Computer Networks and ISDN Systems, February, 1999.

Architectural Considerations in the design of WDM-based Optical Access Networks

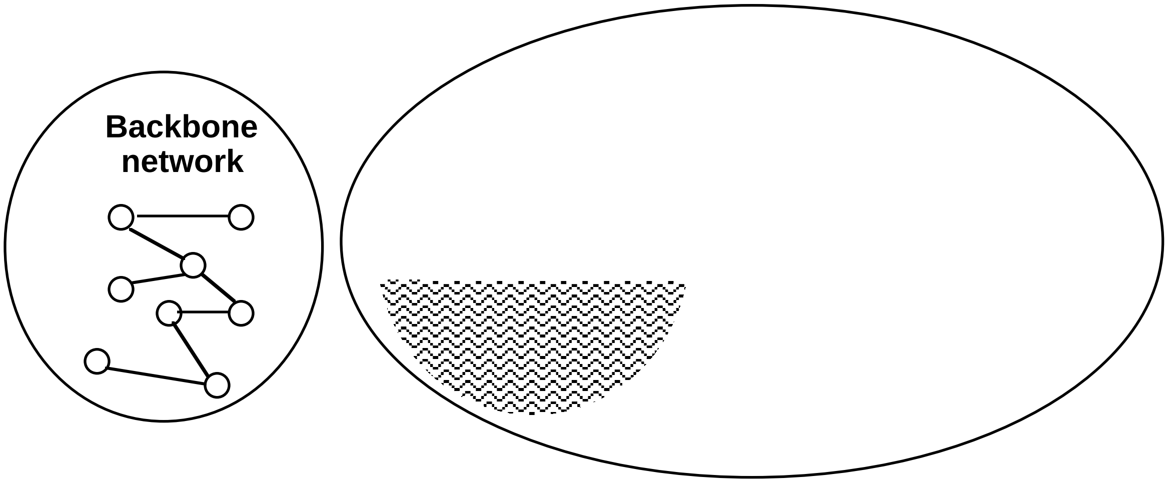
Eytan Modiano and Richard Barry
MIT Lincoln Laboratory
Lexington, MA 02173

Abstract

We describe a WDM-based optical access network architecture for providing broadband Internet services. The architecture uses a passive collection and distribution network and a configurable Feeder network. Unlike earlier papers that concentrate on the physical layer design of the network, we focus on higher layer architectural considerations. In particular we discuss the joint design of the electronic and optical layers including: WDM Medium Access Control protocols; the choice of electronic multiplexing and switching between the IP and WDM layers; joint optical and electronic protection mechanisms; network reconfiguration algorithms that alter the logical topology of the network in response to changes in traffic; and traffic grooming algorithms to minimize the cost of electronic multiplexing. Finally we also discuss the impact of the optical topology on higher layer protocols such as IP routing, TCP flow control and multi-layer switching.

1. Introduction

Over the past decade the growth in the use and capabilities of communication networks has transformed the way we live and work. As we progress further into the information age, the reliance on networking will increase. While today's network traffic is still dominated by voice, there is an increasing demand for data services with broader bandwidth needs and a wide range of Quality of Service (QoS) requirements. These emerging demands offer new challenges and opportunities for the design of *access networks*. The access network refers to the portion of the communication infrastructure responsible for reaching the customer premises. Because of the proximity to the end-user, an access network is quite different from a backbone network and hence offers additional technological and economic challenges. While backbone networks have been able to take advantage of developments in high speed transmission and switching systems to tremendously increase their transmission capacity, access networks have not advanced accordingly, thus the transmission rates available to subscribers are still rather limited.



Earlier efforts on optical access concentrated on the design of PONs for the collection and distribution portion of the access network. The focus on a passive architecture was motivated by the need for low cost, simple maintenance and powering considerations. A number of architectures were developed and systems were demonstrated [FR96, LS89, MC89, FPS89, WL89]. The assumption was that the PON would be used to provide connectivity between the end user premise and the CO, where signals would be carried over an electronic network.

Despite some successful experiments, optical access has failed to materialize, primarily because of the relatively high cost of optical equipment (e.g., lasers). In recent years efforts have been geared toward the development of low cost PON architectures aimed at reducing system costs. Recent progress toward the development of low cost optical technology for use in the local loop have brought the cost of optical local loop to the point that it is becoming competitive with electronic alternatives, especially in new locations where no access infrastructure exists [LU98]. Nonetheless, since in most locations a copper and coax infrastructure already exists, hybrid architectures that utilize the existing copper or coax have emerged and appear to offer a lower cost alternative to an all-fiber solution. A Fiber To The Curb (FTTC) architecture uses a curbside Optical Network Unit (ONU) to serve several subscribers. The ONU is connected to an AN using a PON and connectivity between the ONU and the subscriber is provided over existing twisted pair [PB95]. Similarly, the cable TV industry is adopting a Hybrid Fiber Coax (HFC) architecture where the curb side ONU (typically referred to as the Fiber Node) is connected to the subscribers over existing coax [LU96].

These approaches appear attractive for meeting present and near-term demands of most residential customers. However, they are limited to a transmission capacity of a few tens of Mega-bits-per-second (Mbps). It is widely believed that in the future, applications such as video on demand may require transmission rates in the 100’s of Mbps or even Giga-bits-per-second [RS92]. Furthermore, certain high-end businesses already have needs for these kinds of transmission rates. It is for these applications and users that an optical access solution will be necessary.

Since much has already been written about the design of PONs for access networks, we refer the reader to the literature for more information about the physical layer design of PONs [FR96, PB95, LS89, MC89, FPS89, WL89]. Instead, we focus in this paper on higher layer architectural considerations in the access network. We begin in the next section with the description of an architecture for a WDM-based optical access network.

2. Access network architecture

The proposed architecture consists of a configurable *Feeder* network and passive *Collection and Distribution (C/D)* network, as shown in figure 2. The choice of a passive C/D network is driven by the need for low equipment and maintenance costs at or near the customer premise where equipment is shared among a small number of users. However, in the feeder network the cost of configurable components can be justified because equipment is shared among many more users. Furthermore, in the relatively long distance feeder network, the cost of the fiber can be substantial and hence both electronic and optical multiplexing can be used to make efficient use of the fiber. Lastly, configurability in the feeder network is needed to provide rapid and efficient protection mechanisms, for example through the use of SONET equipment.

As shown in figure 2, the proposed Feeder network is a configurable WDM ring and the C/D network is a WDM PON. Subscribers communicate with an Access Node (AN) over the passive C/D network. At the AN their communication is switched, either optically or electronically, over the feeder network and onto another AN or to the Central Office (CO). As will be discussed in Section 4, the Feeder network uses a combination of optical and electronic techniques, where the electronic layer is aware of the optical layer and vice versa, to make full use of the WDM layer. This is in contrast to existing approaches that separate the optical and electronic layers.

