1998 International Computer Symposium
Workshop on Computer Networks, Internet, and Multimedia
December 17-19, 1998, N.C.K.U., Tainan, Taiwan, R.O.C.

ARCHITECTURE AND ANALYSIS OF TRAFFIC SCHEDULING FOR VOD SERVICES IN BROADCAST-AND-SELECT WDM NETWORKS[†]

Wen-Yu Tseng, Sy-Yen Kuo
Department of Electrical Engineering
National Taiwan University, Taipei, Taiwan, R.O.C.
wayne@lion.ee.ntu.edu.tw, sykuo@cc.ee.ntu.edu.tw

James Wan-Li Cheng
Department of Computer Engineering & Science
Yuan-Ze University, Chung-Li, Taiwan, R.O.C.
wlcheng@saturn.yzu.edu.tw

ABSTRACT

The VOD (video on demand) service supported by B-ISDN over the WDM networks is emerging as one of the most promising services that may gain wide market acceptance. However, if the traffic scheduling in the WDM networks does not work well and the poor design of the VOD server induces frequent multicast transmissions, the performance of the WDM networks will decrease significantly and the QoS (quality of service) of other services will suffer a lot. In this paper, we present the architecture and analysis of traffic scheduling for VOD services in broadcast-and-select WDM networks. A VOD server follows the general structure except that the PVFS (Proxy Video File Server) is separated from the AS (Archive Server). The AS consists of control processors and network interface to maintain the retrieving hierarchy to request and transmit video segments. The AS is directly connected to the local subscribers through the switch and to the remote subscribers through the WDM networks. If the PVFS can have a larger capacity, the traffic among the VOD server nodes will be reduced, the multicast ratio and the size of the multicast group will be kept small to prevent significant channel dominance due to multicast traffic. Numerical results indicate that the multicast scheduling and the capacity of PVFS have significant impact on the VOD services in the broadcast-and-select WDM networks.

1. INTRODUCTION

With the wavelength division multiplexing (WDM) techniques developed to efficiently utilize the huge bandwidth of an optical fiber [1-3], the VOD service, supported by B-ISDN (Broadband Integrated Service Digital Network), is emerging as one of the most promising services that may gain wide market acceptance [4,5]. The VOD service can be considered as an electronic videotape rental service. Instead of physically picking up and returning videotapes, customers can access encoded videos electronically through the high-speed WDM networks. Video information is stored in information storage components distributed over the network. However, the traffic generated by the VOD systems, which is generally processed as the multimedia type ATM cells, is not studied extensively to evaluate the impact on the entire WDM network. It should be noticed that the session length of the VOD data is pretty large (*e.g.* 1 gigabyte for a single hour of compressed VHS quality video). Besides, the multicast transmission to the VOD server nodes is quite frequent to distribute the VOD data because these nodes may request the same video session at the same time. If the traffic scheduling in WDM networks does not work well and the poor design of the VOD server induces frequent multicast transmissions, the performance of the WDM networks will decrease significantly and the QoS of other services will suffer a lot.

Most VOD studies concentrated on the video information storage structure, network architecture, deadline scheduling of information retrieval, *etc* [5,7]. In addition, the VOD service is assumed to be offered by large-scale public switched networks with ADSL (Asymmetric Digital Subscriber Loop) and HDSL (High-speed Digital Subscriber Loop) [8]. In fact, they investigated the design of VOD servers, the transmissions of videos, and the playback between the VOD server and the clients. When the VOD service receives extensive acceptance from the network subscribers, the VOD servers deployed on the WDM network systems will exist everywhere and the traffic among the servers will increase dramatically. Especially the multicast transmissions among the VOD servers will paralyze of the entire network system easily. Traditional multicast scheduling in the WDM networks can be classified as the reservation-based and the pre-allocation-based protocols [18,19], but the performance of the WDM networks decreases rapidly if the VOD traffic is scheduled with these protocols [9].

In this paper we will present a new architecture and analysis of traffic scheduling for VOD services in broadcast-and-select WDM networks. Although the VOD systems are developed on large-scale public switch networks, the VOD servers are interconnected with the broadcast-and-select WDM networks. The VOD server follows the basic structure proposed in [11], except that the PVFS (Proxy Video File Server) is separated from the AS (Archive Server). A video is partitioned into multiple fixed-size segments. Video segments retrieved from other VOD server nodes will be stored in the PVFS intermittently and out of sequence because the video transfer speed is assumed greater than than the playback speed. If the PVFS has a larger capacity, the traffic among the VOD server nodes will be

---

[†] **Acknowledgment**: This research was supported by the National Science Council, Taiwan, R. O. C., under the Grant NSC 87-2213-E259-007.

reduced as well as the multicast ratio and the size of the multicast group will be kept small to prevent significant channel dominance due to multicast traffic [9]. Numerical results show that the multicast scheduling and the capacity of PVFS have significant impact on the VOD services in the broadcast-and-select WDM networks.

