The Network Layer Forwarding Tables and Switching 'Fabric'

Smith College, CSC 249
February 27, 2018

1

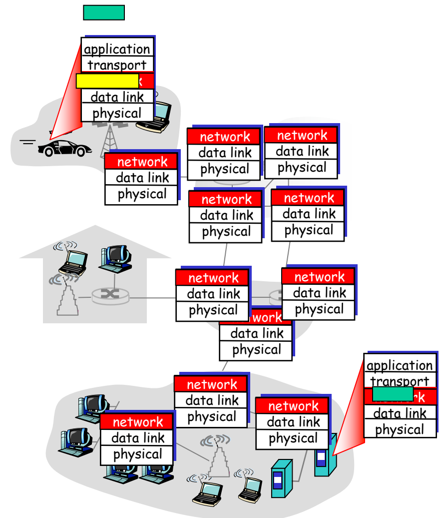
Network Layer Overview

- Network layer services
  - Desired services and tasks
  - Actual services and tasks
- Forwarding versus routing
  - Routing algorithms path selection
  - Routing algorithms creation of forwarding table
- Inside a router: switching 'fabric'
- Three Network Layer protocols
  - IP – for addressing and forwarding
  - Routing protocols – determining the best path
  - ICMP – messaging protocol

2

Network layer

- Transport a segment from sending to receiving host, but implemented in the network core
- The sending side encapsulates segments into datagrams
- The receiving side delivers segments to transport layer
- Network layer protocols run in *every* host & router
  - Router examines header fields in all IP datagrams passing through it



Network Layer Services of IP?

- Guaranteed delivery?
- Guaranteed minimum delay?
- In-order datagram delivery?
- Guaranteed minimum bandwidth to flow?
- Restrictions on changes in inter-packet spacing?

- IP Provides? → "Best-effort service"

Key Network-Layer Functions

1. *routing:* determine route taken by packets from source to destination
   - *Network-wide routing algorithms*

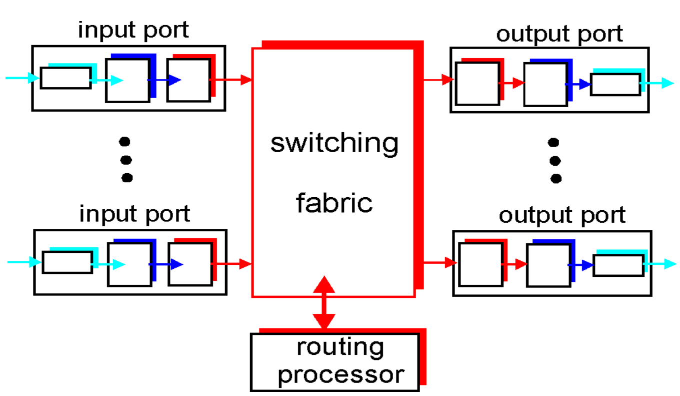
2. *forwarding:* move packets from router's input link to appropriate output link
   - *Internal to a single router*

6

Router Architecture Overview

Two key router functions:

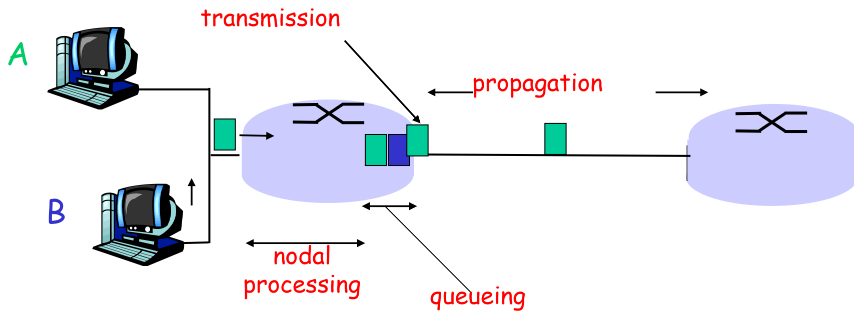
- 1. *run routing algorithms/protocol*
- 2. *forward* datagrams from incoming to outgoing link



