

Delay Estimation through Monitoring the Address Buffers of a Shared Buffer Type ATM Switch*

Seung Yeob Nam¹, Joo Yong Lee¹, Dan Keun Sung¹, and Soo Jong Lee²

¹Department of Electrical Engineering, KAIST, Korea

²Quality Assurance Team, ETRI, Korea

E-mail: synam@cnr.kaist.ac.kr

Abstract

We usually measure cell delays in ATM switches by using time stamps. Alternatively, we here propose a delay estimation method by monitoring the address buffers of a shared buffer type ATM switch. We derive a relation between the address queue length distribution and cell delay distribution and verify this relation by simulation. For a multiple stage switch, we can estimate end-to-end delay characteristics by the successive convolutions of the delay distribution of each stage. We investigate cell delay distributions of multiplexed streams as well as a specific virtual connection (VC) stream. We derive the upper bound of cell delay distribution of each connection. Finally, we study the effect of burstiness on delay, and show that the burstiest traffic yields the worst delay performance under the given traffic parameters when the background traffic load is low.

1. Introduction

An essential feature of Asynchronous Transfer Mode (ATM) based solutions for B-ISDN are their potential to use the same set of network resources to support a variety of user services. However, for network managers or service providers, it is difficult to implement various traffic control schemes including CAC and congestion control because the statistical characteristics of multiplexed traffic is very complex.

Delay measurements are essential to characterize delay characteristics of a specific connection. A conventional approach is to directly measure the delay in ATM networks by using time stamps in OAM cells [1]. However, if time stamps are used for only OAM cells, it is very complicated to analyze the delay characteristics for various connections between input and output ports. Alternatively, we propose a delay estimation method by monitoring the address buffers of a shared buffer type ATM switch. This method yields an upper bound of delay distribution for a specific connection. In this paper, we investigate the cell delay distribution at a single switch module and extend to analyzing it at a multiple stage ATM switch by the successive convolutions of the delay distribution at each stage.

This paper is organized as follows. In Section II, we derive delay distributions from the information obtained by monitoring the address buffers of a shared buffer type ATM switch. In Section III, we obtain the delay distributions of a specific VC as well as multiplexed streams and investigate the effect of burstiness on delay performance. In addition, we characterize the delay performance of a multiple stage switch. In Section IV, we verify the derived relations by simulation. Finally, we conclude in Section V.

2. Prediction of Cell Delay Distribution through Monitoring Address Buffers

2.1 Single ATM Switch Module

Fig. 1 shows a shared buffer ATM switch module. It consists of serial/parallel converters, a multiplexer (MUX) at the input side, a demultiplexer (DMUX) and parallel/serial converters at the output side, one shared memory, address first-in-first-out (AFIFO) buffers, idle address pool (IAP), idle address controller (IAC), priority control & routing decoder, and broadcasting routing memory (BRM).

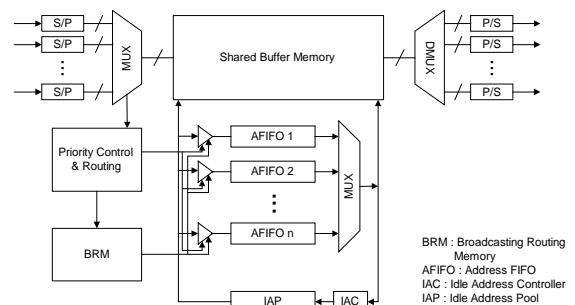


Fig. 1. An $n \times n$ single ATM switch module

2.2 Relation between an Address buffer and Output Link Utilization

We model an address buffer as a single queue system as shown in Fig. 2. For the below single queue system, the write-in address information of cells arrives in the

* This research is supported in part by Electronics and Telecommunications Research Institute (ETRI).

discrete time slot, and the service time is deterministic. If the system is empty when a cell address arrives at the queue, the corresponding output cell time becomes idle. Output link utilization is defined as the ratio of the number of busy slots to the number of total time slots. The probability that an output port is idle at an arbitrary time slot is equal to the probability that the queue is idle at the previous time slot. Let N be the system size observed at a random time slot. Then, we can obtain the following relation between the output link utilization (U_{out}) and the probability that the queue is empty at random times.

$$U_{out} = 1 - P(N = 0) \quad (1)$$



Fig. 2. A single queue system

2.3 Relation between AFIFO Length Distribution and Cell Delay Distribution

When the capacity of the AFIFO is K , we derive a relation between the cell waiting time distribution at the queue and the queue length distribution observed by an arriving cell.

