IP Addressing

Scalable Infrastructure Workshop

AfNOG 2011

Purpose of an IP address

- Unique Identification of:
  - Source
    - So the recipient knows where the message is from
    - Sometimes used for security or policy-based filtering of data
  - Destination
    - So the networks know where to send the data
- Network Independent Format
  - IP over anything

Purpose of an IP Address

- Identifies a machine's connection to a network
- Physically moving a machine from one network to another requires changing the IP address
- Unique; assigned in a hierarchical fashion
  - IANA (Internet Assigned Number Authority)
  - IANA to RIRs (AfriNIC, ARIN, RIPE, APNIC, LACNIC)
  - RIR to ISPs and large organisations
  - ISP or company IT department to end users
- IPv4 uses unique 32-bit addresses
- IPv6 uses unique 128-bit addresses

Basic Structure of an IPv4 Address

- 32 bit number (4 octet number):
  (e.g. 133.27.162.125)
- Decimal Representation:

133	27	162	125

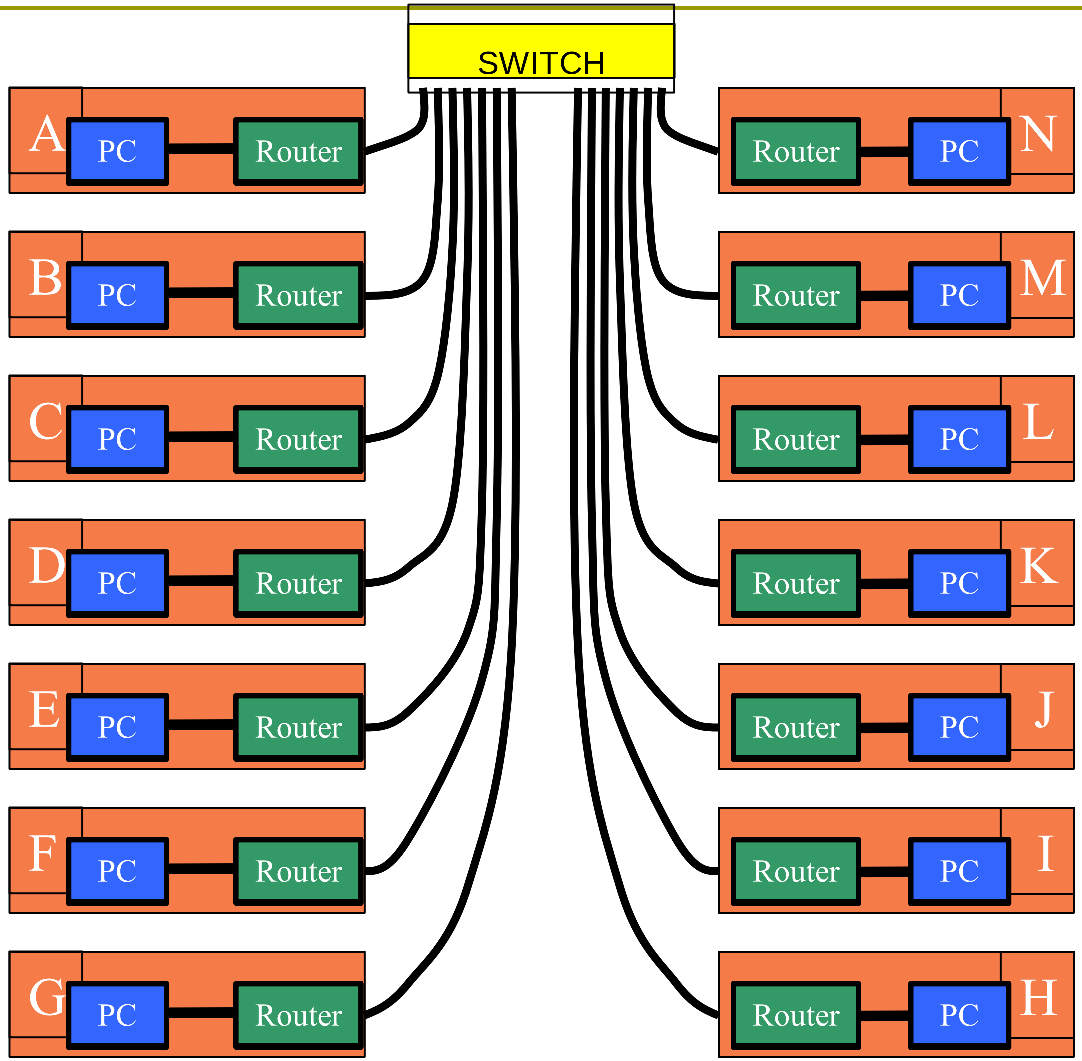
- Binary Representation:

10000101	00011011	10100010	01111101

- Hexadecimal Representation:

85	1B	A2	7D

Address Exercise



Address Exercise

- Construct an IP address for your router’s connection to the backbone network.
- 196.200.220.x
- x = 1 for row A, 2 for row B, etc.
- Write it in decimal form as well as binary form.

Addressing in Internetworks

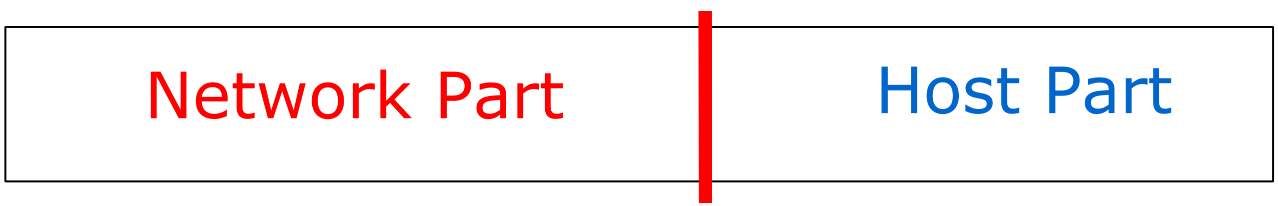
- The problem we have
  - More than one physical network
  - Different Locations
  - Larger number of hosts
  - Need a way of numbering them all
- We use a structured numbering system
  - Hosts that are connected to the same physical network have “similar” IP addresses
  - Often more then one level of structure; e.g. physical networks in the same organisation use “similar” IP addresses

