Computer Networks
Delhi Technological university
Network Layer 3-1
Instructor: Divya Sethia

Objective

•Network Layer

•tnternetworking Architecture

•IP Addressing

•Overview of IPv4

•Overview of IP Routing

•IP Subnetting

Network LAyer

What is Network Layer?

•Network layer is concerned with getting packets from the source all the way to the destination.

•Packets may require to make many hops at intermediate routers while reaching the destination.

•Network layer must know about topology of communication network.

•Choose routes to avoid overloading of some of the communication lines while leaving others idle.

Functions

1. Routing –

- Process of transferring packets received from Data Link Layer of source network to Data Link Layer of correct destination network is called routing.
- Involves decision making at each intermediate node on where to send the packet next so that it eventually reaches its destination. The node which makes this choice is called a router.
- For routing we require some mode of addressing which is recognized by the Network Layer. This addressing is different from the MAC layer addressing.

Functions

2. Inter-networking –
•The network layer is the same across all physical networks (such as Token-Ring and Ethernet).
•If two physically different networks have to communicate, the packets that arrive at the Data Link Layer of the node which connects these two physically different networks, would be stripped of their headers and passed to the Network Layer. The network layer would then pass this data to the Data Link Layer of the other physical network..

Functions

3.Congestion Control –
•If incoming rate of packets arriving at router is more than outgoing rate, then congestion is said to occur.
•Congestion may be caused by many factors:
- Suddenly, packets begin arriving on many input lines and all need same output line, then queue will build up. If there is insufficient memory to hold all of them, packets will be lost.
- But even if routers have an infinite amount of memory, congestion gets worse, because by the time packets reach to front of queue, they have already timed out (repeatedly), and duplicates have been sent. All these packets are dutifully forwarded to next router, increasing the load all the way to destination.
- Router's CPUs slow at performing bookkeeping tasks required of them, queues can build up, even though there is excess line capacity.
- Low-bandwidth lines

Internetworking

•**Internetworking** is the practice of connecting a computer network with other networks through the use of gateways that provide a common method of routing information packets between the networks.

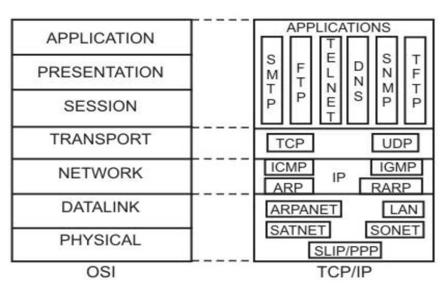
•The resulting system of interconnected networks is called an **internetwork**, or simply an **internet**.

• **Internet**, a network of networks based on many underlying hardware technologies, but unified by an internetworking protocol standard, the Internet Protocol Suite, often also referred to as TCP/IP.

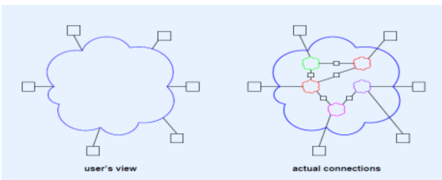
•IP provides unreliable, connectionless best-effort datagram delivery service

•TCP provides reliable, efficient and cost-effective end-to-end delivery of data

Relationship between TCP/IP and OSI



The Two Views Of A TCP/IP Internet



•End-user system is called *host computer*
– Connects to physical network
– Possibly many hosts per network
– Possibly more than one network connection per host

•Dedicated systems called *IP gateways or IP routers* interconnect networks
– Router connects two or more networks

Classful Internet Address

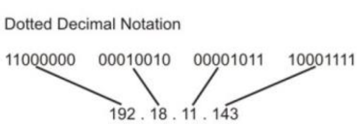
•Every host and router on network has unique standard form of network address, which encodes its network number and host number

