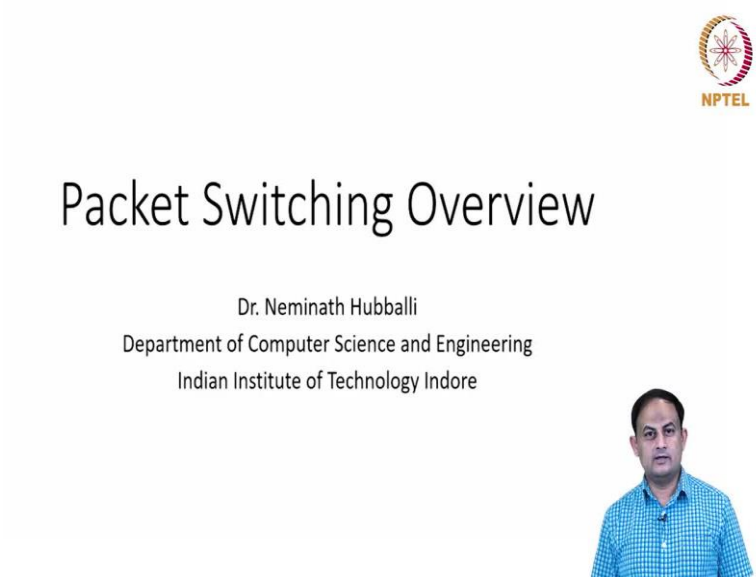
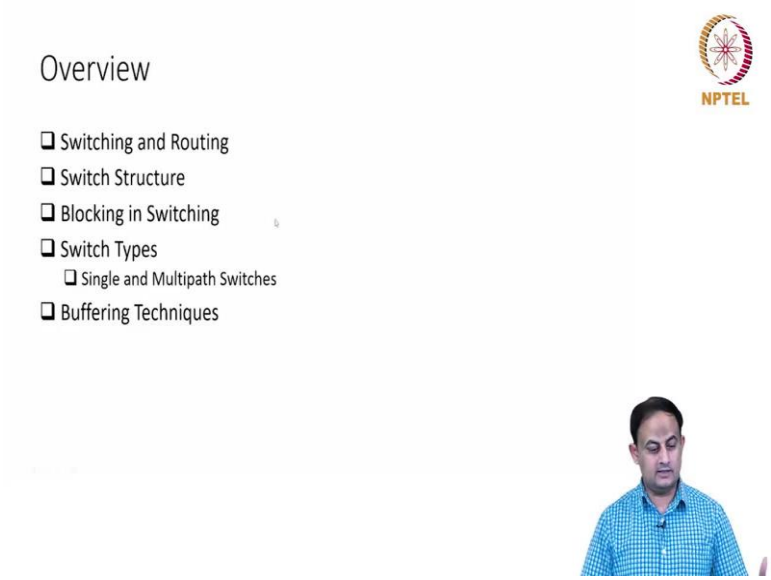
Advanced Computer Networks
Professor Doctor Neminath Hubbali
Department of Computer Science and Engineering
Indian Institute of Technology, Indore
Lecture 21
Packet Switching Fabric Design – Part 1

(Refer Slide Time: 0:20)



Welcome to the last lecture of the first module of this course on Advanced Computer Networking. Today I want to discuss about something a topic called as packet switching. In the last few lectures, what we saw is what is the routers in the networks are supposed to do, basically, the lookup operation and then the, we understood that packet classification and the quality of service and other things are the essential components of the routers in today's networks. So, what we will do in today's lecture is to basically understand the if I had to go about designing a switch, what exactly is there inside that switch.

(Refer Slide Time: 01:04)



So, here is the agenda, we will spend some minutes in understanding the difference between the switching and routing the structure of the switch. And then there is something called as the blocking that happens inside the switch when you put it in when you come up with the design. So, what is that blocking different types of the blocking and then finally, we will venture into we will take a look at the different types of the switch designs available in the market or the at least the strategies that are followed and then we will end the discussion on the something called buffering techniques.

