

## Chapter 4

# Probabilistic Delayed Packet Discard Schemes

Reaction to congestion detection should be started as quickly as possible. However, complicated congestion control schemes take long time to begin to work. Thus, complicated schemes are not recommended to control traffic and to avoid congestion. EPD, as is described in the previous chapter, meets this requirement for simplicity and is an effective way to avoid congestion. Nevertheless, TCP monitors traffic and regulates it at an upper layer than the ATM layer. Thus, EPD does not have as many chances to work as it would without TCP control. Although EPD begins to work after the queue length is larger than the threshold, the data stream must have already been controlled by TCP at the higher layer. Therefore, when TCP packets are transmitted over ATM, EPD may regulate more network traffic at the cell level which reduces the utilization of the buffer in a switch. Next section describes a result of a simple simulation which shows the possibility of improved buffer utilization. Then, the packet discard schemes are generalized. Based on the generalization, a new packet discard scheme named probabilistic delayed packet discard scheme (PDPD) is proposed which is expected to help exploit more available network resources.

### 4.1 Preliminary experiments of the PDPD scheme

To show the potential for improvement of availability of buffer space in a switch, this author will describe the results of a preliminary simulation experiment [MJO00].

In EPD, if a queue length in a switch grows larger than the predefined threshold value, a *discard* (or *D*) flag is set and newly propagated packets are not permitted to enter the buffer in the switch. Instead, they are discarded as long as this *D* flag is set. This discard is quite a hasty reaction. If the application transmission of the higher layer may be bursty, in other words, if the higher layer transmits the burst of packets sporadically, the increase of the queue length may continue for a moment, and the congestion can be resolved by itself. Hence, the above behavior of EPD might be too hasty for the network to use more available buffer resources in a switch.

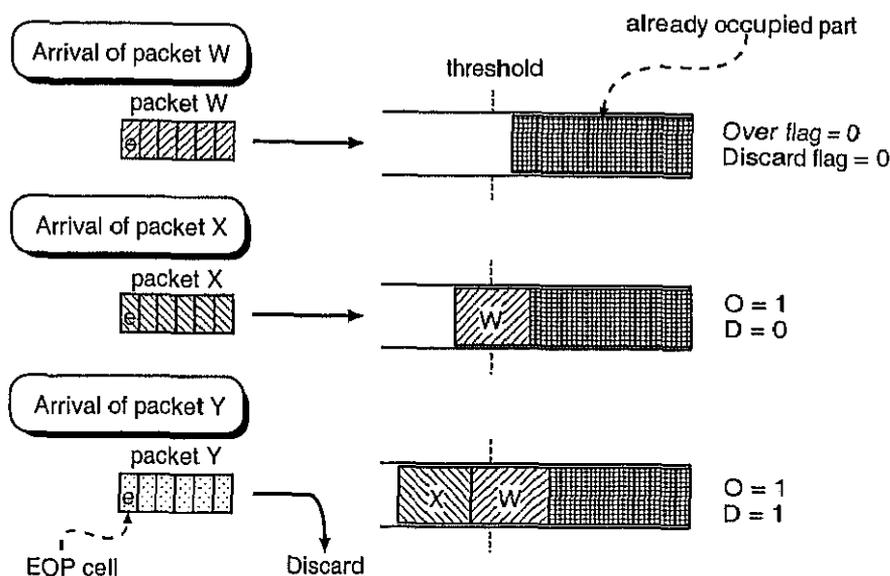


Figure 4.1: Introduction of preliminary PDPD scheme.

Figure 4.1 illustrates the introduction of simple modification of EPD. The scheme is called “pre-PDPD” in this dissertation. Pre-PDPD provides an additional *over* (or  $O$ ) flag to indicate whether the queue length is larger than the threshold. If it is larger than the threshold, an  $O$  flag is set but a  $D$  flag is not. Provided a new packet arrives when an  $O$  flag is set and a  $D$  is not, the packet is allowed to be buffered which then turns the  $D$  flag on. Actual discarding begins when both the  $O$  and  $D$  flags are set. Once both these flags are set, they will be reset when the queue length shrinks to less than the threshold. Pre-PDPD is an optimistic algorithm. It appears to be identical with the EPD scheme with the higher threshold value by one packet size. Obviously, both the EPD scheme and the pre-PDPD scheme are identical in the  $D$ -flag-setting phase. In the  $D$ -flag-resetting phase, however, pre-PDPD must suspend buffering longer than EPD operating with a higher threshold<sup>1</sup>. This is the tradeoff incurred for using this optimistic scheme. The idea of the pre-PDPD scheme is very simple. It temporarily permits one more new packet to enter the buffer even when the queue length is larger than the threshold. Considering the regulation of data flow by TCP congestion control schemes at the higher layer, this author has hypothesized that there remains available space in the buffer for a few more packets.

This author examines this pre-PDPD scheme in a simulation based on a model shown in Figure 4.2. In this simulation, there are ten TCP sources connected to ten TCP receivers via two ATM switches transmitting segmented cells back-to-back. Various Reno version flow controls are executed in the TCP level. All packets from these sources are fragmented into AAL 5 (ATM adaptation layer type 5) format ATM cells. These segmented cells are transmitted over

<sup>1</sup>The  $D$ -flag-setting phase and  $D$ -flag-resetting phase are the phases that the queue length is increasing and decreasing, respectively, according to the difference between the cell input rate and the cell output rate.

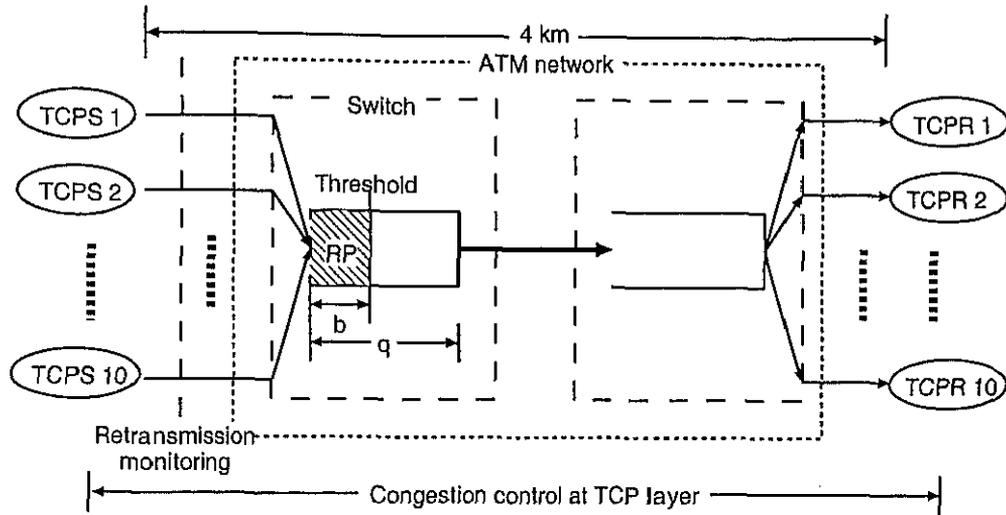


Figure 4.2: Simulation model of preliminary PDPD.

ATM using UBR (Unspecified Bit Rate), best-effort category services. User data arrives at the TCP sources deterministically and ceaselessly. The maximum transmission rate of each link is 48 Mbit/s<sup>2</sup>. However, since TCP sources control their streams, the bandwidth does not seem to be fully used, which leaves available buffer resources for other streams. The total length between a source and a receiver is about 4 km. The author calls the buffer space over the threshold a restricted part (RP) in this dissertation because it restricts the entrance of new packets. A buffer can hold  $q$  cells, and the RP size is  $b$  cells. Hence the threshold value of the buffer is  $(q - b)$ . For TCP over ATM, the size of the MTU (Maximum Transfer Unit) is 9180 octets. The total throughput of all TCP sources is examined for various buffer sizes. All simulation results are verified by a batch-means method [MACD90, pp. 140–151]. All results have a confidence level of 90% with a confidence interval of 2.0 %.

Figure 4.3 shows the total throughput of all TCP sources when the RP size ( $b$ ) is 384 cells, which corresponds to two packets<sup>3</sup>. Each point represents the total throughput of TCP sources for a particular buffer size using pre-PDPD (*pre-pdpd*), EPD (*epd*), and no specific policy (plain TCP or *ptcp*), respectively. In Figure 4.3, as the buffer size increases, throughput is also increased for all strategies. In this simulation, if too many retransmissions occur, the TCP connection is closed by itself. For small buffer sizes less than 1200 cells, the *ptcp* achieves no reliable throughput because the TCP connection is cut off due to too many retransmissions. Figure 4.4 shows the normalization of each throughput compared to that of *epd*. From Figure 4.4, it can be understood that *pre-pdpd* achieves at most 15 % higher throughput than the ordinary *epd* for relatively small buffer sizes.

Figure 4.5 shows an instance of snapshots of queue length marked periodically when EPD

<sup>2</sup>The transmission rate of OC-1 is about 48 Mbit/s.

<sup>3</sup>As the MTU is 9180 octets, the RP space of two packets is 384 cells.  $(9180/48 \sim 192, 192 \times 2 = 384)$

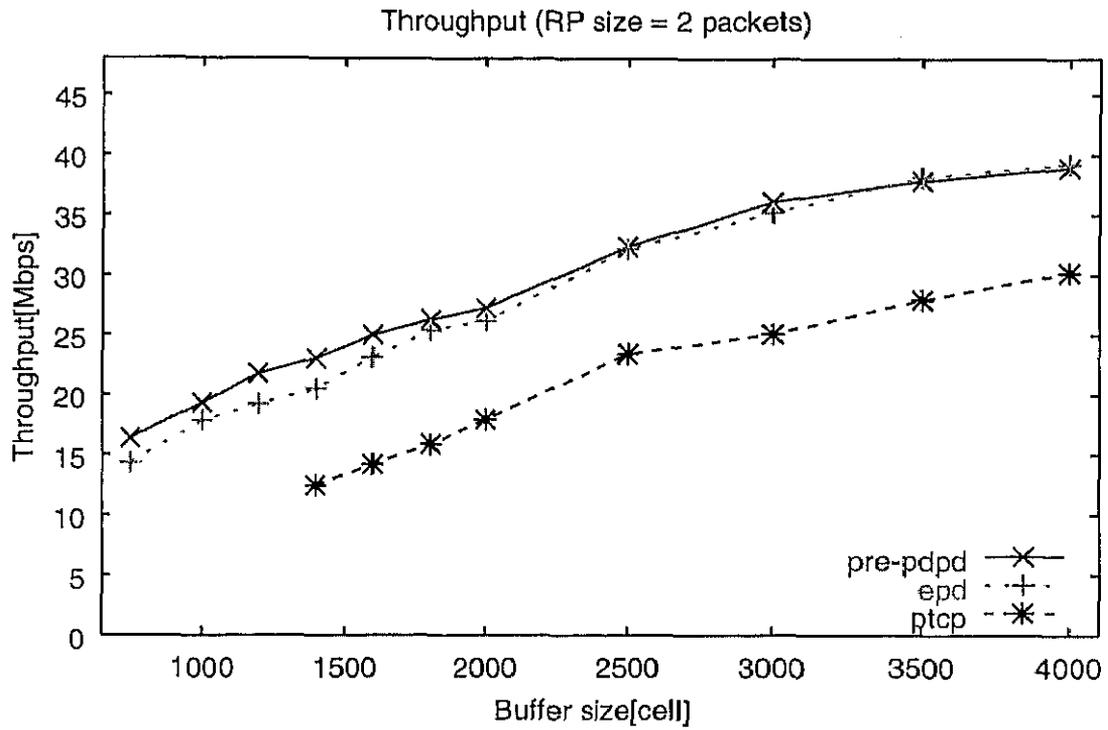


Figure 4.3: Total throughput of TCP sources: RP size of two packets.

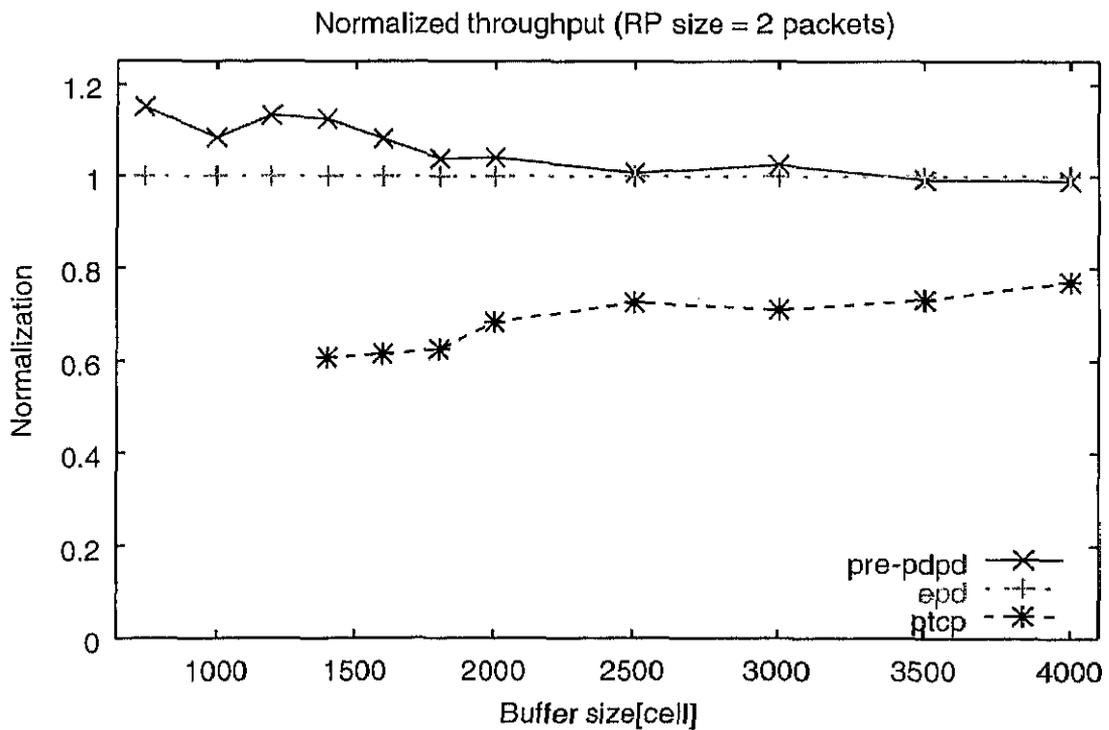


Figure 4.4: Normalized throughput of TCP sources: RP size of two packets.

(upper graph) and pre-PDPD are respectively used. The buffer size  $q$  is 1000 cells and the RP size  $b$  is two packets, hence, the threshold value is 600 cells. The X-axis shows the time slot and the Y-axis shows the queue length. One time slot is one cell transmission time in this simulation. A dotted line in each figure is the average queue length. In Figure 4.5, it seems that the RP (buffer space over 600 cells) is more utilized in pre-PDPD than in EPD. Quantitatively, the average queue length of the buffer is about 331 cells and 369 cells for EPD and pre-PDPD, respectively. Thus, the average queue length is about 11.5% longer by using pre-PDPD than by EPD.

Figure 4.6 and 4.7 show snapshots when the buffer size  $q$  is 2000 cells and 3000 cells, respectively. In Figure 4.6, RP (buffer space over 1600) is more utilized in pre-PDPD than in EPD. Specifically, the average queue length of the pre-PDPD switch is approximately 8% longer than that of the EPD switch quantitatively. When the buffer size  $q$  is 3000 cells, the average queue length of pre-PDPD is approximately 1.4% longer than that of EPD switch as shown in Figure 4.7. From the above figures (Figure 4.5, 4.6, and 4.7), it is known that RP is more utilize in pre-PDPD schemes.

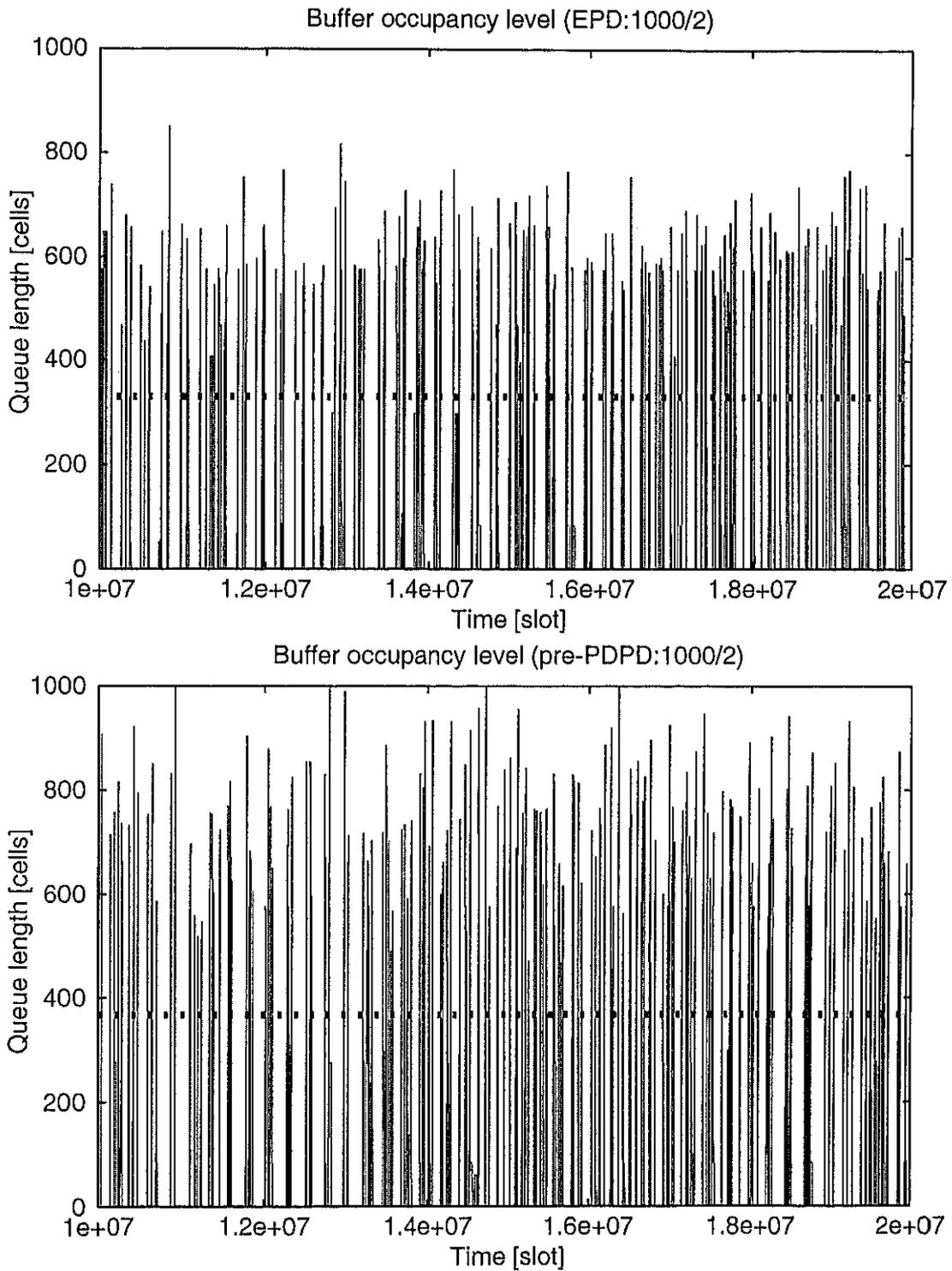


Figure 4.5: Snapshots of buffer occupation in the switch when the buffer size is 1000 cells and the RP size is two packets.

