

Broadband Networks

Prof. Karandikar

Department of Electrical Engineering

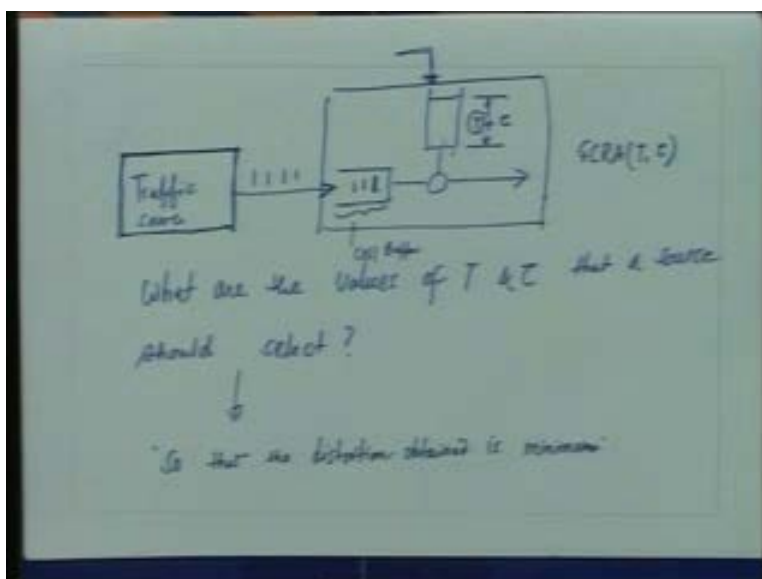
Indian Institute of Technology, Bombay

Lecture - 6

Traffic Descriptor in ATM

Today, now we will discuss the problem of how a source can choose the values of the GCRA (T , τ) parameters such that it accurately represents the traffic characteristics.

(Refer Slide Time: 00:57)



So, as we are seeing in the previous lecture that we have traffic source which is random and bursty and this source is generating the packets randomly and it is then regulated by a leaky bucket regulator or a GCRA (T , τ) regulators. Now, this bucket is initially full with a depth of T plus τ and whenever an ATM cell comes, it takes away capital T units of fluids while the fluid is accumulating in this bucket at a unit rate.

Now, there may be a cell buffer here. So, in case if a cell finds that the bucket does not have adequate amount of fluids that is capital T units of fluids, then the cell is buffered in this cell buffer until the adequate amount of fluid gets accumulated in the bucket. And, when the adequate amount of fluid is available that is the capital T amount of fluid is available, the cell is considered conformant and it is outputted from the GCRA (T , τ) regulator and the bucket is decremented by the capital T amount of fluids.

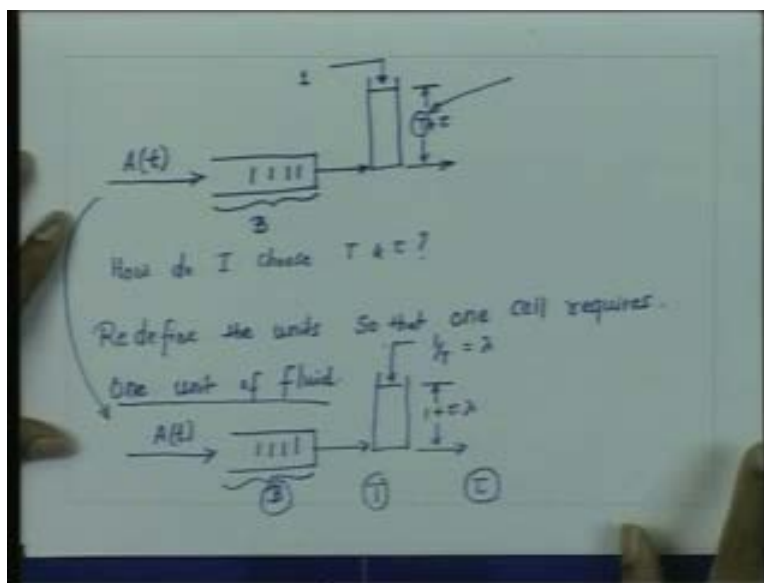
Now, our question is that what are the values of T and τ that a source should select? Because, please note that since the cell is buffered here in the cell buffer till adequate amount of fluid gets accumulated in this bucket; the traffic source actually incurs some amount of distortion. So, we would like to ask this question that what are the values of T and τ that a source should select so that the distortion which is obtained by the traffic source is tolerable.

Now, this distortion could be in terms of acceptable delay at cell buffer or an acceptable loss at the cell buffer, the cell buffer which is present at the traffic source itself. Now typically, in practice, the network will offer a range parameters of T and τ from which the source would be required to choose and depending up on what values of the parameters T and τ have been chosen, a particular call will be priced or charged.

So therefore, the values of the parameters T and τ are not only related to the quality of service guarantees that a network would be able to offer, but it is also related to the pricing of the calls that a network operator is likely to make. Therefore, a traffic source has to make a judicious choice of what values of the parameters T and τ it should choose such that they form the minimal descriptors, minimal traffic descriptors for the traffic source in terms of both describing its traffic characteristics as well as obtaining the required quality of service guarantees and pricing of the calls.

So, I will just give you an illustration of how these parameters should be chosen and related to the effective bandwidth of the traffic source that we had studied in the previous lectures; we will later on come back to this problem of how to choose the parameters T and τ such that a particular cost function related to the distortion is minimized.

