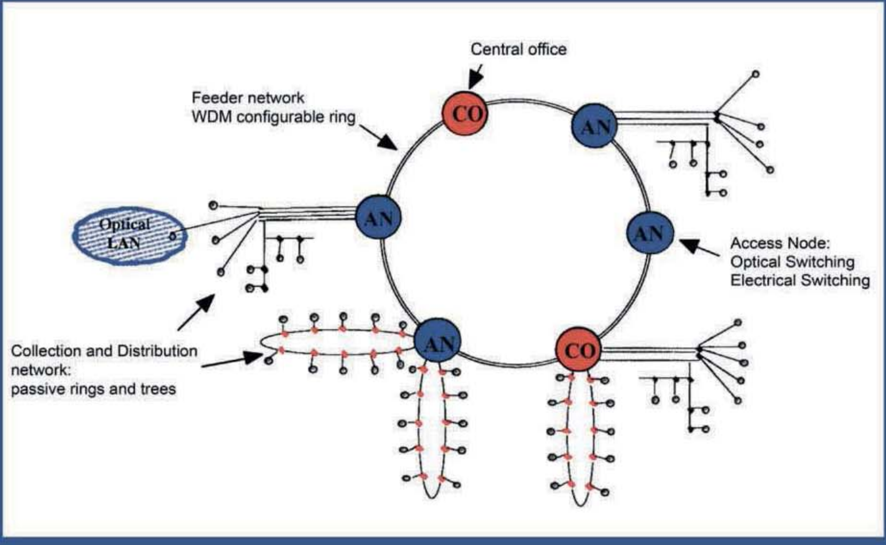
OPTICAL WDM NETWORKS

Principles and Practice



edited by

Krishna M. Sivalingam

Suresh Subramaniam

VI

MISCELLANEOUS TOPICS

Chapter 13

Optical Access Networks for the Next Generation Internet

Eytan Modiano
Massachusetts Institute of Technology

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: 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; traffic grooming algorithms to minimize the cost of electronic multiplexing; and WDM Medium Access Control protocols.

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.

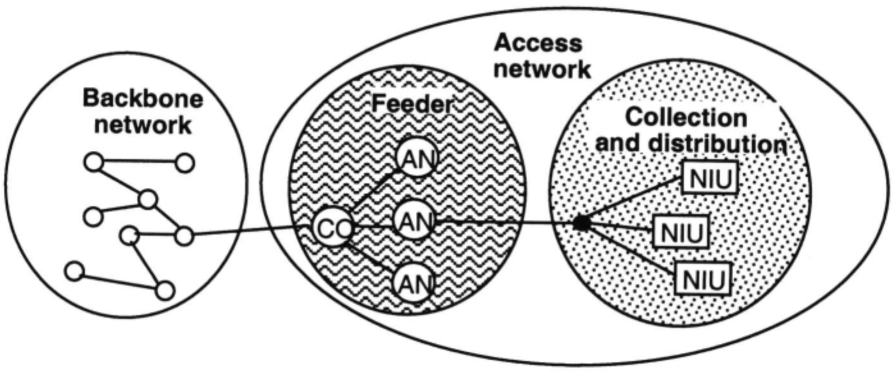


Figure 1. Generic access network architecture

This paper describes an architecture for a WDM based optical access network and examines critical issues in its design. Our discussion will follow the generic access network architecture shown in Figure 1. As shown in the figure the access network consists of two parts: *the collection and distribution* network (C/D) responsible for connectivity between the Network Interface Unit (NIU) at the customer premise, and a remote Access Node (AN), and the *Feeder* network that connects the various ANs to a Central Office (CO). Central offices are connected via a backbone network that is generally not considered part of the access network. Previous efforts in the design of optical access networks focused, almost exclusively, on the design of Passive Optical Networks (PONs) for the *C/D* part of the access network. The PON architecture was motivated by the need for low cost, simple maintenance and powering considerations. Hence, most of that effort was concentrated on the physical design of the PON. A number of PON architectures were developed and systems were demonstrated [FR96, LS89, MC89, FPS89, WL89]. However, little attention has been paid to the overall access network architecture. The focus of this article is on an overall access network architecture that includes the C/D network as well as the *Feeder* network and discusses architectural considerations in their design, paying particular attention to network layer issues.

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 curb-side 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 (Gbps) [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. The challenge for optical access networks is to provide a cost effective interface between end-users and very high capacity WDM-based backbone networks.

This challenge is the focus of the DARPA-sponsored Next Generation Internet (NGI) ONRAMP consortium, a pre-competitive consortium including MIT, JDS Fitel, Nortel Networks, Cabletron systems, AT&T and MIT Lincoln Laboratory. The goal of the consortium is to design and build an optical access network which exploits WDM and other emerging technologies to support next generation Internet services. At least two types of services will be offered: an all-optical service, which establishes an uninterrupted lightpath from source to destination, and an IP service, which connects the source and destination via IP routers. In support of IP and other types of traffic, the network will feature optical flow switching, MAC protocols to share wavelengths among bursty users, dynamic provisioning and reconfiguration, automatic protection switching, and a robust and responsive network control and management system. The ONRAMP test-bed will be described in more detail in section IV of this paper.