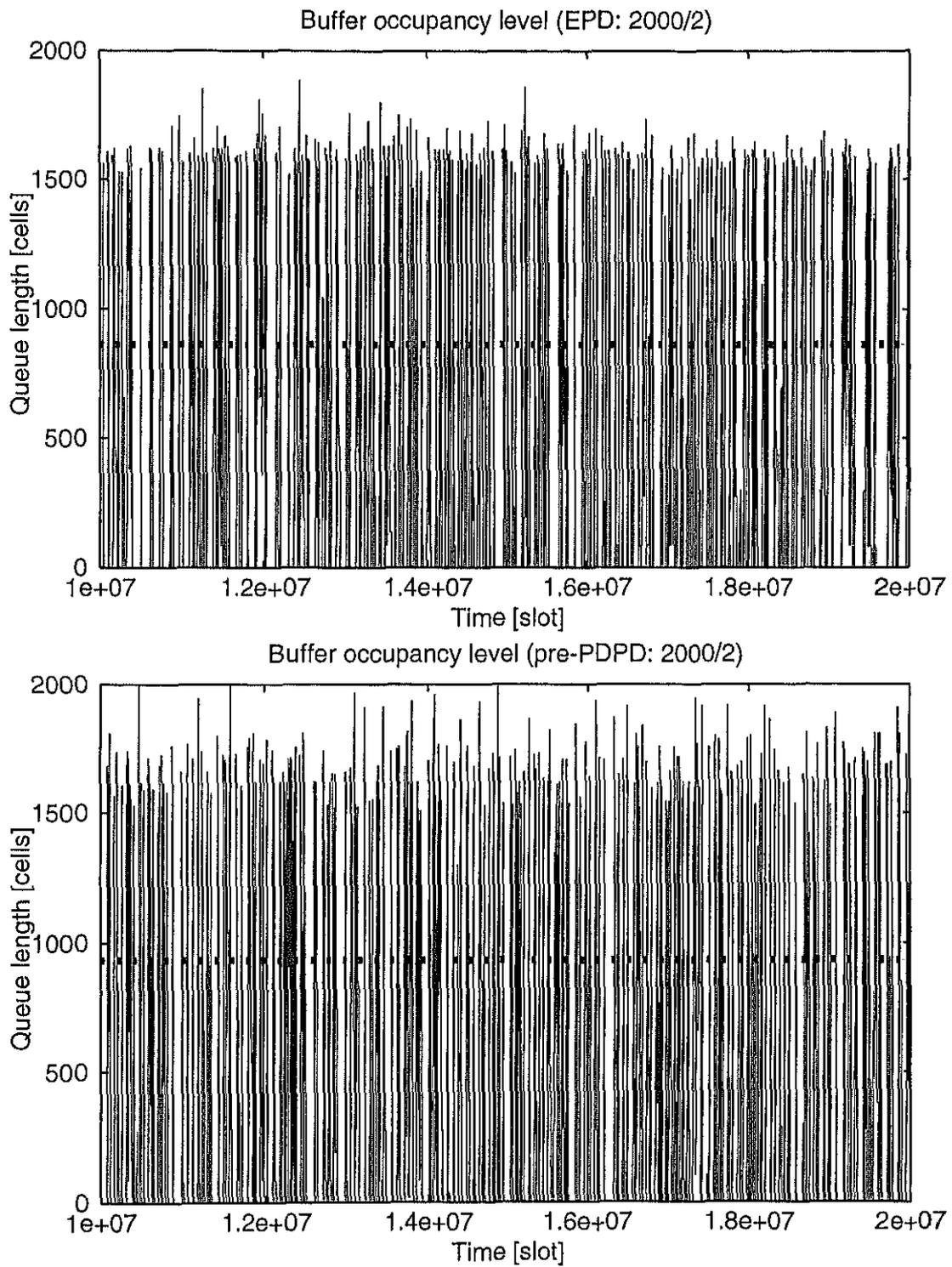


Figure 4.6: Snapshots of buffer occupation in the switch when the buffer size is 2000 cells and the RP size is two packets.

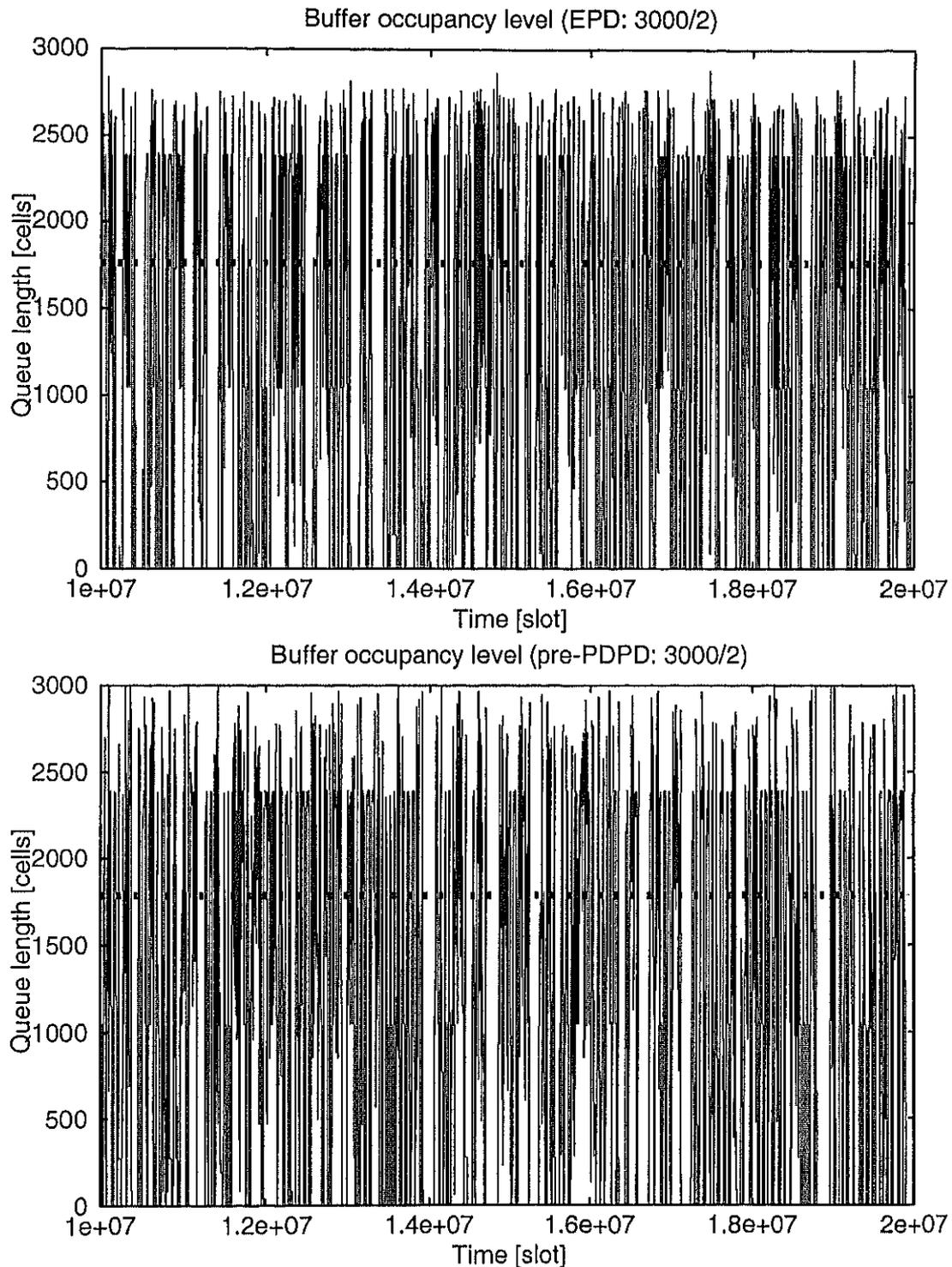


Figure 4.7: Snapshots of buffer occupation in the switch when the buffer size is 3000 cells and the RP size is two packets.

If the RP size is large, it is expected that pre-PDPD can utilize buffer space more effectively than EPD, because EPD starts the packet discarding operation earlier than that of small RP. Figure 4.8 shows the total throughput when the RP size  $b$  is three packets. For a buffer size

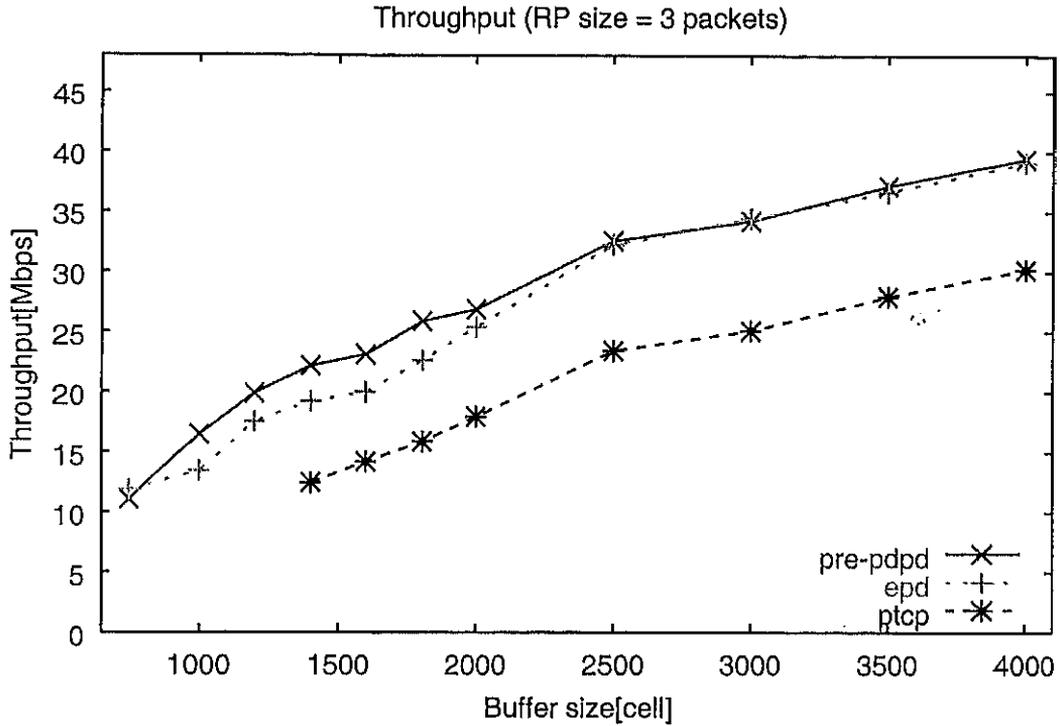


Figure 4.8: Total throughput of TCP sources: RP size of three packets.

of 750 cells, the confidence interval of EPD and pre-PDPD is 2.3 % and 2.6 %, respectively. Normalized throughput is presented in Figure 4.9. In these figures (Figure 4.8 and 4.9), pre-PDPD achieves higher throughput than EPD for small buffer ranges. Specifically, when the buffer size is between 1000 cells and 2000 cells, pre-PDPD improved throughput approximately 10 % compared to EPD. When the buffer size is 750 cells, however, the buffer size is excessively small. In the case of those with quite an insufficient buffer size, the more packets are admitted to be buffered the more packets are dropped due to the buffer overflow. It results in lower throughput via pre-PDPD.

Figure 4.10 shows an instance of snapshot when the buffer size  $q$  is 1000 cells and the RP size  $b$  is three packets. In the Figure 4.10, the average queue length of EPD and pre-PDPD is 260 cells and 277 cells, respectively. Quantitatively, the average queue length of pre-PDPD is 6.5 % longer than that of EPD.

Figure 4.11 shows an instance of snapshot when the buffer size  $q$  is 2000 cells and the RP size  $b$  is three packets. The average queue length of EPD is 723 cells, and that of pre-PDPD is 757 cells, which is 4.7 % longer than that of EPD.

An instance of snapshot when the buffer size  $q$  is 3000 cells with the RP size  $b$  of three packets

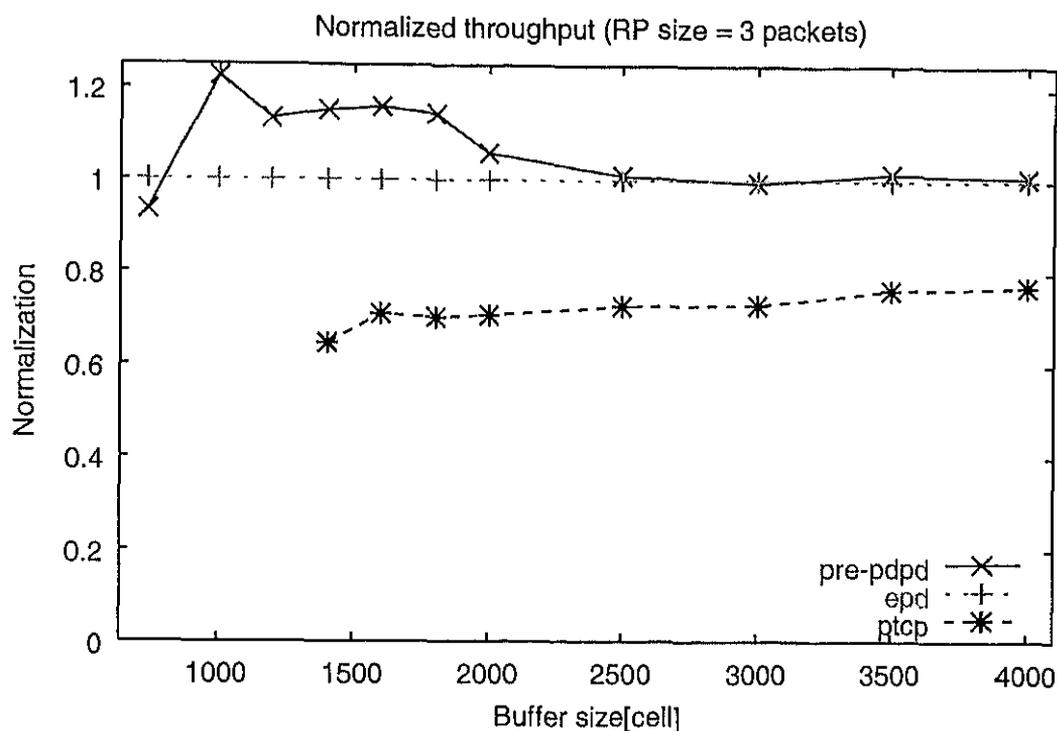


Figure 4.9: Normalized throughput of TCP sources: RP size of three packets.

is shown in Figure 4.7. The average queue length of pre-PDPD is 0.9 % longer than that of EPD in this figure (Figure 4.12). It is true that there is a greater possibility that the buffer overflow will occur in pre-PDPD rather than in EPD. However, since pre-PDPD is basically an optimistic policy to exploit more unused network resources, when the higher layer regulates its streams it is feasible to apply pre-PDPD speculatively.

In TCP over ATM, the retransmission ratio is a very important factor to determine the relationships between TCP and ATM, because the retransmission occurs mainly due to the packet discard in the ATM layer. Figure 4.13 shows the retransmission ratio for moderate buffer ranges when the RP size  $b$  is two packets. The confidence level is 90 % and the confidence interval is 2.7%. From the the Figure 4.13, it can be understood that the retransmission ratio of pre-PDPD is lower than that of EPD for small and moderate buffer sizes.

From the results of these preliminary experiments, it can be concluded that by permitting one more packet to enter the buffer, not only the buffer utilization but also the network throughput and retransmission ratio are improved. Consequently, as TCP regulates the streams, aggressive dropping of cells by EPD makes more available buffer space and leaves it unused. Therefore, if it is possible to alleviate this aggressiveness of EPD, it is also possible to use more buffer space efficiently, thereby increasing network throughput at the same time. The next section describes the generalization of the ordinary packet discard schemes.