(Refer Slide Time: 4:35)



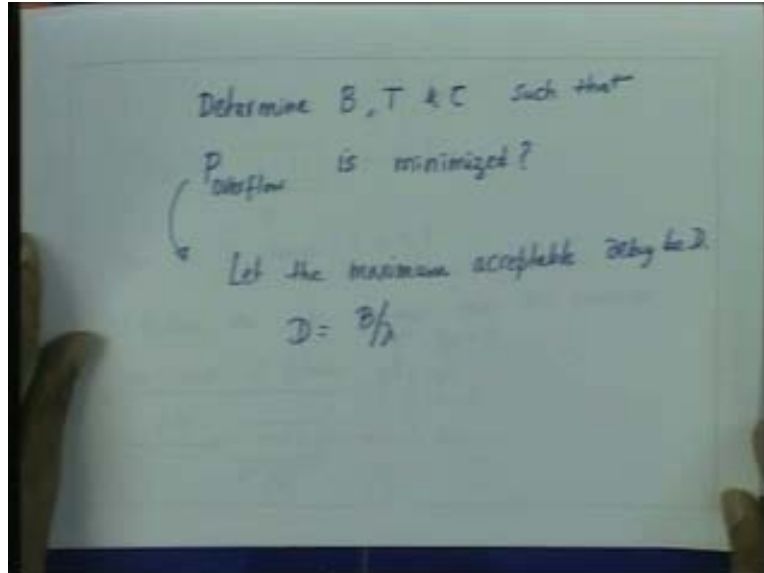
So, let us consider how you can choose this. So, what we are assuming is that a traffic source is

bursty and the traffic source is represented by $A(t)$ and it passes through a cell buffer that we had considered previously and let us assume that the cell buffer has a size of B and then there is a leaky bucket and this leaky bucket as we had considered has a depth of T plus τ and the fluid is coming into this bucket at a unit rate, at a rate of 1 amount of fluids per unit of time.

Now, our question is that how do I choose the values, the value of T and τ ? So, what we do is that we redefine the... let us redefine this above units and we redefined the units in such a manner so that 1 cell requires, 1 unit of fluid, **requires 1 unit of fluid**. Note that in this we are assuming that each cell requires capital T units of fluids, so we redefine this units in such a manner that 1 cell requires exactly 1 unit of fluid. So, we can **you know** rewrite this diagram by saying that we have again a random bursty source $A(t)$ which passes through this buffer. Again, these buffers has a maximum size of B but these leaky buckets; now we assume the fluid is coming at a rate of $1/T$ and let us call this $1/T$ equal to λ and this bucket has a depth of $1 + \tau$ into λ .

So, we have just changed the units in such a manner that each cell requires now 1 unit of fluid instead of capital T units of fluid that it required earlier. So, we have redrawn this picture and our problem now in this picture is that to determine the values of B that is how much amount of cell buffer we should keep here to determine the value of T and to determine the value of τ such that **you know** the probability of buffer overflow here is minimized.

(Refer Slide Time: 7:29)

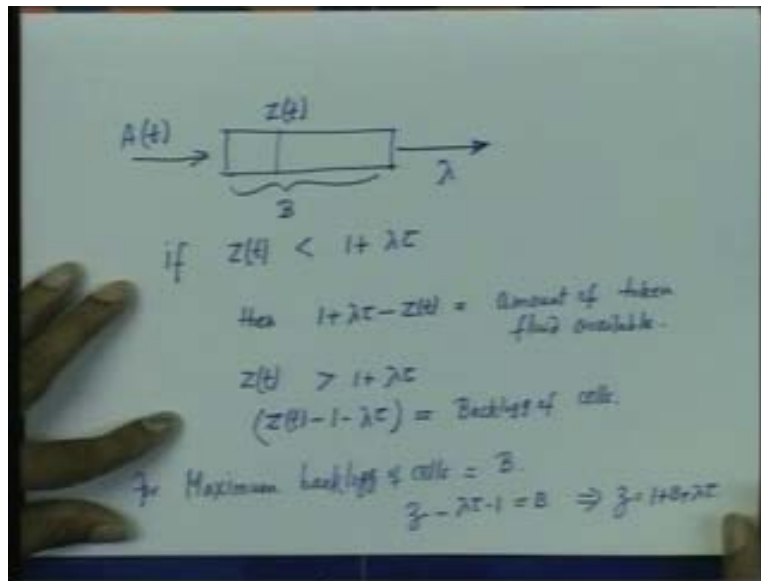


So, we post this problem as to determine B , T and τ such that probability of buffer overflow is minimized. We can also post the problem in other way in which we can say that let the maximum acceptable delay **B B D**. So, if maximum acceptable delay is D , then D is actually given by B by λ , where B will be the maximum acceptable backlog.

So, in this picture if you see, what we are saying is that one problem is to determine the values of B , T and τ in such a manner that the probability of a buffer overflow is minimized here or else the other problem is to assume that there is a maximum acceptable delay D such that there is a maximum acceptable backlog B and then how do we choose the values of T and τ such that the delay D is acceptable?

Now, to consider this problem, we can have an equivalent representation of this entire phenomena and this equivalent representation can be simply given by a single buffer.

(Refer Slide Time: 9:07)



So, we consider a single buffer whose length is capital B and this buffer has an input of the same random source $A(t)$. But this buffer is serving at a rate of λ and let $Z(t)$ represent the buffer length at time t . So, this problem of where we are considering that there is a random traffic source $A(t)$ which is passing the fluid through a GCRA (T, τ) shaper can be equivalently represented by a single buffer where the input is $A(t)$, the maximum buffer length is B and this buffer is being served by λ .

Why we are saying that this is equivalent, because note that each cell is outputted from this GCRA (T, τ) parameters and for each source, it takes 1 unit of fluid and on the other hand, the fluid is getting accumulated at the rate of λ and the cells can get buffered in the cell buffers upto a maximum length of B because the cell buffer is having a length of B . So therefore, you know, the 2 problems are equivalent. Of course, in these equivalent problems, the amount of token fluids which is available in this token bucket will be related to this queue length $Z(t)$ which is there.

