Class Notes

Preamble:

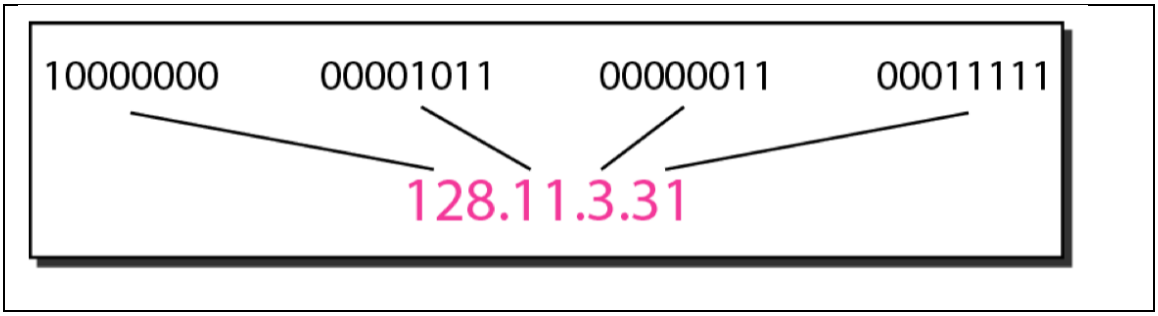
We are now starting our discussion that how IP addresses should be allocated to meet certain goals such as efficient allocation (no waste), contiguous block allocaton (to enable summarization) and to meet customers' IP demands.

(1) One of the main goals of the network layer is to provide host-to-host delivery, where these hosts could be anywhere on the network

(2) Nodes need some ID (with many qualities such as uniqueness, structured, fixed length for performance reasons etc.) so that network layer could effectively do its primary job (host to host delivery)

(3) IPv4 addresses are 32 bit long and application layer gets:
   a. Source address (its own address) either statically by admin or more commonly using DHCP protocol (we will study this protocol later in this chapter)
   b. Destination address is usually got using DNS service

(4) IPv4 addresses are often written as dotted quad, where we put a dot after each octet (8 bits) and consider them as a decimal number. We do so for quickly readability.



(5) IP addresses are strategically assigned by IANA -> Regional Registries -> ISPs -> Other customers because:
   a. We need summarization of IP addresses so that we could write one (or a few) prefixes of the IPs in the routers' forwarding table. Doing so is critical to keep the forwarding tables small so that routers could do address matching (for example: longest prefix match) quickly (at line rate). This summarization is called router aggregation. Doing so is critical to make IP data plane (and control plane) scalable and performant. If, each router could have tons of random IP addresses in its forwarding table, each router would need substantial memory for forwarding tables with special hardware for lookup (which is quite expensive in dollar terms).

Example: Without summarization a forwarding table of a router:

Destination Network	Next-Hop Router
172.20.16.0/24	R2
172.20.17.0/24	R2
172.20.18.0/24	R2
172.20.19.0/24	R2

Now after summarization:

Destination Network	Next-Hop Router
172.20.16.0/22	R2

Note: You need to be comfortable to move back and forth from Decimal numbers to Binary numbers (for IPv4) and Decimal <-> Binary <-> Hexadecimal conversions (for IPv6).

Bit Number	7	6	5	4	3	2	1	0
Bit Value (Weight)	128	64	32	16	8	4	2	1
Designation	*MSB*							*LSB*

MSB = Most sifnificant bit
LSB = Least significant bit

In above example:
172.20.16.0/24 = 10101100.00010100.00010000.00000000 (/24)
172.20.17.0/24 = 10101100.00010100.00010001.00000000 (/24)
172.20.18.0/24 = 10101100.00010100.00010010.00000000 (/24)
172.20.19.0/24 = 10101100.00010100.00010011.00000000 (/24)

We can summarize above as: 172.20.16.0/22

Optional Reading: Why route aggregation and summarization is needed.

Source: Google Gemini

Route aggregation and summarization are primarily done to **improve the scalability and efficiency of routing** on the internet.[1] By consolidating multiple specific routes

into a single, more general announcement, these techniques solve several critical network challenges.[2]

1. Reduce the Size of Routing Tables

This is the most significant reason. Imagine a mail carrier having to know the exact name of every single person in a city to deliver mail. It would be impossible. Instead, they just need to know the street name. Route summarization works on a similar principle.

- **Before Summarization:** A router might have dozens or even hundreds of specific entries for individual networks.[4] For example:
  - 192.168.1.0/24
  - 192.168.2.0/24
  - 192.168.3.0/24
  - ...and so on.
- **After Summarization:** These can be combined into a single, less specific route.
  - 192.168.0.0/16

This consolidation drastically **reduces the memory** required to store the routing table and the **CPU cycles** needed to look up a route, leading to faster and more efficient routers.[7] For large internet backbone routers that handle hundreds of thousands of routes, this is absolutely essential.

2. Improve Network Stability

Route summarization helps to **contain network instability**. When a specific network link goes up or down (an event called "flapping"), the change needs to be advertised to other routers.

- **Without Summarization:** If a single small network link flaps, an update is sent across a large part of the network, forcing many routers to recalculate their routing tables.[10] This can cause a cascade of updates, leading to network instability and high CPU usage on routers.
- **With Summarization:** If the flapping link is part of a summarized block, the instability is **hidden** from the rest of the network.[11] The summary route advertisement remains unchanged as long as at least one of the sub-networks within the summary is still reachable. This localization of issues prevents minor problems from affecting the entire network backbone.

3. Reduce Routing Protocol Overhead

Routers constantly communicate with each other, sending updates about the network topology.

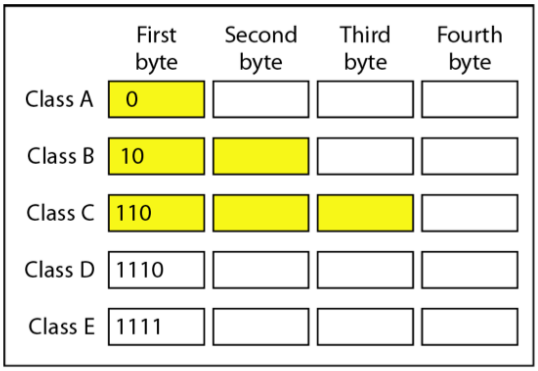
By summarizing routes, the number and size of these routing update packets are significantly reduced.[13] This cuts down on the amount of **bandwidth consumed by the routing protocol** itself, leaving more bandwidth available for actual user data.[14] This is particularly important on slower network links.

b. Since we have scarcity of IPv4 addresses, we need to carefully assign address ranges to meet the demand.

