Chapter 7

MAC PROTOCOLS FOR WDM NETWORKS – SURVEY AND SUMMARY

Bo Li
Department of Computer Science
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Maode Ma
Department of Computer Science
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Mounir Hamdi
Department of Computer Science
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Abstract With the proliferation of the world-wide-web (WWW), current local and wide area networks can hardly cope with the huge demand for network bandwidth. As a result, there is a world-wide effort in upgrading current networks with high-capacity fiber-optic links that can potentially deliver Tera-bits/sec bandwidth. The *Wavelength-Division-Multiplexing* (WDM) is an effective technique for utilizing the large bandwidth of an optical fiber. By allowing multiple simultaneous transmission over a number of channels, WDM has the potential to significantly improve the performance of optical networks. The nodes in the network can transmit and receive messages on any of the available channels by employing one or more tunable transmitter(s) and/or tunable receiver(s). Several topologies have been proposed for WDM networks. Of particular interest to us in this chapter is the single-hop topology where a WDM optical network is configured as a broadcast-and-select network in which all the input nodes are combined using a passive star coupler, and the mixed optical information is broadcast to all destinations. To unleash the potential of single-hop WDM passive star networks, effective medium access control (MAC) protocols are needed to efficiently allocate and coordinate the system resources. In this chapter, we present a comprehensive survey of state-the-art MAC protocols for WDM networks so as to give the reader an overview of the research efforts conducted in this area for the past decade. In addition, it can serve as a starting point for further investigation into ways of coping with the current and the anticipated explosion of multimedia information transfer.

1. INTRODUCTION

Wavelength Division Multiplexing (WDM) is the most promising multiplexing technology for optical networks. By using WDM, the optical transmission spectrum is divided into a number of non-overlapping wavelength bands. In particular, allowing multiple WDM channels to coexist on a single fiber, the task of balancing the opto-electronic bandwidth mismatch, can be implemented by designing and developing appropriate WDM optical network architectures and protocols.

WDM optical networks can be designed by using one of two types of architectures, Broadcast-and-Select networks or Wavelength-Routed networks. Typically, the former is used for local area networks, while the latter is used for wide-area networks. A local WDM optical network may be set up by connecting computing nodes via two-way fibers to a passive star coupler, as shown in Fig. 7.1. A node can send its information to the passive star coupler on one available wavelength by using a laser which produces an optical stream modulated with information. The modulated optical streams from multiple transmitting nodes are combined by the passive-star coupler and then the combined streams are separated and transmitted to all the nodes in the network. A destination's receiver is an optical filter and is tuned to one of the wavelengths to receive its designated information stream. Communication between the source and destination nodes is proceeded in one of the following two modes: single-hop, in which communication takes place directly between two nodes [Mukherjee, 1992a], or multihop, in which information from a source to a destination may be routed through the intermediate nodes of the network [Mukherjee, 1992b; Li and Ganz, 1992].

Based on the architectures of WDM optical networks, Medium Access Control (MAC) protocols are required to allocate and coordinate the system resources. The challenge of developing appropriate MAC protocols is to efficiently exploit the potentially vast bandwidth of an optical fiber to meet the

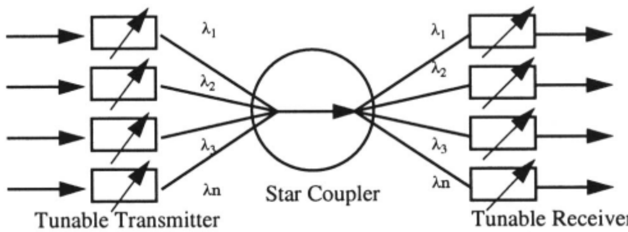

Figure 7.1 Structure of single-hop passive star coupled WDM optical network.

increasing information transmission demand under the constraints of the network resources and the constraints imposed on the transmitted information.

In this chapter, we review the state-of-art MAC protocols in passive star coupler-based WDM networks. Our focus is on the MAC protocols for the single-hop architecture. We will discuss several protocols in some detail to show their importance in the development of MAC protocols for the single-hop passive-star coupler based WDM networks. According to the network service provided to the transmitted information, we roughly divide the MAC protocols into three categories as follows: MAC protocols for packet transmission, MAC protocols for variable-length message transmission, and MAC protocols for real-time message transmission (i.e., MAC protocol with QoS concern).

The reminder of this chapter is organized as follows. Section 2 surveys the MAC protocols for packet transmission. Section 3 reviews the MAC protocols for variable-length message transmission. Section 4 discusses the MAC protocols for real-time service. Section 5 concludes the chapter with a summary.

2. MAC PROTOCOLS FOR PACKET TRANSMISSION

The MAC protocols for packet transmission in single-hop passive-star coupled WDM networks are so called "legacy" protocols, since they are dedicated for fixed length packet transmission, and are often adopted from legacy shared medium networks. In a single-hop network, significant amount of dynamic coordinations among nodes are required in order to access the network resources. According to the coordination schemes, the MAC protocols can be further classified into the following sub-categories.

2.1. NON-PRETRANSMISSION COORDINATION PROTOCOLS

Protocols with non-pretransmission coordination do not need any channels for pretransmission coordination. All the transmission channels are either preassigned to transmitting nodes or accessed by transmitting nodes through contest. These protocols can be categorized accordingly in the following subgroups:

2.1.1 Fixed Assignment. A simple approach, based on the fixed wavelength assignment technique, is a time-division-multiplexing (TDM) extended over a multichannel environment [Chlamtac and Ganz, 1988]. It is pre-determined that a pair of nodes is allowed to communicate with each other in the specified time slots within a cycle on the specified channel. Several extensions to the above protocol have been proposed to improve the performance. One approach, named weighted TDM, assigned different number of time slots to different transmitting nodes according to the traffic load [Rouskas and Ammar, 1995]. Another approach proposed a versatile time-wavelength assignment algorithm [Ganz and Gao, 1994]. Under the condition that a traffic demand matrix is given beforehand, the algorithm can minimize the tuning times and has the ability to reduce transmission delay. Some new algorithms based on [Ganz and Gao, 1994] investigated problems such as the performance of scheduling packet transmissions with an arbitrary traffic matrix; and the effect of the tuning time on the performance [Pieris and Sasaki, 1994; Borella and Mukherjee, 1996; Azizoglu et al., 1996].

