CSCE 463/612
Networks and Distributed Processing
Spring 2024

Network Layer VI

Dmitri Loguinov
Texas A&M University

April 19, 2024

Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

4.5 Routing algorithms

- Link state
- Distance Vector
- Hierarchical routing

4.6 Routing in the Internet

4.7 Broadcast and multicast routing

Hierarchical Routing

Problems in practice:

- Memory: can't store entire Internet graph in router memory
- CPU time: can't overload routers with huge computational expense
- Message overhead: routing table exchanges would overload network
- Competitiveness: ISPs not willing to share their topology with others

Solution: administrative autonomy

- Internet = network of networks
- Network admins control routing in their own networks, export reachable subnets to outside world

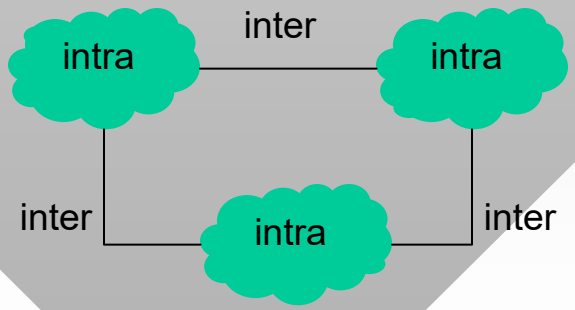
Hierarchical Routing

- Aggregate routers into regions called AS (Autonomous Systems)
- Routers in the same AS run the same algorithm
  - Accomlished via intra-AS routing protocols
- ISPs gain flexibility
  - Routers in different ASes can run different intra-AS protocols that cannot directly speak to each other, which is OK

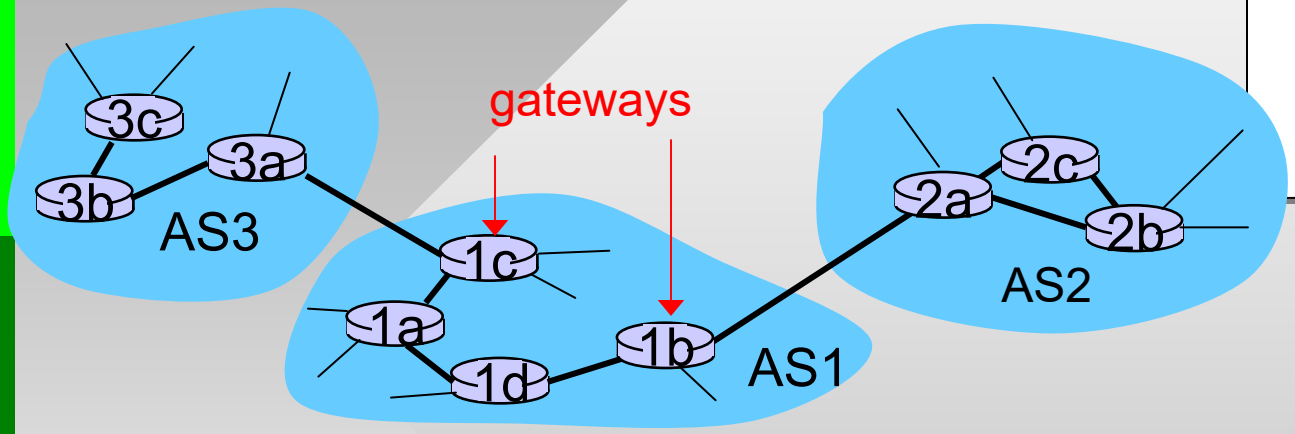
Texas A&M owns AS3794 with two subnets: 128.194/16 and 165.91/16

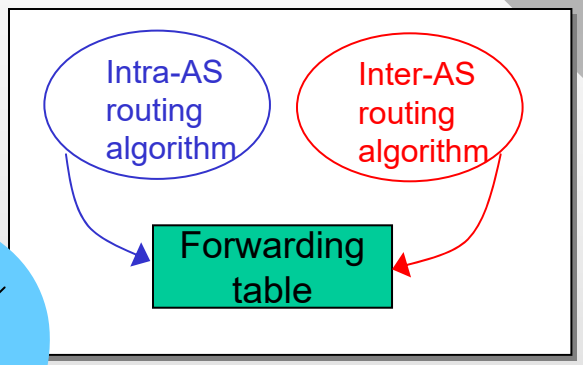
Gateway (border) routers

- Direct links to routers in other ASes
- Exchange routing view of each AS using an inter-AS protocol
  - Summary of subnets to which this AS is willing to route



