# Computer Networks

# IP Forwarding and Address Allocation

Amitangshu Pal
Computer Science and Engineering
IIT Kanpur

# IP forwarding

# Package Forwarding in a Postal System

## Postal analogy:

6 packages need to be sent in these addresses

- UP, Prayagraj, Barrister Mullah Colony, Swati
- UP, Kanpur, Kalyanpur, Amit
- West Bengal, Kharagpur, Inda, Manoj
- West Bengal, Durgapur, Mahatma Gandhi Rd, Prakash
- Maharashtra, Pune, Shivaji Nagar, Vivek
- Maharashtra, Nasik, Old Agra Road, Anand

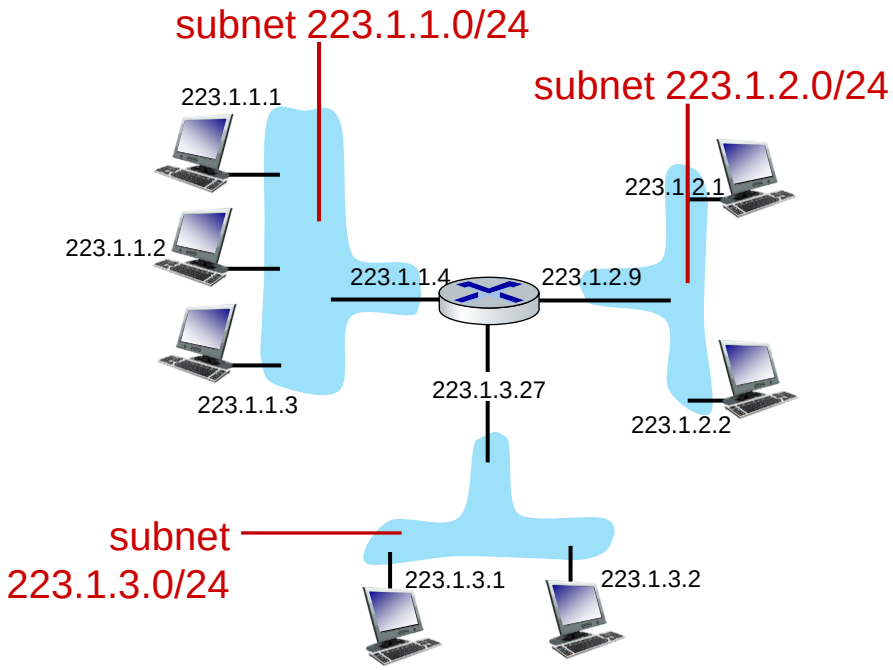
Src: https://pxhere.com/en/photo/1625663

| Destination prefix | Next hop |
|---|---|
| UP | Lucknow flight |
| West Bengal | Kolkata flight |
| Maharashtra | Mumbai flight |

# IP Forwarding

- IP addresses on a network belongs to the same prefix
- Nodes use a forwarding table that lists the prefixes and the corresponding next hop/link interface
  - Improves scalability → reduces the forwarding table size

| Prefix | Next hop/interface |
|---|---|
| 223.1.1.0/24 | 223.1.1.4 |
| 223.1.2.0/24 | 223.1.2.9 |
| 223.1.3.0/24 | 223.1.3.27 |



# IP Forwarding

*forwarding table*

| Destination Address Range | Link Interface |
|---|---|
| 11001000 00010111 00010000 00000000<br>through | 0 |
| 11001000 00010111 00010000 00000100<br>through<br>11001000 00010111 00010000 00000111 | 3 |
| 11001000 00010111 00011000 11111111 | |
| 11001000 00010111 00011001 00000000<br>through<br>11001000 00010111 00011111 11111111 | 2 |
| otherwise | 3 |

But what happens if ranges don't divide up so nicely?

# Longest Prefix Matching

- Prefixes in a forwarding table might overlap

## Longest prefix match

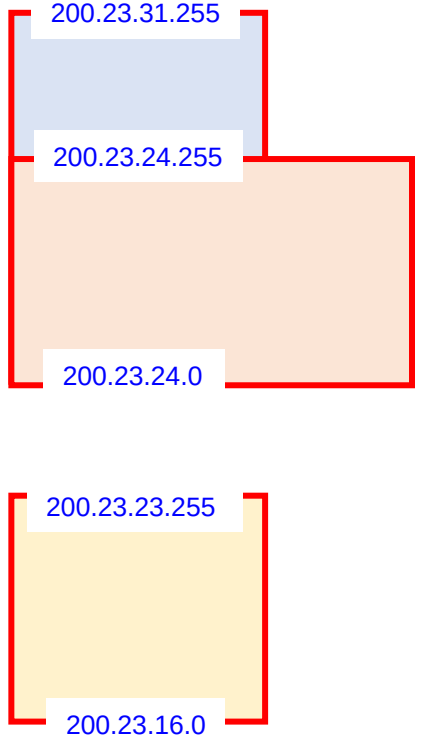
When looking for forwarding table entry for given destination address, use longest address prefix (**most specific entry**) that matches destination address

| Destination Address Range | Link interface |
|---|---|
| 11001000 00010111 00010 *** ******** | 0 |
| 11001000 00010111 00011000 ******** | 1 |
| 11001000 00010111 00011 *** ******** | 2 |
| otherwise | 3 |

Examples:

11001000 00010111 00010110 10100001 which interface?

