

Broadband Networks
Prof. Karandikar
Electrical Engineering Department
Indian Institute of Technology, Bombay

Lecture –7

So, in the previous lectures we were discussing about the token bucket regulators or the rho sigma characterization of the traffic source. So, we will continue the traffic characterization of a traffic source by the rho sigma parameters and we will see some of the properties of such linearly bounded arrival processors.

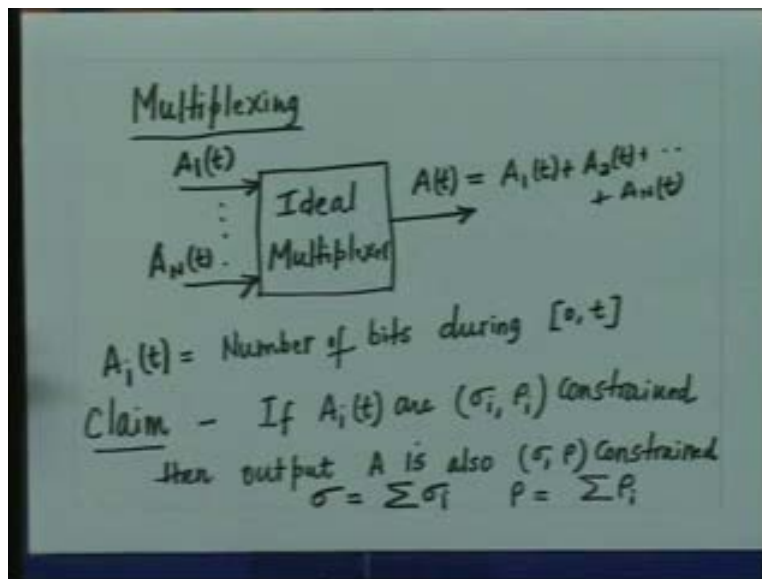
Just to recapitulate what we have discussed in the previous lecture is that what we are saying is that we will represent a traffic source by linearly bounded arrival process or by deterministic traffic descriptors which is sigma and rho and these sigma and rho parameters are such that there are number of bits transmitted by a source during any interval of time t is bounded by sigma plus rho t . So, therefore rho admits the interpretation of an average rate and sigma admits the interpretation of some sort of burst that a source can transmit.

We also saw that such a linearly bounded arrival process can be generated by a token bucket regulator which has a maximum bucket depth of sigma and which has the token generation rate of rho and whenever a packet comes the bucket is decremented by the length of the packet and if a packet comes and if it finds that there are not enough tokens available in the packet; then the packet is queued at the packet buffer at the traffic source site or at the traffic regulators.

We also saw that among all the FIFO controllers that will generate the rho sigma traffic; token bucket regulator delays the traffic the least among all these FIFO controllers and therefore the token bucket regulator is some kind of an optimal traffic regulator for a rho sigma traffic characterizations. So, now we will continue this discussion further of the rho sigma traffic characterizations and study some of the properties of this rho sigma regulated traffics.

So, let us start with first the effect of multiplexing.

(Refer Slide Time: 3:07)



So, multiplexing of several rho sigma traffics: so what we are saying? This is an ideal multiplexer, so we call it to be an ideal multiplexer. The input traffics are represented by $A_1(t)$ $A_2(t)$ so on till $A_N(t)$ and the output is $A(t)$. Now, this is an ideal multiplexer; so therefore, the output $A(t)$ is given by $A_1(t)$ plus $A_2(t)$ plus so on $A_N(t)$.

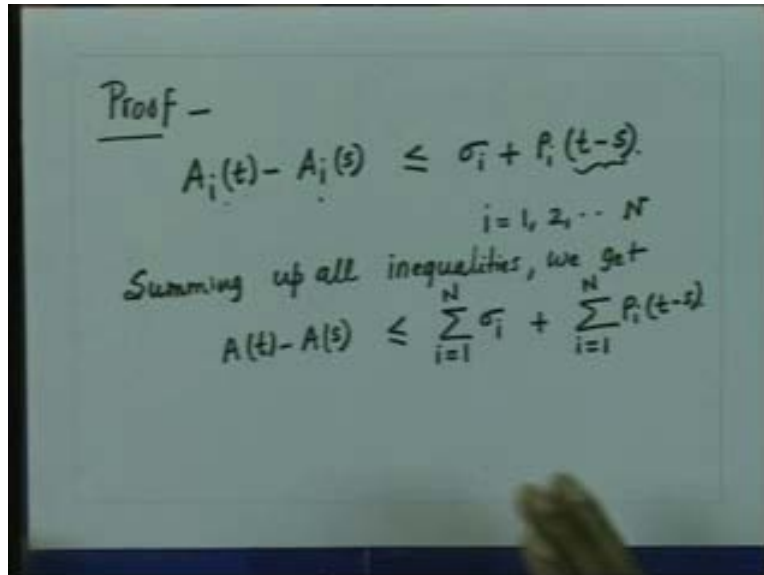
Now here, $A_i(t)$ denotes the number of bits transmitted or the number of arrivals transmitted by the traffic source during the interval of 0 to t or in other words we can say that these are the number of bits or cumulative number of arrivals by the time t. So, this also represents, $A_i(t)$ also represents the cumulative number of arrivals of a traffic source i.

Now, an ideal multiplexer; if you want to characterize an ideal multiplexer, then the input to such multiplexers is rho sigma regulated. An ideal multiplexer gives an output which is the sum of all these traffic sources. So, **our claim is** so we say our claim is that if all these $A_i(t)$ are rho sigma constrained that is σ_i rho ρ_i constrained; then the output A is also sigma rho constrained, where what is sigma? Sigma is the summation of all sigma i's and what is rho? Rho is the summation of all rho i's.

So, we will try to prove this and what we are saying is that suppose if we take an ideal multiplexer, the input to such ideal multiplexers are rho sigma constrained traffic. We are trying to prove that the output will also be a rho sigma constrained traffic, where the sigma will be the sum of the sigmas of all the input traffics and rho will be the sum of the rhos of all the traffics.