2. ACCESS NETWORK SERVICES, SIZE AND CAPACITY

Before designing a high-speed optical access network, it is necessary to know what services the network will provide, what geographical span the access network will cover and what traffic capacity it will support. We start by discussing basic access network services and applications. While it may be difficult, at present, to foresee specific applications that require giga-bit per second transmission rates, it is important to note that past forecasts of new applications have not been terribly accurate. In fact, the World-Wide-Web application was not in anyone's predictions much before its appearance. Hence, we will not attempt to foretell any specific future applications but rather try to provide an abstraction of the type of services that can be offered in the future. We believe that future access networks should be able to provide the following two basic services:

1. *Conventional electronic network services*: - e.g., SONET, ATM, Frame Relay, and IP services.
2. *Switched optical lightpaths* – e.g., point to point optical connections that take up a full wavelength. These lightpaths connections can be used to support a number of applications, such as:
   a) *Large point-to-point circuit-switched trunks on demand* – e.g., OC-48, OC-192 and above to deal with stream traffics but with setup time on the order of tens of milli-seconds.
   b) *Optical flow switching* - for bursty, unscheduled large file transfer (100 Mbyte to 10 Gbyte) at high access rates (> 2 Gbps). This service can be provided by dynamically setting up an end-to-end optical lightpath for the duration of the transaction.
   c) *Analog services* - narrow and broadband analog services with high amplitude, phase and timing fidelity preserving features.

Access networks must be designed for urban, suburban and rural locations. Depending on the type of location that the access network is being designed for, it will a have different size and capacity. For example, an access network for the New York city can be as small as a few miles in diameter, while an access network in Minnesota is likely to have 100's of miles in diameter. These differences will inevitably impact the physical design of a network. In this paper we will consider an access network that is designed to support a metropolitan area with approximately 500,000 people in population and spanning between 100 and 1000 square miles. Clearly, very dense metropolitan areas can be supported by multiple access networks. Using simple arithmetic one can see that approximately 500 such access networks may exist in the entire U.S. While such an access network covers an area with 500,000 people in population, only a relatively small number of users will require the services of a high-speed optical access network. We assume, that at least initially, only high-end businesses will require optical access and that each access network will support somewhere between 100 and 1000 businesses. We will show later, that these numbers have a significant impact on the design of the network.

Another important design criteria is the traffic capacity that such an access network must support. While it is rather difficult to predict future traffic requirements, the U.S. has been experiencing significant growth in internet traffic over the past few years. Furthermore, since optical access networks will not be needed unless such growth materializes, it is reasonable to design a future optical access network that is based on significant traffic growth projections. While today's backbone network traffic in the entire U.S. is less than 1 tera-bits per second, traffic increase at a factor of 2 per year will result in backbone traffic on the order of 1000 tera-bits per second within 10 years. Even, if a factor of two rate of increase is not sustained, it is reasonable to assume that in the foreseeable future a backbone capacity of this order will be required. With 500 access networks in the U.S., each access network will have to support approximately a tera-bit of capacity.

We will focus on the design of an access network that can accommodate these services, geographical span and capacity. In section III we will discuss the physical network architecture and in section IV we will discuss the ONRAMP test-bed.

3. 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 restoration.

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). 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.

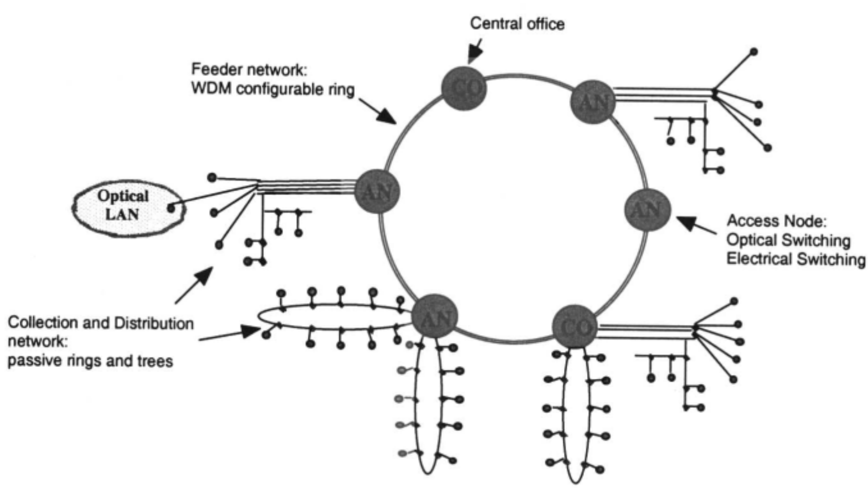


Figure 2. A WDM-based architecture that includes a configurable feeder ring network and passive C/D rings and trees

3.1 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.

3.1.1 Use of electronic multiplexing

One important issue in the design of the feeder network is the form of electronic multiplexing offered at the ANs. 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 AN; however, this approach would require a significant amount of electronic equipment at an AN. The other extreme would be to only offer optical services (e.g., lightpaths) at the ANs 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 inefficiently use the 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 ANs 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.,IP router) can be provided at the AN and alternative services could be carried over IP 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 3a. 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.

