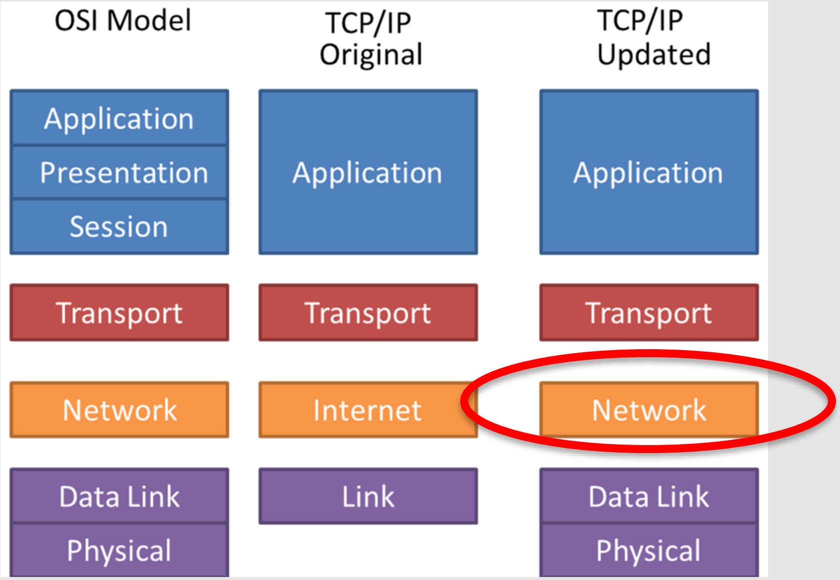
The Network Layer

Forwarding, Routing, Addressing

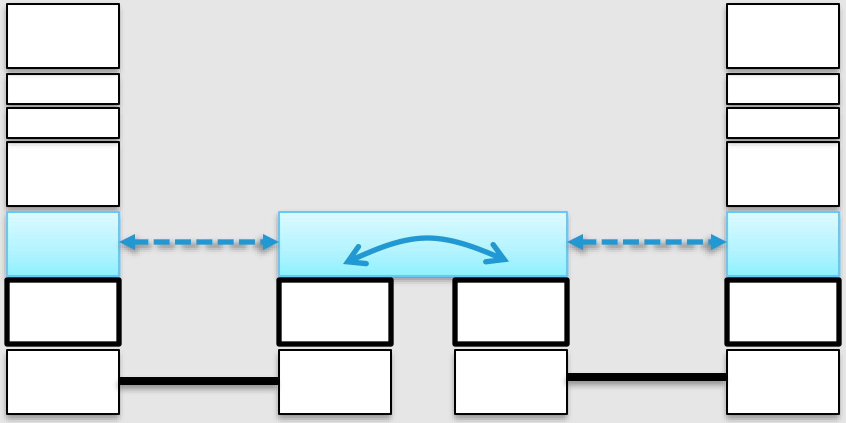
Network Protocols

- Rules of operation of a network



Network Layer

- DLC allows communication between connected computers
- Cooperation can create connectivity not present in physical topology – networking
- Concerns
  - Addressing
  - Forwarding
  - Routing



Concepts - Addressing

- **Unique Identification**: identify devices on a network to ensure data is delivered to the correct destination.
- **Two Versions**: There are two main versions of IP addressing: IPv4 (32-bit addresses) and IPv6 (128-bit addresses).
- **Hierarchical Structure**: IP addresses are hierarchical, consisting of network and host portions to facilitate routing across diverse networks.
- **Dynamic vs. Static**: IP addresses can be assigned dynamically via DHCP or manually configured as static addresses.
- **Network Address Translation (NAT)**: NAT allows private IP addresses to be mapped to public IP addresses, enabling multiple devices to share a single public address for internet communication.

Concepts - Forwarding / Routing

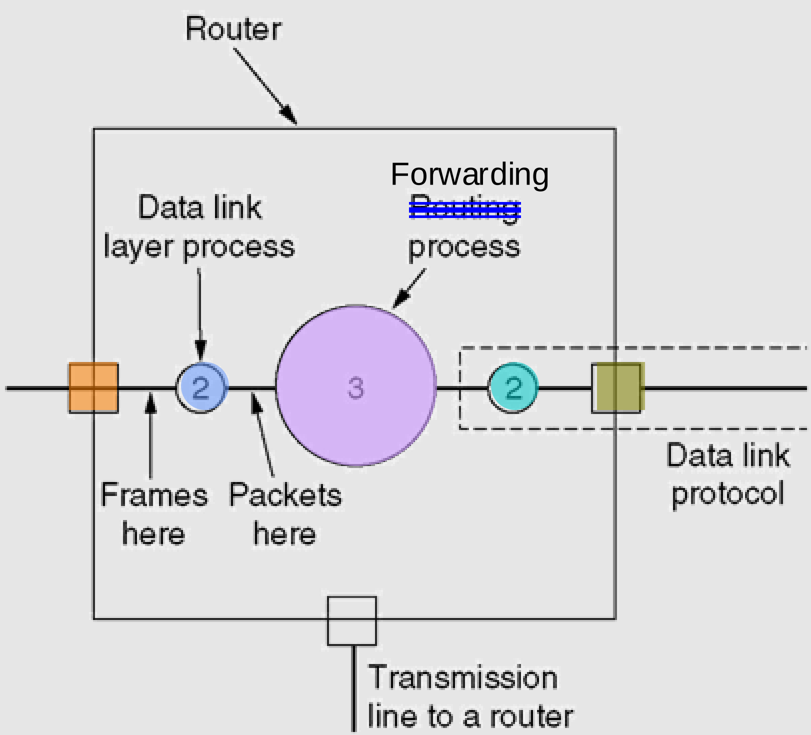
- Forwarding
  - The act of forwarding PDUs (datagrams or packets)
- Where to forward?
  - partially determined by data source/destination
    - Must be embedded in data (header)
  - partially determined by network factors
    - Must be stored at node (topology)
  - Forwarding will combine these two types of information
- Routing
  - Path-finding
    - For any and all paths, determination of best path it takes from source to destination
  - At a node, must translate into:
    - Distilling "network factors" above into easy-to-refer information for forwarding - mapping from data to next-hop