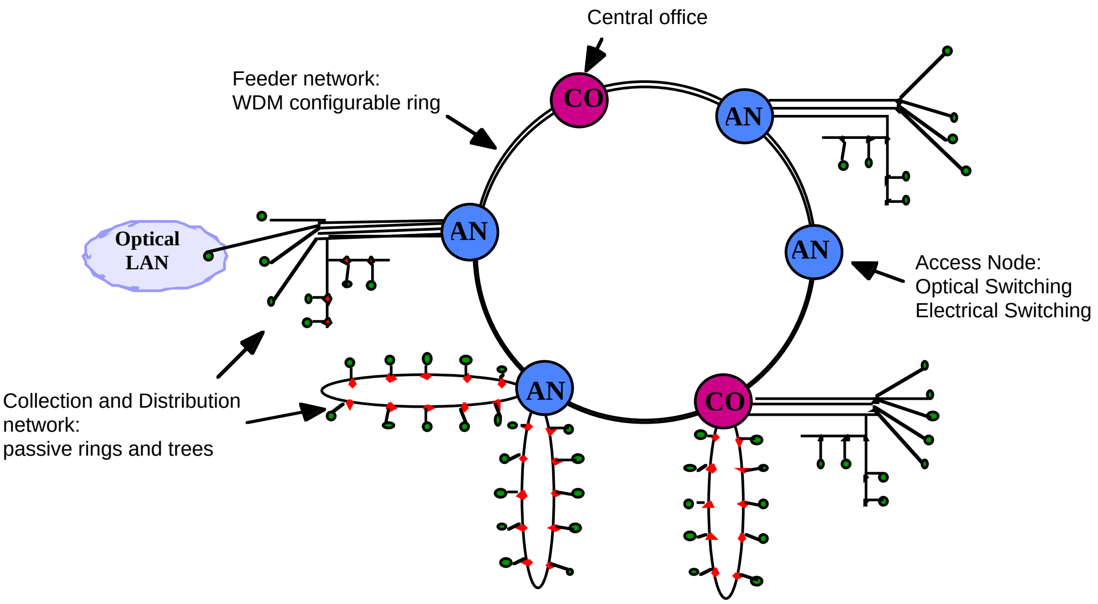


Figure 2. A WDM-based architecture that includes a configurable feeder ring network and passive collection and distribution rings and trees.

For the remainder of this paper we will discuss design considerations for the above architecture. In Section 3 we discuss the C/D architecture and in Section 4 we discuss issues in the design of the Feeder network. Finally, in Section 5 we consider the impact of the WDM architecture on higher layer protocols.

3. C/D Architecture

As stated in the previous section the C/D network is a PON. Many PON architectures have been proposed in the past for use in the access network [LS89, MC89, FPS89, WL89]. Some alternative architectures are shown in figure 3. The simplest architecture would use a dedicated fiber pair for each user. Of course, this approach can be costly because it may require a significant amount of fiber. In addition, a dedicated fiber architecture requires dedicated transceivers at the head-end (located at the access node). An alternative architecture, using a broadcast star at the head-end, would allow the lasers at the head-end to be shared among multiple users, but still requires the same amount of fiber as the dedicated fiber architecture. In order to reduce the cost of fiber, solutions that allow fiber to be shared have been proposed [LS89]. The simplest of which is the double star, with one broadcast star at the head-end and another at a remote location where a cluster of users is located. The most general form of a shared fiber architecture is the tree architecture also shown in figure 3.

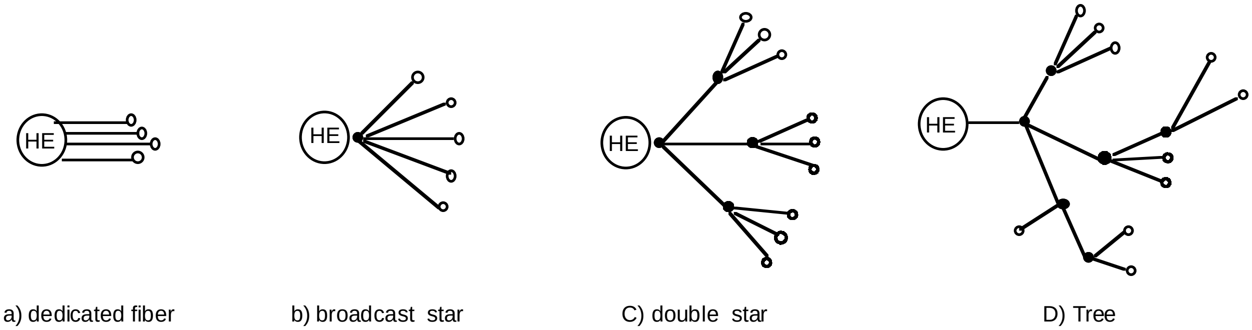


Figure 3. PON architectures.

One shortcoming of all of these shared fiber architectures is the power losses incurred due to the splitting of the fiber. These losses can be particularly significant in older fiber plants where the plant losses are already high. For this reason, many older architectures use dedicated fiber. However, in newer fiber plants, a shared fiber solution is more promising. In addition, recent improvements in passive components significantly reduce the amount of excess loss incurred in the splitters. Consequently, a tree architecture can be used to support more than 100 users, over a 10 KM distance, at 155 Mbps [LS89, RS98]. Furthermore with the additional use of a fiber amplifier at the head-end even higher data rates and many more users can be supported.

An alternative architecture for a WDM PON based on a wavelength router at the head-end was proposed in [IFD95]. With this architecture each user would be communicating on a dedicated wavelength. This architecture has a number of advantages including lower power losses and wavelength isolation. However, it limits the number of users in the PON to the number of wavelengths and it also requires a dedicated fiber pair for each user.

In our proposed architecture, as shown in figure 2, we use passive WDM rings and trees in the collection and distribution network. We choose a tree architecture for both scalability reasons as well as for reducing the amount of fiber needed to connect the nodes (in general, a ring topology may require more fiber miles). The ring architecture is proposed for use in cases where diversity is needed for protection. While these PON architectures have been proposed in the past, the novelty of our approach is in the way in which we propose to use the PON for providing bandwidth on demand services and passive protection, as described below.

3.1 Medium access control (MAC) protocol

Most existing WDM networks employ circuit switching, typically with one connection having exclusive use of an entire wavelength. This approach may not scale to the access network where the number of users may be much larger than the number of wavelengths. Furthermore, this approach is not well suited to bursty data traffic, where even partially aggregated traffic may require very low data rates during periods of inactivity and much higher rates at other times. An access mechanism is needed that provides both scalability and flexibility in provisioning bandwidth.

There are a number of approaches that can be considered for providing scaleable access. One approach is to increase the number of available wavelengths. While present WDM technology provides tens of wavelengths, it is likely that over 100 wavelengths may soon be possible. Nonetheless, even with an increase in the number of wavelengths, it is likely that in certain locations, where fiber is precious, there would not be sufficient capacity to allocate dedicated wavelengths to users. An alternative approach that would allow efficient wavelength sharing is using electronic multiplexing equipment at the fiber merging points (e.g., on poles, pedestals, manholes, etc.). While this approach makes efficient use of the fiber, practical issues regarding the placement of electronic equipment as well as cost considerations and maintenance problems make it infeasible in many circumstances. We therefore propose the use of a Medium Access Control (MAC) protocols for sharing wavelengths in access networks.

Although many WDM MAC protocols for LANs have been proposed and studied in the literature, most of the proposed systems assume a synchronized and slotted system and many require multiple transceivers per node, contributing to their high cost and complexity [MUK92, SKG94]. Furthermore, most of these protocols were designed for a low latency LAN environment and would perform poorly in the access network where propagation delays are relatively high.

We proposed a MAC protocol [MB98a, MB98b] that eliminates the need for slotting and synchronization, uses one tunable transceiver per node, yet results in efficient bandwidth utilization in high latency. The system is based on a simple (and potentially low cost) master/slave scheduler able to schedule transmissions efficiently and overcome the effects of propagation and transceiver tuning delays. As shown in figure 4, a centralized scheduler is located at the access node and responsible for coordinating the transmissions. Users send their transmission requests to the scheduler on a shared control wavelength, λ_C, using a random access protocol, (e.g., Aloha). The scheduler, located at the access node, schedules the requests and informs the users on a separate wavelength, $\lambda_{C'}$, of their turn to transmit. Upon receiving their assignments, users immediately tune to their assigned wavelength and transmit. Hence users do not need to maintain any synchronization or timing information.