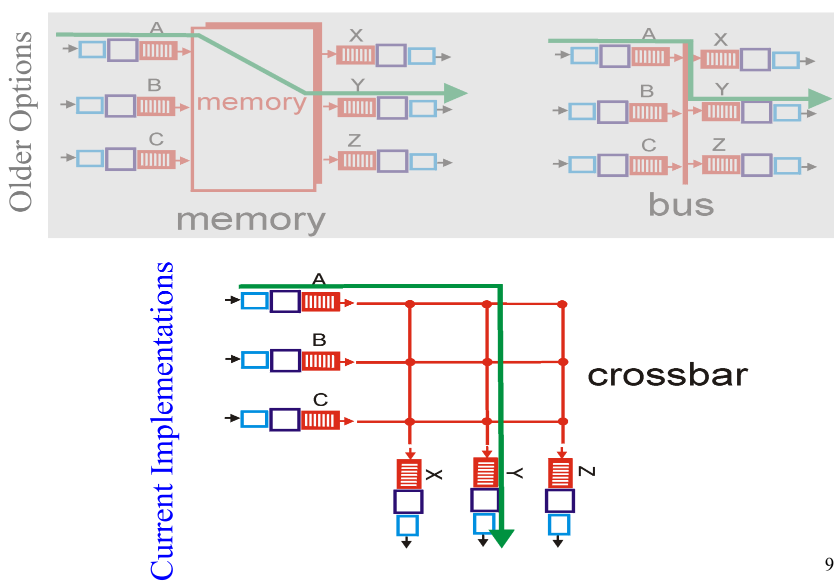
7

Four sources of packet delay

Find an analogy for each category below in the caravan example.



Three types of "switching fabric"

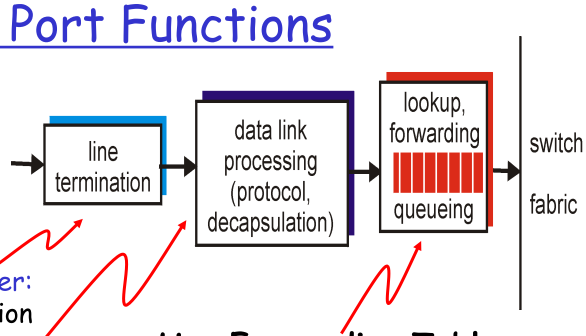


Queuing in Routers

- Where can queuing occur?
- Why does it occur?

10

Input Port Functions



Physical layer:
bit-level reception

Data link layer:
e.g., Ethernet
see chapter 5

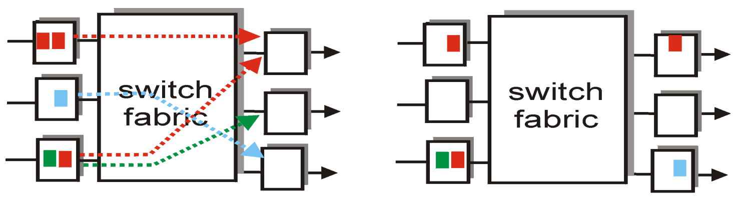
Use Forwarding Table:

- goal: complete input port processing at 'line speed'
- queuing occurs if datagrams arrive faster than forwarding rate into switch circuitry ('switching fabric')

11

Input Port Queuing

- Circuitry slower than input ports combined -> queueing may occur at input queues
- Head-of-the-Line (HOL) blocking: queued datagram at front of queue prevents others in queue from moving forward
- *queuing delay and loss due to input buffer overflow*