Network part and Host part

- Remember IPv4 address is 32 bits
- Divide it into a “network part” and “host part”
  - “network part” of the address identifies which network in the internetwork (e.g. the Internet)
  - “host part” identifies host on that network
  - Hosts or routers connected to the same link-layer network will have IP addresses with the same network part, but different host part.
  - Host part contains enough bits to address all hosts on the subnet; e.g. 8 bits allows 256 addresses

Dividing an address

- Hierarchical Division in IP Address:
  - Network Part (or Prefix) – high order bits (left)
    - describes which physical network
  - Host Part – low order bits (right)
    - describes which host on that network



  - Boundary can be anywhere
    - choose the boundary according to number of hosts
    - very often NOT a multiple of 8 bits

Network Masks

- “Network Masks” help define which bits are used to describe the Network Part and which for the Host Part
- Different Representations:
  - decimal dot notation: 255.255.224.0
  - binary: 11111111 11111111 11100000 00000000
  - hexadecimal: 0xFFFFE000
  - number of network bits: /19
    - count the 1's in the binary representation
- Above examples all mean the same: 19 bits for the Network Part and 13 bits for the Host Part

Example Prefixes

- 137.158.128.0/17 (netmask 255.255.128.0)

  1111 1111 1111 1111 1 | 000 0000 0000 0000

1000 1001	1001 1110	1 | 000 0000	0000 0000

- 198.134.0.0/16 (netmask 255.255.0.0)

  1111 1111 1111 1111 | 0000 0000 0000 0000

1100 0110	1000 0110 |	0000 0000	0000 0000

- 205.37.193.128/26 (netmask 255.255.255.192)

  1111 1111 1111 1111 1111 1111 11 | 00 0000

1100 1101	0010 0101	1100 0001	10 | 00 0000

Special Addresses

- All 0’s in host part: Represents Network
  - e.g. 193.0.0.0/24
  - e.g. 138.37.64.0/18
  - e.g. 196.200.223.96/28
- All 1’s in host part: Broadcast
  - e.g. 193.0.0.255   (prefix 193.0.0.0/24)
  - e.g. 138.37.127.255 (prefix 138.37.64.0/18)
  - e.g. 196.200.223.111 (prefix 196.200.223.96/28)
- 127.0.0.0/8: Loopback address (127.0.0.1)
- 0.0.0.0: Various special purposes

Ancient History: Implied netmasks

- A classful network had a “natural” or “implied” prefix length or netmask:
  - Class A: prefix length /8 (netmask 255.0.0.0)
  - Class B: prefix length /16 (netmask 255.255.0.0)
  - Class C: prefix length /24 (netmask 255.255.255.0)
- Modern (classless) routing systems have explicit prefix lengths or netmasks
  - You can't just look at an IP address to tell what the prefix length or netmask should be. Protocols and configurations need explicit netmask or prefix length.

Classless addressing

- Class A, Class B, Class C terminology and restrictions are now of historical interest only
  - Obsolete in 1994
- Internet routing and address management today is classless
- CIDR = Classless Inter-Domain Routing
  - Routing does not assume that former class A, B, C addresses imply prefix lengths of /8, /16, /24
- VLSM = Variable-Length Subnet Masks
  - Routing does not assume that all subnets are the same size

Classless addressing example

- An ISP gets a large block of addresses
  - e.g., a /16 prefix, or 65536 separate addresses
- Assign smaller blocks to customers
  - e.g., a /22 prefix (1024 addresses) to one customer, and a /28 prefix (16 addresses) to another customer (and some space left over for other customers)
- An organisation that gets a /22 prefix from their ISP divides it into smaller blocks
  - e.g. a /26 prefix (64 addresses) for one department, and a /27 prefix (32 addresses) for another department (and some space left over for other internal networks)

Classless addressing exercise

- Consider the address block 133.27.162.0/23
- Allocate 5 separate /29 blocks, one /27 block, and one /25 block
- What are the IP addresses of each block allocated above?
  - In prefix length notation
  - Netmasks in decimal
  - IP address ranges
- What blocks are still available (not yet allocated)?
- How big is the largest available block?

Configuring interfaces – *ifconfig*

- ifconfig interface [address_family] address [params]
  - interface: network interface, e.g., eth0
  - options: up, down, netmask mask
  - address: IP address