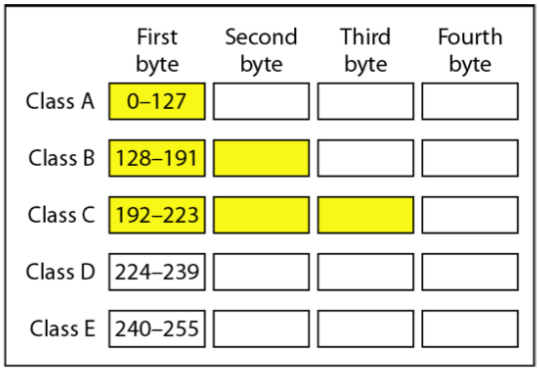
Historical Classful Division of IPv4 Address Range

Originally 32 bits of IPv4 were statically divided as different classes of address for different usecases.

In classful addressing, the address space is divided into five classes: A, B, C, D, and E.



a. Binary notation



b. Dotted-decimal notation

Exercise:

Find the class of each address.

a. 00000001 00001011 00001011 11101111

b. 11000001 10000011 00011011 11111111

c. 14.23.120.8

d. 252.5.15.111

Solution

a. The first bit is 0. This is a class A address.

b. The first 2 bits are 1; the third bit is 0. This is a class C address.

c. The first byte is 14; the class is A.

d. The first byte is 252; the class is E.

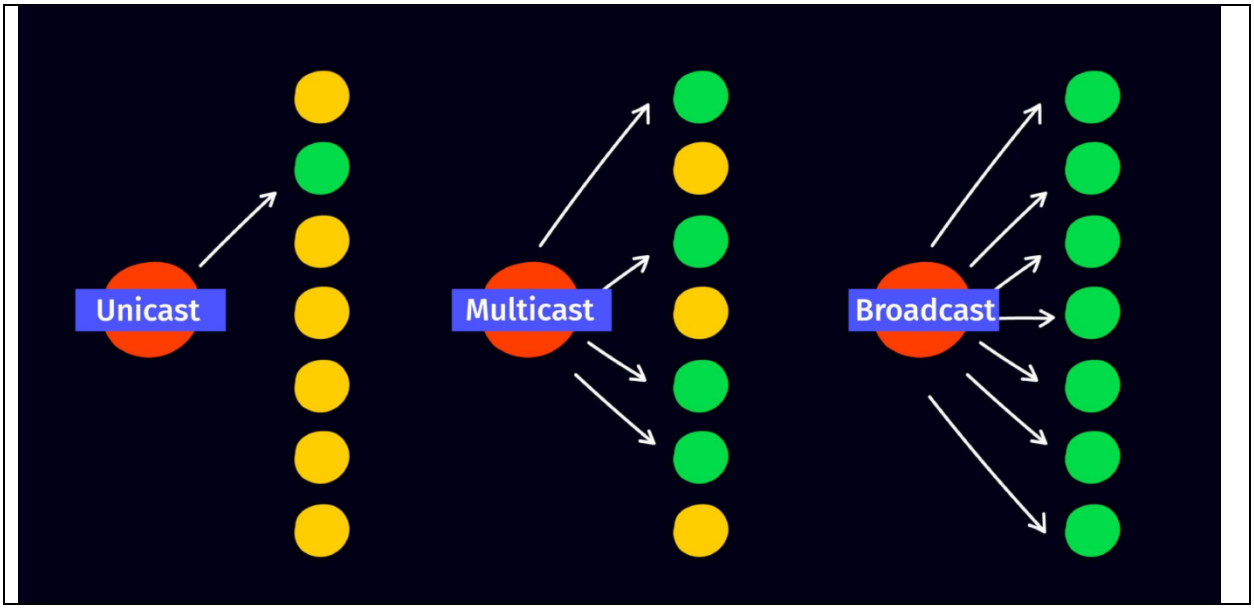
Number of blocks and block sizes in classful IPv4 addresses

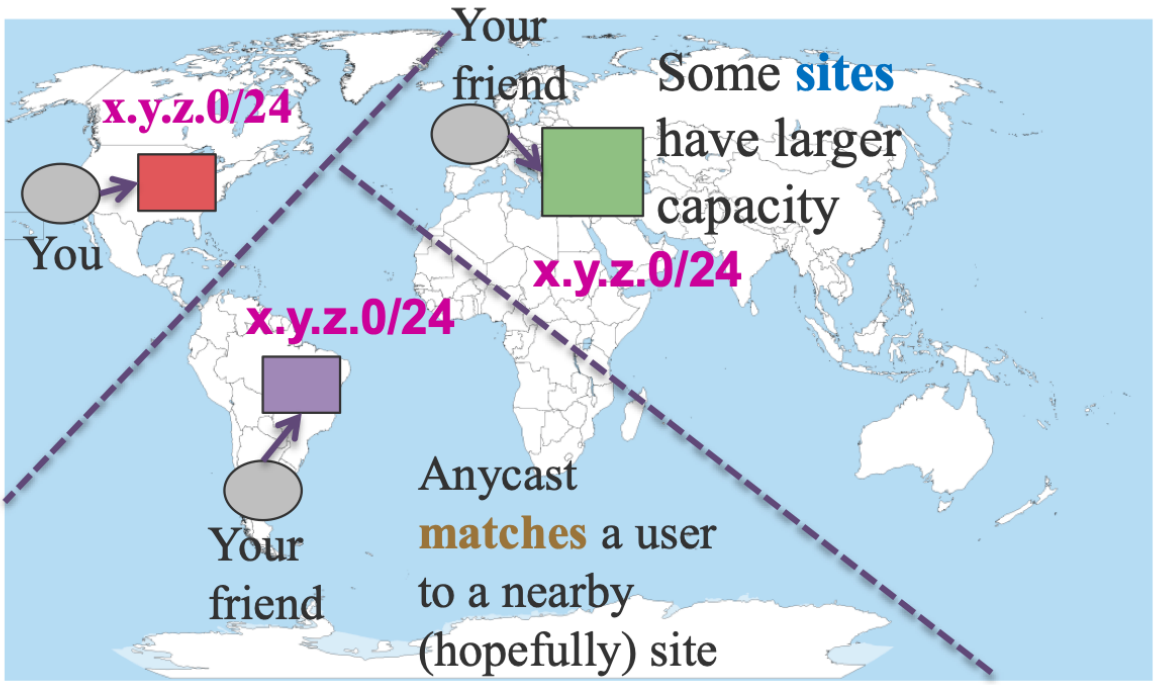
Class	*Number of Blocks*	*Block Size*	*Application*
A	128	16,777,216	Unicast
B	16,384	65,536	Unicast
C	2,097,152	256	Unicast
D	1	268,435,456	Multicast
E	1	268,435,456	Reserved

Unicast, Multicast, Broadcast, and Anycast Communications

Communication Type	Definition	Analogy	Typical Use Case
Unicast	**One-to-One.** A single sender transmits data to a single, specific receiver identified by a unique IP address.	Mailing a letter to a specific home address. 📮	Browsing a website, downloading a file, or sending an email.
Multicast	**One-to-Many.** A single sender transmits data to a group of interested receivers who have "subscribed" to the group.	A magazine subscription sent only to people who have signed up. 📰	Streaming a live video event, online stock tickers, or online gaming.
Broadcast	**One-to-All.** A single sender transmits data to all possible receivers within the same local network segment.	Shouting in a crowded room; everyone in the room hears it. 📣	A device asking "Is there a DHCP server here?" to get an IP address.
Anycast	**One-to-Nearest.** A single sender transmits data to an address that is assigned to multiple devices. The network routes the data to the topologically closest receiver.	Calling a company's single "800" number and being automatically connected to the nearest call center. 📞	DNS root servers or Content Delivery Networks (CDNs) serving content from the closest server to reduce latency.