2.1.2 Partial Fixed Assignment Protocols. Three partial fixed assignment protocols have been proposed in [Chlamtac and Ganz, 1988]. The first one is the Destination Allocation (DA) protocol. In this protocol, the number of source and destination node pairs can be the same as the number of nodes. A Source Allocation (SA) protocol is also defined in which the control of access to transmission channels is further relaxed. Similar to the SA protocol, Allocation Free (AF) protocol has been proposed, in which all source-destination pairs of computing nodes have full rights to transmit packets on any channel over any time slot duration.

2.1.3 Random Access Protocols. Two slotted-ALOHA based protocols been were proposed in [Dowd, 1991]. In the first protocol, transmissions on different channels are synchronized into slots; while in the second protocol, the transmission is synchronized into mini-slots, and each packet transmission can take multiple such mini-slots. Two similar protocols were presented in [Ganz and Koren, 1991].

2.2. PRETRANSMISSION COORDINATION PROTOCOLS

For protocols that require pretransmission coordination, transmission channels are grouped into control channels and data channels. These protocols can be categorized according to the ways to access the control channels into the following subgroups:

2.2.1 Random Access Protocols. The architecture of the network protocols in this subgroup is as follows. In a single-hop communication network, a control channel is employed. Each node is equipped with a single tunable transmitter and a single tunable receiver.

In [Kavehrad and Sundberg, 1987], three random access protocols including ALOHA, slotted-ALOHA, and CSMA are proposed to access the control channel. ALOHA, CSMA, and N-server switch scheme can be the sub-protocols for the data channels. Under a typical ALOHA protocol, a node transmits a control packet over the control channel at a randomly selected time, after which it immediately transmits a data packet on a data channel, which is specified by the control packet.

In [Mehravari, 1990], an improved protocol named slotted-ALOHA/delayed-ALOHA has been proposed. The characteristics of this protocol is that it requires that a transmitting node delays transmitting data on a data channel until it receives the acknowledgment that its control packet has been successfully received by the destination node. The probability of data channel collision can be decreased. And the performance in terms of throughput can be improved.

In [Sudhakar et al., 1991b], one set of slotted-ALOHA protocols and one set of Reservation-ALOHA protocols have been proposed. Further performance gains are observed when comparing to the the protocols in [Kavehrad and Sundberg, 1987].

In [Sudhakar et al., 1991a], a so-called Multi-Control-Channel protocol is proposed, which aims at improving Reservation-ALOHA-based protocols. All channels are used to transmit control information as well as data information. Control packet transmission is contention-based operation; while data transmission follows it.

A different protocol was proposed in [Li et al., 1995], in which it does require separate control channel for control packet dissemination. The reservation packet are transmitted along with data packet using the same channels in different time slot. This is referred as in-band signaling protocol in [Li et al., 1995].

One potential problem in the above protocols is that there can be conflict at the destination node if multiple transmissions from different source nodes and different wavelengths/channels reach the destination at the same time given that there is only one receiver at the destination node. This is referred as receiver collision, in which the issue was resolved in [Jia and Mukherjee, 1993].

2.2.2 Reservation Protocols. In the Dynamic Time-Wavelength Division Multiple Access (DT-WDMA) protocol, a channel is reserved as control channel and it is accessed only in a pre-assigned TDM fashion. It requires that each node has two transmitters and two receivers [Chen et al., 1990]. One pair of the transceivers are fixed to the control channel, while another pair are tunable to all the data channels. If there are N nodes in the network, N data channels and one control channel are required. Although this protocol cannot avoid receiver collisions, it ensures that exactly one data packet can be successfully received when more than one data packet come to the same destination node simultaneously.

One proposal [Chlamtac and Fumagalli, 1991] to improve the DT-WDMA algorithm is to use an optical delay line to buffer the potentially collided packets, when more than one node transmit data packets to the same destination node at the same time. Its effectiveness depends on the relative capacity of the buffer. Another protocol [Chen and Yum, 1991] also tries to improve the DT-WDMA algorithm by making transmitting nodes remember the information from the previous transmission of control packet and combining this information into the scheduling of future packet transmission.

In [Chipalkatti et al., 1992] and [Chipalkatti et al., 1993], another two protocols aiming at improving the DT-WDMA algorithm are proposed. The first one is called Dynamic Allocation Scheme (DAS), where each node runs an identical algorithm based on a common random number generator with the same seed. The second protocol is named Hybrid TDM. Time on the data channels is divided into frames consisting of several slots. In a certain period of time, one slot will be opened for a transmitting node to transmit data packets to any destination receiver.

A reservation-based Multi-Control-Channel protocol can be found in [Humblet et al., 1993]. In this protocol, x channels $(1 < x < (N/2)$ can be reserved as control channels to transmit control information, where N is the number of the channels in the network. The value of x is a system design parameter, which depends on the ratio of the amount of control information and the amount of actual data information. The objective to reserve multiple control channels in the network is to decrease the overhead of control information processing time as much as possible.