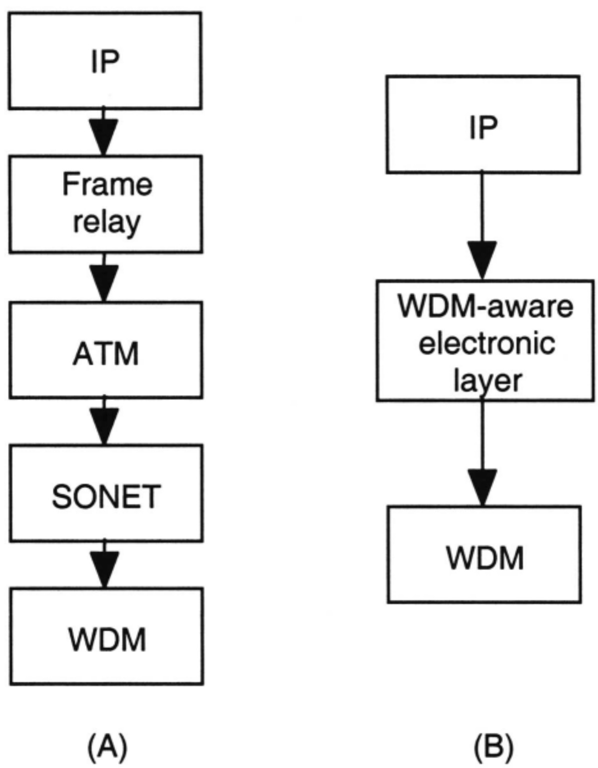


Figure 3. a) Typical protocol stack, b) Simplified protocol stack

One solution is the simplified protocol stack shown in Figure 3b, 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.

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

3.1.2 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. However, 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.

For a simple example of optical loop-back protection, consider the two fiber bi-directional ring shown in Figure 4. 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.

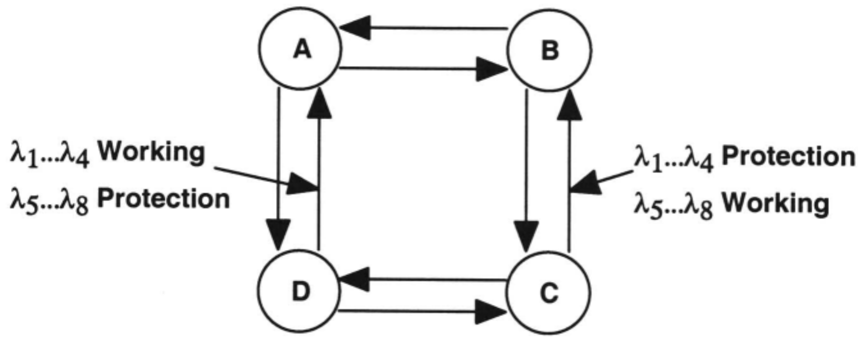


Figure 4. Protection in WDM ring networks

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.

3.1.3 Virtual topology reconfiguration

One of the benefits of having a configurable WDM ring for the feeder network is that the virtual (electronic) topology of the network can be changed in response to changes in traffic conditions. The virtual topology, seen by the electronic layer, consists of a set of nodes interconnected by lightpaths (wavelengths). 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. However, the configurable nature of WDM also allows the electronic topology to be dynamically optimized in response to changes in traffic conditions.

This is achieved by changing the lightpath connectivity between electronic switches and routers, thereby reconfiguring the electronic virtual topology. Lightpaths can be changed via tuning of the transmitter wavelengths in combination with frequency-selective-switches that can alter the route of a wavelength inside the network. For example, consider four nodes physically connected in a ring. Assume that each node has one port (i.e., single receiver and transmitter) and that the fiber supports two wavelengths, λ1 and λ2. A connected, fixed logical topology must take the form of a unidirectional ring, as pictured in Figure 5a. If a full-wavelength call is in progress from node 1 to node 3, and a call request arrives between nodes 2 and 4, then that request must be blocked. In a reconfigurable system, both calls can be supported as shown in Figure 5b.

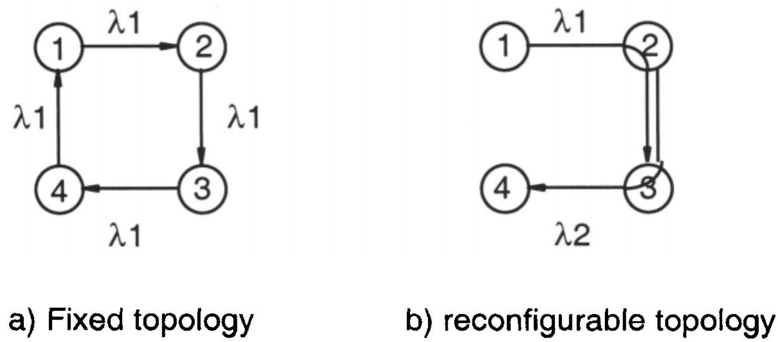


Figure 5. Using WDM to reconfigure the electronic topology

Preliminary studies on reconfiguration of a WDM ring show significant promise [SM00]. The work in [SM00] assumes that calls take up a full wavelength and cannot be rearranged or rerouted. The assumption about rerouting is made in order to eliminate the possibility of calls being adversely affected by reconfiguration. Shown in Figure 6 is the gain that can be achieved through reconfiguration. This gain is defined to be the ratio of the load that can be supported by a reconfigurable system to that of a bi-directional, fixed topology system at a given blocking probability (a blocking probability of 0.01 is used in Figure 6). As can be seen from the figure, the capacity gain due to reconfiguration is most significant when the ratio of wavelengths to ports per node (W/P) is large and the number of ports per node is small. The results from [SM00] indicate that when the number of wavelengths is much larger than the number of available ports per node, a capacity gain on the order of N can be obtained, where N is the number of nodes in the network. When the number of wavelengths approaches the number of ports per node, the benefits of reconfiguration are significantly diminished with approximately a factor of two gain in capacity. The concept of topology reconfiguration can similarly be applied to packet networks (e.g., IP) with the objective of reducing network queuing delays [NM00].