Source: Google Gemini





Anycast **divides** the world into catchments.

In the context of anycast, a **catchment** is the specific set of users, defined by geography or network topology, whose traffic is routed to a single, particular anycast server node.

Think of it as the "service area" for one specific location in an anycast network.

Source: https://www.usenix.org/system/files/sec22_slides-rizvi.pdf

IPv4 Reservations for Different Purposes

Each Local Area Network (LAN) can use private IP addresses for its internal use. Such IP address is not routable on the global Internet.
Such addresses are a way to reduce IPv4 address scarcity.
Commonly used by organizations (school, data centers, enterprises, ISP etc.)
What if a host having a private IP wants to talk to some host on the Internet?

- We need a technology called NAT.
- We will study NAT later on.

Address Block (CIDR)	Full Address Range	Number of Addresses	Purpose and Designation
0.0.0.0/8	0.0.0.0 – 0.255.255.255	16,777,216	"This Network": Used by hosts to refer to their own network, often during the boot process (e.g., DHCP).
10.0.0.0/8	10.0.0.0 – 10.255.255.255	16,777,216	Private Network: Used for local communications within large private networks (RFC 1918).
100.64.0.0/10	100.64.0.0 – 100.127.255.255	4,194,304	Shared Address Space: Used for Carrier-Grade NAT (CGNAT), allowing ISPs to assign private addresses to customers.
127.0.0.0/8	127.0.0.0 – 127.255.255.255	16,777,216	Loopback: Used by a host to send traffic to itself for testing (e.g., 127.0.0.1 is localhost).
169.254.0.0/16	169.254.0.0 – 169.254.255.255	65,536	Link-Local Addresses: Automatically self-assigned by devices (APIPA) when a DHCP server is not available.
172.16.0.0/12	172.16.0.0 – 172.31.255.255	1,048,576	Private Network: Used for local communications within medium-sized private networks (RFC 1918).
192.0.0.0/24	192.0.0.0 – 192.0.0.255	256	IETF Protocol Assignments: Reserved for various internet protocols.
192.0.2.0/24	192.0.2.0 – 192.0.2.255	256	Documentation (TEST-NET-1): Reserved for use in examples and documentation.
192.88.99.0/24	192.88.99.0 – 192.88.99.255	256	Reserved: Formerly used for IPv6 to IPv4 relay but now largely obsolete.
192.168.0.0/16	192.168.0.0 – 192.168.255.255	65,536	Private Network: Used for local communications, most commonly in home and small office networks (RFC 1918).
198.18.0.0/15	198.18.0.0 – 198.19.255.255	131,072	Network Interconnect Device Benchmark Testing: Reserved for testing network devices.
198.51.100.0/24	198.51.100.0 – 198.51.100.255	256	Documentation (TEST-NET-2): Reserved for use in examples and documentation.
203.0.113.0/24	203.0.113.0 – 203.0.113.255	256	Documentation (TEST-NET-3): Reserved for use in examples and documentation.
224.0.0.0/4	224.0.0.0 – 239.255.255.255	268,435,456	Multicast (Class D): Used for sending data to groups of interested receivers simultaneously.
240.0.0.0/4	240.0.0.0 – 255.255.255.254	268,435,455	Reserved for Future Use (Class E): Historically reserved and not used on the public internet.
255.255.255.255/32	255.255.255.255	1	Limited Broadcast: Used to send a message to all devices on the local network segment.

Address Range	Purpose
`0.0.0.0/8`	Used for the "current network" (a source address only).
`100.64.0.0/10`	Reserved for Carrier-Grade NAT (CGNAT), used by ISPs to assign private addresses to their customers.
`127.0.0.0/8`	**Loopback addresses** for testing the network interface card and software on a local machine.
`169.254.0.0/16`	**Link-Local Addresses**, also known as Automatic Private IP Addressing (APIPA). These are self-assigned by devices when they cannot obtain an IP address from a DHCP server.
`192.0.2.0/24`, `198.51.100.0/24`, `203.0.113.0/24`	Reserved for **documentation and examples** to be used in tutorials and manuals.
`224.0.0.0/4`	The entire **Class D range** is designated for **multicast** traffic.
`240.0.0.0/4`	The entire **Class E range** is reserved for **future and experimental use**.
`255.255.255.255`	The **broadcast address**, used to send a message to all devices on the local network.

Classless Interdomain Routing (CIDR) Way of IP Allocation

CIDR is a method for more efficiently allocating and routing IPv4 addresses by allowing variable-length subnet masks, instead of the rigid, predefined classes (A, B, and C). This provides flexible network sizing to exactly match an organization's needs.

The move from classful to classless addressing occurred to combat the rapid exhaustion of IPv4 addresses and to reduce the size of internet routing tables. The old classful system wasted many addresses by assigning entire large blocks to organizations that needed far fewer, which also forced routers to keep track of excessive individual network entries.

Subnet Mask:

A subnet mask is a 32-bit number that helps a device identify which part of an IP address represents the network and which part represents the host device. It is essential for routing network traffic efficiently by determining whether a destination is on the local network or a different one.

Example:

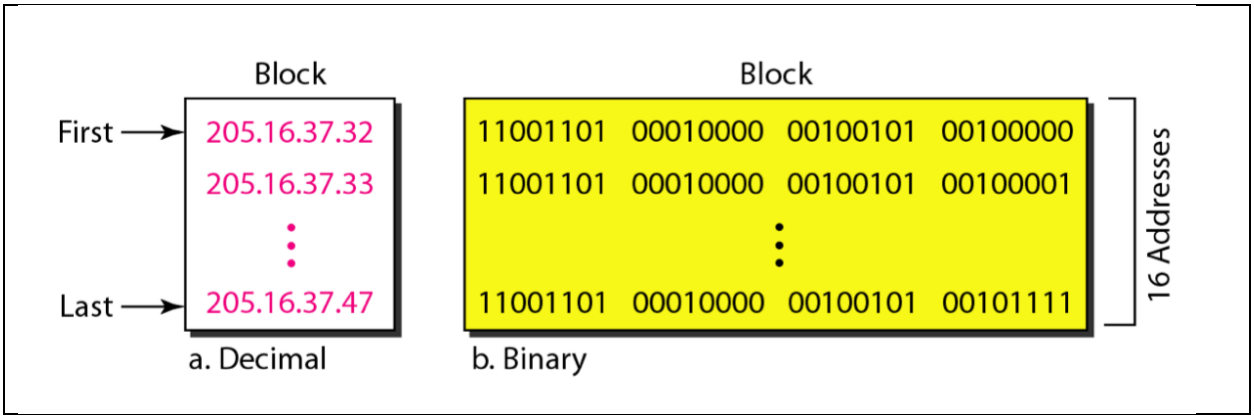
Consider the following common configuration for a home network:
IP Address: 192.168.1.10
Subnet Mask: 255.255.255.0
When converted to binary, the subnet mask
of 255.255.255.0 is 11111111.11111111.11111111.00000000.

