DWDM Network Designs and Engineering Solutions

ByAshwin Gumaste,Tony Antony

Publisher: Cisco Press
Pub Date: December 13, 2002
ISBN: 1-58705-074-9
Pages: 368

A comprehensive book on DWDM network design and implementation solutions.

- Study various optical communication principles as well as communication methodologies in an optical fiber
- Design and evaluate optical components in a DWDM network
- Learn about the effects of noise in signal propagation, especially from OSNR and BER perspectives
- Design optical amplifier-based links
- Learn how to design optical links based on power budget
- Design optical links based on OSNR
- Design a real DWDM network with impairment due to OSNR, dispersion, and gain tilt
- Classify and design DWDM networks based on size and performance
- Understand and design nodal architectures for different classification of DWDM networks
- Comprehend different protocols for transport of data over the DWDM layer
- Learn how to test and measure different parameters in DWDM networks and optical systems

The demand for Internet bandwidth grows as new applications, new technologies, and increased reliance on the Internet continue to rise. Dense wavelength division multiplexing (DWDM) is one technology that allows networks to gain significant amounts of bandwidth to handle this growing need. DWDM Network Designs and Engineering Solutions shows you how to take advantage of the new technology to satisfy your network's bandwidth needs. It begins by providing an understanding of DWDM technology and then goes on to teach the design, implementation, and maintenance of DWDM in a network. You will gain an understanding of how to analyze designs prior to installation to measure the impact that the technology will have on your bandwidth and network efficiency. This book bridges the gap between physical layer and network layer technologies and helps create solutions that build higher capacity and more resilient networks.

DWDM Network Designs and Engineering Solutions

ByAshwin Gumaste,Tony Antony

Publisher: Cisco Press

Pub Date: December 13, 2002

ISBN: 1-58705-074-9

Pages: 368

Index

Copyright

Published by:
Cisco Press
201 West 103rd Street
Indianapolis, IN 46290 USA

Printed in the United States of America 1 2 3 4 5 6 7 8 9 0

First Printing December 2002

Library of Congress Cataloging-in-Publication Number: 619472050743

Warning and Disclaimer

This book is designed to provide information about DWDM networks. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an "as is" basis. The authors, Cisco Press, and Cisco Systems, Inc. shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the authors and are not necessarily those of Cisco Systems, Inc.

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc. cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members of the professional technical community.

Reader feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through e-mail at feedback@ciscopress.com. Please be sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Publisher	John Wait
Editor-In-Chief	John Kane
Cisco Representative	Anthony Wolfenden
Cisco Press Program Manager	Sonia Torres Chavez
Cisco Marketing Communications Manager	Tom Geitner
Cisco Marketing Program Manager	Edie Quiroz
Acquisitions Editor	Amy Moss
Production Manager	Patrick Kanouse
Development Editor	Howard Jones
Project Editor	San Dee Phillips
Copy Editor	Karen A. Gill
Technical Editors	Tim Benner, Wayne Hickey, Arthur Lowery, Steve Wisniewski
Team Coordinator	Tammi Ross
Book Designer	Gina Rexrode
Cover Designer	Louisa Adair
Composition	ContentWorks
Indexer	Tim Wright

CISCO SYSTEMS

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems Europe
11 Rue Camille Desmoulins
92782 Issy-les-Moulineaux
Cedex 9
France

http://www-europe.cisco.com
Tel: 33 1 58 04 60 00
Fax: 33 1 58 04 61 00

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems Australia,
Pty., Ltd
Level 17, 99 Walker Street
North Sydney
NSW 2059 Australia
http://www.cisco.com
Tel: +61 2 8448 7100
Fax: +61 2 9957 4350

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China • Colombia • Costa Rica • Croatia • Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong • Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden • Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Dedications

Ashwin Gumaste: I dedicate this book to my late father, Dr. Anil H. Gumaste. I also dedicate this book to my two beloved dogs, Johnny and Zanjeer. My entire research career would not have been possible without the help of Siddhivinayak Temple in Bombay, India.

Tony Antony: I dedicate this book to my parents, C.P. Antony and Ritha Antony; my wife, Sheela; and my daughters, Chelsey and Melanie.

About the Authors

Ashwin Gumaste received a master's degree in telecommunications and is currently with the Center for Advanced Telecommunications Systems and Services (CATSS) at the University of Texas at Dallas where he is pursuing a Ph.D. in electrical engineering. He is also part of Photonics Networking Laboratory (PNL) at Fujitsu in Richardson, Texas where his research includes network development and design. He has previously worked with Cisco Systems in the Optical Networking Group. He has numerous papers and pending U.S. patents. He was awarded the National Talent Search Scholarship in India in 1991. His research interests include optical and wireless networking and Self-Similar phenomenon in social and networking environments. He has also proposed the first architecture to implement optical burst switching and for multicasting lightpaths called Light-trails.

Tony Antony has more than 11 years of telecommunications/data-networking experience and is currently working at Cisco Systems as a Technical Marketing Engineer in the Optical Networking Group. He received a master's degree in telecommunications from SMU, Dallas and also holds a CCNP and CCIP (Optical) along with multiple other certifications. Tony's previous experience includes engineering positions at Texas Instruments and KPMG. He has authored numerous technical papers at international conferences in the networking area. His research interests include optical Internet and network simulations. Tony can be reached at tantony@cisco.com.

About the Technical Reviewers

Tim Benner has been in the networking telecommunications industry and electronics field for more than 13 years. His first training came while he was in service as an electronics technician (communications and radar) in the United States Navy. He has attained every major telecommunications certification. Tim has worked for a Cisco Silver, Gold, and several Cisco Professional Services Partners as a senior principal network consultant. He has designed, implemented, and had exposure to many different technologies, including VoX technologies, LAN/WAN, and Internet technologies for service provider and enterprise networks. He spends much of his free time reading industry books and doing research in his home lab.

Wayne Hickey has more than 20 years of telecommunications, computer, and data experience. He has product expertise in SONET, SDH, DWDM, IP, ATM, Frame, HFC, Voice, Video, and Wireless. Wayne has held various positions with Cisco Systems as a CSE, SEM, SSEM, and now a product manager for the Optical Technical Business Unit. Previously, he spent 19 years working for Aliant Telecom (NBTel), the third largest telecommunications provider in Canada, where he was focused on transmission network design and the evaluation of emerging access and transmission technologies. Wayne has co-authored and authored several papers on PMD, long-haul transmission systems, and he has applied for several patents on primary and secondary protect for HFC.

Dr. Arthur Lowery earned a first-class honour's degree in applied physics from Durham University, England, in 1983, and a Ph.D. from Nottingham University, where he worked as a lecturer in electrical engineering. In 1990, Arthur joined the Photonics Research Laboratory at the University of Melbourne, Australia as a senior lecturer and later a reader, where he led projects developing CAD tools for photonic devices and systems, optical amplifier applications, and mode-locked lasers. In 1996, he cofounded Virtual Photonics Pty Ltd. to commercialize photonic design automation tools, which merged with BneD (Germany) in 1998 to become VPIsystems Inc. Since then, Arthur has been group technology officer of Melbourne's Optical Design Group, which develops design tools for components, links, and systems. He has written more than 160 papers on photonics and simulation.

Steve Wisniewski, CCNP, has an M.S. degree in telecom engineering from Stevens Institute of Technology. He has more than 10 years of networking experience and is employed as a senior engineer for Greenwich Technology Partners. He has authored two books on networking for Prentice Hall and is presently co-authoring a book for Cisco Press. He lives in East Brunswick, N.J. with his wife, Ellen, and their 16 dogs.

Acknowledgments

Ashwin Gumaste: A very special thanks to my Ph.D. advisor and the distinguished chair in telecommunications, Dr. Imrich Chlamtac, at the University of Texas at Dallas, for the encouragement he bestowed upon me, as well as for being a strong source of motivation in the optical networking field. I would also like to thank Dr. Hideo Kuwahara, senior VP of Photonics Networking Laboratory at Fujitsu, for his constant support during the writing of this manuscript. I would like to thank Dr. George Thomas at the University of Louisiana, Lafayette for showing me the intricacies in DWDM networking. A special acknowledgement to the efforts of Amy Moss and Howard Jones from Cisco Press, as well as the technical reviewers Wayne, Arthur, Steve, and Tim. I would like to note the initiatives of the interfacing members from VPI (Arthur, Elizabeth, and Petter) and Agilent (Sudhir and Randy).

Tony Antony: I would like to thank my parents and family for supporting me during the entire process of this book. A special thanks to my team members and managers, especially Faraz Aladin, Dave Merrell, and Russ Tarpey at Cisco Systems. I would also like to thank Artie Thomas and all my other colleagues for the continuous encouragement they offered me. I extend my thanks to Amy Moss for coordinating the efforts, Howard Jones for developing the book, and Wayne Hickey, Arthur Lowry, Tim Benner, and Steve Wisniewski for their excellent comments. Special thanks to Didi Foster for coordinating Agilent test and measurement information, as well as Sudhir Peddireddy and Randy Foster. Finally, I would like to thank Elizabeth Morgan, Arthur Lowry, and Petter Willdhagan from VPI for compilation of Chapter 10, "Simulations of WDM Systems."

Introduction

The massive growth in Internet traffic has created a surge of bandwidth requirement in today's networks. Wavelength division multiplexing (WDM) is one enabling technology that alleviates the bandwidth issue. This book is an attempt to discuss this nascent technology from an implementation perspective. It is designed for levels from the novice to the expert, keeping in mind the practical relevance of deployment. The objective is to acquaint the reader with the enabling technologies that affect DWDM networking and to enable the reader to design (synthesize) and analyze DWDM networks, from a systems perspective. The book attempts to bridge the gap between physical layer and network layer technologies and creates a solution that absorbs the best of both worlds to demonstrate higher capacity and more resilient networks for the future.

Goals and Methods

This book discusses the technological implementations of DWDM networking from a systems-level approach. DWDM as a technology attempts to coalesce the optical and network layer strategies in creating huge surges in bandwidth that are available to the end user. This book deals with different facets of DWDM networking and creates a resource pool of network-systems knowledge. It can be regarded as a premier source of information to cover a wide spectrum of issues from components to system design. The objective is to make the reader able to design and understand DWDM networks as well as the underlying optical technology.

Some of the most important details covered include the following:

- Study various optical communication principles as well as communication methodologies in an optical fiber. Be able to choose the right fiber for the given application.
- Design and evaluate optical components in a DWDM network.
- Learn the effects of noise in signal propagation in WDM networks, especially from OSNR and BER perspectives.
- Design optical amplifiers.
- Design optical links based on power budget.
- Design optical links based on OSNR.
- Design a real DWDM network with impaired from OSNR, dispersion, and gain tilt.
- Classify and design DWDM networks based on size and performance.
- Understand and design nodal architectures for different classifications of DWDM networks.
- Understand the effects of routing and wavelength assignment in DWDM networks.
- Comprehend the different protocols for transport of data over the WDM layer.
- Learn future trends in WDM networking.
- Test and measure different parameters in WDM networks and optical systems.
- Evaluate performance of WDM networks using simulations.

Who Should Read This Book?

This book is meant to cover the entire spectrum of technologies from physical layer to network layer and is designed to benefit a wide range of people in the networking community from novice to researcher. The objective is not just to explain the underlying technologies but also to make the reader well versed regarding design aspects. This book can be considered a standalone tutorial for DWDM networking.

In general terms, network engineers, network architects, network design engineers, customer support engineers, salespeople, system engineers, and consultants who design, deploy, operate, and troubleshoot WDM networks and who want to provide new-world enabling services on their networks should read this book.

Academicians who need to understand this technology model and theoretical framework as part of their research will also find this book useful.

How This Book Is Organized

The book is partitioned into 10 chapters, and each chapter deals with an important facet of DWDM networking. It is structured in a way that the reader is initially acquainted to premier optical technologies.

DWDM networking is built on the strong fundamentals of enabling optical technologies, and Chapters 1 through 3 discuss these technologies. Following an engineering-level understanding on optical networking, Chapters 4 and 5 detail DWDM network design from optical as well as topological aspects. Those chapters look at different classifications of DWDM networks and how networks are designed and implemented. A system design perspective creates a resource pool of knowledge for industry designers.

Chapter 6 amalgamates network layer issues with DWDM networking, discussing routing and wavelength assignment issues in a DWDM network. From the higher layers of the OSI model, a wide range of protocols are feasible alternatives for transport of data over WDM networks. Chapter 7 throws light on some of these protocols. As DWDM technology matures, newer technologies appear that enable better, more resilient, as well as higher-capacity DWDM networks. These trends of future DWDM networking are discussed in Chapter 8.

As industry begins to deploy DWDM networks in a substantial way, these networks and the associated network gear need to be tested and the parameters measured. Chapter 9 explains some generic methods for measurement of optical parameters, as well as validating techniques to prove the effectiveness of WDM networks. To correctly channel investments in WDM networks, the design needs to be validated. Simulation exercises are an excellent method for such validation.Chapter 10 discusses simulation studies of WDM networks.

The 10 chapters cover the following topics:

- Chapter 1, "Introduction to Optical Networking"— This chapter introduces basic principles that govern optical communication such as reflection, refraction, dispersion, and polarization, as well as the properties of optical fiber such as birefringence and nonlinearity. This chapter further details the methods of communication in a fiber as well as the various impairments to signal flow in a fiber. It also highlights properties of DWDM networking from the optical communication point of view.

- Chapter 2, "Networking with DWDM -1"— This chapter showcases some of the DWDM components and technologies. It introduces optical transmitters and various forms of lasers that are used in DWDM transmission, as well as their characteristics. It also discusses optical receiver design and its importance to system design based on noise and bit error rate (BER). In addition, Chapter 2 does some elementary mathematical analysis of BER and discusses the Q-factor and signal-to-noise ratio (SNR). The chapter also studies components such as couplers, circulators, various forms of filters, and optical switches. Finally,Chapter 2 helps readers map the components and their technology into a fully progressed WDM network.

- Chapter 3, "Networking with DWDM -2"— This chapter discusses an important innovation—optical amplifer—that affects DWDM network design. Also examined are doped fiber, Raman, and semiconductor optical amplifers. Finally, this chapter explains dispersion-compensation techniques for chromatic as well as polarization mode dispersion from a systems level perspective.

- Chapter 4, "WDM Network Design -1"— This chapter discusses optical network system-level design. It initially considers power budget-based design and moves on to more complex optical signal-to-noise ratio (OSNR)-based designs. Chapter 4 also shows the importance of OSNR in estimating BER and the need for evaluation of the Q-factor as an intermediate stage in BER calculation. In addition, it discusses dispersion-based systems and penalties that are associated with dispersion-limited systems. This chapter looks at how dispersion-limited systems can be compensated by using generic schemes and shows the methods of pre- and postcompensation as well as placement of such compensating units. Finally, Chapter 4 studies nonlinear effects from a design point of view and various penalties and their cures from a pure system design perspective.

- Chapter 5, "WDM Network Design -2"— This chapter discusses philosophies of optical network design from a topological point of view. It classifies optical networks into three main areas: access, metro, and long haul. Each network type has different components and design issues associated with it. Access networks have relatively less optical impairments and can be implemented using low-cost technologies. Access networks are more flexible and provide a direct point of attachment to end users. Metro networks require comparatively more stringent optical requirements than access networks. Metro networks are physically larger than access networks and need more specific technologies for implementation. Long-haul networks have the maximum number of optical impairments associated. Long-haul network design involves many different philosophies and components. This chapter explores how to design optical WDM networks and choose the right equipment for the right application.

- Chapter 6, "Network Level Strategies in WDM Network Design: Routing and Wavelength Assignment"— This chapter focuses on routing and wavelength assignment (RWA) strategies in optical networks. RWA analysis is important from a network planning and capacity planning objective. This chapter discusses some algorithms proposed by leading researchers that consider various facets of the RWA problem.

- Chapter 7, "X over DWDM"— This chapter considers various protocols over the DWDM layer for possible implementations. It looks at SONET/SDH, Ethernet, IP, and RPR as some of the protocols that can transfer data over the WDM layer. SONET/SDH seems to be the most common way of sending data, but it comes at a high price. In contrast, Ethernet has a simple implementation and is rapidly growing in popularity. Ethernet is widely considered to be the future replacement of SONET/SDH. RPR is an efficient protocol for ring networks. IP directly over WDM seems to be the most promising protocol, but it is quite difficult to widely implement at this time.

- Chapter 8, "Future WDM Networks and Technologies"— This chapter looks at some of the future trends that will affect DWDM networking. Burst switching, self-similarity in network traffic, 40 G communication, and IP over DWDM implementations are discussed.

- Chapter 9, "Tests and Measurements"— This chapter discusses measurement of parameters that affect the DWDM network. It also highlights the technolgies used in the measurement and testing of these WDM parameters. In addition, this chapter explains principles of operation of test equipment such as optical spectrum analyzer and optical power meter.

- Chapter 10, "Simulations of WDM Systems"— This chapter studies DWDM network simulations and methods. Using the software (VPItransmissionMaker) in the enclosed CD, you can perform network simulations. The CD contains demos and case studies for WDM networking. This chapter details methods and requirements for conducting good simulation study.

Chapter 1. Introduction to Optical Networking

Communication and communication systems have evolved tremendously over the past century. The tremendous growth in demand for bandwidth has led to various premier technologies that facilitate more communication in today's networks.

The need to send more information (data) through a communication medium (channel) is one of the motivating factors for continuous research to invent more efficient communication systems. In the past, both conventional copper and wireless methods were good means of transporting data. They had the limitations of a finite bandwidth and high losses that were proportional with transmitting length. Glass as a possible means of communication was studied and experimentally deployed as early as the 1960s.

It was not until the mid-1980s that commercial deployment of fiber actually occurred. A paradigm shift in fiber manufacturing technologies and new developments in semiconductor lasers and detectors enhanced the speed of this deployment. A gradual shift from the conventional multimode to the more exotic single mode fiber caused a steep increase in communication data-rates. The heavy dependence of Synchronous Optical Network (SONET)/Synchronous Digital Hierarchy (SDH) as a standard transport gear for data and voice coalesced with the demand for high bit rates, ensuring massive deployment of fiber rings.

Three more technologies that revolutionized the optical networking segment included tunable lasers, arrayed waveguides, and optical amplifiers. Tunable lasers could emit light at different wavelengths by using some physical characteristic of the lasing media that facilitated the change in emitted wavelengths. Arrayed waveguides, although introduced almost two decades after tunable lasers, were important for multiplexing and demultiplexing wavelengths of lights and forming a composite signal. Optical amplifiers were genuine optical equivalents of electronic amplifiers, with the same generic functionality.

Conventionally, light propagation is an electromagnetic phenomenon, and its propagation through a fiber can be mathematically studied with Maxwell's equations.

NOTE

Optics is a complicated field; its understanding requires solid physics concepts.

One of the key notions behind communication in a fiber is the total internal reflection (TIR) of light within a transmissive media. This concept is detailed in the following section, "WDM," and it leads to the basic theory that supports optical communication. Further, multiplexing, in time and frequency domains, is quite prominently deployed in various forms of communication systems. This concept is furthered to the field of optical communications, which gives birth to what is commonly termed wavelength division multiplexing (WDM). Later in this chapter, the section titled "Optics: An Update" discusses WDM in more detail.

An enormous bandwidth and a low error rate would seem to place the optical fiber in a category that is superior to other conventional forms of communication media. Theoretically, a single strand of fiber as thin as a single human hair could create bandwidth of 30000 Gbps (1 Gbps =

10^9bps). However, single channel communication systems today have not been able to function higher than 40 Gbps due to the severe opto-electronic mismatch. What makes optical communication so attractive is this capability to transport huge data files and connect optical circuits or *lightpaths* across large distances with minimal communication errors.

The building blocks of an optical network can be categorized into three groups:

- The communication media, which is the fiber.
- The passive and active components that interface with this media, such as lasers, detectors, amplifiers, waveguides, and so on.
- The software-based network management system and the protocols that run through the channels of communication, creating a conducive communication environment.

The fiber is a passive media that is governed by laws of optical physics. Light pulses propagating through the fiber experience optical effects such as reflection, refraction, birefringence, attenuation, dispersion, and polarization to name a few. Forthcoming sections focus on the basic concepts of each of these effects. The mathematical analysis of each effect is beyond the scope of this book, but they are described quite exhaustively in references 1, 2, and 3 at the end of this chapter.

The components that interact with the fiber and are responsible for communication are grouped primarily into passive and active devices. Passive devices—such as fiber couplers, taps, circulators, and gratings mux/demux—are key WDM components and are studied in the first half ofChapter 2, "Networking with DWDM -1." Active devices—such as lasers, photo-detectors, amplifiers, and switches—are described in Chapters 2 and 3, "Networking with DWDM -2." Chapter 4, "WDM Network Design -1," and Chapter 5, "WDM Network Design -2,"describes the actual network design issues, and Chapter 6, "Network Level Strategies in WDM Network Design: Routing and Wavelength Assignment," describes how to handle network-related problems. Chapter 7, "X over DWDM," discusses various protocols over the WDM layer, and Chapter 8, "Future WDM Networks and Technologies," describes various new technologies and focus on new areas of WDM networking with the predominant IP layer imbibed onto the optical layer. Chapter 9, "Tests and Measurements," deals with testing WDM systems as well as measuring WDM system parameters. Chapter 10, " Simulations of WDM Systems," explains simulation techniques and methodologies of WDM networks.

WDM

WDM is the abbreviation for wavelength division multiplexing—a term that has risen to prominence over the past decade. Charles Brackett's seminal paper[13] on wavelength division multiplexing in 1990 set the tone for the recent advances in this sector of networking. Optical communication can be envisaged as transmitting information optically by modulating a carrier frequency that is emitted by an optical generator such as a laser and detecting it at the far end of an optical fiber with a photo-diode. The laser emits light that is characterized by its optical frequency and hence its wavelength. (Wavelength-frequency product is a constant and gives the velocity of light.) For optical communication to occur, this frequency must be subjected to the lowest possible attenuation.

Traditionally, three low attenuation windows are available for communication that is located around the 980 nm, 1310 nm, and 1550 nm bands. By transmitting different data streams on different wavelengths (frequencies) and multiplexing these different frequencies as a composite signal, we can increase the cumulative data rate of the entire fiber. This multiplexing scheme is commonly termed WDM.

The transmitter end of the communication channel has a finite limit to the maximum data that can be modulated onto a single wavelength. Multiplexing many such data streams on different wavelengths not only increases the net data rate but also circumvents the opto-electronic mismatch to a certain extent. (Although the fiber can accommodate up to 40 Tbps of capacity, electronic systems do not currently exist that can modulate bit streams at such a high rate; this fundamentally causes the opto-electronic mismatch.) Each modulated wavelength in the composite signal is called a *channel*, and each channel is generally at a fixed spacing from its neighbors. In today's networks, each channel is usually 100 GHz/50 GHz from its neighbors; this spacing is also the standard for ITU-T grids for WDM systems today. Service providers started most WDM services with 200 GHz spacing. That was a norm for a long while, until 100 GHz became feasible.

There have been reports of up to 40 Gbps per channel of data rate, as well as several experimental demonstrations of lesser channel spacing (typically 25 and 12.5 GHz) between two adjacent channels). Lesser channel spacing and a higher bit rate per channel means that more data than before is in the fiber.

The use of different modulation schemes to increase the data rate has been studied extensively. Although non-return to zero (NRZ) and return to zero (RZ) are the common ones used, experimental demonstration of binary phase shift keying (BPSK)—in which a logical one and logical zero become two phases of an optical signal in the transmitting media—and intensity modulated (IM) schemes have also been proposed. Recent advances in subsidiary optical components—such as Fabry-Perot cavities, Mach-Zehneder interferometers, switches, and tunable components—have made WDM-based networking a commercial reality. Doped fiber amplifiers (Silicon fiber with Erbium and other rare-earths) that can amplify the entire composite WDM signal pave the way for long-haul communication.

Optics: An Update

This section looks at some of the basic phenomena that govern optical communication. Most of these phenomena are quite simplistic and intuitive and can serve as a refresher for basic optical principles. Although some of the analysis of optical communication involves complex mathematics, a simpler geometrical solution can give us an idea of the basic principles in a subtle way.

Reflection, Refraction, Total Internal Reflection, and Snell's Law

For the most part, in optical networking, it is the physics that plays a crucial role in determining the heuristics of the network and, therefore, is of paramount importance. Some of the basic optical phenomena observed in free space as well as within the fiber are reflection, refraction, birefringence, polarization, and dispersion. The first two are simple effects that are easy to understand; whereas the latter three are somewhat complicated and are severe impairments to optical communication and major contributors to attenuating a propagating signal or distorting it beyond recognition value.

Light travels in different media with different velocities. The speed of light in vacuum is approximately 3 x 10^8 meters per second, while the velocity in other media varies. A term that reflects this change is the *refractive index of the media*.

NOTE

The refractive index of a particular media with respect to a vacuum is given by the ratio of the speed of light in vacuum to that in the given media.

The refractive index is 1.5 for Pyrex glass and 1.33 for water. Mathematically, refractive index n is given as in Equation 1-1.

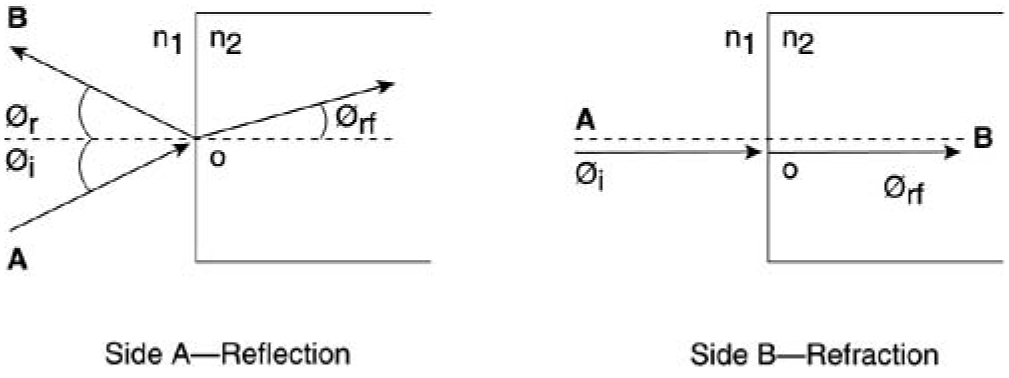
Equation 1-1 Refractive Index n

$$n = c_0/c_m$$

In this equation, c_o is the speed of light in a vacuum and c_m is the speed of light in the concerned media (whose refractive index we want to investigate).

When a ray of light (*light* is referred to as rays of light to facilitate the geometrical optic calculations (see reference 5) traverses from one medium to another, it partially gets reflected back into the incident media where it came from. This principle is known as *reflection* (see section A of Figure 1-1). The light, which is not reflected back into the incident media, is refracted into the second media, and this phenomenon is known as *refraction*. If we draw a line perpendicular to the point at which the ray of light strikes the second media, this line is called a *normal* (shown by NON' in the figure). A ray of light that is parallel to this normal passes through (from one media to another) without a change in path (see section B of Figure 1-1).

Figure 1-1.

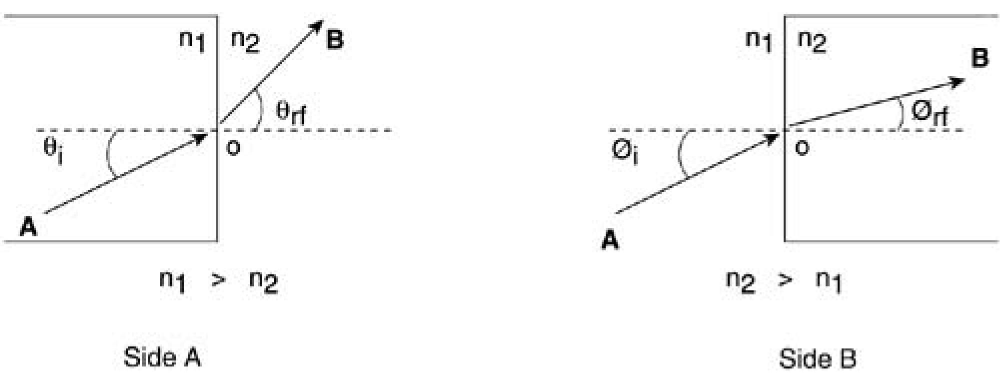


InFigure 1-1, section A, AO represents a ray incident on a denser medium from a rarer medium. A denser medium is one of high material density and also high refractive index; in contrast, a rarer medium is one with lower index and lower density. OB represents the reflected ray. In Figure 1-1, section B, OB represents the ray that passes through when the angle of incidence is 0 degrees to the normal.

Further, the angle of incidence is always congruent (equal to) to the angle of reflection. Moreover, the incident ray, reflected ray, and normal all lie in the same plane.

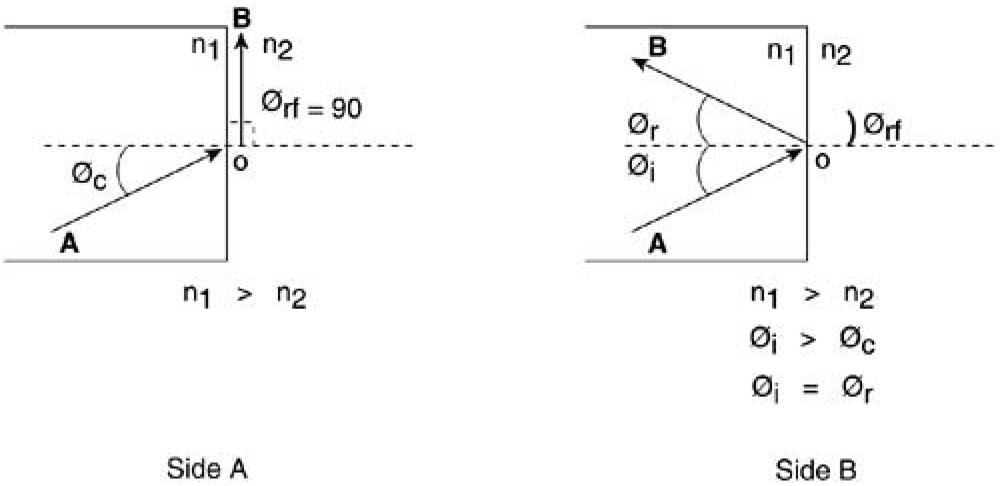
Refraction occurs when a ray of light passes from one medium to another, but its effects are observed when the angle of incidence is greater than zero. For all angles of incidence that are greater than zero, the ray of light while passing from a rarer to a denser medium bends toward the normal; in contrast, when the ray of light passes from a denser to a rarer medium, it bends away from the normal (see Figure 1-2, section B). This bending of light when passing through different mediums of different indices and hence densities is called *refraction*.

Figure 1-2. Refracted Light



If the angle of incidence of a ray of light, in a denser medium is continuously increased, the corresponding refracted ray (in the rarer medium) is bent away from the normal; this "away bending" can be mathematically explained by Snell's law (see Equation 1-2). As we increase the angle of incidence, a point is reached when the angle of refraction is perpendicular to the normal. The ray is submerged and glazes the boundary (see Figure 1-3, section A) of the two media. The minimum angle of incidence for which the angle of refraction is 90 degrees is called *critical angle*, and this value for glass is 41 degrees and 24 minutes.

Figure 1-3. Critical Angle and Total Internal Reflection



InFigure 1-2, section A, ray AO is the incident ray in a denser medium, while OB is the refracted ray in a rarer medium. Note here that in OB, the refracted ray bends away from the normal. Hence, angle of incidence is greater than angle of refraction. Further in Figure 1-2, section B, AO is the incident ray in the rarer medium while OB is the refracted ray in the denser medium. Note here that the refracted ray OB is bent toward the normal. Therefore, angle of incidence is lesser than angle of refraction.

InFigure 1-3, section A, the incident ray AO is at an angle such that the refracted ray OB glazes the surface or, in other words, has angle of refraction 90 degrees (perpendicular to the normal). Any further increase in the angle of incidence leads to total internal reflection—a condition achieved in Figure 1-3, section B. In this case, the incident ray gets reflected in the medium itself and is completely reflected back into the originating medium, thus obeying conventional laws of reflection (see Figure 1-3, section B).

Snell's law of refraction mathematically states the following, shown in Equation 1-2.

Equation 1-2 Snell's Law of Refraction

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

In this equation, n_1 and n_2 are the refractive indices of the two media, and θ_1 and θ_2 are the angles of incidence and refraction.

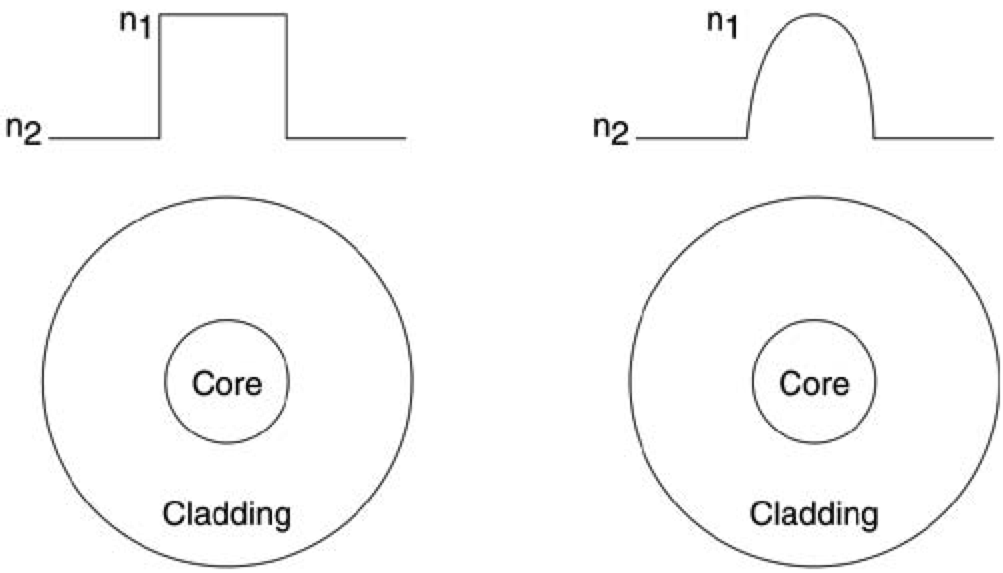
ConsiderFigure 1-5, which is a longitudinal cross section of a fiber. The fiber is a cylindrical waveguide, described by cylindrical coordinates θ,ϕ, and z. The inner conducting medium of a higher refractive index is called the *core*, whereas the outer medium (of a lower index profile) is called the *cladding*. The index profile of the core and cladding are shown in sections A and B of

Figure 1-4 for graded index (gradual shift in refractive index from center to periphery) and step index (discrete shift in refractive index from center to periphery at the demarked core-cladding boundary).

Figure 1-4. Longitudinal Cross Section of a Fiber

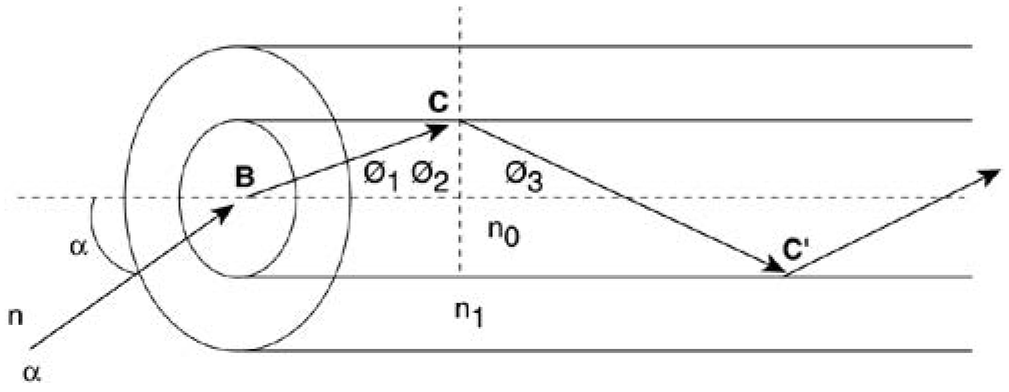
Step Index fiber profile ($n_1 > n_2$) Graded index fiber ($n_1 > n_2$)

n_1 = index of the core and n_2 = index of the cladding.



A ray of light AB incident (see Figure 1-5) to the core hits the core at angle θ_{AB}, and obeying Snell's laws, gets refracted toward the denser medium. This refracted ray BC strikes the core-cladding boundary at C and undergoes total internal reflection, provided that the angle BCC' is greater than the critical angle for that medium. This phenomenon continues in the z-direction (propagation direction) and serves as the basis of fiber-optic communication.

Figure 1-5. Geometric Optics Principle for Optical Communication in a Fiber



From Snell's law, it is possible to calculate the maximum angle that the incident ray can make with the axis OO' so that it remains within the periphery of the core (gets contained in the core).

The light-gathering capability of an optical fiber is called *numerical aperture* (NA). The greater the numerical aperture, the greater the light-gathering capacity. The acceptance angle α determines the amount of light that a fiber collects (see Figure 1-5). The acceptance angle is measured in terms of numerical aperture.

NA can be derived as shown in Equations 1-3 and 1-4.

Equation 1-3

$$n \sin\alpha = n_0 \sin\Phi_1$$

Equation 1-4

$$n_0 \sin\Phi_2 = n_1 \sin\Phi_3$$

In the equations, n, n_o, and n_1 are the refractive indices of each medium.

Substituting n = 1 for air, in Equation 1-3, we get Equation 1-5, we get the following.

Equation 1-5

$$\sin\alpha = n_0 \sin\Phi_1$$

Equation 1-6

$$\Phi_2 = \frac{\pi}{2} - \phi_1 \text{ (geometric property)}$$

$$n_0 \sin\frac{\pi}{2} - \phi_1 = n_1 \sin\Phi_3$$

Substituting preceding trigonometric relation back in Equation 1-6, we get the following result.

$$\sin\frac{\pi}{2} - \phi_1 = \cos\Phi_1 \text{ (trigonometric relations)}$$

we get

$$n_0 \cos\Phi_1 = n_1 \sin\Phi_3$$

Equation 1-7

$$n_0 \cos\Phi_1 = n_1 \text{, or } \cos\Phi_1 = n_1/n_0 \text{ (for critical angle } \Phi_3 = 90 \text{ and } \sin 90 = 1\text{)}$$

$$\sin^2\theta + \cos^2\theta = 1 \text{, so } \sin\theta = \sqrt{1 - \cos^2\theta} \text{ (standard identity)}$$

Substitute this in Equation 1-5.

$$\sin\theta_1 = \sqrt{1 - \left(\frac{n_1}{n_0}\right)^2}$$

$$\sin\propto = n_0\sqrt{1 - \left(\frac{n_1}{n_0}\right)^2}$$

$$= \sqrt{n_0^2 - n_1^2}$$

Birefringence

In certain transparent materials, a refractive index varies as a function of the direction of the incident ray and polarization. *Birefringence* literally means "double refraction." When nonpolarized light falls on birefringent material, it refracts the nonpolarized incident ray into two orthogonally polarized light rays. These rays are horizontally and vertically polarized (a more detailed description on polarization is given in the "Polarization" section as well as in reference[2]). Of these two rays, one ray is called ordinary ray "O," which obeys Snell's law; the other ray is called extraordinary ray "E," and it does *not* obey Snell's law. All crystals that have cubic lattice structures have some birefringent properties in them. Certain materials show birefringent behavior when subjected to mechanical stress.

Birefringence phenomena in fiber-optic communication causes the pulse to spread. In an ideal case, the optical fiber would be a uniform medium that is perfectly cylindrical and free of mechanical stress. A single wavelength of light would propagate in a fiber without any change in its state of polarization. Practically, the fiber core is not a perfect cylinder, which might have undergone nonuniform mechanical stress and hence be deformed. These defects usually give rise to birefringence within the fiber (assuming strong birefringent properties) that causes a single nonpolarized light pulse to split into horizontally polarized and vertically polarized pulses.

Due to differential group-delay (DGD) between vertical and horizontal pulses, the traveling pulse gets distorted during transmission in a fiber. DGD has a significant impact on the maximum bit rate that is possible in an optical fiber. Group delay is a function of the birefringence in the fiber for the entire length and also depends on the temperature and mechanical stress of the fiber. A useful way to characterize group delay is by polarization mode dispersion (PMD). DGD occurs and affects optical signals across the entire transmitting spectra and makes no exception to the wavelength of the signal. Being a universal phenomenon, DGD can be cured by certain compensation techniques discussed later.

Polarization

Light is a form of electromagnetic radiation; it has electric (E) as well as magnetic (H) fields that are orthogonal to each other as its elementary constituents. These time-varying (E) and (H) fields of an electromagnetic wave are said to be *linearly polarized* if the direction of their components and magnitudes is constant over time. This condition of constant proliferation of the axial (X,Y,Z) components is called *circular polarization*. As light propagates through a fiber, the wave constantly interacts with the medium. This interaction leads to a condition in which the individual components are no longer equal in magnitude and direction, which in turn leads to Polarization mode dispersion (PMD) (explained in detail in the "Dispersion" section). The interaction of light with the medium leads to a change in the electric dipole moment per unit volume, or the polarization, producing elliptical or noncircular fields.

The degree of polarization (P) is defined as shown in Equation 1-8.

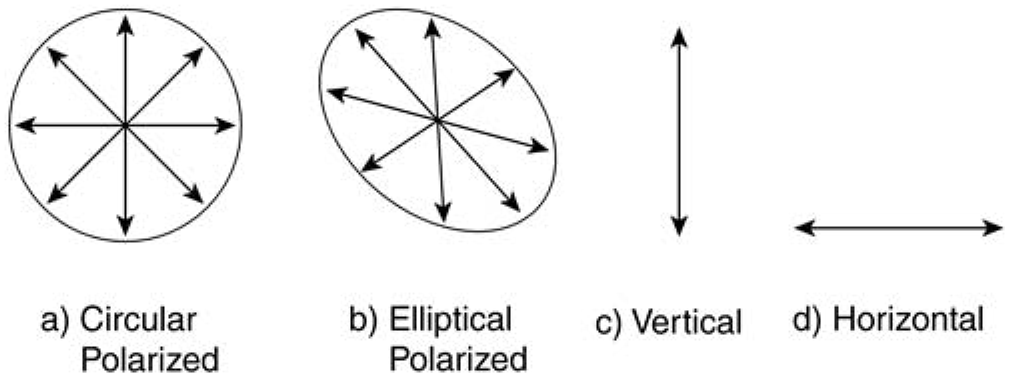
Equation 1-8

$$p = \frac{I_{pol}}{I_{pol} + I_{unpol}}$$

In the equation, I_{pol} equals strength of polarized component, and I_{unpol} equals strength of unpolarized component.

Polarization can be a resultant of reflection, refraction, or scattering. An incident ray that undergoes reflection, refraction, or polarization and is subjected to interaction with the media or with itself gets polarized. The degree of polarization depends on the angle of incidence, the refractive index, and the scattering profile of the media. Figure 1-6 shows different polarization profiles of signals. A is circularly polarized, B is elliptically polarized, C is vertically polarized, and D is horizontally polarized.

Figure 1-6. Different Types of Induced Polarization to an Electromagnetic Pulse Within the Fiber



Dispersion

Dispersion is an inherited property of the fiber that can be attributed to the spreading of an optical pulse in time domain due to the difference in the velocities of the various spectral components that are associated with that optical pulse. We have to note that each optical pulse has different spectral components or multiple frequencies. Each spectral component has its own velocity and can travel through a different path. Because of this, each component reaches the exit end of a communication channel (fiber) at different intervals of time. This difference in time experienced by the various spectral components leads to a longitudinal spreading of the pulse in the z-direction of a cylindrical waveguide. Note that the z-direction of a cylindrical waveguide is the direction of propagation of the optical pulse.

The amount of optical spread depends on two factors: the bit rate and the length of the communication channel (fiber). For high bit rates, two consecutive pulses are close to each other. Over a sufficiently long fiber, the dispersion might lead to intersymbol interference (ISI; that is, a pulse might be so severely distorted that it could spread into the time slot of adjacent pulses, causing difficulty in detecting the pulses). The accumulated dispersion of both the pulses spreads into one another, making it practically difficult for the receiver (at end of the communication channel) to detect the pulses correctly. The velocity at which the different spectral components in a pulse propagate is said to be the group velocity; and it is numerically calculated as shown in Equation 1-9.

Equation 1-9

$$v_g = \left(\frac{\partial\beta}{\partial\omega}\right)^{-1}$$

In the equation, β is the propagation constant and $\omega=2\varepsilon f$ is the angular frequency or simply optical frequency. Another important parameter is the wave-number k given by $k=2\pi/\lambda$.

The differential delay (the amount of pulse spread) is shown in Equation 1-10.

Equation 1-10

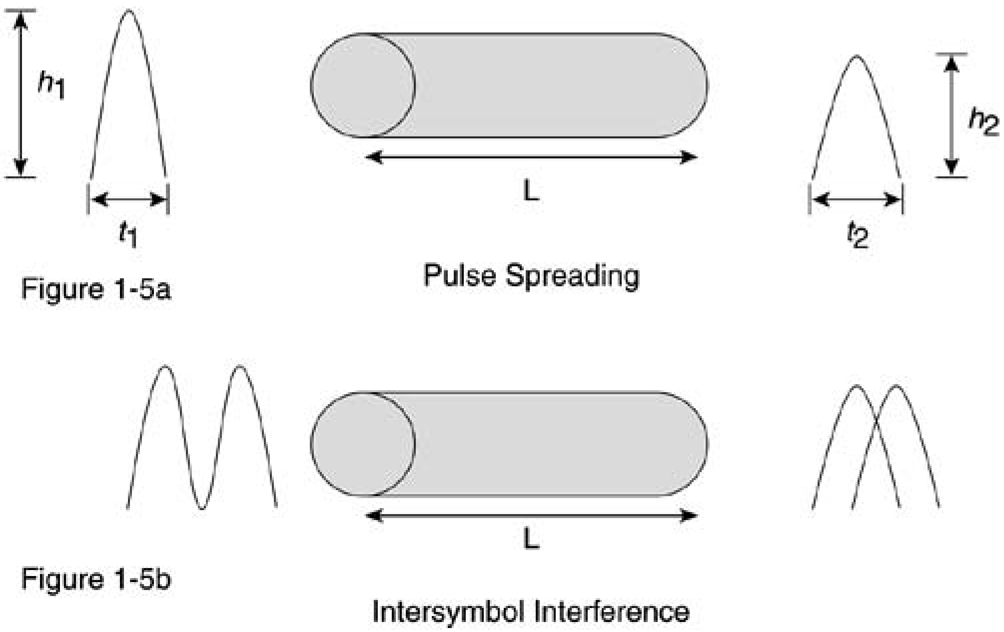
$$\delta\tau = -\frac{L}{2\pi c}\left(2\lambda\frac{d\beta}{d\lambda} + \lambda^2\frac{d^2\beta}{d\omega^2}\right)\delta\omega = L\left(\frac{d^2\beta}{d\omega^2}\right)\delta\omega$$

In this equation, L is the length of the fiber through which we desire to calculate the group delay, λ is the wavelength, and $\beta_2 = d^2\beta/d\omega^2$ is known as the group velocity dispersion (GVD) parameter, which is a measure of actual broadening of the pulse. Finally, the following quantity

$$D = -\frac{2\pi c}{\lambda^2}\beta_2$$

is called the dispersion that is induced in an optical pulse. Technically, it is a measure of the pulse spread of the wavelength in terms of the group delay. Refer to Figure 1-7. This kind of dispersion is also called *chromatic dispersion*.

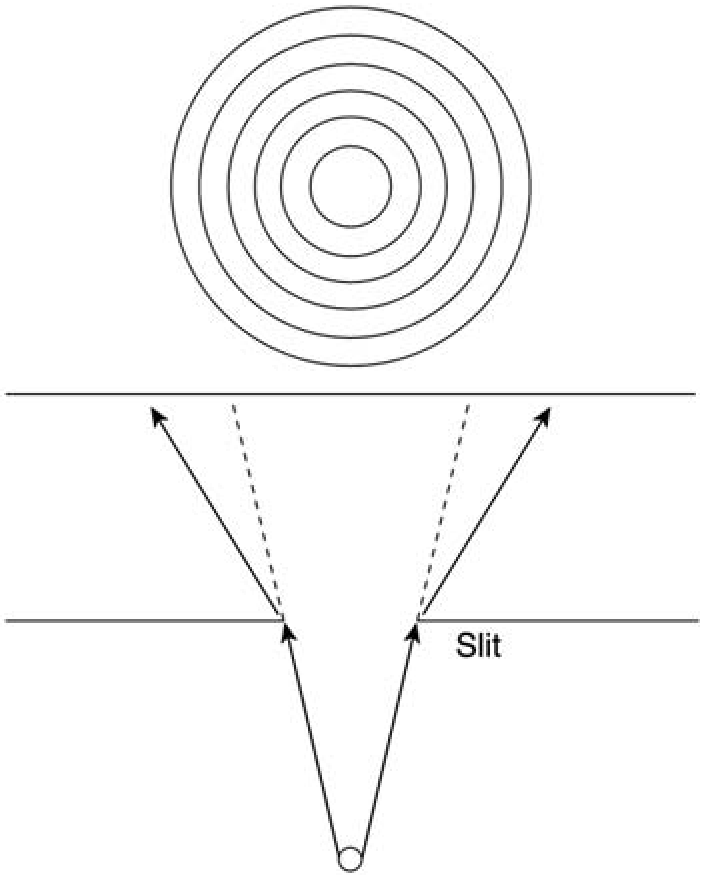
Figure 1-7. Intersymbol Interference. Pulse of width t_1 spreads to t_2 upon traveling through a fiber of length L. Two adjacent pulses spread into each other, resulting in intersymbol interference.



Diffraction

A parallel beam of light incident on the edge of a slit (hole) gets diffracted to a wider angle, and this phenomenon is due to the characteristics of diffraction of light, or Fresnel's effect[5] (see Figure 1-8). There is a wavelength dependence on the amount of diffraction the beam undergoes, and this spatial dependence on wavelength leads to the phenomenon called Bragg's diffraction. Consider a slab of glass on which numerous concentric circles are etched. If a narrow beam of light is incident upon this surface from below, each concentric circle offers a diffraction pattern that refracts each wavelength into a wider egress (output) angle. This phenomenon is known as *diffraction by gratings*. The concentric circles are called *gratings*. We can observe an example of this phenomenon in a manual overhead slide projector. (Diffraction and its relation to optical networking is explained in more detail in Chapter 2.)

Figure 1-8. Diffraction Using Etched Concentric Circles Figure 1-8 is an example of diffraction using etched concentric circles on a slab of glass and projecting a narrow beam of light on them. The beam is diffracted to a wider angle than expected.



Fiber

A *fiber* is a cylindrical waveguide in which light propagates on the basis of modal theory. Modes are solutions of Maxwell's equations for particular boundary conditions. From a layman's perspective, modes can be considered different paths of propagation in a core of a fiber. Maxwell's equations define the relation between the two components of light: electric field E and magnetic field H. Optical pulse propagation within a fiber can be described best by electromagnetic wave-propagation theory. To understand this approach, we need to solve Maxwell's equation for a cylindrical waveguide. If E and H are the electric and magnetic field vectors in space (x,y,z) and time; further B is the magnetic flux density, and D is the electric flux density, then μ_0 and ε_0 are constants of permeability and permittivity.

Maxwell's equations represent one of the most elegant and concise ways to state the fundamentals of electricity and magnetism. From these equations, we can develop most of the working relationships for optical transmission. Because of the equations' concise statements, they embody a high level of mathematical sophistication and are not usually introduced in an introductory treatment of the subject, except perhaps as summary relationships.

Equations 1-11–1-14 reproduce a set of four equations, which are the constituents of Maxwell's equation of electromagnetics. In a nutshell, these equations summarize the various relationships that are associated among electric and magnetic fields producing effects that govern the standard motion of electromagnetic waves in different media.

Equation 1-11 is called Gauss's Law of Electricity. It states that the electric flux that is associated with a closed body is proportional to the total charge of the body without exception. The divergence (the del operator) of the electric field gives the density of the sources.

Equation 1-12 mentions Gauss's Law for Magnetism as a special case of Maxwell's equation. It states that the net magnetic flux emanating from a closed object is zero. This can be understood by considering elementary magnetic theory: The most fundamental unit of magnetism is the dipole and is characterized by a magnetic North and magnetic South pole. The net result is zero due to the cancellation of equal and opposite forces. The same theory works as a basis for Gauss's Law.

Equation 1-13 is called Faraday's Law of Induction, which is stated as follows: The line integral of the electric field around a closed loop is equal to the negative of the rate of change of the magnetic flux through the area enclosed by the loop. This can be explained as the integration of the net change of electric flux over a surface giving the value of magnetic strength in the opposite direction. The line integral gives the net-generated electromotive force (EMF), or voltage, by a body that is subject to variations of magnetic flux. Consider a closed loop with some current associated in it. If this loop is made to rotate in a magnetic zone, the net surface area of the loop that cuts perpendicular lines of magnetic flux is proportional to the rate of rotation (of the loop).

Equation 1-14 gives Maxwell's equation for calculation of magnetic field. It is also called Ampere's Law. Mathematically, it states that the curl of magnetic flux gives the current flowing through a loop. This equation is extremely important for calculation of magnetic field strength.

Equation 1-11

$$\nabla X E = -\frac{\delta B}{\delta T}$$

Equation 1-12

$$\nabla X H = -\frac{\delta D}{\delta T}$$

Equation 1-13

$$\nabla . D = 0$$

Equation 1-14

$$\nabla . B = 0$$

∇ is the curl, and

$$D \cong E + P$$

and

$$B \cong \mu H$$

P is the polarization (electric).

Taking the curl of Equation 1-11 and using the standard vector formula for the associative cross product shown in Equation 1-15:

Equation 1-15

$$\nabla^2 E = \varepsilon\mu\frac{\delta^2 E}{\partial T^2}$$

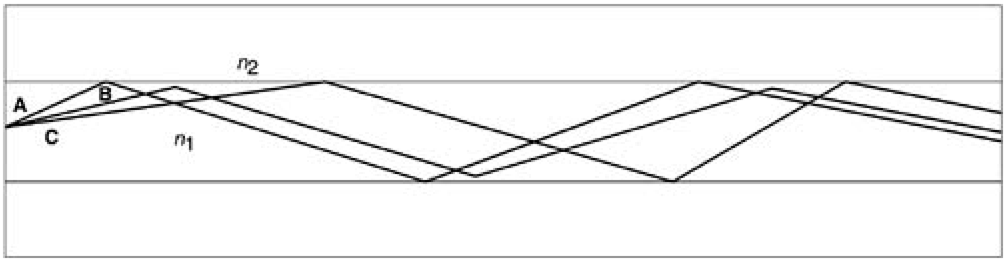
Similarly from Equation 1-12.

Equation 1-16

$$\nabla^2 H = \varepsilon\mu\frac{\delta^2 E}{\partial T^2}$$

Now consider Figure 1-9.

Figure 1-9. Different Modes (A,B,C) Propagating in the Fiber



Rays A, B, and C can be considered as three modes of propagation that are each defined by a propagation constant β. Further, each mode has a wave number.

$$k = \frac{2\pi}{\lambda}$$

We cannot accurately analyze pulse propagation within a fiber by using geometrical optics alone because it might lead to inaccuracy. The geometric optics limit the solution to an approximation of the numerous parallel wavefronts. In addition, the geometric optics do not give an idea of the field distribution or an exact analysis of the same. Finally, the energy flow in the waveguide is not possible by using this approach.

The solution of a waveguide equation for a cylindrical waveguide that is represented by r, ϕ, and z involves finding components E_r, E_f, H_r, and H_f. By considering E and H as a scaled function of time t and propagation constant β, the electric field vector E and magnetic field vector H are both quantitatively dependent on the exponential variation of β with respect to the direction of propagation; thus for a cylindrical waveguide, refer to Equations 1-17 and 1-18.

Equation 1-17

$$E \propto e^{j(\omega t - \beta z)}$$

Equation 1-18

$$H \propto e^{j(\omega t - \beta z)}$$

By normalizing, we scale the solution of preceding equations and then suit it to meet our conditions of extremity that are the boundary conditions; normalizing and equating to the boundary conditions $E = E_0$ and $H = H_0$ such that

Equation 1-19

$$E = E_0(r, \phi)e^{j(\omega t - \beta z)}$$

Equation 1-20

$$H = H_0(r, \phi)e^{j(\omega t - \beta z)}$$

Substitutingequations 1-19 and 1-20 in Maxwell's equations 1-11 and 1-12, we can get values for E_r, E_f, H_r, and H_f in terms of E_z and H_z as well as r, β,ω,ε, and μ.

Further solving, we get the wave equations in cylindrical coordinates, as shown in Equation 1-21.

Equation 1-21

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r}\frac{\partial E}{\partial r} + \frac{1}{r^2}\frac{\partial^2 E_z}{\partial \phi^2} + (k^2 - \beta^2)E_z = 0$$

Cutoff Condition and Single Mode Fiber

For a particular mode to exist and successfully propagate, it must have a field (both E,H) that does not decay outside the core.

The solution of Equation 1-25 for a generalized case by separation of a variables method yields to an analysis of four variables: r, ϕ, z, and t. The solution for z and t is given by Equations 1-23 and1-24, whereas ϕ can be approximated as a harmonic sinusoidal.

Thus,

Equation 1-22

$$E_z = k\alpha(r)\gamma(\phi)\theta(z)\tau(t)$$

Equation 1-23

$$\phi(z)\tau(t) = e^{j(\omega t - \beta z)}$$

Similarly,

Equation 1-24

$$y\phi = e^{j\phi n} n = -\alpha \ldots\ldots\ldots -2, -1, 0, 1, 2, 3\ldots$$

$\alpha(r)$ can be solved as a differential equation for Bessel's Function.[6]

While we are approximating the cutoff condition, we come across an important parameter called the normalized frequency V or V parameter, such that

$$v^2 = \left(\frac{2\pi a}{\lambda}\right)^2 (n_1^2 - n_2^2)$$

In the equation, 'a' is the radius of the core, n_1 and n_2 are the refractive index of the core and cladding, and λ is the wavelength of propagation.

Another important parameter is the normalized propagation constant b, given as follows:

$$b = \frac{\left(\frac{\beta}{k}\right)^2 - n_2^2}{n_1^2 - n_2^2}$$

In the equation, $\propto$ is the propagation constant and

$$k = \frac{2\pi}{\lambda}$$

The cutoff wavelength is an important parameter in single-mode fiber (SMF). In optical fiber with a specific core diameter, we can only transmit light at a wavelength longer than the cutoff wavelengthλ_c. If we decrease the wavelength below λ_c, it begins to exhibit other modes, and the fiber is no longer an single mode fiber for that wavelength. The implication are that a single mode fiber that is manufactured for transmission at 1.3 µm is also a single mode at 1.55 µm because the fiber remains single mode as long as the wavelength is larger than λ_c.

On the other hand, an single mode fiber that is designed to work at 1.55 µm might not remain an single mode fiber at 1.3 µm. We can always refer to the manufacture's specification to find out the cutoff frequency of the fiber for designing DWDM networks.

The higher the value of V, the higher the number of modes that the fiber supports. For single mode fiber, which supports the smallest fundamental mode, v = 2.405. Using v = 2.405 and λ = 1550 nm yields a fiber with an approximately 3–4 um core.

Fiber Losses

This section discusses various impairments that affect signal propagation and limit transmission distances in fibers. Attenuation is the most fundamental impairment that affects signal propagation. It is quite standardized and is given as a specification for a particular fiber type. Attenuation is a property of the fiber, and it is a result of the various material, structural, and modular impairments in a fiber.

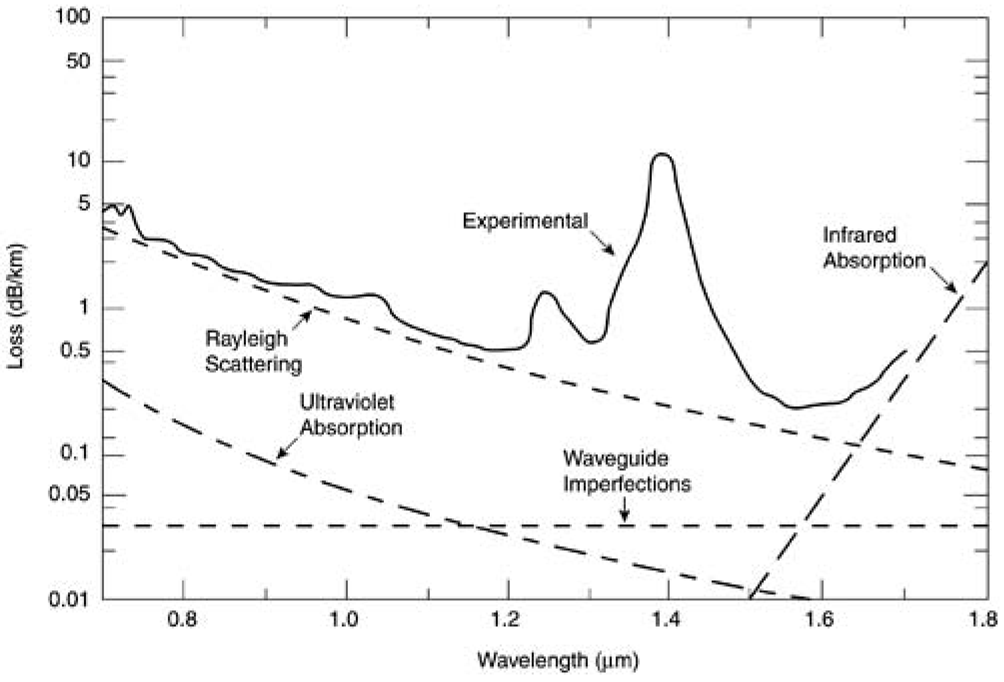
Scattering is another serious source of impairments for a fiber. Among scattering phenomena, Raleigh scattering is the most prominent. Raleigh scattering is quite prominent in optical fibers, and its profile follows a unique wavelength distribution. As signal rates increase, dispersion becomes a serious impairment. Although dispersion does not attenuate the signal as such, it causes severe pulse spreading, leading to difficulty at the receiver end in trying to decode the signal. Dispersion consists of two main types: chromatic dispersion and PMD. The latter is quite prominent at high bit rates. Fiber nonlinearities are another source of severe impairment at high rates. Phase modulation of an optical signal by itself (self-phase modulation, or SPM) or by an adjacent signal on some adjacent wavelength (cross-phase modulation, or XPM) are two sources of penalty in long-haul transmission links. Fourwave mixing, Raman, and Brillouin effects are three more nonlinear effects that affect communication.

Attenuation

Fiber attenuation can be defined as the optical loss that is accumulated from a source to sink along a fiber link. It consists of two components: an intrinsic fiber loss and an extrinsic bending loss. Intrinsic loss can be further characterized by two components: a material absorption loss and a Raleigh scattering.

Material absorption accounts for the imperfection and impurities in the fiber. The most common impurity is the $^{-}$OH molecule, which remains as a residue despite stringent manufacturing techniques. The $^{-}$OH molecule has an absorption peak at 2.73 µm in the optical spectrum, which means that wavelengths near 2.73 µm have high attenuation. Correspondingly, the $^{-}$OH molecule yields harmonics at 0.95 and 1.4 µm. As per the attenuation graph shown in Figure 1-10, the 1.4 µm peak is a severe hindrance to commercial optical communication.

Figure 1-10. Attenuation Curve in a Fiber (Reprinted from IEEE Electronics Letters, 1979)



Absorption also occurs as a result of group 3 (transition) elements being present in the fiber. Lucent Technologies and Corning use a unique manufacturing process to develop fiber types that do not have an $^{-}$OH peak, which almost eliminates the $^{-}$OH molecule. These types of fibers (such as AllWave from Lucent and SMF-28e from Corning) extend the range from 1250 nm to 1700 nm, resulting in more capacity. Attenuation that results from absorption limits the use of wavelengths above 1.7 µm for optical communications. (See the section titled "Fiber Types" at the end of this chapter for more details.)

Raleigh Scattering

Light scatters due to dense fluctuations in the core leading to a phenomenon known as *Raleigh scattering*. This phenomenon results from the collision of light quanta with silica molecules, causing scattering in more than one direction. Depending on the incident angle, some portion of the light propagates forward and the other part deviates out of the propagation path and escapes from the fiber core. The amount of Raleigh scattering a signal is subject to is inversely proportional to the fourth power of wavelength ($R \alpha \lambda^{-4}$). Therefore, short wavelengths are scattered more than longer wavelengths. Any wavelength that is below 800 nm is unusable for optical communication because attenuation due to Raleigh scattering is high. At the same time, propagation above 1.7 µm is not possible due to high losses resulting from infrared absorption.

Bending Losses

Bending of the fiber can be classified as microbending and macrobending. *Microbending* is caused by imperfections in the cylindrical geometry of fiber during the manufacturing cycles. *Macrobending* is the result of bending of fiber in small radius (radius in order of cm). Both bending phenomenon causes attenuation in the fiber.

Attenuation Coefficient

The attenuation coefficient α is expressed in dB per kilometer and represents the loss in dB per kilometer of fiber.(A note on decibel-dB is given later in this section.)

Power attenuation is shown in Equation 1-25.

Equation 1-25

$$\frac{dP}{dz} = -\alpha P$$

In the equation, dP/dz is the change in power with respect to length.

If P is the input power and L is the total length of the fiber, we can express output power P_2 as shown in Equation 1-26.

Equation 1-26

$$P_2 = P_1 e^{-\alpha L}$$

From preceding equation, α can be derived as shown in Equation 1-27.

Equation 1-27

$$\alpha = -\frac{10}{L}\log_{10}\frac{P_2}{P_1}$$

In the equation, α is expressed in db/Km.

Typical values of α for a single mode fiber are 0.25 dB per kilometer in the 1550 nm band and 0.5 db per kilometer in the 1310 nm band. Optical amplifiers (see Chapter 3) can compensate for attenuation based on doped fiber and semiconductor optical amplifiers (SOAs) as well as Raman amplifiers. Fiber manufacturers usually specify the value α in their datasheets.

Dispersion in Fiber

The velocity of propagation of light depends on wavelength. The degradation of lightwaves is caused by the various spectral components present within the wave, each traveling at its own velocity. This phenomenon is called *dispersion*. Several types of dispersion exist, two of which include chromatic dispersion and polarization mode dispersion (PMD). Chromatic dispersion is common at all bit rates. PMD is comparatively effective only at high bit rates. Waveguide and material dispersion are forms of chromatic dispersion, whereas PMD is a measure of differential group delay of the different polarization profiles of the optical signal.

Because of the dual nature of light, we can approximate it as waves as well as quanta (particles). During the propagation of light, all of its spectral components propagate accordingly. These spectral components travel at different group velocities; this observed phenomenon leads to dispersion called Group velocity dispersion or GVD. The velocity of individual groups is called *group velocity* (V_{g}) and is shown in Equation 1-28[7].

Equation 1-28

$$v_g = \left(\frac{d\beta}{d\omega}\right)^{-1}$$

∝ is the propagation constant and ω is the optical frequency. Further,

$$\beta = \frac{\bar{n}}{\omega c}$$

Due to the difference in velocities experienced by various spectral components, the output pulse is time scattered and dispersed in the time domain. The effect of dispersion on bit rate has been approximated by[1] and is given by the condition shown in Equation 1-29.

Equation 1-29

$$BL|D|\,\Delta\lambda < 1$$

In the equation, B is the bit rate, L is the length of communication channel, D is the dispersion parameter, and Δλ is the range of emitted wavelengths (spectral width of source).

From this relation, we can observe a finite limit to both bit rate and propagating length considering physical limits on the narrowness of the spectral source. One way to increase the BL-product is to employ dispersion compensation techniques described in Chapters 3 and 4. Dispersion resulting from GVD is termed *chromatic dispersion* due to the wavelength dependence (chroma is the different colors or wavelengths associated in a spectrum) and is expressed in ps/km-nm.

Polarization Mode Dispersion (PMD)

The fiber is not truly a cylindrical waveguide, but it can be best described as an imperfect cylinder with physical dimensions that are not perfectly constant. The mechanical stress exerted upon the fiber as well as the imperfections resulting from the manufacturing process are the reasons for the variations in the cylindrical geometry. This variation also leads to a phenomenon called birefringence whereby a fiber that acquires birefringence causes a propagating pulse to loose the balance between the polarization components. This leads to a stage where different polarization components travel at different velocities creating pulse-spread, and this spread is PMD.

The degree of birefringence (B_{ire}) is calculated as the difference between the indexes of the polarization component (now termed mode indexes) due to the different magnitude of these components, gaining different modal properties. This can be visualized as the two discrete orthogonal polarization states as two separate modes. B_{ire} (degree of birefringence or just birefringence) is a time-varying phenomenon that carries a state of random polarization of the induced pulse.

The two polarized components (from this point referred to as the two modes due to polarization or just modes) exchange their power over a period, T. The length in which power from one mode is transferred to the other is called the *beat length*. Refer to Figure 1-11.

Figure 1-11. PMD

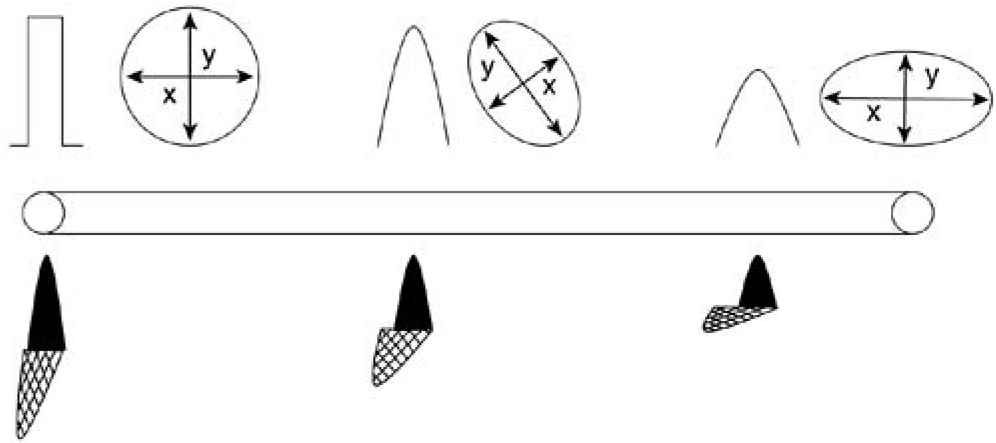


Figure 1-11 shows PMD resulting from the effects of a fiber acquiring birefringence and the energy transfer between the two polarized modes, leading to pulse spread.

The random change in the net polarization of the signal is an issue for short pulses (10–100 ps long). The two polarization components travel at different speeds because of the different power (birefringence) and different group velocities associated with them. The end result is that an induced pulse becomes considerably broader after traveling through a fiber.

The amount of broadening of the pulse is given by $L\delta B_{ire}$ in time units, where L is the length of the fiber and δB_{ire} is the rate of change of modal birefringence with respect to angular frequency ω, normalized by the wave number.

This broadening of the induced pulse due to the different velocities exhibited by induced polarization components leads to a dispersive phenomenon called Polarization Mode Dispersion or (PMD).

PMD compensation techniques are commercial realities today. Dispersion-maintaining fibers are commercially available and are made intentionally by introducing degrees of birefringence in them that negate the effects of PMD over a length of transmission.

Dispersion compensation is a useful technique for long haul as well metropolitan area networks, especially at high data rates (short pulses).

State of polarization and measurement of polarization are two important effects that are mentioned in Chapter 9.

Material Dispersion

Silica, like other materials absorbs electromagnetic radiation at resonant frequencies. Moreover, the refractive index is a function of the frequency and is estimated by the Sellmeier equation (refer to following Note on Sellmeier's Equation) Material dispersion is proportional to the differential of the group index.

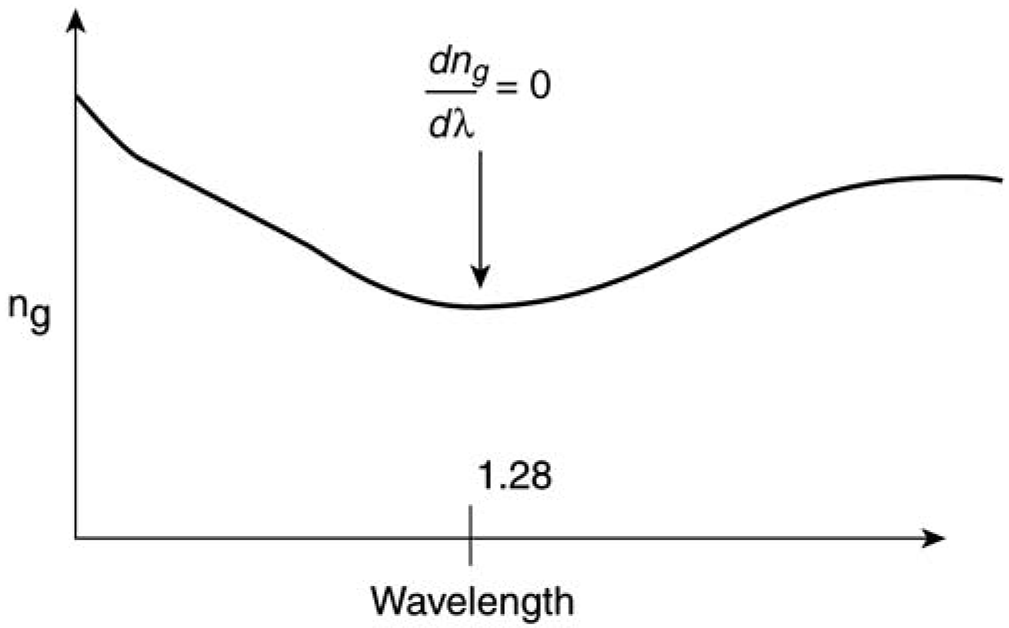
> NOTE
>
> Sellmeier's Equation
>
> $$n^2(\omega) = 1 + \sum_{j=1}^{m} \frac{B_j \omega_j^2}{\omega_j^2 - \omega^2}$$
>
> In the equation, $n(\omega)$ is the frequency (and therefore wavelength)-dependent refractive index. Characteristic resonant frequencies at which the fiber absorbs energy is approximated by Sellmeier's equation. In the equation, B_j is the strength of the j^{th} resonance, whereas m is generally bounded to 3.

Figure 1-12 shows the variation of the group refractive index with wavelength. Equation 1-30 gives differential of group index as a function of wavelength.

$$n_g\left(n_g = n + \omega\frac{dn_g}{d\lambda}\right)$$

with respect to wavelength λ.

Figure 1-12. Variation of Group Refractive Index with Wavelength



Equation 1-30

$$D_m = c^{-1}\frac{dn_g}{d\lambda}$$

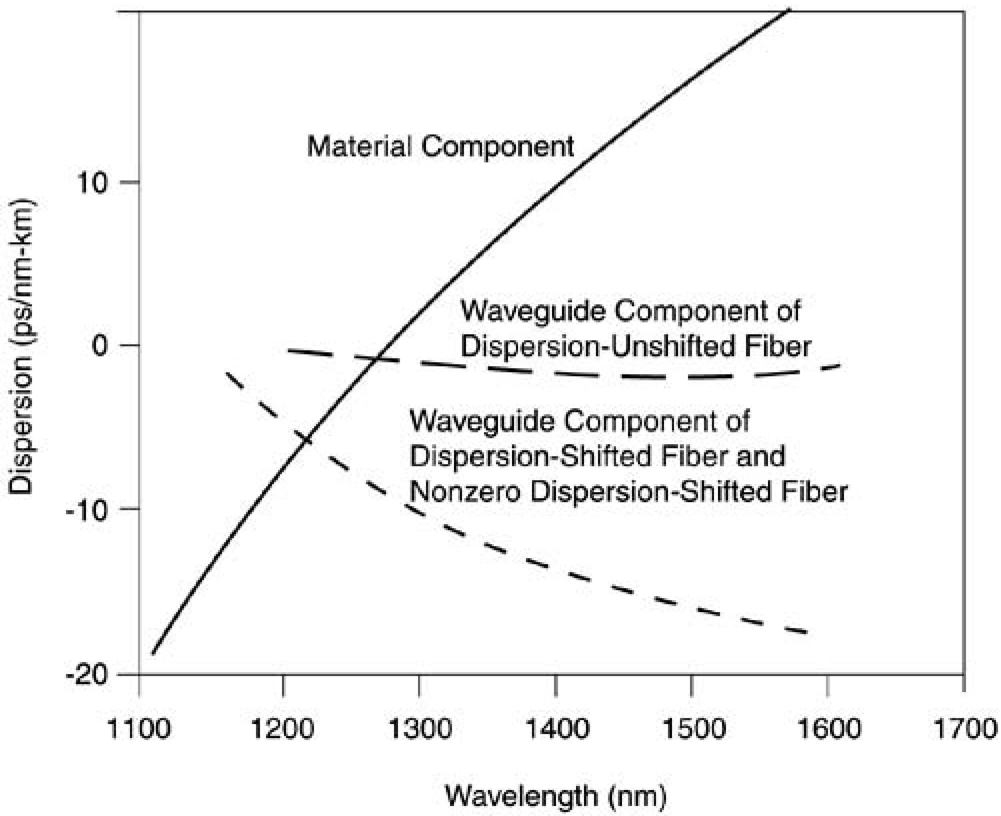
Further

$$\frac{dn_g}{d\lambda} = 0$$

atλ = 1.28µm. This wavelength is called the *zero dispersion wavelength*. $dn_g/d\lambda$ is –ve at wavelengths lower than 1.28µm and +ve above 1.28µm.

Figure 1-13 shows the different dispersion shifted and unshifted fiber profiles.

Figure 1-13. Dispersion Shifted and Unshifted Fiber Profiles



Waveguide Dispersion

We can define wave number k_0 as shown in Equation 1-31.

Equation 1-31

$$k_0 = \omega/c = 2\pi f/c = 2\pi/\lambda \text{ ---}$$

Referring to modal theory, the cutoff condition is defined as shown here.

Equation 1-32

$$v = k_0 r_1 (n_1^2 - n_2^2)^{\frac{1}{2}}$$

In the equation, n_1 and n_2 are core and cladding indexes, and r_1 is the core radius.

V is called the *normalized frequency*, and it is proportional to ω. Dispersion due to the effect of V is called *waveguide dispersion*.

NOTE

Waveguide dispersion depends on Equations 1-34 and 1-35, and therefore, depends on r_1, n_1, and n_2. By manipulating r_1, n_1, and n_2, we can achieve $dn_g/d\lambda = 0$ at 1.55 µm. These fibers are called *dispersion-shifted fibers* because of the lateral shift in the zero dispersion wavelength.

Nonlinearities

Under the influence of electric and magnetic fields, light in optical fibers exhibits nonlinear effects. Primarily, nonlinearity in optical fibers can be traced to the susceptibility $\chi(i)$; this susceptibility directly relates to the polarization vector, P.

Third-order susceptibility $\chi^{(3)}$ is a significant source of nonlinearity in optical fibers. The origination of the nonlinear property stems from the nonlinear component of refractive index. Refer to Equation 1-33.

Equation 1-33

$$\sum \bar{n}_1 = n(\omega_1) + \bar{n}(|E|^2)$$

In the equation, $\bar{n}_1$ is the nonlinear refractive index that is proportional to the real part of third-order linear susceptibility. When light propagates through a medium, the photons interact with the molecules during propagation. Photons also interact with themselves and cause scattering effects such as stimulated Raman scattering (SRS) and stimulated Brillouin scattering (SBS), which are in the +z and –z directions (forward and reverse directions of propagation) along the fiber.

Scattering in this instance means a sporadic distribution of energy in a random direction.

In optics, nonlinearities might conserve the net energy content of a pulse or might not do so depending on whether the nonlinearities are elastic or inelastic. Nonlinear effects that are due to the third-order susceptibility are generally elastic or energy conserving in the sense that the propagating pulse that experiences the nonlinear effect does not loose its energy to the medium.

On the other hand, nonlinear effects exist in which the net content of the energy is scattered into the interacting nonlinear medium. SRS and SBS are inelastic scattering effects. SRS is due to the photon interaction (and hence scattering) with the medium, whereas SBS is due to the acoustic properties of photon interaction with the medium.

SRS and SBS are similar except that SRS scatters in both forward as well as reverse directions, whereas SBS scatters in reverse only. In both SRS and SBS, a wave called Stoke's wave is generated due to the scattering of energy. This could in fact be an amplifying wave of high energy. The gain obtained by using such a wave creates Raman and Brillouin amplification. The Raman gain can extend most of the operating band (C and L band) for WDM networks; therefore, it is an excellent WDM-amplifying technique. The Brillouin gain peaks in a narrow peak near the C band. This amplification phenomenon is discussed in Chapter 3 in detail.

Four-Wave Mixing and Difference-Frequency Generation

Three optical frequencies (f_1, f_2, and f_3) interact in a nonlinear medium giving rise to a fourth frequency (f_4), which is formed by the scattering of the three incident photons producing the fourth photon. This is shown in Equation 1-34.

Equation 1-34

$$f_4 = f_1 + f_2 - f_3$$

When two photons interact with each other in a nonlinear media, they produce a third photon that has an optical frequency based on the difference of the two interacting frequencies. Decreasing channel spacing and high chromatic dispersion will increase FWM effects.

FWM causes interchannel crosstalk and is worst-case for equally spaced WDM channels. This topic is discussed in more detail in Chapter 4.

Window of Operations

Researchers have always argued about the number of operating windows of wavelengths or bands that can exist in an optical communication network. To the designer or systems engineer, this is not much of an issue for argument because practical WDM networks currently function in three discrete bands. These three prominent bands are the C, L, and S bands. The conventional (C) band is approximated from 1525–1565 nm. It has a low loss of about 0.2 dB per kilometer. Most metropolitan as well as long-haul networks use this band. The band is about 40 nm and can accommodate 50 different wavelengths, each 100 GHz (or 0.8 nm) apart or 100 wavelengths at 50 GHz. The spacing between the wavelengths is a standardized value. Currently, for dense division multiplexing, the spacing is standardized at 0.8 nm or 0.4 nm.

The long (L) band starts from about 1570 nm and extends to 1620 nm. It has slightly higher loss than the C band but similar characteristics to the C band. Much research has been carried out in this band, and signs of early commercial deployment are evident. The future will see many vendors positioning their DWDM products and technologies in this band.

The short (S) band is spread around the 1310 nm window. It is of strategic importance due to its close proximity to the zero dispersion wavelength (a wavelength around 1300 nm that has minimal dispersive effects due to the cancellation of material and waveguide dispersions by each other). The S band has a higher loss than the C band at about 0.5 dB per kilometer; therefore, it is not the best solution to long-haul communications. The evolution of wider technologies for the C band—such as doped amplifiers, switch matrices, and filters—make the S band rather underutilized.

Apart from these three standard bands is the traditional 850 nm band, which was used first for optical communication systems. The 850–980 nm band is used most frequently for multimode systems and for short LANs. It has a high loss characteristic of almost 2–3 db per kilometer. Experimental research is being carried out in the 1400 nm segment by new methods to eradicate the $^{-}$OH molecule. The best that a design engineer could hope for is to have a C band from 1300–1650 nm, yielding about 400 wavelengths 0.8 nm apart or 800 wavelengths 0.4 nm apart.

Fiber Types

The most common type of single mode fiber is usually referred to as *standard single-mode fiber*. The *International Telecommunication Union* (ITU), which is a global standardization body for telecommunication systems and vendors, defines different types of fibers. Some of the different fibers described in the standardization process for optical networks include nondispersion-shifted (G.652), dispersion-shifted (G.653), 1550-nm loss minimized (G.654), and nonzero-dispersion fiber (G.655).

Nondispersion-Shifted Fiber (ITU-T G.652 Recommendation)[9]

This type of single mode fiber is also called *standard single-mode fiber*, and it is the most commonly deployed fiber. Nondispersion-shifted fibers are optimized for the 1310 nm region and have zero dispersion wavelength at 1310 nm. We can also use this type of fiber in 1550 nm regions, but it is not optimized for this region. The chromatic dispersion at 1550 nm is high (18ps/nm-km), and for high data-rate applications, dispersion compensations have to be employed. An example of this type of fiber is corning SMF-28.

Dispersion-Shifted Fiber (ITU-T G.653)[10]

In dispersion-shifted fiber, the zero-dispersion wavelength has been shifted from 1310 nm to 1550 nm. The dispersion-shifted fibers are optimized for operating in the region between 1500–1600 nm, and the dispersion coefficient, D, increases with wavelength. When this type of fiber was developed, the assumption was to take advantage of doped amplifiers and operate with multiple channel DWDM systems.

ITU G.654 (Loss Minimized at 1550 nm)[11]

This type of fiber is a special case of standard single mode fiber, which has a low loss at the 1550 nm window. ITU G.654 is optimized for the 1500–1600 nm region. The effective cutoff wavelength λ_{cutoff} is an important parameter in designing this type of fiber. Low loss can be achieved by using a pure-silica core. ITU G.654 fibers are expensive to manufacture and are rarely used. These types of fibers might be best suited for submarine (under the sea fiber cables) and extended long-haul applications.

