Network Layer Services- Packetizing, Routing and Forwarding

**Network layer** is the third layer in the [OSI model](https://www.geeksforgeeks.org/layers-of-osi-model/) of computer networks. It’s main function is to transfer network packets from the source to the destination. It is involved both at the source host and the destination host. At the source, it accepts a packet from the transport layer, encapsulates it in a datagram and then deliver the packet to the data link layer so that it can further be sent to the receiver. At the destination, the datagram is decapsulated, the packet is extracted and delivered to the corresponding transport layer.

**Features :**

1. Main responsibility of Network layer is to carry the data packets from the source to the destination without changing or using it.
2. If the packets are too large for delivery, they are fragmented i.e., broken down into smaller packets.
3. It decides the root to be taken by the packets to travel from the source to the destination among the multiple roots available in a network (also called as routing).
4. The source and destination addresses are added to the data packets inside the network layer.

The **services** which are offered by the network layer protocol are as follows:

1. [**Packetizing**](https://www.geeksforgeeks.org/fragmentation-network-layer/)**–**  
   The process of encapsulating the data received from upper layers of the network(also called as payload) in a network layer packet at the source and decapsulating the payload from the network layer packet at the destination is known as packetizing.

The source host adds a header that contains the source and destination address and some other relevant information required by the network layer protocol to the payload received from the upper layer protocol, and delivers the packet to the data link layer.

The destination host receives the network layer packet from its data link layer, decapsulates the packet, and delivers the payload to the corresponding upper layer protocol. The routers in the path are not allowed to change either the source or the destination address. The routers in the path are not allowed to decapsulate the packets they receive unless they need to be fragmented.

1. [**Routing**](https://www.geeksforgeeks.org/types-of-routing/)**and Forwarding –**  
   These are two other services offered by the network layer. In a network, there are a number of roots available from the source to the destination. The network layer specifies has some strategies which find out the best possible route. This process is referred to as routing. There are a number of routing protocols which are used in this process and they should be run to help the routers coordinate with each other and help in establishing communication throughout the network.

Forwarding is simply defined as the action applied by each router when a packet arrives at one of its interfaces. When a router receives a packet from one of its attached networks, it needs to forward the packet to another attached network ([unicast routing](https://www.geeksforgeeks.org/unicast-routing-link-state-routing/)) or to some attached networks(in case of multicast routing).

Some of the other **services which are expected** from the network layer are:

1. [**Error Control**](https://www.geeksforgeeks.org/error-control-in-tcp/)**–**  
   Although it can be implemented in the network layer, but it is usually not preferred because the data packet in a network layer maybe fragmented at each router, which makes error checking inefficient in the network layer.
2. [**Flow Control**](https://www.geeksforgeeks.org/difference-between-flow-control-and-congestion-control/)**–**  
   It regulates the amount of data a source can send without overloading the receiver. If the source produces a data at a very faster rate than the receiver can consume it, the receiver will be overloaded with data. To control the flow of data, the receiver should send a feedback to the sender to inform the latter that it is overloaded with data.

There is a lack of flow control in the design of the network layer. It does not directly provide any flow control. The datagrams are sent by the sender when they are ready, without any attention to the readiness of the receiver.

1. [**Congestion Control**](https://www.geeksforgeeks.org/congestion-control-in-computer-networks/)**–**  
   Congestion occurs when the number of datagrams sent by source is beyond the capacity of network or routers. This is another issue in the network layer protocol. If congestion continues, sometimes a situation may arrive where the system collapses and no datagrams are delivered. Although congestion control is indirectly implemented in network layer, but still there is a lack of congestion control in the network layer.

**Advantages of Network Layer Services :**

* Packetization service in network layer provides an ease of transportation of the data packets.
* Packetization also eliminates single points of failure in data communication systems.
* Routers present in the network layer reduce network traffic by creating collision and broadcast domains.
* With the help of Forwarding, data packets are transferred from one place to another in the network.

**Disadvantages of Network Layer Services :**

* There is a lack of flow control in the design of the network layer.
* Congestion occurs sometimes due to the presence of too many datagrams in a network which are beyond the capacity of network or the routers. Due to this, some routers may drop some of the datagrams and some important piece of information maybe lost.
* Although indirectly error control is present in network layer, but there is a lack of proper error control mechanisms as due to presence of fragmented data packets, error control becomes difficult to implement.

Introduction of Classful IP Addressing

Last Updated: 09-09-2019

IP address is an address having information about how to reach a specific host, especially outside the LAN. An IP address is a 32 bit unique address having an address space of 232.  
Generally, there are two notations in which IP address is written, dotted decimal notation and hexadecimal notation.

Dotted Decimal Notation:

Hexadecimal Notation:

Some points to be noted about dotted decimal notation:

1. The value of any segment (byte) is between 0 and 255 (both included).
2. There are no zeroes preceding the value in any segment (054 is wrong, 54 is correct).

**Classful Addressing**  
The 32 bit IP address is divided into five sub-classes. These are:

* Class A
* Class B
* Class C
* Class D
* Class E

Each of these classes has a valid range of IP addresses. Classes D and E are reserved for multicast and experimental purposes respectively. The order of bits in the first octet determine the classes of IP address.  
IPv4 address is divided into two parts:

* **Network ID**
* **Host ID**

The class of IP address is used to determine the bits used for network ID and host ID and the number of total networks and hosts possible in that particular class. Each ISP or network administrator assigns IP address to each device that is connected to its network.

**Note:**IP addresses are globally managed by Internet Assigned Numbers Authority(IANA) and regional Internet registries(RIR).

**Note:**While finding the total number of host IP addresses, 2 IP addresses are not counted and are therefore, decreased from the total count because the first IP address of any network is the network number and whereas the last IP address is reserved for broadcast IP.

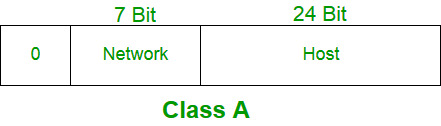
**Class A:**

IP address belonging to class A are assigned to the networks that contain a large number of hosts.

* The network ID is 8 bits long.
* The host ID is 24 bits long.

The higher order bit of the first octet in class A is always set to 0. The remaining 7 bits in first octet are used to determine network ID. The 24 bits of host ID are used to determine the host in any network. The default subnet mask for class A is 255.x.x.x. Therefore, class A has a total of:

* 2^7-2= 126 network ID(Here 2 address is subracted because 0.0.0.0 and 127.x.y.z are special address. )
* 2^24 – 2 = 16,777,214 host ID

IP addresses belonging to class A ranges from 1.x.x.x – 126.x.x.x  
[](https://media.geeksforgeeks.org/wp-content/cdn-uploads/IP_addressing_4.jpg)

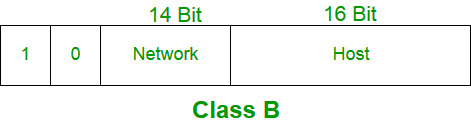
**Class B:**

IP address belonging to class B are assigned to the networks that ranges from medium-sized to large-sized networks.

* The network ID is 16 bits long.
* The host ID is 16 bits long.

The higher order bits of the first octet of IP addresses of class B are always set to 10. The remaining 14 bits are used to determine network ID. The 16 bits of host ID is used to determine the host in any network. The default sub-net mask for class B is 255.255.x.x. Class B has a total of:

* 2^14 = 16384 network address
* 2^16 – 2 = 65534 host address

IP addresses belonging to class B ranges from 128.0.x.x – 191.255.x.x.  
[](https://media.geeksforgeeks.org/wp-content/cdn-uploads/IP_addressing_5.jpg)

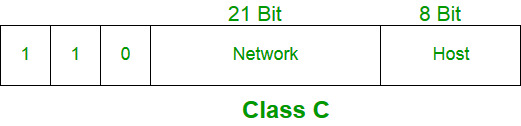
**Class C:**

IP address belonging to class C are assigned to small-sized networks.

* + - The network ID is 24 bits long.
    - The host ID is 8 bits long.

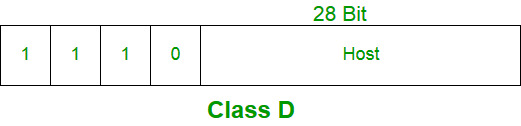
The higher order bits of the first octet of IP addresses of class C are always set to 110. The remaining 21 bits are used to determine network ID. The 8 bits of host ID is used to determine the host in any network. The default sub-net mask for class C is 255.255.255.x. Class C has a total of:

* + - 2^21 = 2097152 network address
    - 2^8 – 2 = 254 host address

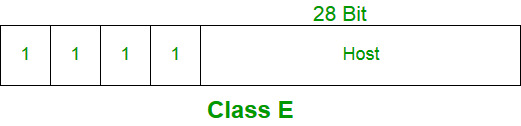
IP addresses belonging to class C ranges from 192.0.0.x – 223.255.255.x.  
[](https://media.geeksforgeeks.org/wp-content/cdn-uploads/IP_addressing_6.jpg)

**Class D:**

IP address belonging to class D are reserved for multi-casting. The higher order bits of the first octet of IP addresses belonging to class D are always set to 1110. The remaining bits are for the address that interested hosts recognize.

