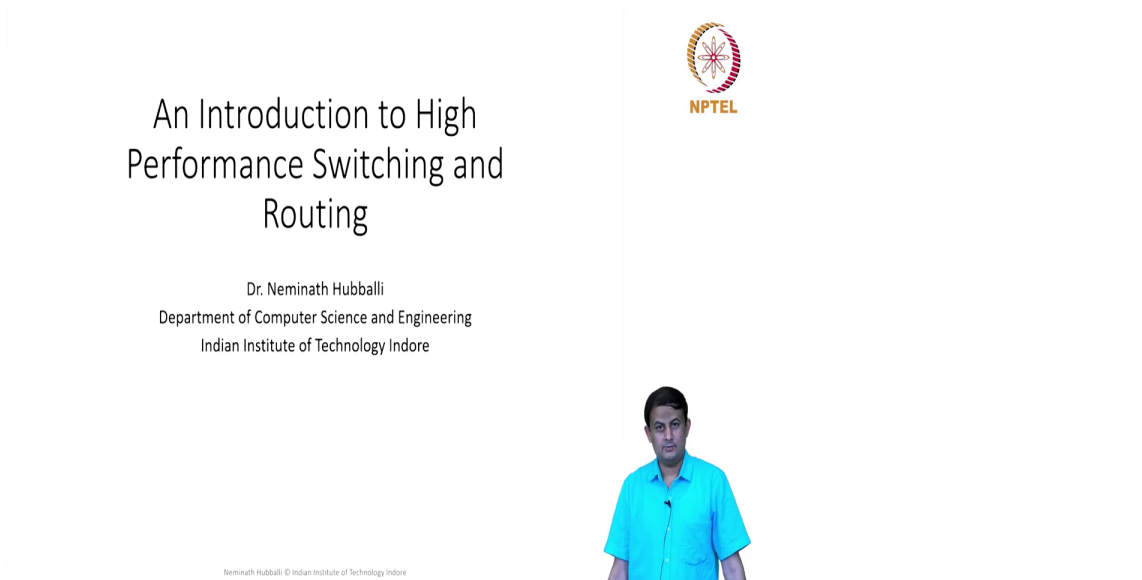


**Advanced Computer Networks**  
**Professor. Dr Neminath Hubballi**  
**Department of Computer Science Engineering**  
**Indian Institute of Technology, Indore**

**Lecture - 1**

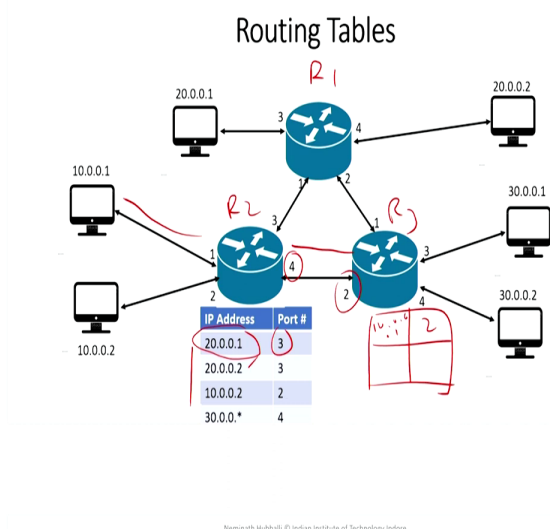
**An Introduction to High-Performance Switching and Routing - Part 1**

(Refer Slide Time: 00:18)



Welcome to this course on Advanced Computer networks. So, today we will be starting with the first module of this course which is high-performance switching and routing.

(Refer Slide Time: 00:33)



So, here is a sample network where there are 3 routers, let us name them R1, R2, and R3. There are 3 routers and then 6 computers, there are some IP addresses assigned to these computers and they are shown here 10.0.0.1 and something along the lines. Typically network router has got something called the routing table.