For example, if we say that $Z(t)$ is less than $1 + \lambda\tau$; if this is so, then $1 + \lambda\tau$ minus $Z(t)$ will be the amount of token fluids which will be available. So, if we say that this buffer $Z(t)$ is less than $1 + \lambda\tau$, then equivalently it means that $1 + \lambda\tau$

minus Z - this much amount of token fluids will be available. Or on other hand, if $Z(t)$ is greater than $1 + \lambda\tau$ that means the backlog is greater than $1 + \lambda\tau$; so on the other hand, in this case, the maximum amount of cells which can be transmitted at a time should be equal to $1 + \lambda\tau$ so because each cell requires one unit of fluid and therefore only these many cells can be sent out.

Therefore, when $Z(t)$ is greater than $1 + \lambda\tau$, obviously $Z(t) - 1 - \lambda\tau$ - this much amount of cells needs to be backlogged. So, this will be as a backlog of cells. Now, in order to consider this problem where we are saying that the maximum acceptable delay is D and therefore the maximum acceptable backlog is B , what it means really is that the maximum backlog is B . So therefore, **if for the maximum backlog of cells** for maximum backlog of cells to be equal to B , let us say that $Z(t)$ has a value is of Z . Then, we say that $Z - 1 - \lambda\tau$ should be equal to B and which means that Z would be equal to $1 + B + \lambda\tau$.

Now, so this is the value of $Z(s)$ which gives us the maximum backlog because the maximum backlog which can be there in this buffer will be equal to B .

(Refer Slide Time: 13:22)

$$\text{Prob}[Z(t) > B+1+\lambda\tau]$$

$$\text{if } \lambda \geq \alpha(B)$$

$$\text{Then } \text{Prob}[Z(t) > B+1+\lambda\tau] = e^{-\alpha(B+1+\lambda\tau)} = e^{-x}$$

$$\lambda \geq \alpha\left(\frac{x}{B+1+\lambda\tau}\right) \Rightarrow \boxed{\frac{1}{\tau} \geq \alpha\left(\frac{x}{B+1+\lambda\tau}\right)}$$

So now, we ask this question that what is the probability that $Z(t)$, **probability $Z(t)$** is greater than $B + 1 + \lambda\tau$. This is the question that we should... we are asking. Now, if you assume that this input $A(t)$ **this input $A(t)$** has an effective bandwidth of $\alpha\delta$ and if you assume that the transmitter is serving at a rate which is greater than the effective bandwidth of the source which is $\alpha\delta$; then we know from the effective bandwidth theory that the probability that $Z(t)$ is greater than $B + 1 + \lambda\tau$ will be equal to e raised to power minus $\alpha\delta(B + 1 + \lambda\tau)$. Now, so let us recap what we are doing.

Essentially, we are saying that there is a random traffic source which we are representing by A

(t). Now, this random traffic source is being shaped by a leaky bucket regulator GCRA (T, τ) regulators and we are trying to ask this question that what are the best possible values of T and τ such that the maximum acceptable delay **such that the maximum acceptable delay** at the traffic source is limited to or is bounded by the capital D.

Now, moment we say that the maximum acceptable delay D is bounded it means that the maximum acceptable backlog B is bounded. So now, we are trying to ask this question that what are the values of the parameters T and τ such that the maximum delay at the traffic source is bounded or the maximum acceptable backlog is bounded which is τ . This backlog is occurring at the cell buffer at the traffic source.

Now, we formulate an equivalent problem of this and the equivalent problem is that it is equivalent to saying that we have a single buffer where the input is our same traffic source and this single buffer is being served by a transmitter which is transmitting with a rate of λ and where this λ is equal to $1/T$ cells per unit of time. So, this transmitter is transmitting with a rate of λ , the input is $A(t)$ and we are asking this question that what is the probability that this $Z(t)$ is greater than $1 + \lambda\tau + B$.

Now, this will ensure that the maximum backlog in our original cell buffer will be bounded by capital B. Now, this problem is equivalent to the problem that we had considered earlier where we had assumed that there is a single buffer transmitter which was transmitting at the rate of C cells per unit of time and there was a random traffic input and we were asking these questions that what is the effective bandwidth of a source.

And, effective bandwidth of a source essentially was equivalent to asking the question that how does the buffer occupancy distribution behave for a large value of the transmitter rate and for the large value of the buffer B and we found out that this buffer occupancy distribution or essentially the loss rate behaves exponentially and that is what gives us the way to calculate the effective bandwidth of the traffic source.

So, that is what we have written here that the probability that $Z(t)$ is greater than $B + 1 + \lambda\tau$ is given by $e^{-\alpha(B + \lambda\tau + 1)}$. Now typically, there will be a certain requirement of how much is this tolerable probability. We can say that this tolerable probability is say 10^{-8} or 10^{-9} . Typical values would be in the range of 10^{-9} to 10^{-10} . So, we can say that let this be equal to some $e^{-\alpha x}$, where x is some given quantity.

In that case, we can show that α will be equal to $x / (B + \lambda\tau + 1)$. Now, this means that the λ , the transmitter rate λ is greater than or equal to $\alpha / (B + \lambda\tau + 1)$ which we can also write to be as $1/T$ is greater than or equal to $\alpha / (B + \lambda\tau + 1)$.