The properties of the "legacy" MAC protocols are summarized as follows:

- Although the protocols using fixed-channel assignment approach can ensure the successful data transmission and reception, they are sensitive to the dynamic bandwidth requirements of the network and they are difficult to scale in terms of the number of nodes. The protocols using contention-based channel assignment approach introduce contention on data channels in order to adapt to the dynamic bandwidth requirements. As a result, either channel collision or receiver collision will occur.

The protocols with contention-based control channel assignment still have either data channel collision or receiver collision because contention is involved in the control channel. While some protocols proposed in [Jeon and Un, 1992; Lee and Un, 1996], have the capability to avoid both collisions by continuously monitoring the network states. The reservation-based protocols, which take fixed control channel assignment approach, can only ensure data transmission without collisions. However, by introducing some information to make the network nodes intelligent, it has potential to avoid receiver collisions as well. It also has potential to accommodate application traffic composed of variable-length messages.

3. MAC PROTOCOLS FOR VARIABLE-LENGTH MESSAGE TRANSMISSION

The "legacy" MAC protocols are designed to handle and schedule fixed length packets. Using these MAC protocols, most of the application level data units (ADU) must be segmented into a sequence of fixed size packets in order to be transmitted over the networks. However, as traffic streams in the real world are often characterized as bursty, consecutive arriving packets in a burst are strongly correlated by having the same destination node. A new idea about this observation is that all the fixed size packets of a burst should be scheduled as a whole and transmitted continuously in a WDM network rather than schedule them on a packet-by-packet basis. Another way of looking at this is that the ADUs should not be segmented. Rather they should be simply scheduled as a whole without interleaving. The main advantages of using a burst-based or message transmission over WDM networks are: 1) To an application, the performance metrics of its data units are more relevant performance measures than ones specified by individual packets; 2) It perfectly fits the current trend of carrying IP traffic over WDM networks; and 3) Message fragmentation and reassembly are not needed.

The first two MAC protocols in [Sudhakar et al., 1991b] proposed for variable-length message transmission are protocols with contention-based control channel assignment. Another two Reservation-ALOHA-based protocols in [Sudhakar et al., 1991b] are presented in order to serve the long holding time traffic of variable-length messages. The first protocol aims to improve the basic slotted-ALOHA-based technique in [Sudhakar et al., 1991b]. The second protocol aims to improve the slotted-ALOHA-based protocol with asynchronous cycles on the different data channels. Data channel collisions can be avoided in the protocols presented in [Sudhakar et al., 1991b].

The protocol in [Bogineni and Dowd, 1991; Dowd and Bogineni, 1992; Bogineni and Dowd, 1992] tries to improve the reservation-based DT-WDMA protocol in [Chen et al., 1990]: The number of nodes is larger than the number of channels; the transmitted data is a variable-length message rather than a fixed length packet; data transmission can start without any delay. Both data collision and receiver collision can be avoided because any message transmission scheduling has to consider the status of the data channels as well as receivers.

There are two protocols called FatMAC in [Sivalingam and Dowd, 1995] and LiteMAC in [Sivalingam and Dowd, 1996] respectively. They try to combine reservation based and preallocation based techniques to schedule variable-length message transmission. FatMAC is a hybrid approach which reserves access preallocated channels through control packets. Transmission is organized into cycles where each of them consists of a reservation phase and a data phase. A reservation specifies the destination, the channel and the message length of next data transmission. LiteMAC protocol is an extension of FatMAC. By LiteMAC protocol, each node is equipped with a tunable transmitter and a tunable receiver rather than a fixed receiver in FatMAC. LiteMAC has more flexibility than FatMAC because of the usage of tunable receiver and its special scheduling mechanism. So that more complicated scheduling algorithms could be used to achieve better performance than FatMAC. Both FatMAC and LiteMAC have the ability to transmit variable-length messages by effect scheduling without collisions. Their performance have been proved to be better than the preallocation based protocols while less transmission channels are used than reservation based protocols. With these two protocols, low average message delay and high channel utilization can be expected.

3.1. A RESERVATION-BASED MAC PROTOCOL FOR VARIABLE-LENGTH MESSAGES

In [Jia et al., 1995], based on the protocols in [Bogineni and Dowd, 1991; Dowd and Bogineni, 1992; Bogineni and Dowd, 1992], an intelligent reservation based protocol for scheduling variable-length message transmission has been proposed. The protocol employs some global information of the network to avoid both data channel collisions and receiver collisions while message transmission is scheduled. Its ability to avoid both collisions makes this protocol a milestone in the development of MAC protocols for WDM optical networks.

The network consists of M nodes and $W + 1$ WDM channels. W channels are used as data channels, the other channel is the control channel. Each node is equipped with a fixed transmitter and a fixed receiver for the control channel, and a tunable transmitter and a tunable receiver to access the data channels. The time on the data channels is divided into data slots. It is assumed that there is a network-wide synchronization of data slots over all data channels. The duration of a data slot is equal to the transmission time of a fixed-length data packet. A node generates variable-length messages, each of which contains one or more fixed-length data packets. On the control channel, time is divided into control frames. A control frame consists of M control slots. A control slot has several fields such as address of destination node, the length of the message, etc. A time division multiple access protocol is employed to access the control channel so that the collision of control packets can be avoided.

Before a node sends a message, it needs to transmit a control packet on the control channel in its control slot. After one round-trip propagation delay, all the nodes in the network will receive the control packet. Then a distributed scheduling algorithm is invoked at each node to determine the data channel and the time duration over which the message will be transmitted. Once a message is scheduled, the transmitter will tune to the selected data channel and transmit the scheduled message at the scheduled transmission time. When the message arrives at its destination node, the receiver should have been tuned to the same data channel to receive the message.