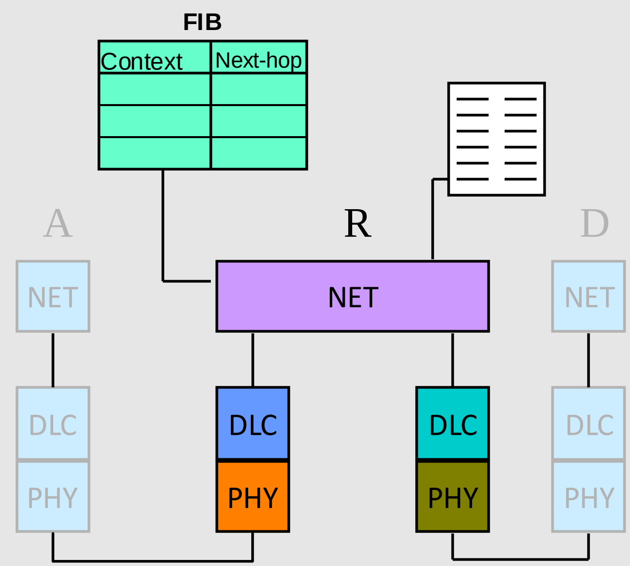
Forwarding & Addressing

Forwarding

Concepts - Forwarding Information Base

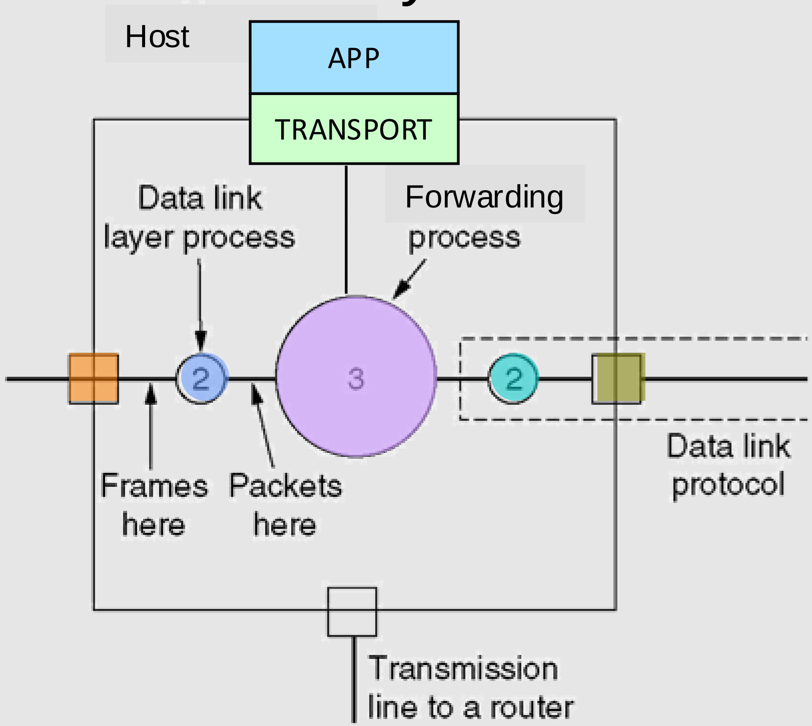
- A “lookup table” for next-hop router
  - Packet arrives with embedded context (full or partial)
  - Context information is extracted and looked up in FIB
  - Packet forwarded to next-hop intermediate node