Nonzero Dispersion-Shifted Fiber (G.655)[12]

Nonzero dispersion-shifted fibers (NZDSFs) are SMFs that have chromatic dispersion that is greater than a nonzero value throughout the C band (1500 nm) region. This dispersion reduces the effect of nonlinearities, such as four-way mixing, self-phase modulation, and cross-phase modulation, which are seen in the DWDM systems. These types of fibers are best suited and optimized to operate between 1500–1600 nm.

Two types of NZ-DSF are available. If the dispersion slope of NZDSF decreases with respect to wavelength, in other words, there is a negative gradient for dispersion as a function of wavelength, the fiber is called –NZDSF. Likewise if the dispersion slope (also called dispersion profile) increases with increase in wavelength, the fiber is called +NZDSF.

Units of Optical Power Measurement: Decibel

The power level in optical fiber communications is too wide ranged to express on linear scale. A logarithmic scale known as decibel (dB) is used to express power in optical communications. The wide range of power values makes decibel a convenient unit to express the power levels that are associated with an optical system. The gain of an amplifier or attenuation in fiber is expressed in decibels. The decibel does not give a magnitude of power, but it is a ratio of two powers. See Equations 1-35 and 1-36.

Equation 1-35

$$(\text{Loss or gain})\ \text{dB} = 10\log_{10}\left(\frac{P_2}{P_1}\right)$$

Example:

Calculate the gain of the amplifier in dB, when 1 watt is applied to the input and 2 watts is measured as the output:

dB = $10\log_{10} 2/1 = 3$ db Measured output is 2W

Gain of this amplifier is 3 db.

dBm is the power level related to 1 mW.

Equation 1-36

$$\text{dBm} = 10\log_{10}\frac{power(mW)}{1mW}$$

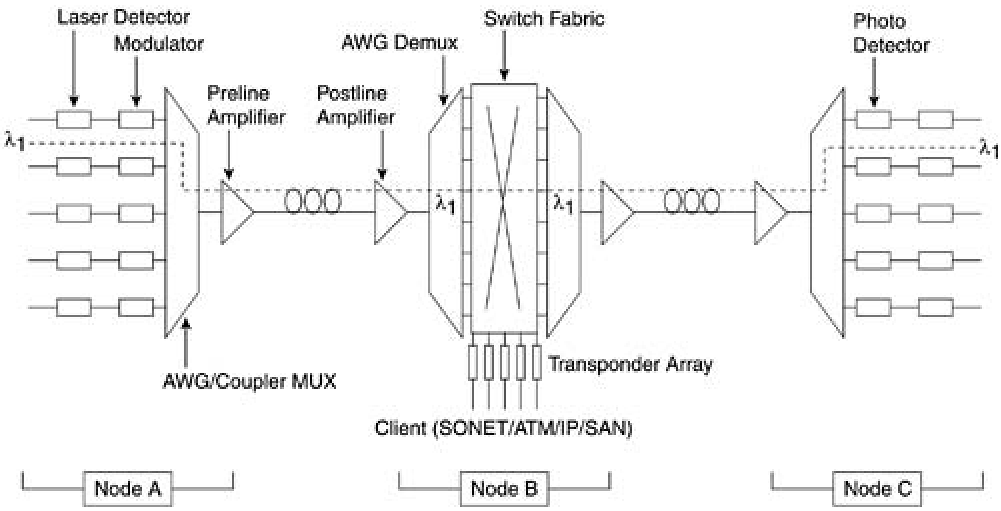
So far, we have studied different effects in optics as well as propagation of light in fibers. We now introduce a point-to-point WDM network and the various parameters as well as components associated with it from a very high level perspective. A more detailed explanation of the components and subsystems can be obtained from Chapters 2 and 3. A design ideology is available in Chapters 4,5 and 6.

A Point-to-Point WDM Network

By definition, WDM is multiplexing different optical information-bearing signals by virtue of the spatial difference in their wavelengths, compositely riding in the same optical fiber.

An electric information bearing signal is modulated onto an optical carrier frequency (signal). Many such optical signals, each at a characteristic wavelength (and hence frequency), are multiplexed into what is called a composite WDM signal. In the point-to-point WDM system as shown in Figure 1-14, node A transmits data to node C through an intermediate node B. Which is to say that we have a point-to-point link ABC consisting of three nodes: A, B, and C. The operating band is assumed to be the entire C band ranging from about 1525–1565 nm, that is, there can be communication wavelengths from 1525–1565 nm. Each channel is spaced 100 GHz / 0.8 nm (or could be 50 GHz / 0.4 nm) apart, and this spacing is defined by the ITU standard.

Figure 1-14. DWDM Point-to-Point Communication System



At node A, an array of electrical data devices injects data into different wavelengths. These devices could be SONET platforms, ATM switches, or even routers. (IP over DWDM is explained inChapter 7.) The electronic data is modulated onto an optical channel at a specified wavelength.

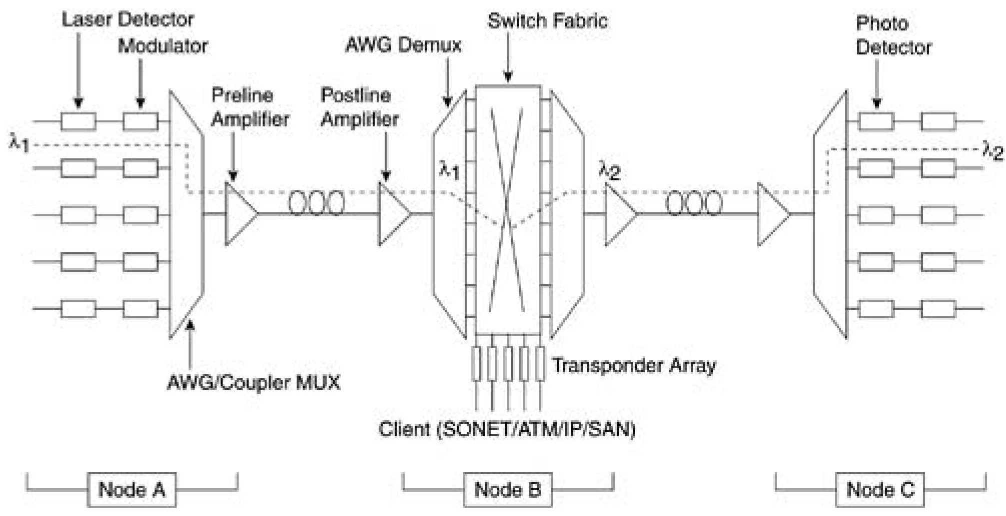
The data is fed to lasers either directly or coupled optically to the light emitted by the laser in a modulating cavity externally (modulation). The power levels of each channel are adjusted using controllable optical attenuators to avoid excess nonlinear effects.

The optical signals are multiplexed into a composite WDM signal in an arrayed waveguide (AWG)- or coupler-based multiplexers.

This composite signal is further amplified just before transmission into the fiber by fiber amplifiers (doped fiber or Raman). The signal is then injected into the transmission fiber. At an intermediate node B, the signal is first amplified by a preline amplifier that could be a doped fiber or a Raman gain amplifier depending on the level of amplification desired and quality of signal (noise figure) as well as quantity of the amplification (gain) required. The amplified composite signal is demultiplexed by an arrayed waveguide demultiplexer.

Each de-multiplexed channel is fed into a switching fabric, which could be an all-optical switch (O-O-O) or an opto-electronic switch (O-E-O). An O-O-O switch has the functionality of switching or routing the channels completely in the optical domain. Refer to Figure 1-15.

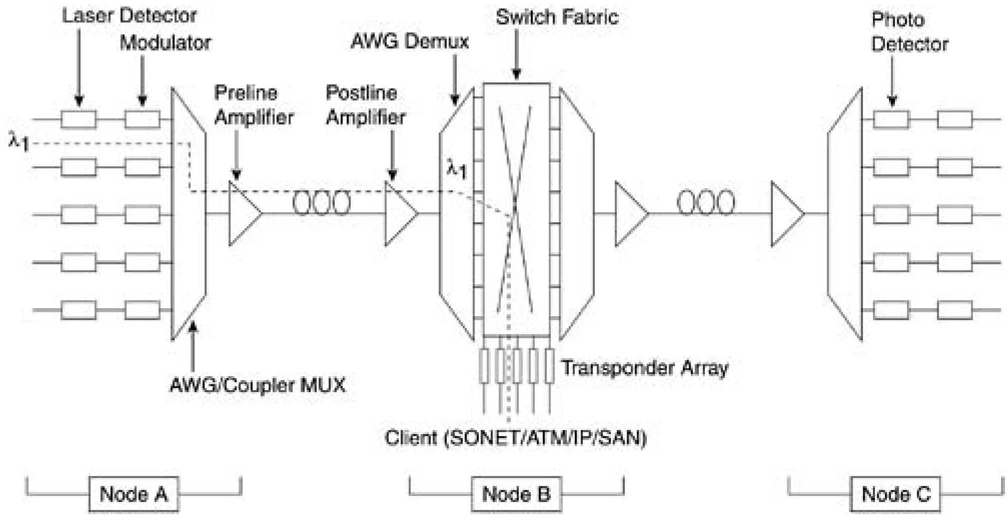
Figure 1-15. A DWDM System That Has Channel Add-Drop Capabilities



An O-E-O switch could do the same as an O-O-O switch but perform the switching function in the electrical domain by converting the optical signals into electrical bit streams and reconverting them back to optical signals after switching them. This could also result in the egress signal to be on different wavelength as compared to the ingress signal creating wavelength conversion explained in Chapter 2 (section on transponders). Individual channels could also be dropped at the desired node if the destination for the channel is that particular node.

Dropping a channel usually involves reconverting a channel into client wavelength (usually shorter wavelength 1310, etc.). This is done by *transponders*. The transponders also facilitate in adding the channel at the intermediate node site. The transponders convert the incoming wavelength into an ITU grid wavelength for WDM applications. Refer to Figure 1-16 for a clear idea on dropping of channels at intermediate nodes.

Figure 1-16. WDM System Demonstrating Dropping of a Wavelength at an Intermediate Node



In summary, the switch fabric does the task of adding, dropping, or switching optical channels or even just passing them through (pass through functionality). The channels are then fed to a multiplexer and recombined into a composite WDM signal. This signal is amplified by the pre-line fiber amplifier.

At node C, the signal is re-amplified by the post-line amplifier and demultiplexed into individual wavelengths or channels. These channels are detected by an array of photo detectors, and the electrical signal is fed to the client side of the network. Transponders usually perform this function of detecting and converting the network signal into client signal. Individual electrical streams could be further demultiplexed in the time domain to give slower rate streams.

A data stream from an ingress (source) node to an egress (destination) node on a wave-length is called a *lightpath*, defined by Chlamtac and otherset al[7]. A lightpath is an all optical connection or channel from source to destination on a given wavelength. A semi-lightpath is an optical channel that can be set up between source and destination nodes on more than one wavelength by converting the light from one wavelength to another and so on, in the course of traversing a source destination link. Assigning wavelengths to different lightpaths dynamically in a physical mesh or ring topology is more challenging than in a point-to-point topology. The reason wavelength assignment in optical networks is challenging is because, there are limited number of wavelengths to choose from, and a lightpath between a source destination pair traversing multiple fiber links (nodes) has to have the same wavelength in each of the links, as wavelength conversion is an expensive as well as infant technology. The other issue is to route these lightpaths in the most efficient manner. The simplest being shortest path routing, and other schemes having more complexity. This is a serious problem in designing optical WDM networks and is known popularly as the routing and wavelength assignment (RWA) problem. (Some references call this the RCA or routing and channel assignment problem. This is discussed in Chapter 6.)

Emerging Technologies: WDM Versus TDM, OCDM, and SCM

Capacity in a fiber can be increased by doing the following:

- Installing new fibers
- Time division multiplexing (TDM)
- WDM
- Code division multiplexing (CDM)
- Subcarrier channel multiplexing (SCM)

Adding new fiber is costly and time consuming; therefore, it is not a preferred choice.

The conventional approach to increasing bandwidth in a single fiber is to use TDM. In TDM, several signals are multiplexed in the time domain to transmit over a single (fast) channel. Interleaving pulses of different signals share a single channel, on a time-sharing basis. Consider three signals at the same bit rate of B bps. If we multiplex these three channels onto one single channel such that the bit rate of the new channel is 3B bps, we obtain a system gain of three. TDM allowed the legacy networks to scale efficiently in the past, and it is the foundation for WANs in the electrical domain, such as SONET/SDH. The main bottleneck for TDM systems is the limitation in obtaining economical fast electronic systems, which can multiplex several bit streams into a single TDM channel.

WDM refers to an optical multiplexing technique in which multiple optical signals that are each on a different characteristic wavelength are multiplexed together by making use of the spatial difference of their wavelengths. Consider four signals at wavelengths λ_1,λ_2,λ_3, and λ_4. If we multiplex these signals onto a single fiber such that the four signals coexist in the time domain, only sharing the spectral frequency (wavelength) domain, we obtain a composite signal that is called a WDM signal. WDM is nothing but multiplexing two or more signals on the basis of their wavelength independence of each channel with respect to the other channels. Unlike TDM, WDM adds capacity by adding wavelengths. We can add the wavelengths one at a time as needed. WDM thus works to be more scalable and flexible (on a desired demand basis) as compared to TDM.

CDM, also known as code division multiplexing, is another method of optical multiplexing. Unlike TDM and WDM, with CDM, each channel transmits its information (bits) as a coded sequence of pulses. This is achieved by transmitting a time-dependent short series of pulses placed within an allotted time period. Channels that have different codes can transmit on the same fiber. At present, it is not economical to generate and modulate ultra-short pulses efficiently. In the meantime, WDM is the technology of choice to alleviate capacity constraints in optical networks.

Optical SCM is another emerging technology that offers many advantages over traditional SONET multiplexing. SCM is a technique that is used to modulate a low frequency data signal to a higher carrier frequency; this, in turn, modulates an optical source that is generating an optical frequency. To obtain the bandwidth efficiency, several high-frequency sub-carriers are multiplexed together before modulating the optical frequency. SCM can be summarized as modulating individual base-band signals at RF frequency and then further modulating the sub-carriers at optical frequency. Subcarrier multiplexing of individual channels can further enhance the system bandwidth by WDM of many such SCM channels. SCM is bit rate independent, unlike

TDM. It can be an important building block for future enterprise and access networks. SCM increases capacity by modulation techniques, without increasing bandwidth spectrum or clock speed.

Summary

In this chapter, we acquainted the reader with optical networking and technologies that are associated with optical networking. Then we glanced through the various optical effects and phenomena associated with optics. From there, the reader is introduced to modal theory for optical pulse propagation in cylindrical waveguides (fiber). This chapter deals with the various losses associated with pulse propagation in an optical fiber, especially dispersion, Polarization mode dispersion, attenuation, four wave mixing, and nonlinearities. We also examined how a WDM network works and saw the discrete components associated with it. Finally, we explored the WDM scheme with its contemporary and predecessor multiplexing schemes: TDM, CDM, and SCM.

References

[1]Agrawal, Govind P. *Fiber-Optic Communication Systems*, Second Edition. Wiley Interscience, 1997.

[2]Keiser, G. *Optical Fiber Communications*, Third Edition. McGraw Hill, 2000.

[3]Stern and Bala. *Multiwavelength Optical Networks: A Layered Approach*. Addison Wesley, 1999.

[4]Kartalopoulos, Stamatios V. *Introduction to DWDM Technology: Data in a Rainbow*. IEEE-Wiley Press, 1999.

[5]Senior, John M. *Optical Fiber Communications*. Prentice Hall, 1992.

[6]Wartikar.*Applied Mathematics*. S. Chand Publishers, 1984.

[7]Chlamtac, I., A. Ganz, and G. Karmi. "Lightpath Communications: An Approach to High-Bandwidth Optical WANs." *IEEE Trans on Communication*. Vol. 40, No. 7, 1992.

[8]Ramaswami.*Optical Networks: A Practical Perspective*, First Edition. Morgan Kaufmann, 1999.

[9]ITU Standard GR 652.

[10]ITU Standard GR 653.

[11]ITU Standard GR 654.

[12]ITU Standard GR 655.

[13]Brackett, C.A. "Dense Wavelength Division Multiplexing Networks: Principles and Applications." *IEEE Journal on Selected Areas in Communications*, Vol. 8, No. 6, 1990.

[14]*IEEE Communications Magazine* Special Issue on Optical Networks. April 2002.

[15]Ramaswami and Sivarajan. "Routing and Wavelength Assignment." *IEEE/ACM Transaction on Networking*, April 1995.

[16]Edminister, Joseph. *Electromagnetics Schaum Series*, Second Edition. McGraw Hill, 1993.

Chapter 2. Networking with DWDM -1

For those readers who are well versed with the fundamentals of optical communication principles, this chapter sheds light on the underlined technologies and components involved in a wavelength division multiplexing (WDM) network. In Chapter 1, "Introduction to Optical Networking," we discussed point-to-point WDM links. Each node in the link is a WDM network element (NE). In this chapter we focus on the individual heuristics of a node or NE. A WDM node has features of adding, dropping, or passing through lightpaths. This kind of lightpath communication is facilitated by O-E-O (opto-electro-opto) or O-O-O (optical to optical) switch fabrics. At the ingress of each node, the composite WDM signal is preamplified by using optical amplifiers (explained in Chapter 3, "Networking with DWDM -2"). This amplified signal is then fully demultiplexed and switched locally. The switching operation is followed by a multiplex section, which regroups the individual wavelengths into a composite signal.

Numerous technologies have matured enough to produce quality optical components, which facilitate these optical operations at each DWDM network element.

At a network element, the various technologies are fused together to produce a response, which ensures a performance from the higher layers so as to interact with the optical layer.

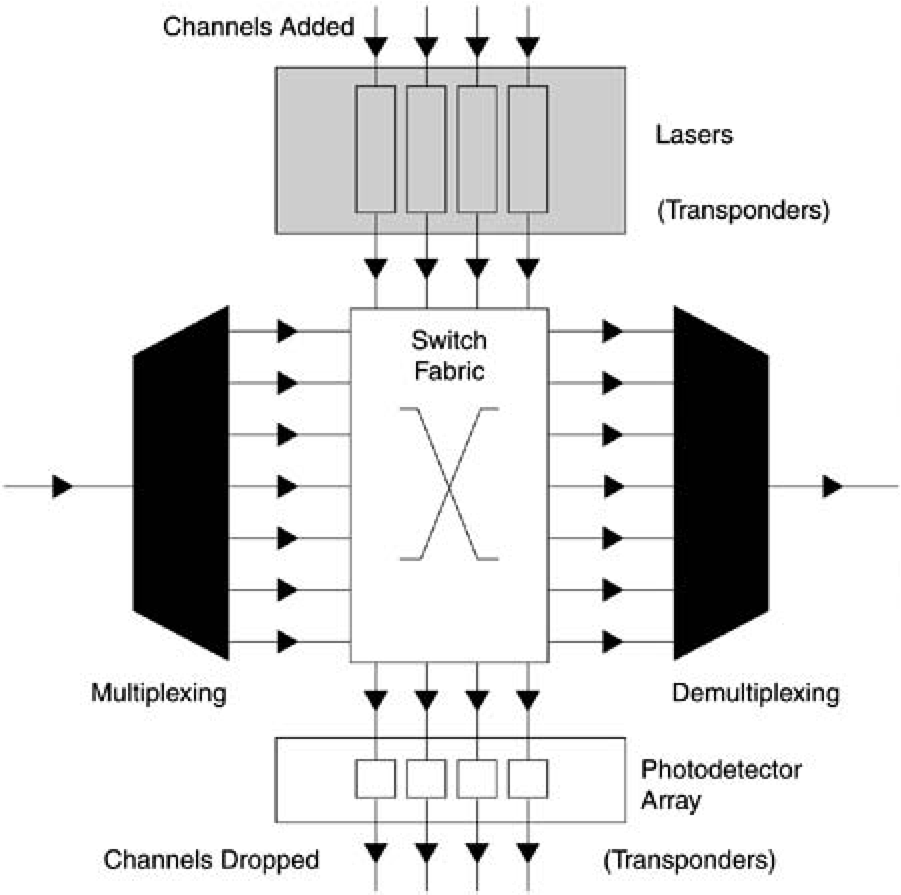
WDM technologies have been developed from classic optical physics. You do not need to know the analysis of the physical phenomena, but you do need to know the applications as well as the effects of these phenomena on a WDM network. A network designer has the liberty to choose from a variety of components and subsystems that broadly perform similar operations. For example, the designer could have a demultiplexer consisting of either arrayed waveguides (AWGs) or fiber Bragg gratings (FBGs). (Both are explained in later sections.) It will be at the designer's discretion to choose a more suitable component for that particular network application. The designer's choice depends on the performance, cost, and requirements of the components as well as the network.

This chapter and Chapter 3 focus on the panorama of technologies and components from a node analysis point of view rather than a system design point of view. This micro-level study helps the designer with the macro-level system design that is explained in Chapters 4, "WDM Network Design -1," 5, "WDM Network Design -2," 6, "Network Level Strategies in WDM Network Design: Routing and Wavelength Assignment," and 7, "X over DWDM." This chapter essentially discusses the details of a WDM node and its implementation perspectives.

A WDM node consists of a multiplex-demultiplex section, a switching section, and a local interface section. The local interface section consists of transponders, which are further broken down into optical sources, optical detectors, and complex electronic circuitry. The multiplex and demultiplex sections, as the name implies, consist of optical multiplexers and demultiplexers. The switching section typically has an array of O-E-O or O-O-O switches in add-drop configuration or cross-connect (any-port-to-any-port) configuration.

A WDM node, as shown in Figure 2-1, can consist of many different components and technologies aligned together to produce a classic WDM system with good performance. The next few sections discuss some technologies and their performances as WDM components. In a more systematic way, the sections address lasers, detectors, switches, passive devices, filters, and transponders as some of the key technologies. The following sections also discuss multiplexing and demultiplexing technologies, such as AWGs, FBGs, and so on. Some of the more active devices, such as amplifiers and dispersion compensators, are addressed in Chapter 3.

Figure 2-1. A Typical WDM Node



Optical Transmitters: Lasers

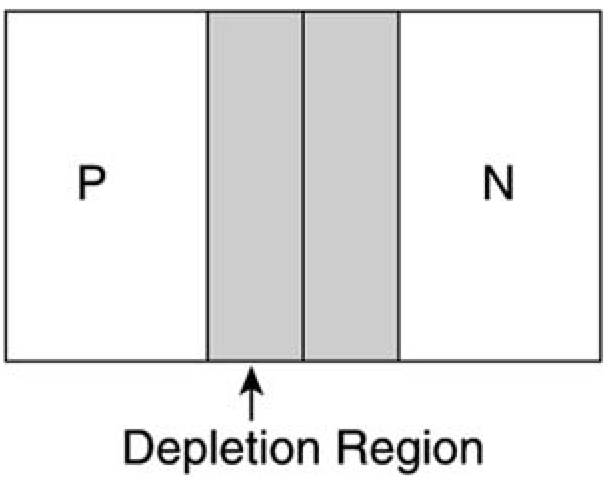
Lasers are used as optical sources for emitting modulated data into an optical fiber. Lasers have a distinct property whereby they can emit a narrow beam of light with a small optical spectra (line width), while having a high output optical power (concentrated beam of photons of approximately the same phase and frequency).

A laser is a semiconductor device (for optical purposes at least, although different forms of lasers do exist) that has an operation that is governed by the population inversion condition. This population inversion condition specifies the numerical superiority in volume of the electrons in the excited state (formed by absorption of energy by normal state electrons) over the electrons in the ground state in a semiconductor junction device.

A laser that is used in optical networking operations should have a narrow spectral line width, in addition to fast response (tunability) and be able to couple a significant amount of optical power into the fiber waveguide. Lasers that are used in optical communications are generally of two types: semiconductor lasers and fiber lasers. Semiconductor lasers are most commonly used in networking applications and are discussed here in detail. Fiber lasers are not so commonly used; therefore, they are not discussed in this chapter.

Semiconductor lasers are based on the optical properties of a p-n junction. Semiconductors as such have intermediate properties as compared to conductors or insulators. Silicon and Germanium have been traditionally used as semiconductor materials. Indium Phosphide (InP) and Gallium Arsenide have also recently discovered applications in lasers. A semiconductor material can be made p-type or n-type by doping the material (adding an impurity) with electrons (n-type) or extracting the material of some of its free electrons (p-type). An electron is a fundamental atomic particle of unit negative charge and negligible mass. Its addition creates an n-type material, whereas its removal creates a p-type substrate. Removal of electrons can also be considered as addition of holes that are positively charged particles in theory, but do not exist in reality. A semiconductor material that contains a region of p-type and a region of n-type with a shared boundary between the p- and n-types is called a p-n junction (see Figure 2-2).

Figure 2-2. P-N Junction Diode



Free electrons in a semiconductor can flow when a voltage difference is applied linearly across the semiconductor; this state is called the *conduction state*. In this state, the electrons are considered to exist in the conduction band. An energy level is associated with the conduction band (the energy of electrons in the conduction band is predetermined), and this energy level is given by Fermi-Dirac distribution equations. (These Fermi-Dirac equations are beyond the scope of this book.)

On passing current across a semiconductor material, free electrons absorb a quanta (an integral multiple of $h\nu$, where h is Planck's constant and ν is the frequency of absorbed radiation) of energy and jump into an excited state. After a period of time, these excited electrons, which have absorbed the excess energy and have risen to a higher excited state, drop back to the original state by emitting the excess absorbed energy in the form of photons at frequency ν.

This random oscillation of electrons from a lower energy level to a higher energy level and the subsequent emission of photonic radiation (light) is called *spontaneous emission*. The set of output optical frequencies is proportional to the energy between the stable and excited states and termed as *bandgap energy*.

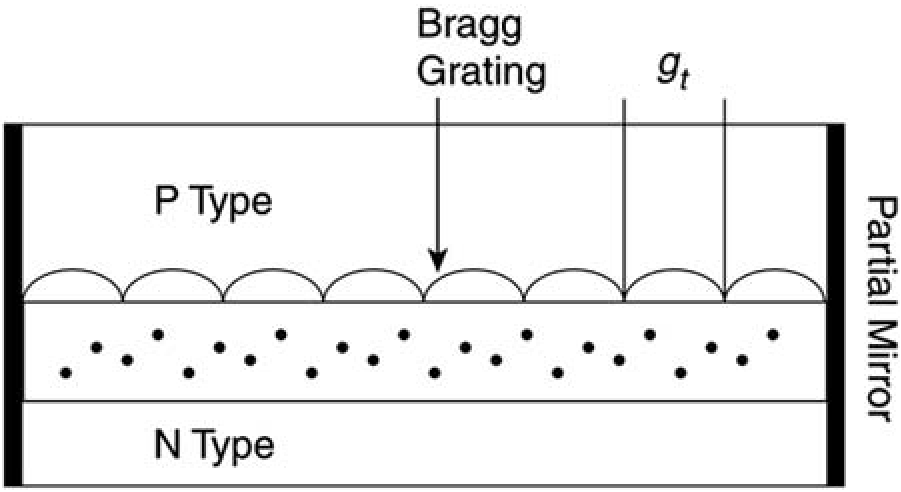
In spontaneous emission, no frequency or phase matching exists between consecutively emitted photons. In other words, every emitted photon has random phase and frequency distribution (*perturbation*). Spontaneous emission cannot sustain optical communication for the simple reason of low power and wide spectra of emission (line-width too large). The signal would be attenuated severely and would smudge into other adjacent channels. For lasers to function, the emission should be stimulated (externally controlled). Consider a case in which a few electrons have absorbed energy and risen to an excited state. Now assume that while these electrons are in the excited state, an external photon is bombarded onto these electrons. These electrons fall from the excited state to the ground state, emitting photons that have the same frequency (as well as phase) as the incident-bombarded photon. In other words, you would get a powerful beam of light at a controllable frequency, which is predetermined by the bombarded photon.

To sustain such a source for a long period of time, you must ensure that at any given time, there is an abundance of photons in the excited state. This kind of emission is called *stimulated emission* because of the external stimulus involved in the emissive process. To sustain an emission of this kind, it is necessary to establish *population inversion*. In other words, the number of electrons in the higher state (excited) should be greater than the number of electrons in the lower (stable) state. If this is not achieved, then the emission is spontaneous (random phase and frequency distribution).

One possible way of achieving population inversion is by having multiple energy levels. The cut-in point, at which stimulated emission is the dominant emission in the system, is called the *lasing threshold*.

As soon as population inversion is established, the system exhibits an optical gain because of the feedback achieved due to the bombardment of photons on the excited electrons creating a beam of high power light at a controllable frequency. This gain amplifies an optical signal exponentially. The optical gain is coupled by one more factor: the optical feedback. By placing the p-n junction inside a cavity that consists of reflecting walls, optical feedback can be achieved (refer to Figure 2-3). Optical gain initiates the stimulated emission into a gain profile that is analogous to the gain profile of an electrical amplifier, while optical feedback ensures the oscillatory function of the p-n junctions.

Figure 2-3. DFB Lasers



Distributed Feedback Lasers

By ensuring feedback, an oscillatory function is realized in the previously discussed laser.

Feedback is realized by placing the p-n junction in a cavity that has fully reflecting walls on all but one side and a partial reflector on the remaining side. By inserting a grating (corrugated surface) within the cavity, as shown in Figure 2-3, optical feedback is achieved. This optical feedback is called *distributed feedback* (DFB) due to its diverse occurrence in the cavity. (The feedback is throughout the length of the cavity.) The feedback is essential for maintaining the lasing threshold, and it is due to Bragg diffraction, which is explained in the section on FBG. Basically, when clusters of wavelengths hit a grating, only wavelengths that correspond to Bragg's condition are reflected back.[3] Bragg's condition is shown in Equation 2-1.

Equation 2-1

$$g_T = a\left(\frac{\lambda_B}{2n}\right)$$

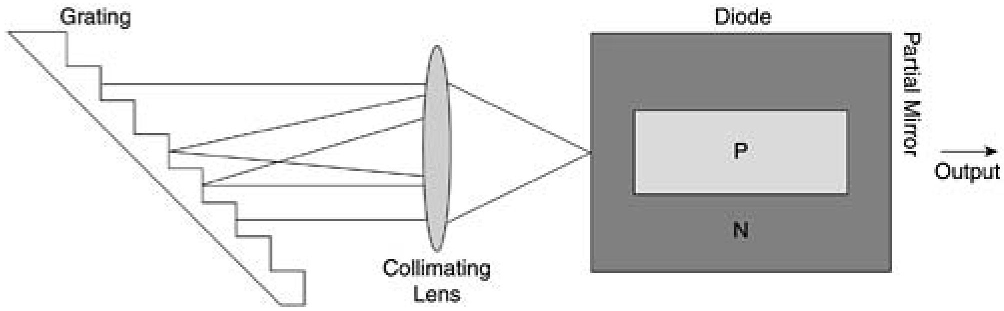
In the Equation 2-1, n equals the refractive index, g_T equals the period of grating, a equals the order of Bragg diffraction, and λ_B equals the Bragg's wavelength of our interest. These kinds of lasers are called DFB lasers (see Figure 2-3) and are of commercial interest in DWDM networks.

The feedback wave adds in phase to the emitted radiation, which is due to the electrons dropping from the excited state to the ground state. The grating is formed by methods of holography. The grating gives best performance for the Bragg's wavelength. The grating can be so formed so that the periodic perturbations can be varied, giving optimum performance for a number of different wavelengths and finding solid application in WDM networks.

Distributed Bragg Reflector (DBR) Lasers

Distributed Bragg reflector (DBR) lasers are a conceptual extension of DFB lasers. The principle is much the same; however, the feedback that is associated through the grating is now extended through the entire region of the cavity (refer to Figure 2-4).

Figure 2-4. DBR Laser



The corrugation or grating now extends to the mirrored walls, thus enhancing tunability. You can achieve wavelength tunability by varying the grating periods outside the gain medium (p-n junction). For a DBR laser, Bragg's condition given by Equation 2-2 needs to be satisfied.

Equation 2-2

$$\left(g_T = a\left(\frac{\lambda_B}{2n}\right)\right)$$

DBR lasers are good candidates for tunable optical sources. DBR as well as DFB lasers are temperature dependent; as a result, they need temperature-controlling elements for stable uniform operations. The temperature-controlling elements add a significant cost to the laser; therefore, thin line-width DBR lasers for C-band applications are quite expensive.

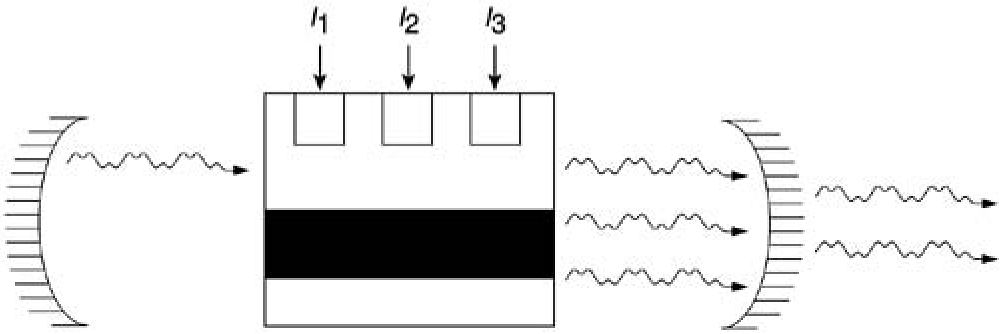
Tunable Lasers

For efficient nonblocking networks, you will probably want a particular optical source to be able to tune to different wavelengths with minimal tuning time. Sectional DBR lasers are one possible candidate for tunable optical sources. The feedback cavity is spread such that different sections have different currents that are needed to create lasing effects.

Lasing effect can be defined as the condition in a cavity or p-n junction when population inversion has been achieved and the junction/cavity can emit a streak of light at the desired wavelength.

The change of current over the grating also changes the Bragg's wavelength and the associated feedback. In this way, a DBR laser can be tuned across several nanometers relatively quickly. Mechanically tunable lasers that have a wider tuning range but slower tuning times have also been demonstrated. One such example is a Fabry Perot (FP) cavity laser (explained later in the section on cavities and filters), whereby the emitted wavelength is a function of the cavity length (FP cavity). Changing the length between the walls of the cavity can change the resonant frequency; this can be brought about mechanically (hence the term *mechanically tuned laser*). Recently, a tunable laser that could tune over the entire C band (approximately 35.2 nm) was demonstrated[4]. Refer to Figure 2-5 for generic diagram mechanically tunable laser.

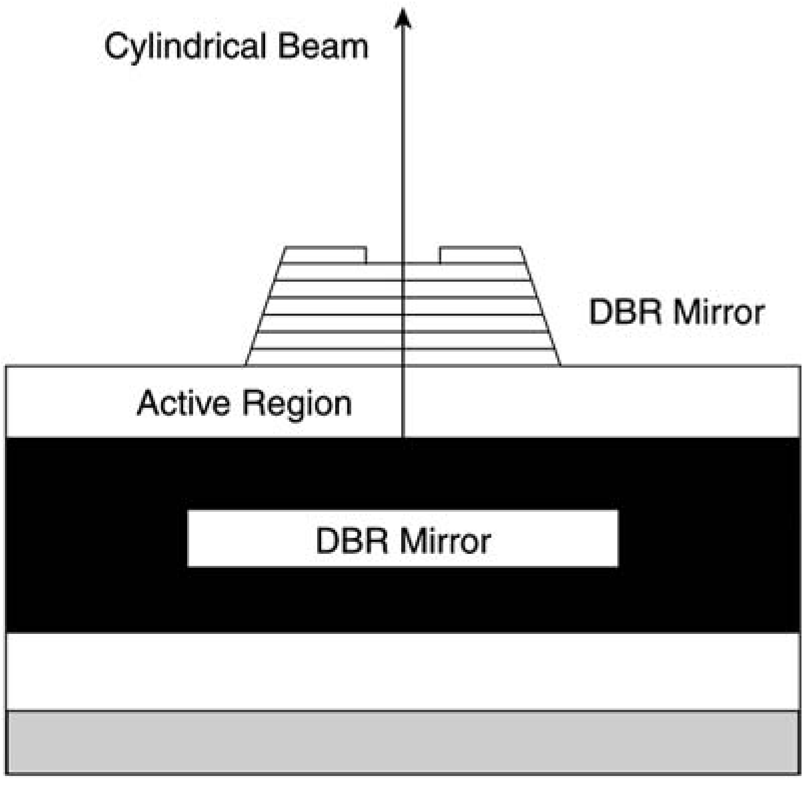
Figure 2-5. Mechanically Tunable Laser



VCSEL Laser

A *vertical cavity surface-emitting laser* (VCSEL), as shown in Figure 2-6, is a semiconductor laser diode that emits light perpendicular to the plane of p-n junction, unlike semiconductor lasers. VCSELs can be integrated with other components without prepackaging. For very short reach (VSR) applications at 850 nm–1310 nm wavelengths, VCSELs have become the preferred choice at even gigabit speeds.

Figure 2-6. VCSEL: Vertical Cavity Surface Emitting Laser



A VCSEL is composed of many specialized layers, analogous to edge-emitting lasers. The main parts of a VCSEL are the active region and the mirrors. A sandwich of active regions (between the mirrors) is created by stacking the subcomponents vertically on top of each other. For fiber-optic communication, VCSEL uses Indium Gallium Arsenide Phosphide (InGaAsP) for wavelengths of 850 and 1310 nm.

When a small current is applied across the device, light is emitted in the active region of the laser (using the same operation as that of semiconductor lasers). This light is reflected back and forth between the mirrors, while a fraction of the light "leaks" through the mirror to form the laser beam. These lasers are efficient and have low voltage requirements due to the high gain and small volume of the VCSEL structures. VCSELs are capable of fast direct modulation speeds (up to 2.5 Gbps), which means that no additional external components, such as modulators, are required. Currently, VCSELs are used in 850 nm wavelength applications as well as 1310 nm.

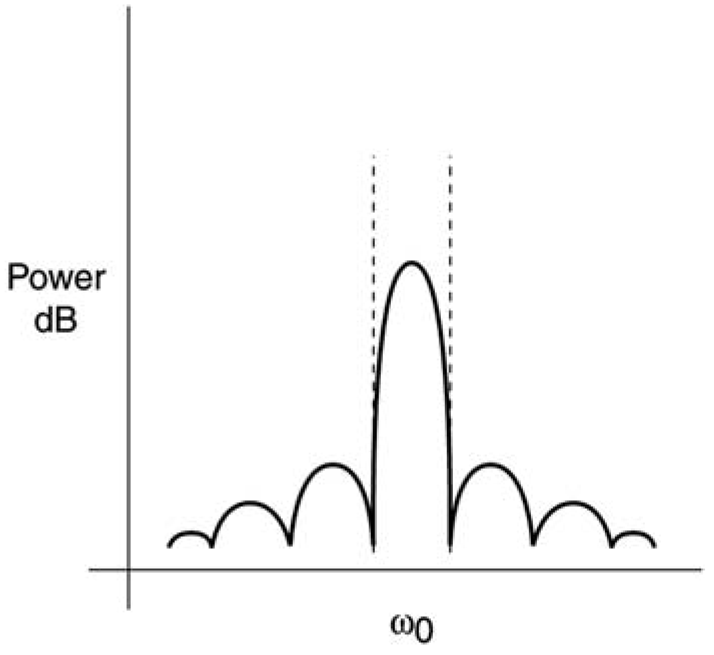
VCSELs offer several advantages over conventional lasers:

- High performance, lower cost, smaller size, and increased power
- The structure's capability to be integrated into a two–dimensional array
- Easier manufacturing and packaging
- Efficient fiber coupling

Line Width of a Laser

The line width of a laser should be as narrow as possible to prevent the data spectrum from overlapping the spectra of adjoining channels. Generally, a laser spectra has a dominant central frequency and multiple smaller sideband frequencies. Refer to Figure 2-7.

Figure 2-7. Line Width



The typical line width for WDM systems is in the range of a few megahertz. Laser line width is a major issue in WDM systems; the greater the line width, the greater the cross-talk with adjacent channels and the greater the nonlinear effects that are associated.

Chirp

When a voltage is applied to a laser, there is a sharp change in electron-hole density results in the active region due to lasing effect. This sudden change in density of electron-hole pairs influences the refractive index. Lasing effect also increases the temperature of the region, which in turn influences the length of the active region. These effects in the active region change the center frequency of the optical pulse that is transmitted. *Chirp* can be defined as the change in frequency of the transmitted optical signal with respect to time.

In semiconductor lasers, the frequency of a pulse shifts to a shorter frequency from the original frequency due to chirp. Nonlinear effects can also introduce chirp in optical communication systems. You can reduce chirping effect that results from lasing effect by using external modulators.

Modulation: Direct and External

Modulation can be defined as superimposing a data stream onto a carrier signal by altering one of the virtues of the carrier signal with respect to a change in the data stream. In other words, you can make a binary data stream superimpose on a carrier frequency. The motive behind modulation is to enable transport of data efficiently and without many errors. In an optical WDM network, data is modulated onto the light that a laser diode emits. One way of modulation is to make the output optical power of a laser diode proportional to the binary sequence of the data stream. You can use two techniques for modulation by using optical lasers: direct modulation, as shown in Figure 2-8, and external modulation, as shown in Figure 2-9.

Figure 2-8. Direct Modulation Technique

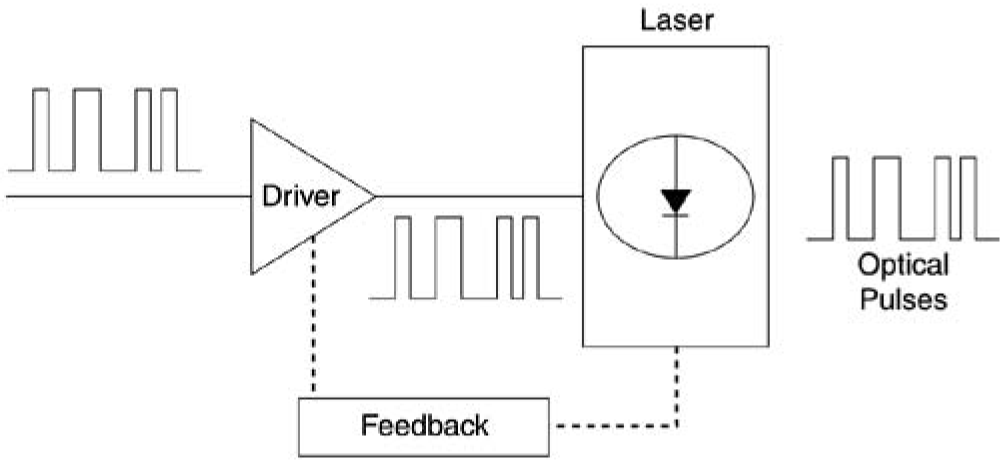
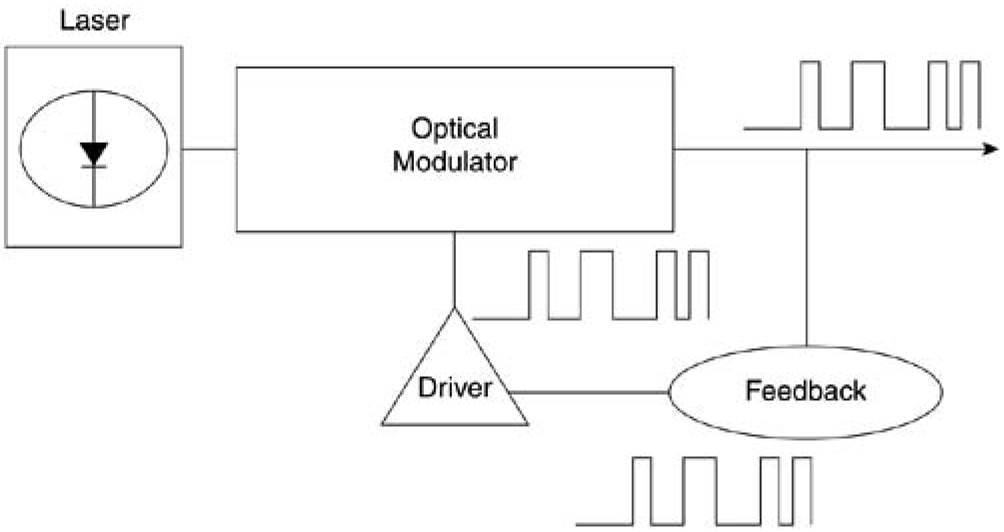


Figure 2-9. External Modulation Technique



In direct modulation, the laser drive current that is needed to cause stimulated emission is varied with the data stream. This causes the output optical power to fluctuate as per the data stream. In other words, a stream of binary data, when made proportional to the optical power, creates a series of isomorphous optical pulses. This modulation technique is termed *direct modulation* because the data is directly coupled with the laser drive current. Direct modulation has severe drawbacks at high data rates. It cannot be used at bit rates that are greater than 2.5 Gbps. Direct modulation—creates non-linearity's especially self phase modulation (SPM) explained later in Chapter 4. Direct modulation also increases the laser chirp. Typically, a binary data stream is made to modulate a laser diode; therefore, the optical power fluctuates between high and low. Due to the return-to-zero type of modulation format, the laser diode switches between ON and OFF for a logical 1 and logical 0 respectively. Turning the laser ON and OFF introduces time dependence. As a result, the bit rate that is transmitted using direct modulation has a maximum limit. Direct modulated lasers are limited more by distance than by bandwidth drawbacks. For short ranges, they are cost effective and useful, especially for metro-optical operations.

External modulation: When a laser source is not directly modulated to feed the data stream, but the output optical frequency is modulated in a separate section by other means (see Figure 2-9), this kind of optical modulation technique is called external modulation.

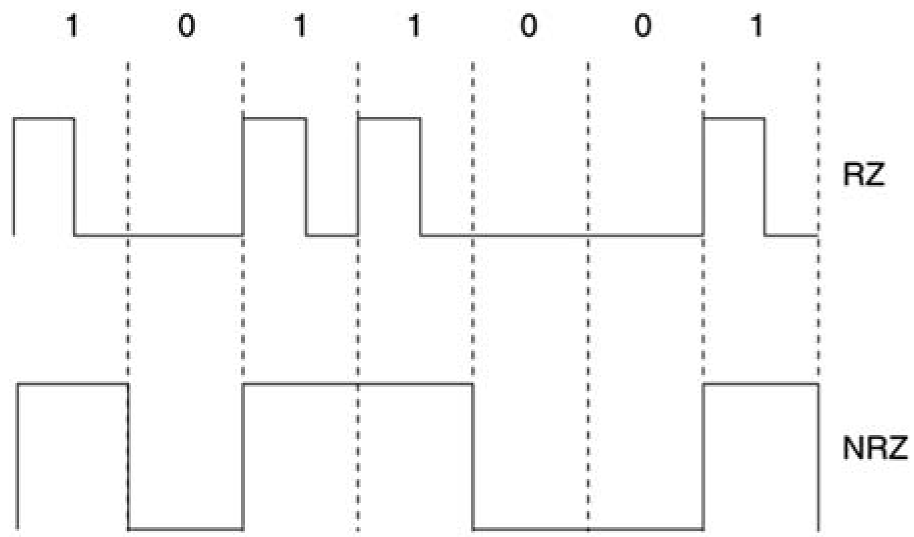
In external modulation, the laser output power is generally modulated in an external cavity. External modulation avoids nonlinearity's and excessive chirp. Fabry Perot cavities or Lithium Niobate-based Mach-Zehneder Interferometers (MZIs) are good candidates for such applications. One possible approach is to vary the voltage across the MZI (explained in the section on filters) arm to change the coupling ratio (between the two arms of a MZI), which produces optical pulses at one of the arms. These pulses are proportional to the variations in data stream (0s and 1s). Most commercial systems contain the laser diode and the modulator as a single unit.

Modulation Formats in the Optical Domain

Typically, you would use ON/OFF keying (*OOK*) of the laser source to communicate data through an optical domain. The data stream looks like a train of pulses, with a high for a logical data 1 and a gap or low for logical data 0. The problem with this technique is the time constraint required to switch a laser between the on state and the off state.

Two forms of OOK exist: return to zero (RZ) and non-return to zero (NRZ). In RZ format for the logical 1 bit, the power level returns to 0 after half the period (pulse slot), whereas for the 0 bit, the power level is 0 continuously. In NRZ, the 1 bit has a signal that is in the high power level throughout the 1-bit period and a 0 power level throughout the 0-bit period (refer to Figure 2-10). The advantage of NRZ over RZ is that NRZ occupies only half the bandwidth of RZ. However, long periods of 1s create a block of high power, making it difficult for the receiver to decode the signal as well as create nonlinear effects. NRZ is the preferred waveform format in optical WDM networks. Studies of intensity-modulated formats with phase shifting techniques are currently being proposed.

Figure 2-10. Modulation Format



[Table 2-1](#) discusses the various forms of lasers that are used for WDM networking as well as the performance of the laser types.

Table 2-1. Laser Types

Laser Type	Tuning Range	Tuning Time
Mechanical laser	10–20 nm	100–500 ms
Acousto-optical or electro-optical laser	10–20 nm	Several tens of microseconds
Injection current tuned laser	4 nm	5–10 ns
DFB arrayed laser	Limited by number of elements in array	1–10 ns

Optical Receivers: Photodetecters

The function of an optical receiver is to decode and interpret the optical signals and generate an electrical data stream proportional to the received optical signal. The main component of an optical receiver is a photodetector, which converts the optical power into electrical current. Photodetectors need to meet stringent requirements to achieve desirable performance. Requirements include good responsivity (sensitivity) to a wide range of wavelengths used for transmission (usually in the 850 nm, 1300 nm, or 1550 nm region), low noise characteristics, low or zero sensitivity to temperature variations, low cost, and extended operating life. Even though several types of photodetectors are available, semiconductor-based photodetectors (photodiodes) are used exclusively for optical communications. The most common photodiodes used in optical systems are PIN photode tectors and Avalanche Photo Detectors (APDs), due to their small size, fast response, high photo-sensitivity and comparably low costs.

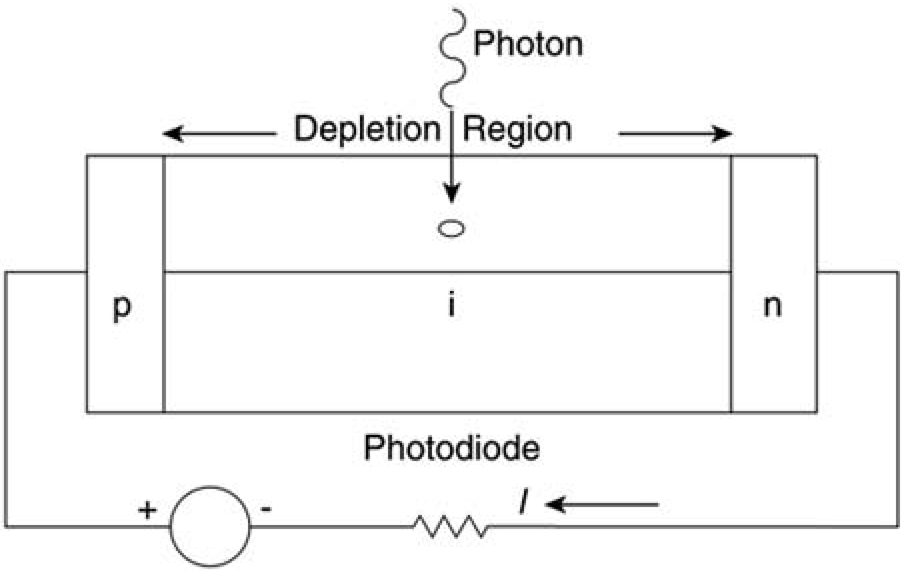
The PIN Photo Detector

The PIN diode is an extension of the P-N junction diode, in which slightly doped intrinsic material (I stands for intrinsic) is inserted in between the P-N junction, thereby increasing the depletion width (region) of the P-N junction.

NOTE

The depletion region is the region in the p-n junction that is formed by some of the electrons from the n type moving over and depleting the holes in the p type, thereby creating a region of neutral charge, upon condition of reverse bias. Refer to Figure 2-11.

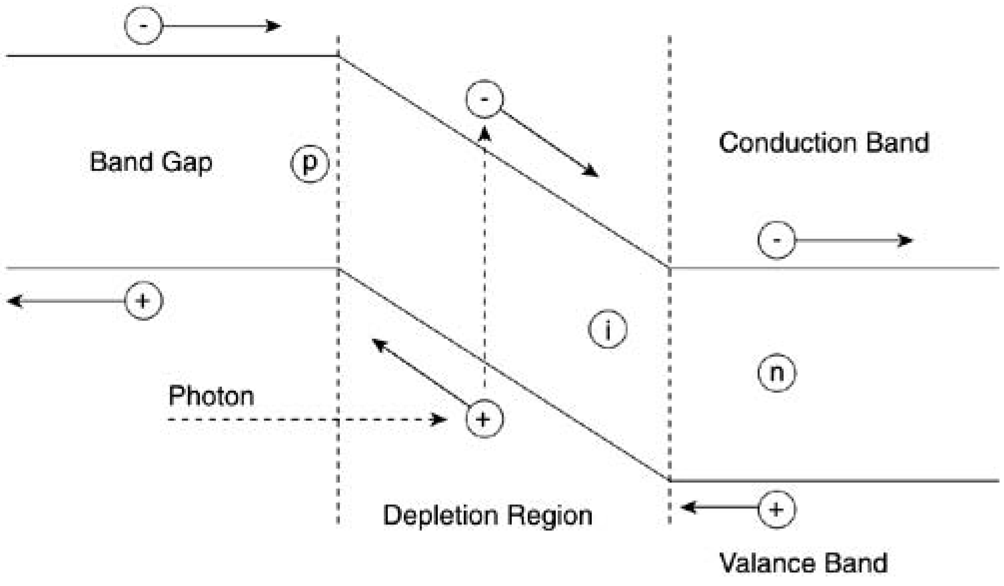
Figure 2-11. PIN Photo-Diode



A high reverse-biased voltage is applied across the PIN diode so that the intrinsic region is completely depleted. Figure 2-11 represents the normal operation of a PIN diode with reverse bias applied across the p-i-n junction. When light (photons) is incident on a semiconductor material, electrons in the valence band absorb it. As a result of this absorption, the photons transfer their energy and excite electrons from the valance band to the conduction band, leaving holes in the valance band.

The design of the PIN photodiode is optimized in such a way that electron hole pairs are generated mostly in the depletion region (see Figure 2-12). After the application of voltage across the depletion region, the formed electron hole pairs induce an electric current flow (also known as *photocurrent*) in an external circuit. Each electron hole pair generates one electron flow[4].

Figure 2-12. Energy Band Diagram for Photodetector



To generate the photocurrent, we must ensure that the energy of the incident photon is equal to or greater than the bandgap energy.

The energy of the incident photon is shown in Equation 2-3.

Equation 2-3

$$hf_c = \frac{hc}{\lambda} \geq eE_g$$

In the equation, λ is the wavelength, E_g is energy of the bandgap, c is the velocity of light, and e is the charge of the electron. λ is the wavelength at which the semiconductor material will function as a photodetector. There is an upper limit of wavelength $\lambda_{cuttoff}$ above which any particular semiconductor material does not generate photocurrent. The cutoff wavelength is about 1.06 µm for silicon. See Figure 2-12.

The analysis of photocurrent that is generated is beyond the scope of this book. For more information, refer to references 2 and 4.

The photocurrent I_p that results from power absorption of photons is shown in Equation 2-4, which provides the equation for photocurrent.

Equation 2-4

$$I_p = \frac{eP_0(1 - EXP(-\alpha_s w))(1 - R_f)}{hv}$$

In the equation, P_o is the optical power, e is the charge of the electrons, h is the Planck constant, ν is the photon frequency, α_s is the absorption coefficient, and R_f is the reflectivity of silicon.

Two important characteristics of photodetectors are quantum efficiency and responsivity.

Quantum efficiency η is the number of electron hole pairs generated per incident photon.

Responsivity is the amount of current produced at a particular input optical power and a measure of the rate of change of electrical current (generated) as per the rate of change of optical power.

Responsivity of a photodetector is expressed in terms of λ and is shown in Equation 2-5.

Equation 2-5

$$R = \frac{e\eta\lambda}{hc} = \frac{\eta\lambda}{1.24} Ampers/Watts$$

Typical PIN photodiode responsivity values are 0.65 A/W for silicon at 850 nm and .45 A/W for Germanium at 1300 nm.

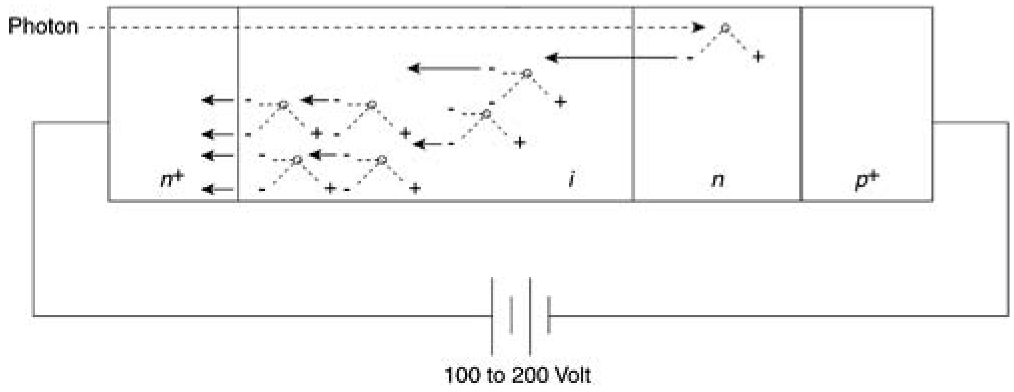
Avalanche Photodiodes

When light is absorbed by a PIN photodetector, only a single electron hole pair is generated per photon. You can increase the sensitivity of the detectors if more electrons are generated, which means that you need less power for photodetection and that the signal can travel longer.

If a high electric field is applied to the generated electrons, enough energy is procured to excite more electrons from the valence band to the conduction band. This, in turn, results in more electron hole pairs being generated. These secondary electron hole pairs that are generated by the preceding process can produce more electron hole pairs if they are subjected to a high electric field (Avalanche effect).

This process of multiplication of electron-hole pairs is called *Avalanche multiplication*, which is demonstrated in Figure 2-13. The photodiode that is designed to achieve this kind of electron-hole pair multiplication is known as *Avalanche photodiode* (APD).

Figure 2-13. Avalanche Multiplication Process



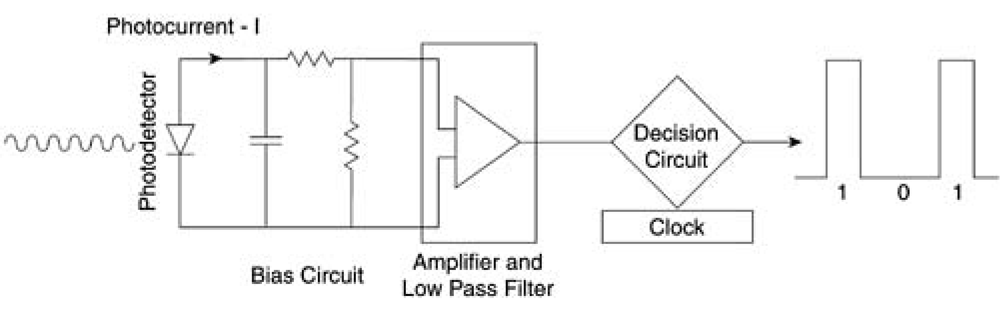
In practice, the avalanche effect is a statistical phenomenon. In other words, electron hole pairs generated by the primary electrons are randomly distributed. The statistical value is termed as *multiplicative factor*, or multiplicative gain (M_f), and is shown in Equation 2-6.

Equation 2-6 Multiplicative Gain

$$M_f = \frac{I_a}{I_p}$$

In the equation, I_a is the average value of the total current (including the current that is generated due to the avalanche effect), and I_p is the value of current due to initial electrons. An APD can be designed with the multiplicative factor equal to infinity, a condition called *avalanche breakdown*. However, large values of M produce unpredictable effects in the generation of photocurrent, which in turn affects the noise performance of APDs. A schematic diagram of an optical receiver is shown in Figure 2-14.

Figure 2-14. Functional Diagram of an Optical Receiver



An optical receiver consists of a photodetector followed by a preamplifier. The function of a preamplifier is to amplify the photocurrent for further processing. The next stage consists of a high-gain amplifier and a low-pass filter. An amplifier gain control circuit automatically limits the amplified output to a fixed level, regardless of the optical power incident on the photodetector.

The low-pass filter reduces the noise level and shapes the pulses. The low-pass filter is designed in such a way that the intersymbol interference (ISI) is minimized. Receiver noise is proportional to receiver bandwidth, and loss-pass filters can reduce noise by having the bandwidth (BW) be lower than the bit rate (B). The electric pulse spreads beyond the bit slot for BW < B and results in ISI, which interferes with proper detection of nearby bits. The final stage of an optical receiver consists of a decision circuit and a clock recovery circuit. The decision circuit compares the output to a threshold level at sampling times that the clocking circuit defines and then decides whether the input signal pulse is a 1 bit or a 0 bit. Due to the noise associated with receivers, it is probable that the decision circuits will detect bits incorrectly. Receiver noise, sensitivity, and performance are explained in the next section. The receivers are usually designed in such a way that the error probability of detecting a 1 for a 0 and a 0 for a 1 is quite small (10^{-9}–10^{-12} in commercial optical links).

Receiver Noise

Noise is a serious problem in detection of optical signals at the receiver. This electrical noise due to current fluctuations affects the receiver performance. There are two major contributions to noise: Shot noise and Thermal noise.

Shot Noise

In simplistic terms, the interarrival rate between electrons flowing is a random phenomenon, which contributes to immense fluctuations in an electrical circuit; this is termed as *Shot noise*. For a photodiode of responsivity R, the current induced is shown in Equation 2-7.

Equation 2-7

$$\text{Current induced } I(t) = RP_{input} + I_s(t)$$

In other words, refer to Equation 2-8.

Equation 2-8

$$I(t) = \langle I_{pd} \rangle + I_s(t)$$

RP_{input} equals the average value of I_{pd}, and $<I_{pd}>$ signifies the average value of the photodetector current.

Shot noise can be statistically depicted as a Gaussian function (mean = 0, variance = 1).

Thermal Noise

The self-random motion of electrons due to the possession of kinetic energy on virtue of the temperature gives rise to uneven fluctuations, or thermal noise. This noise does not need voltage to sustain itself.

Thermal noise, like shot noise, can be approximated as a Gaussian process. The final equation for induced current is provided in Equation 2-9.

Equation 2-9

$$I(t) = \langle I_{pd} \rangle + I_s(t) + I_{thermal}(t)$$

The noise figure (NF) is a figure of merit that is associated with a device. In the receiver, the photodetector is followed by the front end amplifier, and the noise figure gives the amplification ratio of input noise to output noise across the amplifier. See Equation 2-10.

Equation 2-10

$$\sigma^2_{thermal} = \frac{4k_BT}{R}NFB$$

In the equation, B equals the bandwidth of the receiver, k_B equals the Boltzmann constant, R equals the responsivity, T equals the absolute temperature, and NF equals the noise figure.

Receiver Performance

Receiver performance is an important factor in optical system design. The optical system design performance depends on the performance of the receiver in its ability to detect 1s and 0s from an incoming optical signal. Bit error rate (BER) is a figure of merit to measure receiver performance. Receiver sensitivity is another performance-measuring standard for optical detectors that is important for optical system design. Finally, signal-to-noise ratio (SNR) can be regarded as the absolute qualitative measure of the signal at the receiver. It is discussed in subsequent sections.

BER: Bit-Error Rate

During the transmission of data through an optical channel, the receiver should be able to receive individual bits without errors. Errors occur when a receiver fails to detect an incoming bit correctly. Causes for errors generally stem from impairments that are associated with the transmission channel. A receiver fails to detect a bit correctly when it detects a 1 bit for a 0 bit that is transmitted or a 0 bit when a 1 bit is transmitted. The receiver is also bit-rate sensitive. For different bit rates, a receiver has different magnitudes of errors; therefore, BER is a figure of quality in an optical network. Typically, optical end systems should have a BER of 10^{-9} to 10^{-12}; in other words, for every 10^9 bits transmitted, one corrupted bit is allowed.

Mathematically, BER is the sum of probabilities, such that when a 0 bit is transmitted, a 1 bit is received; when a 1 bit is transmitted, a 0 bit is received. This summation of these conditional probabilities gives the BER of the system statistically. This probability is shown in Equation 2-11.

Equation 2-11

$$BER = P(1)P(0 \text{ received for 1 transmitted}) + P(0)P(1 \text{ received for 0 transmitted})$$

This can be expressed as shown in Equations 2-12,2-13, and 2-14.

Equation 2-12

$$BER = P(1)P(0/1) + P(0)P(1/0)$$

Equation 2-13

$$P(0) = \text{probability of a zero bit transmitted} = 1/2$$

Equation 2-14

$$P(1) = \text{probability of a 1 bit transmitted} = 1/2$$

P(0) = P(1) = 1/2. That is because a 1 or a 0 is equally likely to be transmitted, hence probability is half or 0.5. Refer to Equations 2-13 and 2-14.

P(0/1) and P(1/0) depend on the distribution of the current over time while detecting the signal. In other words, the probability density of the noise associated with the system affects the final waveform of the current. That is, if you consider noise as being superimposed on the signal, this superimposed waveform is what determines how many wrong decisions were made at the receiver. Consider Figures 2-15,2-16, and 2-17.

Figure 2-15. Original Signal

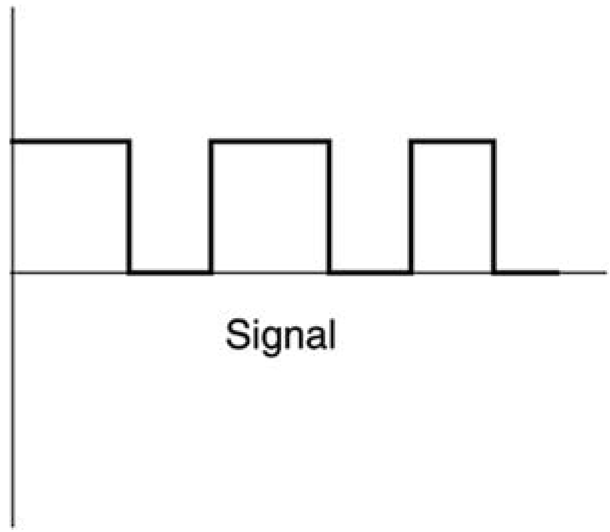


Figure 2-16. Noise Distribution

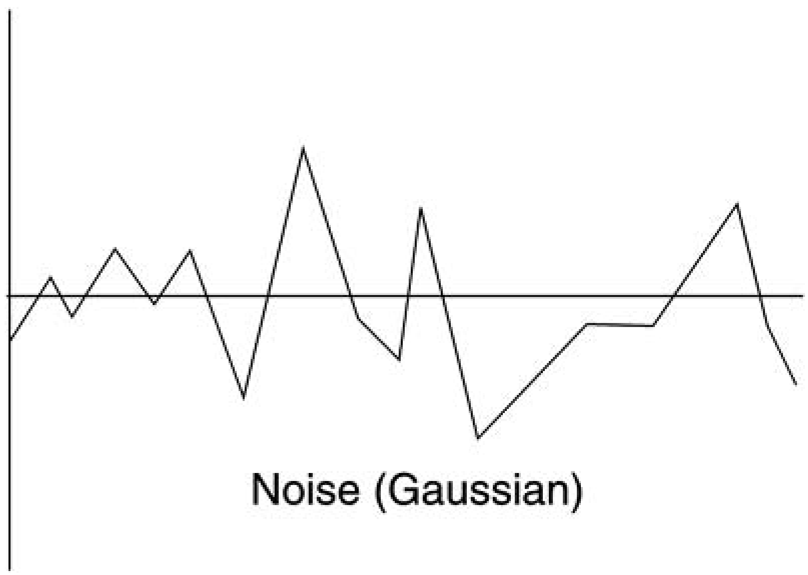
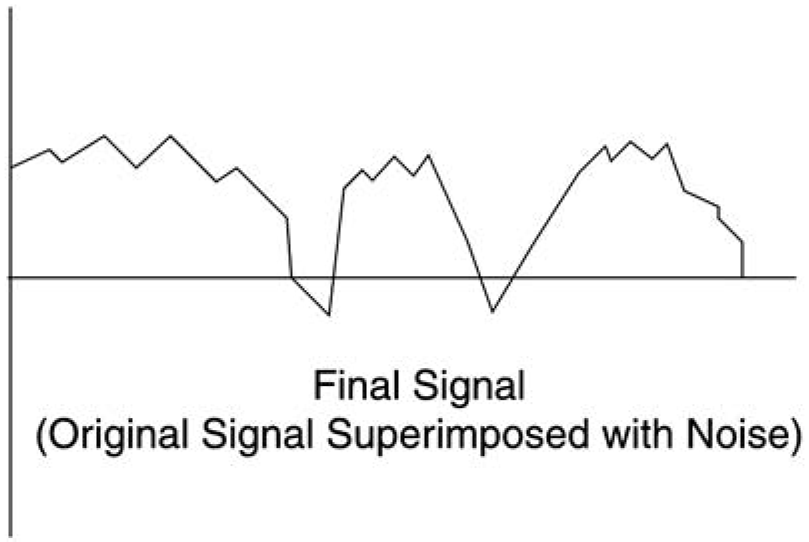


Figure 2-17. Final Signal at the Output of the Channel (Summation of the Signal in Figures 2-15a and 2-15b)



The noise spectra can be given analytically as a summation of the *probability density function* (PDF), which is defined as the first order derivative of the distribution function F(x), shown in Equation 2-15.

Equation 2-15

$$F(x) = P(x < X)$$

In the equation, x is a random variable; therefore, its probability density function is shown in Equation 2-16.

Equation 2-16

$$f(x) = \frac{d}{dx}F(x)$$

Further, the noise is classified into Shot and thermal noise (see the section on receiver noise). Both Shot noise and thermal noise can be approximated as Gaussian density functions.

A Gaussian distribution is defined as shown in Equation 2-17.

Equation 2-17

$$f(x) = \frac{1^{e^{-\frac{(x-m)^2}{2\sigma^2}}}}{\sqrt{2\pi\sigma^2}}$$

In the equation, m is the mean of f(x) and σ^2 is the variance of f(x).

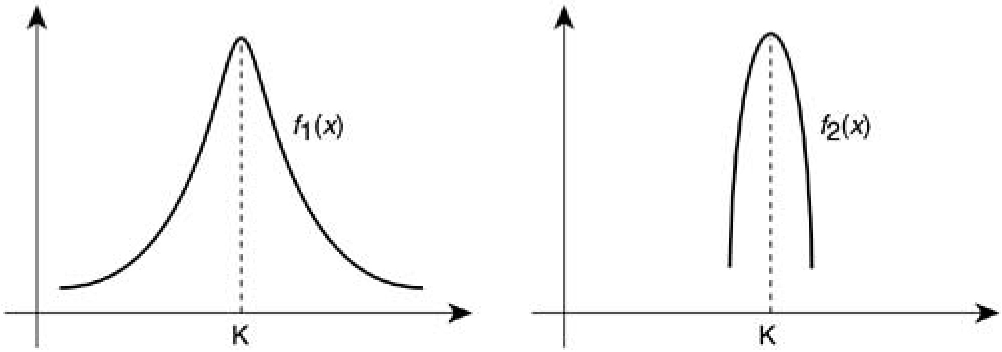
Mean

The mean of a distribution is defined as the average value that the distribution takes.

Variance

Consider[Figure 2-18], which shows two distributions. One distribution has a large variance, and the other has a small variance, but the means are the same. The first one is a broad distribution in which $f_1(x)$ assumes values ranging from quite close to quite far from the line x = k. The second distribution is narrow, and $f_2(x)$ assumes values very close to the line x = k.

Figure 2-18. Distributions of Large and Small Variances



Consider function $f_1(x)$ and $f_2(x)$ such that $f_1(x)$ and $f_2(x)$ are evenly distributed about the line x = k. If you take n random values of $f_1(x)$ and $f_2(x)$ respectively, you will find that on an average, values of $f_2(x)$ are closer to the line x = k than values of $f_1(x)$. The width of $f_1(x)$ is greater than the width of $f_2(x)$, although both have the same mean; k. To describe the width or variance of $f_x(x)$ you need to take the square of each average value about the mean, also termed *second moment* (σ^2). Refer to Equations 2-18 and 2-19.

Equation 2-18

$$\sigma^2 = E(X-K)^2$$

Equation 2-19

$$= E(X^2) - 2E(X)K + K^2$$

Because k equals E(X) (k equals the mean), see Equation 2-20.

Equation 2-20

$$E(X^2) - [E(X)]^2$$

Therefore, a Gaussian distribution has a zero mean and variance provided as in Equation 2-21.

Equation 2-21

$$\sigma^2 = \int_{-\alpha}^{\alpha} \frac{(x-m)^2 e^{-(x-m)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} dx$$

Coming back to derive an equation for BER, the amount of thermal and Shot noise is provided in Equation 2-22.

Equation 2-22

$$\sigma_s^2 = 2q(I_p + I_d)\Delta f$$

In the equation, I_p =R P_{in}, R is the responsivity, I_d is the dark current, and Δf is the bandwidth of the receiver.

I_d is the amount of current flow in absence of an incident light onto a detection circuit of an optical receiver. In other words, it is the dark current.

Equation 2-23 provides the variance of thermal noise.

Equation 2-23

$$\sigma_T^2 = (4k_B T/R_1)F_n \Delta f$$

The distribution of noise is another Gaussian variable, with variance provided as in Equation 2-24.

Equation 2-24

$$\sigma^2 = \sigma_s^2 + \sigma_T^2$$

This variance is a different number for the 1 (high) bit and for the 0 bit(low) given by Equation 2-25 and 2-26.

Equation 2-25

$$P\ (0\ \text{is received for 1 transmitted}) = \frac{1}{\sigma_1\sqrt{2\pi}}\int_{-\alpha}^{I_D}\exp-\frac{(I-I_1)}{2\sigma^2}dI$$

And (see Equation 2-26)

Equation 2-26

$$P\ (1\ \text{is received for 0 transmitted}) = \frac{1}{\sigma_0\sqrt{2\pi}}\int_{I_D}^{\alpha}\exp\frac{(-I-I_0)}{2\sigma_0^{\ 2}}dI$$

Both the preceding equations can be further solved in terms of the error function (erf) and its complimentary error functions (erfc).

Error Function and Its Relationship to BER

The distribution function of a Gaussian distribution is shown in Equation 2-27.

Equation 2-27

$$f(x) = \int_{-\alpha}^{x}\frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}dx$$

This integral can be mapped to error function (erf), which is defined as in Equation 2-28.

Equation 2-28

$$\text{erf}\ K \equiv \frac{2}{\sqrt{\pi}}\int_0^k e^{-k^2}du$$

$$\text{erf}(0) = 0 \quad \text{erf}(\infty) = 1$$

Conversely, complimentary error function (erfc) is given by the compliment of erf, as shown in Equation 2-29.

Equation 2-29

$$Erfc = 1-\text{erf}$$

Therefore, for Gaussian distribution, consider Equation 2-30.

Equation 2-30

$$Fx = \int_{-\alpha}^{x} \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} dx$$

Changing the limits of the integral, as shown in Equation 2-31.

Equation 2-31

$$\int_{-\alpha}^{x} = \int_{-\alpha}^{\alpha} - \int_{x}^{\alpha} \text{ because } \int_{a}^{b} = \int_{a}^{c} - \int_{b}^{c} \quad <> \mathrm{a, b}$$

Applying to F(x), as shown in Equation 2-32.

Equation 2-32

$$F(x) = \int_{-\alpha}^{\alpha} \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} dx - \int_{x}^{\alpha} \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} dx$$

The first term of right hand side (RHS) is 1; the second term on RHS is given by Equation 2-33.

$$\frac{1}{2} erfc\left(\frac{x}{\sqrt{2}\sigma}\right)$$

Therefore, consider Equation 2-33.

Equation 2-33

$$F(x) = 1 - \frac{1}{2} erfc \frac{x}{\sqrt{2}\sigma}$$

Substituting in Equations 2-11 and 2-12, we get the result shown in Equation 2-34.

Equation 2-34

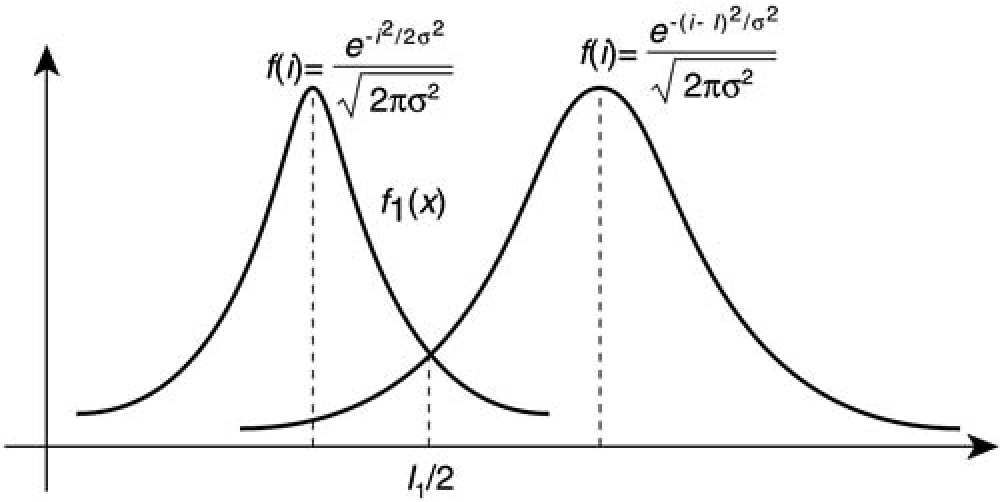
$$BER = \frac{1}{4}\left[erfc\left(\frac{I_1 - I_D}{\sigma_1\sqrt{2}}\right) + erfc\left(\frac{I_d - I_0}{\sigma_0\sqrt{2}}\right)\right]$$

The relation of erfc to practical calculations is that through erfc, we are able to make an estimation of the probability of success (or failure) for a particular random distribution.

In a practical optical system, consider the situation in which we transmit $+I_1$ current for a 1 bit and 0 current for a 0 bit (on/off keying). Now assume that pulses are transmitted over a fiber of finite length. Also assume noise to be present that follows a Gaussian distribution (mean = 0). At the output, the level of instantaneous current $i(t) = I + N$, where $I = 0$ for 0 bit, $I = +I_1$ for 1 bit, and noise is N, whose variance is already shown. Now $i(t)$ is a Gaussian random variable with mean I and variance the same as that of noise N.

Consider the two distributions in Figure 2-19. Due to the symmetry of the two distributions, it is quite natural that the threshold level for separating a logical 1 from a logical 0 is $I_1/2$. In other words, if the current level is $< I_1/2$, the transmitted bit is a 0; otherwise, it is a 1. However, if I_1 is transmitted and the detected level is less than $I_1/2$ (that is, $I_1 + N < I_1/2$), an error is obtained. Similarly, when a 0 current is transmitted and the detected level is greater than $I_1/2$, an error is obtained. Due to the mapping of error probability to this function, it is named as the *error function*.

Figure 2-19. Probability Density Function (PDF) for a 0 and a 1 Bit



The deviation threshold can be further optimized to reduce BER. This is obtained when condition as in Equation 2-35[2] is met.

Equation 2-35

$$(I_1 - I_d)\sigma_1 = (I_D - I_0)/\sigma_0$$

Upon simplifying, we get Equation 2-36.

Equation 2-36

$$I_D = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_0 + \sigma_1}$$

A quantity of interest is the Q factor of the signal, which is provided in Equation 2-37. Q factor gives the absolute quality of an optical signal.

Equation 2-37

$$Q = \frac{I_1 - I_0}{\sigma_1 - \sigma_0} I_1 = 1 \text{ bit current}, \mathrm{I}_0 = 0 \text{ bit current}$$

In the equation, σ_1 is the standard deviation of 1, and σ_0 is the standard deviation of 0.

BER is related to the Q factor, as shown in Equation 2-38.

Equation 2-38

$$BER = \frac{1}{2} erfc\left(\frac{Q}{\sqrt{2}}\right)$$

The higher the Q factor of the signal, the lower the BER will be. (BER is improved with the Q factor.) The Q factor is covered in detail in Chapter 4 for system design and is important because intricate mathematics are involved in determining BER directly. The Q factor represents a simple yet accurate figure of system performance, especially for receiver design considerations and OSNR-based designs.

Receiver Sensitivity

The BER and Q factor are related to the receiver minimum power requirement. For a receiver with a 1-bit current of I_1 and a 0-bit current of $I_0 = 0$ (zero current for logic 0 for practical assumptions), and σ_1,σ_0 are the standard deviations of I_1, I_0, then consider Equation 2-39.

Equation 2-39

$$Q = \frac{I_1}{\sigma_1 + \sigma_0}$$

also if $P_1 = 1$ bit power and $P_0 = 0$ bit power. The Q factor is now provided as in Equation 2-40.

Equation 2-40

$$Q = \frac{2R(P_1 + P_0)}{(\sqrt{\sigma_s^2 + \sigma_T^2}) + \sigma_T}$$

Further, BER and Q are related as per Equation 2-38. Therefore, the received power is proportional to Q as well as BER. Also, the received power is proportional to the data rate. Combining the BER and data rate dependence on receiver sensitivity, we can define *receiver sensitivity* as the mean optical power required to obtain a required BER at a given bit rate. Of course, this does not mean that by increasing the power in an optical link we can have excellent BER. This is because increase in optical power also creates nonlinear effects, and so on.

SNR

The SNR of a receiver is defined as the ratio of signal power to noise power in the electrical domain. Consider Equation 2-41.

Equation 2-41

$$SNR = \frac{(SignalPower)Electrical}{(NoisePower)Electrical}$$

The SNR is proportional to input power squared. We can enhance SNR by increasing the load resistance. See Equation 2-42.

Equation 2-42

$$SNR = \frac{R_L R^2 P_{in}^2}{4K_B TFB}$$

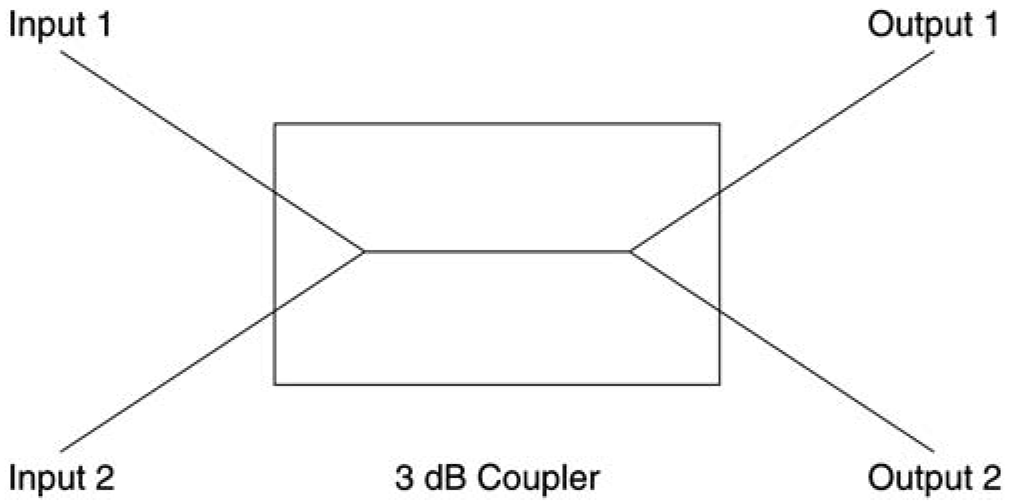
F is a proportionality factor for increasing the thermal noise content of a receiver. B is the bandwidth, P_{in} is the input power, R_L is the load resistance, and T is the temperature.

So far, this chapter has covered lasers and detectors. Of our WDM node, we have broached the transmitter and receiver sections only. Now we will discuss other components and underlying technologies in the WDM network. Starting with some simple passive components such as couplers and circulators, the chapter will move on to more advanced technologies such as filters and waveguides.

Couplers and Circulators

Couplers are the simplest optical devices. They are passive and completely bidirectional in nature in the sense that we can interchange the input and output ports. Couplers are N x M, where N and M are integers. In other words, we can have N input segments (fibers) and M output segments (fibers). The principle (of a coupler) is to fuse the cores of the N input fibers to the cores of M output fibers so as to create a power transfer device. Practically, 2 x 2 couplers are most common and are known as 3dB couplers because of the 3 dB loss in power at each output port due to a signal at one of the input ports. Refer to Figure 2-20.

