

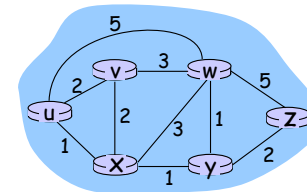
The Network Layer: Routing II

Smith College, CSC 249
March 28, 2008

slides mostly from J.F Kurose and K.W. Ross,
copyright 1996-2007

Routing Algorithms

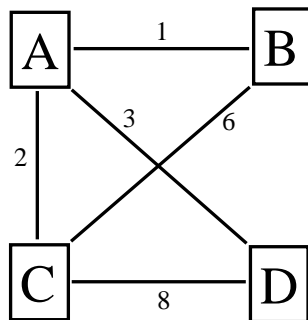
The objective of a **routing algorithm** is to find the least-cost path between all sources and all destinations



2

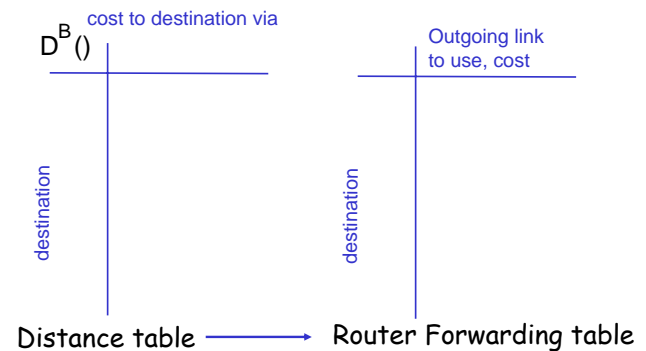
Recap Routing Algorithms

- Compare and contrast:
 - ❖ Global, link state algorithm
 - ❖ Distributed, distance vector algorithm



3

Distance table gives routing table



$$D^X(Y,Z) = \text{distance from X to Y, via Z as next hop} \\ = c(X,Z) + \min \{D_W^Z(Y,w)\}$$

4

Comparison of LS and DV algorithms

- Information requirements
- Message complexity
- Convergence time varies
- Robustness: what happens if router malfunctions?
- Oscillations possible?
- Loops possible?

5

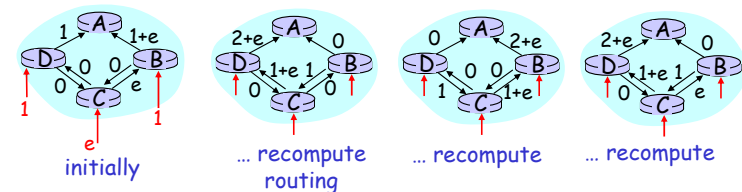
Dijkstra's Algorithm: Problem

Algorithm complexity: n nodes

- each iteration: need to check all nodes, w , not in N
- $n^*(n+1)/2$ comparisons: $O(n^2)$ worst case

Oscillations possible: (ADAPTING TOO QUICKLY)

- e.g., link cost = amount of carried traffic

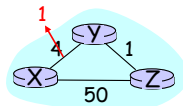


6

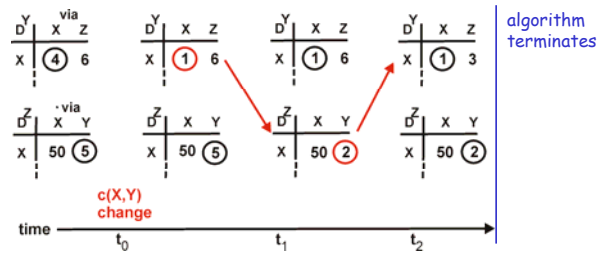
Distance Vector: link cost changes

Link cost changes:

- node detects local link cost change
- updates distance table (line 15)
- if cost change in least cost path, notify neighbors (lines 23,24)

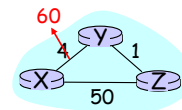


"good news travels fast"