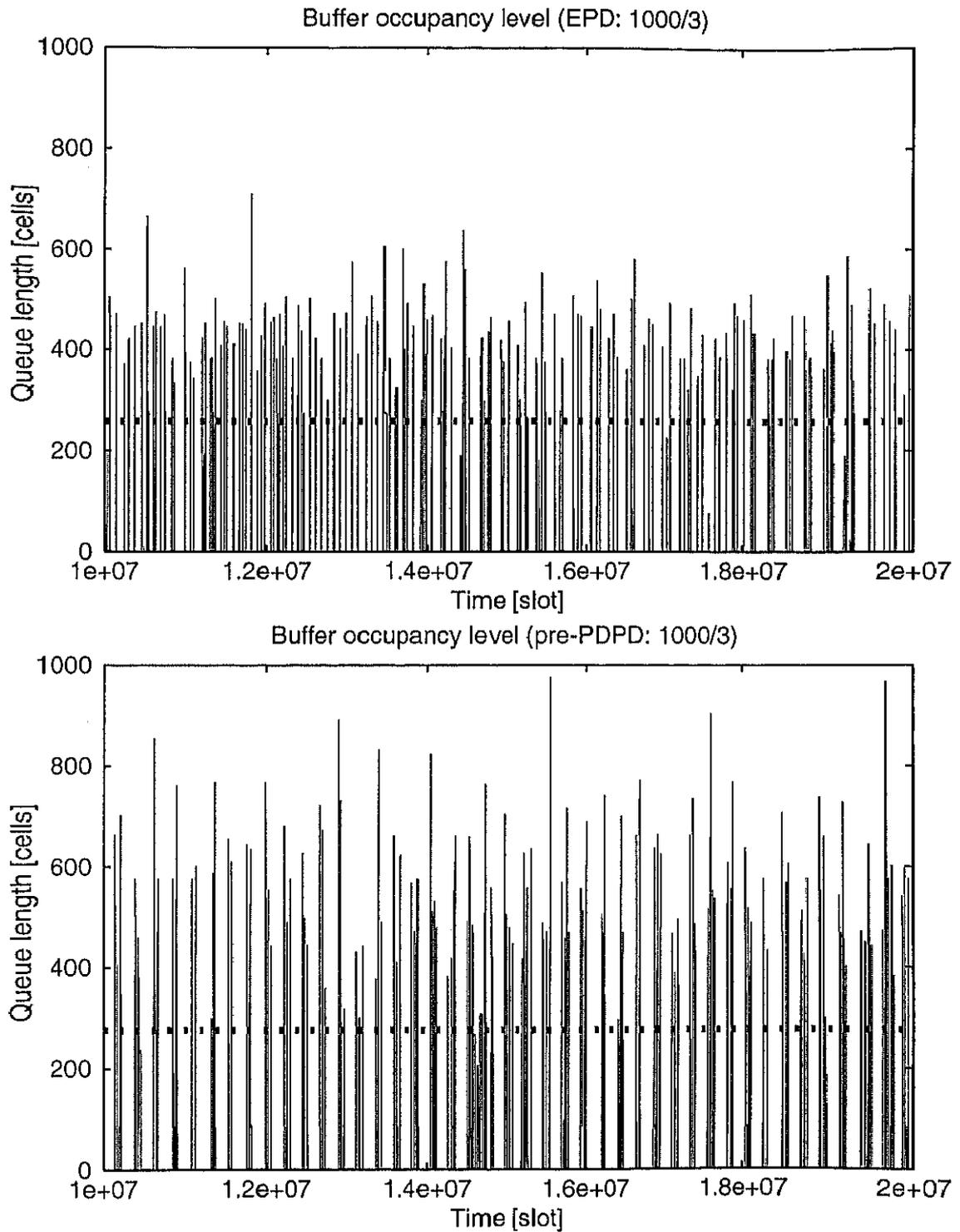


Figure 4.10: Snapshots of buffer occupation in the switch when the buffer size is 1000 cells and the RP size is three packets.

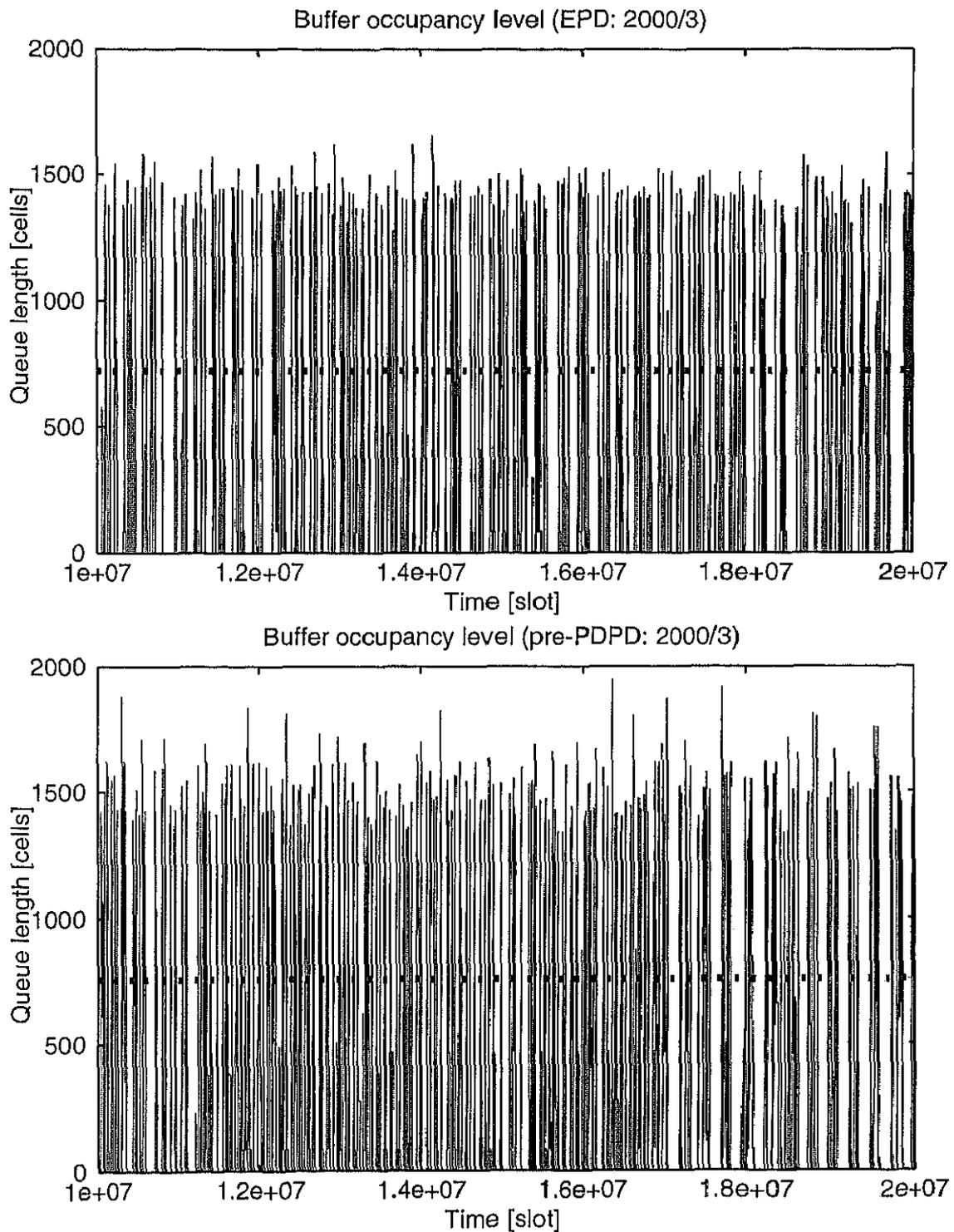


Figure 4.11: Snapshots of buffer occupation in the switch when the buffer size is 2000 cells and the RP size is three packets.

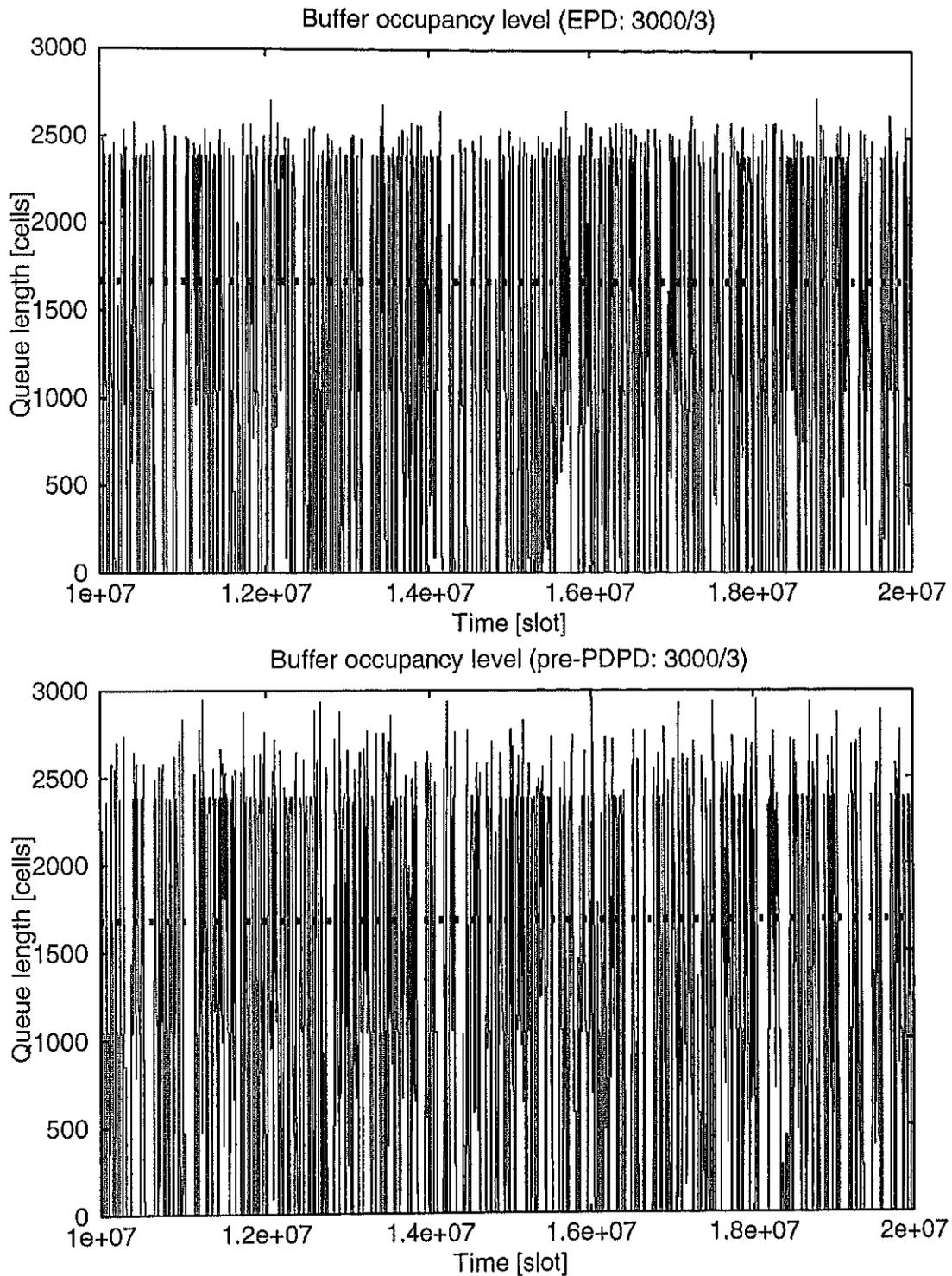


Figure 4.12: Snapshots of buffer occupation in the switch when the buffer size is 2000 cells and the RP size is three packets.

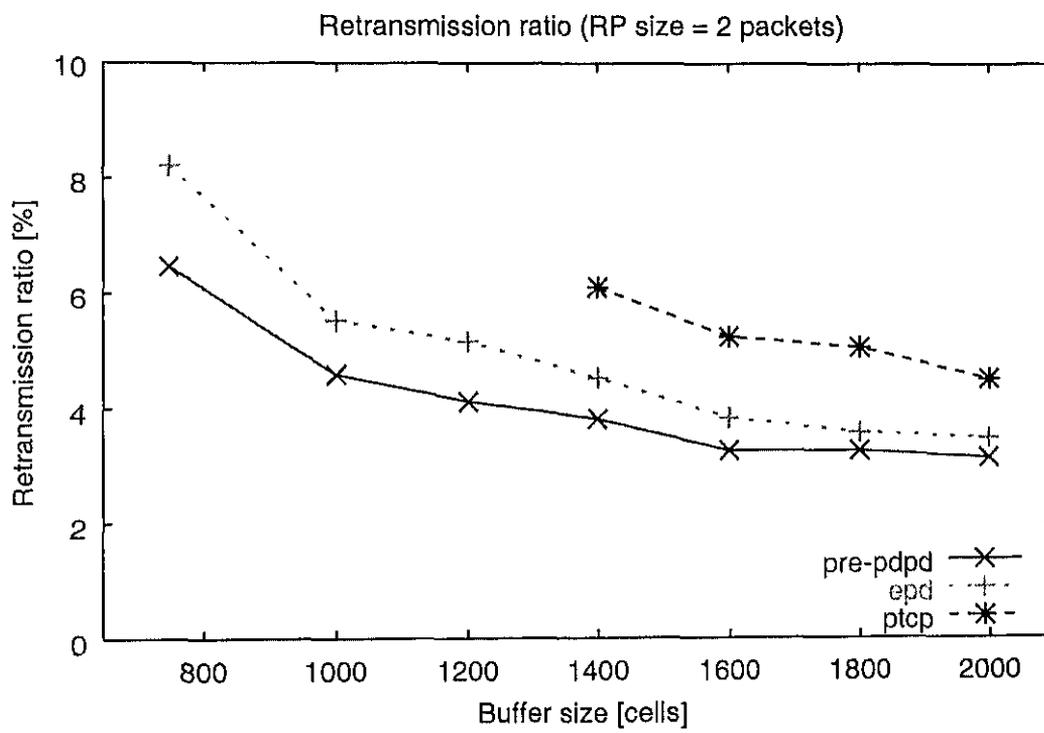


Figure 4.13: Retransmission ratio.

## 4.2 Generalization of packet discard schemes

In EPD, a  $D$  (discard) flag is set in two cases. In the first case, if a queue length gets larger than the predefined threshold value, a  $D$  flag is set, thereby preventing any cells of a new packet from entering the buffer. In the second case, the  $D$  flag is set when any cells are dropped due to buffer overflow, for example. In contrast, PPD does not use this threshold value. It basically sets a  $D$  flag when a cell is dropped due to buffer overflow. In both the PPD and EPD schemes, if the  $D$  flag is set, all the subsequent cells of the corresponding packet are discarded until the flag is reset. Comparing the characteristics of these two methods, it can be understood that PPD is a particular form of EPD whose threshold value is the size of the buffer.

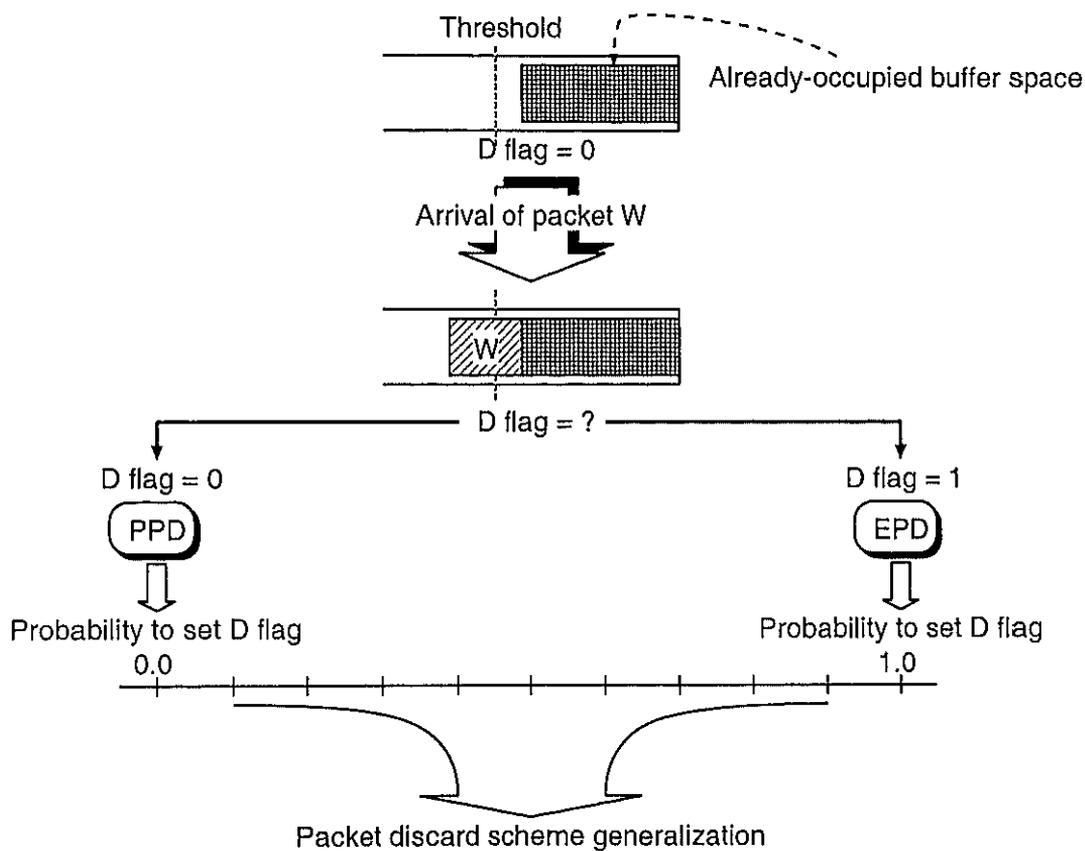


Figure 4.14: Generalization of packet discard schemes.

In addition to this, this author has found that the PPD scheme can be taken for another specific form of EPD from another point of view. Let us assume that PPD also has the same threshold value as that of EPD, and PPD uses the threshold by any means. For the PPD, a  $D$  flag is never set even though the queue length is larger than the threshold. As the buffer becomes full and overflows, cells are dropped and the  $D$  flag is set at last, thereby dropping all cells of the tail part of a packet. Consequently, if a  $D$  flag is always set when a queue length is larger

than the threshold, it is EPD. On the contrary, if the  $D$  flag is never set when a queue length is larger than the threshold, it is PPD. Therefore, if it is possible to set the discard flag with a probability, it is also possible to generalize these ordinary packet discard methods. In other words, both PPD and EPD are two extreme cases of the generalized packet discard algorithm. Figure 4.14 simply shows the generalization of these packet discard schemes. As long as a buffer length is less than the threshold, the  $D$  flag is a zero. Once a buffer length gets larger than the threshold, PPD will set the  $D$  flag with a probability of zero while EPD will set the flag with a probability of one.

