Optical Components for WDM Lightwave Networks

Michael S. Borella
IEEE

Jason P. Jue
IEEE

Dhritiman Banerjee
IEEE

Byrav Ramamurthy
IEEE

Biswanath Mukherjee
IEEE

Follow this and additional works at: http://digitalcommons.unl.edu/csearticles

Part of the Computer Sciences Commons

Borella, Michael S.; Jue, Jason P.; Banerjee, Dhritiman; Ramamurthy, Byrav; and Mukherjee, Biswanath, "Optical Components for WDM Lightwave Networks" (1997). CSE Journal Articles, 73. http://digitalcommons.unl.edu/csearticles/73

This Article is brought to you for free and open access by the Computer Science and Engineering, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in CSE Journal Articles by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Optical Components for WDM Lightwave Networks

MICHAEL S. BORELLA, MEMBER, IEEE, JASON P. JUE, DHритIMAN BANEРJEE, BYRAV RAMAMURTHY, STUDENT MEMBER, IEEE, AND BISWANATH MUKHERJEE, MEMBER, IEEE

Recently, there has been growing interest in developing optical fiber networks to support the increasing bandwidth demands of multimedia applications, such as video conferencing and World Wide Web browsing. One technique for accessing the huge bandwidth available in an optical fiber is wavelength-division multiplexing (WDM). Under WDM, the optical fiber bandwidth is divided into a number of nonoverlapping wavelength bands, each of which may be accessed at peak electronic rates by an end user. By utilizing WDM in optical networks, we can achieve link capacities on the order of 50 THz. The success of WDM networks depends heavily on the available optical device technology. This paper is intended as a tutorial on some of the optical device issues in WDM networks. It discusses the basic principles of optical transmission in fiber and reviews the current state of the art in optical device technology. It introduces some of the basic components in WDM networks, discusses various implementations of these components, and provides insights into their capabilities and limitations. Then, this paper demonstrates how various optical components can be incorporated into WDM optical networks for both local and wide-area applications. Last, the paper provides a brief review of experimental WDM networks that have been implemented.

Keywords—Device issues, experimental systems, lightwave network, optical amplifier, optical fiber, switching elements, tunable receiver, tunable transmitter, wavelength converter, wavelength-division multiplexing.

I. INTRODUCTION

Over the past few years, the field of computer and telecommunications networking has experienced tremendous growth. With the rapidly growing popularity of the Internet and the World Wide Web and with the recent deregulation of the telecommunications industry in the United States, this growth can be expected to continue in the foreseeable future. The next decade may bring to the home and office multiple connections of high-definition television, video mail, and digital audio, as well as full Internet connections via user-friendly graphic user interfaces. As more users start to use data networks, and as their usage patterns evolve to include more bandwidth-intensive networking applications, there emerges an acute need for very high bandwidth transport network facilities whose capabilities greatly exceed those of current high-speed networks, such as asynchronous transfer mode (ATM) networks.

The key to the future of networks rests in the relatively young field of fiber optics. Optical fiber provides the huge bandwidth, low loss rate, and cost effectiveness to enable the vision of a “global village.” Given that fiber has a potential bandwidth of approximately 50 Tbps—nearly four orders of magnitude higher than peak electronic data rates—every effort should be made to tap into the capabilities of fiber-optic networks.

Wavelength-division multiplexing (WDM) is one promising approach that can be used to exploit the huge bandwidth of optical fiber. In WDM, the optical transmission spectrum is divided into a number of nonoverlapping wavelength (or frequency) bands, with each wavelength supporting a single communication channel operating at peak electronic speed. Thus, by allowing multiple WDM channels to coexist on a single fiber, we can tap into the huge fiber bandwidth, with the corresponding challenges being the design and development of appropriate network architectures, protocols, and algorithms.

Research and development on optical WDM networks have matured considerably over the past few years, and a number of experimental prototypes have been and are currently being deployed and tested in the United States, Europe, and Japan. It is anticipated that the next generation of the Internet will employ WDM-based optical backbones.

The success of WDM networks relies heavily upon the available optical components. A block diagram of a
WDM communication system is shown in Fig. 1. The network medium may be a simple fiber link, a passive star coupler (PSC) (for a broadcast and select network), or a network of optical or electronic switches and fiber links. The transmitter block consists of one or more optical transmitters, which may be either fixed to a single wavelength or tunable across a range of wavelengths. Each optical transmitter consists of a laser and a laser modulator and may also include an optical filter for tuning purposes. If multiple optical transmitters are used, then a multiplexer or coupler is needed to combine the signals from different laser transmitters onto a single fiber. The receiver block may consist of a tunable filter followed by a photodetector receiver or a demultiplexer followed by an array of photodetectors. Examples of some WDM transmitters and receivers are shown in Fig. 2. Amplifiers may be required in various locations throughout the network to maintain the strength of optical signals.

Designers of next-generation lightwave networks must be aware of the properties and limitations of optical fiber and devices in order for their corresponding protocols and algorithms to take advantage of the full potential of WDM. Often, a network designer may approach the WDM architectures and protocols from an overly simplified, ideal, or traditional-networking point of view. Unfortunately, this may lead an individual to make unrealistic assumptions about the properties of fiber and optical components, and hence may result in an unrealizable or impractical design.

This paper serves as an introduction to WDM device issues. No background in optics or advanced physics is needed. For a more advanced and/or detailed discussion of WDM devices, we refer the interested reader to [1]–[6].

This paper presents an overview of optical fiber and devices such as couplers, optical transmitters, optical receivers and filters, optical amplifiers, optical routers, and switches. It paper attempts to condense the physics behind the principles of optical transmission in fiber in order to provide some background for the novice reader. WDM network-design issues are then discussed in relation to the advantages and limits of optical devices. Last, this paper demonstrates how these optical components can be used to create broadcast networks for local networking applications and wavelength-routed networks for wide-area deployment. The paper concludes with a note on the current status of optical technology and how test networks have used some of the optical devices described in this paper with a reasonable amount of success.

II. OPTICAL FIBER

Fiber possesses many characteristics that make it an excellent physical medium for high-speed networking. Fig. 3 shows the two low-attenuation regions of optical fiber [1]. Centered at approximately 1300 nm is a range of 200 nm in which attenuation is less than 0.5 dB per kilometer. The total bandwidth in this region is about 25 THz. Centered at 1550 nm is a region of similar size with attenuation as low as 0.2 dB per kilometer. Combined, these two regions provide a theoretical upper bound of 50 THz of bandwidth.¹

The dominant loss mechanism in good fibers is Rayleigh scattering, while the peak in loss in the 1400-nm region is due to hydroxyl-ion (OH−) impurities in the fiber. Other sources of loss include material absorption and radiative loss.

By using these large low-attenuation areas for data transmission, the signal loss for a set of one or more wavelengths can be made very small, thus reducing the number of amplifiers and repeaters needed. In single-channel long-distance experiments, optical signals have been sent over hundreds of kilometers without amplification. Besides its enormous bandwidth and low attenuation, fiber also offers low error rates. Fiber-optic systems typically operate at bit error rates (BER’s) of less than 10⁻¹⁰⁻¹¹.

The small size and thickness of fiber allows more fiber to occupy the same physical space as copper, a property that is desirable when installing local networks in buildings. Fiber is flexible, difficult to break, reliable in corrosive environments, and deployable at short notice (which makes it particularly favorable for military communications systems). Also, fiber transmission is immune to electromagnetic interference and does not cause interference. Last, fiber is made from one of the cheapest and most readily available substances on earth, sand. This makes fiber environmentally sound; and unlike copper, its use will not deplete natural resources.

A. Optical Transmission in Fiber

Before discussing optical components, it is essential to understand the characteristics of the optical fiber itself. Fiber is essentially a thin filament of glass that acts as a waveguide. A waveguide is a physical medium or path that allows the propagation of electromagnetic waves, such as light. Due to the physical phenomenon of total internal reflection, light can propagate the length of a fiber with little loss. Fig. 4 shows the cross section of the two types of fiber most commonly used: multimode and single mode. In order to understand the concept of a mode and to distinguish between these two types of fiber, a diversion into basic optics is needed.

Light travels through vacuum at a speed of \( c_{\text{vac}} = 3 \times 10^8 \) m/s. Light can also travel through any transparent material, but the speed of light will be slower in the material than in a vacuum. Let \( c_{\text{mat}} \) be the speed of light for a given material. The ratio of the speed of light in a vacuum to that in a material is known as the material’s refractive index (n) and is given by \( n_{\text{mat}} = c_{\text{vac}}/c_{\text{mat}} \).

When light travels from material of a given refractive index to material of a different refractive index (i.e., when ¹Usable bandwidth, however, is limited by fiber nonlinearities (see Section II-E).
refraction occurs), the angle at which the light is transmitted in the second material depends on the refractive indexes of the two materials as well as the angle at which light strikes the interface between the two materials. Due to Snell’s law, we have $n_a \sin \theta_a = n_b \sin \theta_b$, where $n_a$ and $n_b$ are the refractive indexes of the first substance and the second substance, respectively; $\theta_a$ is the angle of incidence, or the angle with respect to normal that light hits the surface between the two materials; and $\theta_b$ is the angle of light in the second material. However, if $n_a > n_b$ and $\theta_a$ is greater than some critical value and the rays are reflected back into substance $a$ from its boundary with substance $b$.

Looking again at Fig. 4, we see that the fiber consists of a core completely surrounded by a cladding (both of which consist of glass of different refractive indexes). Let us first consider a step-index fiber, in which the change of refractive index at the core-cladding boundary is a step function. If the refractive index of the cladding is less than that of the core, then total internal reflection can occur in the core and light can propagate through the fiber (as shown in Fig. 5). The angle above which total internal reflection will take place is known as the critical angle and is given by $\theta_{crit}$, which corresponds to $\theta_{clad} = 90^\circ$. From Snell’s law, we have

$$\sin \theta_{clad} = \frac{n_{core}}{n_{clad}} \sin \theta_{core}.$$  

The critical angle is then

$$\theta_{crit} = \sin^{-1} \left( \frac{n_{clad}}{n_{core}} \right). \quad (1)$$

So, for total internal reflection, we require

$$\theta_{crit} > \sin^{-1} \left( \frac{n_{clad}}{n_{core}} \right).$$

In other words, for light to travel down a fiber, the light must be incident on the core-cladding surface at an angle greater than $\theta_{crit}$.

In some cases, the fiber may have a graded index, in which the interface between the core and the cladding undergoes a gradual change in refractive index with $n_i > n_{i+1}$ (Fig. 6). A graded-index fiber reduces the minimum $\theta_{crit}$ required for total internal reflection and also helps to reduce the intermodal dispersion in the fiber. Intermodal dispersion will be discussed in the following sections.

For light to enter a fiber, the incoming light should be at an angle such that the refraction at the air-core boundary results in the transmitted light’s being at an angle for which total internal reflection can take place at the core-cladding boundary. As shown in Fig. 7, the maximum value of $\theta_{air}$ can be derived from

$$n_{air} \sin \theta_{air} = n_{core} \sin (90^\circ - \theta_{crit}) = n_{core} \sqrt{1 - \sin^2 \theta_{crit}}.$$  

From (1), since $\sin \theta_{crit} = n_{clad}/n_{core}$, we can rewrite (2) as

$$n_{air} \sin \theta_{air} = \sqrt{n_{core}^2 - n_{clad}^2}.$$  

The quantity $n_{air} \sin \theta_{air}$ is referred to as the numerical aperture of the fiber (NA) and $\theta_{air}$ is the maximum angle with respect to the normal at the air-core boundary, so that the incident light that enters the core will experience total internal reflection inside the fiber.

### B. Multimode Versus Single-Mode Fiber

A mode in an optical fiber corresponds to one of possibly multiple ways in which a wave may propagate through the fiber. It can also be viewed as a standing wave in
Fig. 4. Multimode and single-mode optical fibers.

Fig. 5. Light traveling via total internal reflection within a fiber.

Fig. 6. Graded-index fiber.

Fig. 7. Numerical aperture of a fiber.

the transverse plane of the fiber. More formally, a mode corresponds to a solution of the wave equation that is derived from Maxwell’s equations and subject to boundary conditions imposed by the optical fiber waveguide.

An electromagnetic wave propagating along an optical fiber consists of an electric field vector $\mathbf{E}$ and a magnetic field vector $\mathbf{H}$. Each field can be broken down into three components. In the cylindrical coordinate system, these components are $E_\rho$, $E_\phi$, $E_z$, $H_\rho$, $H_\phi$, and $H_z$, where $\rho$ is the component of the field that is normal to the wall (core-cladding boundary) of the fiber, $\phi$ is the component of the field that is tangential to the wall of the fiber, and $z$ is the component of the field that is in the direction of propagation. Fiber modes typically are referred to using the notation $E_\phi H_\phi$ (if $H_\phi > E_\phi$), or $E_\rho H_\rho$ (if $E_\rho > H_\rho$), where $x$ and $y$ are both integers. For the case $x = 0$, the modes are also referred to as transverse-electric (TE), in which case $E_\phi = 0$, or transverse-magnetic (TM), in which case $H_\phi = 0$.

Although total internal reflection may occur for any angle $\theta_{\text{core}}$ that is greater than $\theta_{\text{crit}}$, light will not necessarily propagate for all of these angles. For some of these angles, light will not propagate due to destructive interference between the incident light and the reflected light at the core-cladding interface within the fiber. For other angles of incidence, the incident wave and the reflected wave at the core-cladding interface constructively interfere in order to maintain the propagation of the wave. The angles for which waves do propagate correspond to modes in a fiber. If more than one mode may propagate through a fiber, the fiber is called multimode. In general, a larger core diameter or high operating frequency allows a greater number of modes to propagate.

The number of modes supported by a multimode optical fiber is related to the normalized frequency $V$, which is defined as

$$V = k_0 a \sqrt{n_{\text{core}}^2 - n_{\text{clad}}^2}$$

where $k_0 = 2\pi/\lambda$, $a$ is the radius of the core, and $\lambda$ is the wavelength of the propagating light in vacuum. In multimode fiber, the number of modes $m$ is given approximately by

$$m \approx \frac{1}{2} V^2.$$
of propagation. The effect of intermodal dispersion may be reduced through the use of graded-index fiber, in which the region between the cladding and the core of the fiber consists of a series of gradual changes in the index of refraction (see Fig. 6). Even with graded-index multimode fiber, however, intermodal dispersion may still limit the bit rate of the transmitted signal and may limit the distance that the signal can travel.