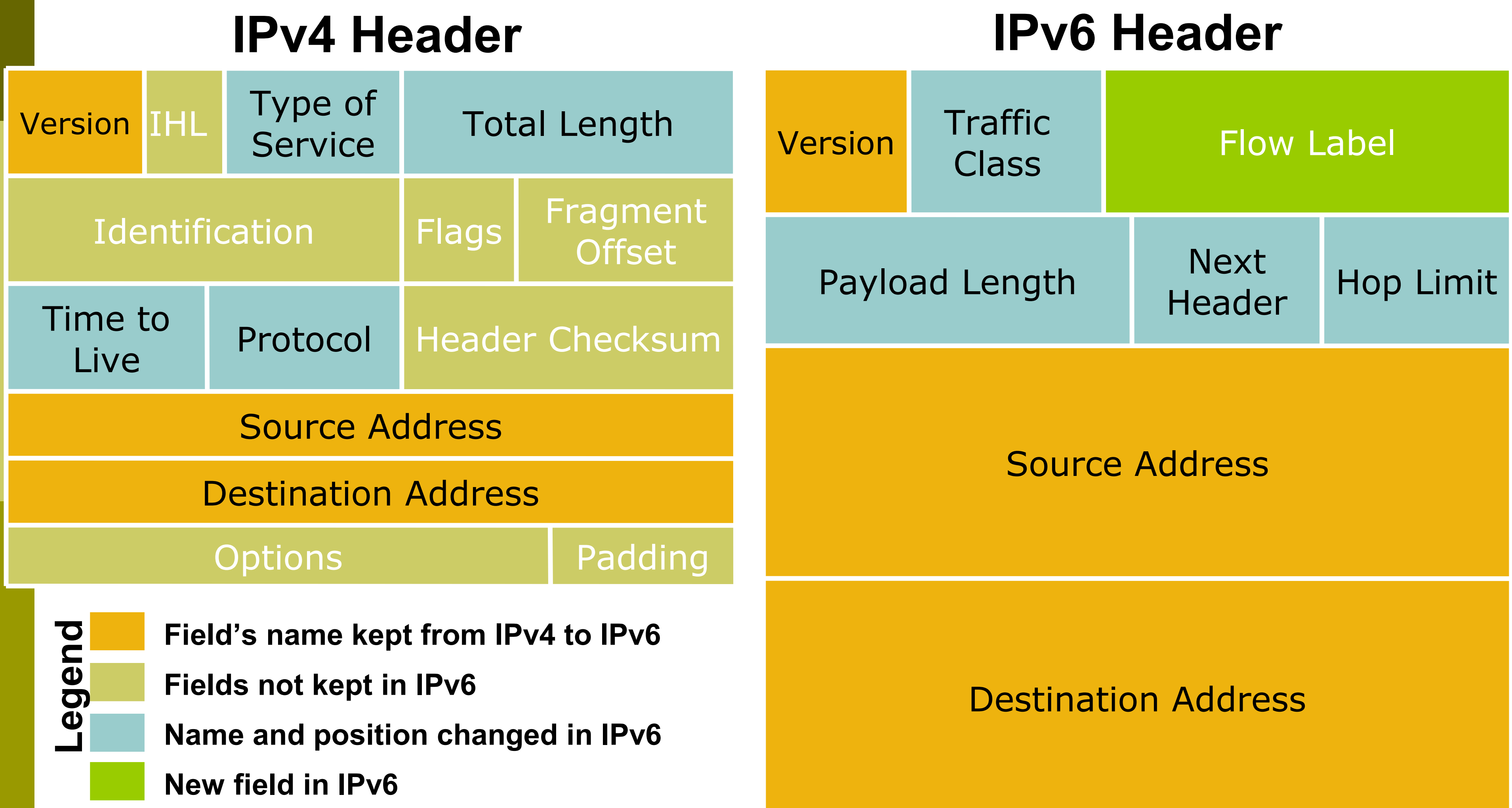
- Examples:
  - ifconfig eth0 192.168.2.2; ifconfig eth1 192.168.3.1
  - ifconfig eth0
  - ifconfig eth0 192.168.2.2 netmask 255.255.255.0
  - ifconfig eth0 inet6 2001:db8:bdbd::123 prefixlen 48 alias

IPv6 Addressing

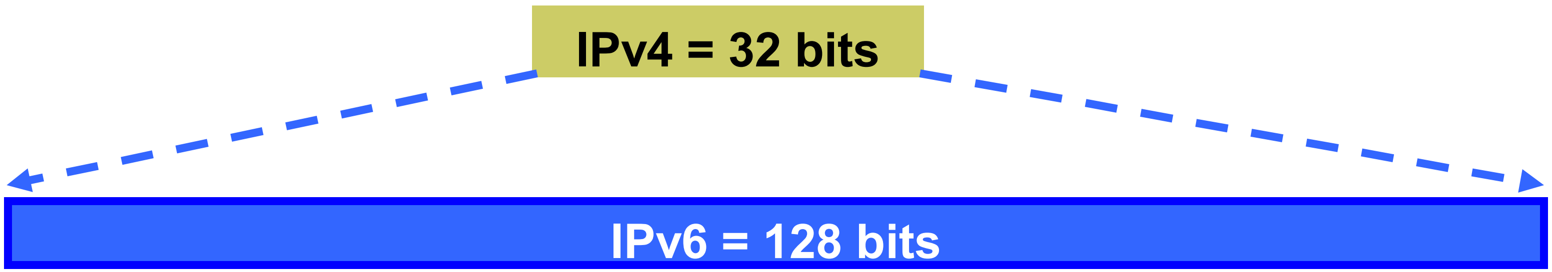
IP version 6

- IPv6 designed as successor to IPv4
  - Expanded address space
    - Address length quadrupled to 16 bytes (128 bits)
  - Header Format Simplification
    - Fixed length, optional headers are daisy-chained
  - No checksum at the IP network layer
  - No hop-by-hop fragmentation
    - Path MTU discovery
  - 64 bits aligned fields in the header
  - Authentication and Privacy Capabilities
    - IPsec is mandated
  - No more broadcast

IPv4 and IPv6 Header Comparison



Larger Address Space



- IPv4
  - 32 bits
  - = 4,294,967,296 possible addressable devices
- IPv6
  - 128 bits: 4 times the size in bits
  - = 3.4×10^{38} possible addressable devices
  - = 340,282,366,920,938,463,463,374,607,431,768,211,456
  - ~ 5×10^{28} addresses per person on the planet

IPv6 Address Representation

- 16 bit fields in case insensitive colon hexadecimal representation
  - 2031:0000:130F:0000:0000:09C0:876A:130B
- Leading zeros in a field are optional:
  - 2031:0:130F:0:0:9C0:876A:130B
- Successive fields of 0 represented as ::, but only once in an address:
  - 2031:0:130F::9C0:876A:130B ← is ok
  - 2031::130F::9C0:876A:130B is NOT ok (two "::")
  - 0:0:0:0:0:0:0:1 → ::1 (loopback address)
  - 0:0:0:0:0:0:0:0 → :: (unspecified address)

