# Network Layer and IP

CSC 343·643

Wake Forest University
Department of Computer Science

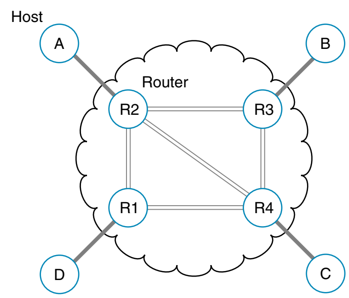
Fall 2013

## Network Layer

- Concerned with getting packets from the source to the destination
- In contrast, the data-link layer
  - Moves frames from one end of the wire to another
  - Assume everyone is **locally** connected
- Network layer deals with **end-to-end** transmission
  - Routing **packets** (or **datagrams**) from one machine to another until destination is reached

  *Token passing required forwarding a frame from one machine to another, is this routing?*

## Network Layer Issues



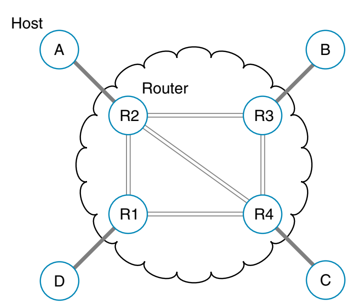

- Routing
  - Given different paths, which should be taken?
  - Should every packet take the same route?
- Congestion control
  - Prevent a *link* (router) from becoming overwhelmed
- Internetworking
  - Interconnect different networks at the network level



## Network Layer Designs

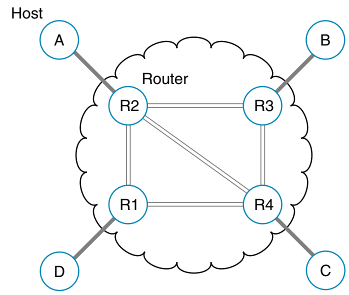
1. **Connectionless** (*Internet community argument*)
   - Network viewed as unreliable
   - Hosts perform error control, flow control, and packet ordering
   - Each packet sent independently
     - Routes taken may change over time

     *Why would a route change? Implications of multiple routes?*



2. **Connection-oriented** (*Telephone company argument*)
   - Network should be *reasonably* reliable
   - Path established before packets sent
     – Negotiate resources (QoS) at each hop