Now, this equation gives us an insight of how to choose values of T and τ for a given value of acceptable backlog B and a given value of x through this equation; because the traffic source A

(t) is known, we can compute the effective bandwidth of this and this equation relates the parameters of the effective bandwidth to the parameters of the GCRA (T, tau) such that we can bound the maximum acceptable backlog at the traffic source.

Now, this concludes our discussion of the ATM traffic descriptors. Just to recapitulate the entire discussion, let me just again tell you that what we started really with the question of how to provide quality of service guarantees in an ATM switch. And, we said that there could be two ways of doing this; either we can do multiplexing without buffering and we can do multiplexing with buffering. Now, in the case of multiplexing without buffering; obviously there are no delays there are no queuing delays but we have to ensure that the probability of a cell loss is kept below at tolerable limit.

In the case of a multiplexing with buffering, there will be delays and therefore we need to ensure that the probability of a delay being bounded or the average delay or the maximum delay being bounded. At the same time, since the buffer is finite, we also need to ensure that the probability of buffer overflow is also kept small.

So, the way to ensure this is we found out that we can ensure this by defining a quantity of a traffic source which we called as the effective bandwidth of a traffic source. So then, the admission control problem becomes simpler. All we need to ensure is that the sum of the effective bandwidths of the traffic sources must be less than the output link capacity. If this is maintained, then we can guarantee both the packet loss and the delays where the effective bandwidth will be a function of the traffic characteristics of the source as well as the quality of service attributes that the source desires from the network.

Now typically, however, we have also seen that it will be difficult to determine the effective bandwidth of a traffic source because it requires the complete traffic characterization of the source. Therefore, the ATM forum standards has resorted to the deterministic description of the traffic source and we say that the traffic source can be completely characterized in terms of its peak cell rate, the sustained cell rate, the burst tolerance and the cell delay variation tolerance - these 4 parameters. And, these 4 parameters can be represented in terms of deterministic traffic descriptors which are represented by generalized cell rate algorithm the GCRA (T, tau).

Now, given the GCRA (T, tau), we have seen that we have a very simple admission control policy and a very simple transmission policy. However, the question remains that what are the best values of the parameters T and tau. Now, that question we try to answer partially. We will take up this question later. So, this concludes my description of how to obtain a quality of service guarantees at the ATM traffic.

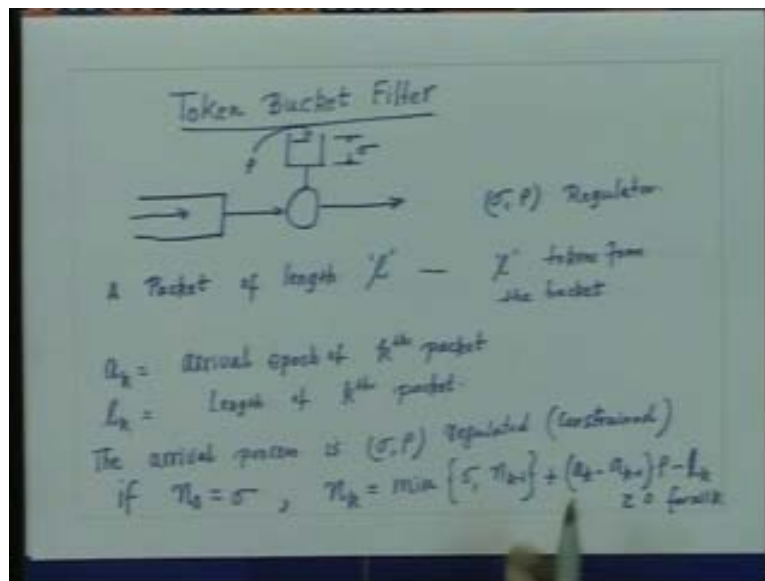
Now here, I would like to point out one thing that in this ATM environment, all packets were of the fixed length. They were all 53 bytes packet as we had seen earlier and therefore these fixed length packets were also called as cells. On the other hand, in the current internet, the packets could be of variable length. The difference between the ATM and the current internet is that in internet, the packets could be of variable length.

Now, when the issue of providing quality of service guarantees in the internet arose, the similar problems which we have seen in the ATM world are also applicable in the internet that is we also have the concept of a statistical multiplexing in the internet and when you want to provide quality of service guarantees in the internet, we also need to have the statistical multiplexing with quality of service guarantees and we therefore also need to invoke the definition of the effective bandwidth in the internet world as well. So, we now come back to the same question that since the effective bandwidth of a traffic source is difficult to determine and we want to represent a traffic source by the deterministic parameters, the similar concepts which were like a leaky bucket regulator have also been applied to the internet world except with an important difference that in the internet, the packets could be of variable length.

So, let me just illustrate now that how the leaky bucket regulator or the GCRA (T, tau) traffic descriptors that we have just studied; how it gets translated to the internet world where the packets could be of variable length. And then, we will study some properties of these kinds of a traffic regulator.

So, let me just explain you how we can define deterministic traffic descriptor for the case of a traffic which has variable packet length. Now, in the internet world, this kind of a traffic descriptor is also called as a token bucket regulator. So, let me just see what is the token bucket regulator.

(Refer Slide Time: 24:32)



In token bucket regulator which is very similar to the GCRA (T, tau) that we have considered. Now, here is a bucket which is just similar to the leaky bucket and this bucket however has a depth of sigma and the tokens accumulate into this bucket with a rate of rho. So, this is also called as sigma rho regulator.

Now, the important difference is that a packet of length l ; so this packet may have a length l and it takes l tokens from the bucket. So, the difference here is that unlike in the GCRA (T, τ) where what we are assuming is that all packets were of the constant length, fixed length, they were having the same size and therefore they were taking away the fixed amount of fluid that is capital T from our GCRA (T, τ).