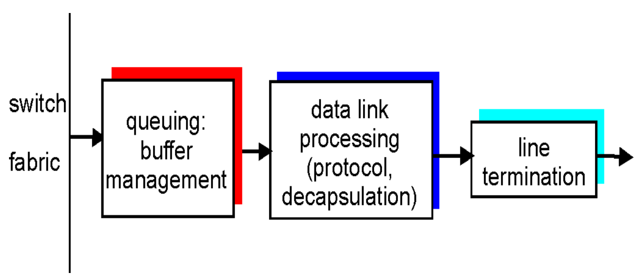


output port contention at time t - only one red packet can be transferred

green packet experiences HOL blocking

12

Output Ports



- *Buffering* required when datagrams arrive from circuitry faster than the line transmission rate
- *Scheduling discipline* chooses among queued datagrams for transmission

13

Output Port Queuing

- Packet scheduler at the output port
  - Select one queued packet for transmission
    - FCFS = "________________"?
    - Weighted-fair-queuing – share the outgoing link "fairly" among connections

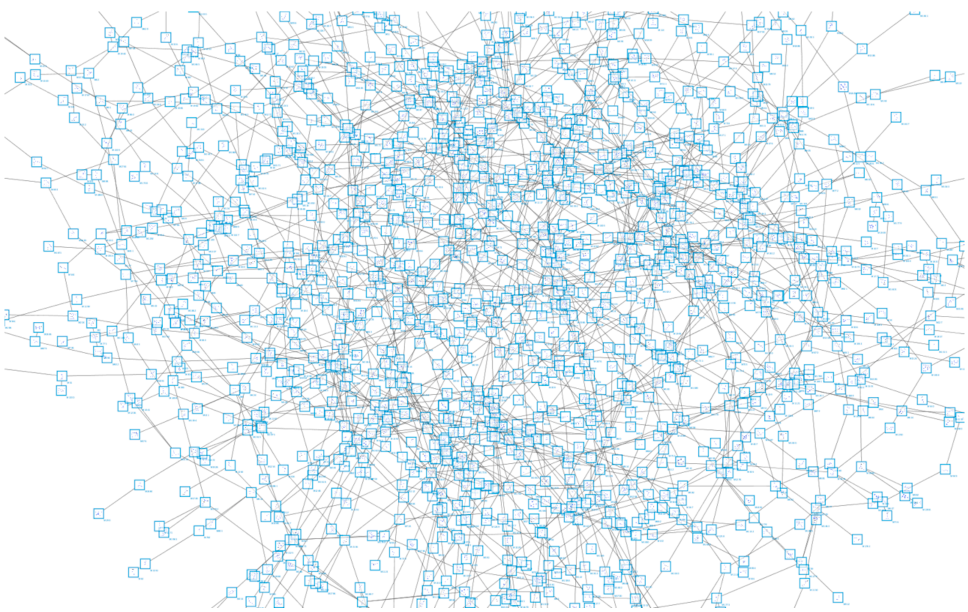
14

Discussion Questions

- Questions on handout…

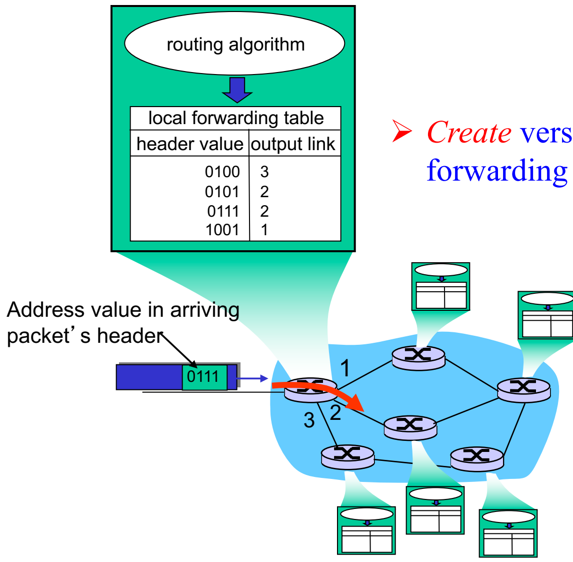
15

16



17

Interplay between routing and forwarding

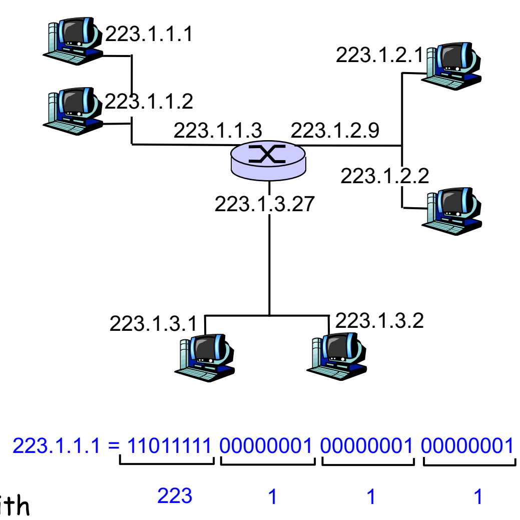


- *Create* versus *use* the forwarding table

18

IP Addressing: Overview

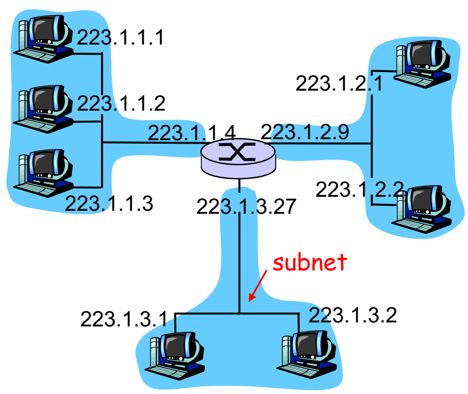
- IP address: 32-bit identifier for each *interface* on a host or router.
  - Dotted-decimal notation
- *Interface:* connection between host/router and physical link