The data channel assignment algorithm determines the data channel and the time duration over which the message will be transmitted. The algorithm schedules message transmissions based on some global information in order to avoid the data channel collisions and the receiver collisions. The global information is expressed through two tables which reside on each node. One table is the Receiver Available Time table (RAT). RAT is an array of M elements, one for each node. $RAT[i] = n$, where $i = 1, 2, ..., M$, means that node i's receiver will become free after n data slots. If $n = 0$, then node i's receiver is currently idle, and no reception is scheduled for it as yet. RAT is needed for avoiding receiver collisions. Another table is named the Channel Available Time table (CAT). CAT is an array of W elements, one for each data channel. $CAT[k] = m$, where $k = 1, 2, ..., W$, means that data channel k will be available after m data slots. If $CAT[k] = 0$, data channel k is currently available. CAT is needed to avoid collisions on data channels. Local and identical copies of these two tables are on each node. They contain consistent information on the messages whose transmissions have been scheduled but not yet transmitted so far. The contents of the tables are relative to current time. Three data channel assignment algorithms have been proposed. The fundamental one is named Earliest Available Time Scheduling algorithm (EATS). This algorithm schedules the transmission of a message by selecting a data channel which is the earliest available.

This reservation-based protocol has been shown to have a quite good performance while it can avoid data channel collisions and receiver collisions.

3.2. A RECEIVER-ORIENTED MAC PROTOCOL FOR VARIABLE-LENGTH MESSAGES

Some related protocols have been proposed to improve the performance of the network based on the same system architecture of [Jia et al., 1995]. In [Muir and Garcia-Luna-Aceves, 1996], the proposed protocol tries to avoid the head-of-queue blocking during the channel assignment procedure by introducing the concept of "destination queue " to make each node maintain M queues, where M is the number of nodes in the network. In [Hamidzadeh et al., 1999], the authors notice that the performance of the network could be further improved by the way of exploiting more existing global information of the network and the transmitted messages. From this point of view, a general scheduling scheme, which combines the message sequencing techniques with the channel assignment algorithms in [Jia et al., 1995], is proposed to schedule variable-length message transmission. In [Ma et al.,], as an example of the general scheduling scheme in [Hamidzadeh et al., 1999], a novel scheduling algorithm is proposed. The new algorithm, named Receiver-Oriented Earliest Available Time Scheduling (RO-EATS), decides the sequence of the message transmission by the information of the receiver's states to decrease message transmission blocking caused by avoiding receiver collisions.

The RO-EATS scheduling algorithm employs the same system structure and network service as those of the protocol in [Jia et al., 1995] to form a receiver-oriented MAC protocol, which is an extension to the protocol in [Jia et al., 1995]. The logic structure of the system model for the RO-EATS protocol can be expressed as in Fig. 7.2. The new protocol proceeds messages' transmission and reception is the same as that of the protocol in [Jia et al., 1995]. The difference between the two protocols is on the scheduling algorithm for message transmission.

The RO-EATS algorithm works as follows. It first considers the earliest available receiver among all the nodes in the network and then selects a message which is destined to this receiver from those which are ready and identified by the control frame. After that, a channel is selected and assigned to the selected message by the principle of EATS algorithm.

The scheme to choose a suitable message to transmit is based on the information of the states of the receivers presented in RAT. The objective of this scheme is to avoid lots of messages going to one or a few nodes at the same time and try to raise the channels utilization. The motivation of the algorithm comes from the observation that two consecutive messages with the same destination may not fully use the available channels when EATS algorithm is employed. The new algorithm enforces the idea of scheduling two consecutive messages away from going to the same destination node. The RO-EATS algorithm always checks the table of RAT to see which node is the least visited destination and

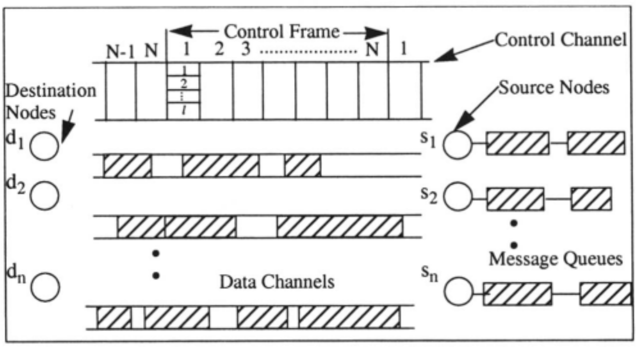

Figure 7.2 The system model of RO-EATS protocol.

to choose the message which is destined to this node to transmit. In this way, average message delay can be shown too be quite low and channel utilization can be shown to be high.

4. MAC PROTOCOLS FOR REAL-TIME SERVICE

An important function of high-speed computer networks such as WDM optical networks is to provide real-time service to time-constrained application streams such as video or audio information. Most of the MAC protocols that provide real-time service on passive star-coupled WDM optical networks are protocols with reservation based pre-coordination. According to the type of the real-time service provided to the transmitted messages, the MAC protocols for real-time service can be classified into three types: protocols with best-effort service, protocols with deterministic guaranteed service, and protocols with statistical guaranteed service