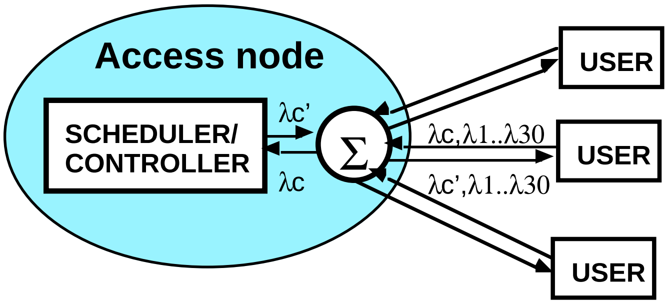


Figure 4. Scheduler based MAC protocol.

The scheduler is able to overcome the effects of propagation delays by measuring the round-trip delay of each user to the access node and using that information to inform users of their turn to transmit in a timely manner. For example consider figure 5. In order for user B's transmission to arrive at the AN at time T, the scheduler must send the assignment to user B at time T-2τ, where τ is user B's propagation delay to the AN. In this way the transmissions of different terminals can be scheduled back-to-back, with little dead time between transmissions. The operation of this MAC protocol, and in particular the ranging process, is somewhat similar to that of the proposed protocol for HFC networks [PG98].

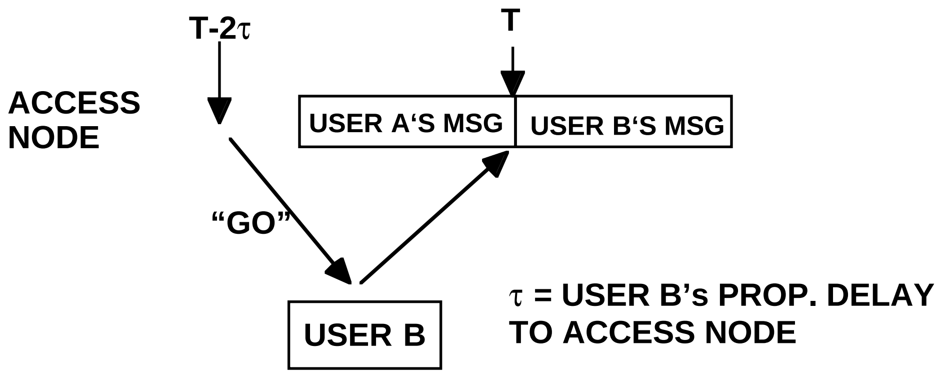


Figure 5. Use of ranging to overcome propagation delays.

An important and novel aspect of this system is the way in which ranging is accomplished. Unlike other systems where terminals need to range themselves to their hubs in order to maintain synchronization [KAM96], here we recognize that it is only the hub that needs to know this range information. Hence ranging can be accomplished in a straightforward manner. The scheduler, ranges each terminal by sending a control message telling the terminal to tune to a particular wavelength and transmit. By measuring the time that it takes the terminal to respond to the request, the scheduler can obtain an estimate of the round trip delay for that terminal. This estimate will also include the tuning time delays. Furthermore the scheduler can repeatedly update this estimate to compensate for fiber inaccuracies. These measurements can also be made by simply monitoring the terminal's response to ordinary scheduling assignments. The significance of this approach is that terminals are not required to implement a ranging function, which simplifies their design.

Other important aspects of this MAC protocol include the control channel access mechanism and the scheduling algorithms used by the scheduler. These issues are addressed in [MB98b]. Scheduling algorithms for transmitting multicast traffic in WDM broadcast-and-select networks are discussed in [MO98]. The performance of the protocol depends primarily on the scheduling algorithm used by the scheduler. In [MB98b] simple scheduling algorithms are described that achieve nearly full utilization. This is a significant improvement over unscheduled WDM MAC protocols that achieve very low channel utilization [MUK92].

3.2 Passive protection and restoration

The passive C/D architecture has the advantage that it is less susceptible to failures because there are no active components in the network. The use of passive components also reduces the maintenance costs of the networks. However, one shortcoming of the passive C/D architecture is that protection or restoration from fiber cuts must be provided optically. Providing protection in the tree C/D network inevitably requires some diversity routing which may eliminate many of the cost benefits of the tree architecture. Therefore the tree architecture is a good choice for users that do not require rapid protection (e.g., homes, small businesses). When rapid protection is more critical, the passive C/D ring can be used. Recently, a number of approaches for providing protection in a passive WDM ring network have been proposed [WW92, GLA96]. A simple example of passive protection in a WDM ring network is shown in figure 6. This passive ring network uses two fiber pairs, one as a primary (P) and one as a backup (B). Within each pair, one fiber is used for upstream traffic to the head end and the other is used for downstream traffic. Consider, first, the downstream operation. Downstream transmission takes place on both the primary and backup fibers. In the event of a fiber cut in either fiber, nodes will receive the transmission on the alternative fiber. If both fibers are cut, at the same location, those nodes before the cut will receive the transmission on the primary fiber and those after the cut will receive the transmission on the backup fiber. Upstream operation is similar, however, it also requires a mechanism to detect which fiber is cut so that it can be switched off (at the head end).

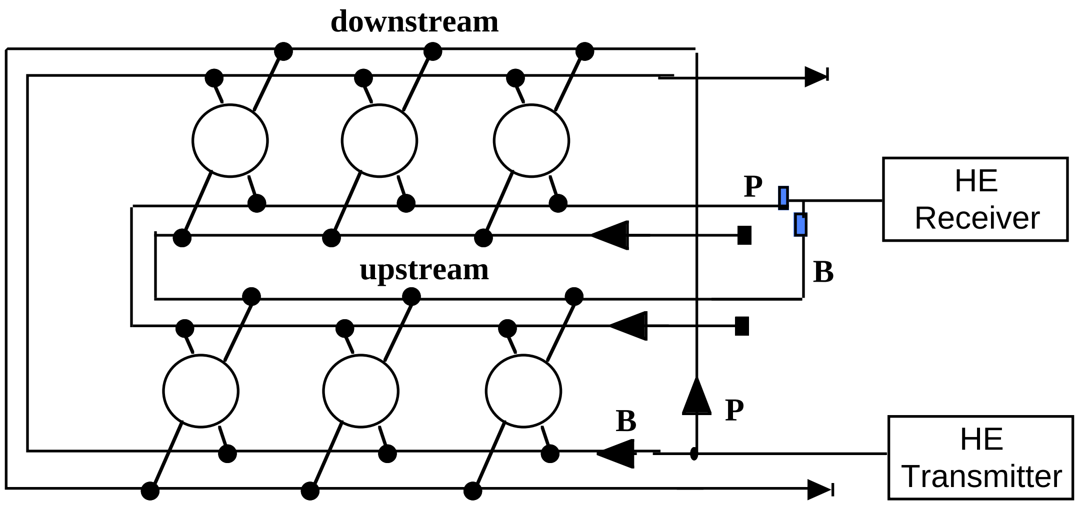


Figure 6. A four fiber passive protection ring.

4. Feeder network architecture

The *Feeder* network has a configurable WDM ring architecture. In order to make efficient use of the fiber in the Feeder network, each node contains a combination of electronic and optical switching equipment. Electronic switches are needed to provide necessary electronic services, such as IP, ATM, SONET or Frame Relay. In addition, electronic switches and multiplexers can also be used to provide efficient statistical multiplexing and fast protection mechanisms. While optical switching is not strictly required in the network, it can significantly simplify the electronic layer by providing optical layer services such as dynamic reconfiguration of the network topology, optical protection and restoration, and traffic grooming. In this section we will discuss issues in the design of both the optical and electronic layers. In particular we will discuss optical layer services that can improve network performance and reduce the cost and complexity of the electronic layer. We will also describe the joint design of the electronic and optical layers of the network so that optical services are used for functions best provided optically and electronic services for functions best provided electronically.

4.1 Use of electronic multiplexing

One important issue in the design of the feeder network is the form of electronic multiplexing offered at the access nodes. Customers may require a variety of electronic services such as SONET, ATM, Frame Relay or IP. One possible solution is to have the network provide all of these services directly at the access node; however, this approach would require a significant amount of electronic equipment at an access node. The other extreme would be to only offer optical services (e.g., lightpaths) at the access nodes and to provide all of the electronic services only at the central office location, which would be accessed by customers optically. This latter option, however, would be inefficient the of use of fiber, since little statistical multiplexing would be done until reaching the central office.