  - routers typically have multiple interfaces
  - hosts typically have one interface
  - IP addresses associated with each interface



19

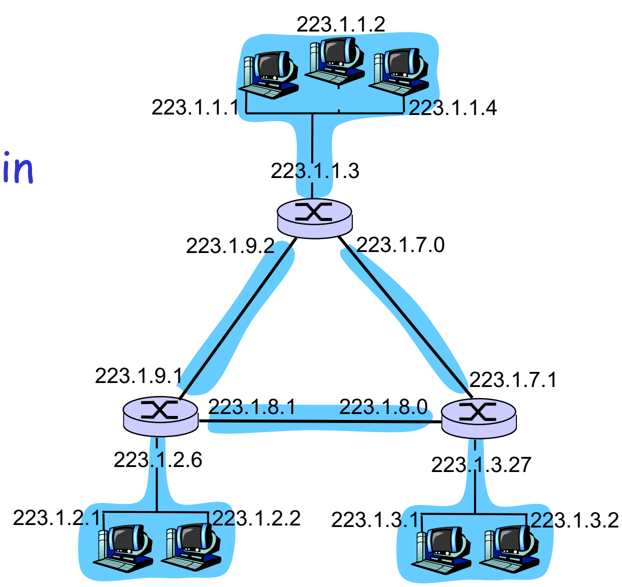
Subnets

- A subnet contains:
  - devices that can physically reach each other without an intervening router
- IP address:
  - subnet portion (high order bits)
  - host portion (low order bits)
- Subnet *mask* notation:
  - Differentiates the network versus host part of the address
  - *e.g.,* the leftmost 24 bits are for the network…
    - 223.1.3.0/24



Subnets

How many subnets are in this figure?



IP addressing: CIDR

CIDR: Classless InterDomain Routing

- Subnet portion of address of arbitrary length
- Address format: a.b.c.d/x
  - x is the number of bits in subnet portion of address
  - These 'x' most significant bits are the 'prefix'
- Addresses of all hosts in the same subnet have the same left most 'x' bits

← subnet part the 'prefix' → ← host part →

11001000 00010111 00010000 00000000

200.23.16.0/23

Forwarding table

2^32 = 4 billion possible addresses
So the table could have 4 billion entries!