So, you might have studied in the undergraduate course where the routing table looks something like this: it has got 2 components one is called destination IP address and the other is called Port number and this is the routing table for the router R2. What it means is: if you receive a packet with the destination IP address of 20.0.0.1, this routing table tells this router that you need to forward it to Port Number 3.

And similarly, if you receive a packet with destination IP address 20.0.0.2, you need to send it to Port Number 3 and so forth. So, such routing tables are there with every router. It is there with the R1, it is there with the R3, and so forth. So, depending on where those destination hosts are connected, the routing table is changing.

So, probably if I take this route from R3 to R2 and then to 10.0.0.1, this routing table R3 would have an entry something like this. If you receive a packet to 10.0.0.1 you need to send it to Port Number 2. Port Number 2 is the portion of this router R3. The router R2 might have a different port number on the same link. So, that is okay.

So, with respect to that router, based on what the routing table is saying, decisions are made. So, this Job of routing is the fundamental job of any router. A router is supposed to do routing; route

the packets to corresponding destinations and that is the primary job, and in addition to that, there are some other additional things that typically routers do. So, you will understand what are those decisions.


This seems to be simple; you have a routing table and you receive a packet and look into the header field of that packet, look at the destination IP address, and then send it to one of the output ports or according to what the routing table actually says.

But it is not as simple with the kind of question we ask. As the link speeds increase, every day now you see that earlier we used to talk relay bandwidth in terms of the Kbps, nowadays we talk in terms of Mbps, Gbps, and several Gbps link speeds. So, if that is the case, then what is it that takes these routers to handle such kind of massive traffic in the live network? So, these are the kind of questions. So basically, these network routers also need to handle such traffic at a very large speed. And we require some algorithms and hardware mechanisms that support handling such massive interrupted traffic in real time. So, that is what is all about high-performance switching and routing.


(Refer Slide Time: 04:42)

### Network Bandwidth Growth Trajectory

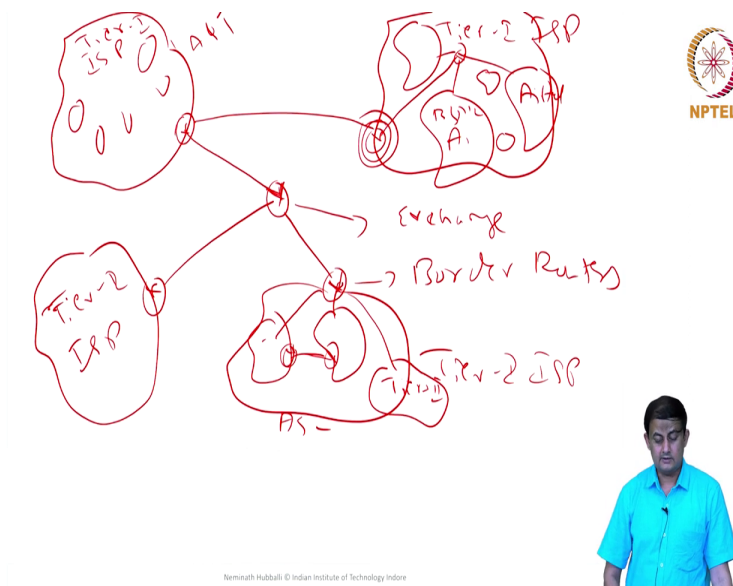
- ❑ ARPANET 1969 -54 kbps
- ❑ NSFNET 1987 1.554 Mbps
- ❑ CISCO - 1990 – 155 Mbps
- ❑ 2000 – 2.5 Gbps
- ❑ 2012 – 40 Gbps
- ❑ 2022 – 250 Gbps Trunk



NPTEL



Neminath Hubballi © Indian Institute of Technology Indore



So, just to put things in context, here is the network bandwidth trajectory. So, earlier when the computer networks came into existence, with ARPANET, people were talking about Kbps, and the first network was operating around 54 Kbps. Back then it was very good bandwidth and then in 1987, people started talking about 1.5 Mbps and the Cisco 1990 said that we have the technology to transmit to 155 Mbps in the backbone.