•The IP addresses are 32-bit long

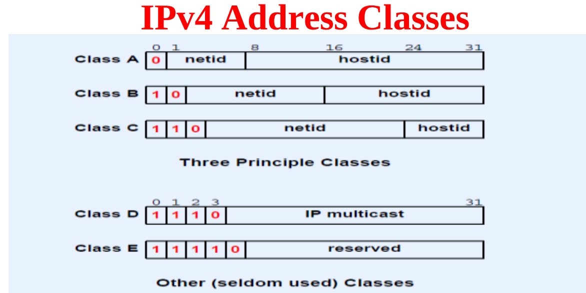
•Address divided into two parts
– Prefix (network ID) identifies network to which host attaches
– Suffix (host ID) identifies host on that network

•Classful because each IP address belongs to a class

IP Address



IPv4 Address Classes



•Class of address determined by three high-order bits
•Class A (1 to 126) – more number of hosts
•Class B (128 to 191) – intermediate size networks
•Class C (192 to 223) – less hosts

IP Address Conventions

•Hostid =0 -> this host on a network
•Netid = 0 -> this network , used when host wants to communicate over a network but does not know the IP address (say before DHCP) 0 netid-> this network

• Broadcast on the local wire
– Network and host fields both contain all *1 bits (*255.255.255.255)
- Broadcast limited to local network only
- used for bootstraping

•Directed broadcast: broadcast on specific (possibly remote) network
– Host field contains all *1 bits*

Special Addresses

Prefix	Suffix	Type of address	Purpose
All 0s	All 0s	This computer	Used during rebooting
Network	All 0s	Network	Identifies a network
Network	All 1s	Directed broadcast	Broadcast on specified net
All 1s	All 1s	Limited broadcast	Broadcast on local net
127	Any	Loopback	Testing

Advantages Of Classful Addressing

• Computationally efficient
– First bits specify size of prefix / suffix

• Allows mixtures of large and small networks

LoopBack IP Address

•Address is 127.0.0.1
•Used for testing
•Refers to local computer (never sent to Internet)
•If destination is loopback address, data traffic not sent out on the network.

IP Address and MAC address

– The Source and Destination IP addresses **NEVER** change

– The Source & Destination MAC addresses **CHANGE** as packet is forwarded from one router to the next

Internet Address Resolution Protocol (ARP)

•Standard for dynamic address resolution in the Internet

•Requires hardware broadcast

•Intended for LAN

•Important idea: ARP only used to map addresses within a single physical network, never across multiple networks

ARP

•Machine A broadcasts ARP request with B's IP address

•All machines on local net receive broadcast

• Machine B replies with its physical address

• Machine A adds B's address information to its table

• Machine A delivers packet directly to B

INTERNET PROTOCOL: CONNECTIONLESS DATAGRAM DELIVERY

•Provides connectionless packet delivery service

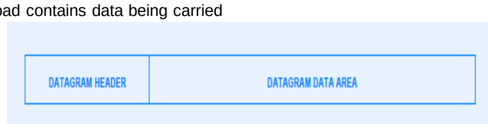
• Defines three important items
  – Internet addressing scheme
  – Format of packets for the (virtual) Internet
  – Packet forwarding

IP Datagram Layout

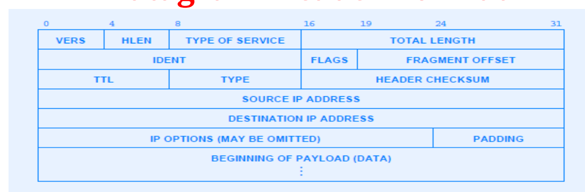
•IP packets is analogous to physical network packet and is known as *IP datagram*

•Header contains
– Source Internet address
– Destination Internet address
– Datagram type field

• Payload contains data being carried



Datagram Header Format



•Vers: Version of IP
•HLEN: Datagram header length since it is variable
•Total Length: Length of header + data
•Type of Service (TOS) : how datagram should be handled

0	1	2	3	4	5	6	7
Precedence			D	T	R	Unused	

Datagram Header Format - TOS

0	1	2	3	4	5	6	7
Precedence			**D**	**T**	**R**	**Unused**	

•**Type of Service (TOS)** : how datagram should be handled
•**Precedence bits** 0 (normal) – 7 (network control) – sender can indicate importance to datagram eg:
- if control info then it must have precedence over data
- routers with precedence val 6,7 to exchange routing info over congested network

•Bit D- requests low delay (eg: keystrokes from user to remote computer)
•Bit T – requests high throughput (eg: bulk file transfer to travel over a high capacity link)
•Bit R requests high reliability
•Bits give a hint to the routing algorithm to choose over multiple paths for the desired request.

