

## Internet Routing

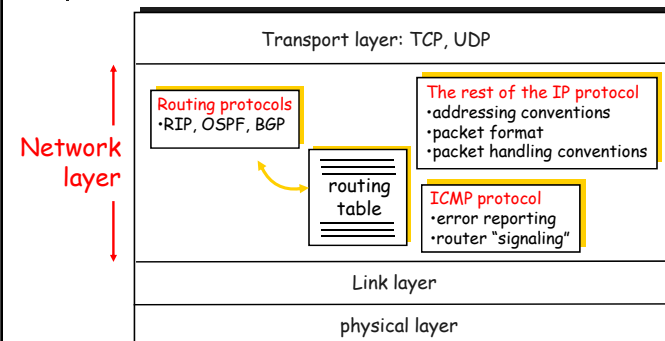
Kai Shen

9/30/2009

CSC 257/457 - Fall 2009

1

## The Internet Network Layer

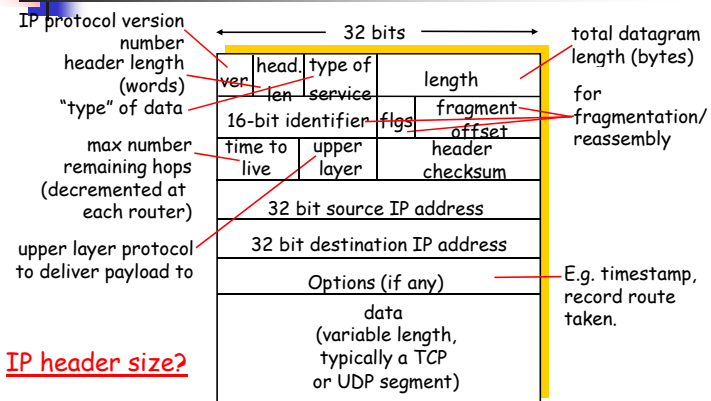


9/30/2009

CSC 257/457 - Fall 2009

2

## IP Packet Format



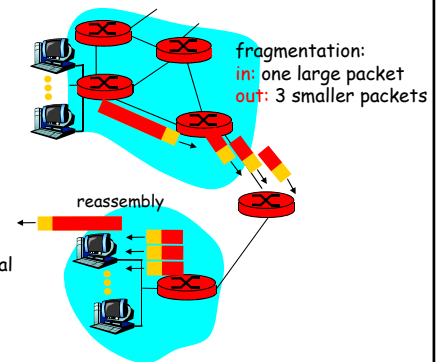
9/30/2009

CSC 257/457 - Fall 2009

3

## 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



9/30/2009

CSC 257/457 - Fall 2009

4

## Principles of Internet Routing

Our routing study thus far - idealization

- uniform routing scheme across whole network
- uniform optimization goal
- ... does **not** work well in practice

**scale:** with hundreds of millions of destinations:

- can't store all dest's in routing tables!
- routing table exchange would swamp links!

**administrative autonomy**

- internet = network of networks
- each network admin may want to control routing in its own network

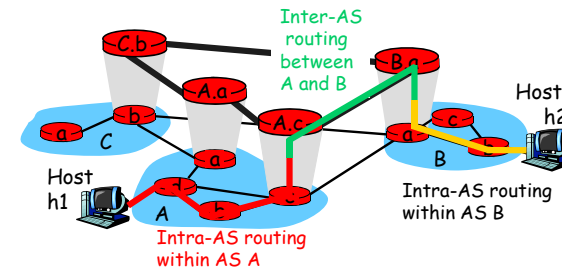
9/30/2009

CSC 257/457 - Fall 2009

5

## Hierarchical Routing

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

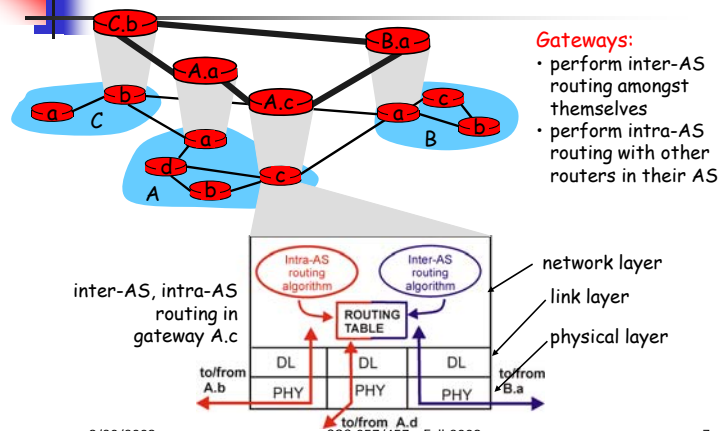


9/30/2009

CSC 257/457 - Fall 2009

6

## Intra-AS and Inter-AS Routing



9/30/2009

CSC 257/457 - Fall 2009

7

## Practices of Internet Routing

- 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

9/30/2009

CSC 257/457 - Fall 2009

8

## RIP: An Intra-AS Routing Protocol

RIP ( Routing Information Protocol)