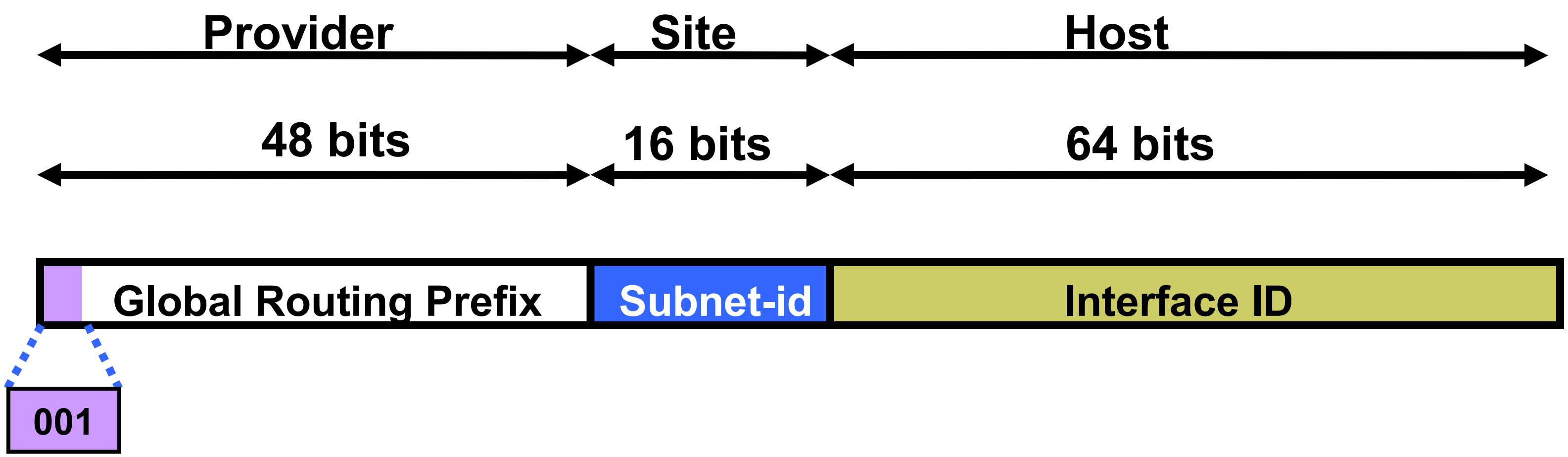
IPv6 Address Representation

- In a URL, it is enclosed in brackets (RFC3986)
  - http://[2001:db8:4f3a::206:ae14]:8080/index.html
  - Cumbersome for users
  - Mostly for diagnostic purposes
  - Use fully qualified domain names (FQDN) instead of this
- Prefix Representation
  - Representation of prefix is same as for IPv4 CIDR
    - Address and then prefix length, with slash separator
  - IPv4 address:
    - 198.10.0.0/16
  - IPv6 address:
    - 2001:db8:12::/40

IPv6 Addressing

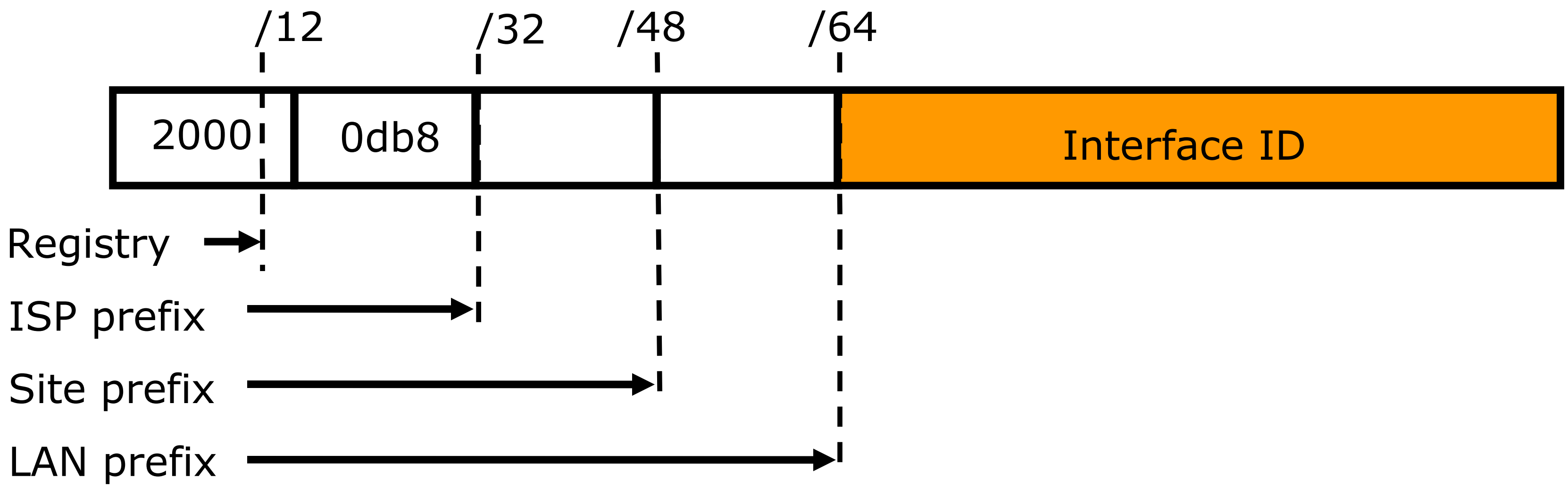
Type	Binary	Hex
Unspecified	0000…0000	::/128
Loopback	0000…0001	::1/128
Global Unicast Address	0010 ...	2000::/3
Link Local Unicast Address	1111 1110 10...	FE80::/10
Unique Local Unicast Address	1111 1100 ... 1111 1101 ...	FC00::/7
Multicast Address	1111 1111 ...	FF00::/8

IPv6 Global Unicast Addresses



- IPv6 Global Unicast addresses are:
  - Addresses for generic use of IPv6
  - Hierarchical structure intended to simplify aggregation