11001000 00010111 00011000 10101010 which interface?



# Longest Prefix Matching

- Prefixes in a forwarding table might overlap

## Longest prefix match

When looking for forwarding table entry for given destination address, use longest address prefix (**most specific entry**) that matches destination address

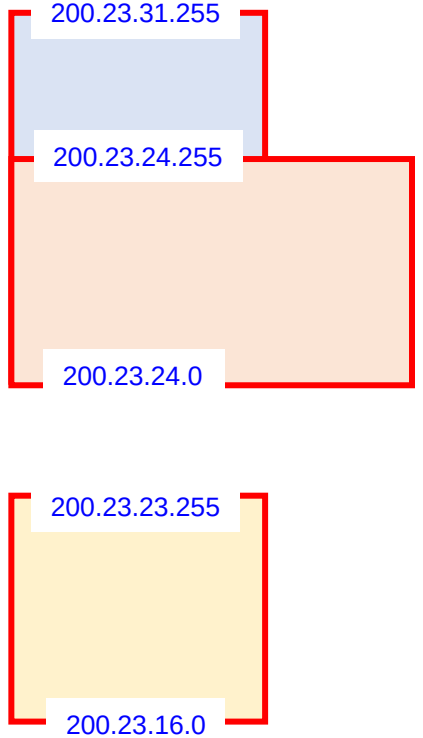
| Destination Address Range | Link interface |
|---|---|
| 11001000 00010111 00010 *** ******** | 0 |
| 11001000 00010111 00011000 ******** | 1 |
| 11001000 00011 *** ******** | 2 |
| otherwise | 3 |

match!

Examples:

11001000 00010111 00010110 10100001 which interface?

11001000 00010111 00011000 10101010 which interface?



# Longest Prefix Matching

- Prefixes in a forwarding table might overlap

## Longest prefix match

When looking for forwarding table entry for given destination address, use longest address prefix (**most specific entry**) that matches destination address

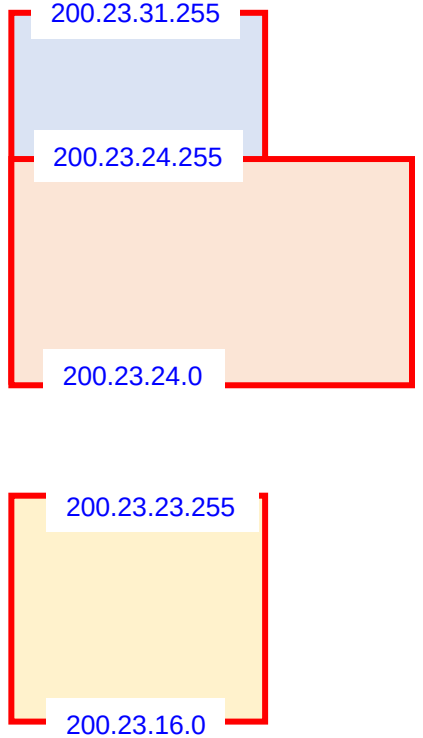
| Destination Address Range | Link interface |
|---|---|
| 11001000 00010111 00010 *** ******** | 0 |
| 11001000 00010111 00011000 ******** | 1 |
| 11001000 00010111 00011 *** ******** | 2 |
| otherwise | 3 |

match!

Examples:

11001000 00010111 00010110 10100001 which interface?

11001000 00010111 00011000 10101010 which interface?



# Longest Prefix Matching

- Prefixes in a forwarding table might overlap

**Longest prefix match**

When looking for forwarding table entry for given destination address, use longest address prefix (**most specific entry**) that matches destination address

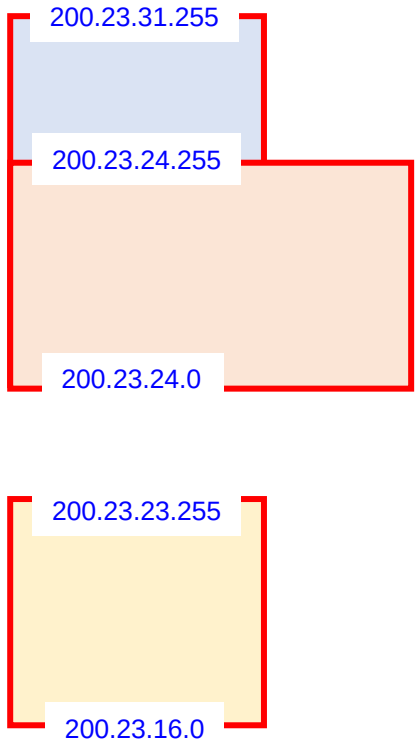
| Destination Address Range | Link interface |
|---|---|
| 11001000 00010111 00010 *** ******** | 0 |
| 11001000 00010111 00011000 ******** | 1 |
| 11001000 00010111 00011 *** ******** | 2 |
| otherwise | 3 |

match!

Examples:

11001000 00010111 00010110 10100001 which interface?

11001000 00010111 00011000 10101010 which interface?



