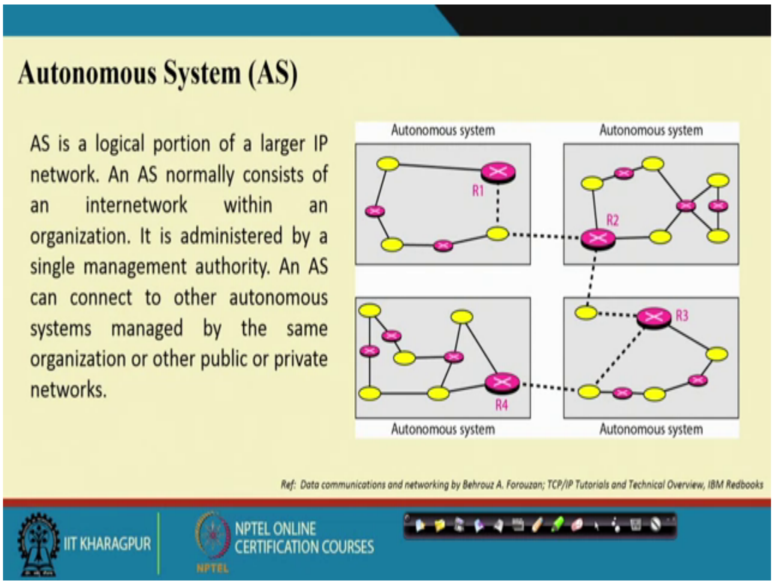
Computer Networks and Internet Protocol
Prof. Soumya Kanti Ghosh
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 40
IP Routing - IV (Border Gateway Protocol - BGP)

Hello, so we will continue our discussion on the course on Computer Networks and Internet. As we are last few lectures we are discussing on IP Routing, so, today we will continue that discussion. Today we will discuss on BGP protocol border gateway protocol. If you recollect that in the previous lectures or previous talks, we discussed on two category of things right: one is interior routing protocols and another is exterior routing protocols. So, the routing protocols, which are within the autonomous systems, and the routing protocols, that is across autonomous systems right. So, BGP or the border gateway protocol is this exterior this routing protocol, which helps in routing packets across autonomous systems right.

So, just a quick slide that is a autonomous system. So, how we define? It is a logical portion of a large IP network or the whole internet. And autonomous system normally consist of an internetwork within an organization control. That means, it is under one administrative control and it is administered by a single management authority like IIT Kharagpur, can be a autonomous systems which is managed by the IIT Kharagpur authority or the cell which is handling those thing.

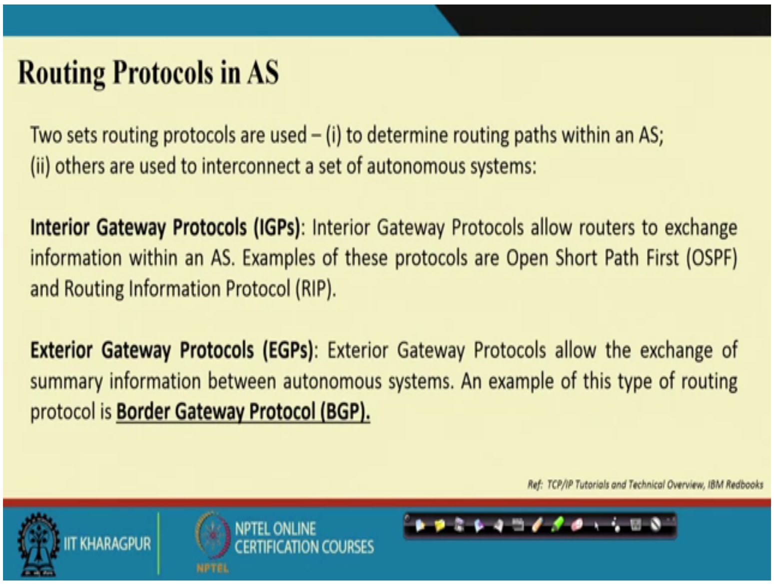
(Refer Slide Time: 01:55)



And an autonomous system can collect connect to other autonomous systems managed by the same organization or other organization public or private. In other sense our whole network is divided into several autonomous systems, which consist of several networks and there are routers or which takes care of the whole information of the autonomous system that, we have seen that there are there is a backbone area router or this backbone router which takes care of those type of things.

So, our basic objective is to how if a packet from a autonomous system here wants to communicate to a autonomous system in some other place, so how it will be routed. So, one is that there should be some internal routing protocol, which will take care of this within the autonomous system and there should be a protocol which goes across these things. If you recollect, so we talked about a path vector protocol which established a path between these different autonomous systems right. So, it is a sequential set of autonomous systems like I say that the packet 1 should travel by autonomous system 6, 8, 9, 10 or 6, 8, 9, 12 and so and so forth to reach the destination autonomous system.

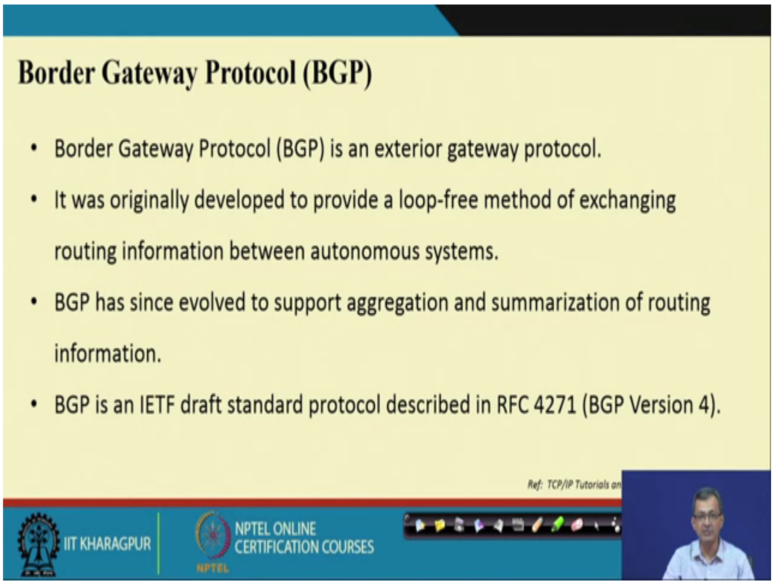
(Refer Slide Time: 03:24)



