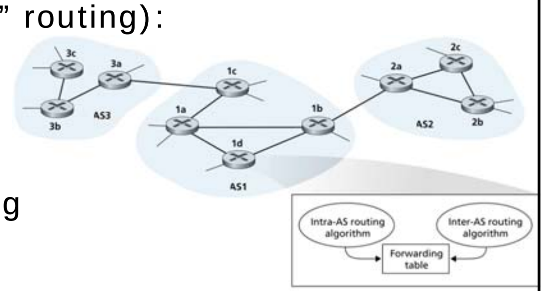
b.

Routing in the Internet

- How many nodes do we have in the Internet?
- How many links do we have in the Internet?
  - At least as many

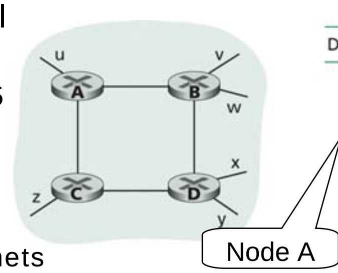
Autonomous Systems

- Scalability becomes a problem
  - Number of nodes/links in algorithm
  - Adding/removing machine could cause global routing update
- Internet is clustered into autonomous systems (AS)
  - Single administrative entity (e.g., company, university)
- Inside an AS ("local" routing):
  - Intra-AS routing protocol
- Between ASs ("global" routing):
  - Gateway routers connect ASs
  - Inter-AS routing protocol
- Combination of routing algorithm determines forwarding table



Intra-AS routing: RIP

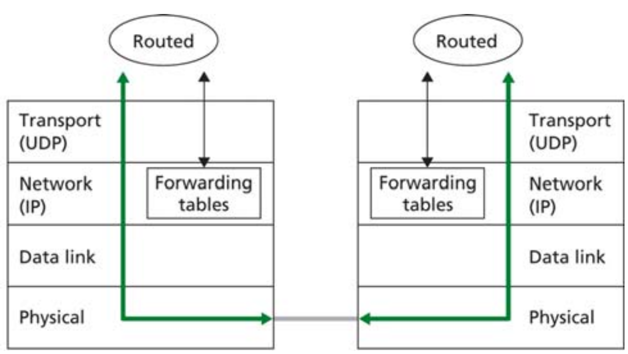
- Routing Information Protocol
  - Originally distributed in 1982 BSD UNIX
  - RFC 2453
- Distance vector protocol
  - “Hop” count as metric
  - Maximum hop count is 15
- Routing updates
  - Every 30 seconds
  - “RIP advertisement”
  - Up to 25 destination subnets
- Link considered down if no update in 180 seconds



Destination	Hops
u	1
v	2
w	2
x	3
y	3
z	2

Intra-AS routing: RIP

- RIP implementation
  - RIP uses UDP packets to exchange data
- Why transport layer for network layer routing?
  - “routed” is routing daemon in OS

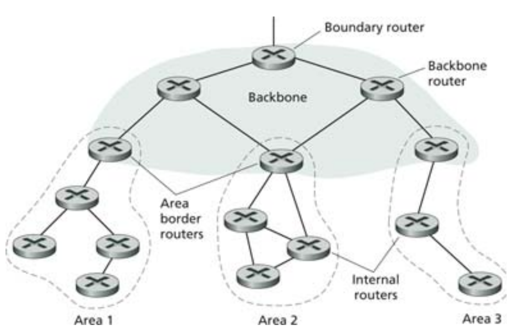


Intra-AS routing: OSPF

- Open Shortest Path First
  - "Open" as in "not proprietary"
  - RFC 2328
  - Designed as successor to RIP
- Link-state protocol
  - Routers have full graph of network
  - Dijkstra's algorithm for shortest path
  - Link weights set by administrator
    - » Difficult to achieve operational goals
- Routing updates
  - HELLO messages every 10 seconds (check if link is alive)
  - Flooding of link-state information
    - » Routers send link-state info to all other routers
  - Route update at least once every 30 minutes

Intra-AS routing: OSPF

- Advanced OSPF features
  - Security: MD5 authentication
  - Multiple same-cost paths
  - Unicast and multicast support
  - Support for hierarchy in single domain
- OSPF areas
  - Details within area not visible to outside
  - Simplifies administration of larger networks



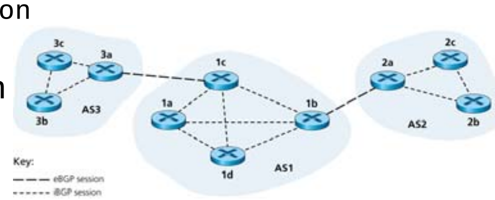
Inter-AS routing: BGP

- Border Gateway Protocol
  - De-facto standard for inter-AS routing in Internet
  - RFC 1771
- Advertisement of reachability

  "A subnet screams "I exist and I am here," and BGP makes sure that all the ASs in the Internet know about the subnet and how to get there. If it weren't for BGP, each subnet would be isolated – alone and unknown by the rest of the Internet."
- BGP provides
  - Information on subnet reachability from neighboring ASs
    - » Propagated to each internal router of AS
  - Means to determine "good" routes to subnets
    - » Based on reachability and AS policy

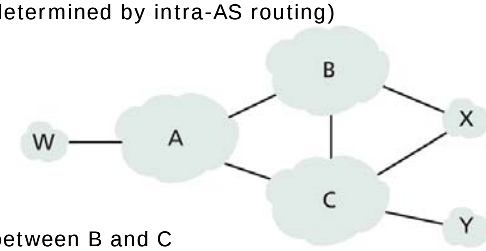
Inter-AS routing: BGP

- BGP sessions
  - Connection between routers to exchange BGP information
  - External BGP (eBGP) session
    - » Session spanning two ASs
  - Internal BGP (iBGP) session
    - » Session within one AS
- Reachability information
  - Reachable subnet (CIDR prefix)
  - BGP attributes
    - » AS-PATH: path to subnet (ASs traversed)
    - » Next-HOP: IP address of advertising router
- Path vector protocol
  - Information to avoid loop or other ASs (import policy)



Inter-AS routing: BGP

- Route selection:
  - Often multiple routes available
  - Elimination procedure:
    1. Local preference value set by administrator
    2. Shortest AS-PATH (=DV with AS hop metric)
    3. Closest NEXT-HOP router (determined by intra-AS routing)
       - "Hot potato routing"
    4. BGP identifiers
- Example
  - Y is "stub" network
  - X is "multihomed" network
    - » X is customer network
    - » X should not forward data between B and C
    - » X advertise as if stub domain (e.g., not XCY to B)
  - B might not want to advertise path to A or W to C



Inter-AS routing: BGP

- Peering agreements between ASs often confidential
  - Administrators are careful what to advertise
  - Avoid free riding of traffic from other ISPs
- BGP issues
  - BGP not always stable
  - Route flapping can cause further instability
    - » Router might get overloaded by BGP messages
    - » If router can't keep up, it might be considered down
  - Various heuristic fixes
    - » Route dampening

Assignments

- Read
  - Kurose & Ross: Chapter 3.6 & 3.7
- SPARK
  - Assessment quiz