One way to limit intermodal dispersion is to reduce the number of modes. From (4) and (5), we observe that this reduction in the number of modes can be accomplished by reducing the core diameter, reducing the numerical aperture, or increasing the wavelength of the light.

By reducing the fiber core to a sufficiently small diameter and reducing the numerical aperture, it is possible to capture only a single mode in the fiber. This single mode is the $HE_{11}$ mode, also known as the fundamental mode. Single-mode fiber usually has a core size of about 10 $\mu$m, while multimode fiber typically has a core size of 50–100 $\mu$m (see Fig. 4). A step-index fiber will support a single mode if $V$ in (4) is less than 2.4048 [7].

Thus, single-mode fiber eliminates intermodal dispersion and can hence support transmission over much longer distances. However, it introduces the problem of concentrating enough power into a very small core. LED’s cannot couple enough light into a single-mode fiber to facilitate long-distance communications. Such a high concentration of light energy may be provided by a semiconductor laser, which can generate a narrow beam of light.

C. Attenuation in Fiber

Attenuation in optical fiber leads to a reduction of the signal power as the signal propagates over some distance. When determining the maximum distance that a signal can propagate for a given transmitter power and receiver sensitivity, one must consider attenuation. Let $P(L)$ be the power of the optical pulse at distance $L$ km from the transmitter and $A$ be the attenuation constant of the fiber (in dB/km). Attenuation is characterized by [2]

$$P(L) = 10^{-AL/20}P(0)$$  \hspace{1cm} (6)

where $P(0)$ is the optical power at the transmitter. For a link length of $L$ km, $P(L)$ must be greater than or equal to $P_r$, the receiver sensitivity. From (6), we get

$$L_{\text{max}} = \frac{10}{A} \log_{10} \frac{P(0)}{P_r}.$$  \hspace{1cm} (7)

The maximum distance between the transmitter and the receiver (or the distance between amplifiers) depends more heavily on the constant $A$ than on the optical power launched by the transmitter. Referring back to Fig. 3, we note that the lowest attenuation occurs at approximately 1550 nm.

3The amplifier sensitivity is usually equal to the receiver sensitivity, while the amplifier output is usually equal to optical power at a transmitter.

D. Dispersion in Fiber

Dispersion is the widening of a pulse duration as it travels through a fiber. As a pulse widens, it can broaden enough to interfere with neighboring pulses (bits) on the fiber, leading to intersymbol interference. Dispersion thus limits the bit spacing and the maximum transmission rate on a fiber-optic channel.

As mentioned earlier, one form of dispersion is intermodal dispersion. This is caused when multiple modes of the same signal propagate at different velocities along the fiber. Intermodal dispersion does not occur in a single-mode fiber.

Another form of dispersion is material or chromatic dispersion. In a dispersive medium, the index of refraction is a function of the wavelength. Thus, if the transmitted signal consists of more than one wavelength, certain wavelengths will propagate faster than other wavelengths. Since no laser can create a signal consisting of an exact single wavelength, material dispersion will occur in most systems.4

A third type of dispersion is waveguide dispersion. Waveguide dispersion is caused because the propagation of different wavelengths depends on waveguide characteristics such as the indexes and shape of the fiber core and cladding.

At 1300 nm, material dispersion in a conventional single-mode fiber is near zero. Luckily, this is also a low-attenuation window (although loss is lower at 1550 nm). Through advanced techniques such as dispersion shifting, fibers with zero dispersion at a wavelength between 1300–1700 nm can be manufactured [8]. In a dispersion-shifted fiber, the core and cladding are designed such that the waveguide dispersion is negative with respect to the material dispersion, thus canceling the total dispersion. The dispersion will only be zero, however, for a single wavelength.

E. Nonlinearities in Fiber

Nonlinear effects in fiber may potentially have a significant impact on the performance of WDM optical communications systems. Nonlinearities in fiber may lead to attenuation, distortion, and cross-channel interference. In a WDM system, these effects place constraints on the spacing between adjacent wavelength channels, limit the maximum power on any channel, and may also limit the maximum bit rate.

1) Nonlinear Refraction: In optical fiber, the index of refraction depends on the optical intensity of signals propagating through the fiber [9]. Thus, the phase of the light at the receiver will depend on the phase of the light sent by the transmitter, the length of the fiber, and the optical intensity. Two types of nonlinear effects caused by this phenomenon are self-phase modulation (SPM) and cross-phase modulation (XPM).

SPM is caused by variations in the power of an optical signal and results in variations in the phase of the signal.

4Even if an unmodulated source consisted of a single wavelength, the process of modulation would cause a spread of wavelengths.
The amount of phase shift introduced by SPM is given by
\[ \phi_{NL} = n_2 k_0 L |E|^2 \]  
(8)

where \( n_2 \) is the nonlinear coefficient for the index of refraction, \( k_0 = 2\pi/\lambda \), \( L \) is the length of the fiber, and \( |E|^2 \) is the optical intensity. In phase-shift-keying (PSK) systems, SPM may lead to a degradation of the system performance since the receiver relies on the phase information. SPM also leads to the spectral broadening of pulses, as explained below. Instantaneous variations in a signal's phase caused by changes in the signal's intensity will result in instantaneous variations of frequency around the signal's central frequency. For very short pulses, the additional frequency components generated by SPM combined with the effects of material dispersion will also lead to spreading or compression of the pulse in the time domain, affecting the maximum bit rate and the BER.

XPM is a shift in the phase of a signal caused by the change in intensity of a signal propagating at a different wavelength. XPM can lead to asymmetric spectral broadening, and combined with SPM and dispersion may also affect the pulse shape in the time domain.

Although XPM may limit the performance of fiber-optic systems, it may also have advantageous applications. XPM can be used to modulate a pump signal at one wavelength from a modulated signal on a different wavelength. Such techniques can be used in wavelength conversion devices and are discussed in Section VII.

2) Stimulated Raman Scattering (SRS): SRS is caused by the interaction of light with molecular vibrations. Light incident on the molecules creates scattered light at a longer wavelength than that of the incident light. A portion of the light traveling at each frequency in a Raman-active fiber is downshifted across a region of lower frequencies. The light generated at the lower frequencies is called the Stokes wave. The range of frequencies occupied by the Stokes wave is determined by the Raman gain spectrum, which covers a range of around 40 THz below the frequency of the input light. In silica fiber, the Stokes wave has a maximum gain at a frequency of around 13.2 THz less than the input signal.

The fraction of power transferred to the Stokes wave grows rapidly as the power of the input signal is increased. Under very high input power, SRS will cause almost all of the power in the input signal to be transferred to the Stokes wave.

In multiwavelength systems, the channels of shorter wavelength will lose some power to each of the higher-wavelength channels within the Raman gain spectrum. To reduce the amount of loss, the power on each channel needs to be below a certain level. In [10], it is shown that in a ten-channel system with 10-nm channel spacing, the power on each channel should be kept below 3 mW to minimize the effects of SRS.

3) Stimulated Brillouin Scattering (SBS): SBS is similar to SRS except that the frequency shift is cause by sound waves rather than molecular vibrations [9]. Other characteristics of SBS are that the Stokes wave propagates in the opposite direction of the input light, and SBS occurs at relatively low input powers for wide pulses (greater than 1 \( \mu \)s) but has negligible effect for short pulses (less than 10 ns) [11]. The intensity of the scattered light is much greater in SBS than in SRS but the frequency range of SBS, on the order of 10 GHz, is much lower than that of SRS. Also, the gain bandwidth of SBS is only on the order of 100 MHz.

To counter the effects of SBS, one must ensure that the input power is below a certain threshold. Also, in multiwavelength systems, SBS may induce cross talk between channels. Cross talk will occur when two counter-propagating channels differ in frequency by the Brillouin shift, which is around 11 GHz for wavelengths at 1550 nm. The narrow gain bandwidth of SBS, however, makes SBS cross talk fairly easy to avoid.

4) Four-Wave Mixing (FWM): FWM occurs when two wavelengths operating at frequencies \( f_1 \) and \( f_2 \), respectively, mix to cause signals at \( 2f_1 - f_2 \) and \( 2f_2 - f_1 \). These extra signals, called sidebands, can cause interference if they overlap with frequencies used for data transmission. Likewise, mixing can occur between combinations of three or more wavelengths. The effect of FWM in WDM systems can be reduced by using unequally spaced channels [12].

FWM can be used to provide wavelength conversion, as will be shown in Section VII.

5) Summary: Nonlinear effects in optical fibers may potentially limit the performance of WDM optical networks. Such nonlinearities may limit the optical power on each channel, limit the maximum number of channels, limit the maximum transmission rate, and constrain the spacing between different channels.

It is shown that in a WDM system using channels spaced 10 GHz apart and a transmitter power of 0.1 mW per channel, a maximum of about 100 channels can be obtained in the 1550-nm low-attenuation region [9].

The details of optical nonlinearities are very complex and beyond the scope of this article. They are a major limiting factor in the available number channels in a WDM system, however, especially those operating over distances greater than 30 km [9]. The existence of these nonlinearities suggests that WDM protocols that limit the number of nodes to the number of channels do not scale well. For further details on fiber nonlinearities, the reader is referred to [11].

F. Couplers

A coupler is a general term that covers all devices that combine light into or split light out of a fiber. A splitter is a coupler that divides the optical signal on one fiber to two or more fibers. The most common splitter is a 1 \( \times \) 2 splitter, as shown in Fig. 8(a). The splitting ratio \( \alpha \) is the amount of power that goes to each output. For a two-port splitter, the most common splitting ratio is 50 : 50, though splitters with any ratio can be manufactured [8]. Combiners [see Fig. 8(b)] are the reverse of splitters, and when turned around, a combiner can be used as a splitter. An input signal to the combiner suffers a power loss of about 3 dB. A 2 \( \times \) 2 coupler [see Fig. 8(c)], in general, is a 2 \( \times \) 1 combiner followed immediately by a 1 \( \times \) 2 splitter,
which has the effect of broadcasting the signals from two input fibers onto two output fibers. One implementation of a 2 x 2 coupler is the fused biconical tapered coupler, which consists of two fibers fused together. In addition to the 50:50 power split incurred in a coupler, a signal also experiences return loss. If the signal enters an input of the coupler, roughly half of the signal’s power goes to each output of the coupler. However, a small amount of power is reflected in the opposite direction and is directed back to the inputs of the coupler. Typically, the amount of power returned by a coupler is 40–50 dB below the input power. Another type of loss is insertion loss. One source of insertion loss is the loss incurred when directing the light from a fiber into the coupler device; ideally, the axes of the fiber core and the coupler input port must be perfectly aligned, but full perfection may not be achievable due to the very small dimensions.

The PSC is a multiport device in which light coming into any input port is broadcast to every output port. The PSC is attractive because the optical power that each output receives $P_{out}$ equals

$$P_{out} = \frac{P_{in}}{N}$$

where $P_{in}$ is the optical power introduced into the star by a single node and $N$ is the number of output ports of the star. Note that this expression ignores the excess loss, caused by flaws introduced in the manufacturing process, that the signal experiences when passing through each coupling element. One way to implement the PSC is to use a combination of splitters, combiners, and couplers as shown in Fig. 9. Another implementation of the star coupler is the integrated-optics planar star coupler, in which the star coupler and waveguides are fabricated on a semiconductor, glass (silica), or polymer substrate. A 19 x 19 star coupler on silicon has been demonstrated with excess loss of around 3.5 dB at a wavelength of 1300 nm [13]. In [14], an 8 x 8 star coupler with an excess loss of 1.6 dB at a wavelength of 1550 nm was demonstrated.

III. OPTICAL TRANSMITTERS

To understand how a tunable optical transmitter works, we must first understand some of the fundamental principles of lasers and how they work. Then we will discuss various implementations of tunable lasers and their properties. Good references on tunable laser technology include [1], [2], [15].

A. How a Laser Works

The word “laser” is an acronym for light amplification by stimulated emission of radiation. The key words are stimulated emission, which is what allows a laser to produce intense high-powered beams of coherent light (light that contains one or more distinct frequencies).

To understand stimulated emission, we must first acquaint ourselves with the energy levels of atoms. Atoms that are stable (in the ground state) have electrons that are in the lowest possible energy levels. In each atom, there are a number of discrete levels of energy that an electron can have; thus, we refer to them as states. To change the level of an atom in the ground state, that atom must absorb energy. This energy can be in many forms but for our purposes here, it can be either light or electrical energy. When an atom absorbs energy, it becomes excited and moves to a higher energy level. At this point, the atom is unstable and usually moves quickly back to the ground state by releasing a photon, a particle of light.

There are certain substances, however, whose states are quasistable, which means that the substances are likely to stay in the excited state for longer periods of time without constant excitation. By applying enough energy (in the form of either an optical pump or an electrical current) to a substance with quasi-stable states for a long enough period of time, population inversion occurs, which means that there are more electrons in the excited state than in the ground state. As we shall see, this inversion allows the substance to emit more light than it absorbs.
Fig. 10 shows a general representation of the structure of a laser. The laser consists of two mirrors that form a cavity (the space between the mirrors), a lasing medium, which occupies the cavity, and an excitation device. The excitation device applies current to the lasing medium, which is made of a quasi-stable substance. The applied current excites electrons in the lasing medium, and when an electron in the lasing medium drops back to the ground state, it emits a photon of light. The photon will reflect off the mirrors at each end of the cavity and will pass through the medium again.

Stimulated emission occurs when a photon passes very closely to an excited electron. The photon may cause the electron to release its energy and return to the ground state. In the process of doing so, the electron releases another photon, which will have the same direction and coherency (frequency) as the stimulating photon. Photons for which the frequency is an integral fraction of the cavity length will coherently combine to build up light at the given frequency within the cavity. Between "normal" and stimulated emission, the light at the selected frequency builds in intensity until energy is being removed from the medium as fast as it is being inserted. The mirrors feed the photons back and forth, so further stimulated emission can occur and higher intensities of light can be produced. One of the mirrors is partially transmitting, so that some photons will escape the cavity in the form of a narrowly focused beam of light. By changing the length of the cavity, the frequency of the emitted light can be adjusted.