   *Any advantages to establishing a path?*

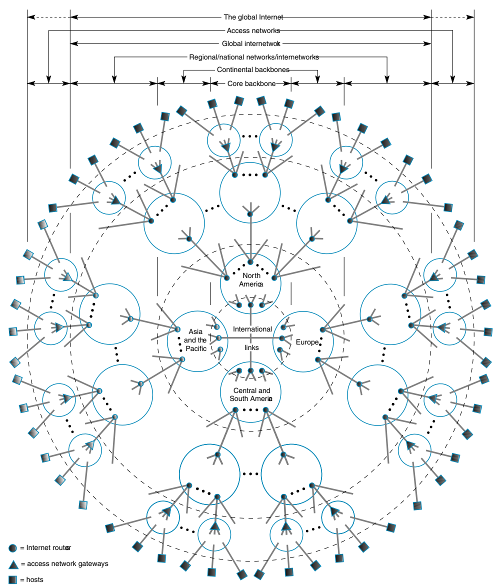



*Any disadvantages to establishing a path?*



## Network Layer in the Internet

- Internet can be viewed as a set of connected Autonomous Systems

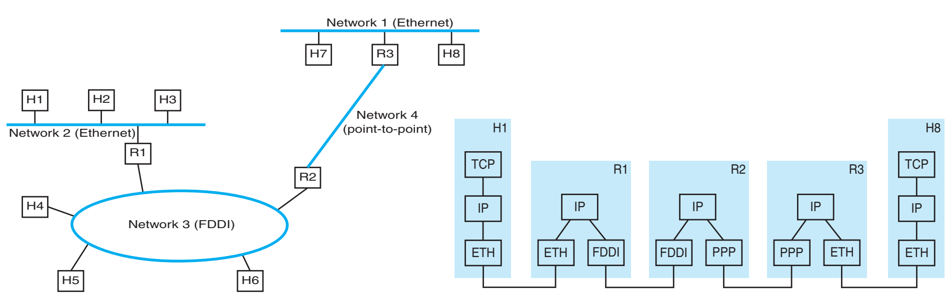


- The network layer is what allows the pieces to interconnect
- The Internet Protocol (IP) provides
  1. **Best Effort** (BE) transport of datagrams
     - Unreliable service
     - Packets may arrive out of order, if at all...
     - No Quality of Service (QoS) guarantees provided
  2. Routing from source to destination
     - Can route to different AS
     - Routes can change based on network conditions

  *Is IP connectionless or connection-oriented?*



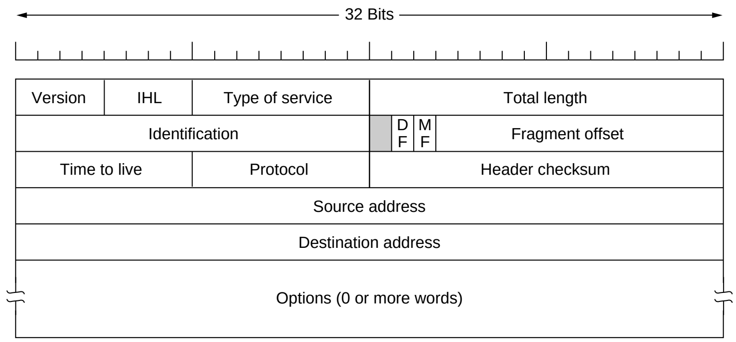
## Internet Operation Overview




- Network layer takes data streams and breaks into datagrams
  - Datagram can be up to 64KB each, average is 1500 bytes
- Each datagram is transmitted through the Internet
  - Possibly fragmented
- Pieces arrive at destination, reassembled into original datagram
- Datagram is passed to the transport layer

## IP Protocol Datagrams

- Datagram (packet) consists of a header part and data part
- Header consists of: 20 byte *fixed part* and an *optional part*




- Big endian order (left $\rightarrow$ right) also called **network byte order**
  - SPARC is big endian, while Pentium is little endian



## IP Header: Version and IHL Fields

- Version field (4 bits)
  - Identifies the version of IP (e.g. IPv4 or IPv6)
- Internet Header Length (IHL, 4 bits)
  - Total length of the IP header, in measured 32-bit words
  - Minimum value is 5 (no options are present)
  - Maximum value is 15, which is a ______ byte header
  - This will limit the usefulness of some options

## IP Header: ToS Field

Type of Service (ToS) is 8 bits

- Indicates the type of service expected, has sub-fields
  1. First three bits are the precedence (priority) sub-field
     - Range from 0 (normal) to 7 (control packet)
     - *"which is ignored today"* - Stevens
  2. Next four bits request different types of service

| Application | Min Delay | Max Throughput | Max Reliability | Min Cost | Hex Value |
|---|---|---|---|---|---|
| Telnet | 1 | 0 | 0 | 0 | 0x10 |
| FTP data | 0 | 1 | 0 | 0 | 0x08 |
| SNMP | 0 | 0 | 1 | 0 | 0x04 |

  3. One unused bit

- ToS feature is not supported by most IP implementations



## IP Header: Total Length and Fragmentation Fields

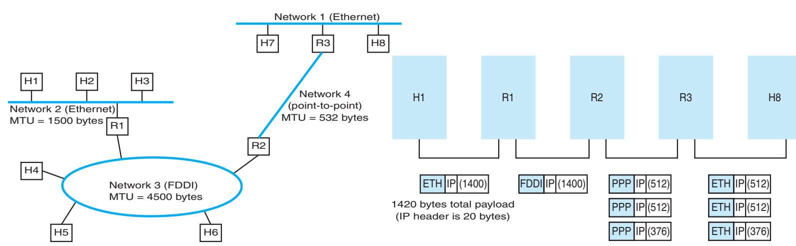
- Total length field (16 bits)
  - Datagram length (header and data), measured in bytes