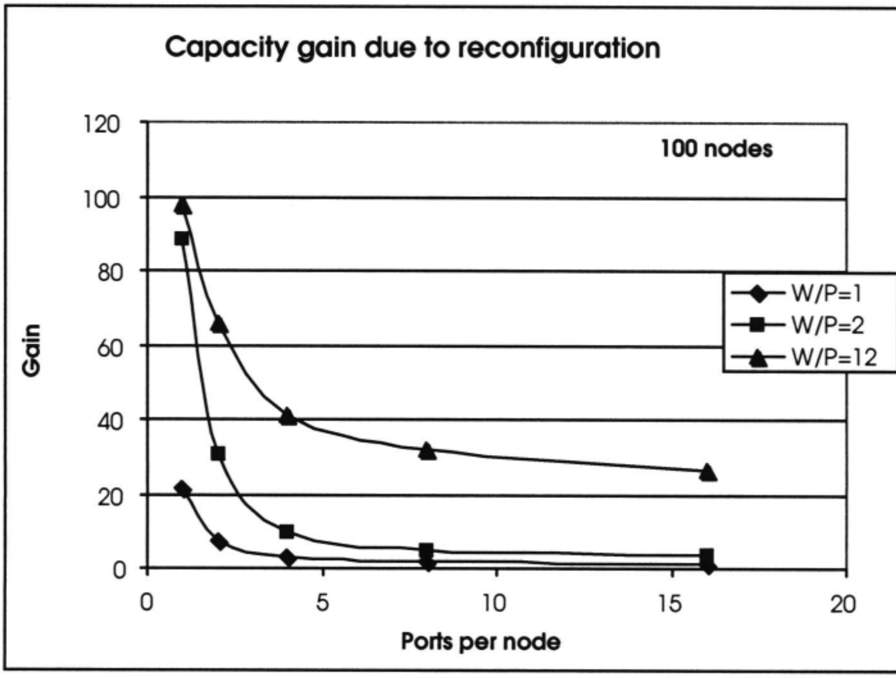


Figure 6. Performance of reconfiguration in a WDM ring network

3.1.4 Traffic grooming

One of the most important functions of the electronic layer of the network is multiplexing lower rate streams into higher rate channels or wavelengths. However, if calls are indiscriminately multiplexed on to wavelengths then each wavelength will have to be electronically processed at every node. Alternatively, if calls are *groomed* with foresight onto wavelengths, then the number of wavelengths that need to be processed at each node can be significantly reduced.

For example, when a SONET ring network is used to provide point-to-point OC-3 circuits between pairs 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. 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). Traffic grooming algorithms can be designed to minimize electronic costs while simultaneously making efficient use of wavelengths. Traffic grooming on SONET ring networks with uniform all-to-all traffic show a significant reduction in the number of required ADMs [MC98, CM98, SGS98,CM00]. Figure 7 shows the significant savings in the number of ADMs needed in a SONET ring network using an algorithm developed in [CM98]. The results from [CM98] assume the use of a unidirectional ring and uniform traffic. Similar results, however, were obtained in [SGS98] for a bi-directional ring. In [CM00] simple algorithms were developed for general, non-uniform traffic, that show similar savings in the number of ADMs. In [BM99] traffic grooming for dynamic traffic is considered, where the number of circuits between node pairs is allowed to change dynamically with the only restriction being that the total number of circuits entering or leaving a node is constrained. Further, while the discussion here focused on the use of SONET equipment, similar optimization can be done for other electronic technologies (e.g., ATM or IP) with the goal of minimizing the number of electronic ports used in the network.

