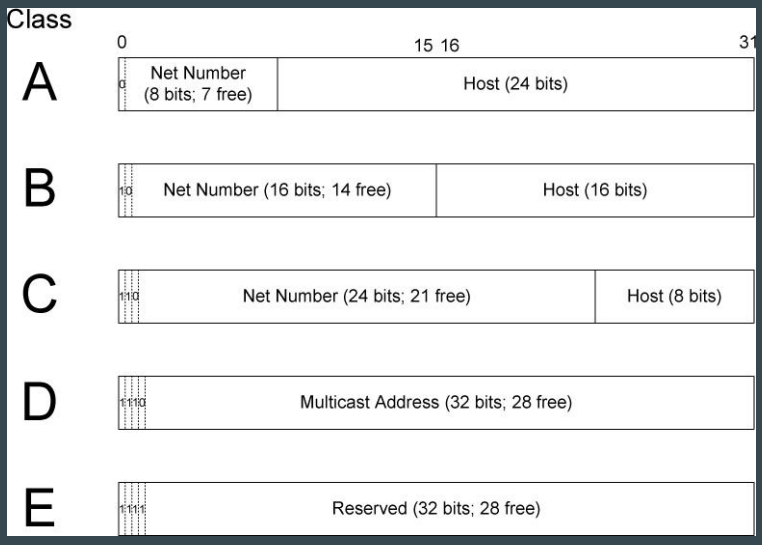
IPv4 Addressing - Chapter 2, Fall & Stevens

- 32-bit
- Written as *a.b.c.d*, where each of *a*, *b*, *c*, *d* ∈ [0,255]

Dotted-Quad Representation	Binary Representation
0.0.0.0	00000000 00000000 00000000 00000000
1.2.3.4	00000001 00000010 00000011 00000100
10.0.0.255	00001010 00000000 00000000 11111111
165.195.130.107	10100101 11000011 10000010 01101011
255.255.255.255	11111111 11111111 11111111 11111111

An IP address is not flat

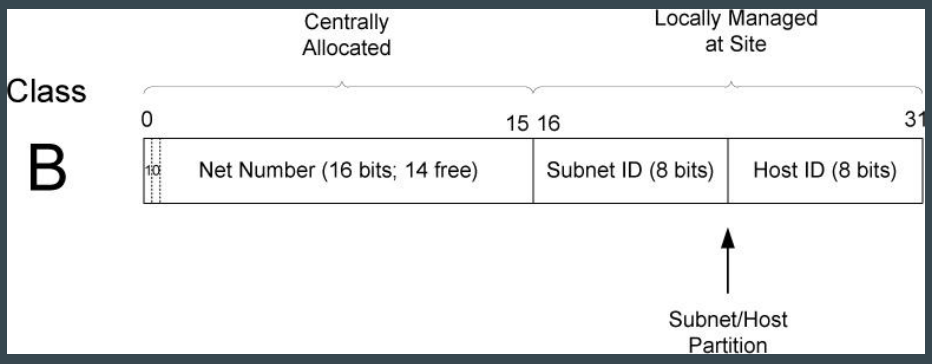
Originally, classful addressing



Originally there was just some bits allocated to the network and some bits allocated to the hosts. The leading bits are used to determine what class of address it is.

Subnetwork Addressing

- Host portion can be subdivided.
- Publicly-routed Internet still sees only classful networks.
- Example:



When you are working with class B addresses you can steal some bits from the hosts to have a subnet.

Need for a “subnet mask”

Identifies network + subnetwork portion of address.