  *What is the maximum size of an IP datagram?*

- Identification field (16 bits)
  - Identifies which datagram the fragment belongs to
  - One number for all the fragments of a packet
- DF (Don't Fragment) 1 bit, if set then don't fragment
- MF (More Fragments) 1 bit, set if not last fragment
- Fragment offset (13 bits)
  - Where in the current datagram this fragment belongs
  - Fragments must be a multiple of 8 bytes (except for last one)

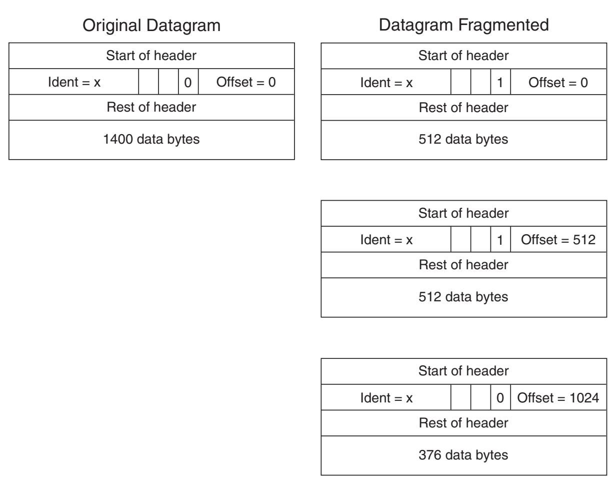
# IP Fragmentation and Reassembly

- Different network technologies have different packet sizes
  - Every network has a **Maximum Transmission Unit** (MTU)
  - If the datagram is larger than the MTU, then it is fragmented
- *"Every internet module must be able to forward a datagram of 68 octets without further fragmentation... Every internet destination must be able to receive a datagram of 576 octets either in one piece or in fragments to be reassembled."* - [RFC791]






- Assume R2 has a MTU (data) of 532 bytes (allows a 20 byte header and 512 bytes of data)
- The original 1420 byte datagram fragmented into 3 pieces at R2




- RFC 1191 gives some example MTU sizes, based on the link layer

## IP Header: TTL and Protocol Fields

- Time To Live (TTL, 8 bits)
  - Counter to limit packet lifetime
  - Maximum lifetime of packet (in seconds)
    *What is the maximum maximum lifetime?*
  - Time spent at every router is subtracted
  - Actually decremented once per hop
  - Once zero is reached, a control packet is sent back
    *What problem does TTL attempt to prevent?*
- Protocol field (8 bits)
  - Which transport process the packet belongs to (e.g. TCP or UDP) numbers are global defined in [RFC 1700]



## IP Header: Checksum and Address Fields

- Header checksum (16 bits)
  - Verifies only the header
  - Add all 16 bit words (one's complement) then take the one's complement of the sum
    *A new checksum is computed and stored in the header at every hop, why?*

    *What happens if an error is detected?*

    *How is the data verified? Do we care at this layer?*
- Source and destination addresses
  - 32 bits each, more later...

## IP Header: Options Field

- Allow subsequent versions of IP to include *new features*
- Option begins with a one byte idenfication code, 5 are defined
  1. Security - Security and handling restrictions [RFC1108]
     *If set, helps a sniffer identify the more interesting datagrams*
  2. Strict source routing - Gives path to follow (*security issue?*)
  3. Loose source routing - List routers not to be missed
  4. Record route - Make every router append IP address
     *Why is this no longer useful?*
  5. Timestamp - Make every router append address and time
- Options field padded out to end on 32 bit boundary
- *"these options are rarely used and not all hosts and routers support all the options"* - Stevens