Now here, what we are saying that the tokens are getting accumulated into the token bucket at a rate of ρ tokens per unit of time and when a packet comes and if the packet is of length l ; it takes away l tokens from the bucket or essentially the bucket is decremented by l tokens.

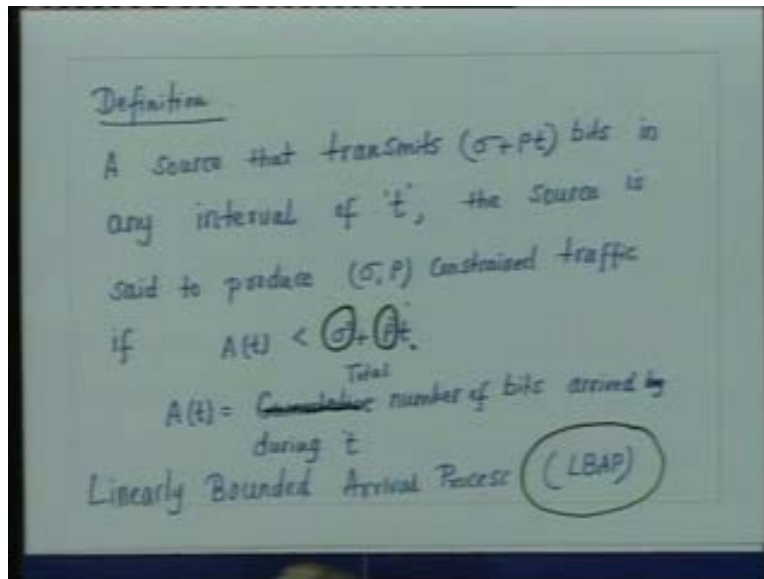
More formally, we can define as that let us say that a_k is the arrival epoch of k^{th} packet and let us say l_k is the packet length of k^{th} packet. Then, we say that the sequence, the arrival process is arrival process $A(t)$ is ρ sigma conferment, ρ sigma regulated or constrained. If initially n_0 is sigma, so we assume initially the bucket is full and n_k is given by minimum of sigma n_k minus l plus a_k minus a_{k-1} into row minus l_k **minus l_k** . This is greater than or equal to 0, for all k .

So, this arrival process then would be called as the sigma rho constrained if this relationship whole would... Essentially n_k , what does n_k says? n_k says, now a_k is the arrival epoch of the k^{th} packet, a_{k-1} is arrival epoch of the $k-1^{\text{th}}$ packet. Now, this is the time between the two packets. Now, between this time, so many tokens would have come. Since, l_k is the length of the k^{th} packet, so many tokens would be decremented and minimum of these would be the fluid that was there when the $k-1^{\text{th}}$ packet left. If you add that, that should be greater than or equal to 0.

Note that therefore, n_k is what? n_k is the number of tokens left in the bucket after the k^{th} packet has left because these many tokens have come between these two intervals, **these many tokens have come**. l_k is the number of tokens that would be decremented and this was a initial token grant at $k-1^{\text{th}}$ instance. So, this represents the number of tokens that would be left after the k^{th} packet has left out of the network. This should be greater than or equal to 0.

Now, two particular cases arise of this that packets are of constant length, **packets could be of constant length**, in that case ρ and sigma; they could be in units of packets and packets per unit time. Essentially, this will degenerate to the case of our GCRA (T, τ) regulator. Another approximation could be fluid approximations, where the input arrival process is basically a fluid process and similarly the bucket also has a fluid and in that case, the bucket depth could be a real number.

(Refer Slide Time: 30:32)



Now, we will concentrate on this rho sigma regulated traffic. Now, before we answer this question, we provide some definition of rho sigma constraint traffic and we say... So, let us define rho sigma constraint traffic. Let us say, a source that transmits sigma plus rho t bits in any interval, in any interval of t, for any possible value of t; then the source is said to produce sigma rho constrained traffic if $A(t)$ is less than sigma plus rho t, where $A(t)$ is the cumulative number of bits arrived by time arrived during the interval of t, so arrived during t. We should say **it should say** total number of bits, cumulative. Total number of bits that are arrived during the t is bounded by sigma plus rho t.

So essentially, what we are saying is that this is a traffic source which is generating the number of bits in such a manner that these number of bits are bounded by sigma plus rho t for any interval t. We also call such arrival processes to be linearly bounded arrival processes. They are also called as linearly bounded arrival processes, **linearly bounded arrival processes** or what we can call it to be LBAP this is an...

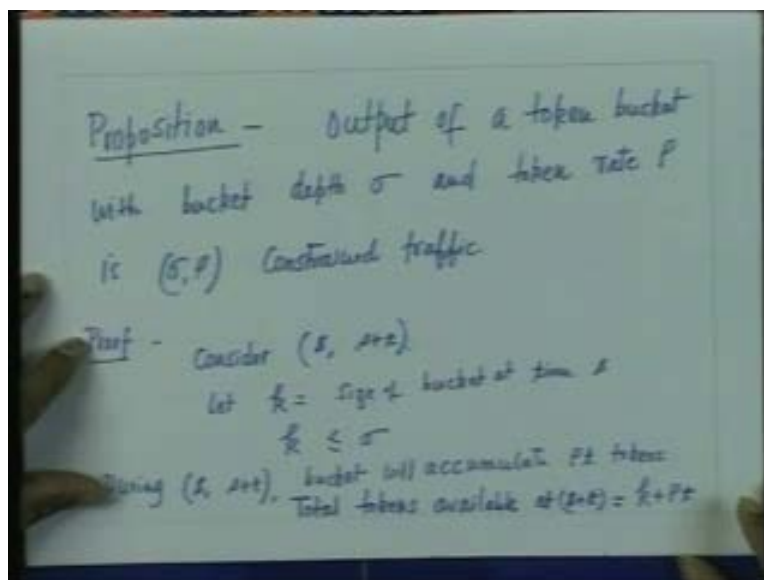
Now, let us look at what is the interpretation of sigma and rho. So, if you consider that this interval t is very small, **this interval t is very small**, then if you look at this equation, then if this interval t is very small; then we can say that the total number of bits that are arrived during time interval t is approximately equal to sigma. So, what is the interpretation of sigma? During a very small interval of time t, the number of bits that can be transmitted is bounded by sigma. So therefore, what is the significance of sigma?