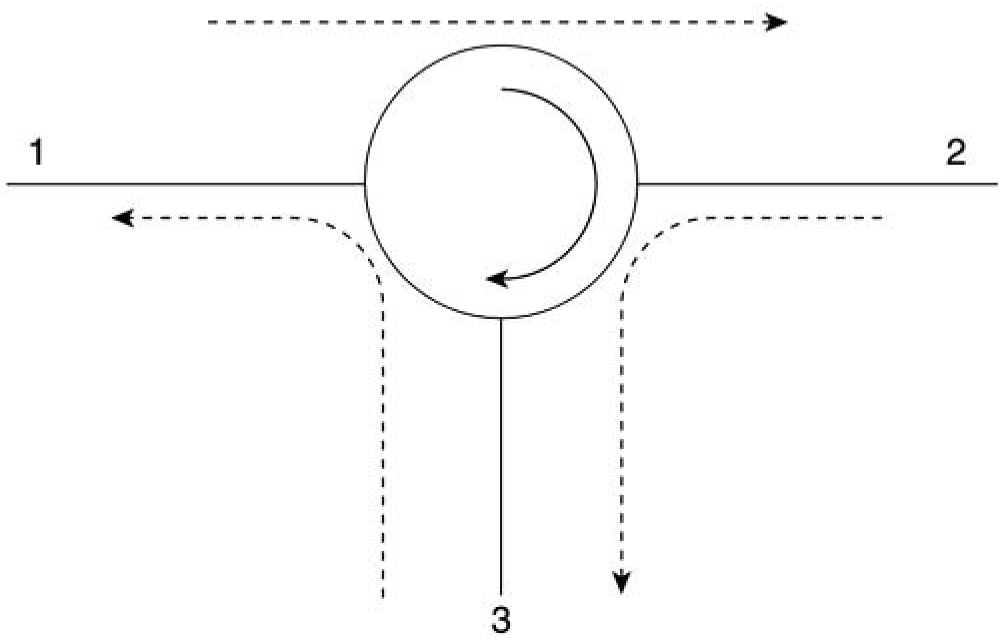
Figure 2-20. Coupler



Couplers find applications for monitoring WDM ports as well as for passively adding channels into a fiber. They are also used in passive optical networks (PONs) as a method to provide shared medium access (Ethernet PON, or EPON).

Circulators are shown in Figure 2-21.

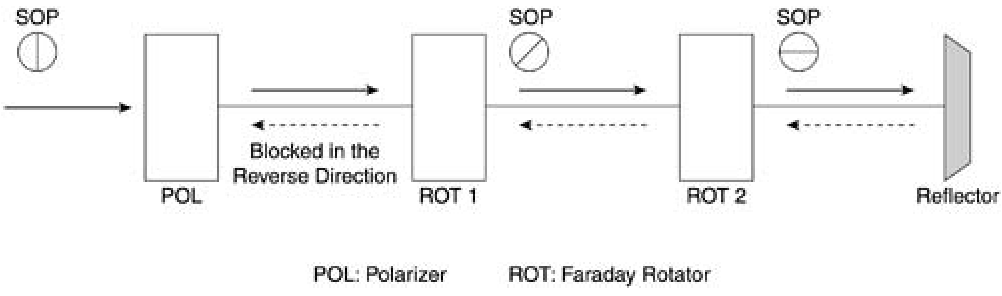
Figure 2-21. Three Port Circulators



A circulator is a multiport device that allows signals to propagate in certain directions based on the port that the signal came from (incident port). The operation is based on an isolator (analogous to an optical valve), which allows unidirectional propagation only. In Figure 2-21, the signal from port 1 moves freely to port 2; while the signal from port 2 cannot go to port 1, but it can go to port 3. Likewise, the signal from port 3 can go to port 1 but not to port 2.

The principle behind this directional communication is *polarization*. A *polarizer* is a device that allows light to pass through only if it is polarized in a certain manifestation. By inducing light to a polarizer, only the light that matches the phase of the polarizer passes through. This light is now subjected to a Faraday rotator, which rotates the state of polarization (SOP) by 45 degrees. A further rotation of 45 degrees by the second rotator makes the output state of polarization at the end of the second rotator 90 degrees as compared to the original input state (SOP). If this light reflects back, it is blocked by the polarizer; this is because its SOP is 90 degrees out of phase with that of the polarizer. Refer to Figure 2-22.

Figure 2-22. Isolator Showing Unidirectional Optical Communication Analogous to an Optical Valve



Cavities and Filter

Optical cavities and filters are important WDM devices that can demultiplex the composite signal. Tunable optical filters are key building blocks that can tune to a desired wavelength and tap a channel or a band of channels. Tunable optical filters are inherently of two types: cavity based and thin-film based. Cavity-based filters are the most common filters available; examples include the Fabry Perot cavity filters and the Acousto-Optic Tunable Filters (AOTF).

A filter is designed to have the following characteristics:

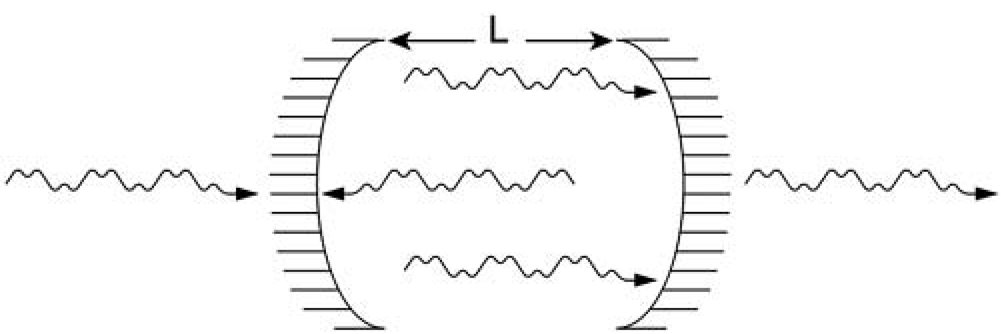
- A clean window of operation (pass band); in other words, minimal cross-talk with adjacent channels
- A wide tuning range that should be able to cover the entire band of operation
- A fast tuning speed that should be dynamically provisioned to facilitate changing traffic requirements
- Should not affect the polarization state of the passing signal.

Most filters are based on the principle of optical cavities, which can be tuned to a resonant frequency. The other technology used in conjunction with optical cavities is that of thin film filters.

Fabry Perot Cavity Filters

A Fabry Perot cavity consists of two reflective surfaces that are separated by a hollow region. The distance between the reflective surfaces can be made to change by changing the current associated with the transducer, responsible for creating the cavity. In general, the cavity has two reflective surfaces with reflectivity that is a function of the operating wavelength. The reflectivity can be made to change for different resonant wavelengths. For a resonating cavity, the resonant wavelength is the only wavelength, and it does not suffer reflection from one of the two mirrored walls (see Figure 2-23).

Figure 2-23. Fabry Perot Cavity Filter



The transmission characteristics or the transitivity of an FP cavity is best for $\lambda = \lambda_{resonance}$, which is generally shown as in Equation 2-43.

Equation 2-43

$$\lambda_{res} = \frac{L}{2n}$$

In the equation, n is an integer and L is the distance between the two walls of the cavity.

Two properties that are important for design of cavities are the free spectral range (FSR) and finesse. For a mirror of reflectivity R, the finesse is provided as in Equation 2-44.

Equation 2-44

$$F = \frac{\pi\sqrt{R}}{1-\sqrt{R}}$$

In the equation, R is a ratio of incident to reflected power of a mirrored surface (reflectivity). Therefore, R is figure of merit for a reflective surface. Finesse of a cavity is a figure of merit that depicts the amount of fine tunability that can be achieved by using this cavity.

The FSR can be defined as the minimum range of two successive filtered peaks; it is the frequency difference between two transmission peaks. See Equation 2-45.

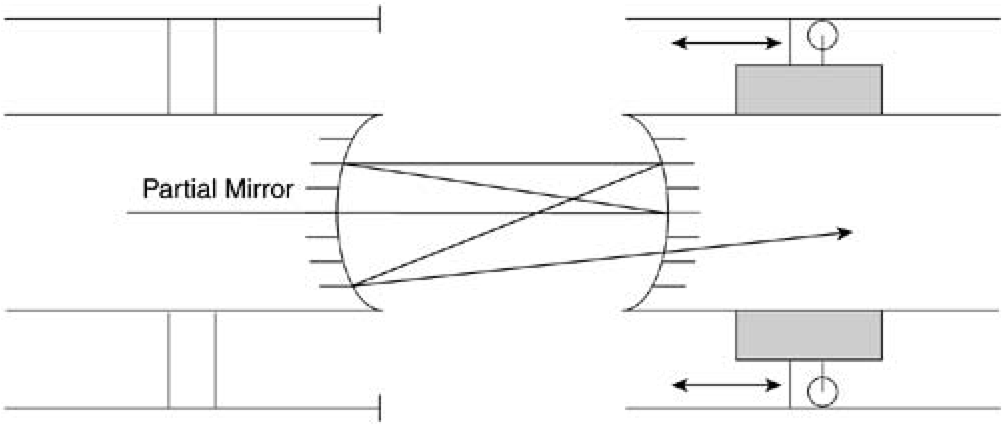
Equation 2-45

$$\Delta f_1 = \frac{c}{2n_g L}$$

In the equation, n_g is the group index and L is the length of cavity.

The length of the cavity can be changed by applying a voltage to a transducer that mechanically shifts the mirrors (closer or further away). The change in length for practical WDM systems is a function of the wavelength. Moreover, the mirrors are not more than 150–200 µm apart, which makes FP cavity fabrication more difficult. A popular approach is to utilize the air gap between the two polished surfaces of two fibers, as in Figure 2-24.

Figure 2-24. Fiber-Based Fabry Perot Cavity



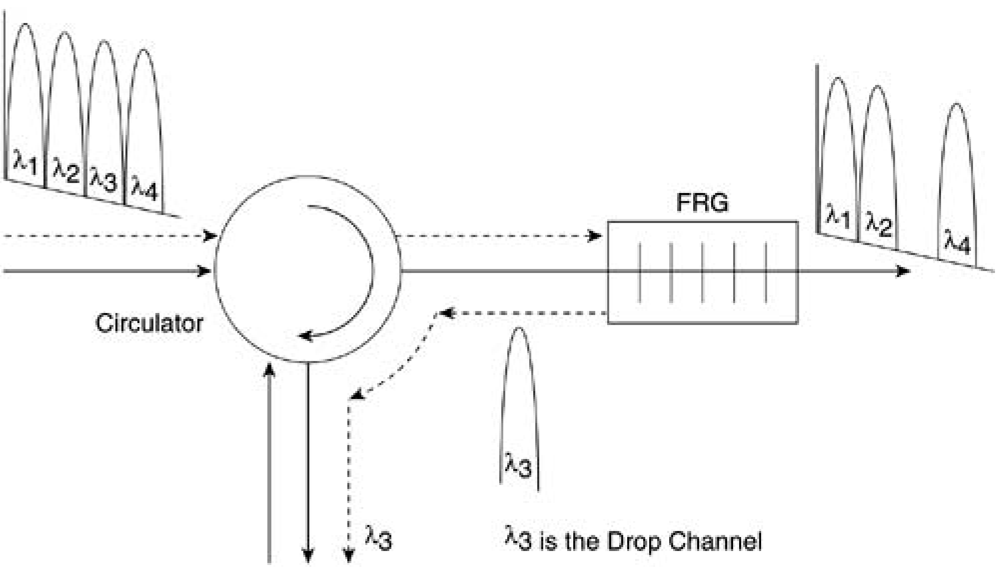
Fiber Bragg Gratings (FBG)

Bragg gratings are cyclic periodic perturbations of the refractive index in the fiber. For an incident WDM composite signal propagating through this perturbation of refractive index, one wavelength of the entire propagating spectra would be reflected back (backward). This wavelength is known as Bragg's wavelength ($\lambda_B = 2n\tau$, where n is the refractive index of the waveguide and τ is the period of perturbation or grating). Therefore, when a composite WDM signal is incident on a Bragg's grating, all wavelengths except Bragg's wavelength pass through while Bragg's wavelength is reflected back.

Gratings can be "written" onto a fiber in numerous ways. One popular method is to use photosensitivity of doped Germanium in fiber and etch a grating pattern by exposing the photosensitive fiber to alternating intensities of ultraviolet light. The FBG can be characterized by low loss $\cong$ 0.1 dB and low channel cross-talk. FBGs find applications in most WDM systems, such as channel drop elements, dispersion compensation devices, and filters.

An example of FBG channel drop units is shown in Figure 2-25.

Figure 2-25. FBG as Channel Drop Units



Acousto-Optical Tunable Filter

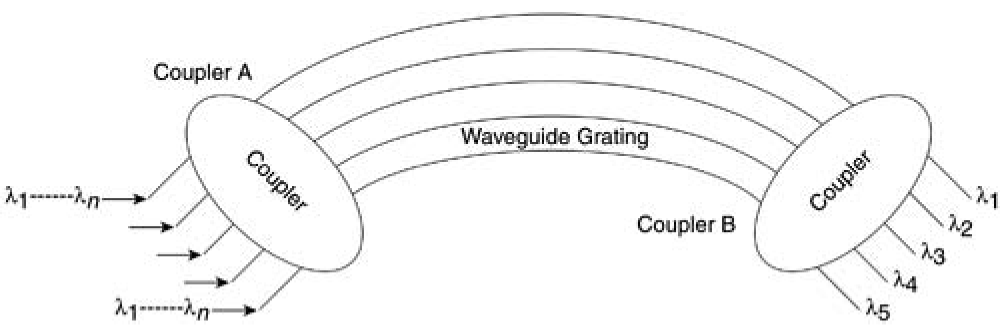
Grating filters discussed previously cannot be dynamically tuned with ease to drop or add any channel. This issue can be solved by creating dynamic gratings by using acoustic waves. By creating a series of acoustic (sound) waves inside a waveguide, an acoustic grating is formed. Sound waves are essentially longitudinal waves with propagation based on formation of compression and rarefaction zones, unlike the crest and troughs as in a transverse wave (lightwave). These compressions and rarefactions are equivalent to regions of high and low refractive index. Light passing through such a disturbance has the same effect as passing through a grating. The interaction of light with the acoustic waves is termed as the *photon-phonon interaction* given by an effect known as the *photoelastic effect*. A photon-phonon interaction can easily be understood as collision under energy conservation.

AOTF can be fabricated best by using Lithium Niobate ($LiNbO_3$) waveguides, producing small polarization-independent filters. AOTFs are characterized by a tuning range in the excess of 100 nm covering both C and L bands. Tuning times of AOTF can be very low in the range of several microseconds. (Some demonstrations have nanosecond tuning also.) One limitation is channel cross-talk, which is currently being investigated.

AWG Arrayed Waveguides

An AWG device consists of many waveguides of different lengths converging at the same point(s). Signals coming through each of these waveguides travel through a length such that they interfere from the signals through the other waveguides (at the converging point) either constructively or destructively, depending on the net phase difference between the signal and its interfering counterpart(s). Such a phased array of waveguides can be used as a multiplexer or demultiplexer (follow Figure 2-26 closely). For demultiplexing, the composite WDM signal is coupled into an array of waveguides using a 1 x N coupler. Each signal in the waveguides gets a different phase shift because of different lengths of each waveguide. The amount of phase shift induced depends on the wavelength. The interference caused at the second coupler (see Figure 2-26) can be controlled such that each channel is separated into each of the output fibers. This is due to the spatial diversity induced by the interference of phase-shifted signals. In this way, a composite signal consisting of many wavelengths can be demultiplexed into individual wavelengths, one in each of the output fibers (ports).

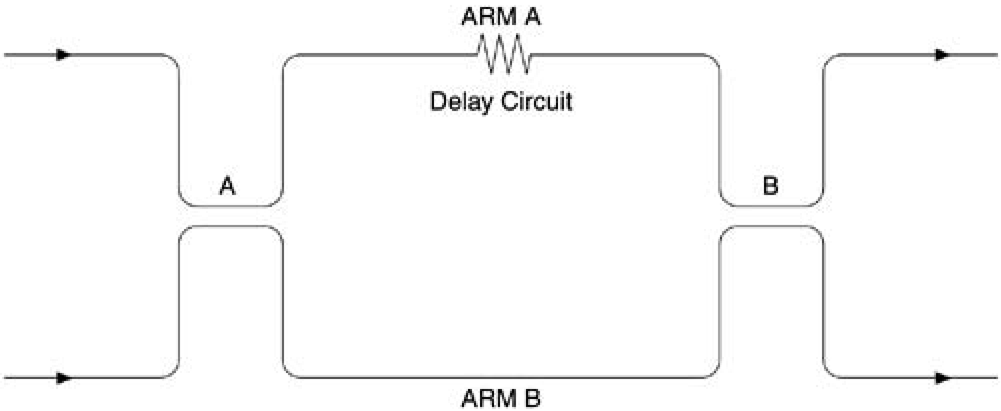
Figure 2-26. AWG (Arrayed Waveguide)



Mach Zehneder Interferometer and Filter

A*Mach Zehneder Interferometer* (MZI) is a two-arm device, such that the signals in the two arms interact with each other twice. Physically, an MZI can be constructed by connecting two passive 2 x 2 couplers in tandem. The couplers are equibalanced; in other words, input power is equally split into the two arms. The first coupler (A) divides the signal into two (see Figure 2-27).

Figure 2-27. MZI- Mach Zehneder Interferometer



The two propagating signals can be made to obtain different phase shifts by varying the lengths of the two arms. The signals, upon interfering with each other at coupler B, might have constructive or destructive interference. The phase shift induced is a frequency (hence wavelength)-dependent quantity ϕ(f), shown in Equation 2-46.

Equation 2-46

$$\phi(f) = \cos^2 \pi f \Delta$$

In the equation, f is the optical frequency, and Δ is the delay induced due to the different lengths. The MZI, by inducing phase shifts that cause either constructive or destructive interference, blocks one or segregates a particular frequency (wavelength). The MZI is built on silica substrate, and these kinds of optical circuits are called *planar lightwave circuits* (PLC) because of their planarity of the substrate. A more mature technology is Indium Phosphide substrate or Lithium Niobate substrate.

Thin Film Filters

Another variety of optical filters is the thin film filter. Thin film filters are similar to cavity filters in the sense that the resonant cavity selects the wavelengths that are allowed to traverse through. The cavity is formed by the thin films with interfaces that act as reflectors. The wavelength or group of wavelengths that is selected depends on the length of this cavity. Thin film filters are commercially available and perform functions of optical band pass filtering as well as single wavelength filtering. Multicavity cascaded thin film filters have excellent response in the sense that due to cascade, the filtered band approximates to more rectangular characteristics (an ideal filter). Thin film filters are typically made of quarter wave ($\lambda/4n$) thick layers of alternating high and low refractive indexes. The principle is that of coalescing multiple layers, forming multiple resonant cavities on top each other. A thin film filter has three main regions: the spacer, the transition layer, and the reflective stack. The spacer consists of multiple quarter wave layers of either high refractive index or low index, but not both. The spacer is between the reflective ends of the formed cavity. The transition layer is composed of a single quarter wave layer, and its function is to produce flat-top filtering. The reflective stack is made of alternating high and low indexes and forms a dielectric mirror. Table 2-2 discusses various filters and their characteristics.

Table 2-2. Filter Classification Table

Filter Type	Tuning Range	Tuning Time	Comment
Fabry Perot	500 nm	1–10 ms	Fiber implementation available
Acousto-optic	250 nm	10 micro sec	Can be used as router
AWG tunable Filter	40 nm	10 ms	Thermo-optic tuning
Light crystal FP	30 nm	0.5–10 micro sec	Low power
Electro optic	16 nm	1–10 ns	Band filter
Fiber Gragg	10 nm	1–10 seconds	Mechanical tuning
Mach Zehneder	4 nm	50 ns	High loss
Semiconductor	5 nm	0.1–1 ns	Small number of channels

Complex Components: Transponders

By definition, a *transponder* is a device that enables end-users to access the WDM channels. A transponder can detect optical signals at various wavelengths and convert them to ITU grid wavelengths. Transponders are considered complex WDM components because they consist of several subsystems—such as lasers and photodetectors—in addition to filters. Different versions of transponders are available depending on the requirement. The simplest is the reshape and reamplify (2R) version in which protocol-independent conversion and detection of the optical signal is carried out. The more complex and expensive version is the reshape, retime, and reamplify (3R) transponder, which is protocol dependent. Typically, such transponders are needed for high bit rate signals. For example, an OC-192 transponder card will not work for a 10 GigE card for the simple reason of protocol incompatibility even though the line rates are almost the same.

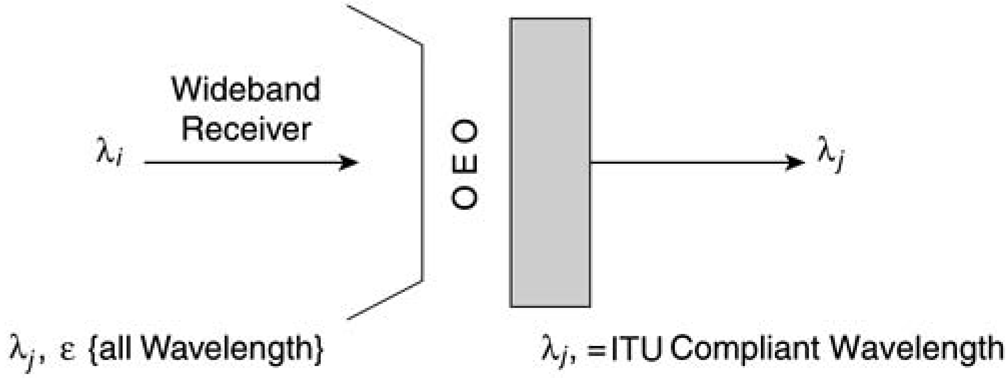
The channels in a WDM network are allocated according to a certain standardized rule. The standardizing body, the ITU, is responsible for allocating fixed wavelengths in a WDM network. These wavelengths are known as ITU-grid wavelengths. Currently in the C band as well as L band, the ITU allots wavelengths with 0.8 nm or 100 GHz as well as 0.4 nm or 50 GHz separations. Previous allocation was 200 GHz spacing.

In a service provider network, it is now imperative to have all lightpaths or wavelengths specified by the ITU standard. In practice, most service provider networks use C and sometimes L bands. However, a client to the service provider might not use a standard wavelength on his lightpath.

The client wavelength needs to be translated into an ITU-defined network-compactable wavelength. This translation of wavelength from a client signal to an ITU-compliant signal is accomplished by transponders. Essentially, transponders have a functionality that allows them to receive any wavelength but transmit only an ITU-compliant wavelength. Furthermore, this functionality can be increased so that transponders transmit a range of tunable (ITU-defined) wavelengths instead of just one fixed wavelength. As shown in Chapter 6, "Network Level Strategies in WDM Network Design: Routing and Wavelength Assignment," tunable transponders increase the throughput of a network by allowing more flexibility in the number of lightpaths administered into the network.

Tunable or fixed transponders receive a client signal, detect the optical signal, and convert it to electronic signal. A laser then modulates the electronic signal by using external modulation techniques. By replacing one laser with an array of pretuned lasers, we can select a wavelength from a band of available wavelengths. Therefore, transponders essentially are wavelength converters that use O-E-O as a means to convert the ingress wavelength to the egress wavelength. Refer to Figure 2-28.

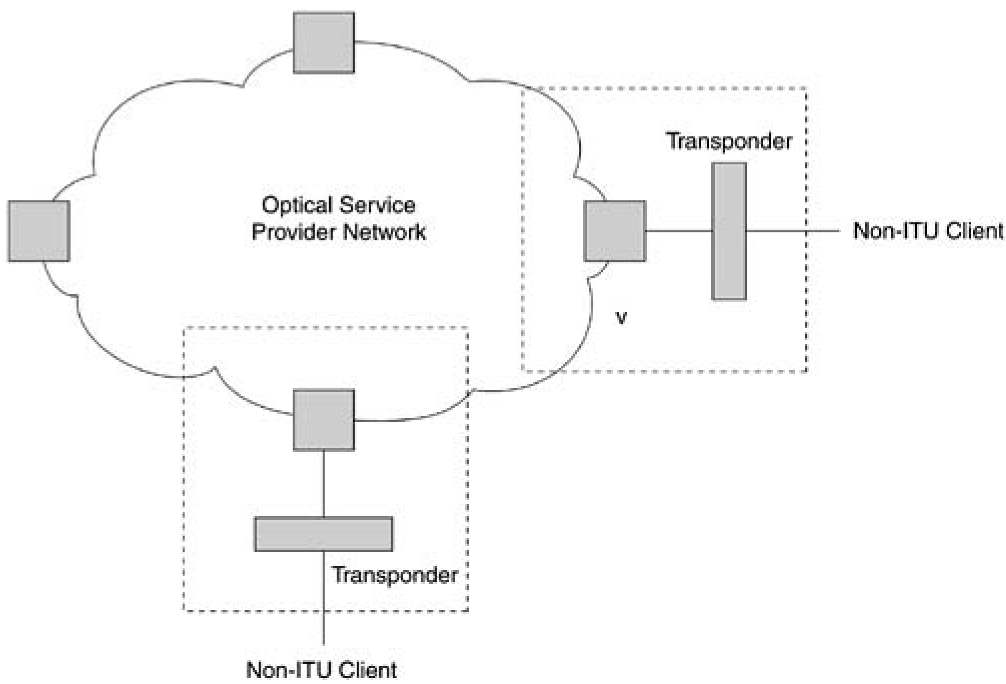
Figure 2-28. Transponder Functional Diagram



Because of the O-E-O symmetry, transponders usually perform 3R regenerations, but 2R generations are also possible. (3R transponders are not bit-rate transparent.) For example, an OC-48 transponder cannot be used for an ESCON or a FICON (SAN interfaces). Configurable 3R transponders are available in the market today; an interface can be configured to OC-48/OC-12/OC-3/GigE.

Transponder-based WDM systems are considered close systems due to the compatibility with the installed client base (SONET/SDH/IP). They allow the legacy client equipment to feed directly into the transponder system. The transponders allow a variety of client interfaces over WDM networks (Ethernet over WDM; storage area network (SAN) over WDM; and so on). Refer to Figure 2-29.

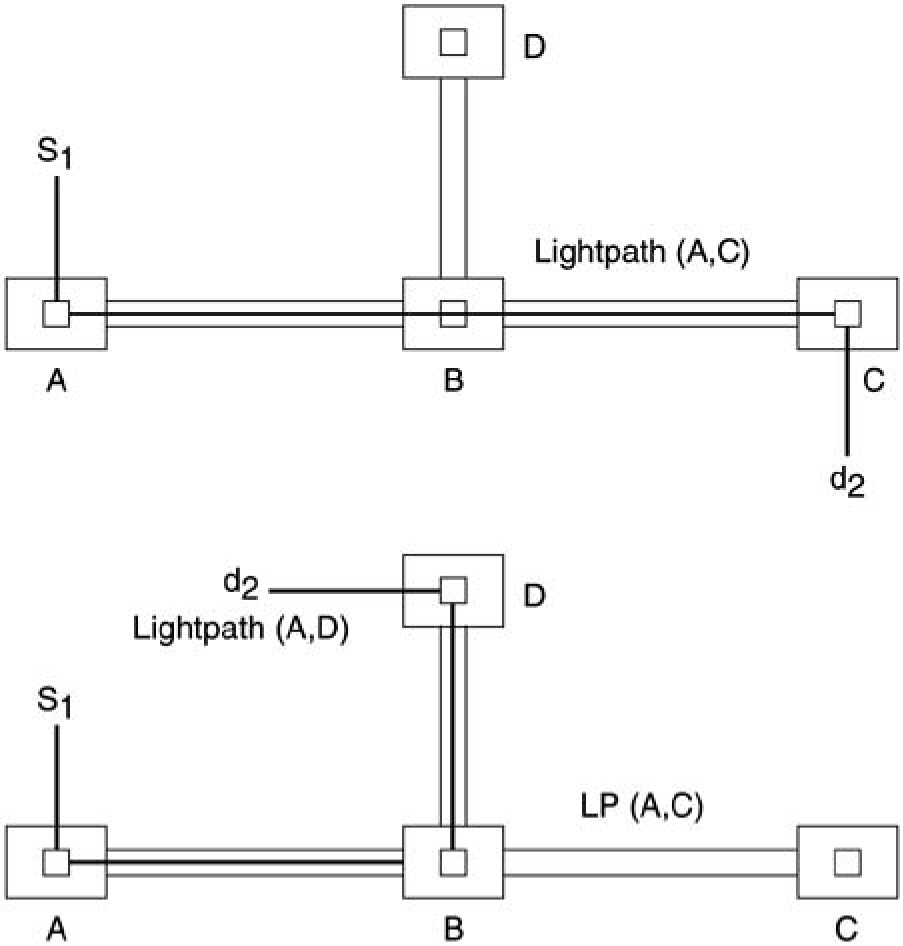
Figure 2-29. Typical Transponder-Based Network



Switches

Optical switches represent the single-most dynamic element in a WDM network. Traditionally, switches can switch data between different ports of a network element. Switches generally cater to two types of data: circuits and packets. A circuit or a lightpath (in the optical domain) is an end-to-end connection (source- destination pair) over which data flows. Packets are discrete messages/datagrams of short sizes. Current technology facilitates circuit switching. The philosophy behind circuit switching is that a lightpath between a particular source and destination pair can be established for a sufficiently long period of time. Lightpath may be switched optically from this destination to another using an optical switch, as shown in Figure 2-30.

Figure 2-30. Lightpath Switching: Predominant in Today's Networks. Note the switch at node B is reconfigured.

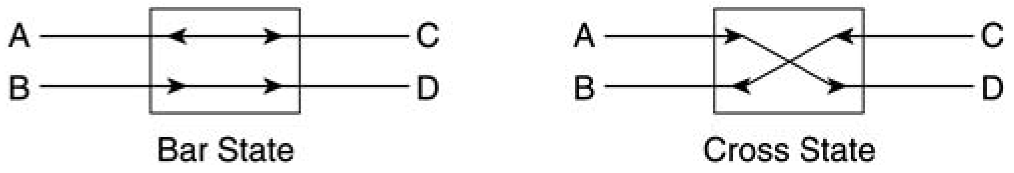


In packet-switched networks, individual packets are switched between source-destination pairs. Packet switching is more dynamic than circuit switching. Considering the present form of the Internet, packet switching is more desirable than circuit switching. This is because data traffic is currently more dominant than voice. However, packet switching in the optical domain is currently only an academic exercise—it's far from being feasible. Circuit switching in the optical layer is more feasible and is known as *lightpath switch*. Optical switches consist of two types: add-drop switches and optical cross-connect.

Mechanical Switches

Micro-mechanical switches, as shown in Figure 2-31, have become a mature technology for switching lightpaths. Switches of small degree (for example, 2 x 2) work by mechanically moving a pair of fibers between corresponding output ports. Due to the mechanical movements involved, such switches are typically slow (5–10 ms). The movements also create dynamic loss variations. Such switches have high insertion loss of about 1–2 dB.

Figure 2-31. Mechanical Switches: Logical Connection Diagram

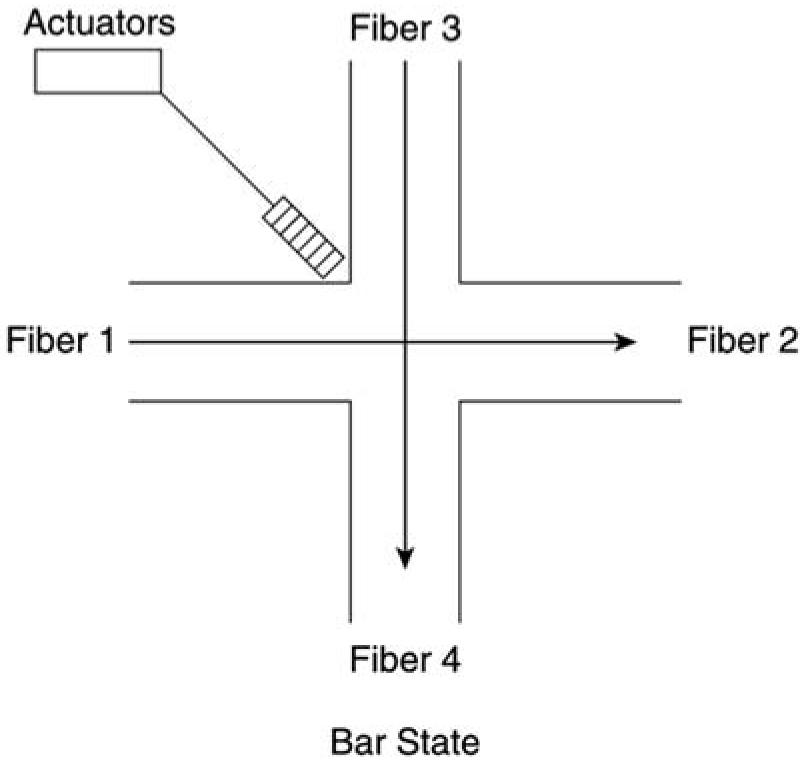


The advantage that micro-mechanical switches offer is that they are fairly robust and inexpensive. Presently, many vendors use such switches in network elements because of their low cost and comparative performance. These switches have negligible wavelength dependent loss and work quite the same for different wavelengths.

Micro-Electro-Mechanical System Technology

Micro-electro-mechanical systems (MEMS) is a fascinating innovation that is applied to optical networking. Minute electro-mechanical systems can be deployed to perform certain switching functions in the optical domain. Consider the example of Figure 2-32 on MEMS.

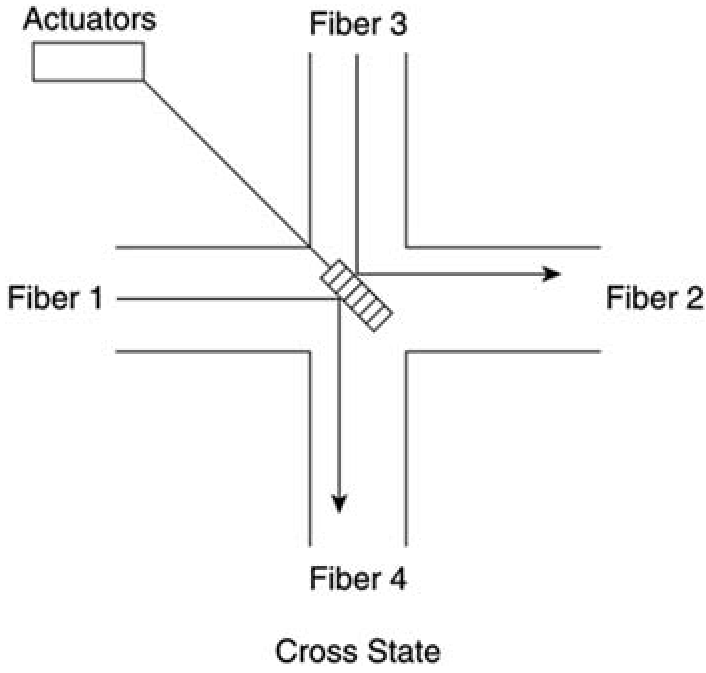
Figure 2-32. 2D MEMS (Bar State)



Four fibers (1, 2, 3, and 4) are coupled together to form a 2 x 2 cross-connect. A double-sided mirror that is perpendicular to the plane of the page is present and held in position by two actuators. In the bar state, light from fiber 1 is coupled to fiber 2, and light from fiber 3 is coupled to fiber 4.

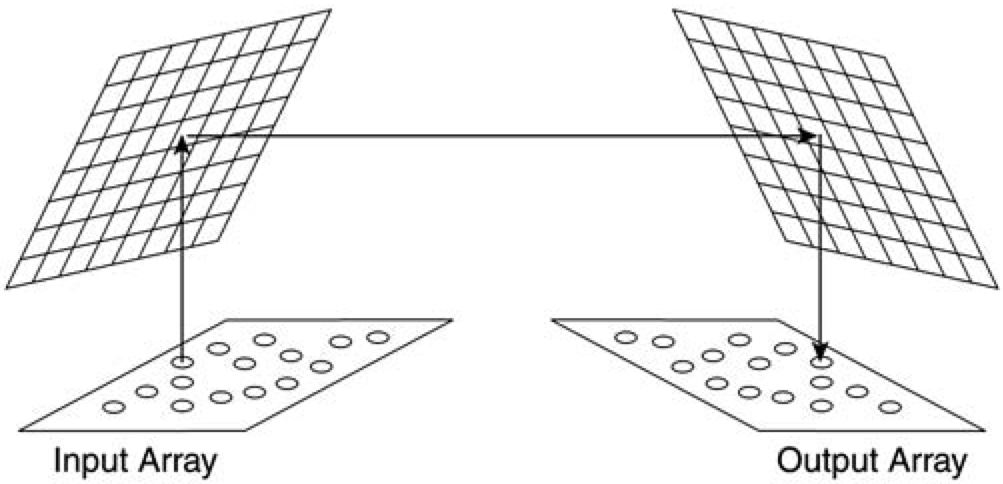
Upon application of current to the lower actuator, the mirror moves "in" to the cavity (gap). Now the light from fiber 1 is reflected to fiber 4 and the light from fiber 3 is reflected to fiber 2 (see Figure 2-33). The movement of the mirror is due to electro-mechanical interaction of the actuators. These kinds of switches can be applied to a WDM network to switch lightpaths dynamically. The construction of such a switch is a tedious process. However, MEMS is a mature technology, and the deployment is quite feasible in today's networks.

Figure 2-33. Cross State



Typical switching times are in millisecond range. The insertion loss of the switch is about 1 dB per port. MEMS switches can be scaled to provide an N x N cross-connect. The one shown in Figures 2-32 and 2-33 is a 2D MEMS (2-dimensional MEMS). In contrast, Figure 2-34 shows a 3D MEMS switch (3-dimensional). The incident light in 3D MEMS is switched in 3D space using collimated lens to provide efficient 3D switching; it is comparable to free space optics.

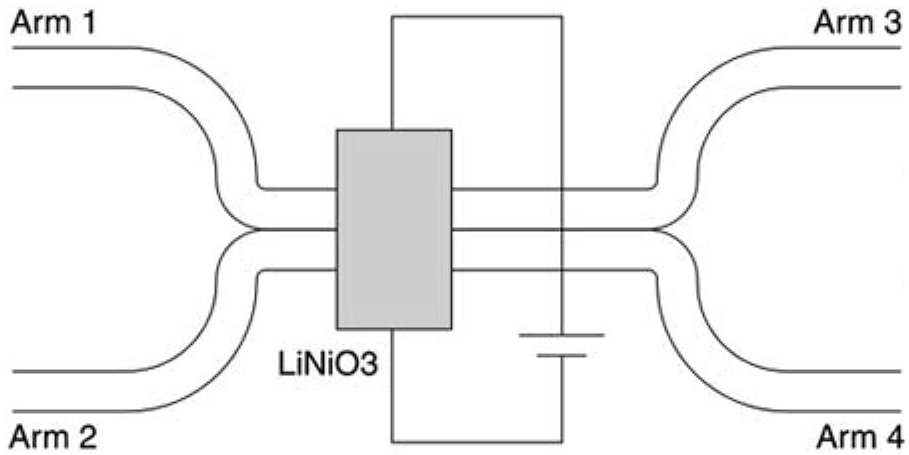
Figure 2-34. 3D MEMS Switch Operation



Electro-Optical and Thermo-Optical Switches

By using a *directional coupler*—a coupler with a coupling ratio between the two output ports that can be made to change—you can fabricate an electro-optic switch. A 2 x 2 coupler can be made to switch from bar state to cross state by changing the refractive index inside the coupling medium. This index change leads to different coupling ratios. The change in refractive index is brought about by inducing different currents in a Lithium Niobate (LiNiO3) modulator region. Lithium Niobate waveguide has the property to change the index of refraction subject to different current conditions. Therefore, the coupling ratio can be made to change, causing the power at one of the ports to vary accordingly and switch the lightpaths. Refer to Figure 2-35.

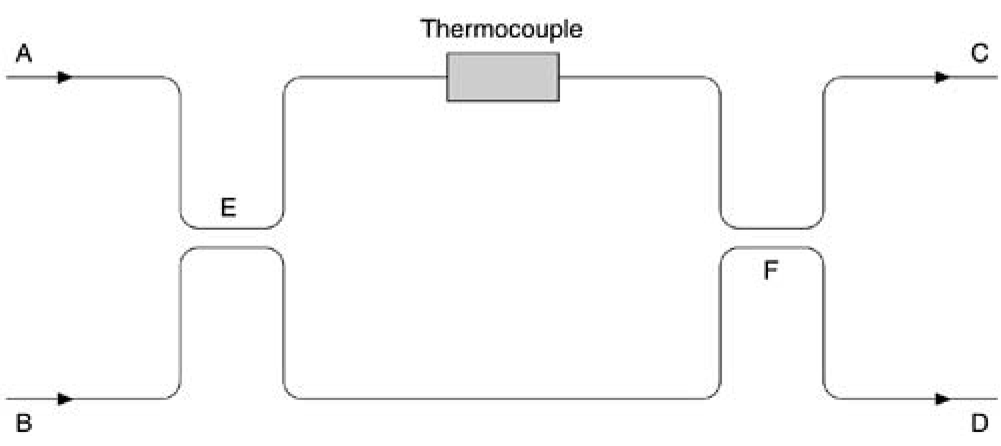
Figure 2-35. Electro-Optic Switch Based on MZI



Switching times is about 1–5 ns, but loss can be almost 2–3 dB because of imperfect coupling ratios.

Thermo-optic switches are based on MZIs. The refractive index of one of the arms can be changed by altering the temperature, which is further controlled by current. This change in length causes a phase difference associated with both arms. The phase difference can be made to constructively or destructively interfere with each other, causing a similar direction coupler-like-environment and the ability to switch lightpaths. Refer to Figure 2-36.

Figure 2-36. Practical Implementation of Thermocouple Switch



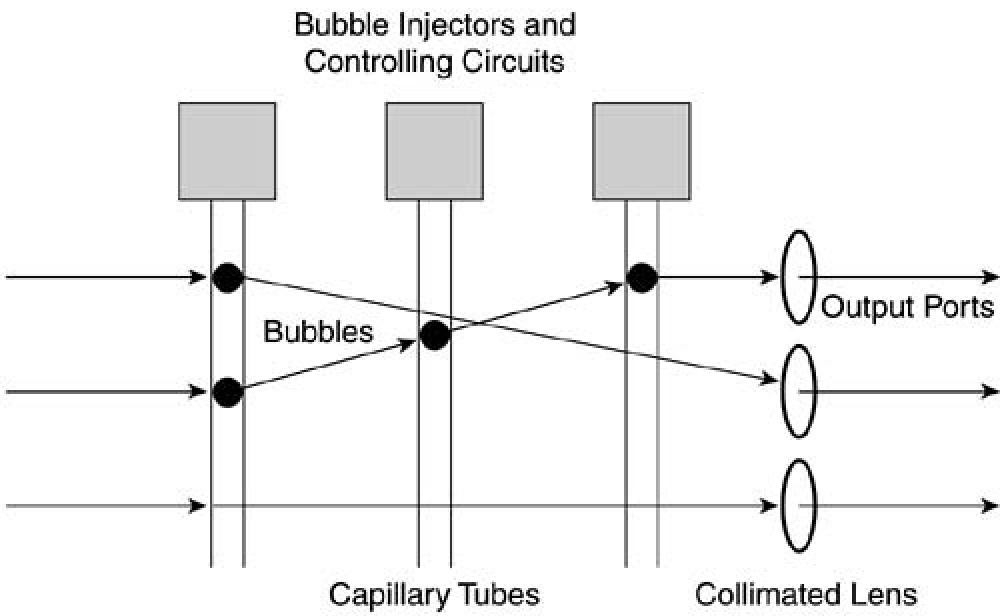
Such devices switch signals as follows. Consider Figure 2-36. Signals A and B are made to interfere at point E such that B interferes destructively; therefore, at point C, only A will appear. Similarly, A can be made to switch to port D by destructively interfering at point F. Such switches are slow and have high cross-talk due to an imperfect coupling effect.

Bubble Technology

Recently, Agilent demonstrated a bubble technology-based optical switch. The principle of bubble switch is demonstrated in the bubble jet printer. Micro bubbles are made to enter a region of interaction of optical beams inside capillary waveguides. The refractive index of the bubbles can be made to vary such that optical beams refract to different ports.

ConsiderFigure 2-37, in which the input signals are switched between output ports by introducing bubbles of refractive index that are capable of deflecting the incident beam. These switches are easy to make. Some severe drawbacks include large loss resulting from a lossy medium and large cross-talk. Such switches have a relatively small lifetime (1,000 hours). The founding companies are conducting extensive research to make the bubble switch more efficient. Scalability issues are also quite prominent in bubble technology. 32 x 32 switches were demonstrated at OFC 02.

Figure 2-37. Bubble Switch

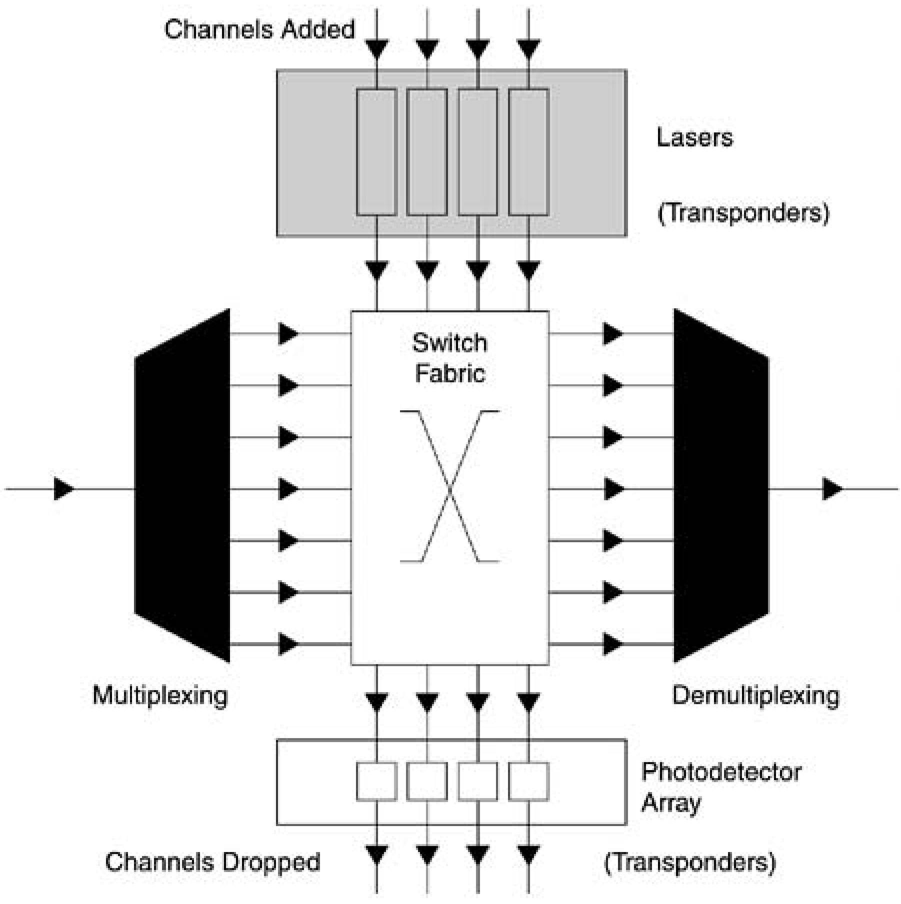


Analysis of the Node

So far, this chapter has discussed various WDM components and subsystems that are used to develop a node. Now the chapter will turn to the positioning of these technologies and components in an actual WDM network element. The most important WDM element is the optical amplifier, which is considered in Chapter 3.

Figure 2-38 shows the functional representation of WDM network elements. For simplicity without loss of generality, a single East-West (direction) fiber is assumed. Consider the propagation of a WDM signal from left to right. The signal before entering the node might be amplified. Most practical WDM nodes have both preline (before the node) and postline (after the node) amplifiers as part of the nodal configuration.

Figure 2-38. Analysis of Node



The composite signal is demultiplexed at the AWG. AWG technology has matured significantly over the past decade. AWGs that have constant loss per channel can now be fabricated. The individual wavelengths demultiplexed from the AWG feed to a switching matrix. The switching matrix is the heart of WDM node. Currently, most switches have add-drop and pass-through functionality for wavelengths. More futuristic designs have wavelength converters, meaning that lightpath switching is embedded into the matrix, demonstrating optical cross-connect architecture. Still more futuristic designs will have optical packet or burst switched/photonic slot-routing architectures. Those designs might not be introduced for yet another decade.

Lightpaths on "wavelengths" can be added/dropped or passed through by the switch. The dropped lightpath is fed to a detector, which might be canvassed inside a transponder (receiver) card. The switch could also add lightpaths emanating from laser diodes (which, like detectors, can be canvassed as transponder-transmitters). Individual wavelengths from the switches are fed to an egress AWG that combines them into a composite WDM signal, ready for transmission. This demuliplex and multiplex regions in a WDM node are also called optical multiplex section.

Numerous technologies are associated with the various interacting components. Although this book emphasizes the network heuristics, these optical technologies play a crucial role in governing the network parameters.

Summary

This chapter showcased some of the WDM components and technologies. It introduced optical transmitters and various forms of lasers used in WDM transmission. It also discussed optical receiver design and its importance to system design based on noise and bit error rate (BER). Some elementary mathematical analysis of BER as well as the Q factor and signal-to-noise ratio relations have been studied. Components such as couplers, circulators, various forms of filters, and optical switches are also studied.

References

[1]Agrawal, Govind. *Nonlinear Optics*, Second Edition. Agrawal Wiley Publications, 1997.

[2]Ramaswami, Rajiv and Kumar Sivarajan. *Optical Networks: A Practical Perspective*, First Edition. Morgan Kauffman Publications, 1999.

[3]Keiser, Gerd. *Optical Fiber Communication*, Third Edition. McGraw-Hill Publications, 2000.

[4]Bouda et al. *Tunable Lasers*. Proc of OFC 2001. Anaheim, CA.

[5]Fukuda, I. *Optical Semiconductor Devices*. New York: Wiley, 1990.

[6]Goodman, J. W. *Physics of Optoelectronic Devices*. Wiley, 1995.

[7]Steeleman, G. E. and C. M. Wolfe. *Avalanche Photodiodes*. Academic Press, 1977.

[8]Taub and Shilling. *Principles of Communication Systems*, First Edition. McGraw-Hill, 1990.

[9]IEEE*Quantum Photonics letters*. March 1984.

[10]Kressel, H. and J.K. Butler. *Semiconductor Lasers and Heterojunctions LEDs*. San Diego: Academic Press, 1977.

[11]*IEEE Journal of Lightwave Technology*. April 1987.

[12]Robinson, F. N. H. *Noise and Fluctuations*. Oxford: Oxford University Press, 1977.

[13]IEEE Special Issue on Optical Technology and Multiwavelength Networks in *Journal of Lightwave Technology*. June 1993.

[14]Liu, M. *Principles and Applications of Optical Communications*. Prentice Hall, 1996.

[15]Green P. *Fiber Optic Networks*. Prentice Hall, 1993.

[16]Zirngibl, M. et al. Photonics Technology Letters. August 1986.

[17]Brackett, C. *IEEE Transaction on Communications*. May 1990.

[18]Agrawal, G. *Fiber-Optic Communication Systems*, Second Edition. New York, NY: Wiley Interscience; 1997.

Chapter 3. Networking with DWDM -2

Signal propagation in an optical fiber is limited because of attenuation. Attenuation, dispersion, nonlinearities, and other harmful effects cause the signal level in an optical fiber to degrade with their accumulation (transmission length). To sustain optical communication, the receiver (photodetector) must be able to detect or read signal pulses and distinguish between 1s and 0s efficiently.

The signal before reaching the detector suffers from multiple attenuations (it adds noise due to decrease in signal level and increase in noise due to accumulation); therefore, its signal-to-noise ratio (SNR) is degraded. In Chapter 2, "Networking with DWDM -1," we saw bit error rate (BER) as a figure of merit for optical signals. The SNR is an important parameter to quantify BER over a given transmission length.

In a communication channel, the signal level can be boosted at intervals with *optical amplifiers*. These monolithic blocks (subsystem modules) can optically amplify a signal (completely in the optical domain—with no electronic regeneration) and, therefore, raise the signal level.

Optical amplifiers are functionally identical to electrical amplifiers or repeaters. Repeaters are 3R (reshape, reamplify, and retime) O-E-O regenerators, which completely detect, amplify, and retransmit a signal. In contrast, optical amplifiers have an advantage over repeaters in the sense that the data streams being amplified are amplified entirely in the optical domain. Moreover, optical amplifiers are completely transparent to signals and protocols. Three basic types of amplifiers are being deployed:

- Doped fiber amplifiers (such as Erbium doped fiber amplifiers, or EDFAs)
- Raman amplifiers (scattering amplifiers)
- Fiber-semiconductor optical amplifiers (SOAs)

Doped fiber and Raman amplifiers were first introduced almost half a century ago. SOAs are comparatively more recent advances. This chapter discusses each of these types of amplifiers in the subsequent sections.

A Typical Optical Amplifier

An optical amplifier works on the same principle as that of a laser. In short, incident light is amplified by sustained stimulated emission (discussed in Chapter 2). The amplification is achieved by a pumping process whereby either electrical or optical pumping boosts the incident signal power in a gain medium or just in a fiber. A pump is a local power source that couples its power to an incident optical signal, thereby amplifying the incident signal by transferring its power either directly or through doped impurities to the optical signal.

For an amplifier to exhibit gain, stimulated emission needs to achieve a condition of population inversion. On achievement of this status, a majority of the electrons are in the excited state (higher state = N_2) as compared to the electrons in the ground or common state (lower state = N_1). Various levels of excitation are possible, each yielding a different gain profile. The gain profile depends on factors such as pump power, gain medium heuristics, and spontaneous emission in the medium. An optical amplifier should have the following characteristics:

- A flat gain spectra (equal gain for all the channels).
- High gain/channel-to-input pump power required ratio. In other words, for every milliwatt (mW) of incident pump power, the amplifier should give the highest possible amplification.
- Low amplified spontaneous (random) emission (ASE), causing noise in the system.
- Negligible wavelength dependence and polarization depended loss (PDL).
- Low gain tilt: The difference in power levels between the lowest amplified and the highest amplified channels should be minimal.

In general, while designing a WDM network, the issues needed to be consider are gain flatness, output power, noise figure, and amplifier bandwidth.

Amplifier Heuristics

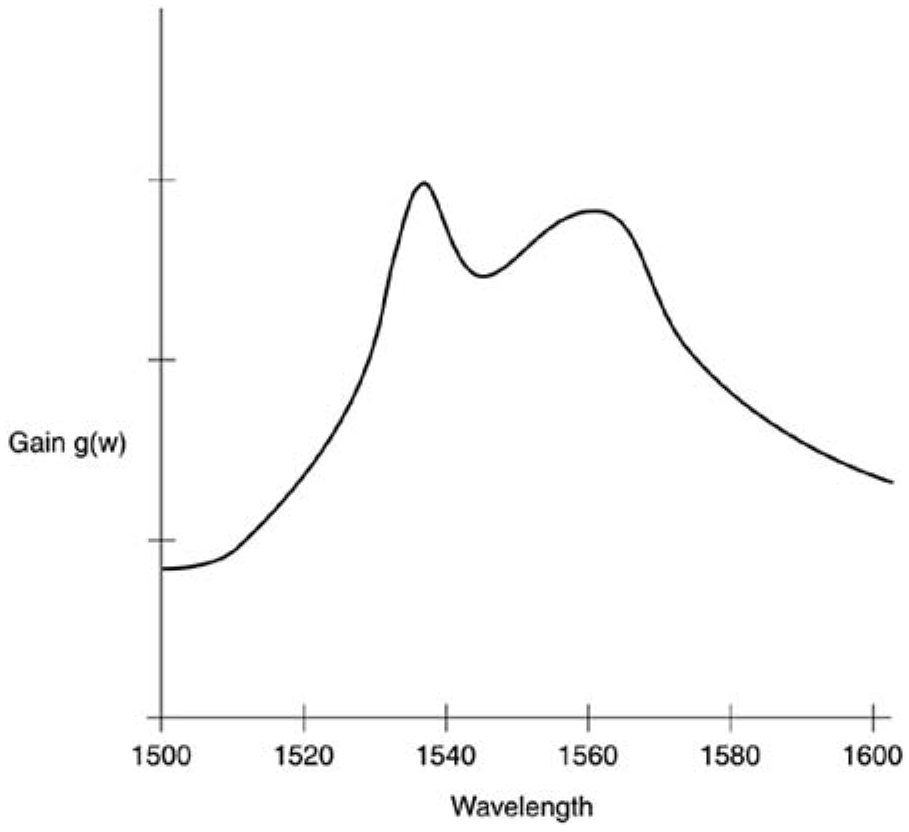
The gain obtained in an optical amplifier is a frequency-dependent quantity. The gain coefficient of an amplifier g(ω) is shown in Equation 3-1[2].

Equation 3-1

$$g(w) = \frac{g_0}{1 + (w - w_0)^2 T_2^2 + \frac{P}{P_s}}$$

In the equation, g(ω) is the gain coefficient, ω is the incident optical frequency, ω_0 is the atomic transition frequency, T_1 is the fluorescence time (through which T_2 is calculated), T_2 is the dipole relaxation time, P is the input power, and P_S is the saturation power for a given amplifier. For an optical amplifier, bandwidth of amplification is an important value. It shows the maximum spectral spread of the cumulative WDM signal that can be amplified. For a particular amplifier exhibiting frequency-dependent gain g(ω), the *amplifier bandwidth* is defined as the cumulative spectral band at full width half maximum (FWHM) of g(ω). The gain profile of g(ω) has a twopeaked curve, as shown in Figure 3-1.

Figure 3-1. Amplifier Gain Profile



In Equation 3-1, it is clear that the gain is maximum for $\omega = \omega_0$ and when P/Ps tends to be zero. It can be shown that the amplifier bandwidth is less than the gain bandwidth. Another factor of interest is the saturation power of an amplifier. The *saturation power* of an amplifier can be defined as the input power at which the gain begins to reduce after reaching a certain value. This phenomenon can also be viewed as *gain saturation*, in which any further increase in pump power produces no more increase in output power of an amplifier. Saturation power is an important quantity because it limits the performance of optical amplifiers. Amplifiers are saturated when the incident power crosses the threshold Ps, resulting in reduced gain.

Output saturation power[2] (P_{out}; see Equations 3-2 and 3-3) is defined as the output power for which the amplifier gain is reduced by a factor of 2 (3 dB) from its unsaturated value (g_0).

Equation 3-2

$$P_{out} = \frac{g_0 L_n(2)}{g_0 - 2} . P_s$$

Equation 3-3

$$P_{out} = 0.7P_s$$

Amplifier Noise

Noise is inherent within an optical amplifier; it is a byproduct of optical amplification. The noise degrades the signal quantity and affects overall signal to noise ratio (SNR). A figure of merit for optical amplifiers is the noise figure (NF). *NF*, shown in Equation 3-4, can be defined as the ratio of input OSNR to the output OSNR in an optical amplifier.

Equation 3-4

$$NF = \frac{SNR_{input}}{SNR_{output}}$$

Noise in amplifiers is essentially due to spontaneous emission. Spontaneous emission in amplifiers is a byproduct of the gain media, whereby the optical amplifier emits electrons that abruptly fall from various levels, emitting uncontrolled light of random phase and frequency distribution. The emitted light is noise or random perturbations and is called *amplified spontaneous emission* (ASE)[4]. ASE in an optical amplifier is a serious cause of concern that severely affects system performance; ASE impairment can create ripples in the power budget. Currently, it is impossible to suppress ASE due to the continuous nature of ASE signal throughout the entire spectral width of the WDM channels. ASE rejection filters can filter ASE outside the WDM band or even the operating wavelength. A figure of merit for ASE is the spontaneous emission factor or population inversion factor shown in Equation 3-5 and denoted by n_{sp}.

Equation 3-5

$$n_{sp} = \frac{N_2}{N_2 - N_1}$$

In the equation, N_2 is the volume of electrons in an excited state, and N_1 is the volume of electrons in the ground state. Because ASE noise cannot be reduced, it accumulates in sequential amplifiers, with each amplifier stage in a WDM system adding its own component of noise in addition to amplifying the noise that is already in the signal due to previous stages. The end result is catastrophic degradation of the signal, which causes SNR to drop and transmission to be limited to a fixed number of amplifier stages.

Doped Fiber Amplifiers

By doping fiber with certain elements, the behavior of the fiber waveguide transforms to a gain media. The dopant is generally a rare earth element, such as Erbium or Pracydymium. EDFAs or Erbium doped fiber amplifiers have been the most popular because of their operation in the 1550 nm C band. Pracydymium-based amplifiers work in the 1310–1380 nm region. Doped fiber amplifiers are analogous to a reservoir and pump system, whereby a pump signal transmits its energy onto an incoming signal through the doped atoms as an intermediate media. Atoms of the doped element absorb the local pump and rise to a higher level. This transfer of energy from the pump to the incoming signal depends on the atomic transition frequency ω_0. As the incoming frequency approaches ω_0, the gain of the system increases until it reaches a peak value governed by Equation 3-1.

EDFA—Erbium Doped Fiber Amplifier

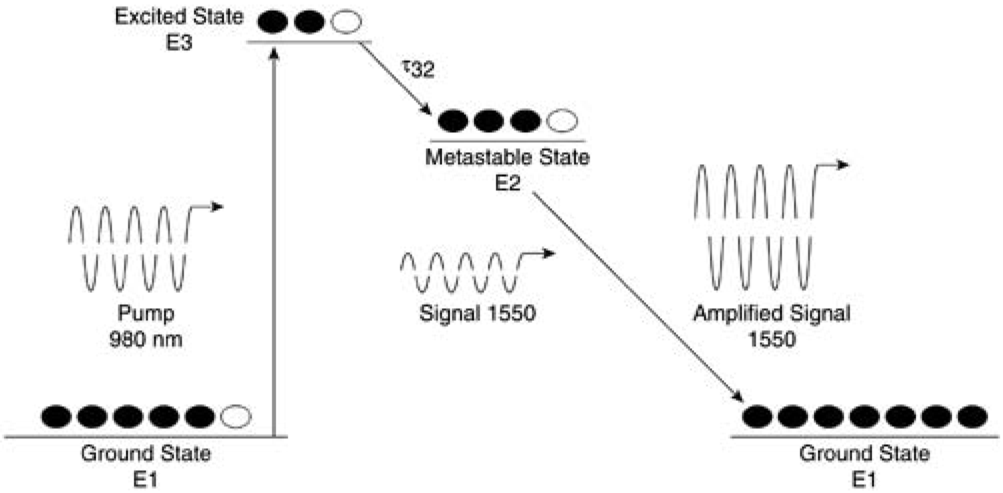
An EDFA is a doped fiber amplifier that is functional in the C band and the dopant used is Erbium ions. With characteristics of a moderately high (yet uneven) gain spectra, doped fiber amplifiers exhibit optical gain because of stimulated emission of the higher excited state. In principle, a doped fiber amplifiers such as EDFA depicts three energy levels.

Upon application of an optical pump (1480 nm/980 nm continuous wave), electrons in the stable ground state-E_1 absorb quantities of energy and rise to a meta-stable state of energy-E_2. The existence of such a state is governed by two material factors: the atomic transition frequency-w_0 corresponding to the longest lifetime state[2], and the transition wavelength on which a major chunk of pump energy is transitioned. The transition wavelength is in the strategic 1550 nm band; therefore, EDFAs are popular in commercial WDM networks.

EDFA Working (Operation of EDFA)

ConsiderFigure 3-2 and note the three energy levels E_1, E_2, and E_3 corresponding to the ground, excited, and meta stable states. A pump laser pumps a continuous wave signal at 1480 nm/980 nm. The ground level electrons (whose instantaneous volume = N_1) absorb this pump energy and get excited, rising to level E_3. The volume of electrons that rise to level of energy E_3 (excited state) is given by N_3. E_2 represents a meta stable state with a lifetime τ_{32} (transition time from E_3 à E_2). Electrons from E_2 drop to E_1, emitting photons at frequency ν. An incident optical signal, upon passing through such a system, is amplified by the instantaneous absorption of these photons.

Figure 3-2. Three-Level Model of EDFA Operation



Good pumping power can enhance the gain of a signal. You can achieve a pumping efficiency of 1 dB/mW with a 980 nm pump; therefore, you can achieve an overall WDM (all channels) gain of 30 dB. EDFAs can be made to operate with pumping in the same as well as opposite directions of the signal. When a pump signal is in the same direction of the WDM signal, it is known as *forward pumping* (see Figure 3-3); if a pump signal is in opposition to the WDM signal, it is known as *reverse pumping* (see Figure 3-4).

Figure 3-3. Forward Pump

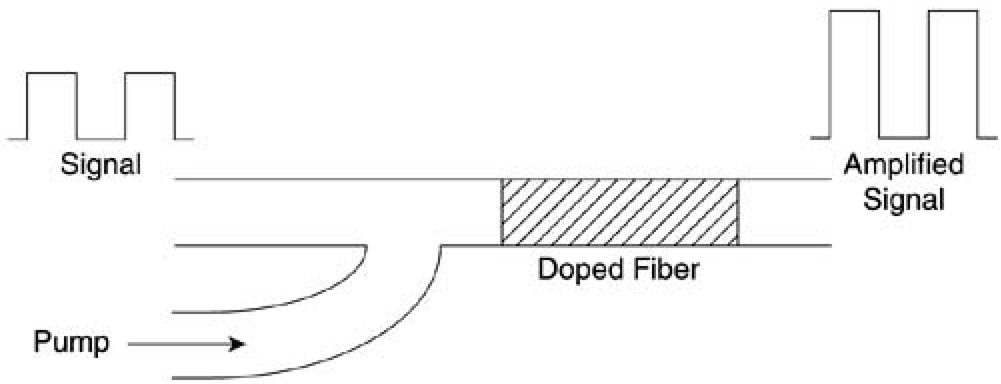
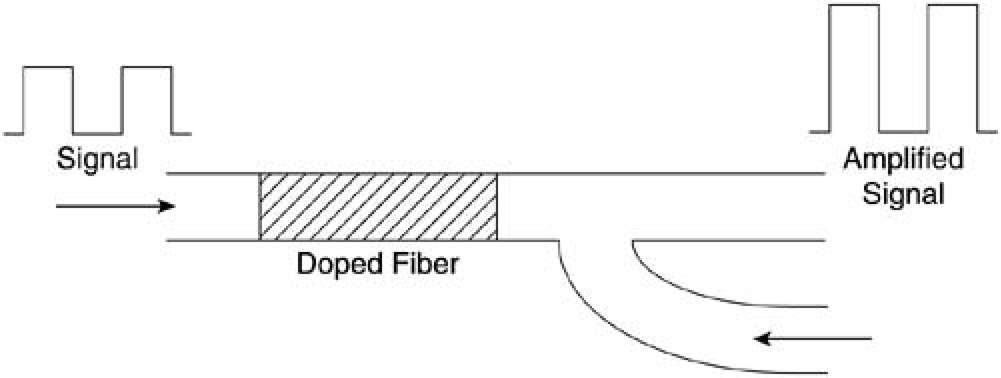
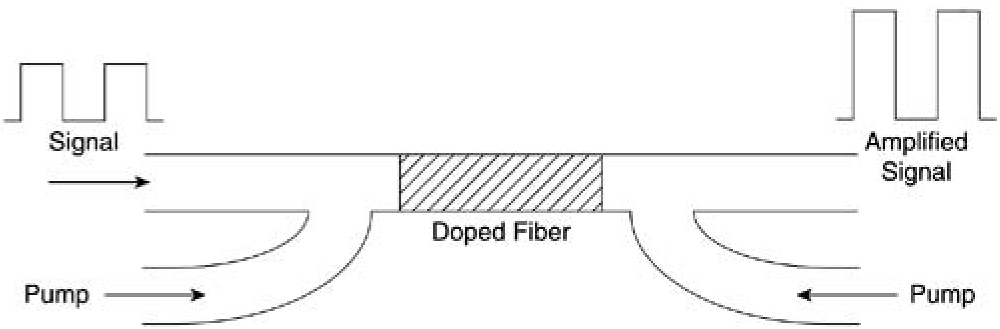


Figure 3-4. Reverse Pump



A third example is when two pumps are applied[7] and are in forward and reverse direction at the same time; this method is known as *bidirectional pumping*. Power conversion efficiency (gain per input pump power) is exhausted in a backward (reverse) pumping direction. In bidirectional pumping, the gain profile throughout the doped region is almost linear, which is an advantage. SeeFigure 3-5 for a view of bidirectional pumping.

Figure 3-5. Bidirectional Pump



The gain spectra of EDFA is quite broadened, and this broadening is proportional to the dipole relaxation time T_2 (see Equation 3-1). The gain profile in the C band exhibits a double-peaked curve; the amplification can be anywhere between 7–30 dB and depends on the pump power, relaxation time, transition frequency, and saturation power. For an amplifier, the gain increases exponentially with the applied pump power. At a certain value, the increase in gain slows down and any further increase in pump power decreases the gain.

EDFA Noise

EDFA noise is important in optical transmission systems because it severely limits the system performance[7]. The NF of an amplifier is given as NF = $2n_{sp}$, where n_{sp} is the spontaneous emission factor shown in Equation 3-6.

Equation 3-6

$$n_{sp} = \frac{N_2}{N_2 - N_1}$$

Noise figure, NF is shown in Equation 3-7.

Equation 3-7

$$NF = \frac{2N_2}{N_2 - N_1}$$

In Equation 3-7, N_2 and N_1 are the volumes of electrons in the excited and ground states, respectively. The spontaneous emission factor causes the amplifier to generate large amounts of noise (4–7 dB). This is a severe impairment for optical signal-to-noise ratio (OSNR), shown in Equation 3-8.

Equation 3-8

$$OSNR = \frac{SNR_{in}}{SNR_{out}}$$

The NF, which can be considered as a figure of merit, is proportionally dependent to the pump power and the length of the doped fiber. One hindrance in achieving good amplification is the spontaneous emission that is amplified in a reverse direction, thus depleting pump power. One solution is to use an isolator (which ensures transmission in one direction only) to prevent backward emission.

When we are considering WDM links with cascaded EDFA stages, the amplifier noise (ASE) becomes a serious issue. The end-to-end system OSNR degrades due to cumulation of noise figures (NFs) of each EDFA stage.

In a WDM system, cross saturation occurs whereby the gain of a WDM channel is saturated by either its own power or cross-talk power from neighboring channels. Cross-talk in EDFAs is not a serious impairment and by ensuring operation in an unsaturated region removes cross-talk to a great extent.

Raman Amplifier

Stimulated Raman scattering (SRS) was briefly discussed in section on Raman and Brillouin effects in Chapter 1, "Introduction to Optical Networking." SRS is a type of inelastic scattering that results in broadband amplification of optical channels. The amplifiers resulting from this effect are called *Raman amplifiers* and have a distinct feature of amplification in a large waveband[2]. Raman amplification occurs when a pump signal is made to propagate through a fiber. The pump signal is at a different (generally lower) wavelength and creates a *Stokes wave* (high amplifying wide band wave) that amplifies the many channels in a WDM system. Raman amplification occurs for multichannel WDM systems only and not for single-channel optical communication. The gain spectra for Raman amplification is quite broad (150–200 nm), thus covering entire operating bands. The final breadth of the amplification band in nanometers depends on the number and power of the pumps used. The pump signal is at a lower wavelength. Raman amplification is pumping of energy into a wide band and creating amplification by phonon-phonon (optical phonon) interaction.

Although analogous to EDFA, the working of a Raman amplifier is not based on stimulated emission. In an EDFA, a pump photon stimulates another photon and the net energy (frequency, phase, and energy) of the two photons is the amplified output. The energy of the incident photon is not lost during this process. On the other hand, in Raman amplification, the pump photon has to lose its energy to create another photon at a lower frequency (higher wavelength) and lower energy. The difference in energy creates optical phonons, which are absorbed by the medium.

Due to the absence of stimulated emission in Raman amplifiers, population inversion does not need to be achieved. Moreover, no energy transfer between energy levels is needed. The pumping can be done in forward as well as reverse directions, where the signal and pump copropagate (forward) or counter-propagate (reverse). The signal is amplified only when both signal and pump are at logical high (both are at 1 bit). Therefore, to amplify a pseudo-random bit sequence, we need a pump that is perpetually high. This might create other impairments, such as self–phase modulation (SPM) and cross-phase modulation (XPM). SPM and XPM are discussed in a design aspect in Chapter 4, "WDM Network Design -1."

Raman Spectra

As mentioned in the previous section, due to the nonlinear interaction of pump and propagating signal, a Stokes wave at a lower frequency (higher wavelength) is produced. This Stokes wave has properties of amplification by transferring energy to the propagating signal. The amplification occurs by direct energy transfer between the pump and the signal rather than by stimulated emission, as with EDFA. The spectral difference between the pump and the generated Stokes wave is called the *Raman gain spectra* $(g_r(\Omega))^2$. Moreover, the gain spectra is inversely proportional to the pump frequency, and it extends over a large range. Generally, $\Omega = f_p \times f_s$ (where f_p is the pump frequency and f_s is the Stokes frequency). Ω is approximately equal to 13 THz.

An incident signal(s) is amplified if it spectrally falls into the gain spectra Ω (the Raman gain spectra). Its gain profile is shown in Figure 3-6, and the working diagram is shown in Figure 3-7. The gain exhibited is highest when Ω = 13.2 THz. In addition, the amount of amplification obtained depends on the input pump power; after a certain threshold, $g_r(\Omega)$ increases exponentially with input pump power until it gets saturated. Raman amplification exhibits a sort of energy conservation.

Figure 3-6. Raman Gain Spectra

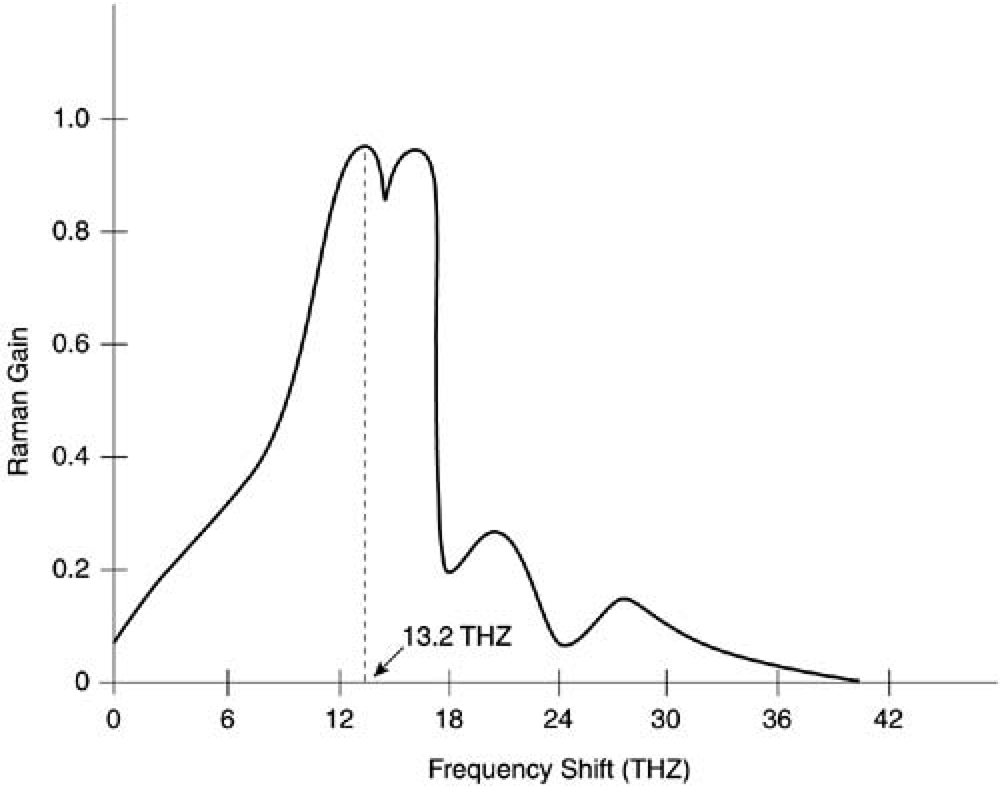
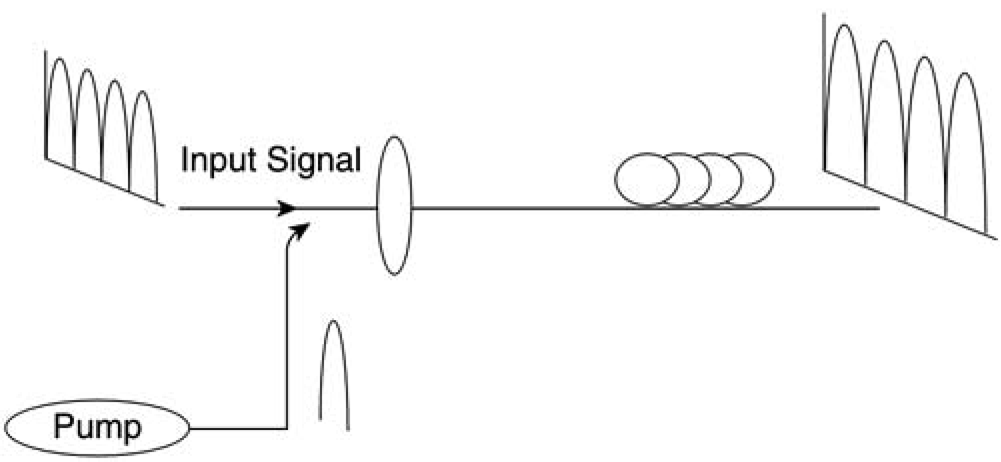


Figure 3-7. Raman Amplification



In other words, energy is transferred from a pump to a Stokes wave, and the number of photons in a pump wave and a Stokes wave remains constant. A figure of merit in the functioning of Raman amplifiers is the *Raman threshold*, defined as the amount of pump power needed to make the pump equal to the power of a Stokes' wave[2]. Raman threshold is also subject to a constraint: the pump and Stokes wave should have the same state of polarization (SOP).

If the SOP of the pump and Stokes wave is not maintained, the threshold (above which good Raman amplification can occur) increases by a factor of about 2. A typical Raman threshold for 1550 nm WDM signals is between 500–700 mW.

As the input pump reaches Raman threshold, all the energy from the pump is transferred to the Stokes wave. The power of the Stokes wave might increase to a new threshold point, whereby a second-order Stokes wave could be created. An experiment whose results were shown in the European Conference on Optical Communication 2001 in Amsterdam [PP] showed a foundation of multiple Stokes waves creating fine wide band amplification of the WDM signal.

Shown in Figure 3-7 is a typical Raman amplifier with forward pumping. Amplification occurs as a result of SRS as long as the input signal is within the gain band. The amount of gain experienced is exponential to the input pump power until the amplifier is saturated and gain begins to stabilize. Typically, Raman amplifiers can produce 20–35 dB gain with 800 mW to 1W pump power. A typical laser used is the Nd:YAG laser. SRS amplification occurs upon fulfilling three main conditions:

- Input pump power exceeds Raman threshold.
- Polarization of signal and pump are the same. If they are orthogonal, no amplification occurs. Amplification also occurs when input signal is at a logical high.
- The signal is within the Raman gain spectra.

Raman amplifiers are quite effective for WDM signals because of their broad gain profiles. The main issue of concern is the high pump power that is required. This high power can create other nonlinear impairments as well.

An important consideration in Raman amplifiers is that they have lower NFs than EDFAs. This results in better OSNR for an optical signal if we use Raman amplifiers rather than EDFAs. However, the Raman amplifier's cost, gain, and pump power requirements do not allow them to be fully exploited.

Distributed Amplification

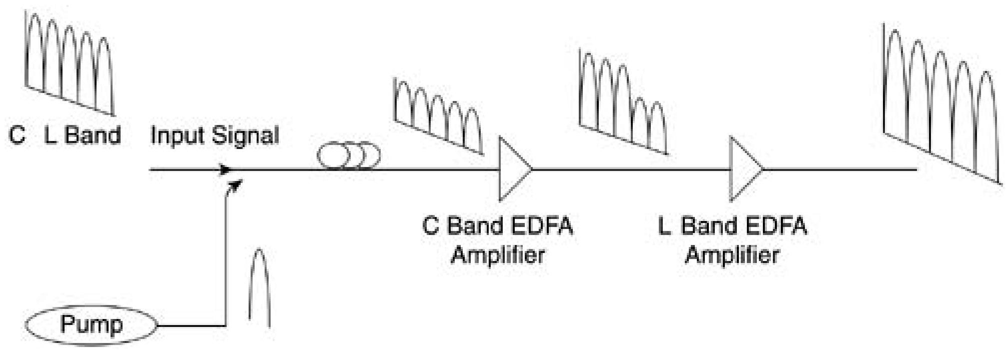
Instead of amplifying the signal at one discrete point in the network, we can amplify it at various points throughout the network. This kind of amplification is called *distributed amplification* as opposed to *lumped amplification*. Distributed Raman amplifiers (DRAs) are created by injecting pumps at various regions or by ensuring a large effective length for Raman amplifiers. The latter is done by controlling the pump power and taking care of the transients that are associated with the signal.

Distributed amplification can be achieved for both Raman amplifiers and EDFAs. One good way to achieve this is to introduce bidirectional pumping, which creates an almost linear gain profile.

Hybrid Amplification

The low noise, broad bandwidth of Raman, and low pump power requirements of EDFAs can be combined into one hybrid amplifier to solve amplification issues in long-haul and ultra long-haul networks. Raman amplifiers provide a gain across a large bandwidth, even though the gain provided might not be high. In contrast, EDFAs provide a substantial gain but across a relatively small band, and the gain provided is not flat (uneven). By using both forms in tandem—providing one Raman and one or more EDFAs (depending on the bandwidth to be amplified)—the amplification achieved is much better and cleaner (better OSNR) than individual configuration. Refer to Figure 3-8.

Figure 3-8. Hybrid Raman-EDFA Amplifiers



Advantages of Raman Amplifiers and Comparison to EDFAs

Raman amplification is an effective technique for optical regeneration (1R); it has a low NF as compared to fiber doped amplifiers. The gain profile for a particular band is also much flatter for Raman amplification as compared to Doped fiber amplifiers. The gain profile is much broader for Raman amplifiers, so they are able to accommodate more channels. Due to the low NF of Raman amplifiers, the OSNR performance is better than that for EDFAs. On an average, up to a 5 dB gain in OSNR is exhibited through Raman amplifiers. Also, due to Raman amplifiers' nondependence on carrier lifetimes of meta stable states, they can be used at higher bit rates.

Raman amplifiers have certain disadvantages as compared to EDFAs. They need high input pump power, which can cause severe nonlinear impairments. In addition, stimulated Raman scattering can cause four-wave mixing, creating harmonic peaks and severe cross-talk.

SOA—Semiconductor Optical Amplifiers

Semiconductor lasers can be made to function as optical amplifiers by removing the associated feedback. Traditionally, semiconductor lasers provide for lasing effect through stimulated emission by creating a region of population inversion between highly reflective facets (mirrors) of a cavity. A semiconductor optical amplifier can be created by removing the feedback cavity and replacing the end facets of a laser with antireflectory coatings.

A signal that is to be amplified enters the p-n junction. There is a condition of population inversion already present in the p-n junction. The signal is now amplified as it knocks off higher energy electrons (excited state) and releases photons, which theoretically should have the same phase and frequency as the incident perturbation (input signal). For optical amplification at a given wavelength or a given band of wavelengths, the p-n junction should operate under the lasing threshold. (Otherwise, it will emit its dominant wavelength. Refer to Chapter 2 for information on how lasers work.)

Two kinds of SOAs are available: the *Fabry Perot cavity-based* SOA and the *traveling wave* SOA. The Fabry Perot cavity-based SOA is similar to a distributed Bragg reflector (DBR) or a distributed feedback (DFB) laser (mentioned in Chapter 2) without feedback. Low gain and low power are the two serious limitations of these types of SOAs. The traveling wave SOA consists of a long, narrow p-n junction. Unlike the FP cavity SOA, there is no need for a cavity, and amplification occurs along the length of the SOA. Typically, SOAs are built on Indium Gallium Arsenide Phosphide (InGaAsP) substrate because of their fast electron-hole reconfiguration times.

SOAs have severe limitations, such as high cross-talk and low output power. When multiple channels are to be amplified, high cross-talk exists between them. SOAs are currently reduced to academic projects and access networks (PON and so on).

Wavelength Conversion Using SOAs

A high-power signal saturates (by depletion of excited electrons) in an SOA; in such a condition, the system has severe cross-talk. This cross-talk can be put to good use to create wavelength converters. A high-powered data signal and a low-powered continuous wave (at a different wavelength) signal are fed together to the SOA. The cross-talk in the SOA ensures that the data from the input signal is transferred to the continuous wave signal. Therefore, wavelength conversion occurs. The new output wavelength has inverted data (1 for 0 and 0 for 1). Refer to Figure 3-9.