A compromise approach is to provide some electronic services (and hence statistical multiplexing) at the access nodes and back-haul all of the traffic to the Central Office where the other electronic services would be available. For example, some electronic multiplexer (e.g., ATM switch) can be provided at the access node and alternative services could be carried over that switch to the Central Office where those services would be available. Of course, with this option a number of problems might arise such as protocol compatibility issues and inefficiencies due to a multi-layered protocol stack. For example, currently a typical high-end customer's Internet connection involves a multi-layered protocol stack as shown in Figure 7a. Such customers typically gain access to service provider networks via Frame Relay where the IP packets are encapsulated in frames. In the backbone network, the frames are sometimes mapped into ATM cells, which, in turn, are carried over Synchronous Optical Networking (SONET) transport frames.[1] The multitude of layers produces gross bandwidth inefficiencies, adds to the latencies of connections, and inhibits providing quality of service assurances. Worse, the layers are largely unaware of each other causing duplication of network services.

---

[1] In some instances, only one of either Frame Relay or ATM is used.

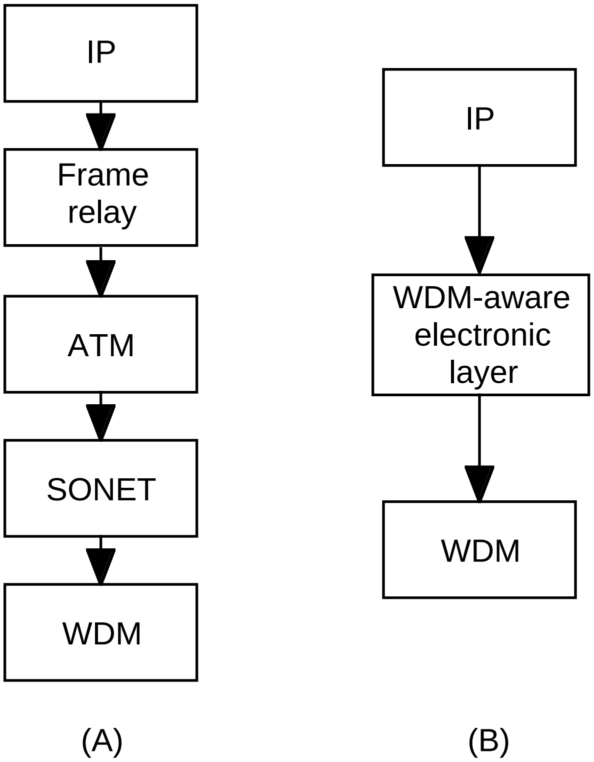


Figure 7. a) Typical Protocol stack. b) Simplified protocol stack.

One solution is the simplified protocol stack shown in Figure 7b, where the IP traffic is carried directly by a simplified electronic layer. Such an arrangement would not only reduce the overhead associated with the different layers but would also allow the electronic layer to be "WDM-aware" and take advantage of network services offered at the optical layer. For the remainder of this section we will describe optical layer services and algorithms that can significantly improve the performance of the network and simplify the design of the electronic layer.

4.2 Topology Reconfiguration of WDM Networks

In WDM networks, the *physical topology* is the one seen by the optical layer. It consists of passive or configurable optical nodes interconnected by fiber. The *virtual topology*, seen by the electronic layer, consists of a set of nodes interconnected by lightpaths. In this way, WDM networks provide a way to interconnect electronic switches with high bandwidth bit pipes without dedicating a fiber pair between each pair of switches. The design of static network topologies has been studied extensively in the past [BFG90, RS96]. However, the configurable nature of WDM also allows the logical topology to be dynamically reconfigured in response to changes in traffic conditions.

WDM networks can reconfigure lightpaths, providing the ability to dynamically optimize the network for changing patterns of externally offered traffic, subject to availability of wavelengths and station equipment. This is achieved by changing the lightpath connectivity between switches, thereby reconfiguring the electronic virtual topology. For example, consider a WDM ring network with one transceiver per node. Shown in Figure 8 are two of the many ways in which the ring can be configured: on the left, nodes are connected in sequence using a single wavelength that is dropped at every node; on the right, nodes are connected in a different order using three wavelengths. Notice that the connectivity on the left does not allow both call 1 and call 2 to be admitted simultaneously, while the one on the right allows both.

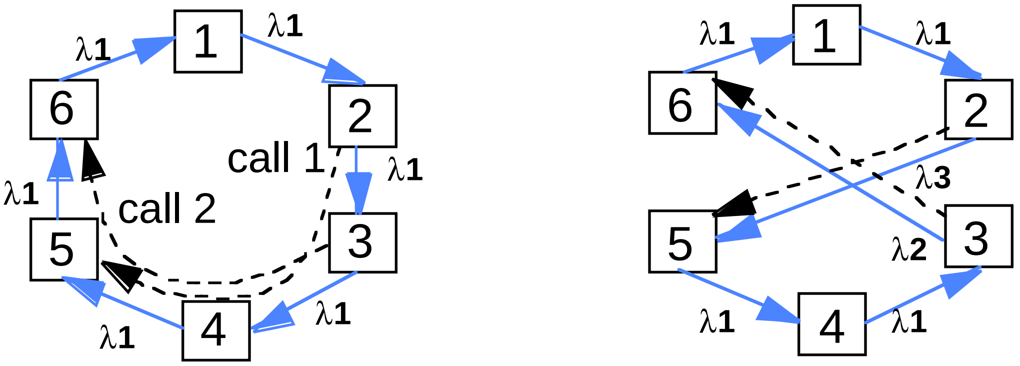


Figure 8. Using WDM to reconfigure the electronic topology.

Reconfiguration effectively increases the overall system capacity by allocating bandwidth only where it is needed. However, there is an overhead associated with reconfiguration in that signaling messages need to be sent between switches to coordinate the topology change. In addition, reconfiguration may impact existing connections and these connections may need to be rerouted in the virtual topology. However, it may be possible to design a reconfiguration strategy that does not allow the rearranging of existing calls. Such a strategy, may provide fewer benefits, but would be simpler to implement and eliminate many coordination and network management problems.

Preliminary studies on reconfiguration of a WDM ring show significant promise [CH97]. Plotted in Figure 9 is the call blocking probability vs. call arrival rate for a WDM ring with 20 nodes and a granularity of 1 call per wavelength. As can be seen from the figure, a reconfigurable WDM topology can support six times the traffic of a fixed WDM topology for the same blocking probability. Similar results were found for various ring and star networks. However, topology reconfiguration is not simple to implement in a network because changes in the network topology can complicate network management and routing. Also, care must be taken to manage reconfiguration in such a way that calls are not dropped or disrupted. Despite the promising results discussed above, the benefits of reconfiguration remain an open research issue.

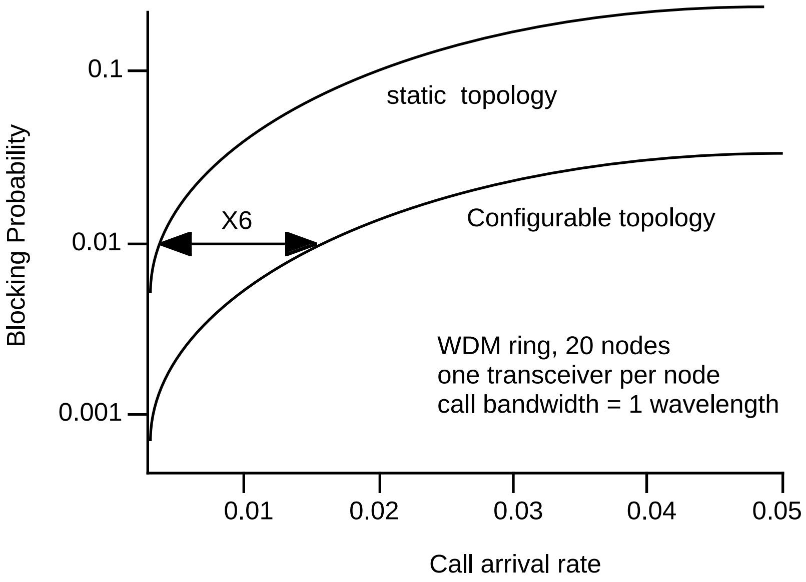


Figure 9. Performance of reconfiguration in a WDM ring network.