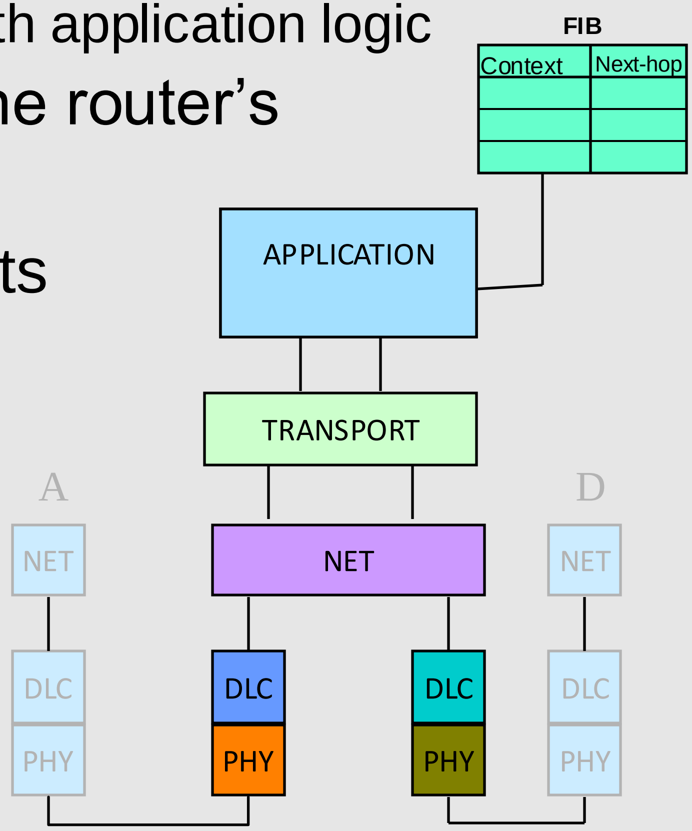




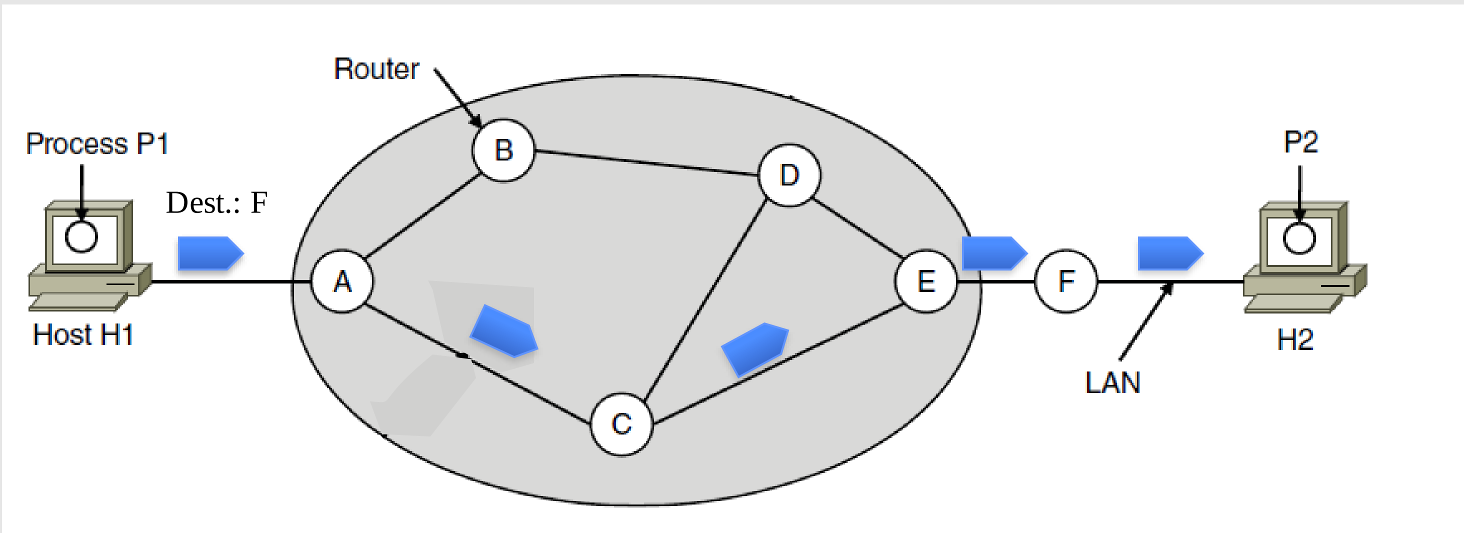
Overlay Networking

- Operation of L3 can be performed by L7
  - Overlay forwarding – integrate with application logic
- Without requiring changes to the router's firmware or physical system
- Better for dynamic environments





Connection-less Forwarding



A's FIB

Dest.	Line
A	–
B	B
C	C
D	B
E	C
F	C

C's FIB

A	A
B	A
C	–
D	E
E	E
F	E

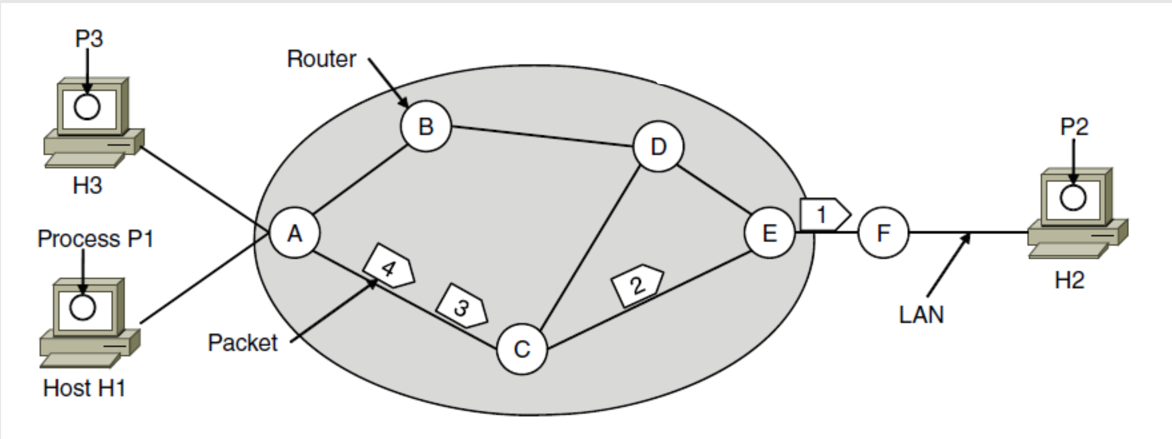
E's FIB

A	C
B	D
C	C
D	D
E	–
F	F

- H1 sends packet 1 to A
- A consults FIB → forward to C
- C consults FIB → forward to E
- E consults FIB → forward to F
- F uses LAN to forward to H2

- If FIB changes, a following packet may traverse different path

Connection-oriented Forwarding



- H1 sends request to A
- A assigns label "1", forwards request to C
- C assigns label "6", forwards request to E
- E assigns label "3", forwards request to F
- F accepts request, replies to E with label "11"
- E notes label, replies to C with assigned label
- C notes label, replies to A with assigned label
- A notes label, replies to H1 with assigned label
- H1 sends packets with label "1" to A on "virtual circuit"