Differentiated Services (DiffServ)

0	1	2	3	4	5	6	7
CODEPOINT						UNUSED	

•Late **1990's redefined meaning of the Service Type field for DiffServ for Qos**
•First six bits form codepoint **– Differentiated services Code point (DSCP),** last 2 bits are unused
•Few services and codepoints under it.
•Backward compatibility – when last three bits of the codepoint are zero precedence bits define 8 class of service xxx000
•Router honors original precedence scheme for high priority traffic for val 6 and 7 even when it is set for Diff services

Pool	Codepoint	Assigned By
1	xxxxx0	Standard
2	xxxx11	Local/exper
3	xxxx01	Local/exper

Quality of Service (Qos)

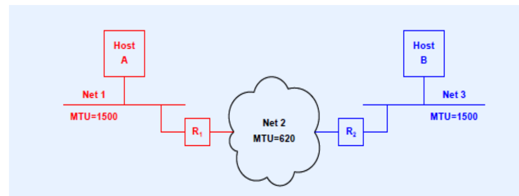
•Internet should give better quality of service to some packets
•Factors:
- Application performance
- Bandwidth requirements
- Complexity/cost

•Providing Better Service
- Routing /Forwarding
- Scheduling/Dropping
- Relative/Absolute – favored packets get lower delay and lower drop rate

•**Diffserv** provides different levels of services
•Edge routers :
- sort packets into classes based on factors
- police/shape traffic
- set DSCP bits in packet header

•Code routers:
- Handle packets based on DSCP

Fragmentation



•Hosts A and B send datagrams of up to 1500 octets

•Router R1 fragments large datagrams from Host A before sending over Net 2

• Router R2 fragments large datagrams from Host B before sending over Net 2

Datagram Fragmentation

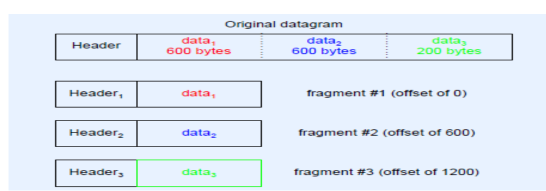
•Performed by routers

•Divides datagram into several, smaller datagrams called fragments

• Fragment uses same header format as datagram

• Each fragment forwarded independently

Fragmentation



•Offset specifies where data belongs in original datagram

•Offset actually stored as multiples of 8 octets

•MORE FRAGMENTS bit turned off in header of fragment #3

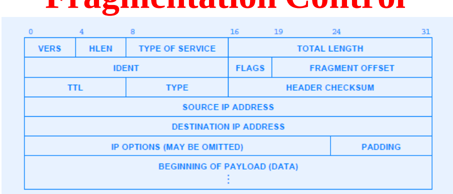
Reassembly

- Ultimate destination puts fragments back together
  - – Key concept!
  - – Needed in a connectionless Internet
- Known as *reassembly*
- No need to reassemble subfragments first
- Timer used to ensure all fragments arrive
  - – Timer started when first fragment arrives
  - – If timer expires, entire datagram discarded

Place of Reassembly

- Should datagram be reassembled intermediately or at the end destination?
- Preserving fragments all way to destination is ideal:
  - allows fragments to be router independently
  - intermediate routers do not require to store and reassemble
- But has two disadvantages:
  i) Larger intermediate networks carry much smaller MTU packets
  ii) Lost fragments need to be retransmitted with a reassembly timer when it gets the initial fragment
- Reassembly at the destination is the usual trend

Fragmentation Control



- IDENT- identification is a unique integer identifying datagram – common to all fragments of the datagram
- FRAGMENT OFFSET –offset in the datagram measured in units of 8 octets. For reassembly require all fragments from 0 to the highest number to be available at destination
- FLAGS

