Chapter

2

Enabling Technologies: Building Blocks

2.1 Introduction

This chapter is an introduction to WDM device issues. The reader needs no background in optics or advanced physics. For a more advanced and/or detailed discussion of WDM devices, we suggest that the reader refer to the references cited in this section. We highly recommend [KaBW96, Hech99, SaLu99, Keis00, RaSi01, DuDF02, Alwa04, Agra04, Pala04, Hech04].

This chapter presents an overview of optical fiber and devices such as couplers, optical receivers and filters, optical transmitters, optical amplifiers, and optical switches. The chapter attempts to condense the physics behind the principles of optical transmission in fiber in order to provide some background for the nonexpert. In addition, WDM network design issues are discussed in relation to the advantages and limits of optical devices. Finally, we demonstrate how these optical components can be used to create various WDM network architectures.

2.2 Optical Fiber

Fiber possesses many characteristics that make it an excellent physical medium for high-speed networking. Figure 2.1 shows the attenuation (and dispersion) characteristics of optical fiber.

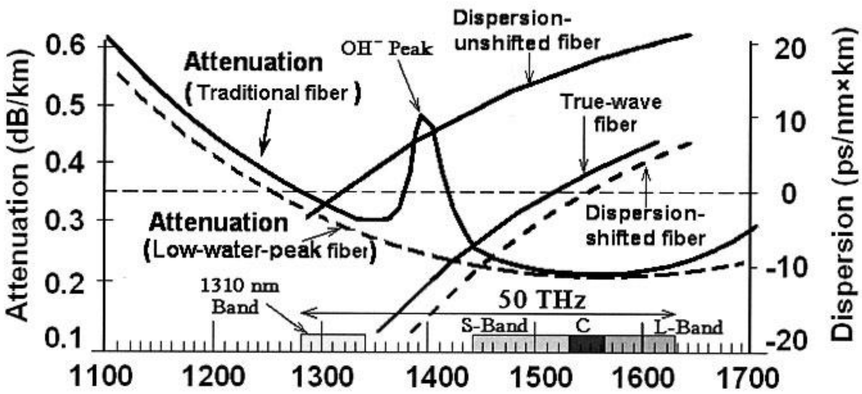


Figure 2.1 The low-attenuation regions of an optical fiber.

For traditional fiber, centered at approximately 1310 nm is a window of 200 nm in which attenuation is less than 0.5 dB/km. The total bandwidth in this region is about 25 THz. Centered at 1550 nm is a window of similar size, with attenuation as low as 0.2 dB/km, which consists of three bands, i.e., *S-band* (1460-1530 nm), *C-band* (1530-1560 nm), and *L-band* (1560-1630 nm). Combined, these two windows 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 neighborhood is due to hydroxyl ion (OH^-) impurities in the fiber. Other sources of loss include material absorption and radiative loss.

Besides traditional fibers, *full-spectrum fiber* has also attracted a lot of attention in the industry, because of its permanently reduced water peak, as well as additional enhanced specifications in the L-band. Full-spectrum ap-

[1]However, usable bandwidth is limited by fiber nonlinearities (Section 2.2.5), spectrum of optical amplifiers (see Section 2.5), etc.

plications involve simultaneous (WDM) transmission in multiple operating windows (1270 to 1610 nm) over a single fiber. Full-spectrum fibers provide more useable wavelengths than standard single-mode fiber and therefore more bandwidth per fiber. Specifically, *low-water-peak fibers* have attenuation specifications in line with the attenuation values in other transmission windows. Fibers with low-water-peak attenuation may use the 1360 to 1480 nm range without the severe loss previously experienced in traditional standard single-mode fibers. Industry standards organizations have established new classes of standard single-mode fibers that require the average attenuation at 1383 nm after hydrogen aging to be less than or equal to the specified attenuation at 1310 nm. The most widely recognized examples of such a standard are ITU-T G.652.C and D.

By using these large low-attenuation windows 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. Considering factors such as *low attenuation, wide window, and availability of optical amplifiers*, the 1550-nm window is preferred for long-haul wide-area applications. In single-channel long-distance experiments, optical signals typically have been sent over 80 km without amplification. Besides its enormous bandwidth and low attenuation, fiber also offers low error rates. Fiber optic systems typically operate at bit-error rates (BERs) of less than 10^{-15}.

The small size and thickness of fiber allows more fiber to occupy the same physical space as copper, a property which is desirable when installing local networks in buildings. Fiber is flexible, light, reliable in corrosive environments, and deployable at short notice (which makes it particularly favorable for military communication systems). Also, fiber transmission is immune to electromagnetic interference, and does not cause signal interference between fibers. Finally, fiber is made from one of the cheapest and most readily available substances on earth, namely glass. This makes fiber environmentally sound, unlike copper.

2.2.1 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 which acts as a waveguide. A waveguide is a physical medium or a path which 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, which is illuminated as following.

Light travels through vacuum at a speed of $c = 3 \times 10^8$ 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_{mat} be the speed of light for a given material. The ratio of the speed of light in vacuum to that in a material is known as the material's *refractive index* (n), and is given by: $n_{\text{mat}} = c/c_{\text{mat}}$. Given that $n_{mat} = 1.5$ approximately for glass, the velocity of signal propagation in a fiber approximately equals 2×10^8 m/s, which corresponds to a signal propagation delay of 5 μs/km.

When light travels from material of a given refractive index to material of a different refractive index (i.e., when refraction occurs), the angle at which the light is transmitted in the second material depends on the refractive indices of the two materials as well as the angle at which light strikes the interface between the two materials. Due to Snell's Law, $n_a sin\theta_a = n_b sin\theta_b$, where n_a and n_b are the refractive indices of the first substance and the second substance, respectively; θ_a is the angle of incidence, or the angle with respect to normal that light hits the surface between the two materials; and θ_b is the angle of light in the second material. However, if $n_a > n_b$ and θ_a is greater than some critical value, the rays are reflected back into substance a from its boundary with substance b.

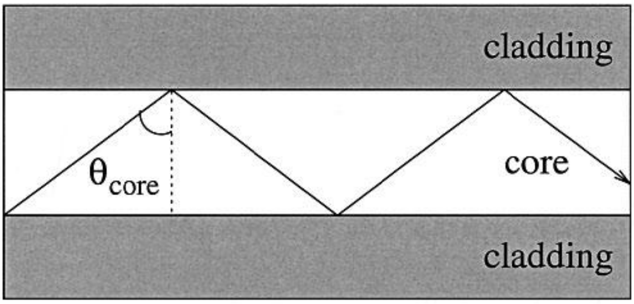


Figure 2.2 Light traveling via total internal reflection within a fiber.

Looking at Figs. 2.2 and 2.5, we see that the fiber consists of a core completely surrounded by a cladding (both the core and the cladding consist of glass of different refractive indices). 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. 2.2). The angle above which total internal reflection will take place is known as the *critical angle*, and is given by θ_{core} which corresponds to $\theta_{\text{clad}} = 90°$. From Snell's Law, we have:

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

The critical angle is then:

$$\theta_{\text{crit}} = \sin^{-1}\left(\frac{n_{\text{clad}}}{n_{\text{core}}}\right). \tag{2.1}$$

So, for total internal reflection, we require:

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

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 θ_{crit}.

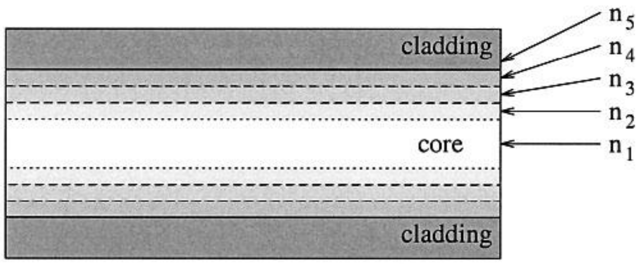


Figure 2.3 Graded-index fiber.

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. 2.3). A *graded-index* fiber reduces the minimum θ_{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.

In order 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 being at an angle for which total internal reflection can take place at

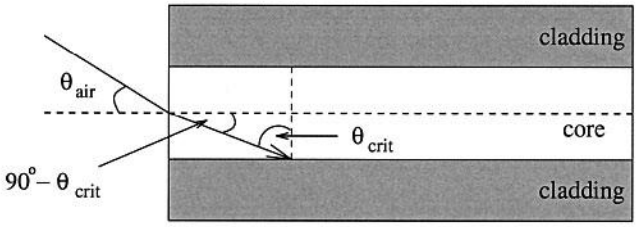


Figure 2.4 Numerical aperture of a fiber.

the core-cladding boundary. As shown in Fig. 2.4, the maximum value of θ_{air} can be derived from:

$$\begin{aligned} n_{\text{air}} \sin\theta_{\text{air}} &= n_{\text{core}} \sin(90^\circ - \theta_{\text{crit}}) \\ &= n_{\text{core}} \sqrt{1 - \sin^2\theta_{\text{crit}}} \end{aligned} \tag{2.2}$$

From Eqn. (2.1), since $\sin\theta_{\text{crit}} = n_{\text{clad}}/n_{\text{core}}$, we can rewrite Eqn. (2.2) as:

$$n_{\text{air}} \sin\theta_{\text{air}} = \sqrt{n_{\text{core}}^2 - n_{\text{clad}}^2} \tag{2.3}$$

The quantity $n_{\text{air}} \sin\theta_{\text{air}}$ is referred to as NA, the *numerical aperture* of the fiber, and θ_{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. According to Snell's Law and fiber refractive index, typical delay of optical propagation in optical fiber is 5 μs/km.

2.2.2 Single-Mode vs. Multimode Fiber

A mode in an optical fiber corresponds to one of possibly many ways in which a wave may propagate through the fiber. It can also be viewed as a standing wave in the transverse plane of the fiber. More formally, a mode corresponds to a solution of the wave equation which is derived from Maxwell's equations and is subject to boundary conditions imposed by the optical fiber waveguide.

An electromagnetic wave propagating along an optical fiber consists of an electric field vector, **E**, and a magnetic field vector, **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 ρ is the component of the field which is normal to the wall (core-cladding boundary) of the fiber, ϕ is the component of the field which is tangential to the wall of the fiber, and z is the component of the field which is in the direction of propagation. Fiber modes are typically referred to using the notation HE_{xy} (if $H_z > E_z$), or EH_{xy} (if $E_z > H_z$), 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 $E_z = 0$, or transverse-magnetic (TM) in which case $H_z = 0$.

Although total internal reflection may occur for any angle θ_{core} which is greater than θ_{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 higher operating frequencies allow 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:

$$V = k_0 a \sqrt{n_{\text{core}}^2 - n_{\text{clad}}^2} \tag{2.4}$$

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

$$m \approx \frac{1}{2} V^2. \tag{2.5}$$

The advantage of multimode fiber is that its core diameter is relatively large; as a result, injection of light into the fiber with low coupling loss[2] can be accomplished by using inexpensive, large-area light sources, such as light-emitting diodes (LEDs).

The disadvantage of multimode fiber is that it introduces the phenomenon of *intermodal dispersion.* In multimode fiber, each mode propagates at a different velocity due to different angles of incidence at the core-cladding boundary. This effect causes different rays of light from the same source to

[2] Coupling loss measures the power loss experienced when attempting to direct light into a fiber.

arrive at the other end of the fiber at different times, resulting in a pulse which is spread out in the time domain. Intermodal dispersion increases with the propagation distance. 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. 2.3). However, even with graded-index multimode fiber, intermodal dispersion may still limit the bit rate of the transmitted signal and the distance that the signal can travel.

One way to limit intermodal dispersion is to reduce the number of modes. From Eqns. (2.4) and (2.5), we observe that this reduction in the number of modes can be accomplished by reducing the core diameter, by reducing the numerical aperture, or by increasing the wavelength of the light.

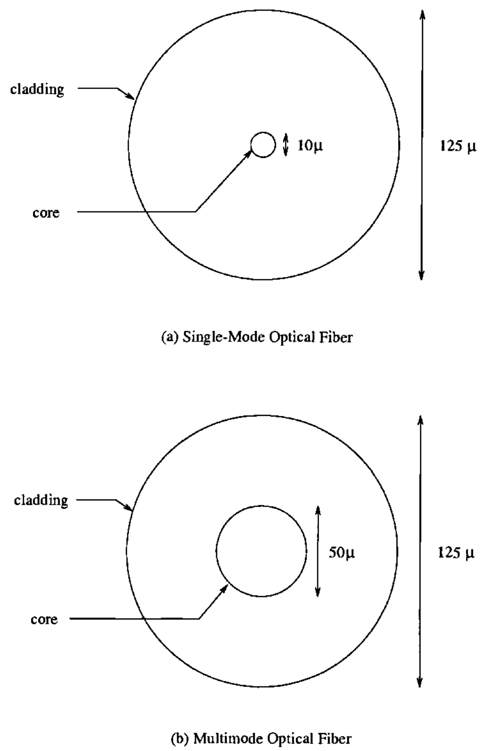


Figure 2.5 Single-mode and multimode optical fibers.

By reducing the fiber core to a sufficiently small diameter and by 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 μm, while multimode fiber typically has a core size of 50 to 100 μm (see Fig. 2.5). A step-index fiber will support a single mode if V in Eqn. (2.4) is less than 2.4048 [Ishi91].

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. LEDs 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.

2.2.3 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. Receiver sensitivity is the minimum power required by a receiver to detect the signal. 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 [Henr85]:

$$P(L) = 10^{-AL/10} P(0) \tag{2.6}$$

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 Eqn. (2.6), we get:

$$L_{\max} = \frac{10}{A} \log_{10} \frac{P(0)}{P_r} \tag{2.7}$$

The maximum distance between the transmitter and the receiver (or the distance between amplifiers[3]) depends more heavily on the constant A than on the optical power launched by the transmitter. Referring back to Fig. 2.1, we note that the lowest attenuation ($\approx$ 0.2 dB/km) occurs at approximately

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

1550 nm. In optical communication system, the traditional value of optical propagation is 80km without amplification[RaSi01]. With developments of new fibers and optical communication systems, the traditional propagation distance is surpassed in both physical system and research field. In March 2005, a vendor announced a new generation of optical amplifiers by increasing amplifier spacing from the traditional 80 km to 160 km.

2.2.4 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 rate 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 *chromatic dispersion* (see Fig. 2.1). Chromatic dispersion represents the fact that different colors or wavelengths travel at different speeds, even within the same mode. 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. Chromatic dispersion is the result of *material dispersion*, *waveguide dispersion*, and/or *profile dispersion*. *Material dispersion* results from the different velocities of each wavelength in a material. Since no laser can create a signal consisting of an exact single wavelength, or more precisely, since any information carrying signal will have a nonzero spectral width (range of wavelengths/frequencies in the signal), material dispersion will occur in most systems[4]. *Waveguide dispersion* is caused because the propagation of different wavelengths depends on waveguide characteristics, such as the indices and shape of the fiber core and cladding. *Profile dispersion* is caused by the variation of refractive index with respect to wavelength. These first two forms of dispersion are universal in optical fibers. Although the single-mode fiber (SMF) can perfectly eliminate several types of dispersion (which the multimode fiber cannot), *chromatic dispersion* and *polarization mode dispersion* (PMD) still need to be dealt with.

[4]Even if an unmodulated source consisted of a single wavelength, the process of modulation would cause a spread of wavelengths.

Polarization mode dispersion (PMD) is another complex optical effect that can occur in single-mode optical fibers. Single-mode fibers support two perpendicular polarizations of the original transmitted signal. If a fiber were perfectly round and free from all stresses, both polarization modes would propagate at exactly the same speed, resulting in zero PMD. However, practical fibers are not perfect. The two perpendicular polarizations may travel at different speeds and, consequently, arrive at the end of the fiber at different times (one perpendicular polarization direction is the fast axis, and the other one is the slow axis). The difference in arrival times between the axes is known as PMD. Like chromatic dispersion, PMD causes digitally-transmitted pulses to spread out as the polarization modes arrive at their destination at different times. For digital high-bit-rate transmissions, this effect of PMD can lead to bit errors at the receiver or limit the receiver sensitivity. The maximum acceptable dispersion penalty is usually 2 dB, though it is possible for a system to tolerate a larger dispersion penalty if the optical attenuation is low.

The above discussion mainly referred to the most prevalent fiber type, namely, dispersion-unshifted SMF. Currently, there are a number of special designs of optical fibers available, which offer lower dispersion than the dispersion-unshifted SMF. For example, in a dispersion-shifted fiber (e.g., non-zero dispersion-shifted fiber (NZDSF), dispersion-compensated fiber (DCF), etc.), the core and cladding are designed such that the waveguide dispersion is negative with respect to the material dispersion, thus the total *chromatic dispersion* is set to zero (see Fig. 2.1).