So, let us try to prove this proposition and then we will see a practical example of that.

(Refer Slide Time: 6:41)



The image shows a whiteboard with handwritten mathematical text. At the top left, it says "Proof -". Below that is the inequality $A_i(t) - A_i(s) \leq \sigma_i + \rho_i(t-s)$. Underneath this, it says "i = 1, 2, ... N". Then, it says "Summing up all inequalities, we get" followed by the inequality $A(t) - A(s) \leq \sum_{i=1}^N \sigma_i + \sum_{i=1}^N \rho_i(t-s)$.

So, the proof of this is; note that the traffic i is rho sigma regulated, therefore $A_i(t)$ minus $A_i(s)$ - the $A_i(t)$ is the number of arrivals by the time t and $A_i(s)$ is the number of arrivals by the times s . Therefore, $A_i(t)$ minus $A_i(s)$ is the number of arrivals or number of bits transmitted during the interval of t to s .

Now, since this traffic is rho sigma regulated, this will be less than or equal to sigma i plus rho i into t minus s . So, this is the interval; note in our previous discussion, we were using the t for the interval, for the time interval. Here now, we are using the variable t to denote a time instant; s is another time instant. So, $A_i(s)$ is the number of bits that have occurred by the time s , $A_i(t)$ is the number of bits that have occurred by the time t . So, $A_i(t)$ minus $A_i(s)$ is the number of bits that have occurred during an interval of now t minus s ; t minus s is the interval.

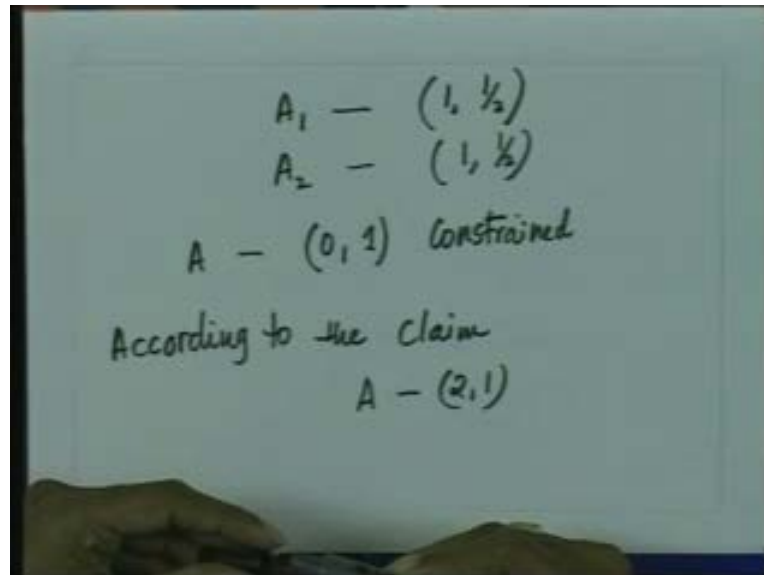
Now, since this traffic is rho sigma regulated, therefore these number of bits which have been transmitted during the interval of t minus s that is from the time instant s to time instant t will be less than or equal to sigma plus rho into t minus s . And now, this is valid for all i 's. So, I will say that this will be valid for all i 's from 1 to N .

Now, we sum up all such inequalities. So, if we are summing up all such inequalities for i equal to 1 to N ; then we get $A(t)$ minus $A(s)$ that is the number of bits which are transmitted during an interval t to s at the output of the multiplexer will be less than or equal to summation of sigma i 's for i is equal to 1 to N plus summation of i equal to 1 to N rho i (t minus s). And hence, we have proved that the multiplexer output is a rho sigma regulated traffic where the sigma is the sum of all sigma's and rho is a sum of all rho's.

Now, it can be seen some times that this multiplexer can reduce the burstiness. Let me just give you an example of this multiplexer. Now, we have this multiplexer which is an ideal multiplexer; let us say, it has 2 inputs - A_1 and A_2 and this is the output. Let us say A_1 is 101010 like this.

So, these are the number of arrivals that are occurring at these time instants. So, this is how the traffic A_1 is; the number of packets or number of bits which are generated. A_2 is 010101 so on. Now, if it is ideal multiplexers, output A will be the sum of these two and we get 1111 so on. Now, what is the input A_1 if we look at this input A_1 ?

(Refer Slide Time: 11:05)



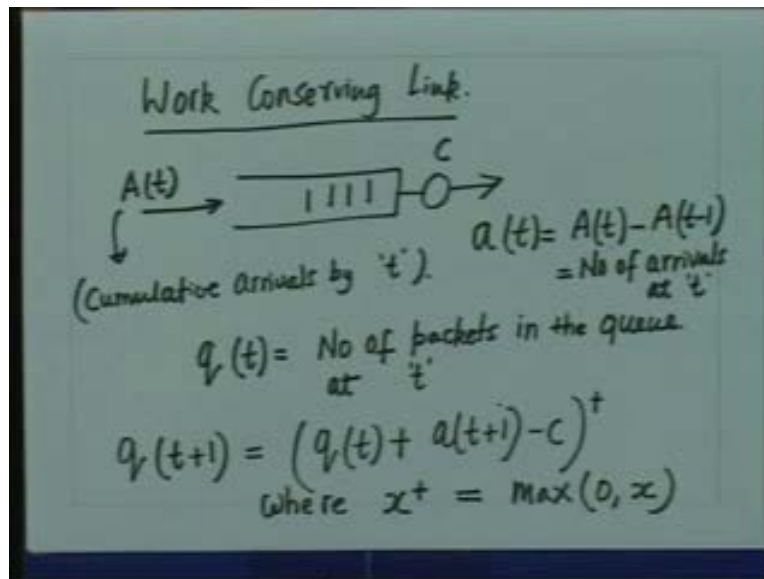
It is, the input A_1 is actually 1 into half constrained traffic. Sigma here is 1 because we can maximum burst here is 1 unit because the traffic is transmitting as 101 so on. Similarly, A_2 is 0101; so it is also 1 and half constrained. So that is the average rate is of half; so, this is also 1 into half constrained.

