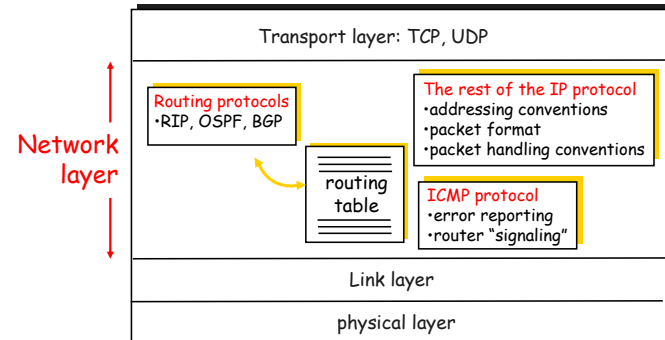


Internet Routing

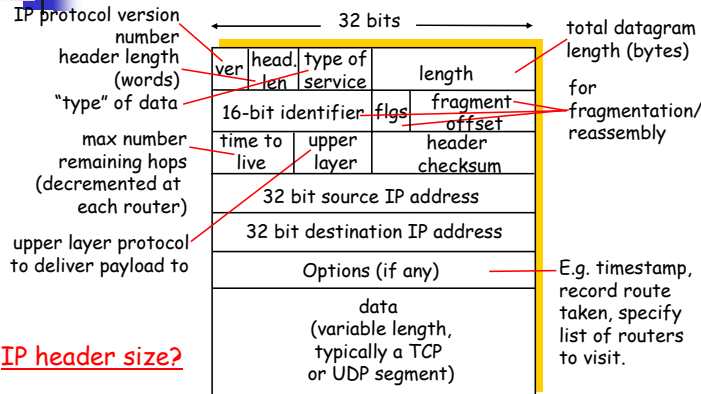
Kai Shen

Dept. of Computer Science, University of Rochester

The Internet Network Layer



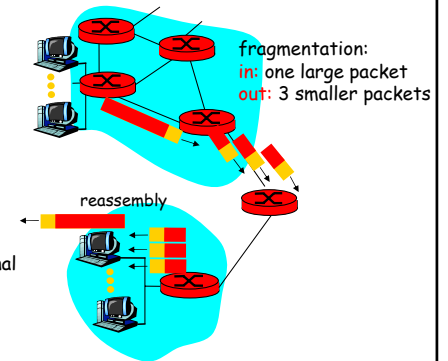
IP Packet Format



IP header size?

IP Fragmentation & Reassembly

- network links have MTU - largest possible frame.
 - different link types, different MTUs
- large IP packet divided ("fragmented") within network
 - one packet becomes several packets
 - "reassembled" only at final destination
 - IP header bits used to identify, order related fragments



Routing in the Internet

- The Global Internet consists of **Autonomous Systems (AS)** interconnected with each other:
 - Stub AS**: small corporation: one connection to other AS's
 - Multihomed AS**: large corporation (no transit): multiple connections to other AS's
 - Transit AS**: provider, hooking many AS's together
- Two-level routing:
 - Intra-AS**: administrator responsible for choice of routing algorithm within network
 - Inter-AS**: unique standard for inter-AS routing: BGP

10/1/2007

CSC 257/457 - Fall 2007

9

RIP: An Intra-AS Routing Protocol

RIP (Routing Information Protocol)

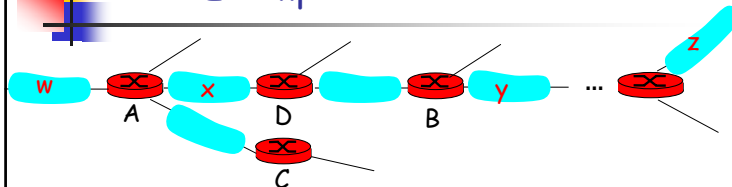
- Distance vector routing algorithm
- For scalability, each entry in routing table typically represents a **network**
- Distance metric: # of hops
- Distance vectors: exchanged among neighbors periodically (e.g., every 30 secs)
 - this exchange is also called route **advertisement**

10/1/2007

CSC 257/457 - Fall 2007

10

RIP: Example



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

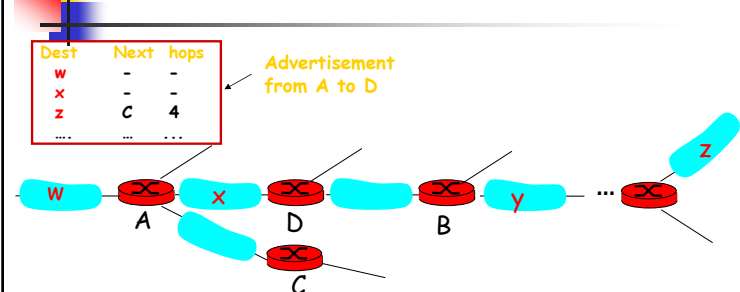
Routing table in D

10/1/2007

CSC 257/457 - Fall 2007

11

RIP: Example



Dest	Next	hops
w	-	-
x	-	-
z	C	4
....

Advertisement from A to D

Destination Network	Next Router	Num. of hops to dest.
W	A	2
Y	B	2
Z	B A	2 5
X	--	1
....