## IPv4 Addresses

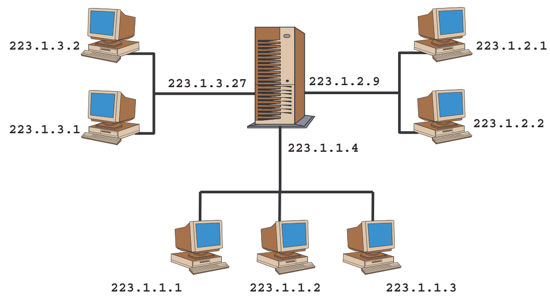
Every host or router (actually interface) has a unique IP address

- IP addresses are 32 bits long (IP version 4) and are used in the source and destination fields of the IP datagram
- *Dotted-decimal notation* is used to represent each address, each byte is represented via a decimal number
  - `193.32.216.9` ⇒ [`11000001 00100000 11011000 00001001`]

  *The data link layer also has an address, what is the difference?*
  *Why is a network address needed?*
- Addresses are hierarchical and encode two numbers, **network** and **host**

## IP Network Example

Consider one router and seven hosts (*one address per interface*)

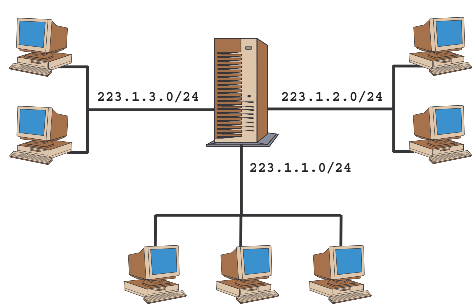



- Three hosts at bottom have similar addresses, 223.1.1.$x$
  - The leftmost 24 bits they share is the **network** portion
  - Remaining 8 bits is the **host** portion

    *How many hosts can connect to the 223.1.1.$x$ network?*



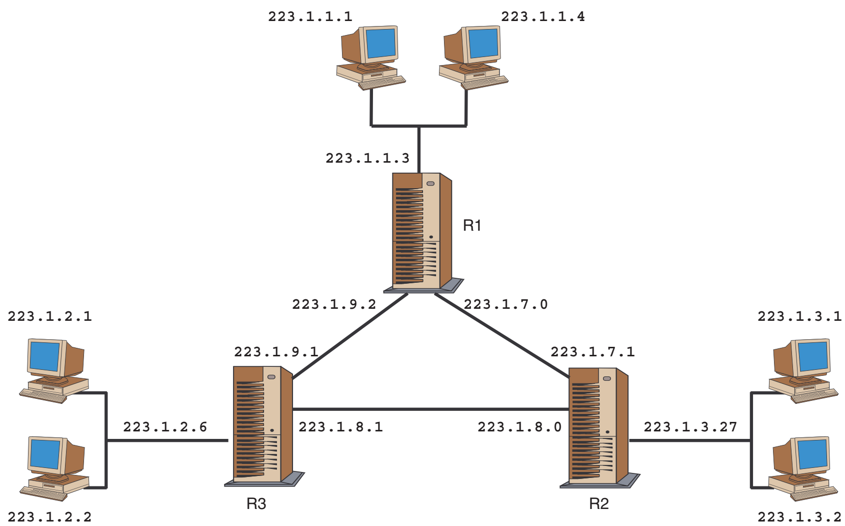
- Hosts of 223.1.1.$x$ form a network, interconnected via a LAN
  - The network address is 223.1.1.0/24
  - The /24 is also called the **network mask** or **network prefix**
    - Indicates the 24 leftmost bits define the network address
  - Any additional host that would attach to this network must have a unique address of the form 223.1.1.$x$
- The remaining networks have a similar structure