But if you see this input traffic which is the output 111; so, how it is constrained, 111? Output is actually is 01 constrained. So, the output if you see, it is 01 constrained. Now if you had applied this theorem, the previous theorem; then the output should have been sum of sigma i's which would have been 2 and sum of rho i's which would have been 1. So, output should have been (2,1)constrained according to this theorem. According to the theorem, output should have been, according to our claim, output A should have been 21 constrained.

So, indeed what we see that this theorem which says that the output of an ideal multiplexer which will be the ideal multiplexer, where the inputs are all sigma rho constrained; the output will also be sigma rho constrained where the sigma will be the sum of all the sigma's and rho will be the sum of all the rhos. This result in fact, gives us the worst case bound and it may be possible to reduce the burstiness where the multiplexing that we just saw in this example that the input one input was 1 and half constrained and other input was also 1 and half constrained and if you add them, you get an output which is actually (0,1) constrained which is actually less bustier than what was the 2, what were the 2 inputs. So, this is one aspect of what happens when these sigma rho traffic gets multiplexed.

Now, let us consider that this sigma rho traffic is served by work conserving link which we have been all through considering which is basically a buffer which is equipped with a transmitter which transmits at the rate of c cells per unit of time; so, we call this to be a work conserving link. So, the server is never ideal whenever there is some traffic or whenever there are some bits in the buffer. So, let us define this work conserving link.

(Refer Slide Time: 14:23)



Now, in this work conserving link, what we are saying that there is a buffer just transmitting at c cells per unit of time. We would like to characterize this work conserving link later on for the rho sigma characterized traffic. But, so what we are saying is right now is $A(t)$ is the input traffic where $A(t)$ denotes the cumulative number of arrivals by the time t ; so, i want to again reemphasize this that cumulative arrival by time t . For this discussion, we may assume that the arrival occurs at discrete time instant, let us say $0, 1, 2, 3, 4$ so on. So, the arrival process $A(t)$ is essentially a stair case function.

So, for this work conserving link, it is easy to see that if $q(t)$ denotes the number of packets in the queue at time t , so if it is number of packets in the queue at time t , so this is essentially a buffer length; then, it can be shown that $q(t)$ plus 1 is given by $q(t)$ plus $a(t)$ plus 1 minus c plus of this, where this is equal to... where x plus - I am using it to denote - **x plus** is max of 0 into x ; so, maximum of this.

Now, what does it say? It says that the number of packets in the q at time t plus 1 is equal to max of either 0 or $q(t)$ plus $a(t)$ plus 1 minus c . Now, what is $a(t)$ plus 1? We were saying that $a(t)$ is given by $A(t)$ minus $A(t-1)$. So, we are assuming discrete time arrival process and $A(t)$ denotes the cumulative number of arrivals by the time t . So, what it is saying is that **$a(t)$ capital** $A(t)$ minus $A(t-1)$; so, this actually is the number of arrival. Small $a(t)$ denotes the number of arrivals at time t .

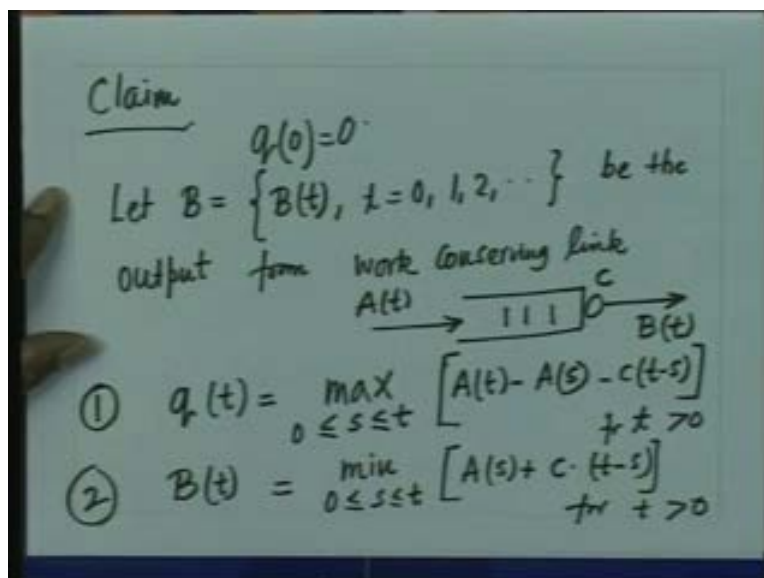
So, now what this equation is trying to say is this equation is trying to say that the q length at time $t + 1$ is given by the q length at time t plus the number of arrivals that occurs at time $t + 1$ minus the number of packets that would have departed at the time $t + 1$. So, this was the previous q length, the q length at time t . These were the number of arrivals that have occurred at $t + 1$, these are the number of arrivals that have departed; so, this becomes the new q length. However, if this is less than c if this is less than c ; then obviously, the q becomes empty at $t + 1$ and the q length at $t + 1$ will be 0.

So, if $q(t) + a(t) + 1$ is greater than c , then this will be equal to $q(t) + a(t) + 1$ minus c and if $q(t) + a(t) + 1$ is less than c , then the q length at time $t + 1$ will be 0. So, this is a simple interpretation of the fact I just explained you here because in our next subsequent discussions, we would like to consider that there is a FIFO scheduler for the time being not necessarily a FIFO scheduler but a single buffer, there is a single buffer equipped with a transmitter which is transmitting at c cells per unit of time and there is an input arrival process which is $a(t)$.

We had considered these processors, we had considered these examples earlier also where we had assumed that this $a(t)$ was a bursty random traffic and it was being served by a transmitter which was transmitting at c cells per unit of time and we characterized the effective bandwidth of a traffic source etcetera previously. But right now we want to consider what we should be calling appropriately as deterministic queuing system. Why deterministic? Because the arrival process is actually deterministic which is a linearly bounded arrival process; instead of a stochastic process, instead of a random bursty process as an input to this queuing system, it is a deterministic arrival process or other deterministically bounded arrival process and such a queuing system, we would like to characterize.