IPv6 Address Allocation

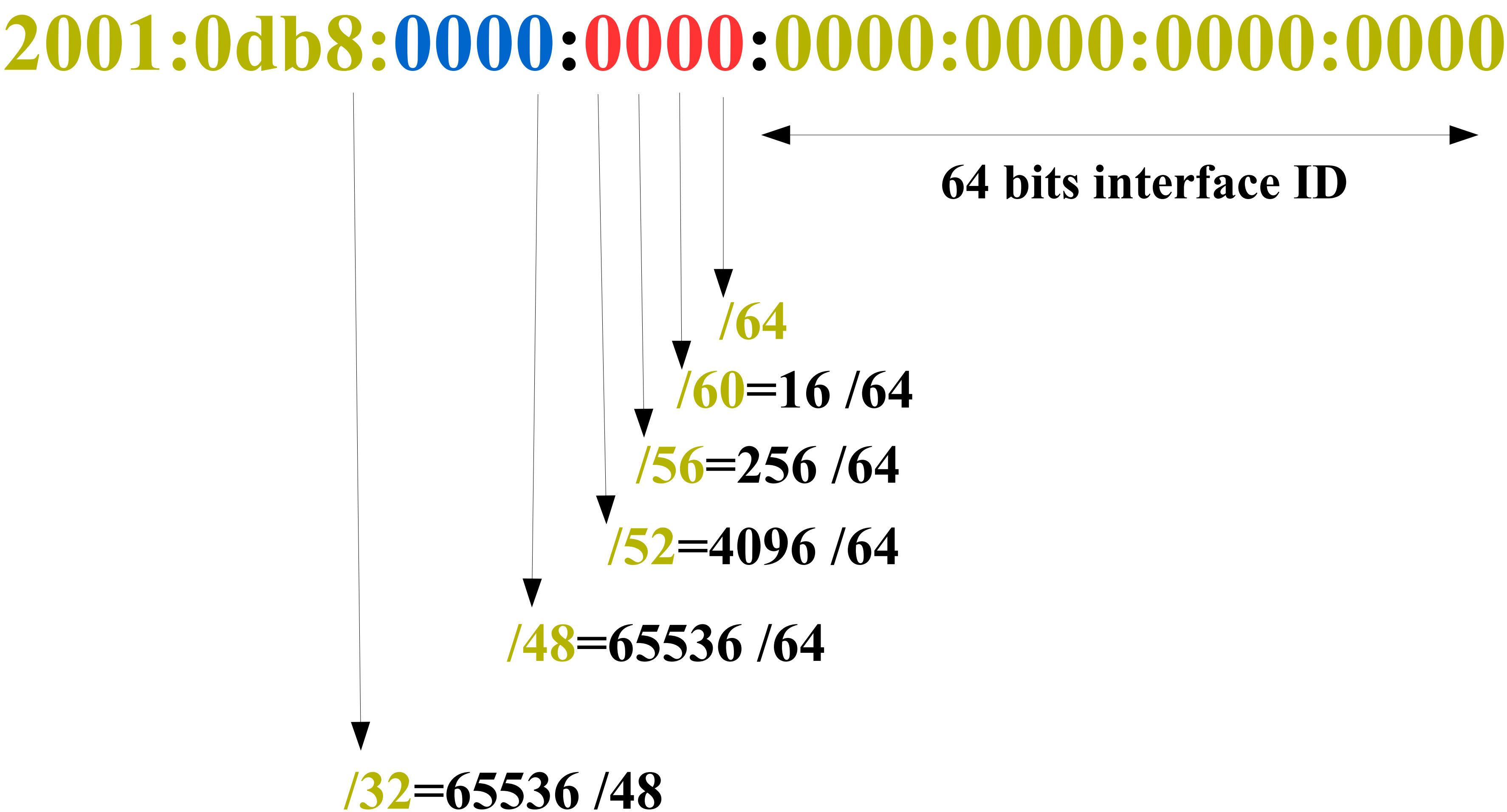


- The allocation process is:
  - The IANA is allocating out of 2000::/3 for initial IPv6 unicast use
  - Each registry gets a /12 prefix from the IANA
  - Registry allocates a /32 prefix (or larger) to an IPv6 ISP
  - ISPs usually allocate a /48 prefix to each end customer

IPv6 Addressing Scope

- 64 bits used for the interface ID
  - Possibility of 2^{64} hosts on one network LAN
  - Arrangement to accommodate MAC addresses within the IPv6 address
- 16 bits used for the end site
  - Possibility of 2^{16} networks at each end-site
  - 65536 subnets

IPV6 Subnetting



Nibble (4 bits) Concept

Decimal	Binary	Hexadecimal
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
10	1010	a
11	1011	b
12	1100	c
13	1101	d
14	1110	e
15	1111	f