So, what we are talking about is, routing protocol is one categories interior gateway protocols IGPs that is interior gateway protocol allows router to exchange information within a autonomous system. Examples of these protocols are OSPF and RIP predominantly use protocol is OSPF. Exterior gateway protocol EGPs that is which are which follow a path vector protocol and it allows exchange of information across autonomous system.

So, one of the popular or the protocol for this EGP or exterior gateway protocols is border gateway protocol or BGP So, today we will be as we are discussing is the basic feature of BGP protocol.

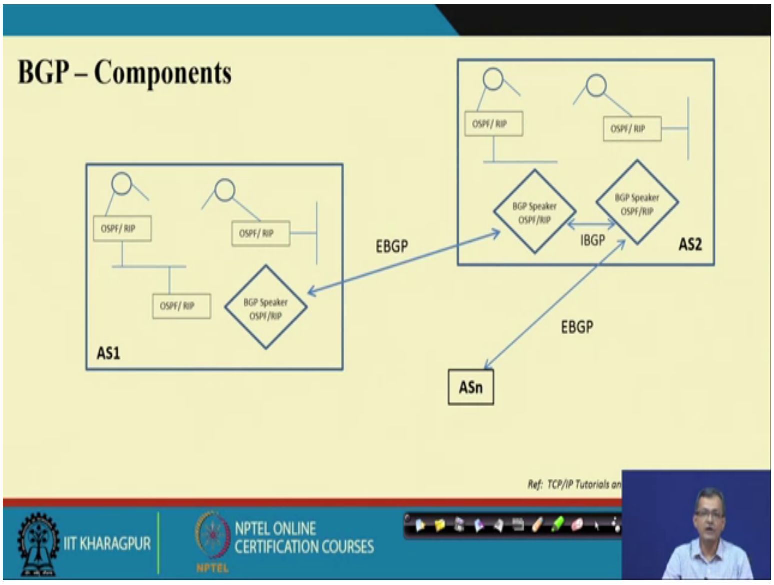
(Refer Slide Time: 04:26)



So, BGP as we are discuss is a exterior gateway protocol right. It was originally developed to provide look free method of exchanging information between autonomous systems, so that is why not be any loop right. BGP has since evolved to support aggregation and summarization of the routing informations right. So, if you see that if we look at the whole internet or network of networks, so there are in order to push this or any forwarding any information from one system or one network to another, it has to hop through several networks and autonomous systems right.

So, there should be some way of forwarding across autonomous system where, we try to look at this BGP. So, BGP presently the popular version is BGP 4 is an IETF standard protocol; that means, all BGP enabled devices also called BGP routers or BGP devices follow a standard which allows it to route packets right.

(Refer Slide Time: 05:35)

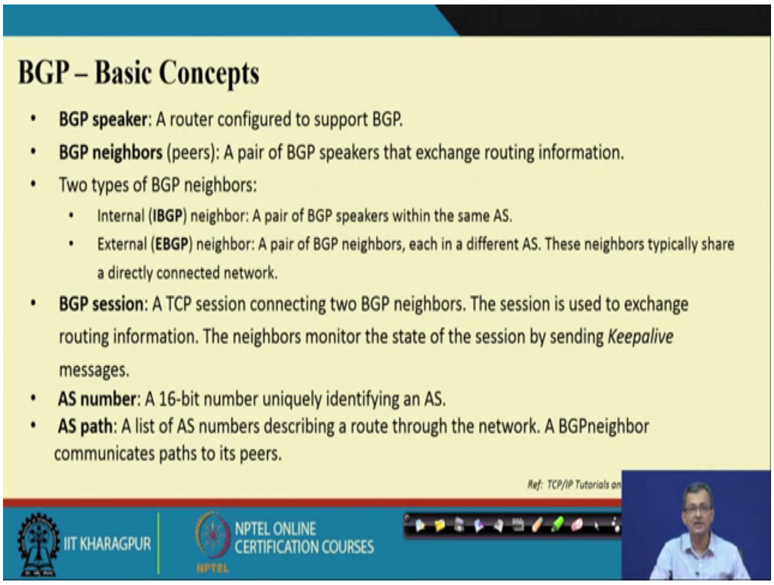


So, if we look at the BGP components, so we see if this is the autonomous system AS1, AS2, so, there are several thing one special device or things or router is called BGP speaker we will come to that. There can be several networks inside the things. Inside the autonomous system there are protocols called OSPF and RIP which can be within the autonomous system, as each autonomous system under single administrative control so, the administrator is free to choose which protocol to use right. And if you again recollect what we talked about the autonomous system is again divided into different area, out of that one area is or area 0 is the; or the backbone area right. And there are other areas, so that this backbone area or this backbone area router collects this information about these whole autonomous systems.

So, it is those designated routers, which takes care of the whole information base of the autonomous systems and it does not mean that that all this all traffic of the area should go through the area, but a router or all traffic should go through that router, but, but it what it means that it collects or it has the information about the whole autonomous systems.

Now, the within across the autonomous system we have the protocol all EBGP or between the within the autonomous systems we have this protocol called IBGP. We will come to those things right.