(Refer Slide Time: 01:41)

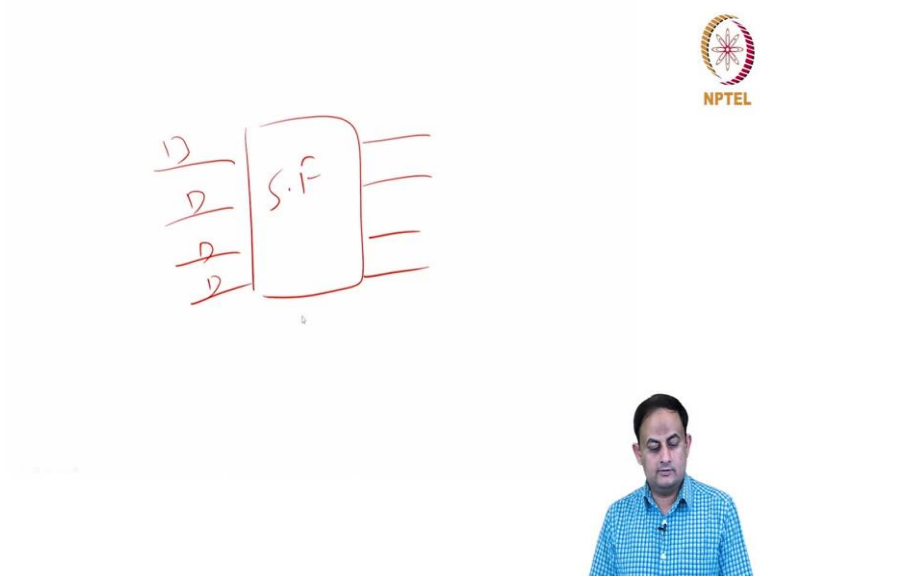
So, the switching and the IP switching whatever it call has a history in telecom networks. So, basically, the switching operation that was happening in the IP networks is similar to that of

what used to happen in the earlier days the telecom networks. So, telecom in the sense the telephone network, there were companies which provided the telephony services had something called as the exchanges, these are telephone exchanges and here is a picture which actually shows such a telephone exchange.

And back then, in earlier days, the switching operation used to happen manually somebody who is sitting at the these centers used to connect the people who for example, if A wants to talk to B there is a user task panel cords they used to connect the wires running from A to B in this board.

So, subsequently such manual connection disappeared and then we came up with the mechanism to do this operation automatically, where the exchange automatically finds out where exactly or who is talking to whom and how to establish the connection between those two points. So, using data as a motivation, the first routers or switches which were designed actually inherited the property or at least the structure or the mechanisms that exist in the telephone switches.

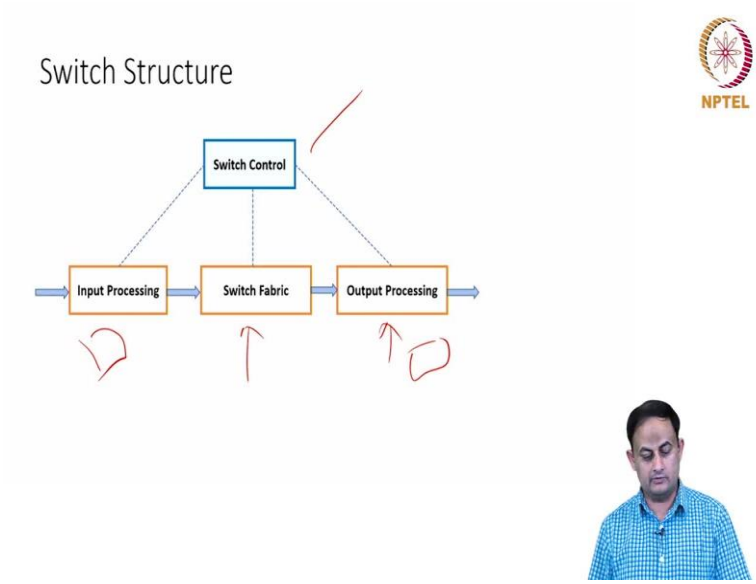
(Refer Slide Time: 03:18)



So, what we said is a switch when I see the switch the switch that operates in the IP network has got these components one is the switch fabric, let me call it as S.F. then mutually we have got a series of input ports and then a series of the output ports, the packet arrives on these input ports and then you need to forward those packets to the appropriate output port numbers, so that is what the switch in very brief it has got. So, switch fabric and some of the

port numbers, so the line cords and other things that we will talk or connecting to is port (())(4:00).