And in 2000 people talking about 2.5 Gbps; I am talking about the backbone and not the end-user bandwidth. And as we speak today in 2022, people are talking about 250 Gbps of trunk speeds. Trunk speeds are between the backbone router, one router to another router, and the backbone of the Internet service towards networks. So, that is the typical speed we are talking about.

So, if you want to handle such a massive amount of traffic, 250 Gbps, then your network routers and switching elements also need to have the capacity to process these packets. Basically, if I want to do the look-up at this speed, there are millions of packets that are coming in per second, Can I handle those millions of packets in real-time? So, if at all your answer is yes, then what is that it takes in the routers and switches to do or perform such kind of massive scale operations in real-time that is the question.

So, in order to understand how this bandwidth increases in the trajectory, and also the performance improvements that are required in the network routers and the switches, let us try to begin by looking at how exactly the Internet service providers' networks are interconnected or in

other words, how the Internet looks like in today's world, how they are actually interlinked. If I signed up even sitting here and accessed a site which is located somewhere in the US, how the routing actually happens at each router; at each network element? what is actually happening?

And so if you look at it in context, you will understand what is the scale of the operations that we are looking at. So, all of us probably will buy an internet connection from the Internet service providers. If I am an institution subscriber, let us say like, IIT Indore, we might go to an Internet service provider and buy some kind of bandwidth and they provide a lease link connection or if I home user, then I might go to the same service provider and then ask the kind of the connection that I get at home which is slightly different from the kind of the connection that I have at the institution. Nevertheless from the service provider's perspective, it is one network they have, the backbone and they interconnect with other Internet service providers.

So, the question is how these Internet service providers themselves are interconnected to each other. Internet in today's world looks something like this: you have got a hierarchy of Internet service providers, and at a larger scale are something called the Tier-I ISPs. And there are a bunch of them, so, Tier-I ISP is here, this is the similar second Tier-I ISP, and the third one here, this is also a Tier-I ISP.

And there are routers, which are located at the boundary of these networks. Either these Tier-I ISP themselves are interconnected through a link or they might go to a third party and then terminate their network links to these third-party locations. So, for example, in between this router or this guy, or there is a third ISP which is again Tier-I ISP, so he might bring one link and then terminate into this point.

So, this location where multiple Internet service providers are terminating their connection is called the exchange, this is also called Internet exchange where multiple links are actually terminating. And these routers which are sitting at the edge of their networks are called border routers. So, you can think of this as one autonomous system, and at the autonomous system boundary, you have got the router, that router is called the boundary router.

And typically these Tier-I ISPs do not provide Internet services to the end users, be it the home user or be it the institution subscriber. For both of them, they typically do not provide the service. But what they actually do is: there might be smaller ISPs located within the territory and to all of

them, they might provide the services, so there are Tier-II ISPs and multiple of them within their geographical area.

So, the Tier-I ISPs have got connections, they have got something called point of presence, these Tier-II ISPs might actually connect to their point of presence locations and then buy bandwidth. So, this could be probably a BSNL network in India, an Airtel network in India, and something like this, these are probably the Tier-II ISPs. They do provide the connectivity to the end users (the subscribers), but they themselves buy the connections from bandwidth from the other Tier-I ISPs.

And you might have even a smaller player located within this area here and there which are called Tier-III ISPs. In earlier days or even now also sometimes cable network operators also used to provide internet services. These were typically the links, we have their presence in one small city, some small districts, and all those things. They cover very limited areas and such Internet service providers are called the Tier-III ISPs.

Tier-I ISPs are anyway connected either directly or through some Internet exchange point. But sometimes, the Tier-II ISPs themselves might decide to connect their own network and build their exchange points. So, that prevents them from going, if I want to send traffic from a host, which is located in this ISP in the absence of this peering link, I need to go to this guy and then come to enter into the other network.