A's FIB

A	⊠
B	B
C	C
D	B
E	C
F	C

C's FIB

A	A
B	A
C	⊠
D	E
E	E
F	E

E's FIB

A	C
B	D
C	C
D	D
E	⊠
F	F

A:

In		Out	
H1	1	C	6
H3	1	C	2

C:

In		Out	
A	6	E	3
A	2	E	2

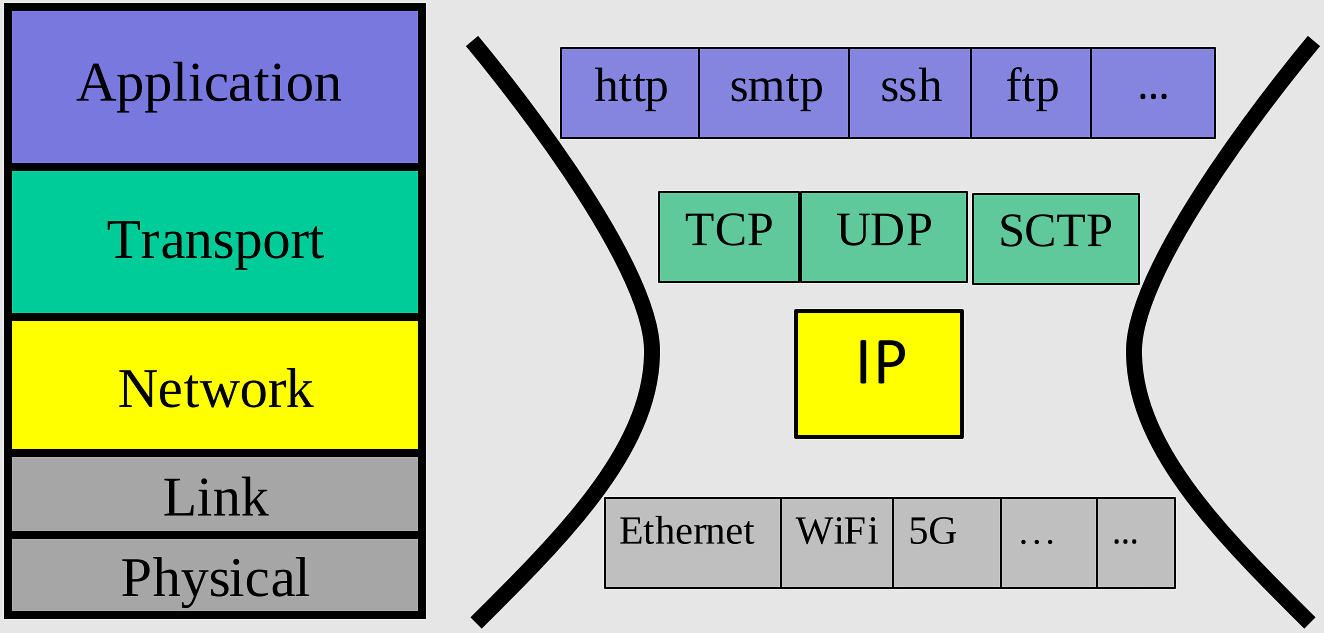
E:

In		Out	
C	3	F	11
C	2	F	2

Comparison

Issue	Datagram subnet	Virtual-circuit subnet
Circuit setup	Not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short VC number
State information	Routers do not hold state information about connections	Each VC requires router table space per connection
Routing	Each packet is routed independently	Route chosen when VC is set up; all packets follow it
Effect of router failures	None, except for packets lost during the crash	All VCs that passed through the failed router are terminated
Quality of service	Difficult	Easy if enough resources can be allocated in advance for each VC
Congestion control	Difficult	Easy if enough resources can be allocated in advance for each VC

IP is the “thin waist”



IP Header

- Various fields have different uses
  - Not all are used by core IP

Total 32 bits

includes header and payload

<table>
<tr><td>Version (4)</td><td>Hdr Len (4)</td><td colspan="2">TOS (8)</td><td colspan="4">Total Length in bytes (16)</td></tr>
<tr><td colspan="4">Identification (16 bits)</td><td>Flags (3)</td><td colspan="3">Fragment Offset (13)</td></tr>
<tr><td colspan="2">Time to Live (8)</td><td colspan="2">Protocol (8)</td><td colspan="4">Header Checksum (16)</td></tr>
<tr><td colspan="8">Source IP Address</td></tr>
<tr><td colspan="8">Destination IP Address</td></tr>
<tr><td colspan="6">Options (if any)</td><td colspan="2">PAD</td></tr>
<tr><td colspan="8">Data Field</td></tr>
</table>

