Communication Networks
Prof. Goutam Das
G. S. Sanyal School of Telecommunication
Indian Institute of Technology, Kharagpur

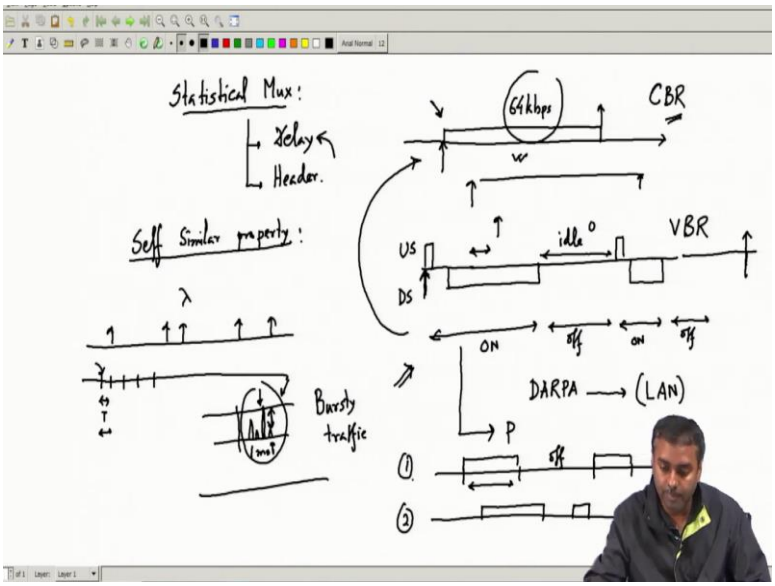
Module - 06
Data Networks
Lecture - 28
Introduction to Data Networks

.In the last lecture we have probably finished our discussion about circuit switch networks. So, what we have promised will be deriving formulas for calculating the blocking probability of the trunk switch. So, that is something we have done and basically, our understanding of queuing theory and finally, the continuous time Markov chain helped us in modeling all those different kinds of queues and two of them were actually helpful towards developing two different formulas Erlang B and Erlang C ok.

So, that kind of concludes our circuit switch network we have seen multiplexing, different kinds of switching trunk switching, and general local subscriber loop switching. And then we have also seen the synchronization issues, we have seen how to design switches, the complexity of it and we have seen how we actually calculate the blocking probability of different kind of switches.

So, all these things have been covered. Now what will be starting in the second half of our course is the part of data networking. So, basically it is we are going from circuit switch network towards another kind of network this is something we have already discussed in our introductory sections.

(Refer Slide Time: 01:55)



So, in that particular part what will be trying to do is something like this. So, we have told that there is a concept called statistical multiplexing. So, this is a concept which we have discussed in detailed. We have given some example also what statistical multiplexing all those things and due to that we have seen two particular things I am just quickly recapitulating our earlier lectures the first week lectures actually.

So, two particular things happens over here. One is that it is no longer the traffic characteristics that is it is no longer a continuous bitrate kind of traffic which was happening in voice traffic. So, circuit switch network was predominantly for voice traffic it was a telephony these things. So, there are the traffic was very smooth it was always a sample taken at every 125 microsecond and that was PCM encoded with 8 bit or 10 bit whatever it is ok.

So, it was a constant bitrate that was coming. So, it was a very smooth traffic whereas, data networks or data traffic that is actually the traffic that we generate interacting with machine ok. So, human machine interaction generates this kind of traffic let us say I means surf internet ok I go through some different W, W, W pages ok. So, and I basically go through some of my emails. So, all these requires human intervention.

So, as you can see if you again try to see the events, we were earlier seeing the events. So, if you try to see the events over here the events are something like this that every time when I will be generating a data ok. So, let us say in circuit switch network what

happens I generate a call that is a, event probably, but after the call has been generated it will be a constant traffic that will be going on ok, till I actually finish the call ok.

So, at that time if it is a voice call it will be always 64 k b p s constant bitrate will be generated in the circuit ok. So, that was the case only thing is that this initiation those are random. So, when there will be another call will be coming when that call will be finished that was all random so, that part was random. Now what is happening I started surfing the internet this morning ok. So, that is one event that is the big event, but inside that there are small event that is how my data will be generated.