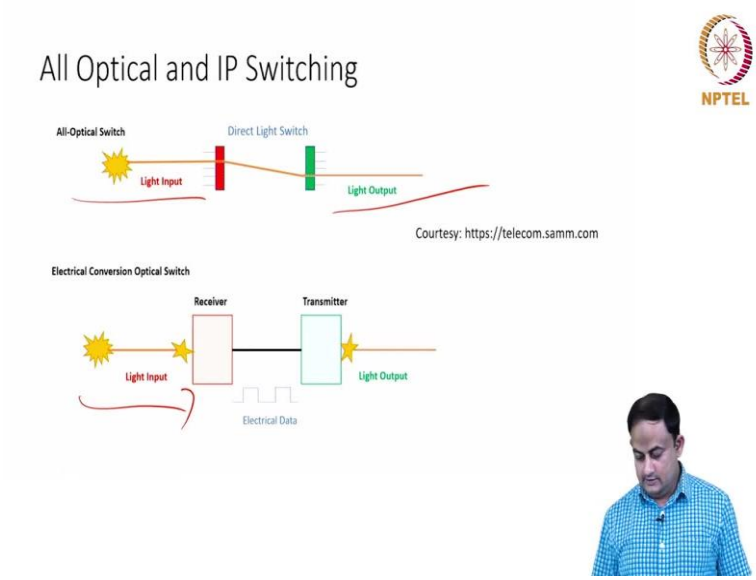
(Refer Slide Time: 04:15)



So, now, the question is, how does the switch operate, what kind of the using what technology the switch actually forwards the packets, how the packet is arriving on the input port number? So, this is what is a picture which is showing the three things, one is the input when the packet arrives at the input mode, there is a processing happening, what is that processing? Basically, you need to strip out the link layer header and then decrement the detail value and then consider the FIB and send it to the appropriate output port number.

And the switch fabric is the one which is actually establishing a connection between the input port and the output port. And once the packet goes to the output port, there is some more crossing happening so that process is basically you are adding a link layer header corresponding to the links that exist between this switch and the next off switch. And switch control is the place where the logic or the how the FIB information is constructed a bunch of routing algorithms run inside that control plane. So, this is what we understood.

(Refer Slide Time: 05:15)



So, now, the switches when I say the switch is actually the router, this is a generic term that is used, anything that is doing the routing and switching all of them, including the layer two switches, the fundamentally using the same mechanism, so that is called as the switch, so you can very well call it as the router as well. So, these routers, these routers are of two types, one is which work completely using the optical fiber technology where the input the packet data is coming in the form of the light, it is a wave and the output that exits from that router is also is in the form of a light and it operates completely inside by processing that light information itself, it actually does the processing.

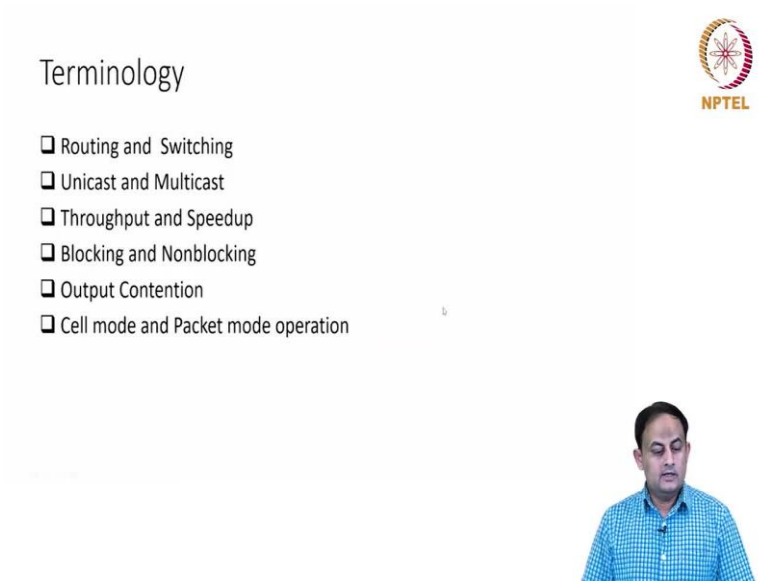
So, be the consulting the FIB information or other things, whatever it is. So, the two entry and exit are in the form of light and then the processing is actually done in the form of the digital data. Basically, you need to interpret that what packet, what data it is carrying, and the header information and what is the IP address and all that you interpret and then you look at the FIB information based on that you would make the forwarding decisions.

And there is a second category of the switch, which are the router which actually received the input or the packet data is arriving in the form of the light, it is an optical fiber on the left-hand side. And it converts that data into which is received in the form of light into an electric signal and then transfer it to the (())(6:54). So, the entry is in the form of the light and the exit is in the form of the electric signal.