You should be comfortable finding network part of the address using explicit subnet mask or slash (CIDR) notation.

Rules to follow while allocating IP addresses:

(1) IP addresses allocated should be contiguous.
(2) The number of IPs allocated should be some power of 2.
(3) The first address of the IP block should be divisible by the block size (number of IPs).

Example: A block of 16 addresses given to a small organization



Rule 1 followed: All 16 addresses are contiguous
Rule 2 followed: 16 is a power of 2 (2^4 = 16)
Rule 3 followed: First address of the block is 205.16.37.32 =
11001101 00010000 00100101 0010**0000** = CD102520 = 3440,387,360

3440,387,360/16 = 215,024,210

Rule 1 and Rule 2 are about efficient allocation of IPs
Rule 3 is so that we could summarize the whole block correctly.
Above block can be summarized as: 205.16.37.32 /28
Subnet Mask = 255.255.255.15

Example:

A block of addresses is granted to a small organization. We know that one of the addresses is 205.16.37.39/28. What is the first address in the block?

Solution

The binary representation of the given address is

11001101 00010000 00100101 00100111

If we set 32−28 rightmost bits to 0, we get

11001101 00010000 00100101 0010000

or

205.16.37.32.

The first and the last IP addresses of an allocated block have special significance:

First IP in our example above is: 205.16.37.32
Last IP in the block is: 205.16.37.47

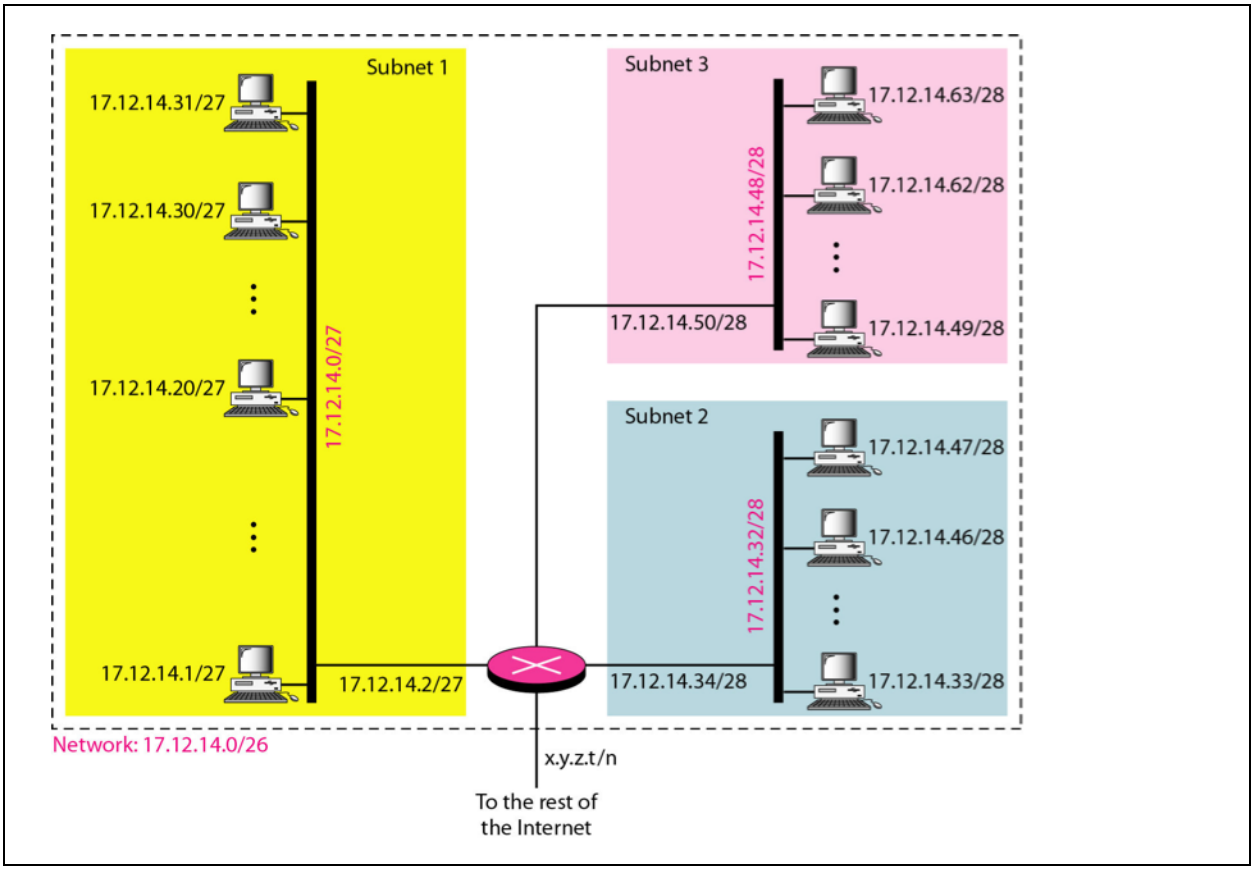
We know it is a /28 block => Host part has 32 – 28 = 4 bits
Last 4 bits (least significant) of the first address = 0000 (last full octet = 00100000)
Last 4 bits (least significant) of the last address = 1111 (last full octet = 00101111)

First IP of the block **represents the whole block**. It is stored as a prefix in routers' forwarding table and it is used when routers announce routes to other routers.
Last IP is a broadcast address of the subnet when used for a packet with that destination IP, that IP packet will go to all hosts connected to that specific subnet.

Due to above reasons no host in the subnet is assigned the first of the last address of the block. (**Note**: In the exam if you are asked to make a subnet with say 16 addresses, then you usually include first and the last in the count. But if question asks for 16 addresses for the hosts, then you

need a block with 16 + 2 = 18 addresses, which will not be a power of 2 and you might need to give out a block of 32 addresses.)

Exercise: Configuration and addresses in a subnetted network:



Find out Rules 1, 2, and 3 are being followed in subnet 1, 2 and 3 above. (I will leave it as an exercise for you to do.)
Note 1: For a point-to-point (say connecting two routers), they need to be part of the same subnet.
Note 2: See above how different links of the router are part of different subnets.