2001:0db8::/32

2001:0db8:0000::/48	2001:0db8:0100::/48	2001:0db8:0200::/48	2001:0db8:0300::/48
2001:0db8:0001::/48	2001:0db8:0101::/48	2001:0db8:0201::/48	2001:0db8:0301::/48
2001:0db8:0002::/48	2001:0db8:0102::/48	2001:0db8:0202::/48	2001:0db8:0302::/48
2001:0db8:0003::/48	2001:0db8:0103::/48	2001:0db8:0203::/48	2001:0db8:0303::/48
2001:0db8:0004::/48	2001:0db8:0104::/48	2001:0db8:0204::/48	2001:0db8:0304::/48
2001:0db8:0005::/48	2001:0db8:0105::/48	2001:0db8:0205::/48	2001:0db8:0305::/48
2001:0db8:0006::/48	2001:0db8:0106::/48	2001:0db8:0206::/48	2001:0db8:0306::/48
2001:0db8:0007::/48	2001:0db8:0107::/48	2001:0db8:0207::/48	2001:0db8:0307::/48
2001:0db8:0008::/48	2001:0db8:0108::/48	2001:0db8:0208::/48	2001:0db8:0308::/48
2001:0db8:0009::/48	2001:0db8:0109::/48	2001:0db8:0209::/48	2001:0db8:0309::/48
2001:0db8:000a::/48	2001:0db8:010a::/48	2001:0db8:020a::/48	2001:0db8:030a::/48
2001:0db8:000b::/48	2001:0db8:010b::/48	2001:0db8:020b::/48	2001:0db8:030b::/48
2001:0db8:000c::/48	2001:0db8:010c::/48	2001:0db8:020c::/48	2001:0db8:030c::/48
2001:0db8:000d::/48	2001:0db8:010d::/48	2001:0db8:020d::/48	2001:0db8:030d::/48
2001:0db8:000e::/48	2001:0db8:010e::/48	2001:0db8:020e::/48	2001:0db8:030e::/48
2001:0db8:000f::/48	2001:0db8:010f::/48	2001:0db8:020f::/48	2001:0db8:030f::/48
2001:0db8:0010::/48	2001:0db8:0110::/48	2001:0db8:0210::/48	2001:0db8:0310::/48
2001:0db8:0011::/48	2001:0db8:0111::/48	2001:0db8:0211::/48	2001:0db8:0311::/48
2001:0db8:0012::/48	2001:0db8:0112::/48	2001:0db8:0212::/48	2001:0db8:0312::/48
2001:0db8:0013::/48	2001:0db8:0113::/48	2001:0db8:0213::/48	2001:0db8:0313::/48
2001:0db8:0014::/48	2001:0db8:0114::/48	2001:0db8:0214::/48	2001:0db8:0314::/48
2001:0db8:0015::/48	2001:0db8:0115::/48	2001:0db8:0215::/48	2001:0db8:0315::/48
2001:0db8:0016::/48	2001:0db8:0116::/48	2001:0db8:0216::/48	2001:0db8:0316::/48
2001:0db8:0017::/48	2001:0db8:0117::/48	2001:0db8:0217::/48	2001:0db8:0317::/48
2001:0db8:0018::/48	2001:0db8:0118::/48	2001:0db8:0218::/48	2001:0db8:0318::/48
2001:0db8:0019::/48	2001:0db8:0119::/48	2001:0db8:0219::/48	2001:0db8:0319::/48
2001:0db8:001a::/48	2001:0db8:011a::/48	2001:0db8:021a::/48	2001:0db8:031a::/48
2001:0db8:001b::/48	2001:0db8:011b::/48	2001:0db8:021b::/48	2001:0db8:031b::/48
2001:0db8:001c::/48	2001:0db8:011c::/48	2001:0db8:021c::/48	2001:0db8:031c::/48
2001:0db8:001d::/48	2001:0db8:011d::/48	2001:0db8:021d::/48	2001:0db8:031d::/48
2001:0db8:001e::/48	2001:0db8:011e::/48	2001:0db8:021e::/48	2001:0db8:031e::/48
2001:0db8:001f::/48	2001:0db8:011f::/48	2001:0db8:021f::/48	2001:0db8:031f::/48
2001:0db8:0020::/48	2001:0db8:0120::/48	2001:0db8:0220::/48	2001:0db8:0320::/48
2001:0db8:0021::/48	2001:0db8:0121::/48	2001:0db8:0221::/48	2001:0db8:0321::/48
2001:0db8:0022::/48	2001:0db8:0122::/48	2001:0db8:0222::/48	2001:0db8:0322::/48
2001:0db8:0023::/48	2001:0db8:0123::/48	2001:0db8:0223::/48	2001:0db8:0323::/48
2001:0db8:0024::/48	2001:0db8:0124::/48	2001:0db8:0224::/48	2001:0db8:0324::/48
2001:0db8:0025::/48	2001:0db8:0125::/48	2001:0db8:0225::/48	2001:0db8:0325::/48
2001:0db8:0026::/48	2001:0db8:0126::/48	2001:0db8:0226::/48	2001:0db8:0326::/48
2001:0db8:0027::/48	2001:0db8:0127::/48	2001:0db8:0227::/48	2001:0db8:0327::/48
2001:0db8:0028::/48	2001:0db8:0128::/48	2001:0db8:0228::/48	2001:0db8:0328::/48
2001:0db8:0029::/48	2001:0db8:0129::/48	2001:0db8:0229::/48	2001:0db8:0329::/48
2001:0db8:002a::/48	2001:0db8:012a::/48	2001:0db8:022a::/48	2001:0db8:032a::/48
2001:0db8:002b::/48	2001:0db8:012b::/48	2001:0db8:022b::/48	2001:0db8:032b::/48
2001:0db8:002c::/48	2001:0db8:012c::/48	2001:0db8:022c::/48	2001:0db8:032c::/48
2001:0db8:002d::/48	2001:0db8:012d::/48	2001:0db8:022d::/48	2001:0db8:032d::/48
2001:0db8:002e::/48	2001:0db8:012e::/48	2001:0db8:022e::/48	2001:0db8:032e::/48
2001:0db8:002f::/48	2001:0db8:012f::/48	2001:0db8:022f::/48	2001:0db8:032f::/48
2001:0db8:0030::/48	2001:0db8:0130::/48	2001:0db8:0230::/48	2001:0db8:0330::/48
2001:0db8:0031::/48	2001:0db8:0131::/48	2001:0db8:0231::/48	2001:0db8:0331::/48

Summary

- Vast address space
- Hexadecimal addressing
- Distinct addressing hierarchy between ISPs, end-sites, and LANs
  - ISPs have /32s
  - End-sites have /48s
  - LANs have /64s
- Other IPv6 features discussed later

Large Network Issues & Routers

The need for Packet Forwarding

- Many small networks can be interconnected to make a larger internetwork
- A device on one network cannot send a packet directly to a device on another network
- The packet has to be forwarded from one network to another, through intermediate nodes, until it reaches its destination
- The intermediate nodes are called “routers”

An IP Router

- A device with more than one link-layer interface
- Different IP addresses (from different subnets) on different interfaces
- Receives packets on one interface, and forwards them (usually out of another interface) to get them one hop closer to their destination
- Maintains forwarding tables

IP router - action for each packet

- Packet is received on one interface
- Checks whether the destination address is the router itself – if so, pass it to higher layers
- Decrement TTL (time to live), and discard packet if it reaches zero
- Look up the destination IP address in the forwarding table
- Destination could be on a directly attached link, or through another router