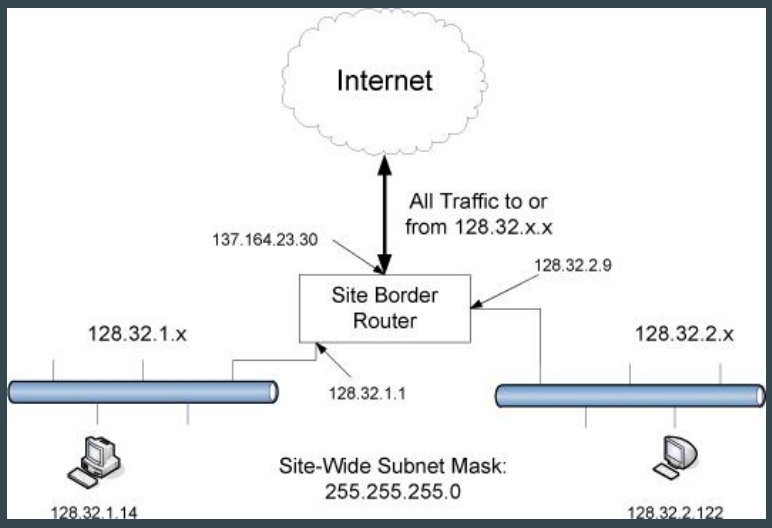
Dotted-Decimal Representation	Shorthand (Prefix Length)	Binary Representation
128.0.0.0	/1	10000000 00000000 00000000 00000000
255.0.0.0	/8	11111111 00000000 00000000 00000000
255.192.0.0	/10	11111111 11000000 00000000 00000000
255.255.0.0	/16	11111111 11111111 00000000 00000000
255.255.254.0	/23	11111111 11111111 11111110 00000000
255.255.255.192	/27	11111111 11111111 11111111 11100000
255.255.255.255	/32	11111111 11111111 11111111 11111111

We can apply a mask to an ipaddress by doing a bitwise and to get prefixes. It can be denoted by / where the number is the number of 1s in the mask. So 1.1.0.0/16 tells us we only care about the first 16 bits of the address.

What's the big deal?

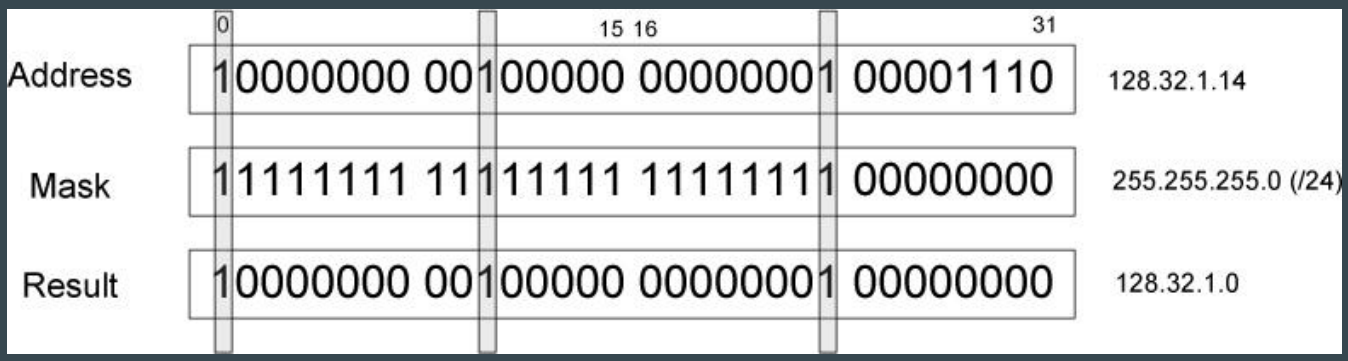
Network equipment (router) support.

Example:



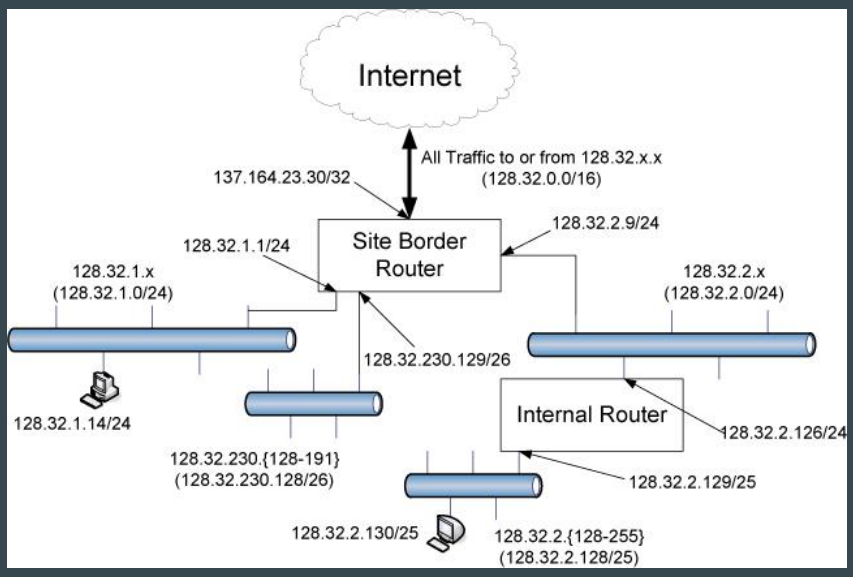
We can see they they have a class B (since 128 is 10000000) with two hosts. Their router has to properly route to those two, it must use a mask to do so. Anything that you have to do with software in a network is bad (mainly for performance reasons). With a routing table we have just the destination address (we dont give a shit about the content). Masks make this harder because they have a concrete destination address and match it against a network prefix. There is no good way to do this. What we do is take the network prefix out and see how many destinations there could be and enumerate the possibilities in a table. This makes the lookup ridiculously fast but uses a bunch of memory. The memory is called TCAM which directly influences your speed but is super expensive.