4.3 Traffic Grooming

Unless users require full wavelength connections, sub-wavelength capacity connections need to be allocated. This can be accomplished through the use of electronic multiplexing equipment that can aggregate low rate calls on to a higher rate channel (e.g., SONET multiplexers). However, if calls are indiscriminately multiplexed on to wavelengths then each wavelength entering or leaving a node will need to be converted to electronics to make drop/forwarding decisions. Alternatively, if calls are cleverly *groomed* onto wavelengths, then the number of wavelengths that need to be processed at each node can be significantly reduced.

Consider, for example, the two SONET ring networks shown in Figures 10a and 10b. Both networks are used to provide point-to-point OC-3 circuits between each pair of nodes; SONET Add/Drop Multiplexers (ADMs) are used to combine up to 16 OC-3 circuits into a single OC-48 that is carried on a wavelength. If a wavelength carries traffic that originates or terminates at a particular node, then that wavelength must be dropped at that node and terminated in a SONET ADM. The network of figure 10a drops each wavelength at every node, thus requiring many ADMs. In order to reduce the number of ADMs used, it is better to groom traffic such that all of the traffic to and from a node is carried on the minimum number of wavelengths (and not dispersed among the different OC-48's), as shown in figure 10b. Notice that this is not the same as minimizing the total number of wavelengths used in the network, a subject that will be discussed in the next section.

In this context, the problem is to design traffic grooming algorithms to minimize electronic costs while simultaneously making efficient use of wavelengths. Optimal traffic grooming has been shown to be an NP-complete problem [MC98]. However, studies of traffic grooming on SONET ring networks with uniform all-to-all traffic show that grooming can result in a significant reduction in the number of required ADMs [SGS98, MC98]. For example, we show in figure 11 the number of ADMs required in a unidirectional SONET ring network using an algorithm developed in [MC98]. As can be seen from the figure, the algorithm can obtain a significant reduction in SONET ADMs. Also, in many cases, the solution that uses the minimum number of ADMs also uses the minimum number of wavelengths. Hence, a good grooming algorithm has the potential of minimizing network cost both in terms of efficient use of fiber and efficient use of electronics.

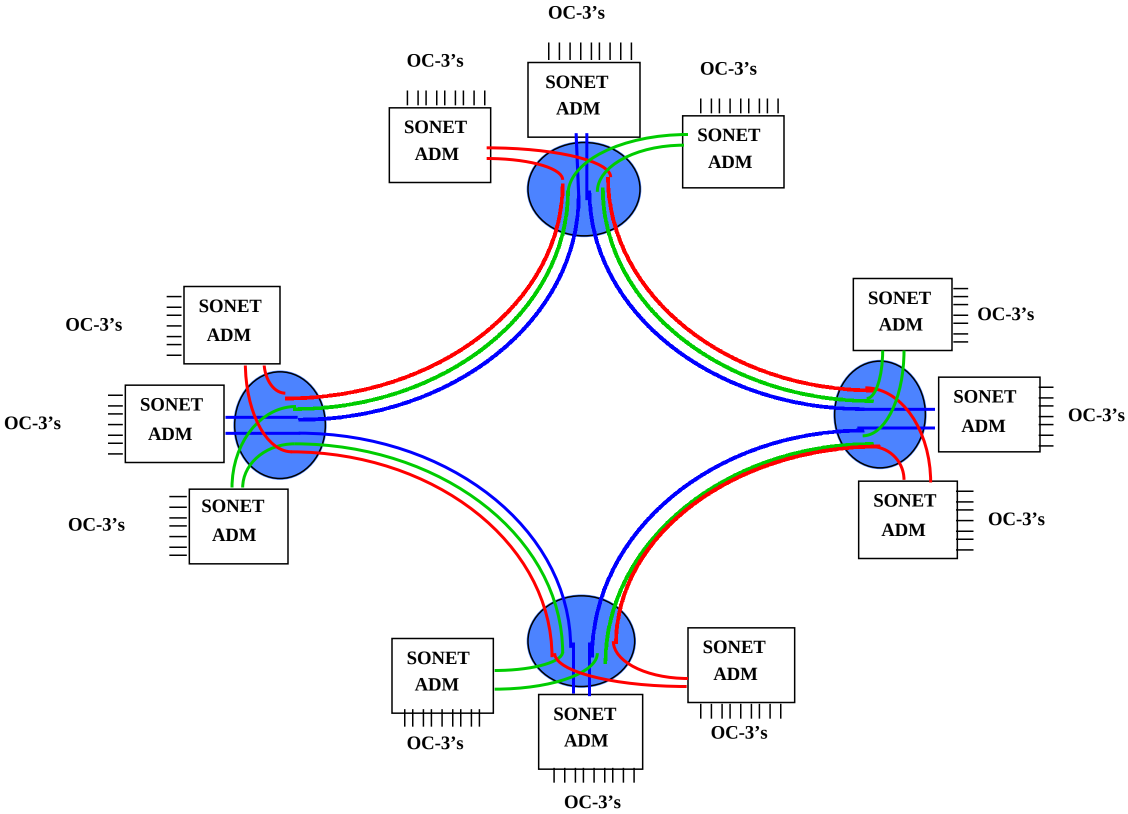
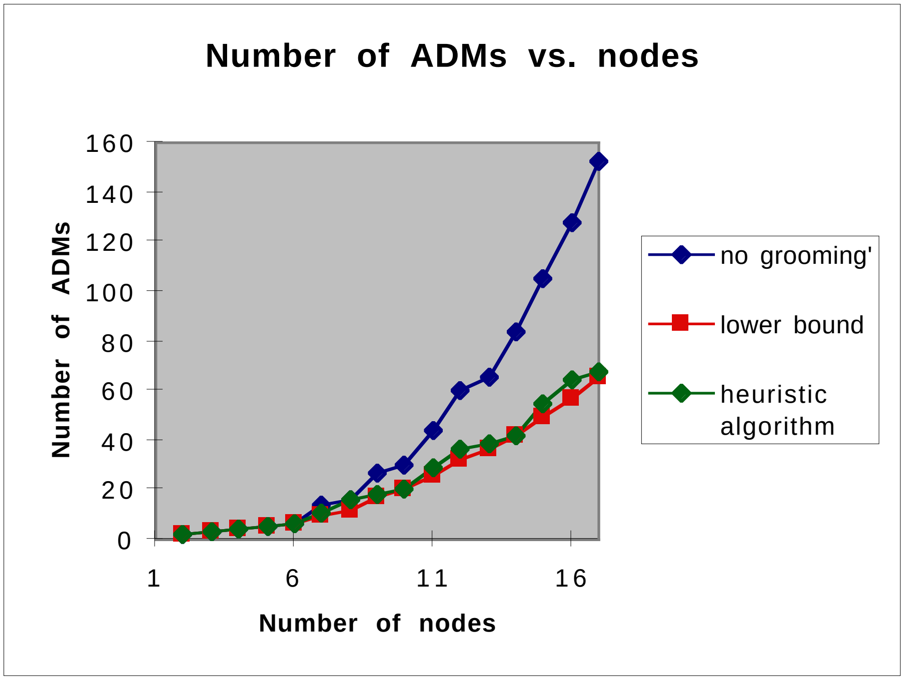

OC-3's
OC-3's
SONET ADM
OC-3's
SONET ADM
SONET ADM
OC-3's
SONET ADM
OC-3's
SONET ADM
OC-3's
SONET ADM
OC-3's
SONET ADM
OC-3's
SONET ADM
SONET ADM
SONET ADM
SONET ADM
OC-3's
OC-3's
SONET ADM
OC-3's

Figure 10a. A SONET/WDM ring network

Figure 10b. The benefits of traffic grooming in a SONET/WDM ring.