(Refer Slide Time: 07:32)



So, let us look at some of the components. So, one thing what we have seen is the BGP speaker. So, it is a router configured to support BGP. So, a device or a router which is configured to support BGP is a; what we says BGP speaker. There are BGP neighbors or peers that a pair a pair of BGP speakers that exchange routing informations are called peers right.

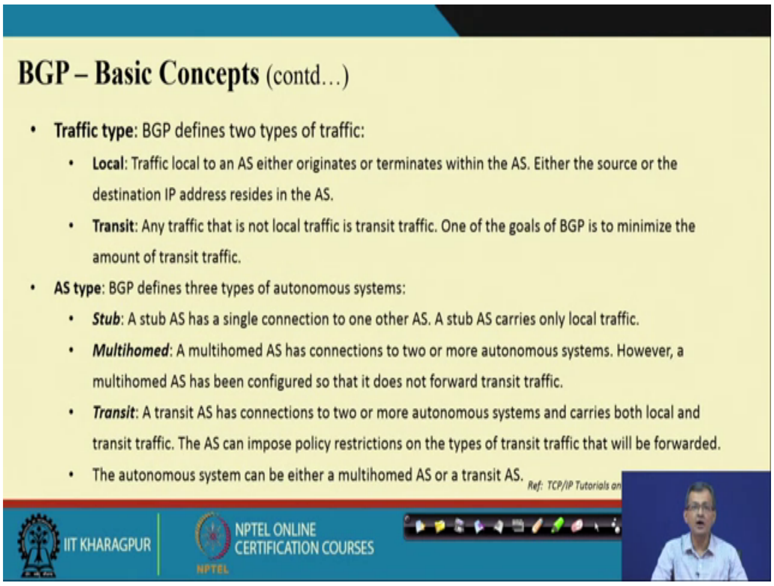
So, there are there can be two type of neighbors, the neighbors within the autonomous system and neighbor across the autonomous systems. There accordingly we have internal neighbor or IBGP a pair of BGP speakers within the autonomous systems and we have external neighbor or EBGP or a pair of BGP neighbors each in a different autonomous systems right So, if we again come back to this figure, so there is IBGP, this is the internal and these are the two BGP speaker or two routers which there can be n number of routers within this particular autonomous systems. So, these are IBGP and there these are EBGP right.

So, a pair of neighbors each in different days these neighbors typically share a directly connected network right. So, that is they are directly connected network scenario. There is a concept of BGP session, a TCP session connecting 2 BGP neighbors right. The season is used to exchange routing information and neighbors monitor the state of session by sending keep alive messages. So, periodically send that keep alive messages, so that monitor the state of the other sessions like. So, it is some sort of a pinging and after every or the coning at every at a regular interval.

There is as number for any autonomous systems, which is a 16 bit number uniquely defined a particular autonomous system. So, it is a so you can see that with the 16 we can have so, many number of AS, right. So, it is say two to the power 16 number of AS are possible that is a quite a large number considering that the AS are having several router, several network source inside the things. This is a terminology or there is a concept of as path list of number of numbers as number describing the route through the network a BGP neighbor communicates the path to its peers right.

So; that means, if I want to go from network say n 1 1 of AS 1 to network 6 2 of a n 6 2 of say a 6 then, what are the path I need to follow right. So, I start with AS 1 to AS 3 to as 4 to as 6 or some other paths type of things. So, it is a set of or sequential set of as numbers, which allow this routing within the network.

(Refer Slide Time: 10:50)



Now, if we look at it, there are if we look at the BGP traffic there are typically two type of traffic, one is local traffic, the traffic local to as either originates or terminates within the as right. So, it originates or terminus within the either the source or the destination resides in the AS. So, if the source or destination is residing in AS then, the traffic for that particular BGP session we say that is a local traffic. Whereas, in a transit any transit that is not local traffic is transit traffic. That means one goal of the BGP is to minimize the amount of transit traffic that means it is neither originating nor terminating the AS, but it is passing through this AS right.

So, in other sense, it is a traffic which is a load on that AS, so to say right. So, it is one of the objective may minimize the type of traffic. So, there are a lot of transit traffic means you overload the things right.

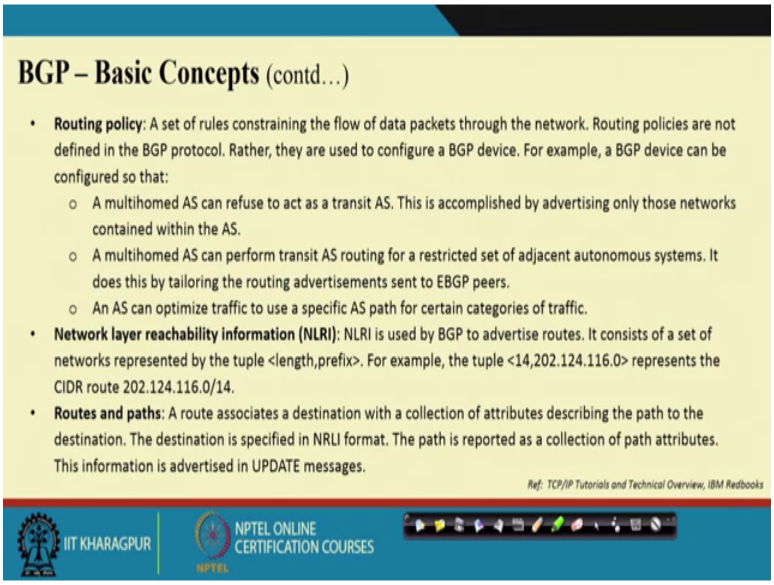
And if we look at the AS types, so BGP defies primarily 3 type of autonomous systems right; one is stub. A stub AS is a single connection to another as right escape stub AS carrying on only local traffic because, it is a one connection to the AS, but there is no other AS across that. That means, it is a stub connection so, that is only have local traffic there is no transit traffic, but say there is a multi homed AS. There is a multi homed AS has connections to 2 or more AS, AS's rather autonomous systems. However, a multi home AS has been configured. So, that it does not forward traffic. So, a multi home AS can be configured that, it does not transmit say what we say transit traffic, it may drop or block and type of things.