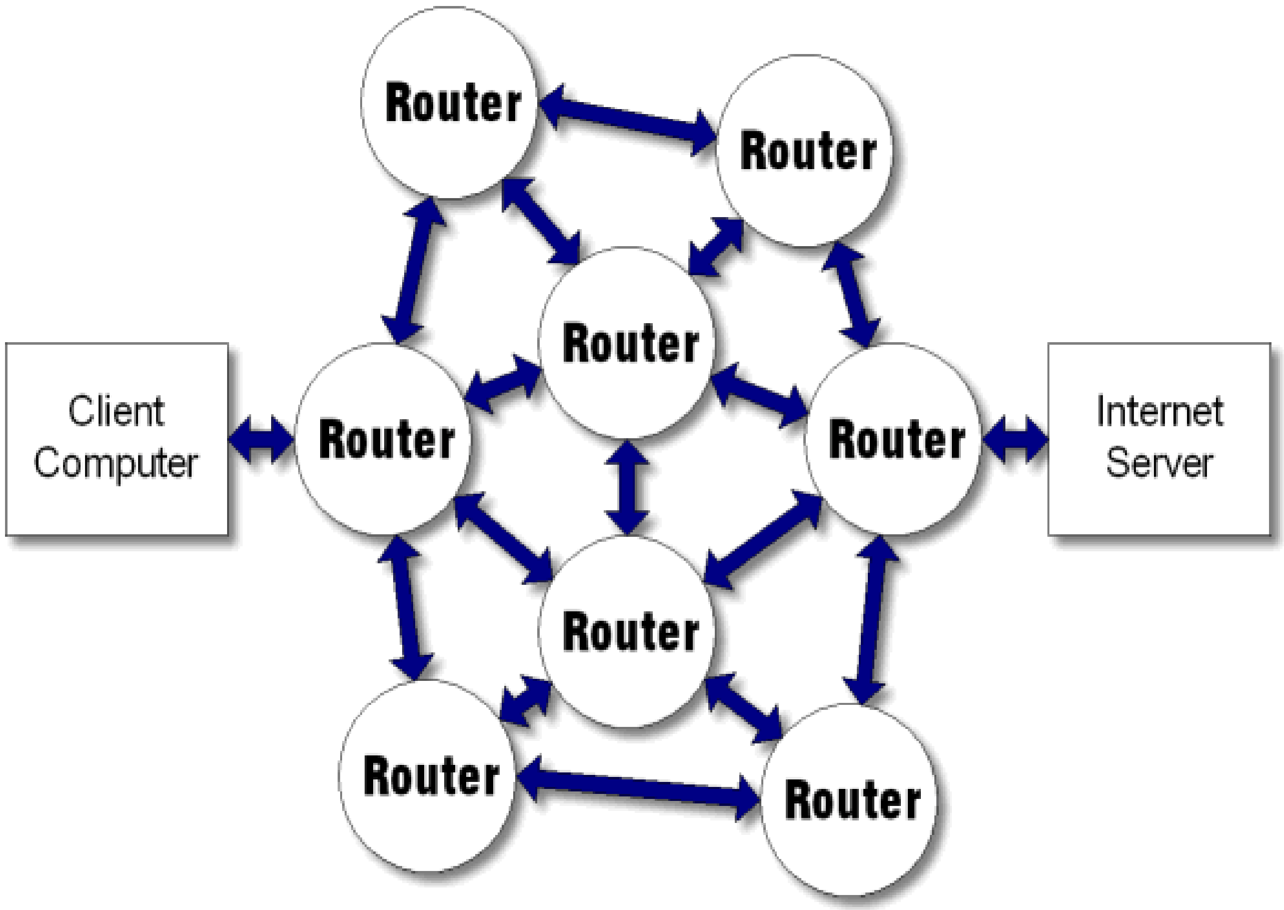
Forwarding vs. Routing

- Forwarding: the process of moving packets from input to output
  - The forwarding table
  - Information in the packet
- Routing: process by which the forwarding table is built and maintained
  - One or more routing protocols
  - Procedures (algorithms) to convert routing info to forwarding table.
- (Much more later …)

Forwarding is hop by hop

- Each router tries to get the packet one hop closer to the destination
- Each router makes an independent decision, based on its own forwarding table
- Different routers have different forwarding tables and make different decisions
  - If all is well, decisions will be consistent
- Routers talk routing protocols to each other, to help update routing and forwarding tables

Hop by Hop Forwarding



Router Functions

- Determine optimum routing paths through a network
  - Lowest delay
  - Highest reliability
- Move packets through the network
  - Examines destination address in packet
  - Makes a decision on which port to forward the packet through
  - Decision is based on the Routing Table
- Interconnected Routers exchange routing tables in order to maintain a clear picture of the network
- In a large network, the routing table updates can consume a lot of bandwidth
  - a protocol for route updates is required

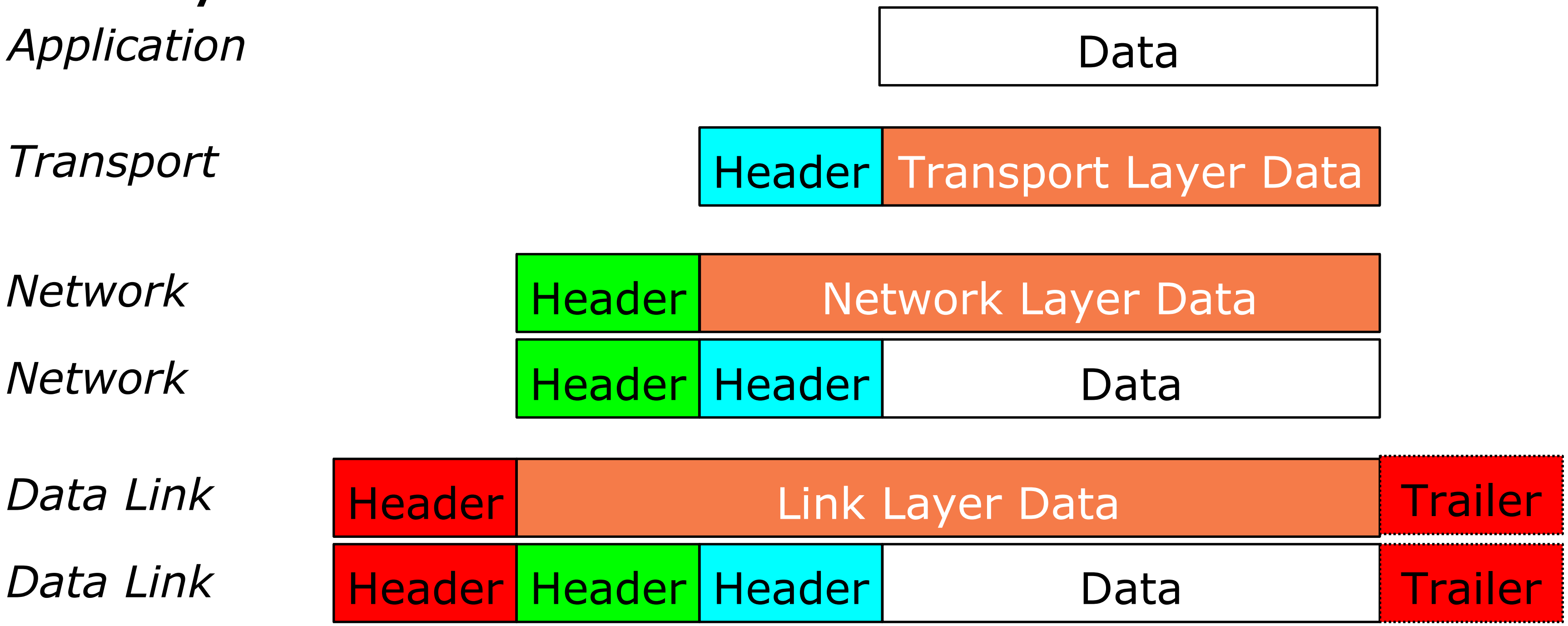
Forwarding table structure

- We don't list every IP number on the Internet - the table would be huge
- Instead, the forwarding table contains prefixes (network numbers)
  - "If the first /n bits matches this entry, send the datagram thataway"
- If more than one prefix matches, the longest prefix wins (more specific route)
- 0.0.0.0/0 is "default route" - matches anything, but only if no other prefix matches