# Package Forwarding in a Postal System

## Postal analogy:

6 packages need to be sent in these addresses

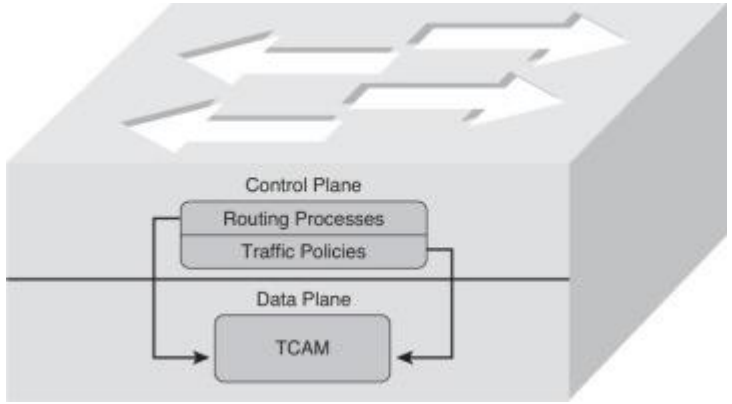
- UP, Prayagraj, Barrister Mullah Colony, Swati
- UP, Kanpur, Kalyanpur, Amit
- West Bengal, Kharagpur, Inda, Manoj
- West Bengal, Durgapur, Mahatma Gandhi Rd, Prakash
- Maharashtra, Pune, Shivaji Nagar, Vivek
- Maharashtra, Nasik, Old Agra Road, Anand

Src: https://pxhere.com/en/photo/1625663

| Destination prefix | Next hop |
|---|---|
| UP | Lucknow flight |
| UP, Prayagraj | Prayagraj flight |
| West Bengal | Kolkata flight |
| Maharashtra | Mumbai flight |

# Longest Prefix Matching

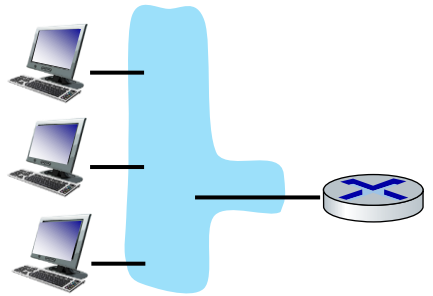
- Longest prefix matching: Often performed using ternary content addressable memories (TCAMs)
  - Content addressable: Present address to TCAM → retrieve address in one clock cycle, regardless of table size
  - Cisco Catalyst:  ~1M routing table entries in TCAM

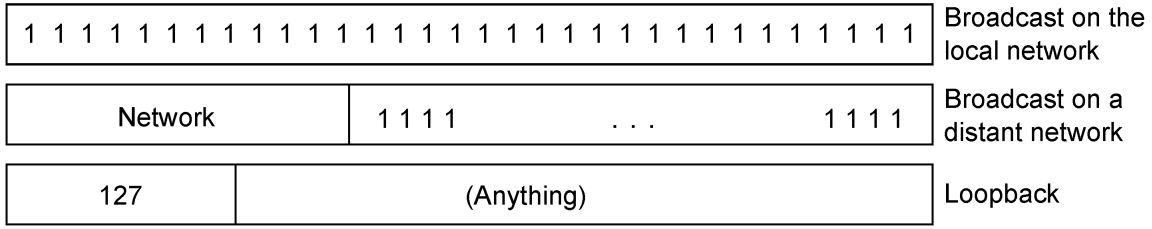



Src: https://www.ciscopress.com/articles/article.asp?p=2264831&seqNum=2

# Host Forwarding

- Hosts send the remote traffic to its nearest router
- 0.0.0.0/0 includes all IP addresses






| Prefix | Next hop |
|---|---|
| My network prefix | Send directly to that destination |
| 0.0.0.0/0 | Send to the router |

```
IPv4 Route Table
===========================================================================
Active Routes:
Network Destination        Netmask          Gateway       Interface  Metric
          0.0.0.0          0.0.0.0    172.27.16.254    172.27.17.157     291
        127.0.0.0        255.0.0.0         On-link         127.0.0.1     331
        127.0.0.1  255.255.255.255         On-link         127.0.0.1     331
  127.255.255.255  255.255.255.255         On-link         127.0.0.1     331
      172.27.0.0      255.255.0.0         On-link    172.27.17.157     291
    172.27.17.157  255.255.255.255         On-link    172.27.17.157     291
    172.27.255.255  255.255.255.255         On-link    172.27.17.157     291
        224.0.0.0        240.0.0.0         On-link         127.0.0.1     331
        224.0.0.0        240.0.0.0         On-link    172.27.17.157     291
  255.255.255.255  255.255.255.255         On-link         127.0.0.1     331
  255.255.255.255  255.255.255.255         On-link    172.27.17.157     291
===========================================================================
```

