The Network Layer: Routing in the Internet

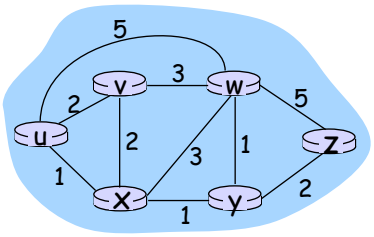
Smith College, CSC 249
March 8, 2017

Overview

- Review of theoretical routing algorithms
  - Link state & Dijkstra's algorithm
  - Distance vector & Bellman-Ford equation
- Routing in the Internet
  - Implementation of Link-state and Distance-vector in actual networks
  - Intra-networking & Inter-networking
  - RIP & OSPF
  - Border Gateway Protocol, BGP

Routing Algorithms

The objective of a routing algorithm is to find the least-cost, and loop-free path between all sources and all destinations (routers, not hosts)



3

Q: Compare & Contrast Routing

- What is the objective of routing algorithms?
- Compare and contrast Link-State and Distance-Vector
  - What does each do?
  - What does each do the same as the other algorithm?
  - What do they do differently?

4

Summary Table Comparing Link-State and Distance-Vector

Name	Algorithm/Equation Description & Overview	Initial Information, at Start-Up	Message Complexity and Volume	Router Malfunctions? Information that is shared? Router calculates what?
Link State				
Distance Vector				
Open Shortest Path First (OSPF)				
Routing Information Protocol (RIP)				
Borger Gateway Protocol (BGP)				

Overview of Routing & Recap…

- The set of routers comprise a distributed database
  - Routers propagate information to other routers
- Distribute State of Links
  - Advertise information about each link to which it is connected (the 'state' of the link)
  - Flood the network with this information
- Distribute Vectors
  - Advertise a vector with information on each destination it can 'reach' (entire network)
  - Communicates only with neighbors

6

A Link-State Routing Algorithm

Dijkstra's algorithm

- Computes the shortest paths in a graph by using weights on edges as a measure of distance.
- Each node has global information on network topology and edge weights
  - Starts with complete information

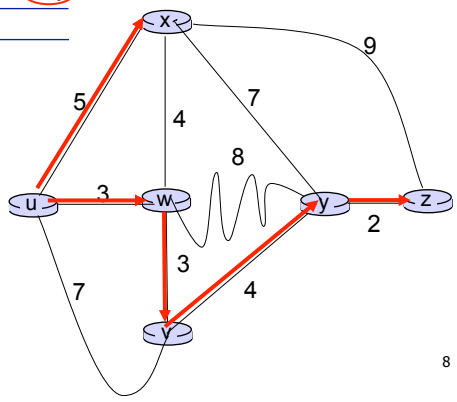
7

Dijkstra's algorithm: example

Step	N'	D(v) p(v)	D(w) p(w)	D(x) p(x)	D(y) p(y)	D(z) p(z)
0	u	7,u	3,u	5,u	∞	∞
1	uw	6,w		5,u	11,w	∞
2	uwx	6,w			11,w	14,x
3	uwxv				10,v	14,x
4	uwxvy					12,y
5	uwxvyz					

notes:

- Construct shortest path tree by tracing predecessor nodes
- Construct the forwarding table by recording the *next hop* to the destination node
- What is the forwarding table??



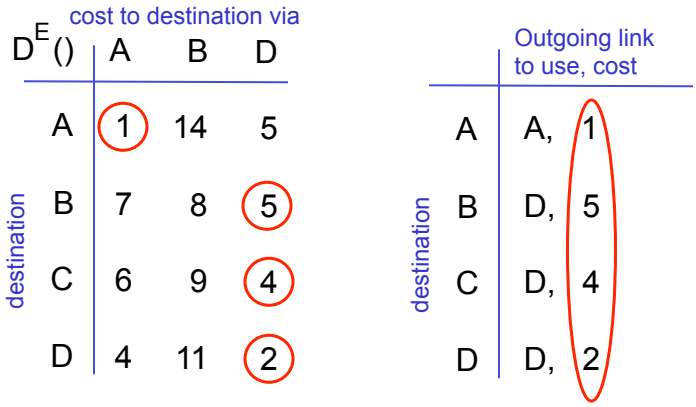
8

Distance vector algorithm

- Each node begins with $D_x(y)$
  - An estimate of the cost of the least-cost path from itself to node y, for all nodes in N (might be ∞)
  - (Some nodes might not be known to exist at the start)
- Each node periodically sends its own distance vector estimate to neighbors
  - ➔ A vector of least costs from itself to all other routers
- When a node x receives new DV estimate from neighbor, it updates its own DV using B-F equation, and sends any update to its neighbors

$$D_x(y) = min_v\{c(x,v) + D_v(y)\} \quad for\ each\ node\ y \in N$$

Distance table to forwarding table



Distance table → Forwarding Table

Distance Vector sent around: (A-1, B-5, C-4, D-2)