The remainder of the paper is organized as follows. In Section 2 we introduce the system architecture of a broadcast-and-select WDM network, including the structure of a VOD server. The protocol to schedule the general traffic and the VOD traffic is presented in Section 3. Performance analysis is depicted in Section 4. Conclusions are given in Section 5.

2. SYSTEM ARCHITECTURE

2.1 Broadcast-and-Select WDM Network

The architecture for the WDM networks, as shown in Figure 1, uses the passive star coupler to connect N nodes. The passive star coupler is an $N \times N$ broadcast-and-select device interconnected through the optical fibers. The number of wavelengths supported by most WDM systems is significantly less than the number of nodes in anything but a trivial network. Thus the medium is assumed to support $\Omega \leq N$ data wavelengths, λ_0, λ_1, ..., $\lambda_{\Omega-1}$ and the control wavelength λ_Ω; the system is also called a wavelength-limited system. Each node contains a VOD server with other servers operating concurrently to provide services to the local clients or subscribers.

At least two pairs of transmitter and receiver are needed to establish a connection between two nodes. In order to implement the pre-allocation-based protocol, either the TT-FR architecture (one tunable transmitter and one fixed receiver) or the FT-TR architecture (one fixed transmitter and on tunable receiver) would be selected. In this paper we select the FT-TR architecture to access the data wavelengths λ_0, λ_1, ..., $\lambda_{\Omega-1}$ for the pre-allocation-based protocol of unicast traffic because the FT-TR architecture can support multicast transmission much easier than the TT-FR architecture. The tunable receiver at node i, $i = 0, 1, ..., N-1$, can tune to the data wavelength γ_i to listen to packets based on the data wavelength/tuning receiver wavelength allocation map. The fixed transmitter at node i is assigned a data wavelength τ_i as the home channel. Moreover, the FT-FR architecture is also able to broadcast the transmission and multicast status of the overall network on the control channel λ_Ω in the reservation-based fashion.

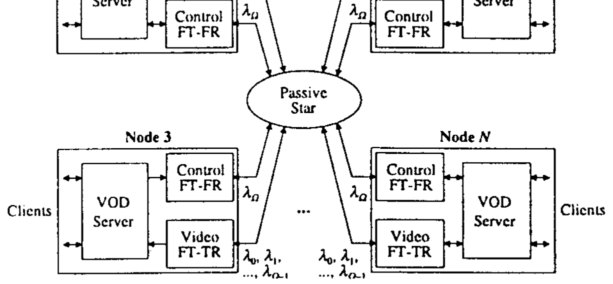


Figure 1: The architecture of the broadcast-and-select WDM network with the VOD server.

The network is packet-switched with fix-sized packets and operates in a slot mode without any packet loss, while the time slot is equal to the packet transmission time plus the latency of the tunable receiver with the *normalized tuning latency* $\delta \ll 1$. The network is assumed that each node has equal distance from the passive star coupler, which makes the propagation delay between all node pairs identical. When two or more fixed transmitters access the same wavelength at the same time slot, also called a *collision*, packet loss is avoided with the arbitration procedure of the protocol. The *destination conflict*, *i.e.* two or more source nodes transmit data to the same node on different channels at the same time slot, will never occur because the wavelength/tuning receiver wavelength allocation map ensures that only one node is allowed to receive on the allocated channel. Besides, all nodes are synchronized to slot boundaries to transmit packets. The buffer at each node is assumed to be infinite and uses link-list addressing to allocate $N-1$ queues, one dedicated queue per node, and to allocate one queue for multicast transmission [24]. This eliminates the head-of-line blocking effect observed in [14], and results in improved network throughput and decreased packet delay.

2.2 VOD Server

The proposed VOD server consists of processing and storage components such as an AS (Archive Server), a LVFS (Local Video File Server), a PVFS (Proxy Video File Server), CPEs (Customer Premise Equipments), and high-speed electronic switch interconnecting these components, as shown in Figure 2. A customer directly interfaces with a set-top box that has capabilities to decode and display video signals. Each set-top box also has a limited-size hard disk for storing prefetched video segments. This makes the store-and-play schemes possible.

