ATM SWITCH PERFORMANCE MODELING WITH DYNAMIC BANDWIDTH ALLOCATION

Mohamed Othman, Chee Heok Loh, Shyamala D., Lily Nurliana Abdullah Department of Communication Technologies and Networks Faculty of Computer Science and Information Technology University Putra Malaysia 43400 Serdang, Selangor, Malaysia email: mothman@fsas.upm.edu.my

ABSTRACT

One new technology that can be used to integrate various types of networking services today is ATM(Asynchronous Transfer Mode) where these services demand different QoS(Quality of Service). Allocating and scheduling the bandwidth for efficient utilisation of this valuable network resource is a critical issue. Dynamic Bandwidth Allocation is a promising scheme for this efficient utilisation. This study models an ATM switch using VBR traffic sources and tests its performance for three dynamic bandwidth allocation strategies. The traffic sources are modeled using the ON-OFF traffic source model.

Keywords: ATM, ON/OFF Source Model, VBR Traffic, Bandwidth Allocation

1.0 INTRODUCTION

This study is an extension of [1] based on some of the issues highlighted below:

1) Sources are identical

A single source model which generated cells was used. In ATM networks, typically different traffic sources such as VBR(Variable bit rates), CBR(Constant bit rates) and ABR(Available bit rates) are integrated in a high speed transmission link.

- 2) Cells are generated following a particular pattern. Only one source generates cells at any particular time. Other sources will only be able to generate cells once the present source completes its cells generation process. This means that there is no overlapping of cells arrival between sources.
- 3) Traffic parameter assumptions.

Using a random number generator, cells are generated using Poisson interarrival process. At most only 4 cells are generated at a time. International standard bodies have standardized realistic traffic parameters for different traffic sources. An example of VBR traffic parameters standard is given in Table 1.

Table 1:	Parameter	Values	for Ty	pical	VBR	Traffic
	Sources, a	as Prop	osed by	CCI	TT	

VBR Traffic Sources	Mean Burst Length(Cells), N	Average Cell Arrival Rate, m
Connectionless Data	200	700 kbit/s
VBR Video	2	25 Mbit/s
Connection Oriented Data	20	25 Mbit/s
VBR Video/Data	30	21 Mbit/s

In this study, the ATM traffic source model used is a general ON/OFF traffic source model. Three bandwidth allocation strategies have been developed. Their performance has been tested and compared. Two strategies were from Lu and Hansson [2], i.e. static bandwidth allocation and bandwidth allocated proportional to expected queue length. Modification of these two strategies is used as the third strategy, where bandwidth is allocated proportional to expected queue length with threshold value.

1.1 ON/OFF Source Model



Fig. 1: ON-OFF source model

Selecting appropriate models of ATM traffic sources is an important issue, since it is closely related to the successful design and efficient performance of the ATM networks to be built [3]. Each traffic source is characterized to precisely identify its behaviour, which will provide network management with the ability to manipulate flexibly the

various services in term of connection acceptance, negotiation of the Quality of Service (QoS), congestion control, traffic enforcement and resource allocation. In ATM networks, there is a general trend to visualize cells generation as a succession of active and silent (also called idle) periods [3]. Cell generation occurs only during active periods; a group of successive cells that are not interrupted by an idle period is called a burst. The most prominent paradigm of source model exhibiting this behaviour is the ON/OFF model which is depicted in Fig. 1.

The ON/OFF source model is described by a set of parameters (p, m, β , t_{on}), where:

- p: peak arrival rate of the cells when the source is in active state (peak rate), or the maximum amount of network resource requested by the source. Also can be defined as number of cells belonging to the same connection measured during a predefined short time interval T, divided by T.
- m: average cell arrival rate or the average amount of network resources requested by the source.
- β: burstiness, defined as the ratio between the peak cell rate and the average cell rate. Can be viewed as a measure of the duration of the activity period of a connection.
- t_{on:} the average duration of the active state.