ARP

Encapsulation Reminder

- Lower layers add headers (and sometimes trailers) to data from higher layers



Ethernet Essentials

- Ethernet is a broadcast medium
- Structure of Ethernet frame:

Preamble	Dest	Source	Length	Type	Data	CRC

- Entire IP packet makes data part of Ethernet frame
- Delivery mechanism (CSMA/CD)
  - back off and try again when collision is detected

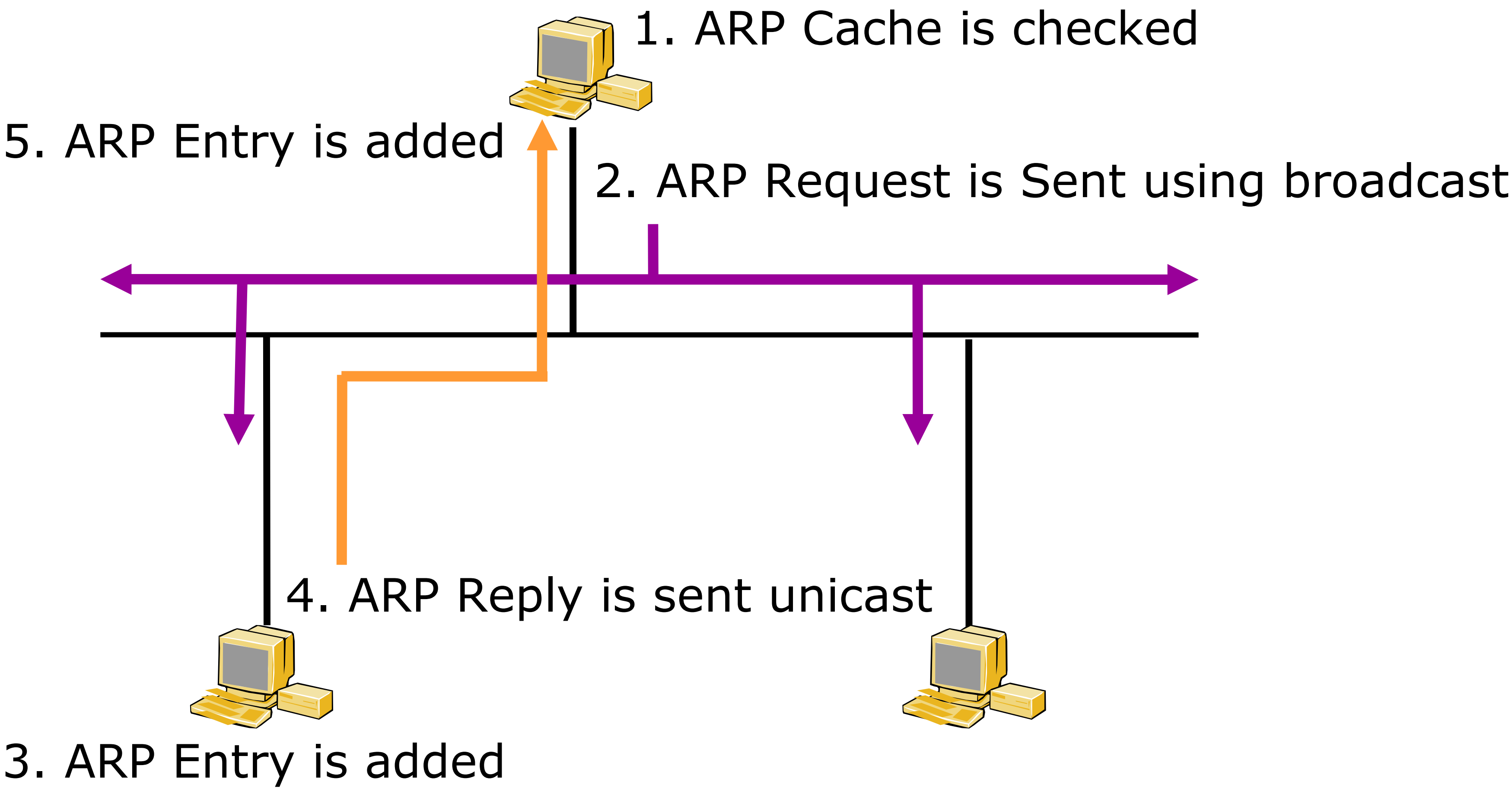
Ethernet/IP Address Resolution

- Internet Address
  - Unique worldwide (excepting private nets)
  - Independent of Physical Network technology
- Ethernet Address
  - Unique worldwide (excepting errors)
  - Ethernet Only
- Need to map from higher layer to lower (i.e. IP to Ethernet, using ARP)

Address Resolution Protocol

- ARP is only used in IPv4
  - ND (Neighbor Discovery) replaces ARP in IPv6
- Check ARP cache for matching IP address
- If not found, broadcast packet with IP address to every host on Ethernet
- "Owner" of the IP address responds
- Response cached in ARP table for future use
- Old cache entries removed by timeout

ARP Procedure



ARP Table

IP Address	Hardware Address	Age (Sec)
192.168.0.2	08-00-20-08-70-54	3
192.168.0.65	05-02-20-08-88-33	120
192.168.0.34	07-01-20-08-73-22	43

Types of ARP Messages

- ARP request
  - Who is IP addr X.X.X.X tell IP addr Y.Y.Y.Y
- ARP reply
  - IP addr X.X.X.X is Ethernet Address hh:hh:hh:hh:hh:hh