2.2.5 Nonlinearities in Fiber

Nonlinear effects in fiber may potentially have a significant impact on the performance of WDM optical communication 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, and they limit the maximum power per channel, the maximum bit rate, and the system reach.

Nonlinear Refraction

In an optical fiber, the index of refraction depends on the optical intensity of signals propagating through the fiber [Chra90]. 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. The amount of phase shift introduced by SPM is given in the equation:

$$\phi_{NL} = n_2 k_0 L |E|^2 \tag{2.8}$$

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 the spreading or compression of the pulse in the time domain and affect the maximum bit rate and the bit error rate.

Cross-phase modulation (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 dispersion, may also affect the pulse shape in the time domain.

Although XPM may limit the performance of fiber-optic systems, it may have advantageous applications as well. 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 2.7.

Stimulated Raman Scattering

Stimulated Raman Scattering (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*[5] 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 shorter-wavelength channels 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 [Chra84], it is shown that, in a 10-channel system with 10-nm channel spacing, the power on each channel should be kept below 3 mW to minimize the effects of SRS. In Section 2.5.4, SRS will be utilized as a basis of Raman amplifier.

Stimulated Brillouin Scattering

Stimulated Brillouin scattering (SBS) is similar to SRS, except that the frequency shift is caused by sound waves rather than molecular vibrations [Chra90]. 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 μs), but has negligible effect for short pulses (less than 10 ns) [Agra01]. 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 crosstalk between channels. Crosstalk will occur when two counter-propagating channels differ in frequency by the Brillouin shift, which is around 11 GHz for wavelengths at 1550 nm. However, the narrow gain bandwidth of SBS makes SBS crosstalk fairly easy to avoid.

---

[5]The Raman gain spectrum typically describes the measured Raman-gain coefficient for silica fibers as a function of the frequency shift at a pump wavelength of 1.0 μm.

Four-Wave Mixing

As the bit rate of optical data streams in fibers increases, four-wave mixing (FWM) is one principal among nonlinear effects in pulse propagation.

FWM causes inter-channel crosstalk and is worst-case for equally-spaced WDM channels. FWM penalty can be mitigated by using fiber with high local dispersion (SMF, NZDSF) or unequally spaced channels [FTCM94]. Even if using NZDSF can mitigate the FWM penalty, the minimum channel spacing has a lower limit.

However, there are some significant motivations that are taking advantages of FWM in WDM networks. For example, FWM can be used to provide wavelength conversion. More information will be shown in Section 2.7.

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, the maximum number of channels, and 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 [Chra90].

However, there are some applications, in which nonlinear effects are exploited, such as Raman amplifier, and wavelength conversion by using FWM.

There are some efforts to combat nonlinearities by using dispersion. An example is an optical soliton[6] system which is suitable for a compensation between dispersion and nonlinear effects. The main limitation of soliton systems is fiber loss. With decreasing power, compensation between dispersion and nonlinearities is no longer achieved [CCEL04].

The details of optical nonlinearities are very complex, and beyond the scope of this book. However, they are a major limiting factor in the available number channels in a WDM system, especially those operating over distances greater than 30 km [Chra90]. The existence of these nonlinearities suggests that WDM protocols which 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 [Agra01].

[6] An optical soliton is a pulse with specific shape, power, and duration.

2.2.6 Optical Fiber Couplers

Coupler is a general term that covers all devices that combine light into or split light out of a fiber. Optic fiber couplers can be either active or passive devices. The difference between active and passive couplers is that a passive coupler redistributes the optical signal without optical-to-electrical conversion. Active couplers, however, are electronic devices that split or combine the signal electrically and use fiber optic detectors and sources for input and output. The passive coupler is the most popular because it requires no external power to operate. Thus, passive couplers are referred in next sections. 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. 2.6(a).

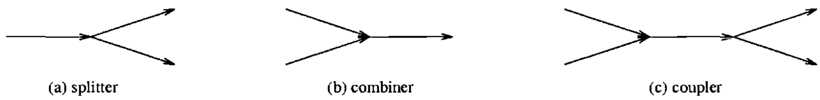


Figure 2.6 Splitter, combiner, and coupler.

The *splitting ratio*, α, 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 [Powe93]. Combiners (see Fig. 2.6(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. 2.6(c)), in general, is a 2×1 combiner followed immediately by a 1×2 splitter, which has the effect of broadcasting the signals from two input fibers onto two output fibers. One implementation of a 2 $\times$ 2 coupler is the *fused biconical tapered coupler* which basically 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, in which the amount of power returned by a coupler is typically 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.

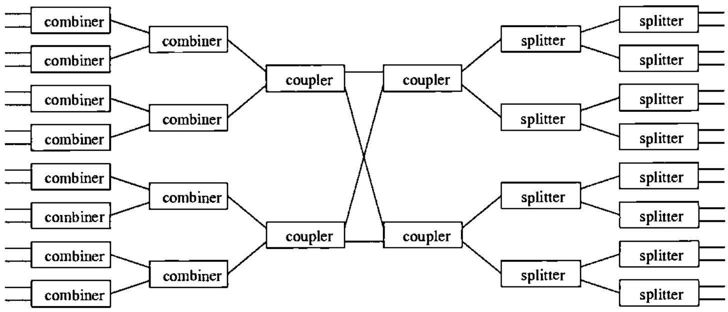


Figure 2.7 A 16 × 16 passive-star coupler.

The passive-star coupler (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:

$$P_{out} = \frac{P_{in}}{N} \tag{2.9}$$

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. 2.7. (However, this implementation of a PSC and the implementation of the 2 × 2 coupler in Fig. 2.6(c) are not the most power efficient. As an exercise, an alternative, power-efficient design of the PSC can be explored.) 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 × 19 star coupler on silicon has been demonstrated with excess loss of around 3.5 dB at a wavelength of 1300 nm [DHKK89], and an 8 × 8 star coupler with an excess loss of 1.6 dB at a wavelength of 1550 nm has been demonstrated in [OkTa91].

2.3 Optical Transmitters

In order 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 [Gree93, Brac90, LeZa89, RaSi01].

2.3.1 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 which contains one or more distinct frequencies).

In order 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. In order to change the level of an electron in the ground state, the atom must absorb energy. When an atom absorbs energy, it becomes excited, and moves to a higher energy level. At this point, the electron is unstable, and usually moves quickly back to the ground state by releasing a *photon*, a particle of light.

However, there are certain materials whose states are *quasi-stable*, which means that the substances are likely to stay in the excited state for longer periods of time. By applying enough energy (either in the form of an optical pump or in the form of 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.

Figure 2.8 shows a general representation of the structure of a laser. The laser consists of two mirrors which 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.

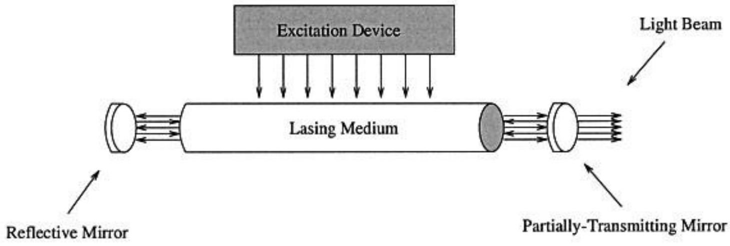


Figure 2.8 The general structure of a laser.

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 coherence 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:

$$f = \frac{E_i - E_f}{h} \tag{2.10}$$

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*, which changes depending on the temperature of the gas. Although many frequencies are possible, only a single frequency, which is determined by the cavity length, is emitted from the laser.

Semiconductor Diode Lasers

The most useful type of laser for optical communications is the semiconductor diode laser. The simplest implementation of a semiconductor laser is the bulk laser diode, which is a p-n junction with mirrored edges perpendicular to the junction (see Fig. 2.9). To understand the operation of the semiconductor diode requires a brief diversion into semiconductor physics.

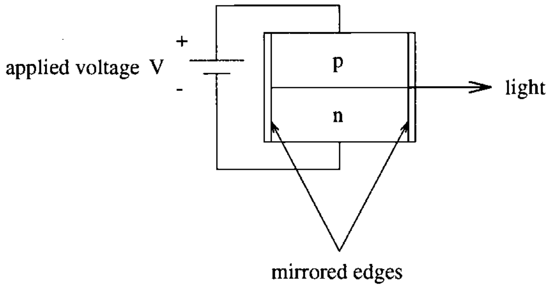


Figure 2.9 Structure of a semiconductor diode laser.

In semiconductor materials, electrons may occupy either the valence band or the conduction band. The valence band and conduction band are analogous to the ground state and excited state of an electron mentioned in the previous section. The valence band corresponds to an energy level at which an electron is not free from an atom. The conduction band corresponds to a level of energy at which an electron has become a free electron and may move freely to create current flow. The region of energy between the valence band and the conduction band is known as the *band gap*. An electron may not occupy any energy levels in the band-gap region. When an electron moves from the valence band to the conduction band, it leaves a vacancy, or *hole*, in the valence band. When the electron moves from the conduction band to the valence band, it recombines with the hole and may produce the spontaneous emission of a photon. The frequency of the photon is given by Eqn. (2.10), where $E_i - E_f$ is the band-gap energy. The distribution of the energy levels which electrons may occupy is given by the *Fermi-Dirac distribution* in [Fuku98].

A semiconductor 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 which provide extra electrons. These electrons will remain in the conduction band. A p-type semiconductor is doped with impurities which 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 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 which 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 which generate light with a very narrow frequency range).

2.3.2 Tunable and Fixed Lasers

The previous section provided an overview of a generic model of a laser, but 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.

Laser Characteristics

Some of the physical characteristics of lasers which 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 2.2.4, 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 examples of which are *mode hopping*, *mode shifts*, and *wavelength chirp* [MoTo93]. *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. In order to avoid large shifts in frequency, methods must be utilized to compensate for variations in temperature or injection current. One approach for temperature compensation is to package with the laser a thermoelectric cooler element which 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 it can amplify. 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 λ 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 which 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* specifies the time required for the laser to tune from one wavelength to another.

Mechanically-Tuned Lasers

Most mechanically-tuned lasers use a Fabry-Perot cavity that is an optical resonator in which feedback is accomplished by two parallel mirrors and 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 [Brac90], 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. Fabry-Perot lasers further break down into Buried Hetero (BH) and Multi-Quantum Well (MQW) types. BH and related styles have ruled for many years, but currently MQW types are becoming very widespread. MQW lasers offer significant advantages over all former types of Fabry-Perot lasers. MQW lasers offer lower threshold current, higher slope efficiency, lower noise, better linearity, and much greater stability over temperature. As a bonus, MQW lasers have great performance margins, give laser manufacturers better yields, and reduce laser cost. One disadvantage of MQW lasers, however, is their tendency to be more susceptible to back reflections.

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 approximately 10-μs 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 allow continuous tuning over a range of wavelengths.

Injection-Current-Tuned Lasers

Injection-current-tuned lasers form a family of transmitters which 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 which match the period and indices of the grating will be constructively reinforced. All other wavelengths will destructively interfere, and will not propagate through the waveguide. The condition for propagation is given by:

$$D = \frac{\lambda}{2n}$$

where D is the period of the grating [MoTo93]. The laser is tuned by injecting a current which 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 DBR-based laser offers operating advantages such as a broad range of biasing conditions and the tunability of the emitted wavelength. There are various kinds of Modified-DBR lasers, such as Sampled-Grating DBR (SG-DBR), Digital-Supermode (DS-DBR) and Super-Structure Grating DBR (DBR) [GoNT04, JHBC05, WRBB05]. A SG-DBR laser with tuning range exceeding 40 nm is demonstrated in [JHBC05]. A tunable DS-DBR laser has been presented in [WRBB05], in which the full continuous tuning range of the laser can reach 45 nm. However, DBR devices suffer from the major drawback of mode hopping that occurs due to the changes in the lasing-region index of refraction with increasing driven current.

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 which 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.

One current laser array technique is MEMS-based vertical-cavity surface-emitting laser (VCSEL). VCSEL is a new laser structure that emits laser light vertically from its surface and has vertical laser cavity. The VCSEL's principles of operation closely resembles those of conventional edge-emitting semiconductor lasers. The core of the VCSEL is an electrically-pumped gain region, called the active region, which emits light. Layers of varying semiconductor materials above and below the gain region create mirrors. Each mirror reflects a narrow range of wavelengths back into the cavity causing light emission at a single wavelength. VCSELs are typically multi-quantum-well (MQW) devices with lasing occurring in layers only 20-30 atoms thick. In VCSELs, the Bragg reflectors with as many as 120 mirror layers form the laser reflectors.

There are many advantages of VCSELs. The small-size and high-efficiency mirrors produce a low threshold current, which is less than 1 mA. The transfer function allows stability over a wide temperature range, which is unique to this type of laser diode. The MEMS-based VCSEL realizes a fast tuning time of 1 to 10 μs. These features make the VCSEL ideal for applications that require an array of devices.

2.3.3 Optical Modulation

In order 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 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 ($\geq$ 2 Gbps) is to have an external modulator which modulates the light coming out of the laser. The external modulator blocks or passes light depending on the current applied to it.

In long-haul, high-capacity WDM systems, the use of advanced modulation formats has been an effective scheme to manage signal impairments arising from amplified spontaneous emission (ASE) noise, fiber nonlinear effects, and polarization-mode dispersion (PMD). The ideal modulation format for long-haul high-speed WDM transmission links is one that has a narrow spectral width, low susceptibility to fiber nonlinear effects, large dispersion tolerance, and a simple and cost-effective configuration for signal generation. There are a number of advanced formats that meet these criteria to varying degrees, including nonreturn-to-zero (NRZ), return-to-zero (RZ), and duobinary coding. There are also a number of variations of the RZ format, including simple RZ, carrier-suppressed RZ (CS-RZ), chirped RZ (CRZ), vestigial sideband RZ, and dispersion-managed soliton-based RZ [WeNL04].

Recently, the differential phase-shift-keyed (DPSK) format has attracted renewed interest due to its 3-dB lower requirement on optical signal-to-noise ratio and better resilience to cross-phase modulation compared with other on-off-keyed (OOK) formats mentioned above. A typical DPSK transmitter consists of a continuous-wave laser followed by two cascaded modulators. The first modulator is used to encode data via binary phase modulation. It can be a phase modulator (PM) or a dual-drive Mach-Zehnder modulator (MZM) biased at the null in its transmission curve with twice the drive voltage swing. The difference between these two methods is that a PM leaves the optical intensity constant and modulates the phase subject to its bandwidth limitations, while a MZM produces instantaneous phase jumps at the expense of some residual intensity modulation, which are of minor importance when using an RZ format. The second modulator was used to carve pulses out of the NRZ-phase modulated signal. Depending on the drive conditions of this modulator, RZ-DPSK with 33% and 50% duty cycle and CSRZ-DPSK with 67% duty cycle can be obtained correspondingly [WeNL04].

The Mach-Zehnder interferometer, described later in Section 2.4.2, can be used as a modulation device. A drive voltage is applied to one of two waveguides creating an electric field which causes the signals in the two waveguides to either be in phase or 180° out of phase, resulting in the light from the laser being either passed through the device or blocked. Mach-Zehnder amplitude modulators which offer bandwidths of up to 45 GHz are currently available [SMOI05]. One of the advantages of using integrated-optics devices such as the Mach-Zehnder interferometer is that the laser and modulator can be integrated on a single structure, which may be cost effective. Also, integrating the laser with the modulator eliminates the need for polarization control and results in low chirp.

2.3.4 Summary

Table 2.1 summarizes the characteristics of the different types of tunable transmitters. Observe that there is a trade-off between the tuning range of a transmitter and its tuning time.

2.4 Optical Receivers and Filters

Tunable optical filter technology is a key in making WDM networks realizable. Good sources of information on these devices include [Gree93, Brac90, KoCh89, RaSi01].

Table 2.1 Tunable optical transmitters and their associated tuning ranges and times.