The AS is the central part of the VOD server. As shown in Figure 2, the AS is directly connected to the local subscribers through a switch and to the remote subscribers through WDM networks. The AS consists of control processors and a network interface to maintain the retrieving hierarchy to request and transmit video segments. When a subscriber requests a particular video segment, the AS first checks whether the requested segment is in the LVFS. If the requested segment is not in the LVFS, the AS checks whether the requested segment is in the PVFS. If the requested segment is either in the LVFS or in the PVFS, the AS retrieves the segment and send it to the subscriber through the switch. If the requested segment is not in the PVFS, the AS generates the requesting signal through the control FT to the VOD server node that contain the segment. The AS of the requested VOD server node then receives the signal, and the segment will be sent back from the LVFS through the video FT. The segment will be

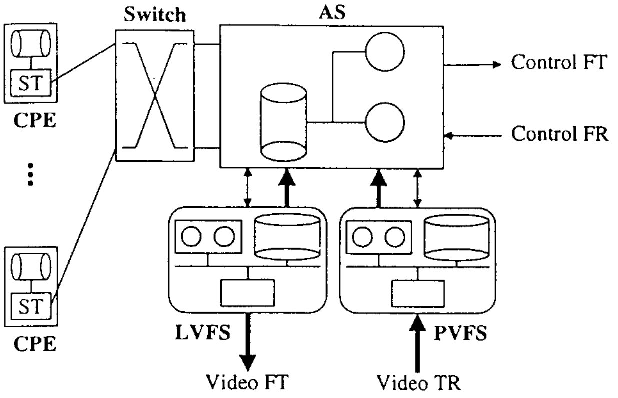


Figure 2: The structure of a VOD server.

stored in the PVFS through the video TR and the AS will retrieve the segment and send to the subscriber.

The storage components, the LVFS and the PVFS, collect large amounts of video segments to supply the resources for VOD service. Both of them can use magnetic disks, magnetic tapes, or high-speed optical disks for storage. However, the resource refresh rates for the LVFS and the PVFS are quite different. The resource refresh for the LVFS may occur according to a static schedule, usually once per week or per month. However, the resource refresh for the PVFS occurs according to the capacity of the PVFS. When a new segment is transferred into the PVFS, an existing segment may have to be deleted to make rooms for the new one. Because the video segments have different popularities [11], segment replacement policies that are best tailored to the requested pattern should be used. Moreover, if the capacity of the PVFS is small, the refresh rate will be very high. Because several VOD server nodes may request the same video segment at the same time, the node that contains the requested segment will induce multicast transmissions. The high refresh rate of the PVFS will increase the multicast ratio on the WDM networks and the size of the multicast group such that this will induce significant channel dominance and intolerable packet delay for unicast traffic [9].

3. MULTICAST SCHEDULING

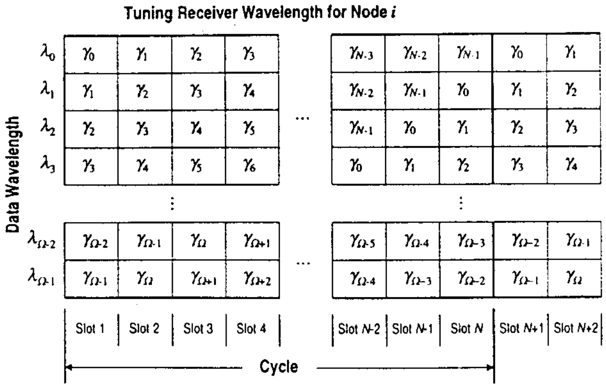
The multicast scheduling protocol originates from the combinational unicast-based protocol [14,9]. The protocol is pre-allocation-based with FT-TR architecture to avoid the architectural requirement of both a tunable transmitter and a tunable receiver. The protocol allocates one channel to node i as the home channel ε_i for the fixed transmitter, while the destination node tunes the receiver to ε_i to receive packets if the packets are transmitted from node i. Because the network is a wavelength-limited system with $\Omega \leq N$, some nodes may be assigned to the same home channel and therefore, collisions need to be avoided by the arbitration procedure.

In order to connect the transmission from the fixed transmitter of the source node to the tunable receiver of the destination node, data wavelength/tuning receiver wavelength allocation map is derived to determine the tuning wavelengths for Ω nodes at each time slot. The wavelength φ_i for the tunable receiver of node i at the time slot t is determined according to the following simple equation:

$$c_i = (i - t + 1) \bmod N \qquad i = 0, 1, \ldots, N-1.$$

$$\varphi_i = \begin{cases} \omega_{c_i} & 0 \leq c_i \leq \grave{U} - 1 \\ \phi & c_i \geq \grave{U} \end{cases} .$$

$\varphi_i = \phi$ means that the receiver is idle. As shown in Figure 3, each node is assigned Ω slots per cycle and idle for the remaining N - Ω slots. The cycle length is N slots to achieve self-routing. However, the source node has to request the access privilege of the home channel to complete the transmission with the FT-TR architecture. The access requests of the home channels are collected through the control channel λ_Ω.