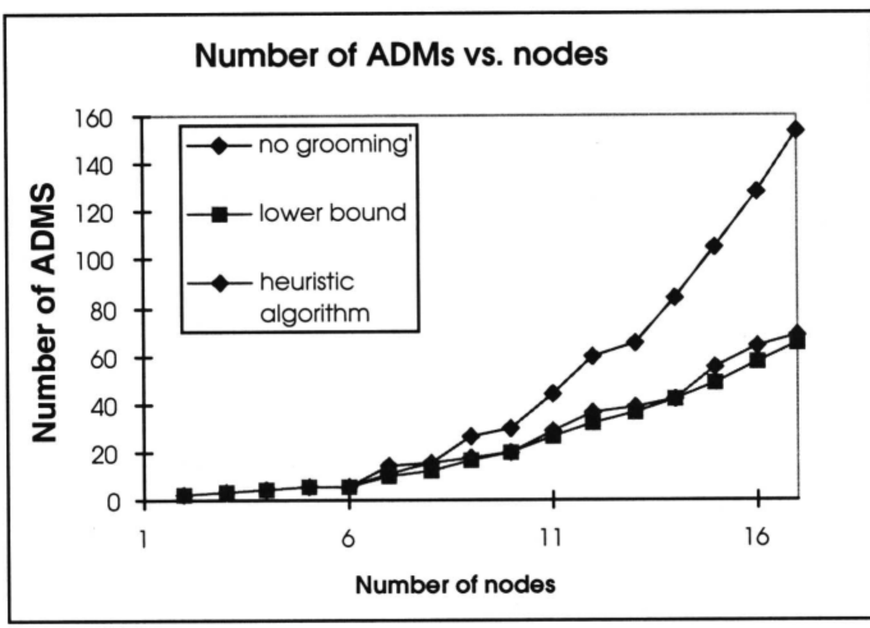


Figure 7. Performance of grooming in a WDM/SONET ring network

3.2 C/D Architecture

The C/D network is responsible for aggregating traffic from user locations to the Access Node (AN). With about 10 AN's per feeder network, each C/D network will cover an area of approximately 100 sq. km and support between 10 and 100 businesses with an aggregate transmission capacity of 100 Gbps. As stated earlier the C/D network should be passive (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 8. The simplest architecture would use a dedicated fiber pair for each node. 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 AN). 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 8.

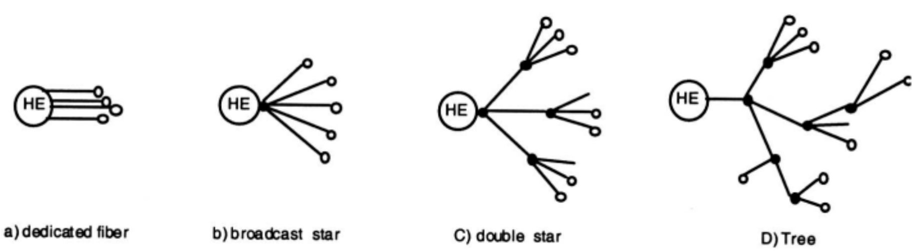


Figure 8. PON architectures

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, FR94]. 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. Alternatively, one could also design a wavelength routed tree PON where the fiber is shared but each user gets its own wavelength. With this approach, at each splitting point on the tree a WDM mux/demux is used to split off the proper wavelengths. This approach has a number of attractive features including wavelength isolation among the different users and reduced splitting losses. However, a wavelength routed PON has one significant disadvantage in that it cannot be used for end-to-end optical flows. In the absence of wavelength changers in the network, setting up end-to-end lightpaths requires the use of the same wavelength along the path. With a wavelength routed PON, setting up end-to-end lightpath between two users becomes rarely possible. For this reason, in our architecture we only consider <u>broadcast</u> PONs where each user can access any of the wavelengths.

In our proposed architecture, as shown in Figure 2, we use passive WDM broadcast rings and trees in the C/D network. As we will show next, a tree architecture provides greater scalability as it can achieve more efficient power splitting. On the other hand, the ring architecture is proposed for use in cases where diversity is needed for protection, as discussed later in this section.

3.2.1 Power budget

Using our assumptions on the size and capacity of an access network, a C/D network must support as many as 100 users, each transmitting at a minimum rate of 2.5 Gbps (OC-48) and traversing a distance of up to 10 km. These requirements impact the power budget in the C/D network and hence the topology of the network. To demonstrate this impact we provide a simple analysis of the power budget for a tree and ring C/D network. Consider a broadcast tap (i.e., coupler) at each of the splitting point in the C/D network, as shown in Figure 9. At each tap (on the tree or ring) a fraction, T, of the power will go to the left side and 1-T to the right side. In addition, each tap will incur in an additional excess power loss of α, for imperfection in the coupler and splicing of the coupler into the fiber.

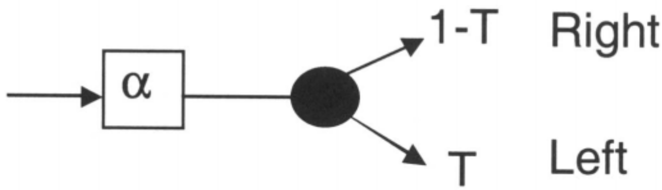


Figure 9. Power splitting at a C/D network tap

When the C/D network is a binary tree, T should be 0.5, because an equal fraction of the power should be sent along either side of the tree. With a binary tree network, each leaf node will be at depth $\mathrm{Log}_2(N)$[2]. When the C/D network is a ring, each leaf node will be at a different depth and hence nodes near the head of the ring will experience different losses than nodes further away. Ideally, one could set the tap values at each node on the ring so that all nodes receive the same amount of power [RS98]. However, designing a network where each node has a different tap is not practical. Instead, one can use a single tap value that would minimize the maximum loss on the ring. For a ring with N nodes the maximum loss is sustained by the last node and is equal to $L_{max} = T^{N-1}(1-T)$. In order to minimize this loss a value of T=(N-1)/N should be used. While the exact value of T depends of the size of the ring, for moderately sized rings (10-20 nodes) a value of T=0.95 can be used. Shown in Figure 10 is the power loss vs. number of nodes for a ring and tree networks where the excess loss in each tap is 0.3 and T=0.5 for a tree and 0.95 for the ring. After accounting for fiber plant losses (2.5 dB for 10 km), transmitter launch power (5 dBm), receiver sensitivity (-25 dBm at 2.5 Gbps and BER=10^{-12}), and a 5 dB link margin, a power budget of about 22dB is available. As can be seen from the Figure with this power budget a ring C/D network can support only about 10 to 20 nodes while a tree can support over 100 nodes. For this reason, when a large number of nodes must be supported on a single C/D network, a tree is the preferred choice. However, as we will show later in this section, a ring network can be used to provide passive protection and so, should be used in cases where protection is essential.