Class D does not posses any sub-net mask. IP addresses belonging to class D ranges from 224.0.0.0 – 239.255.255.255.  
[](https://media.geeksforgeeks.org/wp-content/cdn-uploads/IP_addressing_7.jpg)

**Class E:**

IP addresses belonging to class E are reserved for experimental and research purposes. IP addresses of class E ranges from 240.0.0.0 – 255.255.255.254. This class doesn’t have any sub-net mask. The higher order bits of first octet of class E are always set to 1111.  
[](https://media.geeksforgeeks.org/wp-content/cdn-uploads/IP_addressing_8.jpg)

**Range of special IP addresses:**

**169.254.0.0 – 169.254.0.16** : Link local addresses  
**127.0.0.0 – 127.0.0.8** : Loop-back addresses  
**0.0.0.0 – 0.0.0.8** : used to communicate within the current network.

**Rules for assigning Host ID:**

Host ID’s are used to identify a host within a network. The host ID are assigned based on the following rules:

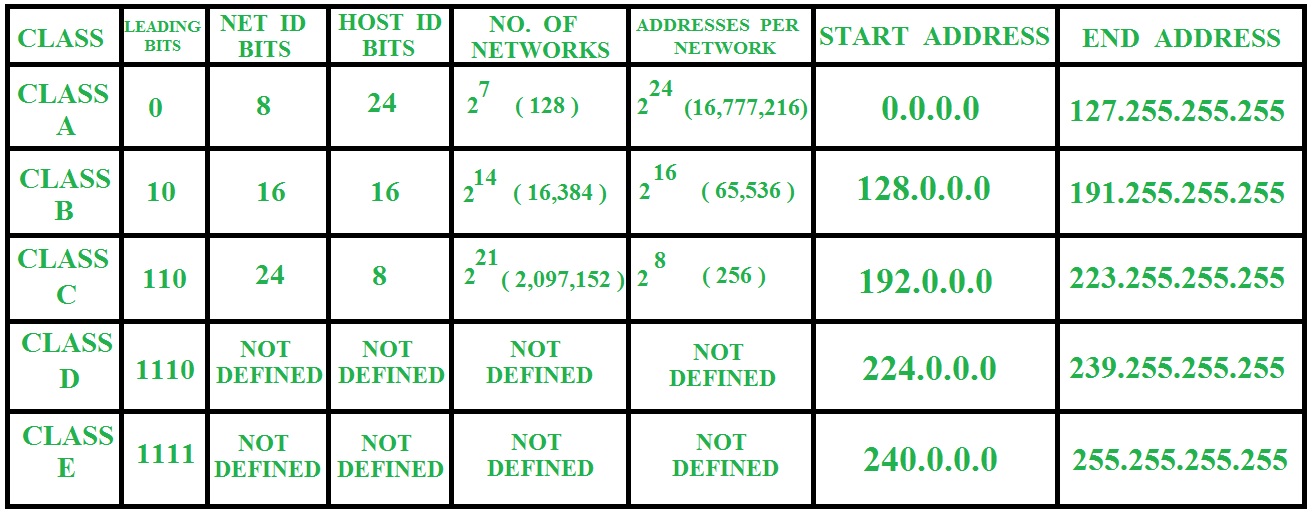
* + - Within any network, the host ID must be unique to that network.
    - Host ID in which all bits are set to 0 cannot be assigned because this host ID is used to represent the network ID of the IP address.
    - Host ID in which all bits are set to 1 cannot be assigned because this host ID is reserved as a broadcast address to send packets to all the hosts present on that particular network.

**Rules for assigning Network ID:**

Hosts that are located on the same physical network are identified by the network ID, as all host on the same physical network is assigned the same network ID. The network ID is assigned based on the following rules:

* + - The network ID cannot start with 127 because 127 belongs to class A address and is reserved for internal loop-back functions.
    - All bits of network ID set to 1 are reserved for use as an IP broadcast address and therefore, cannot be used.
    - All bits of network ID set to 0 are used to denote a specific host on the local network and are not routed and therefore, aren’t used.

**Summary of Classful addressing :**



**Problems with Classful Addressing:**

The problem with this classful addressing method is that millions of class A address are wasted, many of the class B address are wasted, whereas, number of addresses available in class C is so small that it cannot cater the needs of organizations. Class D addresses are used for multicast routing and are therefore available as a single block only. Class E addresses are reserved.

Since there are these problems, Classful networking was replaced by Classless Inter-Domain Routing (CIDR) in 1993. We will be discussing Classless addressing in next post.

Each IP class is equipped with its own default subnet mask which bounds that IP class to have prefixed number of Networks and prefixed number of Hosts per network. Classful IP addressing does not provide any flexibility of having less number of Hosts per Network or more Networks per IP Class.

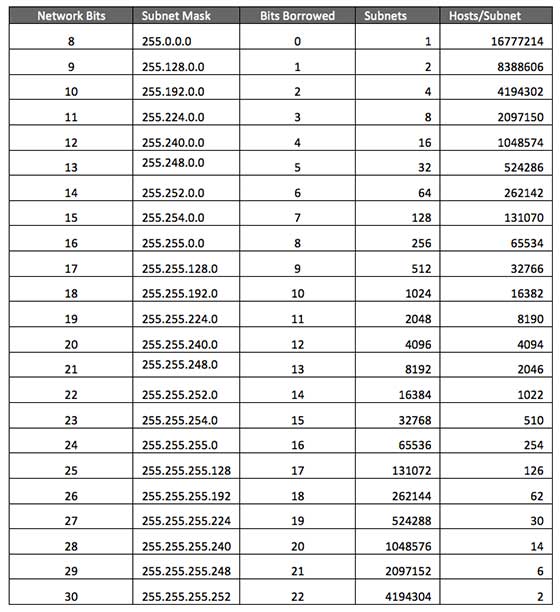
CIDR or **Classless Inter Domain Routing** provides the flexibility of borrowing bits of Host part of the IP address and using them as Network in Network, called Subnet. By using subnetting, one single Class A IP address can be used to have smaller sub-networks which provides better network management capabilities.

## Class A Subnets

In Class A, only the first octet is used as Network identifier and rest of three octets are used to be assigned to Hosts (i.e. 16777214 Hosts per Network). To make more subnet in Class A, bits from Host part are borrowed and the subnet mask is changed accordingly.

For example, if one MSB (Most Significant Bit) is borrowed from host bits of second octet and added to Network address, it creates two Subnets (21=2) with (223-2) 8388606 Hosts per Subnet.

The Subnet mask is changed accordingly to reflect subnetting. Given below is a list of all possible combination of Class A subnets −



In case of subnetting too, the very first and last IP address of every subnet is used for Subnet Number and Subnet Broadcast IP address respectively. Because these two IP addresses cannot be assigned to hosts, sub-netting cannot be implemented by using more than 30 bits as Network Bits, which provides less than two hosts per subnet.

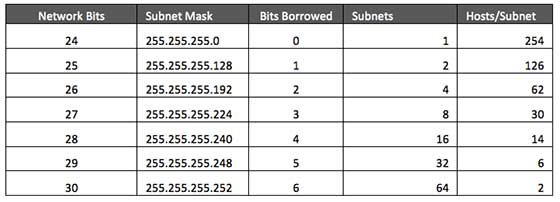
## Class B Subnets

By default, using Classful Networking, 14 bits are used as Network bits providing (214) 16384 Networks and (216-2) 65534 Hosts. Class B IP Addresses can be subnetted the same way as Class A addresses, by borrowing bits from Host bits. Below is given all possible combination of Class B subnetting −



## Class C Subnets

Class C IP addresses are normally assigned to a very small size network because it can only have 254 hosts in a network. Given below is a list of all possible combination of subnetted Class B IP address −



Introduction of Variable Length Subnet Mask (VLSM)

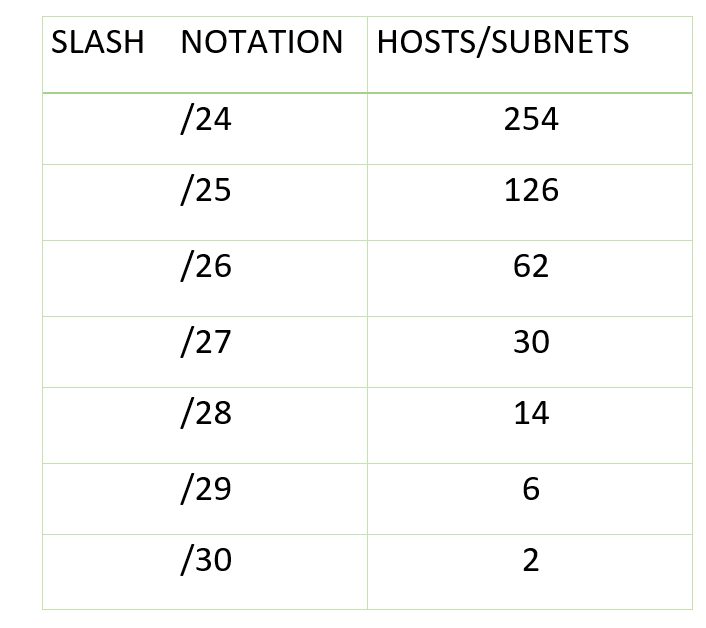
Last Updated: 04-02-2020

VLSM stands for Variable Length Subnet Mask where the subnet design uses more than one mask in the same network which means more than one mask is used for different subnets of a single class A, B, C or a network. It is used to increase the usability of subnets as they can be of variable size. It is also defined as the process of subnetting of a subnet.