Application of mask - example



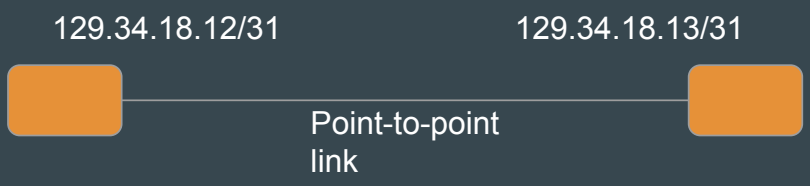
Say we have a prefix of length 24. When you do the bitwise and you get then network id. Just do this example.

Variable Length Subnet Mask (VLSM)



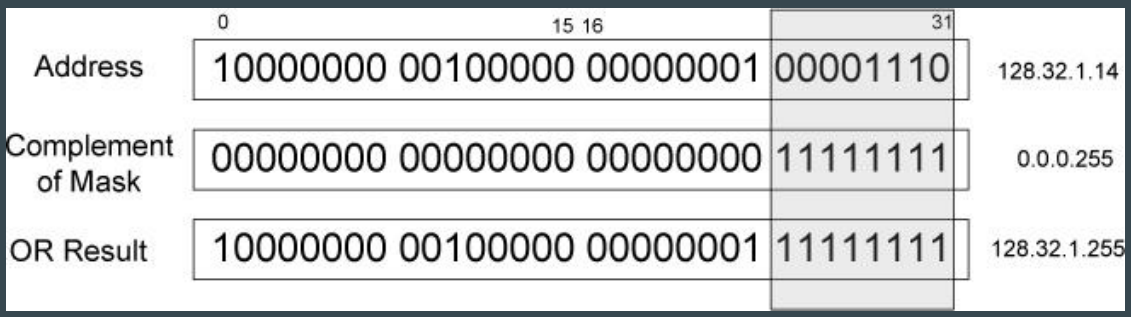
/ What if we want subnets to be different sizes. We want to make the most use of our ipaddresses since they are valuable resources. We can do this by having variable network prefixes. You have to be super careful that nothing collides.

E.g., mask of /31



A special case occurs with a mask of size 32, yes but its weird. A mask of 31 is actually thing when you have a point to point link. It results in the host ids being 0 and 1. Usually we dont use the host of 0 so that we can use it to denote the network prefix and we tend not to use 1 because it is for the ip broadcast.

Broadcast address



- Efficient broadcast - only 1 packet till destination network.
- A big security issue - routers disable “directed broadcast” by default.

The broadcast address is special. If the host id is set to all 1s it is the broadcast address. When we go to send a packet we try to send it but the routing table is going to first send it to somewhere within the network since that is how its routing table works. So your first hop is always in the same network as you. From there it goes to the site boarder router. To get out of our network you need to map an ethernet address to an ipaddress. To do this you package your destination address and broadcast it to the internet to see who knows what ethernet address corresponds to the known destination ipaddress. You can abuse the shit out of this by listening to the request for a corresponding address and just respond to it. Then you get the packet. You do need to get access to the wireless network to do this but that isnt really hard. Another thing you can do is send one packet and it goes to each of the hosts to see who knows where it goes which is a massive security issue since that packet gets cloned to hell.