Sigma represents some kind of a burst. This is the maximum number of bits that sources allowed to transmit at a particular instant during a very small infinitesimal interval of time. On the other hand, if you say that this interval of time t is very large **if this interval of time t is very large**; then if you consider this quantity - sigma plus rho t, then here this quantity will be dominated by rho t.

In that case, we can say that this ρ will be equal to the number of bits that have arrived during t divided by the interval. So therefore, what is the interpretation of ρ ? ρ is some kind of an average rate. So therefore, this admits a notion that the traffic source, the linearly bounded arrival process is such that the maximum number of bits it can generate at an instant is bounded by σ and its long time average rate is going to be ρ . So, we call these processes to be linearly bounded because they are bounded by a linear envelope by a straight line of σ plus ρt .

Now, the question is how do we generate such a ρ σ constraint traffic? It turns out that the output of a leaky bucket... So, we just hold this proposition, simple proposition that output of a token bucket with bucket depth σ and token rate ρ , token rate ρ is ρ σ constrained traffic.

(Refer Slide Time: 35:59)



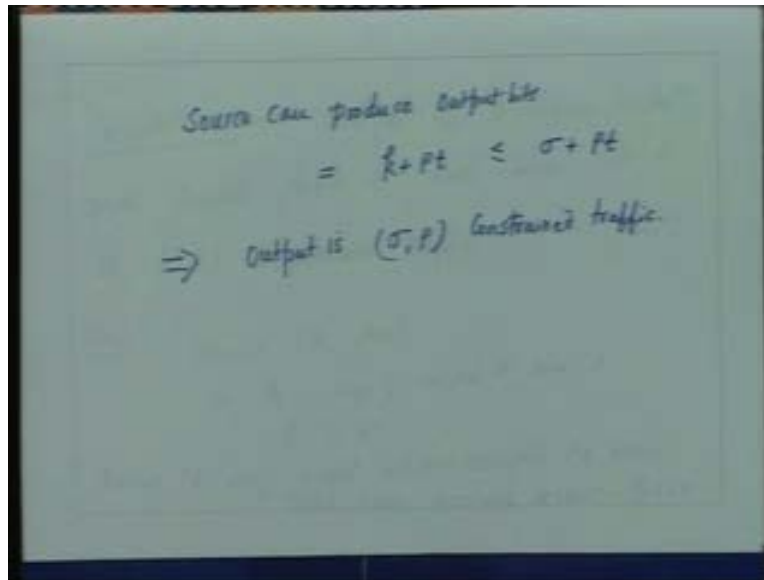
So, what we are saying is that here is a token bucket where our bucket depth is σ , the token rate is ρ ; whenever a packet comes, it takes away the number of tokens which is equal to the length of the packet. We are saying that this output of this traffic is given by this expression where the number of bits will be limited or bounded by σ plus ρt . So, the proof is very simple.

Consider an interval s and s plus t . Now, let k is the size of the bucket at time s . Now, obviously this k is less, will be less than or equal to σ because σ is the maximum bucket depth. So, we consider an interval of s to s plus t . So that total interval if of length t . Now, at time s assume that the bucket depth is k . This k is obviously less than or equal to σ because σ is the total bucket depth. Now, during this during this interval from s to s plus t , the bucket the bucket will accumulate the bucket will accumulate ρ into t tokens. During this interval, the bucket accumulates ρ into t tokens.

So therefore, a total token available at s plus t is how much? The total tokens available at s plus t

will be k plus ρt . Now, these are the total tokens available and therefore during this interval, the source can produce an output bits.

(Refer Slide Time: 39:30)



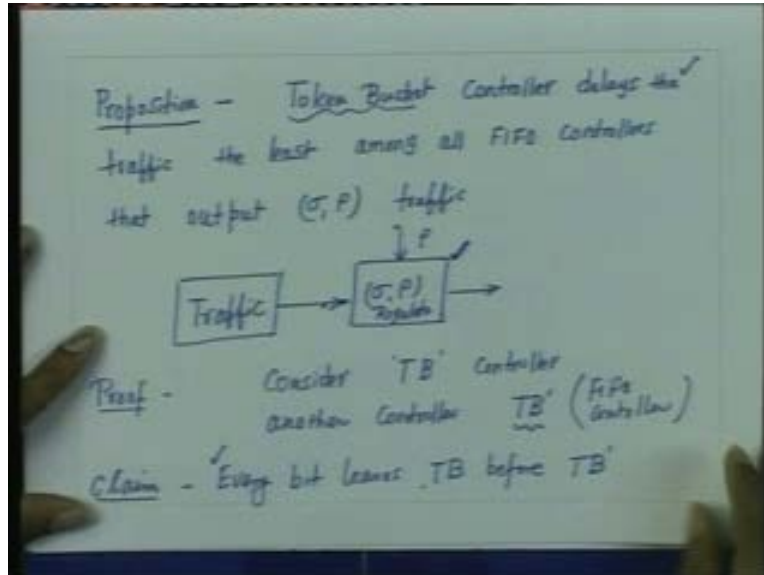
Source can produce output bits
 $= \rho t \leq \sigma + \rho t$
 \Rightarrow Output is (σ, ρ) constrained traffic.