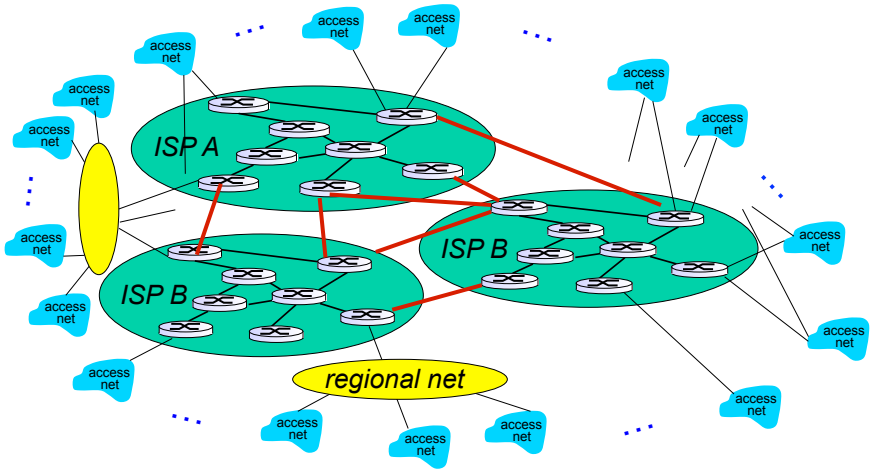
Hierarchical Routing & Autonomous Systems

- aggregate routers into regions, "autonomous systems" (AS)
- routers in same AS run the same routing protocol
  - "intra-AS" routing protocol
- routers in different AS can run different intra-AS routing protocol

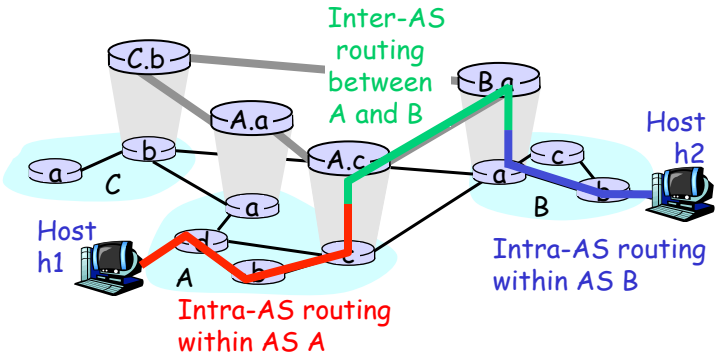
gateway routers

- special routers in AS
- run intra-AS routing protocol with all other routers in AS
- run *inter-AS routing* protocol with other gateway routers
  - *also* responsible for routing to destinations outside AS

Internet structure: network of networks

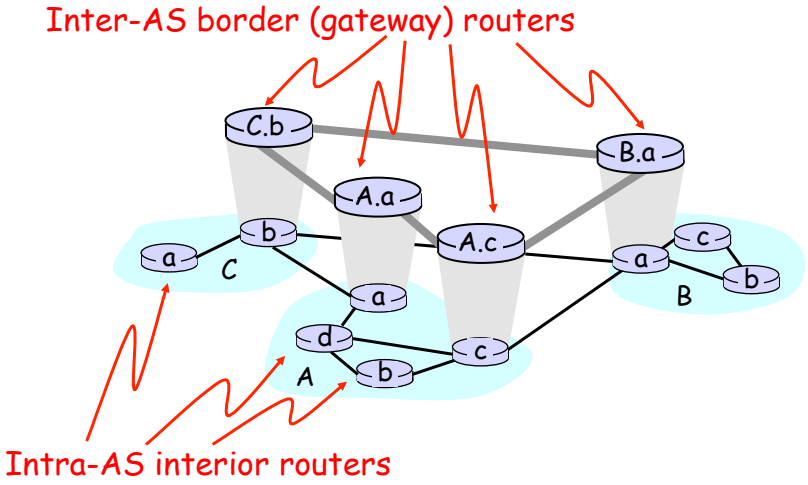


Intra-AS and Inter-AS routing



18

Internet AS Hierarchy



19

Intra-AS Routing

- Also known as *interior gateway protocols (IGP)*
- Most common intra-AS routing protocols:
  - RIP: Routing Information Protocol
    - Distance Vector
  - OSPF: Open Shortest Path First
    - Link State