1.2 VBR Traffic

VBR traffic is expected to become a major traffic component on Broadband integrated services digital networks (BISDN) [4]. Most of the VBR sources are of ON/OFF types [5]. These types of traffic require very high bandwidth on private and public networks. Though today network layout has installed high transmission speed medium such as optic fiber, and encourage the submission of video services, current and future new services will tend to overload them. Therefore, in this study, VBR sources will be modeled using ON/OFF model.

1.3 Bandwidth Allocation

On top of congestion control, the ATM network has to find solutions to bandwidth allocation [6]. Improper bandwidth allocation will degrade network performance, low resources utilization and finally cause congestion to happen. Many suggestions to bandwidth allocation strategies have been proposed, such as Peak Rate Allocation, Equivalent Bandwidth Allocation and Dynamic Bandwidth Allocation. This study concentrates on Dynamic Bandwidth Allocation. Three dynamic bandwidth allocation strategies will be modeled and simulated. There are Static Bandwidth Allocation, Bandwidth Allocated Proportional to Expected Queue Length, and Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value. Research for Static Bandwidth Allocation and Bandwidth Allocated Proportional to Expected Queue Length are carried out. Bandwidth Allocated Proportional to Expected Queue Length has better performance than Static Bandwidth Allocation [2]. The purpose of the current work is to test the performance of the last strategy where bandwidth is allocated proportional to the expected queue length with a threshold value in the buffer, and compare it with the first two strategies.

2.0 SYSTEM DESCRIPTION AND MODEL DEVELOPMENT

The study of ATM covers a wide area of issues such as call admission control, congestion control and buffer policing but this work focuses on bandwidth allocation schemes. As mentioned earlier, bandwidth allocation is a leading issue in ATM traffic management. Studies other than bandwidth allocation schemes will not be included in this work. In addition, ATM integrates a variety of traffic sources and some source types are appropriately modeled only by its dedicated modeling technique, e.g. Autoregressive Moving Average model for video sources. Therefore, it is very difficult to model all ATM supported traffic sources and in this study only VBR traffic sources are modeled by using general source model: ON-OFF model.



Fig. 2: The ATM Simulation Model

The simulation model consists of three components, which are ATM traffic sources, buffer and dynamic time slice server. The model transmission rate consists of a single virtual path which is 622 Mbps. As an ATM cell equals to 48 bytes of data information, thus to transmit a cell over a network, the simulation model considers a time slot is equal to one cell service time. In other words, one cell transmitted per slot or zero cell if there is no cell to transmit. Based upon the given ATM cell size, the service time is calculated as follow:

$$ServiceTime = \frac{PacketSize}{ServiceRate(Bandwidth \operatorname{Re} source)}$$

$$=\frac{48*8bits}{622,000,000bits/sec}$$
$$= 0.000000617$$
(1)

2.1 VBR Traffic Source Model

ATM traffic sources are the flow of ATM cells that have been generated. Different types of traffic sources have their own input parameters (Table 1), which indicate their cell generation pattern. Among the input parameters are average cells arrival rate, peak arrival rate, mean burst length, mean active period, mean idle period and burstiness. Fixed input parameters are average cells arrival rate, mean active period, mean idle period and mean burst length. The rest of the input parameters are varied. Peak arrival rate varied when burstiness varied. Active period and idle period are Poisson distributed. Four types of VBR traffic are modeled in this study, they are VBR Video, Connection Oriented Data, VBR Video/Data and Connectionless Data. Connection Oriented Data is assumed to be most delay insensitive in this study. The total number of sources for each traffic type is shown in Fig. 2. In this simulation, the traffic pattern used for VBR sources will be persistent and bursty source [7]. Persistent sources also known as "greedy" or "infinite" sources, always have cells to send. Therefore, the network is always congested. Bursty sources oscillate between active state and idle state. During active state, sources generate a burst of cells in accordance to their duration which are Poisson distributed. This is followed by an idle period, where no cells are generated, also Poisson Sources are independent of each other, distributed. meaning cells generation of a source are not dependent on other source's state, neither active or idle state.

Using Mean Burst Length (N) and Average Cell Arrival Rate (m), other parameter value can be obtained by assigning a value to Burstiness parameter. In this study, Burstiness (Burst) values under investigation are 2 and 4, and this will illustrate the effect of cell generation frequency to the network performance.