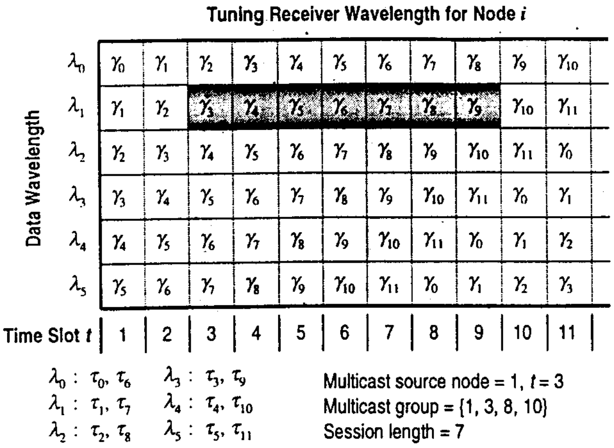


	Slot 1	Slot 2	Slot 3	Slot 4		Slot N-2	Slot N-1	Slot N	Slot N+1	Slot N+2
λ_0	γ_0	γ_1	γ_2	γ_3		γ_{N-3}	γ_{N-2}	γ_{N-1}	γ_0	γ_1
λ_1	γ_1	γ_2	γ_3	γ_4		γ_{N-2}	γ_{N-1}	γ_0	γ_1	γ_2
λ_2	γ_2	γ_3	γ_4	γ_5	...	γ_{N-1}	γ_0	γ_1	γ_2	γ_3
λ_3	γ_3	γ_4	γ_5	γ_6		γ_0	γ_1	γ_2	γ_3	γ_4
			⋮					⋮		
$\lambda_{\Omega-2}$	$\gamma_{\Omega-2}$	$\gamma_{\Omega-1}$	γ_Ω	$\gamma_{\Omega+1}$	...	$\gamma_{\Omega-5}$	$\gamma_{\Omega-4}$	$\gamma_{\Omega-3}$	$\gamma_{\Omega-2}$	$\gamma_{\Omega-1}$
$\lambda_{\Omega-1}$	$\gamma_{\Omega-1}$	γ_Ω	$\gamma_{\Omega+1}$	$\gamma_{\Omega+2}$		$\gamma_{\Omega-4}$	$\gamma_{\Omega-3}$	$\gamma_{\Omega-2}$	$\gamma_{\Omega-1}$	γ_Ω

Figure 3: Data wavelength/tuning receiver wavelength allocation map for the combinational protocol.

The Multicast Slot Reservation (MSR) procedure inside the protocol verifies the multicast packet, makes a reservation of the channel for the packet, and modifies the allocation map on each node. When the node receives a multicast packet with the multicast session length S and the multicast group G (with the size G), the MSR examines first if the packet is qualified to make the reservation according to the packet distance M and the multicast distance M_d. If M is smaller than or equal to M_d, the packet is scheduled as the unicast packet conventionally. The MSR replicates the packet and transmits the replicated packets from the multicast source node to members of the multicast group individually. If M is larger than M_d, the packet is scheduled as the multicast packet. Then the MSR makes the reservation of the home channel of the multicast source node. The number of time slots reserved for the packet is equal to the multicast session length S.

After reserving the home channel, the MSR examines receiver availability of the multicast group and deletes the slots pre-allocated for the receivers in the multicast group during the multicast session. The slots pre-allocated for the receivers not included in the multicast group still follow the original unicast-based protocol. As shown in Figure 4, for instance, the MSR manipulates two qualified multicast packets at different time slots. For the packet at the time slot 3, the MSR reserves ω_1 from the time slot 3 to the time slot 9 as shown in the gradient area. Then the MSR deletes



Tuning Receiver Wavelength for Node *i*

Data Wavelength											
λ_0	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}
λ_1	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}
λ_2	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}	γ_0
λ_3	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}	γ_0	γ_1
λ_4	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}	γ_0	γ_1	γ_2
λ_5	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}	γ_0	γ_1	γ_2	γ_3
Time Slot *t*	1	2	3	4	5	6	7	8	9	10	11

$\lambda_0 : \tau_0, \tau_6$ $\lambda_3 : \tau_3, \tau_9$
$\lambda_1 : \tau_1, \tau_7$ $\lambda_4 : \tau_4, \tau_{10}$
$\lambda_2 : \tau_2, \tau_8$ $\lambda_5 : \tau_5, \tau_{11}$

Multicast source node = 1, *t* = 3
Multicast group = {1, 3, 8, 10}
Session length = 7

Figure 4: The MSR manipulates two qualified multicast packets at different time slots.

10 slots as shown in the shaded areas during the multicast session from the time slot 3 to the time slot 9. This means that no other transmissions to the receivers of nodes 3, 8, and 10 could be accomplished during the multicast session.