7

Distance Vector: link cost changes



$$D^Y(X,X) = c(Y,X) + \min_w \{D^X(X,w)\}$$

$$D^Y(X,Z) = c(Y,Z) + \min_w \{D^Z(X,w)\}$$

...

$$D^Z(X,X)$$

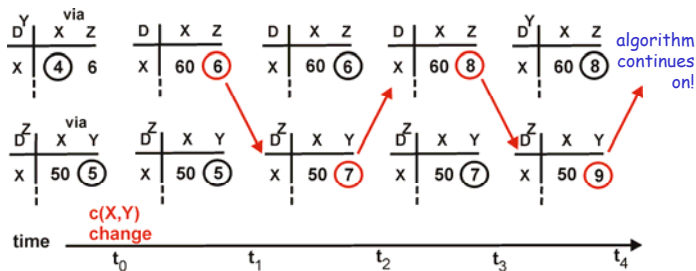
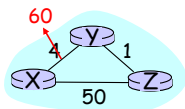
$$D^Z(X,Y)$$

8

Distance Vector: link cost changes

Link cost changes:

- good news travels fast
- bad news travels slowly - "count to infinity" problem!



9

Comparison of LS and DV algorithms

Message complexity

- **LS:** with n nodes, E links, $O(nE)$ msgs sent each
- **DV:** exchange between neighbors only
 - ❖ convergence time varies

Speed of Convergence

- Start-up v. steady-state
- **LS:** $O(n^2)$ algorithm requires $O(nE)$ msgs
 - ❖ may have oscillations
- **DV:** convergence time varies
 - ❖ may be routing loops
 - ❖ count-to-infinity problem

Robustness: what happens if router malfunctions?

LS:

- ❖ node can advertise incorrect *link* cost
- ❖ each node computes its *own* table using own info

DV:

- ❖ DV node can advertise incorrect *path* cost
- ❖ each node's table used by others
 - errors propagate through the network

10

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

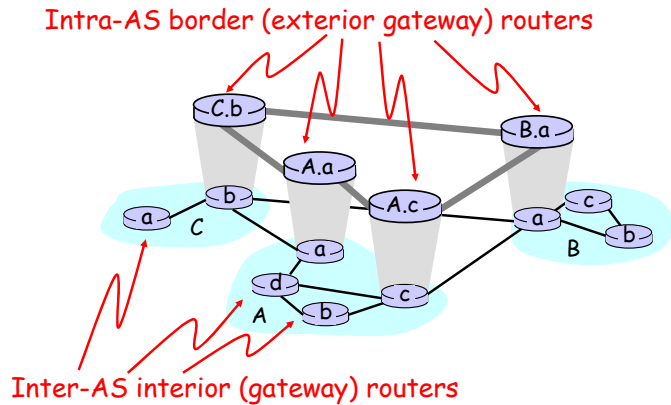
11

Routing Algorithms

- Each host is connected to a default router (the first-hop router & source router)
 - ❖ Each packet transmitted by a host is sent to this source router
- Packet needs to be routed from the source router to the destination router

12

Internet AS Hierarchy



13

Intra-AS Routing

- Also known as **Interior Gateway Protocols (IGP)**
- Most common IGPs:
 - ❖ RIP: Routing Information Protocol
 - ❖ OSPF: Open Shortest Path First
 - ❖ IGRP: Interior Gateway Routing Protocol (Cisco proprietary)

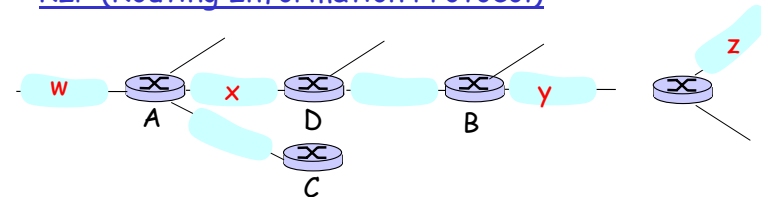
14

RIP (Routing Information Protocol)