There is a transit AS a transit AS has connection of 2 or more autonomous systems like multi homed AS but carries both local and transit traffic. The AS that, impose policy restriction on the types of transit traffic that will be forwarded right. So, autonomous systems can be either multi homed AS or transit AS. So, primarily the autonomous systems are multi homed or transit stub is a special case of those scenarios.

So, what we see that primarily two types of things are there, there is one is multi homed and transit though there are more than one connection but in case of a multi homed you may not allow transit traffic. And there is a stub type of things, where the where there is no there is only local traffic either is originates or terminates.

And as we understand there can be policy as these are all path vector So, I can have policy defined that which can be transit to the things and like that right. So, there can be restriction on which sort of AS it should first type of thing that you allow it to transit the traffic.

(Refer Slide Time: 14:20)



There are few more concepts one is that routing policy that is important the set of rules constraining the flow of data packets through the network right. That is what we look at the routing policy. So, lighting policy are not defined parse in BGP protocol rather, they are used to configure a BGP device right.

In other sense the while, I am configuring the BGP device, I embed that policy into that the thing. For example, a BGP device can be configured, so that a multi home AS can refuse to act as a transit right. So, that is a policy, that I you not allow. This is accomplished by advertising only those networks, which are either trans terminating or relating that AS right So, it is confined within that particular AS.

So, a multi home AS can perform transit AS routing for restricted set of agent autonomous systems it does this by advertising or by tailoring the routing advertisement to send to the EBGP vapors. In other sense what it tries to say that, while advertising that routing information to the other EBGP routers, so what it does it does say tailoring of the things. That means, it modifies the things in a such a way that, which is within its policy paradigm.

Or an AS can optimize traffic to use specific AS path for certain category of traffic. Like I say I get a traffic up streaming video or a particular type of traffic then, it says that it should be channelized to a through a particular path. So, it can be configured or channelized through a particular path. So, that that can be one sort on such policy type of things.

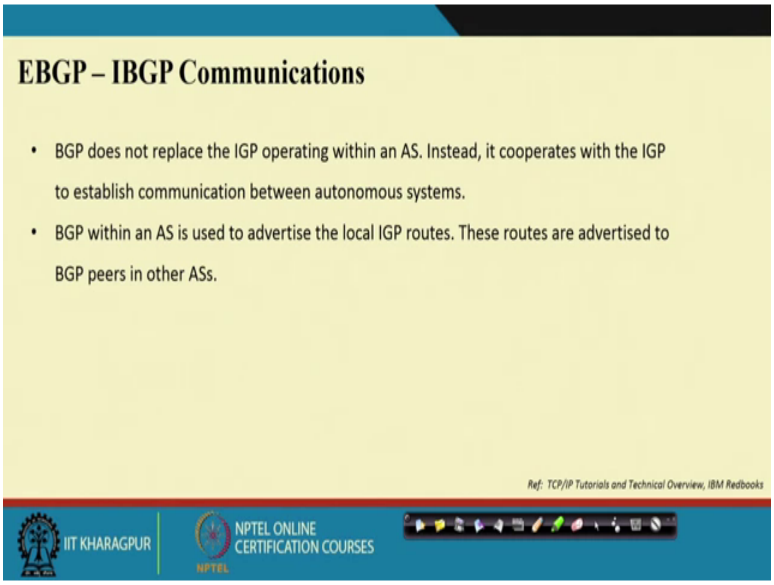
There is a concept of network layer reachability information.Network layer reachability information also known as NLRI, so NLRI is used by BGP to advertise routes right. So, it says that what is the network layer information; so it is consist of a set of network represented by a tuple that is length and prefix like it says that tuple 14 202 write one example represent that the CIDR route as this so and so forth right. So, it is network layer reachability information; that means that which port sent or which domain of the network you can access right. So, in this case that slash 14 is the thing.

Routes and paths with respect to BGP: a route associates a destination with a collection of attributes describing the path to the destination. A destination specified by NLRI format. The path is reported as a collection of path attributes this information advertised by the update things. So, now how do I define, I it is a defined by a set of AS. How this AS is defined is by a set of path parameters right, and the destination of the things where it will go because, when it when it goes through this BGP router it need to know that where the destination is that is in that NLRI format. And it goes on these this it basically, takes this path attributes along with the things right. So, through which the a for through which the which AS it will hops.

And when that this advertisement is for this advertisement, this update message format or the updates message protocol is used in the BGP; we will come to that that what is update. Actually we discussed little bit of what are the different type of BGP formats and that update is there any way we will discuss again.