Figure 3-9. Semiconductor Optical Amplifier in Wavelength Converter Configuration

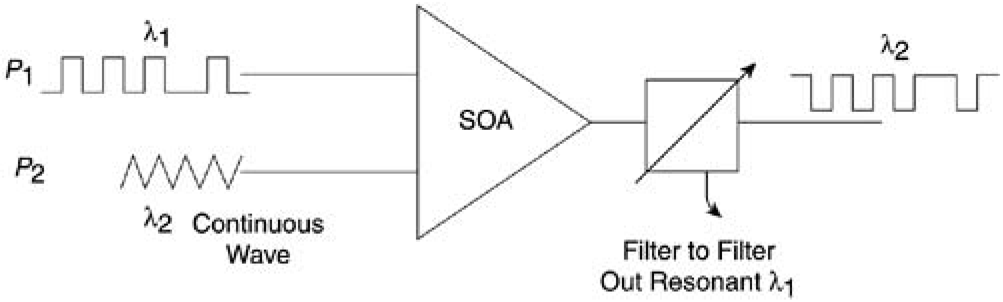


Table 3-1 compares the three basic configurations of optical amplifiers: doped fiber, Raman, and SOA. The table examines the different optical characteristics. A good designer often chooses a particular amplifier based on the requirements of the network.

Table 3-1. Amplifier Comparison

	EDFA	RAMAN	SOA
Gain	~30 dB	~20–25 dB	~ 10–20 dB
Output power	High	High	Low
Input power	Moderate	High	High
Crosstalk	Low	Low	Very high
Gain tilt	High gain tilt	Low	High
Application	Metro, long haul	Typically long haul	Short haul, single channel, wavelength converters

Dispersion Compensation Techniques: Pre- and Postcompensation

Chapter 1 discussed dispersion and its analysis in optical fibers. Dispersion limits the transmission length of an optical link especially at high data rates. Methods are available to compensate for dispersion. This section discusses dispersion compensation techniques and their evolution.

Postcompensation

We can deploy electronic techniques for postcompensation of chromatic dispersion[3]. It is possible to design a receiver (photodetector) with adjoining circuitry, which can reduce the amount of dispersion that a signal undergoes. The main idea behind using a detector circuit to compensate for accumulated dispersion is that a signal that is dispersion impaired can be recovered by a system (in this case a detector circuit) whose response is linear. Group velocity dispersion (GVD) though not entirely a linear phenomenon still can be compensated by this technique. In one embodiment, a receiver converted the optical signal to an intermediate microwave signal. A microwave bandpass filter with a linear response function treated the signal, and the output was the dispersion-corrected signal.

The condition under which this kind of compensation is successful is when we make sure that the microwave filter gets the electrical signal that is an exact replica of the optical signal, corresponding to variations in both amplitude and phase. Most detectors are based on intensity logic detection; therefore, their output is a string of 1s and 0s whose existence depends on the cumulative value of the optical intensity in that clock pulse.

The electrical output of the detectors is generally independent of the phase fluctuations in the optical signal, leading to dispersion-based losses (BER degradation). Some new emoluments that have been proposed deal with nonlinear functions, which can compensate the bit pattern at the receiver. In one such scheme, the decision of whether a bit is a 1 bit or a 0 bit is taken considering the status of the past few 1 and 0 bits, and the characteristics of this bit are compared to some of the characteristics of the preceding bits.

In another scheme, multiple independent decision-making logic circuits look at the signal and decide the status of each evaluating circuit. Overall, post compensation is an attractive scheme for slow signals (less bit rate) and short distances. The need for having electronic circuits that are much faster than the optical line rate makes deployment slightly hard beyond 2.5 Gbps. (Much of the work on postcompensation is explained in Chapter 9, "Tests and Measurements.")

Precompensation

Precompensation is quite the reverse of postcompensation. In precompensation, the pulses are compensated for dispersion[2] even before they are transmitted through the fiber. The pulses are chirped (see Chapter 2) in such a way that the effect of the fiber channel is not sufficient to disperse the pulse out of its intended time slot. This scheme is quite intuitive and simple to demonstrate. Prechirped pulses are given the right amount of chirp, which usually depends on the length of the channel and the dispersion the pulse would undergo at that bit rate.

Prechirping is quite difficult in direct modulated lasers in which the chirp parameter is often negative, leading to high dispersive pulses. On the other hand, externally modulated lasers can have the right amount of prechirp to create pulses that can travel several kilometers of dispersive fiber.

Dispersion Compensation Using Fibers

Group velocity dispersion in optical fibers can be compensated easily by using high dispersion fibers. A fiber of length L_i with a dispersion of D_i ps/nm-km can be compensated by using another spool of fiber of length L_j whose dispersion parameter is D_j such that it satisfies the relationship of Equation 3-9.

Equation 3-9

$$D_i L_i + D_j L_j = 0$$

Each of the two terms gives the total amount of accumulated dispersion that a pulse would undergo after transmission through any of the two fibers. The primary aim now is to keep the total dispersion at the end of the communication channel as low (close to zero) as possible. The second fiber that is used is the high dispersion fiber, which has a strong dispersion profile. The dispersion profile of the second fiber is opposite to the dispersion profile of the first fiber, and over the length of the channel, it is quite supplementary, yielding almost zero dispersion over the entire channel.

The problem of having dispersion compensating fibers (DCFs) is the invariably high loss that the high dispersion fibers go through. The attenuation parameter α for DCF fibers is much greater than the attenuation for normal single-mode fibers (SMFs). The attenuation parameter is typically as high as 0.5–0.8 dB/km as compared to 0.2 dB/km for standard 1550 nm SMFs. A figure of merit[8] for such DCF fibers is the ratio of dispersion parameter D to the attenuation parameterα (alpha). This figure of merit is often expressed in ps/nm-dB and is denoted by M. To compensate a channel that is L_i in length with a dispersion parameter of D_i, we can use a high dispersive fiber of dispersion parameter D_j and a length shown in Equation 3-10.

Equation 3-10

$$L_j = -(D_i / D_j) L_I$$

Therefore, to ensure that the length of the compensating fiber is small enough, its dispersion parameter D_j has to be quite large. Several techniques are illustrated in literature that deal with methods to make efficient DCF fibers. DCF fibers are generally kept at one point in the network—typically in central offices (COs) of carriers; therefore, they are wound in spools, creating high bending losses. Moreover, a power asymmetry exists due to the coupling difference between normal fiber and DCF, causing nonlinearities as well as high insertion losses. Curbing the nonlinearity and reducing insertion loss is an ongoing area of research today.

GVD Compensation by Using Fiber Bragg Gratings

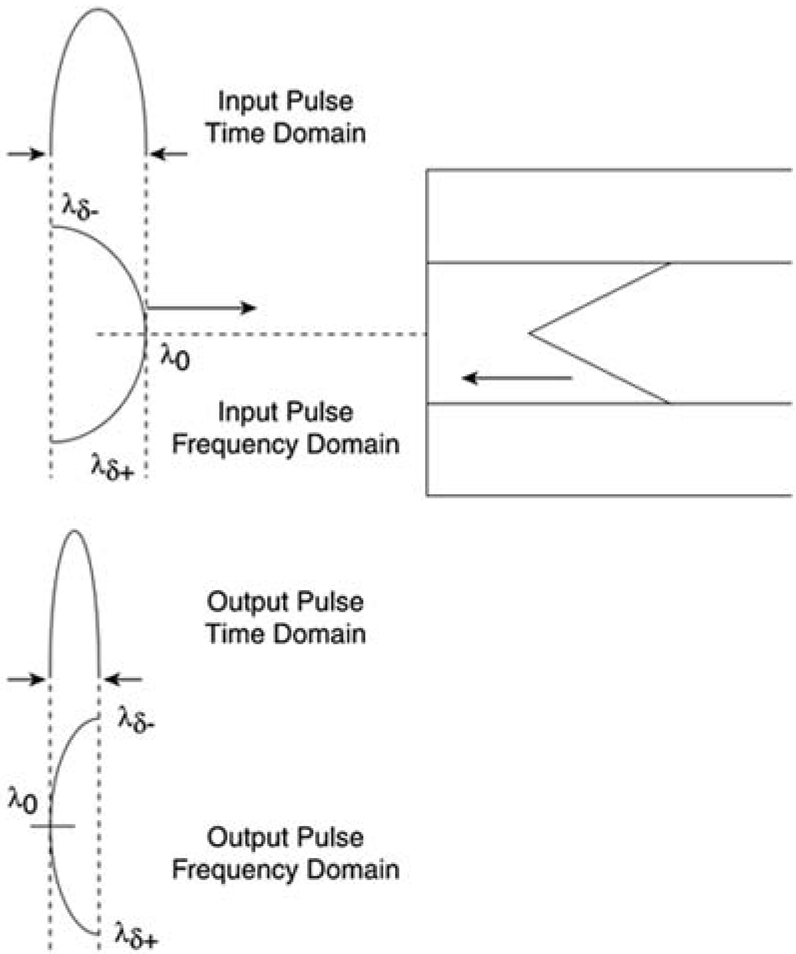
Use of Fiber Bragg gratings (FGBs) for rejection filters in WDM systems was discussed at length inChapter 2. Another use of FBGs is for dispersion compensation. Dispersion causes a signal's spectral components to spread in the time domain[8]. In other words, components at different wavelengths (frequencies) travel different distances (in the same time), and they need to be aligned together so that the pulse does not spread out of its intended slot. FBGs can accomplish this task of alignment.

FBGs are essentially devices that can act as wavelength reflectors for selective wavelengths. By varying the index profile inside a fiber, the fiber can reflect exactly those wavelengths that correspond to some particular profile changes in the core. The same principle is applied for dispersion compensation. Dispersion compensation can be accomplished using FBG by creating a grating profile of the refractive index inside the core in such a way that the faster moving components of the pulse are reflected back from a further distance inside the fiber. Whereas, the slower moving components of the pulse are reflected back from a shorter distance in the fiber. This scheme is particularly useful for single-channel optical communication.

Many such FBG-based dispersion compensators can be cascaded to provide dispersion compensation for a multichannel WDM system. The main issue here is fabricating the FBG inside the fiber for various dispersion profiles. Dynamically compensating dispersion is a key for producing good FBG-based compensators. A 10 Gbps signal would have far more dispersion to compensate than a 2.5 Gbps signal would. Therefore, we have to design FBGs that can be tuned.

One such method for FBG design is by deploying holographic technology. A hologram can be created inside the core of the fiber, which essentially acts like an index grating. This hologram can also be made to change the gain profile as and when required. This scheme is quite efficient and easy to use, but costs are relatively high. Figure 3-10 shows the working of FBGs in fibers to compensate for GVD.

Figure 3-10. Schematic Diagram of Dispersion Compensation in Fibers Using FBG Technology


Input Pulse
Time Domain
λδ-
λ0
Input Pulse
Frequency Domain
λδ+
Output Pulse
Time Domain
λδ-
λ0
Output Pulse
Frequency Domain
λδ+

Polarization Mode Dispersion and Compensation Techniques

Polarization mode dispersion (PMD) is explained in Chapters 1 and 9; however, this section discusses the issues associated with PMD and how to compensate it in high-speed transmission systems (>10 Gbps).

Unlike chromatic dispersion, statistical dimension of PMD is due to the variations in time of the external stress on the fiber. In transmission systems, there is no guarantee of the maximum penalty due to PMD. However, for design considerations, we can choose an upper limit. (A chance called *outage probability* is always available if you go beyond this upper limit). The PMD value of fiber ps/km$^{1/2}$ is the mean value over time or frequency of the differential group delay (DGD). See the following equation.

$$(ps/\sqrt{km})$$

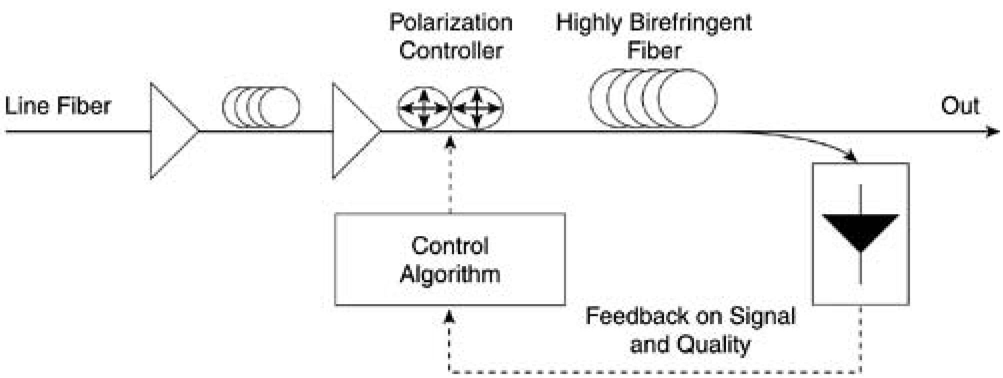
Currently installed fibers have PMD value greater than $1.5 ps/\sqrt{km}$.

The dispersion due to PMD must be compensated in long-haul systems that have bit rates greater than 10 Gbps for error-free transmission.

PMD Compensation Techniques[8]

The principle of operation of a PMD compensator is to reduce the total PMD of the fiber line plus that in the compensator. Figure 3-11 shows the basic scheme of a PMD compensator. It consists of one fixed highly birefringent element (polarization maintaining fiber), polarization controller to orient the axis of the counteractive elements and control algorithm with control circuits[8].

Figure 3-11. PMD Compensation Technique



The control algorithm is in a tracking mode and operates in blind mode-PMD conditions are not dynamically controlled; Control algorithm has to keep track of local optimum feedback signals.

This is the limitation of PMD compensators operating in an optimum fashion. Detailed analysis of local and suboptimum compensation and its effects on transmission is beyond the scope of this book.

Different concepts have been proposed for improvement of the basic PMD compensator. One of the improvements suggested is to increase the number of polarization-maintaining elements. In theory, adding more birefringent fiber segments will help the compensator address PMD issues in the link. This will also increase the number of parameters that need to be controlled. Tunability of the birefringent device is an implementation issue for practical deployments. FBGs allow variable DGDs for use in PMD compensators. Proper setting of the DGD can be accomplished using a polarization scrambler at the input of the line.

Summary

This chapter explored amplifiers for optical communication networks. Amplifiers appear to be a small part of the WDM network design, but they are probably one of the most important subsystems. Amplifiers act as 1R regenerators of data; therefore, they can increase the length of an optical transmission line. An opto-electronic regenerator would indeed be able to enhance the signal power level, but such a regenerator is needed for each channel. However, optical amplifiers can amplify the entire composite WDM signal at once without demultiplexing each channel. This is precisely the reason why WDM networking is becoming so popular. Amplifiers become a low-cost and effective means of regenerating an optical WDM signal entirely in the optical domain without retiming or reshaping it.

The disadvantage associated with optical amplifiers is that they tend to worsen the signal to noise ratio of the composite signal by adding their own noise (often called *amplified spontaneous emission*) because of some spontaneous emission.

Amplifiers are basically classified into three categories: doped fiber amplifiers, Raman amplifiers, and semiconductor optical amplifiers (also called semiconductor laser amplifiers). Doped fiber amplifiers, especially ones using Erbium ions as a dopant, have a good gain profile in the 1550 nm C band. They are easy to fabricate and have a relatively tolerable noise figure of about 5–7 dB. Raman amplifiers are traditionally the oldest known amplifiers and are based on the phenomenon of Raman's scattering of light, which dates back to the 1930s. These amplifiers basically build up a wave at a lower frequency (compared to the signal to be amplified) called the Stokes wave, which is responsible for wide band amplification. Raman amplifiers have a low NF and provide good OSNR. Semiconductor optical amplifiers are more recent innovations and act as semiconductor lasers without feedback. They are being investigated at this time. Due to their low amplification and high cross-talk, semiconductor optical amplifiers appear potential candidates only for metro-access and metro collector networks.

This chapter discussed some of the dispersion compensating schemes for both chromatic dispersions and PMD. Precompensation and postcompensation are two of the most fundamental schemes to reduce GVD. FBG-based schemes are comparatively recent innovations for dispersion compensation. PMD compensation is an issue of tremendous ongoing research.

References

[1]Agrawal, Govind P. *Fiber-Optic Communication Systems*, Second Edition. New York: Wiley Interscience, 1997.

[2]Keiser, G. *Optical Fiber Communications*, Third Edition. McGraw-Hill, 2000.

[3]Senior, John M. *Optical Fiber Communications*. Prentice Hall, 1992.

[4]Agrawal G. *Non-Linear Fiber Optics*, Second Edition. Academic Press, 1995.

[5]Lee, T., ed. *Current Trends in Optical Amplifiers and Their Applications*. World Scientific Publishers.

[6]Kinoshita, Susumu, ed. Proceedings of OAA. Torino, Italy, 2001.

[7]Shimada, S. and H. Ishio. *Optical Amplifiers and Their Applications*. John Wiley and Sons, 2000.

[8]Proceedings of Symposium on Polarization Mode Dispersion. Venice, Italy, 2002.

Chapter 4. WDM Network Design -1

Introduction to Optical Design

A network planner needs to optimize the various electrical and optical parameters to ensure smooth operations of a wavelength division multiplexing (WDM) network. Whether the network topology is that of a point-to-point link, a ring, or a mesh, system design inherently can be considered to be of two separate parts: optical system design and electrical or higher-layer system design. To the networking world, the optical layer (WDM layer) appears as a barren physical layer whose function is to transport raw bits at a high bit rate with negligible loss. Most conventional network layer planners do not care about the heuristics of the optical layer.

However, such lapses can often be catastrophic. Until the bit rate and the transmission distance is under some bounded constraint (for example, small networks), it is often not important to consider the optical parameters.

However, as the bit rate increases and transmission length increases, these optical parameters have the capability of playing truant in the network. A network planner must consider the affecting parameters and build a network that accommodates the impairments caused by the optical parameters. This chapter explores some of the design constraints involved in the WDM network design.

Consider an optical signal as a slowly varying signal of amplitude A (τ, t) (function of distance 'τ' and time 't'), on which various parameters are acting at all times. An optical signal, as discussed inChapter 1, "Introduction to Optical Networking," propagates through a silica fiber with propagation constant β, whose value is obtained from the solution of the basic wave equations. Further, this optical signal is subjected to attenuation, which by virtue of itself, is a property of the propagating medium—silica fiber in this case. Attenuation in a fiber is characterized by the attenuation constant α, which gives the loss (in dB) per traveled km.

Why is attenuation parameter important? First, common thinking says that if the total accumulated attenuation is greater than the signal input launch power P_{in}, a signal will not exist at the receiving end. This, although colloquial, is an important issue for verifying signal reception at the receiving end of a communication channel. Second, for optical communication to happen, a receiver (essentially a photodetector, either a PIN or APD type) needs a minimum amount of power to distinguish the 0s and 1s from the raw input optical signal.

The minimum power requirement of the receiver is called the *receiver sensitivity*, R, and is covered in Chapter 2, "Networking with DWDM -1." Here, we must ensure that the transmit power is high enough so that it can maintain signal power > R at the receiver end, despite the attenuation along the transmission line. That does not mean that if we increase the transmit power to a high level, we can send bits across great distances. High input power also is a breeding ground for impairments (nonlinearities such as cross-phase modulation [XPM], self-phase modulation [SPM], four-wave mixing [FWM] and so on). In addition, an upper limit exists for every receiver (APD type or PIN type) for receiving optical power. This is given by the dynamic range of the receiver, and it sets the maximum and minimum power range for the receiver to function. For example, –7 dBm to –28 dBm is a typical dynamic range of a receiver. Therefore, the maximum input power that we can launch into the fiber is limited. This also limits the maximum transmission distance, L. If P_{inMax} is the maximum input power, the transmission distance is L, and P_r is the minimum receiver power; then Equation 4-1 shows the maximum input power that can be sent into the fiber and Equation 4-2 shows the maximum transmission distance.

Equation 4-1

$$P_{in_{max}}(dB) = \alpha L + P_r(dB)$$

Equation 4-2

$$L = \frac{P_{in} - P_r}{\alpha}$$

> **NOTE**
>
> The optical power at the receiver end has to be within the dynamic range of the receiver; otherwise, it damages the receiver (if it exceeds the maximum value) or the receiver cannot differentiate between 1s and 0s if the power level is less than the minimum value.

For an input power of +5 dB and a receiver sensitivity of –20 dBm at 1550 nm, the maximum transmission distance without amplification is shown in the following equation. (Assume α = 0.2 dB/km at 1550 nm. We usually get α from the manufacture's spec.)

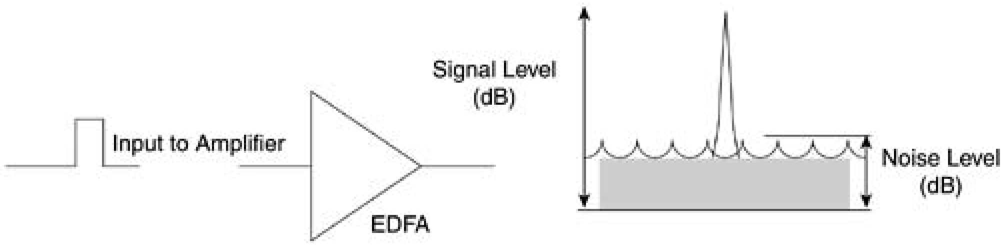
$$L = \frac{5-(-20)}{0.2} = 125\text{ km. (Here we neglected all other losses.)}$$

Further in the preceding calculation, we have neglected dispersion, fiber nonlinearities, polarization, spectral broadening, chirp (source broadening), fiber plant losses (connecters, splices, and aging factors), and so on. If we consider these effects, then the maximum length is reduced further. How can we then have ultra long-haul intercontinental systems? By placing repeaters in cascade, we can enhance the transmission distance.

Two kinds of repeaters exist: opto-electro-opto (OEO) electrical repeaters that detect, reshape, retime, and retransmit (3R) the signal (channel-by-channel), and the fiber amplifiers (1R)(doped fiber, Raman, and SOA) that boost the signal power level (no reshape and no retiming) entirely in the optical domain. A third technique also exists: reshape and reamplify (2R) regeneration. This technique is gaining in popularity due to its protocol independence. This book discussed 2R in Chapter 2, so it is not necessary to consider it here from a design perspective because it proposes a generic alternative to the other design schemes.

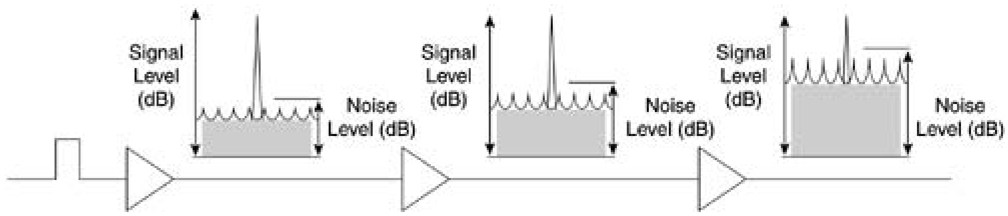
Electrical repeaters have an advantage in that they can completely relaunch the signal by regenerating and further retransmitting it due to opto-electronic conversion and regeneration. To do so, the composite WDM signal needs to be fully demultiplexed, which is neither cost effective nor efficient. Optical amplifiers alleviate that problem by amplifying all the channels together completely in the optical domain; therefore, optical amplifiers can enhance the transmission distance. So, does that mean that optical amplifiers can increase the amplifying distance as much as they wants? Not really! Amplifiers come at a price and induct a trade off; they enhance the signal power level, but at the same time, they add their own complement of noise. This noise is amplified spontaneous emission (ASE), which was introduced in Chapter 3, "Networking with DWDM -2." Please refer to Figure 4-1.

Figure 4-1. Single Stage Amplifier and Noise Associated with Signal



The noise is random in nature, and it is accumulated at each amplification stage. Refer to Figure 4-2.

Figure 4-2. Noise Accumulation Resulting from Multistage Amplification



Amplifier noise is a severe problem in system design. A figure of merit here is the optical signal-to-noise ratio (OSNR) requirement of the system. The OSNR specifies the ratio of the net signal power to the net noise power. It is a ratio of two powers; therefore, if a signal and noise are both amplified, system OSNR still tells the quality of the signal by calculating this ratio. System design based on OSNR is an important fundamental design tool.

NOTE

OSNR is not just limited to optical amplifier-based networks. Other active and passive devices can also add noise and create an OSNR-limited system design problem. Active devices such as lasers and amplifiers add noise. Passive devices such as taps and the fiber can add components of noise. In the calculation of system design, optical amplifier noise is considered the predominant source for OSNR penalty and degradation. That does not imply unimportance to other sources of OSNR penalty.

Figures 4-1 and 4-2 shows the effect of noise on signal as the signal and noise pass through the amplifiers.

Dispersion, mentioned in Chapter 1, causes pulse spreading. The most important form of dispersion is group velocity dispersion (GVD). Group velocity is inversely proportional to the rate of change of propagation constant β with respect to frequency. Isn't β a constant? Not really! β

actually (indirectly) depends on γ, the nonlinear coefficient, and P, the power of the signal. β further depends on the group index, which in turn depends on the GVD parameters. Therefore, dispersion causes severe pulse spreading and leads to intersymbol interference (ISI). The GVD parameter β_2 is the second order differential of the β with respect to change in optical frequency-omega.

Techniques are available to compensate dispersion. Note here that the dispersion discussed so far (GVD) is called chromatic dispersion as opposed to other forms of dispersion, such as polarization mode dispersion, or PMD. (PMD is discussed from a design point of view in Chapter 5, "WDM Network Design -2".) Dispersion flattened or shifted fibers are an example. Dispersion shifted fibers (DSFs) have the zero-dispersion wavelength shifted into the operating band. Further dispersion-compensating fibers can be placed at strategic locations in a network so that we can reshape the broadened pulse as desired. Yet another technique is to use fine fiber Bragg gratings (FBGs)-based dispersion compensators. The question still remains: In a network, where do we place the dispersion compensators? Dispersion compensation is needed only for signals above a certain bit rate.

Another design issue is polarization. Assuming fibers to be polarization preserving is not a good idea. Different polarization states create different levels of PMD. PMD compensation and placement is yet another strong issue at high bit rate signals.

Last but not least, we need to consider fiber nonlinearities. Self-phase modulation and cross phase modulation are two common coupling problems. FWM, stimulated Raman scattering (SRS), and stimulated Brillouin scattering (SBS) are also high bit rate, high power issues.

A system design can be optimized by considering these effects in a strategic manner. The sections that follow consider several steps that need to be considered in an ideal system design case. Initially, let's assume a point-to-point link and then specifically look at ring and mesh networks in Chapter 5.

Factors That Affect System Design

Initially, fiber loss was considered the biggest factor in limiting the length of an optical channel. However, as data rates grew and pulses occupied lesser and lesser time slots, group velocity dispersion and nonlinearities (SPM, XPM, and FWM) became important considerations.

As we will see in the following sections, an optical link is designed by taking into account a figure of merit, which is generally the bit error rate (BER) of the system. For most practical WDM networks, this requirement of BER is 10^{-12} (~ 10^{-9} to 10^{-12}), which means that a maximum one out of every 10^{12} bits can be corrupted during transmission. Therefore, BER is considered an important figure of merit for WDM networks; all designs are based to adhere to that quality.

InChapter 2, we saw the analytical explanation behind BER. It showed BER to be a ratio of the difference of high and low bit levels (power) to the difference in standard deviation of high and low bit levels. As can be observed it is quite difficult to calculate BER instantaneously.

Another plausible explanation of BER can be considered as follows. For a photodetector to detect a 1 bit correctly (assuming nonreturn-to-zero/return-to-zero, or NRZ/RZ modulation; see Chapter 2), it needs a certain minimum number of photons (N_p) falling on it. If N_{TP} is the number of photons launched at the transmitter and Δp is the number of photons lost (hypothetically) due to attenuation, absorption, scattering, and other impairments during transmission, then if $N_{TP} - \Delta p < N_p$, the receiver will not be able to decode the signals properly. To sustain good communication, it is imperative that $N_{TP} - \Delta p > N_p$ over 'L' the desired length of transmission channel. The number of photons translates to the power (which is a function of intensity) of the optical signal.

From the explanation, it becomes evident why optical system design considers power budget and power margins (safety margins for good design) so important.

As far as the dispersion issue goes, we know that dispersion is the spreading of a pulse in time domain, generally due to the large variance of the spectral domain. (Many different spectral components exist in a pulse, each travelling at a different speed.) That means dispersion causes pulse spreading.

The most harmful effect of this pulse spreading is ISI. Even if you assume ISI never to happen (due to good design), still a small amount of dispersion has several harmful effects. The spreading of a pulse lowers its power content, which means that Δp increases. In other words, the number of photons that will strike the photodetector decreases. Therefore, when we are considering dispersion-limited systems, we must consider a power penalty due to dispersion. This power penalty[3] can qualitatively be defined as the net loss in power because of dispersion during transmission of a signal in a dispersion affected/limited system.

Qualitatively, power penalty can also be considered as the net extra power required to pump up the signal so that it reaches the receiver (photodetector) while maintaining the minimum BER requirement of the system. Typically, the power penalty for most networks is in the range of 2–3 dB. ITU specification G957 states that this penalty should not be greater than 2 dB.

Long-Haul Impairments: Nonlinearity

By placing optical amplifiers, we can greatly enhance the power of an optical signal as it reaches the photodetector. Yet another system design consideration is the net fiber nonlinearity that is present in silica fibers. The intensity of the electromagnetic wave propagating through a fiber gives rise to nonlinearities. The refractive index has a strong nonlinear component that depends on the power level of the signal. Nonlinearity produces a nonlinear phase shift denoted by ϕ_{NL}. This is shown in Equation 4-3.

Equation 4-3

$$\phi_{NL} = \bar{\gamma} P_{in} \frac{[1 - e^{-\alpha L}]}{\alpha}$$

In Equation 4-3, $\bar{\gamma}$ is the nonlinear coefficient that is denoted by Equation 4-4.

Equation 4-4

$$\gamma = \frac{n_2 \omega_0}{c A_{eff}}$$

In Equation 4-4, n_2 is the cladding index, and A_{eff} is the area of cross-section of the core. Further ϕ_{NL} being dependent of P_{in} such that P_{in} by itself is a time varying response.

Therefore, the nonlinear phase shift induced in a fast-moving optical pulse is quite dynamic.

The implication is that a frequency chirp is associated with this phase shift. In other words, a pulse at frequency ω_0 would, in time, have components in the frequency range shown in the next equation.

$$\omega_0 \pm \phi_{NL}$$

In the equation, ϕ_{NL} is dynamic. The result is pulse spreading, which is a result of the power dependence on the induced phase shift. Therefore, to keep a check on the maximum phase shift that a pulse can have, it is imperative to set a threshold to the maximum input power. This nonlinear phase shift is Self Phase Modulation (SPM). In optical communication, lightpaths need to be designed, keeping in mind the maximum tolerable phase shift $\phi_{NL} < 1$. Therefore, the maximum power $[P_{in\ max} \mid \phi_{NL} < 1]$ can limit phase shift to less than the system requirement.

SPM does not act alone. In optical communication, GVD and SPM often go hand in hand, acting quasi-simultaneously over a length of the fiber. The input channel power needs to be optimized so that it ensures a net dispersion (at a given bit rate) that is less than the minimum tolerable dispersion, as well as for which the net nonlinear effects are under control. In other words, a tradeoff is involved: We need some more power to take care of the dispersion-induced power penalty, but this additional power leads to fiber nonlinear effects (such as SPM), which creates more spread.

An optimization technique involves simulation, whereby we can correctly design the network by considering all the affecting factors and using the appropriate entities to com-pensate for these factors. In one method, SPM and GVD are both calculated on a split Fourier transform. Using this method, frequency domain analysis of the two effects is accomplished by breaking the cylindrical waveguide (fiber) into infinitesimal overlapping segments, such that SPM is assumed to act on odd segments and GVD is set to act on even segments. The final sum of effects on the last and penultimate segments gives the net impairment in the system.

So far, nonlinearities have been considered on just one channel. What happens when we have a WDM system? Do parallel channels have an effect on each other? Two or more channels have nonlinear effects on each other: XPM and FWM. Cross phase modulation results from the different carrier frequency of independent channels, including the associated phase shifts on one another. Cross phase modulation is severely harmful and is twice as powerful as Self phase modulation. The induced phase shift is due to the "walkover" effect, whereby two pulses at different bit rates or with different group velocities walk across each other. The slower pulse sees the walkover and induces a phase shift because of this walkover effect. The total phase shift depends on the net power of all the channels and on the bit output of the channels. Maximum phase shift is produced when two 1 bits walk across each other due to the high power in both the bits (as opposed to the lower power levels when both bits are not at logical 1).

Mathematically, the phase shift is shown in Equation 4-5.

Equation 4-5

$$\phi_i^{NL} = \gamma\left[\frac{1-e^{-\alpha L}}{\alpha}\right]\left[P_i + 2\Sigma_{\substack{K=1\\K\neq 1}}^{w} P_k\right]$$

InEquation 4-5, w is the total number of channels and P $_k$ is the power of the kth channel. The maximum phase shift (for all 1 bits) is shown in Equation 4-6.

Equation 4-6

$$\phi_{\max}^{NL} = \left[\frac{\bar{\gamma}}{\alpha}\right][2w-1]P_i$$

NOTE

Another method is to solve Schrödinger's nonlinear propagation equation shown in the next equation.

$$\frac{\partial A}{\partial z} + \frac{i}{2}\beta_2\frac{\partial^2 A}{\partial t^2} - \frac{1}{6}\beta_3\frac{\partial^3 A}{\partial t^3} = i\gamma|A|^2 A - \frac{\alpha}{2}A$$

Effect of Chromatic Dispersion on Transmission Length and Induced Power Penalty

Group velocity dispersion (chromatic dispersion) is a primary cause of concern in high bit rate (> 2.5 Gbps) single-mode WDM systems. As explained previously, dispersion in an optical pulse creates pulse broadening such that the pulse spreads into the slots (in the time domain) of the other pulses. This not only causes ISI, but it also introduces a power penalty, which can cause degradation of the system's SNR. As shown in the next section, OSNR is a true figure of merit for optical communications. The power penalty due to dispersion is shown in Equation 4-7.

Equation 4-7

$$Penalty_{DISP} = \frac{10\log\frac{\sigma}{\sigma_0}}{\sqrt{1+\left(D_L\frac{\sigma_\lambda}{\sigma_0}\right)^2}}$$

In the preceding equation, is the spectral width and is the pulse width. Further for SMF fibers, dispersion parameter D = 17 ps/Km – nm. The limit on transmission distance is shown as in Equation 4-8.

Equation 4-8

$$B^2L < 16\frac{\lambda^2 D}{2\pi c} \text{ or } L < \frac{16}{B^2}\frac{\lambda^2 D}{2\pi c}$$

Therefore, the following equation is true.

$$L < \frac{K}{B^2}$$

K is a constant; therefore, as B increases, L decreases with the square root of B.

Design of a Point-to-Point Link Based on Q-Factor and OSNR

To design a network, it is imperative to comply the system design with the BER require-ment of the network. If one carefully considers the preceding criteria, it should be evident that calculating BER instantaneously is an intriguing task given that a designer has tools such as a spreadsheet and calculator. Chapter 2 briefly discussed the Q-factor of an optical signal. The Q-factor provides a qualitative description of the receiver performance because it is a function of the signal to noise ratio (optical). The Q-factor suggests the minimum SNR required to obtain a specific BER for a given signal. Figure 4-3 shows the relationship of Q-factor to BER. As we can see, the higher the value of Q-factor, the better the BER.

Figure 4-3. The Relationship of Q to BER

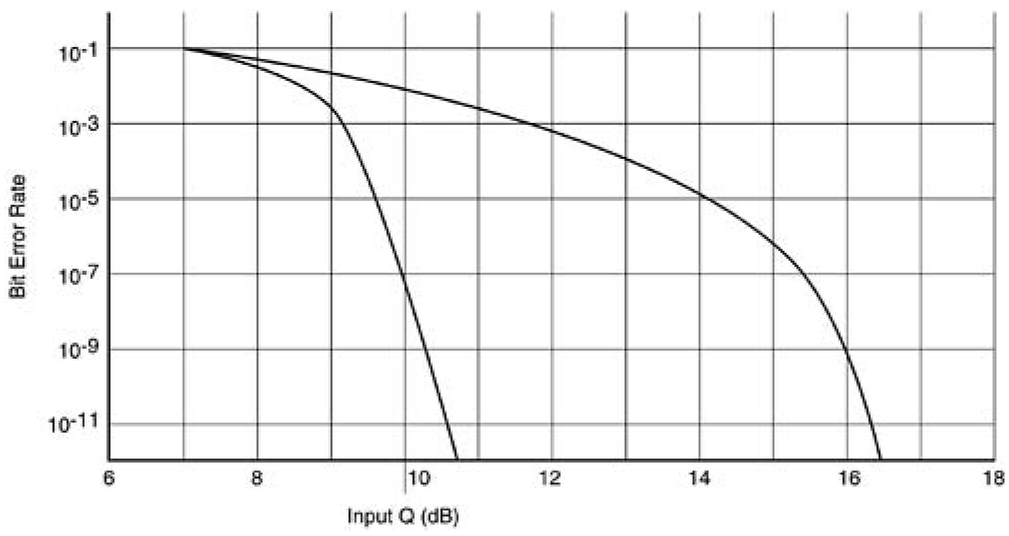
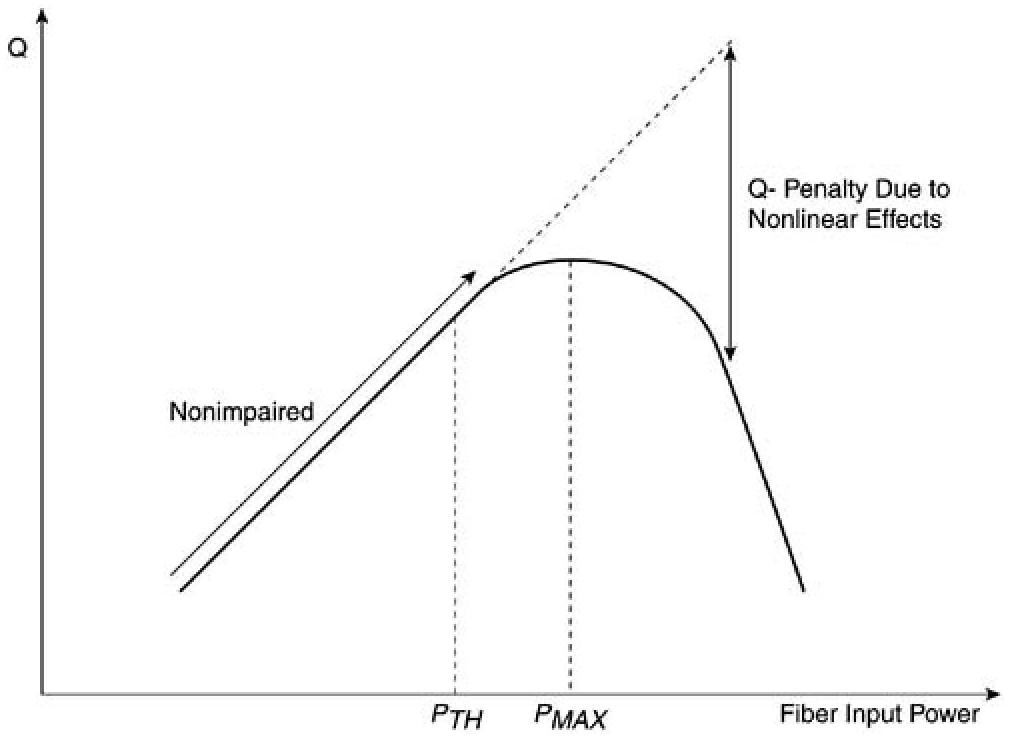


Figure 4-4 shows the penalty of the Q-factor due to nonlinear effects by increase in input power.

Figure 4-4. Q-Factor Penalty



Mathematically,Equation 4-9 gives the Q-factor of an optical signal.

Equation 4-9

$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0}$$

InEquation 4-9, I_1 is the value of the 1-bit current, I_0 is the value of the 0-bit current, σ_1 is the standard deviation of the 1-bit current, and σ_0 is the standard deviation of the 0-bit current. The relationship of Q-factor to BER is shown in Equation 4-10.

Equation 4-10

$$BER = \frac{1}{2} erfc\left(\frac{Q}{\sqrt{2}}\right)$$

A note on error function is provided in Chapter 2.

Why is BER difficult to simulate or calculate? For a given design at a BER (such as 10^{-12} and a line rate of OC-3, or 155 Mbps), the network would have one error in approximately 10 days. It would take 1000 days to record a steady state BER value. That is why BER calculations are quite difficult. On the other hand, Q-factor analysis is comparatively easy. Q is often measured in dB. The next question is how to dynamically calculate Q. This is done from OSNR.

Calculation of Q-Factor from OSNR

The OSNR is the most important parameter that is associated with a given optical signal. It is a measurable (practical) quantity for a given network, and it can be calculated from the given system parameters. The following sections show you how to calculate OSNR. This section discusses the relationship of OSNR to the Q-factor.

The logarithmic value of Q (in dB) is related to the OSNR by Equation 4-11.

Equation 4-11

$$Q_{dB} = 20\log\sqrt{OSNR}\sqrt{\frac{B_0}{B_c}}$$

In the equation, B_0 is the optical bandwidth of the end device (photodetector) and B_c is the electrical bandwidth of the receiver filter.

Therefore, Q(dB) is shown in Equation 4-12.

Equation 4-12

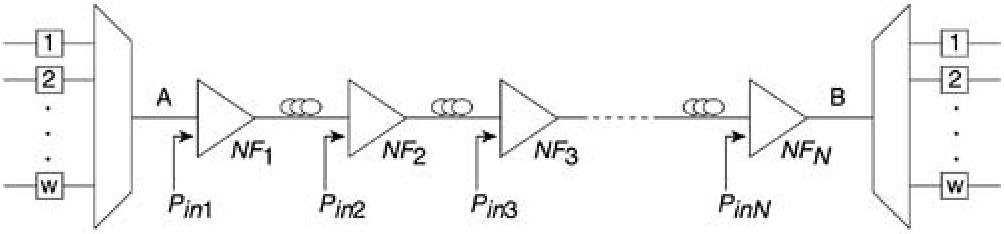
$$Q_{dB} = OSNR + 10\log\frac{B_0}{B_c}$$

In other words, Q is somewhat proportional to the OSNR. Generally, noise calculations are performed by optical spectrum analyzers (OSAs) or sampling oscilloscopes, and these measurements are carried over a particular measuring range of B_m. Typically, B_m is approximately 0.1 nm or 12.5 GHz for a given OSA. From Equation 4-12, showing Q in dB in terms of OSNR, it can be understood that if $B_0 < B_c$, then OSNR (dB)> Q (dB). For practical designs OSNR(dB) > Q(dB), by at least 1–2 dB. Typically, while designing a high-bit rate system, the margin at the receiver is approximately 2 dB, such that Q is about 2 dB smaller than OSNR (dB).

Calculation of OSNR for a Point-to-Point Link

Consider a physical link AB, as shown in Figure 4-5. Assume this to be a long-haul fiber WDM link (a link that is several hundred kilometers). Amplifiers are placed periodically at repeated intervals to boost signal power. Therefore, a signal can reach much farther than the maximum allowable accumulated loss due to the fiber ($\propto$L). However, in doing so, each amplifier stage adds its own component of amplified spontaneous emission (ASE) noise and degrades the OSNR further. Moreover, every amplifier amplifies the already present noise. Note that noise is omnipresent throughout the spectra and almost impossible to be removed. Therefore, it is imperative to devise a method to calculate the OSNR (output) at the end of an N stage-amplified system and see if the value N is still valid.

Figure 4-5. A Multiple Stage Amplified DWDM System Deployed in a Point-to-Point Topology



In an OSNR-based design, we must ensure that OSNR of the final stage is in compliance with system OSNR requirements and hence the BER requirements. To make the system support a particular BER, it is necessary to make the OSNR system design compliant.

The OSNR of each stage is shown in Equation 4-13.

Equation 4-13

$$OSNR = \frac{P_{in}}{NF_{stage} h\nu \nabla f}$$

In Equation 4-13, NF_{stage} is the noise figure of the stage, h is Plank's constant (6.6260 x 10^{-34}), ν is the optical frequency 193 THz, and Δf is the bandwidth that measures the NF (it is usually 0.1 nm).

The total OSNR for the system can be considered by a reciprocal method and is shown in Equation 4-14.

Equation 4-14

$$\frac{1}{OSNR_{final}} = \frac{1}{OSNR_1} + \frac{1}{OSNR_2} + \frac{1}{OSNR_3} \cdots\cdots\cdots\cdots \frac{1}{OSNR_N}$$

for the 'N' stage system. That summarizes to Equation 4-15.

Equation 4-15

$$\frac{1}{OSNR_{final}} = \sum_i \frac{1}{OSNR_i}$$

A slight detailed analysis provides a more appropriate equation for OSNR. For a single amplifier of gain G, the OSNR is shown in Equation 4-16.

Equation 4-16

$$OSNR = \frac{P_{in}}{P_{ASE}} = \frac{P_{in}}{2n_{sp}(G-1)hv\nabla f}$$

InEquation 4-16, n $_{sp}$ is the population inversion parameter that is shown in Equation 4-17 and is the ratio of electrons in higher and lower states.

Equation 4-17

$$n_{sp} = N_2/N_2 - N_1$$

InEquation 4-17, N $_2$ is the number of electrons in a higher state and N_1 is the number of electrons in the lower state. (Refer to Chapter 2 for more details.)

The population inversion parameter is also shown in Equation 4-18.

Equation 4-18

$$n_{sp} = 0.5 \times 10^{\frac{NF}{10}}$$

For an N amplifier stage system, with each amplifier compensating for the loss of the previous span where the span loss in dB is Γ, you have the relationship for final stage OSNR in Equation 4-19.

Equation 4-19

$$OSNR_{final} = \frac{P_{in}}{NF\Gamma hr\nabla f.N}$$

Taking logarithm to the common base (10), we get Equation 4-20.

Equation 4-20

$$OSNR_{db} = 158.93 + P_{in} - \Gamma(db) - NF_{db} - 10\log N - 10\log\nabla f$$

From the previous section, we get ∇f = 0.1 nm, or 12.5 GHz. Substituting this, we get Equation 4-21.

Equation 4-21

$$OSNR_{db} = 58 + P_{in} - \Gamma(db) - NF_{db} - 10\log N$$

The following is assumed:

- The NF of every amplifier is the same. (we assume uniformity of products; therefore, NFs are the same for all amplifiers.)
- Γ is the span loss and is same. (This is a generic assumption and can be changed, as shown later in this section.)
- Noise is totaled over both states of polarization. In short, it is unpolarized noise.

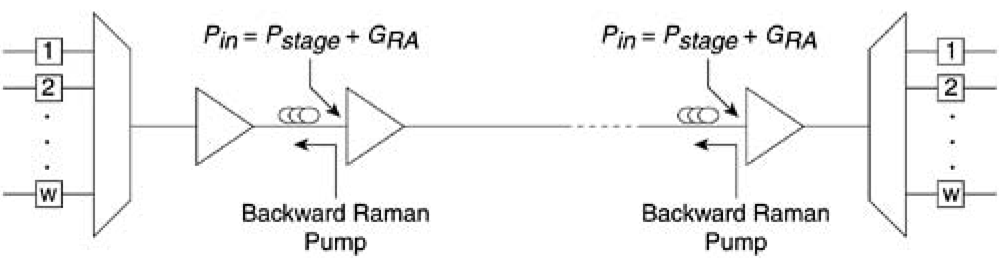
Equation 4-21 provides the actual mathematical calculation of OSNR. This calculation method has quite a few approximations in which we can still find the system OSNR to a great degree of accuracy. In a multichannel WDM system, the design should consider OSNR for the worst channel (the one that has the worst impairment). The worst channel is generally the first or last channel in the spectrum.

OSNR Improvements by Raman Amplification

If we look at Equation 4-21, we can see that the EDFA gain factor G is not considered. That is because OSNR is a ratio, and the gain acts equally on signal and noise, canceling the gain factor in the numerator and denominator. In other words, although EDFAs alleviate the upper bound on transmission length due to attenuation, by cascading EDFAs in a series, the OSNR is continuously degraded with transmission length and ASE (from EDFAs). This degradation can be lessened somewhat by distributed Raman amplifiers (DRAs). As can be seen from Chapter 3, "Networking with DWDM -2," Raman amplification is inherently a result of stimulated Raman scattering of a high intensity pump signal at a different frequency (compared to the signal frequency). This produces a gain because of creation of a Stokes wave, which in turn produces a gain feeding wave of a wide bandwidth.

Consider a hybrid system as shown in Figure 4-6.

Figure 4-6. Multistage Raman Amplifier-Based System



From the preceding system, the OSNR of the final stage is shown as in Equation 4-22.

Equation 4-22

$$\frac{1}{OSNR_{Total}} = \sum_i \frac{1}{OSNR_{i(Stage)}} + \sum_i \frac{1}{OSNR_{RA}}$$

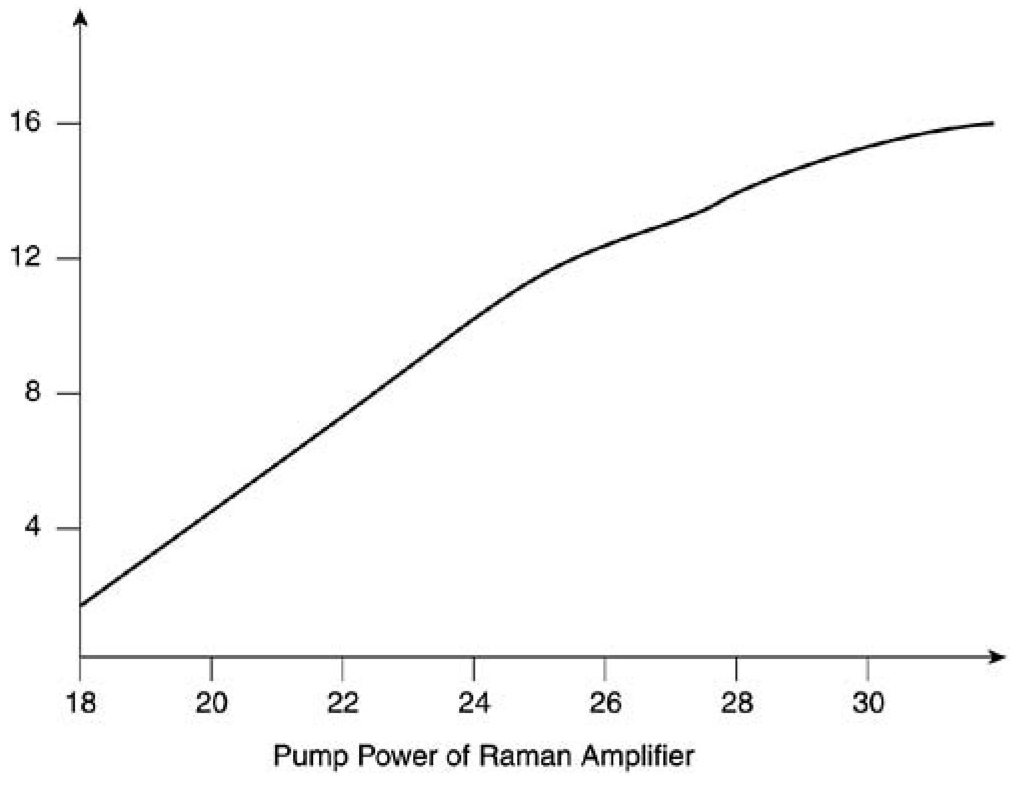
Equation 4-23 gives the OSNR value of each stage.

Equation 4-23

$$OSNR_{i(Stage)} = \frac{P_{in_i} G_{RA_{(i)}}}{NF_i h\nu \nabla f}$$

As we can see from Equation 4-23, the factor G_{RA} in the numerator actually enhances the OSNR of the system. Figure 4-7 shows the variation of Raman gain with pump power.

Figure 4-7. Variation of Raman Gain with Pump Power



Margin Requirements

In a multinode WDM link, the main component of the system loss is not attenuation due to transmission link; but instead, it is the loss associated with the various subsystems. A typical link consists of multiple nodes, each equipped with a variety of components. The loss due to each component is high, which results in a severe penalty for system design. A typical WDM node might have a full optical multiplex section (OMS) that consists of arrayed waveguides (AWGs) and a switching matrix. A typical grating-based AWG has a 5 dB loss (insertion loss) associated with it. An optical signal that is passing through a node with two such AWGs (multiplexer and demultiplexer section) is typically subject to 10 dB loss in addition to the switching fabric loss. An estimate of the loss can be understood with the following argument.

Consider two nodes, each equipped with AWGs (loss = 5 dB) and switching fabric (loss = 3dB) in addition to connector loss (2 dB). If they are separated by 50 km of SMF (α = 0.2 dB/km), the total attenuation due to transmission is 10 dB (.2 x 50). However, at each node, the loss is 5 + 5 + 3 + 2, or 15 dB. In other words, the nodal losses can be higher as in comparison to transmission losses. This affects system designs and OSNR as well. The effect is indirect in the sense that output power from a node is affected due to such losses, which further affects OSNR due to Equation 4-21.

Table 4-1 shows the insertion loss due to typical elements. We have to quantize losses due to impairments in transmission. As mentioned in the introductory section to this chapter, dispersion can be quantified as a penalty in dB. Similar treatment can be done to other phenomena such as polarization and nonlinearities and so on.

Table 4-1. Insertion Loss and Other Losses for 1550 nm Operation

Component	Insertion Loss	Wavelength-Dependent Loss	Polarization-Dependent Loss	Cross-TalkNF
Multiplexer Demultiplex (AWG)	5 dB	< 1 dB	0.1 dB	–40 dB
Optical 2 x 2 add-drop switch	1.2 dB	< 0.2 dB	0.1 dB	–40 dBm
Coupler (2 x 2) passive	3 dB	-	-	-
Filter-Thin-film	1 dB	0.1 dB	-	–40 dBm
Filter- AOTF/MZI	1 dB	0.1 dB	-	–35 dBm
Interleaver	2–3 dB	-	-	-
Optical cross-connect (OXC) Port to port	3 dB typical without AWG loss	< 0.4 dB	0.1 dB	–40 dBm

Table 4-2 presents margin requirements for a good design. These margins adhere to variations in optical signal budgeting issues, especially on a dynamic level. The margins are generally chosen by evaluating a set of readings that represent the pseudo-population of a number of discrete events governing the entire sample space of optical signal design.

Table 4-2. Margin Requirements

Symptom	Loss Margin
Fiber dispersion	1 dB
SPM margin	0.5 dB
XPM margin	0.5 dB
DCU compensation	6 dB
FWM	0.5 dB
SRS/SBS	0.5 dB
PDL	0.3 dB
PMD	0.5 dB
Amplifier gain tilt (due to nonflat gain spectra)	3.0 dB
Receiver sensitivity tilt (wavelength dependence of PMD)	0.5 dB
Transmitter chirp	0.5 dB
AWG cross-talk	0.2 dB
Fiber connectors	0.5 dB

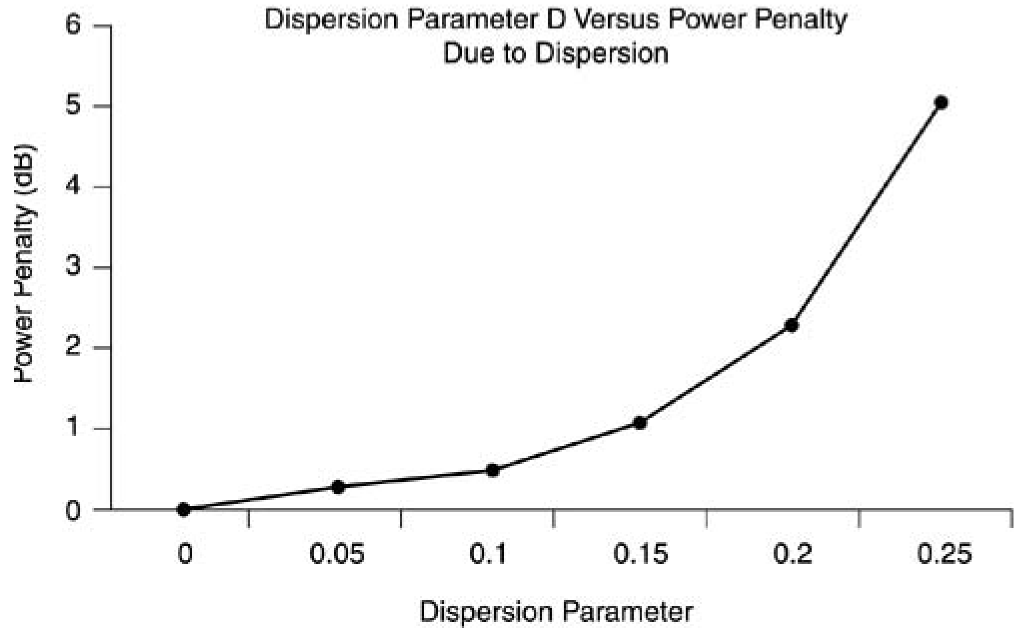
Design Using Chromatic Dispersion Compensation

In a chromatic dispersion-limited system in which the total accumulated dispersion for a traveling pulse is greater than the maximum allowable dispersion, the system cannot function because of tremendous ISI or just pure pulse spread. Therefore, we need to place dispersion compensation units (DCUs) at different positions in a network. Chapter 3 discussed some of the dispersion-compensating schemes, such as dispersion shifted fibers and FBGs, which are the most common. When we are designing a high bit rate WDM link (where dispersion can be considered a major design impairment), we should use dispersion maps to effectively design a system.

Dispersion maps are two-dimensional maps that plot the accumulated dispersion versus the length of transmission. They are particularly useful maps that help designers tell where to place dispersion compensators in a network. Accumulated dispersion is calculated by multiplying the fiber and the laser dispersion specifications for a given bit rate with respect to the length of the fiber. For example, an SMF fiber's typical value of dispersion is 16 ps/nm-km, which means that for every traversed kilometer of SMF fiber, a pulse at 10 Gbps (100 ps pulse width) spreads for about 16 ps from its mean. Ensure that the accumulated pulse spread across 'x' km is less than the maximum dispersion limit (which might be 1600 ps/km-nm for a 10 Gbps signal).

From this discussion, it is obvious that the signal can travel 16x = 1600 km (if x = 100) of SMF fiber at a 10 Gbps bit rate. It is important to note that as the signal traverses a greater distance, the accumulated dispersion also increases. For a given bit rate and at a given operating wavelength (or operating band), the maximum allowable accumulated dispersion is given by a standard specification. At no point in the dispersion map should the value of the curve go higher than the dispersion tolerance limit. Note that the dispersion parameters depend on many factors. The main factors are the bit rate (which gives the pulse width), the length of the fiber, the basic dispersion parameter, and the spectral width of the laser, which qualitatively provides the amount of dispersion induced (GVD). An interesting speculation is that of the variation of power penalty for dispersion-limited systems as a function of the dispersion parameter D, which is derived from the specification of the basic fiber. D can be considered a balancing component between the bit rate, the length of the fiber, and the width of the spectral source that emits the pulse. Refer to Figure 4-8.

Figure 4-8. Variation of Dispersion Parameter D with Power Penalty



Two techniques—precompensation and postcompensation—can compensate dispersion using any of these methods. As the name implies, *precompensation* means compensating for dispersion before the signal is induced in the system. This is a technique of compressing the pulse in advance with DCUs; it takes care of the accumulated dispersion in advance. In contrast, postcompensation uses compensating equipment that is placed at the end of a fiber. In precompensation, we can place the DCU after the postline amplifier. Such units have loops of fiber with dispersion profile opposite to that of the transmission fiber. For example, a transmission fiber would have dispersion parameter of 16 ps/nm-km. A DCU could hypothetically be made to have a dispersion profile of – ~ 50 ps/nm-km. The signal passes through such fiber spools (DCU) and the pulse is precompensated. Conversely, with postcompensation techniques, the DCU modules are placed before the preline amplifier, as shown in Figure 4-9. Table 4-3 shows the dispersion parameter D for different kinds of fiber at 1550 nm.

Figure 4-9. DCU Positioning in a Multistage System (Above) and Dispersion Map for the Same System (Below). Note the fall in dispersion due to the DCU placement and also note that accumulated dispersion never exceeds threshold (dotted line). Preceding technique is for precompensation.

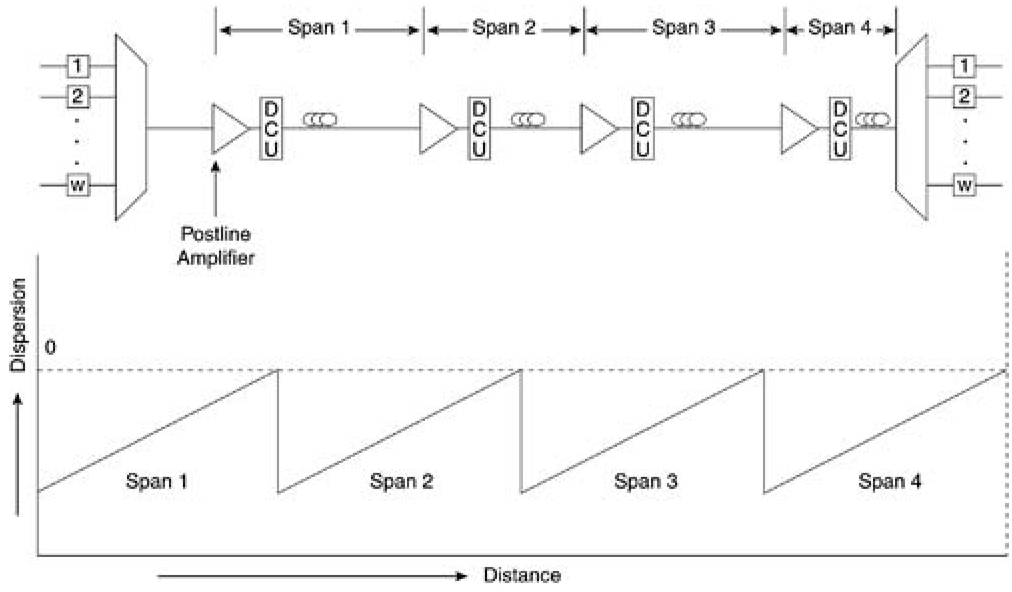
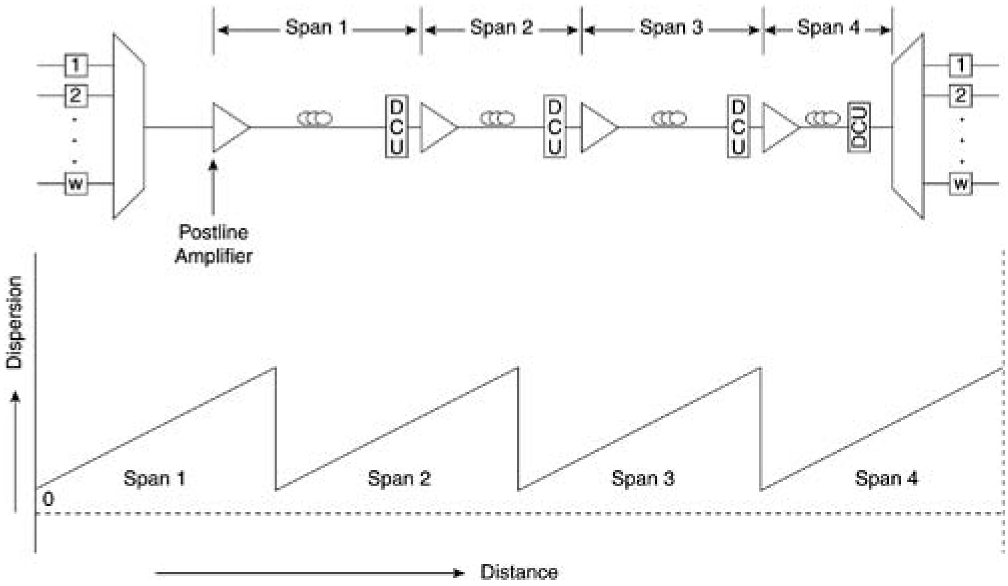


Table 4-3. Dispersion Parameter D

Fiber Type	Normal Dispersion at 1550 nm Measured in ps/nm-km
Single mode fiber (SMF)	17
E-Large Effective area fiber (ELEAF)	4
TrueWave RS (TW-RS)	4.2
Dispersion shifted fiber (DSF)	-0.33

A serious loss (attenuation) occurs when DCUs are added. This is due to a coupling difference between transmission fiber and the DCU. Moreover, different dispersion profiles result in a phase mismatch, which prevents FWM from happening. This is one advantage of DCU in limiting the effects of nonlinearity. Refer to Figures 4-9 and 4-10.

Figure 4-10. Dispersion Maps for Postcompensation Scheme


Span 1
Span 2
Span 3
Span 4
1
2
w
DCU
Postline Amplifier
Dispersion
0
Span 1
Span 2
Span 3
Span 4
Distance

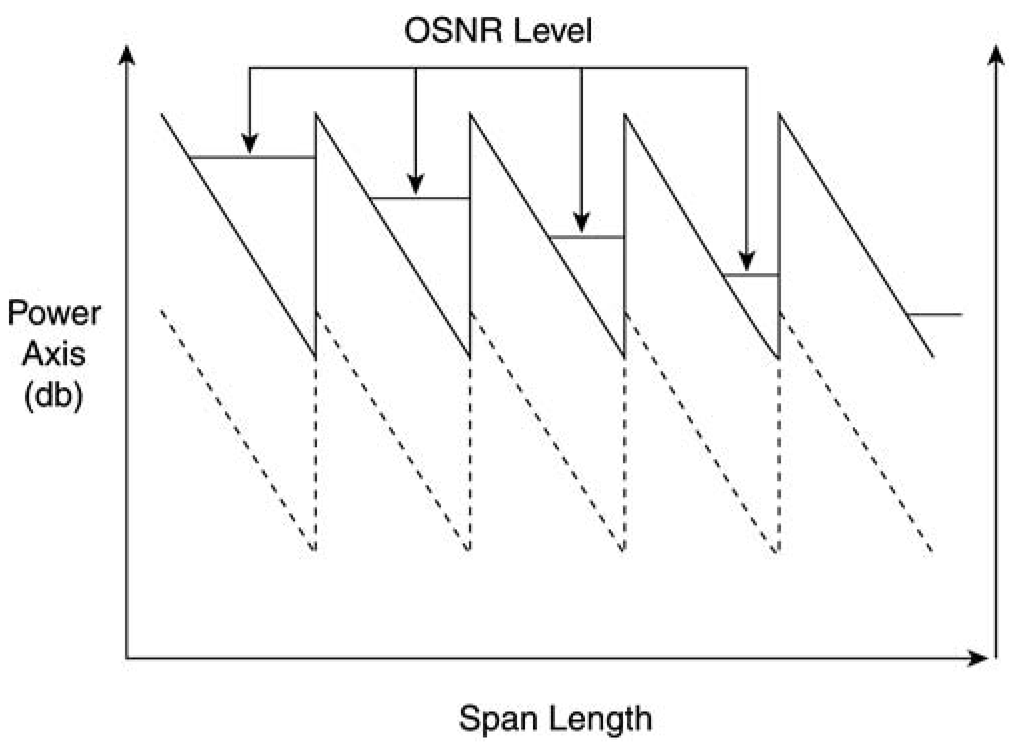
OSNR and Dispersion-Based Design

For a given network, it is important to calculate the OSNR and make a design based on both OSNR and dispersion limitations. It is possible to compensate for dispersion to a great extent. However, OSNR compensation needs 3R (O-E-O) regeneration, which is expensive. In other words, OSNR compensation is almost impossible for multichannel WDM systems. Therefore, when we are designing a WDM link, it is imperative to first consider OSNR's limitations. OSNR-based design essentially means whether the OSNR at the final stage (at the receiver) is in conformity with the OSNR that is desired to achieve the required BER. This also guarantees the BER requirement that is essential for generating revenue.

Following OSNR-based design, dispersion is the next issue to compensate from a design perspective. Dispersion-compensating units are readily available, but an important issue is where to place them. Various algorithms have been suggested depending on the network topology, the transmission length, and the bit rates. For most designs, optimization placements have to be done on a span (per length) basis.

Shown in Figure 4-11 is an OSNR map that carefully disseminates the optical signal level and the noise level as the signal passes through each amplification stage.

Figure 4-11. OSNR Levels in Terms of Signal and Noise Power Levels for Multistage WDM Transmission



Frequency Chirp

When pulses are generated at the transmitted end, intensity modulation causes phase modulation due to the carrier-induced change in the refractive index. This change is inherently due to the laser linewidth. Such optical pulses with a time-dependent phase shift are called *chirped pulses*. The optical spectrum is broadened due to this chirp. Theoretically, the chirp-induced power penalty is difficult to calculate, but it can be approximated to a 0.5 dB margin in system design (Chirp is also defined in Chapter 2.)

Effects of FWM and XPM on Long-Haul Design

FWM is a third-order nonlinearity in optical links that can be compared to the intermodulation distortion in standard electrical systems. FWM is worse for equally spaced WDM systems and at high powers. When three optical channels at frequencies ω_i, ω_j, and ω_k travel such that they are close to the zero dispersion wavelength, they intermingle to produce a fourth signal whose frequency is shown in Equation 4-24.

Equation 4-24

$$w_{ijk} = w_i \pm w_j \pm w_k$$

This ω_{ijk} can mix with another WDM channel, causing severe cross-talk. For W wavelengths in a fiber, the number of FWM channels (N) produced is shown in Equation 4-25.

Equation 4-25

$$N = \frac{w^2}{2}(w-1)$$

Figure 4-12 shows the effects of FWM in equally spaced systems and power considerations, and Figure 4-13 shows the same considerations for unequally spaced systems.

Figure 4-12. Equal Channel Spacing (Three Equally Spaced Channels Generated Nine FWM Signals, Out of Which Three Fall on Top of the Signals)

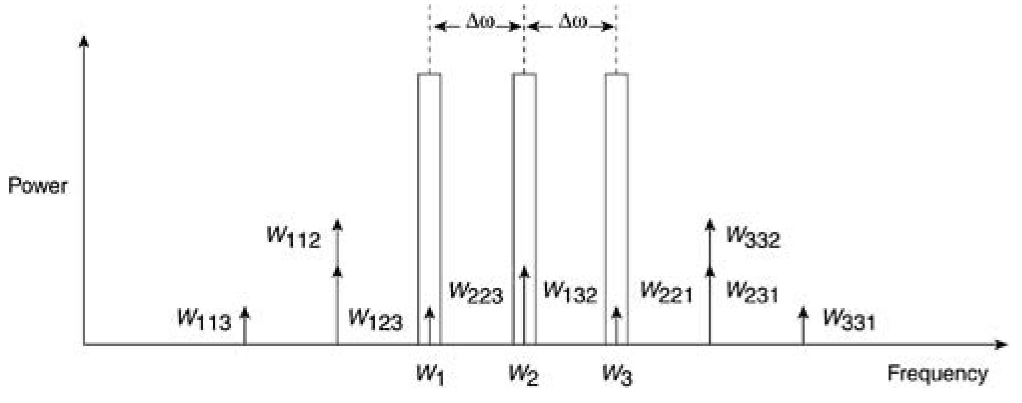
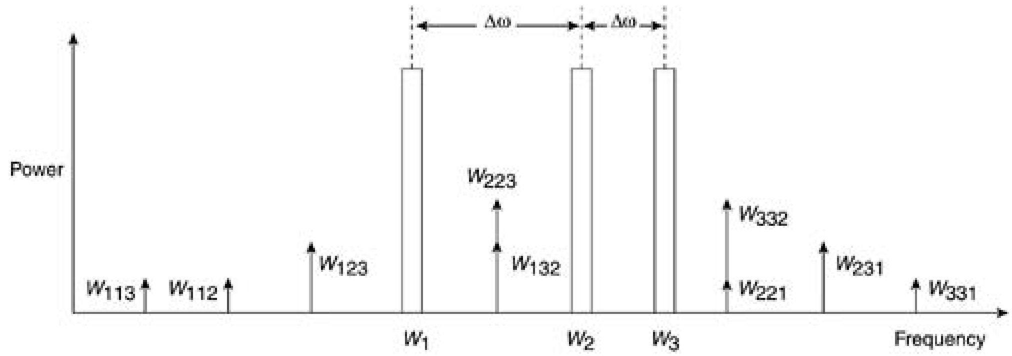


Figure 4-13. Three Unequal Spaced Channels Generating Nine FWM Signals; None of the Generated Signals Falls on Top of the Original

Signals



The solution for minimizing FWM is to use unequal channel spacing in such a way that the generated wavelength does not interfere with the signal channel(s). Use of NZDSF minimizes the effect of FWM.

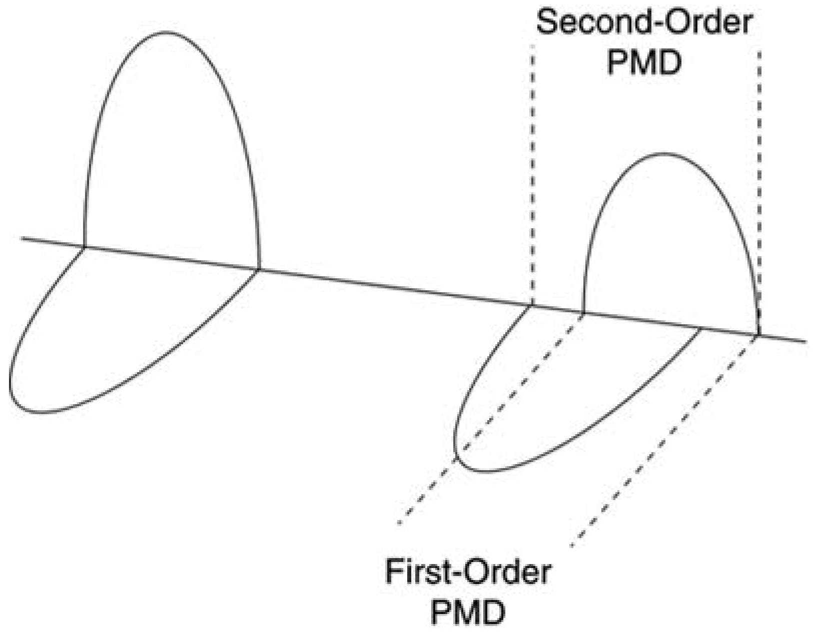
In multichannel WDM systems, XPM causes intensity-based modulation to adjacent frequency channels. XPM causes fluctuations in pulse propagation due to the effect of other channels. Furthermore, if adjacent channels are traveling at the same bit rate, XPM effects are more pronounced. One way to avoid XPM is by carefully selecting bit rates for adjacent channels that are not equal to the present channel bit rate. When designing WDM links, we typically keep a 0.5 dB power penalty margin for both FWM and XPM. XPM has more impact on certain types of modulation formats. Typically, FSK and PSK have a more pronounced impact than pure NRZ and RZ coding.

PMD in Long-Haul Design

PMD is not an issue at low bit rates; it becomes a dominant issue at bit rates in excess of 5 Gbps. PMD is inherently caused by the asymmetry of the fiber. This asymmetry adds to a property called *birefringence*, such that the two principle degenerate modes (polarization modes E_x and E_y) are subject to walkover effect. Due to this walkover effect, the modes are not coupled to each other, which in turn causes the pulse to spread in time.

The main type of PMD considered is second-order PMD, which essentially originates from dispersion due to wavelength dependence of the signal as well as the spectral width of the signal.Figure 4-14 shows the heuristics that create PMD in high bit-rate systems.

Figure 4-14. The PMD in a Pulse



A measure of PMD is the differential group delay (DGD), which can simply be visualized as the time difference in multiple spectral components (at multiple speeds) over a given length of fiber. The polarization axes are no longer joint, and the separation increases as the pulse is transmitted through a fiber. The difference is somewhat proportional to the DGD. Therefore, DGD can be accurately used as a measure of PMD for a given system. Moreover, PMD for a given fiber is defined as the mean of DGD. The mean DGD can be calculated from Equation 4-26.

Equation 4-26

$$DGD = (\text{PMD Coefficient}) \times \text{Length}^{\frac{1}{2}}$$

The typical system margin for PMD is 1 dB for general long haul, but it depends on the transmission length.

Consider a numerical example: If the PMD coefficient of the given fiber is 2 ps and the distance under consideration is 625 Km, calculate the DGD. Refer to Equation 4-27.

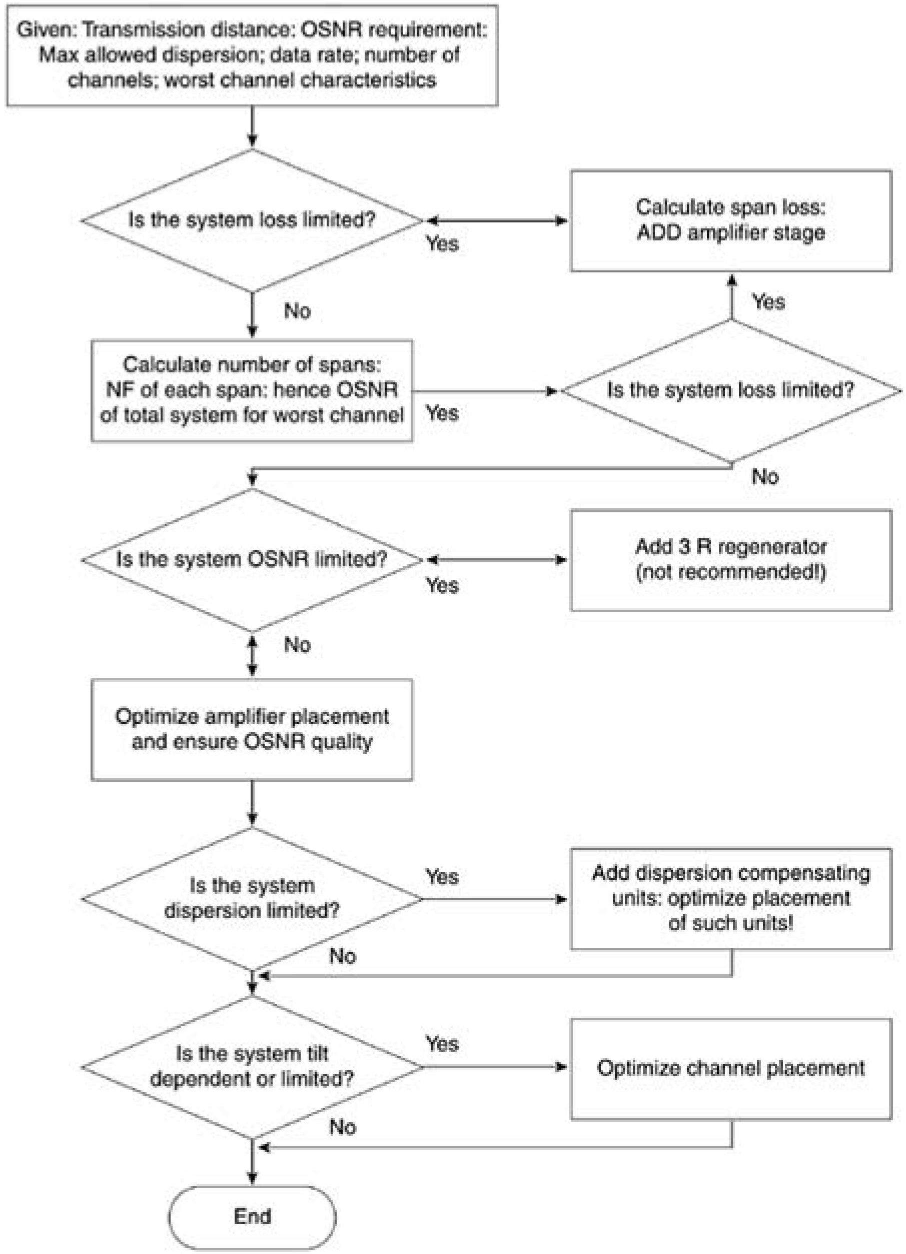
Equation 4-27

$$DGD = 2ps\sqrt{Km} = 2*\sqrt{625} = 2*25 = 50ps$$

In 10 G and 40 G systems, a DGD of this magnitude degrades the performance of the system (causes more BER).

Figure 4-15 shows the basic flow diagram that we would use to design an optical network based on system design principles explained so far. One issue that has not been discussed is the tilt limited system, whereby the gain of the amplifier produces different amplification for different channels (and hence different noise levels), generating different values of OSNR for different channels. Compensating tilt is a hard task despite the availability of flat band filters (called gain flattening filters) as well as band equalizers, which try to create uniformity to some extent in the optical working band.

Figure 4-15. Flow Diagram of a Generic DWDM Design Case



Given: Transmission distance: OSNR requirement: Max allowed dispersion; data rate; number of channels; worst channel characteristics
Is the system loss limited?
Yes
Calculate span loss: ADD amplifier stage
No
Calculate number of spans: NF of each span: hence OSNR of total system for worst channel
Yes
Is the system loss limited?
Yes
No
Is the system OSNR limited?
Yes
Add 3 R regenerator (not recommended!)
No
Optimize amplifier placement and ensure OSNR quality
Is the system dispersion limited?
Yes
Add dispersion compensating units: optimize placement of such units!
No
Is the system tilt dependent or limited?
Yes
Optimize channel placement
No
End

Examples

The following case studies reinforce the principles discussed so far in this chapter.

Case 1

Design a 4 x25 span WDM link with an optical amplifier gain of 22 dB and NF equal to 5 dB.

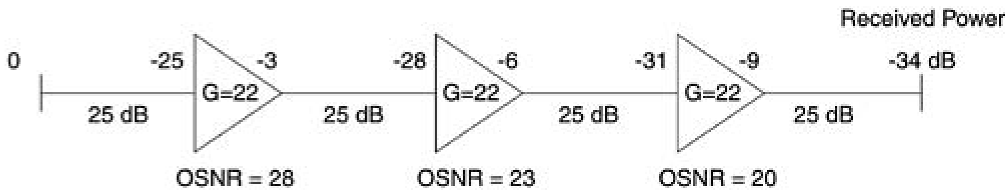
Calculate the final OSNR if the input power is 0 dB. Calculate the signal power at the receiver.

Will this system work if the receiver sensitivity is a minimum of –25 dB?

Will the system work if the input power is 10 dB?

The answer is shown in Figure 4-16.

Figure 4-16. Answer to the Problem



You can use one of the two methods to determine the final OSNR.

Method 1

$OSNR_{Final} = P_0 + 58 - \Gamma - 10\log N - NF$

N = 4 (number of spans)

NF = 5

$\Gamma = 25$ dB

$OSNR_{Final} = 0 + 58 - 25 - 10\log 4 - 5$

$= 58 - 25 - 6 - 5 = 22$ dB (the best estimate value)

Method 2

OSNR stage by stage analysis using the formula:

$OSNR_{stagei} = 1/(1/OSNR_{stage0} + NF.h.v.\Delta f\ /P_{in})\ (1/OSNR_{stage0} = 0)$

OSNRStage 1

NF = 5 dB converting to linear = 3.166 ($10^{NF\ dB/10}$)

h = Plank's constant = 6.6260E – 34

v = frequency of light 1.9350E + 14

Δf = bandwidth (measuring the NF) = 12.5 KHz (.1 nm)

P_{in} = input power at the amplifier

0 – 25 dB = –25 dB

$OSNR_{stage1}$ = 28 dB

Output from the amplifier = –25 + 22 = –3 dB

OSNRStage 2

$OSNR_{stage2} = 1/(1/OSNR_{stage1} + NF.h.v.\Delta f\ /P_{in})$

Input at the second amplifier

= –3 – 25 = –28 dB

$OSNR_{stage2}$ = 23 dB

Output from the amplifier = –28 + 22 = –6

OSNRStage 3

$OSNR_{stage3} = 1/(1/OSNR_{stage3} + NF.h.v.\Delta f\ /P_{in})$

Input at the third amplifier = –6 – 25 = –31 dB

$OSNR_{stage3}$ = 20 dB

Output from the amplifier = –31 + 22 = –9 dB

Power at the receiver

= –9 – 25 = –34 dB

If the receiver sensitivity is –25, the system will not work.

The solution is to (a) increase the gain of the amplifier (b) increase the input power of the transmitter.

The same solution for 10dB input power is shown in Figure 4-16.

Using Method 1: $OSNR_{final}$ = 10 + 58 – 25 – 6 – 5 = 32

Method 2: gives $OSNR_{final}$ = 29 dB

The difference in value is due to the approximation made in the parameters.

Final power at the receiver

= 10(Tr) – 25 (loss1) + 22 (gain1) – 25 (loss2) + 22 (gain 2) – 25 (loss 3) + 22 (gain 3) – 25 (loss) = –24 dB

The receiver sensitivity is given as –25 dB, so the system will work.

