ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Dynamic Bandwidth Allocation Scheme in Support of Video-on-Demand Services over Passive Optical Networks

Ouyang **Yue**

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Master of Engineering

2003

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

To my family, who always give me love, encouragement and support.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Acknowledgements

I am indebted to many individuals for their care and support during my study for a Master degree in NTU. First, I would like to express my deep gratitude to my supervisor Prof. Lu Chao, for his many instructions and continuous encouragement throughout my research. His wise guidance and great concern have provided tremendous help to my professional growth.

I would also like to thank and acknowledge Prof. Zhang Liren and Prof. Shum Ping, who have provided me constant encouragement, insightful comments and invaluable help. Their help not only benefited my research so far, but also my career in the future. Without their continuous encourage and broad vision, my work would not have been possible.

It is also a great pleasure to thank my fellow graduate students in Radar Lab and Network Technological Research Center. I have enjoyed many technical discussions with Hou Fang, Leng Supeng, Zhu Yongqing, Huang Xiaohong, He Jianhua, Shen Gangxiang, etc.. I also wish to thank Hou Zhihua, Song Zhefeng and Zhang Keli for their friendship and all the good time we had together. Especially thanks to Hou Fang, Leng Supeng and Shu Gang, for their care, support and love. And especially thanks also goes to He Jianhua, for his care and encouragement during the first several months I have been here.

Finally, it is with deep appreciation to thank my families whose forever love, sacrifice and encouragement have always enabled me to grow up through all these years - my mother and father, my uncle and aunt, and my grandma.

November 17, 2002 Ouyang Yue

Singapore

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Summary

This thesis examines the techniques to improve the performance of video-on-demand services over passive optical networks. In particular, a dynamic bandwidth allocation scheme is proposed in support of video-on-demand (VoD) services over passive optical networks (PONs). It has achieved significant performance improvement in terms of network utilization through bandwidth management and buffer dimensioning.

Firstly, this thesis explores the state of art technologies and the main design issues related to video-on-demand system over passive optical networks. It has been found that it is necessary to build a dynamic bandwidth allocation management approach in the system in support of VoD over PON. This thesis focuses on the mechanism that is able to efficiently transport pre-stored MPEG VBR-encoded video across network.

To investigate bandwidth management methods, a queuing model of transmitting pre-stored video over passive optical networks is set up. To guarantee the quality of service, it is required that there is neither buffer overflows nor buffer underflows. Hence, a dynamic bandwidth allocation scheme based on the concept of "playback tunnel" are proposed to determine the range of feasible bandwidth solutions that can be allocated to support the transmission of MPEG video streams across passive optical networks. This will ensure that the goal of low transmission cost, easy and flexible traffic control approach, guaranteed quality of service, and high network multiplexing

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

gain are able to be achieved.

To support the dynamic bandwidth allocation scheme based on the playback tunnel, two segmentation schemes including equal segment scheme and dynamic segment scheme are considered with the dynamic bandwidth allocation scheme. The dynamic segment scheme taking into consideration of intrinsic video rate characteristics can make use of dynamic bandwidth allocation more efficiently. The performance of the proposed schemes is evaluated by a set of MPEG VER-encoded long video streams. The numerical results demonstrate the effectiveness of the proposed schemes comparing with the well-known constant rate transmission and transport scheme. The dynamic segment scheme shows a better performance when comparing with the equal segment scheme.

To consider the practical cases for the requirement of delivering multiple video streams, this thesis also studies statistical multiplexing together with dynamic bandwidth allocation scheme. The impact of time correlation multiplexing and content correlation multiplexing of video traffic on network resource management are examined under dynamic bandwidth allocation with both equal segment and dynamic segment schemes. The study of exploiting multiplexing gains demonstrates that, the dynamic bandwidth allocation scheme with statistical QoS guarantees can provide a viable, resource-efficient way to support real-time delivery of pre-recorded video.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table of Contents

Acknowledgements ii

Summary iii

Table of Contents iv

List of Figures vii

List of Tables ix

1 Introduction 1

1.1 Motivation 1

1.2 Objectives 2

1.3 Thesis Outline 4

1.4 Main Contributions 5

2 Video-on-Demand Services over Passive Optical Networks 8

2.1 Introduction 8

2.2 Passive Optical Networks 8

2.2.1 Overview of Passive Optical Networks 9

2.2.2 History of Passive Optical Networks 13

2.2.3 Types of Passive Optical Networks 15

2.2.4 Delivering Video Services over PONs and Its Market View . . 18

2.3 Video-on-Demand over Passive Optical Networks 20

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.3.1 Overview of Video-on-Demand Services 21
2.3.2 Design Issues in Video-on-Demand Services over PONs 22
2.3.3 Models for Video-on-Demand Services over PONs 26
2.4 Dynamic Bandwidth Allocation for Video-on-Demand Services over PONs 28
3 Dynamic Bandwidth Allocation 31
3.1 Introduction 31
3.2 Related Works 32
3.3 Transmission Model of Video-on-Demand over Passive Optical Networks 36
3.3.1 Transmission Model 36
3.3.2 Overflow and Underflow Constraints 40
3.4 Bandwidth Allocation and Buffer Occupancy 41
3.4.1 Minimum Transmission Bandwidth without Buffer Underflow 41
3.4.2 Minimum Buffer Capacity Required to Avoid Overflow 42
3.4.3 Maximum Transmission Bandwidth without Buffer Overflow . 42
3.5 Tunnel Scheme 43
3.6 Segmentation Schemes for Dynamic Bandwidth Allocation 44
3.6.1 Related Works on Video Segmentation Techniques 45
3.6.2 Equal Segment Scheme 47
3.6.3 Dynamic Segment Scheme 49
4 Performance Evaluation 52
4.1 Introduction 52
4.2 Performance Evaluation under Equal Segment Scheme 54
4.3 Comparison with CRTT Scheme 59

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.4 Performance Evaluation under Dynamic Segment Scheme 63
4.5 Comparison of the Two Bandwidth Allocation Schemes 67
4.6 Traffic Characteristics with Bandwidth Allocation Schemes 73
4.6.1 Related Works 73
4.6.2 Peak Rate and Traffic Burstiness 74
4.6.3 Overload and Underload 77
5 Multiplexing Video Streams over Passive Optical Networks 80
5.1 Introduction 80
5.2 Related Works on Statistical Multiplexing Compressed Video 81
5.3 Performance Parameters and Simulation Models 81
5.4 Simulation Results 84
5.4.1 Time Independent Multiplexing 84
5.4.2 Time Correlated Multiplexing 85
6 Conclusions and Recommendations for Future Work 94
6.1 Conclusions 94
6.2 Recommendations for Future Work 96
Author's Publications 99
Bibliography 100

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Figures

1.1 Video-on-demand service over passive optical network 3

2.1 The general structure of a passive optical network 10

2.2 The typical topologies of a passive optical network 11

2.3 The general structure of a VoD system 23

2.4 MPEG2 encoding structure 26

2.5 Video overlay model 27

2.6 In-band VoD model 28

3.1 Transmission of pre-recorded MPEG video over PON 37

3.2 MPEG video sequence "Movie-Lambs" 38

3.3 The relationship between arrival stream, playback stream and buffer occupancy 39

3.4 Playback tunnel 44

3.5 Relationship between arrival stream $S(t)$, playback stream $P(t)$ and buffer occupancy $O(t)$ under equal segment scheme 48

3.6 Playback tunnel under dynamic segment scheme 51

4.1 The playback MPEG video frame trace 54

4.2 The cumulated playback video data 55

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.3 Arrival stream $S(t)$ vs. playback stream $P(t)$ 56
4.4 Zoom-in of Arrival stream $S(t)$ vs. playback stream $P(t)$ 56
4.5 Bandwidth allocation 58
4.6 User's buffer occupancy 58
4.7 Playback buffer capacity vs. segment number 59
4.8 Arrival streams of CRTT and dynamic bandwidth allocation 62
4.9 Buffer occupancy under CRTT and dynamic bandwidth allocation scheme, delay=0.8s 63
4.10 Buffer occupancy under CRTT and dynamic bandwidth allocation scheme. delay=32s 64
4.11 Tunnels under dynamic segment scheme with buffer capacity 2.000.000 bytes and 3.000.000 bytes 65
4.12 Network bandwidth allocation under dynamic segment scheme 66
4.13 Buffer occupancy under dynamic segment scheme 67
4.14 Arrival stream and playback stream with two schemes 69
4.15 Points of segmentation 70
4.16 Bandwidth allocation with two schemes 71

5.1 Multiplexing gain of independent streams 88
5.2 Normalized bandwidth of independent streams 89
5.3 Multiplexing gain of correlated streams: same video streams 90
5.4 Multiplexing gain of correlated streams: different video streams . . . 91
5.5 Normalized bandwidth of correlated streams: same video streams . . 92
5.6 Normalized bandwidth of correlated streams: different video streams 93

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Tables

4.1 Statistical characteristics of the MPEG VBR-encoded video streams 53
4.2 Segment number needed when changing buffer capacity 66
4.3 Comparison of segment number needed 72
4.4 Average rate and peak rate under different schemes 76
4.5 Comparison of traffic burstiness 77
4.6 Comparison of overload and underload 78

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 1

Introduction

1.1 Motivation

Passive optical networks (PONs) have gained a lot of attention as a possible solution for the broadband local access network. The use of PONs is able to give more bandwidth to each end user. This offers greater flexibility in the types and characteristics of the services. Improved bandwidth and security performance make the PON network the best solution to offer services over the multi-service platform. Since it could provide bandwidth on demand, it allows evolution to tomorrow's broadband services such as video-on-demand (VoD), B-ISDN, and broadband multimedia.

The asymmetry of the traffic transported in the two directions of PON reflects the intrinsic asymmetry of the service characteristics. Asymmetric broadband services, such as Internet browsing, switched digital video broadcasting (SDVB), video on demand (VoD) and multimedia information retrieval, require high channel capacity in the network-to-user direction, while a limited amount of bandwidth is needed in the opposite direction. Therefore, PON has become an ideal network in support of such services. Among them, video-on-demand service is believed to be the most

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

demanding application in broadband networks.

To support VoD service over PON, the downstream channel needs to transmit large amount of video data. In particular, when the number of end users increases, it becomes much more important to effectively use the limited downstream bandwidth for VoD service. It is also an important issue to guarantee quality of service (QoS) with minimum data transfer delay and minimum loss probability during the transmission of video data across the networks. By considering these issues and the native burstiness feature of video data, PON must allocate downstream bandwidth dynamically.

The constant rate transmission and transport (CRTT) scheme [30] is well-known for bandwidth allocation in support of VoD service. It is very simple, but it has some weaknesses such as very large start-up delay and very large buffer capacity. This thesis focuses on a dynamic bandwidth allocation scheme in support of VoD service over PON, which has a simple and flexible resource management scheme, and has many important advantages over the CRTT scheme.

1.2 Objectives

The main issues in designing a video-on-demand system over passive optical networks are primarily related to resource allocation. Issues such as disk partitioning and bandwidth reservation are crucial for supporting the system. The other issues include architecture, call admission control, and multiple access to video resource on the server side.

A simple VoD system over PON architecture is shown in figure 1.1. The key elements in the system include:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

- video server and storage
- optical line terminal
- optical splitter
- optical network units
- clients with set-top boxes or specialized personal computers

In this typical VoD system over PON, a set of video databases (content providers) contained in video storage are interconnected to a set of servers which perform routing of VoD traffic to the center office of passive optical networks. Users connect to the VoD system through optical network units of PON.

This thesis focuses on the study of traffic management scheme based on such a system for pre-stored video, with the following objectives:

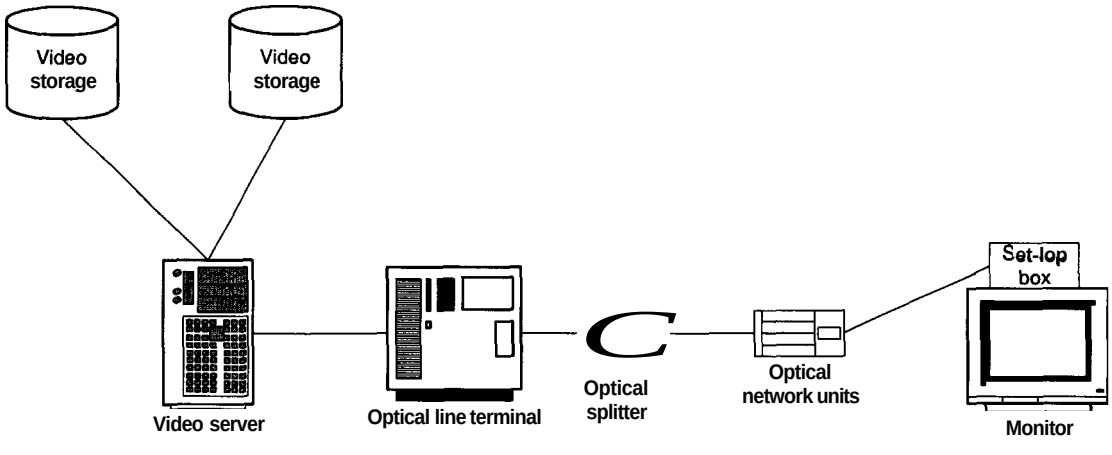


Figure 1.1: Video-on-demand service over passive optical network

1. To set up a practical and efficient dynamic bandwidth allocation scheme for the support of video-on-demand services over passive optical networks.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2. To analyze the impact of the proposed bandwidth allocation scheme on network performance and network resource management.

In designing the traffic management scheme, the following considerations have to be taken:

1. Low network transmission costs;
2. Guaranteed Quality of Service without data loss or delay in real-time transmission;
3. Easy and flexible traffic control mechanism;
4. Efficient bandwidth utilization.

1.3 Thesis Outline

Chapter 1 introduces the motivation, objective and contents of this thesis. The main objectives of the thesis are to find an efficient traffic management scheme for VoD systems over PONs and to study the impacts of the traffic management scheme on network in support of video stream transmission.

Chapter 2 provides a broad outline of VoD systems over PONs. It includes concept, types and physical architecture of PONs. It compares PON with other optical solutions. It also addresses the issues in the design of VoD services over PONs, and the main ideas considered for dynamic bandwidth allocation over PONs for VoD systems.

Chapter 3 presents the dynamic bandwidth allocation scheme for VoD over PONs. The proposed algorithm focuses on an optimized scheme of bandwidth allocation

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

and buffer dimensioning to transport video traffic from video sever to clients over PONs. The scheme is implemented according to the known characteristics of the video traffic to prevent the receiving buffer on the clients' side from underflow and overflow. A tunnel scheme which is very easy for dynamic bandwidth allocation has also been proposed. Two segmentation schemes have been considered for the dynamic bandwidth allocation.

Chapter 4 further evaluates the impacts of the proposed dynamic bandwidth allocation scheme on the characteristics of video traffic over PONs. The impacts under the two different segmentation schemes have been compared and analyzed.

Chapter 5 considers the multiplexing of multiple video streams over PONs and the impacts of the proposed dynamic bandwidth allocation scheme on the multiplexing characteristics of video streams.The new scheme is able to achieve significant traffic multiplexing gain in terms of network link utilization.

Chapter 6 summaries the main findings of this thesis and discusses possibilities for future work. Recommendations are presented for further research work on relevant topics.

1.4 Main Contributions

In this thesis, a novel approach of dynamic bandwidth allocation for video-on-demand service over passive optical networks has been proposed and evaluated. The proposed algorithm focuses on optimized scheme of bandwidth allocation and buffer dimensioning to transport video traffic from video sever to clients over PONs. The scheme is implemented according to the known characteristics of the video traffic since it is

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

pre-stored in the sever to prevent the receiving buffer on the clients' side from underflow and overflow. On the other hand the new scheme is able to achieve significant traffic multiplexing gain in terms of network link utilization.

Based on the dynamic bandwidth allocation scheme, two kinds of segmentation schemes are presented. A dynamic segment scheme taking into consideration of intrinsic video rate characteristics is proposed for efficiently dynamic bandwidth allocation.

The performance of the dynamic bandwidth allocation scheme with the two segmentation schemes is evaluated, by a set of MPEG VBR-encoded video traces in terms of several parameters, such as transmission bandwidth allocation, and buffer occupancy. The numerical results have demonstrated that the proposed schemes are not only able to significantly match the playback stream but also able to largely reduce the playback buffer capacity requirement. Because of the fact that network performance depends fundamentally on peak rate and bursty characteristics of the admitted traffic, this thesis also tests the traffic characterization in terms of peak data rate, traffic burstiness, and traffic load to investigate the impact of the proposed schemes on network traffic. The results show that the proposed dynamic bandwidth allocation scheme associated with both of the two segment schemes has positive effects in reducing the peak rate and burstiness. It also has positive effects in decreasing the video data accumulating in a buffer and the degree to which the network is overloaded during the burst period.

The video streams' statistical multiplexing with dynamic bandwidth allocation scheme is also evaluated by simulation to examine the effect of the proposed approach on the requirement of network resources. The study shows that the dynamic bandwidth allocation schemes whether with equal segment or with dynamic segment,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

have positive impact on the performance of statistical multiplexing.

The research results in this thesis are applicable to the practical implementation of video-on-demand system over passive optical networks.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 2

Video-on-Demand Services over Passive Optical Networks

2.1 Introduction

This chapter provides a broad outline of video-on-demand system over passive optical networks. The concept, types, physical architecture and evolution of PONs are given in detail. The PON is compared with other access network solutions. This chapter also discusses the issues in the design of VoD services over PONs, the related works that have been done and the main ideas considered for dynamic bandwidth allocation over PONs for VoD systems.

2.2 Passive Optical Networks

The present development of new broadband telecommunication services makes the upgrading of the access networks a demanding research goal. To offer video as well as advanced Internet applications to the residential customers requires the availability of a channel capacity that seems hardly achievable with the traditional copper wire-based access networks. Different "wired" solutions for the access network are currently

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

being developed. The most important ones of these solutions are the digital subscriber loop (DSL), hybrid fiber coax (HFC) and fiber-in-the-loop (FITL). It is actually a hotly debated issue which one among these solutions will better meet the needs of the future broad-band telecommunication network. The passive optical network, one kind of FITL solutions, is receiving at present more attention than in the past from the telecommunication operators. This is mainly due to its ability of implementing routing functions entirely by employing passive optical devices.

2.2.1 Overview of Passive Optical Networks

PONs are low cost optical Fiber-to-the Building/Curb/Home (FTTb, FTTc, FTTh or collectively referred to as FTTx) solutions [40]. A PON is a point-to-multipoint optical network that allows service providers to minimize the need for fiber in the outside portion of the network to interconnect buildings or homes.

PONs consist fundamentally of an optical line terminal (OLT) located at the central office (or cable headend), and multiple remote optical network units (ONUs) that deliver broadband voice, data and video services to subscribers. Optical network unit is also called optical network terminal (ONT). Traffic from the ONUs is aggregated back to the OLT using network topologies such as tree, bus or fault-tolerant rings. ONUs are geographically distributed in buildings, on curbs, on utility poles or on the sides of homes. A simple structure of passive optical network is shown in figure 2.1. And the typical PON topologies, i.e., tree, ring and bus, is shown in figure 2.2.

The communication links interfacing the ONUs are dedicated and form the distribution network. The communication link on the central office side is instead shared by all the ONUs of the network and is called feeder network. Between feeder and

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

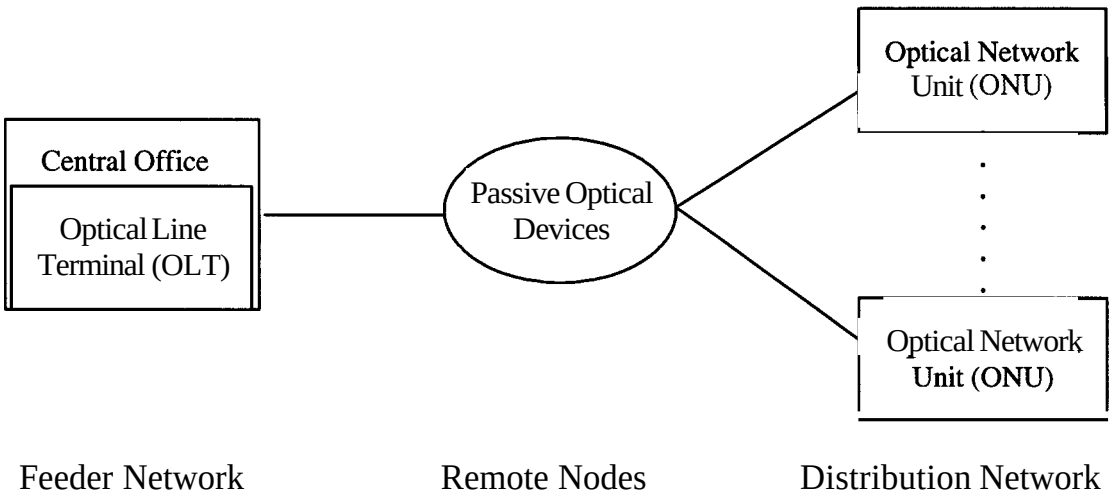


Figure 2.1: The general structure of a passive optical network

distribution network a structure composed by one or more remote nodes (RNs) entirely implemented by passive optical devices like optical splitter (and in some cases intermediate optical links) performs the separation of downstream optical channels as well as the multiplexing of upstream optical channels in a transparent way.

Most networks in the telecommunications networks of today are based on active components at the serving office exchange and termination points at the customer premises as well as in the repeaters, relays and other devices in the transmission path between the exchange and the customers. Active components mean those devices which require power of some sort, and are generally comprised of processors, memory chips or other devices which are active and processing information in the transmission path.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

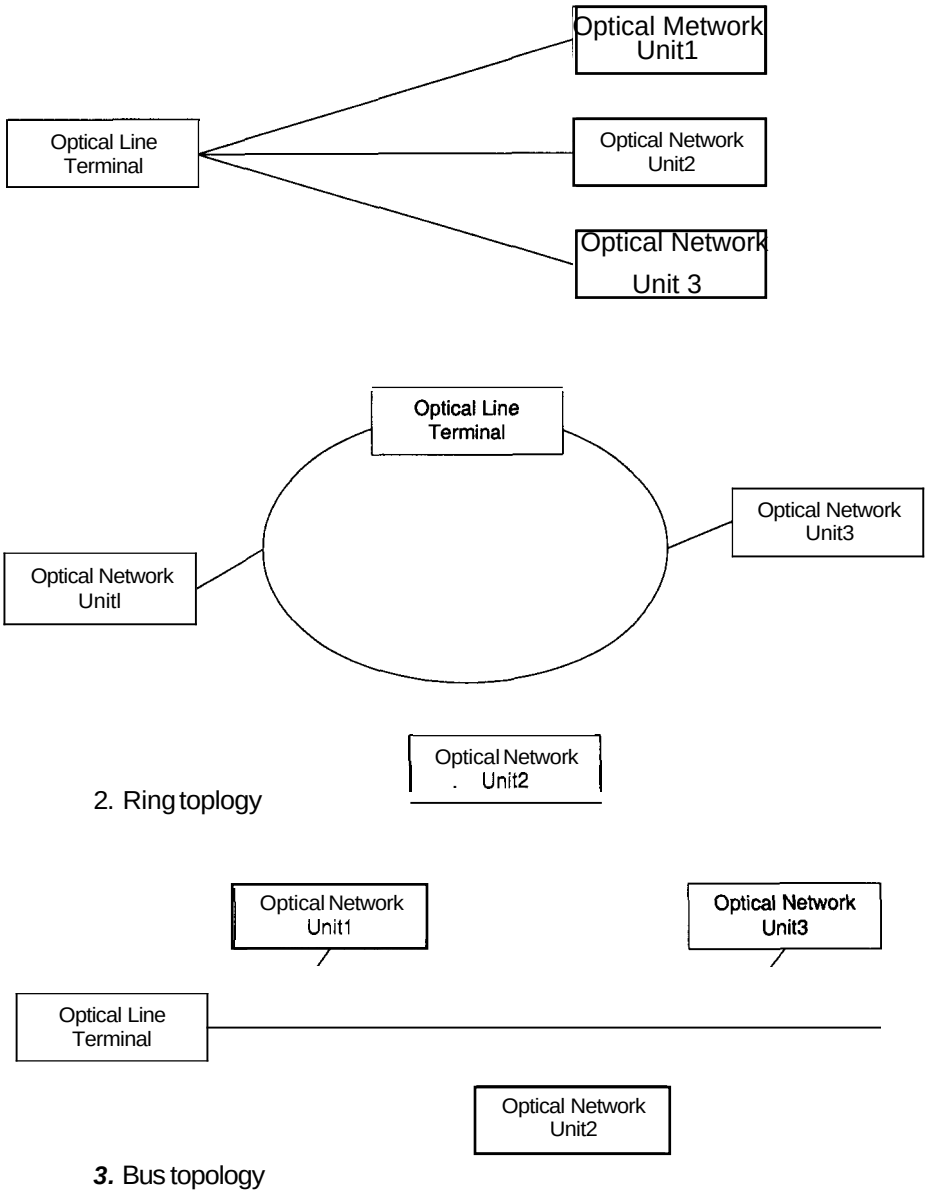


Figure 2.2: The typical topologies of a passive optical network

With passive optical networks, all active components between the central office exchange and the customer premises are eliminated, and passive optical components are put into the network to guide traffic based on splitting the power of optical

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

wavelengths to endpoints along the way. This replacement of active with passive components provides cost-savings to the service provider by eliminating the need of power and service active components in the transmission loop. The passive splitters or couplers are merely devices working to pass or restrict light, and as such, have no power or processing requirements and have virtually unlimited mean time between failures (MTBF) thereby lowering overall maintenance costs for the service provider.