The arbitration procedure determines which node owns the access privilege of the home channel at a specific time slot. The request of the home channel is issued through the control channel λ_Ω, and the arbitration is determined by priorities generated from the traffic type and the queue length. The request for multicast transmission has higher priority to reserve the home channel and can be issued at any time slot. The request for unicast transmission has lower priority to reserve the channel and be issued according to the unicast protocol. If the priorities of the requests are the same, the procedure arbitrates the access privilege according to *M* (the distance of *S* and |*G*|) for multicast transmission and the queue length of the destination node for unicast transmission. For multicast transmission, the procedure selects the node with larger *M*. If the active nodes have the multicast packets with the same *M*, the procedure would randomly select the node. For unicast transmission, the procedure selects the node with the longest queue length for the destination node. The queue lengths of the destination nodes with the same home channel are also transmitted through the control channel λ_Ω. If the queue lengths of the active nodes are zero, the procedure would randomly select the node to own the access privilege of the home channel. Therefore, the arbitration procedure can determine the access privilege of the wavelength according to the traffic type and the queue length.

The control channel λ_Ω, which is used for network management, scheduling of wavelengths and time slots, and clock distribution in the conventional WDM networks, also collects broadcast transmission and multicast status among nodes for the unicast-based protocol and the MSR. As shown in Figure 5, each node transmits a control packet on the control channel λ_Ω via the round-robin TDMA to schedule the transmission of the data packet at the head of the queue. The control packet consists of the source node address field, the multicast session length field, and the multicast group member field. The source node address field encodes the node's identity in $\log_2 N$ bits. The multicast session length field represents *S* with binary coding in $\log_2 N$ bits, while *S* is assumed to be no larger than *N*. If the control packet is for unicast transmission, the multicast session length field represents the queue length for the destination nodes. The multicast group member field is of *N* bits such that the *j*th bit of the field is set if node *j* is one of the destination nodes. The multicast group member field also indicates if the packet is a unicast packet or a multicast packet. If only one bit of the field is set, the packet is a unicast packet. Otherwise the packet is a multicast packet. Thus the control packet length is equal to $N + 2\log_2 N$ bits, which is the length of the control slot. After all control packets are received simultaneously by each node, data transmissions can be scheduled deterministically.

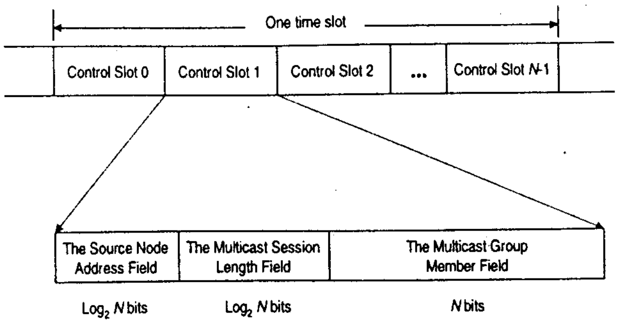


Figure 5: Control channel access via round-robin TDMA.

4. PERFORMANCE ANALYSIS

4.1 Analytic Model

The goal of the performance analysis is to investigate how the design of a VOD server system has an impact on the multicast ratio and the size of a multicast group under multicast traffic in the WDM networks. The analytic model analyzes the impacts on the session length and the group size under multicast traffic, which is proposed in [10]. Multicast traffic can be modeled with the multicast session length *S* and the multicast group *G*, while the size of the multicast group is denoted as *G*. *S* in the given multicast packet has the modified geometric distribution with the mean session length *s*, and the probability of a multicast packet with the session length *k* is

$$\mathrm{Prob}(S = k) = \begin{cases} \dfrac{1}{s} \cdot \left(1 - \dfrac{1}{s}\right)^{k-1}, & k \geq 1 \\ 0, & \text{otherwise} \end{cases}.$$

S is intuitively no less than 1, and assumed to be no larger than *N* in order to maintain the allocation map and to support the protocol described in the previous section.

The size of the multicast group *G* in the given multicast packet has the modified binomial distribution with the mean size $g \cdot (N-3) + 2$, and the probability of a multicast packet with the size of the multicast group *l* is

$$\mathrm{Prob}(G = l) = \begin{cases} \binom{N-3}{l-2} \cdot g^{l-2} \cdot (1-g)^{N-l-1}, & 2 \le l \le N-1 \\ 0, & \text{otherwise} \end{cases}.$$

Because the members of a multicast group contain at least two destination nodes, the random variable G is no less than 2 and no larger than $N - 1$. The assumptions for the model are described as follows:

1. All nodes are assumed to behave independently.
2. A packet arrived at node i, $0 \le i \le N - 1$, follows the Poisson arrival process, with an arrival rate of p packets per unit time slot per node.
3. The packet arrived at node i is the multicast packet with probability q, and the unicast packet with probability $1 - q$.
4. If the packet arrived at node i is the unicast packet, it is destined for j with probability $\lambda = (N-1)^{-1}$ for $j \ne i$, $0 \le j \le N - 1$, and 0 when $j = i$.
5. If the packet arrived at node i is the multicast packet, the session length S and the size of the multicast group G follow the probability distributions described above. The members of the multicast group G are randomly chosen with the uniform probability $\rho = G/(N-1)$ for $i \notin G$.
6. At most one new packet can arrive at each node per time slot.
7. Each queue has the capacity to hold B packets, including the one currently being processed by the transmitter.