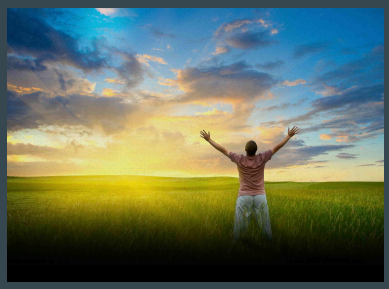
Classless Inter-Domain Routing (CIDR) - motivation

1. By 1994, ≥ half of all class B network IDs already allocated.
2. 32-bit address inadequate.
3. # entries in the global routing table ≅ 65,000.

In the 90s we started running out of ipaddresses. Also all of these tables for looking up classful addresses were getting huge.

Proposed Solution

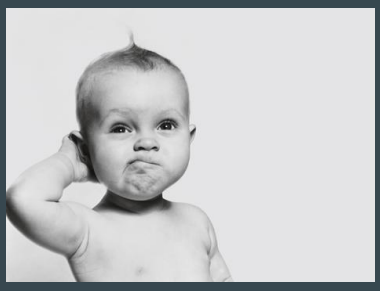
- Simultaneously:
  - Remove class breakdown.
    - E.g., allow network 129.0.18.0/24
  - Allow aggregation in routing tables.
    - E.g., destination 129.0.0.0/8



A proposed solution was to forget classful addressing. You can also agregate adjacent blocks. This decresses the number of entries in the look up table, but increses the complexity of your look up algortihm.

Reducing size of routing table

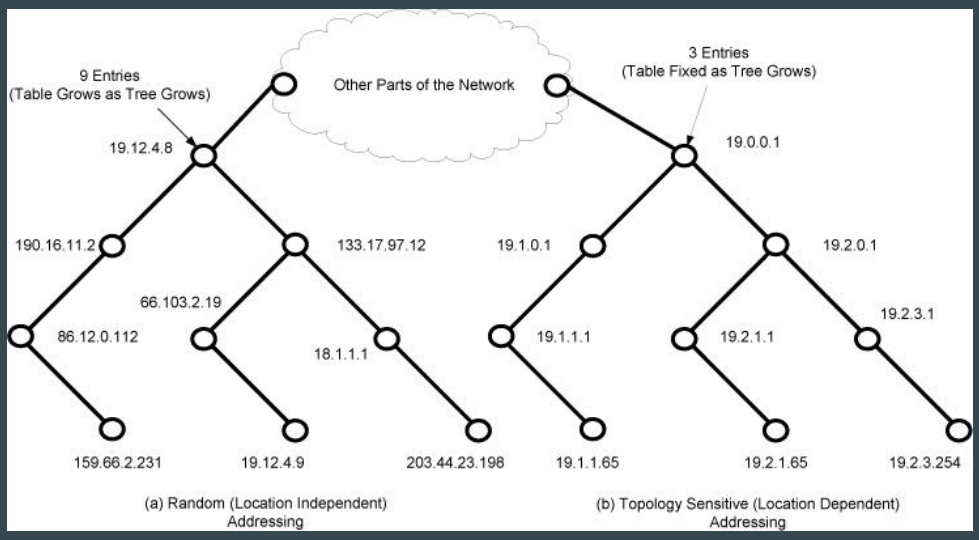
- Removing class aggregation does not help with reducing routing entries.
- But assignment of IP addresses in a way that is sensitive to topology can.
  - E.g., hierarchical routing.



When we removed the class separation it doesnt really help reduce routing table entries.

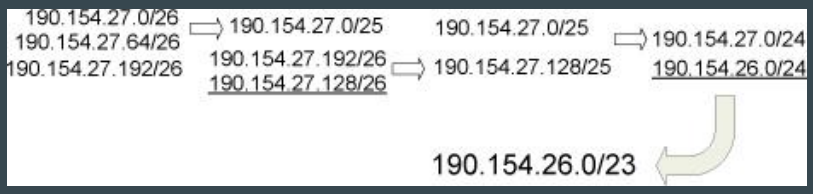
Hierarchical Routing

"if the network topology were arranged as a tree and addresses were assigned in a way that was "sensitive" to this topology, very small routing tables could be used while still maintaining shortest-path routes to all destinations."