**Procedure of implementing VLSM –**  
In VLSM, subnets use block size based on requirement so subnetting is required multiple times. Suppose there is an administrator that has four departments to manage. These are sales and purchase department with 120 computers, development department with 50 computers, accounts department with 26 computers and management department with 5 computers.

If the administrator has IP 192.168.1.0/24, department wise IPs can be allocated by following these steps:

1. For each segment select the block size that is greater than or equal to the actual requirement which is the sum of host addresses, broadcast addresses and network addresses. Make a list of subnets possible:

  
**table –** possible subnets list

1. Arrange all the segments in descending order based on the block size that is from highest to lowest requirement.
2. Sales and Purchase: 120
3. Development: 50
4. Accounts: 26

Management: 5

1. The highest IP available has to be allocated to highest requirement so the sales and purchase department gets 192.168.1.0/25 which has 126 valid addresses that can easily be available for 120 hosts. The subnet mask used is 255.255.255.128
2. The next segment requires an IP to handle 50 hosts. The IP subnet with network number 192.168.1.128/26 is the next highest which can be assigned to 62 hosts thus fulfilling the requirement of development department. The subnet mask used is 255.255.255.192
3. Similarly the next IP subnet 192.168.1.192/27 can fulfill the requirements of accounts department as it has 30 valid hosts IP which can be assigned to 26 computers.The mask used is 255.255.255.224
4. The last segment requires 5 valid hosts IP which can be fulfilled by the subnet 192.168.1.224/29 which has the mask as 255.255.255.248 is chosen as per the requirement. The IP with the mask 255.255.255.240 could be chosen but it has 14 valid hosts IPs and the requirement is less in comparison so the one that is comparable with the requirement is chosen . Thus there is less IP wastage in VLSM as compared to FLSM.

**Advantages of VLSM over FLSM –**

1. In Fixed length subnet mask subnetting (FLSM), all subnets are of equal size and have equal number of hosts but in VLSM the size is variable and it can have variable number of hosts thus making the IP addressing more efficient by allowing a routed system of different mask length to suit requirements.
2. In FLSM there is a wastage of IP addresses but in VLSM there is a minimum wastage of IP addresses.
3. FLSM is preferred for private IP addresses while for public IP addresses VLSM is the best option.

### Differences between FLSM Subnetting and VLSM Subnetting

|  |  |
| --- | --- |
| FLSM (Fixed Length Subnet Masks) Subnetting | VLSM (Variable Length Subnet Masks) Subnetting |
| All subnets are equal in size. | Subnets are variable in size. |
| All subnets have equal number of hosts. | Subnets have variable number of hosts. |
| All subnets use same subnet mask. | Subnets use different subnet masks. |
| It is easy in configuration and administration. | It is complex in configuration and administration. |
| It wastes a lot of IP addresses. | It wastes minimum IP addresses. |
| It is also known as classfull Subnetting. | It is also known as classless Subnetting. |
| It supports both classfull and classless routing protocols. | It supports only classless routing protocols. |

# What is IPv4?

Last Updated: 14-05-2020

**IP** stands for **Internet Protocol** and **v4** stands for **Version Four** (IPv4). IPv4 was the primary version brought into action for production within the ARPANET in 1983.  
IP version four addresses are 32-bit integers which will be expressed in hexadecimal notation.  
Example- 192.0.2.126 could be an IPv4 address.

#### Parts of IPv4

* **Network part:**  
  The network part indicates the distinctive variety that’s appointed to the network. The network part conjointly identifies the category of the network that’s assigned.
* **Host Part:**  
  The host part uniquely identifies the machine on your network. This a part of the IPv4 address is assigned to every host.

For each host on the network, the network part is the same, however, the host half must vary.

* **Subnet number:**  
  This is the nonobligatory part of IPv4. Local networks that have massive numbers of hosts are divided into subnets and subnet numbers are appointed to that.

#### Characteristics of IPv4

* IPv4 could be a 32-Bit IP Address.
* IPv4 could be a numeric address, and its bits are separated by a dot.
* The number of header fields are twelve and the length of the header filed is twenty.
* It has Unicast, broadcast, and multicast style of addresses.
* IPv4 supports VLSM (Virtual Length Subnet Mask).
* IPv4 uses the Post Address Resolution Protocol to map to mack address.
* RIP may be a routing protocol supported by the routed daemon.
* Networks ought to be designed either manually or with DHCP.
* Packet fragmentation permits from routers and causing host.

#### Advantages of IPv4

* IPv4 security permits encryption to keep up privacy and security.
* IPV4 network allocation is significant and presently has quite 85000 practical routers.
* It becomes easy to attach multiple devices across an outsized network while not NAT.
* This is a model of communication so provides quality service also as economical knowledge transfer.
* IPV4 addresses are redefined and permit flawless encoding.
* Routing is a lot of scalable and economical as a result of addressing is collective more effectively.
* Data communication across the network becomes a lot of specific in multicast organizations.

#### Disadvantages of IPv4

* + Limits net growth for existing users and hinders the use of the net for brand new users.
  + Internet Routing is inefficient in IPv4.
  + IPv4 has high System Management prices and it’s labor intensive, complex, slow & frequent to errors.
  + Security features are nonobligatory.
  + Difficult to feature support for future desires as a result of adding it on is extremely high overhead since it hinders the flexibility to attach everything over IP.

Introduction and IPv4 Datagram Header

Last Updated: 13-01-2020

The network layer is the third layer (from bottom) in the OSI Model. The network layer is concerned with the delivery of a packet across multiple networks. The network layer is considered the backbone of the OSI Model. It selects and manages the best logical path for data transfer between nodes. This layer contains hardware devices such as routers, bridges, firewalls, and switches, but it actually creates a logical image of the most efficient communication route and implements it with a physical medium. Network layer protocols exist in every host or router. The router examines the header fields of all the IP packets that pass through it. Internet Protocol and Netware IPX/SPX are the most common protocols associated with the network layer.  
In the OSI model, the network layer responds to requests from the layer above it (transport layer) and issues requests to the layer below it (data link layer).

**Responsibilities of Network Layer:**

***Packet forwarding/Routing of packets:****Relaying of data packets from one network segment to another by nodes in a computer network*

***Connectionless communication(IP):****A data transmission method used in packet-switched networks in which each data unit is separately addressed and routed based on information carried by it*

***Fragmentation of data packets:****Splitting of data packets that are too large to be transmitted on the network*

There are two types of network transmission techniques, circuit switched network and packet switched network.  
**Circuit Switch vs Packet Switch**  
In circuit switched network, a single path is designated for transmission of all the data packets. Whereas in case of a packet-switched network, each packet may be sent through a different path to reach the destination.

In a circuit switched network, the data packets are received in order whereas in a packet switched network, the data packets may be received out of order.

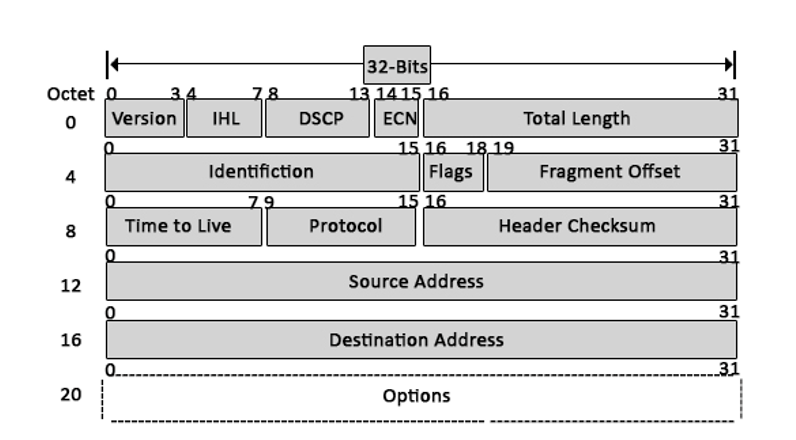
The packet switching is further subdivided into Virtual circuits and Datagram.