Tunable Transmitter	*Approx. Tuning Range (nm)*	*Tuning Time*
Mechanical (external cavity)	550	1–10 ms
Acoustooptic	750	~9 μs
Electrooptic	7	1–10 ns
Injection-Current (DFB and DBR)	45	1–10 ns

2.4.1 Photodetection

In receivers employing *direct detection*, a photodetector converts the incoming optical stream into a stream of electrons. The electron stream (i.e., electrical current) is then amplified and passed through a threshold device. Whether a bit is a logical 0 or 1 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[7] 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. Coherent-detection-based receivers use a monochromatic laser as a local oscillator. The incoming optical stream, which is at a slightly different frequency from the oscillator, is combined with the signal from the oscillator, resulting in a signal at the difference frequency. This difference signal, which is in the

[7]An intrinsic material is a semiconductor material with electrical properties, which essentially has characteristics of the pure crystal, e.g., essentially silicon or germanium crystal with no measurable impurities.

microwave range which frequency range is from 1000 MHz and upward, is amplified, and then photodetected. While coherent detection is more elaborate than direct detection, the former allows the reception of weak signals from a noisy background. However, in optical systems, it is difficult to maintain the phase information required for coherent detection (see [Aziz91]). 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.

2.4.2 Tunable Optical Filters

This section discusses several types of tunable optical filters and the properties of each type, while Section 2.4.3 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 [Gree93] and [RaSi01].

Filter Characteristics

Tunable optical filters are characterized primarily by their *tuning range* and *tuning time*. The *tuning range* specifies the range of wavelengths which 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 network 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 which is a distance of $n \times$ FSR from the selected frequency, where n is a positive integer. For example, in Fig. 2.10, if the filter is tuned to frequency f_1, then all frequencies labeled with a 1 will be passed by the filter; tuning the filter to the next frequency, f_2, will allow all frequencies labeled with a 2 to be passed by the filter; etc. The FSR usually depends on various physical parameters in the device, such as cavity lengths or waveguide lengths.

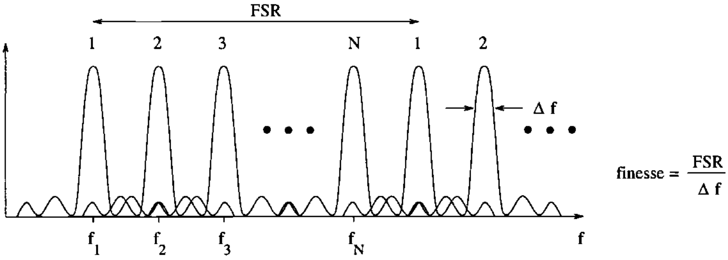


Figure 2.10 Free spectral range and finesse of a tunable filter capable of tuning to N different channels.

The *finesse* of a filter is a measure of the width of the transfer function. It is the ratio of FSR to channel bandwidth, where the channel bandwidth is defined to be the 3-dB bandwidth of a channel.

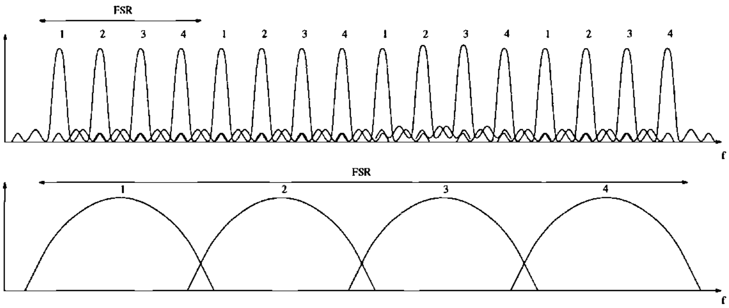


Figure 2.11 Cascading filters with different FSRs.

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 (passband peaks) 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 crosstalk, resulting in fewer channels. One approach to increasing the number of channels is to cascade filters with different FSRs [Gree93]. Figure 2.11 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.

The Etalon

An etalon can be one of the following two types:

- First, a single of cavity is formed by two parallel flat mirrors (planar etalon). Almost all etalons are planar etalons.
- Second, a single of cavity consists of two identical spherical mirrors with their concave sides facing each other and with the distance between the mirrors equal to each mirror's radius of curvature (confocal or spherical etalon). This kind of etalons is much less common and will not be elaborated in this section.

In planar etalon filters, a typical type is the Fabry-Perot etalon, which is the simplest form of a Fabry-Perot interferometer. Its primary optical property is that if a monochromatic light ray travels back and forth between two mirrors and distance between mirrors equals an integral number of wavelengths, then the light passes through the etalon. Many modifications (e.g., multicavity and multipass) to the etalon can be made to improve the number of resolvable channels. In a multipass filter, the light passes through the same cavity multiple times, while in a multicavity filter, multiple etalons of different FSRs are cascaded to effectively increase the finesse.

The Fabry-Perot etalon can be made to virtually access the entire low-attenuation region of the fiber and can resolve very narrow passbands. But, 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.

For applications to optical WDM communication systems, the FSR of the Fabry-Perot Etalon corresponds to the typical channel spacing of 100 GHz or 50 GHz, thus requiring cavity lengths of 1−2 mm. Furthermore, the Fabry-Perot filter ideally covers a whole communication band, which is typically tens of nanometers large. All-fiber Fabry-Perot devices with finesse reaching 240 and covering a spectral band of 26 nm are presented in [SIDL03].

The Mach-Zehnder Chain

In a Mach-Zehnder (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. 2.12). 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° are filtered out. By constructing a chain of these elements, a single desired optical wavelength can be selected.

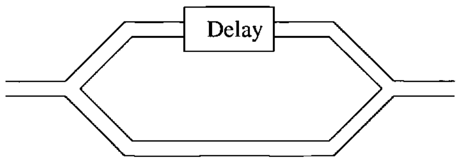


Figure 2.12 Structure of a Mach-Zehnder interferometer.

While the MZ chain may 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 [Gree93]. The high tuning time is due to thermal elements used to implement the delay elements. Recent advances have produced a fast tuning MZ filter, which exhibits the total free-spectral range of 16 nm, an extinction ratio of 20 dB, and a 3-dB transmission bandwidth of 32 GHz [SMOI05].

Acoustooptic Filters

Using acoustooptic filters leads to a fast tuning time. 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 [Gree93]. 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 (SAW)* to about 10 μs [NewF94]. However, the tuning range for acoustooptic filters covers the entire 1300 nm to 1560 nm spectrum. This tuning range potentially allows about 100 channels.

If more than one RF wave is passed through the grating simultaneously, more than one wavelength can be filtered out [CSBH89]. 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, crosstalk will occur. The selection of up to five wavelengths was reported in [CSBH89].

An acousto-optic tunable filter (AOTF) is one of the most suitable tunable filters for wavelength tunable lasers because it has a wide tuning range (100 nm) and fast switching speed (several microseconds). Additionally, an AOTF provides stable operation against shock and vibration owing to its non-mechanical structure in [TTHD04]. A vendor recently announced its Acousto-Optic Filter whose tuning range is up from 380 to 750 nm with rapid spectrum scanning from 6 to 9 μs in March 2005.

One drawback of acoustooptic filters is that, because of their wide transfer function, they are unable to filter out crosstalk 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.

Electrooptic Filters

Since the tuning time of the acoustooptic filter is limited by the speed of sound, crystals whose indices 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 [Gree93]. Since the tuning time is limited only by the speed of electronics, tuning time can be on the order of several nanoseconds.

Liquid-Crystal Fabry-Perot Filters

The design of a liquid-crystal filter is similar to the design of a Fabry-Perot filter, but the cavity consists of a liquid crystal (LC). The refractive index of the LC is modulated by an electrical current to filter out a desired wavelength, as in an electrooptic filter. 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 WDM networks.

In [Hira05], a new filter is presented that uses electrooptic material instead of LC in Fabry-Perot tuning filter. The free spectral range (FSR) of the filters is about 10 nm, tunable range is about 10 nm, loss is 2.2 dB, finesse is 150, and tuning speed takes only 1 μs.

2.4.3 Fixed Filters

An alternative to tunable filters are 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.

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 [Hech04]. 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 us to select the wavelength(s) we want 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 2.3.2).

An alternative implementation of a demultiplexer is the *arranged waveguide grating* (AWG) (also know as *waveguide grating router* (WGR)), in which only one input is utilized. AWGs will be discussed in Section 2.6.5.

Fiber Bragg Gratings (FBG)

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 [ISII95]. 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 [ASWC96]. Fiber Bragg gratings may be used in the implementation of multiplexers, demultiplexers, and tunable filters.

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. However, advances have been made which address some of these issues [DMNM04].

2.4.4 Summary of Optical Filtering Technologies

Table 2.2 Tunable optical filters and their associated tuning ranges and times.

Tunable Receiver	*Approx. Tuning Range (nm)*	*Tuning Time*
Fabry-Perot	500	1–10 ms
Acoustooptic	250	~10 μs
Electrooptic	16	1–10 ns
LC Fabry-Perot	50	0.5–10 μs

Table 2.3 Summary of optical filtering technologies [TzZT04].

Technology	*Loss* (dB)	*Channel Spacing* (GHz)	*Crosstalk* (dB)	*Tunability*	*Maturity*
FBG	0.5	50	30	Yes	High
TFF	1	100	12	Yes	High
AWG	6	50	30	No	High
MZI	1	50	30	No	High
AOF	3	100	15-20	Yes	Low
DCE	5	50	30	Yes	High

FBG = Fiber Bragg Gratings; TFF = Thin-Film Filters;
AWG = Arrayed Waveguide Gratings; MZI = Mach-Zehnder Interferometers;
AOF = Acousto-Optic Filters; DCE = Dynamic Channel Equalizers.

Table 2.2 summarizes the state of the art in tunable receivers. As has been stated earlier, tuning range and tuning time seem inversely proportional, except in LC Fabry-Perot filters.

Table 2.3 is a summary of several technologies used in filters with the corresponding features and specifications.

2.4.5 Channel Equalizers

The WDM systems currently under development incorporate longer spans and/or higher bit rates and/or reconfigurable add/drop, all of which lead to requirements for dynamic compensation using optical amplifiers (see Section 2.5. Two different types of compensation are generally required. To compensate for the residual non-flatness of amplifier gain profiles, resulting from imperfections in gain-flattening filters, changes in the amplifier operating conditions, and changes in channel loading, one needs a dynamic gain equalizer (DGE), which can achieve a smooth, low-ripple spectral attenuation profile that is the negative of the deviations of the amplifier gain profile from the desired profile. To compensate for unequal channel powers, resulting from dropping and adding channels, one needs a dynamic channel-power equalizer (DCE), which can achieve a flat attenuation profile across the full bandwidth of each channel, but can individually adjust the attenuation for each channel.

A DCE can not only provide a variable attenuation for each channel, but it can also achieve a sufficiently high attenuation (perhaps 40 to 50 dB) to effectively eliminate or block selected channels. Thus, it can be used as a key active element in a reconfigurable add/drop node. These devices are commonly referred to as wavelength blockers (WB) [Toml03]. In some cases, the same technology can be used to implement either the DGE functionality or the DCE functionality, but the basic design parameters for these two cases are sufficiently different that different device designs are required to accomplish the two different functionalities. In particular, the DGE plays an important role in WDM networks, because of its ability to control the power profile of the wavelength channels, hence maintaining a high quality of service (QoS) and providing more flexibility in transmission management.

The key requirements of future dynamic WDM equalizers include low insertion loss, wide bandwidth, fast equalization speed, small size, and low cost. Dynamic WDM equalizer structures include micro-opto-mechanical-system (MEMS) filters, Mach-Zehnder interferometer filters, acousto-optic filters, digital holographic filters, and liquid-crystal modulators. These structures utilize few cascaded (or parallel) optical filters, whose weights are dynamically optimized to realize smooth spectral equalization. Usually, it is hard to achieve channel-by-channel equalization unless the number of optical filters used in an equalizer subsystem is made equal to the number of WDM channels.

In [RAAE04], a dynamic WDM equalizer structure that can achieve channel-by-channel spectral equalization is proposed, while maintaining a constant insertion loss, independent of the number of WDM channels. It is based on a reflective, free-space opto-VLSI processor, which generates phase holograms to independently steer/reshape each incident WDM beam, thus realizing channel-by-channel optical attenuation. For more detailed information, we refer the reader to [Toml03, RAAE04].

2.5 Optical Amplifiers

Although an optical signal can propagate a long distance typically 80 km in current deployment before it needs amplification, optical networks, particularly for long-distance links, 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, but not to restore the shape or timing of the signal. This type of amplification is known as *1R (re-amplification)*, and it provides total data transparency (the amplification process is independent of the signal's modulation format).

However, in today's digital networks (e.g., Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH)), 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 a process is referred to as *3R* (*re-amplification*, *re-shaping*, and *re-timing*).

The *re-shaping* of the signal reproduces the original pulse shape, eliminating much of the noise. Reshaping applies primarily to digitally-modulated signals, but in some cases may also be applied to analog signals. The *re-timing* of the signal synchronizes the signal to its original bit timing pattern and bit rate. Re-timing applies only to digitally-modulated signals.

Another approach to amplification is *2R* (*re-amplification* and *re-shaping*), in which the optical signal is converted to an electronic signal which is then used to directly modulate a laser. Research and development on all-optical *3R* regeneration is a very important topic today.

Also, in a WDM system with optoelectronic regeneration, each wavelength would need to be separated before being amplified electronically, and then recombined before being retransmitted. Thus, in order 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 optical noise, as well as the signal, will be amplified. Also, the amplifier introduces spontaneous emission noise.

Optical amplification uses the principle of stimulated emission, similar to the approach used in a laser. Optical amplifiers can be divided into two basic classes: optical fiber amplifiers (OFAs) and semiconductor optical amplifiers (SOAs), which will be discussed in detail in the following section. In Table 2.4, comparison between OFAs and SOAs are presented in general. Besides, there is a new kind of optical amplifier that is *Raman amplifier.* They will be discussed in detail in the following sections.

Table 2.4 Characteristics of OFAs and SOAs.

Features	*OFAs*	*SOAs*
Maximum Internal Gain (dB)	25-30	20-25
Insertion Loss (dB)	0.1-2	6-10
Polarization Sensitivity	Negligible	< 2 dB
Nonlinear Effects	Negligible	Yes
Saturation Output Power (dBm)	13-23	5-20
Noise Figure (dBm)	4-6	7-12
Integrated Circuit Compatible	No	Yes
Functional Device Possibility	No	Yes

2.5.1 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 pump power in dB/mW, where pump is a local source of energy.

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. 2.13). 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*, e.g., in long-distance links. Each subsequent amplifier in the cascade amplifies the noise generated by previous amplifiers.

2.5.2 Semiconductor Laser Amplifier

A semiconductor laser amplifier (see Fig. 2.13) is a modified semiconductor laser, which typically has different facet reflectivity and different device length. A weak signal is sent through the active region of the semiconductor, which, via stimulated emission, results in a stronger signal emitted from the semiconductor.

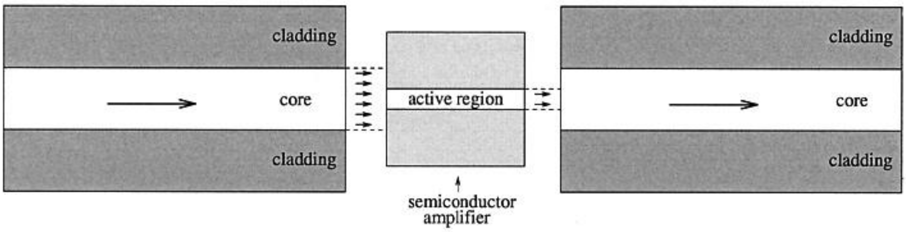


Figure 2.13 A semiconductor optical amplifier.

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 TWAs have a reflectivity of around 0.01% [Maho93]. 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, TWAs are more appropriate than Fabry-Perot amplifiers for WDM networks.

Semiconductor amplifiers based on multiple quantum wells (MQW) have been studied in [LKSK05, MTTK05, LKSK05]. 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.

Currently, SOAs attract more interest in both research and industry fields, and their advances are illustrated in [KaIw05, LOLH05, OYTB05]. One 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 current on and off, the amplifier basically acts like a gate, either blocking or amplifying the signal.