On the other end if I establish a direct link between these two, then the traffic I do not have to rely on the higher-level Internet service provider to route my packets. And this is how the internet is actually structured. So, this is the structure of the internet. And look at the scale, this guy might be an AT&T network in the US and I have some other maybe Bhartiya Airtel or BSNL in India.

And I am talking about the entire network, the entire traffic of BSNL or Airtel going to this border router and then going to the AT&T network, so that is the scale. So, the entire nation's traffic or probably a region's traffic this router needs to handle in real-time, it needs to process that. So, that is the scale. So, it requires a special link meaning this router needs to be capable of handling those many packets.

I said that routers need to be doing this routing operation which is consulting the routing table and then making the decision of where to forward that packet. And in addition to that these routers also need to do some other jobs maybe the quality of service, maybe the other operations like if I want to block certain traffic from going, filtering operations, and all that these all into the picture. So, all these decisions need to happen in real-time.


So, if that is the structure, how do you engineer or build such a scalable router which can perform these operations in the real-time, that is the question we want to ask in this part of the course.

(Refer Slide Time: 14:44)

### Tier-1 ISPs

- ❑ UUNet
- ❑ Sprint
- ❑ Qwest
- ❑ XO Network
- ❑ AT&T
- ❑ CenturyLink
- ❑ Cogent Communications
- ❑ Deutsche Telekom
- ❑ Global Telecom and Technology (GTT)
- ❑ NTT Communications
- ❑ Tata Communications
- ❑ Telecom Italia Sparkle
- ❑ Telia Carrier and
- ❑ Verizon





Neminath Hubballi © Indian Institute of Technology Indore

So, before we move forward, I said that Tier-I ISPs usually do not provide a connection to the end users. Here are a handful number of Tier-I ISPs which are there as we speak on today's date. So, there are only a few of them. So, they are having a large geographical presence probably covering the entire nation. And internally they are distributing the bandwidth to the other service providers, Tier-II and Tier-III networks.

(Refer Slide Time: 15:17)

### Tier-1 ISP Network

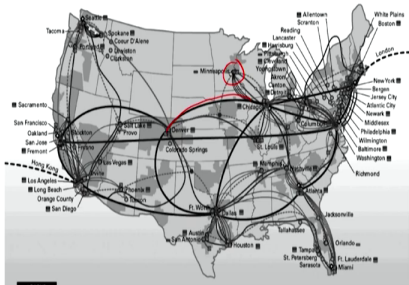




Image Courtesy:  
H. JONATHAN CHAO  
and BIN LIU





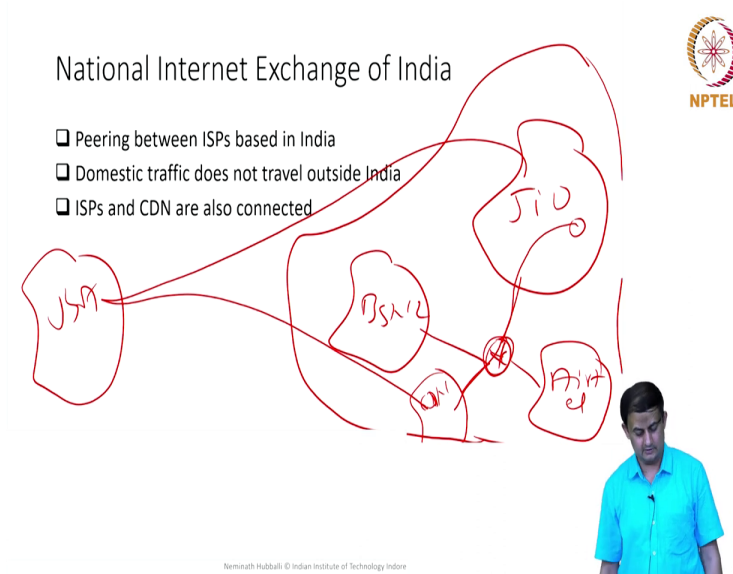
Neminath Hubballi © Indian Institute of Technology Indore

So, here is a sample of connected Internet service providers' networks and what their backbone links look like. So, whatever you see as a dotted line here, is called a point of presence or an



exchange point. And you have whatever you call this one is the link and they have scattered links all over, this is the USA Network and through this point of presence, actually they are providing the connectivity to the other Tier-II service Internet service providers.