In a PON, a single piece of fiber can be run from the serving exchange out to a subdivision or office park, and then individual fiber strands to each building or serving equipment can be split from the main fiber using passive splitters or couplers. This allows for an expensive piece of fiber cable from the exchange to the customer to be shared amongst many customers, thereby dramatically lowering the overall costs of deployment for fiber to the building or fiber to the home applications. The alternative is to run individual fiber or copper strands from exchange to customer premises, which results in much higher serving costs per customer.

Comparing to other fiber optical solutions, passive optical networks reduce or eliminate the number of active components, such as lasers, regenerators and amplifiers, thus cutting costs, reducing maintenance and improving network performance. In contrast to legacy copper solutions, PONs offer an even more compelling case: they make broadband possible while totally eliminating electronics from the outside plant, such as remote terminals and powering systems, thus radically reinventing the economics for the access network.

The main fiber run on a PON network can operate at 155 Mbps, 622 Mbps, 1.25 Gbps or 2.5 Gbps using APON/ BPON, EPON or the emerging GPON standards. Bandwidth allocated to each customer from this aggregate bandwidth can be static

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

or dynamically assigned in order to support voice, data and video applications.

With PONs, a single fiber from the carrier's exchange can service 16, 32 or more buildings through the use of both passive devices to split the optical signal, and PON protocols to control the sending and transmission of signals across the shared access facility.

Most PONs use wavelength division multiplexing (WDM) to put both the downstream and upstream optical signals onto a single fiber at different wavelengths. Upstream data is transmitted according to control mechanisms in the OLT, using a time division, multiple access (TDMA) protocol, in which dedicated transmission time slots are granted to each individual ONT. The time slots are synchronized so that transmission bursts from different ONTs do not collide.

The process of transporting data downstream to the customer premises is different from transporting data upstream from the customer premises. Downstream data is broadcasted from the OLT to each ONT, and each ONT processes the data destined to it by matching the address at the protocol transmission unit header.

2.2.2 History of Passive Optical Networks

The first formal PON activity was initiated in the Spring of 1995 when a group of seven major network operators established the Full Service Access Networks (FSAN) consortium. This group's goal was to define a common standard for PON equipment so that equipment vendors and operators could come together in a competitive market of PON equipment. The result of this first effort was the 155 Mbps PON system specified in the ITU-T G.983 [10] series of standards. This system has become known as the B-PON system, and it uses ATM as its bearer protocol (also known as the APON protocol).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The APON Standards were later enhanced to support 622 Mbps bit rates as well as additional features in the form of protection, dynamic bandwidth assignment and so on.

In order to construct and evaluate an access network that connect a large number of subscribers onto a single line termination, the ACTS PLANET project developed the SuperPON architecture. For the SuperPON an optical splitting factor of 2048 and a range of 100km is visualized. The downstream direction distributes 2.5 Gbps while the upstream direction shares 311 Mbps.

On a parallel track with APON, in early 2001, the IEEE established the Ethernet in the First Mile (EFM) group, realizing the enormous prospect that lies ahead in the optical access market. The group works under the auspices of the IEEE 802.3 group, which also developed the Ethernet standards, and is restricted in architecture and compliance to the existing 802.3 MAC layer. The EFM work is concentrated on standardizing a 1.25 Gbps symmetrical system for Ethernet transport only.

In 2001 the FSAN group initiated a new effort for standardizing PON networks operating at bit rates of above 1 Gbps. Apart from the need to support higher bit rates, the overall protocol has been opened for reconsideration and the sought solution should be the most optimal and efficient in terms of support for multiple services, OAM&P functionality and scalability.

As a result of this effort, a new solution , Gigabit PON (GPON), has emerged into the optical access market place, offering high bit rate support while enabling transport of multiple services.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.2.3 Types of Passive Optical Networks

There are various flavors of PON technologies including ATM-PON/Broadband-PON, SuperPON, Ethernet-PON and Gigabit-PON. Each of these is further explained in the following paragraphs.

ATM-PON and Broadband-PON

ATM-PON(APON) systems are based upon ATM as the bearer protocol. Downstream transmission is a continuous ATM stream at a bit rate of 155.52 Mbps or 622.08 Mbps with dedicated Physical Layer OAM (PLOAM) cells inserted into the data stream. Upstream transmission is in the form of bursts of ATM cells, with a 3 byte physical overhead appended to each 53 byte cell in order to allow for burst transmission and reception.

The transmission protocol is based upon a downstream frame of 56 ATM cells (53 bytes each) for the basic rate of 155 Mbps, scaling up with bit rate to 224 cells for 622 Mbps. The upstream frame format is 53 cells of 56 bytes each for the basic 155 Mbps rate.

The downstream frame is constructed from 2 PLOAM cells, one at the beginning of the frame and one in the middle, and 54 data ATM cells. Each PLOAM cell contains grants for upstream transmission relating to specific cells within the upstream frame (53 grants for the 53 upstream frame cells are mapped into the PLOAM cells) as well as OAM&P messages.

Upstream transmission consists of either a data cell, containing ATM data in the form of VPs/VCs or may contain a PLOAM cell instead when granted a PLOAM opportunity from the central OLT.

APONs are defined within the FSAN organization as well as the ITU-T.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The initial PON specifications defined by the FSAN committee used ATM as their layer 2 signaling protocol. As such, they became known as ATM-based PONs or APONs.

Use of the term APON led users to believe that only ATM services could be provided to end-users, so the FSAN decided to broaden the name to Broadband PON. BPON systems offer numerous broadband services including Ethernet access and video distribution.

BPON networks are defined by the FSAN and ITU committees comprised of both equipment vendors and service providers.

Super-PON

The possible evolution scenarios for the APON system with respect to bit rate, range, and splitting factor have led to the SuperPON concept [14].

APON has a splitting factor of 32, a downstream bit rate of 622 Mbps and an upstream bit rate of 155 Mbps. While, the SuperPON has an optical splitting factor of 2048 and a range of 100 km is visualized. The downstream direction distributes 2.5 Gbps and the upstream direction shares 311 Mbps. Since the power budget of SuperPON increases significantly, optical amplifers are introduced to offset the higher losses.

Ethernet-PON

Ethernet for subscriber access networks, also referred to as "Ethernet in the First Mile", or EFM, combines a minimal set of extensions to the IEEE 802.3 Media Access Control (MAC) and MAC Control sublayers with a family of Physical Layers (PHY). These Physical Layers include optical fiber and unshielded twisted pair (UTP) copper cable Physical Medium Dependent sublayers (PMDs) for point-to-point connections

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

in subscriber access networks.

EFM also introduces the concept of Ethernet Passive Optical Networks (EPONs), in which a point to multipoint (P2MP) network topology is implemented with passive optical splitters, along with optical fiber PMDs that support this topology. In addition, a mechanism for network Operations, Administration and Maintenance (OAM) is included to facilitate network operation and troubleshooting.

EPON is based upon a mechanism named Multi-Point Control Protocol (MPCP), defined as a function within the MAC control sublayer. MPCP uses messages, state machines, and timers, to control access to a P2MP topology. Each ONU in the P2MP topology contains an instance of the MPCP protocol, which communicates with an instance of MPCP in the OLT.

At the basis of the EPON/MPCP protocol lies the point-to-point emulation sublayer, which makes an underlying P2MP network appear as a collection of point to point links to the higher protocol layers (at and above the MAC Client). It achieves this by prepending a Logical Link Identification (LLID) to the beginning of each packet, replacing two octets of the preamble.

EPON as a protocol is still under work within the IEEE EFM group.

Gigabit-PON

Gigabit PON (GPON) came into the optical access market place in 2001. It aims at offering unprecedented high bit rate support as well as enabling transport of multiple services, specifically data and TDM, in native formats and at an extremely high efficiency.

As part of the GPON effort a Gigabit Service Requirements (GSR) document has been put in place based upon the collected requirements from all member service

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

providers, representing the leading RBOCs and ILECs of the world. The document has also been recently submitted as an official recommendation to the ITU-T [18].

GPON systems are fully based on the entire set of requirements laid out by the GSR document, and thus represent a pervasive solution for the service providers' needs.

GPON as a protocol is still under work within the FSAN and ITU organizations.

2.2.4 Delivering Video Services over PONs and Its Market View

Service providers are endorsing next-generation, packet-based integrated voice and data access platforms for fiber-to-the-home networks, but video still presents unique challenges because of its bandwidth and transmission requirements. With the advent of interactive TV (ITV) and video-on-demand, these challenges become even greater. The deployment of video services requires a cost-effective solution that can support evolving services with requirements that may not yet be known.

The market for video services is enormous and growing. There are now over 73 million households subscribing to basic cable TV (CATV) services and over 17 million paying for satellite-delivered TV services in United States. For network operators, video represents a highly attractive broadband service offering in terms of revenue potential, both for content and advertising. This potential is increasing as services such as pay-per-view (PPV) and video-on-demand gain subscribers.

Successful deployment of FTTx networks, from both a technology and financial standpoint, will require the ability to deliver video services effectively along with voice and data services. This "triple-play" of voice, video, and data is an essential revenue component for the future success of service providers. Today, many networks face

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

bandwidth constraints in the last mile, so delivering all of these services is either impossible or requires that service providers operate parallel networks. To compete successfully, service providers will need to transform their access networks to deliver all of these services via a single integrated transport and access solution.

Several network operators plan to deliver all services through a direct fiber connection to residences and businesses at some point in the future. A passive optical network architecture is one way to accomplish this goal. As previously mentioned, the key feature of the PON architecture is the elimination of all active electronics in the distribution network. Low-cost passive optical splitters are used to distribute an optical signal over multiple paths to several subscribers using a single fiber or fiber pair.

Alternative FTTx models require active electronics in outdoor field locations. In these architectures, an intermediate access multiplexer or a digital loop carrier is needed in the distribution portion of the access network. These architectures use an optical transport link from the central office or headend out to an active neighborhood node, from which fibers go to individual businesses and homes. The major disadvantage of these approaches is the use of active electronics. Active electronics require battery backup, decrease overall network reliability, and raise field maintenance costs, even when the electronics are enclosed in a controlled environmental vault or in a protected aboveground enclosure. Because a PON eliminates outside plant electronics, an upgrade to a PON, if required, is much easier and less expensive to accomplish. Moreover, active FTTx models face bandwidth constraints due to the intermediate electronics, while PONs essentially can offer unlimited bandwidth over the fiber itself. Current ATM-PON standards support 622 Mbps symmetrical

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

transport shared among up to 32 subscribers, and work is underway to extend that above 1 Gbps. Ultimately, the PON topology is the most cost-effective and flexible approach to building access networks and delivering VoD services.

From an equipment perspective, PON is rapidly reaching cost parity with traditional access systems based on synchronous optical network (SONET) or hybrid fiber/coaxial (HFC) architecture. The primary barrier to deploying PON is the initial installation cost of the fiber itself. But from a total cost of ownership perspective, PON can provide a much less expensive and more reliable network than traditional topologies in use today. For all these reasons, industry analysts project strong growth for PON over the next few years.

Worldwide PON equipment sales are expected to reach 314 million in 2005, from 66 million in 2001, according to market researcher IDC (Framingham, MA) in an October 2001 report on PON equipment. Once the fiber is there, PON is the lowest-cost approach to delivering services over fiber. Data will be a big driver, but video services will add more revenue and make the value proposition stronger for the service provider.

The delivery of VoD services over PONs can be accomplished using one of two approaches: video overlay or in-band VoD. Both models offer unique benefits and have some disadvantages. They are described in detail in section 2.3.3.

2.3 Video-on-Demand over Passive Optical Networks

It is generally believed that interactive video-on-demand services will be one of the most demanding residential services to be provided in emerging high-speed access

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

networks. This section gives an overview of VoD services, the model of VoD over passive optical networks, and the main design issues for VoD over PONs.

2.3.1 Overview of Video-on-Demand Services

Video-on-Demand is a technology in which a viewer can order a specific video program from a large selection, specify the video's start time, and have it sent over a telecommunication network to the viewer's home. A VoD system integrates the entertainment, telecommunication, and computer industries and provides electronic video rental services to geographically dispersed users from remote video servers on a broadband network. Users are no longer restricted to being passive watchers. They are allowed to choose the program contents, to decide the viewing schedule, and to interact with the programs with such operations as pause, jump forward, speed-up, etc. Upon receiving the demand from user, video server must stream the encoded video data to the user's equipment through a broadband network such as PON.

Current video-on-demand systems can be clssified into three categories [3][25]: (1).Quasi Video-on-Demand (Q-VoD), (2).Near Video-on-Demand (N-VoD) and (3).True Video-on-Demand (T-VoD). Each is explained as follows.

- Quasi Video-on-Demand is a service in which users are grouped based on a threshold of interest. Users can perform rudimentary temporal control activities by switching to a different group. It is suitable for services where interaction is limited. A local controller, like set-top box can filter multiple channels to achieve the service.

- Near Video-on-Demand transmits video data repeatedly over multiple broadcast or multicast channels to enable multiple users to share a single video channel

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

so that system cost can be reduced. Functions like forward and reverse are simulated by transitions in discrete time intervals. There are some trade-offs such as limited interactive control, fixed playback schedule, and limited video selections.

- True Video-on-Demand provides full-function video cassette recorder (VCR) capabilities including fast forward and reverse play, pause, stop, and random positioning. The user has complete control over the session presentation. It requires the use of a bidirectional signal for each user to achieve short response times so that the user can select what video to play, when to pay it and perform interactive VCR-like controls at will. T-VoD is the most difficult to implement where service quality is maximized.

A generic architecture of video-on-demand is shown as figure ***2.3.*** The key elements in such a VoD system includes the video storage and the VoD server on the server side, the telecommunication network and the user's set-top box. Upon request, the VoD server can retrieve the encoded video streams from the video storage, and transport them into the network. The network provides the pathway between the VoD server and the subscribers. It transmits these video data to the subscriber's set-top box. The set-top box decodes the video streams and forwards them to the user's video display.

2.3.2 Design Issues in Video-on-Demand Services over PONs

The main issues in the design of VoD systems over PONs are (1) transmission of video data across the network and (2) multiple access to video resource on the server side. The first topic is searching for economic solutions to transmit video data across

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

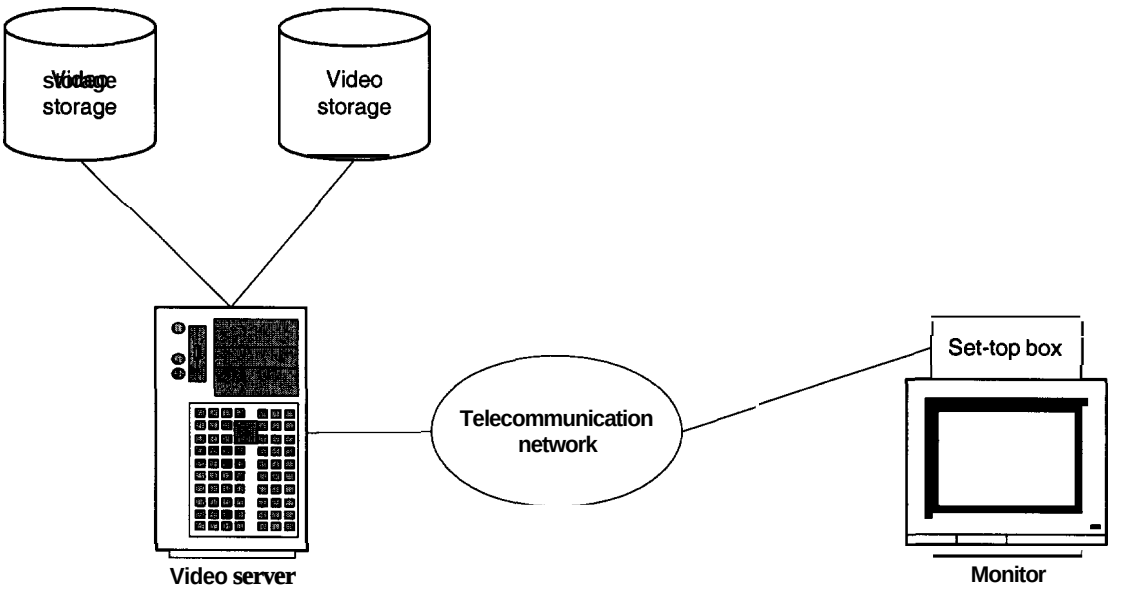


Figure 2.3: The general structure of a VoD system

the network. The requirements on this topic are as shown below.

- Low network transmission costs
- Easy traffic control approach
- Guaranteed QoS
- Efficient bandwidth utilization

The second topic is to research on the algorithm for simultaneously multiple access to video resource to achieve the maximum utilization of the video database.

However the two topics are closely related. Multiple access to video resource on the server side depends on the underlying efficient transmission of video data across the network.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

This thesis focuses on the first topic, transmission of video data across the network, i.e., video retrieval techniques, which is one of the most important issues to support VoD services over PONs.

Issues of video retrieval technique includes two different kinds:

1. Transmission of live real-time video data across the network;
2. Transmission of pre-recorded video data across the network.

The delivery of video data to the subscribers can be provided in several ways. To reserve bandwidth upon request is the most expensive way, but surely safeguard the QoS. However, it results in very conservative estimates of the actual bandwidth resources required. On the other hand, another type is to deliver the data in a best-effort mode. This is the simplest approach but resulting in no guarantees in QoS. These two can be considered as two extreme method. Between them, there are two techniques, statistical multiplexing, and bandwidth smoothing techniques.

The statistical multiplexing technology is often used in the transmission of live real-time video data. It reserves bandwidth for a video channel very near to the mean of the expected video frame size. In this way, when the bit rate of the video streams are at the peak rate simultaneously, some video data will be delayed or even loss, resulting in low video quality at the subscriber. This means there is no safeguard for transmitting a frame.

On the other hand, bandwidth smoothing techniques attempt to achieve a lower data loss rate and a higher multiplexing gain through removing the rate fluctuation, i.e. video burstiness. However, the trade-off is additional delay. Smoothing uses buffering to reduce the variance in necessary bandwidth. It is useful and scalable for VoD system because it does not require that the time between video capture and

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

video playback to be minimized. As a result, bandwidth resource allocation can be designed before the video streams are transmitted. Thus, video channels can have bandwidth guarantees.

Design issues are also related to video compression technique since video are always compressed before they are stored and transmitted. In order to reduce the storage space requirement for archived video information and the bandwidth for transmission of the video information from one point to another, compression technologies have been developed for a variety of video applications, such as H.261 [17], H.263 [16], JPEG [15], and MPEG [1][2]. Among them, MPEG video compression is used in many current and emerging products. It uses several techniques to achieve high compression ratios with minimal impact on the perceived video quality. It has become a key technique for developing digital television set-top boxes, DSS, HDTV decoders, DVD players, video conferencing, Internet video, and other applications. MPEG-1 defines a framework for coding moving video and audio, significantly reducing the amount of storage with minimal perceived difference in quality. In addition, a system specification defines how audio and video streams can be combined to produce a system stream. This forms the basis of the coding used for the VCD format. MPEG-2 builds on the MPEG-1 specification, adding further pixel resolutions, support for interlace picture, better error recovery possibilities, more chrominance information formats, non-linear macroblock quantization and the possibility of higher resolution DC components.

To take advantage of temporal correlation between frames, MPEG compression algorithm uses three kinds of frames: Intra-coded (I frame), Predictive-coded (P frame) and Bidirectionally-predictive-coded (B frame). The use of P and B frames

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

typically causes traffic burstiness, which is one important feature of MPEG video. In this thesis, burstiness is one performance evaluate parameter which is discussed in the latter parts. In MPEG encoding, I frames are coded independently with no reference to any other frame in the sequence. Starting with an I frame, the MPEG encoder can forward predict a future frame. This is commonly referred to as a P frame, and it may also be predicted from other P frames, although only in a forward time manner. The B frames however, are coded based on a forward prediction from a previous I or P frame, as well as a backward prediction from a succeeding I or P frame. As an example of the usage of I, P, and B frames, a group of pictures (GOP) can be I, B, P, B, P, B, I, B, P, B, P, B. Figure 2.4 shows the actual MPEG2 encoding structure.

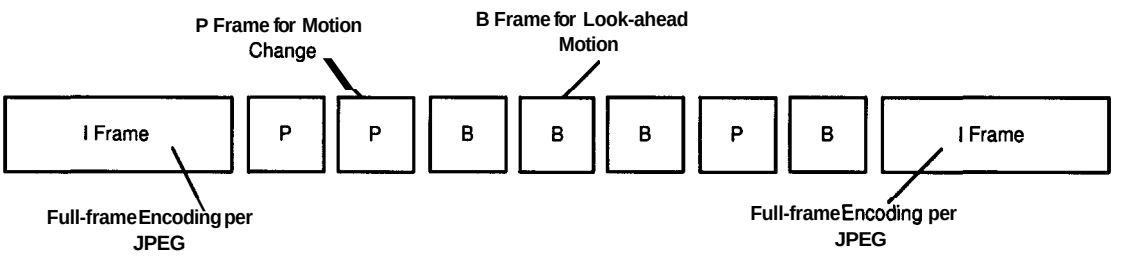


Figure 2.4: MPEG2 encoding structure

2.3.3 Models for Video-on-Demand Services over PONs

Two methods can be used to implement video-on-demand services over passive optical networks. One is video overlay model, the other is in-band VoD model.