### 4.3 Probabilistic Delayed Packet Discard Scheme

Based on the above formalization of packet discard schemes as discussed in the previous section, this author proposes a “Probabilistic Delayed Packet Discard” (PDPD) scheme. As shown in

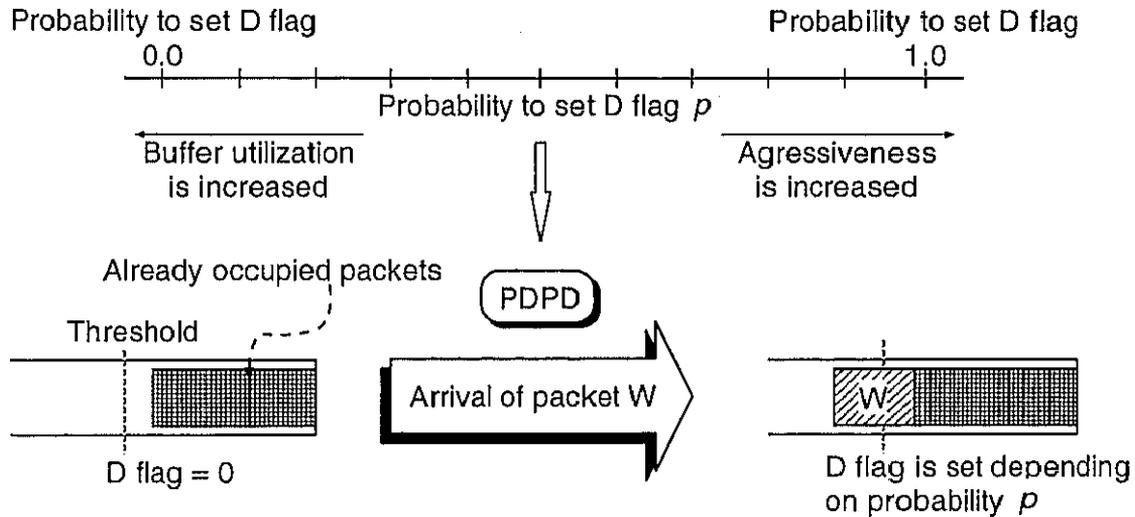


Figure 4.15: Probabilistic Delayed Packet Discard scheme.

Figure 4.15, in the PDPD scheme, when a queue length in a buffer is larger than the threshold, the  $D$  flag is set with a probability  $p$ . In other words, at probability  $(1 - p)$ , the  $D$  flag remains unset. This  $D$  flag is reset when a queue length becomes less than the threshold value. Since the actual dropping of the packet begins after the flag is set, as long as the flag is a 0, a new packet is allowed to be buffered. Provided the  $D$  flag is a 0 due to the probability  $p$ , when a new packet arrives and the queue length is larger than the threshold, the discarding operation is delayed until the  $D$  flag is set. This is a major difference between the original EPD and our PDPD scheme. It is also why we call our algorithm the “Probabilistic Delayed Packet Discard” scheme.

The PDPD scheme is a rather optimistic algorithm which permits some new packets to enter the buffer even when a queue length is larger than the threshold. In terms of implementation,

it is clear that PDPD is slightly more complex than PPD and EPD. Despite this, however, the advantage PDPD offers is that it is still very simple. As in the EPD, the switch has to monitor the queue length. Thus, the decision of the flag setting in cases when the queue length is larger than the threshold can be calculated beforehand using this monitoring duration with one more instruction. The author has examined the PDPD scheme by conducting a simulation.

#### 4.4 Simulation and results

Figure 4.16 is a simple model of the simulation. 20 unidirectional TCP sources (from TCPS 1 to TCPS 20) send packets to 20 TCP receivers (from TCPR1 to TCPR 20) via two ATM back-to-back switches ceaselessly. Application data arrives deterministically to TCP sources with a short interval and all the TCP sources always have data to transmit. Each link is assumed to

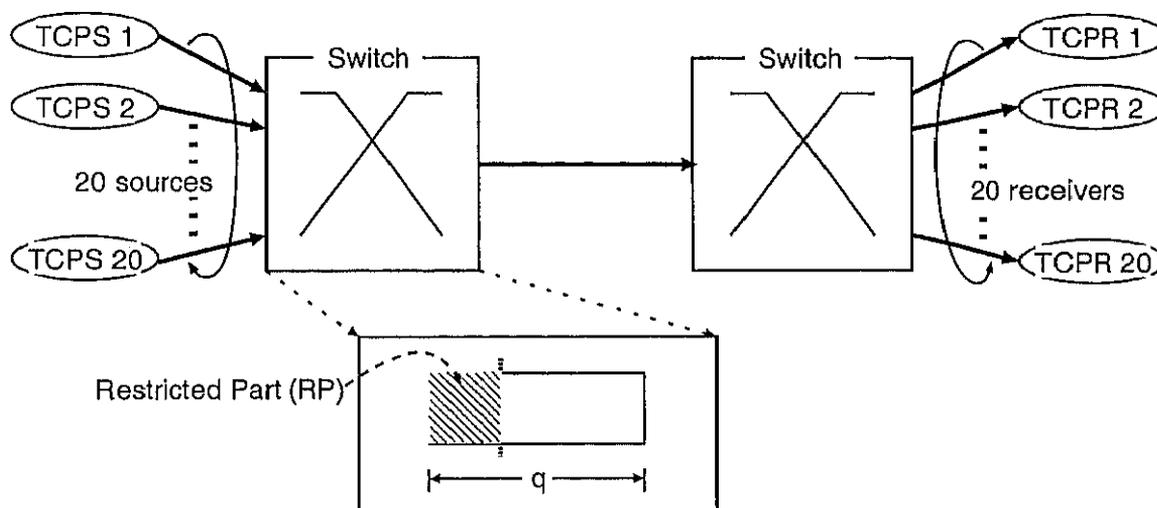


Figure 4.16: Simulation model of Probabilistic Delayed Packet Discard scheme.

be an OC-1 link. Thus, the capacity of each link is approximately 48 Mbit/s. The distance between a source and destination is 4 km. A maximum segment size (MSS) of a TCP packet is 9140 octets, which corresponds to 192 ATM cells. A user buffer size of each TCP source is 64 Kbytes. This comes from the maximum size of the congestion window (64 Kbytes), which can hold seven TCP packets<sup>4</sup>. The TCP sources send data constantly to the receivers through a bottleneck link, L2, between two back-to-back switches. The TCP timer granularity is 500 milliseconds and congestion control policies of the TCP Reno version are also performed in the TCP level. All packets from these TCP sources are fragmented into AAL5 format ATM cells. These segmented cells are transmitted over a UBR (Unspecified Bit Rate). Then the author examines this model for various buffer sizes ( $q$ ) and various probabilities ( $p$ ), which vary from 0.0 to 1.0 in increments of 0.25. As previously described, when  $p$  is 0.0 or 1.0, it corresponds

<sup>4</sup>64 Kbytes / 9140 [octet/packet] = 7.17 packets.

to PPD or EPD, respectively. Buffer size  $q$  varies from 750 cells to 4000 cells. The RP size  $b$  is also a variable. The confidence level is 95 % and the confidence interval is 2 %.

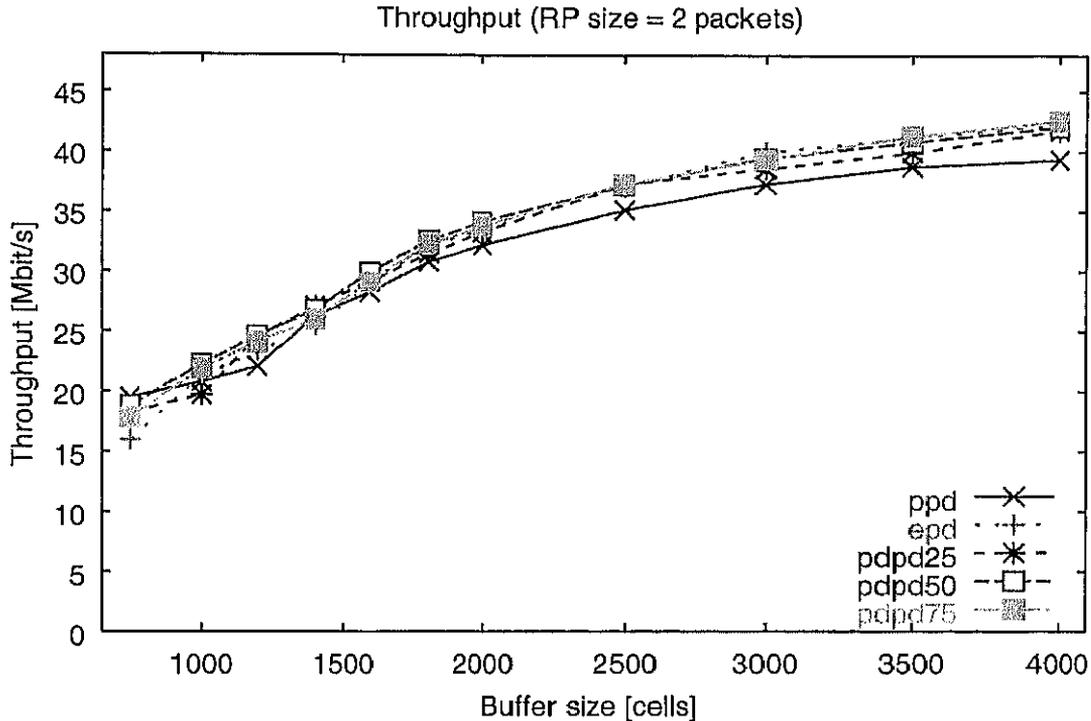


Figure 4.17: Total throughput of TCP sources when RP size is two packets.

Figure 4.17 is the throughput when the RP size  $b$  is two packets. In this figure, throughput is increased as the buffer size increases. The EPD does not achieve the highest throughput for all buffer ranges. Thus, it can be said that the PDPD can improve the throughput higher than the EPD scheme.

The normalized throughput of each parameter of probability  $p$  is shown in Figure 4.18. The throughput of EPD is 1. In the Figure 4.18, PDPD can achieve at most 15 % higher throughput than EPD in small buffer sizes. Particularly, if the probability  $p$  is 0.5 or 0.75 the throughput is consistently higher than that of the EPD in small buffer range. They (*pdpd50* and *pdpd75*) achieve as high a throughput as EPD for larger buffer sizes as well. For the probability  $p$  of 0.0 (i.e., PPD) or 0.25, PDPD  $p$  allows more packets to be buffered, which might consequently bring more packet losses.

The throughput decrement in small buffer ranges can be caused mainly by: 1) aggressiveness of packet discard, or 2) insufficient buffer space compared to the number of incoming packets. The effects of these two causes are reduced as the buffer size is increased. When the buffer size is 750 cells, aggressive packet discard by EPD achieves too poor throughput. However, as the buffer size is increased to 1000 cells, the effect of this aggressiveness is slightly neutralized. Thus, the throughput of EPD is improved much compared to that when the buffer size is 750

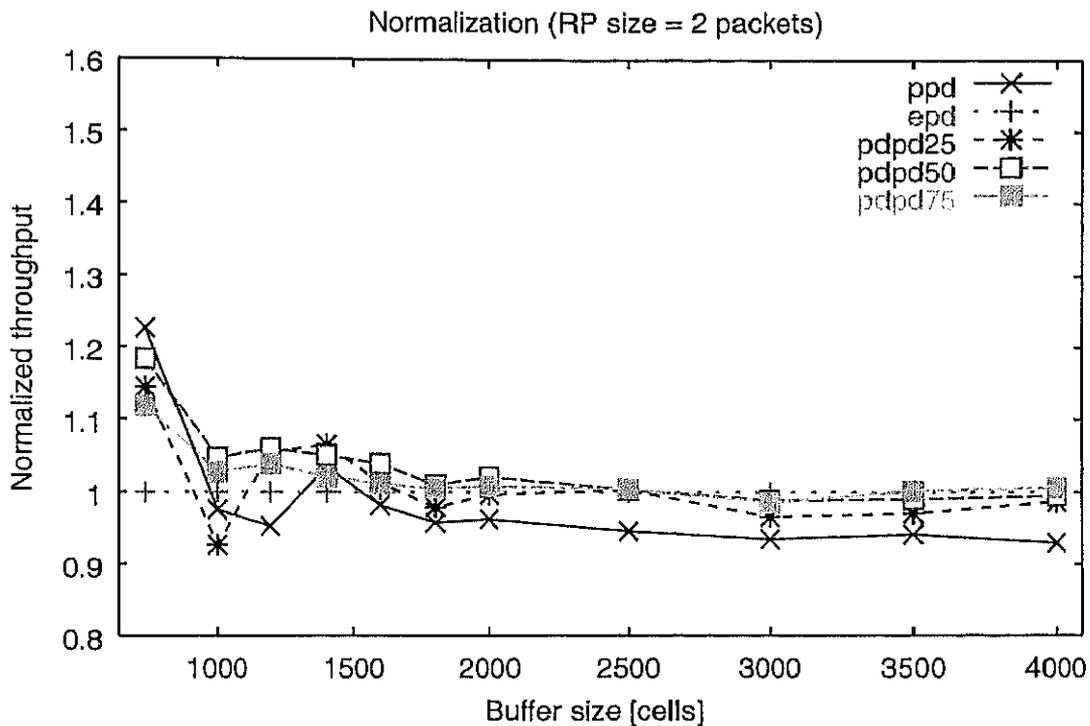


Figure 4.18: Normalized throughput when RP size is two packets.

cells. Compared to this, when the buffer size is 1000 cells, throughput of *pdpd25* and PPD is not so much increased as the buffer size increases. This is because the second reason, insufficiency of buffer space, is still effective for *pdpd25* and PPD. This is the reason why the curves of normalized throughput of *pdpd25* and PPD in Figure 4.18 seem decreased compared to that of EPD. Nevertheless, the throughput with *pdpd25* or PPD is still increasing in that case.

Figures 4.19 through 4.23 show the snapshots of buffer occupation in an ATM switch after some warming up duration. Each value is marked periodically and the threshold is approximately (*buffer size* - 400) cells. The buffer size is 1000 cells (top), 2000 cells, or 3000 (bottom) cells.

In each figure, the horizontal line identifies the average queue length. It can be seen from the figures that the RP (restricted part) is utilized more frequently for smaller  $p$  probabilities. On the other hand, buffer overflow occurs more frequently when the probability  $p$  is small. Based on

Table 4.1: The comparison of average queue lengths when RP size is two packets.

	PDPD0	PDPD25	PDPD50	PDPD75	PDPD100
1000	.57911	.38359	.38776	.35143	.32919
2000	.59109	.51312	.49174	.48904	.47045
3000	.60344	.61169	.58833	.59462	.59205

these figures, Table 4.1 shows an instance of buffer utilization when the RP size is two packets.

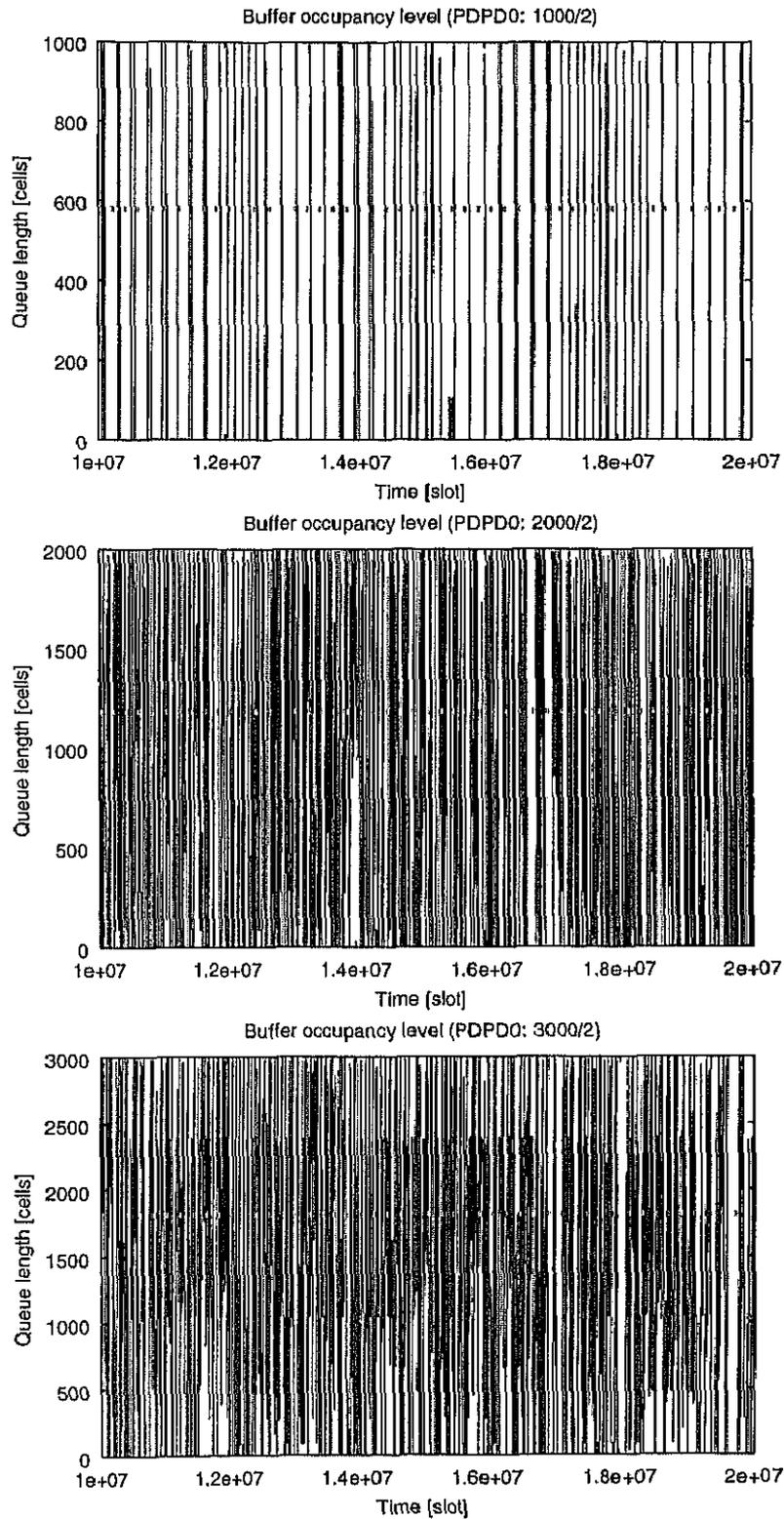


Figure 4.19: Snapshots of buffer occupation in the switch when the probability  $p$  is 0 (i.e., PPD).