4.1. MAC PROTOCOLS FOR BEST-EFFORT REAL-TIME SERVICE

A protocol named Time-Deterministic Time and Wavelength Division Multiple Access (TD-TWDMA) presented in [Jonsson et al., 1997] provides services for both hard real-time messages and soft real-time messages for single destination, multicast, and broadcast transmissions. All channels can be accessed by a fixed-assignment method, which is a TDM approach. Using this approach, each channel is divided into time-slots. Each node has a number of slots for hard real-time message transmissions. Soft real-time messages can be transmitted if there is no hard real-time messages requiring service. Each node is equipped with one fixed-transmitter and tunable receivers. The transmitter is fixed to its assigned channel, while the receiver can be tuned over all channels in the network. Each node has a specified channel because the number of nodes, C, is equal to the number of channels, M in the network. At each node, there are $2 \times M$ queues, M queues for the hard real-time messages, another M queues for the soft real-time messages. For each type of queues, one queue is for broadcast and $M - 1$ queues for the single destination. The messages in the broadcast queue can be either control information or data to be broadcast. The protocol works as follows: First it sends a broadcast slot containing the control information; then invokes the slot-allocation algorithm to determine the slots used to transmit the data information; at last, each node tunes to the specified channel to receive the data. The slot-allocation algorithm follows the static priority approach. The basic idea of the algorithm can be summarized as: 1) The M hard real-time message queues have higher priorities; while the M soft real-time message queues have lower priorities; 2) Each queue in each group has a fixed priority; while the queues for broadcast have the highest priority in each group; 3) Message transmission scheduling is based on the queue priority; and 4) For the hard real-time messages, if transmission delay is over their deadlines, these messages will be dropped; while for the soft real-time messages, they will be scheduled whether they are beyond their deadlines or not.

A reservation-based MAC protocol for best-effort real-time service can be found in [Ma et al., 1999]. This protocol is for the same network structure as that in [Jia et al., 1995]. Both hard real-time and soft real-time variable-length message transmissions have been considered. The scheduling algorithms for the protocol are based on the time related dynamic priority scheme, namely Minimum Laxity First (MLF) scheduling. This protocol employs global information of the network as well as the transmitted messages to ensure zero message loss rate caused by both data channel collisions and receiver collisions and decrease the message loss rate caused by network delay.

4.2. MAC PROTOCOLS FOR DETERMINISTIC GUARANTEED SERVICE

In [Tyan et al., 1996], a pre-allocation based channel access protocol is proposed to provide deterministic timing guarantees to support time constrained communication in a single-hop passive star-coupled WDM optical network. This protocol takes a passive star-coupled broadcast-and-select network architecture in which N stations are connected to a passive star coupler with W different wavelength channels. Each of the W channels is slotted and shared by the N stations by means of a TDM approach. The slots on each channel are pre-assigned to the transmitters. A schedule specifies, for each channel, which slots are used for data transmission from node *i* to node *j*, where $1 \leq i \leq N$, $1 \leq j \leq N, i \leq j$. Each node of the network can be equipped with a pair of tunable transmitters and tunable receivers which can be tuned over all the wavelengths. Each real time message stream with source and destination nodes specified is characterized with two parameters, relative message deadline D_i and maximum message size C_i that can arrive within any time interval of length D_i. A scheme called Binary Splitting Scheme (BSS) is proposed to assign each message stream sufficient and well-spaced slots to fulfill its timing requirement. Given a set of real-time message streams *M* specified by the maximum length of each stream C_i and the relative deadline of each stream D_i, this scheme can allocate time slots over as few channels as possible in such a way that at least C_i slots are assigned to M_i in any time window of size D_i slots. So that the real-time constraints of the message streams can be guaranteed.

In [Ma and Hamdi, 2000], a reservation-based MAC protocol for deterministic guaranteed real-time service can be found. This protocol is for the same network structure as that in [Jia et al., 1995]. In [Ma and Hamdi, 2000], a systematic scheme is proposed to provide deterministic guaranteed real-time service for application streams composed of variable-length messages. It includes admission control policy, traffic regularity, and message transmission scheduling algorithm. A traffic intensity oriented admission control policy is developed to manage flow level traffic. A *g*-regularity scheme based on the Max-plus algebra theory is employed to shape the traffic. An Adaptive Round-Robin and Earliest Available Time Scheduling (ARR-EATS) algorithm is proposed to schedule variable-length message transmission. All of the above are integrated to ensure that the deterministic guaranteed real-time service can be achieved.

4.3. MAC PROTOCOLS FOR STATISTICALLY GUARANTEED SERVICE

A reservation-based protocol to provide statistically guaranteed real-time services in WDM optical token LAN can be found in [Yan et al., 1996]. In the network, there are *M* nodes and *W* + 1 channels. One of the channels is the control channel, while the others are data channels. Unlike the structure in [Jia et al., 1995], the control channel is accessed by token passing. At each node, there is a fixed receiver and transmitter tuned to the control channel. There is also a tunable transmitter which can be tuned to any of the data channels. There are one or more receivers fixed to certain data channels.

The protocol provides transmission service to either real-time or non real-time messages. The packets in the traffic may have variable-length but bounded by a maximum value. At each node, there are *W* queues, each corresponds to one of the channels in the network. The messages come into one of the queues according to the information of their destination nodes and the information of the channels which connect to the corresponding destination nodes. The protocol works as follows: A token exists on the control channel to ensure collision free transmission on data channels. The token has a designated node *K*. Every node can read the contents of the token and updates its local status table by the information in the fields of the token. When node *K* observes the token on the network, it will check the available channels. If there are no channels available, node *K* gives up this opportunity to send its queued packets. Otherwise, the Priority Index Algorithm (PIA) is invoked to evaluate the priority of each message queue on node *K* and then use Transmitter Scheduling Algorithm (TSA) to determine the transmission channel. Also the Flying Target Algorithm (FTA) is used to decide the next destination of the control token. After all these have completed, node *K*'s status and scheduling result will be written into the token. Then the token on the node *K* will be sent out. And the scheduled packets will be transmitted.

4.4. A STATE-OF-ART MAC PROTOCOL FOR STATISTICALLY GUARANTEED SERVICE

A novel reservation-based MAC protocol is proposed in [Li and Qin, 1998] to support statistically guaranteed real-time service in WDM networks by using a hierarchical scheduling framework. It shows that such hierarchical scheduling is essential for achieving scalability such that larger input-output ports can be accommodated.