## Multiple IP Networks

IP definition of a *network* is not restricted to Ethernet segments

- Consider three routers interconnected via point-to-point links



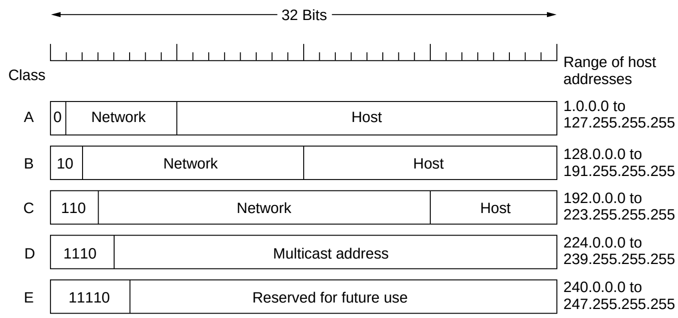



- Each router has three interfaces
  - One for each point-to-point link
  - One for the broadcast link to the hosts
- What are the *networks* in the diagram
  - Three networks interconnecting hosts,
    223.1.1.0/24, 223.1.2.0/24, and 223.1.3.0/24
  - Three additional networks that interconnect routers
    * 223.1.7.0/24 connects R1 $\Leftrightarrow$ R2
    * 223.1.8.0/24 connects R2 $\Leftrightarrow$ R3
    * 223.1.9.0/24 connects R3 $\Leftrightarrow$ R1
- How do we determine what is a network
  - Detach each interface from host or router
  - Resulting *islands* are the networks

## IPv4 Address Classes

The original Internet architecture defined 5 different IP address classes

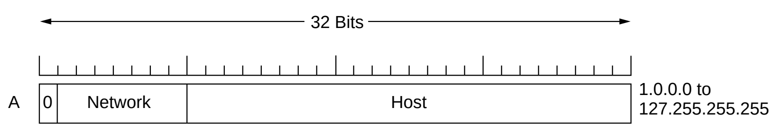
- This is also know as **classful addressing**
- Classes differ on how bits are divided (network versus host)



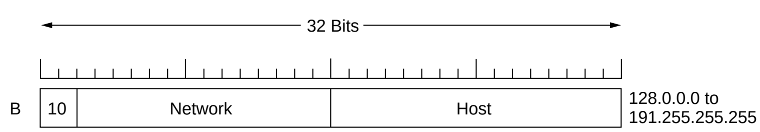

- This creates 3 different classes of networks (A, B, and C)



- For example, consider class A addresses

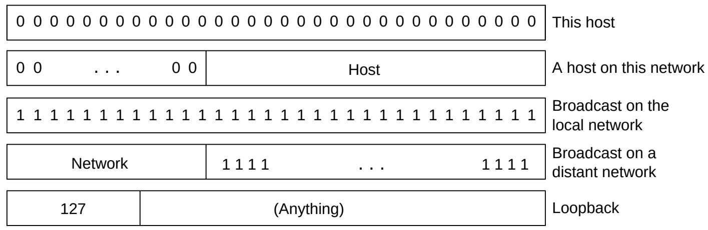



  - First bit is zero, identifies class A
  - Next 7 bits identify the network
  - Last 24 bits identify the host (interface) in the class A network
- In comparison, class B has




*What class is 223.1.1.0/24? As a company, would you prefer an A, B, or C address?*

# IP Addresses with Special Meanings




- `0.0.0.0` only used by a host when booting
- All zeroes for the network number, refers to the local network
  - If `223.1.1.0/24` is the network and I am `223.1.1.52`, locally I can be reached using `0.0.0.52`
- Address of all ones is the broadcast address for the local network

  *What is the dotted-decimal address?*



- Address with the proper network number, and all ones for the host number allows host to broadcast to a different network
  - If `223.1.1.0/24` is a distant network, then `223.1.1.255` broadcasts to all hosts at the network

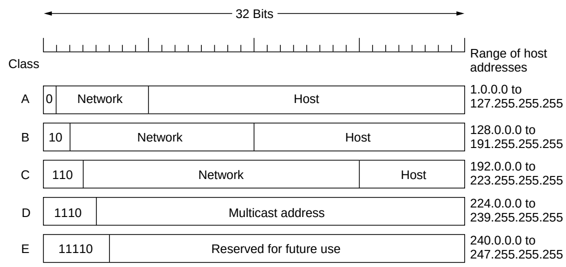
  *This and ping can can be used as a network attack, how?*