PeakArrivalRate(p) = Burst * m⁽²⁾

 $CellGenerationTime(T) = 384 / p \tag{3}$

MeanActivePeriod (1 / a) = T * N

MeanIdlePeriod(1/b) = ((384 * N) / m) - A

CellGenerationTimeInSlot(SR 622) =
$$T / 0.000000617$$

The VBR traffic source algorithm is shown below:

- * TimeSlotNow = current time of the ATM system in slots
- * EndActive = duration of time in slots of active period
- * ElaspedTime = increment of time slot by 1
- * Cell = increment cell generated by 1
- * T = one cell generation time in seconds

2.2 Buffer Model

Next component in the simulation model is buffer or queue. Buffer is the amount of space allocated for the temporary storage of cells flowing in which is important for traffic management purposes. Each traffic source's flow is directed to its specific buffer. This means that VBR video cells has its own buffer and VBR video/data cells will be queued into its own dedicated buffer. Therefore the simulation model itself has four queues. Each queue is serviced using the First-In-First-Out (FIFO) discipline. Every ATM cell that has been generated will be buffered or transmitted directly to network if the buffer length is zero and the bandwidth is available for this queue at the moment. A cell is considered lost if the buffer length reached its maximum limit and a new cell arrived before transmission. Cells will not be dropped even though there is heavy congestion in the network and exceed its delay requirement. When an ATM cell departed from its buffer, cell delay due to how long the cell stayed in the buffer will be calculated. Due to the buffer is also one of the important network resources, the buffer length is also an analysis element. Large buffer length will increase the average cell delay but of course decrease cell loss ratio. On the contrary, if the buffer length is too limited, then cell loss ratio will become a critical issue in measuring the network performance. Therefore there must be a balance between cell loss ratio and average cell delay by adopting appropriate buffer length. In this study, the buffer length will be increased from 200 to 1400 cells for simulation purposes.

(4)

(5)

(6)

Two events occur in the buffer model which are cells arrival and cells departure. Cells arrived will be queued and if the buffer overflows, cell loss will happen. Cells will be transmitted if there are bandwidth available otherwise it will be waiting until the bandwidth is allocated to them.

The Buffer model algorithm is shown below:

```
Cell_Arrived() {
  If queue length < buffer length {
     P = new elem()
     P->Slot = current slot time
     P->SourceI = source number
     If buffer not empty
       Insert cell into buffer, rear->next = P
    Else
       Insert cell as first cell into buffer, front = P
     Accumulate total number of cells buffered
     Accumulate queue length }
 Else
     Accumulate total number of cells lost }
Cell Departed() {
  Elem *p
  Transmitted = false
  If queue length >= 1 {
     If front == NULL
       Error
     Else {
         P = front
         Cell delay += current time slot - P->Slot
         If front == rear {
           Delete(P)
           Front = rear = NULL }
         Else {
           Front = P->next
           Delete(P) }
         Minus queue length by 1
         Accumulated cells departed by 1
         Transmitted = true } }
   Else
     Spare Slots available, transmit lower priority cell }
```

2.3 DTS Server Model

The final component is Dynamic Time Slice Server. This server is used for allocating the bandwidth for the system. Dynamic Time Slice (DTS) uses (T1, T2)-scheme defined in [8]. DTS (Dynamic Time Slice) scheme is based on the principle of having different buffering for different traffic classes [2]. As shown in Fig. 2, there are 4 types of ATM traffic sources, VBR Video Data, Connectionless Data, VBR Video and Connection Oriented Data, each of them is serviced through a separate queue. The server cyclically visits each buffer and devotes a period of its time to it. Assuming there are 2 buffers B1 and B2 in the system, DTS server is allocated T1 ms service time for B1 and T2 ms for B2. Thus B1 will be served until a maximum of T1 ms or

until B1 queue is exhausted, whichever occurs first. Then B2 is served correspondingly to the allocated time slice T2 ms. The transmission of VBR Video and Data cells will occur first after the bandwidth reallocation has happened. This is followed by Connectionless Data cells and then only VBR Video cells.