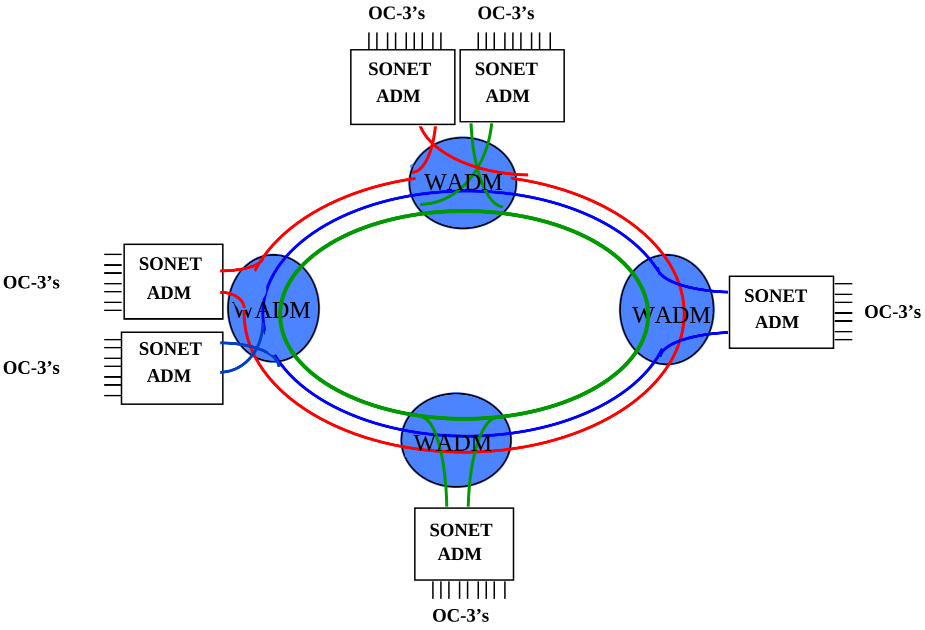


Figure 11. Performance of grooming in a WDM/SONET ring network

4.4 Wavelength assignment algorithms

The nodes in the Feeder ring network use configurable wavelength switches that can switch the data optically from any input fiber to any output fiber. In order to reduce the cost of these switches wavelength conversion is not employed. That is, data is switched from any input fiber to any output fiber on the same wavelength. The elimination of wavelength conversion may lead to reduced network efficiency because the same wavelength must be available on each link of a route in order for a call to be established. Therefore, there may be a route with a wavelength available on each of its links, but unless they are all the same wavelength, the call will be rejected. With wavelength conversion, on the other hand, the call could be accepted.

While the cost of wavelength changers is formidable, a good wavelength assignment algorithm can go a long way in increasing the efficiency of the network. The wavelength assignment algorithm is responsible for selecting a suitable wavelength among the many possible choices for establishing the call. For example, the three calls in figure 12 can be established using 3 wavelengths as shown on the left or just two wavelengths as shown on the right. A number of wavelength assignment schemes have been proposed [SB97], and the subject remains an active area of research.

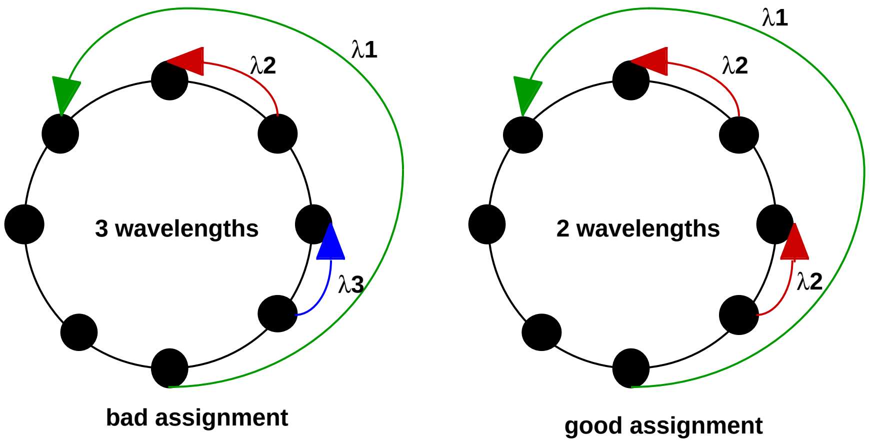


Figure 12. Two possible wavelength assignments for three calls on a ring.

Figure 13 compares the performance of some proposed wavelength assignment algorithms. The simplest algorithm is to randomly select a wavelength, and if it is available along the path the call is accepted and otherwise it is blocked. Clearly such an algorithm would be very inefficient, and , as can be seen from the figure, the random algorithm results in the highest blocking probability. A first-fit heuristic assigns the first available (i.e., lowest index number) wavelength that can accommodate the call. The most used heuristic assigns the wavelength that is used on the most number of fibers in the network and lastly, the max-sum algorithm assigns the wavelength that maximizes the number of paths that can be supported in the network after the wavelength has been assigned. Also shown in the figure is the blocking probability that results when wavelength changers are used. This represents an upper-bound on the performance of any wavelength assignment algorithm. The significance of the figure is that a good wavelength assignment algorithm can result in a blocking probability that is nearly as low as if wavelength changers are employed. Hence, a significant reduction in network costs can be obtained by using a good wavelength assignment algorithm.

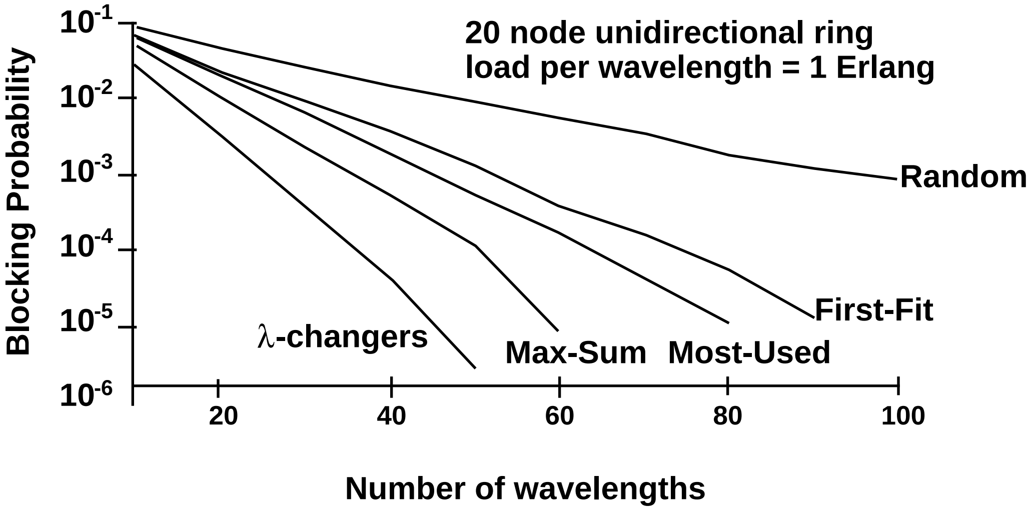


Figure 13. Performance of wavelength assignment algorithms in a ring network.