The parameters analyzed in this paper are described as follows: $N = 100$ network nodes, $\Omega = 40$ wavelengths. The buffer size of the dedicated queue per node B is 100. Packet generation follows the Poisson arrival process with the parameter $p = 0.1$. In order to analyze the behavior of the VOD traffic, S, G and q have to be described in detail. Because the length of a VOD session is also an important research topic, a small video session is assumed with $S = 10$ and a long video session is assumed with $S = 30$. The capacity of the PVFS will change the values of G and q. Therefore, a small multicast group is assumed with $G = 10$ and a large multicast group is assumed with $G = 50$. The multicast ratio q varies from 0.01 to 0.1. Two scheduling strategies are used to compare the performance according to different M_d. The SS (Separate Scheduling) is represented as $M_d = 0$ and the SMU (Scheduling of Multicast traffic as Unicast traffic) is represented as $M_d = 50$.

Two metrics are discussed. Channel utilization for multicast traffic (U_m) is defined as the number of channels used by multicast traffic per time slot. U_m reveals the channel dominance for multicast traffic. The number of channels that can be used by unicast traffic is $\Omega - U_m$. Packet delay for unicast traffic (D_u) is the number of time slots taken from the time slot that a unicast packet is generated at the source node to the time slot that the packet is received at the destination node. D_u reveals the efficiency of the protocol that schedules the unicast and multicast traffic. Packet delay for multicast traffic (D_m) with $M > M_d$ is almost 0 because the protocol preemptively schedules the multicast packet. The transmission time of the multicast packet with $M < M_d$ is $S \times D_u$.

4.2 Numerical Results

First the small session of the VOD systems is examined for $S = 10$. Figure 6 depicts the channel utilization for multicast traffic U_m versus the multicast ratio q. Note that $U_m = 0$ with SMU because SMU schedules all VOD data as unicast traffic. U_m with SS is proportional to q no matter what the value of G is. This means that the channel utilization is only dependent on q and other unicast services can use more channels if q is small.

Figure 7 depicts the packet delay for unicast traffic D_u versus the multicast ratio q. Note that D_u for $G = 10$ and $G = 50$ have different phenomena. D_u with SMU is smaller than that with SS for $G = 10$, but D_u with SMU is larger than that with SS for $G = 50$. This is because almost all VOD data with SMU for $G = 10$ do not increase the number of unicast packets in the buffer. Those for $G = 50$, however, increase the number of unicast packets in the buffer significantly and induce quite large D_u. The transmission time of the VOD data ($S \times D_u$) will be very large for $G = 50$. Besides, D_u with SS for $G = 10$ is almost the same with $G = 50$ and the range of D_u is acceptable.

Second the small session of the VOD systems is examined for $S = 30$. Figure 8 depicts the channel utilization for multicast traffic U_m versus the multicast ratio q. Note that $U_m \cong 0$ with SMU because SMU schedules almost all VOD data as unicast traffic. Some of the VOD data with $M > M_d$ may arrive and SMU schedules them as multicast packets. U_m with SS is still proportional to q no matter what the value of G is. Compared with Figure 6, however, the value of U_m increases rapidly for $S = 30$. This means that if the defined VOD session is very large, other unicast services can only use small number of channels. For instance, the percentage of channels used by other unicast services for $q = 0.1$ is only 25%. 75% of channels are used for VOD transmissions.

Figure 9 depicts the packet delay for unicast traffic D_u versus the multicast ratio q. Different from Figure 7, D_u with SMU is larger than that with SS no matter what G is. This is because almost all VOD data with SMU for $S = 30$ increase the number of unicast packets in the buffer significantly and induce quite large D_u. D_u with SMU for $G = 10$ is much smaller than that for $G = 50$. The same as Figure 7, D_u with SS for $G = 10$ is almost the same with $G = 50$ and the range of D_u is acceptable.

5. CONCLUSIONS

This paper presented the architecture and analysis of traffic scheduling for VOD services in broadcast-and-select WDM networks. Multicast transmissions among VOD servers usually involve large sessions, which may result in significant channel dominance and large packet delay for

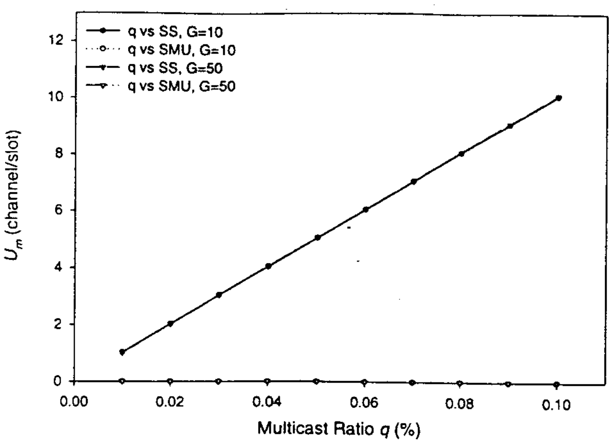


Figure 6: U_m vs. q with $S = 10$.