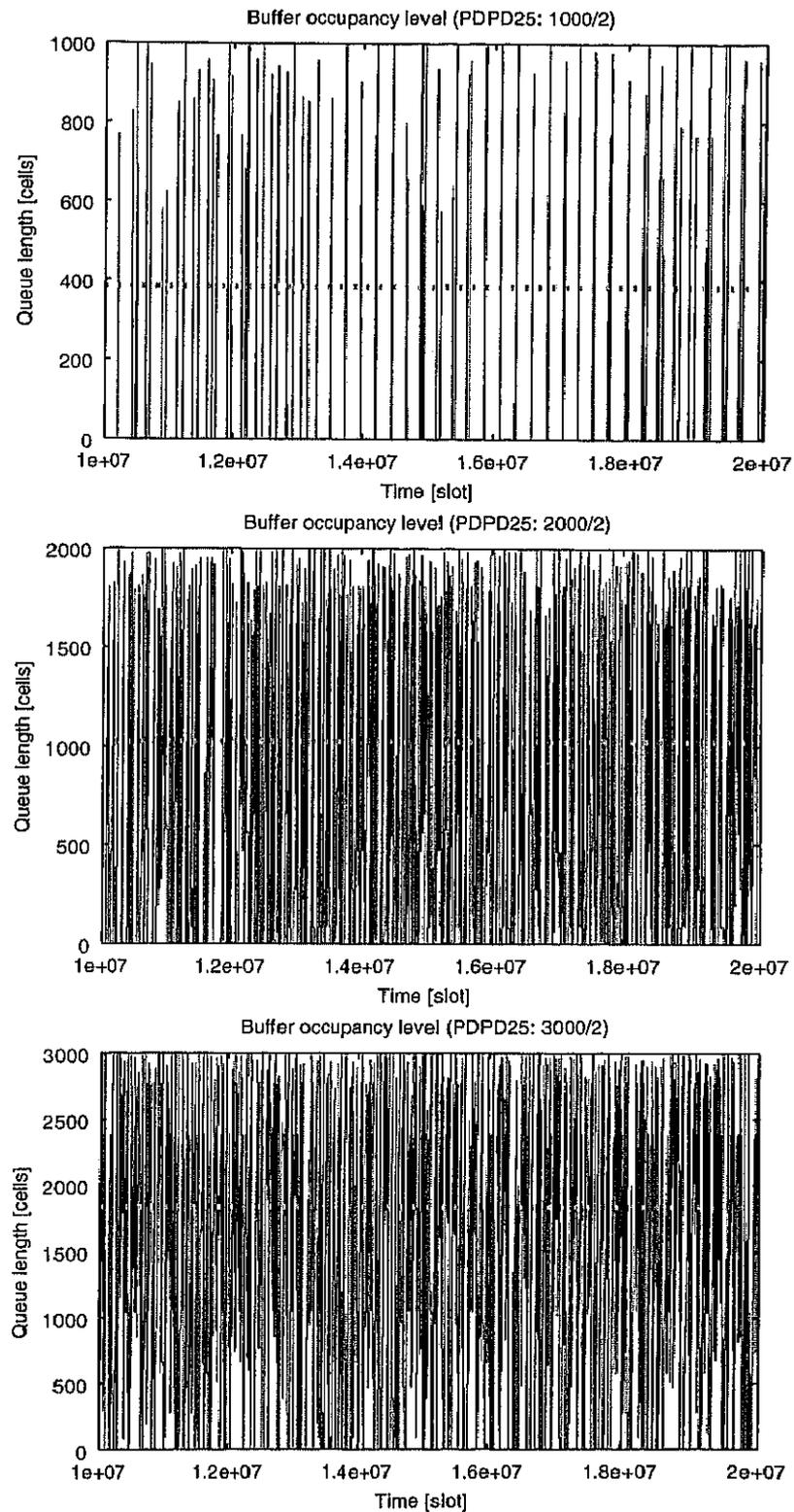


Figure 4.20: Snapshots of buffer occupation in the switch when the probability  $p$  is 0.25.

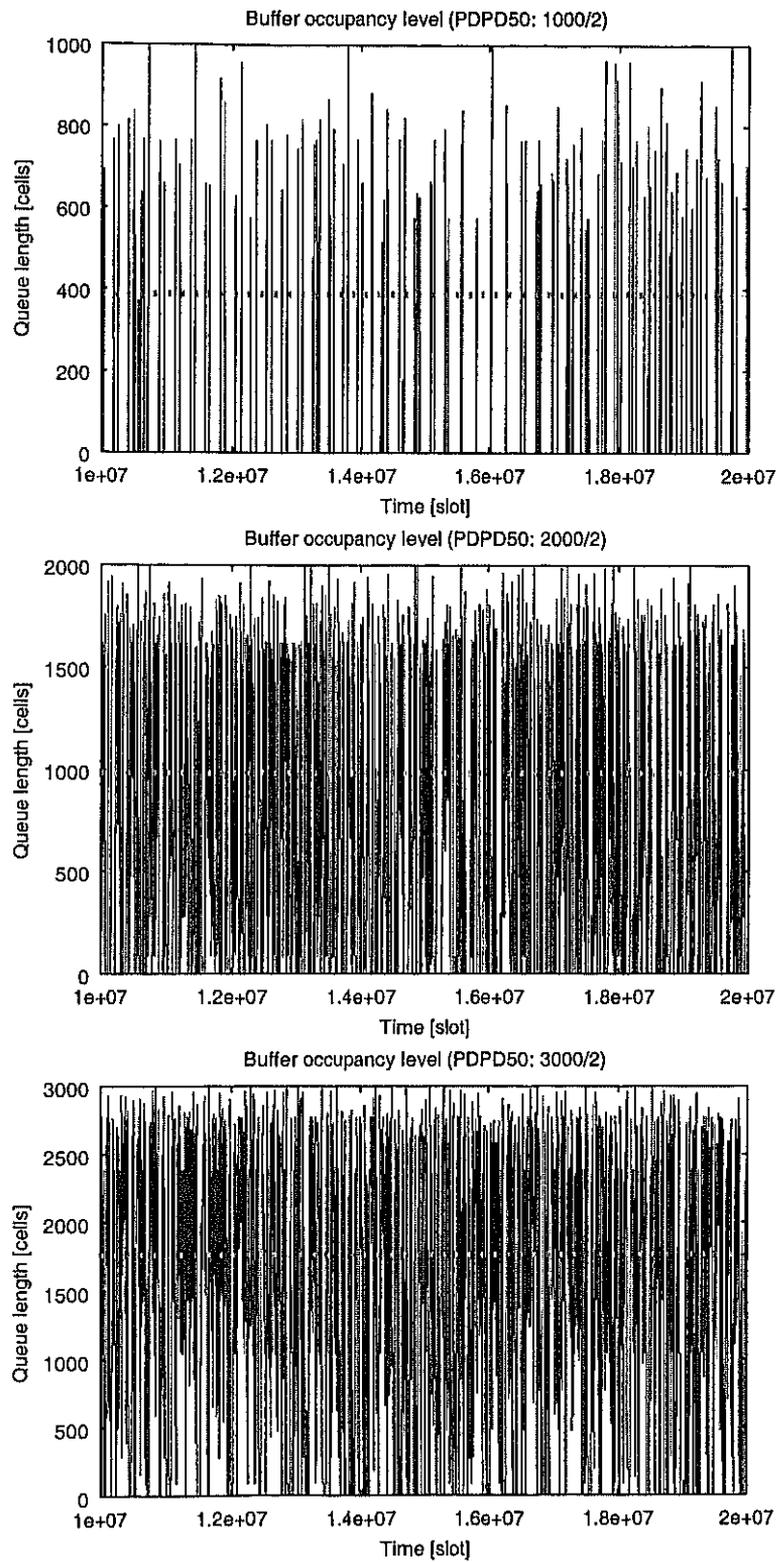


Figure 4.21: Snapshots of buffer occupation in the switch when the probability  $p$  is 0.5.

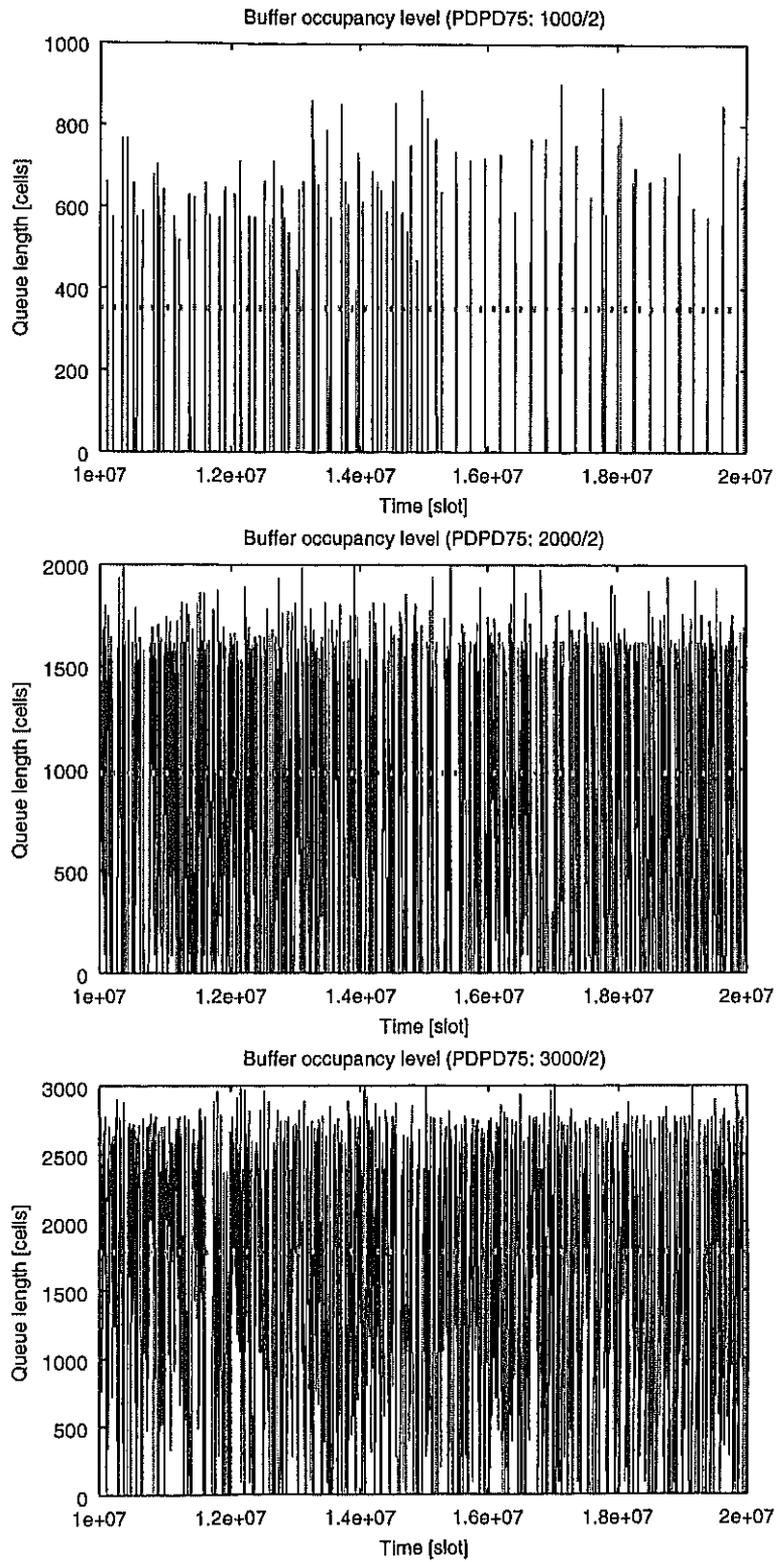


Figure 4.22: Snapshots of buffer occupation in the switch when the probability  $p$  is 0.75.

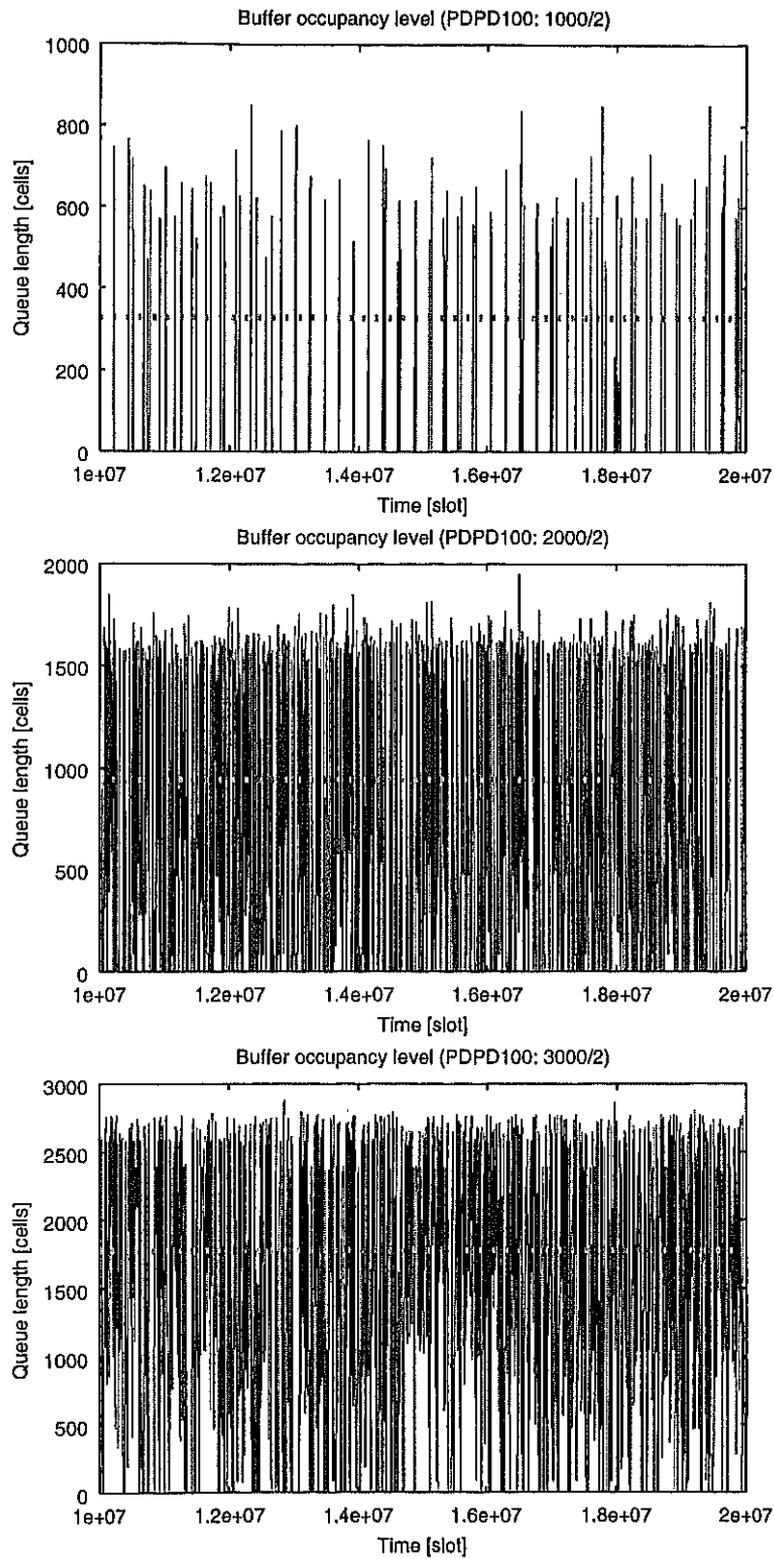


Figure 4.23: Snapshots of buffer occupation in the switch when the probability  $p$  is 1.0 (i.e., EPD).

Each value in this table is the ratio of the average queue length to the buffer size. When the buffer size is small, any kind of regulation makes the buffer utilization limited. Therefore, all values except PDPD0 are less than 50 %. As the buffer size grows, the buffer utilization is also increased for all PDPDs.