[2]We assume, for simplicity, that users are located only at the leaf nodes.

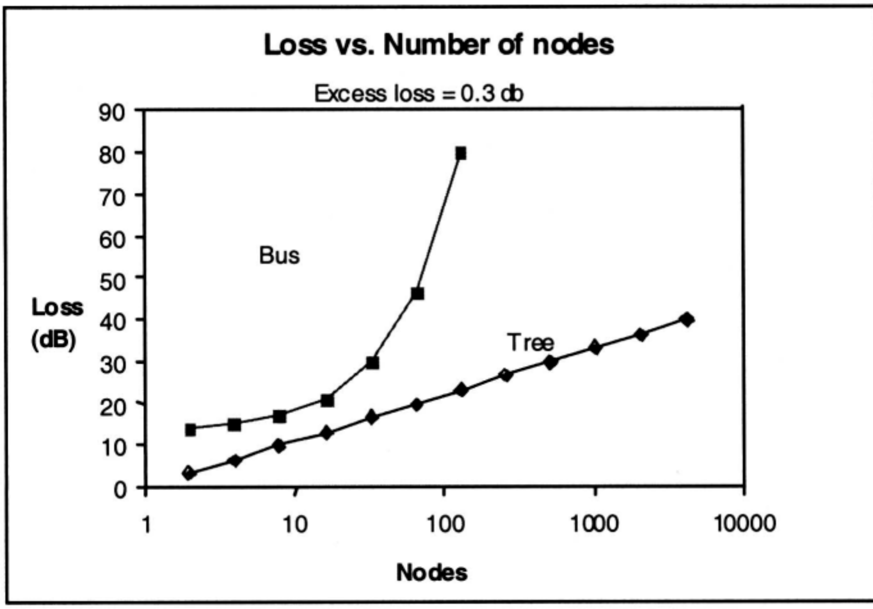


Figure 10. Power loss in the C/D network

3.2.2 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. In the C/D network, such an access mechanism is particularly important to allow users to share access to the IP router located at the AN.

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 C/D 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, BS95]. 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.

Our architecture uses a MAC protocol similar to that proposed in [MB98] 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 11, a centralized scheduler is located at the AN 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 AN, 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.

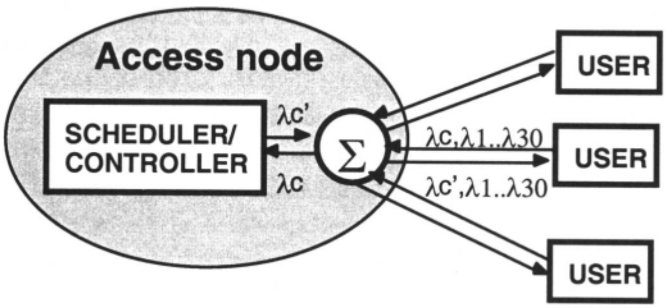


Figure 11. Scheduler based MAC protocol

The scheduler is able to overcome the effects of propagation delays by measuring the round-trip delay of each user to the AN and using that information to inform users of their turn to transmit in a timely manner. For example consider Figure 12. 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].

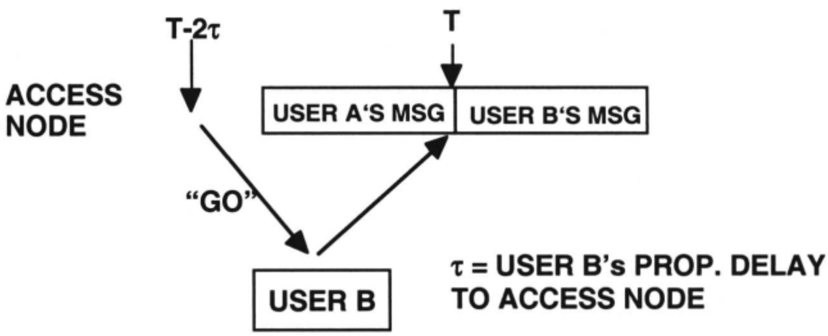


Figure 12. Use of ranging to overcome propagation delays.

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 [MB99]. Scheduling algorithms for transmitting multicast traffic in WDM broadcast-and-select networks are discussed in [MO98, MO99]. The performance of the protocol depends primarily on the scheduling algorithm used by the scheduler. In [MB98] 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].

3.2.3 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 13. 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).