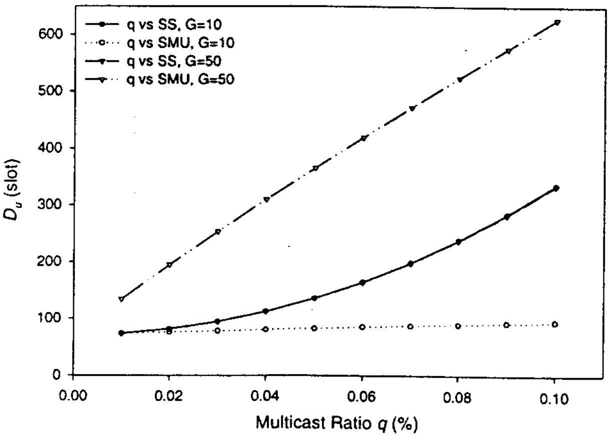


Figure 7: D_u vs. q with $S = 10$.

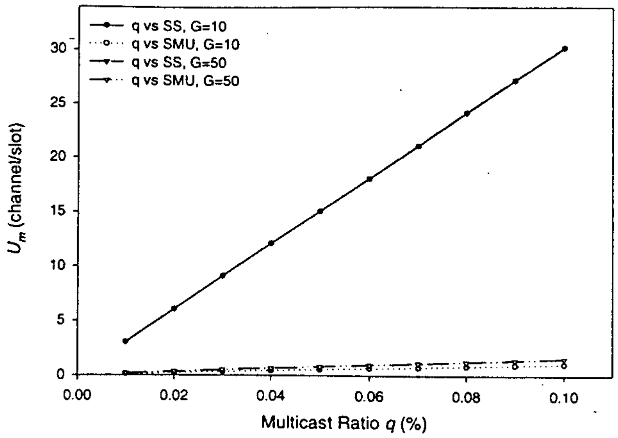


Figure 8: U_m vs. q with $S = 30$.

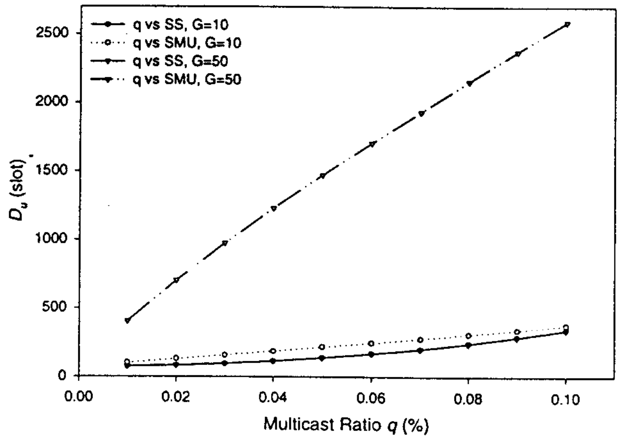


Figure 9: D_u vs. q with $S = 30$.

unicast traffic. Different from other proposed VOD architectures, the PVFS is separated from the AS in order to reduce the multicast ratio and the size of the multicast group for multicast transmissions among the VOD server nodes. The multicast scheduling protocol follows the combinational protocol, which defines multicast distance to verify and reserve the slots for the packets.

According to the numerical results, the VOD service systems in the WDM networks have to define a smaller session length to prevent significant channel dominance. The multicast ratio and the size of the multicast group also have to be kept small to reduce the packet delay. This can be done with enlarging the capacity of the PVFS. For the scheduling protocols, SS and SMU have advantages and disadvantages on scheduling multicast transmissions. SS results in significant channel dominance and smaller packet delay, which is suitable for the real-time VOD systems with enough channels. SMU results in no channel dominance and larger packet delay, which is suitable for the VOD systems with limited channels.

6. REFERENCES

[1] C. A. Brackett, "Dense Wavelength Division Multiplexing Networks: Principles and Applications," *IEEE J. Select. Areas Commun.*, Vol. 8, pp. 948-964, Oct. 1987.

[2] M. H. Ammar and G. N. Rouskas, "On the Performance of Protocols for Collecting Responses Over a Multiple-Access Channel," *IEEE J. Commun.*, Vol. 43, pp. 412-420, Feb. 1995.

[3] P. E. Green Jr., "Optical Networking Update," *IEEE J. Select. Areas Commun.*, Vol. 14, no. 5, pp. 764-779, June 1996.