0	1	2
Reserved	Do no fragment	More Fragments

Time To Live (TTL)

- Allows how long in sec datagram is allowed to remain in the internet system
- Sets max time that datagram should survive
- TTL field of datagram header decremented at each hop (i.e., each router) by the time it remained in the router waiting for service
- If TTL reaches zero, datagram discarded since it cannot travel in the internet forever.
- Prevents datagrams from looping indefinitely (in case forwarding error introduces loop)
- IETF (Internet Engineering Task Force)recommends initial value of 255 (max)

Checksum Field

•16-bit 1's complement checksum

•Over IP header only!

•Recomputed at each hop since ttl is changing

IP Options

•The options field was created in order to allow features to be added into IP as time passes and requirements change
•Seldom used primarily for debugging and testing
•Only *some options copied into fragments*
•Are variable length
•Note: padding needed because header length measured in32-bit multiples
•Option starts with option code octet
•Copy flag set to 1 requires option to be copied in the fragments if 0 then must be present in the first fragment only

0	1	2	3	4	5	6	7
Copy	Option Class		Option number				

IP Options

Option Class	Meaning
0	Datagram/Network control
1	Reserved for future
2	Debugging and measurement
3	Reserved for future

•**Record Route Option**
- Option class = 0, Option number: 7 Meant for tracing a route
- Source enables record route option & destination processes resultant list
- Intermediate machine adds its address to record route list if there is sufficient space
- Length is variable & gives the length of the option as occurs in the header
- Not copied on fragmentation, goes in first fragment only

0 8 16 24 31

Code(7)	Length	Pointer	
First IP Address			
Second IP Address			
….			

Other options

•**Record Route Option…**
- When data gram arrives the destination machine can extract and process the list of IP addresses
- Computer receiving data gram ignores the recorded route

•**Source Route Option**
- Sender dictates path through which datagrams are forced to traverse networks
  - Strict source routing – specify exact path
  - Loose source routing – specifies address that datagram must follow but allows multiple network hops between successive addresses on the list

•**Timestamp Option**
- Each router along the path from source to destination fill in the timestamp information for the IP address of the router.

- Time and date at which a router handles datagram expressed as milliseconds since midnight, universal time

IP Semantics

•IP uses best-effort delivery
– Makes an attempt to deliver
– Does not guarantee delivery

•In the Internet, routers become overrun or change routes, meaning that:
– Datagrams can be lost
– Datagrams can be duplicated
– Datagrams can arrive out of order or scrambled

•Motivation: allow IP to operate over the widest possible variety of physical networks

Forwarding IP Datagrams

Datagram Transmission

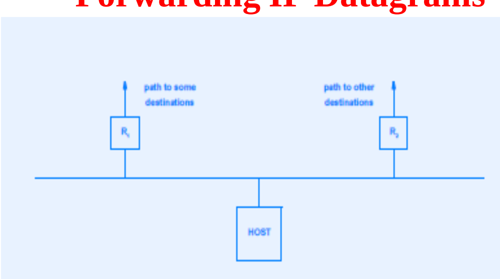
•Host delivers datagrams to directly connected machines

•Host sends datagrams that cannot be delivered directly to router

•Routers forward datagrams to other routers

•Final router delivers datagram directly

Forwarding IP Datagrams



•Host That Must Choose How To Forward Datagrams

•Every interface on a host is singly homed

Forwarding IP Datagrams

•Direct delivery
– Ultimate destination can be reached over one network
– The "last hop" along a path
– Also occurs when two communicating hosts both attach to same physical network
– sender encapsulates datagram in physical frame, binds the destination IP address to a physical hardware address, and sends the resulting frame directly to the destination.