(Refer Slide Time: 15:56)



I said that through the exchange point, these Internet service providers are connected. In India, we have an exchange point that is called the National Internet Exchange of India. So, in short one this is also called the Nixi. So, in the year 2003, the Indian government established this, which is actually a nonprofit organization and it is mandated to provide the exchange services between the service providers which are located in India. Meaning, if I am an Airtel subscriber, if I want to access or send traffic to a BSNL network, I do not have to go to the Tier-I ISPs.

So, this is my BSNL network and this is my Airtel or Jio network. And since BSNL is also connected to this exchange point, and Jio is also connected to this same point, and so is the Bhartiya Airtel and here they also terminate. So, all of them together might be buying the bandwidth from Tier-I ISPs. But these BSNL network does not have to send the traffic to their Tier-I ISPs to exchange packets between these networks. So, this is where the Nixi come into the picture.

So, the idea is the domestic traffic which is moving from one subscriber within the Indian geography to exchange traffic between another computer in the same geography, I do not have to send my traffic to the higher level. Maybe Tier-I ISPs located in some European country or let us

say in the US. So, if BSNL and Jio do not know how to talk or exchange traffic between them and if a subscriber of the BSNL network has to send traffic or packet to another user in the Jio network, all that he needs to do now in the absence of this exchange is going to the Tier-I ISPs who are located in the US and then from there, you come back and deliver the packet to the Jio network, which is actually a lot of work to do. Mainly distance wise also it is far away and also it is costly. There is no need to send domestic traffic to the external service provider.

These Internet service providers are not the only people who are connected to these exchange points. Sometimes there is something called the CDN, CDN stands for Content Delivery Network, maybe Google or YouTube. And they have content. YouTube has a content, lot of content. And these guys might develop or put up their own servers and build their own infrastructure, which is looking as similar to that of an Internet service provider.

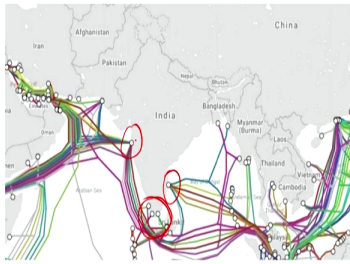
There are multiple servers, they have the optical fiber running between those servers, and they have routers, and all that. They can also come and join these internet exchange points. In fact, in the Nixi network, as we speak today, there is a CDN network which are also connected to this Nixi point. So, they can also pair the connection to this exchange point and then deliver.

So, let us say a subscriber sitting in the Jio network wants to exchange access content located from a server in this CDN network. So again, the same concept works, I do not have to go to the service provider who is in the USA and traffic going all the way and then coming back to this place and relay the traffic between these two systems. Geographically these are closer and if you have an exchange point between them, then you can actually deliver the content. So, that is the role of these exchange points.

So, in summary, what I wanted to say is, these exchange points help to interconnect different Internet service providers. On a global level, the Tier-I ISPs are connected either directly to an undersea cable or some other kind of internet exchange point. They are able to talk to each other, but on the local level within a country, you might still need the exchange points, they then can connect to other local Tier-II and Tier-III ISPs also, or sometimes the CDN networks are also there.