What will be the summarized prefix that router will announce to the rest of the Internet:

Subnet 1 = 17.12.14.0/27 (last octet = 0000 0000)
Subnet 2 = 17.12.14.32/28(last octet = 0010 0000)
Subnet 3 = 17.12.14.48/28(last octet = 0011 0000)

We can summarize above as: 17.12.14.0 /26 . Note that 17.12.14.0/27 has two /28 blocks in it. So 32 addresses in subnet 1, 16 each in subnet 2 and subnet 3. Total becomes 32 + 16 + 16 = 64 Six bits of the host will all be consumed fully for these subnets.

Just like subnetting, supernetting (aggregation or summarization) also has the similar rules:

The Rules of Engagement: Prerequisites for Summarization

A valid and accurate summary route can only be created if a set of strict criteria is met. These rules are not arbitrary but are direct consequences of the binary mathematics that define a network prefix.

1. **The networks must be contiguous:** The address blocks being summarized must form an unbroken, adjacent range. There can be no gaps between the networks.
2. **The number of networks must be a power of two:** A summary can only be created for a group of 2, 4, 8, 16, etc., networks. This is because each bit removed from the network prefix doubles the size of the address block being represented.
3. **The first network address must align on a proper binary boundary:** Specifically, the first address in the block must be evenly divisible by the total size of the summarized range. This ensures the summary route itself represents a valid, natural network block.

Traditionally super netting is used to combine many /24 blocks to make a bigger block.
Example:

A company needs 1000 addresses. Which of the following set of class C blocks can be used to form a supernet for this company?

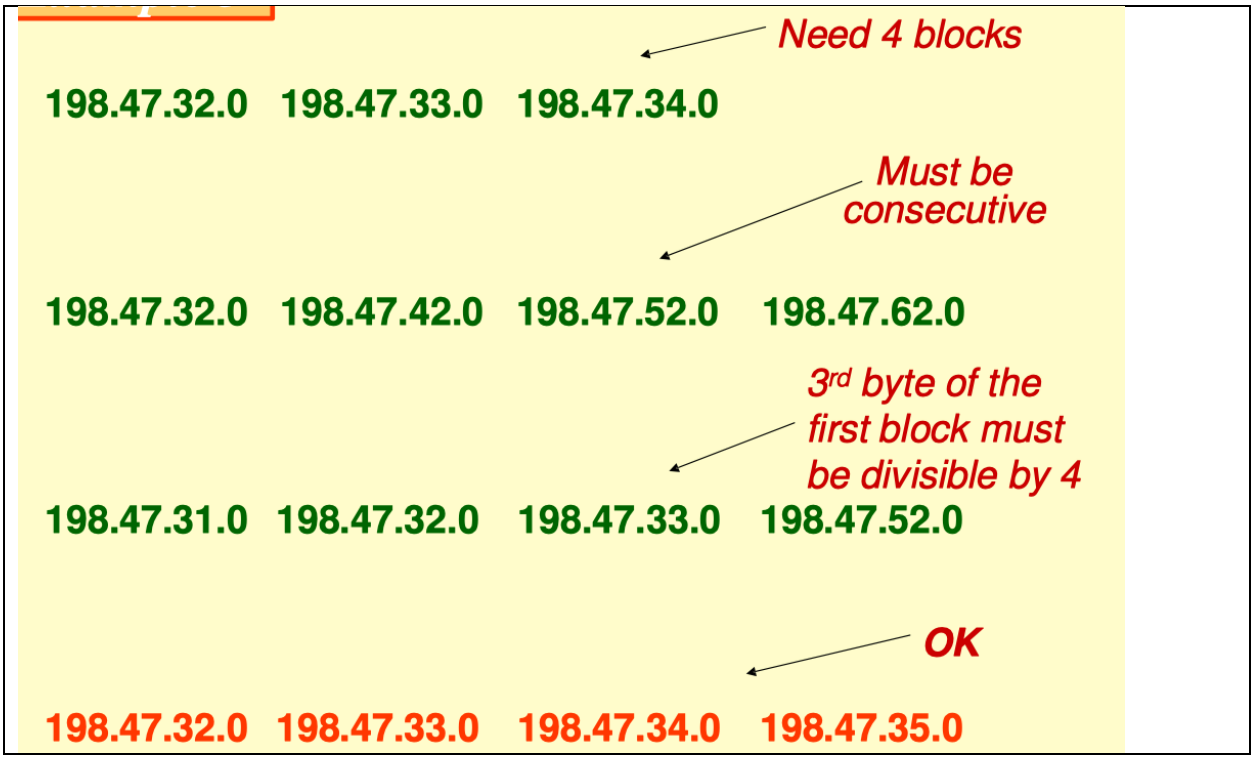
198.47.32.0 198.47.33.0 198.47.34.0

198.47.32.0 198.47.42.0 198.47.52.0 198.47.62.0

198.47.31.0 198.47.32.0 198.47.33.0 198.47.52.0

198.47.32.0 198.47.33.0 198.47.34.0 198.47.35.0

Solution:



More Examples:

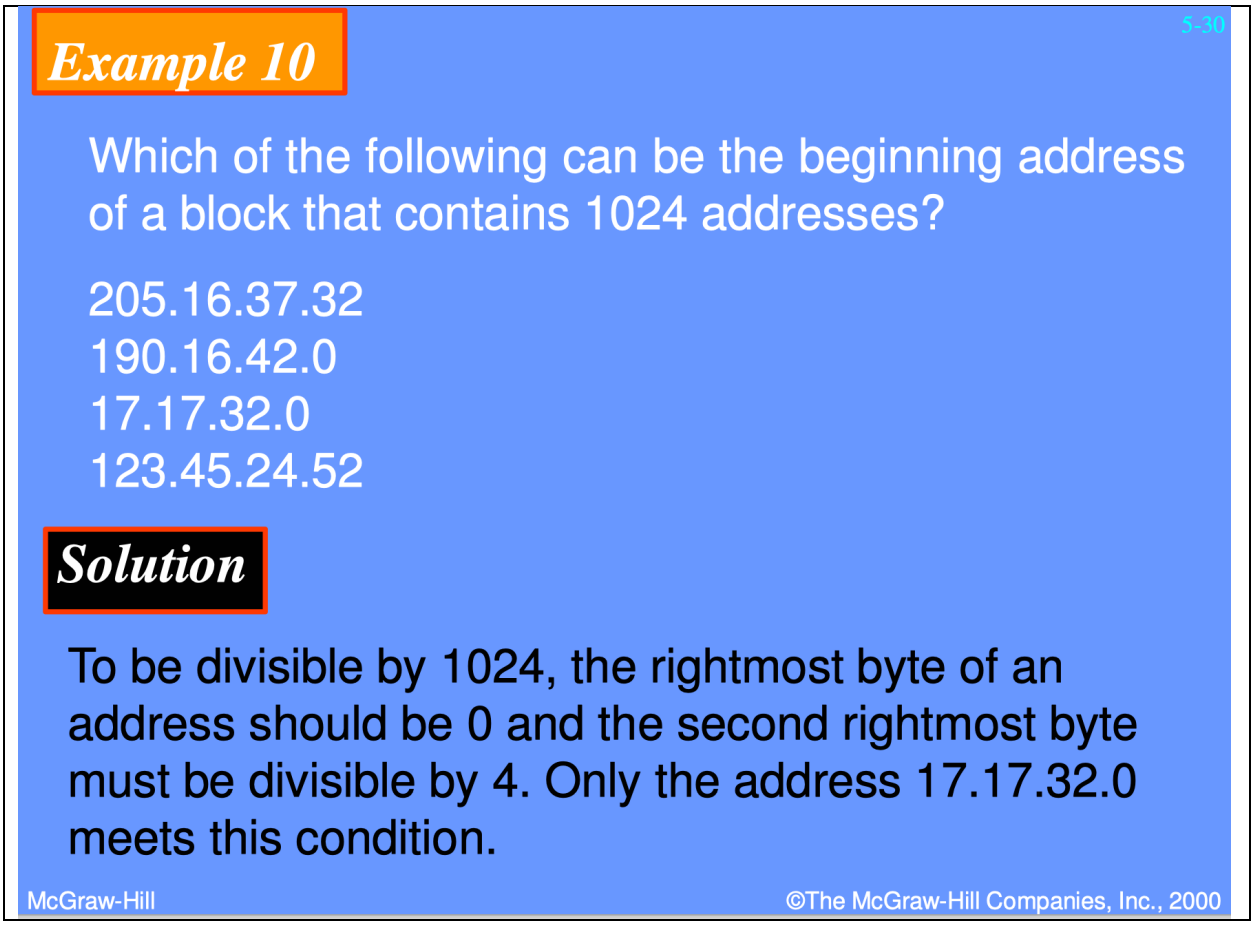
Example 9

Which of the following can be the beginning address of a block that contains 16 addresses?

205.16.37.32
190.16.42.44
17.17.33.80
123.45.24.52

Solution

The address 205.16.37.32 is eligible because 32 is divisible by 16. The address 17.17.33.80 is eligible because 80 is divisible by 16.



Variable Length Subnet Masks

New Rule: Start IP allocation starting with the largest demand.
Why? To avoid fragmenting the IP address space and to keep it contiguous

Example:

An ISP is granted a block of addresses starting with 190.100.0.0/16 (65,536 addresses). The ISP needs to distribute these addresses to three groups of customers as follows:

a. The first group has 64 customers; each needs 256 addresses.

b. The second group has 128 customers; each needs 128 addresses.

c. The third group has 128 customers; each needs 64 addresses.

Design the subblocks and find out how many addresses are still available after these allocations.

Solution:

First sort the demands in descending order.
First group's demand: = 64 * 256 = 16384
Second group's demand: 128 * 128 = 16384
Third group's demand: 128 * 64 = 8192

We can start allocation from first group, then second (we can flip the previous two because both have equal demand), and then the third one.

Group 1

For this group, each customer needs 256 addresses. This means that 8 (log2 256) bits are needed to define each host. The prefix length is then 32 − 8 = 24. The addresses are

1st Customer:	*190.100.0.0/24*	*190.100.0.255/24*
2nd Customer:	*190.100.1.0/24*	*190.100.1.255/24*
. . .		
64th Customer:	*190.100.63.0/24*	*190.100.63.255/24*

Total = 64 × 256 = 16,384

Group 2

For this group, each customer needs 128 addresses. This means that 7 (log2 128) bits are needed to define each host. The prefix length is then 32 − 7 = 25. The addresses are

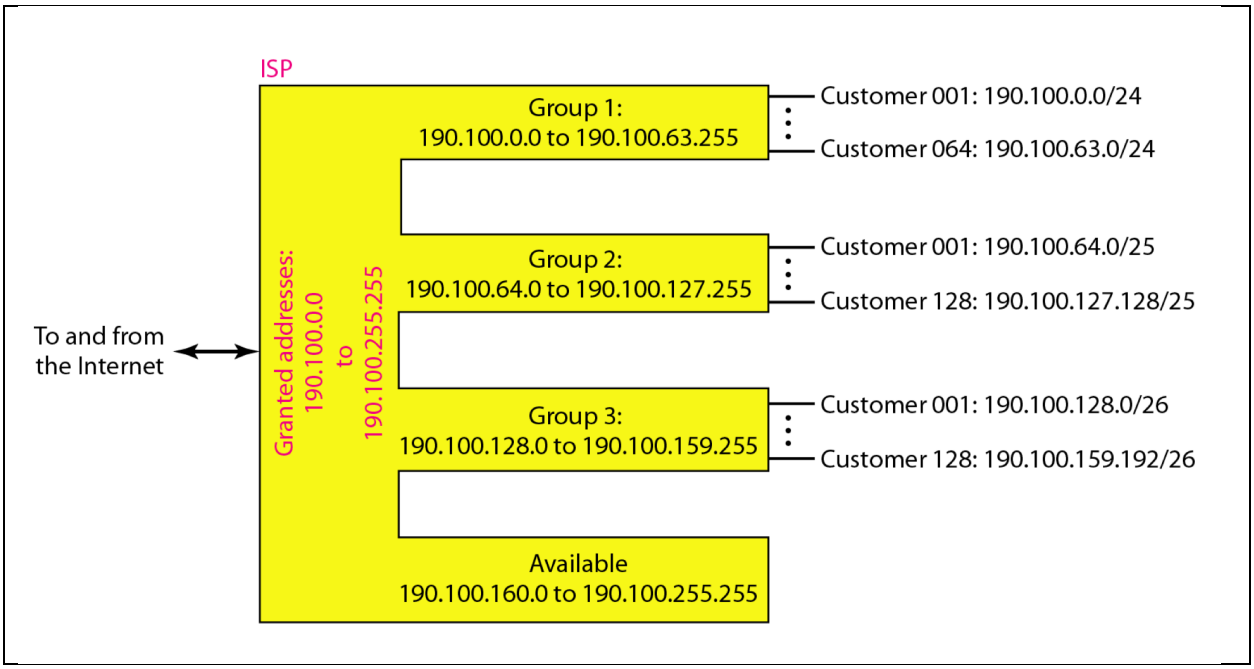
1st Customer:	*190.100.64.0/25*	*190.100.64.127/25*
2nd Customer:	*190.100.64.128/25*	*190.100.64.255/25*
. . .		
128th Customer:	*190.100.127.128/25*	*190.100.127.255/25*
Total = 128 × 128 = 16,384		

Group 3

For this group, each customer needs 64 addresses. This means that 6 ($log_2 64$) bits are needed to each host. The prefix length is then 32 − 6 = 26. The addresses are

1st Customer:	*190.100.128.0/26*	*190.100.128.63/26*
2nd Customer:	*190.100.128.64/26*	*190.100.128.127/26*
. . .		
128th Customer:	*190.100.159.192/26*	*190.100.159.255/26*
Total = 128 × 64 = 8192		

Number of granted addresses to the ISP: 65,536
Number of allocated addresses by the ISP: 40,960
Number of available addresses: 24,576



Can an IP packet be forwarded using layer-2 (local delivery)?