Now, no longer this will be a constant bitrate. So, in between I will be doing a lot of things some page I see ok. So, I fetch that page so lot of things will be downloaded at that time very high data will be coming in some small request. So, if I just instead of giving this let me do that in a more proper way. So, let us say this is the event that I started surfing internet.

So, at that time this is the upstream data, so basically data going from user towards the network ok maybe I am trying to look into some HTTP pages; so, basically or Wikipedia page let us say that is the thing I am doing. So, initially it will be sending some request then after the request is being received by the server. So, this is upstream because I am sending request then there will be a downstream traffic it will be fetching the page ok.

Once the page is fetched then I will be reading. So, that is the time when I will be idled, after I finish reading maybe I will be more interested into doing something else or reading some more things. So, then again I will be clicking somewhere again something will be fetched. So, this kind of thing will be happening if you observe the overall connectivity till I actually disconnect my network connectivity I will be having a lot of this on period well lot of traffic will be either going or coming.

Then followed by idle or off period then another on period followed by another off period and so on. It will be intermitted with on and on structure which was not present in a circuit switch network that is the major difference which is happening.

Whenever I start talking about you take it anything you suppose you are doing gaming, online gaming suppose you are doing online shopping, suppose you are doing any kind

of internet related application you are downloading some video ok. So, in all cases even if you are actually let us say seeing streaming video.

So, there also same thing will happen because of the video and the way we encode they are unlike the voice data it does not constant bitrate, in the video there is an I frame, P frame, and B frame they have different amounts of data when they are transmitted so there will be sometimes high data are going, sometimes low data are going ok.

So, these kind of things for any kind of data fetching always will be happening, Email if you want to do again human machine intervention ok or interaction that is happening. So, basically what will be happening whenever you are trying to see the Email then, that Email will be fetched from the server. So, lot of data will be coming, but then you will be reading so at that time nothing will be happening.

So, again an off period followed by an on period followed by an off. So, this kind of thing will always happen. So, that is the major difference when we start talking about machine-to-machine communication or a human sitting in a machine to another human sitting in another machine so communication between two machines. When this kind of things where mostly some information will be exchanged between machine and then they are means interfaced with some human being.

So, if this is the kind of scenario that is a new kind of data. So, that data is no longer this kind of regular data or we call it constant bitrate CBR. So, from constant bitrate we are going towards some traffic which is called variable bitrate or VBR because it has this on period off period at the on period the bitrate will be very high, off period bitrate will be 0. So, something like that. So, the bitrate will vary or like video, the different frames have different amounts of data.

So, basically, they will be having associated different bitrate requirements. So, this is what will be happening. So, what we can see is that whenever there this paradigm shift was happening that from voice traffic. So, we no longer just want to transmit voice if information sharing is the most important part of networking then why just voice why cannot we share any digital information that we have ok. So, any analog information also can be transferred we can transfer it into the digital domain and then transfer ok.

So, that also we can do ok. So, if it is a digital network ok. So, for any digital information we want to share then, why to just restrict ourselves to voice why cannot we go to all other kinds of data, all other information sources ok. So, once people started expanding it this first happened in the US one DARPA.

So, at that time they were just interested in a very small vicinity a local area network they used to call local area network or LAN later on that was popularized as LAN on campus and everywhere you will be seeing that we construct LAN.

So, will talk about those things what is LAN and all those things. So, initially, it was initiated by DARPA. So, they started constructing this LAN where there were multiple machines at that time it was like computers ok. So, that age computer whatever that was at that time.

So, that earlier age computers and they wanted to share information from one machine to another machine ok. So, once they started looking into that the first thing that came into the picture because in any networking the first thing that comes into the picture is the traffic, is that traffic different? We can now see yes the traffic is different.

So, basically what we can still say is that when a person starts his session and when he ends ok, these things can still be characterized as voice because that is still like a human being intending to generate some session when he will be generating because he is doing it randomly. So, that can still that can have memoryless properties, but the problem happens when we see inside that it is no longer a very smooth regular kind of traffic. So, there was a seminal paper around the 1990s in that year only 1990 or 90 within 1992 3 so, by Sally Floyd and Paxson.