The frequency of the photon emitted depends on its change in energy levels. The frequency is determined by the equation

$$ f = \frac{E_i - E_f}{h} $$

where $f$ is the frequency of the photon, $E_i$ is the initial (quasi-stable) state of the electron, $E_f$ is the final (ground) state of the electron, and $h$ is Planck's constant. In a gas laser, the distribution for $E_i - E_f$ is given by an exponential probability distribution, known as the Boltzmann distribution.

A semiconductor material may be doped with impurities to increase either the number of electrons or the number of holes. An n-type semiconductor is doped with impurities that provide extra electrons. These electrons will remain in the conduction band. A p-type semiconductor is doped with impurities that increase the number of holes in the valence band. A p-n junction is formed by layering p-type semiconductor material over n-type semiconductor material.

In order to produce stimulated emission, a voltage is applied across the p-n junction to forward bias the device.
and cause electrons in the “n” region to combine with holes in the “p” region, resulting in light energy’s being released at a frequency related to the band gap of the device. By using different types of semiconductor materials, light with various ranges of frequencies may be released. The actual frequency of light emitted by the laser is determined by the length of the cavity formed by mirrored edges perpendicular to the p-n junction.

An improvement to the bulk laser diode is the multiple-quantum-well (MQW) laser. Quantum wells are thin alternating layers of semiconductor materials. The alternating layers create potential barriers in the semiconductors that confine the position of electrons and holes to a smaller number of energy states. The quantum wells are placed in the region of the p-n junction. By confining the possible states of the electrons and holes, it is possible to achieve higher resolution, low-linewidth lasers (lasers that generate light with a very narrow frequency range).

B. Tunable and Fixed Lasers

While the previous section provided an overview of a generic model of a laser, the transmitters used in WDM networks often require the capability to tune to different wavelengths. This section briefly describes some of the more popular, tunable and fixed single-frequency laser designs.

1) Laser Characteristics: Some of the physical characteristics of lasers that may affect system performance are laser linewidth, frequency stability, and the number of longitudinal modes.

The laser linewidth is the spectral width of the light generated by the laser. The linewidth affects the spacing of channels and also affects the amount of dispersion that occurs when the light is propagating along a fiber. As was mentioned in Section II-D, the spreading of a pulse due to dispersion will limit the maximum bit rate.

Frequency instabilities in lasers are variations in the laser frequency. Three such instabilities are mode hopping, mode shifts, and wavelength chirp [16]. Mode hopping occurs primarily in injection-current lasers and is a sudden jump in the laser frequency caused by a change in the injection current above a given threshold. Mode shifts are changes in frequency due to temperature changes. Wavelength chirp is a variation in the frequency due to variations in injection current. In WDM systems, frequency instabilities may limit the placement and spacing of channels. To avoid large shifts in frequency, methods must be utilized that compensate for variations in temperature or injection current. One approach for temperature compensation is to package with the laser a thermoelectric cooler element that produces cooling as a function of applied current. The current for the thermoelectric cooler may be provided through a thermistor, which is a temperature-dependent resistor.

The number of longitudinal modes in a laser is the number of wavelengths that are amplified by the laser. In lasers consisting of a simple cavity, wavelengths for which an integer multiple of the wavelength is equal to twice the cavity length will be amplified (i.e., wavelengths $\lambda$ for which $n \lambda = 2L$, where $L$ is the length of the cavity and $n$ is an integer). The unwanted longitudinal modes produced by a laser may result in significant dispersion; therefore, it is desirable to implement lasers that produce only a single longitudinal mode.

Some primary characteristics of interest for tunable lasers are the tuning range, the tuning time, and whether the laser is continuously tunable (over its tuning range) or discretely tunable (only to selected wavelengths). The tuning range refers to the range of wavelengths over which the laser may be operated. The tuning time is the time required for the laser to tune from one wavelength to another.

2) Mechanically Tuned Lasers: Most mechanically tuned lasers use a Fabry–Perot cavity that is adjacent to the lasing medium (i.e., an external cavity) to filter out unwanted wavelengths. Tuning is accomplished by physically adjusting the distance between two mirrors on either end of the cavity such that only the desired wavelength constructively interferes with its multiple reflections in the cavity. This approach to tuning results in a tuning range that encompasses the entire useful gain spectrum of the semiconductor laser [3], but tuning time is limited to the order of milliseconds due to the mechanical nature of the tuning and the length of the cavity. The length of the cavity may also limit transmission rates unless an external modulator is used. External-cavity lasers tend to have very good frequency stability.

3) Acoustooptically and Electrooptically Tuned Lasers: Other types of tunable lasers that use external tunable filters include acoustooptically and electrooptically tuned lasers. In an acoustooptic or electrooptic laser, the index of refraction in the external cavity is changed by using either sound waves or electrical current, respectively. The change in the index results in the transmission of light at different frequencies. In these types of tunable lasers, the tuning time is limited by the time required for light to build up in the cavity at the new frequency.

An acoustooptic laser combines a moderate tuning range with a moderate tuning time. While not quite fast enough for packet switching with multigigabit per second channels, the $10 \mu$s tuning time is a vast improvement over that of mechanically tuned lasers (which have millisecond tuning times). Electrooptically tuned lasers are expected to tune on the order of some tens of nanoseconds. Neither of these approaches allows continuous tuning over a range of wavelengths. The tuning range is limited by the range of frequencies generated by the laser (the laser’s gain spectrum) and the range of wavelengths resolvable by the filter [3].

4) Injection-Current-Tuned Lasers: Injection-current-tuned lasers form a family of transmitters that allow wavelength selection via a diffraction grating. The distributed feedback (DFB) laser uses a diffraction grating placed in the lasing medium. In general, the grating consists of a waveguide in which the index of refraction alternates periodically between two values. Only wavelengths that match the period and indexes of the grating will be constructively reinforced. All other wavelengths will
is the period of the grating [16]. The laser is tuned by injecting a current that changes the index of the grating region.

If the grating is moved to the outside of the lasing medium, the laser is called a distributed Bragg reflector (DBR) laser. The tuning in a DBR laser is discrete rather than continuous, and tuning times of less than 10 ns have been measured [3]. In [17], a tuning time of 0.5 ns is reported for a DBR laser with a tuning range of 4 nm, which is capable of supporting eight wavelengths. Because the refractive index range in the DBR laser is limited, the DBR laser has a low maximum tuning range (around 10 nm), which can provide up to 25 channels [18]. One of the drawbacks of the DBR laser is that it is susceptible to mode hopping. Typical linewidths for injection-current semiconductor lasers are in the range of about 1–50 MHz. MQW lasers offer narrower linewidths, which can be on the order of hundreds of kilohertz [16].

5) Laser Arrays: An alternative to tunable lasers is the laser array, which contains a set of fixed-tuned lasers and whose advantage/application is explained below. A laser array consists of a number of lasers that are integrated into a single component, with each laser operating at a different wavelength. The advantage of using a laser array is that if each of the wavelengths in the array is modulated independently, then multiple transmissions may take place simultaneously. The drawback is that the number of available wavelengths in a laser array is fixed and is currently limited to about 20 wavelengths. Laser arrays with up to 21 wavelengths have been demonstrated in the laboratory [19], while a laser array with four wavelengths has actually been deployed in a network prototype [20].

C. Optical Modulation

To transmit data across an optical fiber, the information must first be encoded, or modulated, onto the laser signal. Analog techniques include amplitude modulation (AM), frequency modulation (FM), and phase modulation (PM). Digital techniques include amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK).

Of these techniques, binary ASK currently is the preferred method of digital modulation because of its simplicity. In binary ASK, also known as on–off keying (OOK), the signal is switched between two power levels. The lower power level represents a “0” bit, while the higher power level represents a “1” bit.

In systems employing OOK, modulation of the signal can be achieved by simply turning the laser on and off (direct modulation). In general, however, this can lead to chirp, or variations in the laser’s amplitude and frequency, when the laser is turned on. A preferred approach for high bit rates (>2 Gb/s) is to have an external modulator that modulates the light coming out of the laser. The external modulator blocks or passes light depending on the current applied to it. The Mach–Zehnder (MZ) interferometer, described in Section IV-B, can be used as a modulation device. A drive voltage is applied to one of two waveguides, creating an electric field that causes the signals in the two waveguides to be either in phase or 180° out of phase, resulting in the light from the laser’s being either passed through the device or blocked. MZ amplitude modulators, which offer bandwidths of up to 18 GHz, are currently available [21]. One of the advantages of using integrated-optics devices such as the MZ interferometer is that the laser and modulator can be integrated on a single structure, which may potentially be cost effective. Also, integrating the laser with the modulator eliminates the need for polarization control and results in low chirp.

D. Summary

Table 1 summarizes the characteristics of the different types of tunable transmitters. We observe that there is a tradeoff between the tuning range of a transmitter and its tuning time.

<table>
<thead>
<tr>
<th>Tunable Transmitter</th>
<th>Approx. Tuning Range (nm)</th>
<th>Tuning Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanical (external cavity)</td>
<td>500</td>
<td>1-10 ns</td>
</tr>
<tr>
<td>Acoustooptic</td>
<td>83</td>
<td>10 µs</td>
</tr>
<tr>
<td>Electrooptic</td>
<td>7</td>
<td>1-10 ns (estimated)</td>
</tr>
<tr>
<td>Injection-Current (DFB and DBR)</td>
<td>10</td>
<td>1-10 ns</td>
</tr>
</tbody>
</table>

IV. Optical Receivers and Filters

Tunable optical filter technology is key in making WDM networks realizable. Good sources of information on these devices include [1], [3], and [22].

A. Photodetectors

In receivers employing direct detection, a photodetector converts the incoming photonic stream into a stream of electrons. The electron stream is then amplified and passed through a threshold device. Whether a bit is a logical zero or one depends on whether the stream is above or below a certain threshold for a bit duration. In other words, the decision is made based on whether or not light is present during the bit duration.

The basic detection devices for direct-detection optical networks are the PN photodiode (a p-n junction) and the PIN photodiode (an intrinsic material is placed between “p-” and “n-” type material). In its simplest form, the photodiode is basically a reverse-biased p-n junction. Through the photoelectric effect, light incident on the junction will create electron-hole pairs in both the “n” and the “p” regions...
of the photodiode. The electrons released in the “p” region will cross over to the “n” region, and the holes created in the “n” region will cross over to the “p” region, thereby resulting in a current flow.

The alternative to direct detection is coherent detection, in which phase information is used in the encoding and detection of signals. Receivers based on coherent detection use a local monochromatic laser as an oscillator. The incoming optical stream is added to the signal from the oscillator, and the resulting signal is detected by a photodiode. The photodiode output is integrated over the symbol duration, and a detection threshold is used to attain the bit stream. While coherent detection is more elaborate than direct detection, it allows the reception of weak signals from a noisy background. In optical systems, however, it is difficult to maintain the phase information required for coherent detection. Since semiconductor lasers have nonzero linewidths, the transmitted signal consists of a number of frequencies with varying phases and amplitudes. The effect is that the phase of the transmitted signal experiences random but significant fluctuations around the desired phase. These phase fluctuations make it difficult to recover the original phase information from the transmitted signal, thus limiting the performance of coherent detection systems.

B. Tunable Optical Filters

This section discusses several types of tunable optical filters and their properties, while Section IV-C examines fixed-tuned optical filters. The feasibility of many local WDM networks is dependent upon the speed and range of tunable filters. Overviews of tunable filter technology can be found in [1] and [3].

1) Filter Characteristics: Tunable optical filters are characterized primarily by their tuning range and tuning time. The tuning range specifies the range of wavelengths that can be accessed by a filter. A wide tuning range allows systems to utilize a greater number of channels. The tuning time of a filter specifies the time required to tune from one wavelength to another. Fast tunable filters are required for many WDM local-area networks (LAN’s) based on broadcast-and-select architectures.

Some filters, such as the etalon (described in the following section), are further characterized by two parameters: free spectral range and finesse. In some filters, the transfer function, or the shape of the filter passband, repeats itself after a certain period. The period of such devices is referred to as the free spectral range (FSR). In other words, the filter passes every frequency that is a distance of $n \cdot \text{FSR}$ from the selected frequency, where $n$ is a positive integer. For example, in Fig. 12, if the filter is tuned to frequency $f_1$, then all frequencies labeled with a one will be passed by the filter; tuning the filter to the next frequency, $f_2$, will allow all frequencies labeled with a two to be passed by the filter; etc. The free spectral range usually depends on various physical parameters in the device, such as cavity or waveguide lengths.

The finesse of a filter is a measure of the width of the transfer function. It is the ratio of free spectral range to channel bandwidth, where the channel bandwidth is defined as the 3-dB bandwidth of a channel.

The number of channels in an optical filter is limited by the FSR and finesse. All of the channels must fit within one FSR. If the finesse is high, the transfer functions are narrower, resulting in more channels being able to fit into one FSR. With a low finesse, the channels would need to be spaced further apart to avoid cross talk, resulting in fewer channels. One approach to increasing the number of channels is to cascade filters with different FSR’s [1]. Fig. 13 shows the filter passbands for a high-resolution filter and a low-resolution filter, each with four channels within a FSR. By cascading these filters, up to 16 unique channels may be resolved.

2) The Etalon: The etalon consists of a single cavity formed by two parallel mirrors. Light from an input fiber enters the cavity and reflects a number of times between the mirrors. By adjusting the distance between the mirrors, a single wavelength can be chosen to propagate through the cavity while the remaining wavelengths destructively interfere. The distance between the mirrors may be adjusted mechanically by physically moving the mirrors or by changing the index of the material within the cavity. Many modifications (e.g., multicavity and multipass) to the etalon can be made to improve the number of resolvable channels [3]. In a multipass filter, the light passes through the same cavity multiple times, while in a multicavity filter, multiple etalons of different FSR’s are cascaded to increase the finesse effectively. An example of a mechanically tuned
etalon is the Fabry–Perot filter. In [23], it was found that the maximum number of channels for a single-cavity Fabry–Perot filter is $0.65F$, where $F$ is the finesse. A two-pass filter was found to have a maximum of $1.4 F$ channels, and two-cavity filters up to $0.44 F^2$ channels.