Destination Address Range	Link Interface
11001000 00010111 00010000 00000000 through 11001000 00010111 00010111 11111111	0
11001000 00010111 00011000 00000000 through 11001000 00010111 00011000 11111111	1
11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111	2
otherwise	3

Longest prefix matching

Prefix Match	Link Interface
11001000 00010111 00010*** *********	0
11001000 00010111 00011000 *********	1
11001000 00010111 00011*** *********	2
otherwise	3

Examples

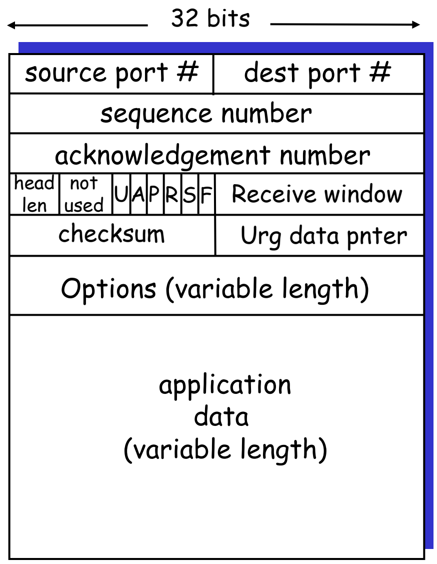
DA: 11001000 00010111 00010110 10100001 Which interface?

DA: 11001000 00010111 00011000 10101010 Which interface?

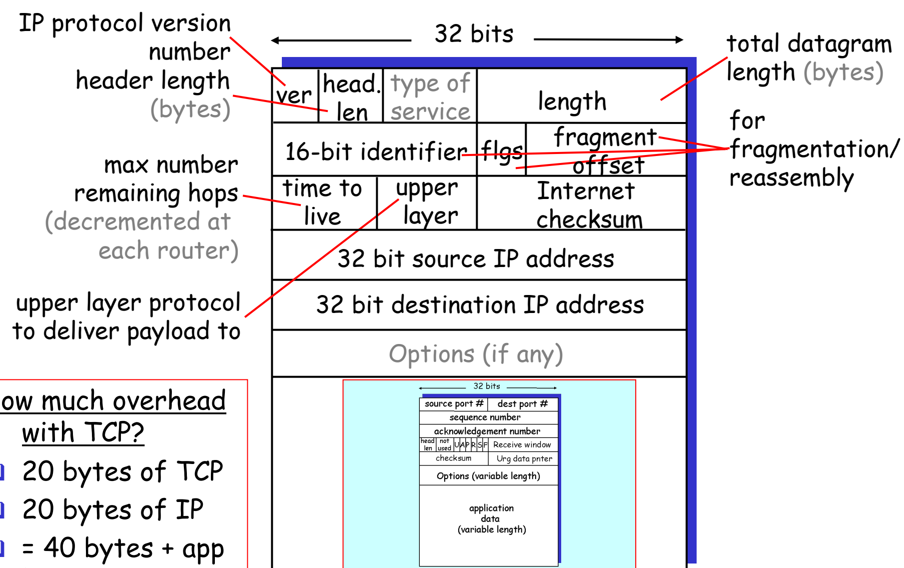
Forwarding Table "Ranges"

- What are the assumptions and implications of having large ranges of IP addresses forwarded to the same outgoing link?
- Why is CIDRized ('classless') addressing an improvement over 'classful' addressing, that restricted the network prefix to complete bytes?