4.5 Protection and Restoration

Various failures can occur that disrupt network services, such as fiber cuts, line card and switch failures, and software failures [WU92]. Protection and restoration are two methods networks use to recover from these failures. Protection refers to hardware-based, pre-planned, fast failure recovery; restoration refers to software-based, dynamic, slow recovery. Protection is generally limited to simple topologies like rings or the interconnection of rings; restoration works on general mesh networks and is typically more bandwidth efficient. Recently, fast protection mechanisms at the optical layer have been proposed for general mesh networks [ES96, FMB97, FMB98], and for ring networks [GR97, MB96].

Failure recovery must be done at the electronic layers in order to recover from line card or electronic switch failures. Electronic recovery mechanisms, e.g., as is done in SONET, can also be used to protect against failures at the optical layer such as a fiber cut or a malfunctioning optical switch. Therefore, at first sight, it might appear that optical layer recovery is not needed.

There are, however, three significant advantages to doing protection and restoration optically. First, optical layer recovery can protect electronic services which do not have built-in recovery mechanisms or whose recovery mechanisms are slow (e.g., IP). Second, in many cases, optical layer recovery is more natural and provides enhanced reliability. For instance, consider the case of 32 SONET rings being supported over a WDM ring network with 32 wavelengths. Without optical protection, each of the 32 SONET rings would need to individually recover from a single fiber cut, e.g., by loop-back in a SONET bi-directional ring network. On the other hand, the fiber cut can be optically restored with a simple 2x2 optical switch, thereby simultaneously restoring service to many electronic connections. A simple example of optical loop-back protection using a 2x2 switch is given below. Protection at the optical layer has the added advantage that the failure is transparent to SONET, allowing each SONET ring to individually respond to additional failures such as a line card failure. If protection were only performed electronically, there would be no guarantees that the SONET ring would be resilient to both a fiber cut and a line card failure. Thus protection at the optical layer provides increased reliability to failures at the electrical layers. Third, optical recovery allows the construction of arbitrary virtual topologies resilient to fiber failures. For instance, consider IP routers connected in a virtual ring topology over a WDM ring. Traffic may dictate that the IP routers be connected such that more than one IP link travels over the same fiber and hence a single fiber cut will disrupt many IP links. Since rings are in general only resilient to a single link failure, a single fiber cut of the optical ring may disconnect the electronic ring. Optical layer restoration solves this problem by restoring the fiber directly.

For a simple example of optical loop-back protection, consider the two fiber bi-directional ring shown in figure 14. On each ring half of the wavelengths are used for working traffic and the other half are reserved for protection against a cut in the fiber on the other ring. In the event of a fiber cut, the wavelengths from the cut fiber can be switched onto the uncut fiber, using a two-by-two switch at the node before the fiber cut. They can then be looped back to bypass the cut fiber and rejoin their original ring using another switch at the node immediately following the fiber cut.

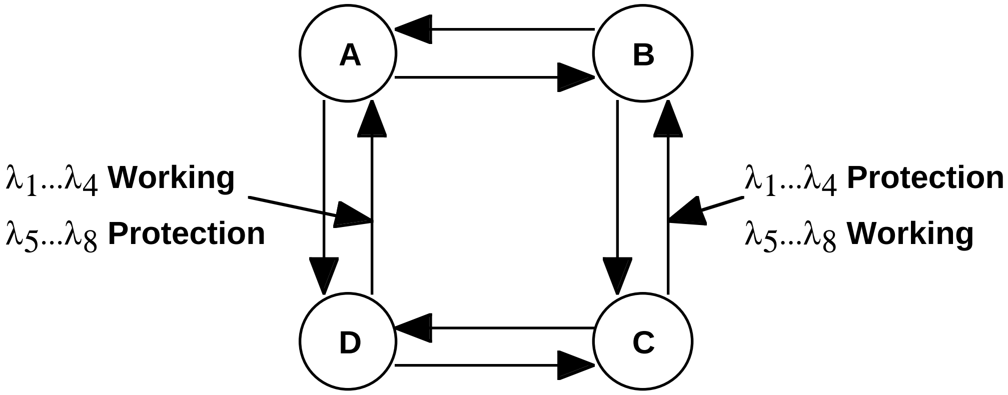


Figure 14. Protection in WDM ring networks.

However, there are problems providing restoration at both the optical and electronic layers if the layers work independently of each other. For instance, if care is not taken, restoration will be duplicated at both the optical and electronic layers leading to a 75% loss in efficiency (assuming 50% efficiency for each layer of protection). In addition, differing time scales may lead to race conditions and topology flapping. Also, in the case of a SONET network, optical protection must somehow be completed before SONET starts its protection process. This may be difficult as SONET starts its protection process as soon as loss of power is detected. Hence, care must be taken to coordinate the protection mechanisms at the electronic and optical layers.

5. High layer protocol issues

The algorithms discussed in the previous section are network layer algorithms, working independently of the higher layer protocols and applications that are running above them. However, these algorithms can, in fact, significantly impact the performance at the higher layers. For example, rapid topology reconfiguration could lead to IP routing table instabilities. Furthermore, many existing higher layer protocols were not designed for the huge capacity and configurability offered by WDM and hence are incapable of taking advantage of the benefits of a WDM physical layer. In this section we focus on internet services and briefly discuss new concepts that can help internet protocols gain greater benefits from the WDM layer.

5.1 Multi-layer Switching

One of the main bottlenecks in the present Internet is routing at the IP layer. Several methods have been proposed to alleviate this bottleneck by switching long duration flows at lower layers [NAG97, NLM96, REK97, VIS97]. Tag switching uses routing protocols to predefine routes within the network; tags are then used to quickly assign flows to these routes. IP switching dynamically sets up layer-2 (e.g., ATM) virtual circuits for those connections that are perceived to be long.

This concept of lower-layer switching can be extended to switching long duration electronic flows at the optical layer. That is, a new lightpath can be established for long duration, and perhaps high bandwidth, flows across the network. To achieve scalability, an optical flow switching protocol may need to aggregate flows with similar characteristics in order to switch them together. This concept gives rise to a multi-layer switching approach where long duration sessions are switched at the ATM layer, and even longer duration and higher bandwidth flows are switched optically. While it appears that such a multi-layer switching approach can reduce bottlenecks and processing delays in networks, many issues in the design of such a protocol remain to be resolved. For example, the aggregation of different sessions, with differing traffic characteristics and service requirements, while providing service guarantees remains an open problem for research. Also, the criteria to be used to determine which flows are to be switched electronically and which optically requires investigation.

5.2 TCP Flow Control

Most Internet data applications today use the Transmission Control Protocol (TCP) as the transport layer protocol. TCP implements a window based congestion control mechanisms where users are allowed to transmit up to a window of packets at a time. TCP attempts to prevent congestion from occurring in the network by regulating the window size given to a connection. Connections are initially given a window of one packet and are allowed to slowly increase their window up to some maximum value through a process called *slow start* [LAK97**,** BRA95]. While this *slow start* mechanism may be appropriate for low rate networks, it will clearly prevent TCP users from taking advantage of the WDM capacity. For example, if a new wavelength becomes momentarily available for use by a connection, that wavelength may no longer be available by the time slow start allows the connection to sufficiently increase its transmission rate. In this case, not only is the connection prevented from effectively using the wavelength, but additional congestion may be caused. Furthermore, current TCP implementations limit connections to very small windows which significantly limit the maximum transmission rate available to a connection [LAK97].

Clearly, mechanisms are needed to allow TCP connections to take advantage of the capacity and configurability of WDM. Currently, TCP standards are being developed that allow much larger window sizes [JAC92]. In addition, mechanisms are needed to allow connections to change their transmission rates more rapidly so that they can dynamically respond to changes in network connectivity. For example, it may be feasible to use a new field in the packet header that allows routers to explicitly notify connections of their allowed transmission rates.