While the Fabry–Perot etalon can be made to access virtually the entire low-attenuation region of the fiber and can resolve very narrow passbands, it has a tuning time on the order of tens of milliseconds due to its mechanical tuning. This makes it unsuitable for many packet-switched applications in which the packet duration is much smaller than the tuning time. The Fabry–Perot etalon is used as a tunable receiver in the RAINBOW optical WDM LAN prototype [24].

3) The MZ Chain: In an MZ interferometer, a splitter splits the incoming wave into two waveguides and a combiner recombines the signals at the outputs of the waveguides (see Fig. 14). An adjustable delay element controls the optical path length in one of the waveguides, resulting in a phase difference between the two signals when they are recombined. Wavelengths for which the phase difference is $180^\circ$ are filtered out. By constructing a chain of these elements, a single desired optical wavelength can be selected.

While the MZ chain promises to be a low-cost device because it can be fabricated on semiconductor material, its tuning time is still on the order of milliseconds and its tuning control is complex, requiring that the setting of the delay element in each stage of the MZ chain be based on the settings in previous stages of the chain [2]. The high tuning time is due to thermal elements used in implementing the delay elements. Recent advances have produced a very fast ($\approx 100$ ns) tuning MZ filter, but the number of channels is limited (to eight) and the insertion loss is high [25].

4) Acoustooptic Filters: A fast tuning time is obtained when acoustooptic filters are used. Radio-frequency (RF) waves are passed through a transducer. The transducer is a piezoelectric crystal that converts sound waves to mechanical movement. The sound waves change the crystal’s index of refraction, which enables the crystal to act as a grating. Light incident upon the transducer will diffract at an angle that depends on the angle of incidence and the wavelength of the light [1]. By changing the RF waves, a single optical wavelength can be chosen to pass through the material while the rest of the wavelengths destructively interfere.

The tuning time of the acoustooptic filter is limited by the flight time of the surface acoustic wave to about $10 \, \mu s$ [3], [26]. However, the tuning range for acoustooptic filters covers the entire 1300–1560 nm spectrum [3]. This tuning range may potentially allow about 100 channels.

If more than one RF wave is passed through the grating simultaneously, more than one wavelength can be filtered out [27]. This allows the filter to be effectively tuned to several channels at the same time. However, the received signal is the superposition of all of the received wavelengths; therefore, if more than one of those channels is active, cross talk will occur. The selection of up to five wavelengths was reported in [27]. One manufacturer has introduced an acoustooptic filter [28] that has an 80-nm tuning range, 10-\( \mu \)s tuning time, and the ability to select up to ten wavelengths simultaneously [26].

One drawback of acoustooptic filters is that because of their wide transfer function, they are unable to filter out cross talk from adjacent channels if the channels are closely spaced. Therefore, the use of acoustooptic filters in a multiwavelength system places a constraint on the channel spacing, thus limiting the allowable number of channels.

5) Electrooptic Filters: Since the tuning time of the acoustooptic filter is limited by the speed of sound, crystals whose indexes of refraction can be changed by electrical currents can be used. Electrodes, which rest in the crystal, are used to supply current to the crystal. The current...
changes the crystal’s index of refraction, which allows some wavelengths to pass through while others destructively interfere [1]. Since the tuning time is limited only by the speed of electronics, tuning time can be on the order of several nanoseconds, but the tuning range (and thus, the number of resolvable channels) remains quite small at 16 nm (on the order of ten channels) [3].

6) Liquid-Crystal Fabry–Perot Filters: Liquid-crystal (LC) filters appear to be a promising new filter technology. The design of a liquid-crystal filter is similar to the design of a Fabry–Perot filter, but the cavity consists of an LC. The refractive index of the LC is modulated by an electrical current to filter out a desired wavelength, as in an electrooptic filter. Tuning time is on the order of microseconds (submicrosecond times are expected to be achievable), and tuning range is 30–40 nm [29]. These filters have low power requirements and are inexpensive to fabricate. The filter speed of LC filter technology promises to be high enough to handle high-speed packet switching in broadcast-and-select WDM networks.

7) Summary: Table 2 summarizes the state of the art in tunable receivers. As has been stated earlier, tuning range and tuning time seem inversely proportional, except for LC Fabry–Perot filters.

C. Fixed Filters

An alternative to tunable filters is to use fixed filters or grating devices. Grating devices typically filter out one or more different wavelength signals from a single fiber. Such devices may be used to implement optical multiplexers and demultiplexers or receiver arrays.

1) Grating Filters: One implementation of a fixed filter is the diffraction grating. The diffraction grating is essentially a flat layer of transparent material (e.g., glass or plastic) with a row of parallel grooves cut into it [4]. The grating separates light into its component wavelengths by reflecting light incident with the grooves at all angles. At certain angles, only one wavelength adds constructively; all others destructively interfere. This allows one to select the desired wavelength(s) by placing a filter tuned to the proper wavelength at the proper angle. Alternatively, some gratings are transmissive rather than reflective and are used in tunable lasers (see DFB lasers in Section III-B4).

An alternative implementation of a demultiplexer is the waveguide grating router (WGR), in which only one input is utilized. WGR’s will be discussed in Section VI-B1.

2) Fiber Bragg Gratings: In a fiber Bragg grating, a periodical variation of the index of refraction is directly photo-induced in the core of an optical fiber. A Bragg grating will reflect a given wavelength of light back to the source while passing the other wavelengths. Two primary characteristics of a Bragg grating are the reflectivity and the spectral bandwidth. Typical spectral bandwidths are on the order of 0.1 nm, while a reflectivity in excess of 99% is achievable [30]. While inducing a grating directly into the core of a fiber leads to low insertion loss, a drawback of Bragg gratings is that the refractive index in the grating varies with temperature, with increases in temperature resulting in longer wavelengths being reflected. An approach for compensating for temperature variations is presented in [31]. Fiber Bragg gratings may be used in the implementation of multiplexers, demultiplexers, and tunable filters.

3) Thin-Film Interference Filters: Thin-film interference filters offer another approach for filtering out one or more wavelengths from a number of wavelengths. These filters are similar to fiber Bragg grating devices with the exception that they are fabricated by depositing alternating layers of low index and high index materials onto a substrate layer. Thin-film filter technology suffers from poor thermal stability, high insertion loss, and poor spectral profile. Advances have been made, however, that address some of these issues [32].

V. OPTICAL AMPLIFIERS

Although an optical signal can propagate a long distance before it needs amplification, both long-haul and local lightwave networks can benefit from optical amplifiers. All-optical amplification may differ from optoelectronic amplification in that it may act only to boost the power of a signal, not to restore the shape or timing of the signal. This type of amplification is known as 1R (regeneration), and provides total data transparency (the amplification process is independent of the signal’s modulation format). 1R amplification is emerging as the choice for the transparent all-optical networks of tomorrow. In today’s digital networks [e.g., Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH)], however, which use the optical fiber only as a transmission medium, the optical signals are amplified by first converting the information stream into an electronic data signal and then retransmitting the signal optically. Such amplification is referred to as 3R (regeneration, reshaping, and reclocking).

The reshaping of the signal reproduces the original pulse shape of each bit, eliminating much of the noise. Reshaping applies primarily to digitally modulated signals but in some cases may also be applied to analog signals. The reclocking of the signal synchronizes the signal to its original bit timing pattern and bit rate. Reclocking applies only to digitally modulated signals. Another approach to amplification is 2R (regeneration and reshaping), in which the optical signal is converted to an electronic signal, which is then used to modulate a laser directly. The 3R and 2R techniques provide less transparency than the 1R technique, and in future optical networks, the aggregate bit rate of even just a few channels might make 3R and 2R techniques less practical.

| Table 2 Tunable Optical Filters and Their Associated Tuning Ranges and Times |
|------------------------|----------------------|------------------|
| Tunable Receiver       | Approx. Tuning Range (nm) | Tuning Time (μs) |
| Fabry-Perot            | 500                  | 1-10             |
| Acoustooptic           | 250                  | 10               |
| Electrooptic           | 16                   | 1-10             |
| LC Fabry-Perot         | 30                   | 0.5 - 10         |
Also, in a WDM system, each wavelength would need to be separated before being amplified electronically, and then recombined before being retransmitted. Thus, to eliminate the need for optical multiplexers and demultiplexers in amplifiers, optical amplifiers must boost the strength of optical signals without first converting them to electrical signals. A drawback is that noise, as well as the signal, will be amplified.

Optical amplification uses the principle of stimulated emission, similar to the approach used in a laser. The two basic types of optical amplifiers are semiconductor laser amplifiers and rare-earth-doped-fiber amplifiers, which will be discussed in the following sections. A general overview of optical amplifiers can be found in [33]. For amplifier characteristics, see Table 3.

A. Optical Amplifier Characteristics

Some basic parameters of interest in an optical amplifier are gain, gain bandwidth, gain saturation, polarization sensitivity, and amplifier noise.

Gain measures the ratio of the output power of a signal to its input power. Amplifiers are sometimes also characterized by gain efficiency, which measures the gain as a function of input power in dB/mW.

The gain bandwidth of an amplifier refers to the range of frequencies or wavelengths over which the amplifier is effective. In a network, the gain bandwidth limits the number of wavelengths available for a given channel spacing.

The gain saturation point of an amplifier is the value of output power at which the output power no longer increases with an increase in the input power. When the input power is increased beyond a certain value, the carriers (electrons) in the amplifier are unable to output any additional light energy. The saturation power is typically defined as the output power at which there is a 3-dB reduction in the ratio of output power to input power (the small-signal gain).

Polarization sensitivity refers to the dependence of the gain on the polarization of the signal. The sensitivity is measured in dB and refers to the gain difference between the TE and TM polarizations.

In optical amplifiers, the dominant source of noise is amplified spontaneous emission (ASE), which arises from the spontaneous emission of photons in the active region of the amplifier (see Fig. 15). The amount of noise generated by the amplifier depends on factors such as the amplifier gain spectrum, the noise bandwidth, and the population inversion parameter, which specifies the degree of population inversion that has been achieved between two energy levels. Amplifier noise is especially a problem when multiple amplifiers are cascaded. Each subsequent amplifier in the cascade amplifies the noise generated by previous amplifiers.

B. Semiconductor-Laser Amplifier

A semiconductor laser amplifier (see Fig. 15) consists of a modified semiconductor laser. A weak signal is sent through the active region of the semiconductor, which, via stimulated emission, results in a stronger signal’s being emitted from the semiconductor.

The two basic types of semiconductor laser amplifiers are the Fabry–Perot amplifier, which is basically a semiconductor laser, and the traveling-wave amplifier (TWA). The primary difference between the two is in the reflectivity of the end mirrors. Fabry–Perot amplifiers have a reflectivity of around 30%, while TWA’s have a reflectivity of around 0.01% [33]. In order to prevent lasing in the Fabry–Perot amplifier, the bias current is operated below the lasing threshold current. The higher reflections in the Fabry–Perot amplifier cause Fabry–Perot resonances in the amplifier, resulting in narrow passbands of around 5 GHz. This phenomenon is not very desirable for WDM systems; therefore, by reducing the reflectivity, the amplification is performed in a single pass and no resonances occur. Thus, TWA’s are more appropriate than Fabry–Perot amplifiers for WDM networks.

Today’s semiconductor amplifiers can achieve gains of 25 dB with a gain saturation of 10 dBm, polarization sensitivity of 1 dB, and bandwidth range of 40 nm [33].

Semiconductor amplifiers based on MQW’s are currently being studied. These amplifiers have higher bandwidth and higher gain saturation than bulk devices. They also provide faster on–off switching times. The disadvantage is a higher polarization sensitivity.

An advantage of semiconductor amplifiers is the ability to integrate them with other components. For example, they can be used as gate elements in switches. By turning a drive
C. Doped-Fiber Amplifier

Optical doped-fiber amplifiers are lengths of fiber doped with an element (rare earth) that can amplify light (see Fig. 16). The most common doping element is erbium, which provides gain for wavelengths of 1525–1560 nm. At the end of the length of fiber, a laser transmits a strong signal at a lower wavelength (referred to as the pump wavelength) back up the fiber. This pump signal excites the dopant atoms into a higher energy level. This allows the data signal to stimulate the excited atoms to release photons. Most erbium-doped fiber amplifiers (EDFA’s) are pumped by lasers with a wavelength of either 980 or 1480 nm. The 980-nm pump wavelength has shown gain efficiencies of around 10 dB/mW, while the 1480-nm pump wavelength provides efficiencies of around 5 dB/mW. Typical gains are on the order of 25 dB. Experimentally, EDFA’s have been shown to achieve gains of up to 51 dB with the maximum gain limited by internal Rayleigh backscattering in which some of the light energy of the signal is scattered in the fiber and directed back toward the signal source [34]. The 3-dB gain bandwidth for the EDFA is around 35 nm (see Fig. 17), and the gain saturation power is around 10 dBm [33].

For the 1300-nm region, the praseodymium-doped fluoride fiber amplifier (PDFFA) has recently been receiving attention. These amplifiers have low cross-talk and noise characteristics while attaining high gain. They are able to operate over a range of around 50 nm in the 1280–1330 nm range. In [35], a PDFFA was developed which had a 40.6 dB gain. Recent developments on PDFFA’s are presented in [36].

A limitation to optical amplification is the unequal gain spectrum of optical amplifiers. The EDFA gain spectrum is shown in Fig. 17 (from [37]). While an optical amplifier may provide gain across a range of wavelengths, it will not necessarily amplify all wavelengths equally. This characteristic—accompanied by the fact that optical amplifiers amplify noise as well as signal and the fact that the active region of the amplifier can spontaneously emit photons, which also cause noise—limits the performance of optical amplifiers. Thus, a multiwavelength optical signal passing through a series of amplifiers will eventually result in the power of the wavelengths’ being uneven.