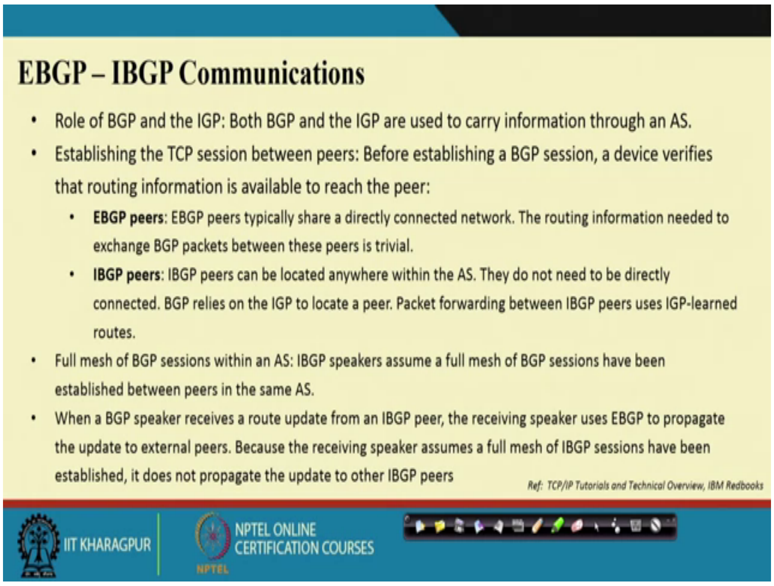
Now, so we have EBGP external BGP to communicate with the acrosses IBGP to communicate with in the AS. So, we have BGP I EBGP IBGP communication. Finally, to make this happen, so if you see it is not only external right, it once it enters it go on hopping with the networks and go through some other router to the type of things right. And this paths is also there are BGP with the for the external and there are OSPF or what do I say that, internal protocols are IGP protocols are running the device p every protocols are running. So, there should be a proper coordination between them right. Otherwise, the packet forwarding will be not will not be possible right.

(Refer Slide Time: 19:19)



So, BGP does not replace IGP operating within AS. So, it is not like that it replace OSPF of RIP type of protocol right instead, what we are saying coordinates or cooperates with the IGP to establish communication between autonomous systems right. So, BGP within AS is used to advertise the local IGP routes right. Within the AS it has to advertise the local IGP routes because, within the AS that is the predominant right. These routes are advertised to BGP peers in other AS right. So, this has to be known to the BGP peers in the other AS.

(Refer Slide Time: 20:04)



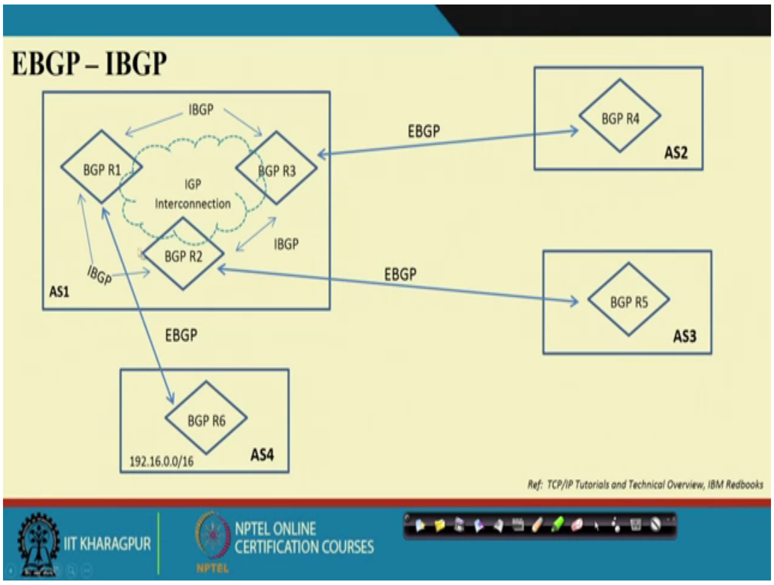
So, what we look at the communication. So, rollup BGP and IGP: so, both BGP and IGP that is border gateway protocol in that the internet gateway protocol. That is, OSPF and RIP are used to carry out information carry information through an AS right, so otherwise how the packet will be forwarded. Establishing TCP sessions between the two pairs, before establishing BGP session the device verifies that BGP devices, the routing information is available in each peer. So, it should be available in each peer BGP, EBGP peers that these EBGP peers statically a typically share a directly connected network.

So, these in case of a EBGP peers as we discussed couple of slides back it typically share a directly connected network. The routing information need to be exchanged by the BGP packets between these peers is trivial right it is directly connected thing. A for IGP IBGP peers can be located anywhere within the AS. So, they do not need to be directly connected. BGP relies on the IG IGP that is, the internal gateway protocols to locate a peer. Packet forwarded between the IBGP peers uses a IBGP learned routes. So that means, the what I what we mean to say that, OSPF or RIP protocols whatever, it is learned that those are with those information are being used by this BGP.

So, full mesh BGP sessions within the AS like IBGP speakers assume that, a full mesh BGP session has been established between the peers in the same AS right. So, it is a full meshed BGP sessions. So, everybody knows or everybody connected to others. So, it is a fully mesh connection. When a BGP speakers receives a router update, from a IBGP peer, the receiving speaker uses the EBGP propagate to update the external peers right. Because, the receiving speaker assume full mesh IBGP sessions have been established, it does not propagate the update to the other BGP peers.

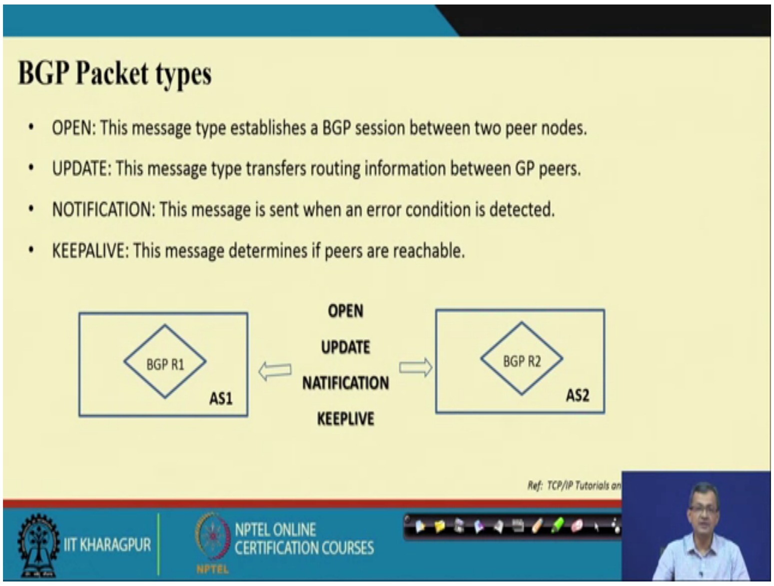
So, as this is a full meshed, so the full mesh IBGP sessions have been establish it does not propagate to the update to the other because, it knows that it the update has been taken care by this full this particular way of connectivity.