Forwarding & Addressing

Addressing

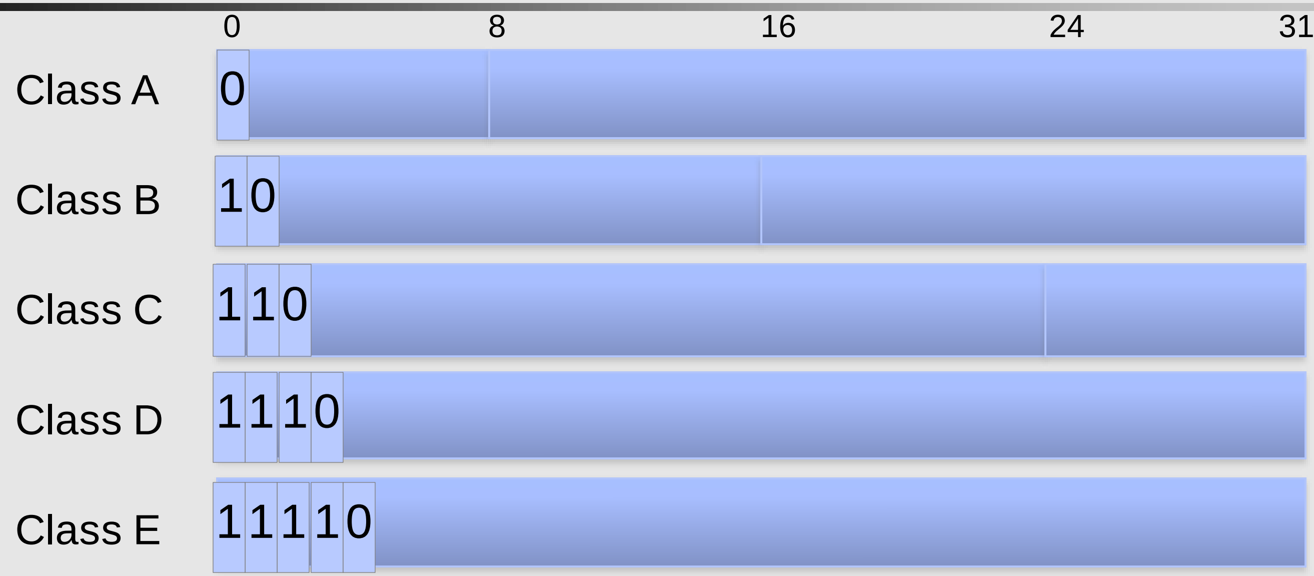
Design Principles (Names and Addresses)

1. No hardcoded addresses
2. Single naming structure
3. Addresses must be unambiguous
4. Upper-layer protocols must be able to identify end-points unambiguously
5. Standardize on compact, binary addresses that make computations (e.g., selection of a route) efficient
6. Name should be case-insensitive ASCII

Internet Addresses

- 32-bit addresses
- Each Internet host has one globally unique IP addresses that is used in all communication with that host. But…
  - the IP address may change
  - may have several addresses (routers ↔ multi-homed hosts)
  - an IP address specifies an interface (network connection), not a host!
- Assignment: Two-level hierarchy
  - 2-level addressing: (network, host)
- Class A, B, or C (initially) → Classless interdomain routing (CIDR) (later)
  - Slice into portions

IP Address Classes (How to Allocate?)



Class	First Byte	# Networks	Hosts per Network	Comments
A	< 128	2^7 (128)	2^{24} –2 (16M)	Mostly used
B	128…191	2^{14} (16384)	2^{16} – 2 (65534)	Mostly used
C	192…223	2^{21} (2M)	2^8 – 2 (254)	
D	224…239	2^{28} (268M)	-	Dynamic, multicast
E	240…255	2^{27} (134M)	-	reserved

Dotted decimal notation

Host Name	myhost.csc.ncsu.edu
32-bit address	1001 1000 0000 0001 0011 0110 0011 0000
Dotted decimal	152.1.54.48

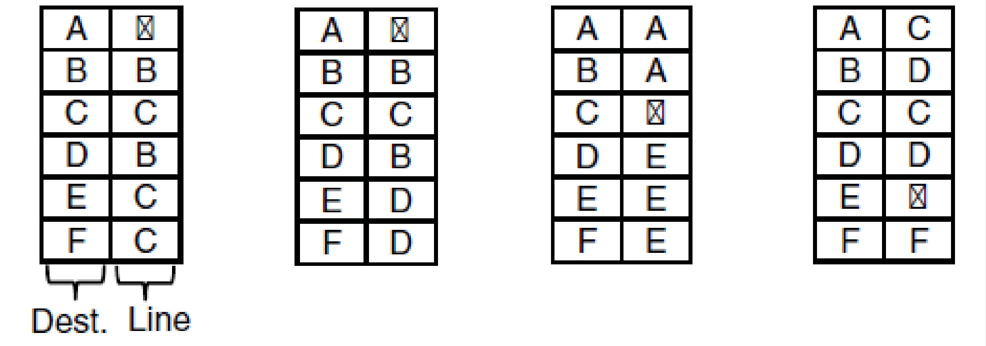
- Dotted decimal is just a different representation of the 32-bit form - 1-to-1 mapping
- Mapping Host Name to IP Address(es) (Domain Naming System)