So, what they did was they tried to measure this traffic earlier for voice traffic people have measured, and they could see that it is all exponential ok. So, they could characterize it that the arrivals process was characterized by exponential sorry Poisson distribution inter arrivals exponential service process was the duration of the call was exponential so all those things they could characterize.

Now they started measuring the LAN traffic once they had installed LAN means once the system was connected they started seeing means before designing it they started seeing what kind of traffic was being generated.

So, basically, what happened was they started measuring this on period and off period and historically it has been seen that these distributions are no longer exponential, they have memory and they are generally called they have long-range dependency we will later on talk about if we get time to discuss about traffic characteristics then we will talk about that.

So, they generally have long-range dependency we will talk about what is long-range dependency and they generally have heavy-tailed distribution. So, heavy-tailed distribution means exponential distribution if you see the tail probability actually goes vanishingly towards 0 whereas, with the heavy tail distribution we cannot really say that their tail will be still heavy. So, there will be still a sufficient amount of a large amount of probability that the corresponding period on period or off period whatever it is might be very large.

So, that distribution will still not be vanishingly 0, it will have some finite value so that that kind of thing. So, this heavy-tailed distribution comes into the picture and generally is characterized by Pareto distribution. We are not going into the details of that distribution, that heavy-tailed distribution how it is characterized. So, we are not going into the details of that particular thing that how the heavy-tailed distributions are characterized and what the statistical properties are that we are not doing, but it is a different distribution and it has memory.

So, all the analysis that we have started doing over here might not be true that exponential distribution and associated queueing theory, the Markovian queueing theory, is something we should understand that is the first thing.

Generally, what happens is the traffic characteristics people that Paxson and Sally Floyd whatever they did they could see that they generally have some heavy-tailed distribution like Pareto or Weibull or log normal this kind of distribution mostly Pareto it is Pareto distribution.

So, they are Pareto distributed, both on and off periods are generally Pareto distributed, and not only that this one on and off period comes from one user there are only multiple such users. So, all will be integrated. So, some if you see user one it might have independent on and off periods which are all independently Pareto distributed, so, this is on. So, where traffic will be there this is off, this is again on something like that then the

second user might have some independent on and off period that might be completely differently distributed.

So, multiple users in the LAN might be present and will have different kinds of on and off periods and this will be aggregated after this aggregation the traffic look looks like a completely different kind of traffic, and that traffic has this particular self-similar property. So, it is called self-similar traffic. So, what is self similar property, I will just give a very brief overview we are not doing traffic engineering over here we will not go we just want to be familiarized with the actual traffic.

So, that later on whatever we do we will be able to characterize it. So, what is self-similarity, self-similarity something like that, will you show exponential traffic? So, let us say we have this exponential arrival, now Suppose I have this time and this is the inter-arrival, now this time actually I divide into small bins of equal spacing and within this time span how many packets have arrived?

So, I am counting ok and that histogram I plot of course, will be looking like a Poisson kind of thing because the count is Poisson. So, if I draw the histograms it will look like that, but what will happen if I keep on increasing this T then it will actually average out ok as T goes bigger and bigger. So, what will happen is it will almost see most of the time this lambda amount of traffic. So, basically, it will even out. So, every time you will be almost seeing lambda amount of traffic every time.

So, this evening out will be happening if you have exponential arrival whereas, for self-similar that is not the case, self-similar actually it means almost constructing the property even if you increase the timeframe for your counting. So, that is the property of self-similarity; this is a very peculiar property and this only happens it can be proven that it only happens because of an infinite number of integrations of independent Pareto on and off, this is something that people have proven already and that is part of the traffic engineering but this is something you will be able to see.

So, basically, the traffic characteristics remain the same. So, in a small amount of time if you try to see there, whatever traffic characters are supposed sometimes it is this much amount of traffic coming, sometimes this much amount of traffic. So, something like that this pattern is there now if you also take a bigger amount of time. So, every millisecond

if you calculate you see this kind of traffic. So, sometimes huge traffic is coming, sometimes no traffic is coming, something like that.