# How To Obtain IP Address?

# IP Addresses: How to Get One?

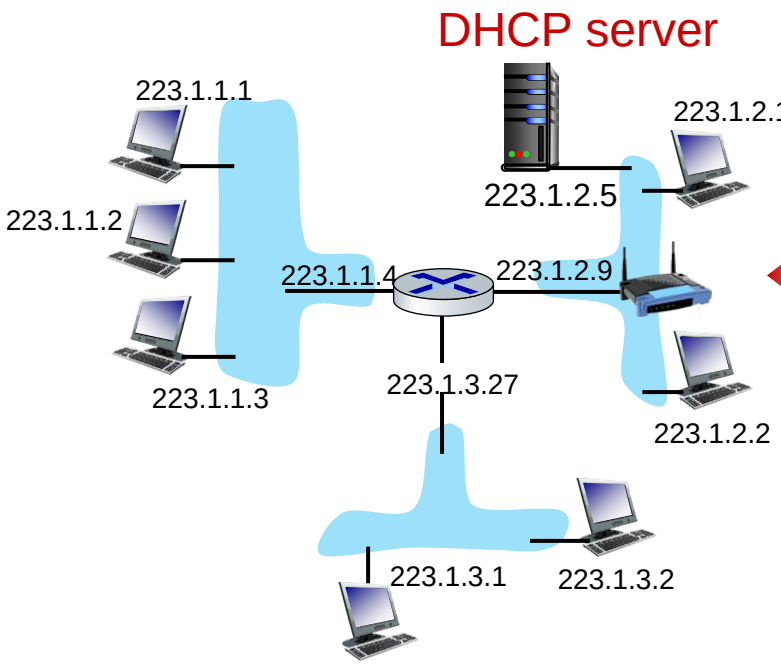
That's actually two questions:

1. How does a host get IP address within its network (host part of address)?
2. How does a network get IP address for itself (network part of address)

How does host get IP address?

- Hard-coded by sysadmin in config file (e.g., /etc/rc.config in UNIX)
- DHCP: Dynamic Host Configuration Protocol: dynamically get address from as server
  - "Plug-and-play"

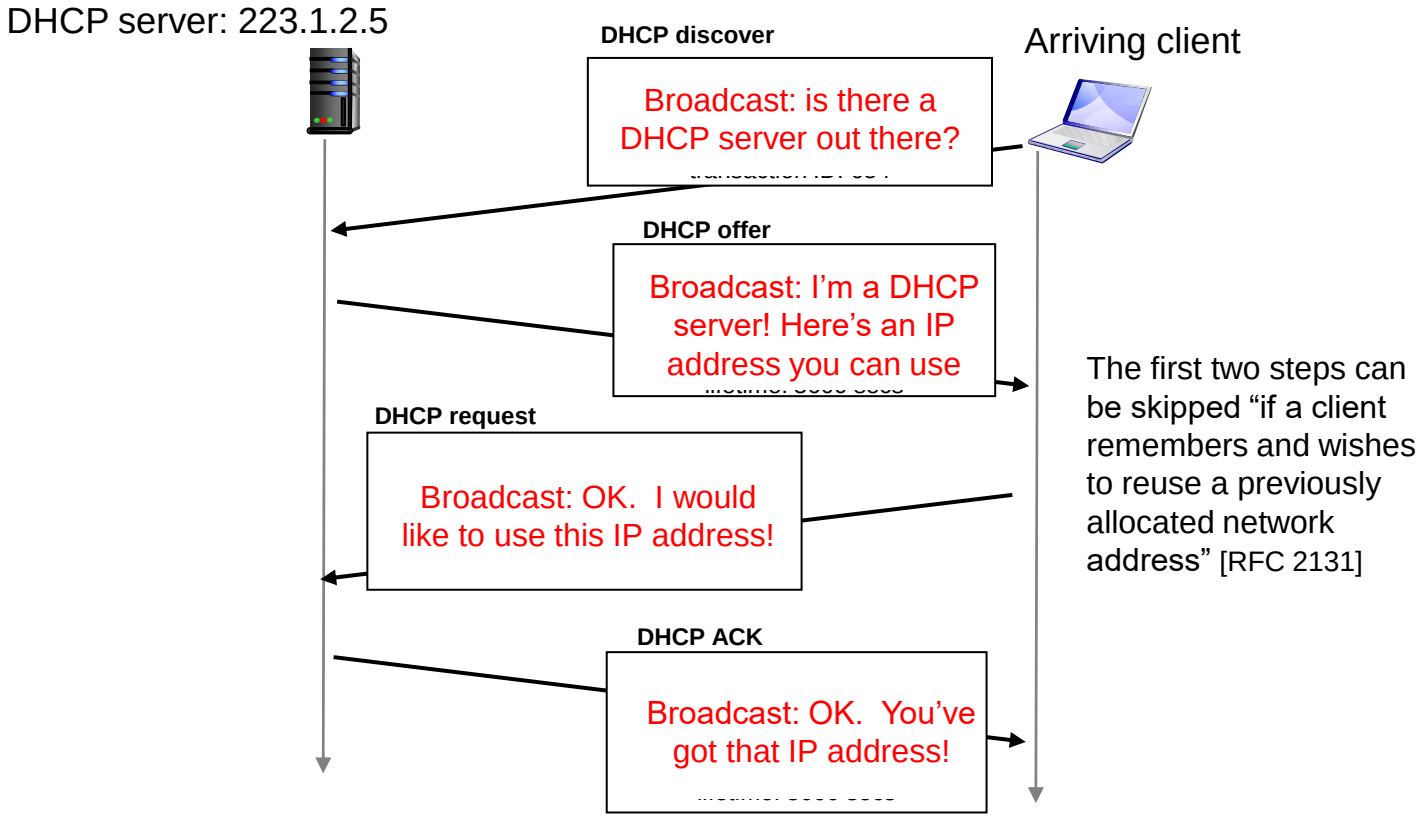
# DHCP Client-Server Scenario




Typically, DHCP server will be co-located in router, serving all subnets to which router is attached

arriving DHCP client needs
address in this network

# DHCP Client-Server Scenario




The first two steps can be skipped "if a client remembers and wishes to reuse a previously allocated network address" [RFC 2131]