The major advantage of the protocol in [Li and Qin, 1998] over that in [Yan et al., 1996], is that it divides the scheduling issue into flow scheduling and transmission scheduling. The former is responsible for considering the order of traffic streams to be transmitted. The latter is to decide the order of the packets transmission. The packets involved in the transmission scheduling are those selected from the traffic streams by the flow scheduling scheme. Compared with the protocol in [Yan et al., 1996], this protocol is expected to diminish the ratio of the packets which are over their deadlines. Another advantage of the protocol is that it employs a multiple retransmission scheme to alleviate the result of unsuccessful packet transmissions. This work is developed for a similar network structure as that in [Jia et al., 1995]. It has further enriched the structure of the network transmitting node to support the proposed protocol. The detailed architecture of the network and transmitting node is presented in Fig. 7.3.

4.4.1 Network Configuration. An $N \times N$ packet-switching system configuration with single hop connection is considered. The switching and

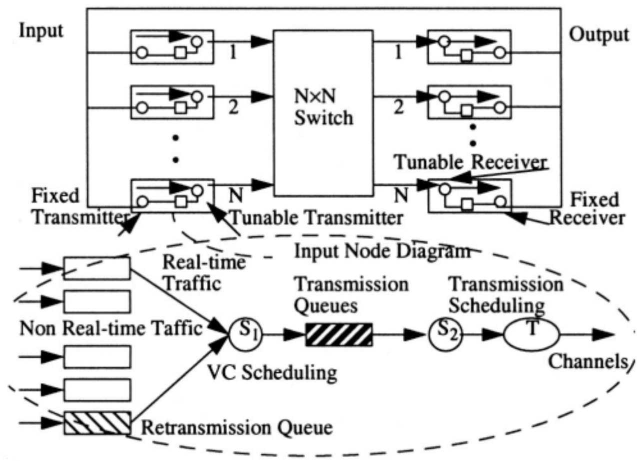

Figure 7.3 The structure of the networks and transmitting node.

routing can possibly be implemented by a recently designed switch, which can support link capacity between 2.4 – $10Gb/s$. In this architecture, a small delay line is employed to separate the data packet and header for allowing the header processing. The queues shown in Fig. 7.3 can be implemented by optical fiber loops. There are $W + 1$ wavelengths in the system, in which W channels are data channels, another channel is control channel. Each input node is equipped with one transmitter fixed to the control channel for control packet transmission, and one data packet transmitter for data packet transmission. Each output node has one control packet receiver fixed to the control channel for control packet reception, and one data packet receiver for data packet reception. Data transmitter and data receiver are tunable, and are assumed to be able to tune to all the channels in the system. Control channels are assumed to be synchronized with data channels. Data channels are divided into data slots and control channel is divided into control slots. The control slot is further divided into mini-slots with each holding a control packet from a node. The number of mini-slots is assumed to equal the number of input/output nodes in the switch, i.e., each input node has its dedicated mini-slot in each control slot then some kind of mini-slot assignment based on either a round robin or contention can be adopted. The hierarchical scheduling frame work is readily applicable to handle variable length packets.

Each input port is fed with multiple traffic streams, and it is noted as a Virtual Circuit. Each VC has its own separate queue. Two generic types of VCs are considered in the switch: one is real-time traffic, the other is non-real-time datagram type of traffic. Real-time traffic has QoS requirements such as delay bound, in-sequence delivery, and packet loss rate, and non-real-time traffic either has no QoS requirement such as the "best effort" UBR or certain minimum QoS such as minimum bandwidth guarantee like ABR.

4.4.2 Scheduling Mechanisms. The are two stages in scheduling procedure: the VC scheduling and the transmission scheduling. Each input node has one transmission queue for scheduling the packet transmission. VC scheduling is responsible for scheduling packets from VC queues to the transmission queue, the transmission scheduling handles the coordination of packet transmissions among multiple input nodes. A simple round robin scheme is adopted in the VC scheduling and a random scheduling with age priority is used in the transmission scheduling.

One of the major complications in an on-demand scheduling algorithm is that the reservation might fail due to either output conflict or channel conflict, thus re-scheduling would be necessary. The outcome of the reservation will be known to all nodes at the end of one round-trip propagation delay. If a scheduling fails, there is a decision that has to be made whether re-scheduling the same packet or scheduling a new packet from another VC. If the failed reservation is from a real-time VC, it certainly makes sense to re-schedule the very same packet as soon as possible. The very same RT reservation packet will be re-transmitted immediately in the next control slot, thus no other new reservation either from RT VCs or non-RT VCs of the same source node can be initiated. Some kind of randomized algorithms, such as the exponential backoff algorithm might be needed in the scheme preferred in which an immediate re-transmission of the reservation packet will take place once a real-time reservation packet fails.

The more intriguing part of the scheduling algorithm is the re-scheduling of non-RT traffic. If real-time traffic has more stringent QoS requirements, in particular the delay, it is worth trading the complication of switch design and the non-real-time best effort traffic performance for the real-time traffic performance. Two scenarios, non-preemptive and preemptive, are considered. Under the non-preemptive case, a failed reservation of non-RT VCs will be re-transmitted at the next control slot, exactly like RT VCs. This is non-preemptive in nature in that once a non-RT VC reservation is made, it can not be interrupted by any other RT VCs. Under the preemptive case, however, at the time when the reservation of a non-RT VC fails (at the end of one propagation delay after the reservation is sent), since there might exist packet(s) from a real-time VC of the same input node ready for transmission, i.e., an RT VC with non empty queues and no HOL blocking, in order to avoid the non-RT traffic blocking the RT traffic, one additional queue called re-transmission queue is introduced for temporarily holding the non-RT packet for later re-transmission if the previous reservation of a non-RT packet is unsuccessful. A number of immediate observations can be drawn on the behavior of the re-transmission queue:

No real-time reservation packet will be put on the re-transmission queue.