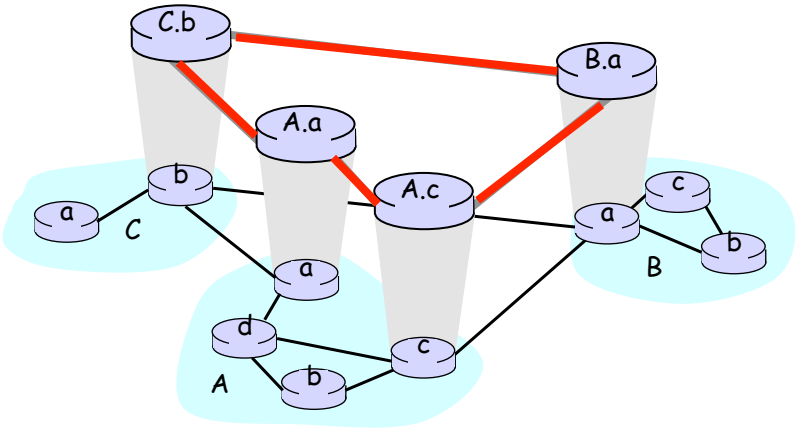
20

Question: RIP vs. OSPF

- Given what we know of LS and DV algorithms, compare the advertisements used by RIP and OSPF
- OSPF – (link state) router periodically broadcasts state of its attached links to **all** other routers in the AS
- RIP – (distance vector) information is sent about all the networks in the AS; is only sent to its neighboring routers

23

Inter-AS routing



24

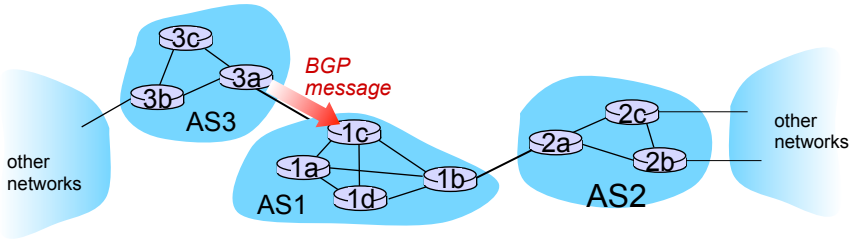
Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): *the* de facto standard
- **Path Vector** protocol:
  - similar to Distance Vector protocol
  - each Border Gateway broadcasts to neighbors (peers) *entire path* (*i.e.*, a sequence of ASs) to destination

25

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



26

Concepts: RIP vs. BGP

- RIP ads announce the number of hops to various destinations while BGP updates announce the _____ to various destinations
  - The sequence of ASs on the routes
- Describe how loops in paths can be detected in BGP.
  - Since full AS path information is available from an AS to a destination in BGP – if a BGP peer receives a route that contains its own AS number in the AS path, then using that route would result in a loop.

27

Different Intra- and Inter-AS routing

Policy:

- Inter-AS: admin wants control over how its traffic is routed, and who routes through its network.
- Intra-AS: single administrative staff, so no policy decisions needed

Scale:

- hierarchical routing saves table size & reduces traffic with update packets

Performance:

- Intra-AS: can focus on performance
- Inter-AS: policy may dominate over performance

28

Traceroute, Ping and ICMP

- Good discussion in text for how Traceroute and Ping work, using ICMP

29

ICMP: internet control message protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP messages are carried in IP datagrams
- ICMP message: type, code plus first 8 bytes of IP datagram causing error

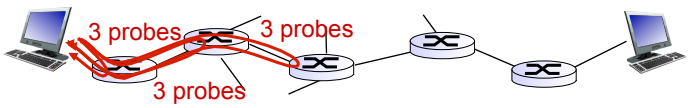
Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Traceroute and ICMP

- source sends series of UDP segments to dest
  - first set has TTL =1
  - second set has TTL=2, etc.
  - unlikely port number
- when *n*th set of datagrams arrives to nth router:
  - router discards datagrams
  - and sends source ICMP messages (type 11, code 0)
  - ICMP messages includes name of router & IP address
- when ICMP messages arrives, source records RTTs

stopping criteria:

- UDP segment eventually arrives at destination host
- destination returns ICMP "port unreachable" message (type 3, code 3)
- source stops



31

Broadcast Routing

- Uses? → Link-state routing algorithms
- Deliver packets from source to all other nodes

32

Multicast Routing

- Uses?
  - Bulk data (software upgrade) transfer
  - Streaming audio-visual media
  - Shared data application (teleconference)
  - Data feeds (stock quotes)
  - Interactive gaming

Summary for Network Layer

Forwarding:

- Leads to questions of addressing
  - Assignment of IP addresses (& DHCP)
  - NAT, IPv6 …

Routing:

- Routing objectives
- Routing notation
- Link state v. Distance Vector
- Hierarchical structure