•Indirect delivery
– Requires intermediary (router)
– Internet addresses of all machines on a single network include a common network prefix and extracting that prefix requires only a few machine instructions
_ Tests efficiently whether a machine can be reached directly

Forwarding IP Datagrams

Datagram Forwarding

•General paradigm
– Source host sends to first router
– Each router passes datagram to next router
– Last router along path delivers datagram to destination host

•Only works if routers cooperate

Forwarding IP Datagrams

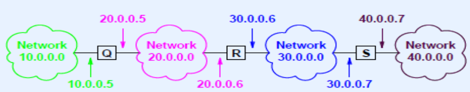
Efficient Forwarding
• Decisions based on table lookup
• Routing tables keep only network portion of addresses (size proportional to number of networks, not number of hosts)

•Extremely efficient
– Lookup
– Route update

•Table used to decide how to send datagram known as *routing table (also called a forwarding table)*
•Routing table only stores address of next router along the path

•Scheme is known as *next-hop forwarding or next-hop routing*

Forwarding IP Datagrams



An example Internet with IP addresses

TO REACH NETWORK	ROUTE TO THIS ADDRESS
20.0.0.0 / 8	DELIVER DIRECT
30.0.0.0 / 8	DELIVER DIRECT
10.0.0.0 / 8	20.0.0.5
40.0.0.0 / 8	30.0.0.7

The routing table for router R

Forwarding IP Datagrams

Special Cases
•Default route
- Special entry in IP routing table
- Matches "any" destination address
- Only one default permitted
- Only selected if no other match in table

•Host-specific route
- Entry in routing table
- Matches entire 32-bit value
- Can be used to send traffic for a specific host along a specific path (i.e., can differ from the network route)

IP Routing Algorithm

Routing Datagram(Datagram, Routing Table)
Extract destination IP address D from Datagram &compute network prefix N:

i)If N matches any directly connected network address send datagram to destination D over that network

ii) else if table contains host specific route for D send datagram to next-hop address

iii) else if table contains route for network N send datagram to next-hop address

iv) else if table contains a default-route send datagram to default router specified in table

v) else declare routing error

IP Error Reporting Mechanism

•Named *Internet Control Message Protocol (ICMP)*

•Required and integral part of IP

•Used primarily by routers to report delivery or routing problems to original source

• Also includes informational (nonerror) functionality

• Uses IP to carry control messages

• No error messages sent about error messages

Classful Addressing

•Three possible classes for networks

•Class C network limited to 254 hosts (cannot use all-1s or all-0s)

•Personal computers result in networks with many hosts

•Class B network allows many hosts, but insufficient class B prefixes

•How can we minimize the number of assigned network prefixes especially class B) without abandoning the 32-bit addressing scheme?

Internet Scaling Problem

•IPv4 defines a 32-bit address.
2 to the power 32 (4,294,967,296) addresses available.
The address shortage problem is aggravated by the fact that portions of the IP address space have not been efficiently allocated

•Disign problem:
- Class C networks are too small (254 hosts).
- Next option is class B, which is too big (65,534 hosts)

Internet Scaling Problem Alternatives

- IPv6
- Subnetting
- **V**LSM
- CDIR
- NAT

Subnetting

- Initial Problems:

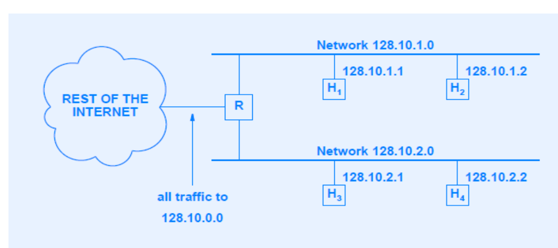
 - Internet routing tables started to grow

 - Local administrators had to request another network number from the Internet before a new network could be installed at their site

Subnetting

- Allows a single network address to span multiple physical networks
- Not part of original TCP/IP address scheme
- Allows an organization to use a single network prefix for multiple physical networks
- Subdivides the host suffix into a pair of fields for physical network and host
- Interpreted only by routers and hosts at the site; treated like normal address elsewhere

Example of Subnet Addressing



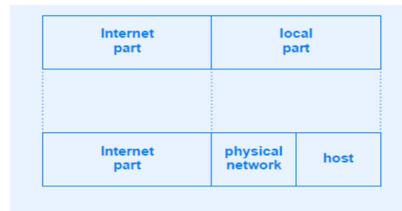
- Both physical networks share prefix 128.10
- Router R uses third octet of address to choose physical net
- Only R is aware of the different physical networks