Case 2

Calculate the number of spans in this link, given P_{in} = 0 dB; $OSNR_{final}$ = 20 dB total length = 300 km Bit rate = 5 Gbps; NF = 5 dB. (Assume fiber type is SMF ∞ = .2 dB/km)

Answer

Total loss over entire length = 300 x 0.2 = 60 dB

Attenuation per span = 60/N (assume N number of spans

$OSNR_{Final} = P_0 + 58 - \Gamma - 10\log N - NF$

20 = 0 + 58 – 60/N –10logN– 5

-33 = –60/N– 10logN

Rearranging, you get this:

60/N– 10logN = 33

Solving for N (for N = 2)

60/2 + 10log2 = 33

Therefore, the number of spans = 2.

Case 3

OSNR = 20 dB, dispersion of the fiber is 17 ps/nm-km, span loss = 22 dB.

What is the total length of the system? (NF of the amplifier = 4 dB, and dispersion tolerance is given as 1600 ps/nm, P_{in} = 10)

Answer

$OSNR_{Final} = P0 + 58 - \Gamma - 10\log N - NF$

20 = 10 + 58 – 22 – 4 – 10logN

–22 = –10 logN

N = 158 spans

Therefore, total length = 158 * 22/0.2 = 17,280 km (theoretical limit)

But due to dispersion, max length = 1600 ps.nm/17 ps /nm.km = 94 km.

Case 4

Customer A wants to build a 200 km OC48 link to transport traffic. Design the link with the following parameters. (Assume SMF-28 fiber with α = .25 dB/km/ and 18 ps/nm/km as the dispersion characteristic.)

Receiver sensitivity: –18 dBm @BER=e^{-12}

Receiver overload: –10 dBm @BER=e^{-12}

Transmit power: +7 to +9 dBm

Dispersion tolerance: 1500 ps/nm

Dispersion penalty: 1.5 dB @ 1500 ps/nm

OSNR tolerance: 20 dB @ Resolution Bandwidth 0.1 nm

EDFA

Input power range: +3 to –25 dBm; Gain: 20 dB to 14 dB

Maximum output power: +17 dBm; Noise Figure: 5 dB

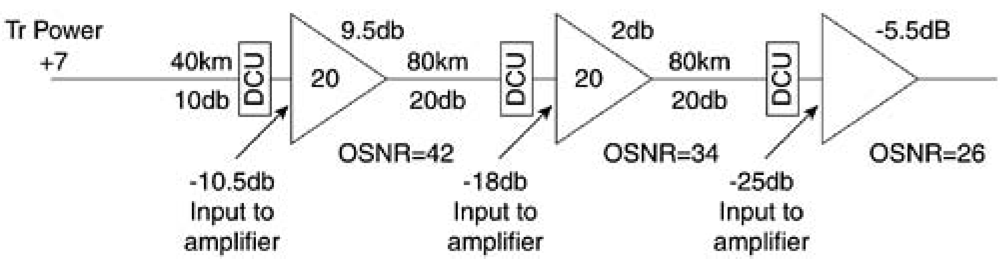
DCU

Loss 5 dB, dispersion compensation –1100 ps

Answer

The answer requires several steps. Refer to Figure 4-17.

Figure 4-17. Case 4 Answer



Step 1. The total distance is 200 km; therefore, total loss is 200 * .25 = 50 dB. You need amplifiers to reach that distances.

Step 2. Total dispersion is 200 km * 18 ps.nm / km = 3600 ps.nm

You need to use dispersion compensators because the given dispersion tolerance is 1500 ps.nm. The maximum distance for this system without DCU is 1500/18 = 83.33 km given DCU = –1100 ps. If you strategically place 3 DCU, you get (3600 – 3 * 1100) = 300 ps, which is well within the limit. The DCU has a 5 dB passthrough loss, so it is best to place it before preamplifiers.

Analysis of the problem

Transmit power = + 7 dB

Stage 1:

loss = –10 (link loss) – 6 (DCU loss) – 1.5 (margin) = –17.5 dB

P_{in} = Power at the end of stage 1 = 7 – 17.5 = –10.5 dB

Stage 1 amplifier gain: 20

Power output from the amplifier: (power input + gain) – 10.5 dB + 20 = 9.5 dB

OSNR calculation

$OSNR_{stagei} = 1/(1/OSNR_0 + NF.h.v.\Delta f / P_{in})$ (for stage 1 $1/OSNR_0 = 0$)

$OSNR_1 = Pin/NF.h.v.\ \Delta f$

Substituting the values Pin = –10.5 dB = 8.9125E – 05 W (convert to watts)

NF = 5 dB converting to linear = 3.166 ($10^{NF\ dB/10}$)

h = planks constant = 6.6260E – 34

v = frequency of light 1.9350E + 14

Δf = bandwidth (measuring the NF) = 12.5 KHz (.1 nm)

Substituting, you get OSNR = 42 dB

Stage 2:

Loss = –20 (link loss – 6 (DCU loss) – 1.5 (margin) = –27.5 dB

Power at the end of stage 2

P_{in} = 9.5 – 20 – 6 – 1.5 (margin) = –18 dB

Stage 2 amplifier gain: 20

Power output from amplifier: –18 + 20 = 2 dB

$OSNR_{stage2} = 1/(1/OSNR_1 + NF.h.v.\Delta f / P_{in})$

OSNR1 is the OSNR of stage 1 = 42 dB and P_{in} = –18 dB

$OSNR_{stage\ 2}$ = 34 dB

Stage 3 loss:

–20(link loss) – 6 (DCU loss) – 1.5 (margin) = –27.5 dB (loss)

Power at the end of stage 3

P_{in} = 2 – 27.5 = –25.5 dB

Stage 3 amplifier gain 20 dB

Power output from the amplifier: = –25.5 + 20 = –5.5 dB

OSNR

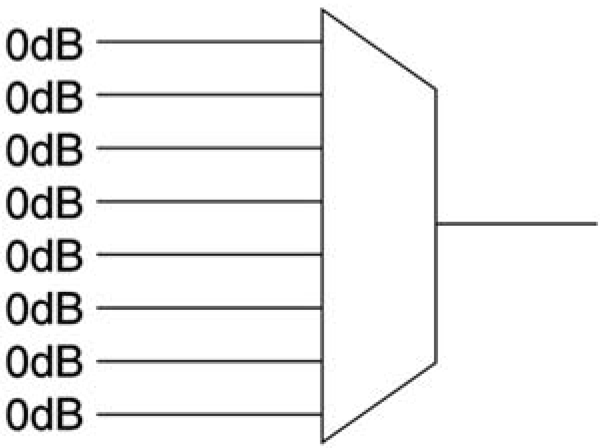
$OSNR_{stage3}$ = 1/(1/$OSNR_{stage2}$ + NF.h.v.Df /P_{in})

OSNR = 26

Case 5

Calculate the composite power at the output of a DWDM 8 channel multiplexor (shown in Figure 4-18) if the input power is 0 dB (insertion loss = 5 dB).

Figure 4-18. Calculate the Composite Power



$P_{composite}$ = $P_{channel}$ + 10logN – Insertion loss (where N is the number of channels)

= 0 + 10log8 – 5 dB = 0 + 9 – 5 dB

= 4 dB

Exercises

The following exercises are left for you to solve. Use previous examples as guidelines.

1. Calculate the dispersion and the receiver power of an optical link that is 90 km in length. The transmit power = +8 dB, the receiver sensitivity = –18 dB, and the dispersion tolerance = 1500 ps/nm (SMF ∞ = .25 dB/km 18 ps/nm-km). Is it dispersion limited? Design the link for proper network operation if you need to use EDFA 15 dB (gain) and DCU

–1100 ps and 5 dB loss).

2. Design a 4 x 25 span WDM link with optical amplifier gain = 18 dB and NF = 6 dB.

   Calculate final OSNR if the input power is 0 dB. Calculate signal power at the receiver.

   (The receiver sensitivity is –25 dB at BER 10^{-15}). Does the system work?

   Now calculate the final OSNR by replacing the following:

   - The transmitter with input power to +7 dB

   - The amplifier gain to 22 dB

3. Calculate power in milliwatts if the input power were 0 dBm.

   Calculate the power in dBm if the input power were 12 mW.

   Hint: XmW power is 10 * $\log_{10}(x)$ in dBm

   Y dBm power is $10^{Y/10}$ in mW

4. Calculate the composite power at the output of a DWDM 16-channel mux if the input power is 4 dB. If the input is plugged into an amplifier, how much attenuation do you need (given NF = 5 dB, input range = 0–25 dB, gain = 22)

Summary

This chapter discussed optical network system-level designs. We initially considered power budget-based design, from which we migrate to more complex OSNR-based designs. This chapter showed the importance of OSNR in estimating BER and the need for evaluation of the Q-factor as an intermediate stage in BER development. We discuss dispersion-based systems and penalties that are associated with dispersion-limited systems. We show how dispersion-limited systems can be compensated using generic schemes and hence show the methods of pre- and postcompensation as well as placement of such compensating units. Nonlinear effects are also studied from a design point of view, as well as the various penalties and their cures from a system design perspective. Finally, we analyzed numerical examples for the studied system design principles.

References

[1]Bononi, A. *Optical Networking*. Springer Publishers, 1999.

[2]Senior, J., C. Qiao, and S. Dixit. *All Optical Networking 1999: Architecture, Control and Management Issues*. Proceedings SPIE, 2000.

[3]Agrawal, Govind P. *Fiber-Optic Communication Systems*, Second Edition. Wiley Interscience, 1997.

[4]Keiser, G. *Optical Fiber Communications*, Third Edition. McGraw-Hill, 2000.

[5]Mukherjee, B. *Optical Communications Networks*. McGraw-Hill, 1997.

[6]Green, P. *Fiber Optics Networks*. Prentice Hall, 1992.

[7]Ramaswami, R. and K. Sivarajan. *Optical Network: A Practical Perspective*, Second Edition. Morgan Kaufmann, 2001.

[8]Agarwal, G. *Nonlinear Fiber Optics*, Second Edition. Academic Press, 1995.

[9]Saleh, B. and M. Teich. *Fundamentals of Photonics*. Wiley, 1991.

[10]Ramamurthiu, B. *Design of Optical WDM Networks, LAN, MAN, and WAN Architectures*. Kluwer, 2001.

Chapter 5. WDM Network Design -2

This chapter focuses on network topology design based on the actual deployment and expected future of optical networks. Practical networks are designed and deployed in topologies that are quite well defined, although the network itself might be spread across vast geographical areas in an extensive way.

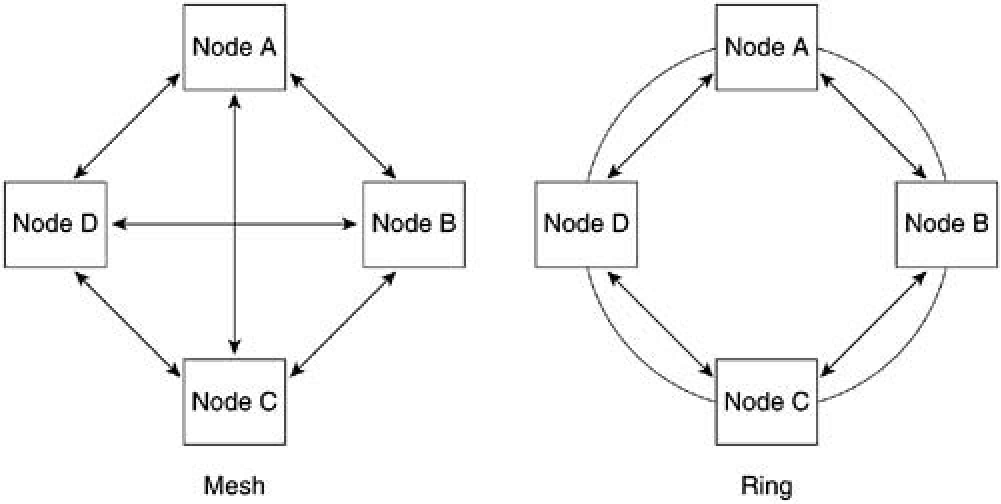
From the perspective of the optical layer, networks are classified by their topology as well as their size. Chapter 4, "WDM Network Design -1," briefly discussed the three topologies: point-to-point networks, rings, and mesh topologies. While designing a network, it is imperative to know which is the most suitable topology for a given cluster of nodes. This decision is important in green field as well as deployed brown field networks, and many factors affect the final outcome of the design.

Consider a given network topology such that the fibers are already deployed (fibers are laid in the field). If we have to optimally use the fibers to provide traffic to the set of nodes and at the same time ensure the least cost and good optical characteristics, we need to make a good design for the network. The easiest way to do this, of course, is to design a topology, which maps each traffic request (source-destination pair) to a fiber link. This is what is called the fully connected mesh. The next obvious question is whether a fully connected mesh (one in which each node is connected to every other node through a fiber) is the most widely deployed topology? The answer is no, for the simple reason that such a network infrastructure would never be economical. What *is* the most feasible topology?

The answer to that question is quite complicated. No topology is the best topology, but different topologies are good enough for different alternative scenarios. The idea is to choose a topology and then to choose architectures that would most suit this topology. We have to choose different topologies and we do so on the basis of geographic area and or size of the network. Topology choices, whether point-to-point, ring, mesh, or any hybrid architectures, depend on the traffic as well as the situational requirement of the network.

ConsiderFigure 5-1a and 5-1b in which we have a mesh and ring network. Note that in Figure 5-1a, we have a fully connected mesh topology, whereas in Figure 5-1b, we have a ring. By definition, a mesh is a group of nodes that are connected to one another by links (fibers in our case) in a fixed or arbitrary manner; these links are further inhabited by WDM channels or wavelengths. The interconnection pattern of nodes in a mesh network can be regular or irregular and can have different degrees of connectivity (explained later). Mesh architectures speak of a large capacity, which is typically equal to the number of links, 'L' times the number of wavelengths 'w' in each link. A mesh is characterized as a planar graph G(V,E) of V vertices and E edges such that each vertex corresponds to a particular network element or a nodal interface unit or just simply a node. The edges correspond to fibers; an edge might be a single fiber or a group of multiple fibers (two in most cases for duplex traffic). The *degree of connectivity* (D of a node) can be defined as the number of nodes that are connected to the present node through the various edges.

Figure 5-1. (a)Mesh Topology (b)Ring Topology



If a particular node, N_1, has three adjoining edges connecting to three nodes, then its degree of connectivity, D is equal to 3. For a fully connected mesh, the degree of connectivity 'D' is N – 1, where N is the number of nodes in the mesh network. A mesh is termed cyclic if cyclic loops are present in the topology. In other words, a cycle can be traced out from the source back to itself through a series of nodes.

A tree is a set of nodes with an originating node (called *root*) and connecting nodes connected as per the network topology. A spanning tree is a tree of nodes that consists of all the nodes such that no node is repeated twice, and all nodes are present in the set corresponding to the tree. The nodes in a tree are called *branches* or *leaves*, and an articulation point in a mesh is defined as a node in the tree, which if removed, breaks the mesh (and hence the tree) into two discrete trees (or two meshes).

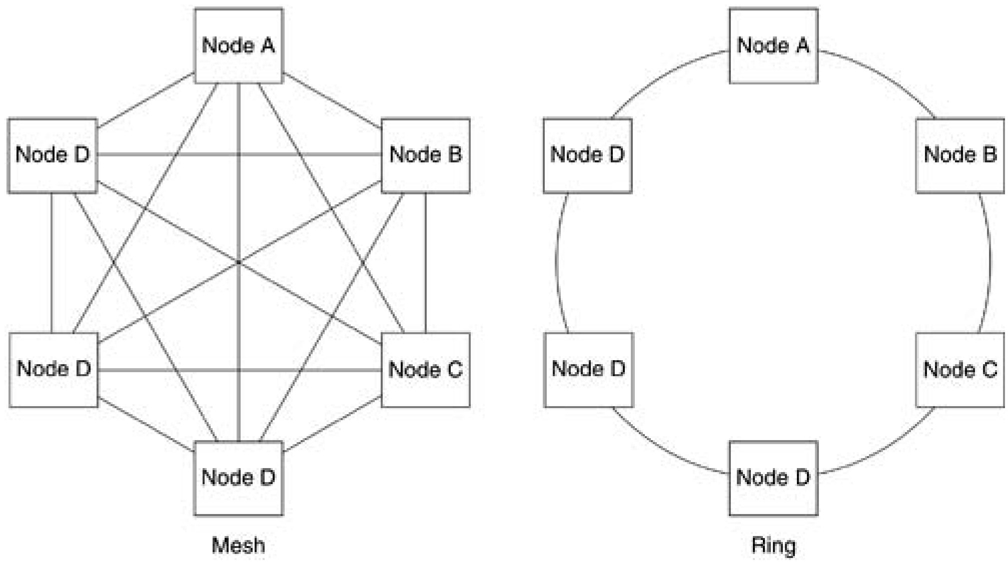
A minimum-weight spanning tree is a tree that includes the shortest path to each node from the root of the spanning tree. Various algorithms in literature can calculate and find the spanning tree in a fixed number of iterations. The number of iterations to locate nodes is often associated with the computational complexity of the algorithm, which is a figure of merit for management systems in optical networks (for example protection and restoration route calculation). A mesh network with all nodes having degree of connectivity 2 is called a *ring*, whereby each node in the ring is connected to only its most adjacent nodes in the network.

In a green-field network, these kinds of graph theory corollaries are beneficial in the overall network design. However, that might not always be the case in brown field or existing network topologies. An issue of tremendous importance is that of mapping these corollaries to different networks and facilitating good interconnection at the optical layer as explained in the sections to come. Most of today's optical networks (especially metropolitan area) are laid out in ring topology due to historical dependence on SONET hierarchy.

WDM Pass-through Case—Virtual or Logical Topology Design

Lightpath communication is an efficient way to increase bandwidth in optical networks. A lightpath is an all-optical path between a given source and destination node such that it can cut through many intermediate nodes. The lightpath is on a given wavelength, and for most cases, wavelength continuity is a constraint. The constraint is such that all the internodal segments of a lightpath (between each pair of nodes from source to destination) have to use the same wavelength. Now consider Figure 5-2a and 5-2b. In Figure 5-2a, we can see that for a 6-node network, we have a fully connected mesh.

Figure 5-2. (a)Fully Connected Mesh (b)Corresponding Ring Topology



Note that the number of links is equal to the number of unique source-destination pairs (assume half–duplex connection per pair). In contrast, in Figure 5-2b, we have a 6-node ring such that the number of links is equal to twice the number of nodes in the network. Obviously, this means that the number of links in a ring is far lesser than the number of links in a mesh network. That means cost saving! How can we map a lightpath for a source destination pair that is quite far apart (in other words, not adjacent)?

Consider a lightpath request from node A to node C. Because of the absence of a direct link between A and C, we have to resort to some logical connectivity. We can establish a lightpath between nodes A and C through node B, such that node B acts as a pure pass-through node. We can do this by choosing a wavelength (often not as easy as we might think due to the limitation on the number of wavelengths) such that the wavelength is free in the links A-B and B-C. Then we can configure the nodal interface unit(s) such that the lightpath is not blocked. Upon completion, we have a lightpath AC, which is quite oblivious to node B. We have done this by using the pass-through character of node B and by intelligently choosing a wavelength between A and C.

In other words, we have also established a logical connection between nodes A and C. This means that although no direct physical link is present between nodes A and C, we still have a logical connection between nodes A and C through node B. The connection is not dropped at node B, which is in fact oblivious to the data flowing through it. This kind of topology design is often termed *virtual* or *logical topology* design. This means that for a given physical topology, we can plot a logical dynamic topology to support the given traffic demands at a given time using a fixed number of resources (usually wavelengths). We can deal with this issue in various ways; one way is to try to design a logical topology that can be mapped onto an existing physical topology for the given number of logical connections to be established.

This chapter studies the method to design networks based on topology, their classification, and their performances, as well as comparison. Before moving on to the basic classification nomenclature, we need to understand the concept of protection. As of today and for the near future, optical circuits or lightpaths are the commodity that carriers offer to customers to generate revenue. Therefore, the service level agreements (SLAs) between customers and carriers or carriers and other carriers are based on the lightpath strength (in numbers and in granularity-bit rate) that is offered.

By time division multiplexing (TDM) and *muxponding* (aggregating several low bit rate streams into a single high bit rate stream in the electrical domain), many low bit rate signals are multiplexed into a lightpath of large granularity, from which the carrier (ISP) is effectively able to make revenue. Most of the traffic in previous years was voice-based traffic, although that is now being replaced rapidly with data traffic. Nevertheless, as of today, voice traffic continues to be the part of the lightpath that earns the most revenue. Voice traffic generally has severe constraints, such as low end-to-end delay and hence latency. If a fiber cut occurs or equipment fails, the lightpath is lost, resulting in loss of revenue.

This issue is somewhat alleviated by the concept of protection, whereby a protection lightpath (generally spatially separated from the 'work' lightpath) is set up. The concept of protection at the optical layer is actually borrowed from legacy SONET networks, whereby one or more of the legacy protection features provided by SONET might protect a work lightpath. This chapter devotes a complete section to studying protection in optical networks and the effects of protection on network heuristics. To pursue the classification of networks in detail, we need to know that protection lightpaths are generally longer than their work counterparts for the simple reason that work lightpaths are established on routes that are based on shortest path algorithms.

Classification of Optical Networks Based on Geographical Sizes and Functionality

The optical networking market is roughly divided into three main segments for ease of analysis as well as customer focus. The division is based on the size of a network because of the underlying physics of the optical layer. In the preceding chapters, we saw that as the transmission distance and line rates increased, there was an almost exponential corresponding increase in optical impairments.

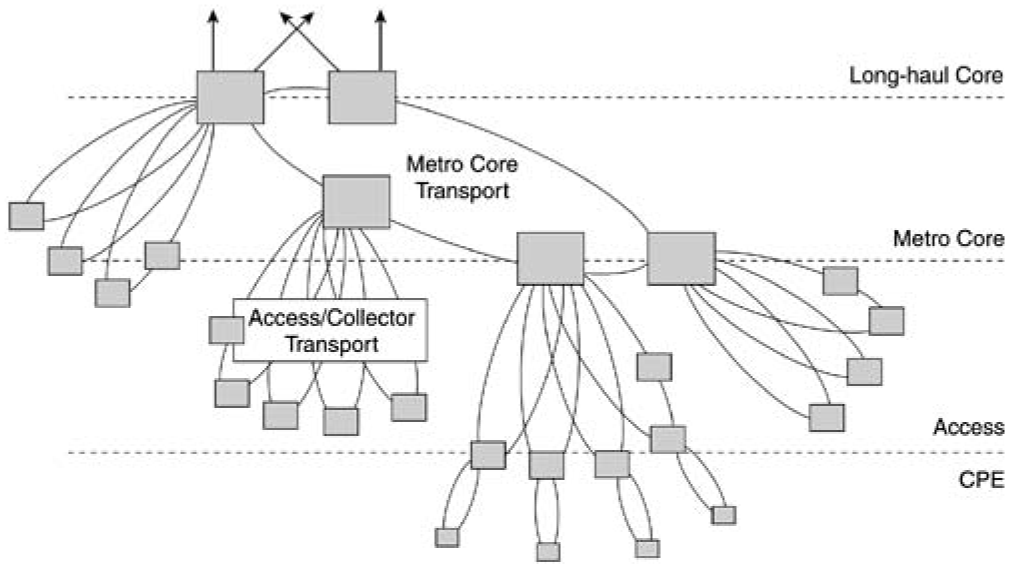
For short WDM links at low bit rates, there was no issue of impairments due to nonlinearities. Similarly, dispersion did not become a potent hazard to transmission until the *bit-rate-transmission distance product* reached a certain cutoff value. To maintain good optical transmission, we need to install dispersion compensation equipment as well as optical amplifiers to ensure that the signal reaches the receiver at a proper power level (bit error rate, or BER, maintenance). Transmission distance plays a critical role in determining network cost. Ultimately, costs guide the whole market, so it makes perfect sense to classify networks based on transmission distances or geographical sizes.

Networks are classified into three main categories corresponding to their size and functions. The first type is that of metro access networks, or simply access networks. The second classification type is that of metropolitan or regional area networks. The last classification type is that of long-haul as well as ultra long-haul (submarine and intercontinental) networks. As we will see later, the network size and requirement greatly enhance the cost of the network.

Classification of Networks

As mentioned before, the classification nomenclature is a three-tiered hierarchy, based primarily on the transmission distances. In the first tier is the access network market. Refer to Figure 5-3.

Figure 5-3. Network Classification: Long-Haul, Metro Core, and Access Hierarchies. Note the topological differences between the different network types as well as the geographical area coverage.



In the next tier is the metro core market, and in the uppermost tier is the long-haul market. The enterprise customers fall under the customer premises equipment (CPE) layer, where the end customers are connected to the service provider's network. Refer to Table 5-1 for more information.

Table 5-1. Classification and Comparison of Optical Networks

	Metro Access Network	Metro Core Network	Long Haul
Topology	Ring	Ring/mesh	Point-to-point/ring and mesh
Traffic flow	Hubbed	Distributed/meshed	Hubbed
Ring circumference/length	40–75 km Smaller rings also possible	< 100 to 250~300 km	< 300+ to 2000 km
Wavelength supported	Up to 16 lambdas	32 to 64 typical	64+
Number of nodes in the ring	6 ~ 8	8 ~ 16	5–12
Distance between the node (span length)	10–30 km	~40–300 km	300+ km
Current market size	20~30%	50~60%	10~20%

DWDM technologies (not limited)	Mux/Demux OADM Electrical switching and routing	Mux/Demux OADM EDFA (amplifier) DCU Wavelength-tunability Electrical switching and routing/limited all optical switching	Mux/Demux OADM EDFA (amplifier) DCU PMD Wavelength-tunability All optical switching and routing

Metro Access Networks

Metro access is the most convenient optical network that is currently deployed. An *access network* can be defined as the edge of a network that a customer can access directly. Access networks can be ring shaped or bus/tree shaped in topology. These networks are generally 40–75 km in longest transmission lengths, which means that if they are deployed as a ring, the circumference is generally less than 75 km.

Access networks are built to reach the end user, who is generally a corporation or an enterprise. Service level agreements are issues to such enterprise businesses. Access networks are currently the most growing topology in the metropolitan area. The market witnessed phenomenal growth even in times of recession due to the ever-increasing Internet traffic and consumer business activity in the enterprise. The main purposes of access networks are routing, aggregation, and transport, although routing can be assumed slightly lower in priority than aggregation because access networks are linked to metropolitan area core networks, which perform the main task of transport.

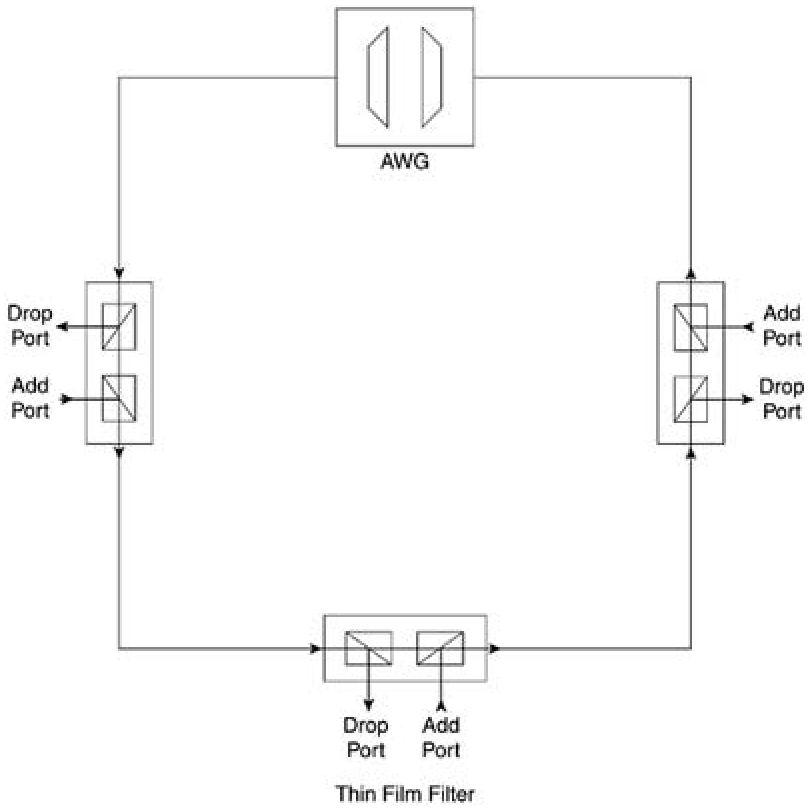
In the access network area, a particular system might consist of a few WDM channels (typically 8–16) provided by redundant protection channels in another fiber and traversing a different path. The metro access network does not consider too many of the optical impairments in system design for two reasons. First, the network is not spread more than 75 km; therefore, issues such as dispersion and attenuation are quite minimal (at low rate). Second, the line rates by themselves are quite low. In one deployed 16-channel WDM access network, the majority of wavelengths were on OC-3 speed, a few were based on OC-12-OC48 speed, and none were based on OC-192 (10 Gbps) speed. This is due to the fact that access networks act as aggregators of traffic from various nodes. Typically, gigabit-switched routers would be positioned with OC-3, OC-12, and Gigabtye E (GigE) interfaces additive into an access network. A rare OC 192 could also be part of the network, but such instances are uncommon.

The traffic schematic for access networks is quite well defined. Although there is a percentage of traffic, which emanates and culminates inside the access ring or tree, a sufficiently large amount of traffic is destined for the core or even the long-haul network. This internetwork traffic is what defines the access topology. Generally due to this kind of traffic matrix demand, the topology is *hubbed*, whereby the individual nodes have a logical connection to the parent hub, which in turn is connected to the core network. Optically, at the hub, the lightpaths can be fully demultiplexed and switched in the electronic layer, or the lightpath can be all optically converted into a metro-core wavelength (can be the same as the access wavelength).

All optical wavelength conversion technology is in its infancy; therefore, the most viable option is to use transponder cards for opto-electro-opto (O-E-O) conversion and regeneration. This kind of 3R regeneration enhances the signal qualities in terms of the optical signal-to-noise ratio (OSNR). In the hubbed architecture, the hub plays a pivotal role in determining which lightpaths are inter ring and which lightpaths are intra ring. The architecture of the nodes (in general) can either be that of a conventional OADM as described in Chapter 1, "Introduction to Optical Networking," with fully multiplex sections decimating the composite signal into its constituent wavelengths, or a more economical architecture with passive devices such as couplers and filters.

Looking at some current access requirements, the most viable architectures are those that can drop and add, a small percentage of traffic at each node and whose optical characteristics (typically insertion loss) are quite low. In one such embodiment, a filter-based architecture was introduced whereby thin film filters were placed (see Figure 5-4) in the line of the composite signal, thus dropping a fraction of the wavelengths to the local multiplex section. The insertion loss of such architectures was typically in the range of 5–6 dB as compared to the insertion loss of almost 14 dB in core Optical Add-Drop Multiplexer (OADMs).

Figure 5-4. Thin Film Filter-Based Network (Metro Access)

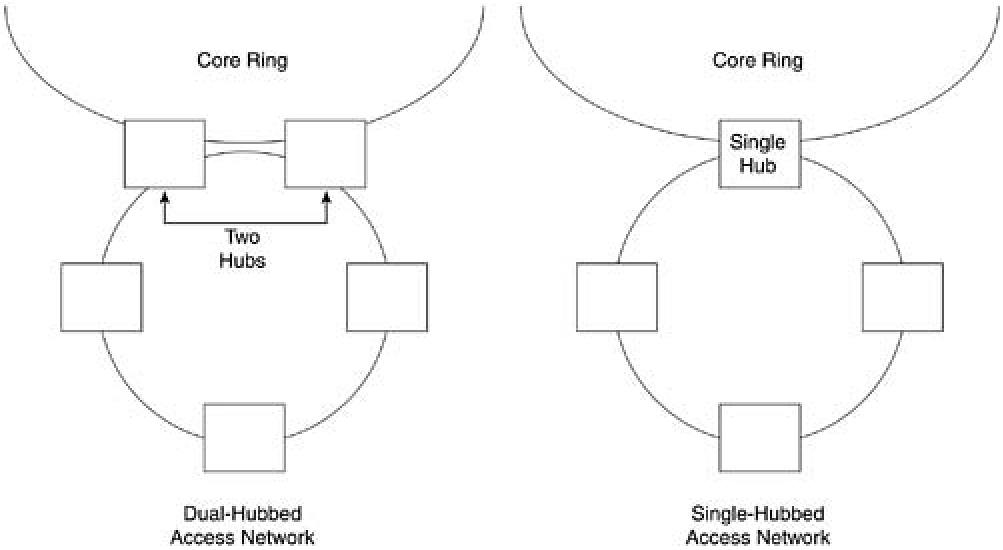


The philosophy of success of such architectures is that the constituent nodes add and drop only a small fraction of the wavelengths; therefore, it is not necessary for a full optical multiplex section to be present. However, at the hub node, there is a need to add and drop most of the traffic; therefore, this architecture might follow conventional models-full demultiplex sections. The advantage of such a hybrid system is not only the lower cost but also an improved system design. This can be attributed to the losses associated with full multiplex sections such as arrayed waveguides (AWGs), whereby the waveguides have a high loss of their own due to phase matching of different wavelength streams. In addition, the wavelengths have considerable detrimental effects on the polarization and pulse shape of the traveling signal(s). In short, we should avoid AWG-based full demultiplex sections wherever possible. Though at times, it becomes imperative to work with AWGs, as shown in the sections to follow.

Because access networks generate the most revenue due to their close proximity to the customers or end users, it becomes imperative to protect the networks with a plausible protection algorithm. Although such protection schemes are discussed in the section later in this chapter titled "Protection in WDM Networks and Protection Switching," it is important to mention now the architectural aspects of protection with respect to hubbed access networks.

It is a good idea to have a dual-homing network or a two-hubbed network (Figure 5-5a and 5-5b) with two independent nodes interconnecting to the core rings. Such dual-hubbed architectures serve as two entry points to the core networks. If a single-hub failure occurs, the second hub can still function as an interconnecting segment between the core and access networks. A great deal of research is currently underway to determine new architectures and topologies for metro access networks. Although this seems to be a good source of revenue for equipment vendors, a loss of vision for most carriers makes deploying new versatile architectures quite difficult.

Figure 5-5. (a)Double-Hubbed Ring (b) Single-Hubbed Ring



Metro Core Networks

Metro core is the next tier in our three-layered approach to classifying optical networks. Because of its revenue-generating capability, this is currently a hot area for carriers. Most equipment vendors have product portfolios in this segment. Metro core is distinguished from the access area by its size. Typical metropolitan area regional networks are 100–300 km in total transmission length. Most deployed metro core networks are ring-based topologies because of their migration from Synchronous Optical Network (SONET) rings.

The issue of why rings have been chosen is a hot debate among academic circles. Although the proponents of mesh argue about the benefits such as greater capacity, it remains a fact that ring-based topologies are more resilient and lower-cost alternatives. Metro core networks have the distinct functionality of bridging the gap between access and long-haul networks. They have the predominant role of transport in the WDM layer. Metro core networks also have a serious intranetwork (within the network) component of traffic compared to the already existing internetwork (between different networks) component. Metro core networks can be described as a multihubbed network with some hubs interfacing to long-haul access points and other hubs interfacing to the metro access rings. Therefore, metro core networks serve as an intermediate layer to both metro access and long-haul networks.

Metro core networks also have subsidiary networks called *collector networks*, which collect traffic from metro access and supply it to the regional rings. The ring size is generally less than 200 km in circumference, and the line capacity is typically up to 40 wavelengths (although demonstrations of greater wavelengths have been sparsely illustrated). Typical metro core networks derive most of their legacy architecture from SONET rings; therefore, they are typically built on two unidirectional (clockwise and counterclockwise) fiber rings.

Metropolitan area core networks assume significant importance because most of the Internet as well as legacy voice traffic is concentrated in this segment of the hierarchy. As mentioned before, these networks are based on ring topologies that have evolved from the traditional SONET rings. Traditional SONET rings had the problem of a single wavelength in the ring causing nonoptimal use of the fiber. By placing an SONET add-drop multiplexer (ADM) at each node, the network functioned extremely well, but the signal had to be dropped at each node, and the ADM had to switch individual bit streams at the STS-1 level.

The traffic demand in the ring increased drastically as more access rings were added to this ring. WDM metro rings alleviate this issue by placing optical ADMs, which can add-drop or pass-through traffic at each site. These ADMs can also be made to behave like wavelength routers, routing wavelengths from one fiber to another entirely in the optical domain. This technique, although cost effective, has the issue of transmission length. Typically, 40 wavelengths can be envisioned in the WDM metro rings, and transmission distances can span the length and breadth of a typical metropolitan city.

Current networks suggest that metro core rings operate in the 100–300 km circumference with about 6–14 nodes in each such ring. A typical WDM metro core ring would interface with both long haul and with access ring at two distinct points, thus serving as access interfacing points to the two networks. Due to the strong emphasis on SLAs, carriers tend to benefit significantly from metro core networks. Until optical packet switching matures, metro core networks will continue to be the main source of revenue. However, after the long-haul market picks up, metro core networks will play the role of silent aggregator of traffic for the access network and will schedule this traffic into the long-haul network.

Metro core networks also have the sincere benefit of having to interface with different carriers. Therefore, standardization is an important issue. Many ITU standards have been proposed for these networks, one of which is wavelength spacing and wavelength allocation. Because of the low loss segment in the fiber situated in the vicinity of 1550 nm, the band commencing from 1525–1565 nm is commonly deployed. This is also called the *Cband* (refer to Chapter 1 for operating bands). Longer L bands (1570+ nm) are being currently investigated. Core rings need to be resilient and often need SONET-like protection features. The switch from a failed lightpath to a protection lightpath is called *restoration*. In metro core networks, typical restoration times have to be in the 50 millisecond range to facilitate excellent responses especially for embedded voice communication.

Salient Features of Metro Core Rings

The ring topology has some strategic finer points, which have made the deployment of ring topology an excellent solution for metro core networks. Intuitively, mesh networks seem like good choices for deployment considering that the objective is to guarantee maximum traffic between any arbitrary source destination pair in a network. Rings seem to be a poor cousin for the higher degree (of connectivity) of the mesh in that regard.

Current network traffic reports that the huge capacity of the mesh is not called for. In one analysis, it was shown that ring networks can do all that a mesh can do but save price for a given load (the key here is that load should be low). Moreover, interconnected rings offer to be a genuine solution for implementing mesh-like architectures, alleviating the tedious routing tables and protection issues. Protection and routing in mesh is quite intricate because of the higher degree of connectivity in the network.

Furthermore, a failed fiber granting a protection path might lead to an algorithm of high computational complexity, whereby the time needed to choose a protection path depends on the network state and that load balancing (dynamic) after restoration can be achieved. In that regard, we can consider a 2-fiber ring such that each fiber is unidirectional; for a fiber cut, we can guarantee protection by switching the signal into a corresponding protection wavelength in the fiber with an opposite direction of propagation. Obviously, the transmission length will increase, but for abrupt and arbitrary failures, this path is predetermined, independent of what the traffic matrix looks like at the given time. Therefore, it is quite independent of load-balancing schemes.

Another serious issue in metro rings is the ability to reuse the wavelengths. When a lightpath is dropped at some destination node, the wavelength (corresponding to the dropped lightpath) can be reused for setting up another lightpath emanating at either the same node or at one of the nodes in the downstream direction. Although mesh networks can also reuse wavelengths based on spatial separation, this is not so pronounced due to the already existing huge capacity in the mesh. Also in mesh, due to strong dependence on cross-connect architectures (explained in the section titled "Nodal Architectures for Different Network Markets," wavelength reuse does not greatly affect system performance. Spatial reuse of wavelengths by add-drop nodes greatly increases the capacity (traffic) of the network. The final amount of traffic (in number of lightpaths) at any point is shown in Equation 5-1.

Equation 5-1

$$Tr = \frac{N * w}{\sum H}$$

In the equation, N is the number of nodes, w is the number of wavelengths, and ΣH is the instantaneous average hop distance. The logical explanation is as follows: Consider the N node ring as shown in Figure 5-2. If w/2 wavelengths are in each of the two fibers, the total number of active wavelengths between any two adjacent nodes is w. If we assume that there are w/2 lightpaths from each node to both its adjacent nodes in clockwise and counterclockwise directions, the total number of lightpaths that can be established is N*w.

Note that for the preceding statement to be true, each lightpath must be exactly one hop long. This is seldom the case. Lightpaths are generally of arbitrary length from 1 to N – 1 hops (including protection LPs). Therefore, we have to scale the value N*w by the average hop distance of the established lightpaths. An interesting argument is to find out how many lightpaths can fit onto a single wavelength. In one analysis using load balancing and embedded topology design, an average of 2.43 lightpaths could be made to fit onto a single wavelength in a two-fiber ring. The traffic distribution was both uniform and random.

In contrast, for a metro core network that has a significant portion of traffic going to the long-haul network such that the metro-to-long-haul traffic is sent through a node that performs hub-like features, the number of lightpaths onto a single wavelength is about 2.05. In other words, wavelength reuse for a metro network with N between 6 and 14 causes capacity to be about twice the number of wavelengths. This analysis can also be verified in Bidirectional Line Switched Ring (BLSR) kind of SONET networks.

Long-Haul Networks

The final tier in our three-tier classification of optical networks is the long-haul network. Traditionally, these long haul networks were the first networks to have fiber as a transmission media on account of the fact that signals can go for a long distance in a fiber and are comparatively error free. Long-haul networks are typically regional or intercontinental networks connecting different cities or even continents together. The networks can span up to a few thousand kilometers. Long-haul networks were first built using fiber because fiber provided for transport of raw data bits with few errors and through higher bit rates. How-ever, as the bit rates increased, different optical parameters had negative effects on their transmission lengths. Therefore, long-haul networks are severely impaired by issues such as attenuation, dispersion, and nonlinearities.

Typically, long-haul systems are point-to-point systems with regenerators at each end and in between; this preserves signal quality because the signal reaches the far end of a transmission link. In some deployments, the point-to-point links have been scaled to form mesh-like architectures such that the irregular mesh topology can facilitate a huge surge in network traffic demands. As the requirement to transfer data increased substantially, the impairments offered due to communication in the optical fiber also increased. In addition, there was a limit to the amount of data (data-rate) that a single optical channel could carry. This brought about the concept of dense wavelength division multiplexing (DWDM), whereby many channels were interleaved or multiplexed together to produce a dense composite signal and have huge bandwidth in the same fiber.

Also, a change was being brought about in the subsystems that facilitate such long-haul networks, such as Raman amplifiers, closely spaced AWGs and so on. Another area of active research was optical cross connect technology, which was used to cut through an intermediate node. Although it was attractive and conceptually brilliant, this vision of optical (wavelength) routing proved quite ineffective for present networks.

Submarine networks are typically undersea networks used to transfer data across continents. They are also point-to-point long-haul links with repeaters to ensure signal power levels and quality. The main issue associated with such networks is that restoring fiber-optic undersea lines after fiber cuts is almost impossible. In addition, fiber-laying costs are the main component to the network cost in such networks.

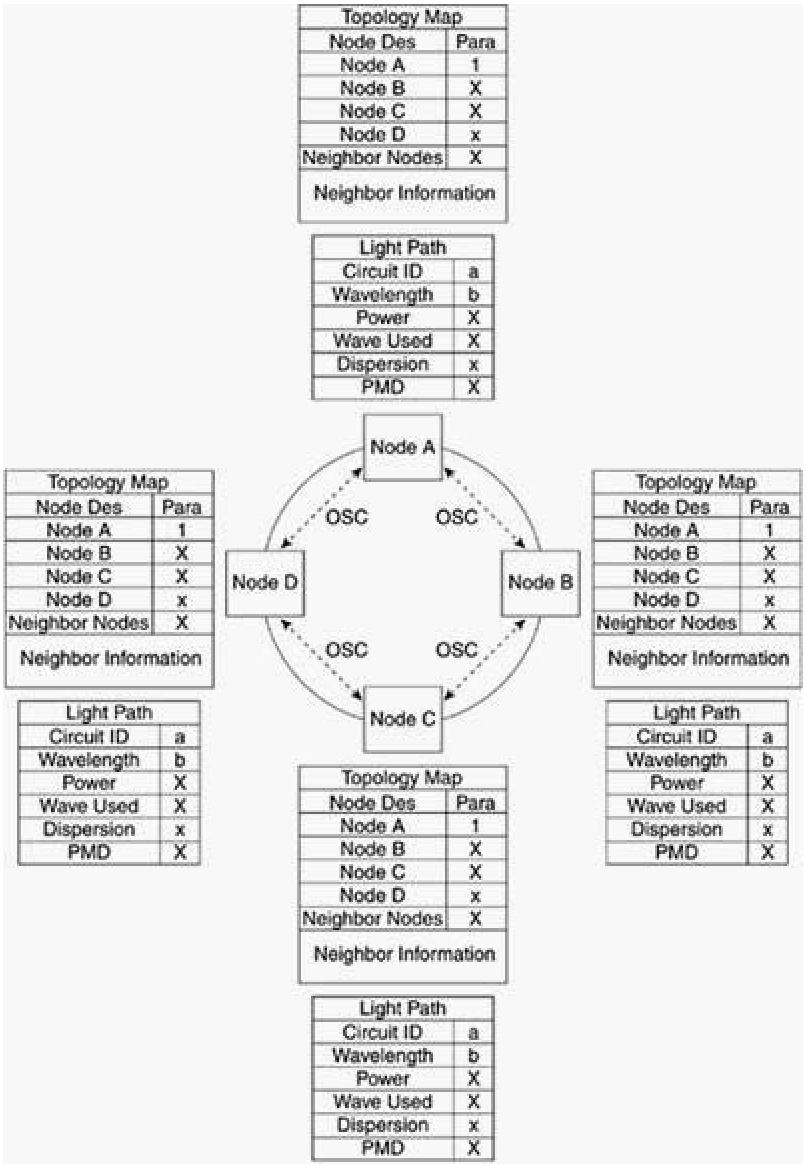
Nodal Architectures and the Optical Service Channel

The most important element in a network is the node. For WDM layered hierarchy, different configurations have been proposed, each adhering to some different set of requirements. In fact, the access, metro, and core networks have their own unique combination of nodal architectures.

A good nodal architecture is characterized by low insertion loss, low polarization-dependent and wavelength-dependent losses, high scalability (to number of wavelengths), and good switching prowess. In addition to the fore-mentioned desired characteristics, a node should also be able to configure and communicate to other nodes about lightpath establishment. The node does this through the network management system (NMS), which is typically built to provide end-to-end *provisionability* (ability to provision the network) in the network. The service user controls the nodal elements by using the element manage-ment system (EMS), which is like an operating tool to operate the node locally in various scenarios.

The NMS accomplishes fast lightpath establishment (setup) and teardown and involves other network-level issues. The node communicates to the other nodes about local position (traffic, states, and so on) through an out-of-band /in-band channel called the *optical service channel* (OSC), which is generally a channel that drops at each node. Shown in Figure 5-6 is the implementation of out-of-band OSC. OSC is like the nervous system of the network; it informs the constituent nodes of the network details at any given point. Typically, OSC carries information pertaining to routing where the routing is based on shortest path algorithms (SPFs). OSC is used to discover topology and identify neighbors. OSC is also used to provision, restore or protect the network. The node builds a topology table and lightpath table using the data collected during the discovery process. Finally, a routing table is derived for automatic routing of lightpaths. OSC also reports the network performance and helps to make decisions during lightpath establishment.

Figure 5-6. Use of OSC in Metro Rings



OSC does this reporting and information broadcasting and makes these decisions by propagating the table of lightpaths, wavelengths used and unused, source, and destination nodes for these lightpaths at all times. Typically, OSC is at a much slower line rate of about OC-3 (155 Mbps), but some OSCs have been reported at an OC-12 speed, although there is no standard for the OSC speed; it depends on the network requirement. The low speed can be attributed to two main causes: first, there is not much information to be conveyed to every node; second, the OSC has to be dropped at each node, so the opto-electronic interface needs to be low cost.

Higher bit rate electronic cards are expensive, so OC-3 is a good alternative speed. The OSC is typically at 1510–1520 nm, and we do not need to define it as per ITU channel spacing. The importance of OSC is growing due to the rapid standardization of signaling procedures for establishment of lightpaths. General multiprotocol label switching (GMPLS) is a new implementation and a paradigm to signaling in optical networks that is implemented using the optical service channel. As optical networks become more mature, the OSC will play a more prominent role in dictating the network heuristics and revenue earning capacity.

Nodal Architectures for Different Network Markets

Long-haul, metro, and access networks have their own distinct nodal architectures suited for their own applications. The architectures are more clearly defined for each market type because of the cost factor involved in them. For example, a long-haul node might have to optimize the signal for all its optical impairments such as loss, dispersion, and nonlinearities, whereas a metro node might have to care only for loss and dispersion. Finally, an access node might have to bother just with loss, not caring for either dispersion or nonlinearities.

This kind of differentiation is the main motivation behind having different kind of architectures in the different markets. Naturally, the functionality in most long-haul networks makes long-haul nodal architecture the most expensive to deploy. Certain important characteristics make nodal functionality an important subject of study. A node is generally desired for its ability to insert traffic into the network and to drop locally oriented traffic to the client layers off the network. In addition, a node might perform certain other functions such as traffic passthrough (act as a passthrough tunnel for traffic) and traffic grooming.

With this generation of WDM networks, we initially envisaged a short time of deployment of optical packet-switching node. The absence of an optical RAM and the sheer uneconomics of the business proposition renders packet switching in the optical domain a wasteful technology at this time. It will probably take a decade for commercial packet-switched networks in the optical domain to be deployed, although small-scale university projects might precede that date on a pilot basis.

A typical node can be broken into four discrete sections. The first section is usually the input section. Here, a signal is fed to the node (often through a preline amplifier—which boosts the signal power—and maybe also through sections of dispersion compensating units, or DCUs, which compensate for the accumulated dispersion). Furthermore, in the input section is a demultiplexer that can segregate the composite WDM signal into individual wavelengths or into bands or clusters of wavelengths.

The next section of the typical node is the switching section. This is by far the most important part of the node. Long-haul nodes can have a complete optical cross-connect device such as an optical switch. An optical cross-connect (OXC), as the name implies, is a cross-connect that is built on some optical technology (such as MEMS, discussed in Chapter 2, "Networking with DWDM -1"). This cross-connect can switch any wavelength from one port to any other port. Of course, this kind of switch also needs wavelength conversion at the ingress or egress ends to be able to support fixed channel-spaced AWG multiplexer sections at the extreme ends.

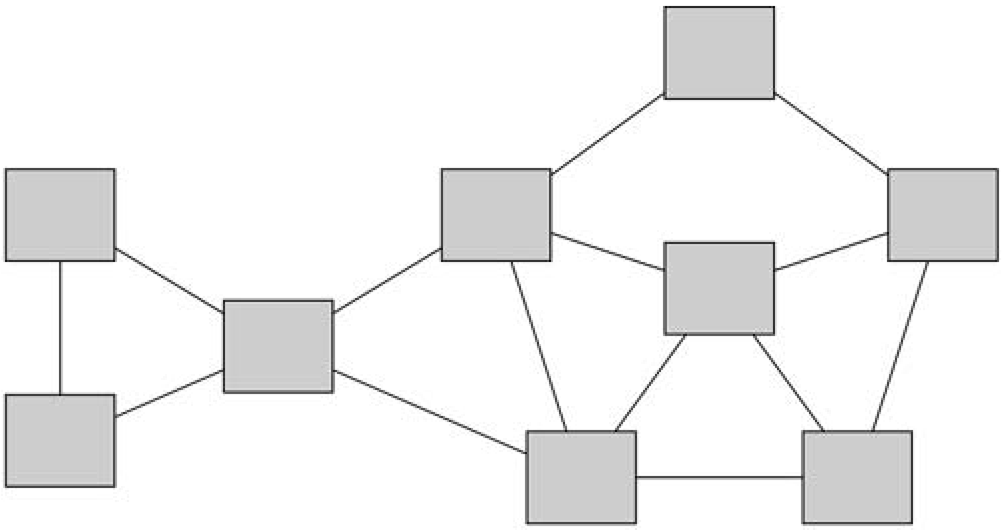
Metro nodes, in contrast, might not have OXC features. Metro architectures are generally ring based; therefore, a simple add, drop, passthrough switch on a per wavelength is enough to provide the necessary flexibility and granularity. Every node has only one pair of fibers entering and one pair leaving the node because of the dual degree of connectivity of the ring. In mesh architectures, the number of fibers that enter/exit a node might vary based on the degree of connectivity of that node. Access architectures might not even need switches; they might be implemented using band filters or passive star couplers. The point to note here is that each architecture has its own advantages and disadvantages, but most importantly, each architecture has its own optical characteristics. Each architecture might pass a different set of wavelengths, have different losses to different wavelengths, different polarization profiles, and different inducements of nonlinearities.

In short, each architecture is versatile in its own right and has some fundamental issues that affect the system design. When an optical system is designed, we have to keep in mind the effect(s) of the node on the system design. Most important is the loss budget. Each node has a severe loss of signal. In the passthrough case, a signal might have up to 16 dB loss as in the case of AWG-based elements. This loss is more than made up for by the postline amplifier, which reboosts the signal power before feeding it into the transmission channel. Therefore, when we are designing a system, we have to consider these factors one by one.

Long-Haul Nodal Architectures

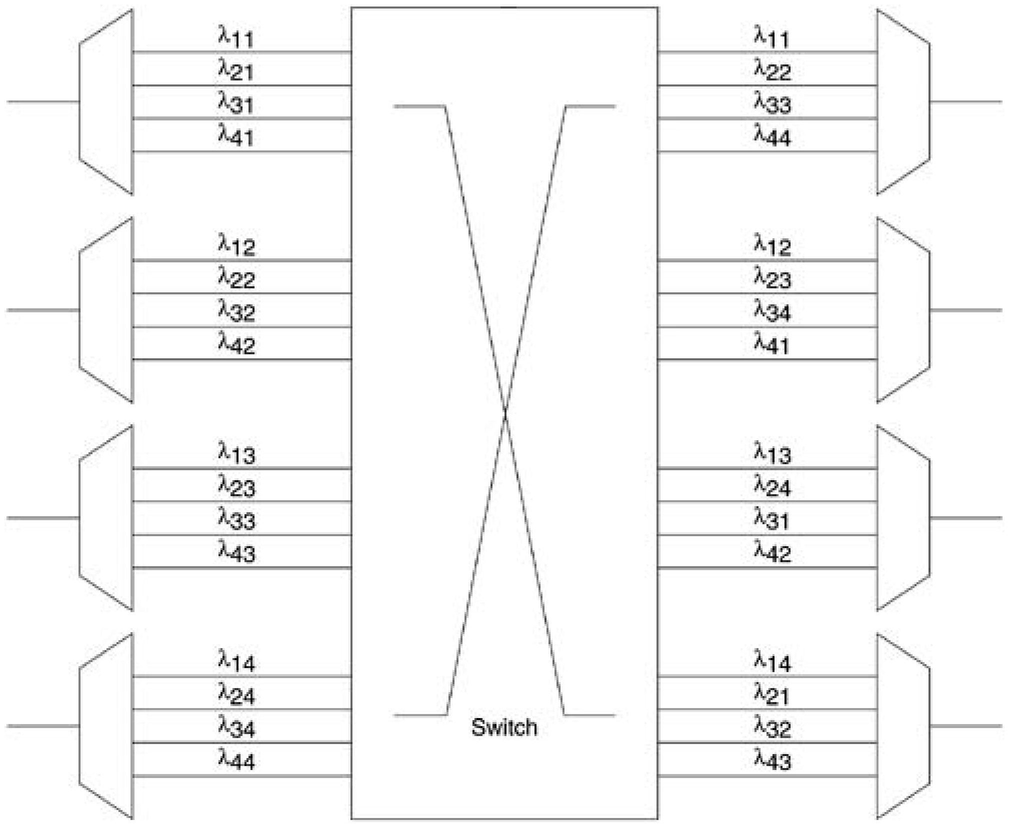
Long-haul nodes are characterized by their nonflexible architectures, which inherently minimize or take care of optical impairments. Typically, long-haul networks are mesh or point-to-point links; therefore, the degree of connectivity of a particular node (the number of neighbor nodes connected to an incumbent node or simply number of pairs of fibers attached to the node in a duplex system) is greater than 2 (as shown in Figure 5-7).

Figure 5-7. A Typical Regional (Long-Haul) Network



This means that a node can be attached to two or more different nodes through the fiber lines. This though is no hard and fast rule, but this high degree of connectivity needs a cross-connect like architecture to facilitate what is called wavelength routing. W*avelength routing* is a routing concept that is extended to the optical layer so that wavelengths (and lightpaths) are routed based on the wavelength granularity and wavelength separation. A wavelength router (see Figure 5-8) is a device that can route a particular lightpath to a particular node (through a predetermined port) based on just its ingress wavelength.

Figure 5-8. Wavelength Router: Note the Static Wavelength Routing as Per the Suffix on the l

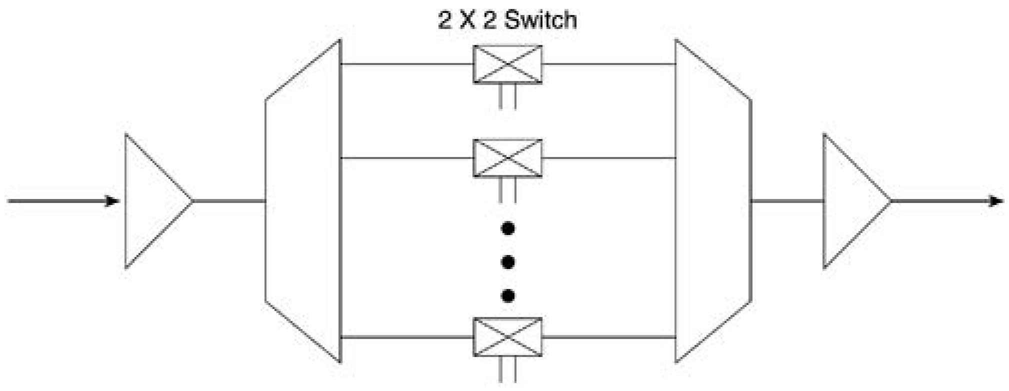


Although wavelength routing is cumbersome to implement, it is a promising approach in optical shortest path routing, and we can easily carry it out by using OXC architectures. One of the key features of OXCs is their all-optical nature. An OXC can switch any lightpath at any port to any other port as long as there is no blocking and wavelength continuity is maintained. A suitable candidate technology for this is MEMS technology. OXCs are also being developed to facilitate new architectures for burst and packet switching. However, the progress in these fields is quite slow. Typical switching time is a few milliseconds, whereas port-to-port loss is about 3–9 dB depending on the OXC configuration and size (remember we have not added multiplex section losses). We can also deploy OXCs using O-E-O technology. Here, the signal is processed and switched in the electrical domain.

Metro Network Nodal Architectures

In metropolitan-area networks (MANs), the main architecture is that of a ring. A ring network does not need cross-connects, although they are always beneficial. Conventionally, ring network nodes are implemented by add-drop or drop and continue kind of devices (like OADMs) such that wavelengths can be added or dropped at each node. Note that no feature exists for all-optical switching (cross connect). Nodes have a typical full demultiplex section that disseminate the composite signal into individual channels. Each channel has a 2x2 switch, as shown in Figure 5-9. The switch can either add-drop or pass-through lightpaths (optical signal). Possible implementations are mechanical switches, which are typically slow but reliable.

Figure 5-9. Metro Node Architecture: Note the Pre- and Postline Amplifiers and 2x2 Switches for Efficient Add-Drop



As of today, most networks are circuit switched, so they work well with mechanical switches. Mechanical switches have a typical switching time (from one state to another) of 10–15 milliseconds. This is enough to ensure 50-millisecond restorations, as in SONET-like networks after fiber cuts. Metro networks can also be built without full demultiplex sections.

This method of implementation is through the use of thin film filters (TFFs). Thin film technology is mature and robust. Moreover, there is some degree of flexibility because any channel can be dropped by using a tunable TFF. The main advantage is in the nodal loss. Nodal loss creates a substantial loss in optical power. In arrayed waveguide (AWG)-based nodes, the nodal loss can be as high as 15 dB.

By using TFFs, this loss can be trimmed to just a few dB. Typically, the insertion loss of one filter is about 1.2 dB. Therefore, a series of TFFs has a lower loss compared to conventional AWG design. Another optimized architecture is by using subbands. By creating bands in the operating spectra, a node can drop an entire band (of wavelengths) from the line (composite signal). This band can be further demultiplexed by client equipment. Therefore, loss is low, and there is some degree of flexibility. This kind of network architecture cannot take large dynamic variations in traffic; in other words, it cannot take excess churn. The tradeoff is the cost involved. Some vendors have managed to make products based on such TFF technology.

Another method is that of broadcast and select networks. Passive stars and broadcast architectures are becoming quite popular because of their cost and performance metrics. However, these (required for efficient sharing of the medium amongst different users) networks are quite academic because medium access protocols are hard to implement successfully over large distances in optical environments.

Although filter-based technologies offer optically better design in terms of low node loss, they are limited in flexibility of operation. We cannot dynamically drop a channel from a band filter. Each band filter is able to drop only that specific band. Therefore, to create a dynamic drop-capable system, we need a large inventory of filters, which might not be an economical solution. Moreover, placing such filters in the transmission channel during active operation can create protection surges, which limit system performance. Also the filtering window is not always a perfect rectangle, but may have a profile which can be lower order Lorentzian or Gaussian creating losses for high rates.

Protection in WDM Networks and Protection Switching

At the optical or WDM layer, a lightpath is set up using a control mechanism that involves the ingress, the egress, the intermediate nodes and equipment, in addition to the fibers to set up the lightpath. Failure of equipment/fiber/nodes can cause the lightpath to be disrupted, resulting in a huge loss in revenue. Protection of lightpaths is a means whereby failure of fiber or equipment can be surpassed by other means. Protection essentially means adding some degree of redundancy or diversity to the network. The excess redundant portion of the network is utilized to provide network capacity in the event of a failure or fault.

In principle, protection is a fast phenomenon such that the failed lightpath is switched onto the excess allocated capacity in the shortest interval of time. An optical network has various kinds of failures. The most common failure is that of equipment. Equipment failure accounts for almost 70 percent of all failures and is a result of various factors such as aging, malfunction, and human error.

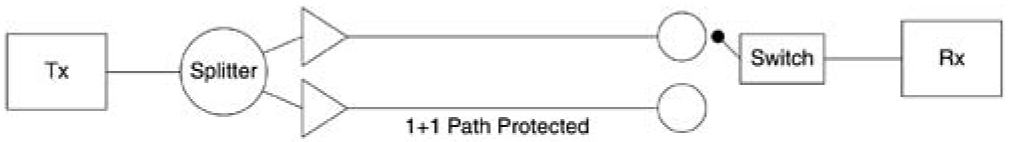
Protecting networks from equipment or subsystem failure is a difficult task. The only way to protect a network from equipment failure is to deploy redundant equipment as protection equipment and to switch from the normal equipment to the redundant gear in the event of a failure. Certain algorithms actually describe and facilitate the way in which the signal is transferred from the normal equipment to the redundant equipment (protection equipment) and also describe when this change is to be done. Equipment failure can also be the entire node failing. This might happen due to power outages or even human errors. Nodal failure is difficult to protect, and unless the failure is partial, the node is almost down until rectification exercises are carried out. A node failure can be regarded as multi-link failure. For example in a ring network a failure of a node, is as good as failure of the two attached links.

The next common failure after equipment failure is fiber failure, which is more commonly known as*fiber cut*. Fiber cuts are failures that are caused by the damage rendered to operational fibers either from physical cuts or from severe bending (thus increasing their losses to unbearable values). We can take care of fiber cuts by using redundant fibers along with work or normal-use fibers and switching signals from the work to the redundant (protection) fibers. Protection in WDM networks can be classified into two types: *line*- and *path-level protection*. Line protection means protecting the entire fiber or the entire band of WDM channels in the event of a fiber cut or failure (see Figure 5-10).

Figure 5-10. A 1:1 Protected Link: Note Both Transmitter and Receiver Have to Switch

In contrast, path protection means according protection to just the single lightpath that has failed (see Figure 5-11). Therefore, path protection is more specific and more difficult to implement, whereas line protection is more generic and easier to implement. For most of the discussion on protection and its mechanisms, this chapter considers fiber cuts as the predominant failure in optical WDM networks because of the simple redundant techniques present to solve the protection issue for equipment failure.

Figure 5-11. 1+1 Protected Link: Note That Only Receiver Has to Switch to Protection Signal

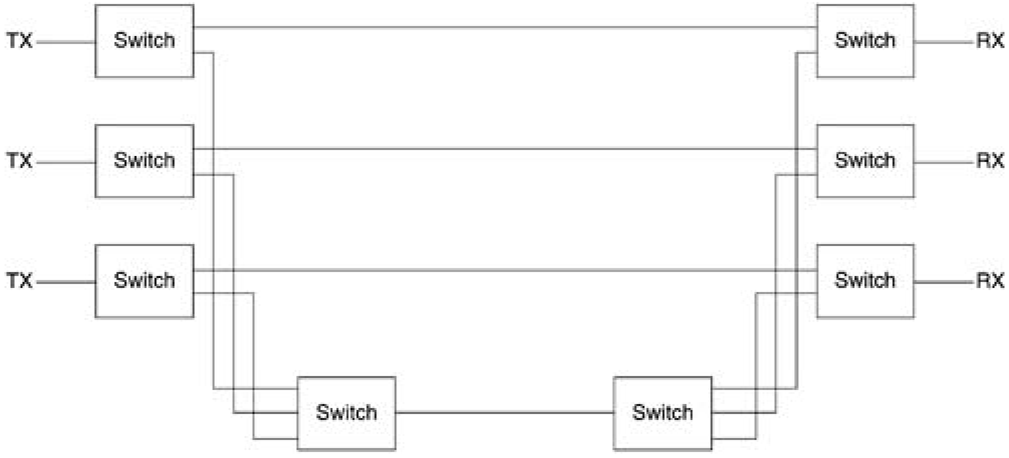


In conventional SONET networks, line and path protection has been incorporated into the ring topology by two distinct schemes: unidirectional path-switched ring (UPSR) and bidirectional line-switched ring (BLSR). We can extend the same scheme to WDM networks over a variety of topologies. For path-switched protection in point-to-point or mesh topologies or even WDM rings, the protection is known as *1+1 protection*; on the other hand, for line-switched protection, the scheme is called *1:1 protection*.

For path-switched 1+1 protection, the transmitter or ingress node transmits the signal (WDM lightpath on specific wavelength-lambda) into the work (normal) as well as the protection (redundant) path simultaneously. At the receiver or egress node, the receiver chooses the signal from either the work or the protection path, depending on the signal quality. Essentially, this kind of scheme is easy to manage because the changeover decision (for choosing either work or protect fiber) is taken only by the receiver section and not be multiple sections. Note that it is best to have the work and the protect paths on different fibers such that exulting physical diversity.

In contrast, in the 1:1 or line-switched scheme as in Figure 5-10, the signal is sent only in the work section or work channel while one protection channel serves as a backup to multiple geographically diverse work channels. In the event of a failure, both the sender and the receiver need to coordinate with one another and switch the signal in the failed section to the protection channel. This involves dual switching at both the transmitter and receiver sections, as shown in Figures 5-10 and 5-12.

Figure 5-12. 1:N Protection (Shared Protection)



Naturally, the 1+1 protection format is a much easier way of guaranteeing fast protection (restoration time is minimal)q, but the cost involved is often more as more bandwidth (resources) is needed to facilitate it. 1:1 protection, in contrast, is not that fast in terms of restoration time (time needed to restore a failed link or lightpath), but the cost involved is often quite lower even though the signaling procedure is often cumbersome because of the fact that there is full duplex signaling (communication) involved.

However, 1:1 protection has the added advantage of optimizing the protection bandwidth among many work paths. This kind of scheme is also referred to as *1:N protection*, whereby one protection path can provide protection of a single fault in any of the N fibers (one fault at a time). See Figure 5-12. When the protection algorithm is 1:1 (that is, every channel has a given resource to ensure protection), the scheme is also called *dedicated protection* or more technically, dedicated 1:1 protection. If the protection algorithm shares many work channels for a single protection channel, this kind of scheme is called *shared protection*. Shared protection envisions one protection channel for N work channels, assuming that only one of the N work channels will be cut (lost) at any given time. Shared path protection is also an interesting concept for according protection to individual lightpaths in WDM metro rings. One issue of ongoing research is that of optimizing the amount of bandwidth allotted for protection as compared to that for work, in a 1:N protection network. This ratio is called the share to work (S/W) ratio.

Protection for Different Qualities of Lightpaths

The simplest form of protection is the dedicated protection scheme in which every lightpath has its own complement of protection bandwidth allocated to it such that in the event of a failure, the lightpath is compactly protected using the basic principles of 1:1 or 1+1 schemes. One major issue in protection is the time needed for protection to take place. When we protect a failed lightpath or link, we need to run an algorithm that actually determines the failure, allocates resources to the failed lightpath, and coordinates the transmitter and receiver on these resources. This takes a finite amount of time.

Previously, we learned that 1+1 protection is much faster and seamless than 1:1 protection. Of course, the tradeoff is the cost involved. In voice networks and SONET models, the protection time must be less than 50 ms. This is difficult to achieve if the protection resources are quite dynamic, in the sense that multiple algorithms need to be run and decisions need to be made to allocate the protection resources. In ring networks, the idea is to have the work lightpath in the shortest path and to have to protection lightpath in the longer path (which is invariably now in an opposite direction to the work path). This means that a good way to deploy WDM rings is to ensure that there are two fibers: one in a clockwise (CW) sense and the other in a counterclockwise (CCW) sense. We can do this by keeping half the channels in a CW direction and the other half in a CCW direction. Then we need to solve the problem of wavelength congruency of work and protection traffic. As it is easier to keep work and protection lightpaths on the same fiber.

Protection channel access (PCA) is a class of traffic that has a lower priority than work traffic. When the network is healthy (no failures), the bandwidth that is allocated to protection can be used to deploy additional lightpaths. However, when a failure occurs, this class of low-priority traffic is the first to be discarded. This class of traffic is also simply considered unprotected traffic. Many low-cost applications today use a PCA traffic scheme. Because fiber cuts are quite rare, PCA-like schemes maximize the capacity of the fiber by efficiently utilizing all the channels.

Bell core document 2979 specifies the requirement for OADM and other equipment from the protection point of view and mentions that we consider only a single failure per network. This is not always the case. Multiple failures can occur because of overloading or human error. Limited algorithms take care of multiple failures. This is an area of constant research especially in mesh networks because two or more failures in ring networks break the basic continuity of the ring network elements.

Mesh Protection

In rings, the protection algorithms are simple. The shortest path is the work path, and the corresponding longest path (on the same λ) is the protection path-in complementary direction. Mesh networks might contain multiple protection paths. It is important to choose the best protection path. This decision might include some aspects of dynamic load balancing, as well as route selection. In one embodiment, protection cycles were created in an N node mesh topology such that for a given cycle, the work path was a chord to the cycle and the protection path was the arc subtending the chord.

Cisco has implemented the path protected mesh network (PPMN) philosophy in its SONET/SDH product, which essentially allocates a protection path to each work path across a series of links that serve as constituents of a lightpath. In ring networks, the protection bandwidth is equal to the work bandwidth. In mesh networks, the protection bandwidth does not need to be as high as the work bandwidth because multiple paths exist. Algorithms have been proposed to optimize the amount of bandwidth placed for the work and protection sections. In one embodiment, the protection bandwidth was 20–60 percent less than the work bandwidth. Although this seems to be quite a specific case, such scenarios do generally exist. However, it is difficult to formulate an optimum number for protection bandwidth for WDM mesh networks; it depends to a great extent on the topology and traffic patterns.

Signaling and Protection in WDM Networks

When a failure occurs in a WDM network, three things need to be done in the least lapse of time. First, if the network fails—whether it is a fiber failure or an equipment failure—we need to identify the fault. This is generally done by using the EMS. Second, upon detecting a failure, we need to broadcast the fault across the network to at least the nodes or terminal equipment that are affected by the fault. This is generally accomplished by using the optical service channel in conjunction with the NMS. Finally, we need to protect the fault. If it is a fiber cut, then line or path or both protection mechanisms need to be invoked. If the fault is an equipment failure, then the redundant equipment is switched to. Finally, after the fault is repaired (which might take a long time, as with a fiber cut), the network is reverted to its original state. The entire procedure is based on the signaling and control mechanism of the network. Different methods of signaling have been demonstrated in literature. In one embodiment, GMPLS-based signaling is used to restore the network from failures. In yet another embodiment, standard SONET-like automatic protection switching (APS) is invoked.

Optical System Design

Chapter 4 considered various aspects of system design for generic WDM networks. We built networks based on BER or OSNR requirements. We also learned about dispersion and attenuation budgets while designing optical networks. So far, this chapter has focused on different topologies such as rings, mesh, and point-to-point networks. Most of the design models inChapter 4 were based on point-to-point networks.

The reader might have noticed that in Chapter 4 we did not take into account losses or, for that matter, the effects of nodes or network elements on the propagating signal. This chapter has classified three basic networks based on the hierarchical model: access, metro, and long-haul networks. Each category has a different nodal structure; therefore, the effects rendered by each network element onto the traversing optical signal is also different.

This chapter considers designing each of the three network types, with different kinds of network elements for each type. It is beyond the scope of this book to actually design networks with each possible architecture for the nodes; therefore, the designs in this book are based on generic architectures. More specific or unconventional architectures are beyond the text.

While designing WDM networks, we need to note the following points:

- Classification of the network (access, metropolitan, and long haul)
- Classification of the network topology (ring, mesh, and point-to-point)
- Number of nodes, type of nodes, and nodal heuristics (passthrough and add/drop loss)
- Maximum span loss and adherence of the network to span budget
- Maximum dispersion on the longest link/placement of dispersion compensation equipment (if any)
- Budget by OSNR/BER for the entire network based on worst link (worst attenuated WDM λ), taking into account nodal losses
- Budget consistency for protection on worst (usually shortest work) path
- Nonlinear effects, modulation formats
- Component design

This list might seem quite intriguing for the beginner to adhere to, but in fact, it is a standard way of designing foolproof WDM networks. Generally, a WDM designer faces two kinds of design scenarios: the green-field design and the brown-field design. The green-field design scenario involves designing networks from scratch, such that even the basic fiber is not laid. The designer then has the liberty of choosing routes to lay fiber based on traffic matrices (traffic variations of nodes with respect to one another), as well as to optimize the network for loss and dispersion budgeting purposes.

In brown-field network design, the fiber is already laid and the designer must optimize the network from there onwards. This kind of design is generally more difficult because the designer is faced with many constraints that need to be fulfilled. This chapter considers both design types, but more emphasis is given on brown-field networks for the practical worthiness that is associated with them. Note, however, that all of the design heuristics cannot be manually computed; therefore, there is a definitive need for simulation software.

Access Network Design

Access network topologies can vary from point-to-point buses (as part of a tree) to normal access rings depending on the network requirements (capacity), number of nodes, and so on. This chapter will first consider access rings. Access rings are generally hub-and-spoke rings whereby a central hub node is situated and acts as a sink of access traffic and a gateway to the metro (or metro core) network. The traffic matrix is generally unipolar in the sense that all nodes have full duplex traffic requirements with the hub node. There might be some intra-ring or node-node communication within the access network, but this is generally quite low.

To facilitate and provision lightpaths in the access network, the network design needs to encompass the hubbed traffic in the normal as well as the protected case. For most access networks in the optical domain, the distances are less than 75 km in circumference, and longest spans are usually not more than 40 km. Access networks are situated in clustered, heavily populated Manhattan-like areas; therefore, typical distances are much less than even 40 km.

Access Node Architecture

The hub node has a completely different architecture than the individual access nodes. This is because it is unnecessary to demultiplex the entire composite signal at each node. On the other hand, the WDM signal needs to undergo full demultiplexing at the hub node; therefore, this architecture conforms to the architecture of a full-fledged OADM as described toward the end of Chapter 1. The nonhub nodes (access nodes) have typical architectures that facilitate for low loss (add/drop or passthrough) as well as low cost.

Typically, the access nodes do not need much capacity for add/drop traffic. Therefore, from the design point of view, the access nodes must offer the least loss for the channels that pass through. One possible implementation of access nodes is using thin film filters (TFFs). Numerous vendors have proposed this configuration, which is shown in Figure 5-4. The TFF is the heart of the nodal apparatus. It has three ports: an input port, a rejection port, and a transmission port. The entire composite signal is fed to the TFF input port. Depending on the filter configuration, some wavelengths (a band of wavelengths or even a single wavelength) are dropped at the rejection port. The remaining signal less the dropped band is available at the transmission port.

Typical passthrough loss for the filter is about 2 dB, which is negligible compared to the OADM counterpart in metro rings. The dropped band can be further demulitplexed by using either a splitter and several filters or a low cost, small range (4–10 nm) demultiplexer (AWG). Note that the demultiplexing operation occurs outside the flow band; that is, the channels that do not need to be dropped are not sent through the walls of the AWG used as a demultiplexer. Generally, the AWG losses are much more than the passthrough losses (of the TFFs); therefore, this does not become a cause of concern for passthrough channels.

By using an arrangement whereby the drop channels are demultiplexed outside the flow of the composite signal, loss is kept to a minimum. This arrangement also lowers cost significantly. This arrangement is good for access networks but cannot be replicated to higher traffic metro core networks because of the low scalability and low flexibility for dropping an arbitrary wavelength (lightpath). In access networks, it is often not necessary to consider dispersion-limited or OSNR-limited designs because there might not be sufficient noise or dispersion impairments, and distances would often be less than what was needed for maximum allowable accumulated dispersion. Even if we consider a 40 km access ring and use few amplifier stages, the dispersion does not become a dominant issue in access system design.

In access networks, the ring topology is not the only dominant topology; bus and tree topologies are also quite normal. If the medium is a bus or a shared bus topology, the network elements (nodes) generally are passive, typically consisting of passive devices such as splitters and taps. A typical access network is characterized by many nodes (4–18), and each node has a small add/drop traffic requirement. Due to the low-capacity requirement of such networks, wavelength reuse is not essential in access networks. Wavelength reuse is a simple yet efficient way of increasing the net capacity of a network given the limited number of wavelengths, by repeating wavelengths for lightpaths such that no two lightpaths on the same wavelength will share the same physical fiber. For access networks—especially rings—wavelength reuse might not be required because the maximum traffic capacity (in number of lightpaths) can always be less than λ_{max}, the total number of wavelengths in the network.

When access networks are implemented in tree/bus topology with passive elements, this kind of network setup is typically referred to as passive optical network (PON). Here, a coupler splits the power from a fiber, and each access node gets a portion of the power. This system creates a multicast group from the central office to the end users. (Refer to Chapter 7, "X over WDM," for more information on PON.) The power splitting can be mathematically explained as follows: If the splitter splits the input fiber to N nodes, and if P is the input power, then the output power to each user node port is shown in Equation 5-2.

Equation 5-2

$$\text{Output power} = P(dB) - 10\log N$$

Metropolitan Area Network (Metro Core) Design

Metropolitan area networks (Metro core) today are the most popular optical networks in the industry because of their tremendous revenue potential. Traditionally, these networks were all voice based, but they are now becoming more data centric. Due to the over-indulgence in SONET/SDH as well as the strong potential in ring topology, metro networks are primarily configured in rings. The topological superiority of rings makes metro networks a truly survivable architecture for long-term applications.

Metro networks can be further classified as metro core and metro collector networks (both ring topologies). Metro core networks can be distinguished by the enormous capacities involved. They are generally linked at one or more nodes to long-haul terminal equipment. In contrast, metro collector networks can be seen as aggregators of metro access traffic. Metro networks have various service classes from legacy SONET on some wavelengths with speeds ranging from OC-3 to OC-192 to Gigabit Ethernet and 40 Gbps experimental lightpaths.

Chapter 6, "Network Level Strategies in WDM Network Design: Routing and Wavelength Assignment," explores the classic problem of routing and wavelength assignment for lightpaths for a given traffic using the minimum resources and under some given constraints. Wavelength assignment by itself is a serious problem not just from the network perspective but also from the physical layer perspective; in a multichannel WDM network, proper placing of different bit rate signals is not an easy procedure.

The complexity involved in assigning different length lightpaths to different wavelengths—with each of them at a bit rate that is not necessarily the same as its spectral neighbor—makes wavelength assignment a challenging proposition, particularly at high bit rates (1 Gbps and beyond) because the channel spectral envelope might spread into the adjacent spectral envelope(s). It is beyond the scope of this book to discuss wavelength assignment due to physical layer constraints, although some reference on the topic can be seen in long-haul network design where such effects (such as walkover due to XPM and FWM) are more pronounced.

A metro network is characterized by OADMs in the periphery of the ring. A typical network would demand the following:

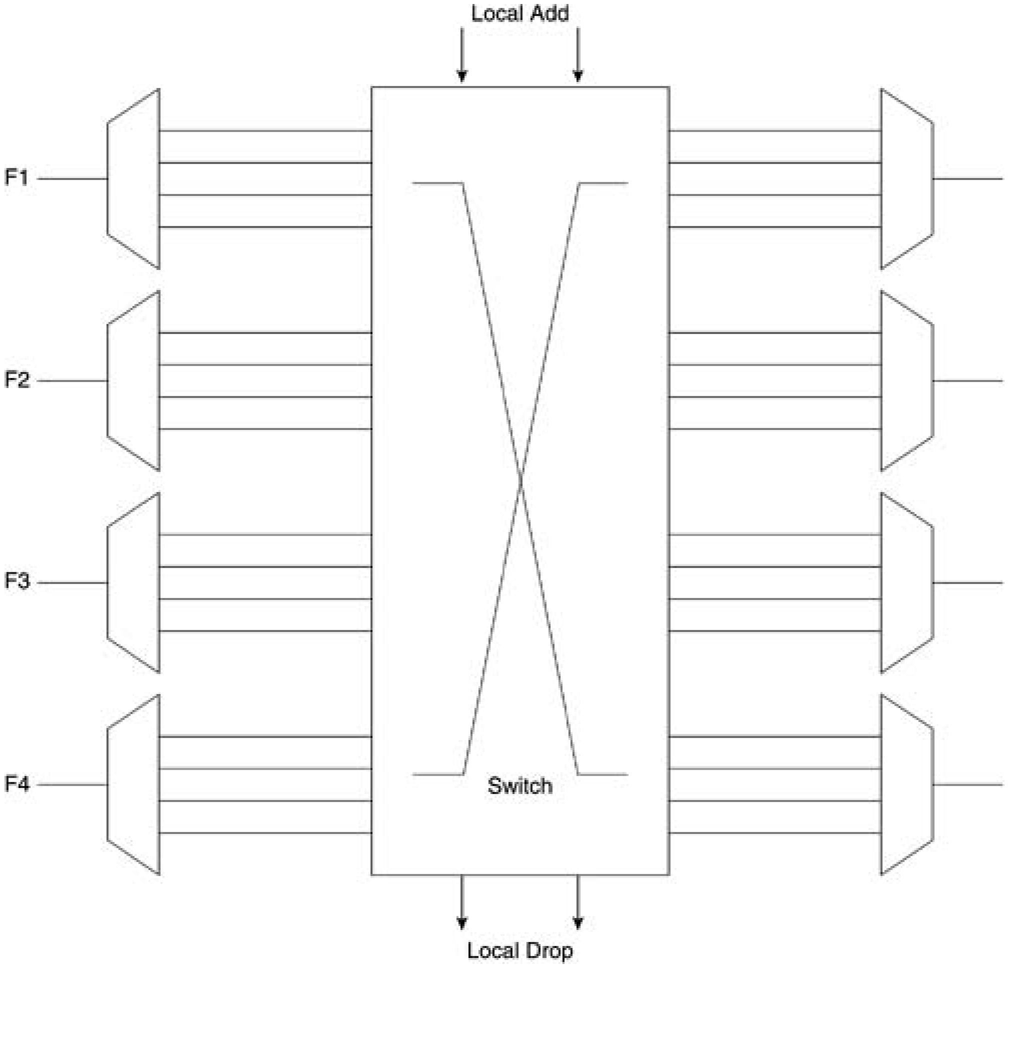
- High capacity (maximum wavelength reuse)
- Good protection and restoration (UPSR/BLSR or even optical shared path protection in rings—OSPPR)
- Network scalability (in-service upgrade)
- Low cost

The OADM architecture can be generically described as an optical multiplex section (OMS) containing some switching elements as well as local access equipment. Chapter 1 briefly described one such OADM configuration (refer to Chapter 1's Figure 1.19). The WDM composite signal is fully demultiplexed and fed to a switching matrix. The switching matrix has various levels of hierarchy such that it can be simple enough to add/ drop channels or complex enough to cross-connect channels and perform some amount of wavelength conversion. Today's metro networks generally need pure add/drop functionality, whereby we can add or drop individual channels depending on the traffic requirement.