When a host want to send an IP packet to the destination IP, sending host needs to figure out if destination is part of its subnet or not. If destination IP is part of sender's subnet, then sender can deliver the IP packet directly using layer-2 technology (no need to forward such a packet to the next hop router).

> When a host with IP address **IP1** needs to send an IP packet to **IP2**, it first checks whether **IP2** is on the same network (subnet) as itself — that is, whether the **network part** of IP2 matches the network part of IP1 (using the subnet mask).
>
> - **"If they are on the same network**, the host knows that IP2 is **directly reachable** at the **data link layer (Layer 2)**."
> - "It then uses **ARP (Address Resolution Protocol)** to find IP2's MAC address and sends the frame directly to it."
> - **"If they are on different networks**, the packet must be sent to a **default gateway (router)** instead, which handles forwarding to other networks."

💡 **Example:**

Suppose:

- IP1 = 192.168.10.5
- IP2 = 192.168.10.20
- Subnet mask = 255.255.255.0

Then:

- Network part for both = 192.168.10
- ✅ Same network → can communicate directly via Layer 2 (Ethernet frame, ARP lookup).

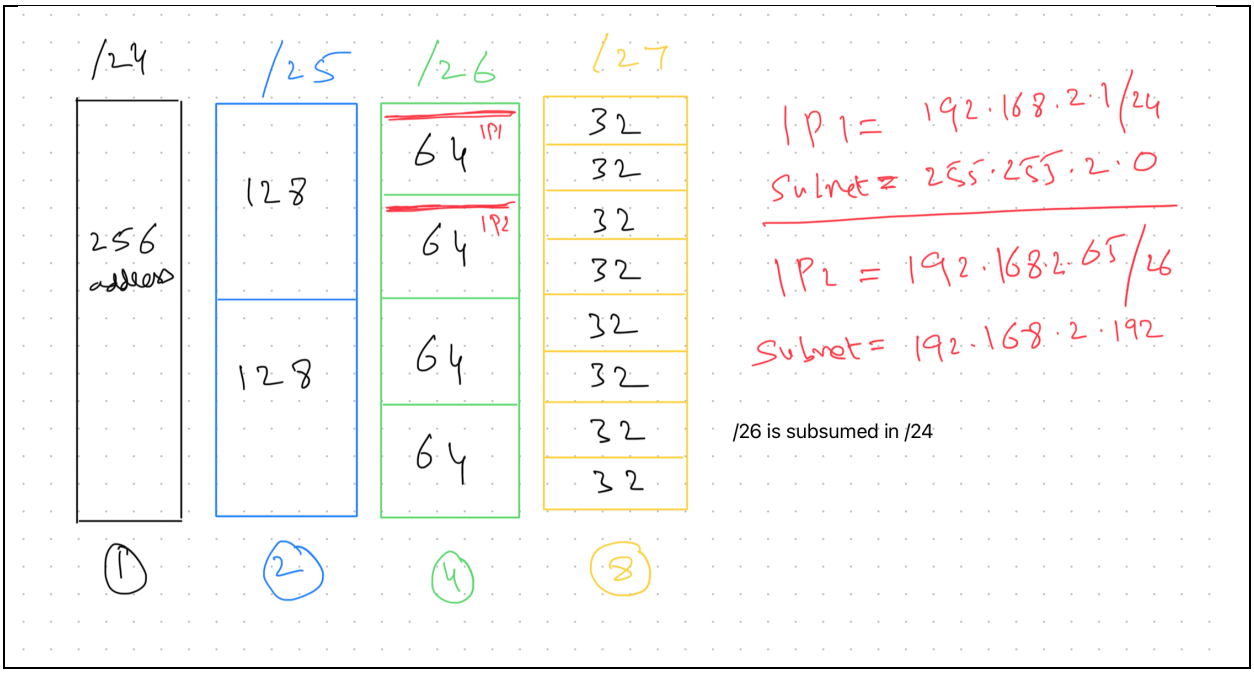
But if:

- IP2 = 192.168.20.7
  → Different network → must send to router (default gateway).

Note in above algo, source host does not know the subnet mask of the destination.
Some of the rules we learned above for subnetting / supernetting ensure that a host could reliably tell if the destination is on the same subnet or not.

Example:

What if IP2 has a /26 address while IP1 is /24 address where first 24 bits of IP2 network happens to be same as IP1? IP2 is on a different subnet, but due to above algo, IP1 host will wrongly assume that IP2 is on the local network.



Exercise:

A supernet has a first address of **205.16.32.0** and a supernet mask of **255.255.248.0**. A router receives three packets with the following destination addresses:

```
205.16.37.44
205.16.42.56
205.17.33.76
```

Which packet belongs to the supernet?

Solution:

Solution

SA =205.16.32.0

```
205.16.37.44 AND 255.255.248.0  ✎ 205.16.32.0

205.16.42.56 AND 255.255.248.0  ✎ 205.16.40.0

205.17.33.76 AND 255.255.248.0  ✎ 205.17.32.0
```

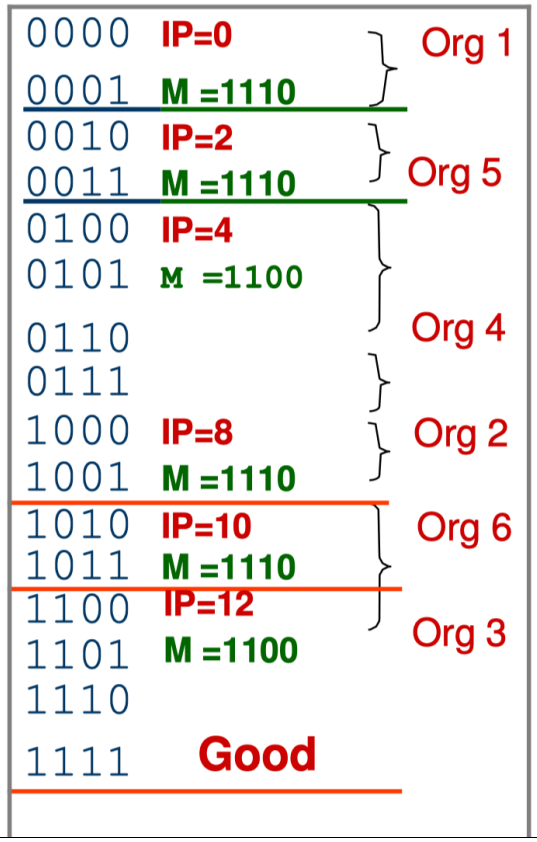
Only the first address belongs to this supernet.