- Distance vector algorithm
- Distance metric: # of hops
 - ❖ each link has a cost of 1
 - ❖ a hop refers to a subnet (not individual links), including the destination subnet
 - ❖ maximum allowed cost of a path is 15, implying what?
- Distance vectors
 - ❖ estimate of the shortest path from router to the subnets in the AS
 - ❖ exchanged every 30 sec via Response Message (also called **advertisement**)

15

RIP (Routing Information Protocol)



Routing table in D

Destination Network	Next Router	Num. of hops to dest.
W	A	2
Y	B	2
Z	B	7
X	--	1
....

16

RIP: Link Failure and Recovery

If no advertisement heard after 180 sec --> neighbor/link declared dead

- ❖ new advertisements sent to neighbors
- ❖ link failure info quickly propagates to entire net
- ❖ poison reverse used to prevent ping-pong loops

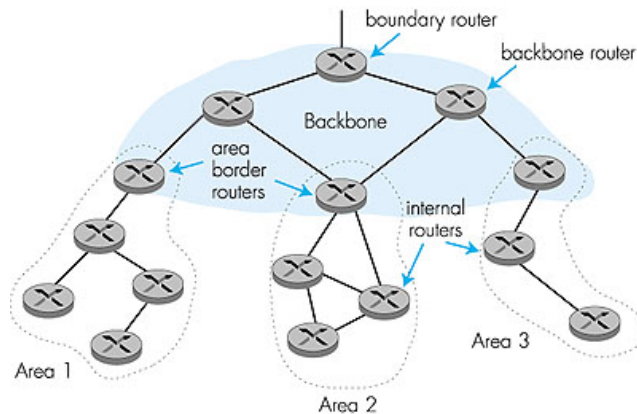
17

OSPF (Open Shortest Path First)

- ❑ "open": publicly available
- ❑ Uses Link State algorithm
- ❑ Advertisements disseminated to **entire AS** (via flooding)
- ❑ **Security**: all OSPF messages authenticated (to prevent malicious intrusion);
- ❑ **Hierarchical** OSPF in large domains.

18

Hierarchical OSPF



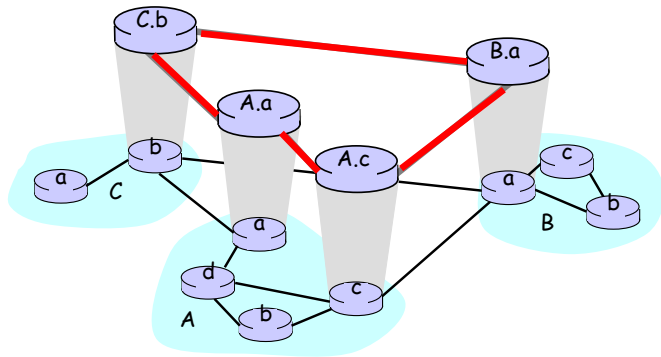
19

Concepts: RIP vs. OSPF

- ❑ Compare the advertisements used by RIP and OSPF
- ❑ OSPF - router periodically broadcasts routing information to **all** other routers in the AS
- ❑ RIP - advertisement sent by a router contains information about all the networks in the AS; is only sent to its neighboring routers

20

Inter-AS routing



21

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., sequence of ASs) to destination

22

Why different Intra- and Inter-AS routing ?

Policy:

- ❑ Inter-AS: admin wants control over **how** its traffic is routed, and **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

23

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.