Typically, each channel passes through a 2x2 switch such that if the switch is configured in the bar state, the channel-lightpath purely passes through. On the other hand, if the switch is configured in the cross state, we can add or drop lightpaths or even perform both functions of adding and dropping at the same time (assuming low cross-talk between the add and drop ports of the 2X2 switch).

In a more complex scenario, wavelength routing can be carried out, whereby lightpaths can be switched between different ports using a cross-connect architecture, such as the one shown in Chapter 2 for 3D MEMS design. However, if the cross-connect is only juxtaposed for channels that are in the same fiber, wavelength conversion is needed to ensure smooth operation. In contrast, a cross-connect architecture that performs cross-connect functions between fibers (and not just on a pure wavelength basis) does not need wavelength conversion (as shown in Figure 5-13).

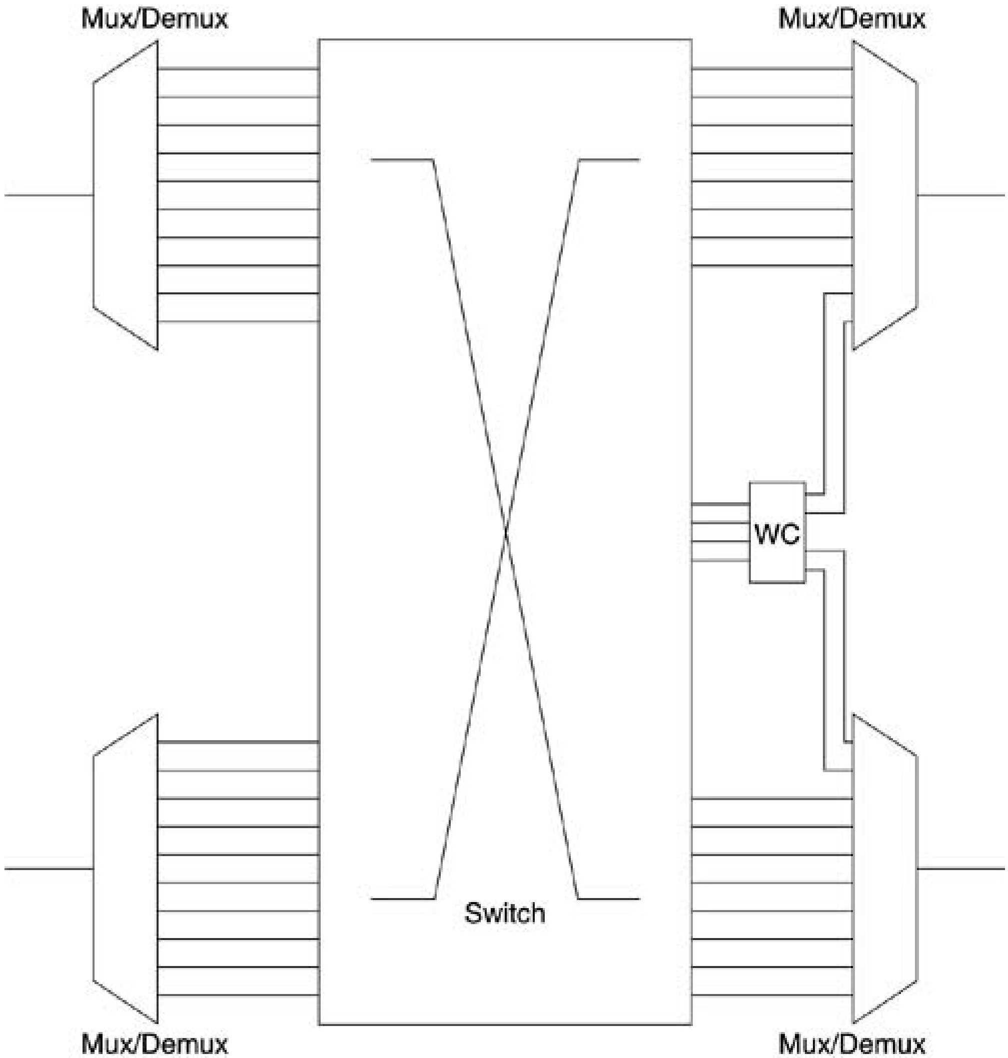
Figure 5-13. Typical Wavelength Router with No Wavelength Conversion



Wavelength conversion as such is quite an expensive technology, and current technology validates only O-E-O wavelength converters, which are protocol dependent as well as bit rate dependent. Some advances have been made in protocol-independent wavelength converters, but implementation is restricted at this time. However, protocol-independent wavelength converters are expected to be more popular in the near future. It is desired that the wavelength conversion technology be transparent to bit rate and protocol and preferably do the conversion entirely in the optical domain. This remains as a serious research topic; semiconductor optical amplifiers (SOAs) that are using XPM and XGM have been cited as possible candidates, but this technology is still years from deployment in commercial systems because of the severe optical penalties in using SOA as a wavelength converter.

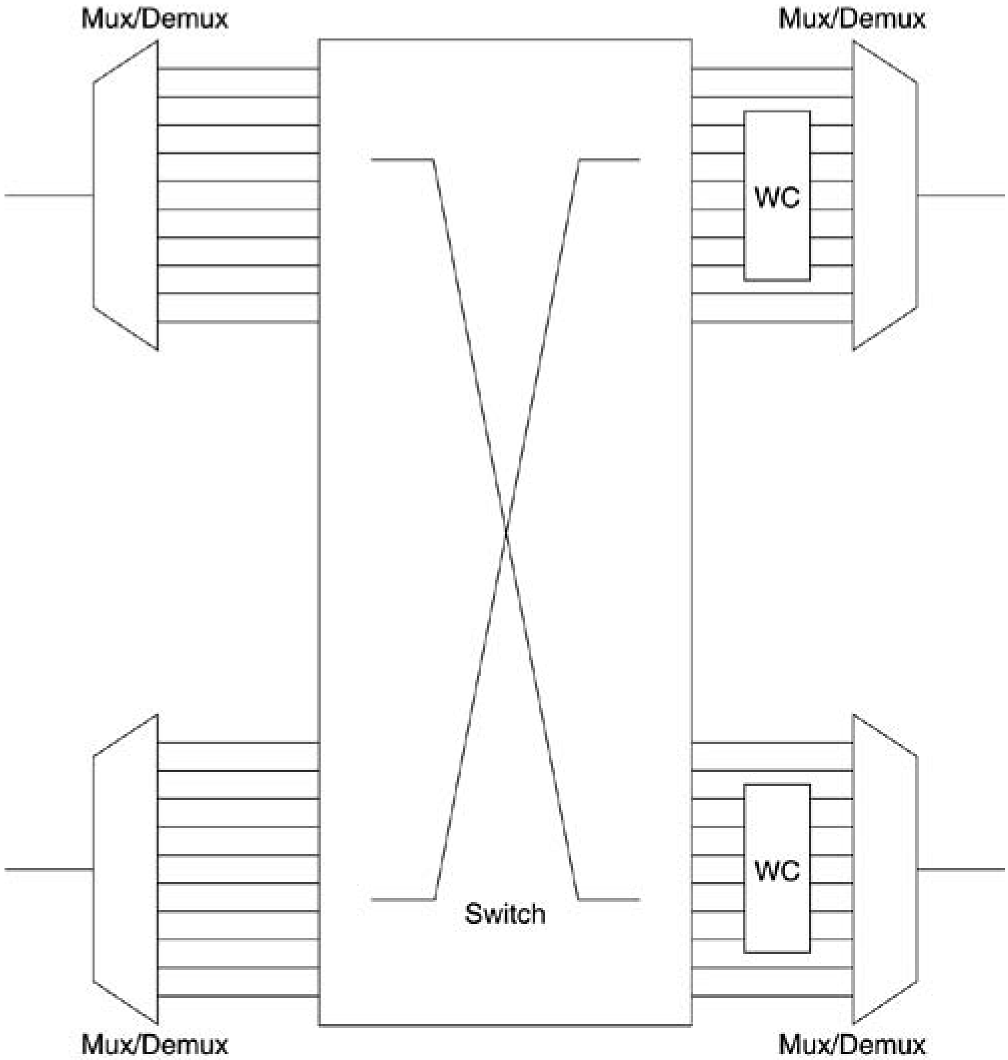
Wavelength conversion initially seemed to be an attractive solution for optimizing the network capacity due to the flexibility associated with it, especially from the routing and wavelength perspective (see Chapter 6), whereby with the use of wavelength converters, wavelength continuity for a lightpath was no longer a constraint. In one such result, it was seen that the benefit of full wavelength conversion (using wavelength converters at each node) was almost identical to partial wavelength conversion (using wavelength converters at few selected nodes). It was also seen that the degree of conversion (the number of wavelengths that a converter can tune to) did not need to be the entire operating band for optimum performance. In one such scheme, the optimum degree of conversion was one tenth the total number of wavelengths in the network. Vendors are muting this kind of limited wavelength conversion for actual deployment. The conversion is done by opto-electronic conversion and regeneration of the signal rather than all-optical conversion. Selective wavelength conversion between fibers or between channels (see Figure 5-14) was also proposed.

Figure 5-14. Wavelength Converters for Sharing Between Two Fibers. A fully nonblocking cross-connect can switch a given wavelength to the WC, which has fixed l outputs connected to the AWG.



Overall, wavelength conversion for metro networks was not a feasible idea. To a great extent, this diminished the advantages of a cross-connect architecture over the conventional OADM add/drop architecture.

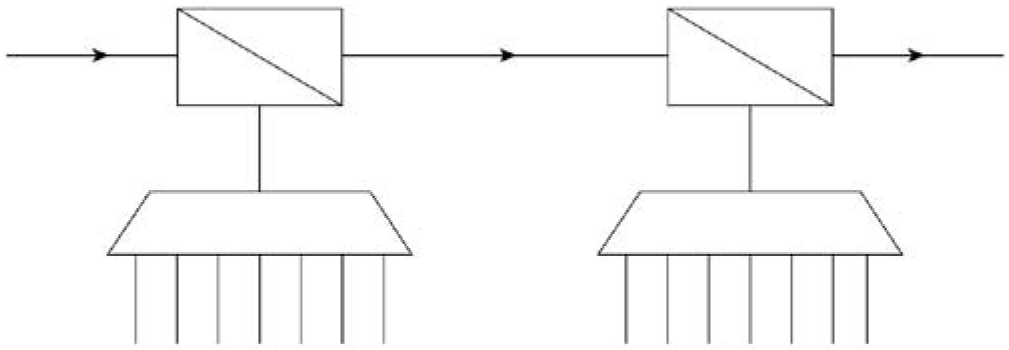
Figure 5-15. Sharing Wavelength Converters (of Small l Dimension) Between Channels in the Same Fiber. This is an economical method of obtaining wavelength conversion.



Yet a third kind of architecture was an extension of access network architectures, whereby the nodes were made by a series of band pass filters. Thin film filter technology is mature enough to be deployed in metro networks. A series of cascaded filters suffices the tasks required for an OADM.

The advantage of such a configuration is the cost involved and the comparatively low loss. As we will see in the design examples later, nodal loss is a major limiting factor for system budgeting and optical design. The apparent drawback is the degree of flexibility in terms of adding or dropping channels. To effectively have flexible add/drop capability for a full C band (about 32–40 channels) WDM system, we need many cascaded filters. This is a management as well as system design disaster. Nevertheless, by deploying fast tunable multichannel TFF, we can have a dynamically configurable OADM architecture (see Figure 5-16).

Figure 5-16. Dynamically Configurable OADM Architecture Note the two band-pass TFFs drop some of the signal to the AWGs, and similarly for adding a band-pass filter can be used. Passthrough loss is only around 4 dB, whereas drop loss is about 7 dB.



When we design OADM architectures for metro networks, we have to consider many issues. The first level of design is generally the power budget. Because there are many wavelengths and many nodes in addition to various bit rates, we need to ensure power budget for the worst channel (lightpath). In Chapter 4, we saw that the upper limit to input (transmit) power is dictated to a great extent by the nonlinear effects that are predominant at high power levels. However, in metro networks, the nonlinear effects do not dictate an upper bound on power level; instead, the maximum power that a receiver (photodetector) can handle sets this margin (receiver sensitivity). Therefore, we must ensure that power levels are within these conformities:

- The power budget needed to ensure that the signal reaches the egress node
- The power level always sustains within the dynamic range of the receiver

The next level of the design is for BER. Chapter 2 explained analytically how to calculate BER. BER by itself is difficult to calculate instantaneously, and a simulation set up on a fast computer would probably take a long time. Chapter 4 described techniques to calculate BER from the OSNR of the signal. BER is directly linked to the Q-factor of the signal. We can perceive the Q-factor as the quality of the signal. See Equation 5-3.

Equation 5-3

$$BER = \frac{1}{Q\sqrt{2\pi}}\exp(-Q^2/2)$$

Further, we can calculate OSNR by using Equation 5-4.

Equation 5-4

$$Q(dbB) = OSNR(dB) + 10\log(B_o/B_e)$$

The only issue in metro network architectures is that span loss and nodal loss might not be constant. A single-line formula (formula 58, which is explained in Chapter 4) is sufficient to calculate the OSNR for the entire network. We have to calculate the OSNR of a lightpath by disseminating the lightpath into its constituent spans. For a given span, the span loss is specified as the actual span loss (attenuation, dispersion penalty, nonlinear penalty, splices, and so on) plus the adjoining nodal loss (node passthrough loss).

Including node loss is important because it greatly dictates the power budget and the OSNR budget. Nodal loss can sometimes be greater than even transmission losses because of the fact that AWG-type demultiplexers and switches have losses amounting to several dB. A typical OADM that consists of full multiplex/demultiplex AWGs separated by a switching matrix may have a typical passthrough loss of 13–14 dB and sometimes as high as 17 dB. The loss per AWG (multiplexer/demultiplexer) is about 6 dB, whereas the loss of the switching fabric is 2 dB for pure add/drop 2x2 switches.

This chapter has not yet considered losses due to connectors, splices, and so on. Therefore, the total passthrough loss might be as high as 14–16 dB. Typically, we can consider lightpath insertion loss at nodes to be approximately half of the passthrough loss, assuming a well-defined granular switch.

One more aspect of design is dispersion-limited system design. In today's networks, dispersion-compensating equipment is quite common. The issue in such networks is in optimizing the costs and placement of the dispersion compensators. In a ring network, depending on the traffic matrix, the designer needs to optimize the placement of the dispersion-compensating equipment for optimum performance. Algorithms can be written to optimize such placements, and this does not pose a serious issue in network design.

Yet another issue that poses considerable problems is the gain tilt of amplifiers. In Chapter 3, "Networking with DWDM -2," the gain profile of an EDFA was shown not to be linear for all the channels; therefore, some channels receive less amplification than others. In a multispan network, this means that the signal-to-noise levels of some channels is different from the signal-to-noise levels of other channels, which creates an issue for optical system design. This issue leads to gain-tilt based design.

The effect of gain-tilt can be understood as follows: For a given WDM network, if channels A through F can exhibit an OSNR of 30 dB, and channels G through M can exhibit an OSNR of just 18 dB because of their negative gain tilt, then the latter channels cannot sustain the BER requirement for the same lightpath length (span budget) as the earlier channels. This issue can be alleviated to some extent by using gain-flattening filters that are cascaded with the EDFAs such that the gain is somewhat uniform. We can also use variable optical attenuators (VOAs) to reduce the extra amplification of the higher amplified channels. We can use equalizers that have feedback signals to adjust some of the gain tilt dynamically. However, despite the use of such equipment, some tilt is always incorporated, and practical scenarios advise us to take into account some extra margin (in decibels) to compensate or work around the gain tilt.

Long-Haul System Design

In long-haul networks, the main thrust is not in optimizing the network level performance; rather it is in transmission characteristics. Transmitting data streams over hundreds or thousands of kilometers is quite problematic in the presence of the various optical impairments discussed in Chapters 1,2, and 3. Long-haul design needs to consider minute details, and power budgets can be quite tight—especially from the OSNR point of view—because the signal undergoes several stages of amplification and amplified spontaneous emission (ASE) noise continues to build up.

Reducing this accumulated ASE noise is a new area of research in today's optical networks. One possible solution is by deploying negative noise filters, although this technology is in its absolute infancy. At such high distances (and such high powers), the optical signal faces severe degradation from nonlinear effects, which were not even considered in metro networks. Self-phase modulation (SPM) and cross-phase modulation (XPM) happen to be two common impairments. This impairment is due to the beating of the signal frequency by itself in the earlier case (SPM) and by an adjoining signal frequency (adjacent wavelength) as in the latter case (XPM).

XPM particularly is severe in long-haul WDM networks, and the acceptable norm in system design to counter this effect is to take into account a power penalty that can be assumed equal to the negative effect posed by XPM. Typically, 0.1–0.5 dB power penalty suffices the design constraints for XPM. Another nonlinearity is four-wave mixing (FWM), which creates several problems in system design and results in implementation of different wavelength assignment strategies to avoid the harmful effects of FWM. In one such assignment, to avoid the harmful effects of FWM, the reactive frequencies are chosen in such a way that the resultant spectral envelopes would not affect each other in a serious way. In other words, different bit rate signals are used to interleave and create a non-FWM system.

We can directly avoid FWM by using uneven channel spacing. This is a good technique that is simple to adopt, but ITU standards have laid down rules for fixed wavelength spacing (for example, 100 GHz in the C band). Long-haul networks typically are mesh or point-to-point topologies, with the latter implementation common in today's networks. Mesh networks are getting some prominence today, especially for long-haul cases. Comparing the two topologies, we can say in a nutshell that rings can do all that mesh can at a lower price as long as the load is bearable by the ring. (The capacity of a mesh far exceeds that of a ring.)

Long-haul nodes are generally composed of multiplexer/demultiplexer equipment at peripheral nodes, separated by cascaded regenerator sites. The regenerator sites might be either optical amplifiers or full 3R regenerators (O-E-O). Furthermore, these sites could be used to add and drop a few lightpaths, which would add flexibility to the network.

Forward Error Correction

Legacy communication networks introduced a concept of coding messages in a lossy communication channel to prevent data loss due to channel impairments. This principle has been extended to optical networks as well. Bits can be coded using a code word such that the new coded word is an extension of the bit in a markup form. The probability of an individual bit being corrupted in a transmission is higher than the probability of an entire code word being corrupted. This is the principle that is applied in FEC.

The tradeoff is that the new signal requires higher bandwidth for transmission. The minimum difference (in words or in vectors) to distinguish between two code words for two adjacent unlike symbols is called the *hamming distance*. The hamming distance is a measure of the accuracy or the foolproofness of the forward error correction (FEC). FEC causes a direct improvement in system gain. For inband FEC, the gain is around 2 dB. The gain experienced for out-of-band FEC can be expected around 6 dB. This does show phenomenal improvement in system performance, but the FEC equipment comes at a price due to high-speed electronic circuitry involved. In today's networks, FEC is implemented using Reed-Solomon codes (RS-codes). The bit rate enhancement due to FEC is shown in Equation 5-5.

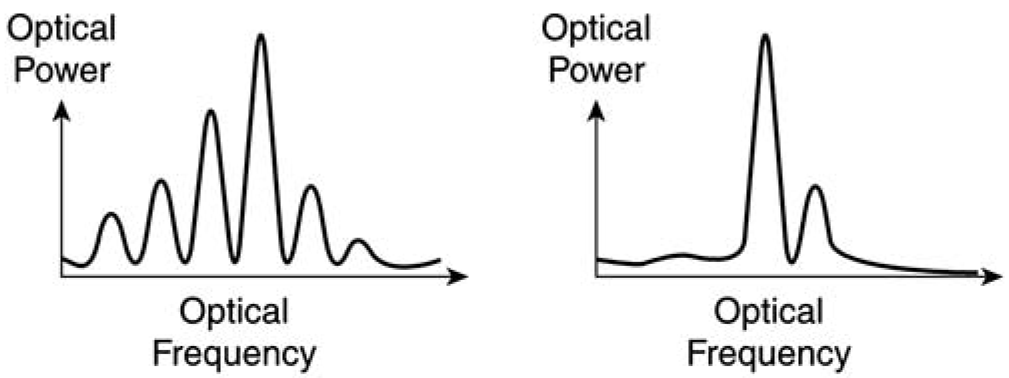
Equation 5-5

$$\text{Bit rate}_{\text{with FEC}} = \text{Bit rate}_{\text{original}} \times 1.07 \text{ for out of band (OOB) FEC}$$

WDM System Design: Components and Subsystem Consideration

The transmitters for WDM systems are made of semiconductor lasers. The lasers are typically single mode and the power that is associated with the output light is governed by carrier-photon dynamics. We want the laser to give power only at the wavelength of operation, which might not happen accurately. Most lasers give powers at various wavelengths due to mode partition noise, whereby multiple modes exist (apart from the dominant mode) and generate a power profile, as shown in Figure 5-17.

Figure 5-17. Laser Profile: Optical Power Versus Frequency Deviation



Typically, we want the dominant mode to have a power level at least 30 dB higher than the side modes. This way, we can minimize the cross-talk in adjacent channels by assuming all output powers are 30 dB greater than the side mode powers. We can prevent the side mode from corrupting the adjacent channels by using a wavelength-selective device such as AWG at the output of the transmitter. This might be an expensive technique, however. Typical output powers of lasers for long-haul operation are 0 to +7 dB. Note here that 30 db ratio corresponds to 1000 times magnification, in other words, the dominant mode power level is 1000 times higher than side mode power level.

Modulating a Laser: Direct and External Modulation and Spectral Efficiency

This book discussed direct and external modulation in Chapter 2. From the system design point of view, direct modulation represents a simple way to send data onto the required wavelength. Direct modulation occurs due to the "direct" modulation of the laser by making the electronic data proportional to the biasing current. Direct modulation has the side effect of producing heavy chirp because of the strong relationship between carrier density and refractive index inside the laser cavity. The frequency is in quadrature to the power during modulation; therefore, chirp results.

The chirped waveform occupies a wider spectral bandwidth than the expected amplitude modulated signal. For direct modulated lasers, an important consideration is between chirp and the extinction ratio. Typically, the extinction ratio (the ratio of a 1-bit power level to a 0-bit power level) is proportional to the chirp. Increasing the extinction ratio also increases the chirp associated with the signal, which is a harmful effect of direct modulation.

In contrast, externally modulated lasers are more stable and are used frequently in today's WDM networks. A DC-biased laser (producing a continuous wave, or CW) feeds the optical wavelength to an external modulator, which modulates the CW signal into desired optical bit streams. Typically, two kinds of modulators are available: electro-absorption modulators (EAMs) and Mach-Zehneder Interferometer (MZI) modulators. Mach-Zehneder interferometers are more common. Demonstrations up to 40 Gbps have been made, but 10 Gbps technology is more common. EAM lasers have the advantage of size over the MZI counterpart because they are much smaller than MZIs.

The single biggest advantage of having external modulators is the reduced chirp, which means that the signal occupies less bandwidth. Typically, MZI modulators are known to have twice the bit rate as their bandwidth. This also means that the spacing between adjacent channels in a WDM system can be greatly reduced.

Today's WDM networks have different modulation formats, as described in Chapter 2. For most practical cases, nonreturn to zero (NRZ) is an efficient format that requires 2.5 Hz of bandwidth for every transmitted bit. Return to zero (RZ) and carrier-suppressed return to zero (CS-RZ) are two additional formats used in WDM transmission links. Different formats are tried because some formats have more efficiency over the others and can have better response to fiber impairments—especially nonlinearity (XPM) and dispersion (CD).

Optical Receiver Design

The receiver (photodetector) happens to be the most important component in the WDM transmission link because system performance is measured as a function of the BER at the receiver. The two main considerations are the noise of the receiver and the receiver dynamic range (for macro-level design). Receiver noise as explained in Chapter 2 is mainly made of shot, white and thermal noise. One way of limiting noise to a certain extent is by placing a preamplifier and a narrow band filter.

The other design consideration is the dynamic range. This gives the minimum and maximum power that a receiver can detect (or work under). Receivers also have clock recovery circuits that can be optimized for various bit rates. Generally, as long as the receiver can detect a particular bit rate and the power falls within the dynamic range, the optical system design will adhere to the given BER assuming that the OSNR budget has been met.

Choosing Fiber and Design Based on Different Fiber Types

Fiber loss and fiber dispersion are the two most important impairments for choosing fiber types. *Fiber loss* is a material characteristic of the fiber, whereas *dispersion* is a characteristic of the fiber at a given bit rate. Typically, the dispersion is proportional to the square of the bit rate. The maximum allowable dispersion D should be less than $104{,}000/B^2L$. Using dispersion-shifted fiber (nonzero) enhances the system performance and increases the transmission length, although it results in higher loss. For example, at 10 Gbps, the maximum transmission length is 65 km using SMF fiber, but it is 500 km using NZ-DSF fiber for dispersion-limited systems only.

The intensity-dependent refractive index creates nonlinear effects in the fiber, such as SPM, XPM, FWM, stimulated Raman scattering (SRS), and stimulated Brillouin scattering (SBS), which have been discussed before. We must leave sufficient margins for these impairments and ensure that no single effect destroys the transmission signal.

SPM is primarily due to the self modulation of the pulses. It is caused generally in single-wavelength systems. At high bit rates, SPM tends to cancel dispersion.

XPM is modulation of pulse power by adjoining pulses (at different wavelengths). We can calculate XPM and dispersion in a fiber by using the split Fourier transform method, whereby the dispersion effects and the XPM effects are taken individually. We also must divide the fiber into minute strips for individual computation.

FWM is directly dependent on the wavelength spacing in WDM systems. It is inversely proportional to the dispersion of the fiber. A 1 dB penalty for XPM and FWM is generally advisable.

SRS and SBS are comparatively unimportant effects on long-haul system design, especially at compact WDM spacing.

PMD is the big source of impairment for ultra long-haul systems due to group delay of the pulse being a function of the state of polarization (SOP) of the signal.

Optical Amplifiers

The performance characteristics of an amplifier to be considered for design are optical gain, gain profile, bandwidth, noise figure (NF), and gain tilt. Optical gain in dB gives the average amount of amplification that can be expected. The gain profile depicts the flatness of the gain over the entire operating spectra.

The width of the operating spectra—that is, the frequency difference between the lowest and highest amplified channels above cutoff—gives the bandwidth of the amplifier. NF gives the noise (in dB) that the amplifier adds to the signal due to the amplified spontaneous emission (ASE). This is one of the most important design constraints in WDM links. The gain tilt can be quantitatively summarized as the difference in highest gain and lowest gain experienced by different channels in the WDM spectra. Different amplifier combinations are used. EDFAs are the most common and provide 22-25 dB gain, but Raman amplifiers based on SRS are becoming quite prominent because of the large operating bandwidth and low noise figure.

SOA technology is gradually maturing, but its use might be restricted for access networks only due to the high cross-talk and low gain profile they offer. The most important design consideration in amplifiers still remains the NF resulting from the addition of ASE. ASE rejection filters are slowly appearing in the industry, but performance is debatable.

OSNR degradation is the most important design constraint in the system because of the noise being added by cascaded amplifier stages. The tradeoff in optical system (wavelength division multiplexing, or WDM) design is between amplifying the power of individual channels and reducing the nonlinear effects. In other words, as power increases, so do nonlinearities; therefore, a tradeoff is needed. To maintain this equilibrium, we need to deploy low noise amplifiers (LNAs), which can create higher OSNR systems at lower power (and therefore, nonlinearities) using the low noise features. Raman amplification is a step in this direction of amplifiers with low NF.

Optical Add/Drop MUX and Cross-Connects

OADMs were described in the previous sections. For long-haul networks, some of the design constraints that we need to look into are low loss, low cross-talk, low configuration (add/drop) times, low wavelength-dependent loss (WDL), low polarization-dependent loss (PDL), and good filter characteristics of the multiplexer/demultiplexer. Optical cross-connects (OXCs) need to have almost the same features as OADM with the exception of wavelength conversion. A low loss per port as well as a low cost are always desirable. The most important considerations of OXCs or OADMs are low insertion loss (low loss for adding a signal) and low passthrough loss.

Questions

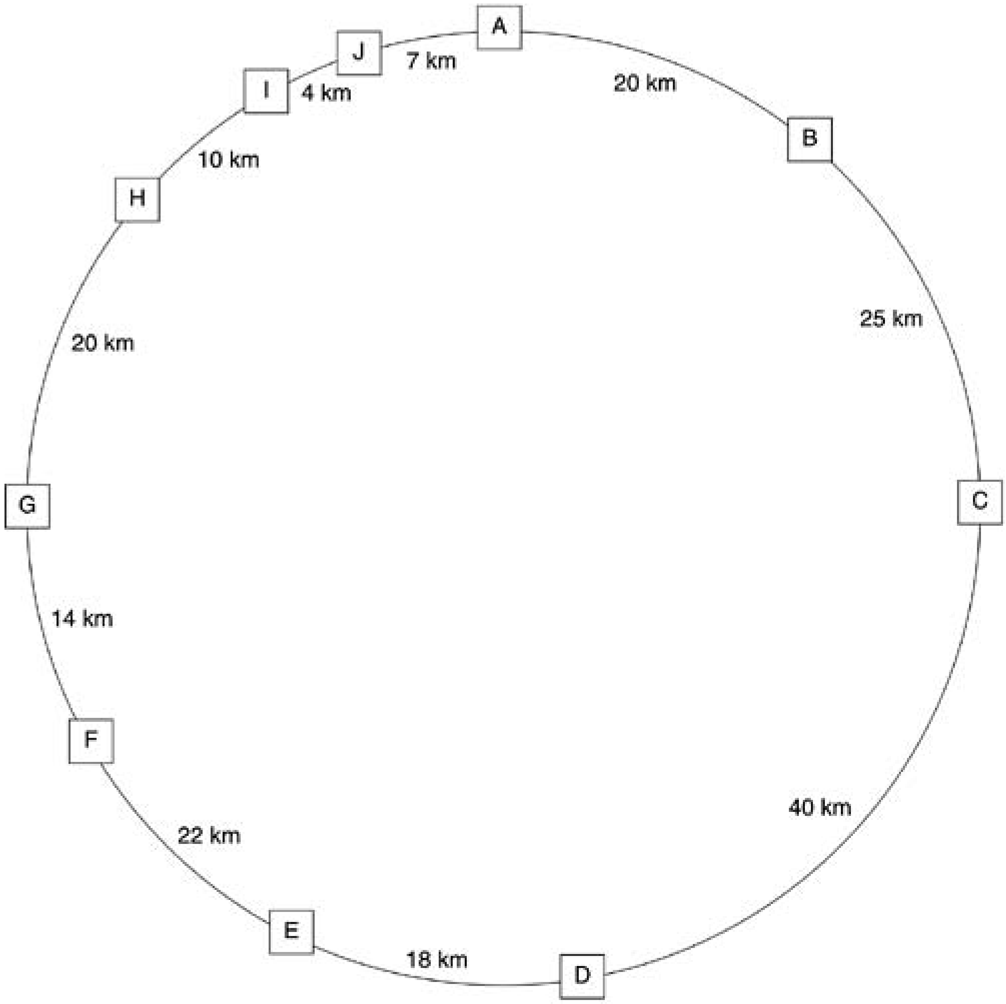
The following questions are left as an exercise for we. Use the principles in Chapters 4 and 5 to solve the problems.

Question 1

Consider a metro core ring (see Figure 5-18) with the following features:

Number of nodes = 10

Figure 5-18. A 10 Node Metro Core Ring



The network has OADM with the following characteristics:

Passthrough loss = 14 dB

Postline amp gain = G1 dB

Add/drop loss = 7 dB

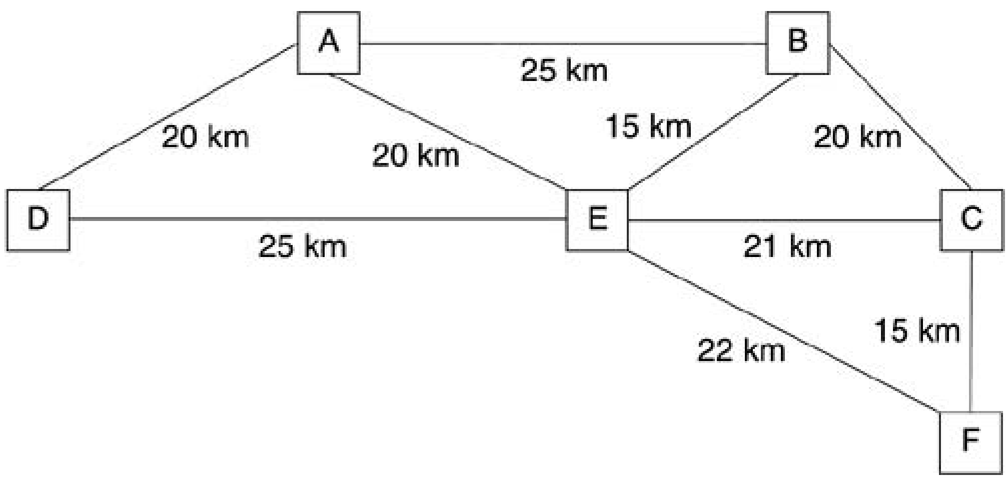
Number of channels = 32, spacing = 100 GHz

For an OSNR of 20 dB, what will be the minimum postline amplifier gain for NF = 4.5 dB? If we assume 1600 ps as the maximum allowable accumulated dispersion, then how many compensators will be needed for SMF fiber? (Dispersion = 16 ps/nm-km) at 10 Gbps line rate.

Question 2

Design a full OADM metro core "mesh" as shown in Figure 5-19.

Figure 5-19. A Sample OADM Metro Core "Mesh"



The OADM node consists of a multiwavelength multifiber cross-connect, as shown in Figure 5-19. Table 5-2 shows its characteristics.

Table 5-2. OADM Node Characteristics

Degree of Connectivity	Passthrough Loss	Add/Drop Loss
2	8	4
3	8.5	4
4	9	4
5	10	4

A postline amplifier has a gain of 22 dB and an NF of 4.5 dB. Assume individual links to be protected. Calculate the OSNR based on worst lightpath case. To use SMF fiber, what is the number of dispersion-compensating equipment to use if the maximum allowable dispersion is 1600 ps? If the average hop distance of the lightpath is 1.8 hops and loading is 0.4, would a ring (A-B-C-F-E-D) suffice the load?

Summary

This chapter discussed philosophies of optical network designs from a topological point of view. Optical networks can be classified into three main areas: access, metro, and long haul. Each network type has different components and design issues associated with it. Access networks have relatively few optical impairments and can be implemented using low-cost technologies. Access networks are more flexible and provide a direct point of attachment to end users. Metro networks require comparatively more stringent optical requirements than access networks. Metro networks are physically larger than access networks and need more specific technologies for implementation. Because of their sheer size, long-haul networks have the brunt of optical impairments. Long-haul network design involves many different philosophies and components. In this chapter, we learned how to design optical WDM networks and choose the right equipment for the right application.

References

[1]Agrawal, Govind P. Wiley Interscience, *Fiber-Optic Communication Systems*.

[2]Keiser, G. *Optical Fiber Communications*. McGraw-Hill; 1996.

[3]Ramaswami, R. and Sivarajan, K. *Optical Networks a Practical Perspective*. San Francisco, CA: Morgan Kauffman; 1999.

[4]Taub, M. and Shilling, S. *Principles of Communication Systems*. McGraw-Hill; 1992.

[5]Bononi, A. *Optical Networking*. Springer Publishers; 1999.

[6]Agrawal, Govind P. Wiley Interscience, *Fiber-Optic Communication Systems*, Second Edition. 1997.

[7]Mukherjee, B. *Optical Communications Networks*. McGraw-Hill; 1997.

[8]Ramamurthiu, B. *Design of Optical WDM Networks, LAN, MAN and WAN Architectures*. Kluwer; 2001.

[9]Narula, A., Lin, P. and Modiano, E. *Efficient Routing and Wavelength Assignment for Reconfigurable WDM Networks*.Journal on Selected Areas in Communication; Jan 2002, pp-75-88.

Chapter 6. Network Level Strategies in WDM Network Design: Routing and Wavelength Assignment

In the preceding chapters, we looked at optical networks from the physical layer perspective. This chapter takes a closer look at optical networks, but from a routing and wavelength assignment perspective. Since the seminal paper on WDM networks by Charles Brackett in 1990, optical networks from the network layer perspective have been actively researched.

The next major breakthrough in network layer technology for fiber-optic communication was discussed in the classic July 1992 paper in *IEEE Transactions on Communication* by Chlamtac et al on Lightpath communication. Lightpath communications seemed to be a convenient way to map data (TDM or plane data frames) onto optical wavelengths. The paper briefly discussed some of the issues that highlight a network philosophy—namely routing in an optical environment and wavelength assignment under the constraints that a limited number of wavelengths were available to choose from.

The most important result of the paper was the mention of virtual topology design in optical networks. The *virtual topology* can be roughly defined as the mapping of lightpaths between source destination pairs for a given traffic matrix. This paper dealt with the optimum ways to map traffic to a given set of wavelengths and construct logical or virtual topology on the embedded physical topology. Chapter 5, "WDM Network Design -2," showed how a logical topology is different from a physical topology.

A third innovation in optical network design was in assigning wavelengths with a constraint of wavelength reuse. In this regard, Acampora and Zhang produced some quality results and analytical explanations to give strong impetus to today's network capacity planning based on wavelength assignment as well as wavelength reuse (optimizing the network for a minimum number of wavelengths by strategic assignment philosophy).

Finally, Ramaswami and Sivarajan formulated the bounds on routing and wavelength assignment in today's optical networks and published a set of results that quantified wavelength assignment strategies based on an integer linear program formulation.

This chapter studies each of these concepts and develops some basic rules that help in optical network design. We must be aware of mathematical intricacies used in this chapter, although care has been taken to minimize the degree of mathematics used. Although routing, wave-length assignment, and virtual topology design might appear to be a promising way to increase the capacity or optimize the performance of a network, note that the physical layer impairments play a strong veto role in the final limitations of the network performance.

When we consider optical networks from the network point of view, we have to consider certain attributes that reflect the quality of the network or that give us the network performance heuristics. From basic packet-switching theory, the most important heuristic that determines the network performance is the throughput of the system. *Throughput* can be defined as the net flow in a channel as compared to the maximum flow possible in that channel. Another way of defining throughput is considering throughput to be the quantity of service that the network provides. Throughput is generally integrated and averaged over the entire network; therefore, it varies from time to time depending on the network traffic load and other associated constraints.

The inverse of throughput is the blocking probability of the network. Blocking probability can be considered as the net packet loss probability in packet-switched networks. It can also be considered as the probability that a lightpath connection gets blocked for a given source destination pair in circuit-switched networks.

The next most important factor to consider in networks is the end-to-end associated latency. *Latency* is basically the end-to-end delay in the network. The delay can be due to processing times of the different nodes involved in the buffering at the nodes, or it can be due to the transmission delay through the fiber. Delay in a network can be considered to be the quality of service that the network provides.

Two kinds of networks are proposed here: the single-hop model and the multihop model. In single-hop networks, the packet or lightpath is set up between source and destination directly without opto-electronic conversion and regeneration at the intermediate nodes. In other words the packet just zooms through from one node to another on a directly connected lightpath. This generally happens if a direct physical link exists between the two nodes. On the other hand, in multihop networks, packets hop from one node to another en route to the destination; therefore, electronic processing does occur at each node.

Multihopping is the preferred communication method in today's Internet because we cannot have fibers connecting every node to every other node in the network. (Full mesh is not possible). However, at high bit rates and numerous channels, multihopping might not always be the preferred approach; processing high volumes of packets at every intermediate node is definitely a difficult task! Therefore, lightpaths are established between source destination pairs through many intermediate nodes, such that the intermediate nodes are oblivious to the data that flows through them (pass-through function).

Multihop network topologies can also be classified as regular and irregular topologies. Regular topologies are those that have physical fibers connected to nodes following the interconnection matrix of a regular graph. Some examples of such topologies are Shufflenet, GEMNET, and TIONet. As of now, regular multihop networks are relegated more to academic projects and pilot access networks only. Irregular multihop networks make up approximately 95 percent of today's Internet.

Routing and Wavelength Assignment: The Basic Problem

Consider an arbitrary network with N nodes and E links. Now assume that the E fiber links are so dispersed; they create a network topology graph that connects the N nodes in some arbitrary manner. In addition, assume that each link has 2 w wavelengths in it, with w wavelengths in each direction. Then the capacity of the network in number of lightpaths is E * 2w / h, where h is the average hop distance of the lightpaths. The objective of a network capacity planner is to optimize the network for some maximum traffic matrix Tr_{max}. In other words, we have to fit onto the physical topology an arbitrary maximum number of lightpaths between the given source destination pairs. In other words, we have to match the traffic matrix to offer the maximum throughput (create a maximum number of lightpaths). While establishing the maximum possible lightpaths, we are subject to the following constraints:

- A lightpath should have wavelength continuity. The same wavelength should be exhibited by the entire lightpath. In other words, even if the lightpath traverses multiple links, it must be on the same wavelength on each of the links. Although each node can have some elements of wavelength conversion, this condition is generally sparse and cannot be assumed.
- The routing should be done so that routes are on the shortest path(s). This ensures that the capacity of the network is maximized. Nonshortest path routes waste bandwidth that other lightpaths would normally use.
- Wavelengths should be assigned to reduce or minimize blocking of additional lightpaths. Wavelength assignment algorithms can be made to optimize a particular quality in a network. In one embodiment, wavelength reuse and capacity were maximized. This was done by allocating wavelengths more probabilistically (heuristically) to shorter hop lightpaths, which maximizes their reuse. This kind of wavelength assignment strategy can be termed *first fit*.
- Wavelength reuse should be maximized. As mentioned in the previous point, wavelength reuse is an important feature of optical networks. We can accomplish wavelength reuse by reusing the same wavelength (lambda) in different (disjointed) segments (links) in the network. Wavelength reuse results from spatial diversity. *Spatial diversity* here means having two lightpaths on the same wavelength, but the two lightpaths are disjoint. In other words, they do not share a common link; therefore, the lightpaths can use the same wavelength. The two lightpaths are spatially diverse.

Routing and wavelength assignment can be considered two separate problems—one in which the physical routes need to be computed, thereby routing lightpaths on these computed routes, and the second whereby the wavelengths need to be assigned to these routes for lightpaths in the most efficient manner. As we can see, although routing and wavelength assignment are two different problems, they solve the same issue: optimizing the network to create maximum benefit. In some embodiments, the two problems are treated as one, and solutions are obtained to solve them in tandem.

Solutions to these problems often lead to algorithms with high levels of computational complexity (multiple iterations). The solution is said to be nondeterministic polynomial time hard (NP-hard). NP algorithms can be characterized by the fact that their solution requires exponentially high times; therefore, their level of computational complexity is extremely high.

We can view the routing and wavelength assignment (RWA) problem as analogous to optical system design, whereby the network resources (in this case, wavelengths and throughput while in the system design case OSNR and BER) are optimized or maximized.

One of the more heuristic solutions in trying to solve the RWA problem is to break the two problems into a set of subproblems, whereby each subproblem is solved to obtain some optimization values. In one approach by Zhang and Acampora, the RWA subproblems were as follows:

- Finding a wavelength assignment algorithm that tended to favor the one optical hop traffic and obtaining a connection (logical) graph for it.
- Developing routing schemes for the obtained logical graph that produces the lowest call-blocking probability.

The constraint here is that no two lightpaths on the same link can have the same color (wavelength). In addition, the second constraint is that of wavelength continuity, whereby the entire lightpath spread across different fiber links must have the same wavelength. The second constraint can be relaxed if we assume wavelength conversion at intermediate nodes.

Formulating the Wavelength Assignment Problem

Consider a topology that we must optimize for a given load. We have to assign wavelengths to set up lightpaths in the most appropriate way. The wavelength assignment algorithm as proposed by Zhang and Acampora essentially maximizes all one-hop traffic and bases this maximization to the condition that no two lightpaths on the same fiber link have the same wavelength. This results in a mathematical formulation that can be defined as a lightpath-link matrix shown in Equation 6-1.

Equation 6-1

$$m = m_{(ij),(lm)}$$

If lightpath(i,j) and lightpath(l,m) have a common link, we have Equation 6-2, where m = 1.

Equation 6-2

$$m = m_{(ij),(lm)} = 1$$

If no common link exists between lightpath(i,j) and lightpath(l,m), we get Equation 6-3.

Equation 6-3

$$m = m_{(ij),(lm)} = 0$$

Note here that lightpath(i,j) signifies that the source node is node i and the destination or sink node is node j. We can generalize this convention for the remainder of the chapter.

Therefore, 'm' would be 1 if a common link exists between the lightpaths i-j and l-m. If no such common link exists, then the value of m is 0. This formulation helps to define and build up the constraint that no two lightpaths in the same fiber link can have the same wavelength.

We also define an array (z), one for each of the N^2 possible lightpath connections in the network, where N is the number of nodes shown in Equations 6-4 and 6-5.

If nodes i and j are connected through wavelength k, we get Equation 6-4 for z.

Equation 6-4

$$z_{ij}(k) = 1$$

If wavelength k does not connect nodes i and j (no lightpath is present between nodes i and j on wavelength k), we get Equation 6-5 for z.

Equation 6-5

$$z_{ij}(k) = 0$$

This means that $z_{ij}(k)$ assumes the value of 1 for all those wavelengths on which there are lightpaths whose source and destination nodes are i and j, respectively.

If C is the line rate (in bits per second), the total capacity between i and j is shown in Equation 6-6.

Equation 6-6

$$\sum_{k=1}^{W} z_{ij}(k).C$$

In the equation, W is the total number of wavelengths in a given fiber link.

We can understand this by using the simple calculation that states that if we add the number of times $z_{ij}(k)$ is 1 for all the wavelengths, we precisely get the number of lightpaths that is set up exclusively between nodes i and j. The number of lightpaths multiplied by the capacity of each lightpath gives us the total capacity of the traffic between nodes i and j.

The wavelength assignment problem can be termed as in Equation 6-7.

Equation 6-7

$$max\sum_{ij} min\left(t_{ij}, \sum_{k=1}^{W} z_{ij}(k).C\right)$$

Equation 6-7 by Zhang and Acampora, although seemingly complicated, is in fact quite straightforward. The term inside the round brackets that we have to minimize is a choice between the total network traffic denoted by t_{ij} (between nodes i and j) and the capacity between nodes i and j on the established lightpaths. This is then maximized over all the source destination pairs throughout the network. It is easy to understand that the summation term provides the capacity of the setup lightpaths between nodes i and j. We initially minimize this with respect to the demand, which is t_{ij}. In other words, we are optimally assigning wavelengths subject to the constraints shown in Equations 6-8 and 6-9.

Equation 6-8

$$\sum_{m=1}^{N} z_{im}(k) \le 1$$

Equation 6-9

$$\sum_{m=1}^{N} z_{mj}(k) \leq 1$$

In the equations, k = 1, 2, ...W, and i and j are elements of the set of nodes given by {1,2,... , N}.

This constraint basically states that for each of the W wavelengths present, there can be at the most one connection (lightpath) emanating (starting) from node i and one lightpath sinking or ending at node j. This ensures that no wavelength is used twice on each link.

The final wavelength assignment condition is shown in Equation 6-10.

Equation 6-10

$$m_{(i,j),(l,m)}(z_{ij}(k) + z_{lm}(k)) \leq 1$$

for all distinct pairs of lightpaths, i,j and l,m:

The third condition states that if two lightpaths—i-j and l-m—cross each other (share) one or more links in the network, then at no time can they have the same wavelength on the same link (wavelength diversity constraint). As per the wavelength diversity constraint, we ensure that no two lightpaths in the same fiber have the same wavelength.

Routing and Wavelength Assignment and Integer Linear Programming Formulations

To maximize the performance (throughput) of an optical WDM network, it is imperative to route the set of demands in the most appropriate way (on the best possible routes and using the proper wavelength assignment). For a given network graph G(V,E), where V represents vertices (nodes) and E represents the links joining these vertices (fiber links), we have a traffic matrix T (given load) that has to be satisfied in the best possible manner. In other words, we have to set up lightpaths for each of the nonzero elements in T, the traffic matrix.

Mathematically, this means that we have to set up lightpaths for every $t_{ij} > 0$ where i,j are elements of the node array (V). While setting up these lightpaths and conforming to the load in the network, we have to consider the constraints that govern the routing and wavelength assignment philosophy. In the absence of wavelength converters at each node, we have to ensure that every lightpath is subject to the wavelength continuity constraint. This constraint ensures that every lightpath has the same wavelength in each of its fiber hops (links).

We also have to ensure that no two lightpaths are on the same wavelength in a given fiber link. Under all these considerations, we have to route the lightpaths to yield the maximum throughput or performance. We might be able to accomplish this in various ways. In one embodiment, we can use shortest path routing. In another embodiment, we can assign wavelengths first to lightpaths on short paths and assign wavelengths later to longer lightpaths (ascending order of number of hops method). Wavelength reuse is a parameter that determines the efficiency of the network. In one RWA algorithm, wavelength reuse is the maximizing quantity.

As we can see from this discussion, routing and wavelength assignment are problems with multiple inputs, multiple working constraints, and various solutions of which only a few are optimal. To solve such an RWA problem, we need to make use of intricate solutions in mathematics. One such solution is to formulate an integer linear program on which literature reference[3] throws light.

In the quest for a solution to the RWA problem, we can expect many possible implementations. The best possible implementation is the one for which the value of the formulated integer problem is the highest. It is beyond the scope of this book to solve or even formulate such a problem. Simulation softwares and Excel graphs are good methods for providing solution to such problems.

Considering a network RWA problem under the constraints of wavelength continuity, wavelength diversity, and shortest path routing, the linear problem yields a near optimum solution for a given traffic load. When the traffic load is subject to rapid changes and the RWA problem needs to cater to these changes on a time-sensitive basis, the RWA problem is known as *onlineRWA*. In contrast, if the required solution needs to be computed for fixed (nonchanging) loads and constraints, the RWA problem is called *offlineRWA*.

The Graph Coloring Approach to the Wavelength Assignment Problem

For a given network and set of traffic demands, it is important to assign wavelengths to lightpaths manually. Some of today's networks are characterized by a small graph (in size) and low loads. An intuitive yet simple way to assign wavelengths to lightpaths is the graph coloring approach. Graph coloring is a legendary mathematical problem whereby nodes of a graph are colored such that no two adjacent (linked) nodes have the same color.

Extending this approach to optical networks, consider the network to be given by a graph G(V,E). In other words, V represents vertices (nodes) and E represents edges (links). Further if we are given a traffic matrix T_{ij}, where i,j are any elements of V, we can route these T demands on the best possible routes (which might be the shortest path). Now we can name each lightpath a unique name such that

$$w = \{w_1, w_2 \ldots\ldots\ldots w_{(i,\,j)max}\}$$

is the set of lightpaths that caters to load T.

We can draw a graph G(W,P) whose nodes are the lightpaths that we named in set {w}. Furthermore, we can draw an undirected edge between any two or more nodes that share a link. Repeating the procedure to draw edges over all the lightpaths that share common links, we get the graph G(W,P). Now color the graph G(W,P) such that no two linked (adjacent) nodes have the same color. The physical meaning is that if two nodes are adjacent, they have at least one common link (lightpath) between them.

Upon completion of the coloring procedure, we get a graph whose nodes are the lightpath connections. The color of these nodes corresponds to the wavelengths that are used. The more colors that exist, the more the wavelengths need to route (T_{ij}) connections over G(V,E).

A more heuristic approach to graph coloring is the sequential approach. Coloring all demands (nodes) simultaneously is a hard procedure, and it is said to be NP complete (meaning that it will take an exponentially long computational time). The sequential approach is quite modular, such that graph coloring is done as the lightpaths arrive. In other words, as new nodes in the graph (lightpaths) arrive, they are colored by taking just the previous state of the graph into account rather than taking the entire graph and reassigning colors in the most appropriate way (using the least number of required colors).

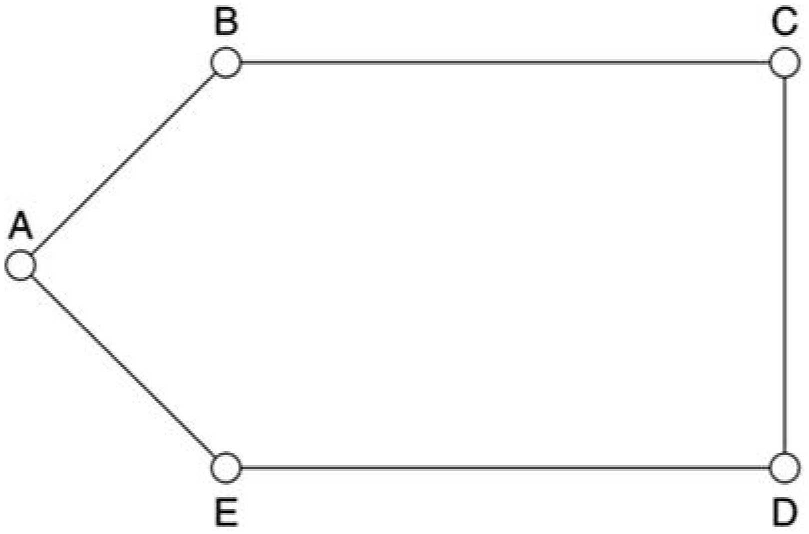
If a graph of the lightpaths (as nodes) has V vertices (V lightpaths) and the degree of these vertices is given by the set D = $\{D_1, D_2, \ldots D_V\}$, then the maximum number of colors needed (wavelengths needed) is shown in Equation 6-11.

Equation 6-11

$$c_{max} = max\{D_i\} + 1$$

Numerical example: Consider the physical graph (network) shown in Figure 6-1.

Figure 6-1. The Basic Topology (physical topology)



If the traffic to be routed is the following (the traffic load):

- AED
- ABC
- BCD
- BAE
- EDC
- DCB
- DEA

then each of the seven demands represents a lightpath from source to destination through an intermediate node. Note, however, that we must establish the lightpath under the constraint of wavelength continuity (explained earlier) and also under the constraint that no two lightpaths in a physical link can have the same wavelength.

Route the requests and color the graph to give the minimum number of wavelengths.

The lightpath graph G or "w_l" light paths is given as

Shown in Table 6-1 is the wavelength assignment scheme based on the results in Figure 6-2, which shows the lightpath graph.

Figure 6-2. The Routed Traffic on the Topology

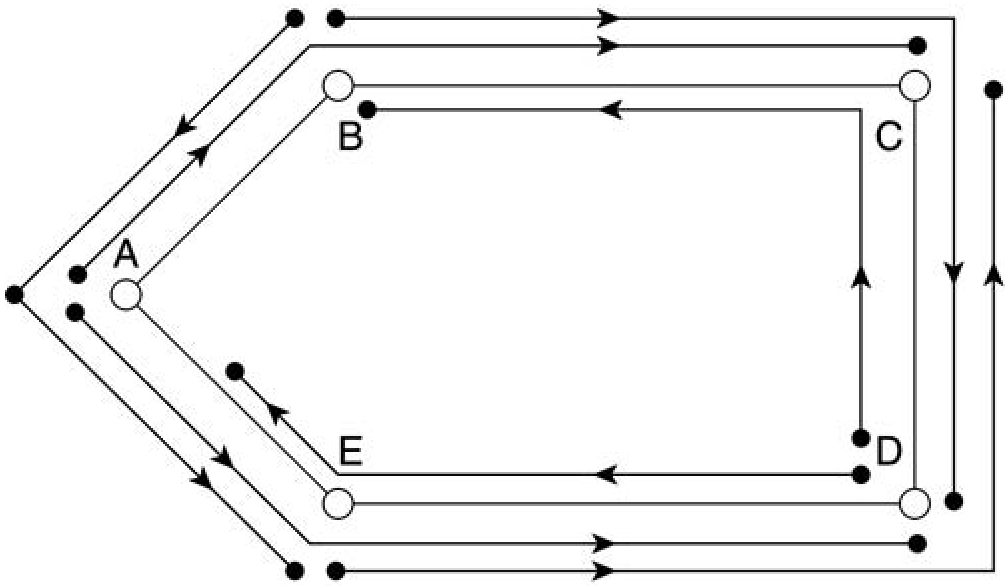


Table 6-1. Obtained Results of the Graph Coloring Approach to Wavelength Assignment

NODE	Color/l	Node	Color/l
AED	λ_1	EDC	λ_4
ABC	λ_1	DCB	λ_3
BCD	λ_2	DEA	λ_2
BAE	λ_3		

From the table, we can see that the number of wavelengths required is 4.

We can verify this solution by considering the node with the highest degree and using the theoretical equation shown in Equation 6-11. Refer to Equation 6-12 for verification.

Equation 6-12

$$Max(D_i) = 3$$

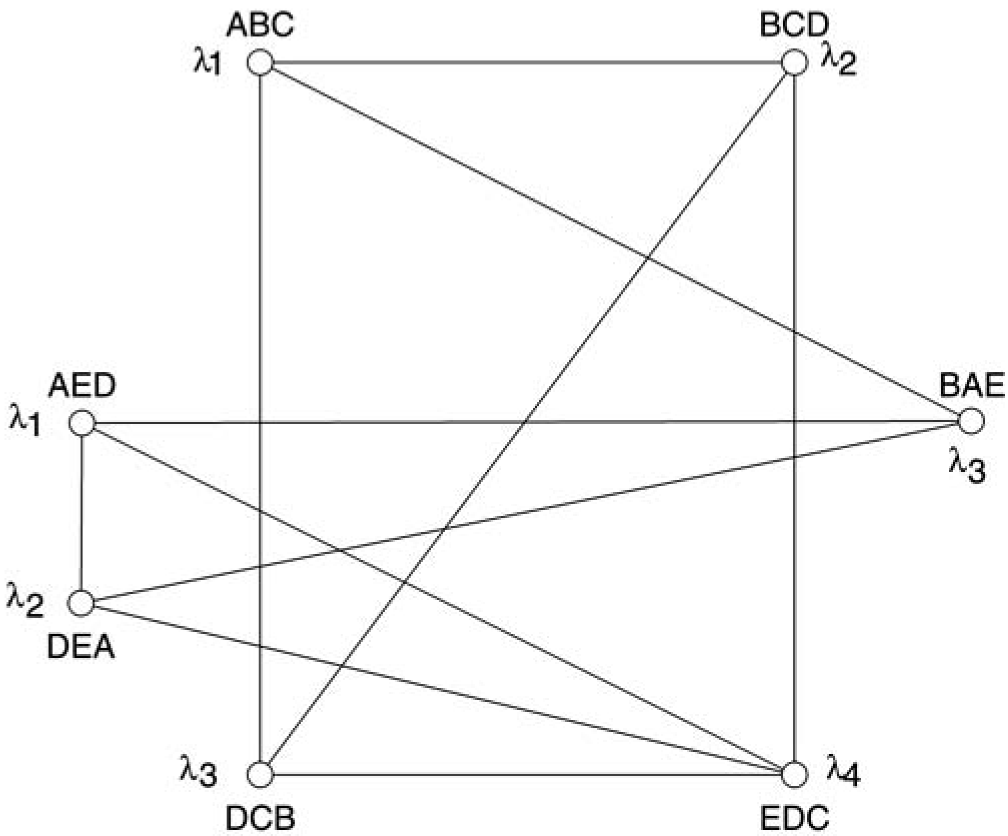
Therefore, the number of Color(s)/λ needed is this.

$$= max(D_i) + 1$$

$$= 3 + 1 = \mathbf{4\ Color(s)/\lambda}$$

From the graph in Figure 6-3, we see that the maximum degree of a node is 3; therefore, the number of wavelengths required for proper wavelength assignment is 3 + 1, or 4.

Figure 6-3. The Lightpath Graph: The nodes in the graph (filled circles) represent established lightpaths. If a link exists between two nodes (lightpaths), the two lightpaths share a fiber.



If we consider 40 color(s) / λs and 100 lightpath connections over 10–16 nodes, we can imagine how hard it would be to draw such a schematic. Therefore, linear programming is a viable solution for large graphs with many wavelengths.

Static and Dynamic Lightpath Establishment

Two kinds of lightpath establishment techniques deserve attention: static lightpath establishment (SLE) and dynamic lightpath establishment (DLE). Both are practical issues in today's network, and their solution has direct effects on network planning mainly in capital expenditure (CAPEX) and operational expenditure (OPEX) reduction.

The first issue, that of static lightpath establishment, is the RWA problem for a preknown set of lightpaths. In other words, the set of demands or traffic matrix is known "off hand." We have to route lightpaths in the most efficient way to minimize the number of wavelengths required.

One possible approach in the static lightpath establishment case is to break up the set of demands into various subsets in descending order of lightpath lengths. This means we can route the longest path first and then route the shorter lightpaths subsequently. The assumption here is that of wavelength continuity, whereby no wavelength conversion at intermediate nodes is assumed.

Another solution for providing static lightpath establishment is by allocating lightpaths with shortest hop distance first. Static lightpath establishment is a practical problem in today's networks; every solution has its own pros and cons.

The second issue is dynamic lightpath establishment. For a network that is subjected to a varying traffic load, we have to assign wavelengths and route the lightpaths so that we have the maximum throughput (number of established lightpaths) over a period of time. This results in optimizing the network for some "churn," which denotes the variation in traffic load.

The optimization parameter to be considered is based on the fact that at a given time, we should be able to route and assign the maximum number of lightpaths in the network; we should be capable of satisfying the maximum number of demands at epoch. In other words, we should be able to take in as many new demands as possible (given the state of the network at present), with old demands still existing and no detrimental damage done to them. This reduces the lightpath blocking probability. *Lightpath blocking probability* is the ratio of lightpaths rejected to lightpaths requested. (Refer to Equation 6-13, blocking probability of lightpaths.)

Equation 6-13

$$BP_{Lightpath(DLE)} = \frac{Number\ of\ lightpath\ rejected}{Number\ of\ lightpath\ requested} \text{ -------[1]}$$

The goal is to minimize the blocking probability, or maximize the likelihood of a lightpath demand being satisfied taking into account the RWA algorithm that is deployed. One way of doing so is by routing lightpaths according to a least congested path (LCP) routing algorithm. In such a scheme, if multiple paths are present between the source and destination pair, a request is routed on the least congested path. The advantage of such an approach is that the overall congestion in the network is under check. That signifies that we can route additional requests as long as the net congestion on a link is less than λ_{max}, the maximum number of wavelengths in the fiber.

Least-cost path algorithms ensure that the loading of the network is uniform as long as the traffic demands are uniform. Least-cost path fails for hubbed traffic. The lower bound on the number of wavelengths needed in a network to route a particular set of demands is often given by the number of wavelengths in the most congested link. The bound is tight for acyclic networks, which does not hold true for cyclic ring networks because of the passthrough traffic in rings.

Ramaswami and Sasaki approximated the ring performance (number of lightpaths established) to be $2\lambda_{max} - 1$ lightpaths[9].

Chlamtac's seminal work in lightpath establishment can be regarded as a fundamental step in lightpath establishment techniques. It is reproduced here for convenience.

Table 6-2 shows Chlamtac's seminal algorithm for assigning wavelengths for a given traffic demand in a WDM network.

Table 6-2. Algorithm for RWA as proposed in "Lightpath Communications: An Approach to High Bandwidth Optical WANs"[1]

lpcm[i,j]	The lightpath collision matrix. lpcm[i,j] = 1 if lightpaths i and j have a link in common. (lightpath i and j collide.)
lpnum[i]	Lightpath ID's ordered by descending lightpath length.
w	Wavelength number currently assigned.
set[i]	Sets of lightpaths ordered according to allocated wavelength.
s,e	Start and end pointers to current set.
lambda[i]	Wavelength definition array. Lambda [i] points to the first lightpath in a set using wavelength i.
lpalloc[i]	Flags indicating whether lightpaths i have already been allocated.
n	Number of lightpaths in set or(set,s,e, lpnum[i], lambda[i]) function; returns true if lightpath lpnum[i]has a link in common with the lightpaths in the set sets[s]……..set[e], based on the lightpath collision matrix l pcm.

The following function gives static lightpath establishment for an arbitrary network:

```
procedure static-establishment

begin

     lambada [1] = w = s =e = 1

     for i= 1 to n do lpalloc[i] = false

    while (e<n) do begin

      for i=1 to n do begin(*)

      if not lpalloc[i] then

      if not or (set, s, e lpnum[i],lpcm) then begin
```

```
        set[e] = lpnum[i]

        e=e+1

        lpalloc[i]=true

        end

end

w=w+1

lambda[w] = s =e

      end

end
```

Virtual Topology Design

Consider a given physical topology depicted by the graph G(V,E), where V stands for the set of vertices and E stands for the set of edges. Also consider a given traffic matrix T, which basically gives the traffic in lightpaths between every source-destination pair in the net-work. We have to create a virtual or logical topology in the most efficient way to map the set of demands in the matrix T most optimally. We need to satisfy some constraints before we create the virtual or logical topology. As an example, a particular constraint is that the number of lightpaths through any link should be less than or equal to the wavelengths in the fiber, and in the absence of wavelength conversion, a lightpath spread across multiple links should have the same wavelength on each link (wavelength continuity constant). The problem can be formulated as mentioned in *Optical Communication Networks*[3] and "Multiwavelength Optical Networks with Limited Wavelength Conversion."[9]

For an N node W wavelength network, consider the network topology to be depicted by graph G(V,E,W), where |V| = N, and E represents the edges or links with W wavelengths in each direction. If F_{mn} is the number of fibers connecting nodes m and n, then let the total number of fiber links in the network can be shown as in Equation 6-14.

Equation 6-14

$$\sum_{m}\sum_{n} F_{m,n} = L$$

Furthermore, θ_{mn} is the length in kilometers from node m to n and

$$\frac{\theta_{mn}}{c} = \Delta_{mn}$$

is the delay between m and n in time. T_{sd} is the traffic matrix, and the maximum capacity of each channel is B_{max}. Let

$$\lambda_{ij}^{sd}$$

denote the traffic between s and d on physical link i, j. Then the issue of virtual topology design is given in Equation 6-15 as a minimization problem.

Equation 6-15

$$minimize \frac{1}{\sum_{s,d} T_{sd}} \sum_{ij}\sum_{sd} \lambda_{ij}^{sd}$$

In other words, we have to minimize the average flow or hop distance of lightpaths over the set of given links in the most optimal manner. As the average hop distance is minimized, the net throughput of the network is maximized and the lightpath blocking probability is minimized. In the process, we will get a matrix V_{ij}, which is called the *virtual topology connection matrix*. The constraints we need to take into account are the wavelength continuity constraint as well as the net routing delay. Other future issues that we might consider are congestion or loading in network behavior and limited wavelength tunability constraints.

Some Corollaries on Routing and Wavelength Assignment

It is important to define the concept of load in optical networks. From legacy telephone (circuit switched) networks, network load was calculated as the product of mean arrival rate and mean call holding time. In optical/lightwave networks, the load can be defined as the product of lightpaths' arrival rate and average lightpath duration. Another factor mentioned previously is the wavelength reuse factor.

A good routing and wavelength assignment scheme will end up maximizing the wavelength reuse factor. In the paper "Efficient Routing and Wavelength Assignment for Reconfigurable WDM Networks,"[7] Lin and Modiano showed that for ring networks, the maximum number of lightpaths per wavelength can be 2.43. (The wavelength reuse was 2.43.) Similarly, in the paper "A Heuristic Wavelength Assignment Algorithm for Multihop WDM Networks with Wavelength Routing and Wavelength Reuse,"[2] by Zhang and Acampora, wavelength reuse was shown as a fraction of the number of nodes in arbitrary topologies.

As we can anticipate, wavelength reuse increases with the increase in nodes. The reason for such an increase can be attributed to an increase in raw capacity (due to more links, and due to more spatial diversity leading to move wavelength reuse). An interesting point of consideration is the wavelength reuse factor compared to the minimum number of wavelengths (or lower bound) for various loads. The minimum number of wavelengths required might be quite high in hubbed architectures, where the traffic tends to sink at one node, causing severe congestion and more wavelengths needed near the periphery of the node.

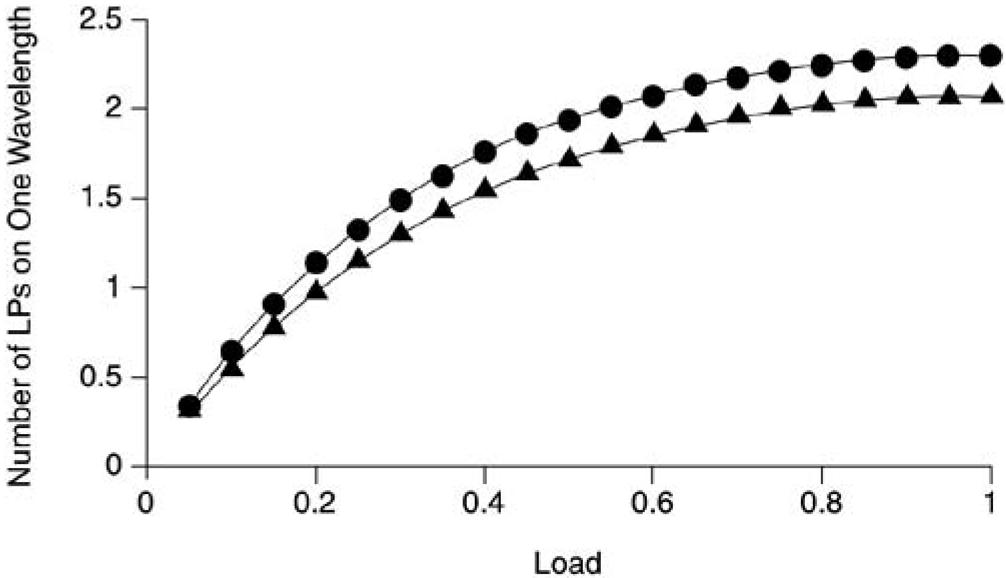
For ring networks with N nodes, a load of L (number of lightpaths) can be supported with a maximum of 2L wavelengths without wavelength conversion. In contrast, if we have limited wavelength conversion of degree 'd', then the maximum number of wavelengths required is shown in Equation 6-16. Here, limited conversion of degree D means that a lightpath on wavelengthλ_i can be converted to a lightpath on any of the wavelengths λ_i to λ_{i+D}. In full wavelength conversion, the converted lightpath can have any of the wavelengths in the operating band that we want.

Equation 6-16

$$max(0, L - d) + L$$

Figure 6-4 shows wavelength reuse against load for two different topologies with the same number of nodes.

Figure 6-4. Wavelength Reuse



Number of LPs on One Wavelength
2.5
2
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
Load

Blocking Probability Computation in Optical Networks

The *blocking probability* of a network is a fundamental metric that judges the quality of the network performance (how much throughput and loss a network is subject to). In packet-switched networks, it can also be defined as packet loss probability. In circuit-switched networks, it can be defined as the ratio of the number of circuits dropped (lightpaths dropped) to the number of circuits set up. In this section, we consider the blocking probability of optical networks with and without wavelength conversion. Much of this work is adapted from a paper by Barry and Humblet in *IEEE JSAC* appeared in 1996[6].

When we compute the blocking probability in an optical network, we consider the circuit-switched approach; therefore, we are actually computing the blocking probability of a lightpath. Consider a network with N nodes and w wavelengths. Assume that the network has no wavelength conversion. For a lightpath request to be honored,[6] it has to satisfy two criteria:

- At least one available wavelength must be present in each of the h links (fiber) for a lightpath of 'h' hops.
- The available wavelength has to be the same in each link (wavelength continuity constraint).

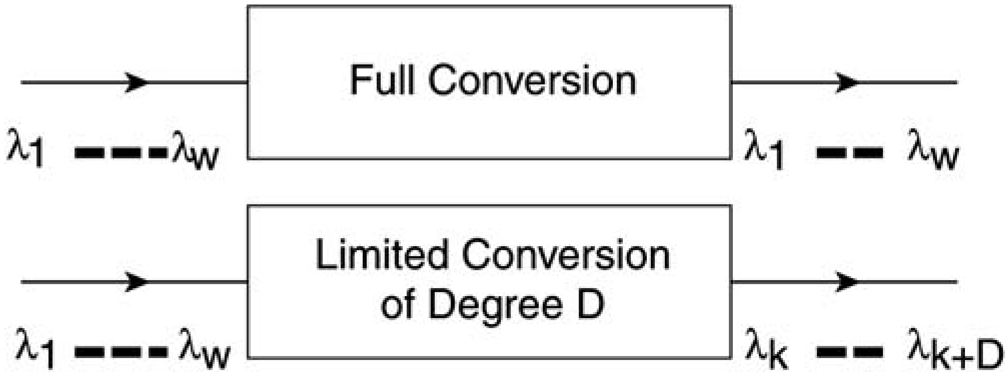
In contrast, if we consider an N node network with wavelength converters, then for a lightpath request to go through, the only criteria that needs to be satisfied is the following:

- At least one available wavelength (free wavelength) must be present in each of the 'h' links through which the lightpath traverses. (Note here that in the presence of wavelength converters, wavelength continuity is not necessary in the h-hop lightpath.)

We can see that by deploying wavelength conversion or translation in optical networks, the blocking probability decreases because only one criterion needs to be satisfied as opposed to the satisfaction of two criteria in the case without wavelength conversion. However, from actual results, the effects of full wavelength conversion on optical WDM networks is in fact minimal. In fact, recent work has shown that a limited wavelength conversion yields similar results to full wavelength conversion under certain generic assumptions. Limited wave-length conversion can be considered as converting any incident wavelength to any of a select band of wavelengths. The size of the band of wavelengths from which the output wavelength is selected is called the *degree of conversion*. For a full wavelength conversion, the degree $D = \lambda_{max}$, the maximum number of wavelengths in the network. An intuitive understanding of blocking probability in networks can be considered from the fact that blocking probability is somewhat proportional to the path length in number of hops. In other words, the greater the number of hops for a lightpath request, the better probability that this request will be blocked because of the difficulty it sees in finding a wavelength (λ_s) for each of the h hops.

Figure 6-5 shows a full wavelength conversion transponder and a limited wavelength conversion (of conversion degree D) transponder.

Figure 6-5. Full and Limited Wavelength Conversion Transponders



Now consider a lightpath request that is h hops long from source to destination. Let there be w wavelengths in each of the network link. If we assume ρ to be the probability that a wavelength is used (in other words, ρ is the probability of congestion on a given link), then ρ.w is the expected mean number of wavelengths that is used in a given link. $(1 - \rho)$ can be considered the probability that a desired wavelength is free on that link. Furthermore, the probability to find the same wavelength free on each of the h links would be $(1 - \rho)^h$. Then the blocking probability of that wavelength would be $(1 - (1 - \rho)^h)$. If we extend this concept over the entire w wavelengths, then the blocking probability for an h-hop link would be like that shown in Equation 6-17.

Equation 6-17

$$P_b = [1-(1-\rho)^n]^w - no\ wavelength\ conversion$$

Reference 6 showed that the utilization for a blocking probability P_{bi} for a WDM network without wavelength conversion is as shown in Equation 6-18.

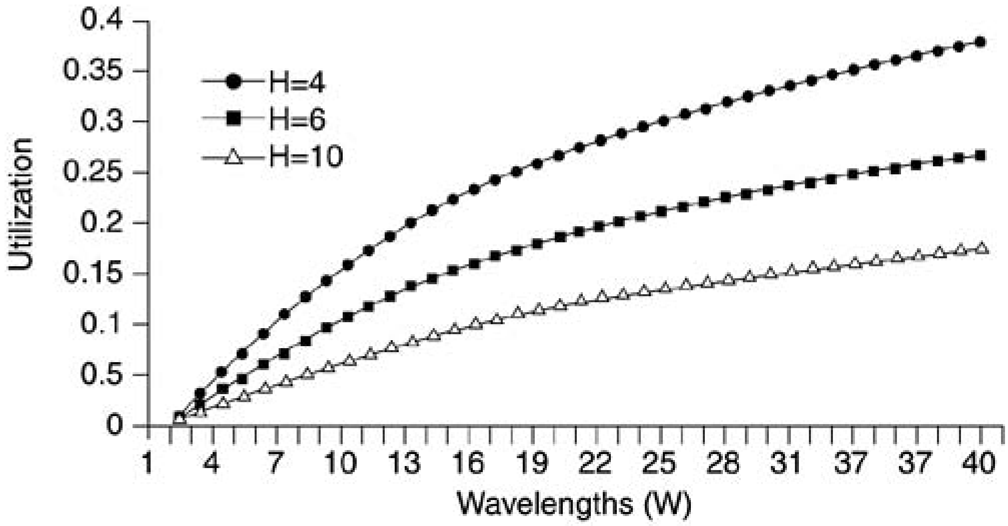
Equation 6-18

$$U_1 = 1-(1-P_{b1}^{1/w})^{1/h}$$

The preceding equation is actually the solution of Equation 6-17.

Figure 6-6 shows utilization verses number of wavelengths for P_{b1} = .001 (10^{-3}) without wavelength conversion.

Figure 6-6. Utilization as a Function of Different Numbers of Wavelengths for Various Values of H at a Blocking Probability of 0.001 with No Wavelength Conversion



Now consider a network with wavelength converters. The immediate effect is that there is no wavelength continuity constraint in the network anymore. We are assuming full wave-length conversion; such converters are present throughout the network at every node.

Again, if ρ is the probability of finding a wavelength busy, then ρ^w is the probability that all the wavelengths are busy. $1 - \rho^w$ is the probability that all the considered wavelengths are free (in the given link). Furthermore, $(1 - \rho^w)^h$ is the probability that all the w wave-lengths are free in all the h links. The blocking probability of a lightpath with wavelength conversion is shown as in Equation 6-19.

Equation 6-19

$$P_{b2} = 1 - (1 - \rho^W)^h$$

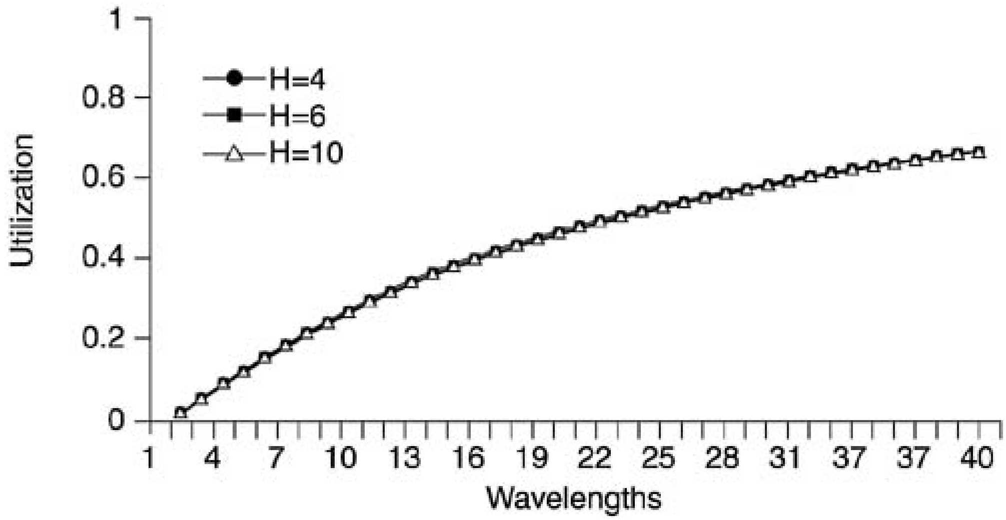
Similar to Equation 6-18, we can now solve Equation 6-19 and plot Equation 6-20 for the net utilization of this network.

Equation 6-20

$$U_2 = \left[1 - (1 - P_{b2})^{\frac{1}{h}}\right]^{\frac{1}{w}}$$

Figure 6-7 shows the utilization U_2 as a function of wavelength w for different path lengths using wavelength conversion.

Figure 6-7. Utilization as a Function for Different Hop Distances with Full Wavelength Conversion



Unlike the previous graph that used no wavelength converters, the hop length has minimal effect on network utilization if we use wavelength converters.

Summary

This chapter focused on routing and wavelength issues in optical networks. RWA analysis is important from a network planning and capacity planning objective. We learned some algorithms proposed by leading researchers about the various facets of the RWA problem. We also considered some manual as well as mathematical solutions to the RWA problem.

References

[1]Chlamtac, I., A. Ganz, and G. Karmi. *Lightpath Communications: An Approach to High Bandwidth Optical WANs' IEEE Transactions on Communications*. Vol. 40, No 7. July 1992.

[2]Zhang, Z., and A. Acampora. "A Heuristic Wavelength Assignment Algorithm for Multihop WDM Networks with Wavelength Routing and Wavelength Reuse." *IEEE/ACM Transactions on Networking*. Vol. 3, No. 3. 1995.

[3]Mukherjee, B. *Optical Communication Networks*. McGraw-Hill, 1997.

[4]Ramaswami, R., and K. Sivarajan. "Routing and Wavelength Assignment in All-Optical Networks."*IEEE/ACM Transactions on Networking*. Vol. 3, No. 5. 1995.

[5]Chlamtac, I. et al. "Lightnets: Topologies for High-Speed Optical Networks." *Journal of Lightwave Technology*. Vol. 11, No. 5/6. May/June 1993.

[6]Barry, R., and P. Humblet. "Models of Blocking Probability in All-Optical Networks With and Without Wavelength Changers." *IEEE Journal on Selected Areas in Communications*. Vol. 14, No. 5. June 1996.

[7]Narula, A., P. Lin, and E. Modiano. "Efficient Routing and Wavelength Assignment for Reconfigurable WDM Networks." *Journal on Selected Areas in Communication*. Jan 2002.

[8]Gibbons, A. *Algorithmic Graph Theory*. Cambridge University Press, 1985.

[9]Ramaswami, R. and G. Sasaki. "Multiwavelength Optical Networks with Limited Wavelength Conversion."*IEEE/ACM Trans. on Networking*. Vol. 6, No. 6. Dec. 1998.

Chapter 7. X over DWDM

Traditionally, the open standards interface (OSI) model assumes seven layers of hierarchy for data communication. The WDM layer can be viewed as the bottom-most layer in this hierarchical model. This layer is quite transparent to data formats or protocols; it just transports simple raw bits. Layers 2, 3, and 4 perform the function of data link control, network layer, and transport layer, respectively. They guide communication sessions between two or more points in a network.

More than 1600 protocols exist in today's Internet that facilitate the transport of datagrams and voice circuits across the network. In this chapter, we shall briefly view some of the protocols over WDM as they facilitate transport of information through a network. Although there is a great deal of literature on each of the individual protocols, there still is minimum detail about the use of these protocols directly over the WDM layer. It is our desire to throw some light on these issues and accustom the reader to the intricacies of these protocols.

It is not possible to consider each protocol as a stack over the WDM layer, but this chapter does consider some of the major protocols, such as Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH), Gigabit Ethernet (GE), Internet Protocol (IP), and Resilient Packet Rings (RPRs) over the WDM layer. Legacy SONET and Packet over SONET (POS)/SDH are the most common implementations for communication over the WDM layer. SONET, with its synchronous hierarchy, seems to be a natural choice for transport. This is supported by the fact that prior to the arrival of WDM networking, SONET and its Eurasian counterpart, SDH, were the primary transport phenomena for global networks. SONET/SDH boasts of a well-defined frame and classic synchronization between user nodes. SONET/SDH networks are also resilient to faults and often provide 50 ms protection times with automatic protection switching (APS)-like features. SONET has its share of drawbacks, the most important of which is the price. Other issues in SONET over WDM are the excess wastage in bandwidth (due to overhead) and problems in synchronization of the nodes with highspeed accurate clocks. A data implementation to SONET networks has been POS or packet over SONET interfaces. Due to the high cost and difficulties in implementing synchronous networks, Ethernet was extended as a possible interface over the WDM layer. The optical layer provides an excellent transport system for Ethernet, which otherwise has severe limitations at high rates above one gigabit per second in copper or coax transport media. Gigabit Ethernet and 10Gigabit Ethernet are soon set to replace conventional SONET/SDH networks, though this change would take a few years due to the enormous resources already invested in SONET/SDH. Ethernet has the distinct advantages of a simple frame format, ease of frame generation, and detection.