TCP segment structure



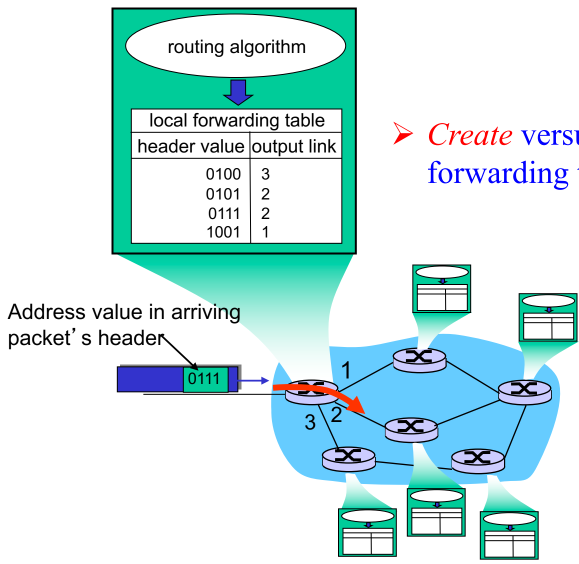
Internet Protocol: IP datagram format



how much overhead with TCP?

- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

Routing and Forwarding



- *Create* versus *use* the forwarding table

Determining the needed submask

- IPv4 address – dotted decimal notation with 4 bytes = 32 bits
  - 2^{32} = 4 billion (= 4,294,967,296)
  - 2^{8} = 256 = numbers 0 through 255

— — — —   — — — —

  - 2^{16} = 65,536 = numbers 0 through 65,535
- If you need addresses for 1000 hosts, what should you request for a subnet mask?
  - xxx.xxx.xxx.xxx/_?_
- ... For 5000 hosts?

Binary Number Sanity Check

- 2^2 = 4 = 0100
- 2^3 = 8 = 1000 (one nibble)
- 2^4 = 16 = 0001 0000
- 2^8 = 256 = 0001 0000 0000 (one byte)
- 2^{10} = 1024 = 0100 0000 0000
- 2^{11} = 2048 = 1000 0000 0000

30

Longest prefix (subnet) matching

Prefix Match	Link Interface
11001000 00010111 00010*** *********	0
11001000 00010111 00011000 *********	1
11001000 00010111 00011*** *********	2
otherwise	3

Examples

Addr: 11001000 00010111 00010110 10100001 Which interface?

Addr: 11001000 00010111 00011000 10101010 Which interface?

31

Discussion Questions

- Back to the questions on handout…

32

Smith College IP Addressing

- Smith uses a variety of masks now, but most of the campus uses 255.255.254.0 rather than the much more common 255.255.255.0.
- The reason goes back to our original subnets using the original Ethernet.
- There weren't many subnets within Smith and the network administrators thought they might need to support more than 256 hosts per subnet.

33

Smith College IP Addressing

- The Science Center is mostly different from the rest of campus, because the CATS move machines around a lot and they are responsible for assigning the IP addresses within the science buildings.
- Ford Hall has a 255.255.248.0 mask to allow for 2048 hosts in the building.
- Bass and McConnell share a subnet of the same size, as do Burton and Sabin-Reed.

34

Smith College IP Addressing

Possible **QUESTIONS**:

1) What mask would you need to support (x) hosts on a subnet?
2) Given a mask of 255.255.254.0, are the machines with IP addresses 131.229.22.50 and 131.229.23.243 on the same subnet?
3) How many hosts are supported in the range 131.229.22.00/23 ?

35

Smith College IP Addressing

- Most people really want to identify a 131.229.23.x address as being on a different subnet from a 131.229.22.y address, whether it is or not.
- **QUESTION**: When would hosts with the above masks be in the same subnet and when would they not be? (how would you specify the network mask in each case?)

36

Booting Up

- How to boot up a computer
  - How does a computer know where to start itself?
- How to enter a computer network
  - How does a computer know how to start communicating with other computers?
  - How does it get its own 'source' IP address
  - Which devices/hosts does an entering computer need to communicate with first, and how does it do this?

37

IP addresses: how to get one?

Q: How does *network* get subnet part of IP address?

A: Is allocated a portion of its provider ISP's address space, which gets that from ICANN (Internet Corp. for Assigned Names and Numbers)

Q: How does a *host* get an IP address?