A number of approaches to equalizing the gain of an EDFA have been studied. In [38], a notch filter (which attenuates the signal at a selected frequency) centered at around 1530 nm is used to suppress the peak in the EDFA gain (see Fig. 17). When multiple EDFA’s are cascaded, however, another peak appears around the 1560-nm wavelength. In [39], a notch filter centered at 1560 nm is used to equalize the gain for a cascade of EDFA’s. Another approach to flattening the gain is to adjust the input transmitter power such that the powers on all received wavelengths at the destination are equal [40]. A third approach to gain equalization is to demultiplex the individual wavelengths and then attenuate selected wavelengths such that all wavelengths have equal power. In [41], this approach is applied to a WDM interoffice ring network.

VI. SWITCHING ELEMENTS

Most current networks employ electronic processing and use the optical fiber only as a transmission medium. Switching and processing of data are performed by converting an optical signal back to its “native” electronic form. Such a network relies on electronic switches. These switches provide a high degree of flexibility in terms of switching and routing functions; the speed of electronics, however, is unable to match the high bandwidth of an optical fiber. Also, an electrooptic conversion at an intermediate node in the network introduces extra delay. These factors have motivated a push toward the development of all-optical networks in which optical switching components are able to switch high bandwidth optical data streams without electrooptic conversion. In a class of switching devices currently being developed, the control of the switching function is performed electronically, with the optical stream’s being transparently routed from a given input of the switch to a given output. Such transparent switching allows for the switch to be independent of the data rate and format.
of the optical signals. For WDM systems, switches that are wavelength dependent are also being developed.

Switches can be divided into two classes: relational devices and logic devices. A switch in the relational devices class establishes a relation between the inputs and the outputs. The relation is a function of the control signals applied to the device and is independent of the contents of the signal or data inputs. A property of this device is that the information entering and flowing through it cannot change or influence the current relation between the inputs and the outputs. An example of this type of device is the directional coupler as it is used in switching applications. Thus, the strength of a relational device, which allows signals at high bit rates to pass through it, is that it cannot sense the presence of individual bits that are flowing through itself. This characteristic is also known as data transparency. The same feature may sometimes also be a weakness since it causes loss of flexibility (i.e., individual portions of a data stream cannot be switched independently).

In a logic device, the data, or the information-carrying signal that is incident on the device, controls the state of the device in such a way that some Boolean function, or combination of Boolean functions, is performed on the inputs. For this class of devices, at least some of the components within an entire system must be able to change states or switch as fast as or faster than the signal bit rate [42]. This ability gives the device some added flexibility but limits the maximum bit rate that can be accommodated.

Thus, relational devices are needed for circuit switching, and logic devices are needed for packet switching.

In the following sections, we review a number of different optical switch elements and architectures.

A. Fiber Crossconnect Elements

A fiber crossconnect element switches optical signals from input ports to output ports. These type of elements are usually considered to be wavelength insensitive, i.e., incapable of demultiplexing different wavelength signals on a given input fiber.

A basic crossconnect element is the 2 × 2 crosspoint element. A 2 × 2 crosspoint element routes optical signals from two input ports to two output ports and has two states: cross state and bar state (see Fig. 18). In the cross state, the signal from the upper input port is routed to the upper output port, and the signal from the lower input port is routed to the upper output port. In the bar state, the signal from the upper input port is routed to the upper output port, and the signal from the lower input port is routed to the lower output port.

Optical crosspoint elements have been demonstrated using two types of technologies: a) the generic directive switch [43], in which light via some structure is physically directed to one of two different outputs, and b) the gate switch, in which optical amplifier gates are used to select and filter input signals to specific output ports.

1) Directive Switches: The directional coupler (see Fig. 19(a) [43]) consists of a pair of optical channel waveguides that are parallel and in close proximity over some finite interaction length. Light input to one of the waveguides couples to the second waveguide via evanescent\(^5\) coupling. The coupling strength corresponds to the interwaveguide separation and the waveguide mode size, which in turn depends upon the optical wavelength and the confinement factor\(^6\) of the waveguide. If the two waveguides are identical, complete coupling between them occurs over a characteristic length, which depends upon the coupling strength. By placing electrodes over the waveguides, however, the difference in the propagation constants in the two waveguides can be sufficiently increased via the linear electrooptic effect so that no light couples between the two waveguides. Therefore, the cross state corresponds to zero applied voltage and the bar state corresponds to a nonzero switching voltage. Unfortunately, the interaction length needs to be very accurate for good isolation, and these couplers are wavelength specific.

Switch fabrication tolerances, as well as the ability to achieve good switching for a relatively wide range of wavelengths, can be overcome by using the so-called reversed delta-beta coupler [see Fig. 19(b)]. In this device, the electrode is split into at least two sections. The cross state is achieved by applying equal and opposite voltages to the two electrodes. This approach has been shown to be very successful [43].

The balanced bridge interferometric switch [see Fig. 19(c)] consists of an input 3-dB coupler, two waveguides sufficiently separated so that they do not couple, electrodes to allow changing of the effective path length over the two arms, and a final 3-dB coupler. Light incident on the upper waveguide is split in half by the first coupler. With no voltage applied to the electrodes, the optical path length of the two arms enters the second coupler in phase. The second coupler acts like the continuation of the first, and all the light is crossed over to the second waveguide to provide the cross state. To achieve the bar state, voltage is applied to an electrode, placed over one of the interferometer arms to produce electrooptically a 180° phase difference between the two arms. In this case, the two inputs from the arms of the interferometer combine at the second 3-dB coupler out of phase, with the result that light remains in the upper waveguide.

The intersecting waveguide switch is shown in Fig. 19(d). This device can be viewed as a directional coupler [see Fig. 19(a)] with no gap between the

5 An evanescent wave is the part of a propagating wave that travels along or outside of the waveguide boundary.

6 The confinement factor determines the fraction of power that travels within the core of the waveguide.
waveguides in the interaction region. When properly fabricated, both cross and bar states can be electrooptically achieved with good cross-talk performance.

Other types of switches include the mechanical fiber-optic switch and the thermooptic switch. These devices offer slow switching (about milliseconds) and may be employed in circuit-switched networks. One mechanical switch, for example, consists of two ferrules, each with polished end faces that can rotate to switch the light appropriately [44]. Thermooptic waveguide switches, on the other hand, are fabricated on a glass substrate and are operated by the use of the thermooptic effect. One such device uses a zero-gap directional-coupler configuration with a heater electrode to increase the waveguide index of refraction [45].

2) Gate Switches: In the $N \times N$ gate switch, each input signal first passes through a $1 \times N$ splitter. The signals then pass through an array of $N^2$ gate elements and then are recombined in $N \times 1$ combiners and sent to the $N$ outputs. The gate elements can be implemented using optical amplifiers, which can be turned either on or off to pass only selected signals to the outputs. The amplifier gains can compensate for coupling losses and losses incurred at the splitters and combiners. A $2 \times 2$ amplifier gate switch is illustrated in Fig. 20. A disadvantage of the gate switch is that the splitting and combining losses limit the size of the switch.

Amplifier gate switches of size $8 \times 8$ are commercially available. In [46], an $8 \times 8$ switch is described that uses semiconductor optical amplifiers to provide lossless switching. It operates around the 1300-nm region, has an optical bandwidth of 40 nm, has low polarization dependence (1 dB), and has a fairly low cross talk (below $-40$ dB). The disadvantages are that the switch is bulky (weight of 50 lb) and expensive.

A possible, less expensive alternative is the integrated-optics amplifier gate switch. Such switches of size $4 \times 4$ have been demonstrated experimentally to operate around the 1550-nm region and have a fiber-to-fiber gain of 5 dB with cross talk levels below $-40$ dB. The main disadvantage, however, is the high polarization sensitivity (6–12 dB).

B. Nonreconfigurable Wavelength Router

A wavelength-routing device can route signals arriving at different input fibers (ports) of the device to different output fibers (ports) based on the wavelengths of the signals. Wavelength routing is accomplished by demultiplexing the different wavelengths from each input port, optionally switching each wavelength separately, and then multiplexing signals at each output port. The device can be either nonreconfigurable, in which case there is no switching stage between the demultiplexers and the multiplexers, and the routes for different signals arriving at any input port are fixed (these devices are referred to as routers rather than switches), or reconfigurable, in which case the routing function of the switch can be controlled electronically. In this section, we will discuss wavelength routers, while Section VI-C will cover reconfigurable wavelength switches.

A nonreconfigurable wavelength router can be constructed with a stage of demultiplexers, which separate each of the wavelengths on an incoming fiber, followed by a stage of multiplexers, which recombine wavelengths from various inputs to a single output. The outputs of the demultiplexers are hardwired to the inputs of the multiplexers. Let this router have $P$ incoming fibers and $P$ outgoing fibers. On each incoming fiber, there are $M$ wavelength channels. A $4 \times 4$ nonreconfigurable wavelength router with $M = 4$ is illustrated in Fig. 21. The router is nonreconfigurable because the path of a given wavelength channel is fixed after it enters the router on a particular input fiber. The wavelengths on each incoming fiber are separated using a grating demultiplexer. Last, information from multiple WDM channels is multiplexed before being launched back onto an output fiber. In between the demultiplexers and multiplexers, there are
direct connections from each demultiplexer output to each multiplexer input. Which wavelength on which input port gets routed to which output port depends on a “routing matrix” characterizing the router; this matrix is determined by the internal “connections” between the demultiplexers and multiplexers.

1) Waveguide Grating Routers: One implementation of a wavelength router is the waveguide grating router (WGR), which is also referred to as an arrayed waveguide grating (AWG) multiplexer. A WGR provides a fixed routing of an optical signal from a given input port to a given output port based on the wavelength of the signal. Signals of different wavelengths coming into an input port will each be routed to a different output port. Also, different signals using the same wavelength can be input simultaneously to different input ports and still not interfere with each other at the output ports. Compared to a passive-star coupler, in which a given wavelength may only be used on a single input port, the WGR with input and output ports is capable of routing a maximum of \( N^2 \) connections, as opposed to a maximum of \( N \) connections in the passive-star coupler. Also, because the WGR is an integrated device, it can easily be fabricated at low cost. The disadvantage of the WGR is that it is a device with a fixed routing matrix that cannot be reconfigured.

The WGR, shown in Fig. 22, can be used as a non-reconfigurable wavelength router or to build a tunable optical transmitter or a tunable optical receiver. It consists of two passive-star couplers connected by a grating array. The first star coupler has \( N \) inputs and \( N \) outputs (where \( N \ll N' \)), while the second one has \( N' \) inputs and \( N \) outputs. The inputs to the first star are separated by an angular distance of \( \alpha \), and their outputs are separated by an angular distance of \( \alpha' \). The grating array consists of \( N' \) waveguides, with lengths \( l_1, l_2, \ldots, l_{N'} \), where \( l_1 < l_2 < \cdots < l_{N'} \). The difference in length between any two adjacent waveguides is a constant \( \Delta l \).

In the first star coupler, a signal on a given wavelength entering from any of the input ports is split and transmitted to its \( N' \) outputs, which are also the \( N \) inputs of the grating array. The signal travels through the grating array, experiencing a different phase shift in each waveguide depending on the length of the waveguides and the wavelength of the signal. The constant difference in the lengths of the waveguides creates a phase difference of \( \beta \cdot \Delta l \) in adjacent waveguides, where \( \beta = \frac{2\pi n_{eff}}{\lambda} \) is the propagation constant in the waveguide, \( n_{eff} \) is the effective refractive index of the waveguide, and \( \lambda \) is the wavelength of the light. At the input of the second star coupler, the phase difference in the signal will be such that the signal will constructively recombine only at a single output port.

Two signals of the same wavelength coming from two different input ports will not interfere with each other in the grating because there is an additional phase difference created by the distance between the two input ports. The two signals will be combined in the grating but will be separated out again in the second star coupler and directed to different outputs. This phase difference is given by \( kR(p - q)\alpha' \), where \( k \) is a propagation constant that doesn’t depend on wavelength, \( R \) is the constant distance between the two foci of the optical star, \( p \) is the input port number of the router, and \( q \) is the output port number of the router. The total phase difference is

\[
\phi = \frac{2\pi \cdot \Delta l}{\lambda} + kR(p - q)\alpha'.
\]

The transmission power from a particular input port \( p \) to a particular output port \( q \) is maximized when \( \phi \) is an integer multiple of \( 2\pi \). Thus, only wavelengths \( \lambda \) for which \( \phi \) is a multiple of \( 2\pi \) will be transmitted from input port \( p \) to output port \( q \). Alternately, for a given input port and a given wavelength, the signal will only be transmitted to the output port that causes \( \phi \) to be a multiple of \( 2\pi \).

Prototype devices have been built on silicon with \( N = 11 \) and \( N' = 11 \), giving a channel spacing of 16.5 nm, and \( N = 7 \) and \( N' = 15 \), giving a channel spacing of 23.1 nm [48]. In [49], a 15 \( \times \) 15 waveguide grating multiplexer on InP is demonstrated to have a free spectral range of 10.5 nm and channel spacing of 0.7 nm in the 1550-nm region. In [50], a 64 \( \times \) 64 arrayed-waveguide multiplexer on silicon is demonstrated with a channel spacing of 0.4 nm. WGR’s with flat passbands also have been developed [51], [52]. Other applications of the WGR, such as tunable transmitters and tunable receivers, are presented in [53]. These tunable components are implemented by integrating the WGR with switched amplifier elements. An amplifier element may either be activated, in which case it amplifies the signal passing through it, or turned off, in which case it prevents any signals from passing through it. By using only a single input port of the WGR, each wavelength on that input port will be routed to a different output port. By placing an amplifier element at each output port of the WGR, we may filter out selected wavelengths by activating or deactivating the appropriate amplifiers. The outputs of the amplifier elements may then be multiplexed into a signal containing only the desired wavelengths.
C. Reconfigurable Wavelength-Routing Switch

A reconfigurable wavelength-routing switch (WRS), also referred to as a wavelength-selective cross connect (WSXC), uses photonic switches inside the routing element. The functionality of the reconfigurable WRS, illustrated in Fig. 23, is as follows. The WRS has $P$ incoming fibers and $P$ outgoing fibers. On each incoming fiber, there are $M$ wavelength channels. Similar to the nonreconfigurable router, the wavelengths on each incoming fiber are separated using a grating demultiplexer. The outputs of the demultiplexers are directed to an array of optical switches between the demultiplexer and multiplexer stages. All signals on a given wavelength are directed to the same switch. The switched signals are then directed to multiplexers which are associated with the output ports. Last, information from multiple WDM channels is multiplexed before being launched back onto an output fiber.