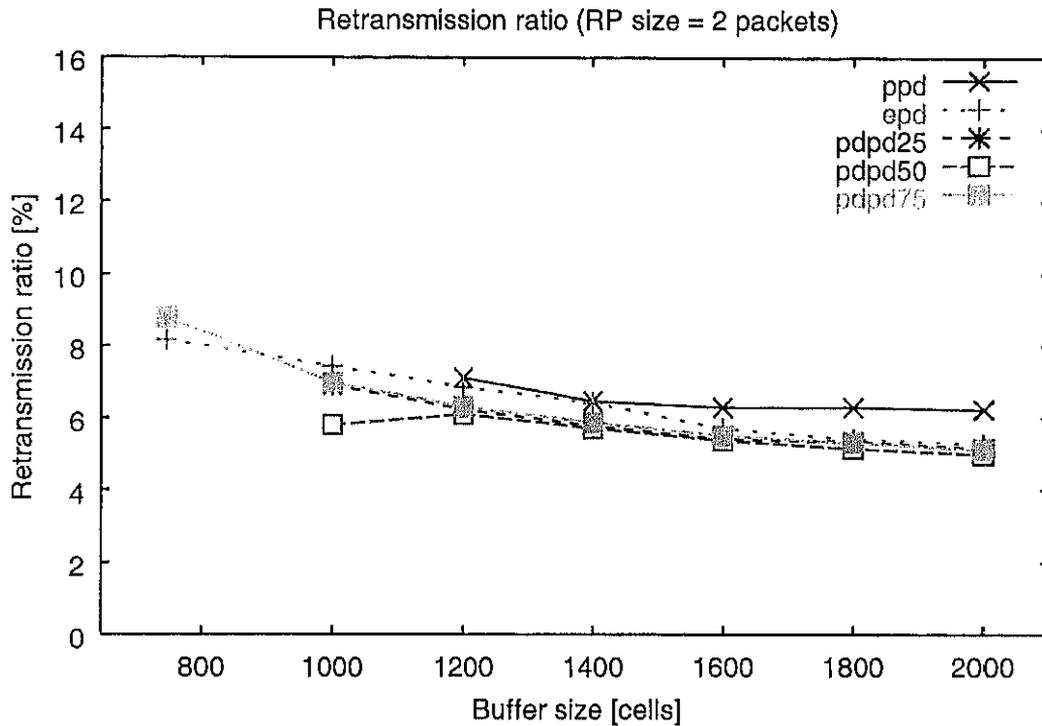


Figure 4.24: Retransmission ratio when RP is two packets.

Figure 4.24 presents the retransmission ratio for small buffer ranges when the RP size  $b$  is two packets. The confidence level is 95 % and the confidence interval is at most 3.5 %. All other results not shown in this figure are those whose margins are above this confidence interval. In Figure 4.24, the retransmission ratios of PPD or EPD are higher than those of almost all PDPDs. This is because, for PPD, excessive buffering incurs buffer overflow. On the other hand, excessive packet discard occurs in EPD. Those features are the reason for the higher retransmission ratio. Therefore, it can be said that PDPD achieves lower retransmission ratio by using appropriate packet discarding as well as appropriate packet buffering. When the buffer size is 750 cells, the retransmission ratio of *pdpd75* is slightly higher than that of EPD. This is the specific case in which the aggressive packet discard by EPD achieves a lower retransmission ratio. By *pdpd75*, because the PDPD allows more packets to be buffered than EPD does, it results in more frequent buffer overflow, hence occasioning more retransmissions. For all methods, the retransmission ratio decreases as the buffer size increases. Moreover, as the buffer size grows, the difference of the retransmission ratio is decreased.

The effect of the aggressiveness of packet discard by EPD is more magnificent when the RP size is large. Figure 4.25 shows the throughput of PDPDs when the RP size  $b$  is three

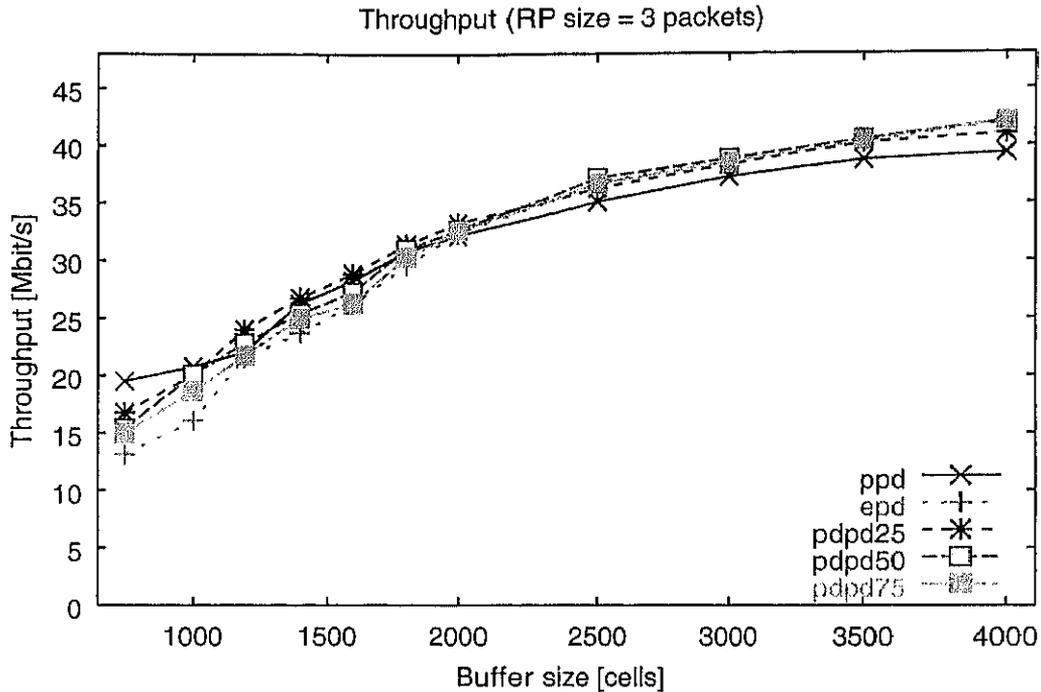


Figure 4.25: Total throughput of TCP sources when RP size is three packets.

packets. Here, again, the throughput is increased as the buffer size increases. When the RP size is larger, the aggressiveness of packet discard in EPD exerts a more negative effect on network throughput. Because of this aggressiveness of EPD, the EPD achieves poorer throughput than PDPD for small buffer sizes. Thus, the PDPD can improve the throughput. Figure 4.26 shows the network throughput with PDPD normalized by those with EPD. In the case of a relatively large RP size, PDPD with more optimistic and speculative probability  $p$  can achieve higher throughput. Quantitatively, the throughput is improved at most 32 % by PDPD25 with a relatively small  $p$  for a small buffer range. In other words, small  $p$  is more advantageous than large  $p$  when the RP size is relatively large. This is why the normalized throughput of probability  $p$  of 0.25 (*pdpd25*) and 0.5 (*pdpd50*) are consistently higher than that of EPD.

With regard to the PPD, as the buffer size is increased, the throughput is also improved. However, enough regulation is not made to the streams by the PPD and it results in a lower throughput compared to others. Therefore, this author concludes that the PDPD scheme can improve network throughput more effectively than EPD by adapting an appropriate probability  $p$ . In Figure 4.26 again, when the buffer size is 1000 cells, aggressive packet discard by EPD achieves poor throughput. However, as the buffer size increases to 1200 cells, the effect of the aggressiveness is decreased, which achieves considerable improvement of throughput compared

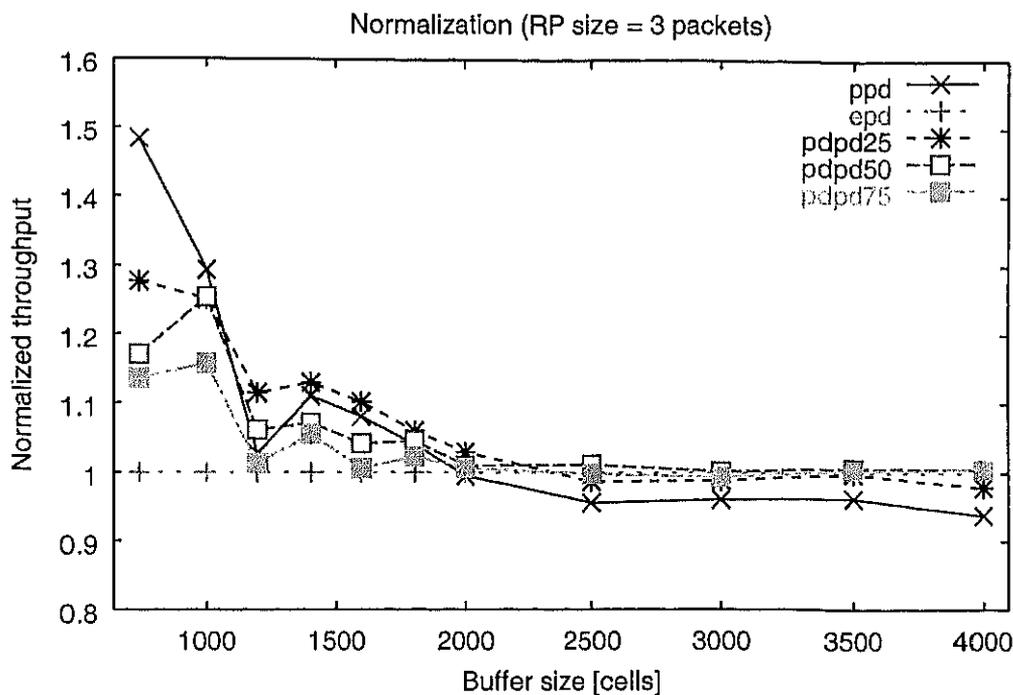


Figure 4.26: Normalized throughput when RP size is three packets.

to that when the buffer size is 1000 cells. Compared to this, even though the throughput of *pdpd25* and PPD is increasing, they are not increased so much when the buffer size is 1200 cells. This is why the curves of normalized throughput of PDPDs and PPD in Figure 4.26 seem decreased compared to that of EPD.

Figures 4.27 to 4.31 show the snapshots of buffer occupation in an ATM switch after some warming up duration. The threshold is approximately (*buffer size* - 600) cells. The buffer size is 1000 cells (top), 2000 cells, or 3000 (bottom) cells. Horizontal lines identify the average queue length. In these figures, again, the RP (restricted part) is more utilized for a smaller  $p$ .

Table 4.2 is an instance of average queue lengths of buffers in a switch when the RP size is three packets. It shows the ratio of the average queue length to the buffer size for different

Table 4.2: The percentile average queue lengths when RP size is three packets.

	PDPD0	PDPD25	PDPD50	PDPD75	PDPD100
1000	.57911	.32204	.27548	.24477	.25251
2000	.59109	.45690	.40778	.39094	.39007
3000	.60344	.57668	.57550	.54919	.57102

buffer sizes and different  $p$  probabilities. In this table, any kind of methods achieve quite poor buffer utilization when the buffer size is small. Thus, it can be said that the input should not be

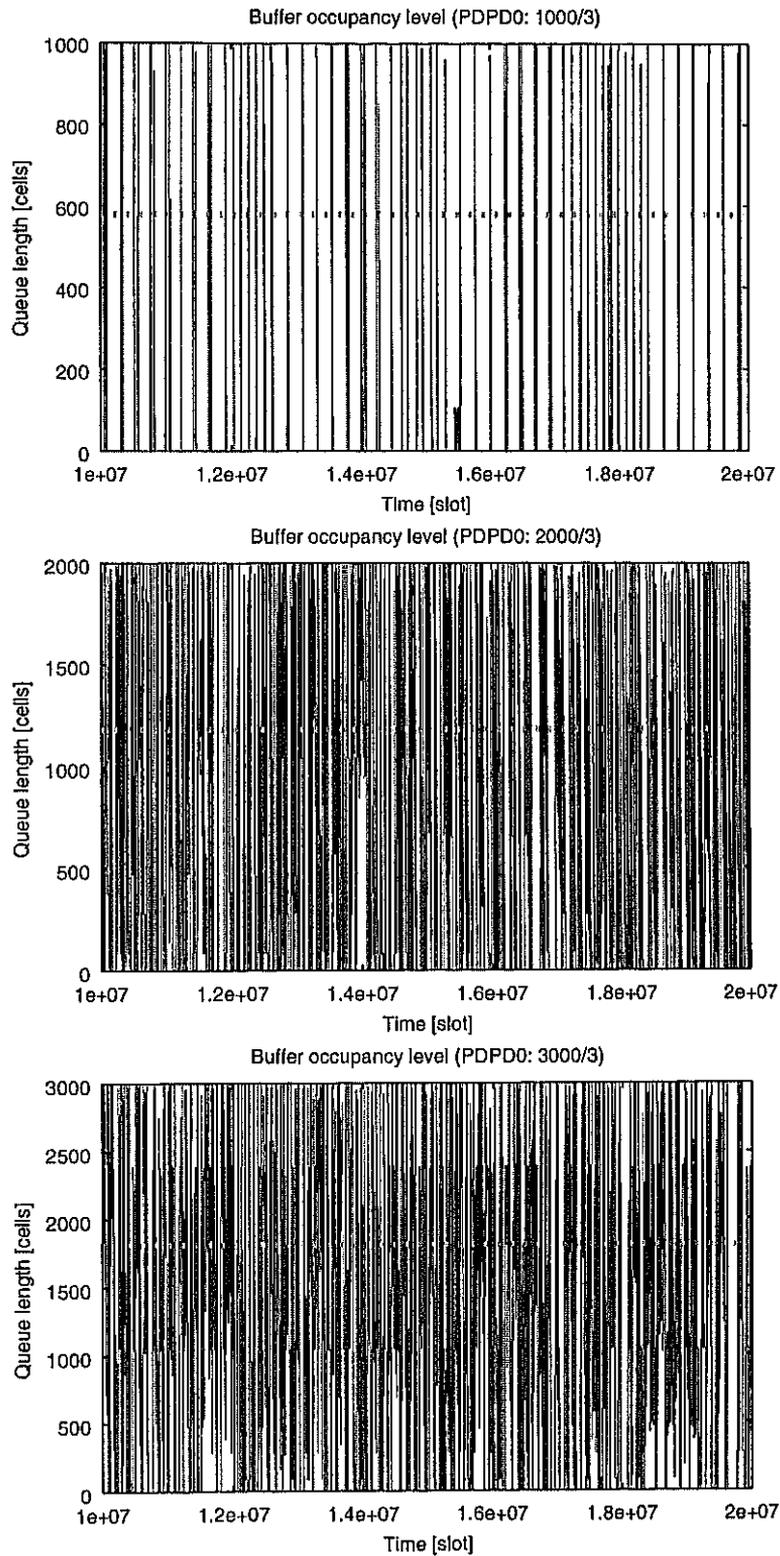


Figure 4.27: Snapshots of buffer occupation in switch when the probability  $p$  is 0 (i.e., PPD).

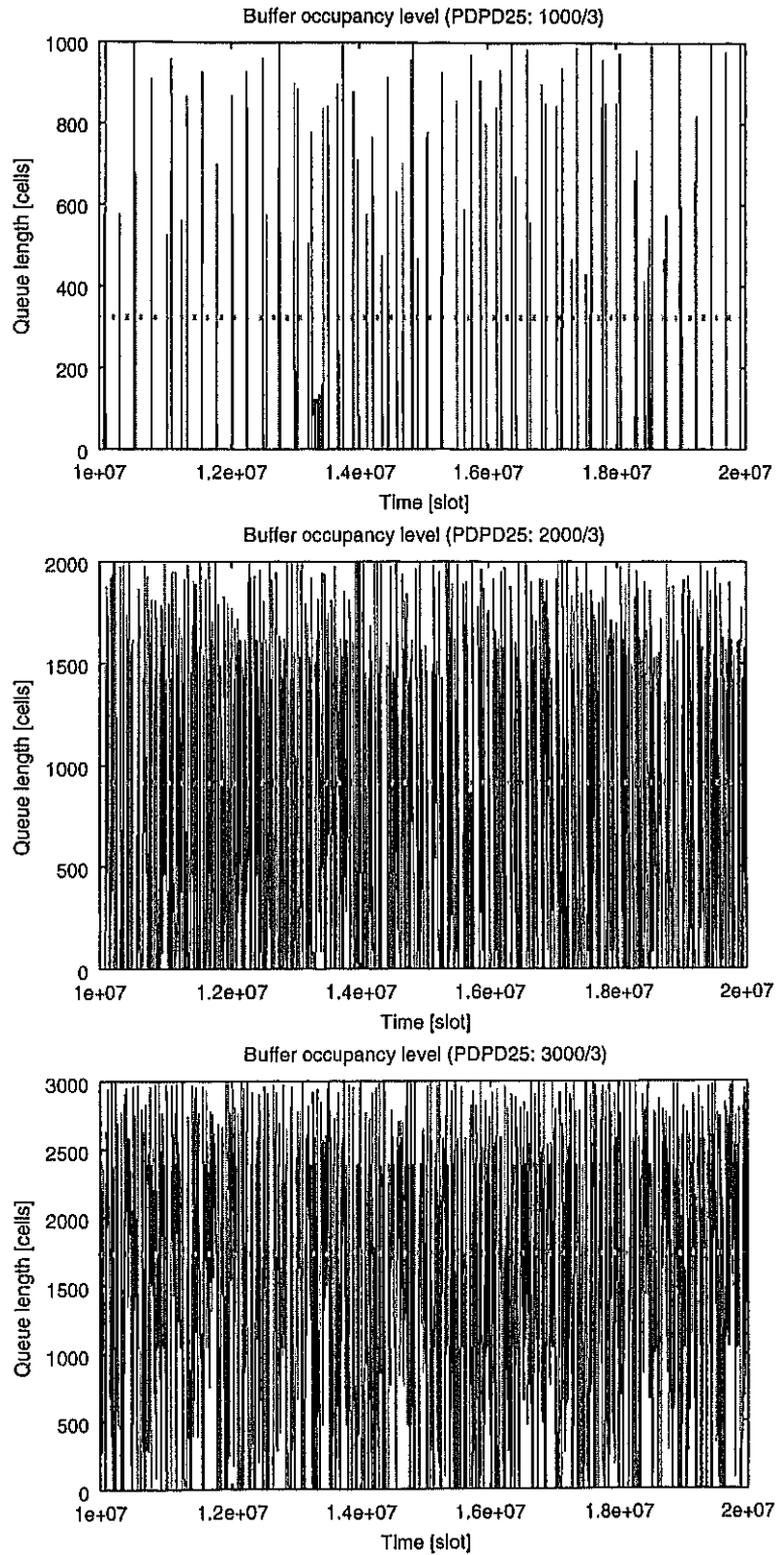


Figure 4.28: Snapshots of buffer occupation in the switch when the probability  $p$  is 0.25.