**IPv4:**  
IPv4 is a connectionless protocol used for packet-switched networks. It operates on a best effort delivery model, in which neither delivery is guaranteed, nor proper sequencing or avoidance of duplicate delivery is assured. Internet Protocol Version 4 (IPv4) is the fourth revision of the Internet Protocol and a widely used protocol in data communication over different kinds of networks. IPv4 is a connectionless protocol used in packet-switched layer networks, such as Ethernet. It provides a logical connection between network devices by providing identification for each device. There are many ways to configure IPv4 with all kinds of devices – including manual and automatic configurations – depending on the network type.

IPv4 is defined and specified in IETF publication RFC 791.  
IPv4 uses 32-bit addresses for Ethernet communication in five classes: A, B, C, D and E. Classes A, B and C have a different bit length for addressing the network host. Class D addresses are reserved for military purposes, while class E addresses are reserved for future use.

IPv4 uses 32-bit (4 byte) addressing, which gives 232 addresses. IPv4 addresses are written in the dot-decimal notation, which comprises of four octets of the address expressed individually in decimal and separated by periods, for instance, 192.168.1.5.

**IPv4 Datagram Header**  
Size of the header is 20 to 60 bytes.



***VERSION:****Version of the IP protocol (4 bits), which is 4 for IPv4*

***HLEN:****IP header length (4 bits), which is the number of 32 bit words in the header. The minimum value for this field is 5 and the maximum is 15.*

***Type of service:****Low Delay, High Throughput, Reliability (8 bits)*

***Total Length:****Length of header + Data (16 bits), which has a minimum value 20 bytes and the maximum is 65,535 bytes.*

***Identification:****Unique Packet Id for identifying the group of fragments of a single IP datagram (16 bits)*

***Flags:****3 flags of 1 bit each : reserved bit (must be zero), do not fragment flag, more fragments flag (same order)*

***Fragment Offset:****Represents the number of Data Bytes ahead of the particular fragment in the particular Datagram. Specified in terms of number of 8 bytes, which has the maximum value of 65,528 bytes.*

***Time to live:****Datagram’s lifetime (8 bits), It prevents the datagram to loop through the network by restricting the number of Hops taken by a Packet before delivering to the Destination.*

***Protocol:****Name of the protocol to which the data is to be passed (8 bits)*

***Header Checksum:****16 bits header checksum for checking errors in the datagram header*

***Source IP address:****32 bits IP address of the sender*

***Destination IP address:****32 bits IP address of the receiver*

***Option:****Optional information such as source route, record route. Used by the Network administrator to check whether a path is working or not.*

Due to the presence of options, the size of the datagram header can be of variable length (20 bytes to 60 bytes).

IPv4 Datagram Fragmentation and Delays

Last Updated: 12-08-2019

**Why IPv4 Datagram Fragmentation required?**  
Different Networks may have different maximum transmission unit (MTU), for example due to differences in LAN technology. When one network wants to transmit datagrams to a network with a smaller MTU, the routers on path may fragment and reassemble datagrams.

**How is Fragmentation done?**  
When a packet is received at the router, destination address is examined and MTU is determined. If size of the packet is bigger than the MTU, and the ‘Do not Fragment (DF)’ bit is set to 0 in header, then the packet is fragmented into parts and sent one by one. The maximum size of each fragment is the MTU minus the IP header size (Minimum 20 bytes and Maximum 60 bytes).

Each fragment is converted to a packet and the following changes happen in the datagram header:

1. The total length field is changed to the size of the fragment.
2. The More Fragment bit (MF bit) is set for all the fragment packets except the last one.
3. The fragment offset field is set, based on the number of fragment that is being set and the MTU.
4. Header Checksum is re-calculated.

**Example:** For a data packet of 4000 bytes and MTU of 1500 bytes, we have actual data of 3980 bytes that is to be transmitted and 1480 bytes is the maximum data size that is permissible to be sent. So, there would be 3 fragments:  
For the first fragment, data size = 1480 bytes, offset = 0 and MF flag = 1  
For the second fragment, data size = 1480 bytes, offset = 1480 and MF flag = 1  
For the third fragment, data size = 1020 bytes, offset = 2960 and MF flag = 0

An important point to be noted here is that all fragments would be having same identification number, thus indicating that all the fragments belong to the same parent data packet.

**Delays –**  
Processing delay: Time taken by the routers to process the data packet header.

Queuing delay: Time taken by the data packet in routing queues.

Transmission delay: Time taken to load a data packet onto the transmission channel  
Dt = N/R,  
N: Number of bits to be transmitted  
R: Rate or transmission speed of the channel

Propagation delay – Time taken by the data packet to reach from source to destination

Dp = D/S,  
D: Distance between the source and the destination  
S: is the speed of propagation

* **Version**: The first header field is a 4-bit version indicator. In the case of IPv4, the value of its four bits is set to 0100 which indicates 4 in binary.
* **Internet Header Length:** IHL is the 2nd field of an IPv4 header and it is of 4 bits in size. This header component is used to show how many 32-bit words are present in the header. As we know, IPv4 headers have a variable size so this is used to specify the size of the header to avoid any errors. This size can be between 20 bytes to 60 bytes.
* **Type of Service:** ToS is also called Differentiated Services Code Point or DSCP. This field is used to provide features related to the quality of service such as for data streaming or Voice over IP (VoIP) calls. It is used to specific how a datagram will be handled.
* **Explicit Congestion Notification:** ECN is used to send notifications to the sender or receive in situations where network congestion happens. This is an optional feature of IPv4 can if one of the endpoints don’t support it, it is not used.
* **Total Length:** Size of this field is 16 bit and it is used to denote the size of the entire datagram. The minimum size of an IP datagram is 20 bytes and at the maximum, it can be 65,535 bytes. Practically, all hosts are required to be able to read 576-byte datagrams. If a datagram is too large for the hosts in the network, fragmentation is used which is handled in the host or packet switch.
* **Identification:** Identification or ID field in a packet is used to uniquely identify fragments of an IP datagram. Some have suggested using this field for other things such as adding information for packet tracing etc.
* **Flags:** flag in an IPv4 header is a three-bit field that is used to control and identify fragments. The following can be their possible configuration:
  + Bit 0: this is reserved and has to be set to zero
  + Bit 1: DF or do not fragment
  + Bit 2: MF or more fragments.
* **Fragment Offset:** This field is 13 bit long in length and it is measured by blocks that units of 8-byte blocks. These are used to specify the offset of a fragment relative to the start of IP datagram which when it was not fragmented. As you can expect, the first offset of a fragment is always set to zero. The maximum possible offset is ( 213-1 ) \* 8 = 65528 but it is more than the maximum possible IP Packet length which is 65,535 bytes long with the length of a header added in.
* **Time to live:** Time to live (or TTL in short) is an 8-bit field to indicate the maximum time the datagram will be live in the internet system. The time here is measured in seconds and in case the value of TTL is zero, the datagram is erased. Every time a datagram is processed, it’s Time to live is decreased by one second. These are used so that datagrams that are not delivered are discarded automatically. TTL can be between 0 – 255.
* **Protocol:** This is a filed in the IPv4 header reserved to denote which protocol is used in the later (data) portion of the datagram. For Example, number 6 is used to denote TCP and 17 is used to denote UDP protocol.
* **The checksum of the header:** The checksum field is of 16-bit length and it is used to check the header for any errors. The header is compared to the value of its checksum at each hop and in case the header checksum is not matching, the packet is discarded. Keep in mind that this is only for the header and the data field is handled by its protocol. UDP and TCP, for example, have their own checksum fields.
* **Source Address:** It is a 32-bit address of the source of the IPv4 packet.
* **Destination Address:** the destination address is also 32 bit in size and it contains the address of the receiver.
* **Options**: This is an optional field of IPv4 header. It is used only when the value of IHL is set to more than 5. These options contain values and settings for things related to security. Record route and time stamp etc. In many cases, you will find that the list of options component ends with an End of Options or EOL.

# What is IPv6?

Last Updated: 26-05-2020

IP address is your digital identity. It’s a network address for your computer so the Internet knows where to send you emails, data, etc.

*IP address determines who and where you are in the network of billions of digital devices that are connected to the Internet.*



IPv6 or Internet Protocol Version 6 is a network layer protocol that allows communication to take place over the network. IPv6 was designed by Internet Engineering Task Force (IETF) in December 1998 with the purpose of superseding the IPv4 due to the global exponentially growing internet users.

#### IPv4 vs IPv6

The common type of IP address (is known as IPv4, for “version 4”). Here’s an example of what an IP address might look like:

25.59.209.224

An IPv4 address consists of four numbers, each of which contains one to three digits, with a single dot (.) separating each number or set of digits. Each of the four numbers can range from 0 to 255. This group of separated numbers creates the addresses that let you and everyone around the globe to send and retrieve data over our Internet connections. The IPv4 uses a 32-bit address scheme allowing to store 2^32 addresses which is more than 4 billion addresses. To date, it is considered the primary Internet Protocol and carries 94% of Internet traffic.  
Initially, it was assumed it would never run out of addresses but the present situation paves a new way to IPv6, let’s see why?