Forwarding & Addressing

Back to Forwarding

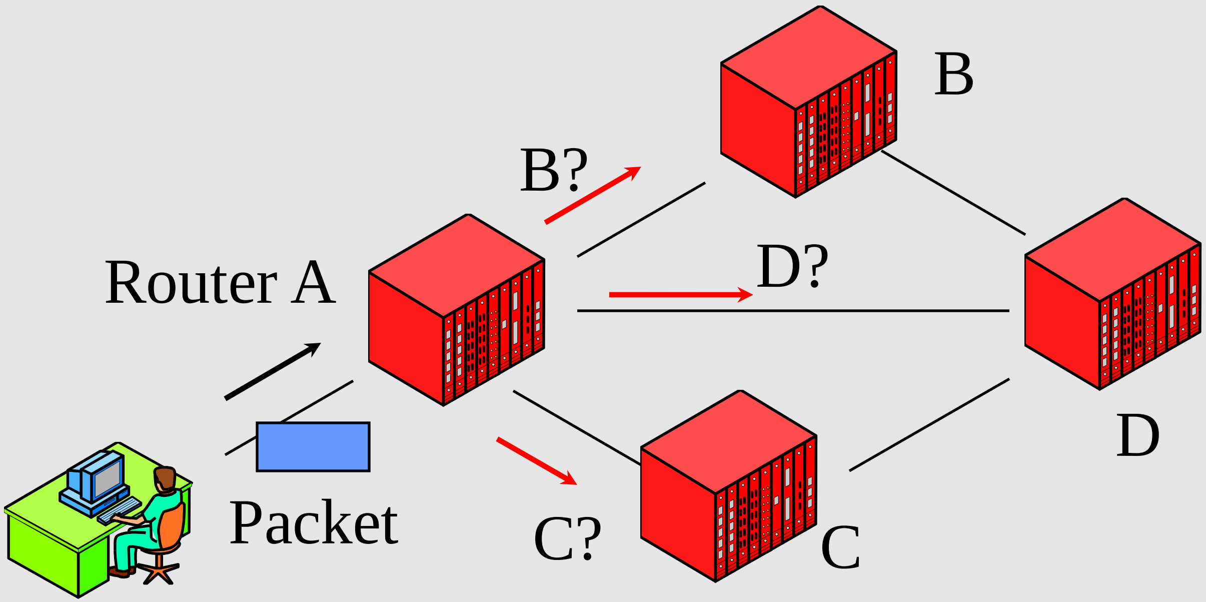
Contextless One-hop Forwarding

- Layer 3 functionality
- Router-to-router, or between router and host
- *Contextless* or *connectionless* paradigm
  - Each packet forwarded independently
  - Low overhead
- *Unreliable*
  - Possible error, loss, reordering, duplication
  - Not necessarily bad



Multihop - IP Forwarding

- Routers look at destination IP address of packet to make decisions
- What to with this packet, based upon its IP destination address *only*



Router Forwarding Tables

- Major duty
  - Provide “destination IP address” → “which router to forward to” mapping
- Index into table by network part of destination IP
- Care about Network rather than Host

Network/Subnet	Delivery
128.171.17.	Local
142.99. (class B)	Next-Hop Router A

prefix

Network Size Problems

- (net_id, host_id) pair, three classes, but
  - Class A (16M hosts) too large for most sites
  - Class C (254 hosts) too small
  - Class B (65,534 hosts) is OK
  - the efficiency problem
- Limited address space
  - Class thresholds, e.g., what if a class C net grows beyond 254 hosts?
- Organizations requested Class B addresses to avoid outgrowing the 8-bit host field of Class C:
  - only 65,534 Class B addresses → not enough
  - more than half of all Class B networks have fewer than 50 hosts
    - inefficient use

Towards Solutions

- Solutions must be backward compatible
- Problem 1: Large number of networks → routing table size
  - IP aggregation: same IP prefix must be shared by multiple physical networks
  - Sharing a large IP network to serve several physical networks
  - "*Subnetting*"
- Problem 2: Exhaustion of Class B addresses
  - temporary fix: use multiple class C addresses instead
  - Using several IP networks to serve one large physical network
  - "*Supernetting*"

- Generalization
  - De-couple network size from IP itself
  - "CIDR"

Subnet Addressing

- Standardized in RFCs 950, 1122
  - Hosts *must* support
- Relates to class B addresses mostly, although it can also be used with class A and C addresses
- Large organizations may have multiple LANs, sharing a single IP network address
  - System administrator may (a) decide to subnet, and (b) how many bits to allocate to subnet ID and host IDs

Class B:

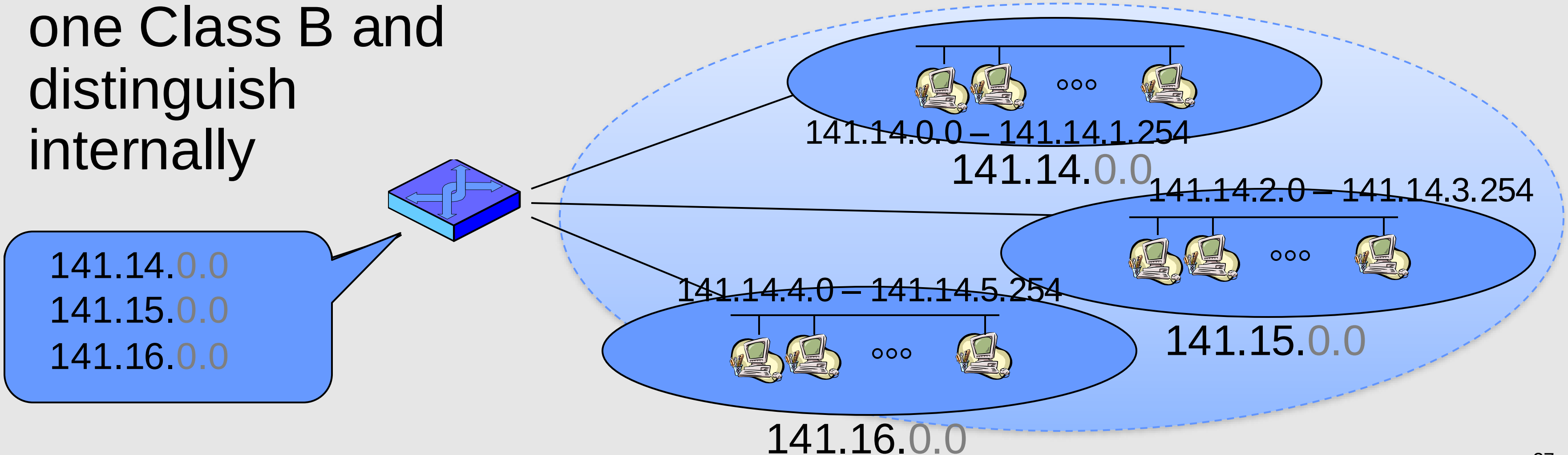
0			16	24 … 31
1	0	Net ID	Subnet ID	Host ID