- 127.$x.y.z$ is reserved for loop-back testing
  - Packet is never placed on the network, processed locally

  *Given a class A address, how many hosts can be connected to the network?*

# IP Addresses and Routing

- We have introduced IP addresses and the concept of a network
  - IP addresses are 32 bits long, and can be divided into classes
  - Each class divides address into network and host portion




  - All hosts in one network have the same network portion, different host portion; therefore, the addresses are hierarchical

*Why is it important to identify the class of an address?*



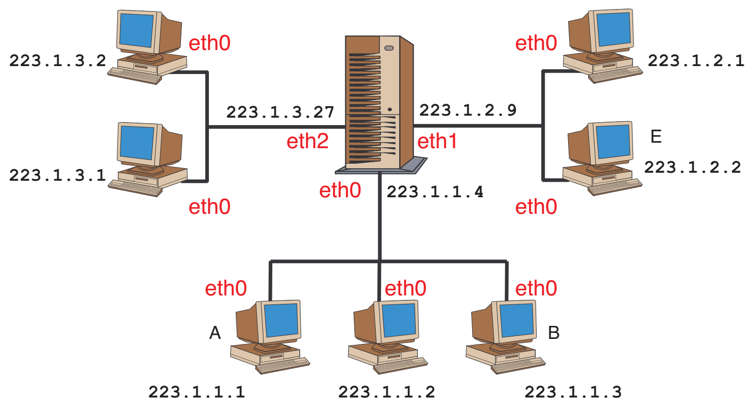
# Routing Tables

How does a source host send a datagram to a destination host?

- The IP layer maintains a **routing table** in memory
  - Remember, routing tables are *next hop* oriented
  - Multiple hop paths are not recorded
- Each entry in the routing table has the following information[a]
  1. Destination address, either *host* or *network* address
  2. IP address of the *next-hop router*
  3. Flags specifying if next hop is host or network
  4. Identification of the interface the datagram should be passed to (e.g. multiple Ethernet cards attached)

[a]Abbreviated list of items, more later.

## Example Routing Tables




- In the diagram, each interface (Ethernet card) is labeled (in red)
- For example, the router has 3 interfaces (eth0, eth1, and eth2)
  - Each interface must be uniquely identified, since it attaches a unique network



- An abbreviated routing table for host A would be

| Routing Table for A | | |
|---|---|---|
| **Destination** | **Next Hop** | **Interface** |
| 223.1.1.0/24 | | eth0 |
| 223.1.2.0/24 | 223.1.1.4 | eth0 |
| 223.1.3.0/24 | 223.1.1.4 | eth0 |

  - First entry indicates `223.1.1.0/24` is the local network
  - The second and third entries indicate datagrams for destinations on network `223.1.2.0/24` or `223.1.3.0/24` must be sent to `223.1.1.4`
  - eth0 is the Ethernet interface (only one card on A)

*Each network is represented with one entry, how many would be required if each host had a separate entry?*

- An abbreviated routing table for the router would be

| Routing Table for Router | | |
|---|---|---|
| **Destination** | **Next Hop** | **Interface** |
| `223.1.1.0/24` | | eth0 |
| `223.1.2.0/24` | | eth1 |
| `223.1.3.0/24` | | eth2 |

  - First entry indicates `223.1.1.0/24` is local on eth0
  - Second entry indicates `223.1.2.0/24` is local on eth1
  - Third entry indicates `223.1.3.0/24` is local on eth2

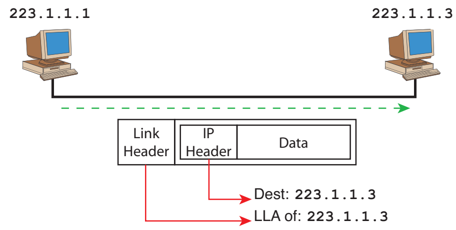


## IP Routing Steps

- IP routing performs the following actions
  1. Search routing table for complete destination address, if found send packet to the next-hop entry
  2. Search routing table for an entry that matches the destination network number, if found send packet to the next-hop entry
     - Must take into account possible subnet mask
  3. Search for *default* entry, if found send to next-hop router
- IP search order is, host address $\rightarrow$ host network $\rightarrow$ default
- If all the steps fail, then the datagram is not deliverable