So, these are the two kinds of switches that you find, most of the backbone networks in today's internet or the networks are basically completely probably the design you are working

using the optical technology. So, maybe the last mile connectivity and the some of the distribution switches might have this second category of the switches. So, these are the two categories of the routers or the switches that you find in the today.

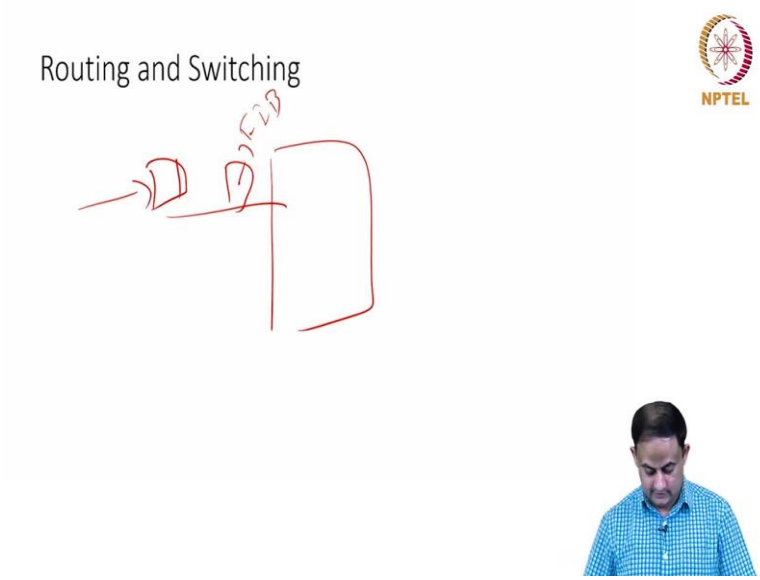
(Refer Slide Time: 07:39)



So, after having understood this, how the data is actually getting transmitted. So, let us try to motivate ourselves in order to look at what exactly is there in the switch inside, or the route inside or how they are actually engineered, how is the switch fabric look like? So, let us define some of the terminologies and after having understood those terminologies, we will take a look at the different types of the switch designs or the fabric designs.

So, the these are the following terminologies. So, routing and switching, what is the difference? So, there are switches which actually work in the form of unicast forwarding and multicast forwarding, what you mean by the throughput and the speedup, blocking and non-blocking, the output contention, and then the finally cell more than the packet mode operation? So, although when most of the discussion in this course is related to the packet mode operation, but nevertheless we did touch upon the ATM network, so we will take a look at the cell mode operation as well. So, we will go one at a time the terminology will we will spend a minute in understanding each of these terminologies in little more detail.

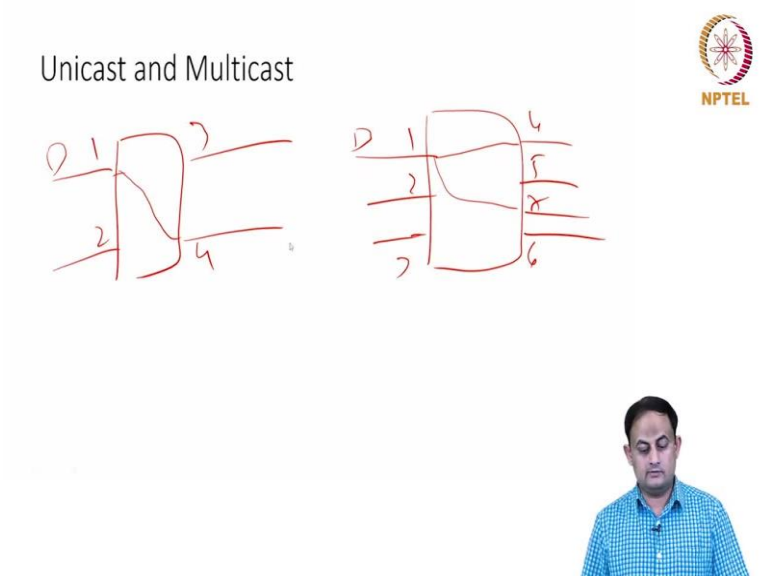
(Refer Slide Time: 08:47)



So, the first terminology is the Routing and Switching. Fundamentally they mean the same thing. So, Routing and Switching what exactly is there inside a switch be the layer two switch or the layer three switch or be it the optical switch or the other type of the switch. So, you have a mechanism, FIB information is sitting on the line cord and there is a switch fabric and you concert the FIB information and the arriving packet, the header information of that packet is taken and then you concert the information and forward it.