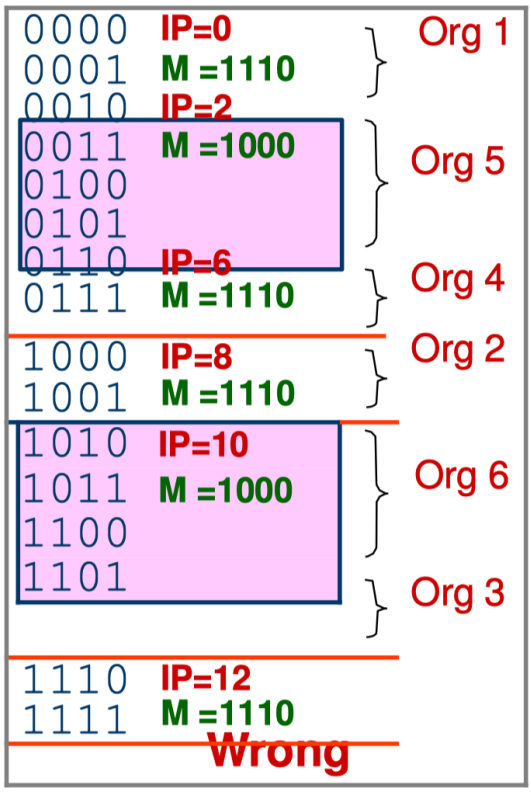
```
00100101 (37)          00101010 (42)
11111000 (248)         11111000 (248)
00100000 (32)          00101000 (40)

00100001 (33)
11111000 (248)
00100000 (32)
```

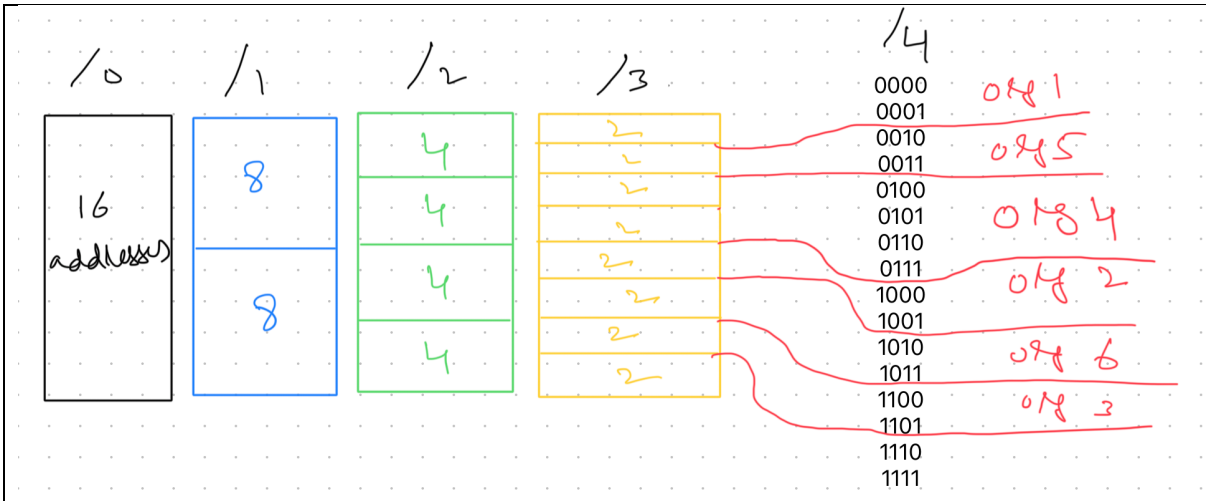
← The third byte of the third IP address even doesn't have to be AND-ed since the second byte is not 16.

Example:

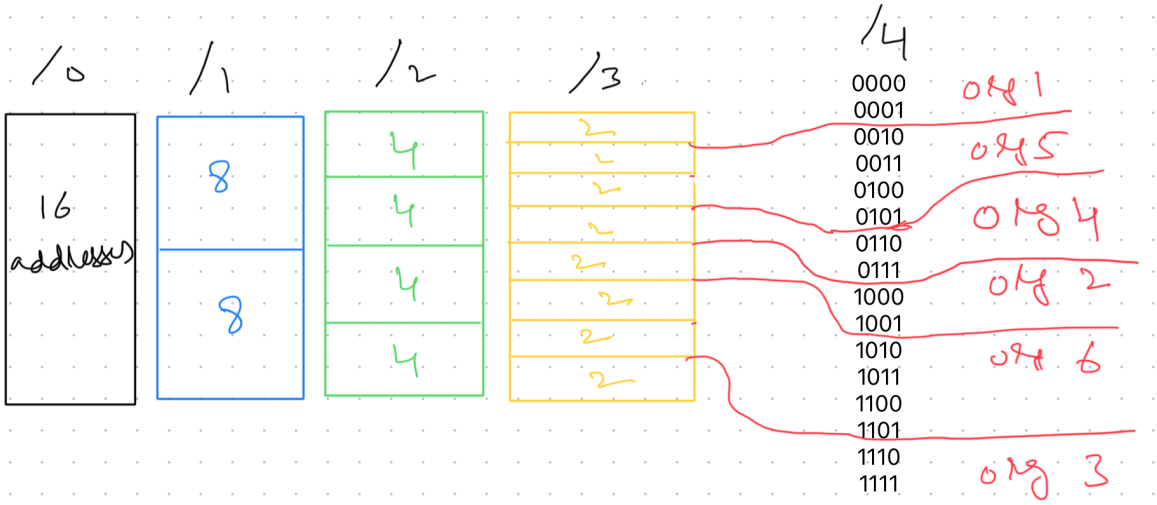
Suppose an IP space of only 16 addresses and six organizations with blocks of sizes 2 or 4.







Explanation for the left side (good)



Explanation for the left side (wrong)

More Practice:

Since subnetting and associated concepts are part of few certifications (such as CCNA), folks have come up with tricks to quickly solve such problems. One such series is the following by practical networking. They have practice questions and its evaluation on their website as well.

What is subnetting: https://youtu.be/BWZ-MHIhqjM?si=1ylTJ9FCcdcZLSQm
Subnetting cheat sheet: https://youtu.be/ljS07YTEJ2I?si=vpPNZ-tF3xP13XJm
Solve subnetting problem in 60 seconds: https://youtu.be/5-wlfAdcmFQ?si=Azog7kRR8A3FrTfj
Practice examples: https://youtu.be/SM0kdVfhxZ0?si=Hz_AXCHdKgShKnwQ
Time-saving tricks: https://youtu.be/Pj_ubhpi0WQ?si=vE_CoFVSa3U0E6ah
More subnetting in 3rd octet: https://youtu.be/p-6BVEbb9q0?si=S9L_FoU0b8v6x5F5
Subnetting in first 2 octets: https://youtu.be/OQ-r_IfeB8c?si=IO2R-B6Um4E1s9jh
Fixed-length subnet mask (FLSM): https://youtu.be/F05sDLXOFh8?si=WuRXa0zeDktl9DZH
Variable-length subnet mask (VLSM):
https://youtu.be/amKyfbg5G2Q?si=09zMZE1LJPkmXURu
Supernetting: https://youtu.be/Q4MArJTbUwk?si=KIhcV8mv81J3Eyag