In the case of video overlay model, the video data are transmitted through a broadband analog radio frequency (RF) signal in a separate optical wavelength. The voice and data services are delivered in different optical wavelength. It is shown as figure 2.5. It uses WDM to combine the video and other optical transport signals onto

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

a single fiber. The typical position for this video signal is in the 1550-nm region. The

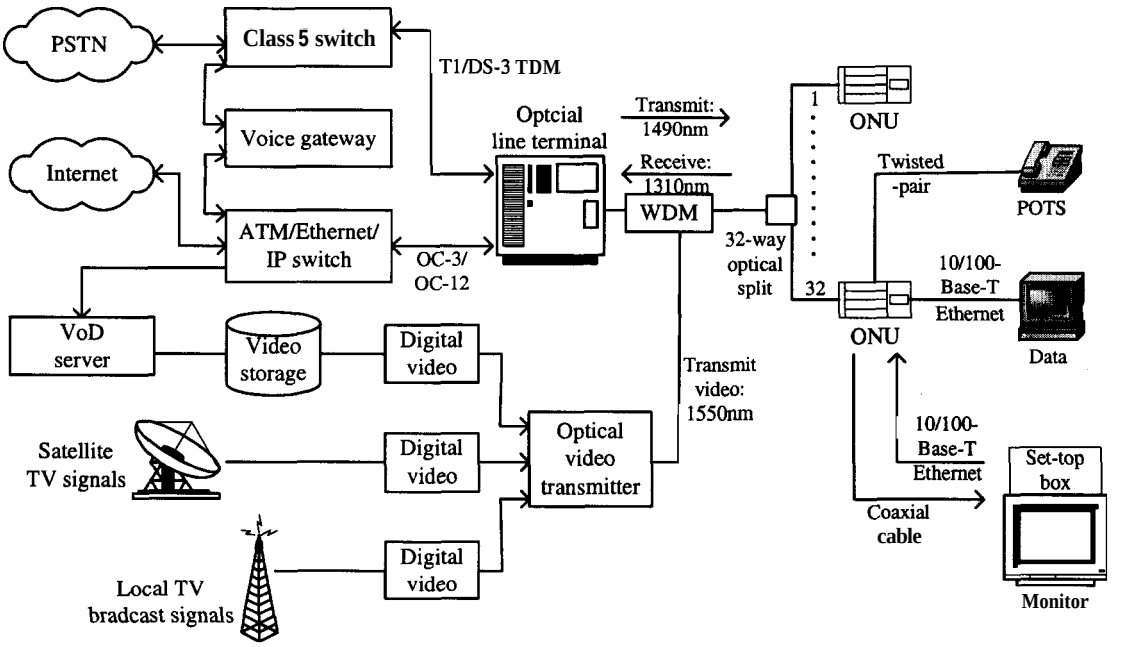


Figure 2.5: Video overlay model

other model, in-band VoD model, has been given as figure 2.6. In this case, the video services is carried in the same digital stream with other voice and data services on the fiber. It supports an integration of all service transporting over a single transport and access technology and does not require a separate optical overlay network and all the RF modulators and ONU receivers associated with it. Therefore, this model is more cost-effective than the above one. In this thesis, the bandwidth allocation scheme is feasible for the in-band VoD model.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

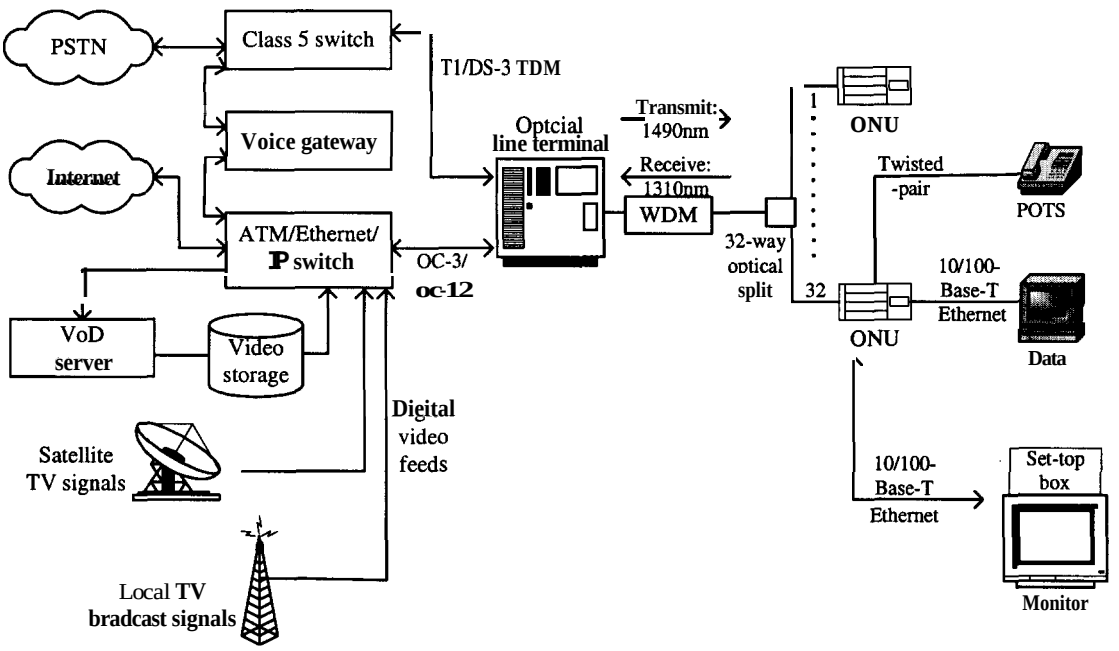


Figure 2.6: In-band VoD model

2.4 Dynamic Bandwidth Allocation for Video-on-Demand Services over PONs

For a video-on-demand system over passive optical network, the downstream of the link should carry large amount of video data, especially when the number of subscribers within the system increases gradually. Although PONs can provide more and more bandwidth with its evolution and have become the most promising solution for access network, the bandwidth allocated for video services is limited. This rises the problem of allocating bandwidth effectively for VoD services over PONs.

Besides the broad-band network, video-on-demand services also rely on data compression technologies. The MPEG algorithm uses two kinds of encoding, VBR and CBR. In CBR encoding, every MPEG frame has a constant number of bytes, through using a hypothetical rate control algorithm. Thus, the video quality varies in time.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

On the other hand, VBR encoding keeps all encoder parameters simply the same all the time. The number of bytes in each encoded frame varies as a function of the complexity of the scene, the extent of the scene change from frame to frame, and the type of MPEG frame used (I,B,or P). As a result, the constant quality VoD service can be achieved. However, for MPEG VBR-encoded video trace, the problem is that video compressed by this mode always exhibits multiple-time-scale rate variability [29]. This problem must be considered in traffic management.

In the case of live real-time video applications, decisions of bandwidth resource management must be made in real time, and the delay between the sender and the receiver must be limited. These constrain the video delivery method critically. While, for prerecorded video services, by using buffering techniques the system can take a rather flexible approach of data transmission. Since the entire video stream is known as a prior, network resource manager can plan the whole transmission method in advance. And thus both the loss of picture quality and wasting of network bandwidth resource due to overestimate of resource requirements can be minimized or avoided. Because the data for any playback frame can be provided either by the network or by a prefetch buffer, it is possible to overcome the burstiness by sending data to the buffer in advance of each burst, transmitting more bytes than needed and draining them when bursts happen. The degree that burstiness can be smoothed depends on the size of the prefetch buffer. The larger the buffer size, the more the possibility of removing the burstiness from the video streams. This, however, needs an entire design of buffer dimension and bandwidth allocation, to gain high bandwidth utilization and guaranteed QoS.

An approach, constant rate transmission and transport (CRTT), as mentioned in

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 1, has been proposed for pre-stored MPEG VBR-encoded video. It establishes a constant bit rate virtual channel to transmit video streams. But the trade-off is a very large prefetch buffer dimension in the viewer's set-top box, long delay before the starting of the playback, i.e., start-up delay and also very low bandwidth utilization.

In order to solve the problems of limited bandwidth resource, bandwidth utilization, and multiple-time-scale rate variability introduced by video compression, it is necessary to consider a dynamic bandwidth allocation for transporting pre-stored MPEG VBR-encoded video trace. In this thesis, we propose a dynamic bandwidth allocation and buffer dimensioning scheme for VoD services over passive optical networks to obtain high bandwidth utilization and to smooth video traffic burstiness. In the approach, pre-recorded MPEG video sequence is divided into M segments. The transmission of each segment is constant bit rate, but the rates vary across segments to match the intrinsic variable frame size in different segments. The algorithm employs the concept of tunnel to determine the transmission bandwidth to minimize the delay and the buffer capacity to avoid the underflow and overflow at the user's buffer. The performance with the new algorithm is evaluated by a number of real MPEG video sequences and the results show that the proposed algorithm can successfully make use of the advantages of both variable bit rate compression and those of constant bit rate transport with simplicity.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 3

Dynamic Bandwidth Allocation

3.1 Introduction

It has been described in chapter 2 that bandwidth allocation is a crucial issue for delivery of pre-stored MPEG video trace. This is because it can greatly reduce bandwidth requirement for transport of video streams, and improve bandwidth utilization. Dynamic bandwidth allocation can also reduce the traffic burstiness introduced by the natural characteristic of compressed video data.

The goal of this chapter is to build up the mathematical model of video-on-demand services over passive optical networks and to analyze the model to propose suitable dynamic bandwidth resource management methods. First, we set up the transmission model to support the delivery of video data from optical line terminal to optical network units. Then the dynamic bandwidth allocation scheme and buffer occupancy are studied with the constrains of no user buffer overflow or underflow. Two segmentation schemes are also provided for dynamic bandwidth allocation. The first one is equal segment scheme, which is easy for calculation. The other one is dynamic segment scheme based on intrinsic video date characteristics.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.2 Related Works

It has been addressed in chapter 2 that compression technology has been widely used for massive video data to reduce the storage space requirement and efficiently make use of transmission resources. Among many compression standards, MPEG is a commonly accepted standardized compression platform for many digital video services. There are basically two encoding modes for MPEG algorithm: constant bit rate (CBR) and variable bit rate (VBR). VBR has a number of advantages over CBR mode in transmitting over passive optical networks. It is determined by the fact that VBR encoding can keep all encoder parameters the same all the time to obtain consistent quality video in an open-loop operation. By using this mode, statistical multiplexing gain can be achieved by multiplexing multiple video streams during transmission on the network. Furthermore, significant quality gain can be provided on the user side.

There are basically two categories of issues related to the MPEG VBR-encoded video data retrieval,

- transmission of live real-time video data across the network
- transmission of pre-recorded video data across the network

The most widely used video transmission technologies in these two categories including,

- best-effort transport
- bandwidth reservation
- statistical multiplexing

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

- batching
- bandwidth smoothing

The simplest method to transport video is in a best-effort fashion. In this mode, there is no QoS guarantees and data can be arbitrarily delayed or even dropped. This method has been studied by K.Jeffay in [20]. Bandwidth reservation at the peak bandwidth requirement is also a simple scheme, but it results in too conservative estimates of the actual resources required and very low bandwidth utilization.

Statistical multiplexing technology has been introduced to take advantage of the law of large numbers of video traces. This approach is commonly used for transporting live real-time video steams. In [24] a comprehensive description of transporting strategies for statistically multiplexing live real-time VBR video traffic was presented. The common problem of statistically multiplexing VBR-encoded video is that the multiplexed traffic may suffer from data delay or even data loss in case that each of the video steams in multiplexing reaches its peak rate at the same time. Many approaches have been proposed in order to overcome this problem and guarantee the video quality at the end user side. Study in [26] proposed an aggregation method to multiplex video data statistically at the server side instead of on the network. This approach requires more processing for multiplexing and demultiplexing. Another paper [6] provided the prioritized transmission for video streams in which data were divided into different priorities according to their importance and transported or discarded selectively based on their priority level if bandwidth resource is deficient. For pre-recorded MPEG VBR video streams, there were also some statistical multiplexing schemes proposed. S.C. Liew have recommended an approach called lossless aggregation (LA) to transmit multiple pre-stored VBR video streams over a shared

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

communication channel in [27]. This scheme can achieve significant receiver-buffer reduction. But it is only applicable for the conditions that video streams are targeted for the same destination.

Batching is another approach commonly used to support video transmission. It simultaneously transport multiple copies of each video that are staggered in time and have different logical playback times by means of multicast technique. Batching can be efficient and cost-effective for video applications, however, the trade off is that the user who requests a programme has to wait until the start of one of the staggered copies. And batching can only apply to support Near-VoD system which provides limited interactive control and limited video selection.

Unlike statistical multiplexing and batching schemes, bandwidth smoothing schemes attempt to manage bandwidth based on a stream-by-stream mode to overcome the drawbacks of other schemes. It reduces traffic burstiness through introducing delay to decrease the requirements for bandwidth. Through pre-buffering some video data before the data enter the network, it reduces the variance in necessary bandwidths. For live real-time video applications over PONs, any bandwidth allocation or traffic shaping must be done on-line in real-time. Since delay between video capture and video playback is limited, live video applications may have to settle for weakened guarantees of services or for some degradation in quality of service when bandwidth smoothing techniques are used. Therefore, it will be difficult for such live real-time video transmission over PONs to be benefited from the bandwidth smoothing mechanisms due to large start-up delay. Smoothing for live video streams has been studied in [41] and [8] and they have shown that smoothing can obtain a lower cell loss rate and is able to achieve a higher multiplexing gain. On the other hand, for pre-recorded video

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

applications, the case is different with live real-time video application. The network bandwidth resources can be scheduled well in advance of the playback of the video stream because the entire video stream is know prior to the playback. Some smoothing mechanisms for pre-stored video have been proposed such as store-and-forward (SAF) [3], constant rate transmission and transport (CRTT) [30], 2-phase smoothing [21], re-negotiation constant bit rate (RCBR) transmission model [43], and re-negotiation deterministic variable bit rate (RED-VBR) transmission model [46]. SAF and CRTT schemes can effectively support video transmission over networks but the problem is they both need a playback buffer with extremely large capacity and introduce long delay before the playback starts. To effectively reduce the playback buffer capacity requirement and start-up latency, S. Kang and H.Y. Yeom have proposed a constant bandwidth allocation scheme which employs 2-phase smoothing technique to covert the VBR video into an on-off signal with fixed bandwidth. This scheme can reduce the buffer capacity largely but it introduces the problem of wasting bandwidth at off states. In order to reduce buffer capacity and long start-up delay while increasing the bandwidth utilization, several algorithms have been proposed in [43] and [46]. RCBR transmission model used in [43] is easy to implement due to the re-negotiation added to the static CBR service. The disadvantage of it is that the admission control is potentially complex. RED-VBR service is based on flexible renegotiation of traffic parameters with graceful degradation of QoS in the case that renegotiations fail. [46] recommended a mechanism in which the video sequence is divided into some segments and each segment are transmitted by RED-VBR mode. By contrast, we propose a novel scheme in this thesis, dynamic bandwidth allocation, to allocate bandwidth dynamically and choose the playback buffer size for delivery per-stored video streams

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

in Video-on-Demand systems over PONs.

3.3 Transmission Model of Video-on-Demand over Passive Optical Networks

3.3.1 Transmission Model

Given a stored MPEG video stream, consisting of N frames, which has totally C bytes. Figure **3.1** shows the diagram of transmission of this pre-recorded MPEG video sequence across a passive optical network adopted in our algorithm. On the server side, the MPEG video sequence, is divided into M segments, with a corresponding length of $L_1, L_2, L_3, \cdots, L_m, \cdots, L_M$ bytes, respectively. A sample MPEG video sequence is shown in figure **3.2.** The bandwidth allocated to the segment L_m is b_m therefore the transmission time for the mth segment is L_m/b_m.

On the user side, a receiving buffer has a capacity of B bytes and $O(t)$ represents the buffer occupancy at time t. $S(t)$ and $P(t)$ represent the cumulate video data being transmitted to the user receiving buffer and the cumulate video data having been played back at time t, respectively. The playback time is evaluated using frames as the unit, because video sequence is normally played back at a fixed rate, say F frames per second. Throughout this report, the frame period **1/F** is used as the unit time.

The relationship between $S(t)$, $P(t)$ and the buffer occupancy $O(t)$ is shown in figure **3.3** for M=4, where we assume that the delay and delay jitter introduced by the network for the arrival stream $S(t)$ is negligible. Clearly we have $O(t) = S(t) - P(t)$. $T(m)$ is the time when the transmission of the first m segments is completed, and frame $l(m)$ is the last video frame contained in the mth ($m = 1, 2, \cdots$, M) segment,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

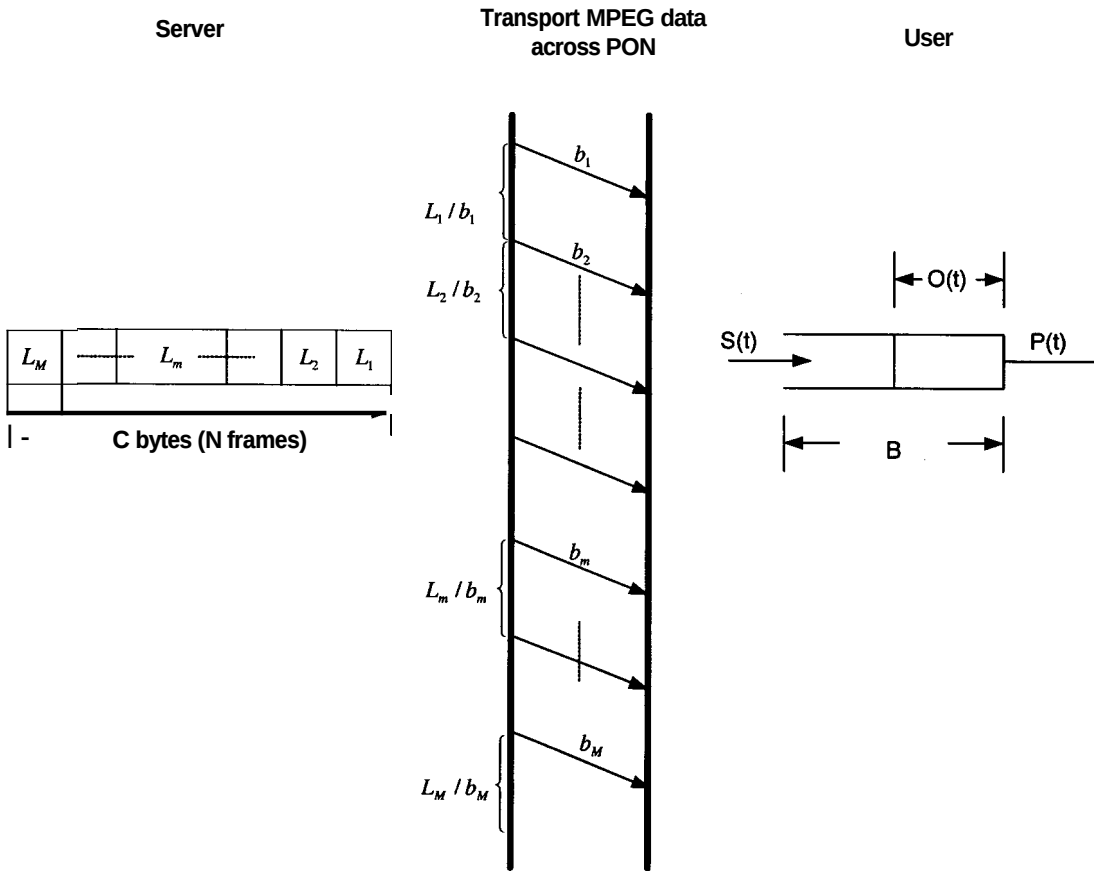


Figure 3.1: Transmission of pre-recorded MPEG video over PON

which satisfies

$$0 = l(0) < \mathbf{Z(1)} < l(2) \cdot \ldots < l(m) < \cdots < l(m - \mathbf{1}) < l(M) = \boldsymbol{N}. \qquad \textbf{(3.3.1)}$$

Therefore, the total number of video frames contained in the mth segment is $l(m) - l(m-1)$. The ith frame requires x_i bytes of storage, then the video data in bytes contained in the mth segment is given by

$$L_m = \sum_{i=l(m-1)+1}^{l(m)} x_i, \; m = 1, 2, \ldots, \boldsymbol{M}. \qquad (3.3.2)$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

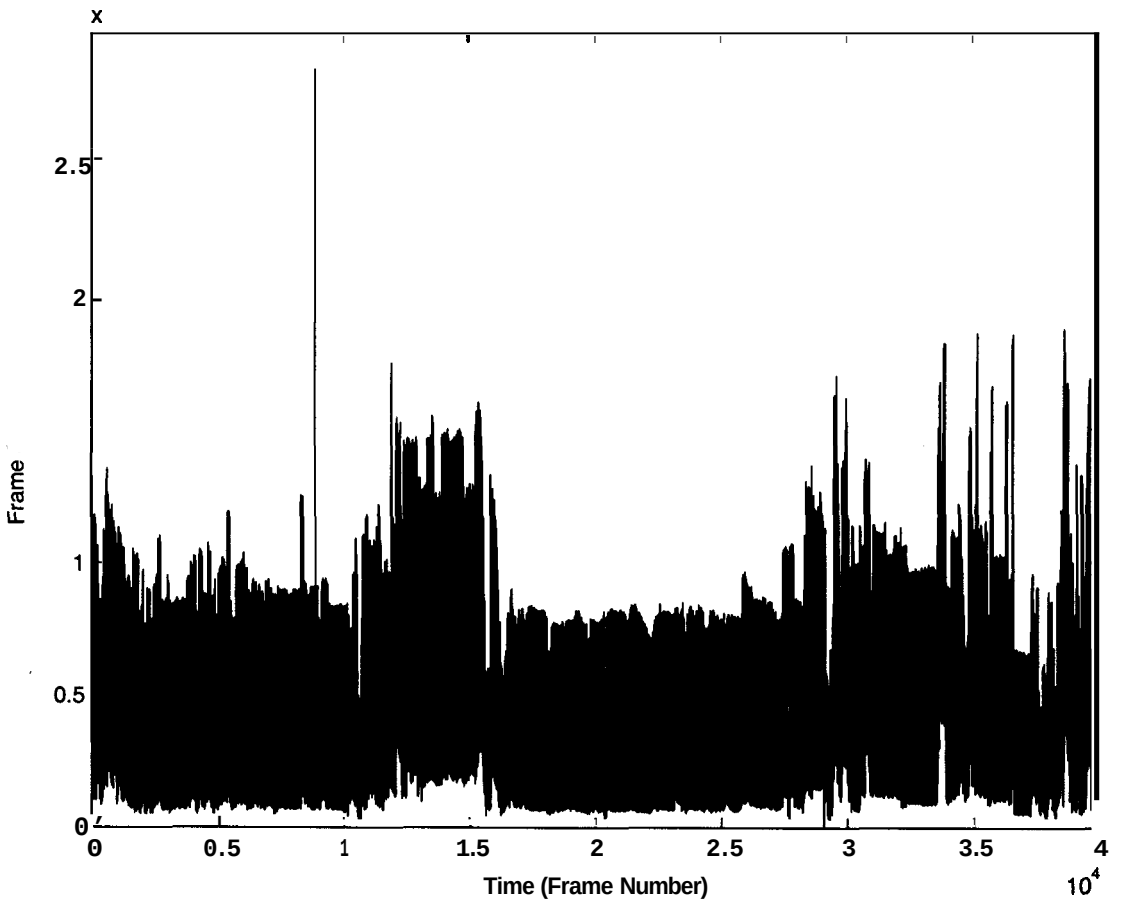


Figure 3.2: MPEG video sequence "Movie-Lambs"

To avoid the underflow of receiving buffer at the beginning of playback, a delay of d frames is considered between $S(t)$ and $P(t)$, say start-up playback delay.

The arrival stream $S(t)$ can be expressed as

$$S(t) = \begin{cases} 0, & t \leq 0, \\ \sum_{i=1}^{m-1} L_i + b_m[t - T(m-1)], & T(m-1) < t \leq T(M),\ m = 1, 2, \ldots, M, \\ \sum_{i=1}^{M} L_i, & t > T(M). \end{cases} \tag{3.3.3}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

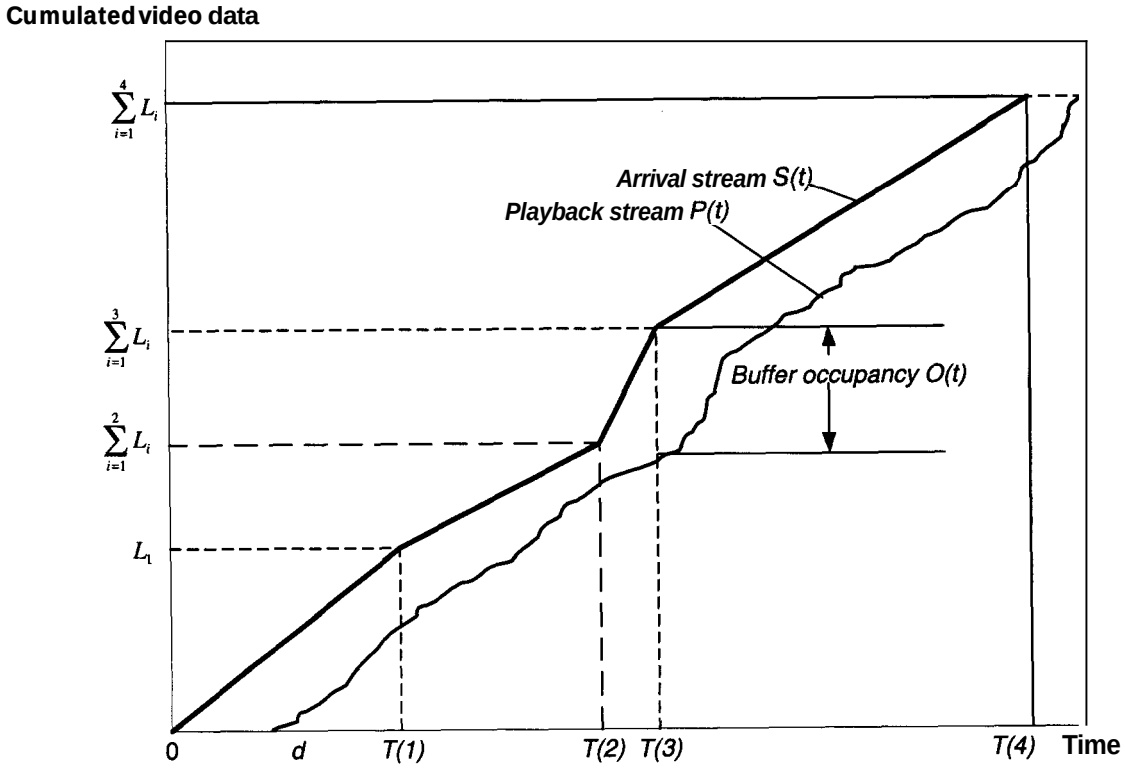


Figure **3.3:** The relationship between arrival stream, playback stream and buffer occupancy

Where term $\sum_{i=1}^{m-1} L_i$ represents the cumulate video data that have been transported to the user buffer during time period $[T(0), T(m-1)]$. The term $b_m[t-T(m-1)]$ represents the data having been transmitted using an allocated bandwidth in the period of the mth segment during $[T(m-1) < t \leq T(m)]$.