# DHCP: Dynamic Host Configuration Protocol

DHCP overview:

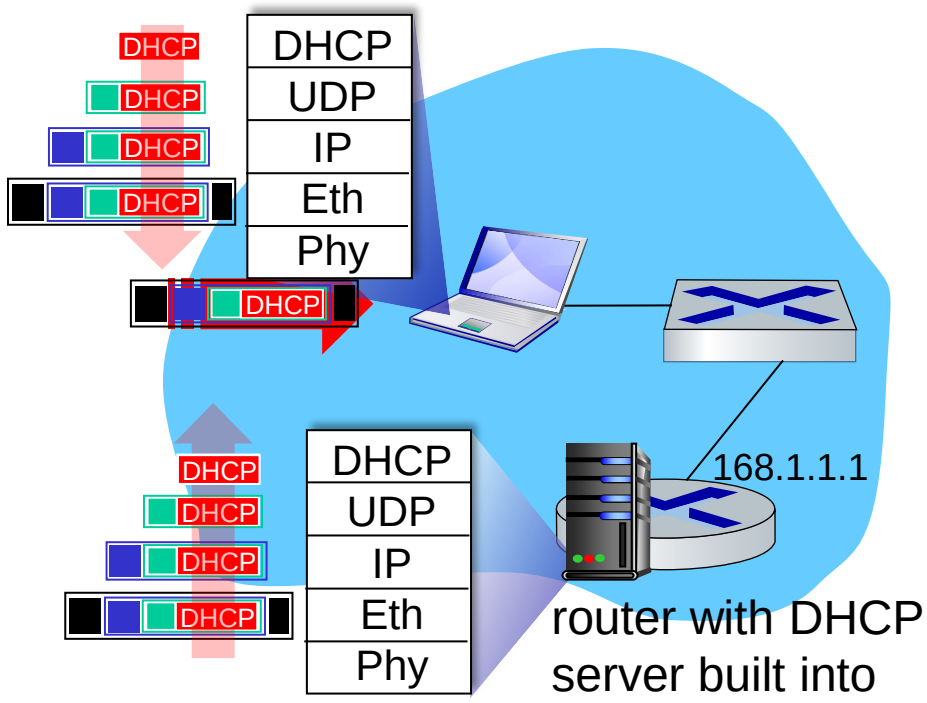
- Host broadcasts DHCP discover msg [optional]
- DHCP server responds with DHCP offer msg [optional]
- Host requests IP address: DHCP request msg
- DHCP server sends address: DHCP ack msg

# DHCP: More Than IP Addresses

DHCP can return more than just allocated IP address on subnet:

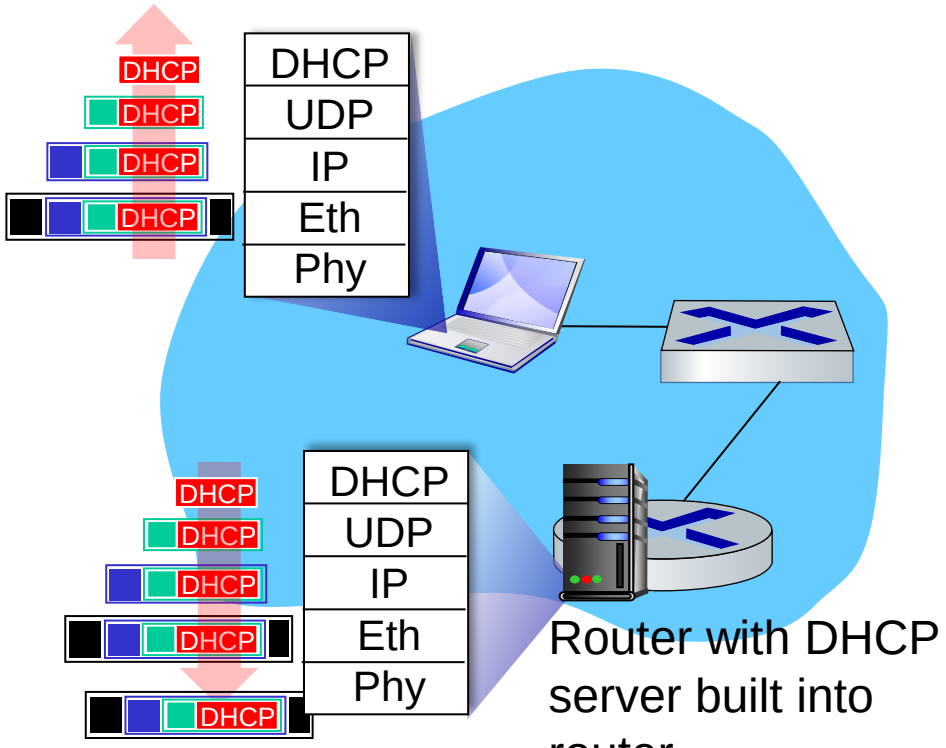
- Address of first-hop router for client
- Name and IP address of DNS sever
- Network mask (indicating network versus host portion of address)