Interconnected ASes





- Intra-AS sets entries for all internal dests
  - E.g., 1a plots shortest path to 1b using link-state alg
- Inter-AS accepts external dests from neighbor ASes
  - E.g., 1b learns 128.194/16 is reachable via AS2
- Inter-AS broadcasts pairs (subnet, exit router)
  - E.g., 1b notifies all routers in AS1 that it can reach 128.194/16

Example: Choosing Among Multiple ASes

- Now suppose AS1 learns from the inter-AS protocol that 128.194/16 is reachable from AS3 *and* from AS2
  - To configure forwarding table, routers in AS1 must determine towards which exit (1c or 1b) they must forward packets
- This is also the job of inter-AS routing protocol!
  - Usually based on ISP policy, SLAs, prior traffic engineering
- Common strategies
  - Hot potato routing: send packet towards closest of the alternative exit points
  - Shortest AS path: fewest ISPs on the way to the target network

Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

4.5 Routing algorithms

4.6 Routing in the Internet

- RIP
- OSPF
- BGP

4.7 Broadcast and multicast routing

Routing Protocols

- Common intra-AS routing protocols:
  - RIP: Routing Information Protocol (DV)
  - OSPF: Open Shortest Path First (LS)
  - IGRP: Interior Gateway Routing Protocol (Cisco proprietary, DV, now obsolete); EIGRP (Extended IGRP, still DV, open sourced in 2013)
  - IS-IS (Intermediate System to Intermediate System, LS)
- For Inter-AS, there is now just one option
  - BGP (Border Gateway Protocol)
  - All ISPs must support it

Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

4.5 Routing algorithms

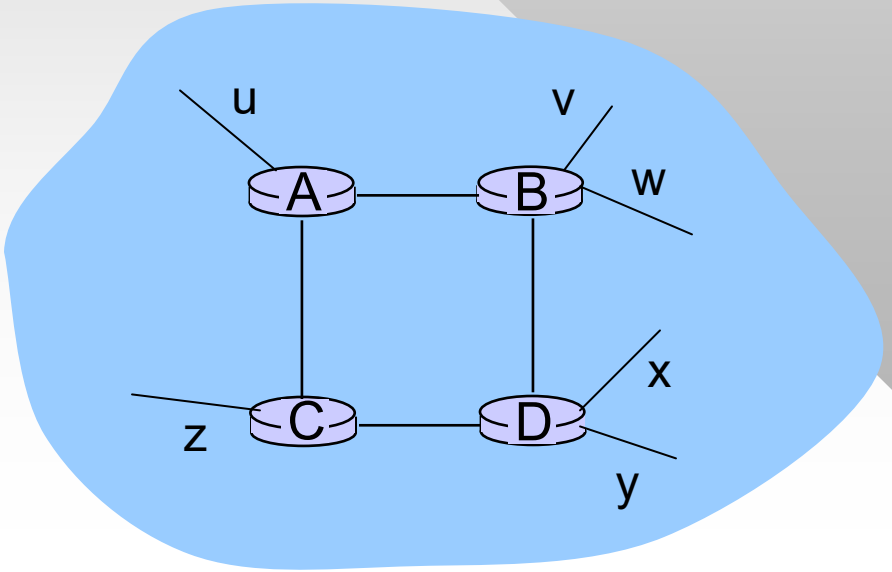
4.6 Routing in the Internet

- RIP
- OSPF
- BGP

4.7 Broadcast and multicast routing

RIP (Routing Information Protocol)

- Included in BSD-UNIX distribution in 1982
  - Distance vector (DV) algorithm
- Distance metric: # of hops (max = 15)
  - Distance vectors: exchanged among neighbors every 30 sec using advertisement messages
  - Each message: lists of up to 25 destination nets within AS



destination subnet	hops from A
u	1
v	2
w	2
x	3
y	3
z	2

Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

4.5 Routing algorithms

4.6 Routing in the Internet

- RIP
- OSPF
- BGP

4.7 Broadcast and multicast routing

OSPF (Open Shortest Path First)