So therefore, during this interval, the source can produce output bits which is equal to k plus ρt and this will be less than or equal to σ plus ρt and therefore we can say that the output is σ ρ constrained traffic. So, this proves our proposition that a linearly bounded arrival processes, linearly bounded arrival process of σ plus ρt can be generated by a token bucket regulator with a bucket depth of σ and token generation rate of ρ .

A packet is transmitted out of the token bucket regulator by taking away the amount of tokens which is equal to the packet length. There could be other regulators also. The question is there could be other regulators also which may also generate this ρ σ constrained traffic which may also generate the linearly bounded arrival processes. But interestingly, it turns out that the leaky bucket is the best among all FIFO controllers that will generate the ρ σ constrained traffic.

Just see, how it happens. So, we will try to see that what is the advantage of having a leaky bucket regulator as compared to other FIFO controllers which may also generate the leaky bucket regulator or ρ σ constrained traffic.

(Refer Slide Time: 41:32)



So, our claim is... So, I just write another proposition. So, our claim is that token bucket regulator, token bucket or a leaky bucket controller delays, delays the traffic the least among all FIFO controllers that will output σ traffic. Now, where is this delay? The delay actually is that here is a traffic source; now this traffic source, in front of this you have put a σ regulator. Now, this σ regulator will output this traffic in such a manner that this output is ρ constrained traffic.

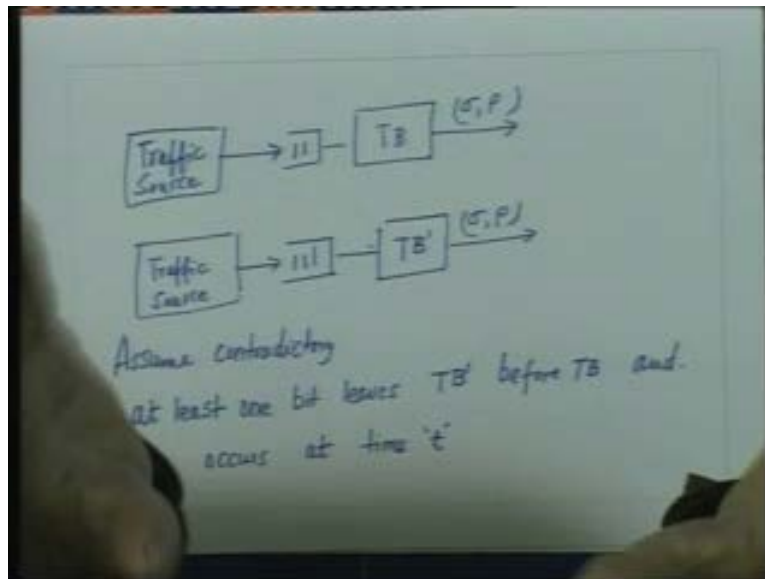
Now, this could be a token bucket regulator. If it is a token bucket regulator, we have a token just generating at a rate of ρ and there is a bucket inside which has a maximum depth of σ . Now, if a packet comes and if it finds that there are not enough tokens in the bucket, then the packet has to wait; very similar to the case we have considered in the ATM network, where a cell has to wait in the cell buffer. So, here also the packet has to wait at the packet buffer. What we are trying to claim here is that the token bucket regulator will delay the traffic the least among all FIFO controllers that will output a ρ constrained traffic.

Now, how does it happen? Let us just let us try to see a proof. Now, consider a token bucket regulator, consider a token bucket controller. So, we denote it by TB and another controller let us call it TB prime. Now, this is not a token bucket regulator, this is another FIFO controller. So, this is you know, another FIFO controller which is also generating a ρ constrained traffic.

Now, our claim is that every bit leaves token bucket regulator because we **have** want to prove that the token bucket controller delays traffic the least. So, our claim is that every bit leaves token bucket regulator before the FIFO controller which is also a TB prime. So, this is what our claim is. Note that we want to prove that ρ regulator when implemented in the form a token bucket will delay the traffic the least. The other FIFO controller may delay it more. So, when we say that it delays the traffic least what we are trying to prove is that every bit will leave

the token bucket regulator first, before another token bucket, another token bucket regulator.

(Refer Slide Time: 45:36)

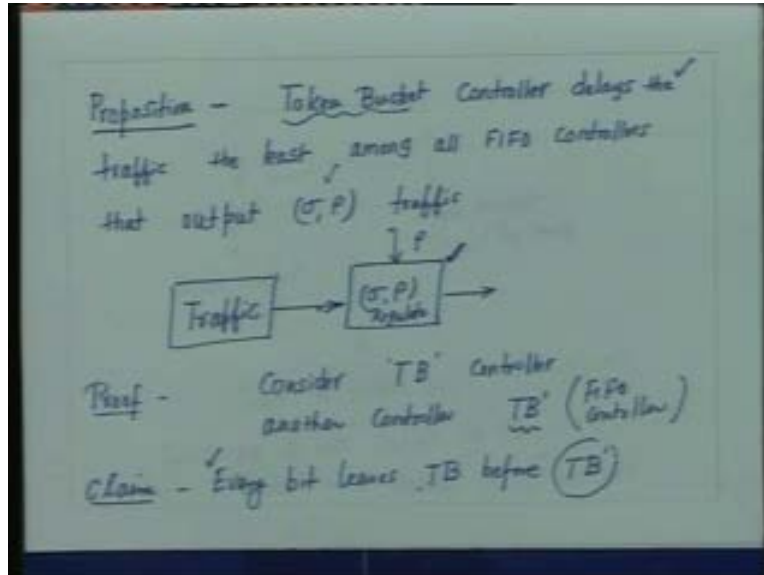


So, the scenarios are like you know, here is a traffic source and here is a buffer, packet buffer and then it is passing through token bucket regulator and the output is ρ sigma. Similar situation is that there is another traffic source and then it is also passing through a buffer and this time, there is a another FIFO controller TB prime and this output is also ρ sigma. But since we are saying that this delay here in this bucket is less, the bit leaves here earlier. So, this is our claim. So, we assume that the contradictory, so we assume the disclaimer is not true. So, the claim is that every bit leaves token bucket, before token bucket prime - that is another controller.