The non-RT reservation packet on the re-transmission queue will be sent in the next available control slot when there is no RT VC reservation packet available. Notice, this includes a new reservation packet from an RT VC, or a previous failed RT reservation packet. Thus, the non-RT reservation packet on the re-transmission queue might have to wait for more than one reservation cycle.

No new non-RT reservation can be initiated if the re-transmission queue is non empty.

The queue length of the re-transmission queue is finite, which can be shown to be bounded by the pipelining depth, defined as ratio between the propagation delay and the the size of control packet.

VC Scheduling Algorithm. The VC scheduling essentially selects a proper candidate if there is any from multiple application VC queues, and moves it into the transmission queue ready for transmission scheduling. Notice there are differences in handling real-time and non-real time traffic. For real-time traffic, the in-sequence delivery mandates that a new transmission cannot be initiated until the successful reservation of current packet transmission. On the other hand, the non-real-time VC does not have this constraint. Depending on the VC type, different scheduling at the VC scheduling level can be used. The non-real-time traffic are assumed as "best effort" UBR type. This results in the real-time traffic having strict priority over the non-real-time traffic. Within each type, a simple round robin policy is employed.

Transmission Scheduling Algorithm. During the transmission scheduling stage, a control packet if any is sent in the corresponding control mini-slot. Depending on whether the transmission queue or/and the re-transmission queue is empty or not, and depending on the VC type, a number of scenarios are possible. The following summarizes the priority order of reservation packet transmission.

A failed real-time reservation packet will be re-transmitted in the next control slot.

A new non-blocked real-time VC's reservation packet will be transmitted.

A non-real-time reservation packet from the re-transmission queue will be transmitted.

A new non-real-time VC's reservation packet will be transmitted.

Within one control slot, each node has one control mini-slot for sending out its reservation. The following is the order for transmission scheduling to avoid both channel conflict and output conflict.

Real-time packet has higher priority than non-real-time packet.

Real-time (non-real-time) traffic with earlier arrival has higher priority than real-time (non-real-time) packet of late arrival (age priority).

Random selection with equal probability further breaks the tie.

5. SUMMARY

This chapter has summarized state-of-the art medium access control protocols for wavelength division multiplexing (WDM) networks. Depending on the characteristics, complexity, and capabilities of these MAC protocols, we have classified them under non-pretransmission coordination protocols, pretransmission coordination protocols, variable-length transmission protocols, and quality-of-service oriented protocols. Detailed architectural, qualitative and quantitative descriptions of various protocols within each category have been provided. Most of these protocols are targeted towards local and metropolitan area environments. This chapter should serve as a good starting point for researchers working on this area so as to give them an overview of the research efforts conducted for the past decade. In addition, it presents the fundamentals for further investigation into ways of coping with the current and the anticipated explosion of multimedia information transfer.

References

Azizoglu, M., Barry, R. A., and Mokhtar, A. (1996). Impact of tuning delay on the performance of bandwidth-limited optical broadcast networks with uniform traffic. *IEEE J. Sel. Areas Comm.*, 14(6):935–944.

Bogineni, K. and Dowd, P. W. (1991). A collisionless media access protocols for high speed communication in optically interconnected parallel computers. *SPIE*, 1577:276–287.

Bogineni, K. and Dowd, P. W. (1992). A collisionless multiple access protocol for a wavelength division multiplexed star-coupled configuration: Architecture and performance analysis. *IEEE/OSA J. Lightwave Tech.*, 10(11): 1688–1699.

Borella, M. S. and Mukherjee, B. (1996). Efficient scheduling of nonuniform packet traffic in a WDM/TDM local lightwave network with arbitrary transceiver tuning latencies. *IEEE J. Sel. Areas Comm.*, 14(6):923–934.

Chen, M. and Yum, T.-S. (1991). A conflict-free protocol for optical WDM networks. In *IEEE GLOBECOM '91*, pages 1276–1291, Phoenix, AZ.

Chen, M.-S., Dono, N. R., and Ramaswami, R. (1990). A media access protocol for packet-switched wavelength division multiaccess metropolitan area networks. *IEEE J. Sel. Areas Comm.*, 8(8): 1048–1057.

Chipalkatti, R., Zhang, Z., and Acampora, A. S. (1992). High-speed communication protocols for optical star networks using WDM. In *Proc. INFOCOM '92*, pages 2124–2133, Florence, Italy.

Chipalkatti, R., Zhang, Z., and Acampora, A. S. (1993). Protocols for optical star-coupler network using WDM: Performance and complexity study. *IEEE J. Sel. Areas Comm.*, 11(4):579–589.

Chlamtac, I. and Fumagalli, A. (1991). Quadro-stars: High performance optical WDM star networks. In *Proc. IEEE GLOBECOM '91*, pages 1224–1229, Phoenix, AZ.

Chlamtac, I. and Ganz, A. (1988). Channel allocation protocols in frequency – time controlled high speed networks. *IEEE Trans. Comm.*, 36(4):430–440.

Dowd, P. W. (1991). Random access protocols for high speed interprocessor communication based on an optical passive star topology. *IEEE/OSA J. Lightwave Tech.*, 9(6):799–808.

Dowd, P. W. and Bogineni, K. (1992). Simulation analysis of a collisionless multiple access protocol for a wavelength division multiplexed star-coupled configuration. In *Proceedings of the 25th Annual Simulation Symposium*, Orlando, FL.