- "Open": protocol specifications publicly available
  - v1 (1989), v2 (1998), and v3 (2008)
- Uses Link State (LS) algorithm
  - LS packet dissemination
  - Topology map at each node
  - Route computation using Dijkstra's algorithm
- Advertisements disseminated to entire AS (via flooding)
  - Carried in OSPF messages directly over IP (rather than TCP or UDP) using protocol number 89
  - Layer 3.5 similar to ICMP
  - Handles own error detection/correction

Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

4.5 Routing algorithms

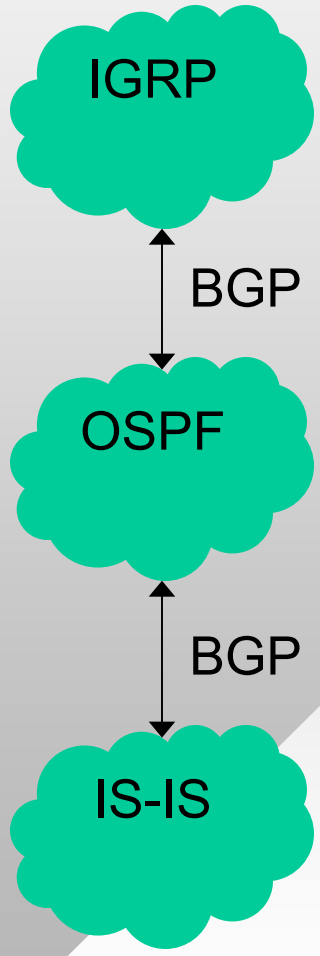
4.6 Routing in the Internet

- RIP
- OSPF
- BGP

4.7 Broadcast and multicast routing

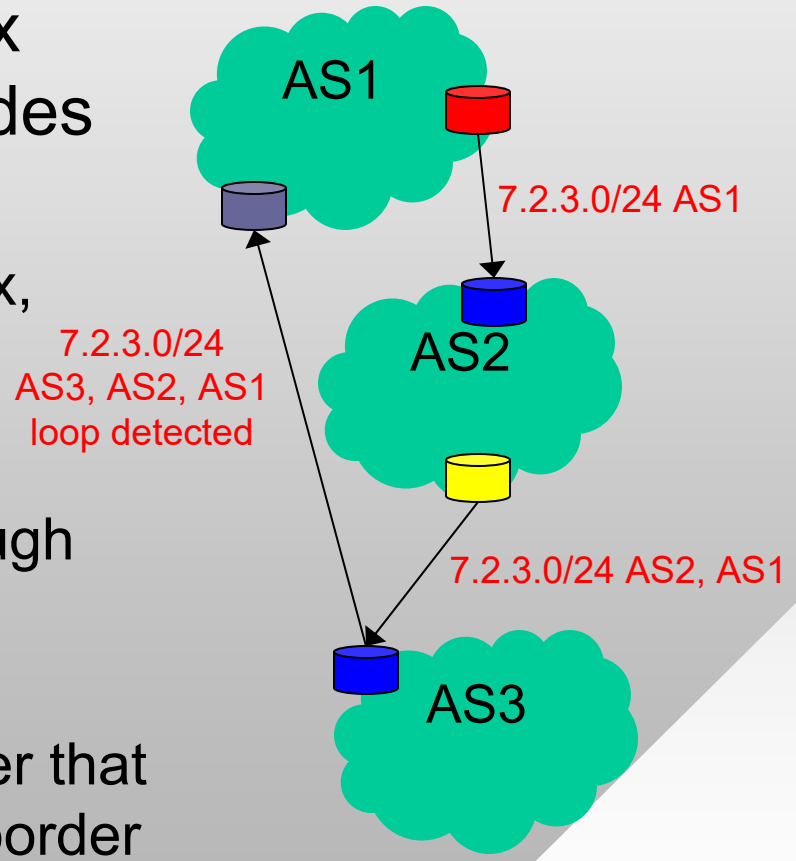
Inter-AS Routing: BGP

- BGP (Border Gateway Protocol): de facto standard for inter-AS (exterior) routing
- BGP provides each AS a means to:
  - Obtain subnet reachability information from neighboring ASes
  - Propagate the reachability information to all routers internal to the AS
  - Determine “good” routes to subnets based on reachability information and policy
- Allows a subnet to advertise its existence to the rest of the Internet: *“I am here”*