So, that is the so although you might have the layer two switch might have a different type of the information but layer three switch might have different kinds of information but fundamentally the operation is to do the lookup and you have a database that is the file FIB information in the network layer or IP layer or at the layer two you might have the link layer addresses and so forth. You pick up address from the packet and then go and look into the database and then do the forwarding operation, so that is the routing and the switching. So, the more or less mean the same thing.

(Refer Slide Time: 10:16)



So, the routers switch are out there in today's networks, operate in two modes one is called as the unicast mode and second one is called a multicast mode. So, all the majority of the traffic that is going inside the network is unicast in nature that is point to point. Meaning, a packet which is arriving on an input port wants to exit the router and let me here is a picture. So, let me name these ports as 1, 2, 3 and 4 and if a packet arrives on the port number one and after consulting the FIB information, if it has to be forwarded to the port number 4, all that you need to do is to establish the link between port number 1 and the 4. So, that is the job of the fast switch fabric that you have.

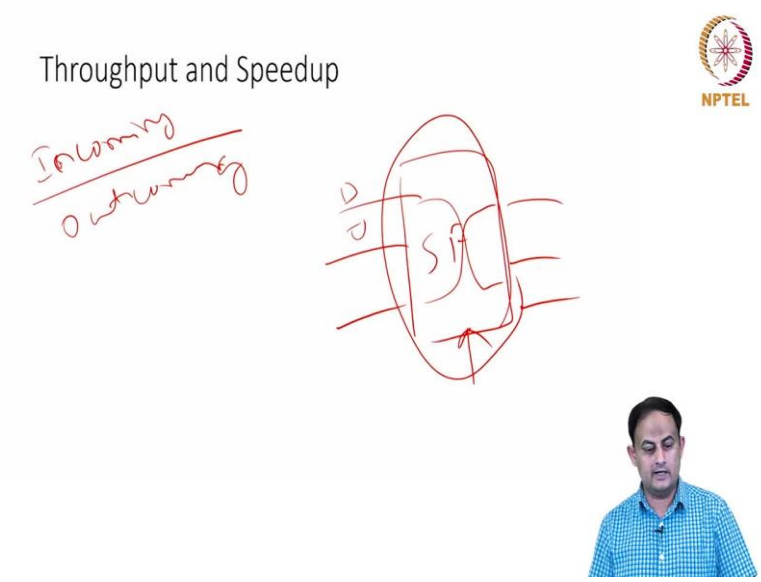
Now, this kind of the operation is called a point to point, meaning you have one input port on which the packet is arriving and the exit to port number is also 1, so there is 1 to 1 mapping. So, normally whenever we browse the content from the internet, it is actually a point to point information. Occasionally, you might also have a different kind of the traffic where you might want to transmit the arriving packet on multiple port numbers. So, for example, if I have 3 input ports and then 4 output ports, sorry yes, 4 output ports something like this and a packet which is arriving a port on port number 1.

So, you may want to pass it on to the port number 4 or port number 5 and together are a subset of the port output port numbers that you have on the switch or the router. So, why do we need to do this? There are some applications which actually demand such a kind of the operation. So, the basically one packet is replicated and sent on two different output port numbers. So, think of this audio video conferencing application where a user multiple number of users are engaged in conversation and you want to have all of them receive the

contents sent to all of them then you may need to replicate the packets and send it on the multiple port numbers.

So, how exactly it impacts the switch? So, basically the router or a switch you need to be aware that such a so and so application is running, who is connected where, and on which port number I need to forward it, all that decisions you need to make it here. So, those are the operations which are required the are of the type multicast type of the switches.

(Refer Slide Time: 13:04)



And the next terminal is something called as a Throughput and the Speedup. So, throughput refers to the amount of the data or the aggregate amount of the data that is incoming or the number of the packets or the bytes whatever are receiving, and how much of that is actually transmitted. So, in the best case, the whatever the data that is coming on your input port number, so when I say the input port, all the input ports, all of them are able to you are able to transmit it, that is the best performance of the switch. So, by virtue of that throughput is at the max is going to be 1. So, you might very well the switch might say that I am not able to handle these many number of the packets or some of them might get dropped right there at the input port, because of the non-availability of the space.