So, now we will prove a very important result for this work conserving link which is transmitting at c packets or c bits per unit of time and that important result I would just like to prove.

(Refer Slide Time: 20:16)



Let us say that I call it to be another claim is, let us say that $q(0)$ is 0 and let us say that these B which is $B(t)$ for t equal to 0, 1, 2; let us say that this denotes the output from the this work conserving link. So, here was our conserving link, here was the input $A(t)$ and let us say that $B(t)$ is the output.

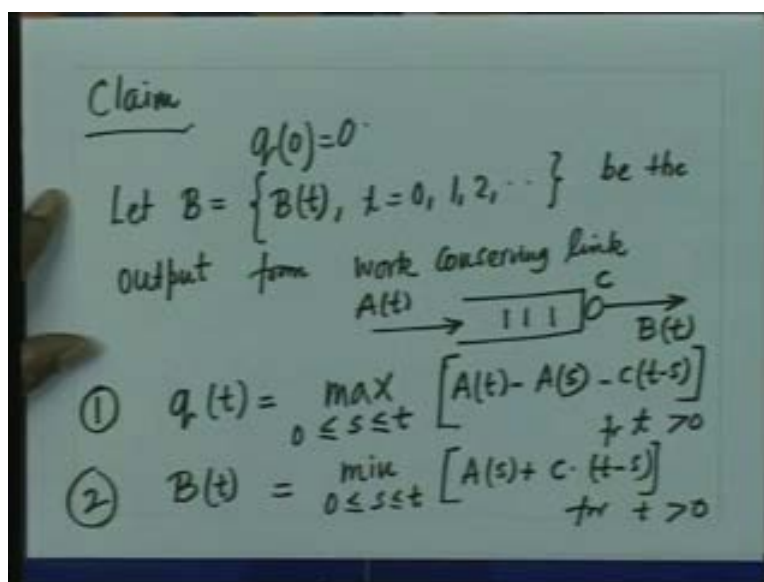
Now, you have to prove 2 results: one result is we want to prove that $q(t)$ is given by maximum over 0 to t and which is maximized by $s A(t) - A(s) - c$ into $t - s$ for all t greater than 0 and the other result that we would like to prove is $B(t)$ which is for the output is **minimum of 0 to** minimum over 0 to $t A(s) + c$ into $t - s$, again for all t greater than 0. So, we would like to prove this result.

Right now, we are not assuming the input traffic $A(t)$ to be a sigma rho constrained traffic but we will soon assume it to be a sigma rho constrained traffic and we will specifically see how the queue length, how the queue length and the output departure process behaves for a deterministic queuing system. That is what we would like to characterize; how the queue length evolves and how the output departure process evolves.

Note that asymptotic results for a random arrival process for such queuing system we had considered and where we had seen that the packet loss rate basically behaves or decays exponentially and that is how we characterize the effective bandwidth of a random traffic source. Now, we would like to consider the deterministically bounded arrival process and we would like to characterize various queuing parameters which is like the queue length and which is like output departure process; how do they behave?

But first consider general result where we have shown that $q(t)$ and $b(t)$ we would like to characterize. So, we would like to prove these 2 results.

(Refer Slide Time: 23:31)

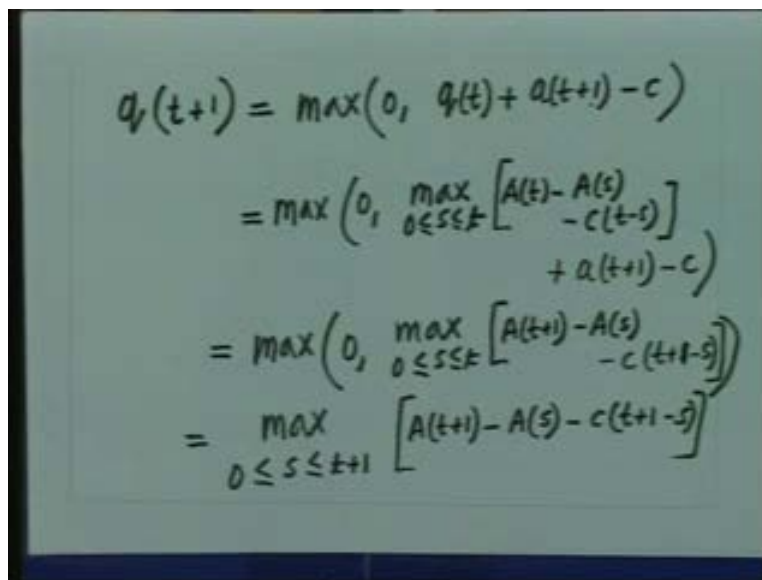


Let us first prove the result which is the first result and this first result is we would like to prove that $q(t)$ is given by max of from 0 to t $A(t) - A(s) - c(t-s)$; this result we would like to prove. So, we prove this result by induction. So, let's prove this by induction.

So, let us see what happens for t is equal to 1. For t equal to 1, this result is trivially true because we know the $q(1)$ which is from the Lindley's equations is given by max of 0, $q(0)$ which was a queue length at time 0 plus $a(1)$ - number of arrivals at 1, minus c that is the number of departures that have occurred. Now, note that this is 0. So therefore, $q(0)$ is 0; so this is 0, $a(1)$ minus c and this is nothing but we can say that this is maximizing over 0 to 1 over an interval of 0 to 1 $a(1) - A(1) - c(1-s)$. So, this result is a trivially true for t equal to 1.

Now, let us assume that this result is true for t and then we will prove that this result will also be true for t plus 1. So, this we have, we would have proved this by induction. So, let us assume that this result is true for t .