If we were to make our topology a tree and leverage it we could fix a bunch of our problems. The example on the left ignores the tree topology and just hands out addresses. Because of this your router needs to keep entries for each node in the tree. If instead we leverage the tree topology only needs to know its address and its children's address. This is now in use. We tend to all get provider agregatable address. They manage these trees of addresses.

Route Aggregation



(Underlined prefixes are those that happen to be added later.)

In this example we can see that each address on the left have the same leading 24 bits and a mask of 26. With the first address the 25th bit is 0 and the other two have 1s so we can make its address mask shorter. We do the same for the other two and on and on to form the tree root.

Important note: in this example more addresses are added as we go, it gets a bit confusing.

Special-use IP addresses

Prefix	Special Use	Reference
0.0.0.0/8	Hosts on the local network. May be used only as a source IP address.	[RFC1122]
10.0.0.0/8	Address for private networks (intranets). Such addresses never appear on the public Internet.	[RFC1918]
127.0.0.0/8	Internet host loopback addresses (same computer). Typically only 127.0.0.1 is used.	[RFC1122]
169.254.0.0/16	"Link-local" addresses—used only on a single link and generally assigned automatically. See Chapter 6.	[RFC3927]
172.16.0.0/12	Address for private networks (intranets). Such addresses never appear on the public Internet.	[RFC1918]
192.0.0.0/24	IETF protocol assignments (IANA reserved).	[RFC5736]
192.0.2.0/24	TEST-NET-1 addresses approved for use in documentation. Such addresses never appear on the public Internet.	[RFC5737]
192.88.99.0/24	Used for 6to4 relays (anycast addresses).	[RFC3068]
192.168.0.0/16	Address for private networks (intranets). Such addresses never appear on the public Internet.	[RFC1918]
198.18.0.0/15	Used for benchmarks and performance testing.	[RFC2544]
198.51.100.0/24	TEST-NET-2. Approved for use in documentation.	[RFC5737]
203.0.113.0/24	TEST-NET-3. Approved for use in documentation.	[RFC5737]
224.0.0.0/4	IPv4 multicast addresses (formerly class D); used only as destination addresses.	[RFC5771]
240.0.0.0/4	Reserved space (formerly class E), except 255.255.255.255.	[RFC1112]
255.255.255.255/32	Local network (limited) broadcast address.	[RFC0919] [RFC0922]

There are some addresses that are special

- 0.0.0.0 - source ip address
- 10.0.0.0, 172.16.0.0, 192.168.0.0 - private addresses
- 127.0.0.0 - loopback address, used to test software without sending shit out
- 169.254.0.0 - used for link local addresses, if you machine doesn't know what ipaddress is you just pick one in that range
  - to keep people from picking the same ipaddress we use gratutious arp (address resolution protocol) chose an ip address, send out packet and see if someone has it, if so pick again

IP Address Allocation

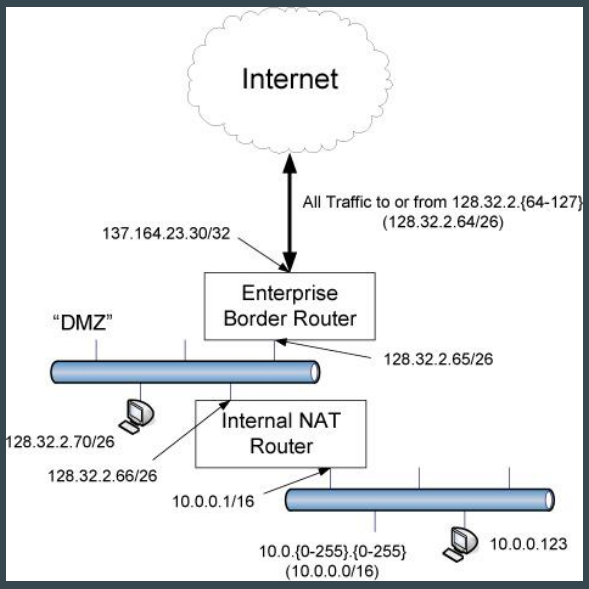
- Internet Assigned Numbers Authority - IANA
- Has delegated entire pool to Regional Internet Registries (RIRs)