# DHCP Example




- Connecting laptop will use DHCP to get IP address, address of first-hop router, address of DNS server
- DHCP REQUEST message encapsulated in UDP, encapsulated in IP, encapsulated in Ethernet
- Ethernet frame broadcast (dest: FFFFFFFFFFFF) on LAN, received at router running DHCP server
- Ethernet demux'ed to IP demux'ed, UDP demux'ed to DHCP

# DHCP Example




Router with DHCP server built into router

- DCP server formulates DHCP ACK containing client's IP address, IP address of first-hop router for client, name & IP address of DNS server
- Encapsulated DHCP server reply forwarded to client, demuxing up to DHCP at client
- Client now knows its IP address, name and IP address of DNS server, IP address of its first-hop router

# How Does Networks Get IP addresses?

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

A: Gets allocated portion of its provider ISP's address space

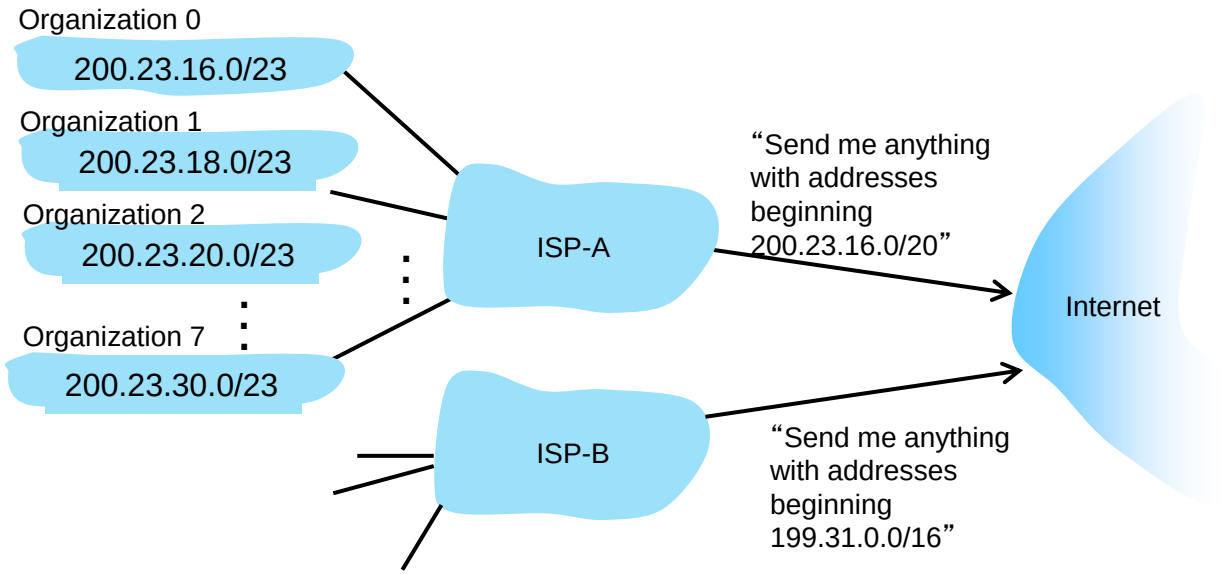
| | | | |
|---|---|---|---|
| ISP's block | 11001000 00010111 00010000 | 00000000 | 200.23.16.0/20 |

ISP can then allocate out its address space in 8 blocks:

| | | | |
|---|---|---|---|
| Organization 0 | 11001000 00010111 00010000 | 00000000 | 200.23.16.0/23 |
| Organization 1 | 11001000 00010111 00010010 | 00000000 | 200.23.18.0/23 |
| Organization 2 | 11001000 00010111 00010100 | 00000000 | 200.23.20.0/23 |
| ... | ..... | .... | .... |
| Organization 7 | 11001000 00010111 00011110 | 00000000 | 200.23.30.0/23 |

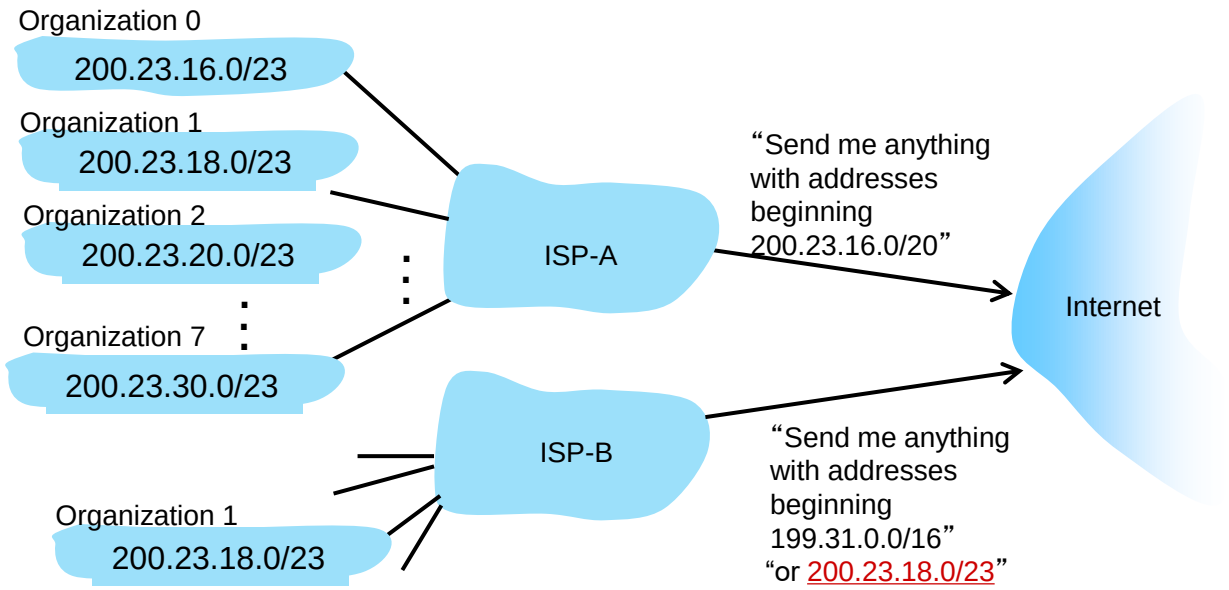
# Hierarchical Addressing: Route Aggregation