2.5.3 Doped-Fiber Amplifier

Doped-fiber amplifiers are lengths of fiber doped with an element (rare earth) which can amplify light (see Fig. 2.14). The most common doping element is erbium, which provides gain for wavelengths between 1525 nm and 1560 nm. At the end of the length of the fiber amplifier, a laser transmits a strong signal at a lower wavelength (referred to as the *pump wavelength*) to back up the fiber. This pump signal excites the doped atoms into a higher energy level. This allows the data signal to stimulate the excited atoms to release photons. Most erbium-doped fiber amplifiers (EDFAs) are pumped by lasers with a wavelength of either 980 nm 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, EDFAs have been shown to achieve gains of up to 51 dB with the maximum gain limited by internal *Rayleigh backscattering (RBS)*, which occurs when a fraction of scattered light is recaptured and back-reflected towards the launch end within the optical waveguide [HaDL92]. The 3-dB gain bandwidth for the EDFA is around 35 nm (see Fig. 2.15), and the saturation power is around 20 dBm[8] [Maho93]. A serial structured wide-band EDFA with 25 dB of flat gain over 77 nm (1528–1605 nm) and dynamic gain clamping over 13 dB of input range (25–483 W) is demonstrated experimentally in [JLWF04].

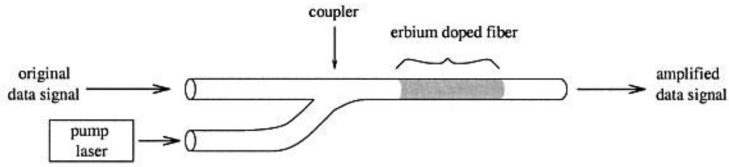


Figure 2.14 Erbium-doped fiber amplifier.

Additionally, the praseodymium-doped fluoride fiber amplifier (PDFFA) is presented in [Whit95], and Yb-Doped Fiber Amplifier is shown in [WrVa05].

A limitation to optical amplification is the unequal gain spectrum of optical amplifiers. The EDFA gain spectrum is shown in Fig. 2.15 (from [Sams97]). 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 facts that optical amplifiers amplify noise as well as signal and that the active region of the amplifier can spontaneously emit photons which also cause noise, limit 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 [TaLa91], a notch filter (a 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. 2.15). However, when multiple EDFAs are cascaded, another peak appears around the 1560-nm wavelength. In

[8]Power is commonly expressed in decibel values regarding to losses and gains. The conversion equation between dBm and mW units is:

$$P(dBm) = 10\log(P(mW))$$

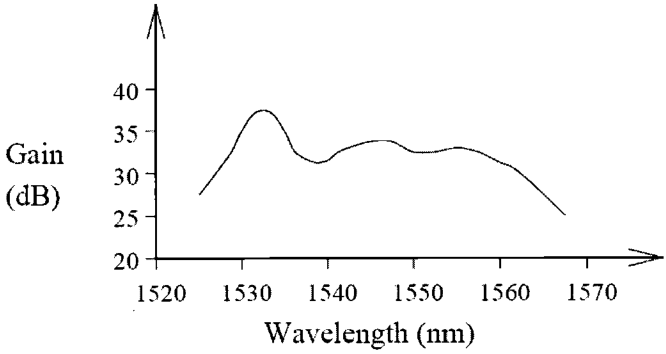


Figure 2.15 The gain spectrum of an erbium-doped fiber amplifier with input power = $-40\,\text{dBm}$.

Table 2.5 Amplifier characteristics.

Amplifier Type	*Gain Region*	*Gain Bandwidth*	*Maximum Gain*
Semiconductor	Any	40 nm	25 dB
EDFA	1525–1560 nm	35 nm	30 dB
PDFFA	1280–1330 nm	50 nm	30 dB

[WiHw93], a notch filter centered at 1560 nm is used to equalize the gain for a cascade of EDFAs. Another approach to flattening the gain is to adjust the input transmitter power such that the power on all received wavelengths at the destination is equal [ChNT92]. 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 [EGZJ93], this approach is applied to a WDM interoffice ring network.

2.5.4 Raman Amplifier

Besides the classical amplifiers above, Raman amplifiers have been deployed for new long-haul and ultra-long-haul fiber-optic transmission systems, making them one of the first widely commercialized nonlinear optical devices in telecommunications. Raman amplifiers have some fundamental advantages. First, Raman gain exists in every fiber, which provides a cost-effective means of upgrading from the terminal ends. Second, the gain is nonresonant, which is available over the entire transparency region of the fiber ranging from approximately 0.3 to 2 μm. The third advantage of Raman amplifiers is that the gain spectrum can be tailored by adjusting the pump wavelengths. For instance, multiple pump lines can be used to increase the optical bandwidth, and the pump distribution determines the gain flatness. Another advantage of Raman amplification is that it is a relatively broad-band amplifier with a bandwidth > 5 THz, and the gain is reasonably flat over a wide wavelength range [Moha02].

However, Raman gain requires more pump power, roughly tens of milliwatts per dB of gain, whereas the tenths of a milliwatt per dB is required for EDFAs for small signal powers. This disadvantage, combined with the scarcity of high-power pumps at appropriate wavelengths, slowed down the development of the Raman amplifier during the commercialization of EDFAs in the early 1990s. Then, in the mid-1990s, the development of suitable high-power pumps sparked renewed interest. Researchers were quick to demonstrate some of the advantages that Raman amplifiers have over EDFAs, particularly when the transmission fiber itself is turned into a Raman amplifier. This, in turn, fueled further advances in Raman pump technology. Now, Raman amplification is an accepted technique for enhancing system performance.

The schematic diagram of a Raman amplifier is shown in Fig. 2.16. When an optical field is incident on a molecule, the bound electrons oscillate at the optical frequency. This induced oscillating dipole moment produces optical radiation at the same frequency, with a phase shift that leads to the medium's refractive index. Simultaneously, the molecular structure itself is oscillating at the frequencies of various molecular vibrations. Therefore, the induced oscillating dipole moment also contains the sum and difference frequency terms between the optical and vibrational frequencies. These terms give rise to Raman scattered light in the re-radiated field. In a solid-state quantum-mechanical description, optical photons are inelastically scattered by quantized molecular vibrations called optical phonons. Photon energy is lost (the molecular lattice is heated) or gained (the lattice is cooled), shifting the frequency of the light. The components of scattered light that are shifted to lower frequencies are called Stokes lines, while those shifted to higher fre-

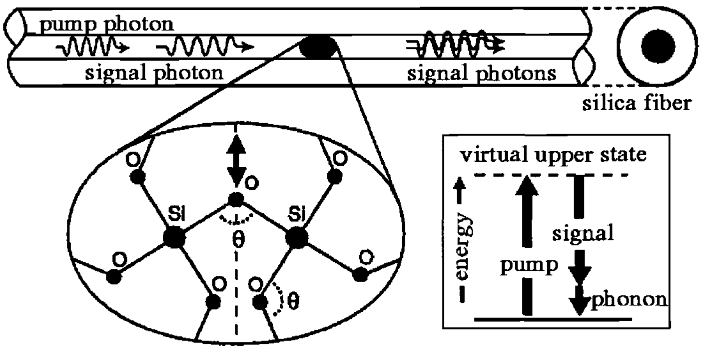


Figure 2.16 Schematic depicting amplification by stimulated Raman scattering in an optical silica fiber. The Raman Stokes interaction between a pump and signal photon and the silica molecules converts the pump into a replica of the signal photon, producing an optical phonon.

quencies are called anti-Stokes lines. The frequency shift is equal to the oscillation frequency of the lattice phonon that is created or annihilated. (The anti-Stokes process is not mentioned further in this book as it is typically orders of magnitude weaker than the Stokes process in the context of optical communications, making it irrelevant.) Raman scattering can occur in all materials, but in silica glass the dominant Raman lines are due to the bending motion of the Si-O-Si bond (see bond angle in Fig. 2.16). Raman scattering can also be stimulated by signal light at an appropriate frequency shift from a pump, leading to stimulated Raman scattering (SRS). In this process, pump and signal light are coherently coupled by the Raman process. In a quantum-mechanical description, shown in the energy-level diagram in Fig. 2.16, a pump photon is converted into a second signal photon that is an exact replica of the first one, and the remaining energy produces an optical phonon. The initial signal photon, therefore, has been amplified. This process is considered nonresonant because the upper state is a short-lived virtual state.

Raman amplification has enabled dramatic increases in the reach and capacity of lightwave systems. Novel Raman pumping schemes that have recently been developed are highlighted in [Jake04].

2.6 Switching Elements

Obviously, *switching elements* are essential component of any network. The concept of switching originated from electronics field. Thus, according to the signal carriers, there are *optical switching* and *electronic switching.* In the switching granularity point of view, there are two basic classes: *circuit switching* and *cell switching.* In optical field, *circuit switching* is corresponding to *wavelength routing,* and *cell switching* is *optical packet switching* and *optical burst switching.* As far as the transparency of signals is considered, there are opaque switching and transparent switching. In the section, switching devices are classified into two basic classes: logic switching and relational switching.

Logic switching is performed by a device in which the data (or the information-carrying signal) 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. In a logic device, format and rate of data would be changed or converted in intermediate nodes, thus, *logic switching* is also sometimes referred to *opaque switching.* Furthermore, some of its components must be able to change states or *switch* as fast as or faster than the signal bit rate [Hint90]. This ability gives the device some added flexibility but limits the maximum bit rate that can be accommodated. Based on the logic device available and ideal performance in electronic field, logic switching is primarily employed in electronic field. But, traditional optical-electronic-optical (o-e-o) conversion in today's optical networks is still widely applied due to the lack of counterpart logic devices in the optical field. It means that most current optical 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, i.e., logic devices. It provides a high degree of flexibility in terms of switching and routing functions for optical networks; however, the speed of electronics 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 and cost. These factors above 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. The emergence of *relational switching* is derived from this push.

Relational switching is to establish 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. In the class of switching devices currently being developed, the control of the switching function is performed electronically with the optical stream being transparently routed from a given input of the switch to a given output. Such transparent switching allows the switch to be independent of the data rate and format of the optical signals. 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 *transparent switching.* Due to the limits of optical hardware, various kinds of optical switching devices basically employ *relational switching. Relational switching* provides more advantages for optical networks in terms of the optical hardware limits. In the following sections, we review a number of different optical switching elements and architectures which are typically developed in today's optical networks.

2.6.1 Optical Add-Drop Multiplexers (OADM)

Optical Add/Drop Multiplexers (OADMs) are elements that provide capability to add and drop traffic in the network (similar to SONET ADMs). They are located at sites supporting one or two (bidirectional) fiber pairs and enable a number of wavelength channels to be dropped and added, thereby reducing the number of unnecessary optoelectronic conversions, without affecting the traffic that is transmitted transparently through the node (see Fig. 2.17).

An OADM can be used in both linear and ring network architectures and in practice operates in either fixed or reconfigurable mode. In fixed OADMs, the add/drop and through channels are predetermined and can only be manually rearranged after installation. In reconfigurable OADMs, the channels that are added/dropped or pass transparently through the node can be dynamically reconfigured as required by the network. These are more complex structures but more flexible as they provide on-demand provisioning without manual intervention; therefore, they can be set up on the fly. The reduction of unnecessary optoelectronic conversions through the use of OADMs introduces significant cost savings in the network.

Reconfigurable OADMs can be divided into two categories: one is partly reconfigurable architecture and the other is fully reconfigurable architecture. In partly-reconfigurable architectures, there is capability to select the chan-

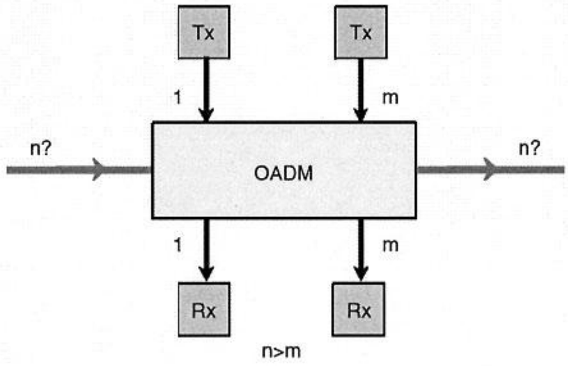
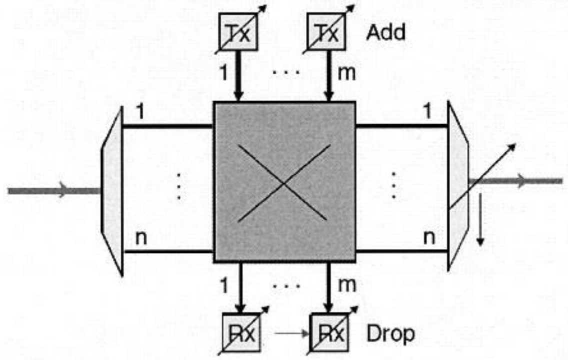


Figure 2.17 Generic OADM architecture.

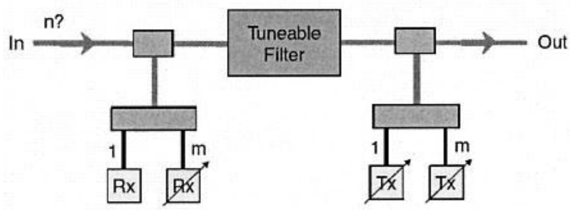
nels to be added/dropped, but there is also a predetermined connectivity matrix between add/drop and through ports, restricting the wavelength-assignment function. Fully-reconfigurable OADMs provide the ability to select the channels to be added/dropped, but they also offer connectivity between add/drop and through ports, which enables flexible wavelength assignment with the use of tunable transmitters and receivers. Reconfigurable OADMs can be divided into two main generations. The first is mainly applied in linear network configurations and does not support optical path protection, while the second is applied in ring configurations and provides optical layer protection.

The two most common examples of fully-reconfigurable OADMs, i.e., *wavelength selective* (WS) and *broadcast selective* (BS) architectures, are illustrated in Figs. 2.18(a) and 2.18(b). The WS architecture utilizes wavelength demultiplexing/multiplexing and a switch fabric interconnecting express and add/drop ports, while the BS is based on passive splitters/couplers and tunable filters. The overall loss introduced by the through path of the BS solution is noticeably lower than the loss of the WS approach, significantly improving the optical signal-to-noise ratio (OSNR) of the node and therefore its concatenation performance in a practical transmission link or ring. In addition, the BS design offers superior filter concatenation performance, advanced features such as drop and continue, and good scalability in terms of add/drop percentage.

For more detailed information, we refer the reader to [TzZT04].



(a) WS OADM architecture



(b) BS OADM architecture

Figure 2.18 Fully-reconfigurable WS and BS OADM architectures [TzZT04].

2.6.2 Optical Cross-Connect (OXC)

An *optical crossconnect (OXC)* 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. 2.19). In the cross state, the signal from the upper input port is routed to the lower 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

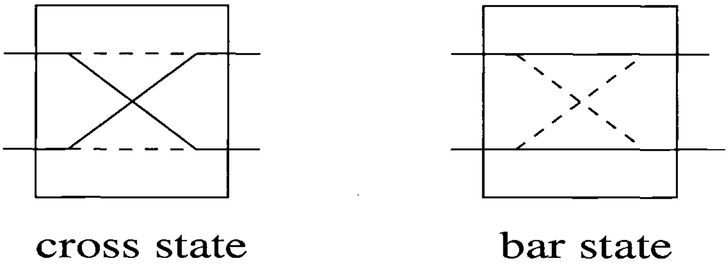


Figure 2.19 2 × 2 crossconnect elements in the cross state and bar state.

upper output port, and the signal from the lower input port is routed to the lower output port.

To support flexible path provisioning and network resilience, OXCs normally utilize a switch fabric to enable routing of any incoming channels to the appropriate output port and access to the local client traffic. The features that an OXC should ideally support are similar to these of an OADM, but OXCs need to additionally provide:

(1) strictly nonblocking connectivity between input and output ports,

(2) span and ring protection as well as mesh restoration capabilities.

A number of OXC solutions based on different technologies have been proposed to date and, depending on the switching technology and the architecture used, they are commonly divided into two main categories: opaque and transparent [TzZT04].

Opaque OXCs are either based on electrical switching technology or on optical switch fabrics surrounded by optical-electrical-optical (OEO) conversions, imposing the requirement of expensive optoelectronic interfaces. In OXCs using electrical switching, subwavelength switching granularities can be supported by providing grooming capabilities for more efficient bandwidth utilization. Opaque OXCs also offer inherent regeneration, wavelength conversion, and bit-level monitoring.

In transparent OXCs, the incoming signals are routed through an optical switch fabric without the requirement of optoelectronic conversions, thereby offering transparency to a variety of bit rates and protocols. The switching granularity may vary and support switching at the fiber level, the wavelength band level, or the wavelength channel level.

In OXC architectures, optical crosspoint elements have been demonstrated using two types of technologies:

(1) the generic directive switch [Alfe88], in which light is physically directed to one of two different outputs, and

(2) the gate switch, in which optical amplifier gates are used to select and filter input signals to specific output ports.