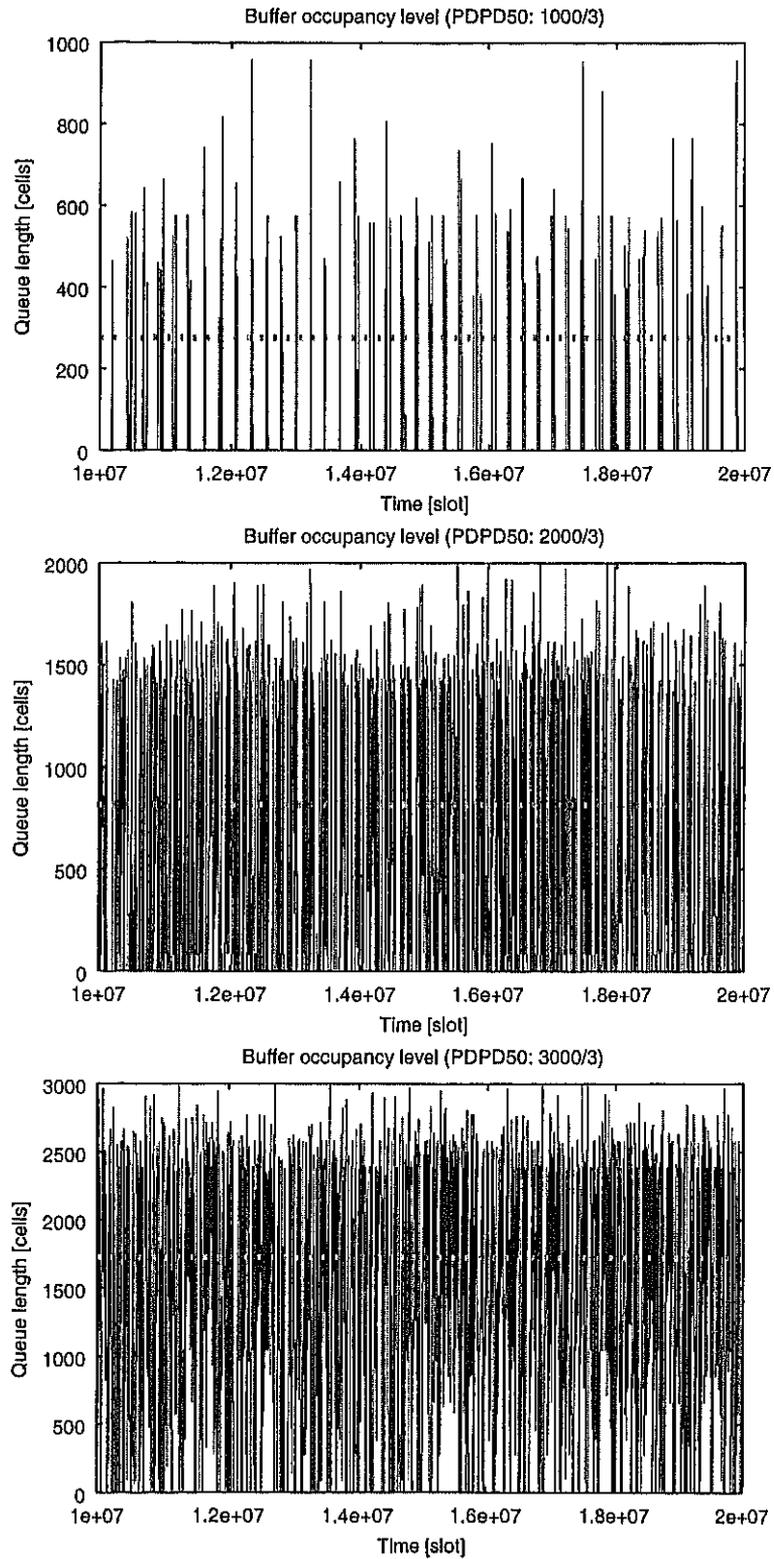


Figure 4.29: Snapshots of buffer occupation in the switch when the probability  $p$  is 0.5.

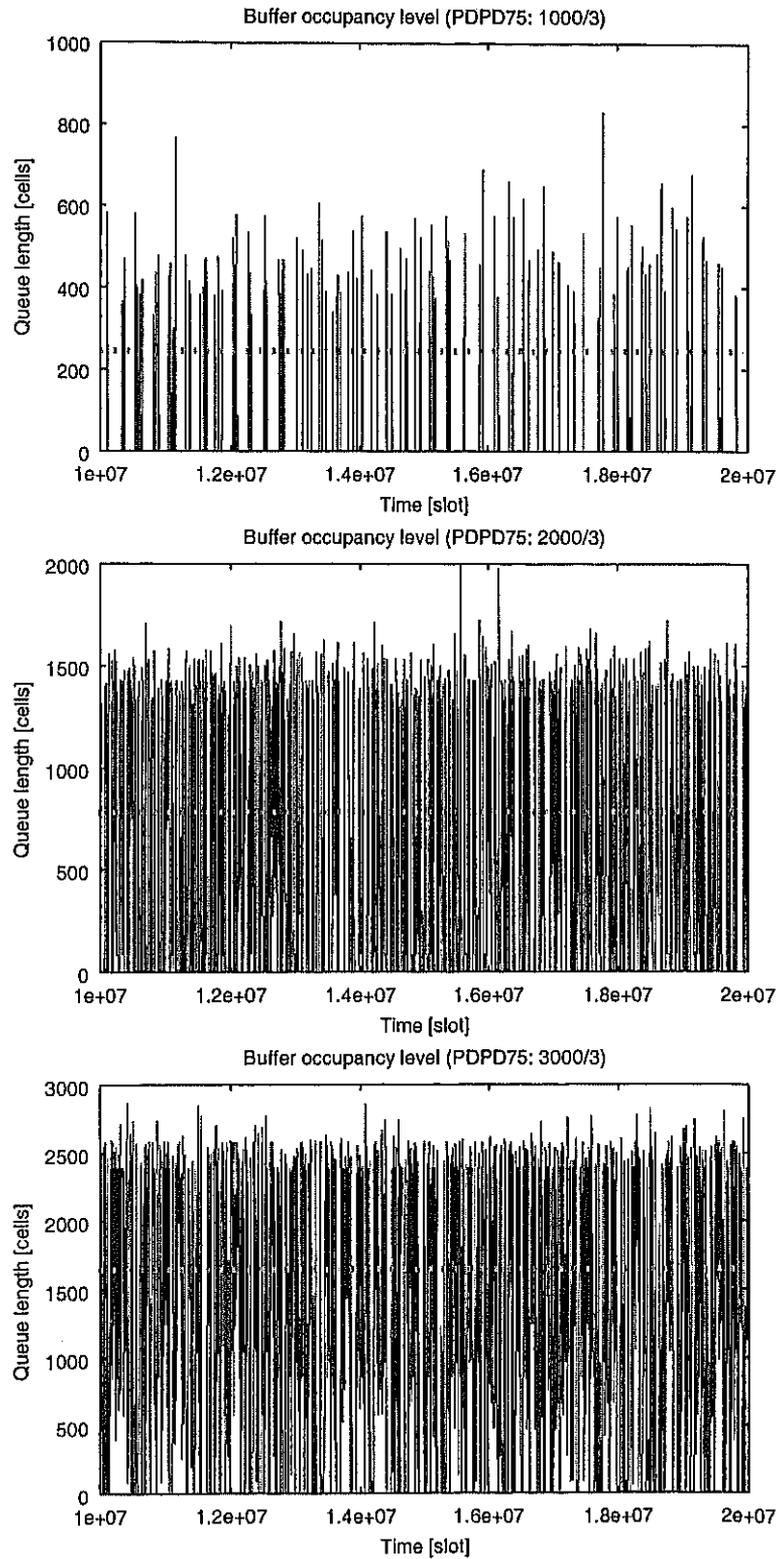


Figure 4.30: Snapshots of buffer occupation in the switch when the probability  $p$  is 0.75.

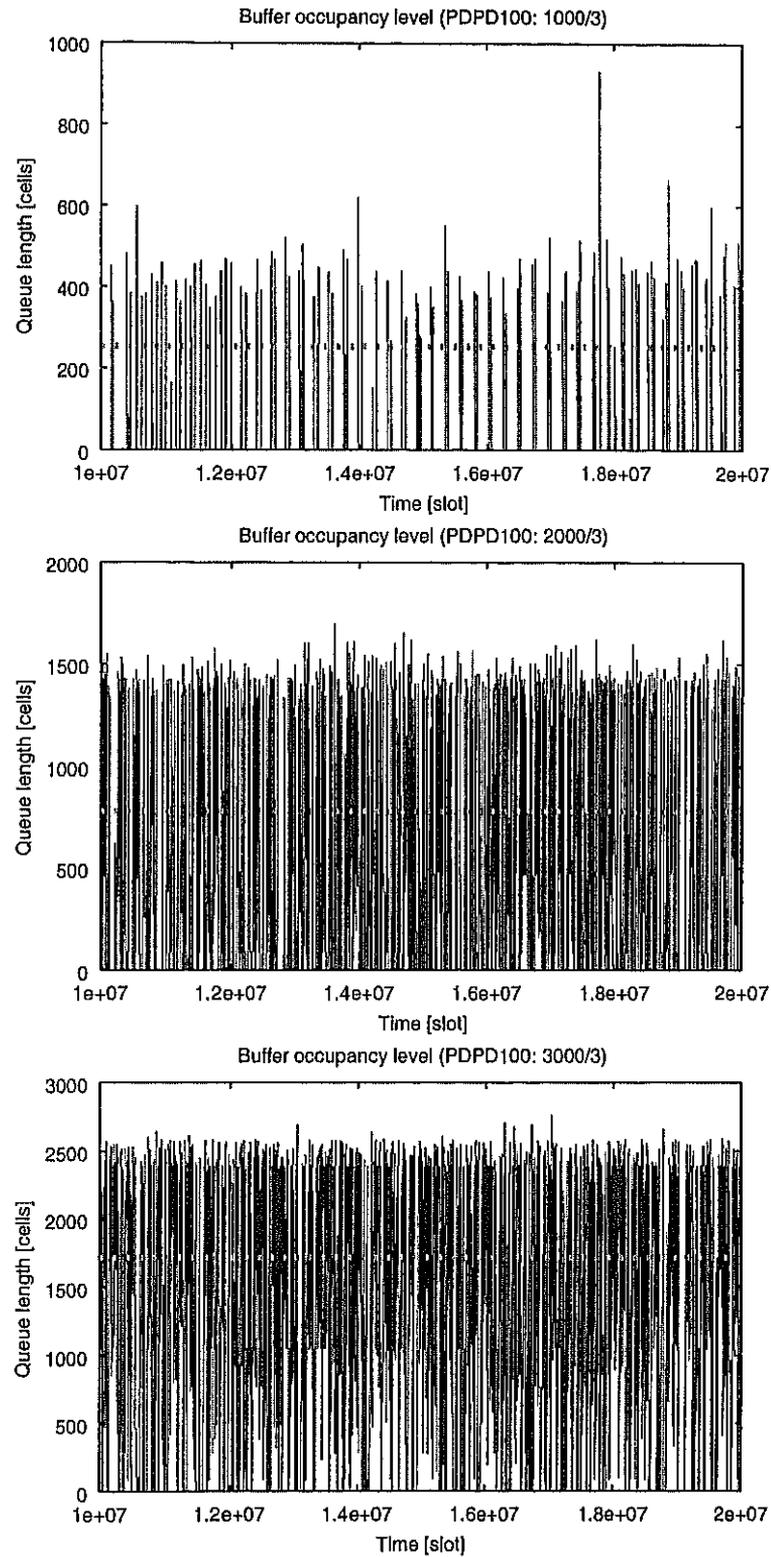


Figure 4.31: Snapshots of buffer occupation in the switch when the probability  $p$  is 1.0 (i.e., EPD).

regulated much for small buffer sizes in order to increase buffer utilization. The buffer utilization of all methods is increased as the buffer size increases. Even though the buffer utilization is increased, 40 % of the buffer space is still not fully utilized.

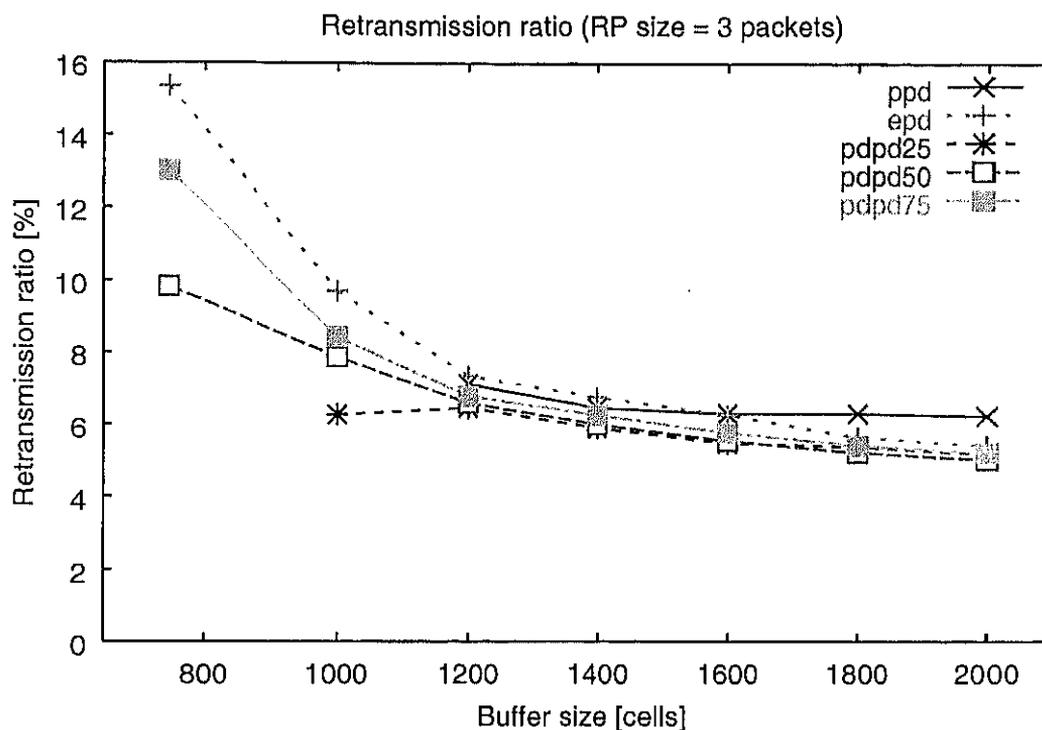


Figure 4.32: Retransmission ratio when RP is three packets.

Figure 4.32 is the retransmission ratio for small buffer ranges when the RP size  $b$  is three packets. The confidence level is 95 % and the confidence interval is at most 3.1 %. Also in Figure 4.32, the retransmission ratios of PPD and EPD are higher than those of PDPDs. Specifically, when the RP size is relatively large, the aggressiveness of EPD augments substantially and the retransmission ratio is considerably increased. Although this tendency is common to PDPD with a large probability  $p$ , it can be said that the PDPD can improve the retransmission ratio compared to those of EPD. When the buffer size grows from 1000 cells to 1200 cells, the retransmission ratio of *pdpd25* is increased. This is because, as the buffer size increases, TCP senders try to transmit more packets, which brings about a slightly increased retransmission ratio. Also in Figure 4.32, the retransmission ratio of all methods decreases as the buffer size increases.

This author experimented with PDPD schemes in another simulation environment. Figure

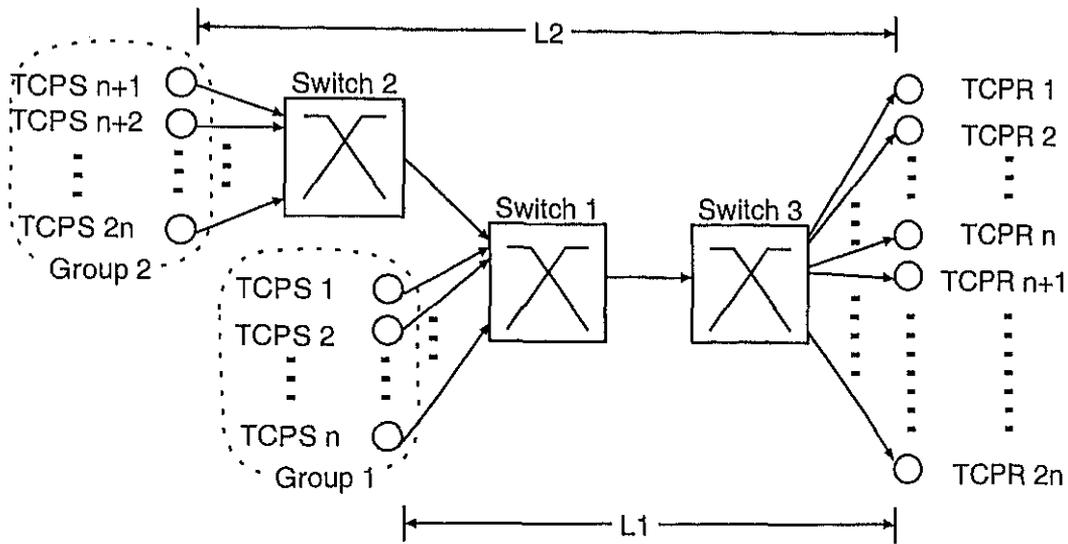


Figure 4.33: Simulation model of multistage switches.

4.33 illustrates the simulation model. Unidirectional TCP sources (Group 1 from TCPS 1 to TCPS  $n$ ) send packets to TCP receivers (from TCPR 1 to TCPR  $n$ ) through two ATM switches (Switch 1 and 3). Other unidirectional TCP sources (Group 2) send packets to other TCP receivers (from TCPR  $n + 1$  to TCPR  $2n$ ) via three ATM switches (Switch 2, 1, and 3). Thus, Switch 1 has to control the tributaries from Switch 2 and from Group 2 TCP sources. Application data arrives deterministically to TCP sources with a short interval and, hence, all the TCP sources always have some data to transmit. Each link is assumed to be an OC-1 link again, thereby the capacity of each link is approximately 48 Mbit/s. The RP size  $b$  is fixed to two packets in this experiment.