The outputs of the demultiplexers are directed to an array of $M \times P \times P$ optical switches between the demultiplexer and multiplexer stages. All signals on a given wavelength are directed to the same switch. The switched signals are then directed to multiplexers which are associated with the output ports. Last, information from multiple WDM channels is multiplexed before being launched back onto an output fiber.

Space-division optical-routing switches may be built from $2 \times 2$ optical crosspoint elements [54] arranged in a banyan-based fabric. The space division switches (which may be one per wavelength [55] can route a signal from any input to any output on a given wavelength. Such switches based on relational devices [42] are capable of switching very-high-capacity signals. The $2 \times 2$ cross-point elements that are used to build the space-division switches may be slowly tunable and may be reconfigured to adapt to changing traffic requirements. Switches of this type can be constructed from off-the-shelf components available today.

Networks built from such switches are more flexible than passive, nonreconfigurable, wavelength-routed networks because they provide additional control in setting up connections. The routing is a function of both the wavelength chosen at the source node and the configuration of the switches in the network nodes.

D. Photonic Packet Switches

Most of the switches discussed above are relational devices, i.e., they are useful in a circuit-switched environment where a connection may be set up over long periods of time. Here, we review optical packet switches that have been proposed in the literature. These switches are composed of logic devices instead of relational devices so that the switch configuration is a function of the data on the input signal.

In a packet-switched system, there exists the problem of resource contention when multiple packets contend for a
common resource in the switch. In an electronic system, contention may be resolved through the use of buffering. In the optical domain, however, contention resolution is a more complex issue since it is difficult to implement components that can store optical data. A number of switch architectures that use delay lines to implement optical buffering have been proposed [56]–[59], and a sampling of these approaches will be reviewed below. A delay line is simply a long length of fiber that introduces propagation delays that are on the order of packet transmission times. A review of photonic packet switches is given in [60].

1) The Staggering Switch: The staggering switch, which is an “almost-all-optical” packet switch, has been proposed in [56]. In an “almost-all-optical” network, the data path is fully optical but the control of the switching operation is performed electronically. One of the advantages of such switching over its electronic counterpart is that it is transparent, i.e., except for the control information, the payload may be encoded in an arbitrary format or at an arbitrary bit rate. The main problem in the implementation of packet-switched optical networks is the lack of random-access optical memory.

The staggering switch architecture is based on an output-collision-resolution scheme that is controlled by a set of delay lines with unequal delays. The architecture is based on two rearrangeably nonblocking stages interconnected by optical delay lines with different amounts of delay. The work in [56] investigates the probability of packet loss and the switch latency as a function of link utilization and switch size. In general, with proper setting of the number of delay lines, the switch can achieve an arbitrary low probability of packet loss. Fig. 24 gives a simple overview of the switch architecture. The staggering switch has yet to be implemented, and it may suffer from cross-talk and power budget limitations.

2) Contention Resolution by Delay Lines (CORD): Another architecture that deals with contention in a packet-switched optical network is the CORD [57]. The CORD architecture consists of a number of $2 \times 2$ cross-connect elements and delay lines (see Fig. 25). Each delay line functions as a buffer for a single packet. If two packets contend for the same output port, one packet may be switched to a delay line while the other packet is switched to the proper output. The packet that was delayed can then be switched to the same output after the first packet has been transmitted.

3) The HLAN Architecture: The HLAN architecture, described in [58], avoids channel collisions by using a slotted system with empty slot markers that indicate when a node can write data into a slot. HLAN is implemented on a helical unidirectional bus in which the physical fiber wraps around twice to visit each node three times, and the fiber medium is considered to be divided into three nonoverlapping “segments” (see Fig. 26). A headend periodically generates equal-sized frames of empty slots and puts them on the bus. The guaranteed-bandwidth traffic is transmitted on the GBW segment, the bandwidth-on-demand traffic is transmitted on the BOD segment, and data is received on the RCV segment. All users receive traffic on the third segment, i.e., all receivers are downstream of the transmitters. Each node in the network is equipped with a header/slot processor, protocol logic units, clock recovery mechanisms, and optoelectronic buffers. Though
the network has been designed to operate at 100 Gb/s, it is scalable in principle to faster media rates.

VII. WAVELENGTH CONVERSION

Consider Fig. 27, which shows a wavelength-routed network containing two WDM cross connects (S1 and S2) and five access stations (A–E). Three lightpaths have been set up (C–A on wavelength $\lambda_1$, C–B on $\lambda_2$, and D–E on $\lambda_3$). To establish a lightpath, we require that the same wavelength be allocated on all the links in the path. This requirement is known as the wavelength-continuity constraint (e.g., see [61]). This constraint distinguishes the wavelength-routed network from a circuit-switched network, which blocks calls only when there is no capacity along any of the links in the path assigned to the call. Consider the example in Fig. 28(a). Two lightpaths have been established in the network: 1) between Node 1 and Node 2 on wavelength $\lambda_1$ and 2) between Node 2 and Node 3 on wavelength $\lambda_2$. Now, suppose a lightpath between Node 1 and Node 3 needs to be set up. Establishing such a lightpath is impossible, even though there is a free wavelength on each of the links along the path from Node 1 to Node 3. This is because the available wavelengths on the two links are different. Thus, a wavelength-routed network may suffer from higher blocking as compared to a circuit-switched network.

It would be easy to eliminate the wavelength-continuity constraint if we were able to convert the data arriving on one wavelength along a link into another wavelength at an intermediate node and forward it along the next link. Such a technique is actually feasible and is referred to as wavelength conversion. In Fig. 28(b), a wavelength converter at Node 2 is employed to convert data from wavelength $\lambda_2$ to $\lambda_1$. The new lightpath between Node 1 and Node 3 can now be established by using the wavelength $\lambda_2$ on the link from Node 1 to Node 2, and then by using the wavelength $\lambda_1$ to reach Node 3 from Node 2. Notice that a single lightpath in such a wavelength-convertible network can use a different wavelength along each of the links in its path. Thus, wavelength conversion may improve the efficiency in the network by resolving the wavelength conflicts of the lightpaths. The impact of wavelength conversion on WDM wide-area network (WAN) design is further elaborated in Section VIII-F.

The function of a wavelength converter is to convert data on an input wavelength onto a possibly different output wavelength among the $N$ wavelengths in the system (see Fig. 29). In this figure and throughout this section, $\lambda_m$ denotes the input signal wavelength, $\lambda_c$ is the converted wavelength, $\lambda_p$ is the pump wavelength, $f_s$ is the input frequency, $f_c$ is the converted frequency, $f_p$ is the pump frequency, and CW is the continuous wave (unmodulated) generated as the pump signal.

An ideal wavelength converter should possess the following characteristics [62]:

- transparency to bit rates and signal formats;
- fast setup time of output wavelength;
- conversion to both shorter and longer wavelengths;
• moderate input power levels;
• possibility for same input and output wavelengths (i.e., no conversion);
• insensitivity to input signal polarization;
• low-chirp output signal with high extinction ratio\(^7\) and large signal-to-noise ratio (SNR);
• simple implementation.

A. Wavelength Conversion Technologies

Several researchers have attempted to classify and compare the several techniques available for wavelength conversion [62]–[66]. The classification of these techniques presented in this section follows that in [65]. Wavelength conversion techniques can be broadly classified into two types: optoelectronic wavelength conversion, in which the optical signal must first be converted into an electronic signal, and all-optical wavelength conversion, in which the signal remains in the optical domain. All-optical conversion techniques may be subdivided into techniques that employ coherent effects and techniques that use cross modulation.

1) Optoelectronic Wavelength Conversion: In optoelectronic wavelength conversion [67], the optical signal to be converted is first translated into the electronic domain using a photodetector (labeled “R” in Fig. 30 from [68]). The electronic bit stream is stored in the buffer (labeled FIFO for the first-in-first-out queue mechanism). The electronic signal is then used to drive the input of a tunable laser (labeled “T”) tuned to the desired wavelength of the output (see Fig. 30). This method has been demonstrated for bit rates up to 10 Gb/s [66]. This method, however, is much more complex and consumes a lot more power than the other methods described below [62]. Moreover, the process of optoelectronic (O/E) conversion adversely affects the transparency of the signal, requiring the optical data to be in a specified modulation format and at a specific bit rate. All information in the form of phase, frequency, and analog amplitude of the optical signal is lost during the conversion process. One form of transparency is digital transparency, in which digital signals of any bit rate up to a certain limit can be accommodated by the system [66].

2) Wavelength Conversion Using Coherent Effects: Wavelength conversion methods using coherent effects are typically based on wave-mixing effects (see Fig. 31). Wave-mixing arises from a nonlinear optical response of a medium when more than one wave is present. It results in the generation of another wave whose intensity is proportional to the product of the interacting wave intensities. Wave-mixing preserves both phase and amplitude information, offering strict transparency. It is also the only approach that allows simultaneous conversion of a set of multiple input wavelengths to another set of multiple output wavelengths and could potentially accommodate a signal with bit rates exceeding 100 Gb/s [66]. In Fig. 31, the value \( n = 3 \) corresponds to four-wave mixing (FWM) and \( n = 2 \) corresponds to difference frequency generation (DFG). These techniques are described below.

\( a)\) FWM: FWM (also referred to as four-photon mixing) is a third-order nonlinearity in silica fibers, which causes three optical waves of frequencies \( f_i, f_j, \) and \( f_k \) \((k \neq i, j)\) to interact in a multichannel WDM system [69] to generate a fourth wave of frequency given by

\[
f_c = n f_p + f_j - f_i.k.
\]

FWM is also achievable in other passive waveguides, such as semiconductor waveguides, and in an active medium, such as a semiconductor optical amplifier (SOA). This technique provides modulation-format independence [70] and high bit-rate capabilities [71]. The conversion efficiency from pump energy to signal energy of this technique is not very high, however, and it decreases swiftly with increasing conversion span (shift between pump and output signal wavelengths) [72].

\( b)\) DFG: DFG is a consequence of a second-order nonlinear interaction of a medium with two optical waves: a pump wave and a signal wave [66]. DFG is free from satellite signals, which appear in FWM-based techniques. This technique offers a full range of transparency without adding excess noise to the signal [73]. It is also bidirectional and fast, but it suffers from low efficiency and high polarization sensitivity. The main difficulties in implementing this technique lie in the phase matching of interacting waves [74] and in fabricating a low-loss waveguide for high conversion efficiency [66].

---

\(^7\)The extinction ratio is defined as the ratio of the optical power transmitted for a bit “0” to the power transmitted for a bit “1.”
3) Wavelength Conversion Using Cross Modulation: Cross-modulation wavelength-conversion techniques utilize active semiconductor optical devices such as SOA’s and lasers. These techniques belong to a class known as optical-gating wavelength conversion [66].

a) SOA in cross-gain modulation (XGM) and XPM mode: The principle behind using an SOA in the XGM mode is shown in Fig. 32 (from [62]). The intensity-modulated input signal modulates the gain in the SOA due to gain saturation. A CW signal at the desired output wavelength ($\lambda_c$) is modulated by the gain variation so that it carries the same information as the original input signal. The CW signal can be launched into the SOA either in the same direction as the input signal (codirectional) or in the opposite direction (counterdirectional). The XGM scheme gives a wavelength-converted signal that is inverted compared to the input signal. While the XGM scheme is simple to realize and offers penalty-free conversion at 10 Gb/s [62], it suffers from drawbacks due to inversion of the converted bit stream and extinction ratio degradation for the converted signal.

The operation of a wavelength converter using SOA in XPM mode is based on the fact that the refractive index of the SOA is dependent on the carrier density in its active region. An incoming signal that depletes the carrier density will modulate the refractive index and thereby result in phase modulation of a CW signal (wavelength $\lambda_c$) coupled into the converter [62], [75]. The SOA can be integrated into an interferometer so that an intensity-modulated signal format results at the output of the converter. Techniques involving SOA’s in XPM mode have been proposed using nonlinear optical loop mirrors (NOLM’s) [76], MZ interferometers (MZI’s) [77] and Michelson interferometers (MI’s) [78]. Fig. 33 shows an asymmetric MZI wavelength converter based on SOA in XPM mode (from [62]). With the XPM scheme, the converted output signal can be either inverted or noninverted, unlike in the XGM scheme, where the output is always inverted. The XPM scheme is also very power efficient compared to the XGM scheme [62].

b) Semiconductor lasers: Using single-mode semiconductor lasers, lasing mode intensity is modulated by input signal light through lasing-mode gain saturation. The obtained output (converted) signal is inverted compared to the input signal. This gain suppression mechanism has been employed in a DBR laser to convert signals at 10 Gb/s [79].

In the method using saturable absorption in lasers (e.g., [80]), the input signal saturates the absorption of carrier transitions near the band gap and allows the probe beam to transmit (see Fig. 34). This technique shows a bandwidth limit of 1 GHz due to carrier recombinations [66].

4) Summary: In this subsection, we reviewed the various techniques and technologies used in the design of a wavelength converter. The actual choice of the technology to be employed for wavelength conversion in a network depends on the requirements of the particular system. It is clear, however, that optoelectronic converters offer only limited digital transparency. Moreover, deploying multiple optoelectronic converters in a WDM cross connect, e.g., in a WRS, requires sophisticated packaging to avoid cross talk among channels. This leads to an increased cost per converter, making this technology even less attractive than all-optical converters [66]. Other disadvantages of optoelectronic converters include complexity and large power consumption [62]. Among all-optical converters, converters based on SOA’s using the XGM and the XPM conversion scheme presently seem well suited for system use. Converters based on FWM, though transparent to different modulation formats, perform inefficiently [62]. Wave-mixing converters, however, are the only category of wavelength converters that offer the full range of transparency while also allowing simultaneous conversion of a set of input wavelengths to another set of output wavelengths. In this respect, DFG-based methods offer great promise. Further
Fig. 34. Conversion using saturable absorption in a laser.