Let W and N^- denote the waiting time of a cell in the queue and the system size observed by an arriving cell, respectively. Cells arrive and depart the queue on slotted time slots. If an arriving cell sees i cells in the system, the cell should wait until i cells are all served. Thus, the following equation is hold:

$$P(W = i) = P(N^- = i). \quad (2)$$

Eqn (2) is valid in the case of infinite queue capacity. In the case of a finite system capacity of K , we can obtain the following modified equation.

$$P(W = i) = \frac{P(N^- = i)}{1 - P(N^- = K)}, \quad (3)$$

for $i = 0, 1, \dots, K - 1$.

2.3.1 Relation between the Queue Length Distribution Observed by an Arriving Cell and the Queue Length Distribution Observed at Random Times

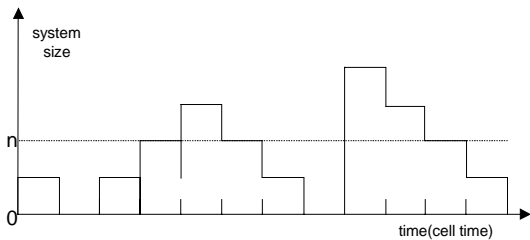


Fig. 3. Change of system size over time

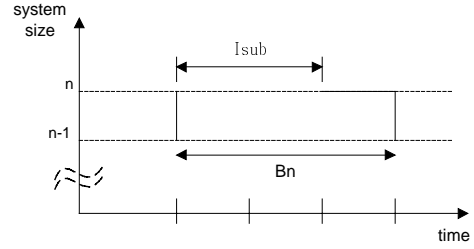


Fig. 4. Busy period of level n

Fig. 3 shows a sample path of system size over time. We derive a relation between the system size observed by an arriving cell and the system size at random times. If we denote the number of cell times when the system size is equal to n during t as $S_n(t)$, the probability that there are n cells in the system at an arbitrary time ($P(N = n)$) can be expressed as

$$P(N = n) = \lim_{t \rightarrow \infty} \frac{S_n(t)}{t}. \quad (4)$$

If an arriving cell sees a system size of $n-1$ ($n=1, 2, \dots, K$), the system remains in the busy period (B_n) of level n , as shown in Fig 4, where interval I_{sub} is defined as B_n , less one cell time. The system size is larger than n during I_{sub} . At the last time slot of busy period (B_n), the system remains at the state of a system size of n . If we denote the number of arriving input cells that see a system size of $n-1$ during t as $A_{n-1}(t)$, the following relation is hold:

$$|S_n(t) - A_{n-1}(t)| \leq 1. \quad (5)$$

If we denote the number of cell arrivals during t cell times as $l(t)$, the probability that an arriving cell sees the system size of $n-1$ can be expressed as

$$P(N^- = n-1) = \lim_{t \rightarrow \infty} \frac{A_{n-1}(t)}{\sum_{j=0}^K A_j(t)} = \lim_{t \rightarrow \infty} \frac{A_{n-1}(t)}{l(t)}. \quad (6)$$

During t cell times $\sum_{j=0}^{K-1} A_j(t)$ cells can enter the queue. For a finite system capacity of K , the difference between the number of input cells and output cells cannot exceed K . Let $Q_{out}(t)$ be the number of output cells during t cell times.