(Refer Slide Time: 20:53)

## How India is Connected ?



Courtesy: Indiatimes.com



Neminath Hubballi © Indian Institute of Technology Indore

So, that brings us to the question: if you take the Indian continent as we speak today, what does the internet connectivity to the Indian subcontinent look like, and how we are actually connected to the rest of the world as we speak today? So, you can see that on the inside, there is a bunch of undersea cables terminating at this location, this is the Mumbai ending and on the eastern side you might have this Chennai, there are a lot of cables terminating, going to the east mainly to the Singapore and Malaysia and other places.

And from this location, we are actually going to the Europe and US continent and some of these locations are in Kerala, Thiruvananthapuram, and Cochin, and, these places highlighted in the red mark are called the landing sites where the undersea cable actually terminate and these are the high capacity, high bandwidth optical fiber cables which connect the entire nation's internet network to the rest of the world.

So, you see these landing sites are because when talking about undersea cable, they actually the presence of these cables actually terminates at the places where you see the sea coast.

(Refer Slide Time: 22:27)

## Internet: How it is Connected Today?



Courtesy: Google Images



Neminath Hubballi © Indian Institute of Technology Indore

So, here is a picture which is actually one of them of how these undersea cables are laid down on the left-hand side diagram you see a set of men who are actually laying such cable, and on the right-hand side we see a ship which is actually carrying that undersea cable, they are doing the laying operation, and here the bottom tool is actually how the cable looks like and the typical machine how actually that undersea cable is actually laid.

So, these undersea cables are now the backbone or the prominent sources of the entire data exchange. So, if there is a fiber cut or if this undersea cable is cut, then probably a part of the entire internet, one of the Internet Service Providers Network might be affected. So, it requires very high availability and maintenance so that this cable is never down, even in the undersea cable also after every kilometer or mile, you put repeaters because the signal deteriorates and then you need to maintain that cable.



So, that is why it becomes essential to see how much traffic is being exchanged between these two service providers based on that whether it makes sense to lay down a cable between or connect these two internet service providers directly or not or go through an intermediate exchange point is a better option or not, that is the business decision you need to make.

(Refer Slide Time: 24:07)

How fast is the backbone ?

- ❑ 100 Gbps trunk speeds
- ❑ Mainly rely on BGP for exchanging routing tables

| Link Type | Speed            |
|-----------|------------------|
| OC-1      | 51.84 Mbit/s     |
| OC-3      | 155.52 Mbit/s    |
| OC-12     | 622.08 Mbit/s    |
| OC-24     | 1244.16 Mbit/s   |
| OC-48     | 2488.32 Mbit/s   |
| OC-192    | 9953.28 Mbit/s   |
| OC-768    | 39,813.12 Mbit/s |
| OC-1536   | 79626 Mbit/s     |
| OC-3072   | 159.252 Gbps     |



Neminath Hubballi © Indian Institute of Technology Indore

So, for the backbone (these undersea cables) there are a bunch of standards and over a period, people have developed how much capacity or bandwidth they can offer. So, right from something called OC-1, OC-1 was able to provide 51.84 Mbps of the backbone data. If you leave one cable of the type OC-1, the maximum bandwidth that we get is about 51.84 Mbps.

Similarly, there is a standard called OC-3 that will give you around 155.52 Mbps of data and so forth. You can see that as time passed and the developments in optical fiber technology brought several Gbps or Mbps of the bandwidth to the backbone. So, a technology called OC-3072 is now able to provide you with 159.252 Gbps of bandwidth at the backbone level. So, that is actually a lot of bandwidth.

So, using this kind of bandwidth, now, if I want to design a router which is doing the lookup operation and routing the packets, how do we actually could design such a massive scalable router, what is that, that goes inside such routers which are high-level, high availability, operational 24x7, able to handle a large amount of the traffic and so forth.

(Refer Slide Time: 25:50)

Routers Everywhere !