Directive Switches

(a)
(b)
(c)
(d)

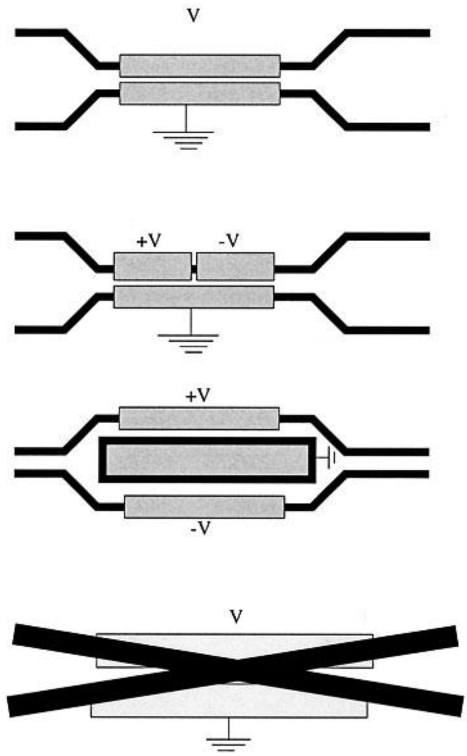


Figure 2.20 Schematic of optical crosspoint elements.

The directional coupler (Fig. 2.20(a) [Alfe88]) consists of a pair of optical channel waveguides that are parallel and in close proximity over a finite interaction length. Light input to one of the waveguides couples to the second waveguide via evanescent[9] coupling. The coupling strength corresponds to the interwaveguide separation and the waveguide mode size which depends on the optical wavelength and confinement factor[10] of the waveguide. If the two waveguides are identical, complete coupling between the two waveguides occurs over a characteristic length which depends on the coupling strength. However, by placing electrodes over the waveguides, the difference in the propagation constants in the waveguides can be sufficiently increased so that no light couples between the two waveguides. Therefore, the cross state corresponds to zero applied voltage, and 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. 2.20(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 [Alfe88].

The balanced bridge interferometric switch (see Fig. 2.20(c)) consists of an input 3-dB coupler, two waveguides sufficiently separated so that they do not couple, electrodes to allow changing 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 electrooptically produce 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. 2.20(d). This device can be viewed as a directional coupler (see Fig. 2.20(a)) with no gap between the waveguides in the interaction region. When properly fabricated, both cross and bar states can be electrooptically achieved with good crosstalk

[9]An evanescent wave is the part of a propagating wave which travels along or outside of the waveguide boundary.

[10]The confinement factor determines the fraction of optical power that travels within the core of the waveguide.

performance.

Other types of switches include the mechanical fiber-optic switch and the thermo-optic 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 [Ande95]. Thermo-optic waveguide switches, on the other hand, are fabricated on a glass substrate and are operated by the use of the thermo-optic effect. One such device uses a zero-gap directional-coupler configuration with a heater electrode to increase the waveguide index of refraction [LeSu94].

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 are then recombined in $N\times$ combiners and sent to the N outputs. The gate elements can be implemented using optical amplifiers which can either be turned 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 2×2 amplifier gate switch is illustrated in Fig. 2.21. A disadvantage of the gate switch is that the splitting and combining losses limit the size of the switch.

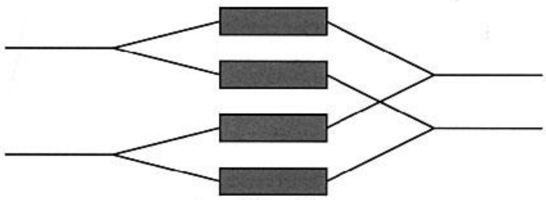


Figure 2.21 A 2×2 amplifier gate switch.

2.6.3 Clos Architecture

For more than 50 years, the Clos architecture [Clos53] has been the hands-down favorite for building multi-stage TDM switching systems. It has maintained this status, independent of the advances in switching fabrics over that same time period. It can realize the fewest switching crosspoints for the largest range of scalability while providing strict or rearrangably non-blocking traffic paths. In Fig. 2.22, 3-stage Clos architecture is presented, which is a typical application of an OXC.

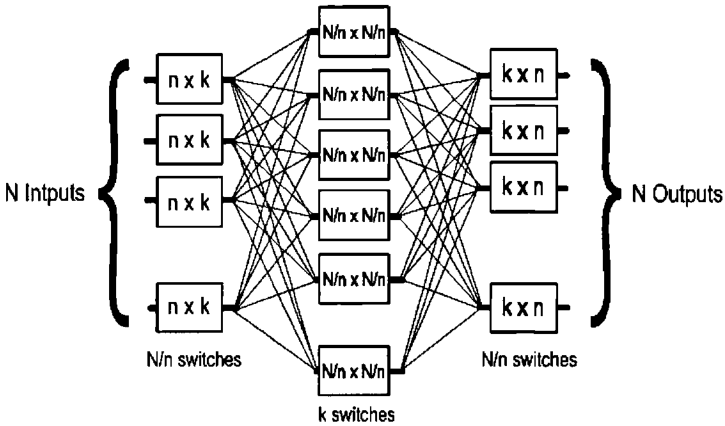


Figure 2.22 A 3-stage Clos architecture.

In Fig. 2.22, the number of 2nd stage switches depends on blocking: in fully non-blocking, $k \geq 2n - 1$; in rearrangeably non-blocking, $k \geq n$.

Advanced development of *3-stage Clos Cross-connect Switch Architecture* with up to 2048 $\times$ 2048 ports and 10 Gbps per port is presented by some vendors in 2005.

2.6.4 Micro-Electro Mechanical Systems (MEMS)

Currently, *micro-electro mechanical systems (MEMS)* is widely believed to be the most promising method for large-scale optical cross-connects. Optical MEMS-based switches are distinguished in being based on mirrors, membranes, and planar moving waveguides. The former two are free-space switches; the latter are waveguide switches. Furthermore, MEMS-based switches are classified into the two major approaches, i.e., 2-Dimensional and 3-Dimensional approaches. Among these classifications, the 3D optical MEMS based on mirrors is popular because it is suitable for compact, large-scale switching fabrics. The ability of this architecture to achieve input- and output-port counts of over one thousand is the primary driver of the large-scale OXCs, in which spatial parallelism is utilized. In particular, the type of switch provides high application flexibility in network design because of low and uniform insertion loss with low wavelength dependency under various operating conditions. Furthermore, this switch exhibits minimal degradation of the optical signal-to-noise ratio, which is mainly caused by crosstalk, polarization dependent loss (PDL), and chromatic and polarization mode dispersions.

Figure 2.23 shows the basic configuration of the 3D MEMS optical switch. The optical signals passing through the optical fibers at the input port are switched independently by the gimbal-mounted MEMS mirrors with two-axis tilt control and then focused onto the optical fibers at the output ports. In the switch, any connection between input and output fibers can be accomplished by controlling the tilt angle of each mirror. As a result, the switch can handle several channels of optical signals directly without costly optical-electrical or electrical-optical conversion. The 3D MEMS-based O-O-O switch has been built in sizes ranging from 256 $\times$ 256 to 1000 $\times$ 1000 bi-directional port machines. In addition, encouraging research seems to show that 8000 $\times$ 8000 ports will be practical within the foreseeable future. The port count, however, is only one dimension to the scalability of an O-O-O switch. An O-O-O switch is also scalable in terms of throughput. A truly all-optical switch is bit-rate and protocol independent. The combination of thousands of ports and bit-rate independence results in a theoretically future-proof switch with unlimited scalability.

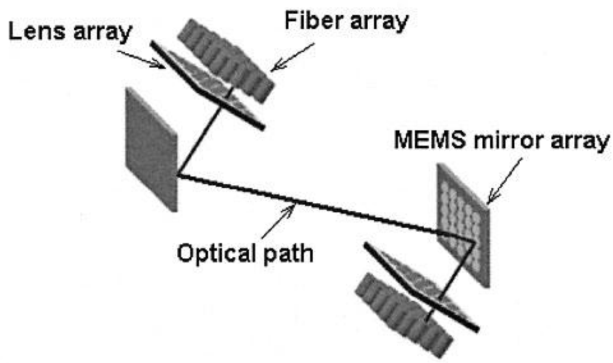


Figure 2.23 Schematic diagram of a 3D MEMS optical switch.

Recent developments of 3D MEMS for optical switching are shown in [YaYT05]. They can lead to compact and stable optical crossconnect switches for simple, fast, and flexible wavelength applications in today's optical networks.

Optical MEMS are miniature devices with optical, electrical, and mechanical functionalities at the same time, fabricated using batch process techniques derived from microelectronic fabrication. Optical MEMS provide intrinsic characteristics for very low crosstalk, wavelength insensitivity, polarization insensitivity, and scalability. Comparison of optical MEMS vs. other OXC switching elements are presented in [MaKu03].

2.6.5 Nonreconfigurable Wavelength Router

In this and next sections, two classes of wavelength routers are introduced, which are in fact circuit switches in the optical field. 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 2.6.6 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 4×4 nonreconfigurable wavelength router with $M = 4$ is illustrated in Fig. 2.24. The router is nonreconfigurable because the path of a given wavelength channel, after it enters the router on a particular input fiber, is fixed. The wavelengths on each incoming fiber are separated using a grating demultiplexer. And finally, information from multiple WDM channels are multiplexed before launching them 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.

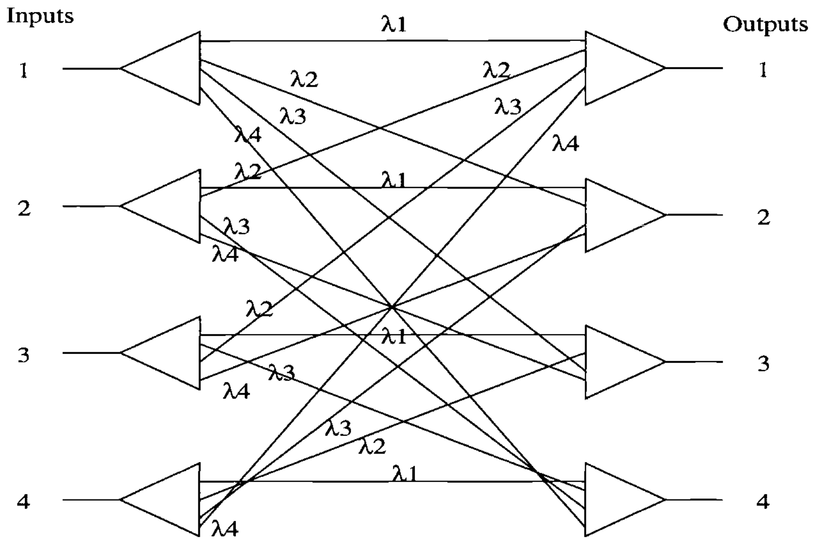


Figure 2.24 A 4 × 4 nonreconfigurable wavelength router.

Arrayed Waveguide Grating

One implementation of a wavelength router is the arrayed waveguide grating (AWG) multiplexer. An AWG 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 AWG with N input and N 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 AWG is an integrated device, it can easily be fabricated at low cost. The disadvantage of the AWG is that it is a device with a fixed routing matrix which cannot be reconfigured.

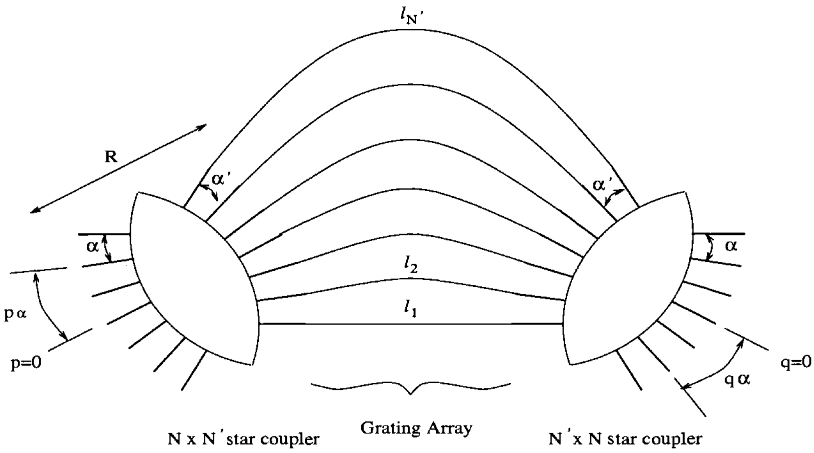


Figure 2.25 The arrayed waveguide grating (AWG).

The AWG, shown in Fig. 2.25, can be used as a nonreconfigurable wavelength-router, or it can be used 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 α, and their outputs are separated by angular distance α'. 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 Δl.

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 $\beta \times \Delta l$ in adjacent waveguides, where $\beta = 2\pi n_{\text{eff}}/\lambda$ is the propagation constant in the waveguide, n_{eff} is the effective refractive index of the waveguide, and λ 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 $kR(p-q)\alpha\alpha'$, where k is a propagation constant which 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:

$$\phi = \frac{2\pi \times \Delta l}{\lambda} + kR(p-q)\alpha\alpha' \qquad (2.11)$$

The transmission power from a particular input port p to a particular output port q is maximized when ϕ is an integer multiple of 2π. Thus, only wavelengths λ for which ϕ is a multiple of 2π 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 which causes ϕ to be a multiple of 2π.

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 [DrEK91]. In [ZiDJ92], a 15 × 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 [OkMS95], a 64 × 64 arrayed-waveguide multiplexer on silicon is demonstrated with a channel spacing of 0.4 nm. WGRs with flat passbands have also been developed [OkSu96, TrBe97]. Other applications of the AWG, such as tunable transmitters and tunable receivers, are presented in [GlKW94]. These tunable components are implemented by integrating the AWG with switched amplifier elements. An amplifier element may either be activated, in which case it amplifies the signal passing through it, or it may be turned off, in which case it prevents any signals from passing through it. By using only a single input port of the AWG, 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 AWG, 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.

2.6.6 Reconfigurable Wavelength-Routing Switch

A reconfigurable wavelength-routing switch (WRS), also referred to as a wavelength selective crossconnect (WSXC), uses photonic switches inside the routing element. The functionality of the reconfigurable WRS, illustrated in Fig. 2.26, 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 M $P \times P$ optical switches between the demultiplexer and the 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. Finally, information streams from multiple WDM channels are multiplexed before launching them back onto an output fiber.

Space-division optical-routing switches may be built from 2×2 optical crosspoint elements [ScAl90] arranged in a banyan-based fabric. The space-division switches (which may be one per wavelength [ShCS93]) can route a signal from any input to any output on a given wavelength. Such switches based on relational devices [Hint90] are capable of switching very high-capacity signals. The 2×2 crosspoint elements that are used to build the space-division switches may be slowly tunable and they 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, as well as the configuration of the switches in the network nodes.

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. Besides *wavelength routing*, *Optical Packet Switching* (OPS) and *Optical Burst Switching* (OBS) have attracted interest in both research and industry. In OPS and OBS, the data path is fully optical, but the control of the switching operation is performed electronically. By employing the electronic control, relationship of data switching is established in optical domain. Essentially, OBS and OPS are based on *relational switching*, which will be studied in Chapters 17 and 18.

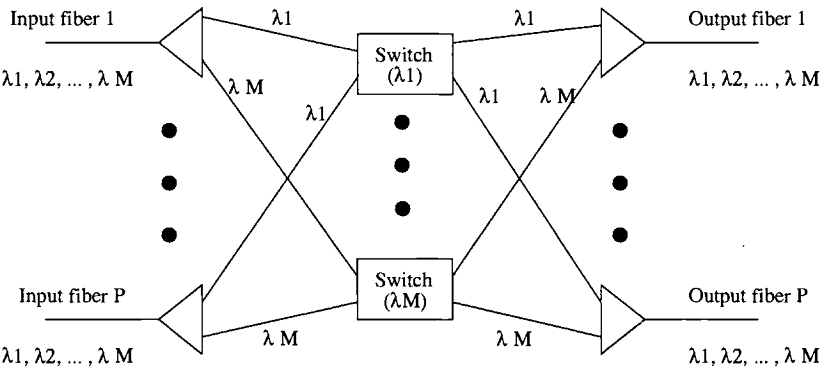


Figure 2.26 A P × P reconfigurable wavelength-routing switch with M wavelengths.