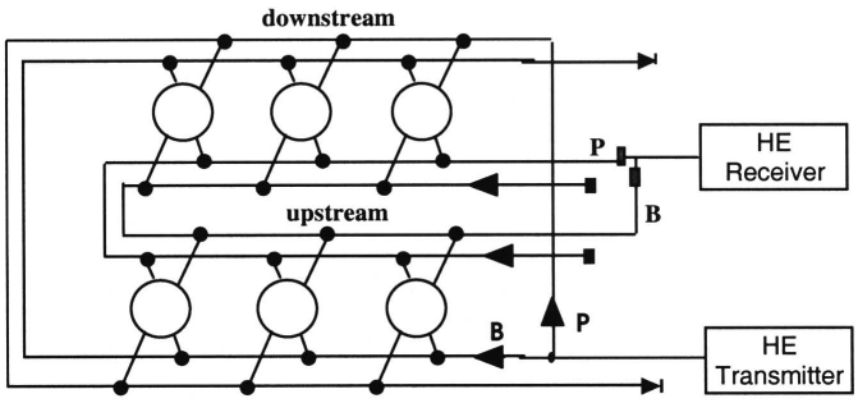


Figure 13. A four fiber passive protection ring

4. THE NGI ONRAMP TEST-BED

The concepts described in this chapter will be demonstrated in the ONRAMP test-bed. The ONRAMP test-bed will consist of a 5 node configurable WDM feeder ring, as shown in Figure 14. A node at the AT&T Boston hub will provide connectivity to backbone networks and other NGI test-beds. The ONRAMP feeder ring is a bi-directional, dual fiber ring with each fiber supporting 16 wavelengths. Eight wavelengths, in each direction, are used for working traffic and eight for loop-back protection. Bi-directionality is used to provide shortest path routing which can increase utilization by a factor of 2 over a uni-directional ring.

Attached to each AN, via a C/D network, are IP routers, workstations and other electronic devices (e.g., ATM switches). Two type of services will be provided; an IP service and a switched optical (wavelength) service. The IP service will be provided via an IP router located at the AN, as shown in Figure 16. Users requiring an IP service can connect to the IP router via the C/D network. Such a connection can be established using a dedicated wavelength, or using a MAC protocol to share wavelengths. In this way, workstations and routers can connect to the IP service at high-rates (e.g., 2.5 Gbps).

In addition to the IP service, ONRAMP will provide a switched optical service. This switched optical service can be used to transparently support other conventional electronic services. For example, it can be used to support a private ATM network, with the ATM switches located at the customer premises and connected via lightpaths. Similarly, SONET, Frame Relay and other legacy equipment can be supported transparently to the access network.

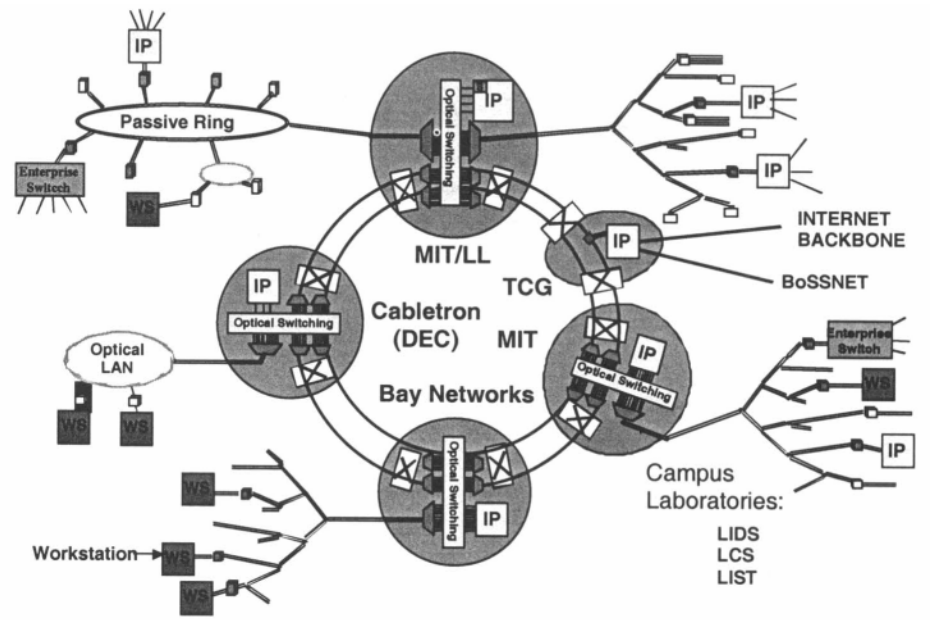


Figure 14. The Onramp feeder ring

The transparent optical service provided by ONRAMP will also enable optical flow switching (OFS). The basic idea behind OFS is very similar to IP switching or Tag switching that is used to bypass IP routing in the internet [NLM96, REK97]. With OFS, a lightpath can be established for large data transactions such as the transfer of large files or long duration and high bandwidth streams across the network. This optical flow will bypass all of the electronics in the network and be switched at the WDM layer as shown in Figure 15.

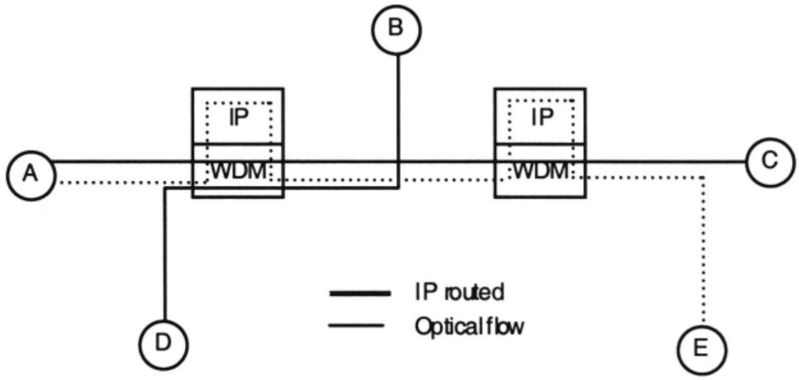


Figure 15. Optical Flow Switching

The AN, shown in Figure 16, has a combination of optical and electronic switching. Electronic switching is provided in the form of an IP router and optical switching in the form of a configurable WDM add/drop multiplexer. Since the feeder network is a configurable WDM ring, each node has the ability to add or drop any of the wavelengths from the ring to the local C/D network. The configurable WDM add/drop multiplexer is designed using a WDM mux and de-mux and a series of 2x2 switches (one per wavelength) that can be configured to either drop the wavelength at a node or bypass that node.