- ❑ Core Router – Provide trunk connections
- ❑ Distribution Routers- Port density
- ❑ Border Router
- ❑ Service Router

Courtesy: Philip Smith

NPTEL

Neminath Hubballi © Indian Institute of Technology Indore

So, the routers in this case, I would say the border routers are one type of router, but there are other types of routers as well. So, the border routers are sitting at the boundary between two networks, two autonomous systems, and one entire internet service provider also can be thought of as an autonomous system. And in between these, there are many routers that are interconnected in some fashion something like this and this guy might be connected to another border router and something like this.

So, the routers which are here, here, here are all the border routers but the routers which are inside the backbone or the internet service provider's internal network, which are not directly providing the service to the end users, meaning if I am home subscriber, my connection from my home to the service provider's network does not terminate in these routers, such routers are called as the core routers.

And again core routers also need to handle a large quantum of traffic, but on the other hand, I might take a connection from this router and then bring it to another router and from there, multiple numbers of users or subscribers are connected to, so such routers are called as the distribution routers or the service provider service routers, where the end users typically connect to these routers.

So, be it my home subscriber or be it the institution subscriber like IIT Indore. So, all these subscribers actually come in and terminate at this router which is called the distribution router or

the service router. And the reason why we are making this differentiation is: the kind of operations that are done by a distribution router or an edge router or core router or a border router are different; they are needed to handle different kinds of operations.

So, the kind of filtering, lookup and all kind of operations that they do at an edge router is completely different than what the core router is supposed to do. So, in nutshell, the core router is typically needed to just do the lookup operation as quickly as possible and then send the traffic to the or packet to the next router so that the packets can actually go through.

So, if the packet comes here, I do not do much of the operation, you just send; find out quickly what is the next hop where I need to send this packet and just send it. These routers need to handle a large quantum of traffic. But on the other hand, this edge router can do some other set of the operations like filtering, the quality of service, traffic policing, how much data who is using, and keeping account of the bandwidth consumption of the end users, and all these things can be done at the edge routers.

So, the point is, depending upon where exactly these routers are located, and what functions they are supposed to do, based on that I need to customize the router's operation. So, this is a core router, it is going to sit in the backbone, it needs to do the lookup operation quite quickly and then you handle the massive amount of traffic. So, I engineer that router for that. This is an edge router connecting the end users, I require accounting information, and I require policing operation, policing means how much bandwidth is in use, and what kind of applications I am using, let us say I am a home user and I have a 10 Mbps connectivity at my home, but if I send traffic more than 10 Mbps, then the service provider usually throttle that, okay! you are sending too much traffic but you are allowed to send only 10 Mbps. Now, I am not able to handle this excess traffic, you drop the packet then and there itself at the edge router. And then transmit only 10 Mbps connecting the traffic from that router to the next router. So, these kinds of operations are done at the edge so that customization is actually done in the routers.

(Refer Slide Time: 30:46)

## Number of Prefixes are Growing



| Date     | Prefixes | CIDR Aggregated |
|----------|----------|-----------------|
| 16-08-22 | 929299   | 505463          |
| 17-08-22 | 929384   | 507128          |
| 18-08-22 | 929579   | 508444          |
| 19-08-22 | 930070   | 508937          |
| 20-08-22 | 929550   | 509291          |
| 21-08-22 | 929527   | 509713          |
| 22-08-22 | 929943   | 503955          |
| 23-08-22 | 930468   | 510491          |

Handwritten notes in red and blue ink to the right of the table:

- 10.10.\*
- 20.15.\*.10
- 20.15.1.1



Neminath Hubballi © Indian Institute of Technology Indore

As the days pass on, the network capacity, and the bandwidth available at the backbone keeps on increasing and more number of the subscribers join the network. As the internet subscribers increase, as the bandwidth increases, it puts more pressure on the backbone routers and the border routers to deal with things quickly.