Subnet Addressing (cont'd)

- Result: *three-level* hierarchical addressing
  - Enables more hierarchical routing
  - Hides internal network structure from outside world
- Does not increase the size of routing tables
  - External routers need not separately represent the different subnet routers, just the single router (network) that the IP class represents
- Helps with address space exhaustion, but
  - Only for big, responsible organizations

Subnetting Example

- Consider an organization with three physical networks, each too big for Class C
  - But collectively smaller than a Class B
- Organization must
  - Purchase three Class B address spaces
  - Or connect all three into one physical network (may be impossible)
- With subnetting, may purchase one Class B and distinguish internally



Subnet Masks

- *Masking*: the process of extracting the address of the physical network from an IP address
- A *mask* is a 32-bit number. Bits are set to…
  - 1, indicating the corresponding bit of the IP address is part of the network address (net ID or subnet ID)
  - 0, indicating the corresponding bit of the IP address is part of the host ID
- Based on first bits of the IP address, you can always figure out how long (how many bits) the network ID is
- With a subnet mask, you can *also* figure out what the subnet ID is

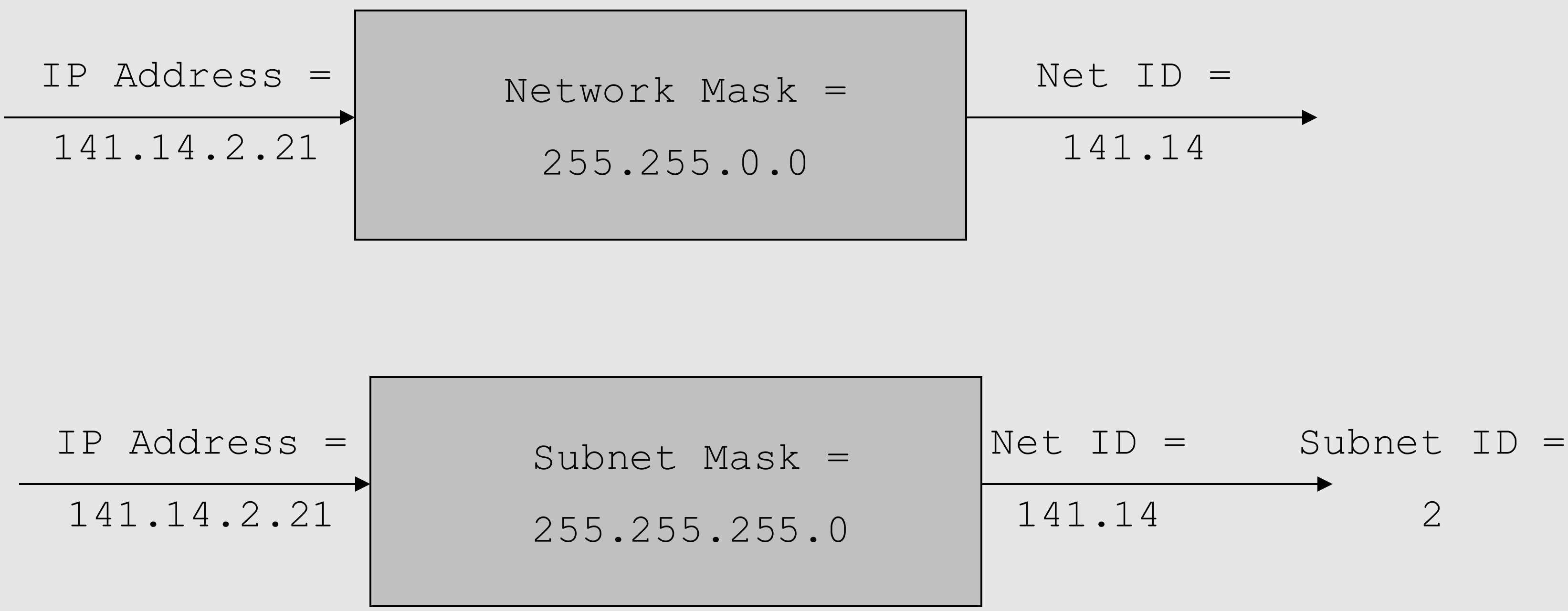
Applying Bitwise-AND Operator

Dotted decimal IP address	152 . 1 . 54 . 48				
32-bit IP address	1001 1000	0000 0001	0011 0110	0011 0000	
Implied (Classful) "Mask"	1111 1111	1111 1111	0000 0000	0000 0000	
Network address	1001 1000	0000 0001	0000 0000	0000 0000	
Dotted decimal network address	152 . 1 . 0 . 0				152.1.0.0/16

Dotted decimal IP address	152 . 1 . 54 . 48				
32-bit IP address	1001 1000	0000 0001	0011 0110	0011 0000	
Subnet Mask	1111 1111	1111 1111	1111 1111	0000 0000	
Subnetwork address	1001 1000	0000 0001	0011 0110	0000 0000	
Dotted decimal subnet address	152 . 1 . 54 . 0				152.1.54.0/24

Subnet Masks

- Apply bitwise-AND operation on IP address and mask to find the net ID and subnet ID
- How can you tell if a destination IP address is…
  - On the same subnet as you?
  - On the same network as you?