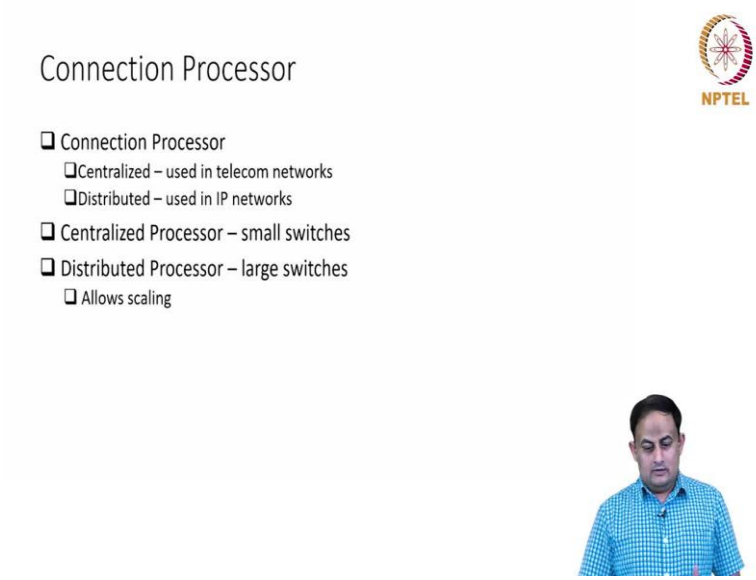
So, the best case is everything that is coming to the input port is actually being transmitted to the output port, but otherwise, a fraction of what is coming to your input what is actually forwarded to the output port. So, that is called the throughput. So, in a nutshell, so this number throughput is maximum has a ceiling of value 1. So, occasionally, what might also happen that the switch fabric that is there inside your switch, might be capable enough to

transmit more number of the bytes or the bits that are arriving together inside to your input ports.

So, let us say cumulatively I have maybe a 10 Mbps of the data is arriving on the input port and the switch fabric itself, when I say the 10 Mbps that is per unit of the time. So, the switch fabric itself might be able to forward more than 10 Mbps of the data, although you cannot really forward what we more than what exactly is arriving on your input port, but it has got the capability to connect the input ports and quickly take that data and put it onto the output port. So, if that happens, then we see that there is a speed of operation.

So, remember, what I am referring to is only to the ability of the switch fabric to take the content or the packets coming to the port and putting it into the output port number, so only the operation of this switch fabric nothing else. So, that if it is got an ability more than what the cumulatively able to receive on your input port numbers so then we call it as the speedup. So, technically, you can have a speedup of more than 1, 1.5 to whatever is the switch fabric is allowing you by that factor it can actually speedup the operation.

(Refer Slide Time: 16:13)



So, we said that the switch actually has two components, one is the port connecting to the switch and the second is the switch fabric. And the way switch fabric makes these decisions or establishes the connection is through something called as the connection processor. So, there is a processor which that is the intelligence that is built into the switch fabric which knows how to establish the connection between one input port to the output port.

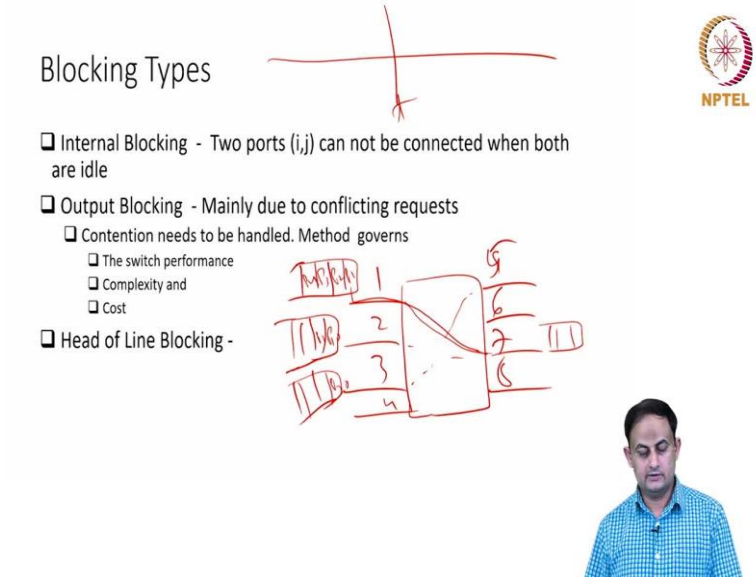
So, the connection processor itself are of the two types, one is called the centralized processor and this was the technology that was used in the telecom networks earlier days where one entity is able to make the decision or how to establish the connection for every two port numbers, incoming and outgoing port numbers in the switch and the second kind category is something called as a distributed processor.