As the MPEG video is played back at a fixed rate of F frames/s, the playback stream $P(t)$ can be determined by the data contained in video frames. Then, $P(t)$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

can be described as

$$P(t) = \begin{cases} 0, & 0 < t \leq d, \\ \sum_{i=1}^{t-d} x_i, & d < t \leq N + d, \\ \sum_{i=1}^{N} x_i, & t > N + d. \end{cases} \tag{3.3.4}$$

Where term $\sum_{i=1}^{t-d} x_i$ represents the data in bytes that has been played back during time period $[d, t]$. Then we can obtain the formula for $O(t)$ by $O(t) = S(t) - P(t)$, that is

$$O(t) = \begin{cases} 0, & t \leq 0, \\ \sum_{i=1}^{m-1} L_i + b_m[t - T(m-1)], & 0 < t \leq d, \\ \sum_{i=1}^{m-1} L_i + b_m[t - T(m-1)] - \sum_{i=1}^{t-d} x_i, & d < t \leq T(M) \\ \sum_{i=1}^{M} L_i - \sum^{t-d} x_i, & T(M) < t \leq N + d, \\ 0, & t > N + d \end{cases} \tag{3.3.5}$$

3.3.2 Overflow and Underflow Constraints

It is the most important constraint for transmission of video data over PONs to avoid overflow and underflow during the whole playback period at the user side. To ensure continuous playback at the user side, the server must always transmit quickly enough to avoid buffer underflow. That is to say, $P(t)$, which indicates the amount of data consumed at the user side by $(t - d)$ frames, must be always less than the amount of data arrived. On the other hand, the clients usually have limited buffer resources. In such situations the bandwidth allocation plan should be created so that the client should not receive more data than $P(t) + B$, thus to prevent playback buffer overflow.

On the other hand, to avoid both underflow and overflow during the transmission, the receiving buffer occupancy must be always greater than zero and less than the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

maximum buffer capacity B, that is

$$0 \leq O(t) \leq B. \tag{3.3.6}$$

In the following part of this chapter we are trying to find a traffic management scheme of pre-stored video from server to user side across passive optical networks under the constrain of underflow and overflow, by means of the completely known traffic characteristics of the pre-stored video streams, in order to support high network utilization.

3.4 Bandwidth Allocation and Buffer Occupancy

3.4.1 Minimum Transmission Bandwidth without Buffer Underflow

To avoid the receiving buffer underflow, it requires $O(t) \geq 0$ during the playback time $t \in [d, T(M)]$, according to equation 3.3.6. From equation 3.3.5, we can get

$$\sum_{i=1}^{m-1} L_i + b_m[t - T(m-1)] - \sum_{i=1}^{t-d} x_i \geq 0, \ 0 \leq t \leq T(M). \tag{3.4.1}$$

Thus we can obtain that

$$b_m \geq \frac{\sum_{i=1}^{t-d} x_i - \sum_{i=1}^{m-1} L_i}{t - T(m-1)}, \ m = 1, 2, \cdots, M. \tag{3.4.2}$$

Let $k = (t-d)F$, and $y(k) = \sum_{i=1}^{k} x_i$, then the minimum transmission bandwidth for the transmission of the mth segment of the video stream to guarantee no receiving buffer underflow is given by

$$b_{m,min} = \max_{T(m-1)+1 \leq k \leq l(m)} \left\lceil \frac{y(k) - \sum_{i=1}^{m-1} L_i}{k + d - T(m-1)} \right\rceil, \ m = 1, 2, \cdots, M. \tag{3.4.3}$$

The symbol $\lceil z \rceil$ denotes the smallest integer that is greater than z, where z is a real number.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.4.2 Minimum Buffer Capacity Required to Avoid Overflow

To avoid user receiving buffer overflow, the buffer capacity $\boldsymbol{B}$ must satisfy:

$$\boldsymbol{B} \geq \max_{0 \leq t \leq N+d} \{O(t)\}. \tag{3.4.4}$$

From equation 3.4.4, it follows that,

$$B_{min} = \max_{d \leq t \leq T(M)} \{\sum_{i=1}^{m-1} L_i + b_m[t - T(m-1)] - \sum_{i=1}^{t-d} x_i\} \tag{3.4.5}$$

Here B_{min} is the minimum buffer capacity required to avoid overflow.

3.4.3 Maximum Transmission Bandwidth without Buffer Overflow

From equation 3.4.5, the transmission bandwidth to guarantee no receiving buffer overflow is given by:

$$b_m \leq \frac{B + \sum_{i=1}^{t-d} x_i - \sum_{i=1}^{m-1} L_i}{t - T(m-1)} \tag{3.4.6}$$

Let $k = t - d$, and $y(k) = \sum_{i=1}^{k} x_i$. Therefore, the maximum transmission bandwidth $b_{m,max}$, for the transmission of mth segment in the video stream to guarantee no receiving buffer overflow is given by

$$b_{m,max} = \min_{T(m-1)+1 \leq k \leq h(m)} \left\lfloor \frac{B + y(k) - \sum_{i=1}^{m-1} L_i}{k + d - T(m-1)} \right\rfloor, \quad m = 1, 2, \cdots, M. \tag{3.4.7}$$

where $\lfloor z \rfloor$ rounds z to the largest integer that is smaller than z and $\min_{T(m-1)+1 \leq k \leq h(m)} \lfloor z \rfloor$ is the minimum $z(k)$ when the variable k increases from $T(m-1)+1$ to $h(m)$, where $h(m)$ is the cross point of the mth segment with the curve $P(t) + B$.

Finally, combining equation 3.4.3 and 3.4.7, to guarantee neither underflow nor overflow at the user receiving buffer, the required transmission bandwidth for the mth

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

video segment must satisfy,

$$b_{m,min} \leq b_m \leq b_{m,max} \tag{3.4.8}$$

3.5 Tunnel Scheme

There are three curves illustrated in figure 3.4, i.e., arrival stream $S(t)$, playback stream $P(t)$ and the maximum amount of cumulate video data $P(t) + B$ that can be delivered to the receiving buffer at time t. The area between $P(t)$ and $P(t) + B$ is defined as a tunnel, where the lower bound is the playback stream $P(t)$ and the upper bound is $D(t) + \mathrm{B}$. The arrival stream $S(t)$ must fall within the tunnel so that underflow and overflow of the user–receiving buffer can be avoided. It can be seen that the mth segment has three cross points of $h(m), s(m)$ and $l(m)$ with $P(t)+B, S(t)$ and $P(t)$, respectively. Obviously, the transmission bandwidth to the following $(m+1)$th segment can be determined by the slope of the line from $s(m)$ to $s(m+1)$. The slope of the line from $s(m)$ to $h(m+1)$ represents the maximum bandwidth without buffer overflow, i.e., $b_{m,max}$. Likewise, the slope of the line from $s(m)$ to $l(m+1)$ is the minimum bandwidth without buffer underflow, i.e.,$b_{m,min}$. If the $s(m+1)$ falls within the range of $h(m+1)$ and $l(m+1)$, then both underflow and overflow can be avoided. Hence, the bandwidth allocated to the mth segment can be dynamically determined by the slope of the line from $s(m)$ to $s(m+1)$ which must be located within the angle between $b_{m,min}$ and $b_{m,max}$.

The selecting of b_m can be improved by the following two approaches. When playback buffer capacity B increases, the tunnel becomes wider and more selective space for b_m is available. On the contrary, when segment size L, is reduced, the angle between $b_{m,min}$ and $b_{m,max}$ becomes wider which also provides more selective space

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

for ϕ .

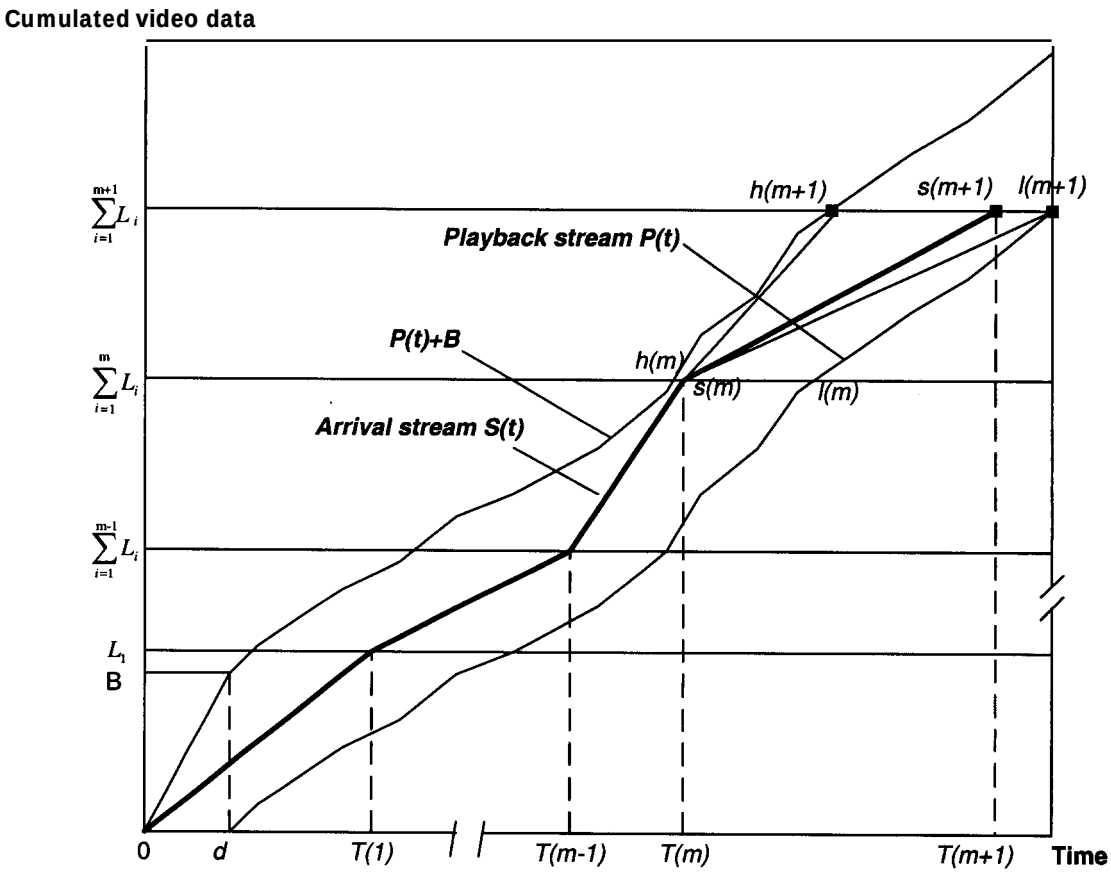


Figure 3.4: Playback tunnel

3.6 Segmentation Schemes for Dynamic Bandwidth Allocation

Although we have derived bandwidth allocation method for general cases, however, the segment length L, is a random variable due to the nature of MPEG video. This may raise a problem of how to determine the optimum segment size. In this section, we first conduct a survey on intrinsic video rate characteristics and video

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

segmentation schemes, and then we propose two segment division methods applicable for our dynamic bandwidth allocation scheme: equal segment, and dynamic segment.

3.6.1 Related Works on Video Segmentation Techniques

To review the video segmentation schemes that have been recommended, it is important to have a look at the research on intrinsic video rate characteristics first. Video information source characteristics have crucial impact on network performance. When the number of sources is sufficiently large and the traffic intensity is low, the performance of network depends mainly on encoded information distribution. On the other hand, when the number of information sources is relatively not very large or traffic intensity is high, then network manager must consider the time domain characteristics of encoded information such as distribution, autocorrelation, coefficient of variation, and so on. Some other parameters, like duration of peaks and intervals of scene changes, which characterize the unsteadiness of encoded information, are also commonly used as important characteristics measures.

The earliest investigation about statistical characteristics of video traffic was taken by B.G. Haskell [11]. It studied the probability density of frame differences for single source and multiple sources of traffic from conditional replenishment-type picture-phone coders. In [42], the authors evaluated the bit rate of VBR-coded video traffic from video-conference and broadcast TV scenes. H.S. Chin et al. studied the statistics of frame difference and the lengths of moving area cluster for viewphone-type video signals in their paper [13]. Y. Yasude et al. [44] evaluated scene duration distribution and bit-rate variation from scene to scene. In [36] the statistics of packet generation interval for a VBR video codec were achieved for small-motion, medium-motion, and large-motion activity scenes. Some other related work about statistical

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

characterization of VBR video can also be found in [37], and [38].

The earlier research mentioned in the previous paragraph have shown that traffic burstiness is a significant characteristic of compressed video. A burst source occasionally transmits at peak rate which is much higher than its long term average rate. In this case, small number of indexes may be sufficient for characterizing video source, such as peak rate, average rate, deviation-to-average ratio, and peak-to-average ratio. Besides traffic burstiness, the presence of traffic variations over multiple time-scales is also an important characteristic of compressed video source. Intuitively, there is a variation in source rate not only over a period from milliseconds to seconds, corresponding to variations within scene, but also over a period from tens of seconds to minutes, corresponding to scenes with considerable motion or rapid chrominance and luminance changes. The traffic variations over multiple time-scales of videos can be reflected from both the video trace and its corresponding playback stream.

Based on the study of intrinsic video rate characteristics, several video segmentation mechanisms have been provided to detect scene changes in VBR video traffic. P. Manzoni et al. [33] have proposed a scheme that scene change occurs when the difference between adjacent frames is greater than a threshold. In order to avoid noise, a variant on the previous technique was recommended in [35]. In this paper, the difference between adjacent frames was compared with the mean value of the W last frames. The main idea in [12] was that the video trace is divided into N clusters which represent N types of scenes. The division was implemented by means of a minimization of the distances to N centroids within the state space. An iterative search algorithm was used to optimize the position of the N centroids. Some other techniques on video segmentation can be found in [31],[9].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The above techniques that have mentioned are commonly proposed for setting up the model of variable bit rate (VBR) video traffic. However, those schemes are based on the video trace, so that they can not effectively overcome the limitation and disturbance of the MPEG encode format using **IBP** frames. In this section, we present two segmentation schemes. The first one, equal segment scheme, is simply divided video traces into equal length segments. While, the other one, dynamic segment scheme, is based on the playback stream instead of the video trace. Intuitively, the dynamic segment scheme should obtain more efficiency in using dynamic bandwidth allocation.

3.6.2 Equal Segment Scheme

Under the equal segment scheme, video sequence is divided into M equal length segments, that is, each segment contains fixed video data of L bytes.

$$L_1 = L_2 =, \cdots, = L_M = L = C/M. \tag{3.6.1}$$

In this case, the minimum bandwidth allocated to the mth segment to guarantee no user buffer underflow can be obtained by replacing the term $\sum_{i=1}^{m-1} L_i$ in equation 3.4.3 with the new term $(m-1)L$, that is,

$$b_{m,min} = \max_{T(m-1)+1 \le k \le l(m)} \lceil \frac{y(k) - (m-1)L}{k + d - T(m-1)} \rceil, \ m = 1, 2, \cdots, M. \tag{3.6.2}$$

Likewise, the maximum bandwidth that can be allocated to the mth segment to guarantee no user buffer overflow can also be obtained from equation 3.4.7, that is,

$$b_{m,max} = \min_{T(m-1)+1 \le k \le h(m)} \lfloor \frac{B + y(k) - (m-1)L}{k + d - T(m-1)} \rfloor, \ m = 1, 2, \cdots, M. \tag{3.6.3}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The minimum buffer capacity to guarantee no user buffer overflow under this scheme can be get from equation 3.4.5, i.e.,

$$B_{min} = \max_{d \le t \le T(M)} \{(m-1)L + b_m[t - T(m-1)] - \sum_{i=1}^{t-d} x_i\} \qquad (3.6.4)$$

Figure ***3.5*** shows the relationship between arrival stream $S(t)$, playback stream $P(t)$ and buffer occupancy $O(t)$ under equal segment scheme.

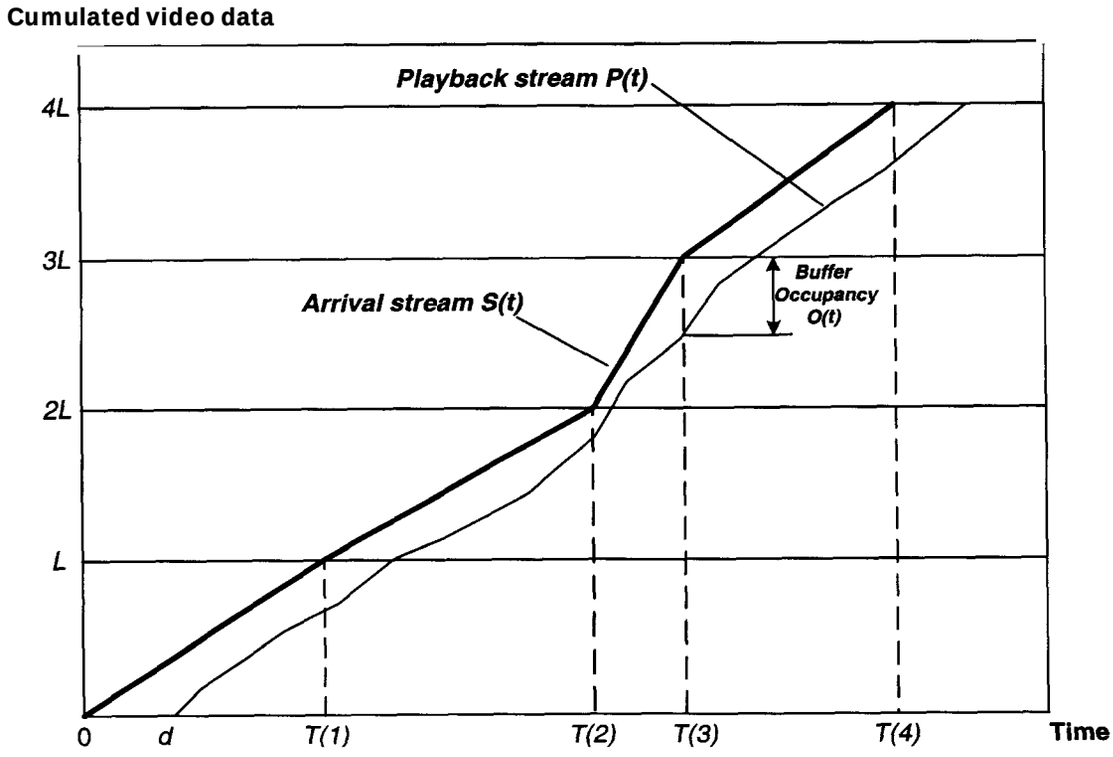


Figure 3.5: Relationship between arrival stream $S(t)$, playback stream $P(t)$ and buffer occupancy $O(t)$ under equal segment scheme

Based on the analysis as before, we get the following procedure for dynamic bandwidth allocation with fixed segment size.

1. Select the segment length $L_m, m = 1, 2, \cdots, M$. Let each segment length equal to C/M;

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2. Set $h(m), l(m)$ corresponding to L_m, $m = 1, 2, \cdots, M$;

3. Set $T(0) = 0$;

4. For $m = 1, 2, \cdots, M$, calculate $b_{m,min}$ over $[T(m-1), l(m)]$ using equation 3.6.2, and $b_{m,max}$ over $[T(m-1), h(m)]$ using equation 3.6.3;

5. Select bandwidth b_m between $b_{m,min}$ and $b_{m,max}$ according to the network conditions;

6. Output the mth segment with b_m;

7. Set transmission time with $T(m) = T(m-1) + \frac{L_m}{b_m}$;

8. The end.

3.6.3 Dynamic Segment Scheme

In this part we try to use a dynamic segment scheme to guarantee the lowest traffic burstiness and the least number of segment points by the information of burstiness of the video traffic contained in the playback stream. And in the same time, overflow and underflow are both avoided. The traffic burstiness is defined as a ratio of the peak rate to the average rate, which indicates the grade of difficulty to satisfy the QoS requirement. This segmentation scheme is proposed relying on the following insight: the segmentation is based on the long-term bursts of the video trace to match the intrinsic rate characteristics of the video playback stream much better, so that less buffer capacity is needed at the user side. Long-term burstiness in the video trace can be reflected from its corresponding playback stream. In detail, corresponding to a long-term burstiness in video trace, the slope of the playback stream $P(t)$ maintains

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

constant for a relatively long period, but jumps to another constant corresponding to another long-term burstiness.

Based on the above considerations, it can be seen that if the changing points of the long-term burstiness where big jumps happen can be found, and used as the segmentation points, then the video arrive trace $S(t)$ can match the video playback stream $P(t)$ very well and therefore the buffer size needed is minimized.

Figure 3.6 shows the technical principle of dynamic segment scheme, in which the arrival stream $S(t)$ represented using the wider line is located between the lower bound and the upper bound, i.e., the playback stream $P(t)$ and $P(t) + B$, where B represents the user buffer capacity. It can be seen from the figure that each segment of the arrival stream is the possible longest line in the playback tunnel under the conditions of without buffer underflow or buffer overflow.

In such a scheme, the subscriber can choose the kind of buffer with some capacity according to his or her economic conditions. When the user buffer capacity is smaller, the dynamic segment scheme uses more segments to make the arrival stream matching the tunnel. In this case the arrival stream $S(t)$ has more similar shape with the playback stream $P(t)$. On the other hand, if the user buffer capacity is larger and the tunnel is much wider, then the arrival stream will much more like a direct line and the segment number will be less.

The following procedure is for dynamic bandwidth allocation with dynamic segment scheme.

1. Set $T(0) = 0$;

2. For k from $s(m-1)$to the total frame number of the video data, calculate each $b_{m,min}$ using equation 3.4.3 and $b_{m,max}$ using equation 3.4.7 for each value of k,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

and compare the maximum one of $b_{m,min}$ with the minimum one of $b_{m,}$ until the former is larger than the latter, and set the division point $s(m)$ as $k-1$;

3. Select $b_{,}$ between the maximum value of $b_{m,min}$ and the minimum value of $b_{m,min}$ for the mth segment;
4. Output the mth segment with b_m;
5. Set transmission time with $T(m) = T(m-1) + \frac{L_m}{b_m}$;
6. The end.