Fig. 35. A switch that has dedicated converters at each output port for each wavelength.

details on comparisons of various wavelength-conversion techniques can be found in [63], [65], and [66].

In Section VII-B, we examine various switch architectures that may be employed in a wavelength-convertible network.

B.P Wavelength Conversion in Switches

As wavelength converters become readily available, a vital question comes to mind: Where do we place them in the network? An obvious location is in the switches (cross connects) in the network. A possible architecture of such a wavelength-convertible switching node is the dedicated wavelength-convertible switch (see Fig. 35, from [81]). In this architecture, referred to as a wavelength-interchanging cross connect (WXSC), each wavelength along each output link in a switch has a dedicated wavelength converter (shown as boxes labeled WC in Fig. 35), i.e., an $M \times N$ switch in an $N$-wavelength system requires $M \times N$ converters. The incoming optical signal from a link at the switch is first wavelength demultiplexed into separate wavelengths. Each wavelength is switched to the desired output port by the nonblocking optical switch. The output signal may have its wavelength changed by its wavelength converter. Last, various wavelengths combine to form an aggregate signal coupled to an outbound fiber. The switch architecture shown in Fig. 35 is similar to that of the reconfigurable WRS (also WSXC) shown in Fig. 23, with additional wavelength converters added after the switching elements.

The dedicated wavelength-convertible switch is not very cost efficient, however, since all the wavelength converters may not be required all the time [82]. An effective method to cut costs is to share the converters. Two architectures have been proposed for switches sharing converters [81]. In the share-per-node structure [see Fig. 36(a)], all the converters at the switching node are collected in a converter bank. A converter bank is a collection of a few wavelength converters (e.g., two in each of the boxes labeled WC in Fig. 36), each of which is assumed to have identical characteristics and can convert any input wavelength to any output wavelength. This bank can be accessed by any wavelength on any incoming fiber by appropriately configuring the larger optical switch in Fig. 36(a). In this architecture, only the wavelengths that require conversion are directed to the converter bank. The converted wavelengths
are then switched to the appropriate outbound fiber link by the second optical switch. In the share-per-link structure [see Fig. 36(b)], each outgoing fiber link is provided with a dedicated converter bank, which can be accessed only by those lightpaths traveling on that particular outbound link. The optical switch can be configured appropriately to direct wavelengths toward a particular link, either with conversion or without conversion.

When optoelectronic wavelength conversion is used, the functionality of the wavelength converter can also be performed at the access stations instead of at the switches. The share-with-local switch architecture proposed in [81] (see Fig. 37) and the simplified network access station architecture proposed in [83] (see Fig. 38) fall under this category. In the share-with-local switch architecture, selected incoming optical signals are converted to electrical signals by a receiver bank. A signal can then be either dropped locally or retransmitted on a different wavelength by a transmitter bank. In Fig. 38, an optical signal on wavelength W1 can be switched to a network access station, where it is converted to an electronic signal. The signal can then be retransmitted by the network access station on a new wavelength W2.

VIII. DESIGNING WDM NETWORKS: SYSTEMS CONSIDERATIONS

In designing a WDM network, it is important to keep in mind not only the desired functionality of the network but also the capabilities and limitations of available optical network components. In this section, we present some of the issues involved in designing optical networks, describe some of the physical constraints that must be considered, and discuss how various optical components may be used to satisfy networking requirements.

A. Channels

An important factor to consider in the design of a WDM network is the number of wavelengths to use. In some cases, it may be desirable to design the network with the maximum number of channels attainable with the current device technology, subject to tuning time requirements and cost constraints. Another approach is to assign a different wavelength to each node, although this type of network does not scale very well. In WAN’s, the objective often is to minimize the number of wavelengths for a desired network topology or traffic pattern. In any case, the maximum
number of wavelengths is limited by the optical device technology. The number of channels is affected primarily by the total available bandwidth or spectral range of the components and the channel spacing.

The bandwidth of the fiber medium, as mentioned in Section II, is limited to the low-attenuation regions around 1300 and 1550 nm. These regions have bandwidths of approximately 200 nm (25 THz) each. Optical networks may not necessarily be able to take advantage of this entire range, however, due to the bandwidth limitations of optical components. Amplifiers have an optical bandwidth of around 35–40 nm and injection-current tunable lasers have a tuning range of around 10 nm, while the tuning range of tunable receivers varies from the entire low-attenuation region of fiber for Fabry–Perot filters to around 16 nm for electrooptic filters.

Some factors that affect channel spacing are the channel bit rates, optical power budget, nonlinearities in the fiber, and resolution of transmitters and receivers. We now illustrate how some of these parameters may relate to the maximum number of channels in a WDM system. We will assume that tunable transmitters and receivers are being used and that we would like to design a WDM passive-star-based network for \( N \) nodes.

Let \( \Delta T \) be the tuning range of the transmitters and let \( \Delta R \) be the tuning range of the receivers (both are measured in nm). The available transceiver bandwidth, \( \text{BW}_T \) is given by the frequency range in which the transmitter tuning range intersects with the receiver tuning range.

Using the identity
\[
\Delta f = \frac{c \Delta \lambda}{\lambda^2}
\]
the frequency needed for \( \text{BW}_T \) is
\[
\Delta f = \frac{c \text{BW}_T}{\lambda^2}
\]
where \( \lambda \) is the center wavelength.

If we want each channel to have a bit rate of \( B \) Gb/s, \( 2B \) GHz of bandwidth will be needed for encoding. According to [3], a channel spacing of at least six times the channel bit rate is needed to minimize cross talk on a WDM system. Thus, if we want \( W \) channels, we need
\[
2B \cdot W + 6B(W - 1) \text{ GHz}.
\]

Thus, the maximum number of resolvable channels for this network is
\[
W = \frac{\Delta f + 6B}{8B}.
\]

Although a maximum of \( W \) channels may be accommodated, in some cases it may be desirable to use fewer than \( W \) channels, e.g., in a shared-channel WDM optical LAN [84], [85]. A higher number of channels may provide more network capacity but it also results in higher network costs and, in some cases, it may require more complex protocols.

### B. Power Considerations

In any network, it is important to maintain adequate SNR in order to ensure reliable detection at the receiver. In a WDM network, signal power can degrade due to losses such as attenuation in the fiber, splitting losses, and coupling losses. Some of the losses may be counteracted through the use of optical amplifiers, and an important consideration in designing a WDM network is the design and appropriate placement of amplifiers.

There are three main applications for optical amplifiers in a lightwave network [1]. The first application is as a transmission power booster, which is placed immediately after the transmitter to provide a high power signal to the network. This allows the signal to undergo splitting at couplers or to travel longer distances. The second application is as a receiver preamplifier, which boosts the power of a signal before detection at a receiver photodetector. The third application is as an in-line amplifier, which is used within the network to boost degraded signals for further propagation. Each of these situations requires the amplifier to have different characteristics. A discussion of the requirements and design of multistage EDFA’s for various applications is given in [86]. Table 4 summarizes some of these requirements.

For in-line amplifier applications, there is the additional issue of amplifier placement. Amplifiers need to be placed strategically throughout the network in a way that guarantees that all signals are adequately amplified while minimizing the total number of amplifiers being used. A study of this problem in LAN’s is reported in [87].

When utilizing cascades of in-line amplifiers, one must also consider issues such as ASE noise introduced by the amplifiers, and the unequal gain spectrum of the amplifiers. The accumulation of ASE noise in a cascade of amplifiers may seriously degrade the SNR. If the input signal power is too low, ASE noise may cause the SNR to fall below detectable levels; if the signal power is too high, however, the signal combined with the ASE noise may saturate the amplifiers. The unequal gain spectrum of the EDFA places limitations on the usable bandwidth in WDM systems. When multiple EDFA’s are cascaded, the resulting gain bandwidth may be significantly reduced from the gain bandwidth of a single EDFA. An initial bandwidth of 30 nm can potentially be reduced to less than 10 nm after a cascade of 50 EDFA’s [88].

Although recent developments in amplifier technology have solved many of the power-loss and noise problems

<table>
<thead>
<tr>
<th>EDFA Application</th>
<th>Gain (dB)</th>
<th>Noise Figure (dB)</th>
<th>Saturated Output (dBm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmitter power booster</td>
<td>MODERATE</td>
<td>MODERATE</td>
<td>HIGH</td>
</tr>
<tr>
<td>Preamplifier</td>
<td>HIGH</td>
<td>LOW</td>
<td>MODERATE</td>
</tr>
<tr>
<td>In-line amplifier</td>
<td>HIGH</td>
<td>LOW</td>
<td>HIGH</td>
</tr>
</tbody>
</table>
in optical networks, network designers should not rely solely on amplifiers for resolving power issues but should also consider other options. For example, to avoid splitting losses in network interconnections, it might be worthwhile to consider using wavelength-routing devices, such as the wavelength-routing switch or the waveguide grating router, instead of wavelength-independent devices, such as the amplifier gate switch or the PSC.

C. Cross Talk

Cross talk may be caused by either interference from signals on different wavelengths (interband cross talk) or interference from signals on the same wavelength on another fiber (intraband cross talk) [89]. Interband cross talk must be considered when determining channel spacing. In some cases, interband cross talk may be removed through the use of appropriate narrow-band filters. Intraband cross talk usually occurs in switching nodes, where multiple signals on the same wavelength are being switched from different inputs to different outputs. This form of cross talk is more of a concern than interband cross talk because intraband cross talk cannot easily be removed through filtering and may accumulate over a number of nodes. The degree of intraband cross talk depends in part on the switch and node architectures, and one approach for reducing intraband cross talk is to introduce additional wavelength selective filters within a node [89].

D. Additional Considerations

Other device considerations in the design of WDM networks include dispersion and architectural issues.

As mentioned in Section II-D, dispersion in an optical communications system causes a pulse to broaden as it propagates along the fiber. The pulse broadening limits the spacing between bits and thus limits the maximum transmission rate for a given propagation distance. Alternatively, it limits the maximum fiber distance for a given bit rate.

Apart from the device considerations mentioned above, there are architectural considerations in designing a WDM network. The topology of the physical optical fiber buried in the ground may influence the choice of which transmitter-receiver pairs to operate on which wavelengths. The need for fault tolerance and reliability affects the choice of the network architectures. Moreover, the emerging standards on optical wavelengths and channel spacing (e.g., ITU-T, MONET) will influence the design of the network components.

E. Elements of Local-Area WDM Network Design

A local-area WDM network typically will consist of a number of nodes that are connected via two-way optical fibers either to some physical network medium or directly to other nodes. In this section, we will present some of the issues involved in selecting the hardware for both the network medium and the nodes.

1) Network Medium: In some cases, such as in a network with a physical ring or bus topology, the network medium simply consists of optical fiber. In other cases, the network medium may consist of a number of optical components such as couplers, switches, and routers.

The simplest and most popular interconnection device for a local-area WDM network is the PSC, which provides a broadcast medium (see Fig. 39). The broadcast capability of the star coupler combined with multiple WDM channels allows for a wide range of possible media-access protocols [90], [91]. Also, since the star coupler is a passive device, it is fairly reliable. The drawback of having a passive network medium is that the network nodes may be required to handle additional processing and may require additional hardware in order to route and schedule transmissions. The broadcast capability of the star coupler also prevents the reuse of wavelengths to create more simultaneous connections.

2) Network Nodes: Another important consideration in the design of a WDM network is the hardware at each node. Each node in a network typically consists of a workstation connected to the network medium via optical fiber, and the node may potentially access any of the available wavelength channels on each fiber. In designing the network interface for each node, one must select the number of transmitters and receivers as well as the type of transmitters and receivers—fixed or tunable—to place at each node. These decisions usually depend on the protocol, degree of access, and connectivity desired in the network.

A WDM network protocol may be either a single-hop protocol [90], in which communication takes place directly between two nodes without being routed through intermediate nodes, or a multihop protocol [91], in which information...
from a source node to a destination node may traverse intermediate nodes in the network. In general, multihop networks require less tuning than single-hop networks.

At a minimum, each node must be equipped with at least one transmitter and one receiver. When both the transmitters and the receivers are fixed tuned to certain wavelength channels, and there is more than one channel, then a static multihop topology must be established over the PSC. The topology is created by establishing connections between pairs of nodes on given wavelengths. An overview of multihop protocols and topologies is provided in [91].

A more flexible approach would be to use either a tunable transmitter and/or a tunable receiver. The tunability allows the network to be dynamically reconfigured based on traffic patterns, and it also allows the implementation of single-hop protocols. A number of single-hop WDM protocols based on nodes with tunable transmitters and/or tunable receivers are presented in [90]. Additional transmitters and receivers at each node may help to increase the connectivity of the network and may also be used to help coordinate transmissions. In some cases, the network may have a control channel, which may be used for pretransmission coordination (which allows a node to preannounce its transmission so that the receiving node may appropriately tune its receiver). Each node may then be equipped with an additional fixed transmitter and an additional fixed receiver, each permanently tuned to the control channel.

The tuning latency of tunable transmitters and receivers may be an important factor in choosing components, depending on the type of network being implemented. A single-hop network generally requires tunable components to create connections on demand and requires some amount of coordination in order to have the source node’s transmitter and the destination node’s receiver tuned to the same channel for the duration of an information transfer. In this case, the tuning time of transmitters and receivers may have a significant impact on the performance of the network. On the other hand, most multihop networks require tunability only for infrequent reconfigurations of the network based on changing traffic patterns; thus, the tuning time of components in a multihop network is not as critical as in the case of a single-hop network.

3) Node Separation in PSC WDM LAN’s: Given the output power of the transmitters and the receiver sensitivity, we can compute the maximum distance separating nodes. Assume that all nodes are $D$ meters from the PSC and that the input-to-output power splitting ratio of the PSC is given by (9). Then, the maximum value of $D$ such that the optical signals reaching each receiver are strong enough to be received ($D_{\text{max}}$) can be computed by combining (7) and (9), so that

$$D_{\text{max}} = \frac{10}{A} \log_{10} \frac{N P_r}{P_t} - 1$$

where $P_t$ is the optical power of the transmitter and $P_r$ is the minimum amount of power that the receiver needs to resolve the optical signal.