2.7 Wavelength Conversion

Consider the network in Fig. 2.27. It shows a wavelength-routed network containing two WDM crossconnects (S1 and S2) and five access stations (A through E). Three lightpaths have been set up (C to A on wavelength λ_1, C to B on λ_2, and D to E on λ_1). 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 [BaMu96]). 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. 2.28(a). Two lightpaths have been established in the network: (i) between Node 1 and Node 2 on wavelength λ_1, and (ii) between Node 2 and Node 3 on wavelength λ_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-continuity network may suffer from higher blocking as compared to a circuit-switched network.

It is 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 referred to as *wavelength conversion*. In Fig. 2.28(b), a wave-

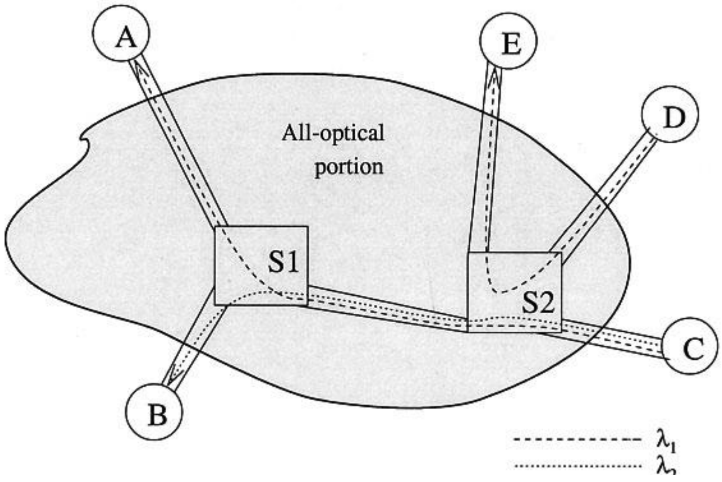


Figure 2.27 An all-optical wavelength-routed network.

length converter at Node 2 is employed to convert data from wavelength λ_2 to λ_1. The new lightpath between Node 1 and Node 3 can now be established by using the wavelength λ_2 on the link from Node 1 to Node 2, and then by using the wavelength λ_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 2.8.6.

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. 2.29). In this figure and throughout this section, λ_s denotes the input signal wavelength; λ_c, the converted wavelength; λ_p, the pump wavelength; f_s, the input frequency; f_c, the converted frequency; f_p, the pump frequency; and CW, the continuous wave (unmodulated) generated as the pump signal.

An ideal wavelength converter should possess the following characteristics [DMJD96]:

- transparency to bit rates and signal formats,

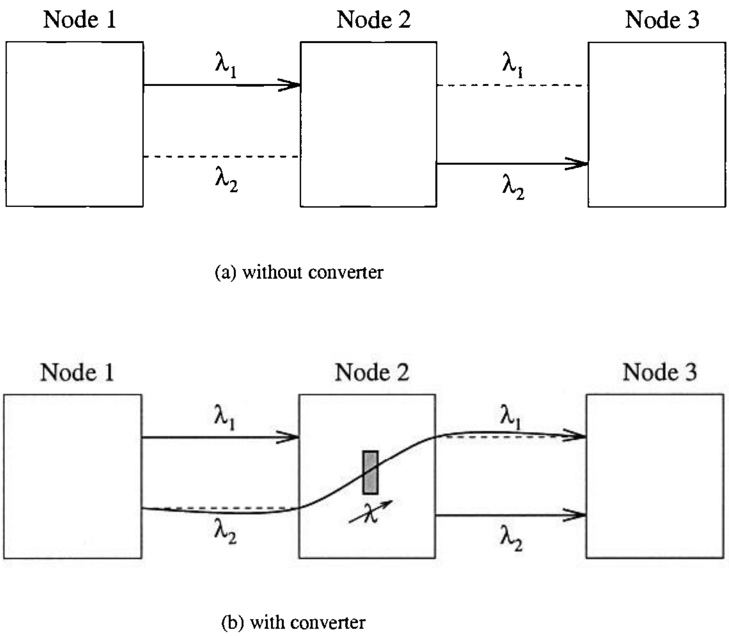


Figure 2.28 Wavelength-continuity constraint in a wavelength-routed network.

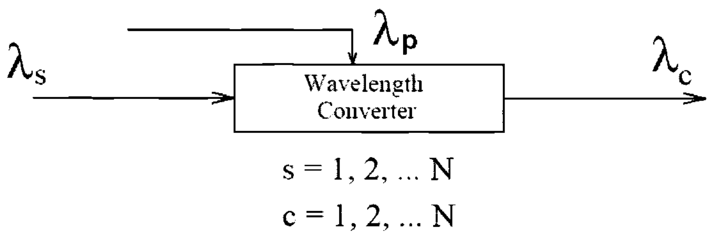


Figure 2.29 Functionality of a wavelength converter.

- 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[11] and large signal-to-noise ratio, and
- simple implementation.

2.7.1 Wavelength Conversion Technologies

Several researchers have attempted to classify and compare the several techniques available for wavelength conversion [DMJD96, MDJD96, SaIa96, Wies96, Yoo96]. The classification of these techniques presented in this section follows that in [Wies96]. Wavelength conversion techniques can be broadly classified into two types: *opto-electronic 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 which employ *coherent effects* and techniques which use *cross modulation*.

Opto-Electronic Wavelength Conversion

In opto-electronic wavelength conversion [Fuji88], the optical signal to be converted is first translated into the electronic domain using a photodetector (labeled R in Fig. 2.30, from [Mest95]). 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. 2.30). This method has been demonstrated for bit rates up to 10 Gbps [Yoo96]. However, this method is much more complex and consumes a lot more power than the

[11] The *extinction ratio* is defined as the ratio of the optical power transmitted for a bit "1" to the power transmitted for a bit "0."

other methods described below [DMJD96]. Moreover, the process of opto-electronic (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.

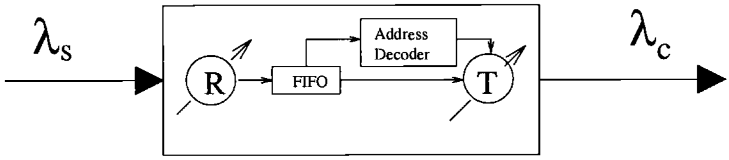


Figure 2.30 An opto-electronic wavelength converter.

Wavelength Conversion Using Coherent Effects

Wavelength conversion methods using coherent effects are typically based on wave-mixing properties(see Fig. 2.31). 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 signal with high-bit rates. In Fig. 2.31, the value $n = 3$ corresponds to Four-Wave Mixing (FWM) and $n = 2$ corresponds to Difference Frequency Generation (DFG). These techniques are described below.

- *Four-Wave Mixing (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 [TCFG95] to generate a fourth wave of frequency given by:

$$f_{ijk} = f_i \pm f_j \pm f_k$$

  FWM is also achievable in other passive waveguides such as semiconductor waveguides and in an active medium such as a semiconductor

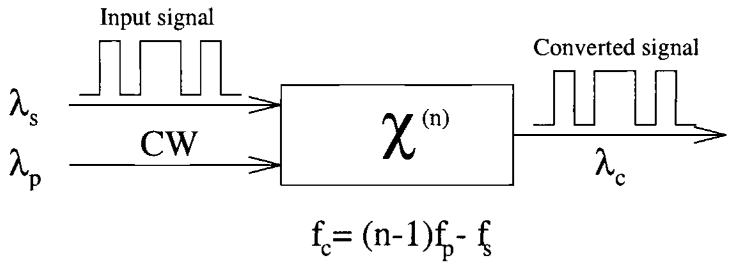


Figure 2.31 A wavelength converter based on nonlinear wave-mixing effects.

optical amplifier (SOA). Four-wave mixing (FWM) is a promising technique for wavelength conversion in optical networks owing to its ultrafast response and high transparency to bit rate and modulation format. In [CSLB05], a 3-dB conversion range over 40 nm (1535C1575 nm) is obtained with a flat conversion efficiency of -16 dB and a polarization sensitivity of less than 0.3 dB.

- *Difference Frequency Generation (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 [Yoo96]. 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 [YCBK95]. 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 [YCBK96] and in fabricating a low-loss waveguide for high conversion efficiency [Yoo96].

Wavelength Conversion Using Cross Modulation

Cross-modulation wavelength conversion techniques utilize active semiconductor optical devices such as semiconductor optical amplifiers (SOAs) and lasers. These techniques belong to a class known as optical-gating wavelength conversion [Yoo96].

- *Semiconductor Optical Amplifiers (SOAs) in XGM and XPM mode*: The principle behind using an SOA in the cross-gain modulation

(XGM) mode is shown in Fig. 2.32 (from [DMJD96]). The intensity-modulated input signal modulates the gain in the SOA due to gain saturation. A continuous-wave (CW) signal at the desired output wavelength (λ_c) is modulated by the gain variation so that it carries the same information as the original input signal. The CW signal can either be launched into the SOA in the same direction as the input signal (co-directional), or launched into the SOA in the opposite direction as the input signal (counter-directional). 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 Gbps [DMJD96], it suffers from the drawbacks due to inversion of the converted bit stream and extinction ratio degradation for the converted signal. But, new advanced developments demonstrate that no signal inversion occurs in [KLKH05].

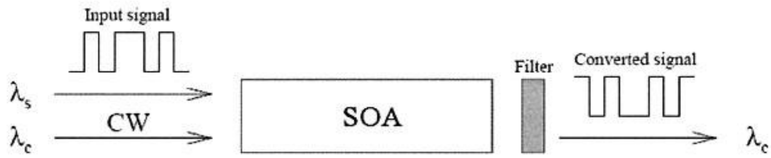


Figure 2.32 A wavelength converter using co-propagation based on XGM in an SOA.

The operation of a wavelength converter using SOA in cross-phase modulation (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 results in phase modulation of a CW signal (wavelength λ_c) coupled into the converter [DMJD96, LaPT96]. The SOA can be integrated into an interferometer so that an intensity-modulated signal format results at the output of the converter. Techniques involving SOAs in XPM mode have been proposed using nonlinear optical loop mirrors (NOLMs) [EiPW93], Mach-Zender interferometers (MZI) [DJMP94] and Michelson interferometers (MI) [MDJP94]. Figure 2.33 shows an asymmetric MZI wavelength converter based on SOA in XPM mode (from [DMJD96]). 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 [DMJD96]. A signal up-conversion utilizing an XPM effect in an all-optical SOA-MZI wavelength converter is demonstrated in [SoLS04], in which this scheme not only shows high conversion efficiency, polarization immunity, and no increase in phase noise, but also linear signal up-conversion with a low optical power requirement.

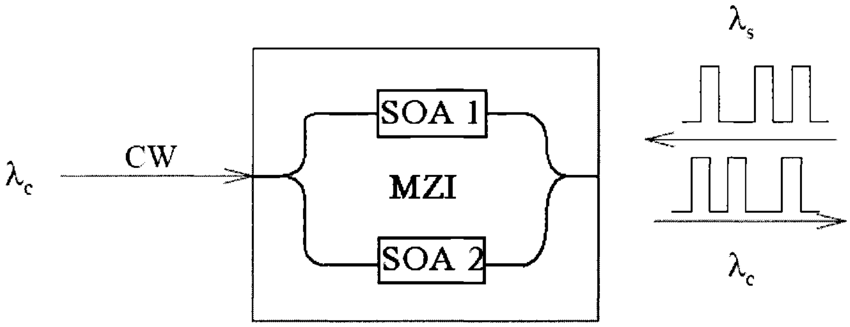


Figure 2.33 An interferometric wavelength converter based on XPM in SOAs.

- *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 Distributed Bragg Reflector (DBR) laser to convert signals at 10 Gbps [YSIY96]. In the method using saturable absorption in lasers (e.g., [KOYM87]), the input signal saturates the absorption of carrier transitions near the band gap and allows the probe beam to transmit (see Fig. 2.34). This technique shows a bandwidth limit of 1 GHz due to carrier recombinations [Yoo96].

- *All-Optical Wavelength Conversion Based on XPM in Optical Fiber*: In particular, besides the classical methods described above, wavelength

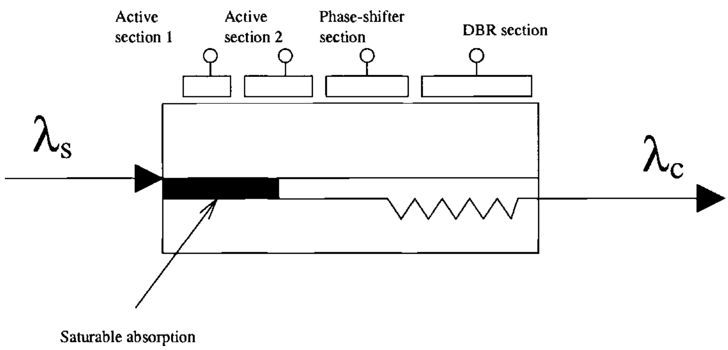


Figure 2.34 Conversion using saturable absorption in a laser.

conversion based on the XPM effect in fibers and subsequent optical filtering has drawn special attention recently.

The principle of the new technology relies on the generation of a frequency comb through cross-phase modulation (XPM) in an optical fiber with subsequent optical filtering of the desired tone. When incoming data copropagates through the fiber with a sinusoidally intensity modulated high-power optical pump signal, the data acquires a sinusoidal phase modulation from the pump through XPM, which generates multiple wavelengths spaced at the modulation frequency on both sides of the incoming signal wavelength, i.e., generating sidebands around the incoming signal wavelength. By filtering out a portion of the sidebands, a wavelength-converted signal at the output is obtained. These new components, the strengths of which depend on the phase-modulation index (i.e., the optical pump power), are imprinted with the information carried by the original wavelength without any distortion, assuming the signal bandwidth is much smaller than the pump modulation frequency. Compared to other techniques, this method has several important advantages: it is environmentally stable, operates in a wide signal wavelength range, and provides a short switching window with relatively broad signal pulses. Moreover, the scheme offers additional attractive functions such as waveform reshaping and phase-reconstruction.

A simple wavelength converter based on sinusoidal cross-phase modulation in optical fibers, which offers format- and bit-rate independence is demonstrated in [MaAT05]. Conversion to more than 20 channels is demonstrated and error-free conversion of data with bit rates of up to 10 Gbps is also reported. In [TSKK05], polarization-independent all-optical wavelength conversion is demonstrated by using cross-phase modulation in a twisted fiber and subsequent optical filtering. By twisting a 1-km-long fiber at a rate of 15 turns/m and aligning the probe light to circular polarization, the polarization sensitivity is successfully reduced from 3.5 to 0.3 dB. Error-free operation with only 1-dB penalty is realized at 40 Gbps with the input signal polarization scrambled. The demonstrated scheme also offers a novel function of restoring the degree-of-polarization of the signal light. The experimental setup is shown in Fig. 2.35.

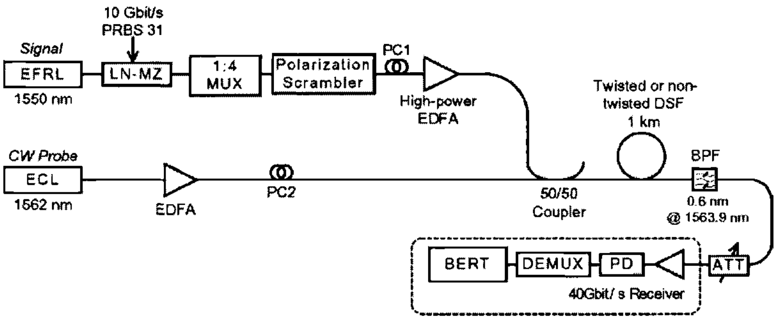


Figure 2.35 Experimental setup for demonstrating the polarization-insensitive 40-Gbps WC based on fiber XPM and optical filtering. ATT: Variable attenuator, PD: Photodetector, DEMUX: Electrical demultiplexer, BERT: BER tester.

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. However, it is clear that opto-electronic converters offer only limited digital transparency. Moreover, deploying multiple opto-electronic converters in a WDM intermediate node requires sophisticated packaging to avoid crosstalk among channels. This leads to increased cost per converter, further making this technology less attractive than all-optical converters [Yoo96]. Other disadvantages of opto-electronic converters include complexity and large power consumption [DMJD96]. Regarding all-optical wavelength conversion, there are a lot of various schemes developed in which each has own strength and week in application. Besides all-optical converters mentioned above, another kind of wavelength conversion is based on crystal material, which has also attracted more attention currently. Specially, *periodically-poled* $LiNbO_3$ *(PPLN)* waveguide is preferential in this kind of all-optical wavelength conversion scheme in [YuGu04].