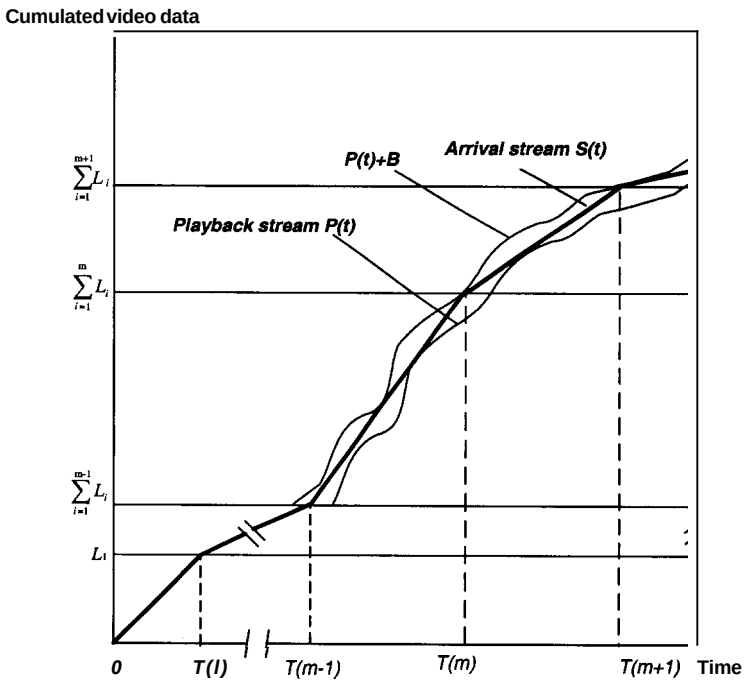


Figure 3.6: Playback tunnel under dynamic segment scheme

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 4

Performance Evaluation

4.1 Introduction

This chapter evaluates the performance of the dynamic bandwidth allocation scheme over passive optical networks. Through performance evaluation, the impact of the proposed schemes on video traffic and network resource management is studied in terms of parameters like traffic peak rate, traffic burstiness. The two schemes that have been provided in chapter 3 are evaluated and compared. The comparison of the proposed dynamic bandwidth allocation scheme under equal segmentation and the constant rate transmission and transport (CRTT) scheme are also studied. Real -life MPEG video traces with different contents including movie, sport, talk show, episode and cartoon are used in the performance evaluation. These MPEG video traces are encoded with the following parameters:

- Encoder input: 384x288pixels
- Color format: YUV (4:1:1, resolution of 8 bits)
- Quantization value: I=10, P=14, B=18

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

- Pattern: IBBPBBPBBPBB
- GOP size: 12
- Motion vector search: 'Logarithmic'/'Simple'
- Reference frame: 'Original'
- Slices: 1
- Vector/range: half pel/10
- Frame rate: 25 frames/second
- Video length: 40,000 frames

MPEG video streams	average rate (bytes/frame)	minimum rate (bytes/frame)	maximum rate (bytes/frame)	covariance	standard deviation
Mr.Bean	2205.9	43.0	28634.0	138720.0	2580.4
Lambs	913.9	36.0	16778.0	40801.0	1399.4
Soccer	3138.7	369.0	23787.0	147140.0	2657.6
Asterix	2793.6	38.0	18422.4	132090.0	2518.0
Talk show	1817.1	260.0	13346.0	88834.0	2064.9

Table 4.1: Statistical characteristics of the MPEG VBR-encoded video streams

The statistical characteristics of the MPEG VBR-encoded video streams are shown in table 4.1. It provides the video average rate, minimum rate and maximum rate in bytes/frame, as well as covariance and standard deviation of the video rate.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.2 Performance Evaluation under Equal Segment Scheme

In this section, the following numerical results are obtained based on a MPEG-1 video trace "Episode: Mr.Bean". The video trace of $N = 40000$ frames is divided into $M = 32$ equal segments. The video trace is played back at a rate of $F = 25$ frames/s. The delay is set at 0.48 seconds. The measured real-time video frame trace in bytes and the corresponding cumulated playback stream in bytes are shown in figure 4.1 and figure 4.2.

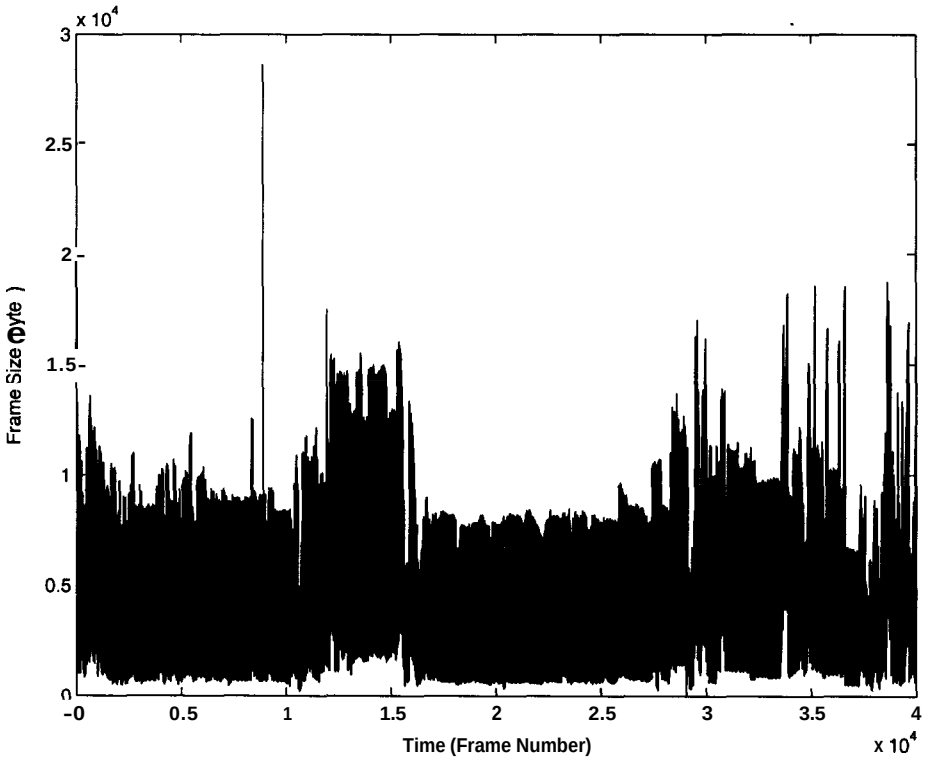


Figure 4.1: The playback MPEG video frame trace

Figure 4.3 shows the measured arrival stream $S(t)$ versus the playback stream $P(t)$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

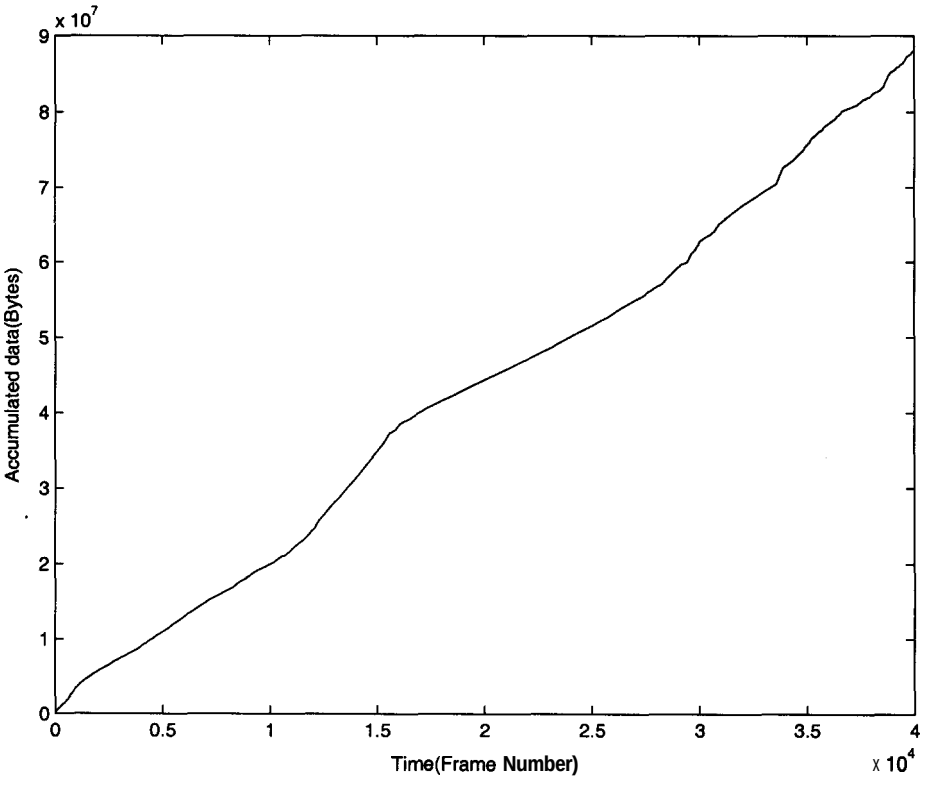


Figure **4.2:** The cumulated playback video data

where the equal segmentation is used. The broken line stands for playback stream, and the real line stands for arrival stream. The zoom-in of figure **4.3** is shown in figure 4.4. It can be seen that when the bandwidth allocated to $S(t)$ is chosen using equation 3.4.3, the playback stream $P(t)$ is always below the arrival steam $S(t)$, so that the underflow of user playback buffer is certainly avoided.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

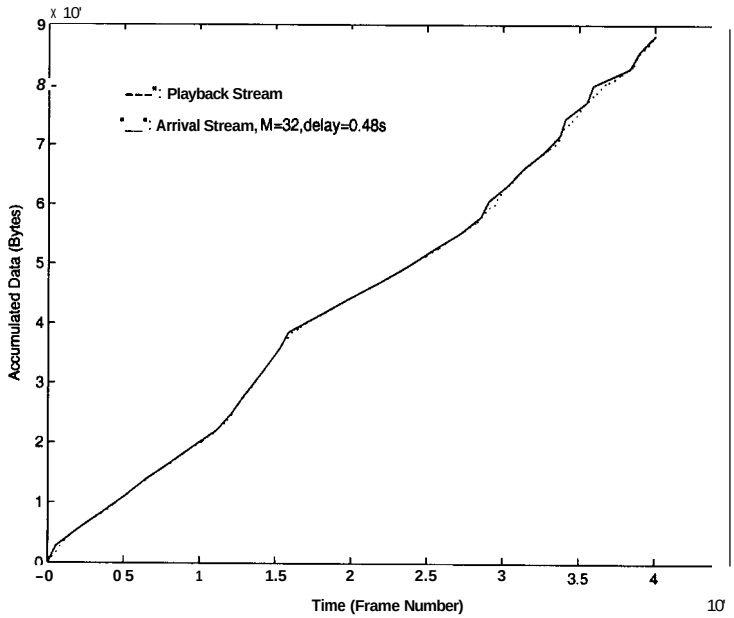


Figure 4.3: Arrival stream $S(t)$ vs. playback stream $P(t)$

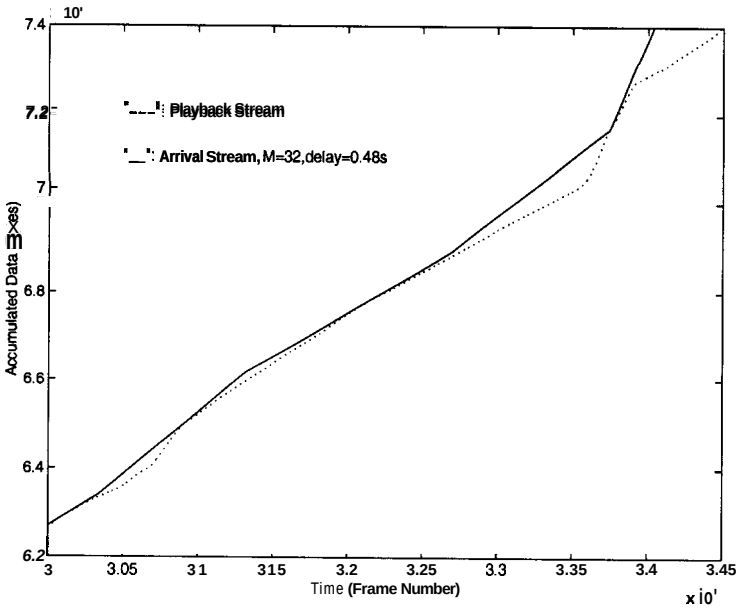


Figure 4.4: Zoom-in of Arrival stream $S(t)$ vs. playback stream $P(t)$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Figure 4.5 and figure 4.6 show the measured bandwidth allocation and corresponding buffer occupancy with the equal segmentation scheme. Figure 4.5 shows that the allocated bandwidth within one playback segment is a fixed value, and the bandwidth in different segment is variable according to the changing of data rate of playback stream. Comparing figure 4.5 with figure 4.1, it shows that the allocated bandwidth has matched real video traffic stream very well.

Figure 4.7 shows the minimum capacity required for the guarantee of no playback buffer overflow versus the number of segments being used for transporting of video data. It can be seen that the required buffer capacity decreases when the number of segment increases. This can be explained using the playback tunnel diagram as shown in figure 3.4. When the capacity of playback buffer is fixed, smaller segment size corresponds to a wider angle between the two lines whose slopes represent $b_{m,min}$ and $b_{m,max}$. On the contrary, if the angle between these two lines is fixed, then small segment size is corresponding to less buffer capacity. When the number of segments increases, the arrival stream matches the playback stream much more, i.e., there are a much wider angle between the two lines whose slopes represent $b_{m,min}$ and $b_{m,max}$, therefore the buffer size needed decreases.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

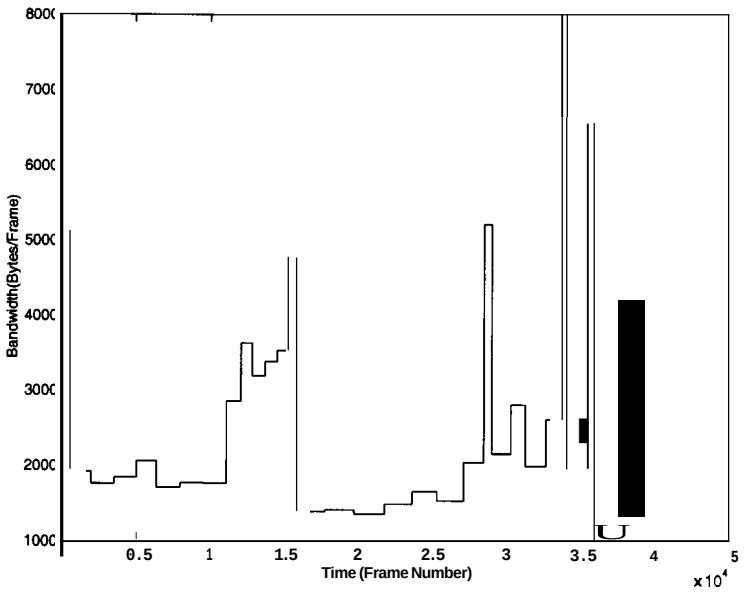


Figure 4.5: Bandwidth allocation

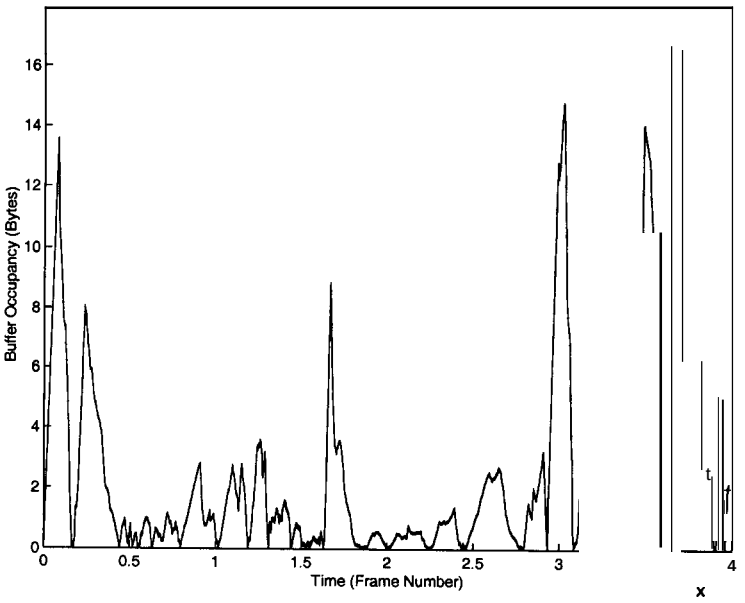


Figure 4.6: User's buffer occupancy

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

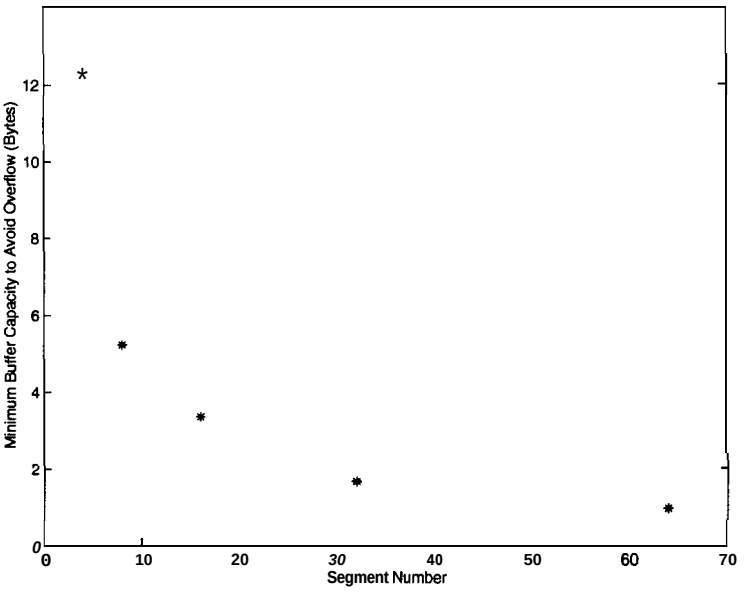


Figure 4.7: Playback buffer capacity vs. segment number

4.3 Comparison with CRTT Scheme

In this section we present a comparison of performance between the proposed dynamic bandwidth allocation scheme under equal segmentation and the constant rate transmission and transport (CRTT) scheme, proposed by Jean M.McMaunus and Keith W.Ross [30]. The basic idea of CRTT is to pre-transport a large amount of video data to the user playback buffer before the playback is commenced. Then CRTT establishes a constant bit-rate channel between the video provider and the viewer's set-top box and transmits data at a constant rate. The brief mathematical analysis is given below.

The main issue in CRTT scheme is to determine the transmission bandwidth as well as the user buffer capacity to avoid buffer overflow and buffer underflow. It specifies a given video stream $x_n, (n = 1, \cdots, N)$ with N frames, a fixed playback

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

rate of F frames per second and a user buffer with capacity B. For a fixed per-play start-up delay d, the minimum transport rate $b_{min}(d)$ which guarantees no starvation during playback is given by,

$$b_{min}(d) = \max_{d \le n \le N-1} \frac{F}{n} \left(\sum_{i=1}^{n+1} x_i - \sum_{i=1}^{d} x_i \right), i = 1, 2, \cdots, N. \quad (4.3.1)$$

The maximum transmission rate $b_{max}(d, \mathrm{B})$ with the given buffer capacity B and t fixed start-up delay d to guarantee no buffer overflow can be determined by,

$$b_{max}(d, B) = \min_{0 \le n \le N(B)-1} \frac{F}{n+1} \times \left(B + \sum_{i=1} x_i - \sum_{i=1} x_i \right), \ i = 1, 2, \cdots, N. \quad (4.3.2)$$

where N(B) is defined as,

$$N(B) = max\{n : \sum_{i=l}^{n} x_i + B < C\} \quad (4.3.3)$$

Moreover, $b_{m,min}$ and $b_{m,max}$ are both decreasing in d.

The minimum user buffer capacity required to avoid overflow under CRTT is given y,

$$B_{min} = min\{B : D(B) \neq \phi\} \quad (4.3.4)$$

where $D(B)$ is the set of all feasible frame build ups, and defined as,

$$D(B) = \{d : d \le d(B), b_{min}(d) \le b_{max}(d, B)\} \quad (4.3.5)$$

$d(B)$ is the set of all feasible start-up delay which is constrained by the user buffer ze, and it is defined as,

$$d(B) = max\{d : \sum_{i=1}^{d} x_i \le \mathrm{B}\} \quad (4.3.6)$$

The performance comparison between the constant rate transmission and trans- ort scheme and dynamic bandwidth allocation scheme with equal segment is given follows.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Figure 4.8 illustrates the arrival streams of CRTT and dynamic bandwidth allocation under equal segment scheme with a fixed start-up delay of 0.8s (20 frames). The dotted line in each figure represents the arrival stream, and the solid line represents the playback stream. Comparing the two results, it can be seen that the arrival stream of the dynamic bandwidth allocation scheme matches the video playback stream much better. In CRTT scheme, a very high bandwidth is set to avoid buffer overflow at the beginning of the transmission and this bandwidth is fixed during the whole transmission period. This causes the problem of very large buffer capacity requirement, and high bandwidth consumed over the network. By contrast with CRTT scheme, dynamic bandwidth allocation scheme shows more flexibility and effectiveness. It matches the playback stream better so that the buffer capacity required must be smaller. And it does not need a high bandwidth all the time during the transmission so that the bandwidth utilization can be greatly improved.

Figure 4.9 and figure 4.10 show the comparison of buffer occupancy between the proposed dynamic bandwidth allocation scheme and CRTT scheme. As the performance of CRTT scheme varies greatly with different values of start-up delay, we evaluate two situations. Figure 4.9 is the result with a delay of 0.8s (20 frames). Figure 4.10 is the result with 32s (800 frames) delay. It is easy to see that the buffer occupancy under dynamic bandwidth allocation scheme can be significantly reduced comparing to CRTT. Under CRTT scheme with delay of 0.8s, the maximum buffer occupancy is 45,097,000 bytes at the frame of 19144. While, with the same delay the maximum buffer occupancy under dynamic bandwidth allocation scheme is not more than 2,000,000 bytes which can be seen from figure 4.9. However, if with delay of 32s, the maximum buffer occupancy is 9,255,000 bytes at the frame of 29072 under

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

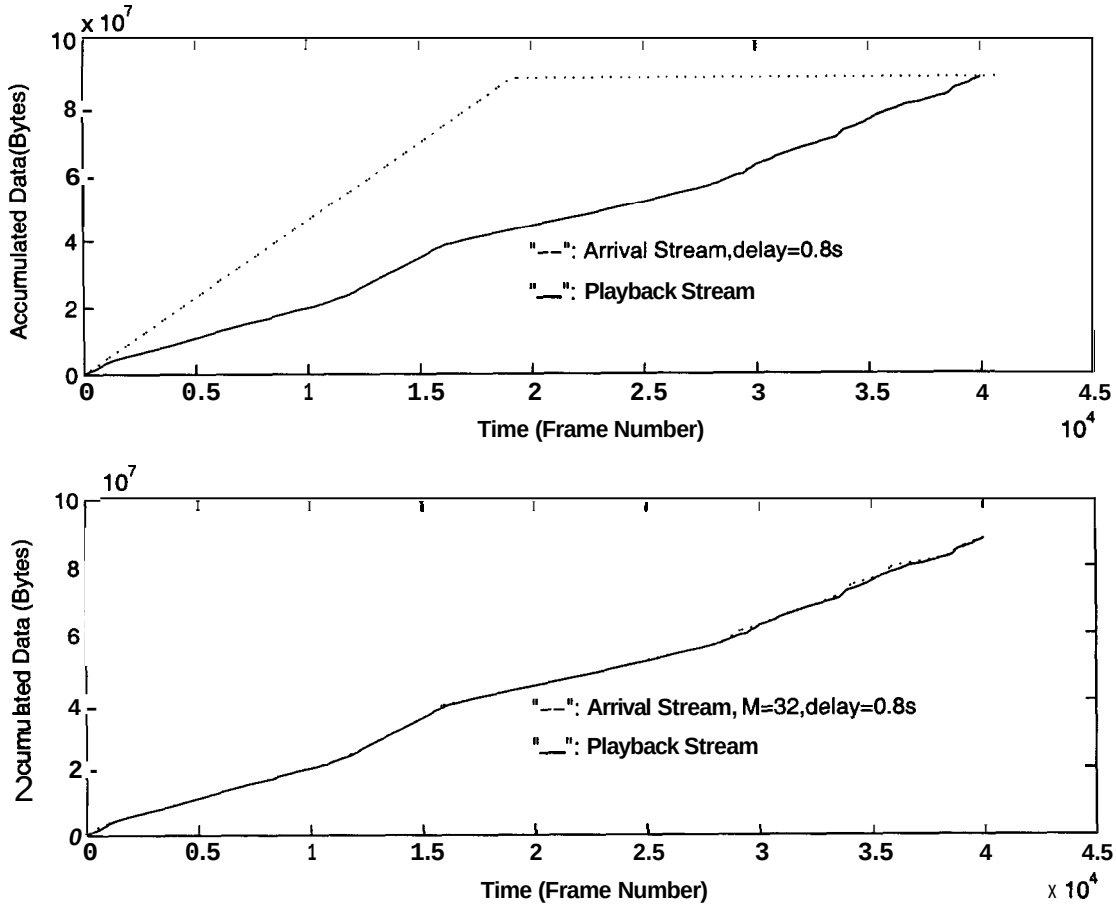


Figure 4.8: Arrival streams of CRTT and dynamic bandwidth allocation

CRTT scheme. This is much less than the one under CRTT with 0.3s delay, but it is still much larger than the buffer occupancy under dynamic bandwidth allocation scheme.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