Path Attributes & BGP Routes

- When advertising an IP prefix (i.e., subnet), message includes BGP attributes
  - Notation: combination (IP prefix, attributes) = route
- Two important attributes:
  - AS-PATH: contains ASes through which the advert for the prefix passed (latest AS first)
  - NEXT-HOP: indicates the router that should receive traffic (usually border router of the AS that advertised prefix; multiple values possible)



Chapter 4: Roadmap

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

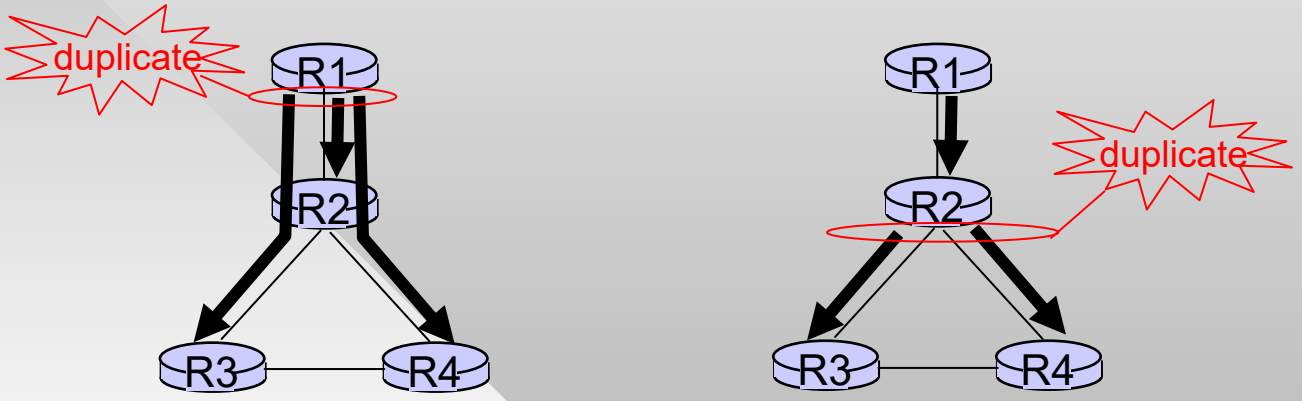
4.5 Routing algorithms

4.6 Routing in the Internet

4.7 Broadcast and multicast routing

Multicast and Broadcast

- Broadcast: send a packet to all hosts in the network
- Multicast: send to a certain subset of nodes
- Unicast: one sender - one receiver



(a) unicast (5 pkt-links)

(b) multicast (3 pkt-links)

- Example: video distribution to 1M receivers via unicast
  - First link R1-R2 carries each packet 1M times
  - 5 Mbps stream requires a 5-Tbps link!

Internet Multicast Routing: DVMRP

- DVMRP: Distance Vector Multicast Routing Protocol, RFC 1075 (1988)
- Builds a distribution tree rooted at source
  - Based on DVMRP's own routing tables constructed by communicating DVMRP routers
  - No assumptions about underlying unicast
- IGMP broadcasts proceed between neighbor routers
- Multicast IP addresses are in 224.0.0.0/4
  - To join a particular group, use setsockopt with IP_ADD_MEMBERSHIP

Multicast Future

- Wide-area multicast deployment has been traditionally slow, now practically dead
  - Mbone was one such endeavor, worked via tunnels
- One issue is scalability
  - Flooding all Internet receivers is dangerous/expensive
  - Opens loopholes for DoS attacks
- Another is ISP unwillingness to accept multicast traffic
  - Who pays for a single packet being replicated 1M times?
- Finally, multicast congestion control is hard
  - Mbone had 30-40% loss, which is much more than most applications can tolerate (typically below 1%)
  - How to recover lost packets?