$$\lim_{t \rightarrow \infty} \frac{Q_{out}(t)}{t} = \lim_{t \rightarrow \infty} \frac{Q_{out}(t) - K}{t} \leq \lim_{t \rightarrow \infty} \frac{\sum_{j=0}^{K-1} A_j(t)}{t} \leq \lim_{t \rightarrow \infty} \frac{Q_{out}(t) + K}{t} = \lim_{t \rightarrow \infty} \frac{Q_{out}(t)}{t}. \quad (7)$$

Since at most one cell can be served in a cell time, $Q_{out}(t)/t$ is the output link utilization U_{out} . Thus, the following relation can be obtained from Eqn (7).

$$\lim_{t \rightarrow \infty} \frac{l(t)}{t} = \frac{U_{out}}{1 - P(N^- = K)}. \quad (8)$$

From Eqns (4), (5), (6) and (8), we obtain the following result.

$$P(N = n) = \lim_{t \rightarrow \infty} \frac{l(t) S_n(t)}{t l(t)} = U_{out} \frac{P(N^- = n-1)}{1 - P(N^- = K)}, \quad (9)$$

for $n = 1, 2, \dots, K$.

Eqn (9) shows a relationship between the system size distribution observed by an arriving cell ($P(N^- = n-1)$) and the system size distribution at random times ($P(N = n)$) for a system of capacity K .

Combining Eqns (1), (3), and (9) yields the following equation:

$$P(W = i) = \frac{P(N^- = i)}{1 - P(N^- = K)} = \frac{P(N = i+1)}{1 - P(N = 0)}, \quad (10)$$

where W is the waiting time of an arriving cell. The above Eqn (10) shows a relationship among the distribution of the system size observed by an arriving cell, that of system size observed at random times, and that of cell waiting time.

3. Cell Delay Performance of a specific VC

3.1 Relation of Cell Delay Distributions of a VC and Multiplexed Streams

We consider the relation between the distribution of the system size observed by arriving cells of all connections and that of the system size observed by cells of each connection.

Let N_j^- be the system size observed by an arriving cell of connection j , and a_j be the number of cells of connection j , and r be the number of connections. Then, we can obtain the following relation.

$$P(N^- = i) = \sum_{j=1}^r \gamma_j P(N_j^- = i), \quad (11)$$

where $\gamma_j = a_j / \sum_{i=1}^r a_i$ is the relative load of

connection j compared to aggregate connections. Therefore, the distribution of the system size seen by arriving cells of all connections is the weighted average of that of each connection.

In case that cell losses do not occur, we can obtain the result that the aggregate delay of all connections is the weighted average of delay of each connection from Eqns (3) and (11).

Theorem 1.

$$P(W_n = i) \leq \frac{1}{\gamma_n} \frac{1 - P(N^- = K)}{1 - P(N_n^- = K)} P(W = i), \quad (12)$$

for $n=1, 2, \dots, r$

where W_n denotes the waiting time of cells of connection n . We can derive the following relation from Eqn (11).

$$P(N_n^- = i) \leq \frac{1}{\gamma_n} P(N^- = i) \quad (13)$$

From Eqns (3) and (13), the result can be derived. \square

Theorem 1 implies that if we know the delay distribution of aggregate traffic, we can obtain the upper bound of delay distribution of each connection. In Section II, we derived the relationship between the AFIFO length distribution and the cell delay distribution. We can estimate the upper bound of cell delay distribution of a specific connection by just monitoring the corresponding AFIFO. The relative load γ_n of Eqn (12) can be evaluated either by monitoring per connection or by calculating from traffic parameters. Theorem 1 can be used to estimate the delay QoS of a specific connection.

3.2 Effect of Burstiness on Delay Performance

Various factors affect the delay distribution of a selected connection, and they include load from other connections, its own traffic characteristics, and so on [3]. The characteristics of background traffic except the connection of interest is very complicated in real network. We assume the arrival process from other connections is Bernoulli process to focus on analyzing the effect of burstiness on delay performance. This assumption is reasonable for low load case. Burstiness is usually defined as the ratio of peak cell rate to average cell rate [5], but it has a limitation in describing various traffic sources. We define the following as a new measure of burstiness.