(Refer Slide Time: 22:37)



So, the same thing what we were discussing; so if we have so there are there are several BGP routers. So, there is a IBGP protocols which are connected across the thing. And this when this protocol BGP this BGP device R6 connected to the a BGP R1 through these IBGP updated to BGP R3 and goes like that and so and so forth right. So, that is the way of connectivity.

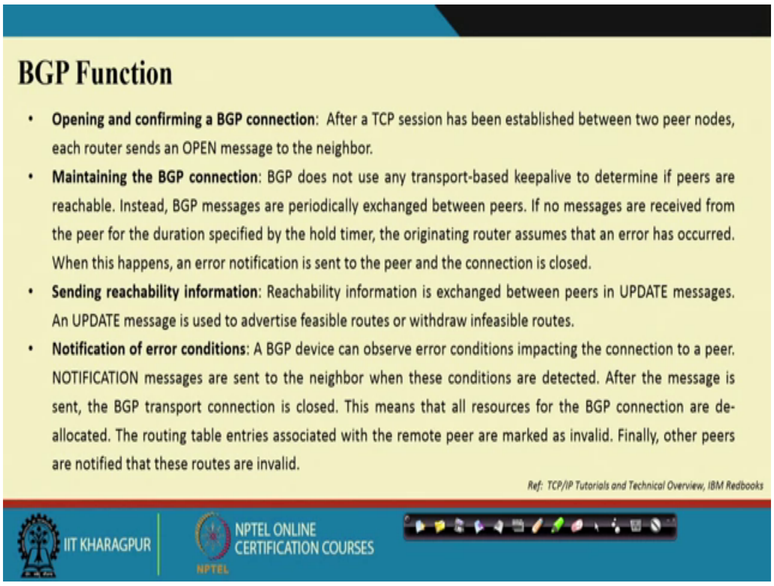
(Refer Slide Time: 23:10)



Now, as we mentioned earlier, so what are the different BGP packet types right. So, one is open, other is update, notification, keep alive. So, these are the typical four types of packets with the BGPs. The open is that this message type establish a BGP session between two peer nodes right. So, this is through establishing a BGP session between two peer nodes that is the open. Update this message type transfer, routing information between the BGP peers, so, there is a typo it should be BGP peers. So, this is this message information that is update is basically, routing information from between the BGP peers.

Notification this is when a error occurs in is detected. So, if there is any abnormal situation, so it requires a notification of the thing. Keep alive, this determines if the peers are reachable or not. So, it is some sort of a beckoning type of things which goes on things. So, one is the open opening the season between the two BGP peers another is the update, updating routing information across the BGP peers right. Another is what we so there is again a typo it should be notification in this here it should be notification. So, then another message is notification, this message is sent when a error occurs your error is detected. And then, other is a keep alive that is the beckoning say that the peers are reachable or not to determine whether the peers are reachable or not, we have those things.

(Refer Slide Time: 24:54)



Now, if we look at the different functionalities, one is opening and confirming BGP connection is the functionality. Primarily after a TCP session has been established between the two BGP peers each router sends a open message to the neighbor to its neighbor. That is the opening and confirming BGP connection as we mentioned earlier that is establishing the connection. Another is maintaining BGP connection so on it is open and thing is then next is maintaining the BGP connection. That BGP does not use any transport layer keep alive to determine if the peers are reachable or not. Instead BGP message are periodically exchanged between the peers right.

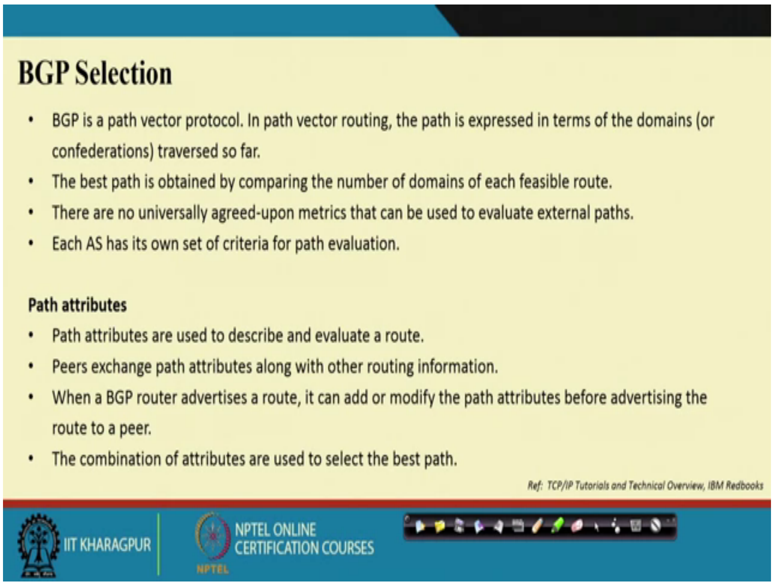
If no messages are received from the peer for duration specified by the whole time that hold timer the originating router assumed that an error has occurred right. When this happens the error notification is sense to the things. So, it is go on beckoning between the peers whenever, whenever it is not receiving within a particular time period So, it goes on a it thinks that there is the BGP that the peer BGP device is not responding and accordingly a error notification or error or a condition is generated and the information is sent to the things.