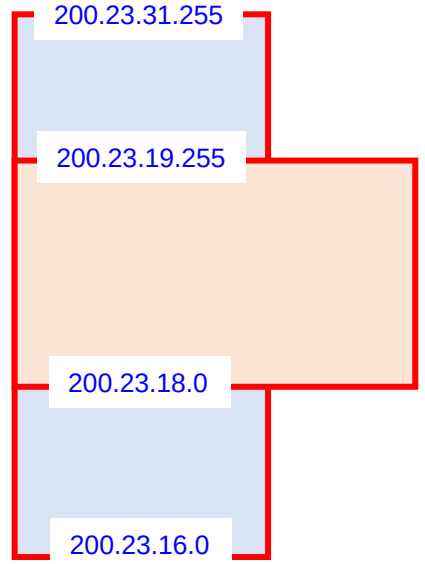
Hierarchical addressing allows efficient advertisement of routing information:



# Hierarchical Addressing: More Specific Routes

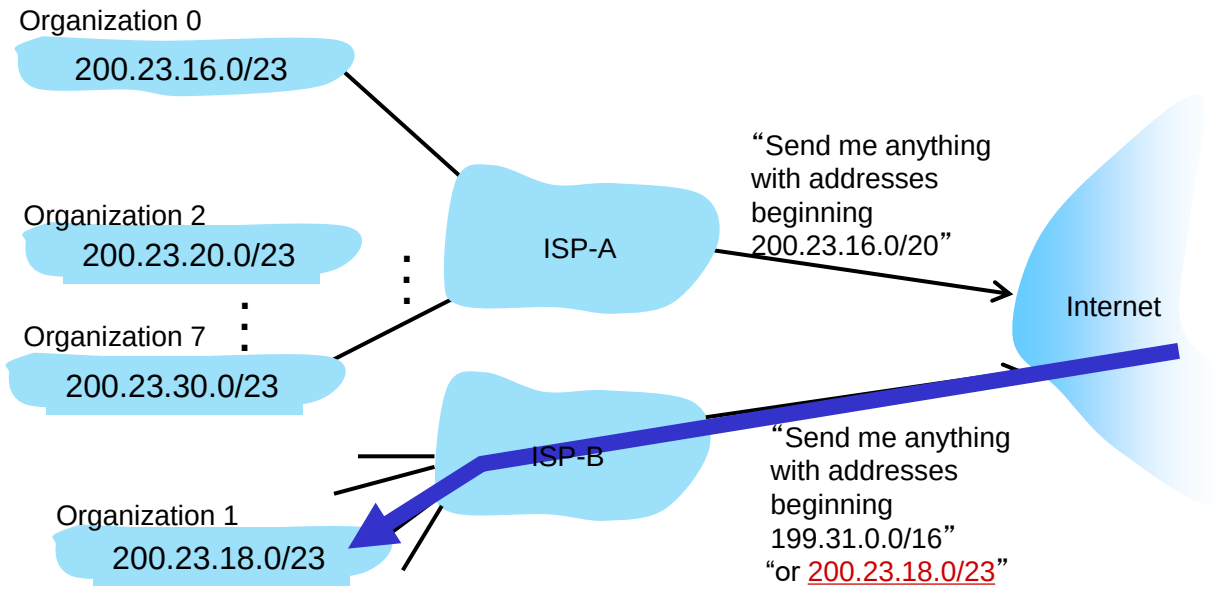
- Organization 1 moves from ISP-A to ISP-B
- ISP-B now advertises a more specific route to Organization 1

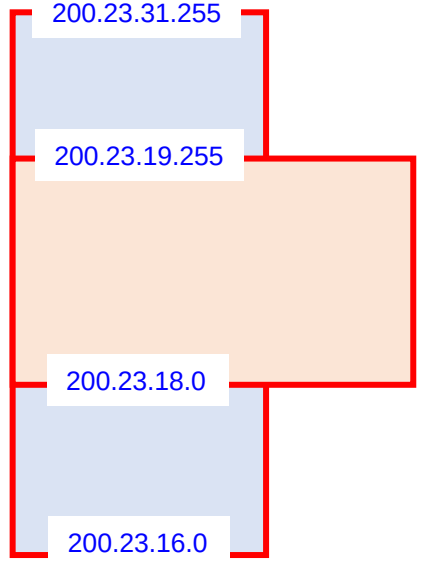


# Hierarchical Addressing: More Specific Routes

- Organization 1 moves from ISP-A to ISP-B
- ISP-B now advertises a more specific route to Organization 1