Definition. The Mean Reciprocal of Interarrival time (MRI) of VC s is defined as

$$MRI_s = \frac{1}{m-1} \sum_{i=1}^{m-1} \frac{1}{\tau_i^s}, \quad (14)$$

where m is the total number of cells belonging to connection s and τ_i^s is the interarrival time between the i -th and $(i+1)$ -th arriving cells of connection s . \square

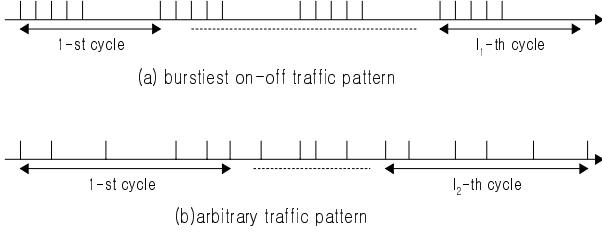


Fig. 5. Comparison of the burstiest traffic and other traffic patterns

We can show that if traffic parameters, peak cell rate (PCR), sustainable cell rate (SCR), and maximum burst size (MBS) are given, the traffic pattern that maximizes the MRI is a on-off traffic, which consists of an on-period when a number of cells, with a size of MBS, are arriving at the rate of PCR during a period of $\{MBS/PCR\}$ and an off-period whose length is $\{MBS/SCR \square MBS/PCR\}$. We will show the traffic pattern which maximizes MRI yields the worst delay performance when the background traffic is Bernoulli process.

Fig. 5 shows two cell arrival patterns. In Fig. 5(a) we consider the traffic that maximizes MRI, given parameters PCR, SCR, and MBS. Fig. 5(b) shows a different pattern. One cycle is a period when the number of cell arrivals is equal to the value of MBS for both the case of Fig. 5(a) and Fig. 5(b). First we compare one cycles of Fig. 5(a) and Fig. 5(b).

For the connection of interest, we can set up a relation which relates the cell interarrival time and the load from other connections to the system size observed by arriving cells of that connection.

We introduce the following variables:

t_n : arriving time slot of the n-th cell belonging to connection s.

$N_s^-(t_n)$: system size observed by the n-th cell of connection s.

$A^s(t)$: number of arriving cells during time slot t from all connections except connection s.

The system size observed by the n-th cell of connection s is given by

$$N_s^-(t_n) = f\{f\{\dots f\{f\{N_s^-(t_{n-1}) + 1 + \overline{A^s}(t_{n-1}) - 1\} + \overline{A^s}(t_{n-1} + 1) - 1\} \dots\} + \overline{A^s}(t_n - 1) - 1\}, \quad (15)$$

$$\text{where } f(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

If the off period is sufficiently long in Fig. 5(a), the effect of adding load by its own bursty traffic decreases and the queue state becomes stable. We assume that the system size observed by the first cell of a cycle has an

identical distribution with that observed by the first cells of other cycles.

We consider a connection s_1 and s_2 for Fig. 5(a) and Fig. 5(b), respectively. We will compare the delay distributions of connections s_1 and s_2 under the condition of identical background traffic.

The assumption that arrival process from the background traffic is Bernoulli process means that

$$\overline{A^{s_1}}(t) \leq 1, \quad \overline{A^{s_2}}(t) \leq 1, \quad \text{for } \forall t. \quad (16)$$

From Eqn (15), we obtain the following relations:

$$P(N_{s_1}^-(t_m) \geq k) = P\left(N_{s_1}^-(t_1) + \sum_{i=1}^{m-1} \overline{A^{s_1}}(t_i) + \sum_{i=1}^{m-1} \sum_{t_i < s_{t_i} \lfloor \frac{\text{link rate}}{\text{PCR}} \rfloor - 1} (\overline{A^{s_1}}(t) - 1) \geq k\right), \quad (17)$$

$$P(N_{s_2}^-(t_m) \geq k) = P\left(N_{s_2}^-(t_1) + \sum_{i=1}^{m-1} \overline{A^{s_2}}(t_i) + \sum_{i=1}^{m-1} \sum_{t_i < s_{t_i} \lfloor \frac{\text{link rate}}{\text{PCR}} \rfloor - 1} (\overline{A^{s_2}}(t) - 1) + \sum_{i=2}^m \sum_{t_{i-1} \lfloor \frac{\text{link rate}}{\text{PCR}} \rfloor \leq t \leq t_i - 1} (\overline{A^{s_2}}(t) - 1) \geq k\right), \quad (18)$$

for $k \geq m$.