- hard-coded by system administrator in a file, ***or***
- DHCP: Dynamic Host Configuration Protocol: dynamically get address from as server
  - "plug-and-play"

38

DHCP: Dynamic Host Configuration Protocol

Goal: allow host to *dynamically* obtain its IP address from network server when it joins a network

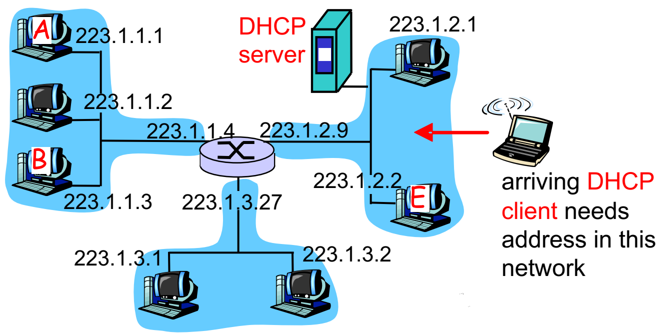
- Can renew its lease on the IP address it is using
- Allows reuse of addresses once one host leaves
- Support for mobile users to join networks

DHCP overview:

1) host broadcasts "DHCP discover" msg
2) DHCP server responds with "DHCP offer" msg
3) host requests IP address: "DHCP request" msg
4) DHCP server sends address: "DHCP ack" msg

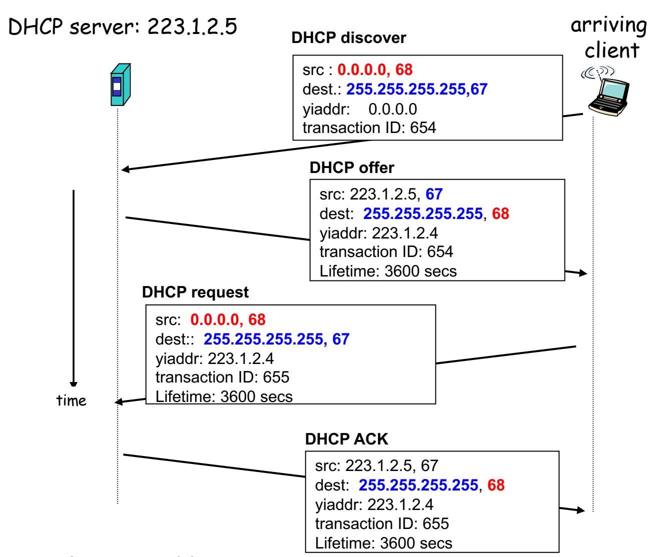
39

DHCP client-server scenario



40

DHCP client-server scenario



yiaddr = 'your internet address'
broadcast address, 255.255.255.255 → sent to every host *in the subnet*

41

Summary

- There are many possible network layer services → IP provides none
- Forwarding vs. Routing
  - Forwarding tables
- Inside a router
  - The internet in miniature
  - Switching 'fabric' (circuitry)
- The network IP datagram
- IP addressing structure

42

Midterm Topics

- Know the principles for each layer
  - Know what the actual Internet implementation is (versus what might be desired)
- Chapter 1 – mainly sources of packet delay,
  - How (and which ones) to calculate
  - Understanding of all sources of delay
- Applications
  - Architectures, mainly client-server
  - Types of connections – parallel, FTP-data and control, etc.
  - What we did with Telnet
  - Socket programming

43

Midterm Topics

- Transport Layer
  - Transport layer desired services
  - UDP and TCP actual services
  - Multiplexing/demultiplexing
  - Checksum
  - Connection management
  - Elements of reliable data transport
  - Timing diagrams with SEQ, ACK, SYN, etc
  - Congestion control – elements and algorithm
  - Flow control

44

Midterm Topics

- Network layer
  - The sources of packet delay, at the router
    - Queueing at the input and output ports, why, how, terms...
  - The definition of / difference between forwarding and routing
  - Basic structure of IPv4 addresses
  - DHCP and NAT (today's topics)

45