(Refer Slide Time: 25:33)



$$\begin{aligned}
 q(t+1) &= \max(0, q(t) + a(t+1) - c) \\
 &= \max\left(0, \max_{0 \leq s \leq t} \left[A(t) - A(s) - c(t-s) \right] + a(t+1) - c\right) \\
 &= \max\left(0, \max_{0 \leq s \leq t} \left[A(t+1) - A(s) - c(t+1-s) \right]\right) \\
 &= \max_{0 \leq s \leq t+1} \left[A(t+1) - A(s) - c(t+1-s) \right]
 \end{aligned}$$

So, if this result is true for t , then we know that we can write $q(t)$ plus 1 again from a Lindley's equations. It will be given by max of 0 $q(t)$ plus $a(t)$ plus 1 minus c . Now since, this result is true for t , then we can write this to be max of 0 into maximum over 0 to t given by $A(t) - A(s) - c(t-s)$ plus this term - $a(t)$ plus 1 minus c . Now, this I can write to be max of 0.

Note that this maximization is over 0 to t , so I can take this $a(t)$ plus 1 inside and if I take $a(t)$ plus 1 inside, then this becomes 0 to s and similarly I can take c also inside because this is independent of s . So, this will not affect our maximizations.

So, if I take that inside, then this will be $A(t)$ plus 1 minus $A(s) - c(t-s)$ plus 1 minus s and this; again, I can write this to be t plus 1 $A(t)$ plus 1 minus $A(s) - c(t-s)$ plus 1 minus s . So, this proves the result. So, this we have proved the result that if you assume that the queue

length evolution result holds for t ; we have proved that it will also be valid for t plus 1. So, among this, we have proved the first claim that is for the deterministic queuing system; if $A(t)$ is the input and $B(t)$ is the departures that is where $A(t)$ denotes the cumulative number of arrivals by the time t and $B(t)$ denotes the cumulative number of arrivals by the time t , then the q length will evolve as $A(t) - A(s)$ where $A(t)$ is the number of arrivals by the time t and $A(s)$ is the number of arrivals by the time s . And, $A(t) - A(s)$ therefore denotes the number of arrivals during the interval t to s minus t minus s . We maximize it over the interval 0 to t .

(Refer Slide Time: 28:43)

The image shows a handwritten derivation on a whiteboard. It starts with a circled '2' and the equation $B(t) = q(0) + A(t) - q(t)$. Arrows point from $q(0)$ to 0 , from $A(t)$ to $A(t)$, and from $q(t)$ to $q(t)$. The next line is $= A(t) - q(t)$. The third line is $= A(t) - \max_{0 \leq s \leq t} [A(t) - A(s) - c(t-s)]$. The final line is $= \min_{0 \leq s \leq t} [A(s) + c(t-s)]$.

Now, let us prove about the $b(t)$. We try to prove this result about $b(t)$ and so we try to prove the claim 2 which is; note that the departures by the time t is given by the q length which was at time 0 , the initial q length plus $a(t)$ that is the number of arrivals by the time t minus $q(t)$ which is the q length at time instant t .

So, this was the q length at time t , this many number of arrivals have occurred by the time t and this was the initial q length; so obviously, $b(t)$ will be equal to the number of departures that have occurred. Now since $q(0)$ we assumed it to be 0 , this we can also write to be equal to $a(t)$ minus $q(t)$. And, we have already proved this result for $q(t)$, we have already proved this result for $q(t)$; so I will write this result for $q(t)$ here and then I can say that this is also given by $a(t)$ minus max over 0 to t $a(t)$ minus $a(s)$ minus c into t minus s .

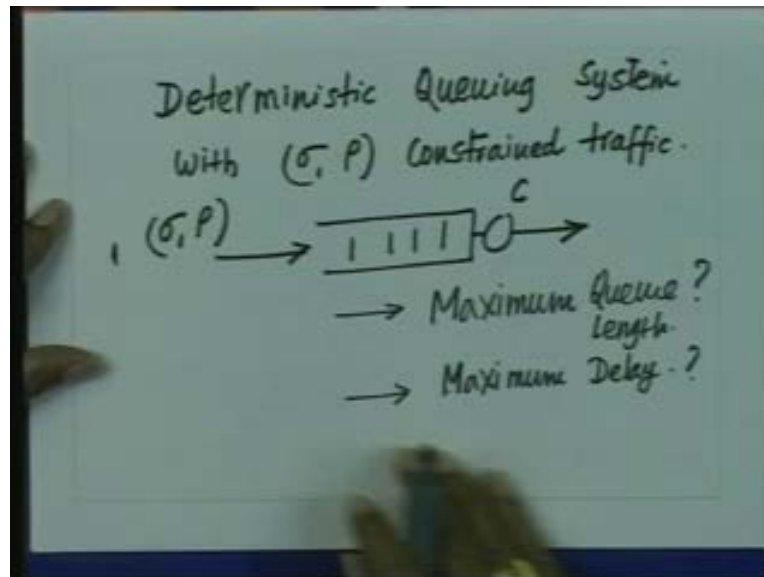
Now, note this is $a(t)$ and this $a(t)$ here; so, this I can take inside and this gets cancelled. But there is a minus sign here, so therefore this is equivalent to saying that we were minimizing it over 0 to s plus t $a(s)$ plus $c(t) - s$ because I take this minus sign out and this will be equivalent to then minimizing over this **0 to capital 0 to capital oh sorry** 0 to t .

So, this way we have proved that how the q length evolves of a deterministic queuing system which has this input arrival process $a(t)$ which is deterministic. Discrete time we assuming for this specific case, discrete time arrival process and the number of departures which is $b(t)$ where

$b(t)$ denotes the cumulative number of departures by the time t , again discrete time and we have proved how the number of departures behave and how the q length evolves.

Now, let us consider the specific cases where this $a(t)$ happens to be a σ ρ constrained traffic and then we would try to characterize this deterministic queuing system for a σ ρ constrained traffic. So, let us see for the case when the traffic is ρ σ constrained. So, we would like to see what happens when the traffic ρ σ constrained traffic.