6. Conclusion

This paper discusses architectural considerations in the design of an optical access network. In contrast to previous efforts that were focused primarily on the physical layer design, the focus of this paper is on architectural considerations at higher layers. Our proposed architecture is aimed at providing very high bandwidth connectivity that can only be achieved through the use of optics. The architecture uses an optical passive Collection and Distribution network with a Medium Access Control protocol to allow bandwidth sharing, and a configurable Feeder network that uses a combination of optics and electronics.

The goal of this architecture is to jointly design the optical and electronic layers so that the best features of each technology can be employed. Toward that end, a number of optical layer services and algorithms that are critical to the design of a WDM-based optical access network are described. These include a Medium Access Control protocol for providing bandwidth on demand in WDM PONs; optical protection and restoration algorithms; topology reconfiguration algorithms that increase the network capacity by reconfiguring the virtual topology of the network in response to changes in traffic conditions; traffic grooming algorithms for reducing the amount of electronic equipment needed in the network; and wavelength assignment algorithms for making efficient use of wavelengths. Lastly, we discuss the need for higher layer protocols that can take advantage of the huge capacity and configurability of the WDM architecture.

References

[BFG90] J. Bannister, L. Fratta, and M. Gerla, “Topological Design of WDM Networks,” Proc. Infocom ‘90.

[BRA95] L.S. Brakmo, and L. Peterson, “TCP Vegas: End-to-End Congestion Avoidance on a Global Internet,” IEEE JSAC, October, 1995.

[CH97] Angela Chiu, “Reconfiguration in WDM Networks,” Unpublished manuscript.

[ES96] G. Ellinas and T.E. Stern, "Automatic Protection Switching for Link Failures in Optical Networks with Bi-directional Links," Globecom '96.

[FMB97] S.G. Finn, M. Medard, and R.A. Barry, ``A Novel Approach to Automatic Protection Switching Using Trees", ICC ‘97.

[FMB98] S.G. Finn, M. Medard, and R.A. Barry, "A New Algorithm for Bi-directional Link Self Healing for Arbitrary Redundant Networks," OFC ‘98.

[FPS89] D. W. Faulkner, et. al., “Optical Networks for Local Loop Applications,” Journal of Lightwave Technology, November, 1989.

[FR94] N.J. Frigo, et. al., “A Wavelength Division Multiplexed passive optical network with cost shared components,” IEEE Photonic Technology Letters, June, 1994.

[FR96] N. J. Frigo, “ Local Access Optical Networks,” IEEE Network, Dec., 1996.

[GLA96] B. Glance, et. al., “Novel optically restorable WDM ring network,” OFC ‘96.

[GR97] O. Gerstel and R. Ramaswami, "Multiwavelength Ring Architectures and Protection Schemes," Photonic Networks, Advances in Optical Communications, pp. 42-51, Springer-Verlag, New York, 1997.

[IFD95] P.P. Iannone, N.J. Frigo, and T.E. Darcie, “WDM passive optical network architecture with bidirectional optical spectral slicing,” OFC ‘95.

[JAC92] V. Jacobson, R. Braden and D. Dorman, “TCP Extensions for High Performance,” Internet Engineering Task Force RFC-1323, May, 1992

[KAM96] I.P. Kaminow, et. al., “A Wideband All-Optical WDM Network”, IEEE JSAC, June, 1996.

[LAK97] T.V. Lakshman and U. Madhow, “The Performance of TCP/IP for Networks with High Bandwidth-Delay Products and Random Loss,” IEEE/ACM Transactions on Networking, 1997.

[LS89] Y. M. Lin and D. R. Spears, “Passive Optical Subscriber Loops with Multiaccess,” Journal of Lightwave Technology, November, 1989.

[LU96] X. Lu, et. al., “Mini-Fiber Node Hybrid Fiber/Coax Networks for Two way Broadband Access,” OFC ‘96.

[LU98] Angelo Luvison, “The architecture of full service access network,” OFC ‘98, San Jose, CA, March, 1998.

[MB96] J. Manchester and P. Bonenfant, "Fiber Optic Network Survivability: SONET/Optical Protection Layer Interworking," National Fiber Optic Engineers Conference 1996.

[MB98a] Eytan Modiano, Richard Barry and Eric Swanson , “A Novel Architecture and Medium Access Control (MAC) Protocol for WDM Networks,” OFC ‘98, San Jose, CA, February, 1998.

[MB98b] Eytan Modiano, Richard Barry and Eric Swanson , “A Novel Architecture and Medium Access Control Protocol for WDM Networks,” TOPs Vol. 20, Optical Networks and their applications, Optical Society of America, 1998.

[MC98] Eytan Modiano and Angela Chiu, “Traffic Grooming Algorithms for Minimizing Electronic Costs in Unidirectional SONET/WDM Ring Networks,” CISS‘98, Princeton, NJ, March, 1998.

[MC89] I. M. McGregor, “Implementation of a TDM Passive Optical Network for Subscriber Loop Applications,” IEEE JLT, Nov., 1989.

[MO98] Eytan Modiano, “Unscheduled Multicasts in WDM Broadcast-and-Select Networks,” Infocom ‘98, San Francisco, CA, March 1998.

[MUK92] B. Mukherjee, “WDM-Based Local Lightwave Networks Part I ,” IEEE Network, May, 1992.

[NAG97] K. Nagami, *et. al.,* ``Toshiba's Flow Attribute Notification Protocol Spec." IETF RFC 2129, April 1997.

[NS97] P. Narvaez and K.-Y. Siu, ``An Acknowledgment Bucket Scheme to Regulate TCP Flow over ATM," Proceedings of IEEE Globecom, Nov. 1997.

[NLM96] P. Newman, T. Lyon, and G. Minshall, ``Flow Labeled IP: A Connection Approach to ATM." IEEE Infocom, 1996.

[PB95] W. Puh and G. Boyer, “Broadband Access: Comapring alternatives,” IEEE Communication Magazine, August, 1995.

[REK97] Y. Rekhter, *et. al.,* ``Cisco Systems' Tag Switching Architecture Overview," IETF RFC 2105, Feb. 1997.

[PG98] S. Perkins and A. Gatherer, “Two-way Broadband CATV-HFC Networks: State of the art and Future trends,” this issue.

[RS92] M.N. Ransom and D.R. Spears, “Applications of public Giga-bit networks,” IEEE Network, March, 1992.

[RS96] R. Ramaswami and K. Sivarajan, “Design of logical Topologies for Wavelength Routed Optical Networks,” IEEE JSAC, June, 1996.

[RS98] R. Ramaswami and K. Sivarajan, *Optical Networks*, Morgan Kaufmann, San Francisco, CA, 1998.

[SB97] S. Subramanian and R. Barry, “Wavelength Assignment in Fixed Routing WDM Networks,”, Proc. of ICC ‘97.

[SGS98] Simmons, J.M., Goldstein, E.L., Saleh, A.A.M, “On The Value of Wavelength- Add/Drop in WDM Rings With Uniform Traffic,” OFC’98, San Jose, CA, February, 1998.

[SKG94] G. N. M. Sudhakar, M. Kavehrad, N.D. Georganas, “Access Protocols for Passive Optical Star Networks,” Computer Networks and ISDN Systems, pp. 913-930, 1994.

[VIS97] A. Viswanathan, et al., ``ARIS: Aggregate Route-Based IP Switching." IETF Internet Draft, March 1997.

[WL89] S.S. Wagner and H. L. Lemberg, “Technology and System Issues for a WDM-Based Fiber Look Architecture,” IEEE JLT, Nov., 1989.

[Wu92] T.-H. Wu, Fiber Network Service Survivability, Artech House, 1992.

[WW92] T.-H. Wu, W.I. Way, ``A Novel Passive Protected SONET Bidirectional Self-Healing Ring Architecture", IEEE Journal of Lightwave Technology, vol. 10, no. 9, September 1992.