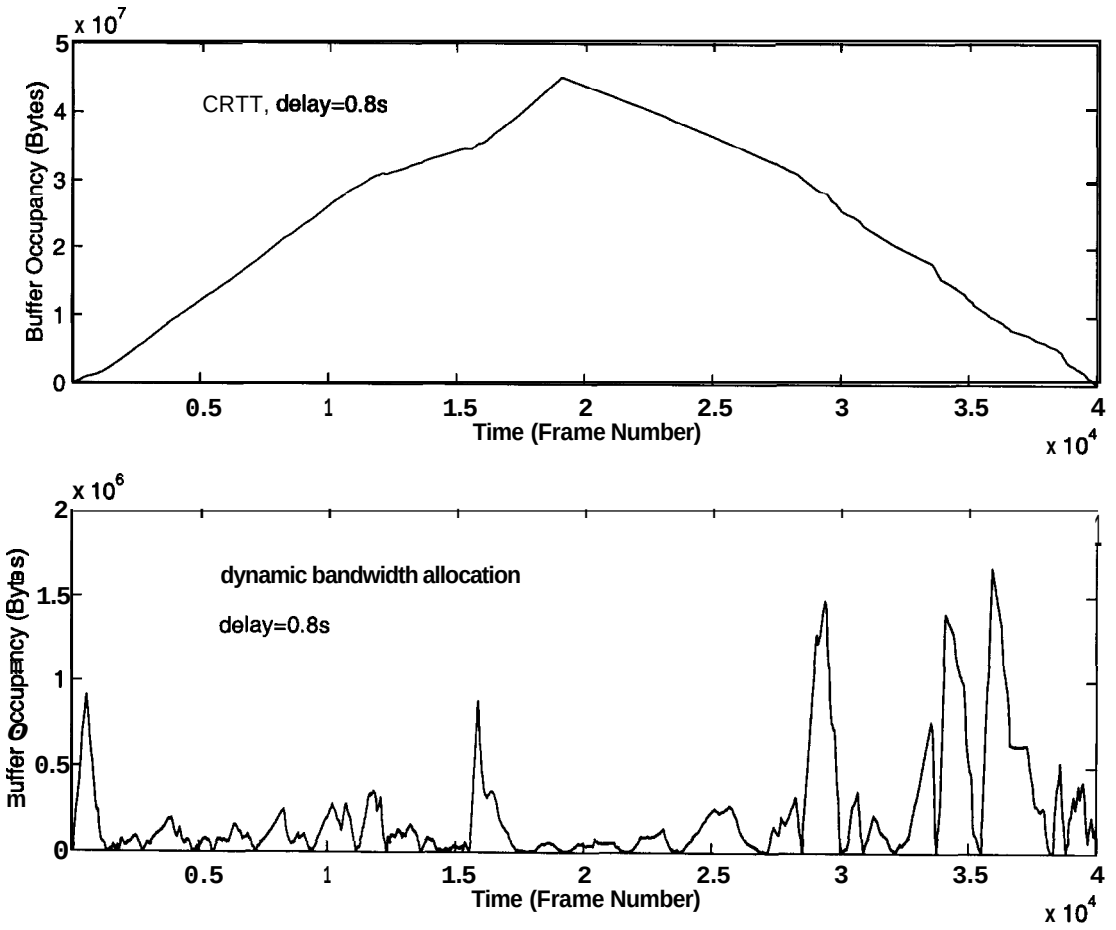


Figure **4.9:** Buffer occupancy under CRTT and dynamic bandwidth allocation scheme, delay=0.8s

4.4 Performance Evaluation under Dynamic Segment Scheme

In this part, we use the dynamic segment scheme proposed in section ***3.6.3*** to guarantee the lowest traffic burstiness and the least number of segment points by the information of burstiness of the video traffic contained in the playback stream. And in the same time, overflow and underflow are avoided. Under this scheme, we assume

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

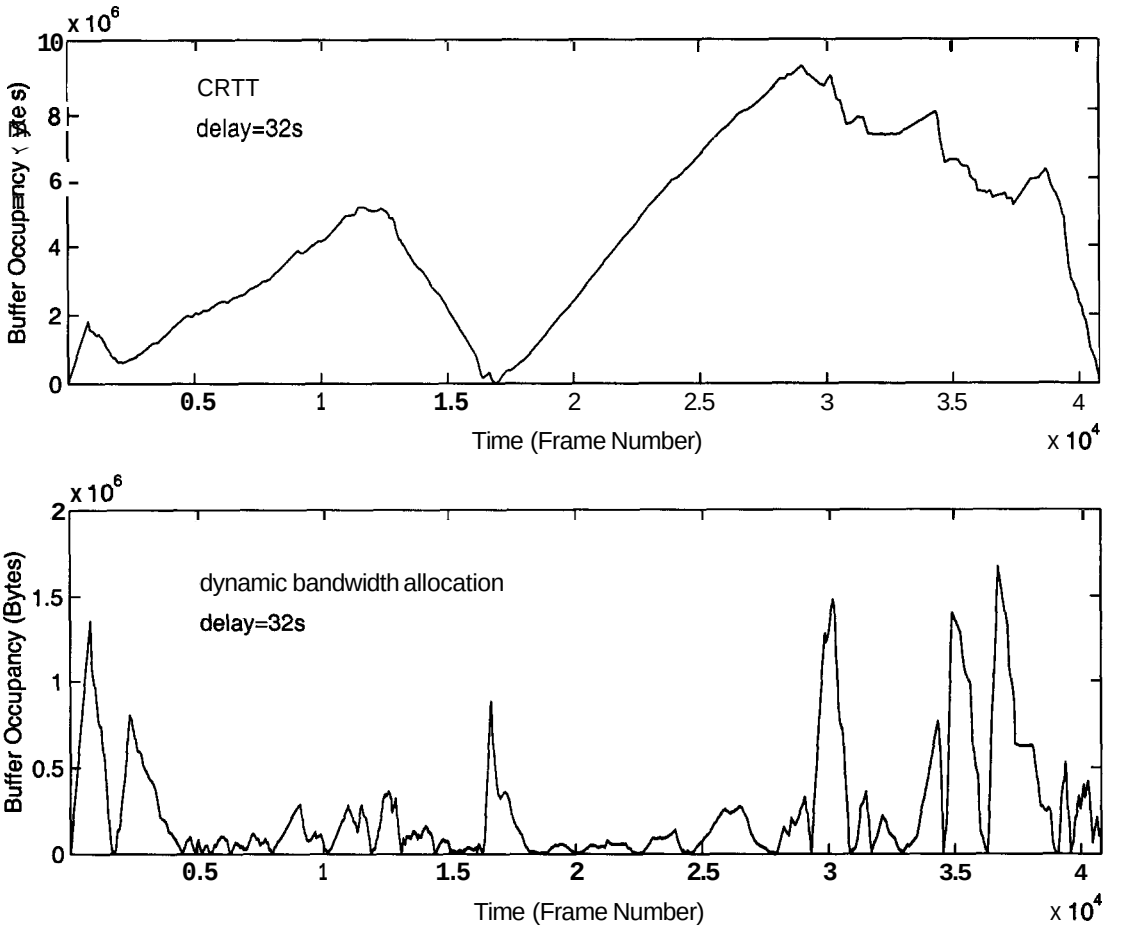


Figure 4.10: Buffer occupancy under **CRTT** and dynamic bandwidth allocation scheme, delay=32s

the buffer capacity is given. The following evaluation results are also obtained based on the MPEG-1 video trace "Episode: Mr.Bean".

Two tunnels have been shown in figure 4.11. Each tunnel in the figures **has** a buffer capacity of 2,000,000 bytes and 3,000,000 bytes respectively. In each tunnel, the dotted line between the two solid lines represents the arrival stream $S(t)$. The lower real line stands for the playback stream $P(t)$, and the higher real line stands for $P(t)$+buffer size. It can be seen that each one satisfies the constrains of no underflow

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

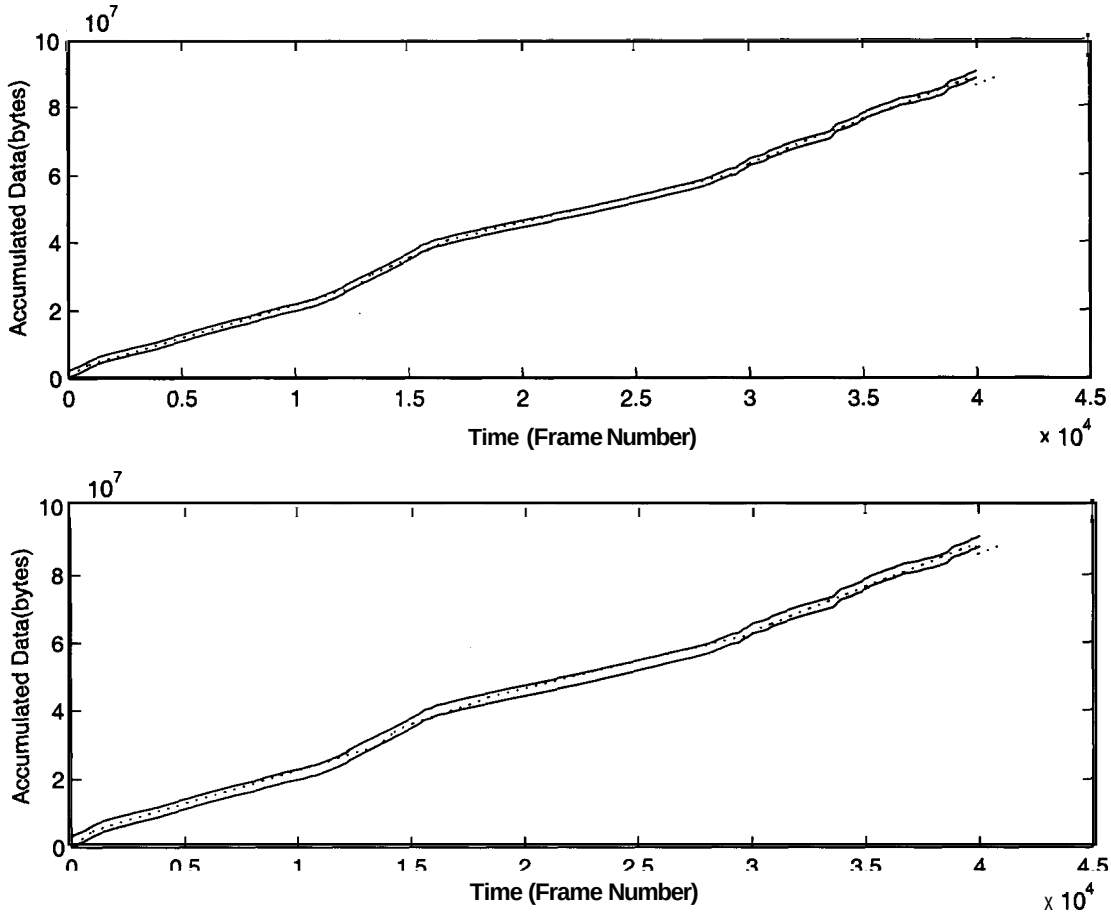


Figure 4.11: Tunnels under dynamic segment scheme with buffer capacity 2,000,000 bytes and 3,000,000 bytes

and no overflow for video transmission. When the buffer capacity becomes larger, the tunnel turns to be much wider. Therefore there are more selective space for the arrival stream. This results in less segment points under dynamic segment scheme.

Table 4.2 shows the segment number needed with different size of buffer capacity given. When the buffer size is 2,000,000 bytes, the segment number is only 7. However, if the user buffer size is much less, for an example, 800,000 bytes, then using this scheme, the segment number changes to 15. Under this scheme, the buffer size can

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

be very small, and the trade-off is more segment number, i.e., more complex traffic plan on the OLT side.

buffer capacity(bytes)	segment number
2,000,000	7
1,700,000	8
1,400,000	12
1,100,000	14
800,000	15

Table 4.2: Segment number needed when changing buffer capacity

Figure 4.12 and figure 4.13 show the bandwidth allocation and corresponding buffer occupancy with the buffer size of 2,000,000 bytes. From the figure 4.12 it can be seen that the transmission bandwidth, b_m allocated to the mth segment is changing by jumps. The transmission rate in each segment is constant. The length of each segment is variable.

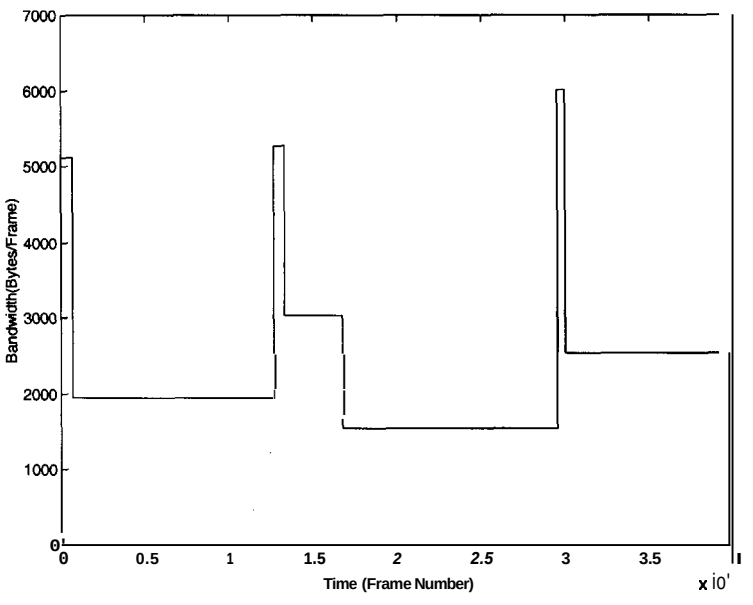


Figure 4.12: Network bandwidth allocation under dynamic segment scheme

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

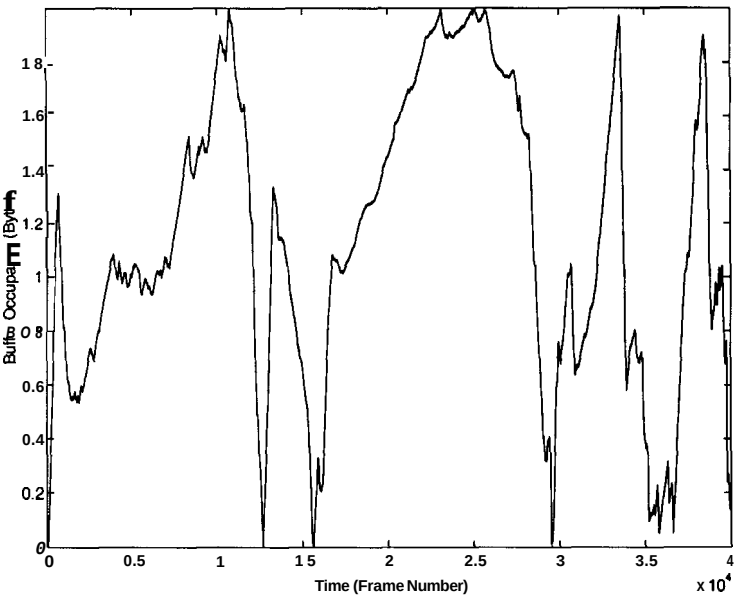


Figure 4.13: Buffer occupancy under dynamic segment scheme

4.5 Comparison of the Two Bandwidth Allocation Schemes

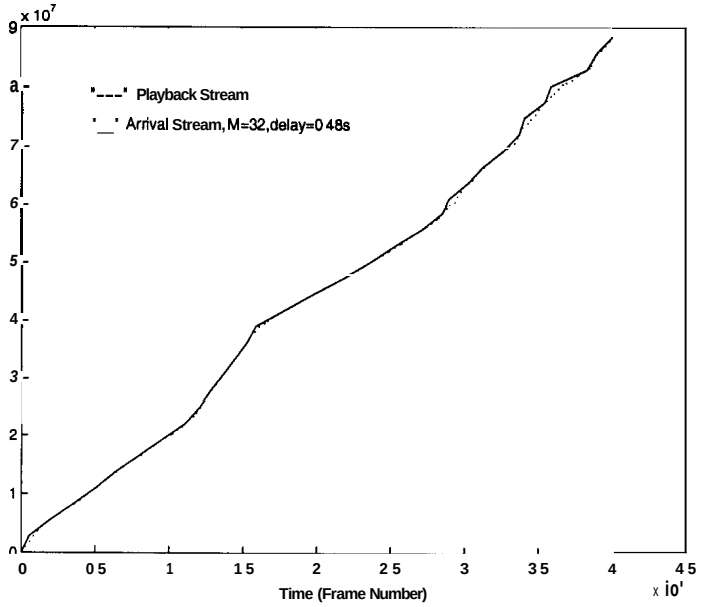
We have proposed two kinds of bandwidth allocation methods with different segmentation schemes in section 3.6.2 and section 3.6.3. The first one, equal segment scheme is much easier in calculation but less flexible comparing with the other one, dynamic segment scheme. In dynamic segment scheme, the computational requirement is larger than in the equal segment scheme. This consume is spent to calculate the segmentation points in the procedure of dynamic segment scheme.

The following figures are simulation results by using MPEG-1 video trace "Episode: Mr.Bean" with the two different segment schemes. In order to compare the two schemes, we set the buffer size similar to the minimum buffer capacity that should

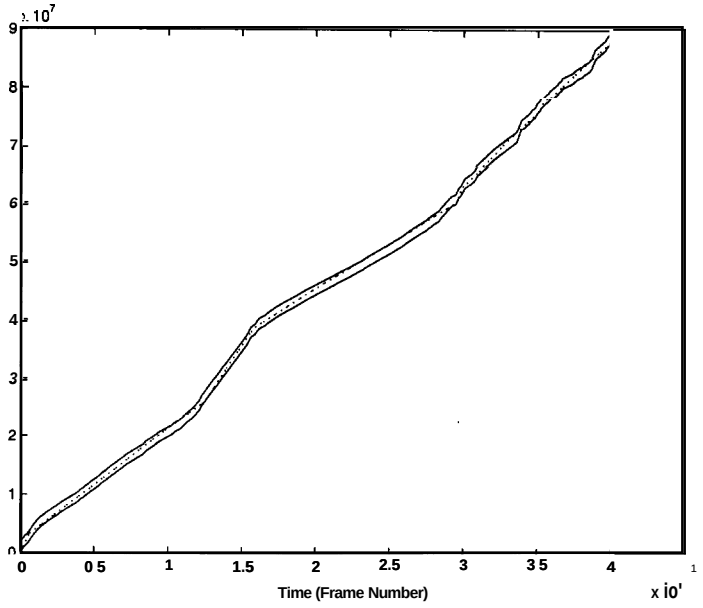
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

be used in the scheme of 32 equal segments (1,700,000 bytes). Figure 4.14 (a) shows the arrival stream and the playback stream under equal segment scheme. Figure 4.14 (b) shows the arrival stream $S(t)$ in the tunnel under dynamic segment scheme. The broken line stands for the arrival stream $S(t)$.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) arrival stream $S(t)$ vs. playback stream $P(t)$ with equal segment scheme



(b) arrival stream in playback tunnel with dynamic segment scheme

Figure 4.14: Arrival stream and playback stream with two schemes

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Figure 4.15 shows the points of segmentation for the arrival stream under dynamic segment scheme. It can be seen that this scheme needs only 8 segments to satisfy no overflow and no underflow, which is much less than **32** segments that have been used under equal segment scheme.

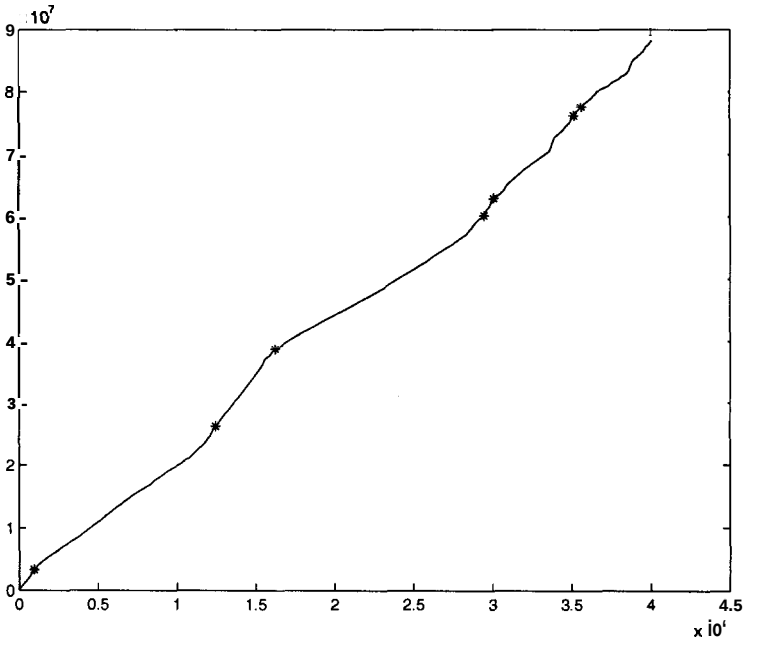
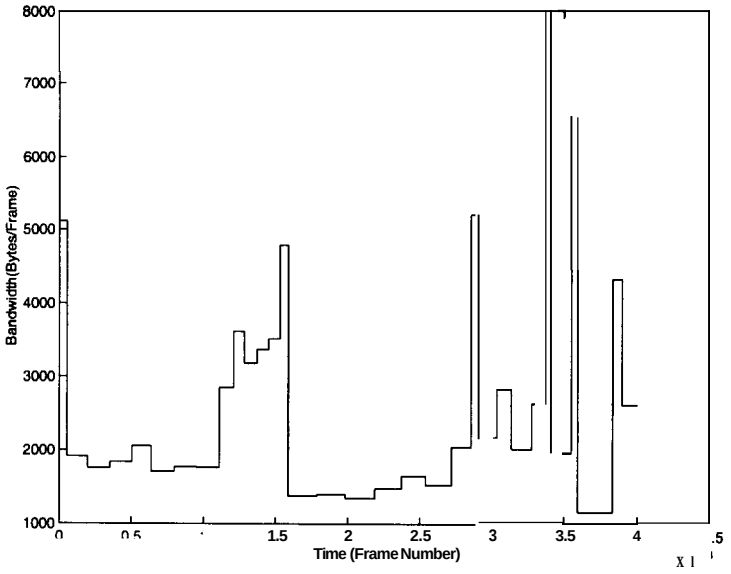


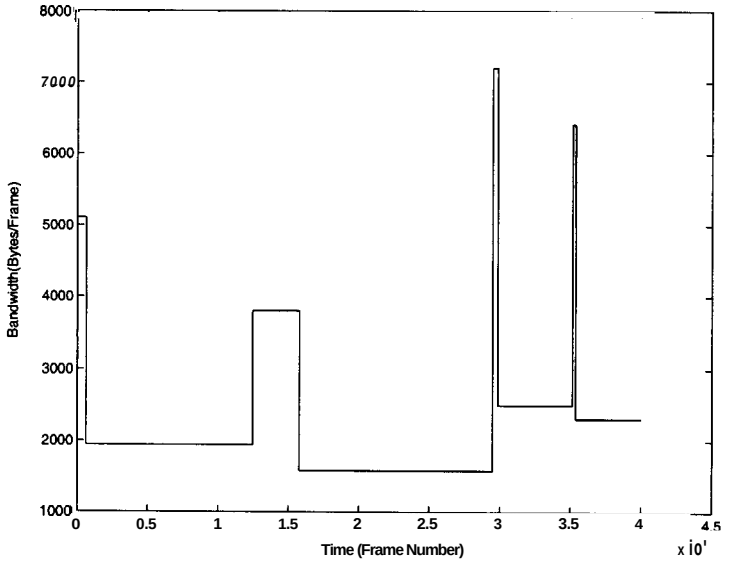
Figure 4.15: Points of segmentation

Figure 4.16 are the results of bandwidth allocation under two segment scheme respectively. Comparing these two figures with figure 4.1, it shows that both of the two bandwidth allocation match the true video trace very well.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) bandwidth allocation with equal segment scheme



(b) bandwidth allocation with dynamic segment scheme

Figure 4.16: Bandwidth allocation with two schemes

In table **4.3** we use a set of different video streams to compare the segment number

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

needed under two schemes. The statistical characteristics of these MPEG VBR-encoded video streams are shown in table 4.1 in section 4.1. The number of segments under equal segment scheme is fixed at 32. The buffer capacity is set at the minimum buffer size required to avoid buffer overflow under equal segment scheme. The results shows that the number of segments needed under the dynamic segment scheme is much smaller than under the equal segment scheme. For example, to transmit the video programme "Soccer" with user buffer size equal or approximate to 2,672,496 bytes, it only needs 7 segments under dynamic segment scheme, but 32 under equal segment scheme. This is because we have made segmentation according to the natural characteristics of the video streams when presenting the dynamic segment scheme. This smaller segment number needed pr ves that the dynamic segment scheme is more efficient.

MPEG video stream	buffer capacity (bytes)	segment number with DBA of equal segment scheme	segment number with DBA of dynamic segment scheme
Mr.Bean	1,700,000	32	8
Lambs	640,992	32	10
Soccer	2,672,496	32	7
Asterix	2,134,368	32	8
Talk show	1,232,112	32	15

Table 4.3: Comparison of segment number needed

To make a summary, dynamic segment scheme which is based on the intrinsic video rate characteristics outperforms equal segment scheme in matching the arrival stream with playback stream, in reducing buffer capacity requirement, as well as in reducing the number of segments.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.6 Traffic Characteristics with Bandwidth Allocation Schemes

To analyze the impact of the two kinds of dynamic bandwidth allocation schemes on the video traffic, it is necessary to evaluate several network performance parameters. In this section, we examine four important parameters as shown below.