Routing table in D:

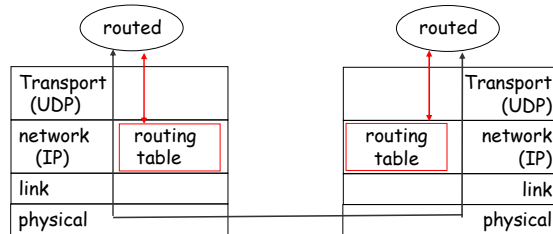
10/1/2007

CSC 257/457 - Fall 2007

12

RIP Table processing

- RIP routing tables managed by user-level process called route-d (daemon)
- advertisements sent in UDP packets, periodically repeated



10/1/2007

CSC 257/457 - Fall 2007

13

RIP Table Example

Check routing table at your own workstation:
netstat -rn

Destination	Gateway	Interface
128.151.67.0/24	0.0.0.0	eth1
0.0.0.0	128.151.67.100	eth1

- A single attached class C IP network
- Routing tables in routers are more complex

10/1/2007

CSC 257/457 - Fall 2007

14

OSPF: Another Intra-AS Routing Protocol

OSPF (Open Shortest Path First)

- Uses Link State algorithm
 - Disseminate link state information to entire AS
 - Complete topology map at each node
 - Route computation using Dijkstra's algorithm
- Solve the scalability problem in OSPF
 - Hierarchical OSPF in large AS

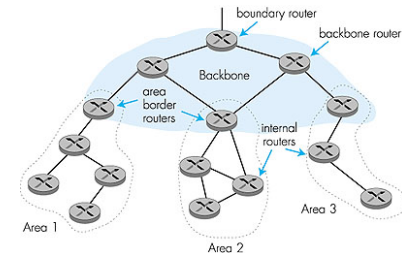
10/1/2007

CSC 257/457 - Fall 2007

15

Hierarchical OSPF

- Two-level hierarchy: local area, backbone.
 - Link-state broadcasts only in area
 - each node has detailed area topology



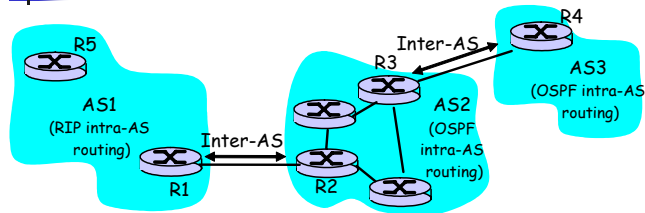
- Routers:
 - **Area border routers:** run intra-area OSPF routing as well as backbone OSPF routing.
 - **Backbone routers:** run OSPF routing limited to backbone.
 - **Boundary routers:** connect to other AS's.

10/1/2007

CSC 257/457 - Fall 2007

16

Inter-AS Routing in the Internet



- There should be just one inter-AS routing protocol
 - BGP (Border Gateway Protocol)
- Inter-AS routing
 - Across multiple administrative domains, which can have conflicting interests

10/1/2007

CSC 257/457 - Fall 2007

17

BGP Routing Protocol

- Path Vector protocol:
 - similar to Distance Vector protocol
 - each Border Gateway broadcast to neighbors (peers) entire path (i.e., sequence of AS's) to destination

Policy freedom in BGP

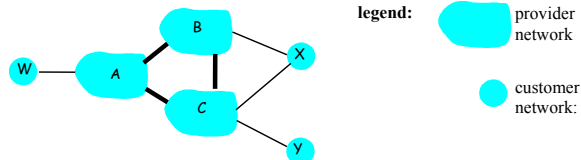
- A router selects routes offered by neighbors based on
 - cost
 - policy (don't route via competitors AS)
- A router can control incoming traffic by controlling its route advertisements to neighbors:
 - e.g., don't want to route traffic from Z -> don't advertise any routes to Z

10/1/2007

CSC 257/457 - Fall 2007

18

BGP: controlling who routes to you



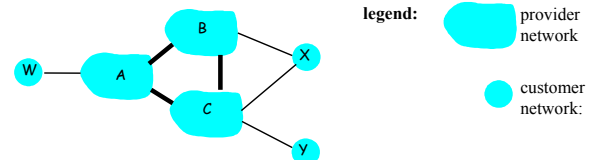
- A, B, C are **provider networks**
- X, W, Y are **customers** (of provider networks)
- X is **dual-homed**: attached to two networks
 - X does not want to route transit traffic between B and C
 - ⇒ X only advertises routes to itself

10/1/2007

CSC 257/457 - Fall 2007

19

BGP: controlling who routes to you



- A advertises to B the path AW
- B advertises to X the path BAW
- Should B advertise to C the path BAW?
 - B wants to route **only** to/from its customers!
 - B doesn't want to route CBAW since neither W nor C are B's customers.

10/1/2007

CSC 257/457 - Fall 2007

20



BGP routing table ⇒ whole-internet topology

- BGP routing table at an AS records complete paths to all other ASes
- Possible to construct (almost) complete internet topology (AS-level) with BGP routing tables from several routers



Why two-level Intra- and Inter-AS routing ?

Scale:

- hierarchical routing saves table size, reduced update traffic

Policy vs. performance:

- Inter-AS: admin wants control over how its traffic routed, who routes through its network. Performance in a global sense is not a major concern for individual service provider.
- Intra-AS: single admin, can focus on performance.



Disclaimer

- Parts of the lecture slides contain original work of James Kurose, Larry Peterson, and Keith Ross. The slides are intended for the sole purpose of instruction of computer networks at the University of Rochester. All copyrighted materials belong to their original owner(s).