The Figure 4.34 shows the throughput and normalized throughput when the number of sources in each group,  $n$ , is five. The length of link L1 from Group 1 to the group of receivers is approximately 4 km and the other of link L2 from group 2 to the receivers is approximately 6 km. The confidence level is 95 % and the confidence interval is 1.5 %. When the buffer size is relatively small, the PPD achieves the highest throughput while the EPD achieves the lowest. This is because as the  $n$  is relatively small, it is not necessary to regulate the input streams too strictly in order to achieve higher network throughput. For small buffer sizes in this Figure 4.34, the throughput of the PDPD  $p$  is decreased as the probability  $p$  increases. Thus, it can be concluded that when the number of sources is relatively small and network congestion is not expected to occur so frequently, small  $p$  is more advantageous than larger  $p$  or EPD.

Next, the throughput and normalized throughput of the PDPD when  $n$ , the number of sources in a group is ten is shown in the Figure 4.35. The confidence level is 95 % and the confidence interval is 1.5 %. In Figure 4.35, the EPD does not achieve the highest throughput

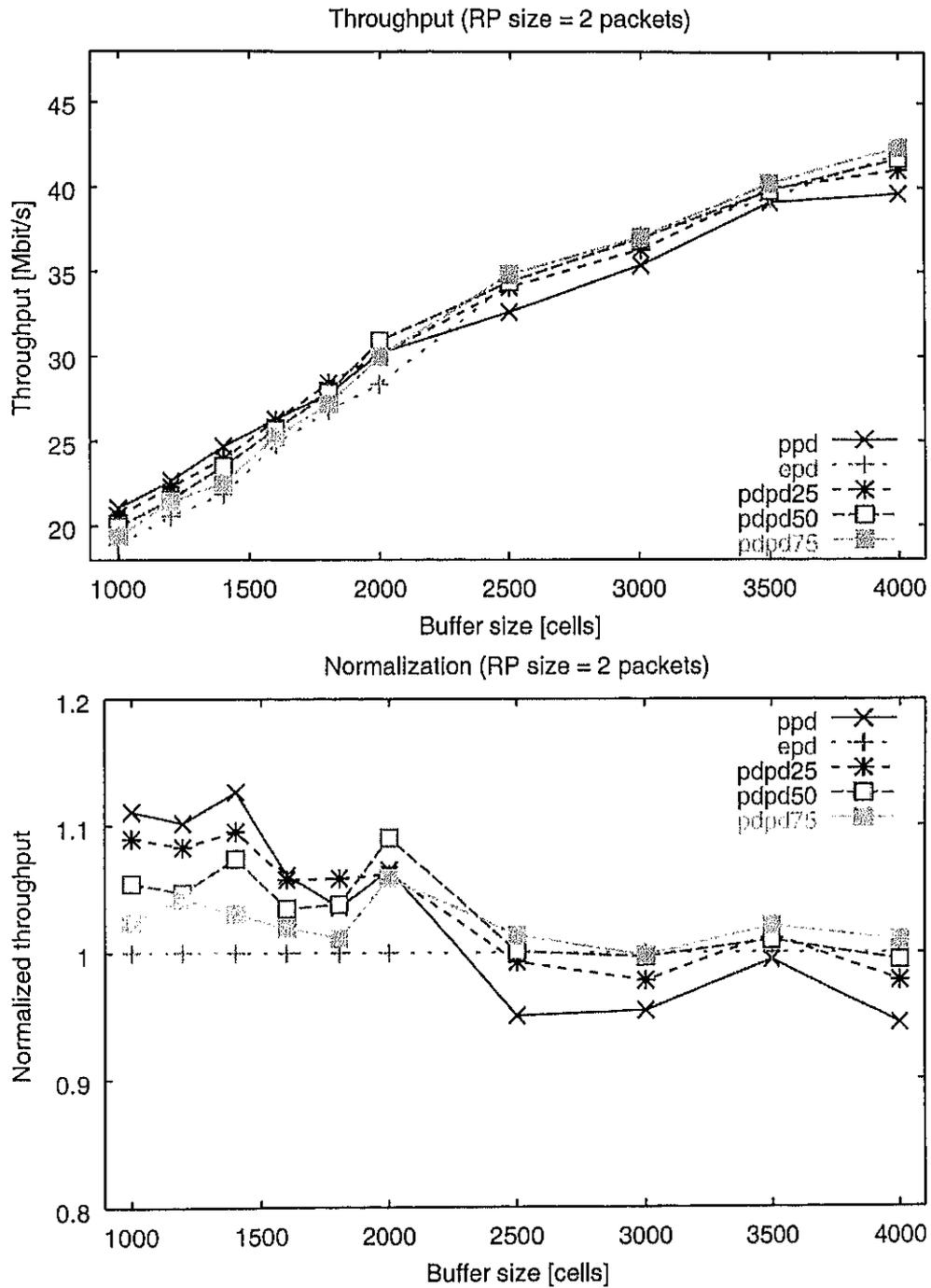


Figure 4.34: Throughput and normalized throughput with PDPD when  $n$  is five.

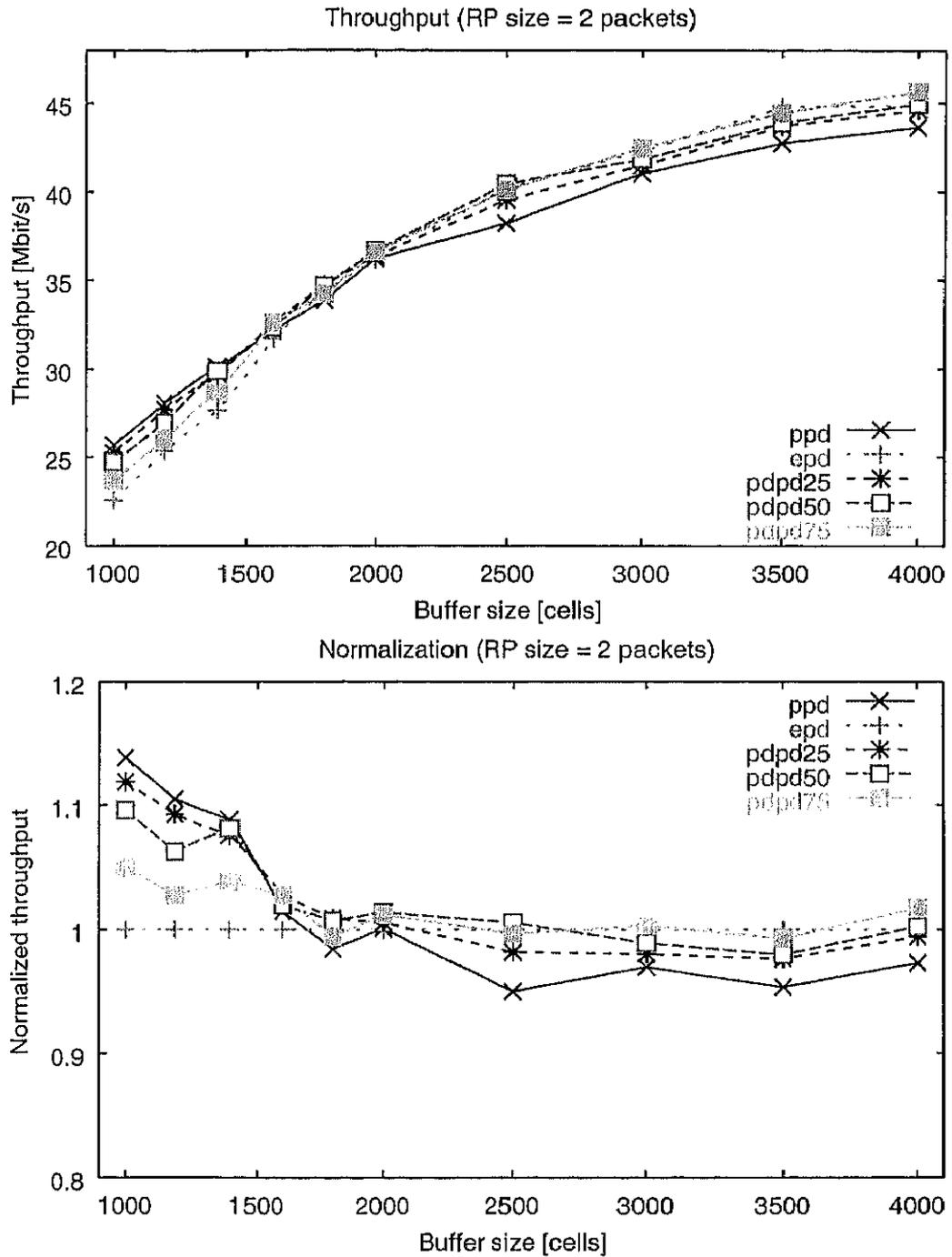


Figure 4.35: Throughput and normalized throughput with PDPD when  $n$  is ten.

for all buffer ranges as in the previous experiments. As can be seen, a smaller probability  $p$  is more advantageous than the larger  $p$  for small buffer sizes. However, as the buffer size grows enough, PDPD with small  $p$  achieves lower throughput than EPD. Comparing to the EPD, when  $p$  is 0.75 (*pdpd75*), the PDPD can improve the network throughput compared to that of EPD in almost all cases.

The next two figures (Figure 4.36 and 4.37) are the throughput and normalized throughput when the length of L2 is 200 km. The number of the Group 1 is 5 or 10, respectively. The confidence level is 95 % and the confidence interval is 1.5 % at most. When the distance between the source and the receiver is long, the round trip time is also elongated. Thus a TCP source has to wait for a longer time for a timeout or the next transmission. This will result in regulation of the input stream by itself. Thus, PDPD with a large  $p$  or EPD happens to impose double regulations on the tributaries. Therefore, PDPD with a small  $p$  achieves quite high throughput for small buffer sizes, because it decreases the effect of double regulation more than those with a large  $p$ . As the buffer size increases, the effect of this double regulation is lessened and the throughput of PDPD with large  $p$  or EPD can achieve higher throughput than the others without losing an appropriate limitation.

From the results of all of the above experiments, when the buffer size is relatively small, the PDPD achieves higher throughput than the EPD. For small buffer sizes, specifically, throughput of the PDPD with a relatively small  $p$  is higher than those with a relatively large  $p$ . Thus, in cases when the buffer space is not enough, small  $p$  is more advantageous than larger  $p$  or EPD. For relatively large buffer sizes also, PDPD with a large  $p$  achieves moderate throughput. Therefore, it can be concluded that the network throughput can be improved by a more lenient and optimistic packet discard approach.

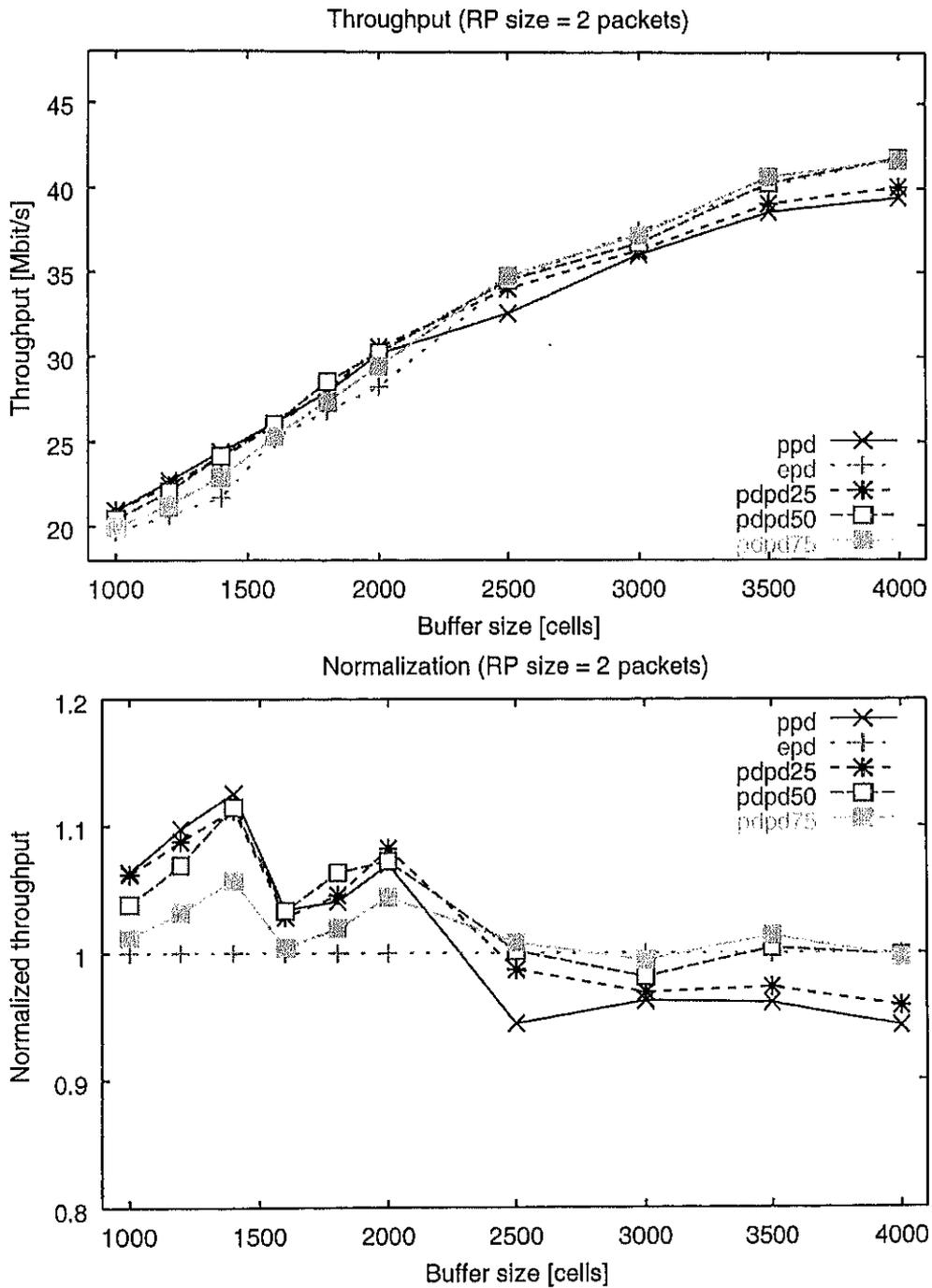


Figure 4.36: Throughput and normalized throughput with PDPD when  $n$  is five and the distance is long.

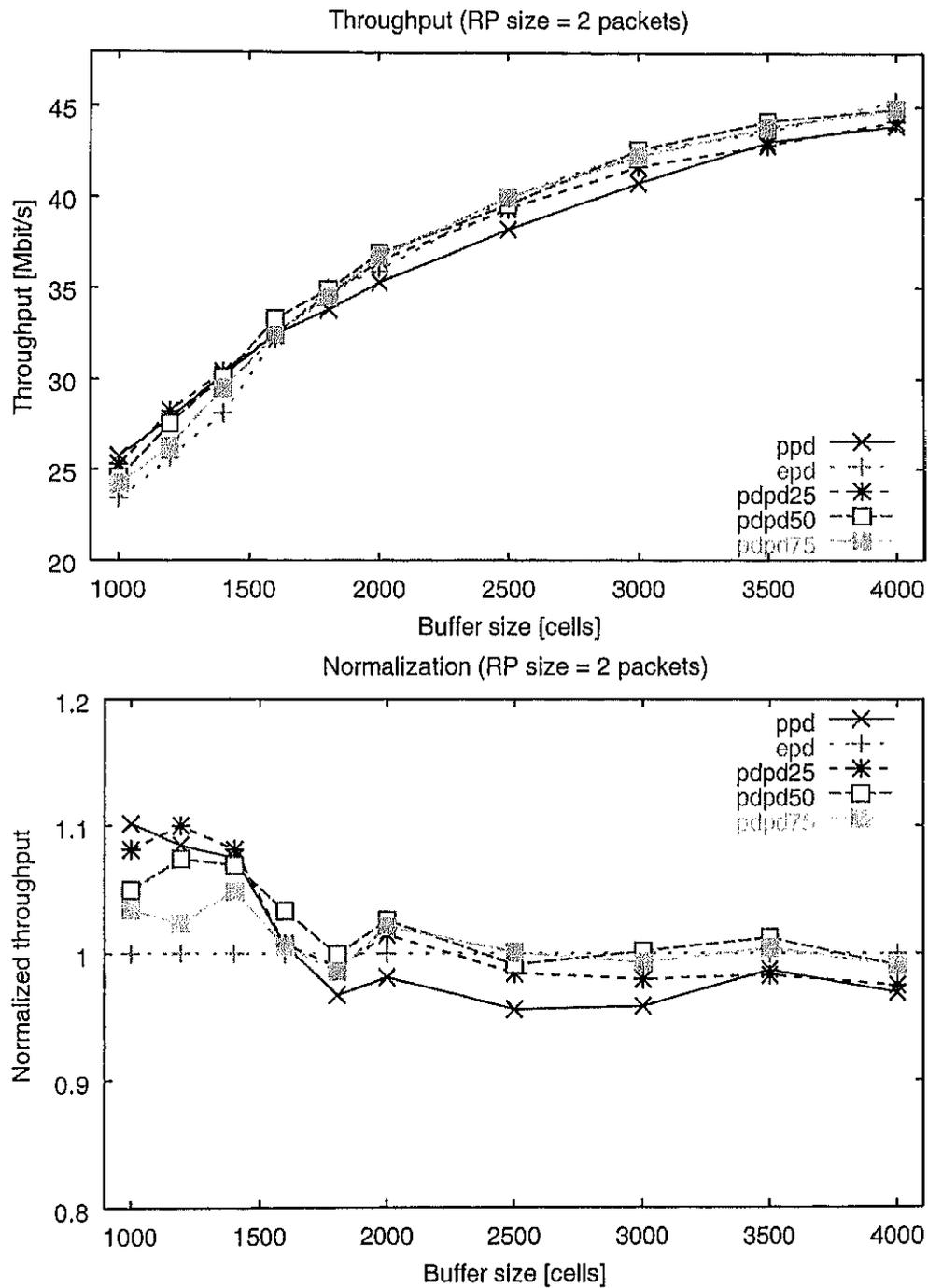


Figure 4.37: Throughput and normalized throughput with PDPD when  $n$  is ten and the distance is long.