An IPv6 address consists of eight groups of four hexadecimal digits. Here’s an example IPv6 address:

3001:0da8:75a3:0000:0000:8a2e:0370:7334

This new IP address version is being deployed to fulfill the need for more Internet addresses. It was aimed to resolve issues which are associated with IPv4. With 128-bit address space, it allows 340 undecillion unique address space. IPv6 also called IPng (Internet Protocol next generation).

*IPv6 support a theoretical maximum of 340, 282, 366, 920, 938, 463, 463, 374, 607, 431, 768, 211, 456. To keep it straightforward, we will never run out of IP addresses again.*

#### Types of IPv6 Address

Now that we know about what is IPv6 address let’s take a look at its different types.

* **Unicast addresses**  
  It identifies a unique node on a network and usually refers to a single sender or a single receiver.
* **Multicast addresses**  
  It represents a group of IP devices and can only be used as the destination of a datagram.
* **Anycast addresses**  
  It is assigned to a set of interfaces that typically belong to different nodes.

#### Advantages of IPv6

* Reliability
* **Faster Speeds:** IPv6 supports multicast rather than broadcast in IPv4.This feature allows bandwidth-intensive packet flows (like multimedia streams) to be sent to multiple destinations all at once.
* **Stringer Security:** IPSecurity, which provides confidentiality, and data integrity, is embedded into IPv6.
* Routing efficiency
* Most importantly it’s the final solution for growing nodes in Global-network.

#### Disadvantages of IPv6

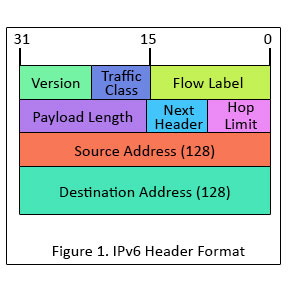
* **Conversion:** Due to widespread present usage of IPv4 it will take a long period to completely shift to IPv6.
* **Communication:** IPv4 and IPv6 machines cannot communicate directly with each other. They need an intermediate technology to make that possible.

## Introduction IPv6 Header Format

IPV6 header format is of 40 bytes in length, contains information essential to routing and delivery, consist of 8 fields, Version, Traffic Class, Flow Label, Payload length, next header, HOP limit, Source address and destination address, where each has its own features and provides essential data required to transmit the data. Version is 4 bit field, traffic class is of 8 bit, flow label is of 20 bits, payload length is 2 byte field, next header is of 8 bit, hop limit is 8 bit field, the source address is of 16 bytes and the destination address is of 16 bytes.

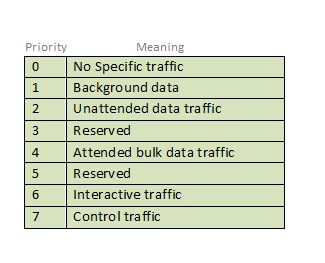
### List of IPv6 Header Format Component

IPv6 Header Format Component, the data packet of IPv6 encompasses two main parts i.e. header and the payload. The IPv6 consists of 40 bytes long fixed header which contains the following fields.



**Version (4-bits) :** Indicates version of Internet Protocol which contains bit sequence 0110.

**Traffic Class (8-bits) :** The Traffic Class field indicates class or priority of IPv6 packet which is similar to *Service Field* in IPv4 packet. It helps routers to handle the traffic based on priority of the packet. If congestion occurs on router then packets with least priority will be discarded.  
As of now only 4-bits are being used (and remaining bits are under research), in which 0 to 7 are assigned to Congestion controlled traffic and 8 to 15 are assigned to Uncontrolled traffic.

Priority assignment of Congestion controlled traffic :  


Uncontrolled data traffic is mainly used for Audio/Video data. So we give higher priority to Uncontrolled data traffic.  
Source node is allowed to set the priorities but on the way routers can change it. Therefore, destination should not expect same priority which was set by source node.

**Flow Label (20-bits) :**Flow Label field is used by source to label the packets belonging to the same flow in order to request special handling by intermediate IPv6 routers, such as non-default quality of service or real time service. In order to distinguish the flow, intermediate router can use source address, destination address and flow label of the packets. Between a source and destination multiple flows may exist because many processes might be running at the same time. Routers or Host that do not support the functionality of flow label field and for default router handling, flow label field is set to 0. While setting up the flow label, source is also supposed to specify the lifetime of flow.

**Payload Length (16-bits) :** It is a 16-bit (unsigned integer) field, indicates total size of the payload which tells routers about amount of information a particular packet contains in its payload. Payload Length field includes extension headers(if any) and upper layer packet. In case length of payload is greater than 65,535 bytes (payload up to 65,535 bytes can be indicated with 16-bits), then the payload length field will be set to 0 and jumbo payload option is used in the Hop-by-Hop options extension header.

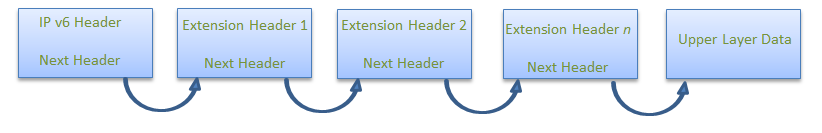
**Next Header (8-bits) :** Next Header indicates type of extension header(if present) immediately following the IPv6 header. Whereas In some cases it indicates the protocols contained within upper-layer packet, such as TCP, UDP.

**Hop Limit (8-bits) :** Hop Limit field is same as TTL in IPv4 packets. It indicates the maximum number of intermediate nodes IPv6 packet is allowed to travel. Its value gets decremented by one, by each node that forwards the packet and packet is discarded if value decrements to 0. This is used to discard the packets that are stuck in infinite loop because of some routing error.

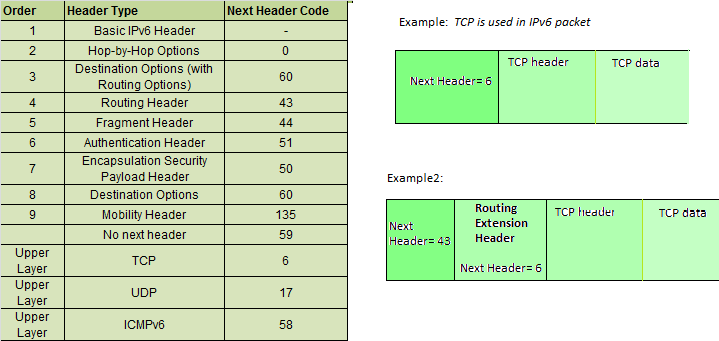
**Source Address (128-bits) :** Source Address is 128-bit IPv6 address of the original source of the packet.

**Destination Address (128-bits) :** Destination Address field indicates the IPv6 address of the final destination(in most cases). All the intermediate nodes can use this information in order to correctly route the packet.

**Extension Headers :** In order to rectify the limitations of *IPv4 Option Field*, Extension Headers are introduced in IPversion 6. The extension header mechanism is very important part of the IPv6 architecture. Next Header field of IPv6 fixed header points to the first Extension Header and this first extension header points to the second extension header and so on.



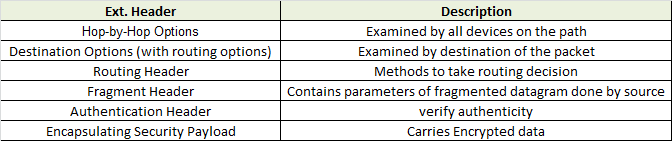
IPv6 packet may contain zero, one or more extension headers but these should be present in their recommended order:



**Rule :** Hop-by-Hop option header(if present) should always be placed after IPv6 base header.  
**Conventions :**

1. Any extension header can appear at most once except Destination Header because Destination Header is present two times in above list itself.
2. If Destination Header is present before Routing Header then it will be examined by all intermediate nodes specified in routing header.
3. If Destination Header is present just above Upper layer then it will be examined only by Destination node.