(Refer Slide Time: 31:49)



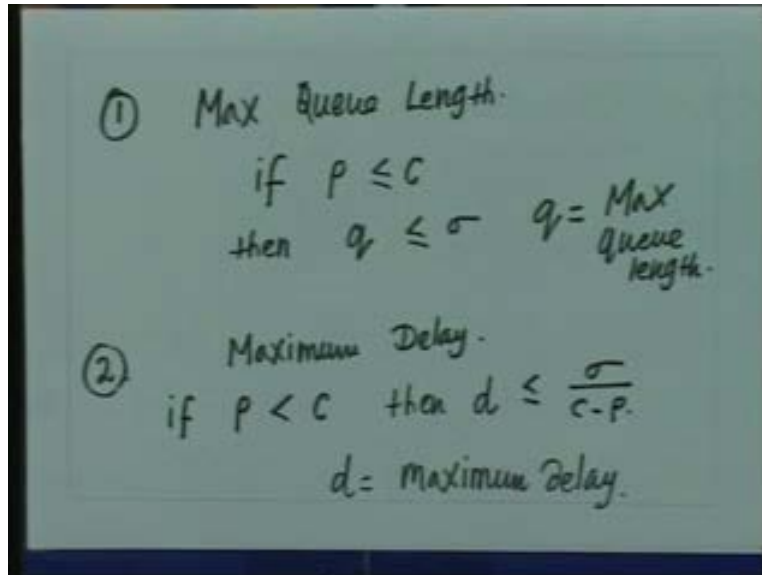
So, I should say here characterization, deterministic queuing system with σ ρ constrained traffic. So, two things we want to prove here: one is so what we were saying essentially is that there is a buffer which is serving at the rate of c , input is now σ ρ constrained traffic. And, we would like to know 2 things: one is we would like to know is what is the maximum queue length and the other thing that we would like to know is what is the maximum delay and queue length. It can be proved; so there two claims which you would want to prove of such a **deterministically arrival at** deterministic arrival process with ρ σ constrained traffic which is being served by a work conserving link with a transmitter rate c and there being a buffer and we would like to know what is the maximum queue length and what is maximum delay.

So, 2 results we can prove here about the maximum queue length and the delay. One is that about max q length. **We can prove that if a ρ is less than or equal to c that means** if the transmitter rate **if the transmitter rate** is more than the average arrival rate what we are assuming because the input has an arrival process which is σ ρ constrained; σ is the maximum burst size and ρ is the average arrival rate.

So, what we are saying is if the input, average arrival rate ρ is less than or equal to the transmitter rate that means the transmitter has a arrival rate more than the **sorry** the transmitter has the output link rate more than the average arrival rate; then we can prove that the maximum

queue length of this queuing system will be bounded by the sigma that is the maximum burst size.

(Refer Slide Time: 34:39)



So, we can prove that the maximum q length will be less than or equal to sigma where q I am saying is to be max q length. So, this is more important result. The second more important result we would like to know about the maximum delay and maximum delay is, for this we need to ... if ρ is less than c , then the maximum delay d will be less than or equal to sigma upon c minus ρ where, so this d is my maximum delay.

Note that in finding out of the maximum delay; we are not assuming that is work conserving link is necessarily a FIFO schedulers. We are assuming that it could be scheduling the packets in some other order also, may be last come first serve also or some other scheduling order also. But the only assumption that we are making is that this link is work conserving. The scheduler is never ideal if there is a traffic in the queue or if there is a packet in the queue, the scheduler is never ideal. So therefore, it is a work conserving link.

We can improve this result if we further put a restriction on the scheduler that it is a first in first out scheduler, a FIFO scheduler or a first come first serve schedulers. So, we will see that result also but for the time being we assume that we are not imposing any restriction on the scheduler as long as we know that it is a work conserving scheduler. So, let us prove the first the result of the maximum queue length. So, to prove this result, so what is the claim?

(Refer Slide Time: 36:55)

① $q \leq \sigma$ if $c \geq \rho$.

$$A(t) - A(s) \leq \sigma + \rho(t-s)$$

for all $0 \leq s \leq t$.

$$q(t) = \max_{0 \leq s \leq t} [A(t) - A(s) - c(t-s)]$$
$$\leq \max_{0 \leq s \leq t} [\rho(t-s) + \sigma - c(t-s)]$$
$$= \sigma + \max_{0 \leq s \leq t} [(\rho - c)(t-s)]$$

The first claim was that this q length is bounded by σ if c is greater than or equal to ρ . So, the proof is since the traffic is ρ σ constrained, $a(t) - a(s)$ is clearly given by σ plus ρ into t minus s for all... Now, we had proved for the deterministic queuing system that the queue length $q(t)$ evolves as max of, we had already proved that it evolves as the max of $a(t) - a(s) - c(t-s)$. This is what the result earlier that we had proved. So, this result is true.

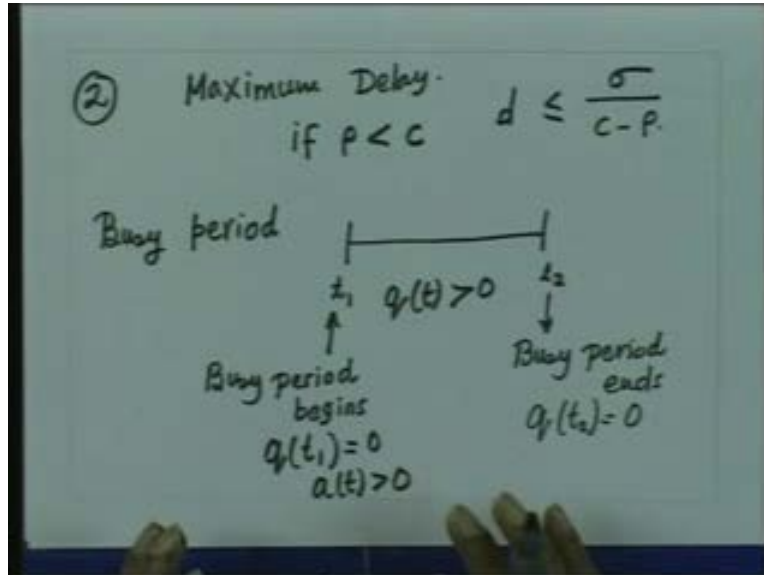
Now, note that $a(t) - a(s)$ in our case is ρ σ constrained traffic. So therefore, we can write that this is equal to max of... and which is therefore should be... this expression is therefore equal to σ plus... now, but ρ is less than or equal to c and therefore we can say that q length, $q(t)$ will always be less than or equal to σ . So therefore, we can say that $q(t)$; since ρ is less than or equal to c , it proves that $q(t)$ is less than or equal to σ . That means max queue length will be equal to σ . So, this is what we have proved from this result.