1. peak data rate
2. traffic burstiness
3. overload
4. underload

4.6.1 Related Works

There have been some related works in studying the characteristics and performance of compressed, smoothed video streams within the networks.

A study of E. Knightly, et al. [22] examined the impact of traffic smoothing at the network's edge on both the client's QoS , i.e., delay, and on the network's utilization. They explored the scenarios where smoothing has positive effect to the networks and to the clients with considering the case of end-to-end deterministic QoS guarantees. Furthermore, the potential benefits of smoothing were quantified with a set of experiments by means of several streams of MPEG-compressed VBR video in heterogeneous multi-hop networking environments. Particularly, the D-BIND video traffic mode proposed in [23] was shown to obtain higher network utilization than a traffic model based on peak rate, average rate, and burst length like leaky-bucket

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

which was recommended by **ATM** forum user-network interface specification version 3.1.

Another research of A.R. Reibman, et al. [39] was conducted to evaluate smoothing techniques within the video encoder which reduce the stream's peak rate in order to increase the number of streams that can be carried over a CBR network services. In their work, the authors implemented the transport of delay-sensitive applications like teleconferencing to study the characteristics of compressed, online-generated video. They introduced delay at the encoder, and performed work-ahead smoothing over the range of delays which is usually from 0 to 4 frame times and can be tolerated by the application. In contrast, as we consider pre-recorded video and much larger buffer capacity in the client's side, we can perform smoothing on a broader time scale, and thereby achieving much more efficient reduction in the video stream's slow time-scale rate variability.

4.6.2 Peak Rate and Traffic Burstiness

From the network point of view, peak data rate and traffic burstiness are important performance parameters to evaluate bandwidth allocation schemes, because of the fact that network performance depends fundamentally on peak data rate and bursty characteristics. Peak data rate is the maximum rate of the video stream. The peak rate of a video stream determines the worst-case bandwidth requirement across the path from the video storage on the server, the route through the network, and the buffer size requirement at the client site. Thus, to minimize the peak rate is a key issue to increase the likelihood that the server, network and the client have sufficient

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

resources to handle the streams. Reducing the maximum bandwidth requirement permits the server to multiplex a larger number of streams and it is especially important if the user has a low bandwidth connection with the network. The total cost of the data transfer can be reduced with a low peak rate. The second parameter, traffic burstiness, indicates the grade of difficulty to satisfy the QoS requirement. We calculate the original peak rate and traffic burstiness of a set of video traces without using dynamic bandwidth allocation and with dynamic bandwidth allocation scheme under the two different segment schemes to compare them. Traffic burstiness is defined as the ratio of the peak data rate to the average data rate, as below,

$$Traffic\ Burstiness = \frac{Peak\ rate}{Average\ rate} \tag{4.6.1}$$

where, for original video trace,

$$Peak\ rate = \max_{1 \le i \le N} \{x_i\} \tag{4.6.2}$$

and,

$$Average\ rate = \frac{\sum_{i=1}^{N} x_i}{N} \tag{4.6.3}$$

where $x_i (i = 1, 2, \cdots, N)$ represents the video data contained in the ith frame, N is the total number in a video trace.

For the video traces controlled by bandwidth allocation scheme,

$$Pealcrate = \max_{1 \le m \le M} \{b,\} \tag{4.6.4}$$

and,

$$Average\ rate = \frac{\sum_{m=1}^{M} b_m \times t_m}{\sum_{m=1}^{M} t_m} \tag{4.6.5}$$

Table 4.4 shows the calculation results of average data rate and peak data rate of several different video traces. It can be seen that the peak data rate can be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

effectively reduced after using dynamic bandwidth allocation scheme. Comparing to the original data, the peak rates with equal segment scheme have been greatly reduced. Comparing to the peak rates with equal segment scheme, the rates with using dynamic segment scheme can even be reduced more. To take an example, the original peak rate of "soccer" trace is 23787.0 bytes/frame, the peak rate with equal segment dynamic bandwidth allocation scheme is only 10929.0 bytes/frame, however, the peak rate with dynamic segment dynamic bandwidth allocation scheme decreases to 9246.1 bytes/frames.

MPEG video traces	Average rate (bytes/frame)			Peak rate (bytes/frame)		
	Original	With DBA of equal segment scheme	With DBA of dynamic segment scheme	Original	With DBA of equal segment scheme	With DBA of dynamic segment scheme
Mr.Bean	2205.9	2205.1	2205.5	28634.0	7969.0	7205.6
Lambs	913.94	913.88	913.9	16778.0	3841.9	3570.5
Soccer	3138.7	3137.7	3138.5	23787.0	10929.0	9246.1
Asterix	2793.6	2793.5	2793.5	18422.4	8185.0	7709.3
Talk show	1817.1	1817.0	1817.1	13346.0	4393.7	4380.0

Table 4.4: Average rate and peak rate under different schemes

The comparison on traffic burstiness among original streams and streams with dynamic bandwidth allocation under the two segment schemes is given in the table 4.5. It can be seen that the traffic burstiness of each video stream with dynamic bandwidth allocation is much lower than the original one. The traffic burstiness under the dynamic segmentation scheme is lower than that under the equal segmentation scheme

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

MPEG video traces	Traffic burstiness		
	Original	With DBA of equal segment scheme	With DBA of dynamic segment scheme
Mr.Bean	12.9806	3.6139	3.2671
Lambs	18.3578	4.2039	3.9068
Soccer	7.5786	3.4831	2.9460
Asterix	6.5944	2.7600	2.6662
Talk show	7.3447	2.4181	2.4105

Table 4.5: Comparison of traffic burstiness

4.6.3 Overload and Underload

Overload and underload are two parameters indicating how much the network and the user buffer are overloaded. They are defined as,

- overload: the degree to which the network is overloaded during the burst period and the overloaded data accumulating in a buffer;
- underload: the degree to which the data in the buffer are drained when the network is underloaded during an interburst period.

Overload and underload can be obtained from the following equations:

$$Overload = \frac{\sum_{i \in \{b_i > Mean\}} b_i \times t_i}{\sum_{i \in \{b_i > Mean\}} b_i \times t_i + \sum_{j \in \{b_j < Mean\}} b_j \times t_j} \qquad (4.6.6)$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$Underload = \frac{\sum_{j \in \{b_j < Mean\}} b_j \times t_j}{\sum_{i \in \{b_i > Mean\}} b_i \times t_i + \sum_{j \in \{b_j < Mean\}} b_j \times t_j} \quad (4.6.7)$$

Overload and underload have the relationship of,

$$Overload + Underload = 1 \quad (4.6.8)$$

MPEG video traces	Overload			underload		
	Original	With DBA of equal segment scheme	With DBA of dynamic segment scheme	Original	With DBAof equal segment scheme	With DBA of dynamic segment scheme
Mr.Bean	0.6119	0.4375	0.4965	0.3881	0.5625	0.5035
Lambs	0.6925	0.4936	0.6711	0.3075	0.5064	0.3289
Soccer	0.6341	0.4892	0.4055	0.3659	0.5108	0.5945
Asterix	0.6464	0.5478	0.5421	0.3536	0.5422	0.6464
Talk show	0.5672	0.5567	0.2090	0.4328	0.4433	0.7910

Table 4.6: Comparison of overload and underload

The performance evaluation results of overload and underload are shown in table 4.6. To make the comparison, the buffer capacity of each stream under dynamic segment scheme is set at the minimum buffer size required to avoid buffer overflow under equal segment scheme. It can be seen that both kinds of dynamic bandwidth allocation schemes can reduce the overload and increase the underload. For example, video trace "Mr.Bean" has an original overload of 0.6119, and it changes to 0.4375 and 0.4865 respectively with using dynamic bandwidth allocation under equal segment and under dynamic segment. Furthermore, note that it is not deterministic which segment scheme can reduce overload more. For video traces "Mr.Bean" and "Lambs",

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

equal segment scheme can decrease the overload more than dynamic segment scheme. For the other three traces, the case is just opposite. However, both of the two schemes can effectively decrease the degree to which the network is overloaded during the burst period and reduce the overloaded data accumulating in a user buffer.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 5

Multiplexing Video Streams over Passive Optical Networks

5.1 Introduction

Chapter 4 have studied the performance of a single video stream with the proposed dynamic bandwidth allocation schemes. This chapter discusses the statistical multiplexing of video streams over passive optical networks. After a survey of related works in the area, the effect of dynamic bandwidth allocation schemes on statistical multiplexing is examined for two multiplexing models: time independent videos and time correlated videos. These simulations focus on two network parameters: statistical multiplexing gain, and normalized bandwidth. Intuitively, it should be the case that bandwidth allocation schemes can have a good impact on the statistical multiplexing performance, and indeed this is what we are able to show in this chapter.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.2 Related Works on Statistical Multiplexing Compressed Video

Many studies have examined the multiplexing performance of compressed video streams. [4][5][7][32][33][28][34] have tested compressed, unsmoothed video streams. However, relatively few has examined the impact of the smoothed video on network traffic and resource requirement.

The work by Z. Zhang, et al. [45] is a notable exception. The authors investigated the effect of their proposed optimal smoothing algorithm on network resource requirements. The results showed that the algorithm can effectively reduce rate variability and end-to-end resource requirements. The advantage of VBR service over CBR service in supporting real-time transport of pre-stored video was examined, and the potential statistical multiplexing gains of smoothed video traces under VBR service was evaluated through a simulation-based empirical study. The impact of time correlated video streams on statistical multiplexing gains was investigated and the requirement for the network to support multiple QoS service levels with varying robustness was illustrated.

5.3 Performance Parameters and Simulation Models

In this section, we provide the definitions of performance parameters and simulation models of statistical multiplexing. Statistical multiplexing gain G is defined as,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$G = (1 - \frac{P_W}{\sum_{w=1}^{W} p_w}) \times 100 \tag{5.3.1}$$

where P_W represents the aggregated bandwidth required to satisfy a given QoS requirement value for the multiplexed video streams and p_w represents the peak rate of the aggregate load for each individual stream, hence, $\sum_{w=1}^{W} p_w$ is the total aggregate bandwidth (which is the sum of the peak rate of the individual streams) required to satisfy the QoS for all MPEG video streams, if these W MPEG video streams are transported separately. Therefore, the statistical multiplexing gain G represents the fractional reduction in the aggregate bandwidth requirement comparing to without multiplexing. Hence, it quantifies the potential utilization improvement that can be achieved by bandwidth allocation schemes.

Normalized bandwidth B_N is defined as the bandwidth requirement for each video comparing to the mean rate of the video,

$$B_N = \frac{P_W}{\sum_{w=1}^{W} Mean_w} \tag{5.3.2}$$

where $Mean_w (w = 1, 2, \cdots, W)$ is the mean rate of aggregate load of each individual video stream.

We consider two models of network multiplexing: multiplexing of time independent videos and multiplexing of time correlated videos. These two models are the most particular topics involving in the transport of multiple video in support of video-on-demand services. These two process models are defined as shown below.

1. Time independent multiplexing. To simulate time correlated multiplexing, we assume that the W video streams arriving at the multiplexer are randomly displaced from each other. In other words, for each video stream, the starting

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

frame is equally likely to be any one of the video frames, with appropriate ”wrap-around” to ensure that the video streams are of the same length;

2. Time correlated multiplexing. To investigate the impact of time correlated multiplexing of video streams on the network management, we consider that all video streams are constrained to start within a short interval of several minutes. Within the short time interval, start times are uniformly, independently and identically distributed.

To simulate the two multiplexing processes, we consider two types of video streams for each model.

1. Same video streams. For this type, all multiplexed video streams are from the same video “Mr.Bean”. In this case, we simulate the condition that different users watch the same video simultaneously but the start times are different at the playback period.

2. Different video streams. For this type, all multiplexed video streams are generated from 8 different videos. The number of streams from each type of video is increased uniformly as the number of streams increases. Therefore, an aggregation of 80 streams consists of 10 sources from each type. In this case, we simulate the condition that different users watch different videos simultaneously, while the start times are different at the playback period.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.4 Simulation Results

5.4.1 Time Independent Multiplexing

In this subsection, we investigate multiplexing video streams with the assumption that video start times are independent of each other. The study is conducted by considering the two types of video streams.

Figure 5.1 shows the simulation results of multiplexing gain for original streams, streams using dynamic bandwidth allocation (DBA) with equal segment, and streams using DBA with dynamic segment scheme. The QoS requirement for this example is that no loss occurs at the multiplexer during the entire transmission of the aggregate video streams.

Figure 5.1 (a) illustrates the multiplexing gain as a function of the number of video streams under the condition that different clients independently access the same video simultaneously but at different time of playback period. It can be seen that, by comparing with the video streams without using DBA, both of the two DBA schemes can effectively improve the statistical multiplexing gain, and the dynamic segment scheme has a better result. These can be found especially when the video source number is less than 10. For example, a potential statistical multiplexing gain of no more than 30% for the original trace is achievable when only two users access the video, while a potential statistical multiplexing gain of about 75% can be gained by DBA with equal segment, and about 90% is obtained by DBA with dynamic segment.

The other result, figure 5.1 (b), is achieved under the condition that different clients independently access different video simultaneously but at different time of

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

playback period. A potential statistical multiplexing gain can be improved from 70%–80% for original streams to 85%–95% by DBA with dynamic segment. In addition, the figure also shows that greater statistical multiplexing gain can be achieved when more videos are multiplexed.

Figure 5.2 shows the normalized bandwidth as a function of the number of video streams for original streams, streams using dynamic bandwidth allocation (DBA) with equal segment, and streams using DBA with dynamic segment scheme. Since the mean rates for the video streams under the three conditions are the same, this figure reveals the impact of DBA on bandwidth reduction.

Figure 5.2 (a) is the results for same video streams. The figure shows that both DBA with equal segment and DBA with dynamic segment can significantly reduce the bandwidth required to support a given QoS level, especially when the number of the multiplexing videos is less than 10.

In figure 5.2 (b), video streams are from different video streams. It also shows that both of the two DBA schemes have a good effect on reducing the bandwidth requirement. The normalized bandwidths with DBA are between 1 to 1.75, while those without using DBA are between 1.4 to 3.25. Furthermore, the results of normalized bandwidth of different video streams are reduced largely comparing with the normalized bandwidths with the same video streams in figure 5.2 (a).

5.4.2 Time Correlated Multiplexing

In the last subsection, we assume that the video start times are independent of each other. However, in practice, this assumption may be violated. In a video–on–demand system, many users may start watching videos within a short time span. In such

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

cases, we need to consider time correlated multiplexing. This subsection investigates how the time correlated multiplexing have impacts on the network parameters.

The effect of time correlated arrivals of video streams on the statistical multiplexing gain is shown in figure 5.3 and figure 5.4 , where (a), (b), and (c) show the simulation results with original traces, with equal segment DBA, and with dynamic segment DBA, respectively.

Figure 5.3 shows that for the same video streams the time correlation has obvious impact on statistical multiplexing gain whether using equal segment or using dynamic segment. For example, when 80 end users access the same video without using DBA within 1 minute, a potential statistical multiplexing gain of 81% is achievable. However, when 80 end users access the same video without using DBA within 10 minute, a potential statistical multiplexing gain of about 90% is achievable. Moreover, comparing figure ***5.3*** (a), (b) and (c), it also shows that both DBA schemes can improve the statistical multiplexing gain with correlated time condition, especially when the number of multiplexed streams are less than 20.

Figure 5.4 for different video streams can be achieved as those in figure 5.3 for same video streams. Comparing figure 5.4 (a)-(c) with figure 5.3 for same video streams (a)-(c), it can be observed that time correlated arrivals have a relatively greater impact on aggregation of the same video streams than on different video streams. The reason is that the heterogeneity of the video streams can alleviate the adverse impact of correlated arrivals on the statistical multiplexing gain.

The impact of time correlated arrivals on normalized bandwidth is shown in figure 5.5 for same video streams and figure 5.6 for different video streams, where (a), (b), and (c) show the simulation results with original traces, with equal segment DBA,

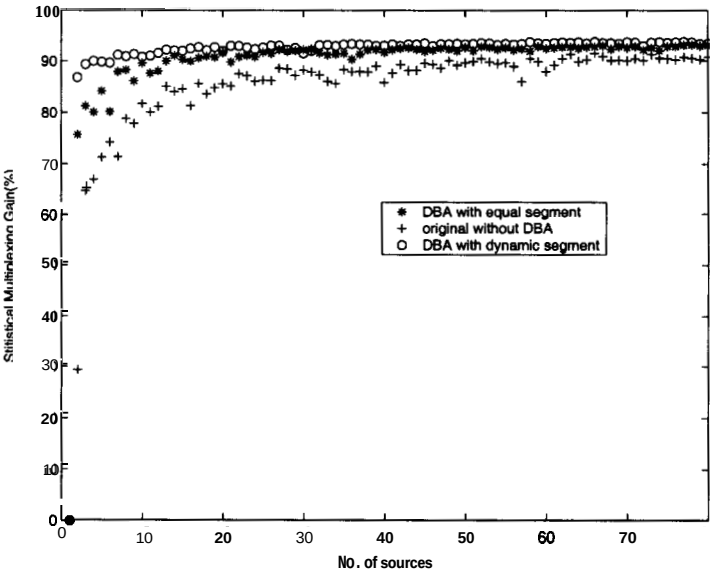
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

and with dynamic segment DBA, respectively.

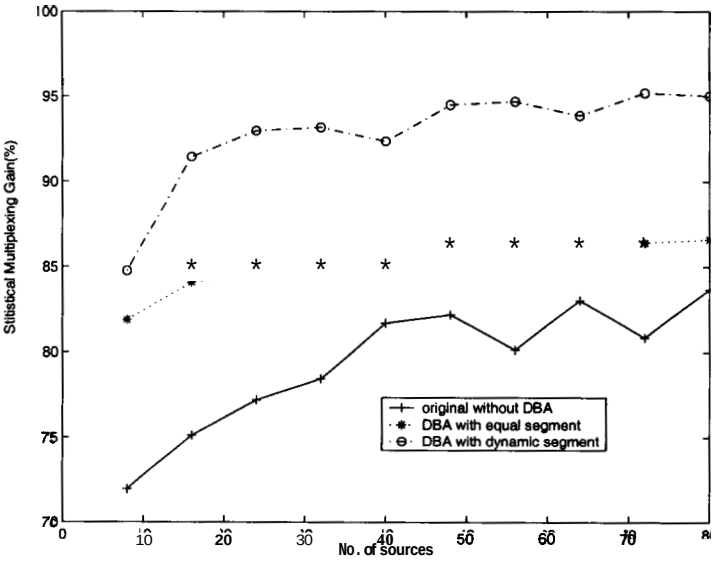
From figure 5.5 (a), (b), and (c), it can be seen that the time correlation has severe impact on the normalized bandwidth of same video streams for whether using DBA or using DBA with the two kinds of segmentation schemes. For example, a normalized bandwidth of 2.8 is required when 80 end users access the same video stream within 1 minute using DBA with dynamic segment scheme. While, when the same number of end users access the same video stream using DBA with dynamic segment scheme within 10 minutes, a normalized bandwidth of only 1.3 is needed. In addition, smaller normalized bandwidth can be achieved when more videos are multiplexed, which can also be revealed from figure 5.3 (a), (b) and (c). For example, as shown in figure 5.3 (c), when 10 end users access the same video streams using dynamic segment DBA within 10 minutes, a normalized bandwidth of 1.7 is needed. However, the normalized bandwidth decreases to 1.35 when 80 end users access the same video stream using dynamic segment DBA within 10 minutes. Moreover, comparing figure 5.3 (a), (b) and (c), for the original streams, the normalized bandwidth needed is mostly in the range of 1.5 to 18. However, for streams using DBA with equal segment scheme, the range is from 1.5 to 4.2, and for streams using DBA with dynamic segment scheme, the range is from 1.5 to 3.8. This proves that both of the two DBA schemes can effectively reduce the bandwidth requirement to support a given QoS level.

In figure 5.6 for different video streams, the result shows that the different arrival intervals have less effect on the normalized bandwidth. This also proves that heterogeneity can reduce the bandwidth requirement significantly which can be seen from figure 5.4.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



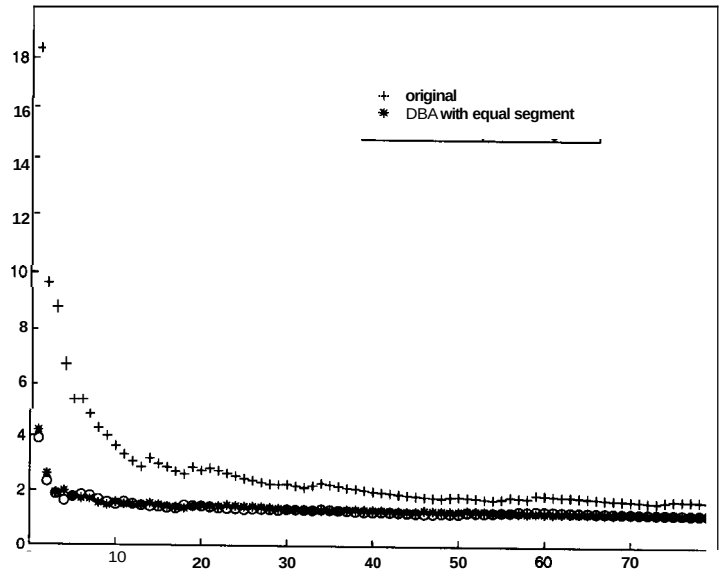
(a) Same video streams



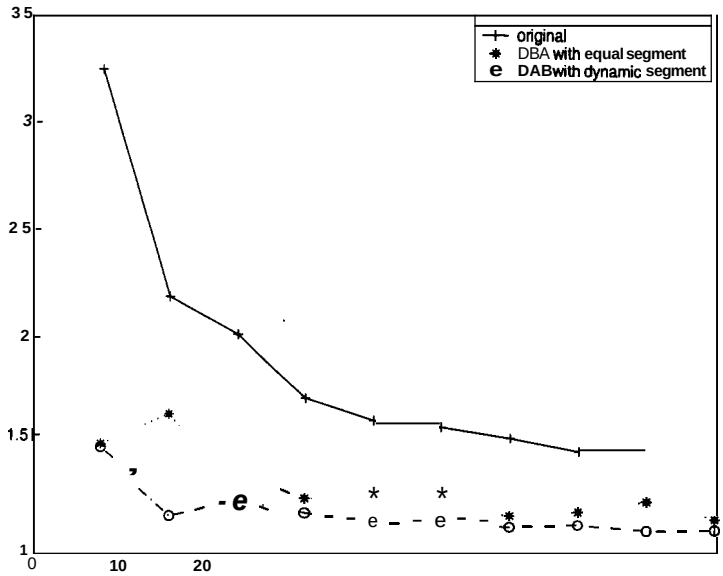
(b) Different video streams

Figure 5.1: Multiplexing gain of independent streams

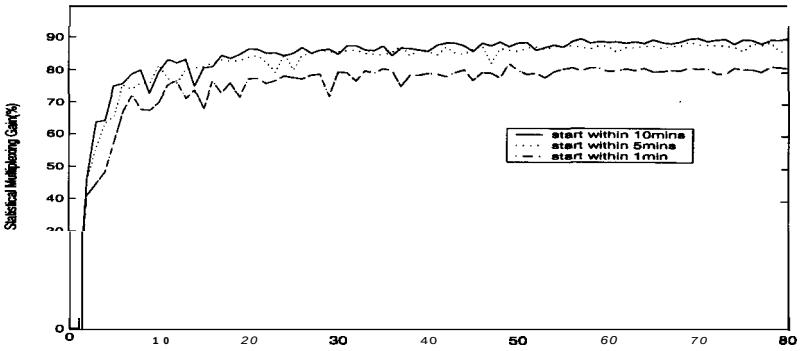
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