Given order in which all extension header should be chained in IPv6 packet and working of each extension header**:**



| **Basis for differences** | **IPv4** | **IPv6** |
| --- | --- | --- |
| Size of IP address | IPv4 is a 32-Bit IP Address. | IPv6 is 128 Bit IP Address. |
| Addressing method | IPv4 is a numeric address, and its binary bits are separated by a dot (.) | IPv6 is an alphanumeric address whose binary bits are separated by a colon (:). It also contains hexadecimal. |
| Number of header fields | 12 | 8 |
| Length of header filed | 20 | 40 |
| Checksum | Has checksum fields | Does not have checksum fields |
| Example | 12.244.233.165 | 2001:0db8:0000:0000:0000:ff00:0042:7879 |
| Type of Addresses | Unicast, broadcast, and multicast. | Unicast, multicast, and anycast. |
| Number of classes | IPv4 offers five different classes of IP Address. Class A to E. | lPv6 allows storing an unlimited number of IP Address. |
| Configuration | You have to configure a newly installed system before it can communicate with other systems. | In IPv6, the configuration is optional, depending upon on functions needed. |
| VLSM support | IPv4 support VLSM (Virtual Length Subnet Mask). | IPv6 does not offer support for VLSM. |
| Fragmentation | Fragmentation is done by sending and forwarding routes. | Fragmentation is done by the sender. |
| Routing Information Protocol (RIP) | RIP is a routing protocol supported by the routed daemon. | RIP does not support IPv6. It uses static routes. |
| Network Configuration | Networks need to be configured either manually or with DHCP. IPv4 had several overlays to handle Internet growth, which require more maintenance efforts. | IPv6 support autoconfiguration capabilities. |
| Best feature | Widespread use of NAT (Network address translation) devices which allows single NAT address can mask thousands of non-routable addresses, making end-to-end integrity achievable. | It allows direct addressing because of vast address Space. |
| Address Mask | Use for the designated network from host portion. | Not used. |
| SNMP | SNMP is a protocol used for system management. | SNMP does not support IPv6. |
| Mobility & Interoperability | Relatively constrained network topologies to which move restrict mobility and interoperability capabilities. | IPv6 provides interoperability and mobility capabilities which are embedded in network devices. |
| Security | Security is dependent on applications - IPv4 was not designed with security in mind. | IPSec(Internet Protocol Security) is built into the IPv6 protocol, usable with a proper key infrastructure. |
| Packet size | Packet size 576 bytes required, fragmentation optional | 1208 bytes required without fragmentation |
| Packet fragmentation | Allows from routers and sending host | Sending hosts only |
| Packet header | Does not identify packet flow for QoS handling which includes checksum options. | Packet head contains Flow Label field that specifies packet flow for QoS handling |
| DNS records | Address (A) records, maps hostnames | Address (AAAA) records, maps hostnames |
| Address configuration | Manual or via DHCP | Stateless address autoconfiguration using Internet Control Message Protocol version 6 (ICMPv6) or DHCPv6 |
| IP to MAC resolution | Broadcast ARP | Multicast Neighbour Solicitation |
| Local subnet Group management | Internet Group Management Protocol GMP) | Multicast Listener Discovery (MLD) |
| Optional Fields | Has Optional Fields | Does not have optional fields. But Extension headers are available. |
| IPSec | Internet Protocol Security (IPSec) concerning network security is optional | Internet Protocol Security (IPSec) Concerning network security is mandatory |
| Dynamic host configuration Server | Clients have approach DHCS (Dynamic Host Configuration server) whenever they want to connect to a network. | A Client does not have to approach any such server as they are given permanent addresses. |
| Mapping | Uses ARP(Address Resolution Protocol) to map to MAC address | Uses NDP(Neighbour Discovery Protocol) to map to MAC address |
| Combability with mobile devices | IPv4 address uses the dot-decimal notation. That's why it is not suitable for mobile networks. | IPv6 address is represented in hexadecimal, colon- separated notation. IPv6 is better suited to mobile networks. |

IPv4 and IPv6 cannot communicate with other but can exist together on the same network. This is known as**Dual Stack.**

Distance Vector Routing (DVR) Protocol

Last Updated: 06-11-2019

A **distance-vector routing (DVR)** protocol requires that a router inform its neighbors of topology changes periodically. Historically known as the old ARPANET routing algorithm (or known as Bellman-Ford algorithm).

**Bellman Ford Basics –** Each router maintains a Distance Vector table containing the distance between itself and ALL possible destination nodes. Distances,based on a chosen metric, are computed using information from the neighbors’ distance vectors.

Information kept by DV router -

* Each router has an ID
* Associated with each link connected to a router,
* there is a link cost (static or dynamic).
* Intermediate hops

Distance Vector Table Initialization -

* Distance to itself = 0
* Distance to ALL other routers = infinity number.

**Distance Vector Algorithm –**

1. A router transmits its distance vector to each of its neighbors in a routing packet.
2. Each router receives and saves the most recently received distance vector from each of its neighbors.
3. A router recalculates its distance vector when:
   * It receives a distance vector from a neighbor containing different information than before.
   * It discovers that a link to a neighbor has gone down.

The DV calculation is based on minimizing the cost to each destination

Dx(y) = Estimate of least cost from x to y

C(x,v) = Node x knows cost to each neighbor v

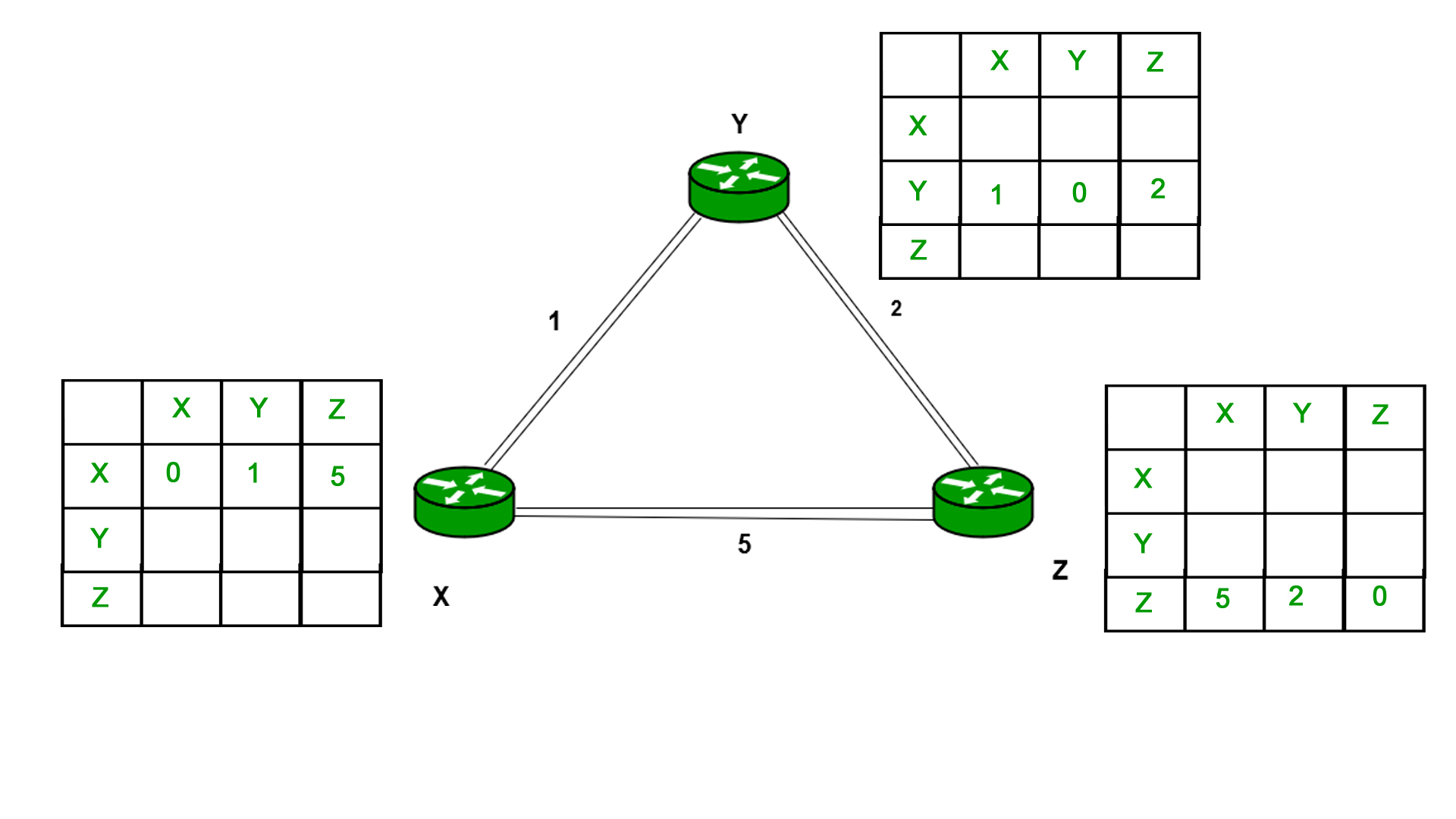
Dx = [Dx(y): y ∈ N ] = Node x maintains distance vector