We obtain the following inequality by comparing Eqns (17) and (18) because $\sum_{i=2}^m \sum_{t_{i-1} \lfloor \frac{\text{link rate}}{\text{PCR}} \rfloor \leq t \leq t_i - 1} (\overline{A^{s_2}}(t) - 1) \leq 0$ from Eqn (16).

$$P(N_{s_1}^-(t_m) \geq k) \geq P(N_{s_2}^-(t_m) \geq k), \quad (19)$$

for a sufficiently large value of k .

If we denote the system size observed by the n-th cell of cycle a of connection s as $N_{s;a}^-(t_n)$, the following relation can be obtained from (19):

$$P(N_{s_1}^- \geq k) = \frac{1}{l_1 \cdot MBS} \sum_{a=1}^{l_1} \sum_{i=1}^{MBS} P(N_{s_1;a}^-(t_i) \geq k) \quad (20)$$

$$\geq \frac{1}{l_2 \cdot MBS} \sum_{b=1}^{l_2} \sum_{i=1}^{MBS} P(N_{s_2;b}^-(t_i) \geq k) = P(N_{s_2}^- \geq k)$$

for a sufficiently large value of k.

The above Eqn (20) shows that given traffic parameters PCR, SCR, and MBS, the traffic of maximum MRI yields the worst delay performance MRI when cell arrival process from other connections is Bernoulli process.

In this section, we investigate the effect of burstiness on delay considering autocorrelation of the traffic of interest and found that burstiness has influence on delay. But the study is confined to the case of low traffic load. It is necessary to study more to analyze more general case. Given parameters PCR, SCR, and MBS, we can find a traffic pattern that maximizes MRI when background traffic load is low. Using this traffic pattern, we can estimate the delay distribution of this traffic before allocating a connection to the requesting source.

3.3 Cell Delay Performance of a specific VC for a Multiple Stage Switch

Thus far, we have considered cell delay performance at a single switch module. We now extend to analyzing the relation for a multiple stage switch.

If delays experienced in consecutive multiplexing nodes are almost uncorrelated, the convolution of the delay distribution of each multiplexing node becomes a good approximation for the distribution of the end-to-end delay [4]. But if a small positive correlation exists, the convolution slightly underestimates the variance of end-to-end delay of that connection. The experiment [4] indicated that delays introduced in consecutive multiplexing nodes are almost uncorrelated.

For a multiple stage switch, if we can obtain the delay distribution of the connection of interest at each stage, we can estimate the end-to-end delay distribution in the multiple stage switch by the successive convolutions of the delay distribution at each stage. Obtaining the end-to-end delay distribution, we can check whether its requested delay QoS such as CDV and CTD is guaranteed [2].

There can be several methods to obtain the delay distribution of a specific connection at each stage by monitoring AFIFO. First, we can differentiate the system size information observed by arriving cells of a specific connection from the information of other connections and manage the system size information separately, then we can estimate the delay distribution of each connection very clearly. Second, we can use Theorem 1 of Section 3.1. By Theorem 1 we can obtain the upper bound of delay distribution with little information such as the relative load of the connection. At each stage we can evaluate the upper bound of the specific connection, and it is possible to obtain the upper bound of end-to-end delay distribution by convolutions.

4. Simulation Results

4.1 Input Traffic Model

We evaluate the performance of a switch shown in Fig. 1. The number of input ports and the number of output ports are both 32. The delay is measured between input and output ports. Input and output ports are based on Synchronous Optical Network (SONET) STS-3c.