24

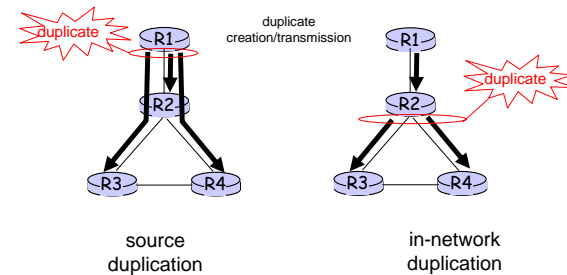
Chapter 4: Network Layer

- ❑ 4.1 Introduction
- ❑ 4.2 Virtual circuit and datagram networks
- ❑ 4.3 What's inside a router
- ❑ 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- ❑ 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- ❑ 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- ❑ 4.7 Broadcast and multicast routing

25

Broadcast Routing

- ❑ Uses?
- ❑ Deliver packets from source to all other nodes
- ❑ Source duplication is inefficient:



26

Multicast Routing: Problem Statement

- ❑ Uses?

- ❑ **Goal:** find a tree (or trees) connecting routers having local multicast group members

Network Layer Summary

- ❑ Services offered and received?
- ❑ Implementation?
 - ❖ Terms? / Principles?
- ❑ Structure of the network layer?

Getting a datagram from source to destination - first glimpse of the link-layer (for next class)

29

Getting a datagram from source to dest.

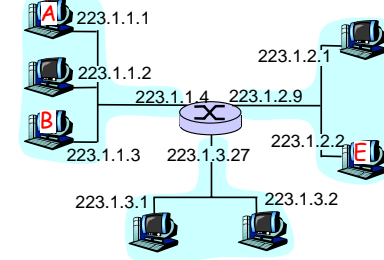
IP datagram:

misc fields	source IP addr	dest IP addr	data
-------------	----------------	--------------	------

- datagram remains unchanged, as it travels source to destination
- addr fields of interest here

routing table in A

Dest. Net.	next router	Nhops
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



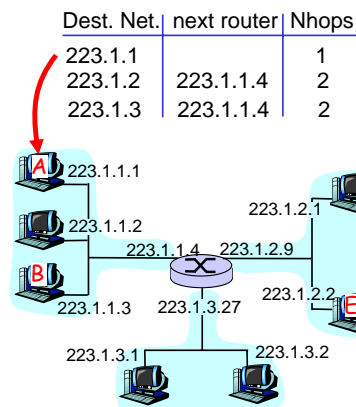
30

Getting a datagram from source to dest.

misc fields	223.1.1.1	223.1.1.3	data
-------------	-----------	-----------	------

Starting at A, given IP datagram addressed to B:

- look up net. address of B
- find B is on same net. as A
- link layer will send datagram directly to B inside link-layer frame
 - B and A are directly connected



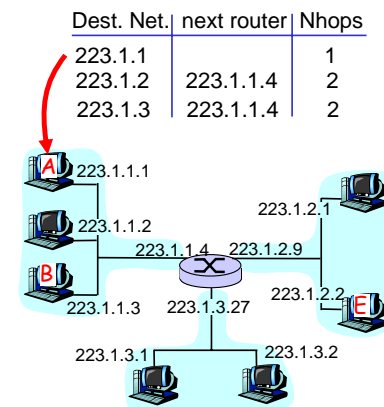
31

Getting a datagram from source to dest.

misc fields	223.1.1.1	223.1.2.2	data
-------------	-----------	-----------	------

Starting at A, dest. E:

- look up network address of E
- E on different network
 - A, E not directly attached
- routing table: next hop router to E is 223.1.1.4
- link layer sends datagram to router 223.1.1.4 inside link-layer frame
- datagram arrives at 223.1.1.4
- continued....



32

Getting a datagram from source to dest.

misc	223.1.1.1	223.1.2.2	data
fields			

Arriving at 223.1.4,
destined for 223.1.2.2

- ❑ look up network address of E
- ❑ E on *same* network as router's interface 223.1.2.9
 - ❖ router, E directly attached
- ❑ link layer sends datagram to 223.1.2.2 inside link-layer frame via interface 223.1.2.9
- ❑ datagram arrives at 223.1.2.2!!! (hooray!)

Dest. network	next router	Nhops	interface
223.1.1	-	1	223.1.1.4
223.1.2	-	1	223.1.2.9
223.1.3	-	1	223.1.3.27