The amount of time slice allocated to a queue is dependent on the bandwidth allocation strategy and algorithm used. This study will concentrate on three bandwidth allocation strategies, Static Bandwidth Allocation, Bandwidth Allocated Proportional to Expected Queue Length and Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value. The cycle length for all these strategies is decided to be 200 time slots ([8] give the example of 450 time slots). For every 200 time slots, a new reallocation of bandwidth will be calculated, thus three priority queues will be provided among these 200 time slots. The sum of the allocated bandwidth on the three priority queues, VBR Video and Data Buffer, Connectionless Data Buffer and VBR Video Buffer equals the maximum cycle length (200 slots). One time slot can only transmit one ATM cell or is equal to one ATM cell transmission time.

Any bandwidth allocated and unused momentarily by 3 higher priority buffer cells will be declared as spare slot and allocated to delay insensitive buffer. This study assumes the Connection Oriented Data flow is insensitive to delay, and cells in this buffer are transmitted immediately when only there are spare slots available.

2.3.1 Static Bandwidth Allocation (Model 1)

Static bandwidth allocation is a strategy that guarantees bandwidth to high priority traffic classes. This allocation is defined in the server and does not change with time. Thus this strategy provides no congestion control. In the study, in order to utilize all bandwidth in the network, VBR Video and Data buffer is allocated 300 Mbps, Connectionless Data buffer is allocated 20 Mbps and VBR Video buffer is allocated 300 Mbps, giving the total 620 Mbps allocated. This pattern of bandwidth allocation is fixed for every 200 time slots. Thus this strategy does not have bandwidth allocation recalculation mechanism. Bandwidth allocated (BA) for the three traffics are shown as below:

$$VBRVideoData _ BA = \frac{300Mbps}{622Mbps} * 200Slots \tag{7}$$

$$ConnectionlessData_BA = \frac{20Mbps}{622Mbps} * 200Slots \quad (8)$$

$$VBRVideo_BA = \frac{300Mbps}{622Mbps} * 200Slots \tag{9}$$

2.3.2 Bandwidth Allocated Proportional to Expected Queue Length (Model 2)

This model is based on measurements of the queue lengths. The point of measuring is after the server has visited it as shown in Fig. 3.



Fig. 3: Measurement of the queue lengths

When the server is calculating the bandwidth allocation for the next cycle it will predict the number of cells in the queue after the server has visited this certain buffer. A proportioned slice of bandwidth is allocated to each queue. The formula for buffer Q1, bandwidth allocation is as follows:

$$QueueLengthAll = QueueLength(Q1 + Q2 + Q3)$$
(10)

$$ExpectedAll = Expected(Q1 + Q2 + Q3)$$
(11)

$$BAQ1 = \frac{Abs (QueueLengt h(Q1) + Expected (Q1))}{\sum (QueueLengt hAll + ExpectedAl l)}$$
(12)

where

= average cells arrival in previous 200 slots

2.3.3 Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value (Model 3)

This strategy is the same as the previous strategy (model 2). The additional control mechanism is added, if buffer VBR Video or VBR Video/Data or Connectionless Data exceed a certain threshold value, 90% of buffer size, (meaning congestion had occurred or being occurred, need more attention) then the amount of bandwidth allocated to congested buffer will be greater. If buffer VBR Video or VBR Video/Data or Connectionless Data does not exceed the threshold value or more than two buffers exceeding the threshold value, then the strategy will be the same as the previous strategy without control mechanism.

Bandwidth allocated will be calculated as follows: bandwidth allocated to non congested buffer will be deducted by 10% calculated from strategy 2 and these deducted bandwidth will be allocated to congested buffer. For example, VBR Video buffer is congested and exceeding the threshold value, under strategy 2, bandwidth allocated to VBR Video/Data and Connectionless Data are 80 slots and 20 slots respectively. Thus by reformulating the allocated bandwidth, bandwidth allocated to VBR Video/Data is 72 slots, Connectionless Data 18 slots and VBR Video is (200-72-18) = 110 slots.

2.4 Performance Metrics

There are two performance parameters that will be discussed in this paper, they are cell loss ratio and average cell delay (cycle).