Router Forwarding Tables

- Enable subnetting by including a “mask” column
  - Provide mask with *each* table entry
  - Automatic implementation of masking by bit-AND
- Destination ID and Mask are both 32-bit numbers in the actual table

Net/Subnet	Mask	Destination
128.171.17.0	24	Local
142.0.0.0	8	Next-Hop Router A

Router Forwarding Tables

- Example
  - Net/subnet in first rule: 128.171.17.13, 24 1's in mask
    - Also written as 128.171.17.13/24
  - Destination IP Address of a packet is 128.171.17.13
  - Mask is /24, so only look at 128.171.17.0
  - Matches row's network/subnet bits, so use Local (direct) delivery

Net/Subnet	Mask	Destination
128.171.17.0	24	Local
142.0.0.0	8	Next-Hop Router A

Default routes

- Strictly speaking, need next hop information for every network in the Internet
- Instead, keep detailed routes only for local neighborhood
  - For unknown destinations, use a *default* router
  - Reduces size of routing tables at the expense of non-optimal paths
  - Move towards hierarchy

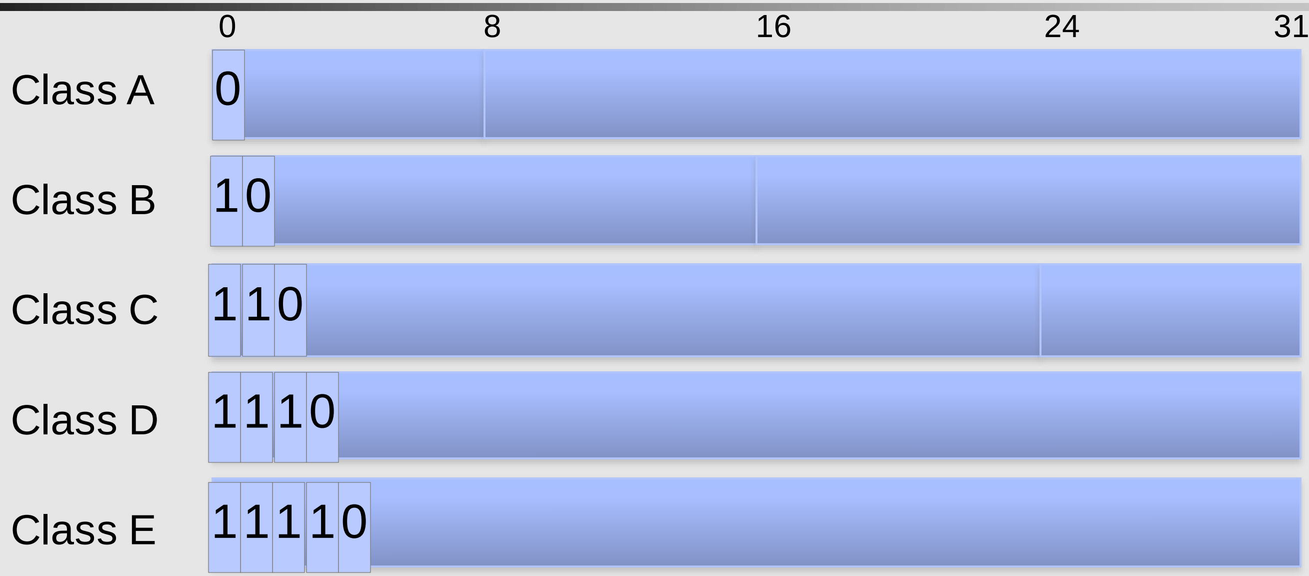
Classful IP Address

Classful IP Address

Given the following IP address, which class does it belong to?
10001100.10110010.11101101.00011001

A	Class A
B	Class B
C	Class C
D	Class D or E

IP Address Classes



Class	First Byte	# Networks	Hosts per Network	Comments
A	< 128	2^7 (128)	2^{24} –2 (16M)	Mostly used
B	128…191	2^{14} (16384)	2^{16} – 2 (65534)	Mostly used
C	192…223	2^{21} (2M)	2^8 – 2 (254)	
D	224…239	2^{28} (268M)	-	Dynamic, multicast
E	240…255	2^{27} (134M)	-	reserved

Forwarding Table Selection Rules

- Compare destination IP address of an arriving packet against **ALL** rows within the router forwarding table because there may be multiple matches
- Select a single row that matches
- If multiple rows match, select **the longest match**
- If multiple rows tie on the longest match, select the row with the largest or smallest metric, depending on the specific metric (e.g., lowest delay)
- (If there is no match, select the default row)
  - Syntactic solution - always include 0.0.0.0 as a net / mask

Example

Line	Destination Address	Netmask	Metric (Cost)	Interface	Next-Hop Router
1	152.19.0.0	16	47	2	B
2	152.15.11.0	20	1	1	Local
3	152.1.0.0	16	12	2	B
4	152.40.0.0	16	33	2	B
5	152.229.0.0	16	34	1	D
6	152.40.6.0	24	47	3	E
7	152.19.17.0	24	55	4	H
8	152.229.0.0	16	20	3	E
9	152.40.8.0	24	23	1	D
10	152.15.12.0	20	9	2	Local
11	152.15.122.0	20	3	3	Local
12	0.0.0.0	0	5	3	H

- Default router?
- 152.1.1.211 – which router/rule?
- 152.40.8.44 – which router/rule?
- 152.15.13.99 – which router/rule?
- 125.1.2.3 – which router/rule?