Sending reachability information, reachability information is exchanged by the update message right. As we seen that readability informations is primarily finding that more that is the routing information. If there is a information change in the information or updating the reachability information it is exchanged between the peers in the update message.

So, update message is used to advertise feasible routers to or withdraw infeasibility routers. So, both advertising the possible feasible routers, and withdrawing infeasible routers. So, notification of error conditions, A BGP device can observe error can observe error condition impacting the connection to a peer. So that means it is it may not be responding or some other error situation arises. So, notification message has sent to the neighbor or sent to the neighbor, when the conditions are detected. After the message is sent the BGP transport connection is closed. This means, all the resources for the BGP connection are de allocated right. So that means, once it is notification since the particular connection to that particular BGP peer of the device is closed and the resource says, provisioned or allocated for that things are released right.

So, the routing table entries associated with the remote peer are mark invalid finally, other peers are notified. So, you see so once we see that the connection with one BGP router or the peer device is there, this is notified and the connection is closed, all resources are released. And now it has a new update right. So, it has to find a new path and it is being advertised to the other or marked as invalid that particular path and it is advertised to the other peers right. So, that is important.

(Refer Slide Time: 28:22)

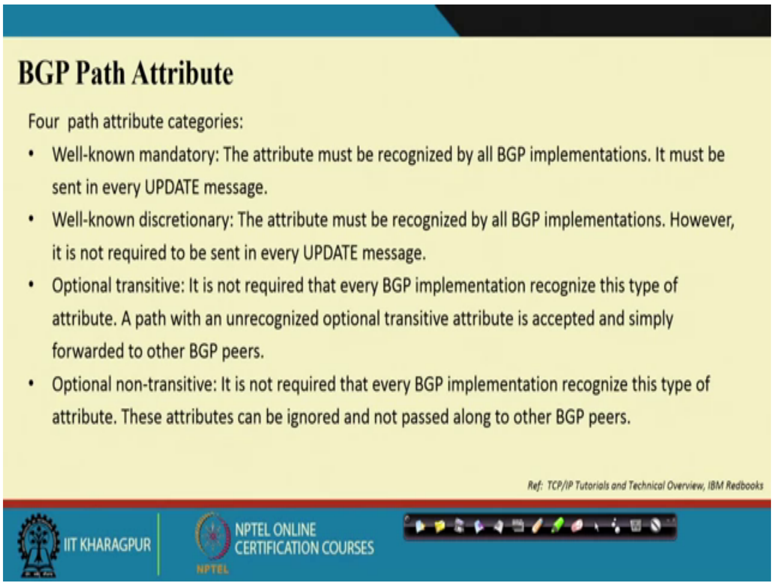


So, there are other things like BGP selection, BGP is a path vector protocol as we discussed. So, in a path vector the path is expressed in terms of domains or configurations transferred. So, path has to be defined how. So, path is defined a series of ways within the properties and type of things or which domain it is hopping through. The best path is obtained by comparing the number of domains of each feasible routes right. So, one is that how many domains need to be hopped. There is no universal agreed upon metric that can be used to evaluate the external path So, that is important that is difficult because, each as had his own criteria path evaluation right So, there is no universal metric because, it is a large network right it say large networks of networks and these there are several AS, so each as has the own way of finding the optimal path.

So, which is based on the path attributes. So, there are several path attributes are used to describe and evaluate a path, peers exchange path attributes along with other routing informations So, when a BGP router advise a route, it can add or modify the path attributes before advertising the route to a peer So, once a BGP router receives a this update then, while updating the things it can basically update the path attributes before transmitting it to the peer.

The combination of attribute, are used to select the best path right. The combination of attributes, are used to select the best path of the things. That means, what we want to say this your path is defined by these several AS which are in turn defined by the attribute set of this AS right; or that particular path and which can be used this attributes values are used to find the optimal path based on the particular by based on the policy of the that given AS.

(Refer Slide Time: 30:39)

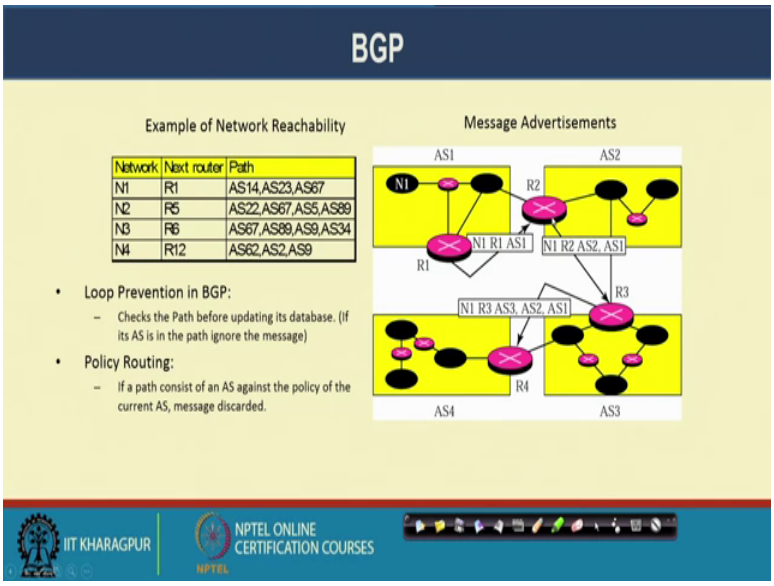


So, there are we also discussed in earlier thing, we will go quickly that is there are four path attribute categories, as we have discussed in the path vector protocol. One is well known mandatory the attribute must be recognized by all BGP implementation, is must be same for every update message. This is a well known and mandatory well known discretionary, that is the attribute must be recognized by all BGP implementation. However, it is not required to send for every message. So, it is a discretion or discretionary optional transitive that is, it is not required for every BGP implementation to recognize this type of attribute right it is of optional. By a path with a unrecognized optional transitive attribute is accepted and simply forwarded to the BGP peers. So, if it is unrecognized optional attribute transitive attribute; that means, it is being transmitted to the peer without doing any analysis or without taking any call on it right N.