RIR Name	Area of Responsibility	Reference
AfriNIC—African Network Information Center	Africa	http://www.afrinic.net
APNIC—Asia Pacific Network Information Center	Asia/Pacific Area	http://www.apnic.net
ARIN—American Registry for Internet Numbers	North America	http://www.arin.net
LACNIC—Regional Latin America and Caribbean IP Address Registry	Latin America and some Caribbean islands	http://lacnic.net/en/index.html
RIPE NCC—Réseaux IP Européens	Europe, Middle East, Central Asia	http://www.ripe.net

Address assignment

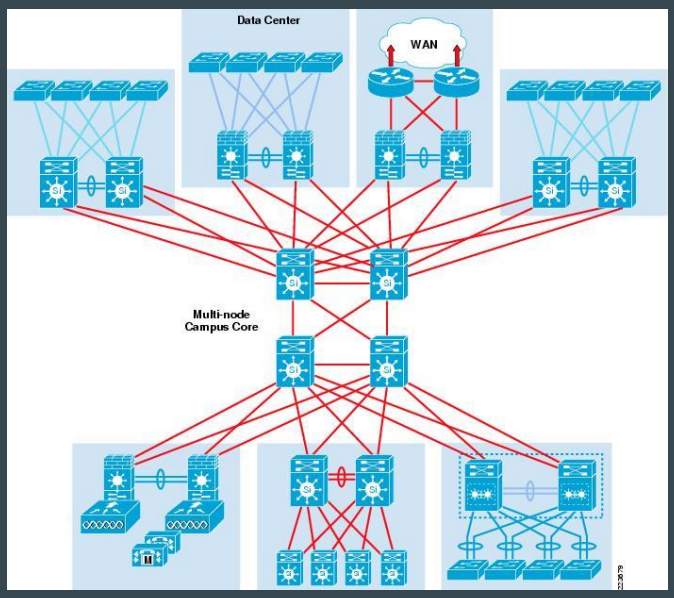
1. Single Provider/No Network/Single Address
2. Single Provider/Single Network/Single Address
3. Single Provider/Multiple Networks/Multiple Addresses
4. Multiple Providers/Multiple Networks/Multiple Addresses (Multihoming)

3. Single Provider/Multiple Networks/Multiple Addresses



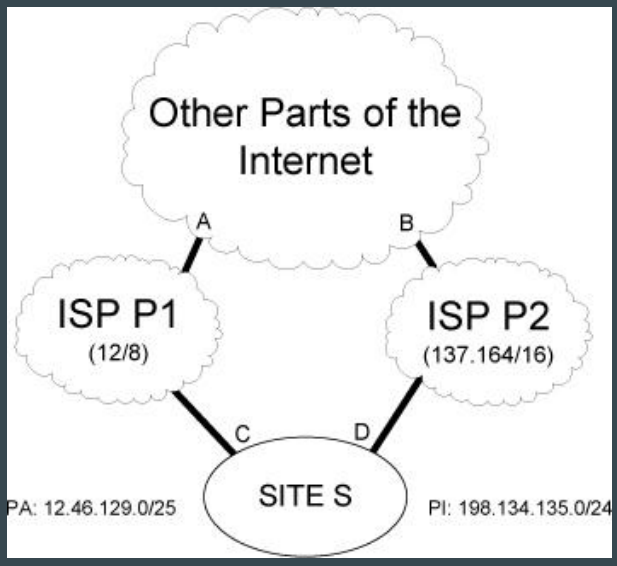
We have a dematerialized zone that separates your internal network from the enterprise boarder where you can put your public facing things.

“Campus” network design



Cisco, Inc.,
“*Enterprise Campus 3.0 Architecture*.”

4. Multiple Providers/Multiple Networks/Multiple Addresses



IP address is used for both:

- Identification
- Location

This can be a bad thing.

This picture is crazy complicated in real life. When you send a packed to a network with multiple providers it applies a best fit algorithm when looking. In this example an packed with desination 12 it will actually go through isp p2 which isnt fair since isp p1 isnt getting its money worth. A solution is to not advertise that address on isp p2. Another solution is to keep them separate.