So, in order to understand the scale of the operation that is required at a typical border and the core router, I took this diagram, which is a set of something called prefixes. The IP address that we saw in the previous example in the first slide, that is a complete IP address. But it is not necessary that you need to have the complete IP address. But there is something like this can also be there. Anything that is destined for 30.0.0.\* star does not matter whether it is 1 or 2, you still need to use Port Number 4 which is what it says. So, if you recollect the CIDR operations classless inter-domain, the way the routers handle this is: they do not have the complete IP address of every destination stored in their routing tables. You merge several of the IP addresses, and get something called the prefix. I might have a prefix for something like the 10.10.\* I might have 20.15.\* something like this.

Let us say, I have an IP address of 10.15.1.1. So, this precisely falls under this prefix and if it is telling you to route to port number 10, send it to port number 10, so I get such a kind of the IP address, complete IP address, but I am able to send that to.

So, I forward this IP address. If an IP packet comes to any of the routers and this particular IP address, I am going to forward that by consulting these prefixes that are available to port number



10. So, this is called the prefix structure and the routers typically keep such prefixes, not the complete IP address. You might have one or a set of the complete IP addresses and few others may be in this prefix format. By aggregating some of these IP addresses such prefixes are actually generated.

And what you see on the left-hand side is: number of the prefixes that are available with a typical backbone router as seen by, this is one of the routers that is located in Japan, and on a particular date on the sixteenth of 8 2022, on that route, 9,29,299 prefixes were there. These are distinct prefixes something like 10.10.\* and 20.15.\* something like that.

And these prefixes distributed, sometimes are aggregated to minimize the number of entries in the routing table. So, on the rightmost column, what we see is CIDR aggregator. If I am able to merge two prefixes and get one so that will decrease my total number of the prefixes inside the routing table.

See, if the large, or smaller the number of the prefixes you have, the better it is, because given one packet, you pick up the IP address, the destination IP address from that packet, and then you go and consult that table, whether prefix number 1 is matching or prefix number 2 is matching or prefix number 3, and so forth up to n. So, the larger the number of prefixes you have, the more amount of work you need to do to do the search because of the look-up operation.


So, that is why the aggregation actually helps, keeping a smaller number of entries will actually give you performance improvements. So, as you can see, over a period of time from sixteenth to twenty third. Twenty third, prefixes increased to seek from 9,29,299 on the sixteenth August to seek from 9,30,468 as of twenty third date 2022. So, what it means is the number of the prefixes over a period of time is going to increase. This increased number of the prefixes will put pressure on the router.


So, as the time passes, you actually need to handle larger IP tables, larger routing tables and more amount of searching you need to do, so that is going to be actually a complex thing. How do we actually handle such things? We know that bandwidth is going to increase. We know that prefixes are going to increase over a period of time. So, we need algorithms that can actually do this for us.

(Refer Slide Time: 36:09)

## IP Router Operations

- ❑ Data Path Operations
  - ❑ Forwarding the packet
  - ❑ On every packet – speed of operation
- ❑ Control Plane Operations
  - ❑ Exchange routing information (RIP, OSPF, BGP)
  - ❑ Software implementation





Nitin Kumar © Indian Institute of Technology Indore

So, any typical router that you think of performs two operations, one is called the data path operation. So, data path operation is basically you are given a routing table, look at that routing table and the packets that are coming on your interface, pick up those packets, go to a routing table, and then make forwarding decisions. These forwarding decisions need to be quickly made.

And the second lookup operation is called the control plane operation where we construct the routing table, whatever the routing table we constructed, we used to do operations in real-time, but in the background, I keep on updating my links. New people are joining, links are breaking or something is happening, and topology changing.

So, I keep learning, what is happening, and what things that are happening inside my network, and keep updating that routing table, and that is typically done with a bunch of routing algorithms. So, you might have studied algorithms like RIP OSPF and BGP, these algorithms bring the status information of the links and connectivity between the routers, you aggregate that and construct that routing table. So, this operation is called the control plane operation.

So, the control plane operation is not real-time. Constructing the routing table is not real-time but on the other hand, forwarding the packets to the next hop needs to be done in real-time because of the fact a large amount of the packets is coming.