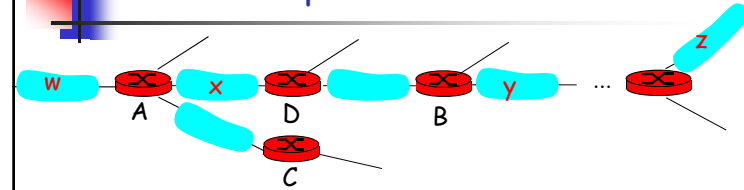
- Distance vector routing algorithm
- For scalability, each entry in routing table typically represents a network
- Distance metric: # of hops
- Routing tables exchanged among neighbors periodically (e.g., every 30 secs)
  - this exchange is also called route advertisement

9/30/2009

CSC 257/457 - Fall 2009

9

## 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

9/30/2009

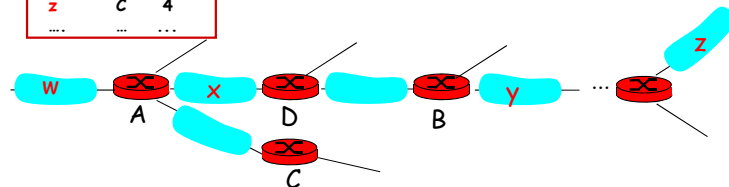
CSC 257/457 - Fall 2009

10

## RIP: Example

Dest	Next	hops
W	-	-
X	-	-
Z	C	4
...	...	...

Advertisement from A to D



Routing table in D:

Destination Network	Next Router	Num. of hops to dest.
W	A	2
Y	B	2
Z	<del>B</del> A	<del>7</del> 5
X	--	1
...	...	...

9/30/2009

CSC 257/457 - Fall 2009

11

## Routing Table Example

Check routing table at your own workstation:  
netstat -rn

Destination	Gateway	Interface
128.151.67.0/24	0.0.0.0	eth0
169.254.0.0/16	0.0.0.0	eth0
0.0.0.0	128.151.67.100	eth0

- Local IP network contains two segments of IP addresses
- Routing tables in routers are more complex

9/30/2009

CSC 257/457 - Fall 2009

12

## 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

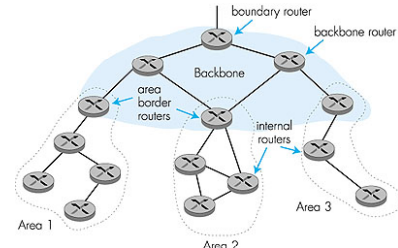
9/30/2009

CSC 257/457 - Fall 2009

13

## 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.

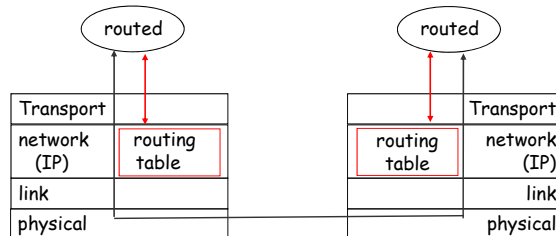
9/30/2009

CSC 257/457 - Fall 2009

14

## Network Routing Architecture within A Host System

- Routing table and routing policy implementation must be protected
  - Kept in the operating system
  - Or managed by protected user process (root-owned daemon)

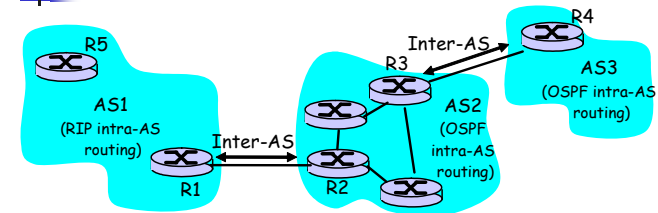


9/30/2009

CSC 257/457 - Fall 2009

15

## 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

9/30/2009

CSC 257/457 - Fall 2009

16

## BGP Routing Protocol

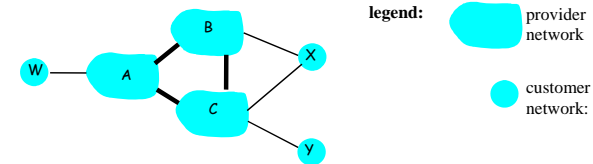
- **Path Vector** protocol:
  - similar to Distance Vector protocol
  - each Border Gateway broadcast to neighbors (peers) entire path (i.e., sequence of AS's) to destination
- Policy freedom in BGP
  - A router selects routes offered by neighbors based on
    - cost
    - policy (don't route via competitors AS)
  - A router can control incoming traffic by manipulating its route advertisements to neighbors:
    - e.g., don't want to route traffic from Z → don't advertise any routes to Z

9/30/2009

CSC 257/457 - Fall 2009

17

## BGP: routing policy management



- A, B, C are **provider networks**
- X, W, Y are **customers** (of provider networks), X is **dual-homed**
- X does not want to route transit traffic between B and C
  - ⇒ X only advertises routes to itself
- A wants to help its customer (w) but only anybody else
  - ⇒ A tells B, C about link to w but not link to C, B

9/30/2009

CSC 257/457 - Fall 2009

18

## BGP Routing Table

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

9/30/2009

CSC 257/457 - Fall 2009

19

## 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.

9/30/2009

CSC 257/457 - Fall 2009

20



## 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).