2.8 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.

2.8.1 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 doesn't scale very well. In wide-area networks (WANs), the objective is often 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. In addition, channel spacing is standardized by the ITU grid.

The bandwidth of the fiber medium, as mentioned in Section 2.2, is limited to the low-attenuation regions around 1310 nm and 1550 nm. These regions have bandwidths of approximately 200 nm (25 THz) each. However, optical networks may not necessarily be able to take advantage of this entire range due to the bandwidth limitations of optical components. Amplifiers have an optical bandwidth of around 35 to 40 nm, 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 which affect the channel spacing are the channel bit rates, the optical power budget, nonlinearities in the fiber, and the 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 we would like to design a WDM passive-star-based network for N nodes.

Let ΔT be the tuning range of the transmitters and let ΔR be the tuning range of the receivers (both are measured in nm). The available transceiver bandwidth, 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 BW_T is

$$\Delta f = \frac{cBW_T}{\lambda^2}$$

where λ is the "center" wavelength.

If we want each channel to have a bit rate of B Gbps, $2B$ GHz of bandwidth will be needed for encoding, assuming a modulation efficiency of 2 Hz/bps. According to [Brac90], a channel spacing of at least 6 times the channel bit rate is needed to minimize crosstalk 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 [TrMu96, TrMu97]. A higher number of channels may provide more network capacity, but it also results in higher network costs, and in some cases, may require more complex protocols.

2.8.2 Power Considerations

In any network, it is important to maintain adequate signal-to-noise ratio (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 countered 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 [Gree93]. The first application is as a *transmitter power booster*, which is placed immediately after the transmitter in order 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 EDFAs for various applications is given in [DeNa95]. Table 2.6 summarizes some of these requirements.

Table 2.6 Requirements for EDFA applications.

EDFA Application	*Gain*	*Noise Figure*	*Saturated Output*
Transmitter power booster	Moderate	Moderate	High
Preamplifier	High	Low	Moderate
In-line amplifier	High	Low	High

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 which guarantees that all signals are adequately amplified while minimizing the total number of amplifiers being used.

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; however, if the signal power is too high, the signal combined with ASE noise may saturate the amplifiers. The unequal gain spectrum of the EDFA places limitations on the usable bandwidth in WDM systems. When multiple EDFAs 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 EDFAs [Maho95]. That is just the reason that we proposed the elimination of all-optical cycle in [Mukh97], which will be further discussed in next section.

Although recent developments in amplifier technology have solved many of the power-loss and noise problems 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 (WRS) or the arranged waveguide grating (AWG), instead of wavelength-independent devices, such as the amplifier gate switch or the passive-star coupler.

2.8.3 All-optical Cycle of Elimination

Obviously, all-optical cycle of elimination is a key issue in transparent optical networks. *All-optical cycle* is referred to a situation in which there exists the possibility of setting up unintended all-optical cycles in the optical network (i.e., a loop with no terminating electronics in it).

A transparent (wide-area) wavelength-routed optical network may be constructed by using wavelength cross-connect switches connected together by fiber to form an arbitrary mesh structure. The network is accessed through electronic stations that are attached to some of these cross-connects. These wavelength cross-connect switches have the property that they may configure themselves into unspecified states. Each input port of a switch is always connected to some output port of the switch whether or not such a connection is required for the purpose of information transfer.

The presence of these unspecified states result possibly in such an all-optical cycle. If such a cycle contains amplifiers [e.g., Erbium-Doped Fiber Amplifiers (EDFA's)], there exists the possibility that the net loop gain is greater than the net loop loss. The amplified spontaneous emission (ASE) noise from amplifiers can build up in such a feedback loop to saturate the amplifiers and result in oscillations of the ASE noise in the loop.

Such all-optical cycles as defined above must be eliminated from an optical network in order for the network to perform any useful operation. Furthermore, for the realistic case in which the wavelength cross-connects result in signal crosstalk, there is a possibility of having closed cycles with oscillating crosstalk signals.

Some algorithms are proposed in [Mukh97], which attempt to find a route for a connection and then (in a postprocessing fashion) configure switches such that all-optical cycles that might get created would automatically get eliminated. In addition, call-set-up algorithms are proposed, which avoid the possibility of crosstalk cycles in [Mukh97].

2.8.4 Additional Considerations

Other device considerations in the design of WDM networks include *crosstalk*, *dispersion*, and so on.

Crosstalk may either be caused by signals on different wavelengths (interband crosstalk or hetero-wavelength), or by signals on the same wavelength on another fiber (intraband crosstalk or home-wavelength) [Maho95]. Interband crosstalk must be considered when determining channel spacing. In some cases, intraband crosstalk may be removed through the use of appropriate narrowband filters. Intraband crosstalk usually occurs in switching nodes where multiple signals on the same wavelength are being switched from different inputs to different outputs. This form of crosstalk is more of a concern than interband crosstalk because intraband crosstalk cannot be removed through filtering, and may accumulate over a number of nodes. The degree of intraband crosstalk depends in part on the switch architecture.

As was mentioned in Section 2.2.4, dispersion in an optical communication 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 standards on optical wavelengths and channel spacing (e.g., ITU-T) will influence the design of the network components.

2.8.5 Elements of Local-Area WDM Network Design

A local area WDM network will typically consist of a number of nodes which 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.

The Network Medium

The simplest and most popular interconnection device for a local-area WDM network is the passive-star coupler which provides a broadcast medium (see Fig. 2.36). The broadcast capability of the star coupler combined with multiple WDM channels allows for a wide range of possible media access protocols [Mukh92a, Mukh92b]. 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.

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-tuned or tunable – to place at each node. These decisions usually depend on the protocol, degree of access, and connectivity desired in the network, as well as on practicality and cost considerations.

A WDM network protocol may either be a single-hop protocol [Mukh92a], in which communication takes place directly between two nodes without being routed through intermediate nodes; or a multihop protocol [Mukh92b], in which information from a source node to a destination node may be routed through the electronics at 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

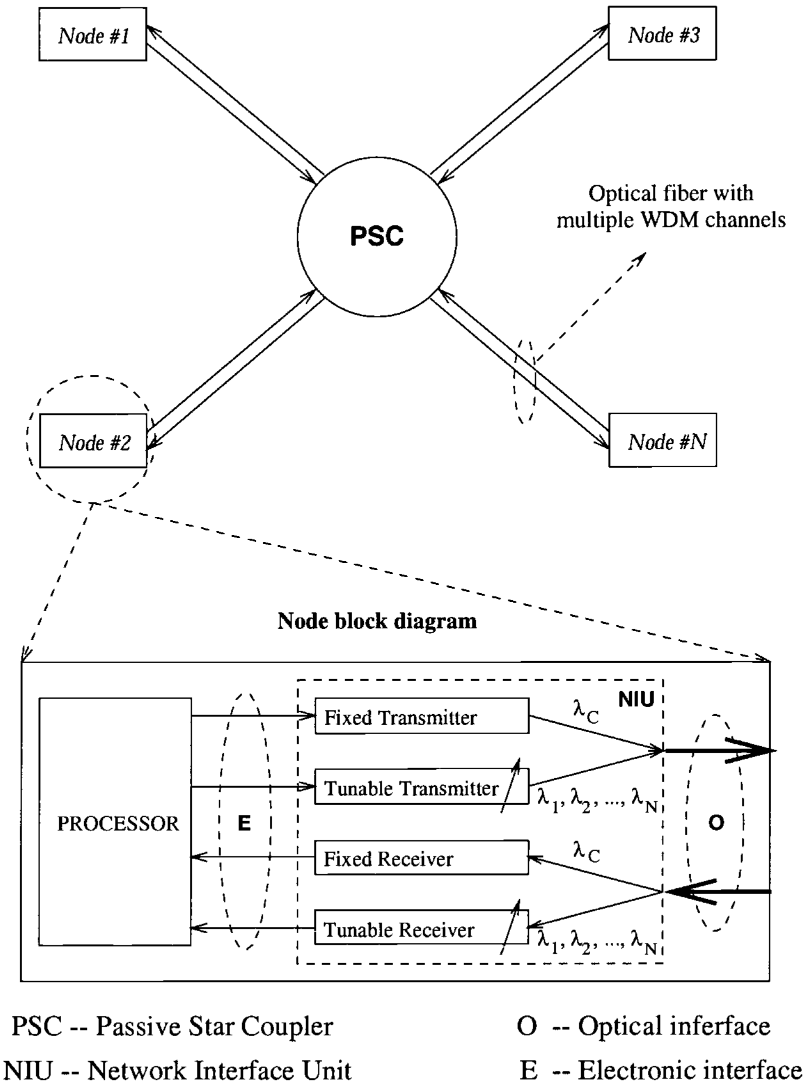


Figure 2.36 Broadcast-and-select WDM local optical network with a passive-star coupler network medium.

fixed tuned to certain wavelength channels, and there is more than one channel, then a static multihop topology must be established over the passive-star coupler. The topology is created by establishing connections between pairs of nodes on given wavelengths. An overview of multihop protocols and topologies is provided in [Mukh92b].

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 [Mukh92a]. 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 (pretransmission coordination allows a node to preannounce its transmission so that the receiving node may get ready for reception, e.g., by appropriately tuning 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.

Chronologically, the multihop concept came first precisely because fast device tuning was difficult to achieve in the early stages of developments in this field.

Node Separation in Passive-Star Coupler WDM LAN

Given the output power of the transmitters and the receiver sensitivity, we can compute the maximum distance allowable between network nodes. Assume that all nodes are D meters from the passive-star coupler (PSC), and that the input-to-output power splitting ratio of the PSC is given by Eqn. (2.9). Then, the maximum value of D such that the optical signals reaching each receiver are strong enough to be received (D_{max}) can be computed by combining Eqns. (2.9) and (2.7), so that

$$D_{max} = \frac{10}{A} \log_{10} \frac{P_t}{NP_r}$$

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.

2.8.6 WDM Wide-Area Network Design Issues

Due to the limitations of optical hardware development currently, today's optical networks may not be able to take full advantage of the bandwidth provided by optical fiber. It is anticipated that the next generation of optical networks will make use of optical 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 which 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 which do not share any links may use the same wavelength. For example, in Fig. 2.37, wavelength λ_1 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 crossconnects used in the access nodes or switching nodes. If wavelength-insensitive crossconnect 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

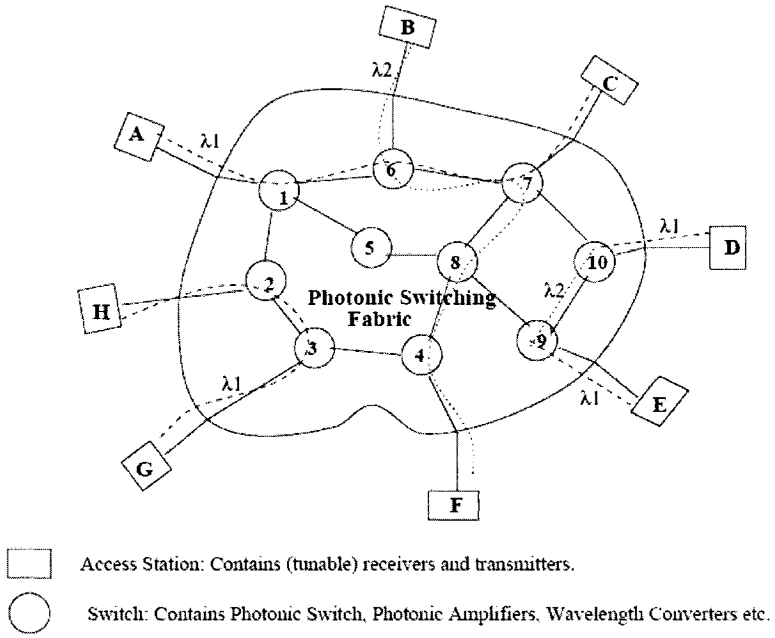


Figure 2.37 Lightpath routing in a WDM WAN.

on the same wavelength arriving on different input ports need to be routed to the same output port. The conflict can be resolved by incorporating wavelength converters at each node, and converting one of the incoming signals to a different wavelength (see Section 2.7). 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. 2.28). 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.

2.8.7 WDM Metro Network Design Issues

Today's metropolitan area networks (MANs) are mostly synchronous optical network (SONET)/synchronous digital hierarchy (SDH) ring networks which suffer from a number of drawbacks. Due to their voice-centric TDM operation and symmetric circuit provisioning, bursty asymmetric data traffic is supported only very inefficiently. Furthermore, SONET/SDH equipment is quite expensive and significantly decreases the margins in the cost-sensitive metro market. With the quickly increasing speeds in the local access networks (due to Gigabit Ethernet and similar emerging technologies) and the provisioning of very-high-capacity backbone WDM networks, the inefficiencies of SONET/SDH ring networks create a severe bandwidth bottleneck at the metro level. The resultant so-called *metro gap* prevents high-speed clients, e.g., Gigabit Ethernet, from tapping into the vast amounts of bandwidth available in the backbone. In order to: 1) bridge this bandwidth abyss between high-speed clients and backbone, 2) enable new applications benefiting from the huge amounts of bandwidth available in the backbone, and 3) stimulate revenue growth, more efficient and cost-effective metro architectures and protocols are needed. *Wavelength-division-multiplexing (WDM)* networks have been extensively investigated as solutions to the *metro gap* in [YHMR04].

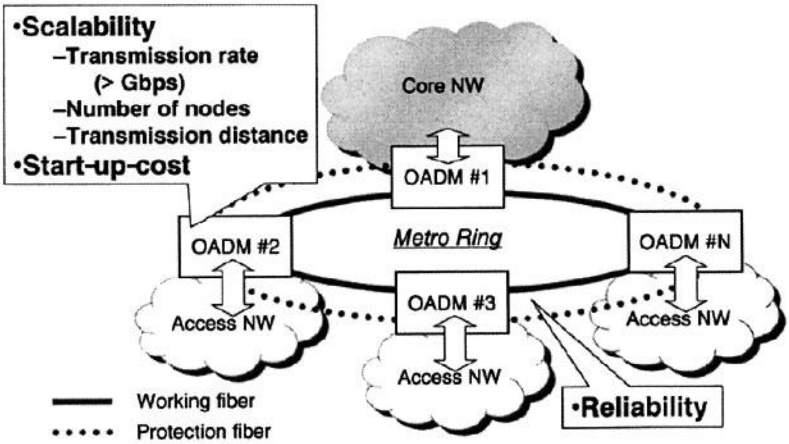


Figure 2.38 A typical WDM metro network.

As for the metro network, there are some technical issues. One is how to make best use of the scalability offered by WDM. We should design networks scalable in terms of transmission rate, number of nodes accommodated, and transmission distance (i.e., span length and total ring length). Another issue is how to reduce the start-up-cost. Some service providers are already starting to deploy scalable, cost-effective value-added metro services. But it is difficult to accurately estimate the future traffic, especially in metro networks. Therefore, an optical add/drop multiplexer (OADM) architecture with both superior scalability and low start-up-cost is necessary to install ring networks in metro areas. A typical configuration of the WDM metro network with OADMs is shown in Fig. 2.38 in [IKSF04]. The other major challenge in metro networks is to guarantee the quality of service (QoS) using intelligent traffic engineering (TE) and management schemes. By using working and protection fibers, automatic protection and supervision functions based on simple OADM are also needed to ensure service reliability. Other challenges include *operations*, *administration*, *maintenance (OAM)*, and so on.

2.8.8 Optical Access Network Design Issues

The recent explosive growth of the Internet has triggered the introduction of a broadband access network based on fiber-to-the-office (FTTO) and fiber-to-the-home (FTTH). This trend will dramatically accelerate from now on due to further progress in e-businesses such as contents delivery and TV services. To deal with the various demands, access networks requires scalability in terms of capacity and accommodation, and the flexibility with regard to physical topology. The introduction of technically mature WDM technologies that are also cost-effective in core networks to support access networks is expected to yield scalable and flexible networks.

Figure 2.39 shows a typical WDM access network in [IKSF04]. A wavelength multi/demultiplexer MUX is used to connect plural *optical network units (ONUs)* placed in homes/buildings (end users) to an *optical line terminal (OLT)*; splitting loss of the MUX can be much lower (i.e., 5 dB for 16 splits) than that of the power splitter used in TDM access (i.e., 10 dB for 8 splits). The main technical issue is to realize colorless ONUs; the ONUs should be colorless (in other words, no ONU is wavelength specific) to decrease the costs of operation, administration, and maintenance (OADM) functions, as well as the production cost since mass production becomes possible with just one specification.