So, since ρ is less than or equal to c , your $q(t)$ will be less than or equal to σ . So, what we have proved is that if we have a deterministic queuing system where the input is a ρ σ constrained traffic and it is being served by a work conserving link which is transmitting at the rate of c and this transmitter rate we are assuming it to be greater than or equal to the average arrival rate of this deterministically bounded process; then we can prove that the maximum queue length will never be more than σ .

So, this also gives you an idea of buffer dimensioning in a internet router that how much we should keep the size of the buffer if we have actually shaped or regulated the traffic by some kind of a ρ σ regulated traffic.

Now let us prove about the maximum delay. Now, we had shown that the maximum delay will be given by this result; the maximum delay is if ρ is less than c , then the maximum delay is less than or equal to σ upon $c - \rho$. So, let us try to prove this result now.

(Refer Slide Time: 41:02)



Now, I will define the busy period. So, we now proved here the result for the maximum delay that if ρ is a sort of less than c , then you want to prove that the delay d is less than or equal to σ upon c minus ρ . Now, let us say define the busy period; to prove this result we define the busy period from t_1 to t_2 . So, we are saying that a busy period begins at t_1 and at t_2 the busy period ends. Now, the busy period begins at t_1 that means $q(t_1)$ was actually 0, that q length at time t_1 was actually 0 and $a(t)$ the number of arrivals were greater than 0 and busy period ends at t_2 because $q(t_2)$ also becomes equal to 0 and during t_1 to t_2 , $q(t)$ is greater than 0. So, this is the definition of the busy period of this deterministic queuing system which is being served by a ρ σ constrained traffic and which is being transmitted by a scheduler which is transmitting at c per unit of time.

So, now the busy period begins at t_1 and ends at t_2 . At t_1 , the q length is 0; at t_2 also the queue length has become 0, during the interval of t_1 to t_2 the q length is greater than 0. So, there was continues backlog here and all the backlog has been served at t_2 . So, that means by the time t_2 , all the traffic that arrived during the interval of t_1 to t_2 has been served.

So obviously, the maximum delay that a traffic source or a packet will encounter will also be bounded by the maximum length of this busy period. So, if we can determine what is the maximum length of the busy period, we would also have proved what is the maximum delay that a packet would encounter. So, let us prove what is the maximum delay.

(Refer Slide Time: 43:36)

Let T = length of the busy period.

$$T \leq \frac{\sigma}{c - \rho}$$
$$A(t_2) - A(t_1) \leq \sigma + \rho(t_2 - t_1) = \sigma + \rho T.$$

No. of packets transmitted = $c \cdot T$

$$c \cdot T = A(t_2) - A(t_1) \leq \rho T + \sigma$$
$$\Rightarrow T \leq \frac{\sigma}{c - \rho}$$

Let T - capital T with the length of this busy period; so let capital T be this length of the busy period. Then, all we need to show is that the capital T is less than or equal to σ upon c minus ρ . Now, note that the number of packets arrived during the interval t_1 to t_2 that is $A(t_2)$ minus $A(t_1)$ is definitely less than or equal to σ plus ρ into t_2 minus t_1 which is equal to σ plus ρ capital T .

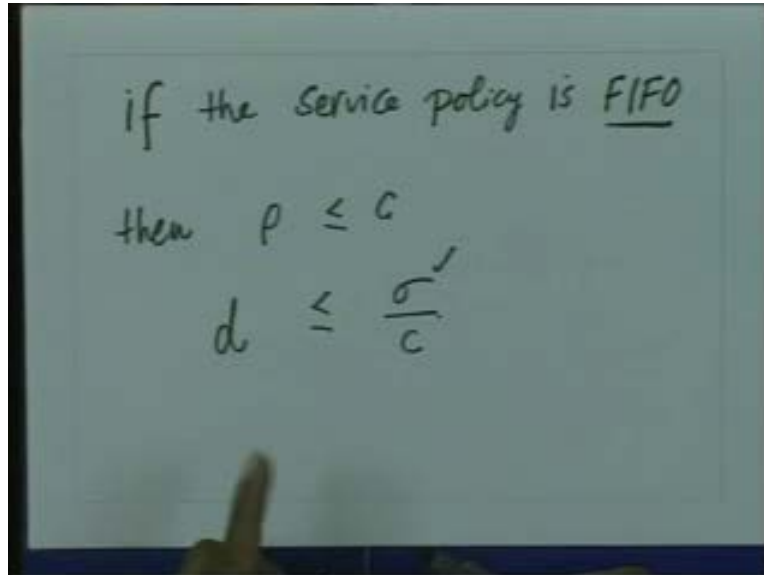
What is the number of packets which are transmitted? Number of packets which are transmitted during the interval is c into capital T . So obviously, c into capital T is equal to $A(t_2)$ minus $A(t_1)$ which is less than or equal to ρ t plus σ and this gives that T is less than or equal to σ upon c minus ρ . Of course, if we are assuming a discrete time arrival processes where the t has to be integer; then we need take the integer part of σ upon c minus ρ .

So, that is what we have proved here that the delay of a packet a in a deterministic queuing system will be bounded by σ upon c minus ρ where **the transmitter's average rate of serving** the transmitters rate of serving is more than the average arrival rate of the input traffic. So, in that case we have shown that the q length is also bounded, the delay is also bounded.

That means we are gravitating towards the fact that if we shape this input traffic or if we regulate this input traffic **by this rho sigma** by this ρ σ characterization that is a token bucket regulator; then downstream network nodes will be able to offer certain quality of service guarantees because in that case the maximum queue length and then the maximum delay can be bounded.