Ganz, A. and Gao, Y. (1994). Time-wavelength assignment algorithms for high performance WDM star based systems. *IEEE Trans. Comm.*, 42(2-3-4): 1827–1836.

Ganz, A. and Koren, Z. (1991). WDM passive star protocols and performance analysis. In *Proc. INFOCOM '91*, pages 991–1000, Bal Harbour, FL.

Hamidzadeh, B., Ma, M., and Hamdi, M. (1999). Message sequencing techniques for on-line scheduling in WDM networks. *IEEE/OSA J. Lightwave Tech.*, 17(8):1309–1319.

Humblet, P. A., Ramaswami, R., and Sivarajan, K. N. (1993). An efficient communication protocols for high-speed packet-switched multichannel networks. *IEEE J. Sel. Areas Comm.*, 11(4):568–578.

Jeon, H. and Un, C. (1992). Contention-based reservation protocols in multiwavelength optical networks with a passive star topology. In *Proc. ICC*, pages 1473–1477.

Jia, F. and Mukherjee, B. (1993). The receiver collision avoidance (rca) protocol for a single-hop lightwave network. *IEEE/OSA J. Lightwave Tech.*, 11(5/6):1052–1065.

Jia, F., Mukherjee, B., and Iness, J. (1995). Scheduling variable-length messages in a single-hop multichannel local lightwave network. *IEEE/ACM Trans. Networking*, 3(4):477–487.

Jonsson, M., Borjesson, K., and Legardt, M. (1997). Dynamic timedeterministic traffic in a fiber optic WDM star network. In *Proceedings. Ninth Euromicro Workshop on Real Time Systems*, pages 25–33, Toledo, Spain.

Kavehrad, I. M. I. H. M. and Sundberg, C.-E. W. (1987). Protocols for very high speed optical fiber local area networks using a passive star topology. *IEEE/OSA. J. Lightwave Tech.*, 5(12):1782–1794.

Lee, J. H. and Un, C. K. (1996). Dynamic scheduling protocol for variable-sized messages in a WDM-based local network. *IEEE/OSA J. Lightwave Tech.*, 14(7):1595–1600.

Li, B. and Ganz, A. (1992). Virtual topologies for WDM star LANs - the regular structure approach. In *Proc. INFOCOM*, Florence, Italy.

Li, B., Ganz, A., and Krishna, M. (1995). A novel transmission coordination scheme for single hop lightwave networks. In *Proc. GLOBECOM*, Singapore.

Li, B. and Qin, Y. (1998). Traffic scheduling in a photonic packet switching system with qos guarantee. *IEEE / OSA J. Lightwave Tech.*, 16(12):2281–2295.

Ma, M. and Hamdi, M. (2000). Providing deterministic quality-of-service guarantees on WDM optical networks. Submitted to ICC.

Ma, M., Hamidzadeh, B., and Hamdi, M. A receiver-oriented message scheduling algorithm for WDM lightwave networks. Accepted by Computer Networks and ISDN Systems.

Ma, M., Hamidzadeh, B., and Hamdi, M. (1999). Efficient scheduling algorithms for real-time service on WDM optical networks. *Photonic Network Communications*, 1(2).

Mehravari, N. (1990). Performance and protocol improvements for very high-speed optical fiber local area networks using a passive star topology. *IEEE/ OSA J. Lightwave Tech.*, 8(4):520–530.

Muir, A. and Garcia-Luna-Aceves, J. J. (1996). Distributed queue packet scheduling algorithms for WDM-based networks. In *Proc. INFOCOM '96*.

Mukherjee, B. (1992a). WDM-based local lightwave networks – Part I: Single-hop systems. *IEEE Network*, 6(3):12–27.

Mukherjee, B. (1992b). WDM-based local lightwave networks – Part II: Multihop systems. *IEEE Network*, 6(4):20–32.

Pieris, G. R. and Sasaki, G. H. (1994). Scheduling transmissions in WDM broadcast-and-select networks. *IEEE/ACM Trans. Networking*, 2(2):105–110.

Rouskas, G. N. and Ammar, M. H. (1995). Analysis and optimization of transmission schedules for single-hop WDM networks. *IEEE/ACM Trans. Networking*, 3(2):211–221.

Sivalingam, K. M. and Dowd, P. W. (1995). A multilevel WDM access protocol for an optically interconnected multiprocessor system. *IEEE/OSA J. Lightwave Tech.*, 13(11):2152–2167.

Sivalingam, K. M. and Dowd, P. W. (1996). A lightweight media access protocol for a -based distributed shared memory system. In *Proc. INFOCOM '96*, pages 946–953.

Sudhakar, G. N. M., Georganas, N., and Kavehrad, M. (1991a). Multi-control channel for very high-speed optical fiber local area networks and their interconnections using passive star topology. In *Proc. IEEE GlOBECOM '91*, pages 624–628, Phoenix, AZ.

Sudhakar, G. N. M., Kavehrad, M., and Georganas, N. (1991b). lotted aloha and reservation aloha protocols for very high-speed optical fiber local area networks using passive star topology. *IEEE/OSA J. Lightwave Tech.*, 9(10): 1411 –1422.

Tyan, H.-Y., Hou, C.-J., Wang, B., and Han, C.-C. (1996). On supporting time-constrained communications in WDMA-based star-coupled optical networks. In *Proceedings. 17th IEEE Real-Time Systems Symposium*, pages 175–184, Los Alamitos, CA, USA.

Yan, A., Ganz, A., and Krishna, C. M. (1996). A distributed adaptive protocol providing real-time services on WDM-based LANs.
IEEE/OSA J. Lightwave Tech., 14(6):1245–1254.