Interpretation of IP Address

•Classful interpretation is two-level hierarchy
– Physical network identified by prefix
– Host on the net identified by suffix

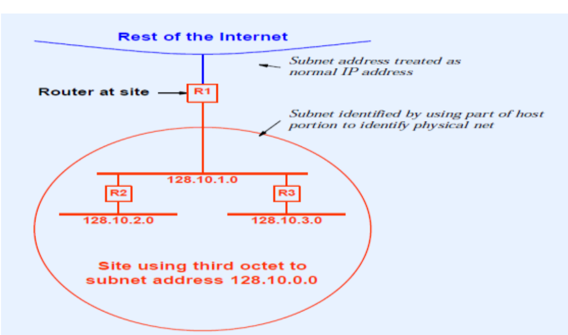
•Subnetted interpretation is three-level hierarchy
– Site identified by network prefix
– Physical net at site identified by part of suffix
– Host on the net identified by remainder of suffix

Subnet Address



Example Class B address , 16-bit host portion is divided into two 8-bit fields

Subnet Address



Address Mask

•Each physical network is assigned 32-bit *address mask (also* called *subnet mask)*
•One bits in mask cover network prefix plus zero or more bits of suffix portion
•Logical *and between mask and destination IP address* extracts the prefix and subnet portions

•32-bit subnet mask for a class B address
11111111 111111111 11111111 00000000
|__________________|_________|__________|
network part subnet host part

255.255.255.0

Variety of Route

•Forwarding must accommodate
– Network-specific routes
– Subnet-specific routes
– Host-specific routes
– Default route
– Limited broadcast
– Directed broadcast to network
– Directed broadcast to specific subnet

• Single algorithm with address masks can accommodate all the above

Use of Address Masks

•Each entry in routing table also has address mask

•All-1s mask used for host-specific routes

•Network mask used for network-specific routes

•Subnet mask used for subnet-specific routes

•All-0s mask used for default route

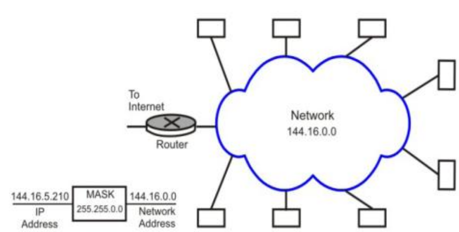
Unified Forwarding Algorithm

Algorithm:

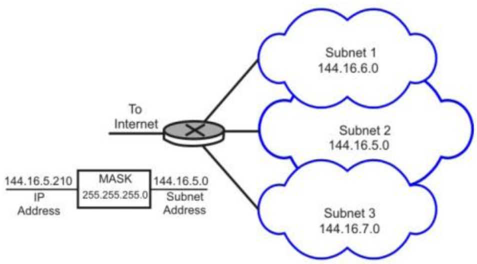
Forward_IP_Datagram (datagram, routing_table)

```
Extract destination IP address, ID, from datagram;
If prefix of ID matches address of any directly connected
   network send datagram to destination over that network
   (This involves resolving ID to a physical address,
   encapsulating the datagram, and sending the frame.)
else
   for each entry in routing table do
      Let N be the bitwise-and of ID and the subnet mask
      If N equals the network address field of the entry then
        forward the datagram to the specified next hop address
   endforloop
If no matches were found, declare a forwarding error;
```


Routing without subnet



Subnet masking with the help of router



Advantages of Subnetting

•Divides the bigger network into smaller individual networks and it becomes easy to handle the smaller networks than handle one huge network

•The network administrators can assign IP numbers to specific departments in an organization so that when they can easily identify the departments on the network using the same numbers.

•Security

•Size of global Internet routing table does not grow because site administrator does not need to obtain additional adress space and routing advertisments for all subnets are combined into a single routing table entry.