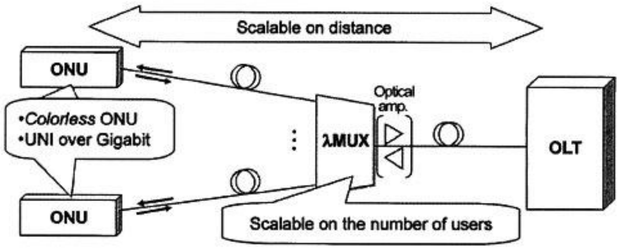


Figure 2.39 A typical WDM-access network.

Enabling technologies include: 1) employing a modulator only in the ONU; 2) employing a light source with broadband optical spectrum at each ONU; and 3) employing a tunable light source at each ONU. Since approach 3) requires wavelength setting and control in each ONU and increases the burden of OADM, approaches 1) and 2) are more attractive. Note that approach 1) has several variants; i.e., the supplied optical carriers can be generated by spectrum-slicing broadband ASE light sources, and/or a *Fabry-Perot* laser can be employed instead of the modulator in the ONU so that its wavelength is locked into that of the supplied optical carrier.

Another issue is how to make the best use of the large scalability offered by WDM. We should design scalable networks in terms of transmission distance, the number of users accommodated, and bandwidth for each user.

The another major challenge is load balancing which is highly desirable in WDM access networks. Balanced traffic distribution can be obtained using an efficient and fast access provisioning tool.

2.9 Summary

Recent advances in the field of optics have opened the way for the practical implementation of WDM networks. In this chapter, 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.

Exercises

2.1. An IP-based network application will be built on top of a fiber-optic communication network. The application programmer considers two options for error correction. One option is to simply use the TCP/IP protocol. The other option is to use the UDP/IP protocol, which the programmer calculates will have a 10% higher bandwidth when transmissions are error free. The programmer has included a cyclic redundancy check in the UDP packets, but does not have a good scheme for retransmitting individual packets. Thus, when an error is detected, an average of 125 megabytes of data will have to be retransmitted. Which scheme will yield the higher average bandwidth if the fiber bit error rates is: (a) 10^{-9}? (b) 10^{-15}?

2.2. Consider a step-index fiber which has a core refractive index of 1.495. What is the maximum refractive index of the cladding in order for light entering the fiber at an angle of 60 degrees to propagate through the fiber? Air has refractive index of 1.0.

2.3. Find the numerical aperture in a graded-index fiber with two layers shown in Fig. 2.41. Compare the answer with the numerical aperture of the step-index fiber shown in Fig. 2.40. Can we use geometric optics to deal with situations where the wavelength and core diameter are of the same order of magnitude (e.g., single-mode fiber)?

2.4. Consider a step-index multimode fiber in which the refractive indices of the cladding and core are 1.35 and 1.4, respectively. The diameter of the core is 50 μm. Approximately how many modes are supported by the fiber for a signal at a wavelength of 1550 nm?

2.5. Find the approximate number of modes in a 100 μm core step-index multimode fiber at 850 nm. Assume the refractive index of glass to be 1.5 and that of the cladding to be 1.47.

2.6. Consider an optical link in which power at the transmitter is 0.1 mW and the minimum power required at the receiver is 0.08 mW. The attenuation constant for the fiber material is 0.033 dB/km. What is the maximum length of the optical link, assuming that there are no amplifiers?

2.7. Describe the various types of dispersion and explain how the effects of each type of dispersion can be reduced.

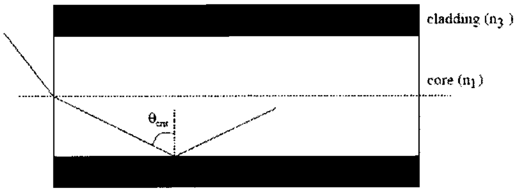


Figure 2.40 Critical angle in a step index fiber.

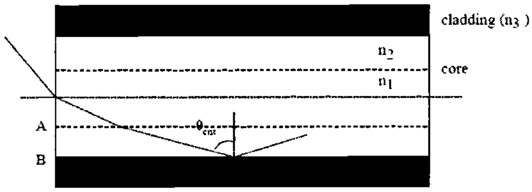


Figure 2.41 Critical angle in a graded index fiber.

2.8. Consider a single mode optical fiber with an attenuation of 0.2 dB/km and a dispersion limit of 200 Gbps·km. The transmitter power is 1 mW and the receiver sensitivity is 10^{-5} mW. The link operates at a rate of 2.4 Gbps. Assume a 10 dB power margin for losses in connectors. Calculate the maximum length of the optical link.

2.9. Suppose we have a system which has 3 channels operating at 1549.32 nm, 1554.13 nm, and 1558.98 nm. At which frequencies will we have sidebands as a result of four-wave mixing? (Use $c = 2.998 \times 10^8$.)

2.10. Consider a broadcast star network with $N = 2^k$ nodes where k is an integer. The network is built out of 2×2 couplers with excess loss β and coupling coefficient α. Each transmitter has a laser with power P_t.

(a) Find the power levels received by the receivers when a single transmitter (say transmitter T_1) transmits. That is, determine how many different power levels are received by the N receivers, and how many receivers receive each of these levels.

Hint: Construct a tree with the transmitter at the root.

(b) Suppose your goal as the network designer is to maximize the minimum power received by any of the receivers from a transmitter. Find the optimal value of α for this design criterion. Explain your answer.

(c) Now suppose you have a different design criterion. Your new goal is to maximize the expected value of the power between a random transmitter/receiver pair. Assume that each such pair is equally likely. How would you select your couplers? Explain your answer. *Hint:* You can exploit the symmetry to fix the transmitter.

2.11. Suppose a 1 mW, 1550 nm signal is transmitted across a 5 km fiber, through an 8×8 passive star coupler, and through another 15 km of fiber before reaching its destination. No amplifiers are used. What is the power of the signal at the destination?

2.12. A 16×16 passive-star hub has been constructed from combiners, couplers, and splitters as in Fig. 2.7. Each combiner, coupler, and splitter results in a 3 dB power loss. Each host is up to 10 km away from the star, with a signal attenuation of 0.2 dB/km. If each host must receive signal power of at least 0.01 mW to clearly recognize signals, how strong must each host's transmission signal be?

2.13. Consider a unidirectional fiber bus with N nodes. (Assume that N is even.) All the couplers have an excess loss β ($\ll 1$). (Let $\gamma = 1 - \beta$ and use γ in your solution instead of β.) The i^{th} coupler has a coupling coefficient α_i for $1 \leq i \leq N$. The coupling coefficients can be independently selected. Optimize the coupling coefficients to maximize the worst-case power transfer between a transmitter and a receiver. Compare the resulting worst-case power with the case where all couplers are identical and optimized. Hint: First assume some transmitter k is the worst-off. Argue why, in the optimal solution, all the receivers to the right of k will see the same attenuation from transmitter k. Similarly all the transmitters to the left of k should be equally badly off. Obtain a recursion for α_i assuming k is known and find the value of k.

2.14. Consider the following simplified model of a direct detection binary FSK system. By using a pair of optical filters and a pair of photodetectors in the receiver, we observe two Poisson distributed photon counts: X_0 and X_1. When the data bit is 0, the parameter of X_0 is $\lambda + \lambda_d$ while that of X_1 is λ_d. (Here λ_d models the dark current.) Conversely, when the data bit is 1, X_0 has parameter λ_d and X_1 has $\lambda + \lambda_d$. X_0 and X_1 are statistically independent when conditioned on the value of the data bit.

(a) Obtain the Maximum Likelihood processing of X_0 and X_1 explicitly.

(b) Find the probability of a bit error as a function of λ and λ_d. You may leave your answer as a double series.

(c) Repeat part (b) when there is no dark current ($\lambda_d = 0$). Now your answer must have a simple form.

2.15. In this problem, you will investigate the relationship between the finesse F and the reflection coefficient R of a Fabry-Perot filter.

(a) Show that for $R \simeq 1$, the finesse can be approximated as

$$F \simeq -\frac{\pi}{\ln R}$$

(b) Find the exact and the approximate expression for R in terms of F. Evaluate the accuracy of the approximation for $F = 10$ and $F = 100$.

2.16. In this problem, you will consider the worst-case crosstalk in a WDM environment with Fabry-Perot (FP) filters. Assume that the filter is lossless ($A = 0$).

(a) Show that the power transfer function $T(f)$ of the FP can be written as

$$T(f) = \frac{1-R}{1+R} \sum_{m=-\infty}^{\infty} R^{|m|} e^{j2\pi m f/P}$$

where P is the free spectral range.

(b) Using the result in (a), show that the worst-case interference

$$C_{\max} = \sum_{i=1}^{M-1} T\left(\frac{iP}{M}\right)$$

is given by

$$C_{\max} = M \frac{1-R}{1+R} \frac{1+R^M}{1-R^M} - 1$$

(c) Using the approximation in Problem 2.15, show that when $F \gg 1$

$$C_{\max} \simeq \frac{\pi M}{2F} \coth\left(\frac{\pi M}{2F}\right) - 1$$

where $\coth(x) = (e^x + e^{-x})/(e^x - e^{-x})$. Find, numerically, the maximum value of M/F such that $C_{\max} \leq 1$. Comment on your result.

2.17. Consider a Mach-Zehnder filter chain of K cascaded filters with

$$\Delta L_i = 2^{i-1}\Delta L \quad i = 1, 2, \ldots, K$$

Show that the power transfer function of this chain is given by

$$T(f) = \frac{\sin^2(\pi M f/P)}{M^2 \sin^2(\pi f/P)}$$

where $M = 2^K$ and $P = c/\Delta L$.

2.18. Optical amplifiers saturate at high levels of output power. Suppose the saturation power of an erbium-doped fiber amplifier is 20 mW, and the amplifier gain is 5 dB/mW of pump power. The pump power is set to 5 mW. What is the largest amount of input power that can be amplified without driving the amplifier into saturation.

2.19. An optical amplifier delivers an output power P_{out} in response to an input power P_{in} as described by the following equation

$$P_{out} = a\,(1 - \exp(-bP_{in}))$$

where a and b are constants.

(a) What is the saturation power P_{sat} of this amplifier?

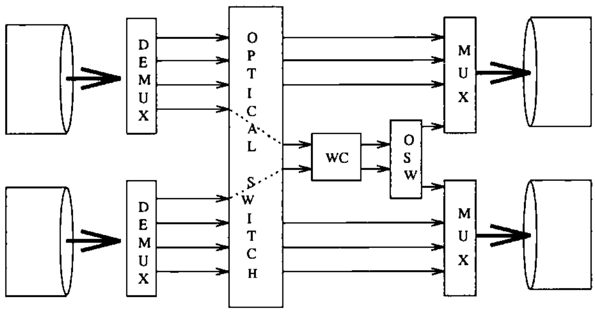
(b) Find the power gain of the amplifier in the linear operating region (i.e., small input power).

(c) Suppose this amplifier is to be placed in a transmission link of length L. The fiber has an attenuation factor of α per unit length, i.e., after a distance l, a factor $e^{-\alpha l}$ of the original power remains. The transmitter has a laser with power P_t. The goal is to maximize the received power P_r. Find the optimal position x (measured from the transmitter) of the amplifier. Comment on your result.
Hint: The only root of the equation $u = e^u - 1$ is at $u = 0$.

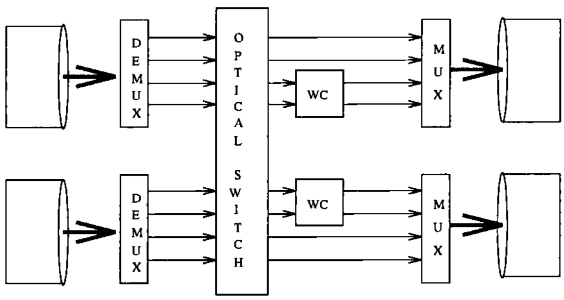
2.20. (a) Use four 2×2 optical crosspoint elements to construct a 4×4 Banyan interconnect.

(b) How many rows and how many columns of 2×2 crossbars would be required for an $N \times N$ Banyan interconnect.

(c) Suppose we have an 8×8 space-division Banyan optical routing switch. Label the inputs from 0 to 7 and label the outputs from 0 to 7. Suppose we need to simultaneously route Input 5 to Output 2 and Input 7 to Output 0. Can this routing be accomplished? If yes, give the routes through the switch, otherwise explain why the routing isn't possible.



(a) Share-per-node wavelength-convertible switch architecture



(b) Share-per-link wavelength-convertible switch architecture

Figure 2.42 Two architectures for wavelength convertible routers: (a) share-per-node, (b) share-per-link.

2.21. What are the uses of wavelength conversion in a WDM network?

Consider the two wavelength-convertible switch (WCS) architectures shown in Fig. 2.42. Construct a set of connection requests that can be routed by the share-per-node WCS and cannot be routed by the share-per-link WCS, and vice-versa.

2.22. Suppose we want to design a system with 16 channels, each channel with a rate of 1 Gbps. How much bandwidth is required for the system?

2.23. Suppose we have a fiber medium with a bandwidth of about 20 nm. The center wavelength is 0.82 μm. How many 10 GHz channels can be accommodated by the fiber? Calculate the maximum number of channels for a center wavelength of 1.5 μm.

2.24. Consider an optical communication system in which the transmitter tuning range is from 1450 nm to 1600 nm, and the receiver tuning range is from 1500 nm to 1650 nm. How many 1 Gbps channels can be supported in the system?

2.25. Consider a WDM passive-star-based network for N nodes. Let the tuning range of the transmitters be 1550 nm to 1560 nm, and let the tuning range of the receivers be 1555 nm to 1570 nm. Assume that the desired bit rate per channel is 1 Gbps. Also assume that a channel spacing of at least 10 times the channel bit rate is needed to minimize crosstalk on a WDM system. Find the maximum number of resolvable channels for this system.

2.26. In which type of network, single-hop or multihop, is a smaller tuning latency more critical? Why?

2.27. In a WDM network node, if two signals on the same wavelength arriving from different input ports need to go to the same output port, then a conflict may occur. Describe two or more methods for resolving this conflict. Discuss the advantages and disadvantages of each solution.

2.28. Figure 2.43 shows a WDM WAN constructed using AWGs at each node. Assume that there are sufficient number of fibers between the node pairs (not shown in Fig. 2.43). Show how the following connection requests can be satisfied (you may have to write a program which tries out various possibilities).

- *node* 3 $\rightarrow$ *node* 1

- *node* 3 → *node* 1
- *node* 1 → *node* 2
- *node* 1 → *node* 2
- *node* 2 → *node* 3
- *node* 2 → *node* 3

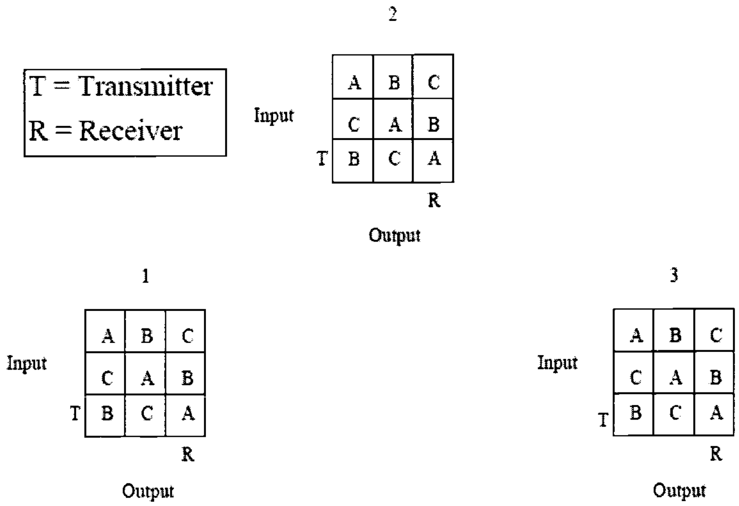


Figure 2.43 A WDM WAN constructed using AWGs at each node. (All connections begin at transmitters and end at receivers.)

2.29. Suppose we are given the network in Fig. 2.37 and have two wavelengths available. We wish to set up the following connections:

i. H-2-3-4-8-9-E
ii. C-7-8-4-F
iii. B-6-7-8-9-E
iv. D-10-7-C

At which nodes are wavelength converters required, and how many conversions are required at these nodes? Explain.