So assume contradictory, assume the contradiction. So, the contradiction is that to assume this let us assume that at least 1 bucket, at least 1 bit leaves this token bucket, another FIFO controller TB prime, before the our token bucket controller and let us say that this occurs for the first time, **at this occurs** at time t for the first time. So at time t , what we are saying... So, the picture is something like this that we have the traffic source and this traffic source is being regulated by a token bucket regulator.

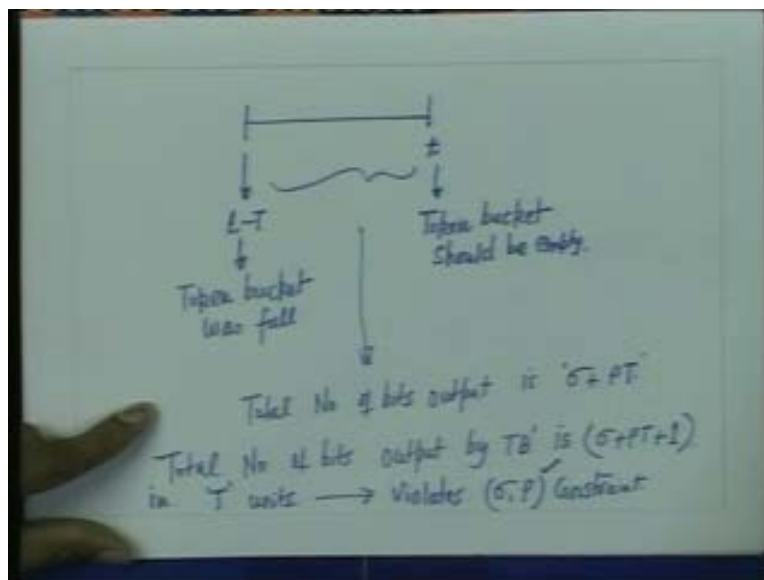
Same traffic source let us say in another situation **is all** is regulated by another FIFO controller. Both of them are generating ρ sigma constrained traffic only. Now assume that 1 bit leaves our another FIFO controller earlier then it is left in the token bucket regulator and let us say that this happens for the first time at time t .

(Refer Slide Time: 47:51)



Now at time t , if this happens; then at time t , the token bucket should have been empty because since it is empty, it is not able to leave. So at time t , the token bucket should be empty. Now, let us assume that the last time when the token bucket was empty is t minus capital T .

(Refer Slide Time: 48:32)



So, this was the last time when the bucket was full **sorry**, when the bucket was full. So, what we are saying is that if a bit is leaving another FIFO controller earlier for the first time at t and this bit, same bit has not left the token bucket regulator, that means the token bucket is empty at a time t .

So, let us assume at time t minus capital T , the token bucket is full. So, during this interval, **during this interval**, the total number of bits that would have been outputted, total number of bits which is output is $\sigma + \rho t$. Now, the total number of bits that would be outputted by the other FIFO controller will be $\sigma + \rho t - 1$, plus one more bit because that is the bit that is going at time t . So, the total number of bits output by the token bucket regulator TB prime is $\sigma + \rho t + 1$ in t units of time. But then, if the other FIFO controller is outputting $\sigma + \rho t + 1$ - this many bits in t seconds; obviously, this traffic is not ρ σ regulated.

Now, this violates ρ σ constrained. So, if it violates ρ σ constraint, then our claim, our contradictory claim that at least 1 bit leaves token bucket from **before token bucket TB prime** before token bucket TB; this is not true. So, this proves our earlier claim that we had made, we have proved it by contradiction that every bit leaves token bucket earlier than any other FIFO controller which is also trying to ensure that the traffic is ρ σ constrained traffic.

So in other words, to summarize what we are saying is that we would like to consider a deterministically bounded traffic source and this traffic arrival process is called a linearly bounded arrival process and this linearly bounded arrival process is such that the number of bits transmitted during an interval of time t is bounded by $\sigma + \rho t$, where σ admits an interpretation of a burst and the ρ admits an interpretation of a long term average rate.

We see that such a linearly bounded arrival process can be generated by a token bucket regulator with the bucket depth of σ , maximum bucket depth of σ and a token generation rate of ρ . This ρ σ traffic can also be generated by some other FIFO controller. However, it is better to use the token bucket regulator because the token bucket regulator is the best among all FIFO controllers that will generate a ρ σ traffic, in the sense that it delays the traffic the least at the source side among all FIFO controllers that are going to generate ρ σ constrained traffic.

Now, we will see in our next lectures some discussions on this ρ σ constrained traffic and how by constraining the traffic to ρ σ ; we can ensure the quality of service guarantees in a packet switched node.

(Refer Slide Time: 53:16)

REFERENCES

1. J. Walrand & P. Varaiya,
'High Performance
Communications Network
Morgan Kaufman , 2000
2. S. Keshav; ' An engineering
approach to computer Networking'
Addison Wesley 1997