If you now put the time bigger you will again see almost a similar pattern which is not the case for exponential it will smooth out ok, but over here if you go to a bigger amount also you will again see a similar pattern that again is actually a huge amount of traffic then no traffic again a huge amount of traffic no traffic. So, these particular characteristics of traffic people call them bursty traffic. So, what do we mean by bursty traffic; that means, it arrives in burst no matter how long you observe there will be still aberration. For exponential if you increase the time frame probably you will see a very smooth amount of traffic coming.

So, it will almost look severe at that point, but for bursty traffic that is a problem that no matter how long period you observe, it will still have this burstiness; that means, in some timeframe huge amount of traffic will come in other timeframe less amount of traffic will come. So, there is some average value of course, over time and it might be still stationary; that means, over time whatever average numbers are coming also will be similar probably, but the problem is no matter how much bigger you make the time still it will look like bursty ok.

So, that is a danger for any kind of design because what is happening whenever means earlier for telephony what used to happen 64 kbps I reserve 64 kbps I am fined nothing will happen to it call will go very smoothly over here what do I reserve I do not know, for how much amount of time what amount I will reserve I do not know because sometimes huge amount of data are coming sometimes low amount of data are coming I might design it for the peak.

So, whatever the peak amount will be coming I might reserve that much bandwidth for him, but then I am wasting because most of the other time it is actually not supplying that amount of traffic and I am actually wasting this entire amount for the entire duration, but that is a huge wastage; that means, I will be over-provisioning the network and that comes with a cost. So, the network cost will be very high this is where statistical multiplexing is even more important.

So, whenever traffic if you now go back to our class second class or third class you will see that over there we have discussed that if the input traffic has sometimes traffic,

sometimes no traffic if these kinds of gaps are there how do you take account of those gaps.

So, basically, what you try to do as you go inside the network you want to because they are the aggregation will be more and more, more number of streams will be coming if you design the inside network for everybody's peak then it will be a disaster because huge amount of capacity will have to give either or otherwise your capacity will be restricted not many people can simultaneously connect, most of the time they will be blocked whenever they wish to connect their means data transfer will be either blocked, their packets will be dropped or they will be having huge delay ok.

So, this is something you do not want to do means either the transfer costs will be very huge or you will not be even able to physically do that and then you will be actually paying a penalty towards user experience. So, I want to watch a movie and I will see [FL] it keeps on rotating and I do not get anything ok? So, this is what will happen if I cannot design it properly.

So, this is where you will see that a lot of under-provisioning is being done as you go inside the network and this under-provisioning is facilitated by the concept called statistical multiplexing.

As you can see, so therefore, if you remember those slides that we discussed earlier when we did statistical multiplexing what was happening at the output, was very smooth. So, basically, as you keep on doing statistical multiplexing the traffic becomes smoother and smoother at the cost of what; at the cost of wherever you do statistical multiplexing there you will have to sometimes allow delay.

So, we have discussed that. So, whenever we do statistical multiplexing we need to allow delay that is something we have to do and we need to also add a header because no longer because we do statistical multiplexing, we have discussed that very in detail that we do not have this time location as the address because time location is now random, I do not know when I will be putting it depends on all other things ok, all other presence of others traffic.

So, therefore, I need to put a header that identifies the packet where it will be going. So, the header must carry all the information about it such as transportation. So, that network can just be reading the header transfer the packet.