Node x also maintains its neighbors' distance vectors

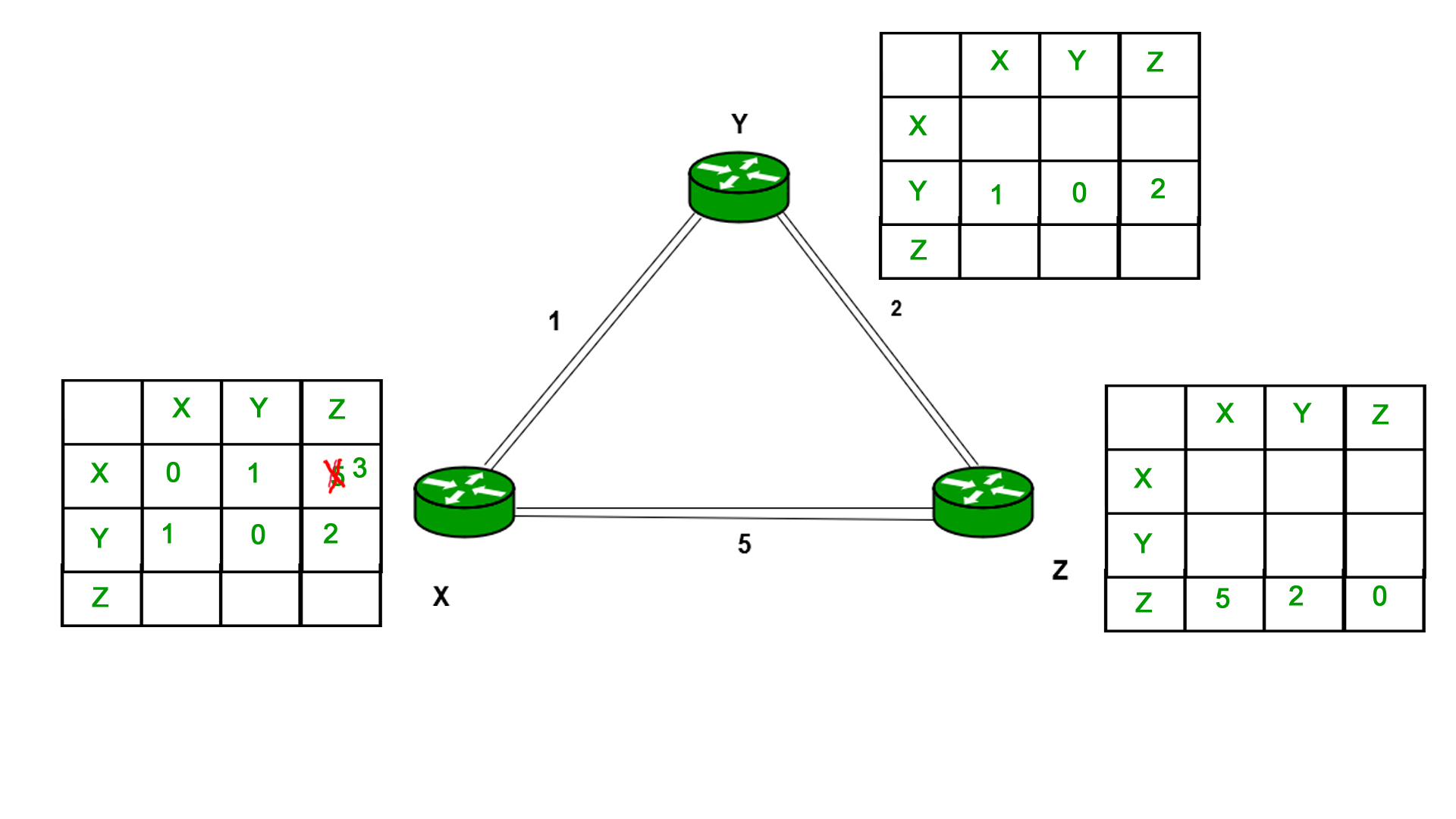
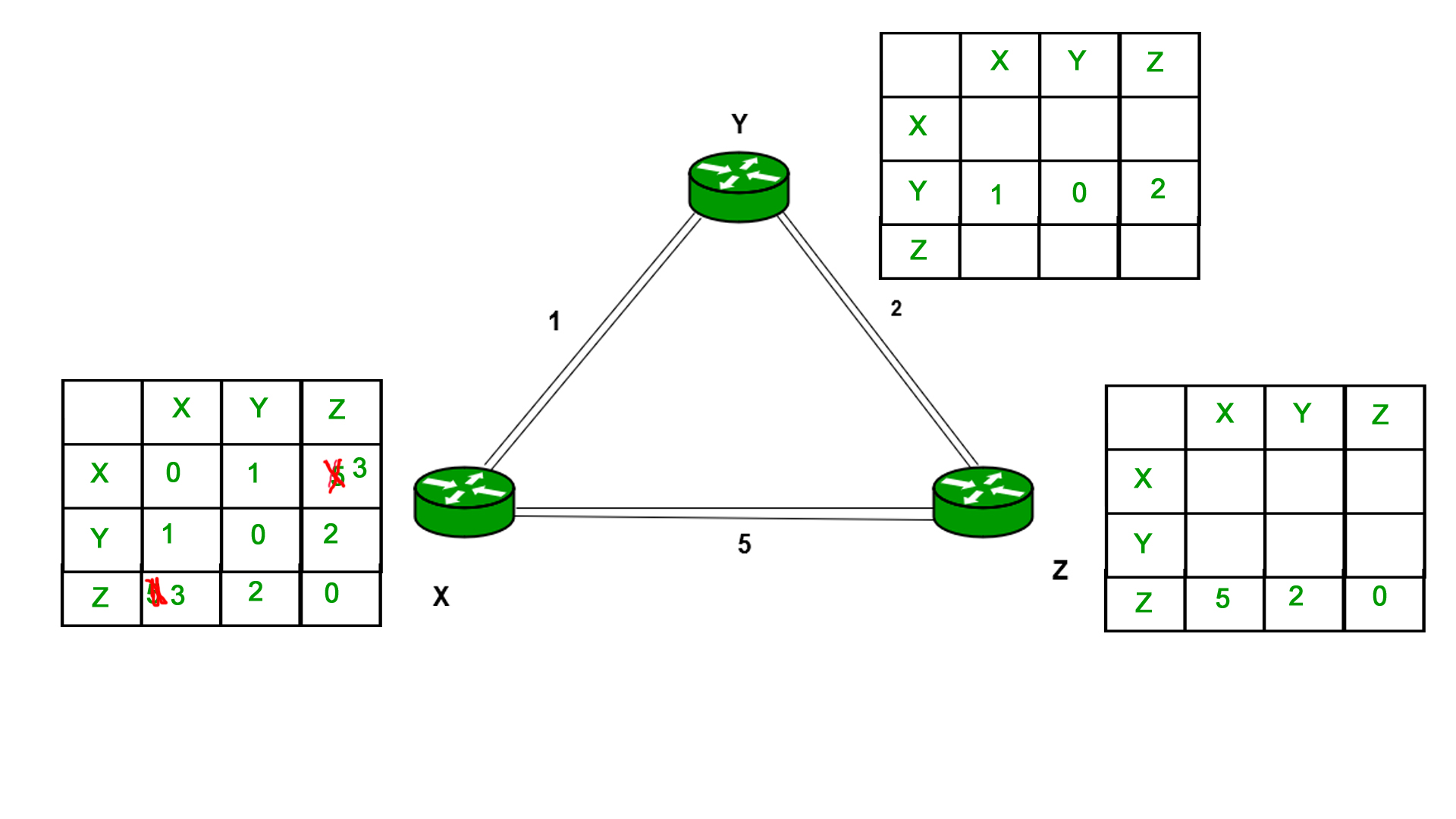
– For each neighbor v, x maintains Dv = [Dv(y): y ∈ N ]