## Routing Example: A $\rightarrow$ B

Assume A (223.1.1.1) sends datagram to B (223.1.1.3)




- There is no host entry for 223.1.1.3
- There is a network entry for 223.1.1.0/24
- A link layer frame (containing the datagram) is created and addressed to the link layer address of 223.1.1.3

  *We are at layer 3, how do we get a layer 2 address?*

- Ethernet frame is sent and received by host B



## Routing Example: A $\rightarrow$ E

Assume A (223.1.1.1) sends datagram to E (223.1.2.2)

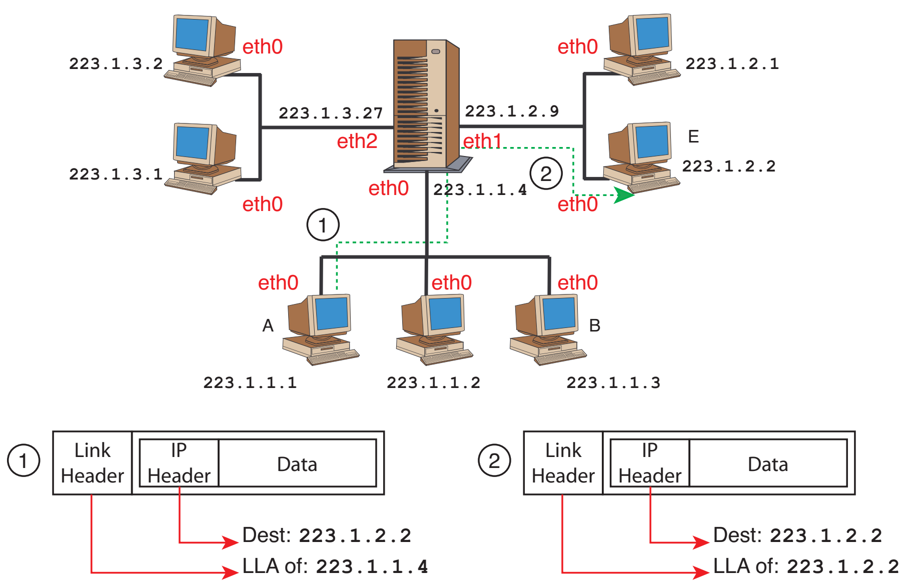
| Routing Table for A | | |
|---|---|---|
| **Destination** | **Next Hop** | **Interface** |
| 223.1.1.0/24 | | eth0 |
| 223.1.2.0/24 | 223.1.1.4 | eth0 |
| 223.1.3.0/24 | 223.1.1.4 | eth0 |

| Routing Table for Router | | |
|---|---|---|
| **Destination** | **Next Hop** | **Interface** |
| 223.1.1.0/24 | | eth0 |
| 223.1.2.0/24 | | eth1 |
| 223.1.3.0/24 | | eth2 |

- Host A finds entry for 223.1.2.0/24 network
  - Requires sending packet to 223.1.1.4
- Host A creates and sends link-layer frame (containing datagram) addressed to the link-layer address of 223.1.1.4
  - Therefore, the next-hop entry is used for the link-layer address
  - IP destination address remains unchanged

- Router 223.1.1.4 receives frame and removes datagram
  - Destination address is 223.1.2.2
  - Router is allowed to forward datagrams
- Router finds entry for 223.1.2.0/24 network
  - This is directly connected via eth1
  - Datagram will be forwarded
- Router creates and sends link-layer frame (containing datagram) addressed to the link-layer address of 223.1.2.2 on eth1
- Frame received by host E, datagram removed and processed
- N.B. operation of host and router are equivalent, except routers are allowed to forward datagrams





# Another Routing Example

Assume 140.1.1.1 sends a datagram to 152.24.25.5