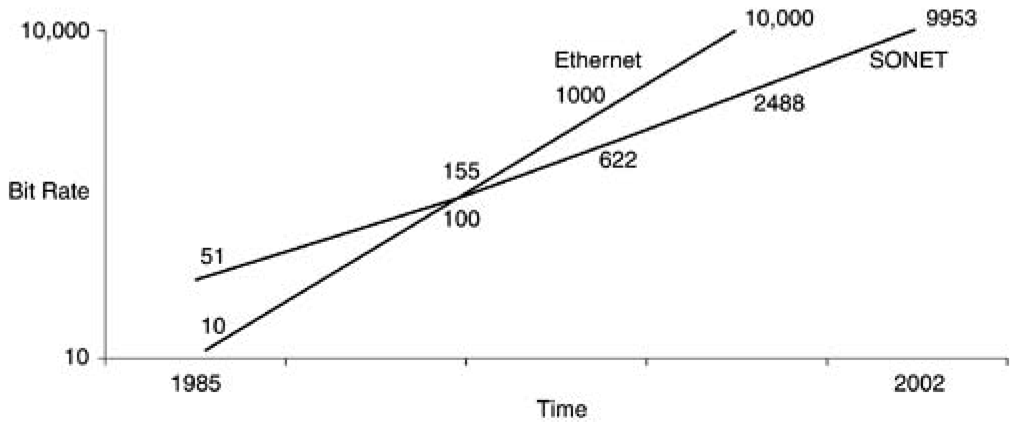
Raw IP over WDM is being muted as a possible approach to sending data over the WDM layer. The main issue is the lack of layer 2 framing. In one approach, HDLC-like framing was considered. In another approach, label-based framing like MPLS is also considered to serve as an approach to IP over WDM. Yet a third approach sends IP packets directly over WDM using CSMA CA (Carrier sense multiple access with collision avoidance) like scheme. The main issue is the slow progress of optical technologies. The need to have fast detectors, tunable lasers, and switches underlines much of the advance in IP over WDM. From the network perspective, IP over WDM becomes one of the most successful protocols in the entire stack. IP is a standard protocol that has easy implementation for transport, whereas WDM means cheap bandwidth. The marriage of IP and WDM is an excellent source to generate revenue. Although IP over WDM might be five years or more from strong implementation, some initial work has already been done on this approach.

Using resilient packet rings (RPRs) is another approach to transport packets over a ring network without the need for synchronization yet providing for strong protection, and an efficient overlay model. RPR and the Spatial Reuse Protocol (SRP) find vast applications in metro networks.

Gigabit Ethernet/ 10Gigabit Ethernet (Optical Ethernet)

Ethernet was defined in the 1980s as an open standard by a consortium, which was com-prised of Digital Equipment Corporation, Intel, and XEROX. Ethernet fits in layer 2 of the OSI model. By 1985, IEEE adopted the Ethernet standard and released the 802.3 standard. Since then, Ethernet has proven to be flexible. In the past 20 years, Ethernet has evolved from 10 Mbps half duplex to 10 Gbps full duplex, serving applications from local area networks (LANs), metropolitan area networks (MANs), and wide area networks (WAN). Figure 7-1 shows the evolution of Ethernet.

Figure 7-1. Evolution of Ethernet from 10 Mbps ➠ 100 Mbps ➠ 1 GE ➠ 10 GE



The initial implementation of Ethernet used a shared Media Access Control (MAC) mechanism known as carrier sense multiple access with collision detection (CSMA/CD). The bit rate deployed was typically 10 Mbps on a shared half-duplex copper medium and supports almost all LAN applications at that time. Ethernet then transformed to 100 Mbps full duplex—also known as Fast Ethernet—to support more bandwidth-demanding applications. With the introduction of switches, Ethernet could support point-to-point full-duplex links (unshared) at 100 Mbps over both copper and fiber. The medium of transport was primarily copper, which is used extensively in LAN applications. Fiber was used in applications that needed to support greater distances.

The next step in the evolution of Ethernet technology is Gigabit Ethernet (GE). The standard supports gigabit Ethernet frames to be transmitted on both copper and fiber medium. GE technology was soon adopted by enterprise networks, campus networks, and even by service provider networks very rapidly due to its low cost and good reliability.

The explosion of bandwidth-intensive applications combined with the growth of Internet data paved the way for 10 GE interfaces and applications. At 10 GE speeds, the real challenge is to make electrical signals to travel significant distance. This is not possible using copper media. The IEEE standard (802.3ae) is used for optical fiber medium as the transport option (no 10 GE over copper standards). 10 GE supports only full duplex operations over fiber medium; therefore, it does not need to implement collision-detection circuits. Note that detection of collision at the optical layer is not feasible, although it is possible. Apart from these differences mentioned, 10 GE is identical to 802.3-defined protocols.

Like SONET/SDH, optical Ethernet can also be transported over WDM networks. Basically, optical Ethernet can be transported using one of the following methods: over a dark fiber, over SONET/SDH (EoSONET) or pure Ethernet over WDM. As the Internet grows and IP traffic increases, Ethernet is an ideal technology for transport of data across WAN applications. The advanced Quality of Service (QOS) mechanisms enable Ethernet to offer voice, video, and data services. Ethernet is also low cost and simple (plug and play) compared to other technologies that are currently available.

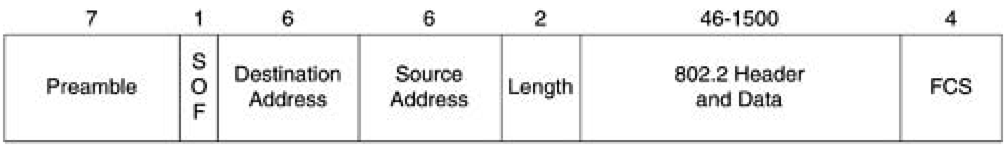
The following Ethernet standards have been established and defined by IEEE. Some of the relevant Ethernet-related standards are given here for reference: 802.3—10BASE-T, 10BASE-5, and 10BASE-2:

- 802.3u— Fast Ethernet (100 Mbps)
- 802.3z— Gigabit Ethernet (1000BASE-SX/LX/CX)
- 802.3ab— Gigabit Ethernet (1000BASE-T)
- 802.3ae— 10 GE
- 802.3af— Power over LAN (for example, VOIP applications)
- 802.3ah— Ethernet Passive Optical Network (EPON)

Ethernet Frame

A typical Ethernet 802.3 frame is shown in Figure 7-2.

Figure 7-2. IEEE 802.3 Ethernet Frame



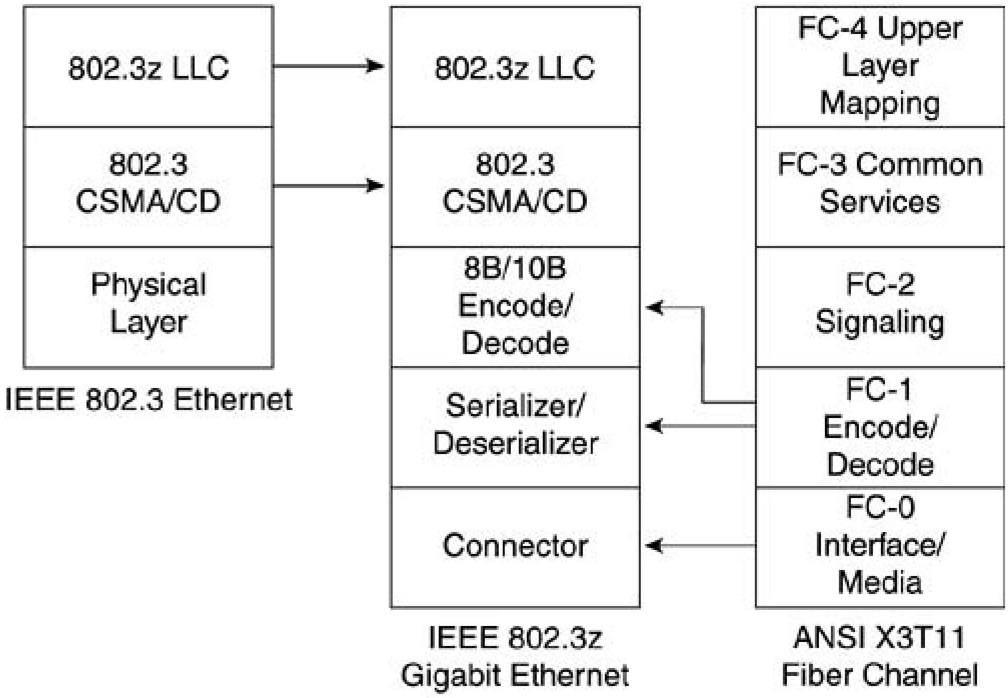
Preamble is an alternating pattern of ones and zeros that is 7 bytes long. It informs the receiver that a frame is coming. *Start of frame* (SOF) serves to synchronize the frame; it is 1 byte long. The next 6 bytes are the *destination address*, and the following 6 bytes are the *source address*. The*length field* is 2 bytes, and it indicates the number of bytes of data that is in the *data field. Frame check sequence* (FCS) is a 4-byte field that the sending device creates. The receiver recalculates the FCS to verify whether the frame arrived has an error. (The system discards a frame with an error.)

Gigabit Ethernet (GE)

Gigabit Ethernet (802.3z) is an extension of the 802.3 standard, and is built on the same Ethernet protocol and frame structure but the speed is increased ten times compared to 100 Mbps Ethernet (Fast Ethernet). The physical interface needs to be modified to move from 100

Mbps to 1 Gbps. Gigabit Ethernet has to look like Ethernet from the data link layer and up (it has to keep the basic Ethernet framing structure), and at the same time, it should be able to interface with high-speed optical interfaces. The merging of the IEEE 802.3 Ethernet frame format and the ANSI X3T11 Fiber channel high-speed physical interface technologies alleviates the challenges involved in architecting gigabit Ethernet in a timely manner. A Gigabit Ethernet stack structure is shown in Figure 7-3.

Figure 7-3. Gigabit Protocol Stack (Source: Internetworking Technologies Handbook, Second Edition)



The serializer/deserializer is responsible for encoding schemes and communication to the upper layers. The encoding scheme used in 802.3z is 8B/10B, which is similar to fiber channel encoding.

Gigabit Physical Layer

The gigabit Ethernet specification addresses three forms of transmission media:

- Long wave (LW) laser on SM/MM (single mode/multi mode) fiber known as LX
- Short wave (SW) laser on SM/MM (single mode/multi mode) fiber known as SX
- Transmission over copper known as CX

Many more gigabit interfaces are available in the market, such as interfaces for 1550 nm and wavelength division multiplexing (WDM) applications. The formal standard, Gigabit Media Independent Interface (GMII), is taken over by the Gigabit Interface converter (GBIC) standard and supports a wide array of applications. Table 7-1 shows different GBIC types and the distances they support.

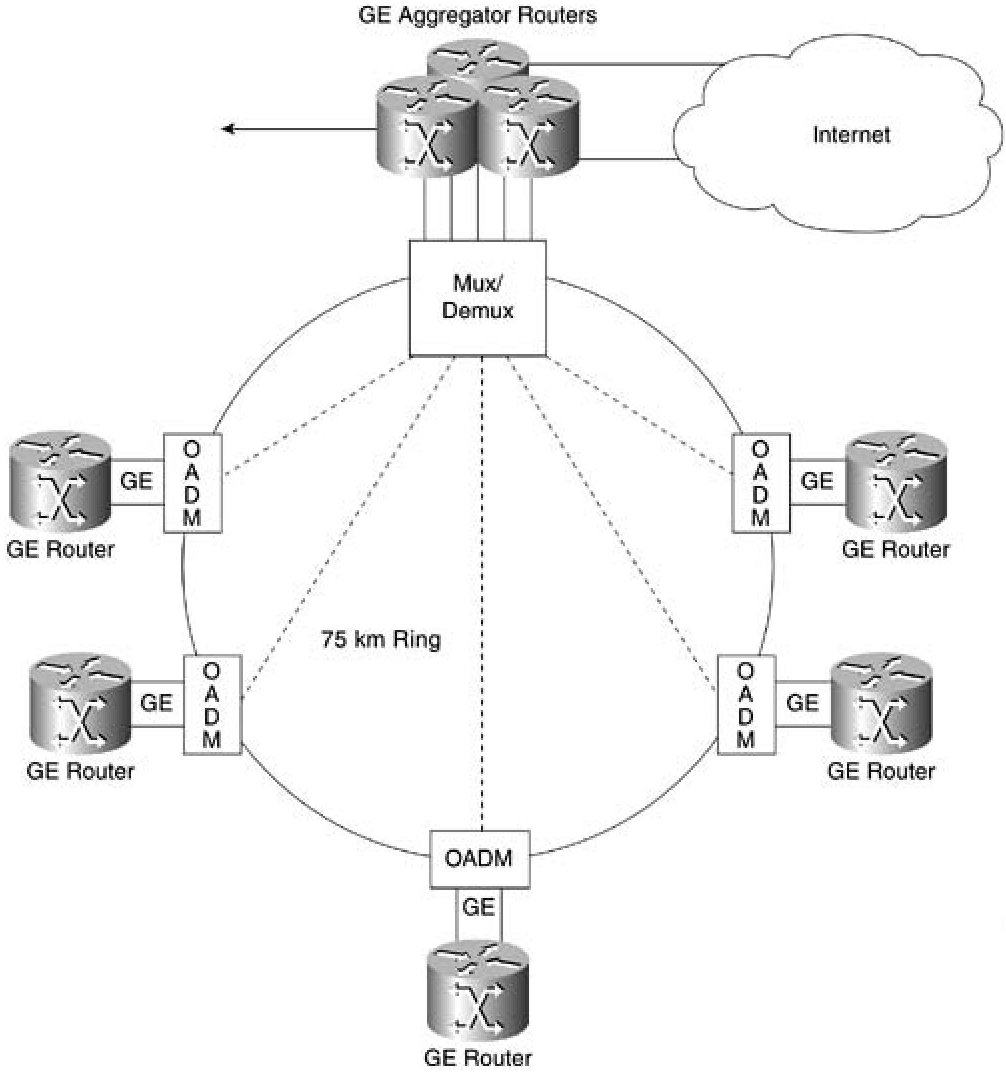
Table 7-1. GIBIC

Type	Cabling Distance
1000BASE-SX (850 nm) on SM/MM	Up to 550 m
1000BASE-LX/LH (1310 nm) MM	Up to 550 m
1000BASE-LX/LH (1310 nm) SM	Up to 10 km
1000BASE-ZX (1550 nm) SM	Up to 100 km
1000BASE-CX (RJ 45)	Max 100 m

GBIC provides hot swappable modules that we can install into any port of a GE device and configure per port basis at the network manager's interest. For example, an SX can be used in short-range applications and an LX GBIC can be used in long-range campus applications on the same gigabit Ethernet switch on different ports.

Gigabit Ethernet finds applications in LANs, storage area networks (SANs), campus applications (as campus backbone), and in metro. A typical gigabit Ethernet metro ring over DWDM is shown inFigure 7-4. This is a gigabit Ethernet GE hub-and-spoke architecture, where four gigabit Ethernet switches are hubbed into a master node. At the hub-end usually a high end GE switch like Cisco 7600 or Cisco 6500 resides, while at the satellite nodes a medium or low-end switch like Cisco 3550 or Cisco 4000 resides. The traffic is all hubbed at the center location and may be switched or routed depending upon the need.

Figure 7-4. GE over WDM (Hub-and-Spoke Architecture)



10 Gigabit Ethernet

10 gigabit Ethernet is the natural evolution of Ethernet from 1GE in speed and distance. The main purpose of 10 gigabit Ethernet is to extend the Ethernet application into WAN and long-haul applications. 10 gigabit Ethernet is a fiber only (optical) technology and is full duplex. Currently, 10 GE is standardized under the IEEE 802.3 ae task force. The task force has defined five criteria that the standard should meet: (Refer to http://grouper.ieee.org/groups/802/3/ae/criteria.pdf for more details on the task force)

- Broad market application, with multiple vendors supporting it along with multiple classes of services
- Compatibility with IEEE 802.3 standards and opens system interconnection (OSI) model
- Identity (do not want alternative problems; should be a solution for a problem)
- Technical possibility (should be able to technically demonstrate the technology)
- Economic feasibility (should justify cost)

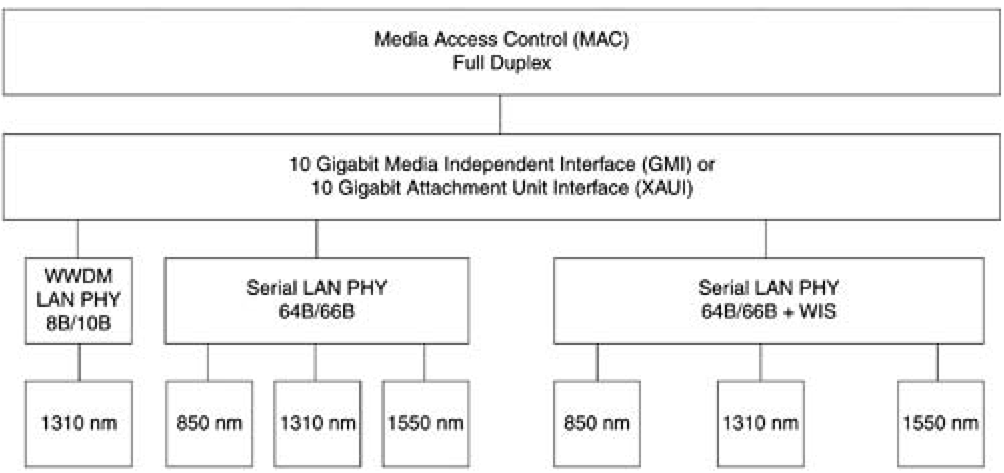
10 gigabit, Ethernet uses the IEEE 802.3 Media Access Control, frame format, and minimum and maximum frame size. Both 1 GE and 10 GE Ethernet physical layer devices (PHY) map to layer 1 of the OSI model, which is connected to the MAC layer (layer 2 of the OSI). The physical layer (PHY) is divided into the physical media dependent (PMD) sublayer and the physical coding sublayer (PCS). The transceiver (both optical and copper) is defined by the PMD sublayer. The PCS sublayer deals with coding and serializer/deserializer functions that are associated with the interface.

The 802.3ae specification defines two PHY types: LAN PHY and WAN PHY:

- LAN PHY— This is a serial interface that uses 64/66B encoding, with a data stream of 10 Gbps and a clock rate of 10.3 Gbps.
- WAN PHY— This is also a serial interface that uses 64/66B encoding, and it is compatible with SONET OC-192 steams with date streams of 9.953 Gbps. The advantage is that this interoperates with SONET OC-192 networks. The WAN Interface sublayer (WIS) takes care of the SONET/SDH framing.

The WAN PHY and LAN PHY are distinguished by the Physical coding sub-layer (PCS). Both of these PHYs operate over the PMD; therefore, they support the distance defined by the PMD. The architectural layer of WAN and LAN PHY is shown in Figure 7-5.

Figure 7-5. The Architectural Layout of the 10 GE (802.3ae) Standard (Source: www.10GEA.org 10 GE White Paper)

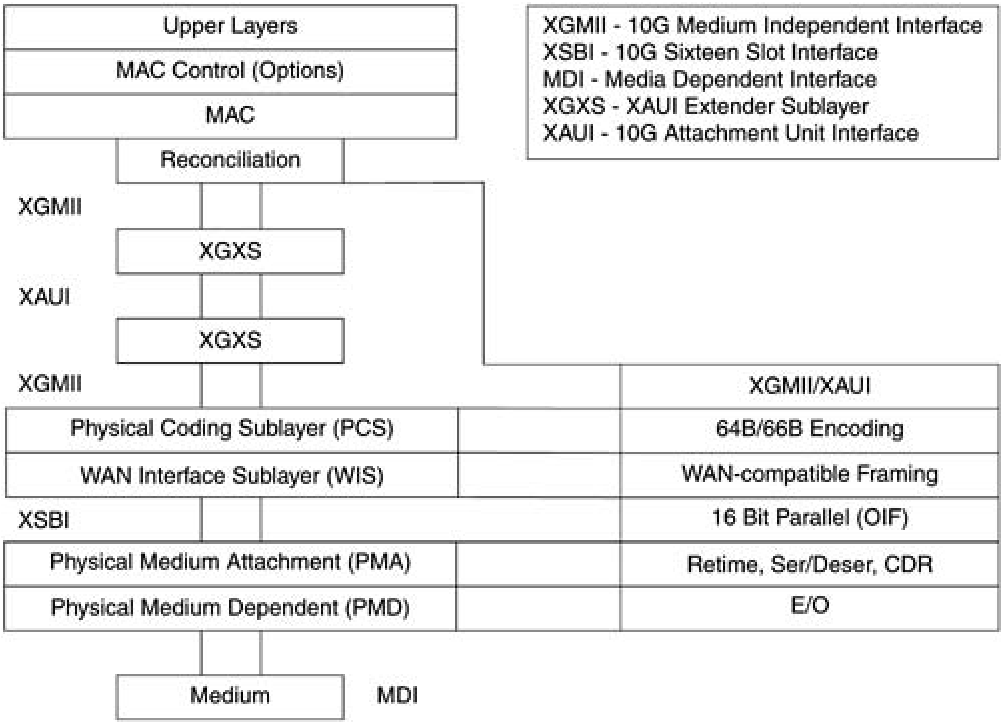


Between the MAC and the PHY is the 10 Gigabit Media Independent Interface (XGMII) or the 10 Gigabit Attachment Unit Interface (XAUI). The XGMII provides full duplex operation at a rate of 10 Gbps between the MAC and PHY. Each path (direction) is independent and contains a 32-bit data path and clock/control signals (74 bits wide in total). The separate transmission of clocking and data along with the timing requirements make the architecture challenging and limit the number of 10 GE ports on a line card.

To overcome these challenges, the 10 GE task force developed the 10 Gigabit attachment unit interface (XAUI). XAUI simplifies the routing of electrical connections and also helps to overcome the length limitation of the XGMII. XAUI also supports full duplex operations at line rate speed

(10 Gbps). XAUI helps the fan out of interfaces on multiport cards on the same line card or any 10 GE line card in the system chassis. The self-managed interface also helps to overcome the clocking issues that are associated with XGMII interfaces (http://www.10gea.org/Tech-whitepapers.htm).Figure 7-6 shows the function of the XAUI interface.

Figure 7-6. XAUI Function as an Extender Interface (Source: www.10GEA.org 10 GE White Paper)

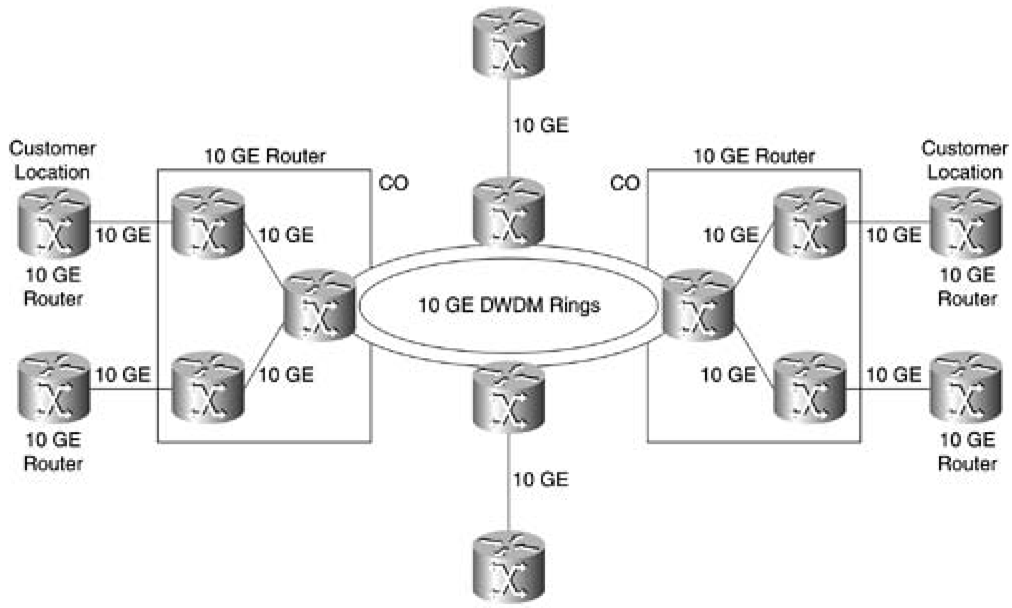


Applications 10 GE can be used in the following:

- Metro— The Internet service providers (ISPs) use 10 GE for low-cost transport of IP traffic across the metro networks.
- LANs— Enterprise backbone upgrade to alleviate traffic demand and connect between different sites.
- Storage— Disaster recovery, data backup, and video servers use 10 GE.
- WANs— 10 GE between central offices and long-haul networks using WAN PHY. This also ensures the interoperability with existing OC-192/STM-64 networks.

A typical wide area application is shown in Figure 7-7 (10 GE over DWDM).

Figure 7-7. GE over DWDM Network Connecting Multiple Central Offices (COs)



10 GE
Customer
Location
10 GE Router
CO
CO
10 GE Router
Customer
Location
10 GE
10 GE
10 GE
10 GE
10 GE
Router
10 GE DWDM Rings
10 GE
Router
10 GE
10 GE
10 GE
10 GE
10 GE
Router
10 GE
Router
10 GE

Ethernet Passive Optical Networks (EPON)

As the demand for voice, video, and data in the local loop (last mile) has increased, technologies such as cable modem and digital subscriber line (DSL) have begun deployment to alleviate the bandwidth problem. The data rate that DSL and ADSL provides is typically 128 kbps to 10 Mbps, and cable companies claim to provide data rates up to 10 Mbps on a shared basis. The problem is that these technologies cannot satisfy the new bandwidth demand for services such as video-on-demand (VOD), streaming video, and videophone, along with other regular Internet applications. Optical fiber is an ideal technology that is capable of delivering bandwidth to high-end applications at greater distances (>20 km) and at higher speeds. (Both cable modem and DSL are distance and bandwidth limited).

Passive optical networks (PONs) are one choice for connecting both residential and business access networks. A PON (passive optical network) is a point-to-multipoint optical network with passive elements in the signal path from source to destination. There are several versions of PON being muted, out of which, Ethernet-PON has gained a lot of attention due to the logical choice for carrying IP data. The QOS techniques that are available lead Ethernet-based networks to support voice, video, and data traffic efficiently. The advantage of using PON for subscriber loop (access network) makes it accommodate a larger coverage area, a higher bandwidth, reduced fiber deployment, multicast and broadcast support, ease of upgrade, and low cost of operation.

IEEE 802.3ah Ethernet in the First Mile (EFM) task force has been set up to define standards for Ethernet to the local loop access. The IEEE 802.3ah task force is focused on both Ethernet to the home (ETTH) as well as Ethernet to the business (ETTB) applications. At present, 802.3ah is focused on the following operations:

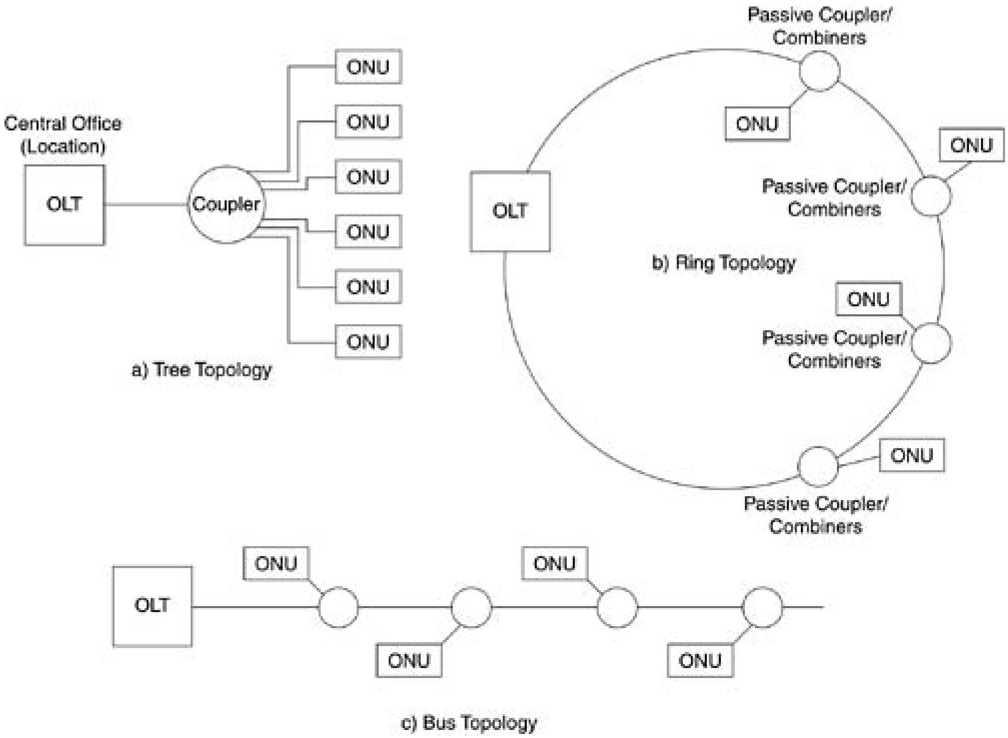
- Ethernet over copper
- Ethernet over point-to-point fiber
- Ethernet over point-to-multipoint fiber
- Operation Administration And Maintenance (OAM)-like features

The Ethernet over point-to-multipoint is concentrating over lower layers of the EPON. This includes defining control messages, slight modification of MAC (802.3), and physical (PHY) layer specification.

PON Topologies

Several multipoint topologies are defined for the first mile (access network), which includes bus topology, ring topology, tree topology, and tree-and-branch topology. PON can be deployed in any of these topologies by using optical splitters and optical combiners (see Figure 7-8). We can add redundancy to a part of the network or to the network as a whole. PON defines two types of network elements. One element resides at the central office and is known as optical line terminal (OLT). The other element is the customer location or the curb, which is known as the optical network unit (ONU). In the downstream (OLT to ONU), the PON is a multipoint solution; in the upstream (ONU to OLT), PON is a multiple point-to-point solution.

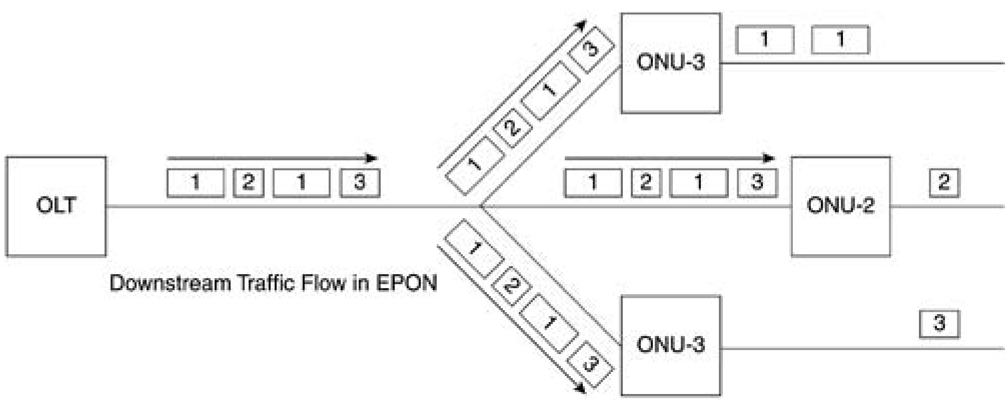
Figure 7-8. PON Topologies[4]



Ethernet Passive Optical Network

The property of EPON is such that it cannot be considered either a shared medium or a point-to-point network. EPON can be expressed as a mixture of both of these; in a downstream direction, it is point-to-multipoint. The passive splitters with splitting ratio 1: R are used (where R can be anywhere from 4–64 depending on the requirement), which splits the main fiber into 'R' arms. The signal that is passed into the main fiber shows up in all the arms due to splitter technology. This broadcast nature of Ethernet fits well with the downstream flow. For downstream transmission, an Ethernet packet is broadcasted and the packet is accepted by targeted ONUs by comparing the MAC address, much like in the Ethernet shared media approach. The flow of downstream traffic is shown in Figure 7-9.

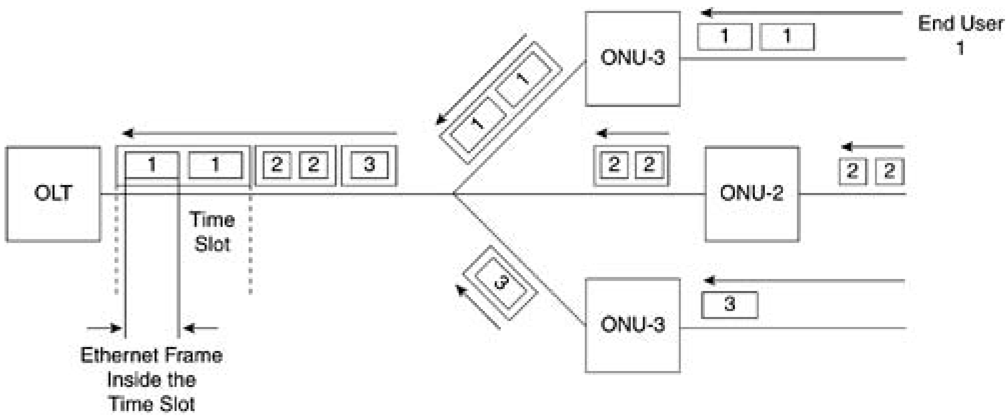
Figure 7-9. Downstream Traffic Flow in EPON



The upstream is much like a multiple point-to-point solution. The combiner combines all the optical signals into one fiber. Due to the nature of the combiner (passive coupler), an ONU can see only the OLT; ONUs cannot see other ONUs. This makes it hard to use Ethernet broadcast as an upstream technique because CSMA/CD works only in a shared media.

An Ethernet frame that is transmitted by two or more ONUs simultaneously collides. The ONU does not know that collision occurred unless a mechanism is available at the OLT to notify ONU of the collision. This kind of mechanism reduces the efficiency of the network by 30–40 percent. Refer to Figure 7-10 for upstream traffic flow.

Figure 7-10. Upstream Traffic Flow



We can also overcome the upstream communication issue by using WDM, where each ONU has its own upstream channel. The solution is simple and straightforward, but it requires tunable lasers and multiple ONU types, which increase the cost of the equipment and operation. At present, this is a viable solution even though it is not the best approach. In one embodiment, spectrally spliced SOAs are set up as ONUs, each at different frequencies.

Another way to solve the upstream communication issue is by time-sharing the optical channel (TDMA like), where each ONU is synchronized to a common time reference. Here, each ONU is allocated a time slot. The ONU stores all the frames until the time slot arrives and sends all the packets at full-line speed. During idle times (in which there is no packet to send), ONU sends an idle packet. Other schemes are available, but at present, there is no standardized way to multiplex packets (frames) efficiently. The basic idea when developing the protocol is to keep in mind that downstream communication is in broadcast mode and upstream communication is in multiple point-to-point mode. We face several challenges apart from the upstream optimization issue. Due to the unequal distance between the OLT and the ONU, the ONU should be able to handle different power requirements. Depending on the distances, ONU should be able to adjust the attenuation to handle received power and vary the transmit power to reach the required distance. Security is an issue in the downstream because every ONU can access all the packets in the downstream. We must implement MAC layer encryption or PHY layer encryption. The ONU should be able to handle voice, data, and video and should be able to migrate from the TDM network to the IP network seamlessly. The advance QoS and prioritizing techniques can handle all these services.

SONET/SDH

Synchronous Optical Network (SONET) and Synchronous Digital hierarchy (SDH) Network are two closely allied standards that define interface rates, formats, multiplexing methods, operations, administrations, maintenance, and provisioning (OAM&P) for transmission systems. SONET is a set of standards for North America, whereas SDH standards are used in Europe and Asia. Both SONET and SDH can transfer signals that are currently defined such as T1/E1 and so on, and they can accommodate any of the future requirements. In this chapter, we briefly describe the technology and discuss how SONET/SDH is used to transport signals over WDM networks (SONET/SDH over WDM).

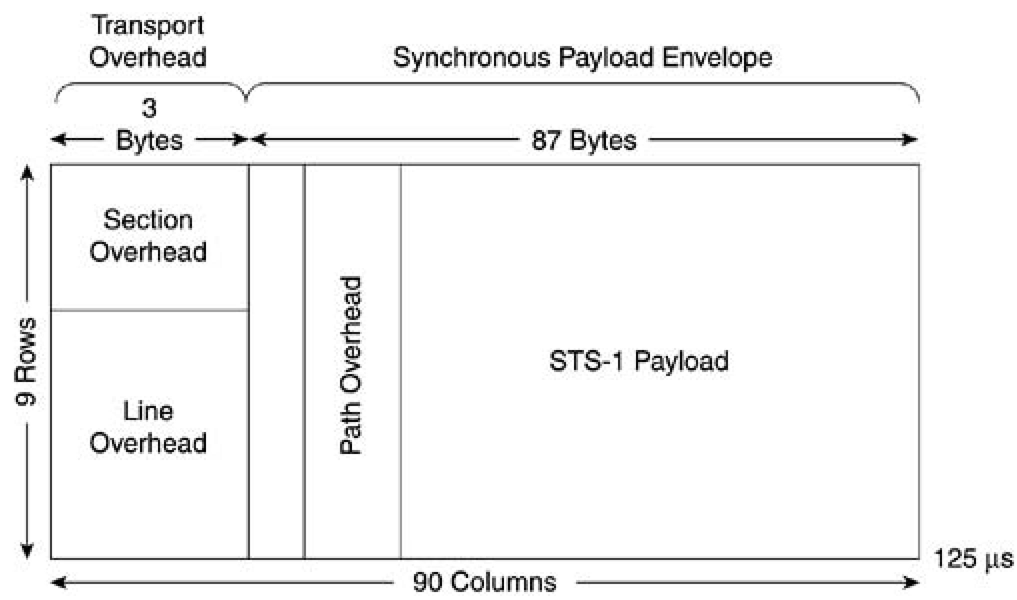
SONET/SDH Date Rate

As mentioned before, SONET and SDH are closely related standards. Telecommunication carriers widely use SONET and SDH for transport of voice, video, and data across their network.

SONET and SDH are considered as legacy technologies, but they are not going to vanish from carriers' networks in the near future. Even though the two technologies are closely related, there are some differences that make them distinct. In this section, we discuss the basic building block of SONET and SDH, the layered protocol, the multiplexing structure, and the common architectures (including SONET/SDH over WDM).

The basic building block of a SONET signal is called the *Synchronous Transport SignalLevel* -1 (STS-1). The STS-1 has a bit rate of 51.84 Mbps and is made up of 90 columns and 9 rows of octets per frame, as shown in Figure 7-11. Lower-rate payloads are mapped into STS-1s, and higher-rate signals are obtained by multiplexing (bit interleaved synchronous multiplexer) N frames of STS-1s to form an STS-N, where N = 1, 3, 12, 24, 48, 192. The transmission rate of STS-N is N * 51.84 Mbps. The order of transmission of the bytes is row by row, from left to right.

Figure 7-11. Basic SONET Building Block



The line rate of an STS-1 is calculated as shown in Equation 7-1.

Equation 7-1

$$\text{Line rate} = 9rows*90columns*\frac{1}{125\mu s}*8(bits/byte) = 51.84Mbps$$

(The line rate of STS-N is N times the line rate of STS-1; for example, STS-3 = 51.84 * 3 = 155.52 Mbps.)

Out of the 810 octets (9 * 90), 27 octets (the first 3 columns and 9 rows) are called the *transport overhead*, which is divided into two portions: the first three rows of section overhead and the next six rows of line overhead. The remaining 87 columns carry payload (synchronous payload envelope, or SPE), and a column of path overhead lies within the STS payload (SPE).

Due to this overhead, the transmission capacity decreases and is calculated as shown in Equation 7-2.

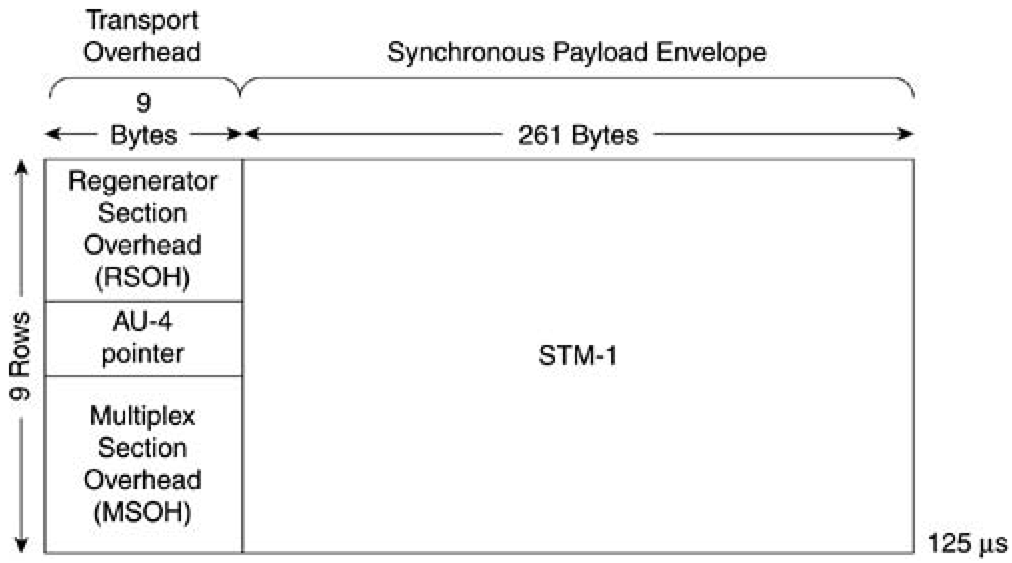
Equation 7-2

$$\text{Transmission capacity} = 9rows*87columns*\frac{1}{125\mu s}*8(bits/byte) = 50.112Mbps$$

Basic SDH Building Block

The basic building block of an SDH signal is called *Synchronous Transport Module Level* -1 (STM-1), and it has a bit rate of 155.52 Mbps (and uses 125us frame length). Each frame is made up of 270 columns and 9 rows of octet per frame. The lower-rate payloads are mapped into an STM-1, and higher-rate signals are obtained by multiplexing N STM-1 signals to form an STM-N signal, where N = 1, 4, 16, and 64 (defined by G.707).

Figure 7-12. Basic SDH Building Block



The line rate for an STM signal is calculated as shown in Equation 7-3.

Equation 7-3

$$\text{Line rate} = 9rows * 270columns * \frac{1}{125\mu s} * 8(bits/byte) = 155.52Mbps$$

Line Rate of STM-N is line rate of STM-1 * N, for example line rate of STM-4 = 155.52 * 4 = 622.08 Mbps.

Out of 2430 (9 rows and 270 columns) octets per frame, the first 9 columns are set aside for section overhead (SOH), and the section overhead is further divided into 3 areas. Of the first nine columns, the first three rows are known as *regeneration section overhead (RSOH)*, the next row contains pointers (H1, H2, and H3 bytes), and rows 5 through 9 are identified as *multiplexing section overhead* (MSOH). The remaining 261 columns carry the payload, which is identified by the virtual container (VC), and a column of path overhead lies along with the VC payload. The VC combined with row four (H1, H2, and H3 pointers) forms the administrative unit (AU). Like STS frames, the STM frame is transmitted row by row in a byte serial fashion.

Transmission capacity is the capacity of SDH STM-1 payload; it is calculated as shown in Equation 7-4. For further details, refer to Table 7-2.

Equation 7-4

$$\text{Transmission capacity} = 9rows * (270-9)Columns * \frac{1}{125\mu s} * 8(bits/byte) = 150.336Mbps$$

Table 7-2. SONET/SDH Digital Hierarchy

Optical Level	SONET (Electrical)	SDH	Line Rate (Mbps)	Payload Capacity (Line Rate – Overhead (Mbps)	SONET Capacity	SDH Capacity
OC-1	STS-1	–	51.840	50.112	28 DS-1s / 1 DS-3	21 E1s
OC-3	STS-3	STM-1	155.520	150.336	84 DS-1s / 3 DS-3s	63 E1s / 1 E4
OC-12	STS-12	STM-4	622.080	601.344	336 DS-1s / 12 DS-3s	252 E1s / 4 E4s
OC-48	STS-48	STM-16	2488.320	2405.376	1,344 DS-1s / 192 DS-3s	1,008 E1s / 16 E4s
OC-192	STS-192	STM-64	9953.280	9621.504	5,376 DS-1s / 192 DS-3s	4,032 E1s / 64 E4s

SONET/SDH Layers

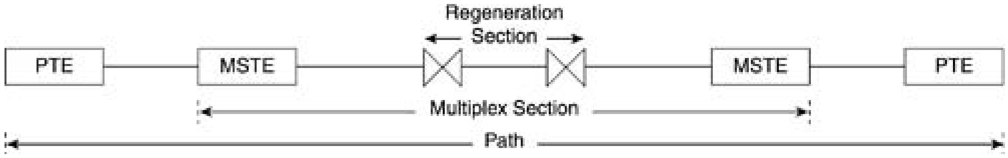
SONET/SDH (protocol) layer models are identical in concept, but they use different terms to define them. Four layers are defined in the protocol stack and are hierarchical in nature. Table 7-3 explains the function of each layer briefly.

Table 7-3. Some Generic Definitions of SONET/SDH

SONET	SDH	Definition
Path	Path	Map signals into an STS (SONET)/STM (SDH) frame
Line	Multiplex Section	Synchronization and multiplexing function for the path layer
Section	Regenerator Section	Transport of STS/STM frames across the physical medium
Photonic	Photonic	Conversion between electrical (STS/STM) optical signal (OC)

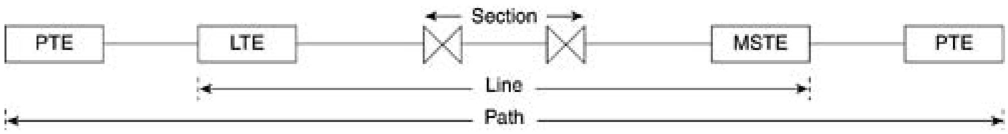
The primary responsibility of the path layer is to map the signals into the required format that is specified by the line layer (SONET)/Multiplex Section (SDH). This layer also defines the different transport services between the path-terminating equipment (PTE). At this layer, the path overheads are read, modified, and interpreted for performance and for automatic protection switching features. The line layer in SONET or the multiplex section in SDH provides multiplexing and synchronization for the path layer. It also deals with the transport of the overhead and payload of the path layer, in addition to performing error monitoring and protection switching at the line level. Figure 7-13 shows a generic SDH end-to-end connection.

Figure 7-13. SDH End-to-End Connection



The section layer (SONET/Regeneration Section, or SDH) is responsible for transport of STS-N /STM-N frames across the network media. This layer's functions include framing, scrambling, maintenance, and error monitoring. The photonic layer is where the conversion of the signals happens before the signal is transported through the fiber. This layer determines the launch power, shape, and frequency of the optical pulse.

Figure 7-14. SONET End-to-End Connection

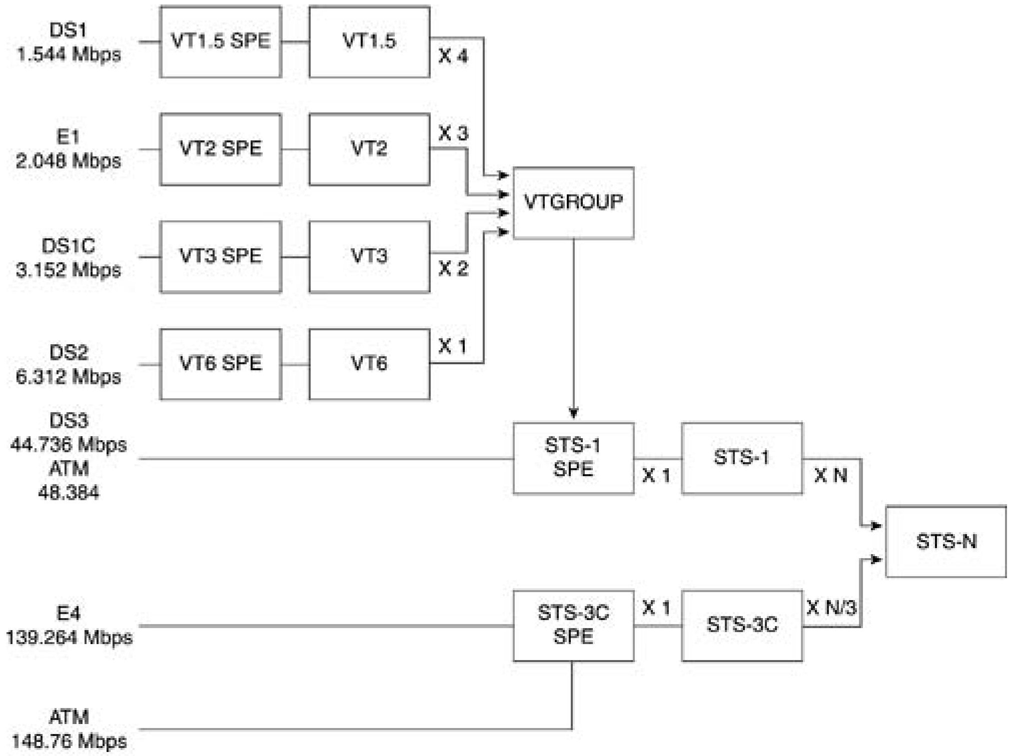


SONET and SDH carry a comparatively high percentage of overhead in each layer. A detailed description of overheads is beyond the capacity of this book; readers are advised to refer to SONET/SDH standards or any SONET/SDH books that are available in the market for more details.

SONET/SDH Multiplexing Structure

The principle of multiplexing SONET and SDH are the same, but the mapping schemes for SONET and SDH are different. The basic SONET signal (STS-1) can carry large payloads (>50 Mbps). At the same time, it can accommodate the lower-speed signals such as T1/E1 and so on. To achieve this, STS payload is divided into smaller structures called *virtual tributaries* (VTs). VTs are used for switching and transporting payloads smaller than STS-1. T3 (DS3) signals are carried in an STS-1 payload. (The remaining bits in STS-1 are stuffed with dummy bits.) N numbers of STS-1 signals are then multiplexed to form STS-N. Figure 7-15 shows the basic multiplexing scheme for SONET frames.

Figure 7-15. SONET Multiplexing Structure[1]



The SDH multiplexing scheme is shown in Figure 7-16. The lower-speed signals are first mapped into one of the five containers (Cs) according to the bandwidth needs. Path Over Head (POH) is added to the Cs to get VCs. Two types of VCs are defined: higher-order and lower-order VCs. The higher-order VCs (VC-3 and VC-4) are mapped either to Administrative Unit – 3 (AU-3) or Administrative Unit –4 (AU-4). Lower-order VCs (VC-3, VC-2, VC-11, and VC-12) are mapped to VC-3/VC-4 along with tributary unit (TU) pointers, which are used to locate the VCs. One TU-2, three TU-12, or four TU-11 forms are mapped into a tributary unit group (TUG). One AU-4 or three AU-3s are mapped to an administrative unit group (AUG), which is the "SDH payload envelope" of the STM-1. N STM-1 signals are byte-interleaved multiplexed to create STM-N.

Figure 7-16. SDH Multiplexing Structure

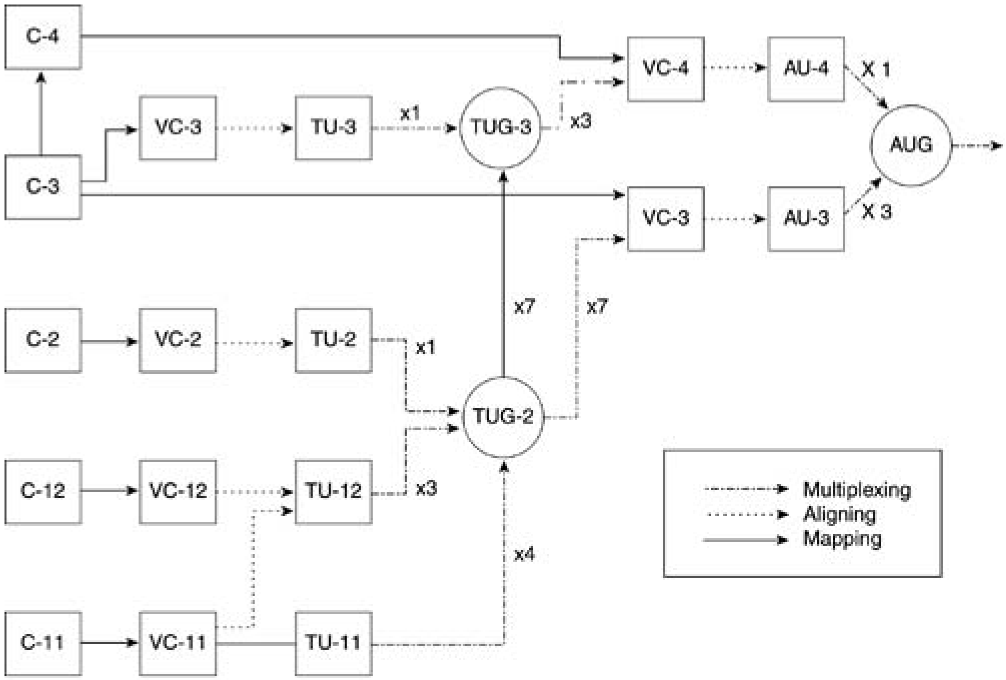
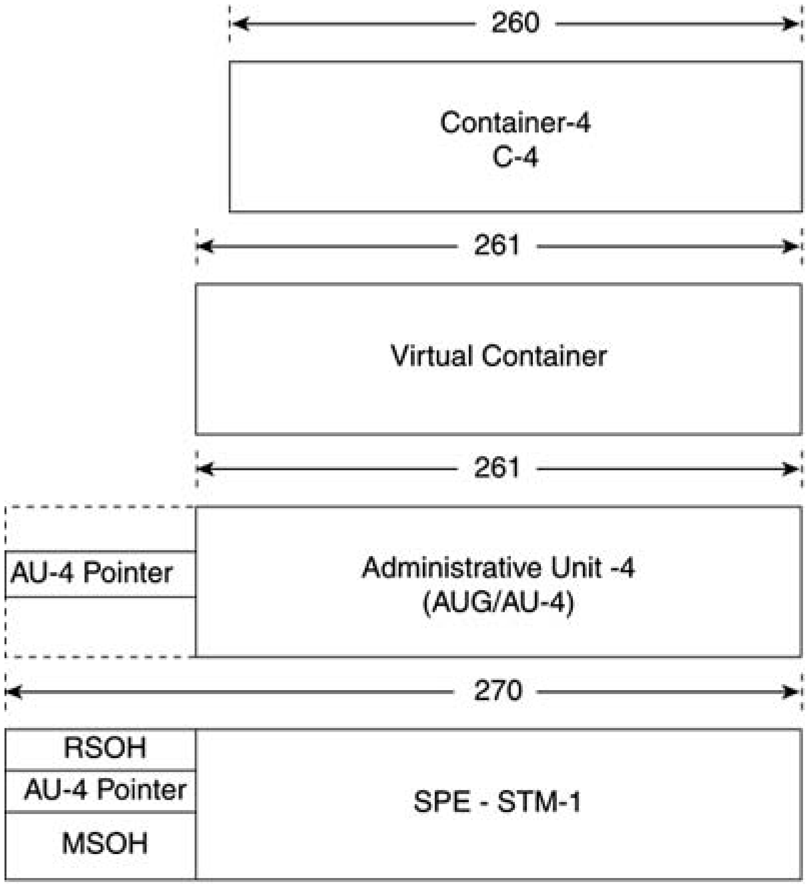


Figure 7-17 shows how a C-4 container is mapped and how an STM-1 signal is formed.

Figure 7-17. STM-1 Signal Is Derived by Adding Overheads[1]



The C-4 container occupies 9 rows and 260 columns. As explained before, VC-4 is obtained by adding a row of path over head (POH) to C-4 signal. To VC-4 signal, a 9 byte of AU-4 pointer is added to get the AU-4 signal. Transport overhead (RSOH and MSOH) is added to AU-4 to get the STM-1 signal.

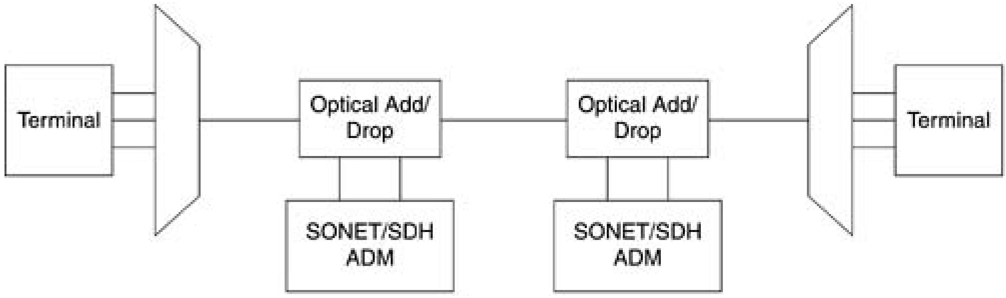
SONET and SDH Architectures

SONET and SDH support linear add/drop and ring architectures, both of which are discussed in detail in this section.

Linear Add/Drop

An initial implementation of linear SONET/SDH architecture is shown in Figure 7-18. In this type of application, traffic between the nodes is carried across multiple nodes that are cascaded in series. Service traffic can originate and terminate between any nodes in this cascaded chain. The end nodes in this configuration are called terminal nodes, and intermediate nodes are called add/drop nodes (ADM). As bandwidth demand increases, WDM can be added to alleviate the capacity constraints by placing multiple SONET streams on different wavelengths.

Figure 7-18. SONET/SDH Linear Architecture over DWDM



Ring Architectures

As the requirements for reliability increased, ring architectures were introduced to support various survivable network topologies. Ring architectures are differentiated by the direction of the routing and protection mechanism. A unidirectional ring carries service traffic in only one direction (clockwise). In contrast, a bidirectional ring carries service traffic in both a clockwise (CW) and a counterclockwise (CCW) direction. Furthermore, a *path-switched ring* protects traffic based on the conditions of the entire path. (A path is an end-to-end service connection.) A *line-switched ring* switches based on the conditions between each pair of node. When an error condition or fault is located in a line, an entire line is switched to the protected line. Based on these routing and protection schemes, the following four types of rings have been defined for SONET/SDH networks:

- A unidirectional line-switched ring (ULSR)
- A bidirectional line switched rings (BLSR)
- A unidirectional path-switched ring (UPSR)
- A bidirectional path-switched ring (BPSR)

A detailed analysis is beyond the scope of this book; therefore, readers are advised to refer to the standards. A WDM network can support all of these architectures.

Packet over SONET/SDH (POS)

Even though SONET/SDH is optimized for efficient transport of voice traffic, its capability to transport high bandwidth data makes it an excellent choice to connect packet (IP) networks at high speeds. POS technology provides efficient transport of data over SONET/SDH networks, and it can be used in a variety of applications. POS places the IP layer (layer 3) on top of the SONET/SDH layer (layer 2) in the OSI model. POS can offer IP service and is used primarily in Internet backbones, data aggregation, and metro applications. It is designed for point-to-point links and is typically used over SONET/SDH over DWDM networks.

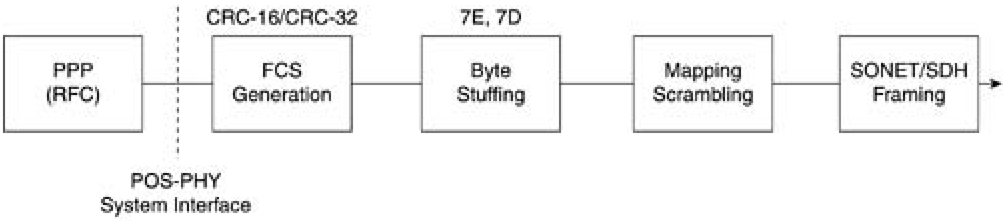
POS uses Point-to-Point Protocol (PPP) encapsulation in high-level data link control (HDLC)-like framing to encapsulate data at layer 2. PPP consists of two protocols: link control protocol (which establishes the connection and tests the data link) and network control protocol (which identifies layer 3/IP protocol). PPP in HDLC like-framing is shown in Figure 7-19.

Figure 7-19. PPP in HDLC-Like Framing

The basic data rate for POS specified is STS-3c for SONET and STM-1 for SDH. Out of 155 Mbps bandwidth, the useable bandwidth is about 149.76 Mbps. The POS frames are mapped into STS/STM-N frames and occupy the payload envelope.

The process of mapping IP into SONET/SDH using PPP is shown in Figure 7-20.

Figure 7-20. Mapping of IP into SONET/SDH



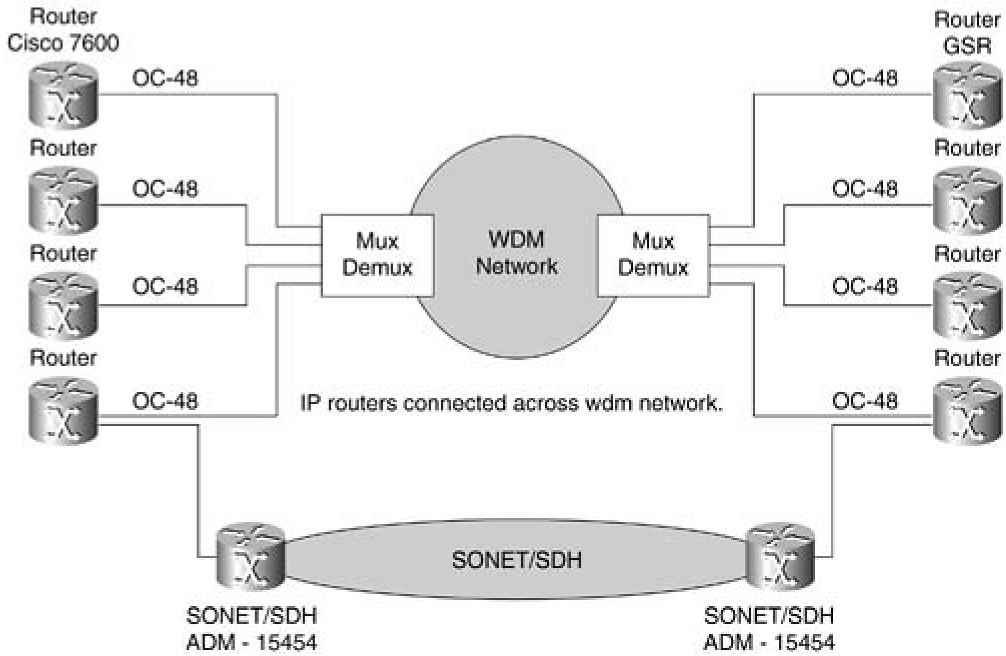
In addition to high bandwidth, POS offers reliable and secure data transmission. The data is scrambled using an ATM-like scrambler, self-synchronized payload scrambler, which is defined by RFC 2615. When POS is connected to the SONET/SDH network, the timing is derived from the SONET/SDH network. When POS is connected through dark fiber or DWDM, timing is derived from the internal clock. (Links are independently timed by 20-ppm internal clock built into the POS interface.)

POS implementation documents are available from the IETF web site. For advanced reading, please refer to the following documents:

- RFC 1619/RFC 2615, "PPP over SONET/SDH"
- RFC 1661, "Point-to-Point Protocol"
- RFC 1662, "PPP in HDLC-Like Framing"
- RFC 1548, "The Point-to-Point Protocol"

POS interfaces are implemented on routers and Layer 3 switches with interfaces speeds ranging from OC-3/STM-1 to OC-192/STM-64. Implementing POS over OC192/STM-64 is challenging because byte processing at 10 G speeds is a complex process. IETF proposed a 32-bit word-oriented approach to alleviate this problem. (Please check with www.ietf.org for more details.) POS is typically used in network backbone infrastructures, data aggregation, and network applications. This is accomplished with the POS interface (router) connected to a SONET/SDH add-drop multiplexer (ADM) and then transported across the network. POS over dark fiber and DWDM systems is also becoming popular. An example of transport of POS over WDM and over SONET is shown in Figure 7-21.

Figure 7-21. POS over DWDM



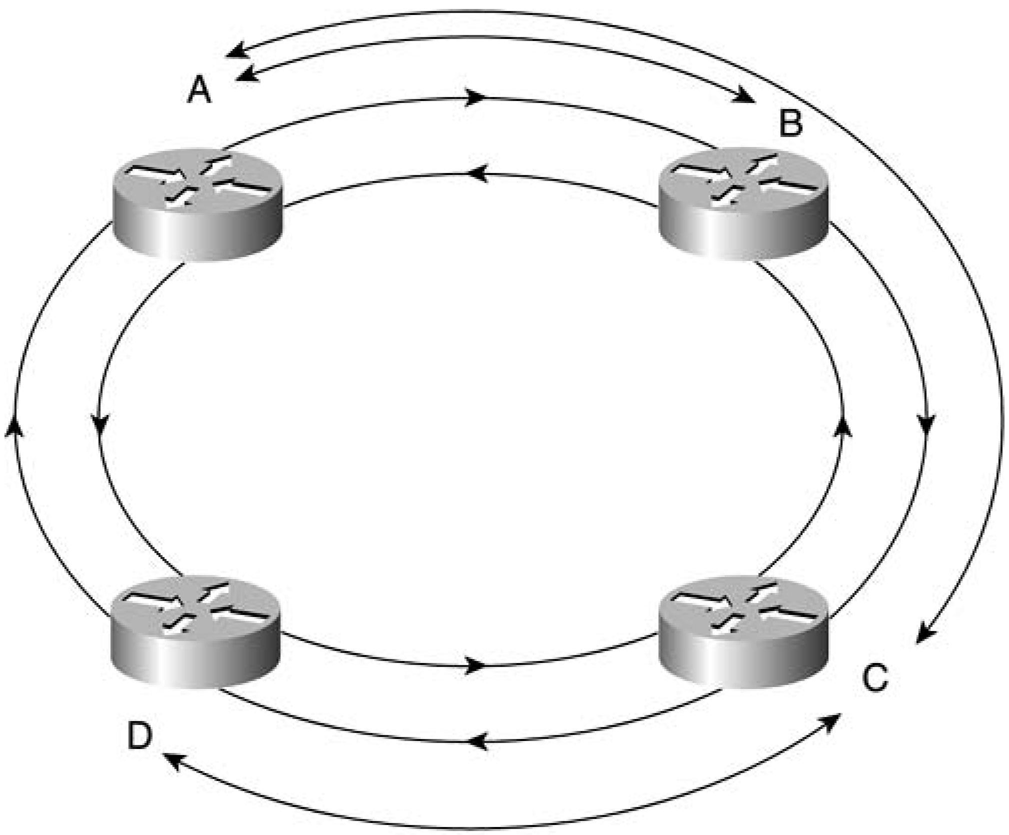
Dynamic Packet Transport (DPT) / Resilient Packet Ring (RPR)

Dynamic Packet Transport (DPT) defines a new technology for a packet-optimized transport system. DPT is based on the Spatial Reuse Protocol (SRP), a MAC-layer protocol for ring-based packet networking. SRP is open and freely available as IETF Informational RFC 2892. DPT/RPR combines the bandwidth-efficient rich service capabilities of IP and protection of SONET rings, thereby presenting a cost-effective solution to transport packet.

DPT/RPR is a dual-counter rotating ring. (These rings are referred to as inner ring and outer ring). Both of these rings can be used to pass data and control traffic. If data is sent in a clockwise direction, a control packet is sent in a counter-clockwise direction. (There is no corresponding control frame for every data frame. Control frames are sent whenever necessary for control information and have no relationship with data packets).

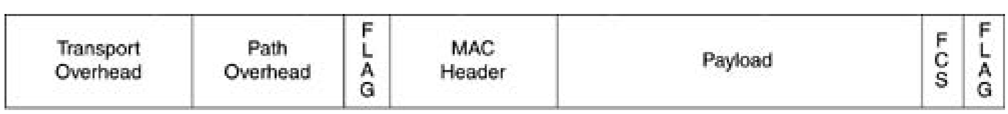
As mentioned previously, DPT uses Spatial Reuse Protocol (SRP). SRP derives its name from the spatial reuse property and is possible because of the destination stripping capability. The control packets are used for sending keepalive messages, protection switching, topology discovery, and bandwidth control. SRP supports multicast, quality of service (QOS), and intelligent layer 3 protection. Fairness is enforced by SRP fairness algorithm (SRP-fa) and is used to control access to the SRP ring. Refer to Figure 7-22.

Figure 7-22. SRP Ring Showing Spatial Reuse. While A is talking to B, A is also talking to C, and D is talking to C.



SRP is a media-independent media access control (MAC) layer protocol. The initial implementation utilizes SONET/SDH framing so that interworking with current SONET/SDH infrastructure is possible. A typical SRP frame is shown in Figure 7-23. Each SRP interface has a unique 48-bit IEEE MAC address (Ethernet-like MAC address). A multicast bit is also defined to support multicasting. In multicasting mode SRP, the source strips the multicast packet (not destination stripping).

Figure 7-23. SRP Packet in SONET/SDH Frame



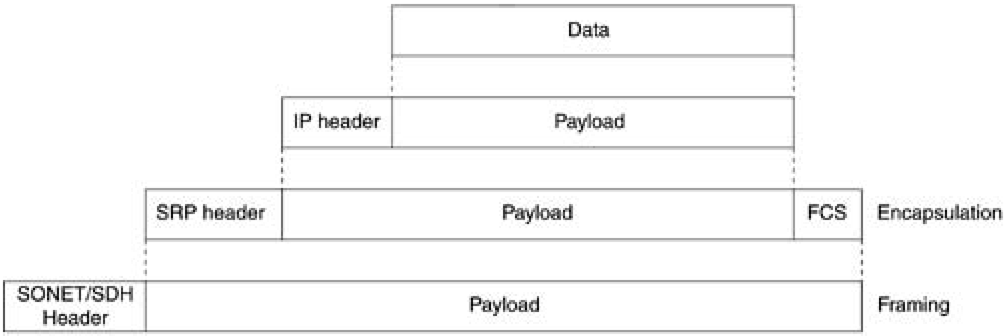
Both outer and inner rings use the same IP addressing schemes to enable fast reoptimization of the ring (path) selection and to avoid router flaps when ring wraps occur. (Ring wraps happen in layers 1/2.)

SRP uses intelligent protection switching (IPS), which is analogous to SONET/SDH protection/switching schemes. However, no protection bandwidth is allocated here, thereby making SRP more efficient. IPS can recover from nodal failure, fiber failure, and signal degradation. The SRP nodes can switch in 50 ms or less, similar to SONET/SDH switching.

As mentioned previously, the initial implementation of SRP is based on a concatenated

SONET/SDH frame. The SRP/DPT defines a bit interface to the physical layer (SONET/SDH), and the bit streams are mapped into the SONET/SDH payload (see Figure 7-24).

Figure 7-24. Figure 7-24

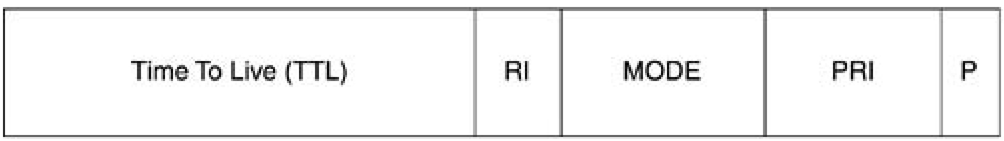


The transmission rates currently defined are OC-12/STM-4, OC-48/STM-16, and OC–192/STM-64 for the SRP/DPT interfaces. Most of the implementation currently is based on OC-12/STM-4 or OC-48/STM-16, but OC-192/STM-64 is gaining popularity. The OC-192/STM-64 interface is deployed on the GSR 12400 series routers.

SRP Generic Frame Header

Two types of packets are available: data and control packets. A generic packet header is 16 bits long and is shown in Figure 7-25.

Figure 7-25. SRP Frame Generic Header



- Time To Live (TTL)— This uses 8 bits. Like IP TTL, each node decrements the TTL field by one when forwarding the packet. When the TTL value reaches 0, the packet is stripped off the ring. Theoretically, 8 bits of TTL allows for 255 hops, and the theoretical maximum number of nodes in a ring is 255. In practice, the maximum number of nodes is 50–60 in a ring. (Note that this is pure layer 2 TTL and has no relationship with IP TTL.)
- Ring identifier (RI)— This is 1 bit long and designates the original source ring. It is used to make forwarding decisions and acceptance of packets, especially in a ring wrap situation. RI has a value of 0 for the outer ring and 1 for the inner ring.
- Mode— This uses the next 3 bits and is used to identify the packet type. The mode values are given in Table 7-4.

Table 7-4. Mode Values

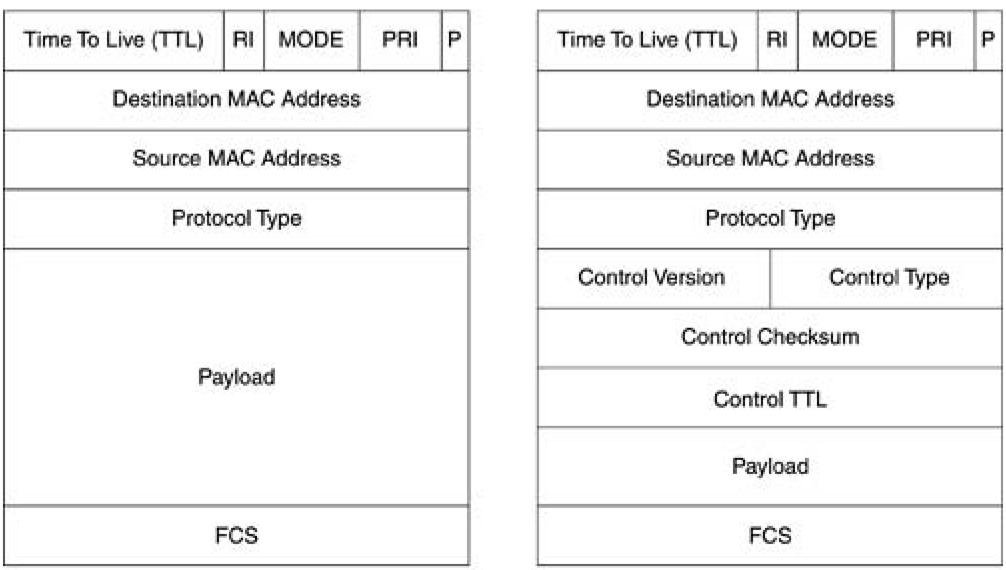
Value	Description
000–010	Reserved
011	ATM data cell
100	Control message (pass to host)
101	Control message (locally buffered for host)
110	Usage packet
111	Data packet

- PRI— This is the priority bit that uses the next 3 bits. Values can be from 0–7; higher values represent higher priority.
- Parity bit (P)— This is 1 bit long and is the odd parity over the previous 15 MAC header bits.

Generic Data and Control Packet

A generic SRP data pack is shown in Figure 7-26. The data pack contains the header, the destination MAC address, the source MAC address, the protocol type, and the payload. The protocol type field is 16 bits long and indicates what type of data is in the payload (IPv4, ARP, or SRP protocol).

Figure 7-26. Generic Data Pack for SRP Implementation



Control packet is similar to data packet with extra fields (see Figure 7-26b) like control version, control type checksum, and control TTL. The control version field indicates the version number that is associated with the control type. The control type field defines the control message type: 0x01 - Topology discovery, 0x02 - IPS message, and 0x03 - 0xFF - Reserved. Control TTL field is decremented each time a node relays a control packet. If the node receives a control packet with TTL <= 1, it is accepted but is not relayed (forwarded).

Topology Discovery and Ring Selection

Each node sends topology discovery packets on the outer ring; every node on the ring appends its MAC address and sends it to the next node on the ring. (If there is a wrap on the ring, the wrapped node indicates a wrap when appending its MAC). The control packets have to be received on the same ring that is sent out before the node accepts this as a valid packet. Once the node receives two consecutive and identical topology packets, the node builds a topology map. Therefore, the topology map includes information such as MAC address and wrap-status of each node.

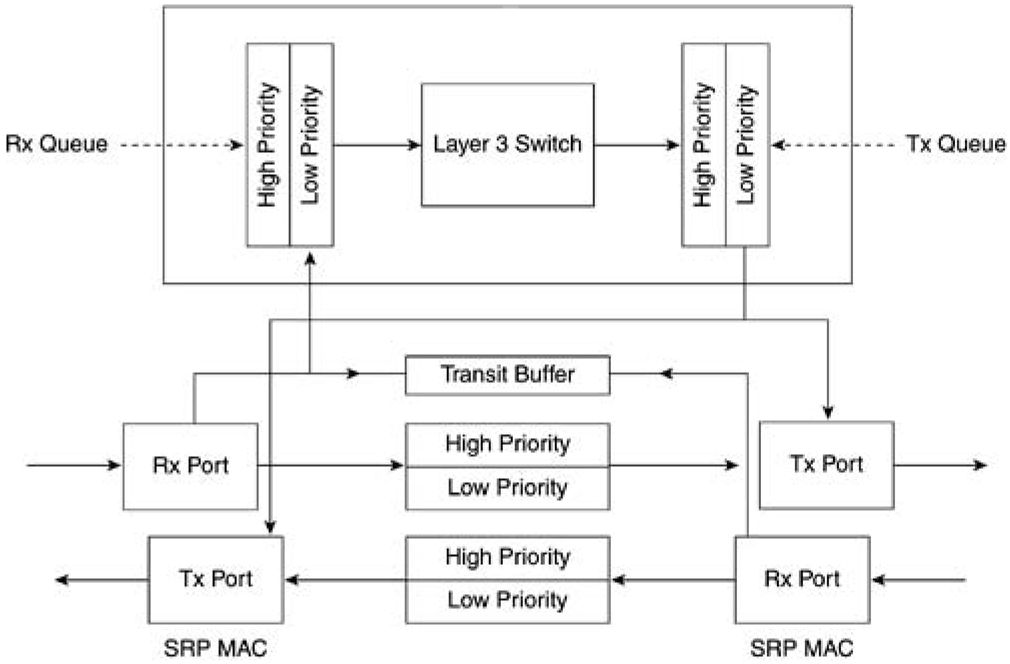
Ring selection is done by ARP (address resolution protocol) techniques (ARPing). ARP requests are sent out alternating on both rings. The nodes that respond utilize the topology map and answer to the requested node by using the shortest hop count as a metric for ring selection.

Packet Processing and Ring (Packet) Flow

At each SRP node, the incoming packets are either accepted or processed or transmitted to the next node without layer 3 processing. The MAC (logic) layer consists of a transit buffer, a receiver queue (Rx), and a transmitter queue (Tx). This layer is responsible for implementing fairness and scheduling traffic. The control packets are point-to-point (dropped at every node) and processed. The control packet information is extracted and forwarded to the upper layers for processing. The data packets are checked against a content-addressable memory (CAM) table to see whether they are destined for that node. If they are, they are stripped from the ring and forwarded to the host for processing. If the data is not destined to that node, the TTL field is checked. If the field is greater than unity, the TTL field is then decremented and put in the high or low priority transit buffer depending on the PRI field.

The locally generated traffic goes into a high or low transmit queue, depending on the PRI value. The scheduler selects the packet from the transmit/transit queue; the priority from top to bottom is the transit high priority packets, the transmit high priority packets, the transit low priority packets, and finally the transmit low priority packets. This whole process is managed by SRP fairness algorithm (SRP-fa). Figure 7-27 illustrates SRP node packet processing.

Figure 7-27. SRP Node Packet Processing



Under normal conditions, packets on the outer ring remain on the outer ring until a destination node removes them. If a wrap occurs, the packets on the outer ring show up on the inner ring.

SRP Bandwidth Fairness (SRP-fa)

SRP rings implement the fairness algorithm to regulate bandwidth. Because the ring bandwidth is a shared resource, it is possible that a single node or user in the ring can take over the whole bandwidth. The SRP fairness algorithm is the underlining mechanism that ensures fairness. The following high-level fairness is defined in the SRP ring:

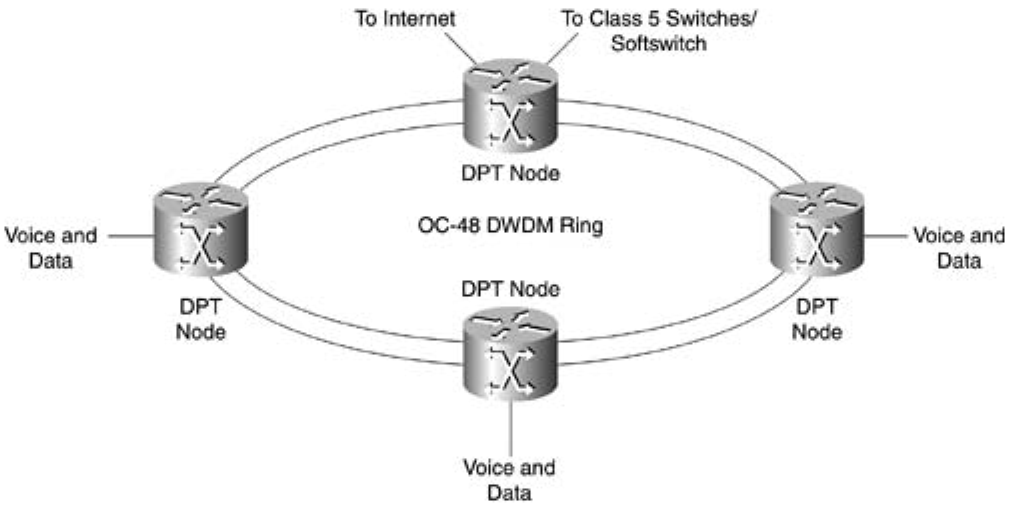
- Global fairness— Each node get its fair share of bandwidth; other nodes cannot create a starvation condition in the network.
- Local optimization— By use of the spatial reuse property of SRP, any node can use more than the fair share of bandwidth when no other nodes are contesting for bandwidth.
- Scalability— This is the ability to deploy large rings (32–64 nodes) in a disperse geographical location.

Every node monitors the number of packets that are sourced and forwarded. When a node experiences congestion, it sends the transmit usage counter to the upstream neighbor via the opposite ring. When the upstream neighbor receives the usage information, it throttles the transmit rate to the advertised usage rate. The SRP fairness applies only to low-priority packets. High-priority packets do not obey fairness rules and can be rate limited by other rate-limiting techniques like committed access rate (CAR). More information on how SPR-fa works can be found at the Cisco web site at http://www.cisco.com/warp/public/cc/pd/rt/12000/dptlc/index.shtml.

SRP Architectures

As explained, SRP/DPT can be transported across dark fiber, over SONET/SDH, and over DWDM. Multiple SRP rings over DWDM are shown in Figure 7-28. DPT can also be deployed in a star configuration. Here, the physical topology is star, but logically, it's a ring.

Figure 7-28. DPT/SRP over OC 48 DWDM Ring

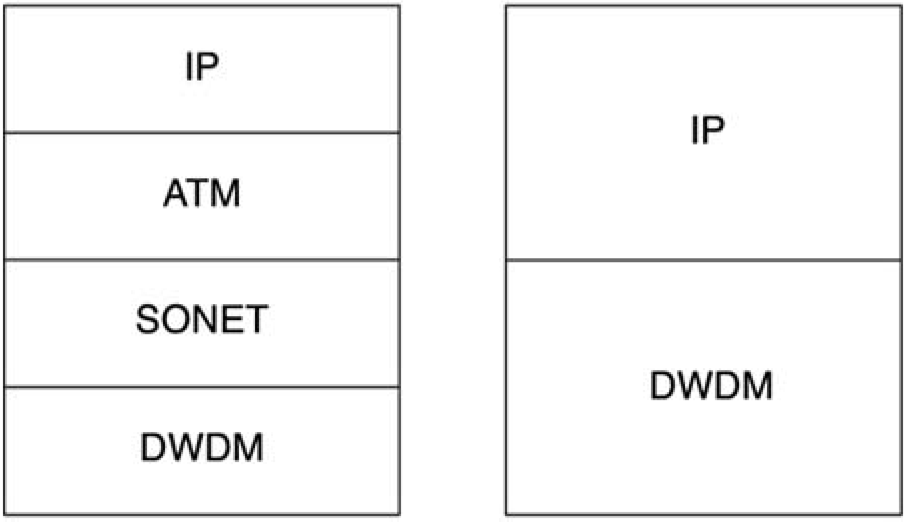


(Using DPT/SRP concentrators, such as Cisco 15194, can accomplish this.)

IP over DWDM

SONET/SDH networks are built for voice traffic and cannot handle the bursty nature of the data traffic. Currently, SONET is used to transport IP such that IP is carried over SONET and finally over DWDM. If we need QoS, another layer of ATM or MPLS is added between the IP and SONET layers, as shown in Figure 7-29. Typically, we need different management functions for each technology; therefore, we need four different management and provisioning software packages to support this overlay network (IPoverATMoverSONEToverDWDM).

Figure 7-29. Current Layout Protocol Stack and Future Vision



An alternative is to implement IP over DWDM, porting function-like QoS and traffic engineering from ATM/MPLS, SONET-like reliability, and protection (50 ms). This process eliminates the overhead of SONET and ATM, thereby reduces the total cost of transporting data. The main issue of implementing such a protocol is the lack of framing and detecting at the IP packet layer. IP packets on a pure optical layer seem to be just pure bits without any way to detect the start and end of frames, leaving other issues, such as correcting transport errors, flow control, etc. Optical framing technology is in its infancy, and its implementation seems to be a distant reality. In that scenario, there are not many ways to implement direct IP over the WDM layer, although such implementation would benefit the entire communication industry tremendously. There is a need for layer 2 framing for IP packets to transport them on the WDM layer.