F. WDM WAN Design Issues

Current WAN’s are designed as electronic networks with fiber links. These networks may not be able to take full advantage of the bandwidth provided by optical fiber, however, because electronic switching components may be incapable of switching the high volume of data that can be transmitted on the fiber links. It is anticipated that the next generation of optical networks will make use of optical routers and switching elements to allow all-optical lightpaths to be set up from a source node to a destination node, thus bypassing electronic bottlenecks at intermediate switching nodes. Also, WDM will allow multiple lightpaths to share each fiber link. The concept of WDM lightpaths is analogous to a multilane highway that can be used to bypass stoplights on city roads. Another concept in WAN design is wavelength reuse. Since each wavelength may be used on each fiber link in the network, multiple lightpaths that do not share any links may use the same wavelength. For example, in Fig. 40, wavelength $\lambda_2$ is used to set up one lightpath from node A to node C and another lightpath from node G to node H. (Such wavelength reuse is not possible in a passive-star-based WDM network.)

The issue of setting up lightpaths and routing the lightpaths over the physical fibers and switches in a wide-area WDM network is an optimization problem in which the overall network performance must be balanced against the consumption of network resources. The degree of freedom in designing the lightpaths depends in part on the type of switching elements or cross connects used in the access nodes or switching nodes. If wavelength-insensitive cross-connect devices are used, then each signal on a given input fiber must be routed to the same output fiber.

Wavelength-sensitive switching devices offer more flexibility, allowing different wavelengths arriving on a single input fiber to be directed independently to different output fibers. However, this approach may still result in conflicts at the nodes if two signals on the same wavelength arriving on different input ports need to be routed to the same output port. The conflict may be resolved by incorporating wavelength converters at each node and converting one of the incoming signals to a different wavelength (see Section VII). If wavelength-conversion facilities are not available at switching nodes, then a lightpath must have the same wavelength on all of the fiber links through which it traverses; this is referred to as the wavelength-continuity constraint (see Fig. 28). Another approach for resolving conflicts is to find an alternate route in the network for one of the two conflicting lightpaths, and in some cases, an alternate wavelength.

In designing an optical network, it is important to recognize what can and cannot be accomplished by optical switching devices. Current optical cross connects may be able to switch optical information based on input ports or wavelengths but they cannot demultiplex time-division-multiplexed (TDM) data within an optical data stream. Also, in a reconfigurable optical switch, the time required to reconfigure the switch is often long when compared to
Fig. 40. Lightpath routing in a WDM WAN.

the speed of data passing through the switch. Therefore, optical switches seem more appropriate for switching large streams of data or for setting up somewhat static routes in the network based on wavelengths rather than for switching individual packets.

IX. EXPERIMENTAL WDM LIGHTWAVE NETWORKS

While WDM is anticipated to be the technology for future networks, it is important to know what currently has been accomplished using WDM technology. This section describes a sampling of experimental WDM testbeds and projects in both the local-area and wide-area domains.

A. LAN Testbeds

Two examples of WDM LAN testbeds are Bellcore’s LAMBDANET [92] and IBM’s Rainbow [24]. Like in many practical networks, simplicity has been chosen over “smart” protocols. Both LAMBDANET and Rainbow are single-hop networks.

1) LAMBDANET: In LAMBDANET, each node is equipped with one fixed transmitter consisting of a DFB laser and $N$ fixed receivers, where $N$ is the number of nodes in the network. The incoming wavelengths are separated using a diffraction-grating wavelength demultiplexer, and each individual wavelength is sent to an optical receiver. Each node’s transmitter is fixed on that node’s home wavelength. This allows single-hop connectivity without requiring tunable components or design of smart protocols. In [92], an aggregate throughput of 36 Gb/s was demonstrated for a 16-node system. While the advantages of LAMBDANET include simplicity of design and architectural support for multicasting, the architecture is not considered scalable because it requires $N$ data wavelengths. The practicality of requiring $N$ receivers per node is also a limitation, as for systems with large $N$, the cost per node might become large. Recent advances in receiver array technology, however, may help reduce the system’s cost [92].

2) Rainbow: IBM’s Rainbow-I testbed was designed to support up to 32 IBM PS/2’s connected via 32 200-Mb/s WDM channels (for an aggregate throughput of 6.4 Gb/s) [93]. Each node is equipped with a fixed DFB laser transmitter and a tunable Fabry–Perot filter receiver. Each node’s transmitter is fixed to a home channel. When a node has a packet to send, it tunes its receiver to the destination’s home channel and then transmits a connection setup request by repeatedly sending the destination node’s address. When a receiver is idle, it scans all wavelengths in round-robin fashion until it finds a channel with a setup request containing the receiver’s own address. It then transmits an acknowledgment to the source node, and two-way communication can begin. Though a simple protocol, the setup time may be too long for packet switching. Like LAMBDANET, Rainbow does not scale well. Rainbow was the first WDM testbed to demonstrate tunable WDM components working in a real environment.

The Rainbow-II optical network [94] is an extension of the Rainbow-I network. It consists of similar optical hardware and the same media-access protocol as the Rainbow-I network but also incorporates some higher layer protocols. The Rainbow-II network supports 32 nodes, each transmitting at rates of up to 1 Gb/s.

B. WAN Testbeds

A number of government-funded programs have been established in Europe and the United States to investigate WDM WAN’s. Some of the more notable projects include the Research and Development in Advanced Communications Technologies in Europe (RACE) Multiwavelength Transport Network (MWTN) project and the Multiwavelength Optical Networking (MONET), Optical Networks Technology Consortium (ONTC), and All-Optical Network
(AON) projects sponsored by the Defense Advanced Research Projects Agency. These programs focus on the design of WDM networks as well as the design of the optical components used in the networks.

1) RACE MWTN: The European RACE-MWTN program [95], [96] involves a consortium of European companies and universities. Its objective is to design an all-optical transport network layer employing optical switches, cross connects, transmitters, receivers, and amplifiers. The RACE-MWTN demonstration network spans 130 km and has been demonstrated with four wavelengths in the 1550-nm band. The project has developed two basic optical networking elements: the optical cross connect (OXC) and the optical add/drop multiplexer (OADM).

The OXC is an optical switching element for switched networks. It is constructed using 8×8 digital switches and 4×4 laser amplifier gate switches and uses EDFA’s at its input and output ports. The OXC is able to perform wavelength-based switching through the use of tunable filters. Four types of tunable filters—acoustooptic filters, integrated multilayer thin-film filters, and Fabry–Perot filters—have been used. Wavelength conversion takes place by converting a signal to electronics and retransmitting on a different wavelength.

The OADM adds or drops one or more wavelengths from a fiber while allowing the other optical signals to pass through. It makes use of 2×2 space switches and 2×2 acoustooptic switches. The OADM is intended primarily for ring or bus networks.

2) MONET: The MONET consortium [97] includes the American Telephone and Telegraph Company (AT&T), Bellcore, Lucent Technologies, Bell Atlantic, BellSouth, Pacific Telesis, and Southwestern Bell Communications. The goal of the program is to develop a transparent multiwavelength optical network, define the required enabling technologies, and explore the potential of WDM networks.

Currently, three testbeds are being constructed in the NJ area—local exchange, cross connect, and long distance (see Fig. 41, from [98]). The purpose of the long-distance testbed is to study the transmission of various optical signals over distances of 2000 km. The cross-connect testbed will connect the long-distance testbed with the local exchange and will allow researchers to study various cross-connect devices and network-management software. The local-exchange testbed will demonstrate various LAN topologies and some of their properties, such as scalability and interoperability. As of February 1997, MONET offers eight wavelengths (“MONET-compliant”) in the 1550-nm region, supporting analog and digital formats, and supports interconnection between the three testbeds. Transmissions of 2.5 Gb/s channels over 2290 km are currently being demonstrated. The MONET program employs laser arrays, consisting of DFB arrays with integrated modulators and integrated coupler; switch fabrics, including 4×4 LiNbO3 switches (WSXC), 2×2 liquid crystal switches with bulk optics (LCXC), and InP digital switch technologies; receiver arrays comprising eight photodetectors with integrated preamplifiers; wavelength converters based on cross-phase modulation with integrated DFB laser and parametric conversion; and optical amplifiers based on gain-flattened EDFA’s for multiwavelength operation.

3) ONTC: The ONTC [99] program includes members such as Bellcore, Columbia University, Hughes Research Laboratories, Northern Telecom, Rockwell Science Center, Case Western Reserve University, United Technology Research Center, Uniphase Telecommunications Product, and

Fig. 41. MONET NJ-area network.
Lawrence Livermore National Laboratories. The goal of the project is to construct a testbed to explore various WDM components and to study ATM/WDM networks.

The testbed consists of four access nodes connected to two fiber rings via $2 \times 2$ WDM cross-connects (see Fig. 42). The two fiber rings are also connected to each other using a $2 \times 2$ WDM cross connect.

Each access node is equipped with an ATM switch and SONET optical interfaces. The node may insert data into or remove data from the network, and may also provide packet forwarding for multihop communications.

Key network elements include EDFA’s, acoustooptic tunable filters, hybrid wavelength selective cross-connect switches, wavelength add/drop multiplexers, multiwavelength transmitters, and multiwavelength receivers.

4) AON: AT&T, Digital Equipment Corporation, and the Massachusetts Institute of Technology Lincoln Laboratory formed a precompetitive consortium for ARPA to address the challenges of utilizing the evolving terahertz capability of optical-fiber technology to develop a national information infrastructure capable of providing a flexible transport layer [100]. The main motivations for this architecture are the following.

- The architecture should scale gracefully to accommodate thousands of nodes and provide for a nationwide communication infrastructure.
- The architecture should be “future-proof,” i.e., modular and flexible to incorporate future developments in technology.

The architecture, shown in Fig. 43, is based on WDM and provides scalability through wavelength reuse and TDM techniques. The architecture employs a three-level hierarchy. At the lowest level are Level-0 networks, each consisting of a collection of LAN’s. Each Level-0 network shares wavelengths internally but there is extensive reuse of wavelengths among different Level-0 networks. A Level-1 network, which is a metropolitan area network, interconnects a set of Level-0 networks and provides wavelength reuse among Level-0 networks via wavelength routers. The highest level is the Level-2 network, which is a nationwide backbone network that interconnects Level-1 networks using wavelengths routers, wavelength converters, etc.

The services provided by the architecture are classified as follows.

- Type A service provides a dedicated optical path for point-to-point or point-to-multipoint communication. This would provide for Gb/s circuit-switched digital or analog sessions.
- Type B service uses TDM over an optical path to provide circuit-switched sessions, with bandwidths in the range from a few Mb/s to the full channel rate of a few Gb/s.
- Type C service is packet-switched and would be used internally for control, scheduling, and network management and for user datagram services.
X. CONCLUSION

Recent advances in the field of optics have opened the way for the practical implementation of WDM networks. In this paper, we have provided a brief overview of some of the optical WDM devices currently available or under development as well as some insight into the underlying technology. As optical device technology continues to improve, network designers need to be ready to take advantage of new device capabilities while keeping in mind the limitations of such devices.

ACKNOWLEDGMENT

The authors wish to thank the reviewers and the editor for their constructive criticism, which greatly served to improve the paper.

REFERENCES


BORELLA et al.: WDM LIGHTWAVE NETWORKS


Michael S. Borella (Member, IEEE) was born in Long Island, NY. He received the B.S. degree in computer science and technical communication (with distinction) from Clarkson University, Potsdam, NY, in 1991, and the M.S. and Ph.D. degrees in computer science from the University of California, Davis, in 1994 and 1995, respectively. In 1990, he was a Technical Writer for IBM, Poughkeepsie, NY. In 1992, he was a Software Engineer for Rolm, Santa Clara, CA. During 1992–1995, he was a Member of the University of California, Davis Networks Research Lab, where his research was supported by ARPA and the NSF. He currently is an Assistant Professor in the Computer Science Department, DePaul University, Chicago, IL. His research interests include internetworking, multimedia, performance evaluation and modeling of computer networks, optical networks, multicastransmission, and computer security. His research currently is supported by U.S. Robotics.

Dr. Borella is a member of the Association of Computing Machinery Special Interest Group on Data Communication.

Jason P. Jue received the B.S. degree in electrical engineering and computer science from the University of California, Berkeley, in 1990 and the M.S. degree in electrical engineering from the University of California, Los Angeles, in 1991. He is working toward the Ph.D. degree at the University of California, Davis.

He currently is a Research Assistant with the Networks Research Lab, University of California, Davis. His research interests include optical WDM networks, multicasting, mobility management, and wireless networks.

Dhrithiman Banerjee received the B.Tech (Hons.) degree from the Indian Institute of Technology, Madras, in 1992 and the M.S. and Ph.D. degrees from the University of California, Davis, in 1994 and 1996, respectively. His Ph.D. dissertation investigated the design and analysis of protocols and architectures for WDM wide-area optical networks. He currently is on the technical staff of Hewlett-Packard, Roseville, CA.

Byrav Ramamurthy (Student Member, IEEE) was born in Madras, India. He received the B.Tech. degree in computer science and engineering from the Indian Institute of Technology, Madras, in 1993 and the M.S. and Ph.D. degrees from the University of California, Davis, in 1994 and 1995, respectively. His Ph.D. dissertation investigated the design and analysis of protocols and architectures for WDM wide-area optical networks.

Mr. Ramamurthy received the Indian National Talent Search scholarship and was a fellow of the Professors for the Future program at the University of California, Davis.

Biswaath Mukherjee (Member, IEEE) received the B.Tech(Hons.) degree from the Indian Institute of Technology, Kharagpur, India, in 1980, and the Ph.D. degree from the University of Washington, Seattle, in 1987.

In 1987, he joined the University of California, Davis, where, since 1995, he has been a Professor of Computer Science. In September 1997, he will become Chairman of the Computer Science Department. His research interests include lightwave networks, network security, and wireless networks. He is the author of Optical Communication Networks (New York: McGraw-Hill, 1997).

Dr. Mukherjee was a co-winner of paper awards presented at the 1991 and 1994 National Computer Security Conferences. He is on the editorial boards of IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE NETWORK MAGAZINE, and the Journal of High-Speed Networks. He was the Technical Program Chair of IEEE INFOCOM '96.