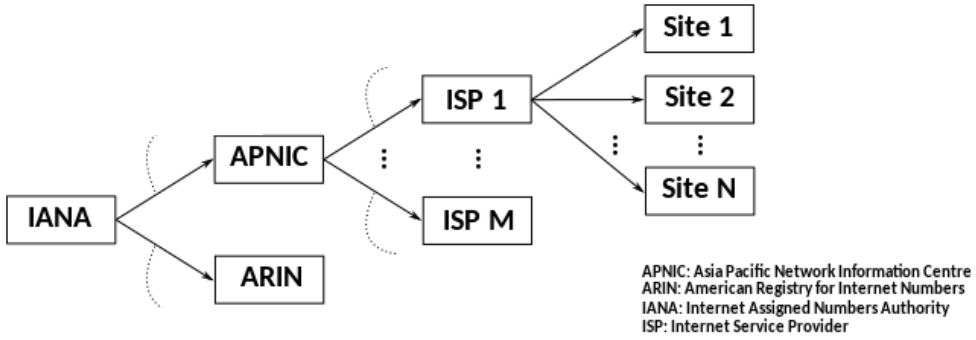


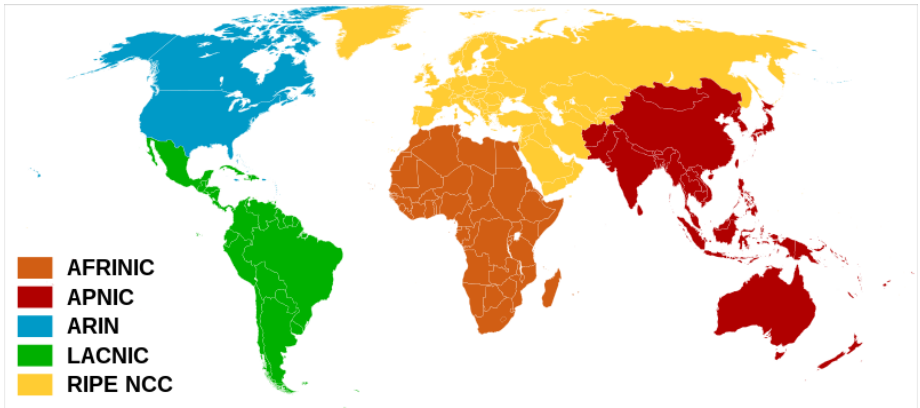
# How ISPs Get IP Addresses?

ICANN: Internet Corporation for Assigned  Names and Numbers http://www.icann.org/

- Allocates IP addresses, through 5 regional registries (RRs) (who may then allocate to local registries)
- Manages DNS root zone, including delegation of individual TLD (.com, .edu , …) management




Src: https://commons.wikimedia.org/wiki/File:IPv6_Prefix_Assignment_Example-en.svg

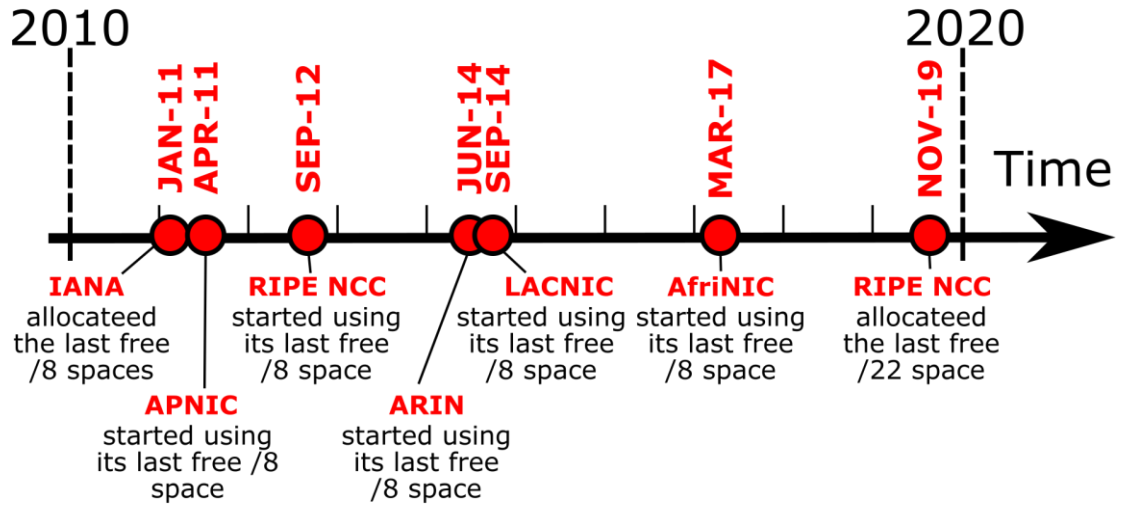



Src: https://commons.wikimedia.org/wiki/File:Regional_Internet_Registries_world_map.svg

# IPv4 Addressing: Concluding Remarks

Are there enough 32-bit IP addresses?

- ICANN allocated last chunk of IPv4 addresses to RRs in 2011
- NAT helps IPv4 address space exhaustion
- IPv6 has 128-bit address space




Src: https://commons.wikimedia.org/wiki/File:IPv4_IANA_and_RIR_exhaustion_time_line-en.svg