Cell loss happens when the buffer is full and a new cell arrives at the particular queue. A queue is full when the number of cells in queue equals the maximum queue length. A cell loss is most likely to happen in a buffer before the server is serving it. Buffer type-2 has no cell loss due to its infinite buffer size.

$$CellLossRatio = \frac{\sum CellLoss}{\sum ArrivedCells}$$
(14)

Cell delay is the waiting time of an ATM cell in the buffer to switch through the link or the waiting time of a cell after residents in buffer until transmission. The waiting time is measured in slot time and can be converted to seconds by multiplying service rate of 0.000000617 seconds. As Connection Oriented Data cell is assumed to be the most tolerable to cell delay, it may be transmitted during slack period. So, cells from buffer of this type are excluded from the calculation of cell delay.

$$AverageCellDelay = \frac{\sum CellDelay}{\sum DepartedCell}$$
(15)

3.0 RESULTS AND DISCUSSIONS

The input parameters considered in this study are total time slots of 2 millions, total number of sources for each type of traffic where Connectionless Data 22 sources, VBR Video 12 sources, Connection Oriented Data 24 sources and VBR Video The number of sources used in this Data 14 sources. simulation study is at optimum value that the network can admit. Even with the increase of only one VBR Video/Data source, the overall network performance will be degraded a lot. The buffer sizes for VBR Video/Data and VBR Video varied from 200 cells to 1400 cells, whereas Connectionless Data remained 100 cells and Connection Oriented Data infinite buffer size. The threshold value of 90% of buffer length would be chosen. Output parameters are the performance metrics discussed earlier in section 2, they are cell loss ratio and average cell delay. The network performance would be grafted out with buffer size as x-axis and output parameters as y-axis. The simulation was carried out on a Pentium II 400 MHz system using Visual C++ 6.0.

(13)

Only the results for VBR Video/Data buffer and VBR Video traffic are given in the discussions. The abbreviation of SVBRViD, EVBRViD, TVBRViD, SVBRVi, EVBRVi and TVBRVi are explained below. S stands for Static Bandwidth Allocation. E stands for Bandwidth Allocated Proportional to Expected Queue Length and T stands for Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value. VBRViD gives the meaning VBR Video/Data sources and VBRVi gives the meaning VBR Video sources.

The comparisons of cell loss ratio performance in VBR Video/Data buffer for three bandwidth allocation strategies are depicted in Fig. 4. Fig. 4 shows that when buffer sizes increase, cell loss ratio performance increases. Buffer sizes of 200 cells show the highest cell loss ratio for all three strategies and gradually when buffer sizes increase, cell loss ratio performance for all three strategies improves. Static Bandwidth Allocation strategy is among the worst cell loss ratio performance cases. Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value strategy shows better cell loss ratio performance among three strategies. From the graph, buffer sizes of 200 and 400 cells give too excessive cell loss. Buffer sizes beginning from 600 cells will give a better working environment for ATM network.

The comparisons of cell loss ratio performance in VBR Video buffer for three bandwidth allocation strategies are depicted in Fig. 5. It is shown that Static Bandwidth Allocation strategy has the worst cell loss performance. Its performance almost remains constant at 0.02. Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value strategy shows cell loss ratio higher than Bandwidth Allocated Proportional to Expected Queue Length strategy from buffer sizes of 200 cells to 600 cells. But after buffer sizes of 600 cells, the cell loss ratio performance for these two strategies does not show a significant difference. The comparisons of average cell delay performance in VBR Video/Data buffer for three bandwidth allocation strategies are depicted in Fig. 6. Static Bandwidth Allocation strategy has the worst average cell delay performance. Although the average cell delay performances for Bandwidth Allocated Proportional to Expected Queue Length is better than Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value at buffer sizes of 800 cells, it is still considered that the later strategy has better performances. This situation might have resulted from the generation of random numbers for cell arrival, active and idle period during the simulation processes.

The comparisons of average cell delay performance in VBR Video buffer for three bandwidth allocation strategies are depicted in Fig. 7. It shows that average cell delay for Static Bandwidth Allocation increases as buffer sizes increase, contrary, other two bandwidth allocation strategies remain constant. Average cell delay performance for Bandwidth Allocated Proportional to Expected Queue Length and Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value does not show a significant difference.