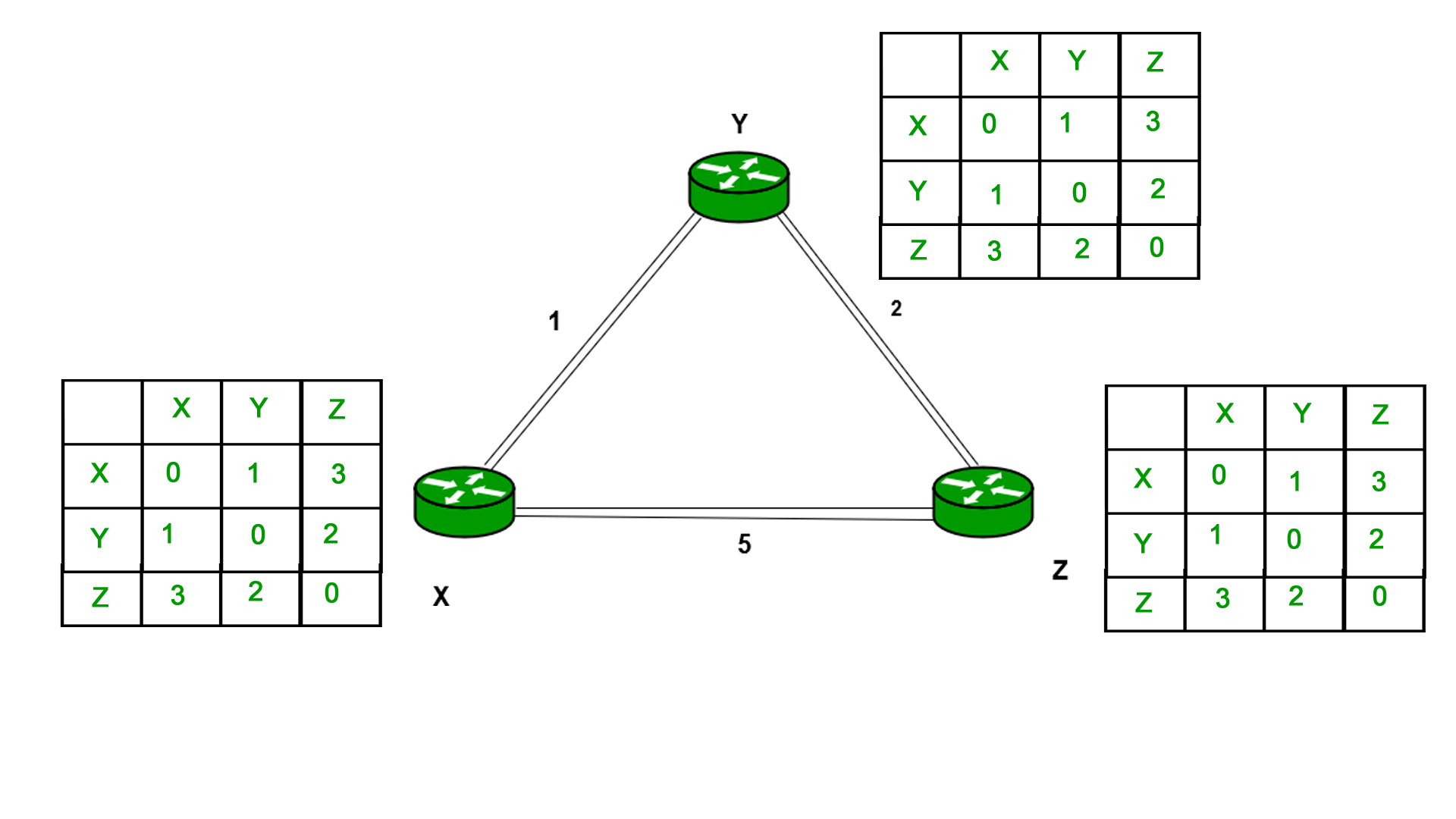
**Note –**

* From time-to-time, each node sends its own distance vector estimate to neighbors.
* When a node x receives new DV estimate from any neighbor v, it saves v’s distance vector and it updates its own DV using B-F equation:
* Dx(y) = min { C(x,v) + Dv(y), Dx(y) } for each node y ∈ N

**Example –** Consider 3-routers X, Y and Z as shown in figure. Each router have their routing table. Every routing table will contain distance to the destination nodes.  
  
Consider router X , X will share it routing table to neighbors and neighbors will share it routing table to it to X and distance from node X to destination will be calculated using bellmen- ford equation.

Dx(y) = min { C(x,v) + Dv(y)} for each node y ∈ N

As we can see that distance will be less going from X to Z when Y is intermediate node(hop) so it will be update in routing table X.  
  
Similarly for Z also –  


Finally the routing table for all –  
  
  
  
**Advantages of Distance Vector routing –**

* It is simpler to configure and maintain than link state routing.

**Disadvantages of Distance Vector routing –**

* + It is slower to converge than link state.
  + It is at risk from the count-to-infinity problem.
  + It creates more traffic than link state since a hop count change must be propagated to all routers and processed on each router. Hop count updates take place on a periodic basis, even if there are no changes in the network topology, so bandwidth-wasting broadcasts still occur.
  + For larger networks, distance vector routing results in larger routing tables than link state since each router must know about all other routers. This can also lead to congestion on WAN links.

**Note –**Distance Vector routing uses UDP(User datagram protocol) for transportation.

# Route Poisoning and Count to infinity problem in Routing

Last Updated: 12-08-2019

The main issue with **D**istance Vector **R**outing (DVR) protocols is Routing Loops, since [Bellman-Ford Algorithm](https://www.geeksforgeeks.org/dynamic-programming-set-23-bellman-ford-algorithm/) cannot prevent loops. This routing loop in DVR network causes Count to Infinity Problem. Routing loops usually occur when any interface goes down or two-routers send updates at the same time.

**Counting to infinity problem:**

So in this example, the Bellman-Ford algorithm will converge for each router, they will have entries for each other. B will know that it can get to C at a cost of 1, and A will know that it can get to C via B at a cost of 2.

If the link between B and C is disconnected, then B will know that it can no longer get to C via that link and will remove it from it’s table. Before it can send any updates it’s possible that it will receive an update from A which will be advertising that it can get to C at a cost of 2. B can get to A at a cost of 1, so it will update a route to C via A at a cost of 3. A will then receive updates from B later and update its cost to 4. They will then go on feeding each other bad information toward infinity which is called as **Count to Infinity problem**.

Routing Information Protocol (RIP)

Last Updated: 09-08-2019

**Routing Information Protocol** (RIP) is a dynamic routing protocol which uses hop count as a routing metric to find the best path between the source and the destination network. It is a distance vector routing protocol which has AD value 120 and works on the application layer of OSI model. RIP uses port number 520.

**Hop Count :**  
Hop count is the number of routers occurring in between the source and destination network. The path with the lowest hop count is considered as the best route to reach a network and therefore placed in the routing table. RIP prevents routing loops by limiting the number of hopes allowed in a path from source and destination. The maximum hop count allowed for RIP is 15 and hop count of 16 is considered as network unreachable.

**Features of RIP :**

1. Updates of the network are exchanged periodically.  
2. Updates (routing information) are always broadcast.  
3. Full routing tables are sent in updates.  
4. Routers always trust on routing information received from neighbor routers. This is also known as *Routing on rumours*.

# Open shortest path first (OSPF) router roles and configuration

Last Updated: 01-05-2018

Prerequisite – [Open shortest path first (OSPF)](https://www.geeksforgeeks.org/computer-network-open-shortest-path-first-ospf-protocol-states/)  
Open shortest path first (OSPF) is a link-state routing protocol which is used to find the best path between the source and the destination router using its own SPF algorithm.

**Open shortest path first (OSPF) router roles –**  
An area is a group of contiguous network and routers. Routers belonging to same area shares a common topology table and area I’d. The area I’d is associated with router’s interface as a router can belong to more than one area.

**Unicast –** Unicast means the transmission from a single sender to a single receiver. It is a point to point communication between sender and receiver. There are various unicast protocols such as TCP, HTTP, etc.

* TCP is the most commonly used unicast protocol. It is a connection oriented protocol that relay on acknowledgement from the receiver side.
* HTTP stands for Hyper Text Transfer Protocol. It is an object oriented protocol for communication.

There are three major protocols for unicast routing:

1. Distance Vector Routing
2. Link State Routing
3. Path-Vector Routing

**Link State Routing –**  
Link state routing is the second family of routing protocols. While distance vector routers use a distributed algorithm to compute their routing tables, link-state routing uses link-state routers to exchange messages that allow each router to learn the entire network topology. Based on this learned topology, each router is then able to compute its routing table by using a shortest path computation.

**Features of link state routing protocols –**

* **Link state packet –** A small packet that contains routing information.
* **Link state database –** A collection information gathered from link state packet.
* **Shortest path first algorithm (Dijkstra algorithm) –** A calculation performed on the database results into shortest path
* **Routing table –** A list of known paths and interfaces.

**Calculation of shortest path –**  
To find shortest path, each node need to run the famous **Dijkstra algorithm**. This famous algorithm uses the following steps:

* **Step-1:** The node is taken and chosen as a root node of the tree, this creates the tree with a single node, and now set the total cost of each node to some value based on the information in Link State Database
* **Step-2:** Now the node selects one node, among all the nodes not in the tree like structure, which is nearest to the root, and adds this to the tree.The shape of the tree gets changed .
* **Step-3:** After this node is added to the tree, the cost of all the nodes not in the tree needs to be updated because the paths may have been changed.
* **Step-4:** The node repeats the Step 2. and Step 3. until all the nodes are added in the tree

Link State protocols in comparison to Distance Vector protocols have:

1. It requires large amount of memory.
2. Shortest path computations require many CPU circles.
3. If network use the little bandwidth ; it quickly reacts to topology changes
4. All items in the database must be sent to neighbors to form link state packets.
5. All neighbors must be trusted in the topology.
6. Authentication mechanisms can be used to avoid undesired adjacency and problems.
7. No split horizon techniques are possible in the link state routing.

Link State Routing is an Intra-domain Routing Protocol ie it is used for routing within an autonomous system.

An autonomous system is a collection of networks or routers working under single administration.

In this protocol, each node of the network maintains entire topology of the network. To update the routing table ie for the formation of shortest path tree for each node, Dijkstra Algorithm is used which is a greedy single source shortest path algorithm.

Principle of LSR:- Tell to everyone what you know about your neighbors.

Moreover, It also uses flooding of LSR Packets for exchange of information.

**Open Shortest Path First (OSPF)** is an implementation of Link State Routing Protocol.