Classless Addressing: Address Blocks

To overcome address depletion.
•No classes, but the address are still granted in blocks.
• The size of the block(the number of addresses) varies based on the nature and size of the entity.
• Household: 2 addresses
• Large organization: thousands of addresses.
• ISP: thousands or hundreds of thousands based on the number of customers it may serve.

Classless Addressing: Address Blocks

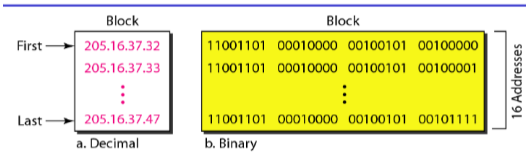
Restriction
The Internet Authorities impose three restrictions:

1. **The address in a block must be contiguous, one after another.**

2. The number of addresses in a block must be a power of 2 (1,2,4,8,…)

3. The first address must be evenly divisible by the number of addresses.

A block of 16 addresses granted to a small organization



We can see that the restrictions are applied to this block.

- The addresses are contiguous.
- The number of addresses is a power of 2 ($16 = 2^4$)
- the first address is divisible by 16. The first address, when converted to a decimal number(use base 256), is 3,440,387,360, which when divided by 16 results in 215,024,210.

Classless Addressing: Mask

Mask: /n
•32- bit
• can take any value from 0 to 32, for ex /24
•The n leftmost bits are 1s
• 32-n rightmost bits are 0s
In IPv4 addressing, a block of addresses can be defined as: x.y.z.t /n
in which x.y.z.t defines one of the addresses and the /n defines the mask.
Example: 172.31.16.42/26

Classless Addressing: Mask

The address and the /n notation define the whole block:
- First address
- Last address
- Number of address

The first address in the block (network address) can be found by setting the rightmost 32 − n bits to 0s.

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

205.16.37. 0010|0111

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

205.16.37. 0010|0000

or

205.16.37.32.

This is actually the block shown in example 8

The first address in a block is normally not assigned to any device; it is used as the network address that represents the organization to the rest of the world.

The last address in the block can be found by setting the rightmost 32 − n bits to 1s.

A block of addresses is granted to a small organization. We know that one of the addresses is 205.16.37.39/28. Find the last address for the block

Solution
The address is
205.16.37. 0010 0111
If we set 32 − 28 = 4 rightmost bits to 1,
we get
205.16.37. 0010 1111
or
205.16.37.47

Classless Inter-Domain Routing (**CIDR**)
- http://en.wikipedia.org/wiki/Classless_Inter-Domain_Routing#CIDR_blocks

IPv4 CIDR blocks

IPv4 CIDR

IP/CIDR	Δ to last IP addr	Mask	Hosts (*)	Class	Notes
a.b.c.d/**32**	+0.0.0.0	255.255.255.255	1	1/256 C	
a.b.c.d/**31**	+0.0.0.1	255.255.255.254	2	1/128 C	d = 0 ... (2n) ... 254
a.b.c.d/**30**	+0.0.0.3	255.255.255.252	4	1/64 C	d = 0 ... (4n) ... 252
a.b.c.d/**29**	+0.0.0.7	255.255.255.248	8	1/32 C	d = 0 ... (8n) ... 248
a.b.c.d/**28**	+0.0.0.15	255.255.255.240	16	1/16 C	d = 0 ... (16n) ... 240
a.b.c.d/**27**	+0.0.0.31	255.255.255.224	32	1/8 C	d = 0 ... (32n) ... 224
a.b.c.d/**26**	+0.0.0.63	255.255.255.192	64	1/4 C	d = 0, 64, 128, 192
a.b.c.d/**25**	+0.0.0.127	255.255.255.128	128	1/2 C	d = 0, 128

References

- TCP/IP volume 1 by Douglas Comer
- Chapter 19 Forouzan
- http://www.networkcomputing.com/netdesign/ip101.html
- http://www.exforsys.com/tutorials/networking/subnetting.html
- http://www.scribd.com/doc/26828247/Variable-Length-Subnet-Masks
- http://www.tcpipguide.com/free/t_toc.htm

THANKS