Protection is provided using optical loop-back using the dual fiber ring. In the clockwise direction a band of wavelengths (A) is used for working traffic and in the counter-clockwise direction band (B) is used for working traffic. The working bands are reserved for protection in the opposite direction. A pair of 2x2 switches at each node is used to provide loop-back protection in the event that a loss of power (light) is detected (due to a fiber cut or node failure). An A/B multiplexer at each node is used to allow the protection bands to bypass that node. This bypass helps reduce the cost of the node, because those wavelengths do not have to be demultiplexed further.

Wavelengths from the C/D network can either terminate at a port on the AN's IP router (for an IP service) or be switched onto the feeder ring to provide a switched optical service. The AN's IP router can access the feeder ring using a number of tunable interfaces (3 shown). These ports are used to connect the routers to one another and to the backbone network. Having tunable ports, combined with the configurable optical switches will allow for reconfiguration of the electronic topology of the network.

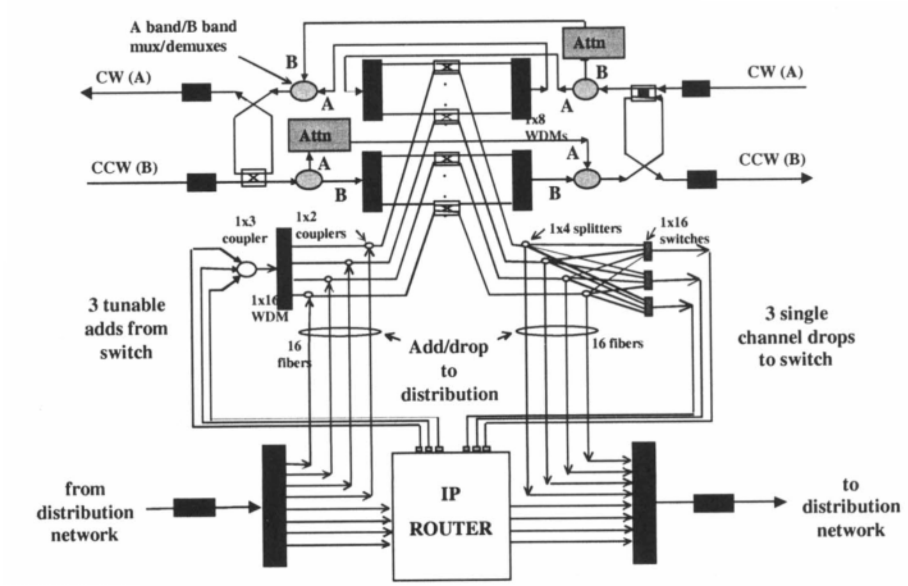


Figure 16. Feeder network AN

ACKNOWLEDGEMENTS

This work was supported by Darpa under the Next Generation Internet (NGI) initiative. The author wishes to acknowledge contributions from: Rick Barry, Vincent Chan, Angela Chiu, Nan Froberg, Mark Kuznetsov, Aradhana Narula, Kristin Rauschenbach , Adel Saleh, Brett Schein, Jane Simmons, and Eric Swanson.

REFERENCES

[BM99] Randy Berry and Eytan Modiano, “Minimizing Electronic Multiplexing Costs for Dynamic Traffic in Unidirectional SONET Ring Networks,” ICC ’99, Vancouver, CA, June 1999.

[BS95] Bo Li and K. M. Sivalingam, "Channel Access Protocols for High Speed LANs Using WDM: A Comparative Study", in Proc. SPIE Conference on All-Optical Communications Systems: Architecture, Control, and Network Issues, (Philadelphia, PA), pp. 283--294, Oct. 1995.

[CM98] A. L. Chiu and E. Modiano, “Reducing Electronic Multiplexing Costs in Unidirectional SONET/WDM Ring Networks via Efficient Traffic Grooming,” Globecom ’98, Sydney, Australia, November, 1998.

[CM00] Angela Chiu and Eytan Modiano, "Traffic Grooming Algorithms for Reducing Electronic Multiplexing Costs in WDM Ring Networks," *IEEE Journal on Lightwave Technology*, January, 2000.

[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.

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

[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.

[MB99] E. Modiano and R. Barry , "Design and analysis of an asynchronous WDM local area network using a master/slave scheduler," Infocom '99, New York, NY, March, 1999.

[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.

[NM00] Aradhana Narula-Tam and Eytan Modiano, "Load Balancing Algorithms for WDM-based IP networks," Infocom, 2000, Tel Aviv, Israel, April, 2000.

[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.

[MO99] Eytan Modiano, "Random Algorithms for scheduling multicast traffic in WDM broadcast-and-select networks," IEEE/ACM Transactions on Networking, August, 1999.

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

[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.

[PG98] S. Perkins and A. Gatherer, "Two-way Broadband CATV-HFC Networks: State of the art and Future trends," Computer Networks, Vol. 31, No. 4, February, 1999.

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

[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.

[SM00] Brett Schein and Eytan Modiano, "Quantifying the benefits of configurability in circuit-switched WDM ring networks," Infocom, 2000, Tel Aviv, Israel, April, 2000.

[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.