Now, we will improve for the this result by showing that if we assume that this scheduler is actually a FIFO scheduler, not necessarily a last come first serves or any other scheduler; then we can also show that this maximum delay bound can be improved further. So, what is that result?

(Refer Slide Time: 46.44)



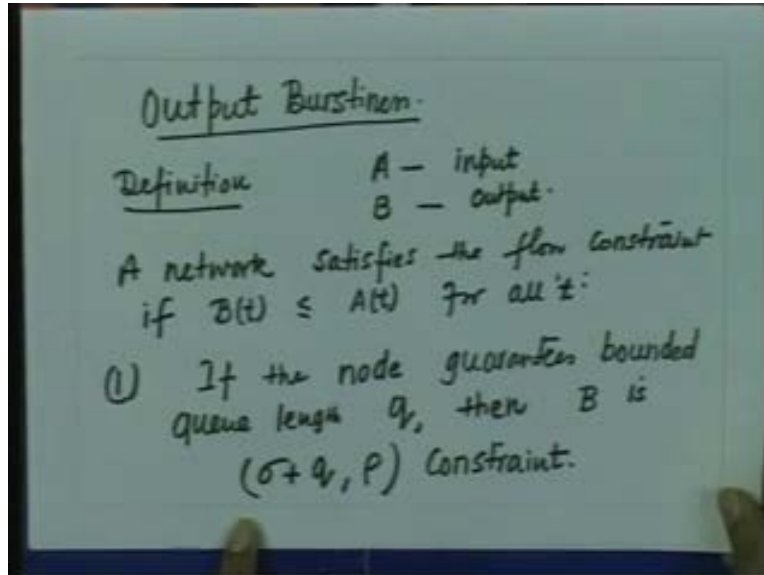
For the FIFO scheduler, we can show that if in this case if the service policy is FIFO that is first in first out and then rho is less than or equal to c. Now, in this case the rho being strictly less than c is not required, it can be less than or equal; then maximum delay will be less than or equal to sigma upon c. This is because the maximum q length happens to be sigma and the packets are being served in the first in first out manner. So obviously, the delay can never be more than the maximum queue length divided by the output link rate of this work conserving scheduler which is given by c. So, this is a result of the scheduling policy happens to be a first come first served.

Now, **we will so we have to** so if the input traffic happens to be a rho sigma characterized traffic; then first of all we have shown that what is the effect of this multiplexing on these input traffic which is rho sigma traffic, so that one part we have seen. The other part that we have seen, what is the effect of if this traffic is served by a work conserving link, then what can we say about the maximum queue length and the maximum delay that also we have seen.

Now, we would like to see that when such an input traffic is served by a work conserving link with a transmitter, then what can we say about the output burstiness; whether output is more bustier and if it is more bustier, then by what amount it is bustier, whether we can reduce the effects of the burrstones and so on and that is also we would like to see.

So now, let us characterize the output burstiness of such a work conserving link whose input is rho sigma constrained traffic.

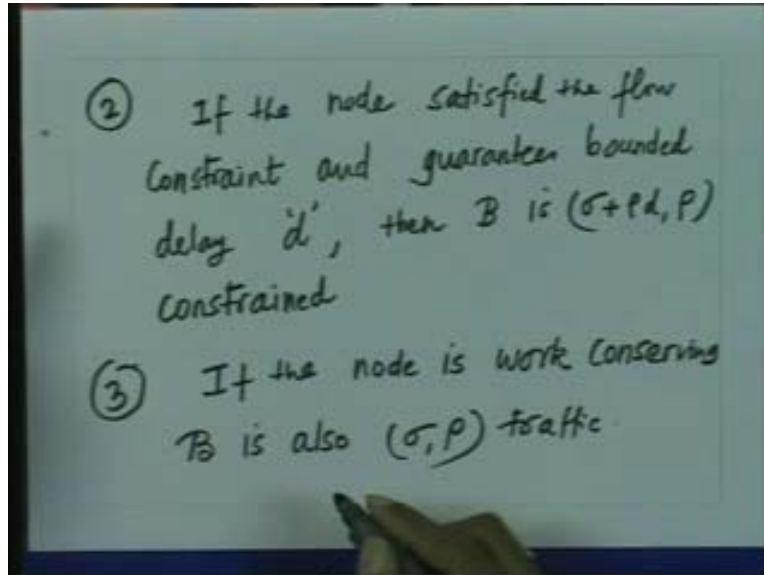
(Refer Slide Time: 49:01)



So now, we would like to now characterize the output burstiness. Now, some definitions are required; let us say A is the input and B is the output. So, we say that a network satisfies the flow constraint if $B(t)$ is less than or equal to $A(t)$ for all t . So, this is our definition.

Now, we will prove 2 results. So, claim one we would like to prove that if the node, the network node guarantees bounded queue length - q the network node guarantees the bounded the queue length q ; then the output B is $\sigma + q$ into ρ constraint. So, what we are saying is if the node can guarantee the bounded queue length q ; note that the node can guarantee the bounded queue length if it is a work conserving link and the input happens to be a ρ σ constrained traffic. We have already seen that a node indeed can guarantee the bounded queue length. Then what we are trying to see that the output of such a node will also be a deterministically bounded arrival process. But there, the σ will be $\sigma + q$ and ρ .

(Refer Slide Time: 51:36)



The second result that we would like to prove is that a second conjecture is that **sorry** the second result is that if the node satisfies the flow constraint and guarantees **if the node satisfies the flow constraint and guarantees** the bounded delay of d ; then the output d is σ plus ρd into σ constrained.

And, the third result is you would like to see is that if the node, the network node is work conserving; then B is also σ ρ traffic. So, what we are trying to say that **even** the network node can give us the bounded delay and the bounded q length; then we can completely characterize the output burrstones of the traffic. So, what we have essentially seen today is that if we have a ρ σ constrained traffic; then what is the effect of multiplexing, what was the effect on the output burrstones and what are the effect on maximum q length and the maximum delay.

(Refer Slide Time: 53:40)

REFERENCES

Cheng - shang Chung
' Performance Guarntess in
Communications Network'
Springer Verlag 2000