4.0 CONCLUSIONS

The performances of cell loss ratio and average cell delay for three bandwidth allocation strategies have been evaluated. It is concluded that when buffer size is greater than 600 cells, Bandwidth Allocated Proportional to Expected Queue Length with Threshold Value strategy will give better performance compared to other two strategies. [9] also gives the conclusion that by using a relatively large buffer sizes (>500 cells), it will improve the multiplexing efficiency.



Fig. 4: Cell Loss Ratio vs Buffer Length for 14 VBR Video/Data Sources for 3 Different Bandwidth Allocation Strategies



Fig. 5: Cell Loss Ratio vs Buffer Length for 12 VBR Video Sources for 3 Different Bandwidth Allocation Strategies When VBR Video/Data Sources = 14



Fig. 6: Average Cell Delay vs Buffer Length for 14 VBR Video/Data Sources for 3 Different Bandwidth Allocation Strategies



Fig. 7: Average Cell Delay vs Buffer Length for 12 VBR Video Sources for 3 Different Bandwidth Allocation Strategies When VBR Video/Data Sources = 14

REFERENCES

- A. K. Ramani, Loh Chee Heok and Keng Seng Oon, "ATM Switch Performance with Dynamic Bandwidth Allocation", *Malaysian Journal of Computer Science*, Vol. 11 No. 1, June 1998, pp. 1-8.
- [2] Lu Wei, Peter Hansson, "Further research on dynamic time slice in ATM switch", *Computer Networks and ISDN Systems*, 28 (1996) 789-798.
- [3] G. D. Stamoulis, M. E. Anagnostou, A. D. Georgantas (1994), "Traffic Source Models For ATM Networks: A Survey", *Computer Communications*, Vol. 17 No. 6, June 1994, pp. 428-438.
- [4] Basil Maglaris, Dimitris Anastassiou, Prodip Sen, Gunnar Karlsson, John D. Robbins, "Performance Models of Statistical Multiplexing in Packet Video Communications", *IEEE Transactions On Communications*, Vol. 36, No. 7, July 1998.
- [5] J. W. Roberts, "Traffic Control In The B-ISDN", Computer Networks and ISDN Systems, 25 (1993) 1055-1064.
- [6] Jean-Yves Le Boudec, "The Asynchronous Transfer Mode: A Tutorial. Computer Networks and ISDN Systems, 24 (1992) 279-309.
- [7] Raj Jain, "Congestion Control And Traffic Management In ATM Networks: Recent Advances And A Survey", *Computer Networks and ISDN Systems* 28 (1996) 1723-1738.

- [8] K. Sriram, "Methodologies For Bandwidth Allocation, Transmission Scheduling, And Congestion Avoidance In Broadband ATM Networks", Computer Networks and ISDN Systems, 26 (1993) 43-59.
- [9] Flavio Bonomi, Sergio Montagna, Roberto Paglino, "A Further Look At Statistical Multiplexing in ATM Networks", *Computer Networks and ISDN Systems*, 26 (1993) 119-138.

BIOGRAPHY

Mohamed Othman completed his PhD from University Kebangsaan Malaysia in 1999. Currently, he is a lecturer at the Department of Communication Technology and Networks, Faculty of Computer Science and Information Technology, University Putra Malaysia. His research interest includes parallel computing, high speed network computing (cluster computing), artificial intelligence, expert system design and programming in logic (parallel).

Chee Heok Loh obtained his Bachelor of Computer Science degree from University Putra Malaysia, in 1998. Currently he is pursuing his Master degree at University Putra Malaysia. His research interests include system development and high speed network.

Shyamala D. is currently lecturing at the Department of Multimedia, Faculty of Computer Science and Information Technology, University Putra Malaysia. Her research interests include multimedia networking and audio technology for multimedia systems.

Lily Nurliana Abdullah is currently lecturing at the Department of Multimedia, Faculty of Computer Science and Information Technology, University Putra Malaysia. Her research interests include multimedia networking.