[4] B. Amin-Salehi, G. D. Flinchbaugh, and L. R. Pate, "Implications of New Network Services on BISDN Capabilities," *INFOCOM'90*, pp. 1038-1045, 1990.

[5] W. D. Sincoskie, "Video On Demand: Is It Feasible?" *GLOBECOM'90*, pp. 201-205, 1990.

[6] H. Y. Woo and C. K. Kim, "Multicast Scheduling for VOD Services," *Multimedia Tools and Applications*, Vol.2, pp. 157-171, 1996.

[7] S. M. Mok and T. S. P. Yum, "Optimal Scheduling of Delayable Video Distribution Networks," *GLOBECOM'92*, pp. 230-234, 1992.

[8] W. Y. Chen and D. L. Waring, "Applicability of ADSL to Support Video Dial Tone in the Copper Loop," *IEEE Commun. Mag.*, Vol. 32, no. 5, pp. 102-109, 1994.

[9] W. Y. Tseng and S. Y. Kuo, "A Combinational Media Access Protocol for Multicast Traffic in Single-Hop WDM LANs," to appear in *GLOBECOM'98*.

[10] W. Y. Tseng, C. C. Sue and S. Y. Kuo, "Performance

Analysis for Unicast and Multicast Traffic in Broadcast-and-Select WDM Networks," submitted to *INFOCOM'99*.

[11] Y. N. Doganata and A. N. Tantawi, "A Cost/Performance Study of Video Servers with Hierachical Storage," in *Proc. Int. Conf. on Multimedia Computing and Systems*, Boston, MA, pp. 393-402, 1994.

[12] B. Mukherjee, "WDM-Based Local Lightwave Networks Part I: Single-Hop Systems," *IEEE Network Mag.*, pp. 12-27, May 1992.

[13] K. M. Sivalingam, K. Bogineni, and P. W. Dowd, "Pre-Allocation Media Access Control Protocols for Multiple Access WDM Photonic Networks," *ACM SIGCOMM'92*, pp. 235-246, Aug. 1992.

[14] K. Bogineni, K. M. Sivalingham, and P. W. Dowd, "Low-Complexity Multiple Access Protocols for Wavelength-Division Multiplexed Photonic Networks," *IEEE J. Select. Areas Commun.*, Vol. 11, pp. 590-604, May 1993.

[15] R. Chipalkatti, Z. Zhang, and A. S. Acampora, "Protocols for Optical Star-Coupler Network Using WDM: Performance and Complexity Study," *IEEE J. Select. Areas Commun.*, Vol. 11, pp. 579-589, May 1993.

[16] P. A. Humblet, R. Ramaswami, and K. N. Sivarajan, "An Efficient Communication Protocol for High-Speed Packet-Switched Multichannel Networks," *IEEE J. Select. Areas Commun.*, Vol. 11, pp. 568-578, May 1993.

[17] M. S. Borella and B. Mukherjee, "Efficient Scheduling of Nonuniform Packet Traffic in a WDM/TDM Local Lightwave Network with Arbitrary Transceiver Tuning Latencies," *IEEE J. Select. Areas Commun.*, Vol. 14, No. 5, pp. 923-934, June 1996.

[18] M. S. Borella and B. Mukherjee, "A Reservation-Based Multicasting Protocol for WDM Local Lightwave Networks," *ICC'95*, pp. 1277-1281, 1995.

[19] G. N. Rouskas and M. H. Ammar, "Multidestination Communication Over Tunable-Receiver Single-Hop WDM Networks," *IEEE J. Select. Areas Commun.*, Vol. 15, No. 3, pp. 501-511, Apr. 1997.

[20] Z. Ortiz, G. N. Rouskas, and H. G. Perros, "Scheduling of Multicast Traffic in Tunable-Receiver WDM Networks with Non-Negligible Tuning Latencies," *ACM SIGCOMM'97*, Cannes, France, Sep. 1997.

[21] E. Modiano, "Unscheduled Multicasts in WDM Broadcast-and-Select Networks," *INFOCOM'98*, 1998.

[22] L. Sahasrabuddhe and B. Mukherjee, "Probability Distribution of the Receiver Busy Time in a Multicasting Local Lightwave Network," *ICC'97*, pp. 116-120, 1997.

[23] Z. Ortiz, G. N. Rouskas, and H. G. Perros, "Scheduling Combined Unicast and Multicast Traffic in Broadcast WDM Networks," *Technical Report TR-97-09, North Carolina State University*, Raleigh, NC, 1997.

[24] T. Kozaki, *et al.*, "32 x 32 Shared Buffer Type ATM Switch VLSIs for B-ISDNs," *IEEE J. Select. Area Commun.*,, Vol.9, pp. 1234-1247, Oct. 1991.