So, when I say the processor, I mean to the inclusion built into the switch fabric on how to establish the connection between the port, two port numbers, nothing more than that. So, the other operations which require some kind of processing are done at the input port and the output port. So, the distributed switch processors are nowadays used in the IP networks whatever the switch that we use in the TCP IP networks are more or less built at a distributed processing framework.

And it is obvious to argue that centralized processing is actually good it does not scale well so only if at all if you have to incorporate that kind of process so you need to do that inside switches which are of the small-scale switches. So, meaning the number of input ports and output ports are less than number. And also, you do not require the throughput and the scaling by many factors.

So, on the other hand, if you use the distributed processor, the scaling you can add more number of the ports to the switch fabric and then the (())(18:28) parallel connections can be established on the using those switch fabric. So, because the intelligence itself is not placed in the one product does not act as a bottleneck for the processing.

(Refer Slide Time: 18:45)



So, because it said when you make an establishment to connection establishment, we are using the switch fabric from one input port to the one output mode there are different ways with which the connection establishment itself can be stalled. So, here is a picture let us assume this the switch fabric is here and then there are some port numbers let me name them as 1, 2, 3 and 4 and this 5, 6, 7 and 8 are the output port numbers.

And one of the simplest way with which the packet the connection establishment can restore this let us assume the 1 and 7 are connected to connect with each other right now. And as we understood in our earlier discussion, there is a queue that is out there at the each input port and the packets are placed inside this queue.

So, let me call these packets are P 1, P 2, P 3, and P 4. And let us assume at the time being the packet P 1 is being transmitted on this switch fabric the bits corresponding to the packet number one are flowing through this link and they are going in sitting inside probably a buffer or the queue that exist at the output port.

So, now, so remember every port input port also has got a queue. So, there are some packets waiting probably waiting at the port number two as well. So, maybe those packets let us call them are P 10, P 11 something like this. And what can happen is because P 1 is already getting transmitted on the switch fabric and the port number 1 and 7 are already connected another packet waiting at the another input port number will not find a place or a pre slot because the switch fabric can connect the port number 7 to only one of the input port, so parallelly it cannot do the transmission.

So, such kind of the blocking is called as the output blocking where there is a contention more than one packet arriving at the different port numbers wants to take an exit route to the same output port number and if there is an ongoing conversation so that to blocking because now the P 10 has to wait it has to go to the port number 7, that is what you have determined after consulting your FIB permission, so that has to wait. So, that kind of blocking is called as the output blocking.

And there is another kind of the blocking that might happen, so that blocking is called the internal blocking. So, the way internal blocking occurs although the two port numbers let us say in this figure now, the port number 1 and 7 are engaged in the exchange of the data packet 1 is being transmitted, but the think of this case. So, maybe the port number 3 on the queue corresponding to port number 3 has a packet.

So, let me call that packet as P 10. And it wants to go to port number 5 an output port number 5, because of the connection that is already established between the input port number 1 and 7 this connection which has nothing to do with neither the input port or nor the output port is now blocked for transmission. So, if that happens, then that blocking is called as the internal blocking.

So, basically two ports numbers i and j cannot be connected to each other when both are idle. So, they are not involved or engaged in any kind of data exchange or packet exchange. So, if that happens, then we call it as the internal blocking. So, internal blocking is a phenomenon of the switch fabric design itself, how it has been designed. So, it is not able to transmit these packets or establish the connection between these two port numbers.