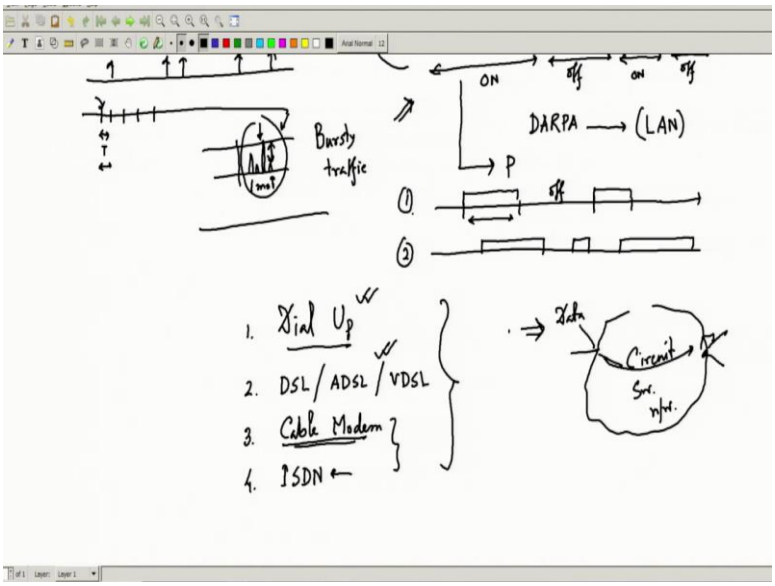
So, this is where the concept of packets comes into the picture earlier it was a circuit, now the concept of packets comes into the picture ok. So, we will see in our data network these concepts are very important this concept of packet that is what we will be actually discussing over the next course of these things that how the packets are being constructed and what is the role of this header, how do we manage this delay ok. These are the things which are very important and how do we design a network.

So, this is one part and the second part is the traffic that we have discussed because of traffic we are doing this. So, everything is getting connected now. So, because of that different kind of traffic now we will have to do these kinds of things no longer this circuit switching we can do, but what we can see more and more statistical multiplexing we are doing towards the core of the network we are getting more smoother traffic.

So, basically, the circuit switch network should be thrown out probably not, we can still apply them in the core part of the network and that is how the whole data networking actually started. The genesis of that was that only the data portion would be there in the access part rest of the things will be equivalently carrying data as well as your voice traffic.

That was the era when we had access we did not have at that time broadband access, fiber to the home, or 5G or 4G so these kinds of things were not there. So, at that time what we were having is something if you remember. So, initially, who used to use them was dial-up, then came something like DSL Digital Subscriber Loop, ADSL, or VDSL ok?

(Refer Slide Time: 24:50)



So, that was the second thing then there was the third part which was cable modem and then there was a standardization which was ISDN. So, these initiatives were there so that actually it has 2 advantages. One is already existing infrastructure of the circuit switch network, which we can use and only in access we put data inject data how do you do that we will see that.

We will try to see the technology dial up the technology of this one and very briefly probably discuss these two of course they were not very popular, some of these cable modems are still being used, but that is a different thing that was not on voice actually, this was for with the video distribution analog video distribution they used to do through this means this satellite TV actually.

So, that network they used to take and ISDN was a, afford to standardize it. It was not that popular because in between the entire data network came into the picture, so happened.

So, basically, this is what will be happening only in the interface I will be somehow converting these things to a circuit switch network, I know, I can do that because the data input will be bursty. I do statistical multiplexing over here, I do under-provisioning over here, we will see all these things later on and then I do circuit switch networking over here.

So, that is why you will be seeing any campus like let us say IIT Kharagpur. So, the campus takes a kind of lease line; this leased line is taken as a circuit switch network ok. Some amount of a particular circuit or particular connectivity I take a very high data rate connectivity probably let us say some initially they used to take some Mbps connection now they are going even higher than that because the data overall data rates are getting increased ok.

So, you take a very high connection, not 64 kbps or something like that not like dial-up. So, you take a very high data rate connectivity and that is the circuit actually goes to the circuit switch network, and then the entire data gets routed through the circuit switch network only in the other end of the network where you want to actually transfer the data that is where the circuit delivers the data over there and then it gets converted to data packetized data and then you do means this statistical multiplexing whatever you have done you do demultiplex them and then deliver the data ok.

So, we will see how these things coexist actually that is the beauty of networking in networking you do not throw anything and you design them in such a manner that nothing has to be thrown. The circuit switch network was designed for voice, there is no problem in that this can still be handy for transferring data. We will see those things, we will appreciate those things, and to date that has only happened internal core backbone switch networks are all circuit switch networks till date ok.

So, what we will try to do in the next class, is probably we will try to see how these things come in handy and how the data network actually works on top of this. This is something we will try to appreciate. So, we must understand first how the data network works by itself and then how it works on top of these kinds of infrastructure, combined infrastructure.

Thank you.