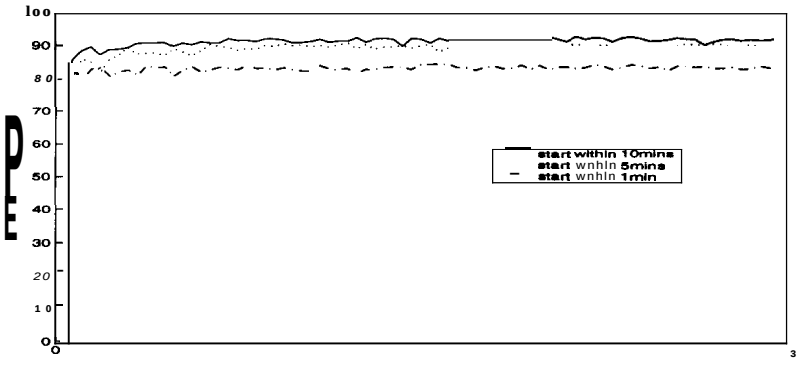


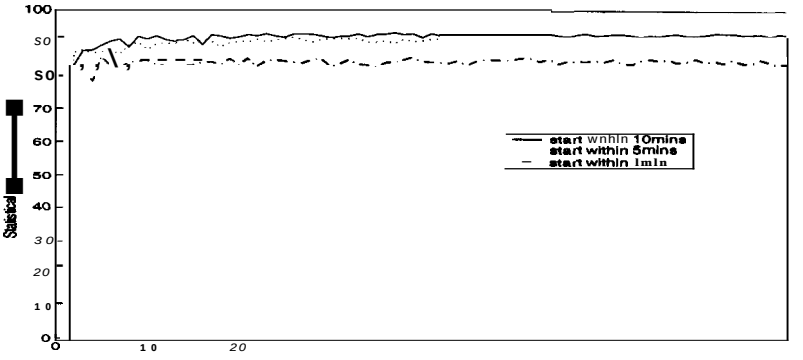
(a) Same video streams



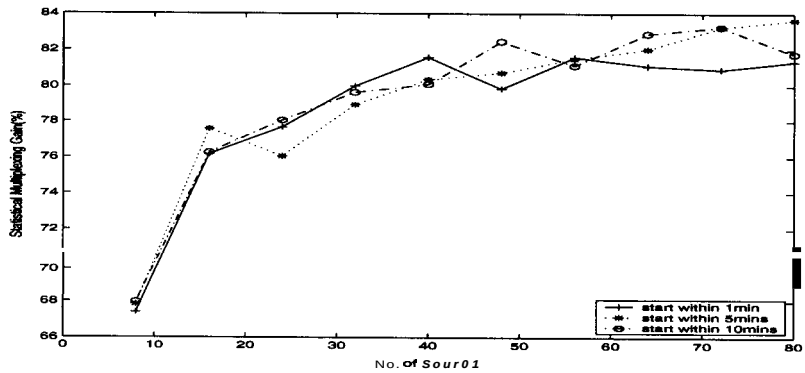
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



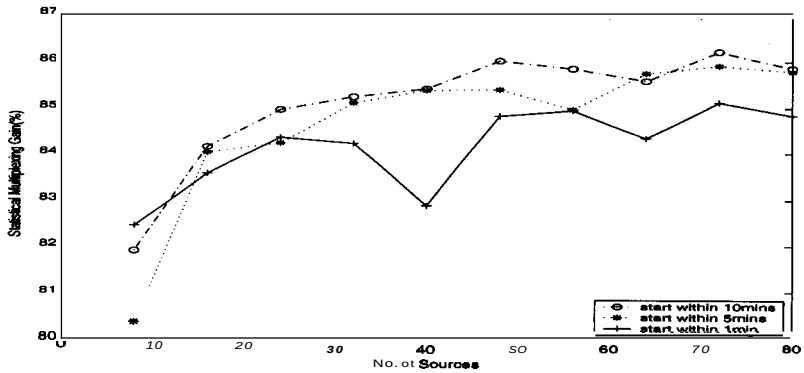




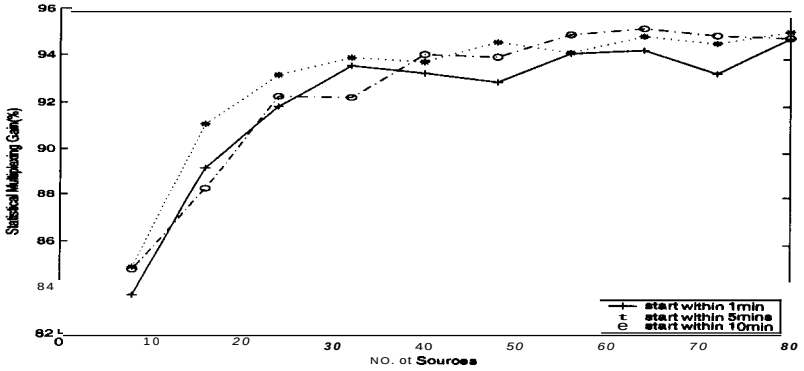
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) original



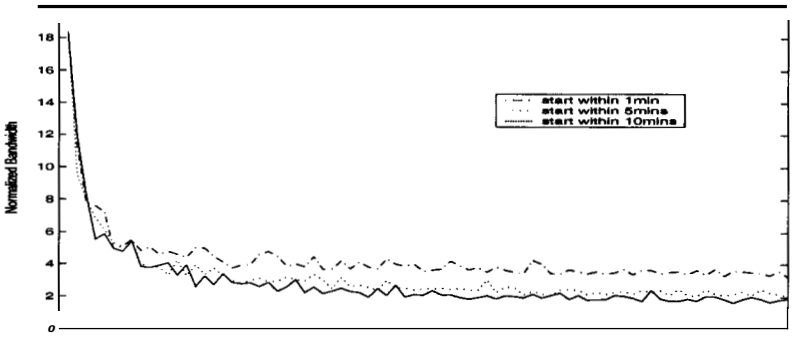
(b) **DBA with equal segment**

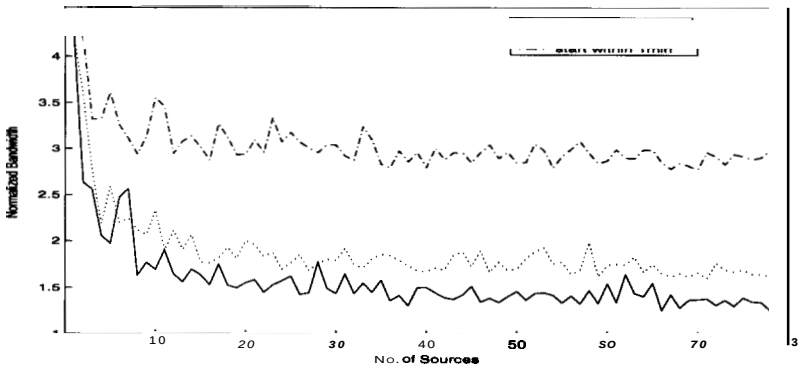


(c) DBA with dynamic segment

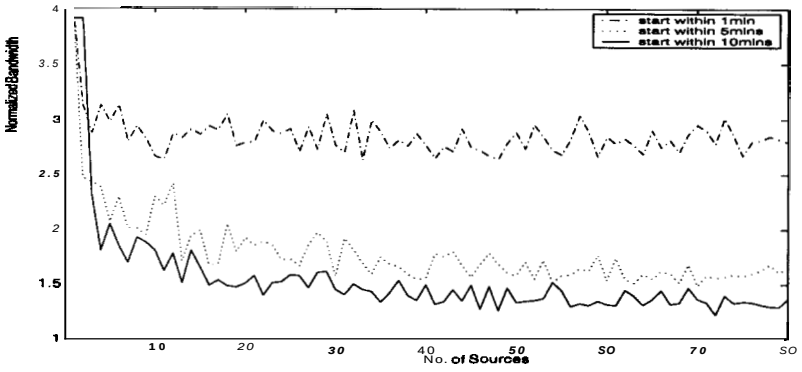
Figure 5.4: Multiplexing gain of correlated streams: different video streams

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library





(b) DBA with equal segment



(c) DBA with dynamic segment

Figure 5.5: Normalized bandwidth of correlated streams: same video streams

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

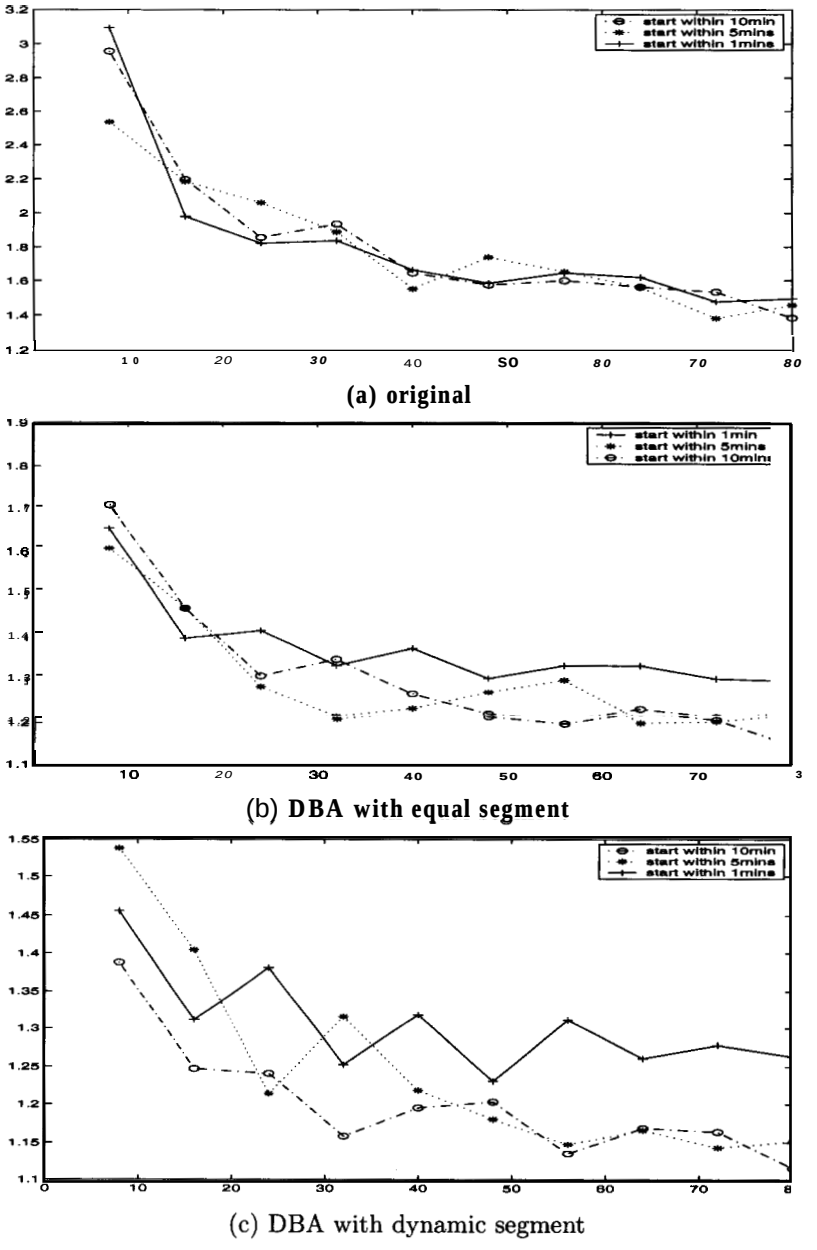


Figure 5.6: Normalized bandwidth of correlated streams: different video streams

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 6

Conclusions and Recommendations for Future Work

6.1 Conclusions

In this thesis the architecture of video-on-demand system over passive optical networks was discussed. The objective of this research has been identified as designing of traffic management scheme which is capable of effectively support the transport of pre-stored video data over passive optical networks. Based on the the characteristics and architecture of video-on-demand system over PON, the main design issues for video-on-demand service over passive optical networks have been studied in detail. In particular, dynamic bandwidth allocation in support of video-on-demand systems over PONs has been considered.

The upstream vod traffic from users to servers over PONs is managed utilizing TDM technology. The time slots are synchronized so that upstream packets from the ONUs do not interfere with each other once the data is coupled onto the common fiber. A dynamic bandwidth allocation scheme was proposed for transmitting pre-stored video from optical line terminal to optical network units across passive optical networks. A transmission model was set up for the downstream transport of VoD

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

data over PON. With this model we explored ways to manage the video to be played back in real-time at the user side without either underflow or overflow. In particular, a dynamic bandwidth allocation mechanism was studied to meet the management requirements of low network transmission costs, guaranteed QoS, easy traffic control method, and efficient bandwidth utilization. A tunnel scheme, which is very easy for dynamic bandwidth allocation, has been proposed. For the established dynamic bandwidth allocation approach, two segmentation schemes were used to support it. The first one is equal segment scheme, in which video streams are simply divided into equal byte segment. As it was noticed that the performance of bandwidth allocation scheme can be further improved by segmenting the video sequence appropriately, a video segmentation method based on intrinsic video rate characteristics was presented, i.e., dynamic segment scheme.

We demonstrated the performance of the proposed dynamic bandwidth allocation approach by a set of MPEG VBR-encoded video streams. The system performance under equal segment scheme has been evaluated in terms of bandwidth allocation, buffer occupancy and the change of playback buffer capacity with different segment number being used. The results showed that the allocated bandwidth matched the real video traffic stream very well. The comparison of the proposed scheme with the well-known constant rate transmission and transport(CRTT) scheme was carried out. The results proved that the performance of dynamic bandwidth allocation with equal segment obviously outperformed the performance of CRTT. The performance of dynamic bandwidth allocation with dynamic segment scheme was also evaluated. The comparison of dynamic segment scheme and equal segment scheme was conducted. The results demonstrated that the dynamic segment scheme had a better

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

performance in terms of matching between the arrival stream and playback stream, buffer capacity requirement, and segment number used. To investigate the benefit of the two kinds of dynamic bandwidth allocation schemes for the video traffic over PON, traffic characteristics with several network performance parameters were evaluated. Both two schemes were shown to be effective in reducing the peak rate and traffic burstiness. Dynamic segment scheme showed a better performance in terms of these two parameters. In evaluating the overload and underload, we found that both schemes can reduce the overload and increase the underload, but it was not conclusive which scheme was more effective.

The multiplexing of multiple video streams under the dynamic bandwidth allocation scheme was also considered in this thesis. The impact of time correlation and content correlation of video traffic on network resource management under dynamic bandwidth allocation scheme were evaluated. We obtained the conclusion from the results that dynamic bandwidth allocation scheme with statistical *QoS* guarantees can provide a viable and resource-efficient way to support real-time delivery of pre-recorded video.

6.2 Recommendations for Future Work

Based on the research conducted in this thesis, there are several promising areas for future works.

One interesting avenue of future works is to investigate the selection of optimal bandwidth in the playback tunnel, which was proposed in chapter 3. With the playback tunnel, the high computing cost of complicated dynamic programming may be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

reduced due to the conciseness of tunnel scheme. On the other hand, under the dynamic bandwidth allocation with equal segment scheme, a feasible range of bandwidth allocation can be found to meet the requirement of receiving buffer neither underflows nor overflows. Obviously, the optimization of the selection of feasible bandwidth to make use of bandwidth resource and receiving buffer more efficiently can be a valuable research area.

In this thesis, the results are obtained through trace-driven simulation. Therefore, to conduct a fully experimental evaluation for the proposed traffic management schemes using pre-stored videos can be an excellent topic for future work. By actual implementations, interesting and important issues will be exposed. For example, the techniques presented in this thesis assume that all the video data is delivered. However, in practice, an end user may be willing to tolerate a small amount of frame loss in exchange for a reduced delivery charge. In this case, the server may have additional latitude in transmitting video streams. On the other hand, it is also a valuable research to combine the adaptive quality techniques with our proposed schemes to further reduce the resource requirements and traffic burstiness for the transport of pre-stored video.

Another prospective study area lies in the transport of delayed video [19], which indicates those applications that fall in between live real-time video and pre-stored video. They are the broadcast videos, such as a newscast or sports event. These kind of video applications do not constrained by the short delay requirement of live video, and also may not need the entire video known as a prior as pre-stored video. To transport delayed videos, it is necessary to implement hybrid bandwidth management schemes. For very large delays of these videos, it is obviously that the techniques

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

similar to our developed approaches for pre-stored video can be conducted.

Finally, interactivity is a key feature in applications such as video-on-demand, distance learning, and catalog browsing.It is widely assumed that interactivity with pause, rewind and fast-forward functions will be demanded in passive optical networks. While we have restricted our attention to user video playback, it is an important question of how to accommodate interactivity by our proposed traffic management schemes.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Author's Publications

[I] Ouyang Yue, Lu Chao: *"Dynamic Network Bandwidth Allocation for VoD over Passive Optical Networks"*, 2002 International Conference on Software, Telecommunications and Computer Networks ,2002.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Bibliography

[1] ISO/IEC IS 11172-2, *Coding of moving pictures and associated audio for digital storage meida up to 1.5 mbit/s - part2: video.*

[2] ISO/IEC IS 13818-2, *Generic coding of moving pictures and associated audio - part2: video.*

[3] L.S. Smoot S.B. Weinstein M. Fortier A.D. Gelman, H. Kobrinski and D. Lemay, *A store-and-forward architecture for video-on-demand service,* ICC'91 (1991), 842–846.

[4] Q. Karlsson P. Sen B. Maglaris, D. Anastassiou and J. D. Robbins, *Performance models of statistical multiplexing in packed video communications,* IEEE Trans. on Commun. (1988), 834–843.

[5] B. N. Bashforth and C. L. Willianmson, *Statistical multiplexing of self-similar video streams: simulation study and performance results,* Sixth International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (1998), 119–126.

[6] B.DeCleene, *Comparison of priority partition methods for vbr mpeg,* INFOCOM'94 (1994), 689–696.

[7] B. Melamed B. Sengupta D. Reininger, D. Raychaudhuri and J. Hill, *Statistical multiplexing of vbr mpeg compressed video on atm networks,* INFOCOM'93 (1993), 919–926.

[8] S.S. Dixit and P. Skelly, *Video traffic smoothing and atm multiplexer performance,* GLOBLECOM'91 (1991), 239–243.

[9] A. Reyes E. Casilari and A.D. Estrella, *Heavy-tailed distribution of scene duration in vbr video,* Electronics Letters **35** (1999), no. 2, 134–135.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[10] ITU-T Recommendation G.983.1, *Broadband optical access systems based on passive optical networks (pon),* (1998).

[11] B.G. Haskell, *Buffer and channel sharing by several interframe picturephone coders,* The Bell Syst. Tech. J. **51** (1972), no. 1, 261–289.

[12] D. Heyman and T. Lakshman, *Source models for vbr broadcast-video traffic,* IEEE/ACM Transactions on Networking **4** (1996), no. 1, 40–48.

[13] R. Griffiths H.S. Chin, J.W. Goodge and D.J. Parish, *Statistics of video signals for viewphone-type pictures,* IEEE Journal on Selected Areas in Communications **7** (1989), no. 5, 826–832.

[14] J. Vandewege I. Van de Voorde, C. M. Martin and X.Z. Qiu, *The superpon demonstrator: an exploration of possible evolution pasths for optical access networks,* IEEE Communciation Magazine (2000).

[15] ISO/IEC, *Jpeg* 2000 *verification model version 1.2.*

[16] ———, *Recommendation h.263, video codec for low bit rate communication.*

[17] ITU-T, *Recommendation h.261, video codec for audivisual services at px64kbit/s.*

[18] G.GPON.GSR **ITU-T,** *Gigabit-capable passive optical networks (gpon): Service requirements.*

[19] J. Dey W. Feng J. Kurose J. Stankovic J. Rexford, S. Sen and D. Towsley, *Online smoothing of live, variable-bit-rate video,* Proceedings of International Workshop in Network and Operating Systems Support for Digital Audio and Video (1997).

[20] T. Talley F.D. Smith K. Jeffay, D.L. Stone, *Adaptive,best-effort, delivery of audio and video data across packet-swithced networks,* Third International Workshop on Network and Operating System Support for Digital Audio and Video (1992), 3–14.

[21] S. Kang and H.Y. Yeom, *Transmission of video streams with constant bandwidth allocation,* Computer Communications ***22*** (1999), no. 2, 173–180.

[22] E. Knightly and P. Rossaro, *On the effects of smoothing for deterministic* qos, Distributed Aystems Engineering Journal: Sepcial Issue on Quality of Service **4** (1997), no. 1, 3–15.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[23] E. Knightly and H. Zhang, *Traffic characterization and switch utilization using a deterministic bounding inerval dependent traffic model,* Proc.IEEE INFOCOM'95 (1995).

[24] M. Krunz, *Bandwidth allocation strategise for transporting variable bit rate video trafic,* IEEE/ACM Transactions on Networking **5** (1997), no. 6.

[25] Jack Y.B. Lee, *On a unified architectum for video-on-demand services,* IEEE Transctions on Multimedia **4** (2002), no. 1.

[26] S.C. Liew and C.Y. Tse., *Video aggregation: adapting video traffic for transport over broadband networks* by *integrating data compression and statistical multiplexzng,* IEEE/ACM J. Select. Areas Commun. 14 (1996), no. 6, 1123–1137.

[27] S.C. Liew and C.Y. Tse, *Lossless aggregation:a scheme for transmitting multiple stored vbr video streams over a shared communications channel without* loss *of image quality,* IEEE J, Select. Areas Commun. **15** (1997), no. 6, 1181–1189.

[28] E. Gregori M. Conti and A. Larsson, *Study of the impact of mpeg-1 correlations on video-sources staistical multiplexing,* IEEE J, Select. Areas Commun. **4** (1996), no. 1, 3–15.

[29] S. Keshav M. Grossglauser and D. Tsc, ***A*** *simple and efficient service for multiple time-scale trafic,* IEEE/ACM Transactions on Networking **5** (1997), no. 6.

[30] Jean M. McManus and Keith W. **Ross,** *Video-on-demand over atm:constant-rate transmission and transport,* IEEE JOURNALL ON SELLETED AREAS IN COMMUNICATIONS 14 (1996), no. 6.

[31] B. Melamed and D. Pendarakis, ***A*** *tes-based model for compressed 'star-wars' video,* GLOBECOM'94 (1994), 120–126.

[32] Y.Zhang N. M. Marafih and R. L. Plickholtz, *Modeling and queueing analysis of variabel-bit-rate coded video sources in atm networks,* IEEE Trans. on Circuits and Systems for Video Technology **4** (1994), no. 2, 121–128.

[33] P. Cremonesi P. Manzoni and G. Serazzi, *Workload models of vbr video traffic and their use in resource allocation polices,* IEEE/ACM Transactions on Networking ***7*** (1999), no. 3, 387–397.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[34] N. E. Rikli P. Sen, B. Maglaris and D. Anastassiou, *Models for packed switching of variable-bit-rate video sources,* IEEE J, Select. Areas Commun. (1989).

[35] A.A. lazar P.R. Jelenkovic and N. Semret, *The effect of multiple time scales and subexpotentiality in mpeg video streams on queueing behavior,* IEEE on Selected Areas in Communications **15** (1997), no. 6, 1052–1071.

[36] Y. Ogate R. Kishimoto and F. Inumaru, *Feneration interval distribution characteristics of packetized variable rate video coding data streams in an atm netwrok,* IEEE on Selected Areas in Communications **7** (1989), no. 5, 833–841.

[37] E. Rathgeb, *Modeling and performance comparison of policing mechanisms for atm network,* IEEE J, Select. Areas Commun. **9** (1991), no. 3, 325–334.

[38] ———, *Policing of realistic vbr video trafic in an atm network,* IEEE J, of Digital and Analog Communications Systems 6 (1993), 213–226.

[39] A.R. Reibman and A. Berger, *Trafic discriptors for vbr video teleconferenceing over atm networks,* IEEE/ACM Transactions on Networking **3** (1995), no. 3, 329–339.

[40] ITU release, *Bring broadband to the home: more stream for internet access,* ITU Press Release (2001).

[41] T.V. Lakshman T. Ott and A. Tabatabai, *A scheme for smoothing delay-sensitive traffic offered to atm networks,* IEEE INFORCOM'92 (1992), 776–785.

[42] L. Pinnoo W. Verbiest and B. Voeten, *The impact of the atm concept on video coding,* IEEE Journal on Selected Areas in Communications 6 (1988), no. 9, 1623–1632.

[43] F. Jahanian W.C. Feng and S. Sechrest, *An optimal bandwidth allocation strategy for the delivery of compressed prerecorded video,* ACM Journal of Multimedia Systems **5** (1997), no. 3, 164–176.

[44] N. Ohta Y. Yasude, H. Yasude and F. Kishino, *Packet video transmission through atrn networks,* GLOBECOM'89 (1989), 876–880.

[45] J. Salehi Z. Zhang, J. Kurose and D. Towsley, *Smoothing, statistical multipexing, and call admission control for stored video,* IEEE J. Select. Areas Commun. **15** (1997), no. 6, 1148–1166.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[46] H. Zhang and E. knightly, *Red-vbr:a renegotiation-based apporach to support delay-sensitive vbr video,* ACM Journal of Multimedia Systems 5 (1997), no. 3, 164–176.