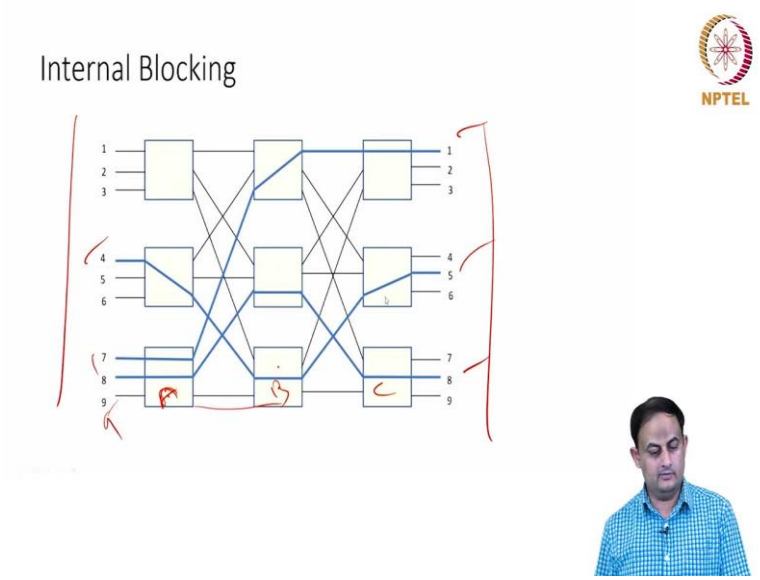
So, on the other hand output blocking is basically because of the contention originating from the packets or sitting at the head of the every input port number sorry input port. So, this contents an output blocking, it needs to be handled in some way. So, there are some mechanism it requires an arbitration who is going to replace, let us say exactly at the same point of time. So, along the time axis exactly, let us say at this time t the packets on the input port number 1, 2 and 3, all of them are scheduled for transmission to the same output port.

Now, you need to determine, so the packets are scheduled on the input port and 1, 2 and 3, all of them simultaneously asked the switch fabric I want to go to port number 7, I want to go to port number 7. Now, you need a mechanism to determine who is going to get this slot who is going to get the connection and not able to transmit the packet to you what port. So, there are different ways of handling that and of the blocking, and the mechanism that we use to do the handling or arbitration can impact the switch performance.

So, basically the throughput that you get out of that switch and the complexity of the switch design itself and also it might also has implication on the cost. So, these are the 2 main types of the blockings that we can find and occasionally there is we also talk about something called head of line blocking. So, this happens because of the IP packets or the variables size and the packet, which is of the larger size might be getting transmitted on the switch fabric right now from one input port to the one output port, because of that a packet which is having are scheduled sometime later in the transmission might have to wait for a longer period of time.

So, that is called as head of line blocking, where somebody who is bigger in size is actually occupying the larger duration for retransmission and thereby blocking someone else who is behind the queue. So, that although this has nothing to do with the switch design or the fabric design itself, but nevertheless that is also one kind of blocking that can take place inside the switch or the router.

(Refer Slide Time: 25:29)



So, here is a picture which actually shows how the internal blocking might happen. So, without worrying about the what type of the switch it is. So, what you can think of is on the left-hand side, there are some port numbers numbered from 1 to 9, and the output port numbers are also numbered from 1 to 9. So, technically the input port and output port number can be same. So, although we did differentiate them in our previous discussion, but they can be saved as well.

So, what is shown along these lines, the thick blue color lines in the switches the current ongoing conversations which port number is actually connected to which of the output port number. So, we can see that 7 is connected to the port number 1 through the intermediate block here, and it is connected to the port output port number 8 and the 4 is actually talking to the port number 5.

So, these are the 3 ongoing conversations and now, what you need to say it is once I establish a link between our transmission between one block to another block. So, let me call this block as A and this block as B and this block as C. So, once I establish a connection between A and B, so for the block A there are 3 ports, which are connected, and there are 3 incoming lines

for the block B, when A and B are connected. So, what it means is only one of the input ports data or the packet being transmitted through that link to the block B.

So, maybe I might transmit the data from the port number 9 to using this link to the block B, so that is what it means. So, if that happens, what it means is I can transmit, I may not be able to transmit to more than one port input port numbers from one block to the another block. So, you are looking at the parallel conversations that are happening right now. So, 8 is connected to the output port number 8, 4 is connected to 5, and 7 is connected to 1.

So, if let us say the port number 9 wants to talk to someone maybe port number 4 or 5, so that cannot be done because so the output the block A is engaged in the conversation. So, the connection can go only through either of this block B or through this blog. So, those 2 are actually blocked now. So, 9 will probably will not be able to transmit the data from whatever the packet want to transmit to the port number 4, we may not be able to establish the connection. So, this is an example of the internal blocking.

So, it is apparent that there are many-many other possibilities which can which can have the such internal blocking. So, that is the difference internal blocking is because of the constraint of the switch design itself something is actually blocked. And output blocking is because of the contention that originate from the poor packets which are actually scheduled for transmission on the same output to port number.