We use four types of CBR source and three types of VBR source [6].

Table 4.1 Characteristics of CBR Test sources

	CBR I	CBR II	CBR III	CBR IV
PCR	4,140	16,560	119,910	173
Load	0.0117	0.04688	0.33333	0.00049

When more than one connections of the same traffic pattern are generated, the phase of each connection is randomized over the interarrival time of that traffic.

VBR test sources are characterized as follows:

- A two-state Markov process consisting of an active state during which the source generates information-carrying cells and a silent state during

which cells are not emitted.

- The duration of an active phase has an integer number of cell times with a geometric distribution with a mean of M_a . The silent state lasts for an integer number of cell times, which is geometrically distributed with a mean of M_s .
- During the active state, a VBR test source emits a synchronous burst of cells, with a period of T , where T is an integer number of cell slots. The first cell of the burst occurs at the beginning of the active state. Thus, the mean burst size is given by $B = M_a / T$ cells.

Table 4.2 Characteristics of VBR Test Sources

VBR Source Type	VBR I	VBR II	VBR III
M_a (cell slots)	240	500	210
T (cell slots)	6	25	1
M_s (cell slots)	720	2500	2500
Load	0.04167	0.00667	0.07749

4.2 Relation between the Distribution of Address Buffers and the Distribution of Cell Delay

We consider two methods to obtain the distribution of address buffer lengths. One is to monitor the AFIFO when cells arrive, and the other is to monitor the AFIFO every cell time to obtain the queue length distribution at random times. We also investigate the relation between the AFIFO length distribution obtained by each method and the cell delay distribution. The result is given in Eqn (10), which indicates that using AFIFO length distribution either by monitoring AFIFO every cell time or by monitoring AFIFO on each cell arrival, we can obtain the cell delay distribution at the AFIFO.

We here compare the distributions estimated by the queue length distribution observed by arriving cells and by the system size distribution at random times with the cell delay distribution measured by time stamps for each cell.

Figs. 6 and 7 compare the estimated cell delay distributions with the measured cell delay distribution. In Fig. 6, CBR III traffic is served by AFIFO no.29, and the traffic load is 0.454544. In Fig. 7 VBR II traffic is served by AFIFO no.10, and the traffic load of that server is 0.447572. Both distributions are nearly identical.

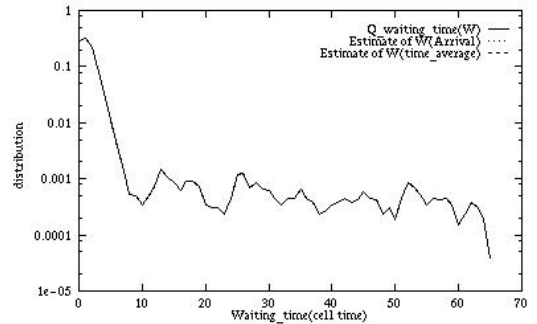


Fig. 6. Comparison of estimated and measured delay distributions (AFIFO 29)

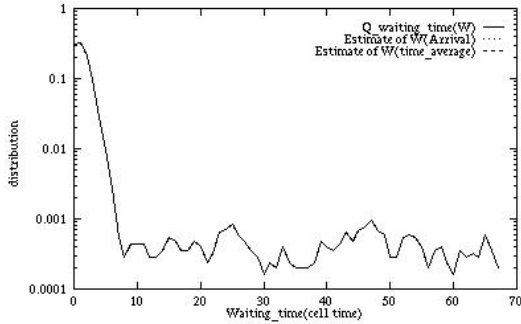


Fig. 7. Comparison of estimated and measured delay distributions (AFIFO 10)

4.3 Cell Delay Distributions of Multiplexed Stream and Each Connection

We now consider a case that seven connections of different traffic pattern are multiplexed into one address buffer. Under such a condition we investigate the delay characteristics of each connection and aggregate traffic. First, we verify Eqn (11) of Section 3.1 by simulation. In Fig. 8 the delay distribution of aggregate traffic is almost the same as the weighted average of each distribution.