Optional non transitive if it is not required for every BGP implementation recognize this type of attribute. These attributes can be ignored and not passed along with the things.

So, it is one is optional transitive, that means, it is optional, but it may not recognize, but it can transmit. Another is optional in non transitive, if it is if it is not recognized or it can be ignored and may not be transmitted to the other side.

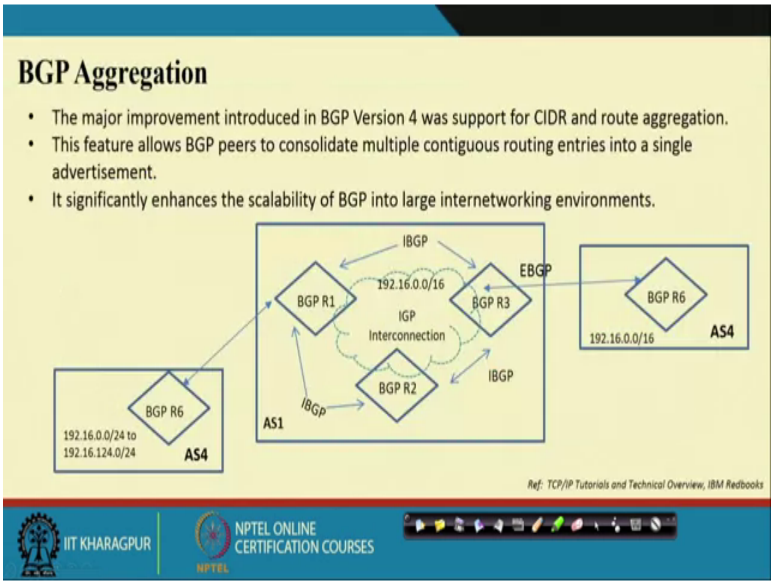
(Refer Slide Time: 32:13)



And if we look at that a previous lecture or previous to previous lecture we discussed. So what we see that this is the N1, R1, AS1 this is the particular reachability of the things. So, in order from here if I want to see N1, R2, N1, R2, AS2, AS1 is the path and so and so forth.

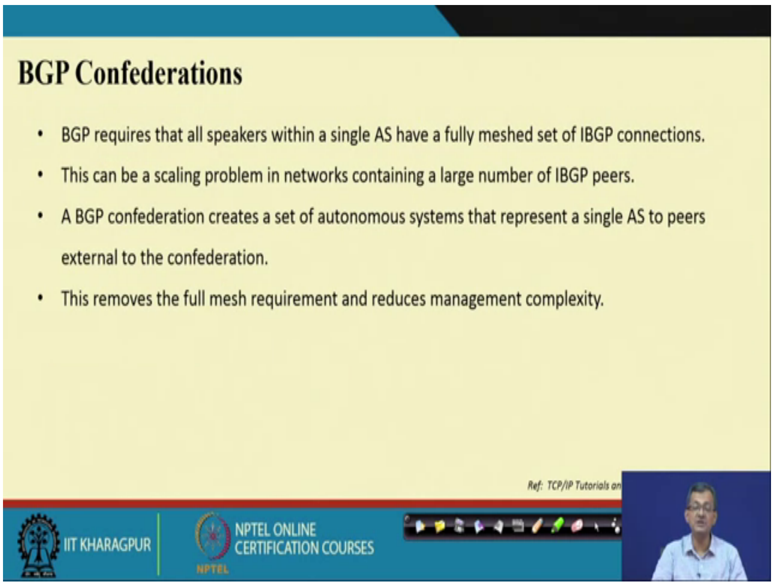
So, where from these R4 we see this is the AS3 AS2 and AS1 and then a next router is router 3 and the destination network is N1. So, next router 3 then AS3 AS2 AS1 and then you have the router. So, these are way the paths can be defined and for different network what is the next router and paths can be defined.

(Refer Slide Time: 33:06)



So, 2 quick concept, one is the BGP aggregation, so a major improvement in BGP version four for CIDR and route aggregation. So that means, the feature allows BGP peers to consolidate multiple contiguous routing entries into a single advertisement and it significantly enhances the scalability of the BGP to the large network. So, I can contiguous routing advertisement in a single entry and in doing so, it basically helps in scaling the things.

(Refer Slide Time: 33:43)



There is another concept called BGP configuration. So, BGP requires all speakers within a single as have fully meshed set of IBGP connections right.

So, if it is basically, this creates a problem in scalability. If there are a large number of speakers within the AS making this connection and if it is dynamic then making and breaking these connections become a major challenge. So, I can have different sub a sort of things within the AS right or what we say that is a configurations of AS is there right. So, a BGP configuration creates a set of autonomous systems that represent a single as right. So, AS1 can have different AS 1 1 2 type of things and two peers external to the configuration. This removes the full mesh requirement and reduces the manageability. Because, the full mesh requirement is now confined within that particular AS where, or sub AS type of things or configuration is maintained.

So, in way in this way the manageability may be made much or scalability is facilitated So, what we see in try to what we discussed today is that, different feature of BGP routing protocols, how it can be defined how path is defined and so and so forth. How, it helps in sending a packet from a particular network, in a particular AS autonomous system to a other AS in the autonomous systems right.

So, we will continue our discussion on this computer network and internet protocols in the subsequent lecture. Slowly we will look at other layers other as we are following a, drop down approach now we will look at the other layers on the TCP IP model.

With this let us conclude our discussion today.

Thank you.