In Fig. 9 for the same background traffic with CBR I, CBR II, CBR IV, VBR I, and VBR II, the delay characteristics of three foreground traffic classes are observed. The load of background traffic is 0.107. Traffic pattern #1, #2, and #3 are different traffic patterns of the same traffic parameters of PCR, SCR, and MBS. The traffic #1 has the traffic pattern of maximum MRI under the given parameters. In Fig. 9 the delay for this traffic shows the worst delay performance, and thus, it also supports the consequence of Section 3.2.

5. Conclusion

In this paper we obtained a relation among the AFIFO size distribution observed by arriving cells, AFIFO length distribution at random times, and cell delay distribution obtained by using the time stamp of each cell. Using this relation, we can estimate the cell delay distribution at a single switch module by monitoring the address buffer without using time stamp of each cell. We showed the validity of this relation by simulation. For a multiple stage ATM switch, we can estimate the end-to-end delay characteristics by the successive convolutions of the delay distribution of each stage.

We also derived a relation that the cell delay distribution of aggregate traffic is the weighted average of cell delay distribution of each connection. From this result we also derived a relation that gives the upper bound of delay distribution of a specific connection, which can be used to test whether the requested delay QoS is satisfied or not.

We showed that if traffic parameters PCR, SCR, and MBS are given, on-off traffic of maximum MRI yields the worst delay performance when the background traffic load is low. It consists of an on-period when a number of cells, with a size of MBS, are generated at the rate of

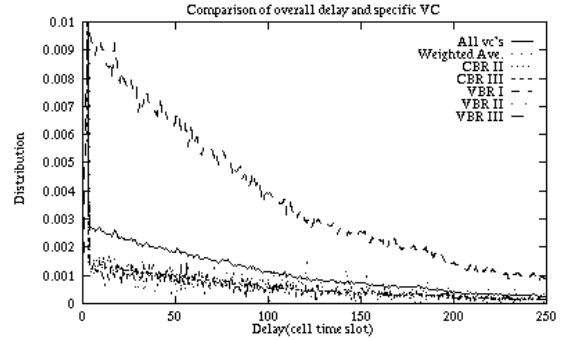


Fig. 8 Comparison of the delay distributions of aggregate traffic and each connection

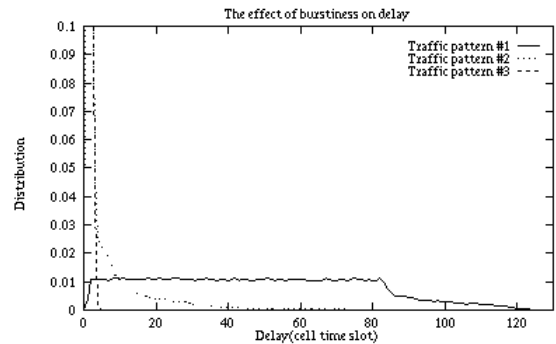


Fig. 9 Effect of burstiness on delay

PCR, and an off-period of the length determined by PCR, SCR, and MBS. This property can be used to CAC if it is extended to more general case.

References

- [1] M. Prycker, *Asynchronous Transfer Mode*, 3rd Ed., Prentice Hall, 1995.
- [2] ATM Forum, *Traffic Management Specification*, Version 4.0, 95-0013R10, 1996.
- [3] J. Roberts, U. Mocci, and J. Virtamo, *Broadband Network Teletraffic*, Springer, 1996.
- [4] D. Vleeschauer, "Experimental Verification of the Statistical Independence of Cell Delays Introduced in Consecutive Switches," in *B-ISDN Teletraffic Modelling Symposium*, pp. 105-116, Feb. 1995.
- [5] D. McDysan and D. Spohn, *ATM Theory and Application*, McGraw-Hill, 1994.
- [6] Bellcore, *Broadband Switching System (BSS) Generic Requirements, GR-1110-CORE*, Issue 1, Sep. 1994, Revision 3, Apr. 1996.