One possible way of implementing IP over WDM is by using MPLS labels. Multi-protocol label switching is a good paradigm to implement fast or ultrafast protocols. Originally intended to relax processing times at routers, labels can be used to indicate start and end of coagulated IP packets, and such label-distributed packets can be made to move across the WDM layer. The IP over WDM paradigm also paves the way for new approaches to protection and restoration. In addition, optical wavelength routing can be considered to be a transport media for such raw packets on the WDM layer. Optical implementations of OSPF and BGP are being considered in IETF drafts [kompella RFC]. The vision of an Optical Internet is now quite realistic with such innovations in technology. All-optical label swapping and label distribution have also been experimentally demonstrated [Post deadline paper in OFC 2002], paving the way for easier implementation of IP over WDM. Control plane algorithms are being optimized to control chunks of packets or bursts from ingress to egress nodes over an optical link. Unified control plane and Optical control plane (UCP and OCP) are being considered as part of the OIF efforts in finding possible solutions to optical network control and management issues.

Summary

In this chapter, we considered various protocols over the WDM layer for possible implementations. We looked at SONET, Ethernet, IP, and RPR as some of the protocols that fit on the WDM layer. SONET seems to be the most common way to send data, but it comes at a high price. Ethernet, in contrast, is a simple implementation, and it is rapidly growing in popularity. Ethernet is widely considered to be the replacement of SONET in the near future. RPR is an efficient protocol for ring networks. Finally, IP directly over WDM seems to be the most promising, but it is quite difficult to widely implement at this time.

Regardless of which protocol is stacked on the WDM layer, the functionality always remains to try to achieve the best throughput at the lowest possible cost. It is difficult to compare the different techniques of transporting data over the WDM layer; each technique has its advantage and disadvantages. Keep in mind that except for 10 GE, all the other protocols have been developed prior to the arrival of WDM. This clearly means that none of the protocols are quite optimized for transport over WDM. What is needed is a good data link or higher protocol that can successfully map data to the optical layer, keeping in mind the finer points of the optical layer as well as the truancies played by the optical layer. In terms of costs, SONET is the most expensive and Ethernet is the least.

References

[1]Chow, Ming-Chwan. *Understanding SONET/SDH Standards and Applications*, First Edition. Andar Publisher, 1996.

[2]http://grouper.ieee.org/groups/802/3/ae/criteria.pdf.

[3]http://www.10gea.org/Tech-whitepapers.htm.

[4]Kramer, Glen, and Gerry Pesavento. "Ethernet Passive Optical Network (EPON): Building a Next-Generation Optical Access." *IEEE Communications Magazine*. Vol. 40, No 2. February 2002.

[5]*Internetworking Technologies Handbook*, Second Edition. Cisco Press Fundamentals series, 1998.

[6]www.cisco.com.

[7]Siller, Jr., A. Curtis, and Mansoor Shafi. *A Sourcebook of Synchronous Networking SONET/SDH*. IEEE Press, 1996.

[8]Rau, L., et al. OFC Post Deadline paper Anaheim CA, 2002.

Chapter 8. Future WDM Networks and Technologies

As networks gradually migrate from legacy SONET to WDM and from circuit- to packet-switched networks, the showcased technologies that underline these networks is changing. Newer technologies are far more efficient in transporting bits across dark fiber with minimum error and maximum flexibility.

In this chapter, we will discuss some of the new technologies, especially from the net-working perspective, that facilitate seamless communication across the fiber. So far, we have only dealt with optical circuit or lightpath-switched networks and have not focused on optical packet-switched networks. In this chapter, we shall to some extent deal with optical packet-like networks as well as technologies that further the communication content of light bits across a fiber. Until today, the role of the optical layer has been purely to transport bits across the fiber. Gradually, this role might change to a more proactive role of packet switching or wavelength routing (which is already being implemented). In this regard, it is important to understand that packet switching, however promising, is still viewed as a technology in its infancy. At line rates of 10 Gbps, a 1500 byte Ethernet frame takes only 1.2 microseconds to pass through. Therefore, switching times for individual frames need to be less, especially considering the fact that there is no optical equivalent of optical random access memory, which is essential in the making of buffers.

Buffering packets entirely in the optical domain can be accomplished to some extent by using fiber delay lines (FDLs). Fiber delay line is a concept whereby long lengths of fibers are used to keep the packet in a transit path and hence "time" is bought, which is proportional to the length of the delay line. The practical difficulty of such delay lines is the length of the delay line itself. To storing an Ethernet frame for 1 ms takes 200 km of fiber! It's also necessary to consider the loss and other optical impairments inside such a buffer. The FDL concept is not a viable solution even though it is attractive.

To implement packet switching, some kind of RAM is always desired for implementation. Another serious aspect that has prevented packet switching from taking place is the speed of optical switches. Optical switching technology—be it micro-electro mechanical systems (MEMS), acousto-optic tunable filters (AOTFs) technologies, or any of the conventional means as described in Chapter 2, "Networking with DWDM -1,"—is currently slow, and it is cumbersome to implement fast switching of various packet streams. Although this might not be considered a serious problem, the absence of optical RAM is also an issue of active research. There have been some demonstrations in OFC 2002 of ultra-fast switches based on Mach Zehneder Interferometer-based lithium niobate modulators and capillary switches, but both appear far from actual deployment. Another new area of active research is tunable lasers. In NFOEC 2002, microsecond tunable VCSELs have been demonstrated.

The telecommunication industry would benefit from the deployment of optical packet switching technology for various reasons. First, the transportation cost per bit would be reduced substantially, and most importantly, the opto-electronic mismatch causing a bottleneck would cease to exist. In the absence of true technology to support optical packet switching, some trade-off technologies are also being investigated. One such paradigm is *burst switching*. The origin or exact definition of burst switching is quite vague because it is a transformed technology from yesteryear's burst-switched networks. A burst is considered an aggregated stream of packets, and this burst is set to cut through a network element en route to a destination. Because it is an aggregation of many packets, the switching requirements for burst switching are considerably relaxed as compared to the switching requirements (time) of pure packet-switched networks.

Different algorithms have been proposed in literature to implement burst-switching heuristics. Optical burst switching (OBS) is now being muted as a possible deployable technology, especially with the advances in the signaling of the bursts—MPLS and its optical counterpart G-MPLS technologies.

Another interesting technology that is worth mentioning is photonic slot routing (PSR). From the pioneer of Lightpath Communication, PSR seems to be an easy way to implement packet-based networks at the optical layer. The idea is to have wavelength-independent switching of slots rather than wavelength-selective switching. Slot routing combined with burst switching might be yet another promising approach to implementing high-speed networks.

A serious issue in optical networks is contention resolution. Contention between two packets/bursts/lightpaths occurs due to the requirement of two or more packets for the same network resource—namely, the same output port in a router. Contention can lead to blocking or even dropping of packets. Several methods to avoid contention at the optical layer have been proposed. One of the most fundamental methods is called *contention resolution using delay* (CORD) lines. Contention resolution using delay lines is a project that the University of Boston initiated in association with Stanford University and GTE Laboratories in 1996.

The idea is to allow one of the contending packets to recirculate in an optical buffer while the resource it requested becomes free. Deflection routing is also an alternative mechanism to work around contention resolution. A recent attractive method to contention resolution is using loopless deflection. This scheme deploys a distributed routing approach whereby the network avoids cycles of the deflected packet by intelligently deflecting it.

Multiprotocol label switching and its optical variant, the generalized multiprotocol label switching, is gaining momentum in today's standard committee meetings. It appears to set the trend for future edge and core routers to have strong optical interfaces and deploy more functionality at the optical layer than just pure transport. These kinds of higher layer protocols also ensure better management of optical networks. Optical label swapping is one such scheme being demonstrated by the University of California at Santa Barbara.

The state of optical packet switching also depends on traffic shaping mechanisms today. Internet traffic can be described as being *bursty*. Burstiness, or *self-similarity*, is one attribute of traffic distribution that we need to seriously consider before using packet switching as a possible mechanism for optical routing. Legacy models have assumed Poisson-like distribution of traffic streams in the network; however, this is valid only for arrival patterns of voice streams. Data streams, in contrast, are more bursty in nature. Therefore, Pareto, or self-similar traffic, distributions are used to characterize data streams. This result was first spotted by Bell Labs in the early 90s.

Advances in 40 Gbps data communication are coming along in optical networks. OC-768 or 40 Gbps data on a single wavelength might soon be a reality, although this might take a year or two to actually happen. At high rates such as 40 G, the optical impairments are enormous. Dispersion and PMD limit a typical system to just a few kilometers; therefore, we have to use dispersion-preserving fiber. Nonlinearities at 40 G are a potent problem, and it is difficult to work around the effects of four-wave mixing (FWM) and cross-phase modulation (XPM).

Other areas with a great deal of research activity include polarization mode dispersion (PMD), PMD compensators, as well as tunable lasers. Low-cost dispersion compensators based on fiber Bragg grating (FBG) or photonic crystals are appearing in the market. High-speed tunable lasers, with phased arrayed laser technology, are also being seen as a possible candidate in alleviating some serious routing bottlenecks in the network.

Burst Switching

The paradigm of burst switching is not new to the routing and forwarding phenomenon in the Internet. Burst switching's application to optical networks was proposed effectively by Chunming Qiao from the University of SUNY Buffalo.

The idea behind burst switching is to aggregate large chunks of data at the edge of the network into bursts and then send these bursts through the all-optical core at high speeds. The burst goes through the core network at the high speed, all-optically without opto-electronic conversion and regeneration at the intermediate core routers. This means that the burst goes through the network cutting through the intermediate core nodes, with the core nodes quite oblivious to the data in the bursts as it goes through the nodes. The protocol that actually enables burst switching does so by assuming a control channel that carries control information for every burst.[1]

Obviously, the control information precedes the burst so that the intermediate node's switches are configured to pass the burst through. The main idea behind burst switching is to keep the data in an optical domain throughout the physical layer and avoid expensive opto-electronic conversion and regeneration.

Burst-Switched Network Algorithms

The following are some important burst-switched network algorithms: Just Enough Time (JET), Tell and Go (TAG), and Reserve for a Fixed Duration (RFD).

The main idea behind burst switching is to ensure seamless transport of data packets through an optical core in the most efficient manner possible. As a technique, burst switching can be considered analogous to packet switching except for the fact that the switching occurs only in the ingress and egress nodes while there is no switching at the intermediate nodes. Burst switching can also be viewed as a short lightpath-established network, where bursts can be considered as lightpaths for a short duration. The main problem in packet switching at the optical layer is the inability to process information—namely, the packet header in the optical domain. There is no feasible solution to this problem at this time. Optical gates, SOAs, and so on have all been proposed as alternatives, but none is worth serious mention due to feasibility considerations.

In burst switching, we can view the system comprising two different types of channels: a high bit-rate data channel and a low bit-rate control channel. Bursts are aggregated at the edges of a network, and a burst that is destined for a particular destination node is transmitted as explained in the following paragraph.

Upon aggregation of data and burst formation, a control packet is created at the ingress node. This control packet is transmitted to the egress node along the best-chosen route. The control packet contains information that pertains to the burst heuristics. The control packet is, as mentioned before, broadcast on a different channel (wavelength) as compared to the desired wavelength on which the burst travels. The control packet reserves resources for the lagging in time burst. These resources could be the bandwidth (or wavelength) along the path, the configuration of switches at intermediate nodes, or the resolution of contention at the intermediate nodes. After the resources are reserved, the burst is sent through the network. The burst, being a large packet, darts through the network, cutting through the intermediate nodes, which do not buffer or process this burst. The burst is then fully demultiplexed at the egress node by detecting and converting it back into an electrical domain. Three important attributes that deserve some attention are as follows:

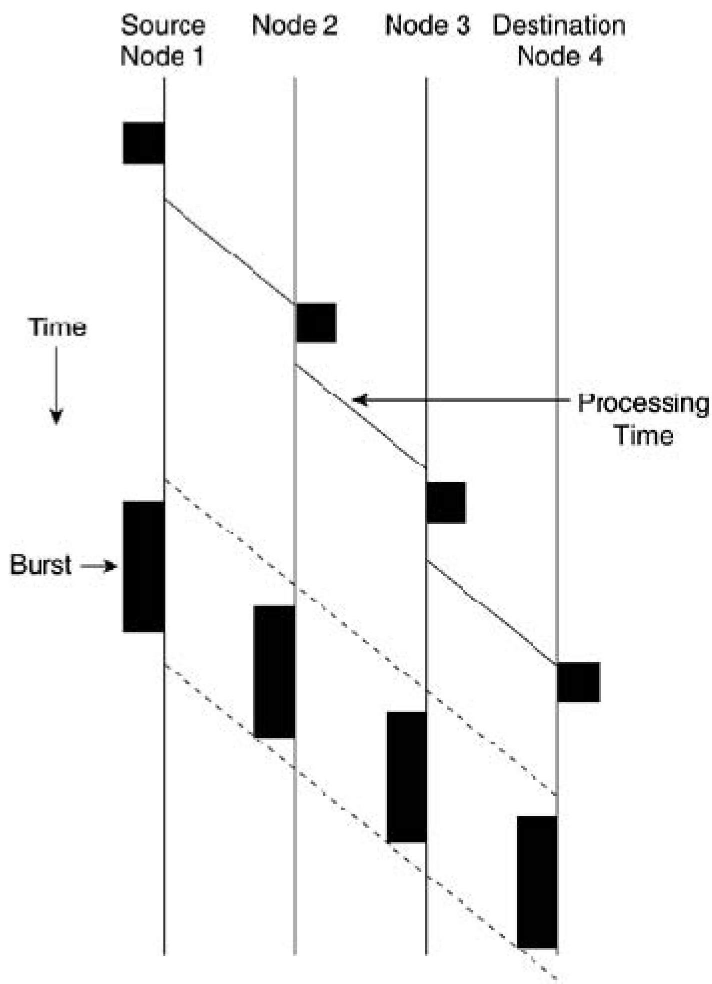
- No processing occurs at the intermediate nodes; the burst just cuts through the nodes.
- The control packet reserves the resources for the burst ahead in time. This kind of delayed reservation facilitates multiplexing of several bursts.
- There is a delay between the transmission of the control packet and the transmission of the burst. The efficiency of the system depends on this delay; a lower bound can be formulated to express this delay.

In context of optical burst switching, some algorithms were proposed to facilitate burst-switched networks in the optical domain. One of the first protocols that discussed burst switching was the Tell and Go protocol or simply TAG. A burst header that contained the burst heuristics, such as burst size, burst rate, and destination, was transmitted on the control channel. After a brief delay, the main burst was transmitted on the data channel. Note that no acknowledgement took place after the control packet's success in reserving the resources along the path. This often led to bursts being dropped at intermediate nodes due to contention. The TAG protocol was simple and easy to implement.

Chunming Qiao and others proposed the now famous JET protocol for optical burst switched networks. JET or Just Enough Time protocol was an optimized version of the original TAG protocol. JET has two interesting features: offset time and delayed reservation. Offset time is the time difference between sending the control packet and actually sending the burst. In other words, it is the mean time required for the control packet to be sent en route to the destination and for the resources to be allocated correctly. Although, it should be noted that this is static because there is no acknowledgment that proclaims successful reservation of resources.

Delayed reservation, on the other hand, signifies the time from where resources (namely, bandwidth and switch configuration) are set to the time that they are de-allocated. Delayed reservation initiates resource allocation by optimizing the time the burst is sent to the time the burst passes through with highest efficiency (see Figure 8-1). Delayed reservation is an efficient method that eliminates the need for optical monitoring of the channel to see when the burst passes through. Delayed reservation also increases the likelihood of good statistical multiplexing of bursts, given the fact that the network knows about the burst heuristics well in advance.

Figure 8-1. The Working of the JET Algorithm[1]



Prior to implementation of JET and delayed reservation, the scheme proposed was in-band termination (IBT), whereby an in-band signal indicated the end of a burst and therefore de-allocated resources. This is not an efficient method because the technology to decimate the details of a burst at the optical layer is still in its infancy. By sending a packet (control) ahead in time with the burst size information, the intermediate nodes know exactly when to reconfigure their switches and reallocate the bandwidth after the burst is through.

Working of the JET Algorithm

Consider the Just Enough Time algorithm, which is depicted in Figure 8-1. We consider a three-hop path from source to destination. Upon the arrival of a burst at the source node (node 1), a control packet is sent to the destination node (node 4) en route to nodes 2 and 3. If the propagation delay for every hop is, say, d, and the processing time at each node (for the control packet) is p, then the control packet reaches the destination node after time interval h(p+d), where h is the number of hops from source to destination (in this case, 3). We are assuming equal propagation delay for every hop and equal time for processing at each node.

If we send the burst after interval T following the transmission of the control packet, we can see fromFigure 8-1 that the burst seamlessly propagates through the network without buffering or processing at the intermediate nodes. It is obvious that the offset time is equal to the product of the number of hops and processing time at any given node, which can be mathematically represented by Equation 8-1.

Equation 8-1

$$T_{offset} = p\Sigma_{i=1;H_{max}} h$$

We can greatly increase the efficiency of the OBS network by utilizing the concept of delayed reservation. Unlike the norm in TAG algorithm, the bandwidth in JET is reserved from the time the burst actually arrives and not from the time the control packet arrives. As we can see, this increases the efficiency as a result of increasing the probability of statistical multiplexing of bursts along the route. For a simple analysis, this benefit can be quantitatively summarized as the ratio of time needed for bandwidth allocation in a TAG network and the time needed for bandwidth allocation in a JET network. Various research groups are currently investigating burst switching as a possibility for implementation in today's optical networks. The notion that burst switching might be needed only in core networks is also fast being replaced by the need of burst switching in metropolitan networks because of the severe nature of bursty data-centric traffic.

QoS issues as well as burst deflection policies are being investigated. An algorithm for deflection routing is also proposed, but its efficiency is questionable. We can safely assume that OBS networks would help migrate today's SONET-based optical networks to IP over WDM types of networks. OBS paradigm can be considered as opening the doors of photonic packet-switching technology for the future.

MPLS and Burst Switching

The multiprotocol label switching approach was ceremoniously touted by several research groups for handling of bursty data traffic at efficient routers. The approach was simple in the sense that data packets were coagulated into larger packets and labels were given to such packets, which were to be routed to a common destination. These labels often carried information about the router input and output ports, in addition to delay (QoS) requirements, and so on. This did facilitate core routers to route these clusters of packets by just looking at the labels rather than looking at the header of each packet and decoding the address field. The same principle can be extended to optical networks, too. Labels can be added to optical bursts, and these labels could carry information such as wavelength number, route information, WDM port information, and so on. This kind of IP over WDM using optical label switching is often called label optical burst switching or LOBS and is under serious study.

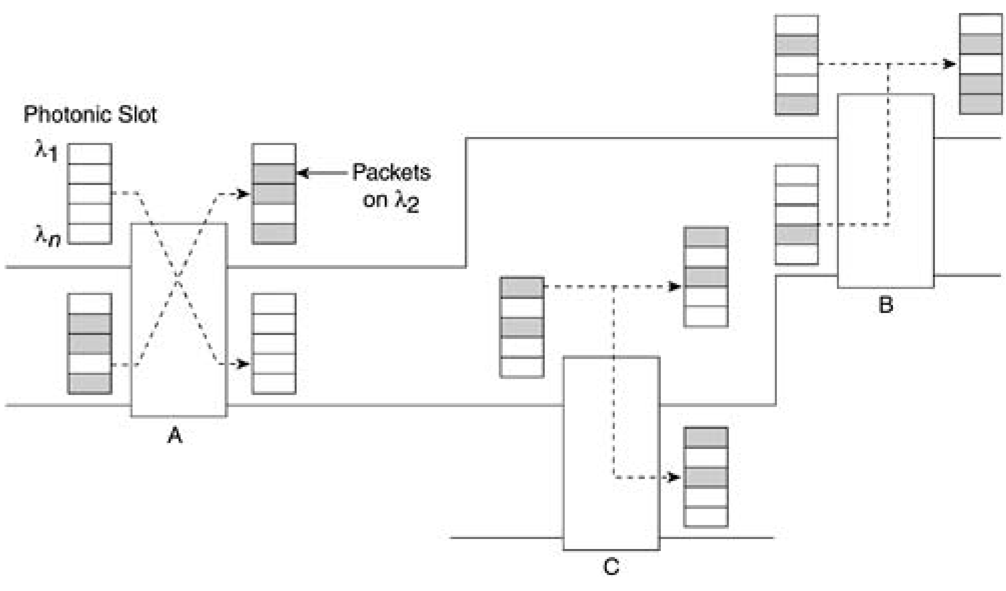
Photonic Slot Routing

A new technique, which tends to bridge the gap between packet switching and lightpath circuit switching, is *photonic slot routing* (PSR)[2]. Optical packet switching is difficult to implement in today's networks because current technology is not mature enough. A typical WDM network element demultiplexes the entire composite signal into individual wavelengths. These wavelengths carry packets in the optical domain. Switching of packets in the optical domain would mean either switching between different ports of a space switch or adding or dropping packets at the network element interface. In both cases, the need of the hour is optical signal processing of the packet header to determine what is to be done with the incumbent packet. Also, if a packet needs to be switched, the concerned switch would be required to be configured in the configuration-state as desired. This would take some finite amount of time.

In electrical networks this delaying is provided by electrical buffers or RAMs whereas in optical networks, no such technology is available. In addition, it is impossible to buffer up packets in the optical domain. Moreover, if we consider the volume of packets that would arrive on a 10 Gbps line, and note that there are several tens of wavelengths or lines, the number of times a nonblocking space switch would have to change its state in a second would be quite high. Therefore, it would be a requirement that the non-blocking switch has a fast switching time, preferably in the picosecond or lower nanosecond range. Today's technology does not showcase such high switching speeds. Even acousto-optic and other technologies do not allow speeds above the microsecond range.

To circumvent some of the problem, photonic slot routing (PSR) was introduced in "Scalable WDM Network Architecture Based on Photonic Slot Routing and Switched Delay Lines."[2] PSR is a novel concept whereby the entire optical operating band (it might be the C band, for instance) is divided into time division slots, and each slot has the entire complement of wavelengths. Consider the photonic slot shown in Figure 8-2. It consists of streams of packets at different wavelengths, with individual subslots at each wavelength. The slots are of a fixed length and can accommodate packets.

Figure 8-2. Implementation of Photonic Slot Routing (PSR)[2]



Typically, we would envision fixed-length packets, but with bit stuffing, variable-length packets can be used as well. A node, which gets an empty slot, puts its packets on the subslots at different wavelengths and sends the photonic slot through the network. An intermediate node can use some of the remaining vacant subslots. The key here is that the photonic slot terminates at some given node. Therefore, PSR can be termed *destination-based slot assembly*. The main advantage in PSR is the complete elimination of wavelength selective switches. Slots can be switched all-optically by a wavelength-independent space switch. Furthermore, this space switch does not need to be fast; its speed depends on the length of the photonic slot. For seamless photonic slot routing, we also need a control channel that informs other nodes about the status of the slot (which subslots are empty) and the destination of the slot. In one embodiment, a node that receives a slot is then eligible to send a slot. The only issue in such a network is that fairness degrades: For a burst traffic matrix, some nodes might hog all the slots, leaving other nodes empty handed. Some key features of PSR are as follows:

- Subslot merging— Slots that arrive at the same time on a number of input ports can be switched to the same output port, thus overlapping each other to form a single slot. This can happen only when the subslots are compatible. In other words, no two subslots on the same wavelength can be filled with packets. (Only one subslot can have a packet, and the other subslot has to be empty.) This allows packets from different input ports to be forwarded together to a common destination.
- Slot Copying— A slot that arrives on a particular input port can be duplicated and forwarded to various output ports. This procedure can be used to form multicast trees. This procedure also allows packets to be transmitted together even if they are not destined for the same destination.

In the presence of non-uniform traffic, slot merging and slot copying improves utilization and throughput. Due to its simple architecture, the PSR node can be built on proven technologies. Moreover, the PSR architecture can be implemented using proven mature optical technology. Its functional simplicity lies in the fact that there is no wavelength selective device—not even an AWG in the network line. The network is scalable in the sense that new wavelengths can be added without disturbing the basic optical line. In a conventional network to add new wavelengths, the entire AWG (multiplex section) would have to be replaced. This does not need to be done in PSR, facilitating in-service upgrade capabilities.

When a photonic slot travels through a network over a length, the individual subslots at various wavelengths undergo different amounts of dispersion. This is primarily because pulses at each wavelength have different group velocities, which causes different amounts of dispersion for each subslot. Therefore, there can be dilation (dispersion) of the original pho-tonic slot. To prevent dilation, we can stuff the subslots with extra bits. This takes care of the dispersion issue in PSR. In reference 2, PSR was implemented (simulated) as interconnected rings. The interconnection points had electronic architectures called *bridges* to align and feed slots[2]. This proved to be a scalable and resilient architecture for practical implementation.

Contention Resolution Using Delay Lines

In an all-optical packet networks, the fundamental problems are contention of resources, signaling, and synchronization of all-optical packets (time slots). Contention occurs when multiple optical packets that arrive on different fibers or wavelengths try to access one or more of the optical network resources like transponders, switches, or any particular wavelength or that matter ports. It is imperative that the solution for this problem, in order to be all-optical, needs to be in the optical domain. (No optical-to-electronic-to-optical conversion should be allowed.)

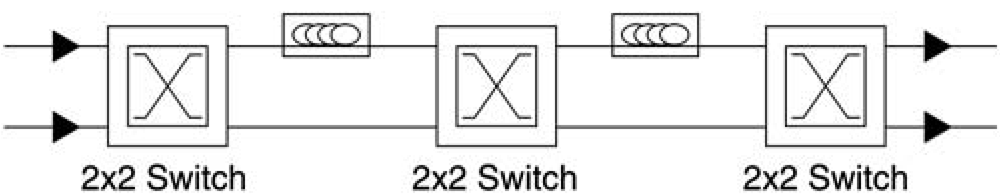
One possible way to surmount this issue in a packetized optical network is to use the concept[3] of switched optical delay lines (SDLs). A switched optical delay line unit consists of an optical fabric based on a combination of optical switches (switching matrix) and fiber delay lines.[3] The packets that are contenting for the same resource are redistributed using the SDL over time and space. Rescheduling packets can also be implemented to give overall fairness.

The next major issue in all-optical packet-switched networks is signaling. An optical router should be able to retrieve header information from multiple optical packets (WDM–multiplexed) simultaneously[3]. Subcarrier multiplexing of packet headers has been proposed to deal with this issue, but several implementation issues need to be worked out before full-scale deployments are realized. Synchronization of both packet time slot and clock is a critical aspect of all-optical packet networks. To explore the practical feasibility and effectiveness of the contention resolution using delay lines, the University of Massachusetts, Stanford University, and GTE Laboratories formed a consortium[3] to carry out research in this area.

Optical Contention Resolution: Principles

The principle of optical contention resolution is based on the use of switched optical delay lines (SDLs), which shift contenting packets in space and time, thereby resolving contention at optical resources. This resource includes receiver ports, fibers, temporary storage for optical packets, and channels (WDM). The basic principle behind SDL is as shown in Figure 8-3.

Figure 8-3. Block Diagram of Two-Stage SDL Architecture[3]



Consider packets arriving at a node on multiple channels. These packets are statistically distributed and can be destined for this node or switched to other nodes. The contention occurs when arriving packets compete for the same resource. Contention types vary with topologies and network type. For example, in a multihop topology, packets must deal with receiver contention while passthrough packets also need to contend for out-bound links on the intermediate nodes.

In a star topology, the main contention is at the receiver end. The SDL approach, which is used to reduce this contention, consists of optical 2 x 2 switches (reconfigurable) and delay lines for optically storing the contenting packets. The 2 x 2 switch is reconfigurable on a packet-by-packet basis. The switching operations are controlled by local nodes, whereas signaling is achieved through in-band or out-of-band methods, depending on the implementation. Figure 8-3 shows a two-stage SDL structure.

40 Gbps Systems

Currently, the most talked about cutting-edge technology under development is 40 Gbps. Even though the cost savings of 40 Gbps systems is not clear, it is anticipated that as the bandwidth demand increases, this will become a viable alternative. Many technical challenges need to be surmounted before 40 Gbps becomes feasible and economical. The systems that are currently being developed are based on SONET/SDH (OC-768/STM-256) at bit rates around 39.8 Gbps. Several impairments affect 40 Gbps or equivalent bit rate transmission. One possible way to alleviate this problem is by using forward error correction (FEC) given by the G.709 standard.

The 40 Gbps systems with Forward error corrections (FEC) help the signal to travel longer distances; therefore, they are used in long-haul transport. FEC adds extra bits to the signal, and the bit rate is increased to 43 Gbps. The constraints that are imposed by the insertion loss and polarization-dependent loss (PDL) are much higher (specification) than those for 10 Gbps systems. Chromatic dispersion and polarization mode dispersion (PMD) need to be compensated at and above 10 Gbps speeds. PMD compensations are expensive and challenging in 40 Gbps transmission systems. (Impairment changes due to PMD over a short distance are unpredictable.) The main issues in 40 Gbps transmission is optical impairments. The exponential rise of these impairments beyond 10 Gbps speeds creates massive problems for 40Gbps implementation. The most persistent problem being dispersion of the optical signal. Both chromatic (GVD) and polarization mode dispersion are severe at such speeds. XPM and FWM too are quite limiting factors for such speeds.

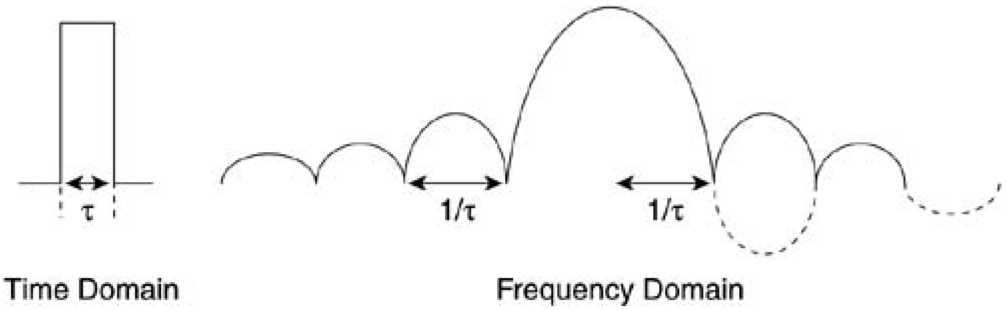
For dispersion, we have to place multiple compensating sites, in the most optimized way (correct placement). Another method to alleviate some of the issues faced by 40 Gbps communication, is by choosing the right fiber. A fiber whose zero dispersion wavelength is shifted in the operating band, creates less chromatic dispersion of the optical signal. For polarization mode dispersion, we have to place the compensators in the right position. Strict power budgeting takes care to some degree of nonlinear effects.)

Return to Zero in 40 G Systems

NRZ modulation and RZ modulations are explained in Chapter 2, "Networking with DWDM -1." To understand NRZ and other modulation formats, we need to study the frequency response of signals that can be well represented by Fourier analysis.

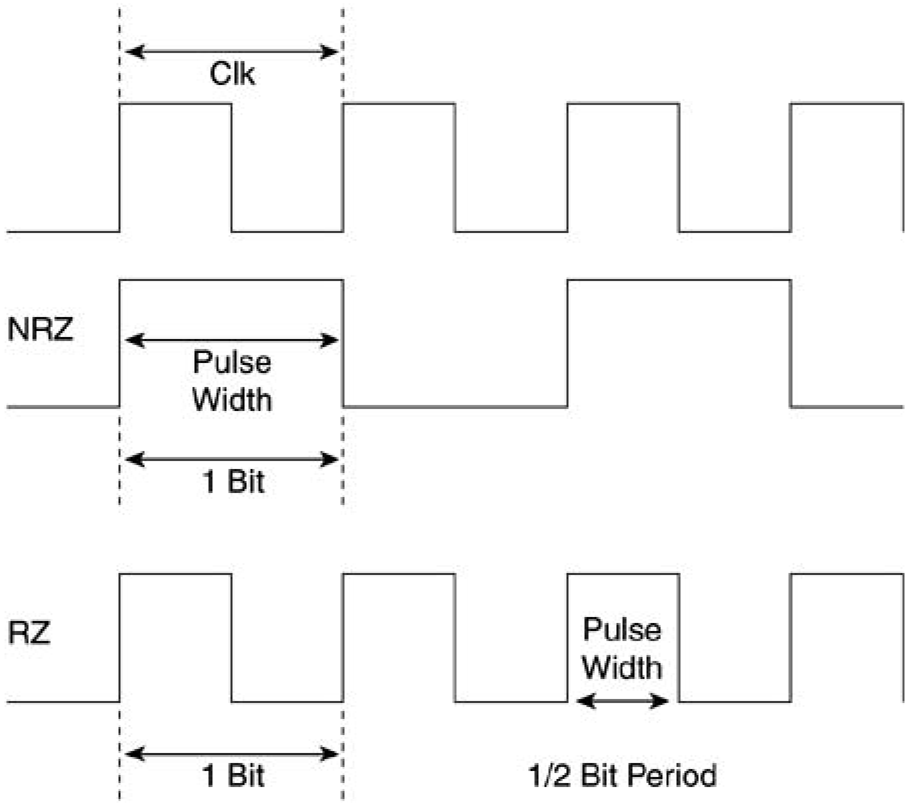
A sine wave in time is a single line on the frequency axis. The frequency representation of a signal is the Fourier transform of the time domain representation. A spectrum is defined by the SIN (X)/X function. A rectangular pulse in the time domain transforms to a sine (X)/X in the frequency domain. The frequency nulls occur at 1/pulse-width frequency (τ) and are shown in Figure 8-4.

Figure 8-4. Wave Representation in Time and Frequency Domain



For the same data rate Return to Zero (RZ) has half the pulse width of that of NRZ (see Figure 8-5).

Figure 8-5. NRZ and RZ



The RZ signal is half the pulse width of NRZ, which means it takes twice the bandwidth and twice the switching time than that required for NRZ. This has a significant effect on component designs, which leads to longer development cycles and expensive product. Then why would we use RZ format? When designing the overall system, a system architect must deal with other significant challenges that can be a barrier for deployment. Even though it is costly and more difficult to build RZ-based components, they do help to overcome other challenges, such as chromatic dispersion (CD) and polarization mode dispersion (PMD).

For a 40 Gbps signal, the bit period is only 25 picoseconds in duration (NRZ). CD- and PMD-based design are challenging at 10 Gbps and even harder at 40 Gbps. With the RZ format, dispersion is less likely to cause an RZ pulse to interfere with subsequent pulses. In addition, clock recovery is simplified due to the strong spectral content at the given line rate. *Soliton* is a type of RZ signal that takes advantage of fiber nonlinearity to counteract the effects of dispersion. It can be considered a shape-preserving pulse, whose shape is retained despite going through a fiber line.

Resource Reservation Protocol and Traffic Engineering in the Optical Layer

As we know, IP traffic is essentially best effort; packets (IP) do not always arrive in a systematic way, and the packets are not guaranteed to provide Quality of Service (QoS) requirements. As data-centric networks grew in size, it was realized that future revenue would not only come from voice-based and video-based circuits but also from data packets. To account for the huge surge in Internet traffic, data networks needed to provide QoS issues. To provide QoS in data networks it was essential to provide some service parameters in the network.

The most important threshold that dictates network performance and revenue is a network's QoS. QoS can be of various forms from delay-based QoS where the end-to-end delay of a packet is kept below a particular threshold, to synchronized QoS where the packets arrive in the same order in which they were sent. The former is the most common QoS scheme. Synchronized QoS is particularly useful in Voice over IP (VOIP), in which hearing a word out of order makes it nonsensical. Imagine if the speaker at one end says "Hello there" and the speaker at the other end hears "There hello." QoS plays an important role in the network.

Broadcasting or multicasting is also becoming an important function in today's Internet, with video broadcasting assuming high proportions. We need to consider two main issues when providing seamless quality to optical networks. First, the QoS of the traversing packets must be maintained. Second, some algorithm that shapes the traffic needs to be inculcated. The second point is quite important. Internet traffic is bursty, meaning that it comes in bursts or spurts. This means that the network requires uneven bandwidth at different times. As a result, utilization suffers. If a particular source were an IP source (is bursty), it would emit large bursts of data followed by intervals in which there is no data. To ensure QoS, we might accord bandwidth to the source corresponding to the maximum input into the network. Although no blocking or packet loss would result from such a scheme, for a long time the line would not have data or would operate at a less than normal data rate, which translates to underutilization.

In contrast, if we keep the line rate at some lesser (mean) value of the maximum input rate, it would cause excessive buffering at the input to the network. This would mean large buffers and large delays, and QoS would suffer. Therefore, we need a scheme that intelligently allocates resources (wavelengths, as in the optical domain) for different times corresponding to network heuristics. This is called *traffic engineering* in the optical layer. Traffic engineering has a strong impact on optical networks; it dictates network line rates and other optical parameters. One algorithm that implements traffic engineering is the resource reservation protocol, or RSVP.

RSVP is a backward reservation protocol in the sense that the path set up is from the destination to the source node, although the request is made in the forward direction, that is from source to destination. A source node that needs to reserve some bandwidth to dedicate a flow to a destination node, and accord some QoS does so by using RSVP. The node sends an RSVP reservation packet to the destination node (RSVP request packet), following which the destination node and the intermediate nodes (along the desired route) comply with an RSVP acknowledgement packet, acknowledging the possible reservation path[4].

Optical Cross-Connect Technology

As networks grow in size and capacity, it becomes increasingly difficult to keep converting data back and forth from the electrical to the optical domain. Therefore, all-optical networks (AON) have been envisioned as the future of today's Internet. To facilitate an AON, individual nodal architectures need to deploy all-optical switches or all-optical cross-connects. Cross-connects are cumbersome switches that can switch signals between different ports.

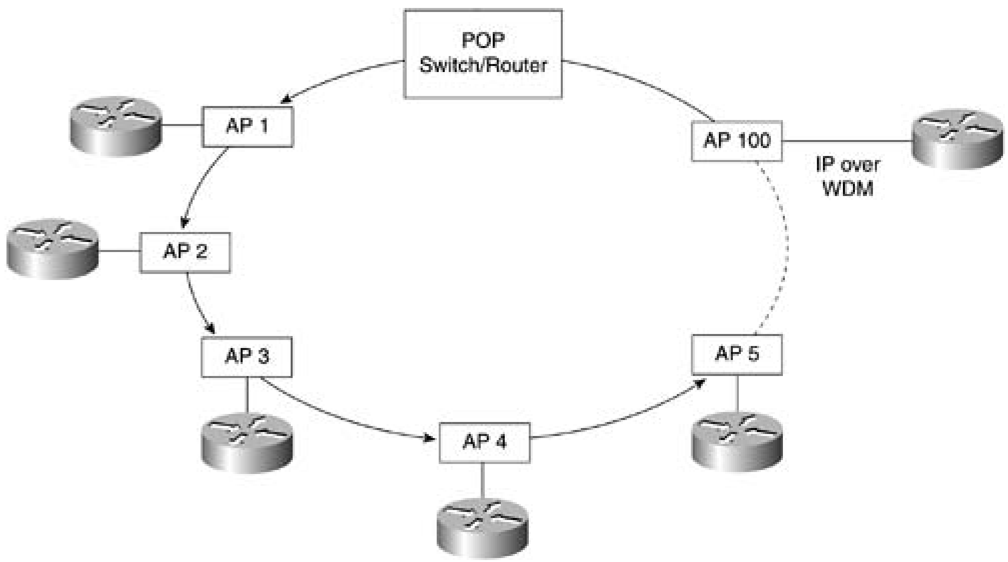
Nonblocking cross-connect architectures suffer from issues such as high loss and slow speeds, and they are limited by wavelength conversion requirements. As mentioned previously, wavelength conversion entirely in the optical domain is a fancy technology that might take decades to implement. Therefore, we have to work around some ways to implement such architectures.

One possible workaround is using nonblocking three-dimensional micro-electro-mechanical system (MEMS) technology. MEMS are a fundamental invention that caters to all-optical switching and routing. MEMS switches can easily be deployed in producing cross-connect architecture; this configuration is gaining popularity. In one embodiment, cross-connect architecture was developed using a full demultiplex section for each of the wavelengths arriving from each fiber, and the cross-connect was a MEMS-based cross-connect. Recently, cross-connects up to 1048 input ports and 1048 output ports were demonstrated. As the need for mesh optical networking grows, cross-connect requirement will also increase. The idea of a lightpath cutting right through the node without serious O-E-O issues is certainly fascinating to the designer[5].

HORNET: (Hybrid Opto-Electronic Ring Network)

Hybrid Opto-Electronic Ring Network (HORNET)[6] is a joint effort between the optical communication research laboratory at Stanford University and Sprint Advance Technology Laboratories. HORNET addresses the inefficiencies of SONET/SDH networks by transporting IP packets directly over WDM networks. HORNET nodes are called *access points*[6] and can drop a fixed WDM channel. The access points use fast tunable transmitters to transmit packets on any wavelength. The access points also implement a novel MAC protocol that governs the network controls and wavelength tuning. The HORNET architecture is shown in Figure 8-6.

Figure 8-6. HORNET Architecture[6]

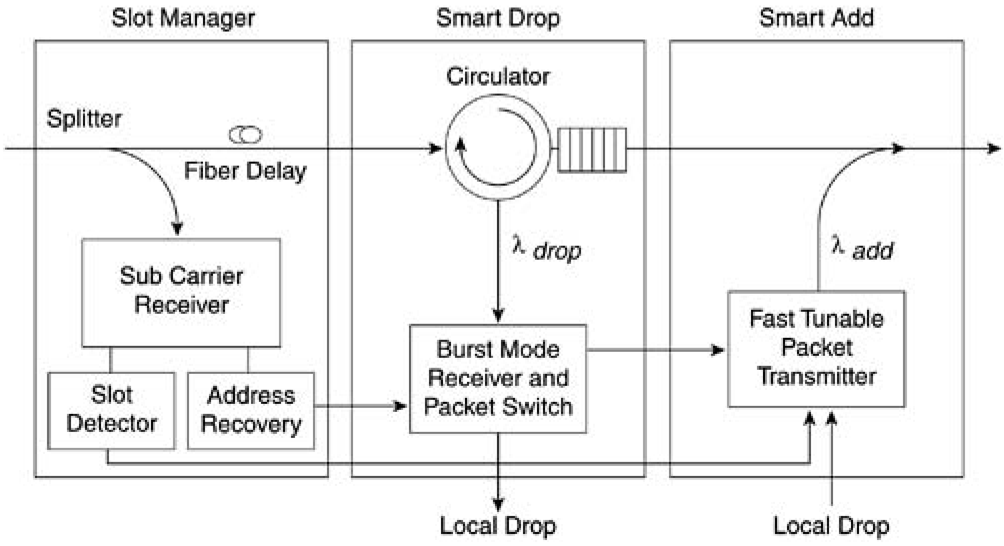


The network is designed to scale to 100 access points and ring circumference of around 100 km with a bit rate of 2.5 Gbps. A subcarrier multiplexing technique is used to carry control information, which includes access point addressing and wavelengths. Multihopping is supported with this architecture, which also helps the network scale efficiently. Access points can use any wavelength, and wavelength bandwidth is shared statistically among various access points.

HORNET uses a carrier sense multiple access with collision avoidance (CSMA/CD) protocol to govern the access of the wavelength and avoid collision. Each network wavelength is associated with a subcarrier; when an access point transmits, a corresponding subcarrier is multiplexed. By monitoring the subcarrier, the access points determine the occupancy of all wavelengths.

The access points act as an interface with an IP router or ATM switch and consist of three subsections: slot manager, smart drop, and smart add[6] (see Figure 8-7). The slot manager taps (with a 10/90 coupler) a portion of the signal for "carrier sensing" and to determine which wavelength is present. The slot manager also performs the demodulation (FSK) of the subcarrier. Address recovery and drop information are then extracted from the subcarrier.

Figure 8-7. Schematic of a HORNET Access Point[6]



The smart drop module consists of a burst mode receiver that recovers the packet bit clock. The module then uses the address information that the slot manager provides to switch packets. The smart add services are responsible for transmit and retransmit queue(s). The transmission wavelength is determined by the wavelength availability information from the slot manager (collision avoidance) and also depends upon the destination access point. The smart add then tunes the fast tunable transmitter to the determined target wavelength and modulates the packet on to the optical carrier[6]. Refer to Figure 8-7.

The CSMA CA MAC protocol is designed to handle variable IP packet sizes, and the burst mode receiver is capable of recovering packet clock within a few nanoseconds. The access points do not maintain synchronization; they use an embedded clock transport technique in which the transmitter frequency multiplexes its bit clock with the base band data and modulates the optical carrier with negligible dispersion[6]. The fast tunable transmitter needs to achieve nanosecond tuning. HORNET access points use a commercially available grating coupler sampled reflector (GCSR) tunable laser. Novel digitally controlled drivers control the tuning aspect.

Burst-Mode Receivers

Conventional receivers are not suitable for burst-mode operations in all-optical multiaccess networks because they cannot instantaneously handle the phase and power variations in the arriving packets. Due to fiber attenuation and dispersion caused by variation of transmitter wavelengths, power and phase variations between two bursty optical data packets can be up to 20 dB and 360^0 out of phase[7]. Burst-mode receivers are designed to adapt to the amplitude and phase variations on a packet-by-packet basis (ultra fast changes).

Two fundamental issues that degrade the performance of BER in a burst-mode receiver are random Gaussian noise and the limitation in the finite charging/discharging time of the detection circuitry in the receivers. Random Gaussian noise, which is always present at the receiver, affects the accurate determination of the decision threshold and introduces a sensitivity penalty. To detect ultra fast variable size packets the receiver needs to have a fast detection circuit.

We can use burst-mode receivers in a variety of applications such as all-optical multiaccess networks, passive optical networks (PONs), and in a wide range of network topologies. The requirements for these applications are a large dynamic range, fast response time, and robust architecture to improve network reliability.

Burst-Mode Receiver Operations

In a conventional receiver, AC coupling is used. The coupled circuitry can provide high receiver sensitivity, but the charging and discharging times that are associated with the capacitor and the receiver's AC-coupled signal path do not allow the average amplitude to vary rapidly with time. Burst-mode receivers use DC-coupling circuits. The DC circuits adapt well to the varying amplitude of received signals in a short time interval. The clock and phase recovery of this type of receiver needs to be performed within a fraction of packet transmission time for proper operations of receiver circuits. Now if we consider a Ethernet packet of 1500 bytes long, at 1 Gbps speed, the time it needs for passing through is in the picosecond range and hence the requirements of such receivers are very stringent.

Burst-switched receivers are one of two types: feed back and feed forward. Detailed discussion of these is beyond the scope of this book. For more information on how burst-switch receivers operate, please refer to reference 7.

Vision of an Optical Internet

The future of the Internet can be visualized as a network of networks that is built on various cores, edges, and individual networks composed of optical and wireless technologies. The skeleton of the Internet—the backbone—would in all possibility be composed of a high bandwidth multiwavelength optical core. An all-optical Internet core would facilitate fast and efficient switching of lightpaths and dynamic adding and dropping of lightpaths to ensure seamless communication between end users.

The optical Internet should have efficient protocols for data transfer. Gigabit and 10 Gigabit Ethernet (GE) are two important protocols that we can consider for direct IP over WDM networks. In addition, a future version of the optical Internet would in all possibility have a strong control and management scheme. Generalized Multi-protocol Label Switching would be an efficient protocol for controlling and signaling the routing and other features in a future all-optical network. GMPLS to a great extent would mean an overlaying control plane that facilitates communication between network elements. Setting up and tearing down lightpaths dynamically would be its primary function. This can be extended to packet switching, whereby optical packets can be switched and routed by some kind of out-of-band signaling. Although burst switching is a promising paradigm, no architectures currently supports the same.

Self-Similarity in Internet Traffic and Its Effect on Optical Networks

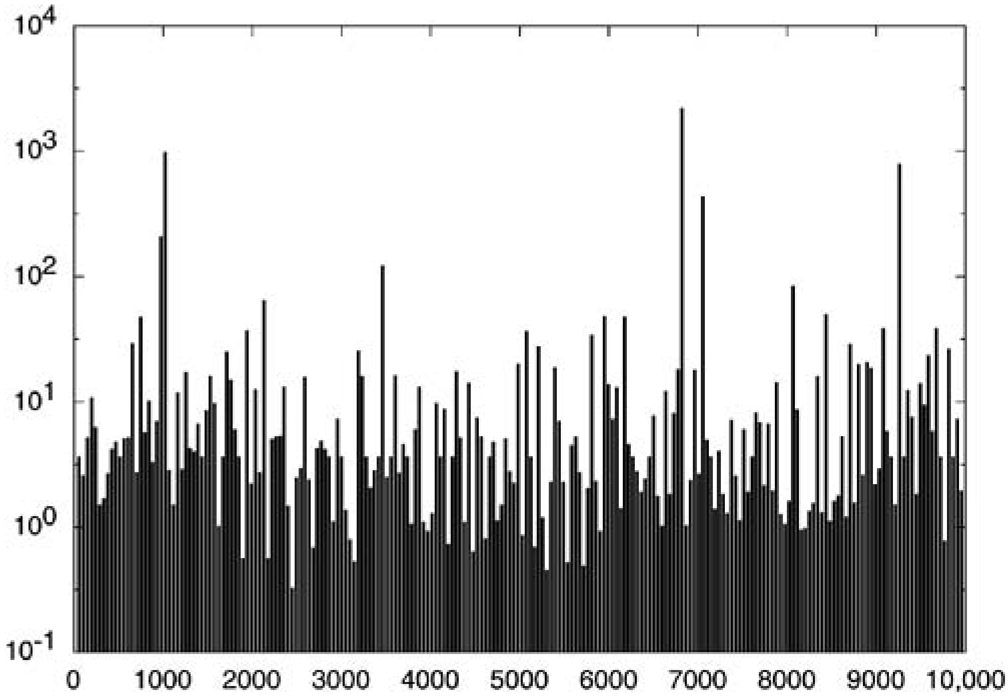
Conventional network modelers assumed that data traffic was Poisson or exponential in nature. This meant that the inter-arrival times of packets were exponentially distributed with some parameter called λ. This assumption was quite valid for voice traffic and could be intuitively understood. On the other hand, as networks migrated from voice to data, it was observed that data traffic was not necessarily Poisson. Data traffic could best be described as bursty traffic with long intervals of ON and OFF periods for a given data source.

In 1992, Bell core scientists observed that this traffic was Self-Similar, or long-range variant. This meant that the traffic streams correlated with themselves over some period of time. A data sample typically consisted of many small files, a few large files, and even fewer extremely large files. This kind of distribution can be termed *bursty traffic distribution*.

Self-Similarity, in contrast, is a natural phenomenon that occurs in many social systems. Self-Similarity can be characterized by the long-range dependence of some abstract portion of a network statistics with the same abstract portion of the network statistics magnified over a period of time. In other words, if we were to consider the arrival process of packets over a small interval of time ($t_1 – t_2$) and further consider the arrival process of packets from the same source over a larger interval of time ($t_3 – t_4$,), we would observe some sort of arrival symmetry between the two processes as long as one of the intervals is a dilated version of the other interval and the distribution is Self-Similar. This sort of symmetry can be termed as Self-Similarity and is a great cause of worry in modeling today's Internet traffic. (Refer to Figure 8-8). Following are some of the properties exhibited by Self-Similarity:[8,9,10]

- Scale invariant burstiness, or Self-Similarity, is a ubiquitous phenomenon in IP networks.
- Self-Similarity has been shown at multiplexing points in a network. Self-similarity is a robust phenomenon that is quite oblivious of the network topology or the control actions that the protocol stack carries out.

Figure 8-8. Self-Similar Process and Long-Range Dependence. Note that the vast majority of samples are small, but a few exist that are large.



Consider a discrete time stochastic process of time X(t), where X(t) represents the network traffic at time t from initialization. Furthermore, X(t) would be termed stationary if it were time invariant. Measurements at different values of t would not cause changes in X(t). X(t) would be called strictly stationary if { $X(t_1)$, $X(t_2)$,......, $X(t_n)$}, and {$X_{(}t_1+k_{)}$, $X_{(}t_2+k_{)}$,.........$X_{(}t_n+k_{)}$} possessed the same distribution.

Denoting this k–shifted process X_k, we can conclude that X and X_k are finite dimensional distributions, such that $X =_d X_k$. Although this kind of stationarity (stationary phenomenon) is good, it is still too stringent for modeling network traffic. We can now consider a weaker stationarity [W] called second order stationarity, which requires that the autocovariance function (relation of a function with itself over some time period) satisfies a translational invariance. In other words, we are trying to imbibe a time-invariant process that gives us various aggregated versions of the basic network process. Each version is a magnified version of the basic process. This kind of scaling in variance can be formulated as the aggregate process $X^{(m)}$ of the basic process X at aggregation level m. This is shown in Equation 8-2.

Equation 8-2

$$x^{m(i)} = \frac{1}{m} \sum_{t=m(i-1)+1}^{mi} X(t)$$

We can now say that X(t) is partitioned into nonoverlapping files each of duration m. ('i' is used to number these files.) Using this basic scheme of second order stationarity, we can now define Self–Similarity.

Second-order Self-Similarity is a process X(t) is exactly second-order Self-Similar with Hurst parameter H, if H > 0.5, m < 1. See Equation 8-3.

Equation 8-3

$$\gamma(k) = \frac{\sigma^2}{2}((K+1)^{2H} - 2K^{2H} + (K-1)^{2H})$$

X(t) is asymptotically second-order self-similar if (refer to Equation 8-4):

Equation 8-4

$$\underset{m \to \alpha}{Lim} \cdot \gamma^{(m)}(k) = \frac{\sigma^2}{2}((k+1) - 2k^{2H} + (k-1)^{2H})$$

Second-order self-similarity ensures that the process is preserved under various conditions of time aggregation. Hurst parameter can be considered as the governing factor that determines the degree of burstiness accorded to a particular traffic flow. H = 0.5 means nonbursty or Poisson flow as H➨ 1 (H tends to 1) the level of burstiness increases.

To understand the particular form of

$\gamma(k)$

fromEquation 8-2:

If we consider the cumulative process,

$\gamma(t)$

thus by definition of Self Similarity, we get Equation 8-5;

Equation 8-5

$$\gamma(t) = {}_d a^{-H} Y(at). \text{ for } t \geq 0, a > 0$$

where Y(at) is Self-Similar with Hurst parameter, H > 0.5.

Therefore,γ(t) is the time-scaled version of Y(at) normalized by a $^{-H}$. In other words, γ (t) and Y(at) look the same, in other words, follow likewise distributions.

Long-Range Dependence

Consider a second-order Self-Similar process with the autocovariance function

$\gamma(k)$

LetEquation 8-6 denote the auto-correlation function.

Equation 8-6

$$r(k) = \frac{\gamma(k)}{\sigma^2}$$

For

$$0 < H < 1, \quad H \neq 0.5$$

we get Equation 8-7.

Equation 8-7

$$\gamma(k) \sim H(2H-1)K^{2H-2} \quad K \to \infty$$

Furthermore, if β = 2 – 2H, then we get Equation 8-8.

Equation 8-8

$$\sum_{k=-\infty}^{\infty} r(k) = \infty$$

In other words, the auto-correlation function decays very slowly. This kind of decay is termed *long-range dependence*. If

$$\sum r(k)$$

is finite, the process is called *short-range dependentdependence*.[8,9,10]

Heavy-Tailed Distribution

We can define *heavy-tailed distribution* in the following way: A random variable X has a heavy-tailed distribution (given by Equation 8-9) if

Equation 8-9

$$P_r(X > x) \sim cx^{-\alpha}$$

0 < ∝ <2 is called tail index and c some positive constant. Heavy-trailed distributions are

characterized by infinite variance for $0 < \alpha < 2$. The sign ~ denotes asymptotically equal to. For network modeling, a common distribution for heavy-tailed sequences is the Pareto distribution, whose distribution function is given by Equation 8-10.

Equation 8-10

$$P_r\{z \leq x\} = 1 - \left(\frac{b}{x}\right)^{\alpha} . b \leq x.$$

In Equation 8-10, α (alpha) is the shape parameter and b is the location parameter. The mean is given by Equation 8-11.

Equation 8-11

$$\text{Mean} = \frac{\alpha b}{\alpha - 1}$$

The main characteristics of the Pareto distribution are that a random variable obeying the heavy tailed distribution exhibits extreme variability. This distribution gives rise to sequences with many of the values small, some values large, and a few values very large. This is in conformity with today's arrival process of Internet traffic or data traffic.

The purpose of introducing Self-Similarity in this chapter is to acquaint the reader with the process of data traffic flow. The next phase is to map such bursty flow on to optical networks. This might not be an easy task considering the fact that it is difficult to estimate the burstiness of a source. That is so say it is very difficult to calculate or predict the behavior of a source transmitting data by stochastic means. This means that the bandwidth of a channel is quite variable, in addition to the fact that any compromise in bandwidth requirement would create higher end-to-end latency for IP packets over the WDM layer. This is in addition to the buffer lengths and queue sizes at the interface of IP routers, which groom this traffic into higher layers. A more detailed study of IP over WDM from a Self-Similar traffic point of view is necessary for building scalable, efficient, and resilient optical networks for the future.

Transparent Optical Networks

As networks grow and the amount of data that is pumped into the networks reaches new highs, it will become more difficult to provision, manage, and support such volumes of traffic. A lightpath can be envisioned to move across multiple networks and multiple topologies from one client to another. As the lightpath traverses through the network, it might be subject to several O-E-O conversions at various network elements.

O-E-O conversion is an important aspect of optical communication for two reasons. First, an optical signal cannot go on endlessly; there is a limit after which this signal cannot be accepted (detected correctly). Therefore, we have to do a reshape, retime, and reamplify (3R) this signal at some finite distance from the source. Secondly, 3R regeneration is essential for traffic grooming. Traffic grooming can be defined as the aggregation of several low speed traffic streams into a single high speed traffic flow. Imagine that a node has 32 input interfaces each carrying OC-3 traffic. If no scope exists for traffic grooming at the node, this node would require 40 wavelengths to cater to this traffic. On the other hand, we can have a wavelength grooming 3-R device (digital cross connect) which can multiplex the 32 OC-3s into just 2 OC 48s on 2 separate wavelengths thereby conserving the other 30 wavelengths as free.

By intelligently keeping traffic grooming functionality in the 3R regenerators, we can multiplex the traffic streams into a few high bit rate streams. (For example, 16 OC-3s can be made to an OC-48.) This is the advantage of 3-R regenerators (O-E-O type) in the network.

However, there is a serious flip side to having O-E-O regenerators: they are expensive. An estimate for a ring network suggests that 75% of the network costs were for transponders (O-E-O devices) alone. Such devices, although handy for good throughput, are extremely expensive and are difficult to obtain in bulk. Moreover, their exact placement to obtain an optimum configuration can be the object of a non-polynomial time solutions. The cost of 3R transponders is severely high because the retiming of high bit rate 10 Gbps and beyond signals requires precision electronics, which are very expensive. In addition to that, the transponders may have some wavelength tunable lasers, which need to have characteristics of narrow line width and good tunable performance.

Overall, the cost of transponders is the highest cost amongst all the network components. Reducing such costs is imperative for a good optical network. One way is to keep the transponders out of the network. Keeping transponders just at client sites and not in the network (allowing all-optical transparent flow of data) is definitely a good solution.

The issue then is physics related: In the absence of 3R regenerators, can the signal reach the destination? That depends on the transmission distance. Not all lightpaths are long, and not all need regeneration within the transmission distance. Nodal architectures have to be designed keeping in mind full flexibility of the traffic matrix. Reconfigurable OADMs are one such alternative. By keeping a small amount of transponders at each node, we can still achieve almost the same performance as that by a large number of transponder cards. The system has to perform well under different conditions of traffic. Transparent networks are a good solution as long as the optical rules are adhered to. By keeping a minimum number of transponders inside the network, the cost of the network decreases substantially.

Transparent networks for long haul are already under a mature state of deployment. Currently, a high bandwidth optical pipe can be made available from New York to Los Angles. Transparent networks are being actively considered for metropolitan areas as well. In summary, transparent networks are an attractive means for optical communication. Although they are not fully mature, this is one sure area of intense commercial and research activity in the next few years.

Summary

In this chapter, we looked at some of the future trends that will affect WDM networking. Although it is almost impossible to cover each area of research, we highlighted some of the areas that affect WDM networking. In this regard, technologies that can set the future of optical communication, such as burst switching and slot routing, were mentioned. We also mentioned some research projects that look at feasible methods to translate circuit-switched networks of today to packet-switched networks of tomorrow. HORNET in this regard can be viewed as a technology demonstrator. Optical burstmode receiver technology is another emerging technology for tomorrow's WDM packet switched application.

The impact on network protocols for traffic engineering (such as RSVP for providing QoS to traffic) and its impact to optical networks is both interesting and intriguing. The revelation that Internet traffic is not adhering to Poisson distribution but is Self-Similar makes traffic engineering over WDM that much more important now. The migration from lightpath networking to packet switching will not be an easy migration. Migratory technologies will become more common. Photonic slot routing is one such candidate that removes the constraint of wavelength selective switches. The final vision is to have a multiwavelength optical network that can provide packet switching, good reconfiguration, and massive capacity on an entirely transparent core.

References

[1]Yoo, M., C. Qiao, and S. Dixit. "Optical Burst Switching for Service Differentiation in the Next-Generation Optical Internet." *IEEE Communications Magazine*. Feb. 2001.

[2]Chlamtac, I., et al. "Scalable WDM Network Architecture Based on Photonic Slot Routing and Switched Delay Lines." IEEE Proc of INFOCOM. 1997.

[3]Chlamtac, I., et al. "CORD: Contention Resolution Using Delay Lines." *Journal on Selected Areas in Communication*. Special Issue on Optical Networks. June 1996.

[4]Zhang, Lixia, et al. *IEEE Communications Magazine*. May 2002.

[5]Ramaswami, R., and K. Sivarajan. *Optical Networks: A Practical Perspective, Second Edition*. Morgan Kauffman, 2001.

[6]Kazovsky, L., et al. Hornet Architectures OCRL Laboratory Web Site. University of Stanford. www.wdm.stanford.edu.

[7]*Su, Chao, Lian-Kuan Chen, and Kwok-Wai Cheung.* "Theory of Burst-Mode Receiver and Its Applications in Optical Multiaccess Networks." Journal of Lightwave Technology. Vol. 15 Issue 4. April 1997.

[8]Park, K., and W. Willinger, ed. "Self Similar Network Traffic and Performance Evaluation." Wiley Interscience Series. 2000.

[9]Crovella, M., and A. Bestavros, "Self Similarity in World Wide Web Traffic: Evidence and Possible Causes." *IEEE/ACM Transactions on Networking*. Vol. 5. 1997.

[10]Leland, W.E., et al. "On the Self-Similar Nature of Ethernet Traffic (Extended Version)." *IEEE/ACM Transactions on Networking*, Volume 2. 1994.

[11]Black, U. *MPLS and Label Switching Networks*. Prentice Hall, 2000.

Chapter 9. Tests and Measurements

A typical optical fiber network is composed of many passive and active components working in conjunction to form a complete communication system. The performance characteristic of each component (subsystem) affects the total performance of the entire system. To effectively design and implement a fiber network, system engineers need to know the exact characteristics of the fiber links. The attenuation and dispersion characteristics of a fiber change during the installation phase, and they also limit the maximum distance and bandwidth of the fiber. Measuring these parameters and testing for conformance is crucial during the manufacturing, installation, and maintenance phases. Cracks in fiber can degrade the overall performance, whereas breaks in the fiber distort the entire communication channel; therefore, identifying those fiber breaks and cracks is an integral part of the fiber system's ongoing maintenance along with measuring and testing other parameters such as attenuation, dispersion, and so on.

While designing a fiber, a fiber cable vendor is interested in measuring properties such as core and cladding diameters, refractive index profile, fiber attenuation, dispersion, and cut-off wavelength, among others. Most of these parameters do not change during the handling or installation phase.

System engineers, in contrast, are interested in dispersion, attenuation, launch power, receiver sensitivity, amplifier gain, and so on. (Fiber characteristics such as dispersion and attenuation change slightly during the installation phase.) Therefore, the test and measurement requirements for system engineers/installers are different from vendor's testing needs.

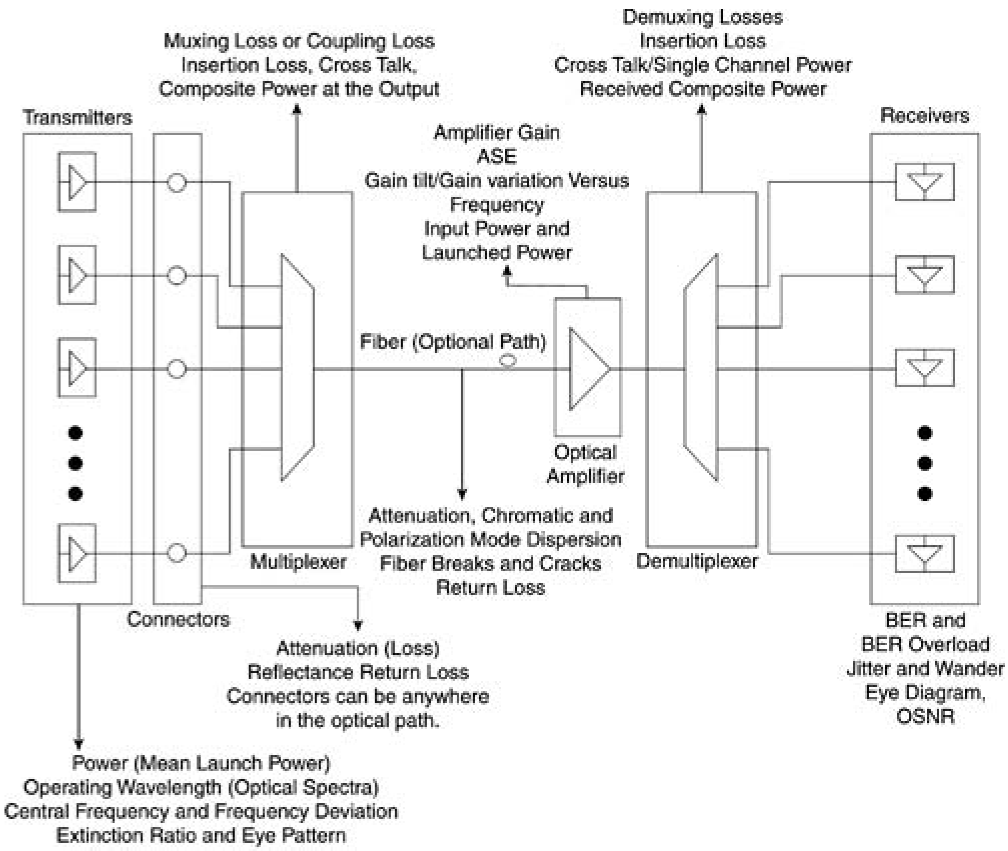
Because an optical network is made up of passive components (splitters, connectors, and couplers) and active ones (transmitter, receiver, and optical amplifiers), a system designer needs to know the characteristics of each component in the network. These passive and active components should operate in a given optical spectrum without degradation in performance.

The vendors usually conduct the performance and conformance testing before launching their products. During the operational and maintenance phase, the operators/installers need to monitor or troubleshoot problems: if there is any break in the fiber, if the transmitter is launching enough power into the fiber, if attenuation is well within the specifications and status of amplifiers in the link.

The performance characteristics that the vendors specify degrade over time even though the networks are designed with these adverse affects in mind. A system designer/installer needs to measure/test the system during the troubleshooting phase. This chapter covers the test and measurement requirements that help system architects and installers/operators design and maintain WDM/optical networks.

A typical WDM network is shown in Figure 9-1. A WDM network is composed of transmitters (source/transponders), connectors (to connect different subsystems), splices, couplers/multiplexers/demultiplexers, optical fibers, amplifiers (EDFAs etc.), and receivers (photodiode/transponders). Figure 9-1 also shows the characteristics/parameters that need to be tested for performance or conformance analysis.

Figure 9-1. Figure 9-1 WDM Network Parameter That Needs to Be Tested



The whole network is divided into multiple subelements, and each network subelement has characteristics that need to individually conform to support an end-to-end system. After the performance and conformance of individual network subelements are complete, the final step is to perform a system-level test or network-level test. This is usually done by testing the bit error rate (BER; receiver sensitivity at 10^{-8} to 10^{-12}), optical signal-to-noise ratio (OSNR), spectral analysis, and eye-masking analysis.

Several standards bodies around the world are involved with optical network component testing requirements (systems testing and measurements standardizing), out of which the International Telecommunication Union (ITU) defines most of the system standards. For component testing and conformance, ITU-T, the Telecommunication Industry Alliance (TIA), along with the Electronic Industries Association (EIA) define fiber-optic test standards. These standards are also known as *fiber-optic test procedures* (FOTPs). Basic measurements and standards of fiber characteristics, such as dispersion, attenuation, cut-off frequency; mode field diameter, and so on have been characterized primarily by the National Institute of Standards and Technology (NIST) in the U.S. (based at Boulder, Colorado). Other countries have similar organizations (such as the National Physical Laboratory (NPL) in Britain) that oversee basic standards and Department of Telecommunication (DoT) in India.

ITU-T G.692 can be used as a good benchmark for evaluating multichannel optical transmitters and receivers. No specific industry standards are currently available to measure/evaluate multiplexer/demultiplexer and optical amplifier (OSNR and noise figure, or NF) specifications. (The standardization process has just begun.) Choosing the right passive component is important to supporting a particular application. Because each vendor reports the specifications it feels are the best, it is impossible to compare passive components from different vendors unless inhouse testing is done. Refer to Table 9-1 for more explanations.

Table 9-1. Required Characteristics/Parameters Defined by ITU-T G.692

Transmitter	Mux (Output)	Amplifier	Optical Path	Demux (Input)	Receiver
Spectral characteristics	Optical transmit-side cross-talk	Multichannel gain Variation	Attenuation	Mean channel input power	Receiver Sensitivity
Mean launched power	Channel output power	Multichannel gain tilt	Dispersion	Mean total input power	Receiver overload
Extinction ratio	Total launched power	Multichannel gain-change difference	Maximum discrete reflectance	Channel signal-to-noise ratio (SNR)	Optical path penalty
Eye pattern mask	Channel SNR	Total received power	Minimum return loss	Optical signal cross-talk	OSNR
Central frequency	Maximum channel power difference	Total launched power		Maximum channel power difference	Minimum receiver wavelength
Channel spacing		OSNR and NF			Maximum receiver wavelength
Central frequency deviation					

- Specifications that can be tested for passive components— This includes insertion loss, polarization dependent loss (PDL), center frequency, bandwidth, passband, ripple, return loss, cross-talk and operating temperature. With multiplexer/demultiplexer testing, in addition to the preceding specifications, channel spacing, channel uniformity, OSNR, spectral power density, and pass-band shape are important and need to be tested/characterized.
- Specifications that are to be tested for amplifier evaluations— This includes gain, polarization-dependent gain, noise figure (NF), OSNR, amplified spontaneous emission (ASE), gain flatness, gain tilt, and input/output power automatic gain control (AGC), automatic balancing control (ABC) and automatic loss control (ALC). The following section discusses how these measurements are made and what equipment is used to test these specifications.

Test and Measuring Devices

Basic test equipment to test power, dispersion, OSNR, and other spectral content is necessary for installing and maintaining a WDM network. The recommended basic test equipment includes optical power meters, attenuators, optical spectrum analyzers (OSAs), and optical time domain reflectometers (OTDRs). This equipment helps the installation team troubleshoot and complete the installation more quickly. Numerous test equipment vendors manufacture these test equipments. More sophisticated equipment is usually for laboratory use, whereas more rugged, small handheld units are preferred for field use. Among other things, a test engineer/system installer might need variable attenuators, power meters, tunable lasers, BER analyzers, and polarization analyzers in his tool belt for ongoing maintenance.

Power Meters

An optical power meter is a vital tool for measuring absolute power in an optical fiber communication system. Power measurement is one of the most basic fiber-optic measurements. It accounts for loss measurements along with the power variations of the transmitter (source) or at the receiver end. Optical power meters typically use semiconductor photodetectors such as Silicon (Si), Germanium (Ge), or Indium Gallium Arsenide (InGaAs) depending on the application wavelength.

Si detectors are used in the 850 nm regions, whereas Ge and InGaAs detectors are typically used in the 1310 and 1550 nm regions due to the typical responsivity of these detectors in these regions. InGaAs detectors have the same responsivity as that of Ge detectors and better noise characteristics (less noisy), but InGaAs is more expensive than Germanium. Optical power meters are calibrated using a reference standard that is administrated by NIST. (Calibration is done at 850, 1310, and 1550 nm regions.) A handheld optical power meter from Agilent technologies is shown in figure 9-2.

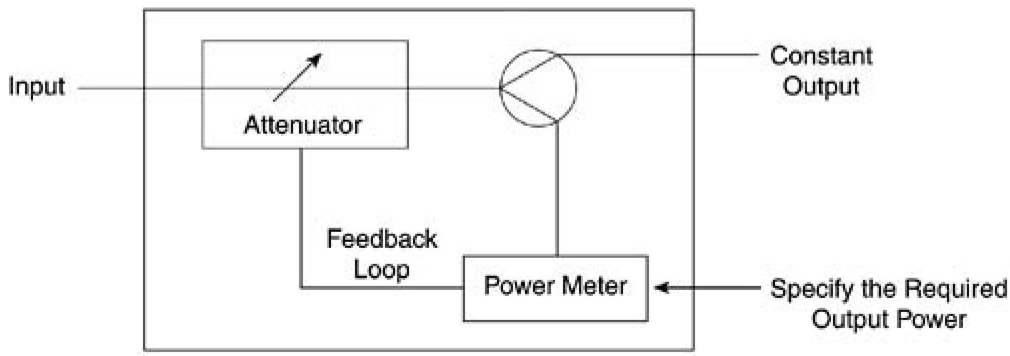
Figure 9-2. Optical Power Meter (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)

Attenuators (for Testing Applications)

Signal power launched from transmitters is usually high and needs to be attenuated before plugging into analyzers to prevent damage. Attenuators come in several form factors, with size varying from small handheld to large desktops. Variable attenuators with variable ranges (up to 60 dB) are available and are ideal for most applications. The wavelength (band) of typical operations is 1310 and 1550 nm.

In some application testing, we might need the power level to be constant at the input of the device under test (DUT). For these types of applications, variable attenuators with feedback are used (see Figure 9-3). Here, regardless of the input power, the output power is always constant; the output power can be specified through software control or knob settings. Variable attenuators find applications in the bit error rate tester (BERT); characterization of receivers, transmitters, and line cards; DWDM channel equalization; power loss simulation in single channel fiber-optic links; testing and calibrating the linearity of power meters; and optical amplifier testing (characterization of EDFAs and Raman amplifiers).

Figure 9-3. Schematic Diagram of a Variable Optical Attenuator with Constant Output



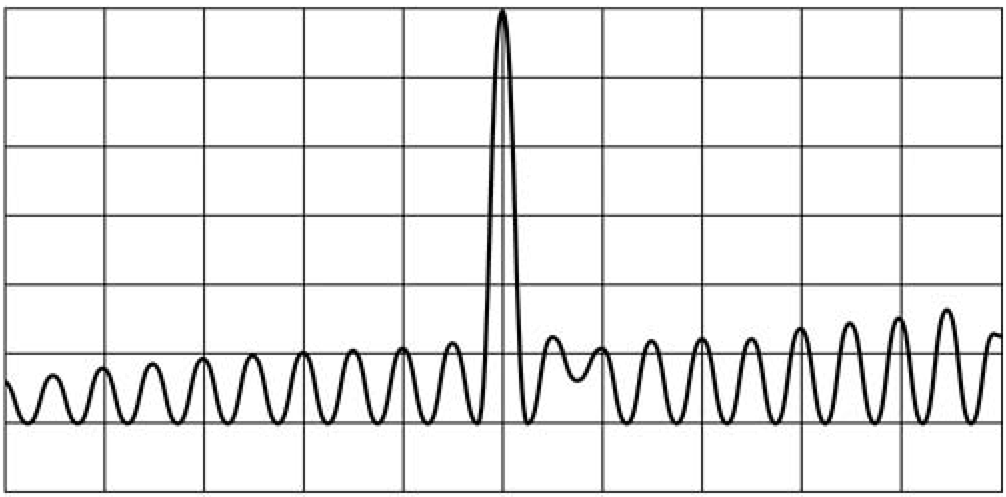
A typical optical attenuator from Agilent is shown in Figure 9-4.

Figure 9-4. A Handheld Optical Attenuator, Agilent N3977A (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)

Laser Sources

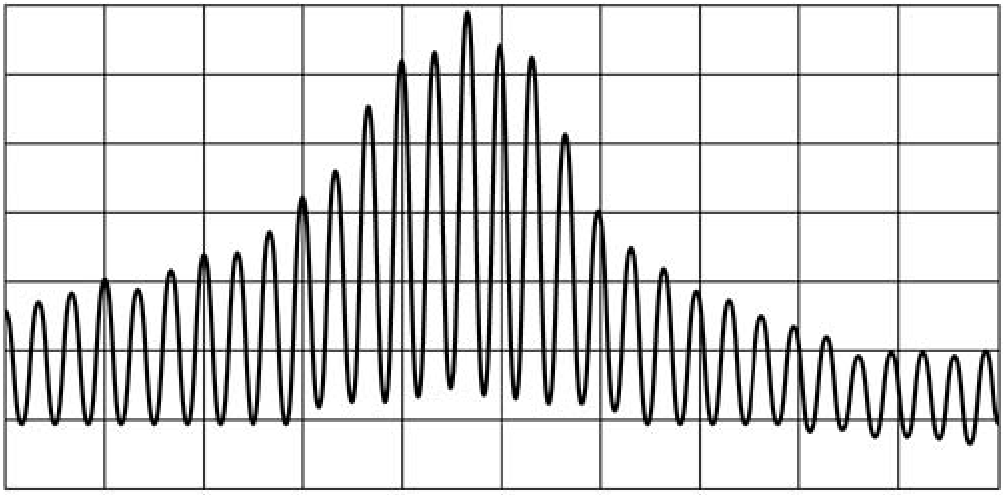
Laser sources are used as input devices to the WDM device under tests. Lasers and their operations are explained in Chapter 2, "Networking with DWDM -1." The distributed feedback (DFB) laser sources are used in applications such as WDM multiplexer/demultiplexer testing, amplifier testing, and system testing. Output of a DFB laser is shown in Figure 9-5. Note the suppressed side modes and the single dominant mode producing a thin wavelength.

Figure 9-5. Output of a DFB Laser



Fabry Perot (FP) lasers are also used in testing optical amplifiers, WDM systems, and so on. The output of an FP laser is shown in Figure 9-6. Both types of lasers are explained in Chapter 2.

Figure 9-6. Output of FP Lasers



Tunable Lasers

Tunable lasers are preferred when it is necessary to change wavelengths during system tests. It is cheaper to use tunable lasers than DFB lasers in certain WDM application tests, where the test has to be done on multiple wavelengths. Tunable lasers find applications in the testing of amplifiers, multichannel optical systems, and so on. Tunable lasers that have tuning ranges from 1525–1600 nm are commercial available. Generally these lasers are built on Distributed Bragg Reflector (DBR) technology, though other newer technologies are also gaining prominence.

Broadband Light Source

Broadband light source is generated using edge-emitting LED (EELED) technology. EELEDs emit more power source and white light into a single-mode fiber (SMF). This is used to characterize long fiber cables, used as a noise source in amplifier characterization (noise gain profile). It can also be used with other passive devices (testing) in characterization of those devices.

Optical Spectrum Analyzers

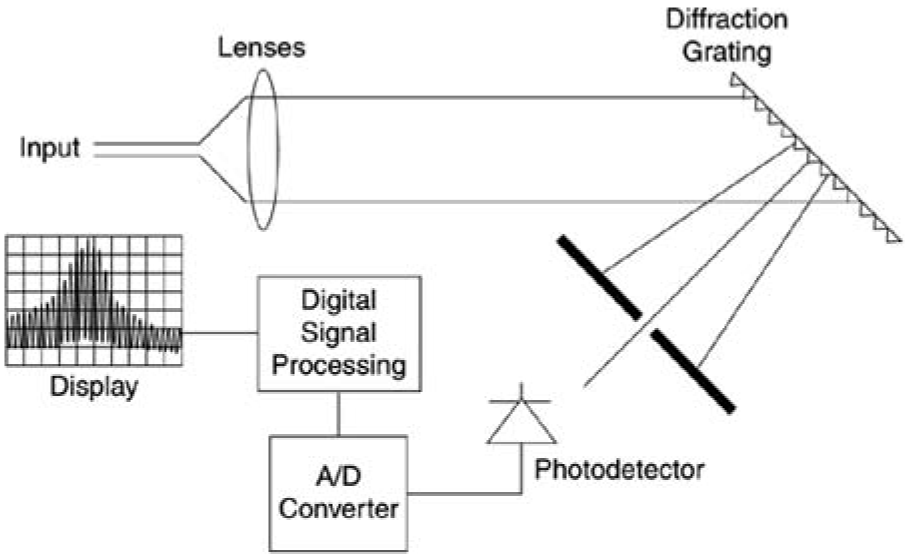
Optical spectrum analysis is the measurement of optical signal power as a function of frequency and time. It is important to analyze the spectral width, power, and pulse spreading (dispersion) of an optical light source in optical communications systems to determine the quality of the optical signal. The optical spectrum analyzers can display the signal characteristics along with noise characteristics, which help in the calculation of OSNR. Optical spectrum analysis is commonly used in testing lasers for spectral width and power distribution, as well as testing transmission characteristics of optical devices.

Several different types of spectrum analyzers are available, out of which diffraction grating-based and interferometer-based analyzers are the most common. Each test equipment vendor has its own method to architect and analyze the optical signal. Diffraction grating-based optical spectrum analyzers are used to analyze lasers and LEDs and have a typical resolution range of .1–10 nm. The interferometer-based analyzers are of two types: Fabry Perot (FP) interferometer- and Michelson interferometer-based analyzers. The FP-based analyzers have a narrow fixed resolution and are used to measure laser chirp. Michelson interferometer-based optical spectrum analyzers display the spectrum by calculating the Fourier transform of a measured interference. They are typically used for direct coherence length measurements.

Operation of a Spectrum Analyzer

The signal is first passed through a tunable filter (grating based or interferometer based), which resolves the signal into individual spectral components (see Figure 9-7).

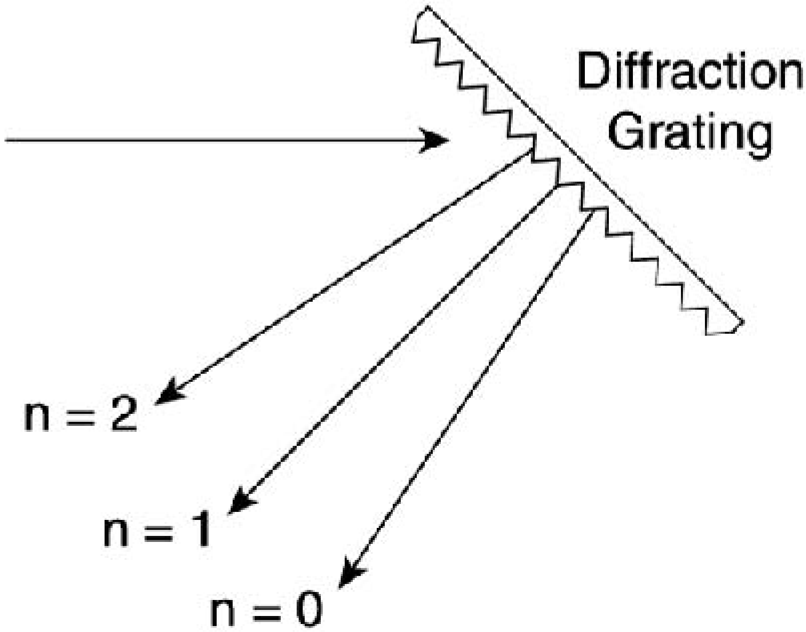
Figure 9-7. Block Diagram of a Typical Spectrum Analyzer (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



The filtered signal falls on a photo detector, which then converts the optical signal into an electrical signal. The signal is then fed into an Analog to Digital (A/D) converter. Finally, the signal is processed by a digital signal processor and displayed on the scope. A simplified block diagram of a diffraction based spectrum analyzer is shown in Figure 9-7. A real spectrum analyzer is more complex with circuits, analyzing and controlling functions.

Most of the optical analyzers use a monochromator as an optical tunable filter. A *monochromator* is made up of diffraction gratings, which separate the input signal into its component wavelengths. As explained in Chapter 2, diffraction grating is a diffractor of light with gratings and is shown in Figure 9-8. The spacing between the gratings is narrow and is approximately equal to the wavelength we want to test. When the input signal (optical light) falls on this grating, the light is reflected in all directions.

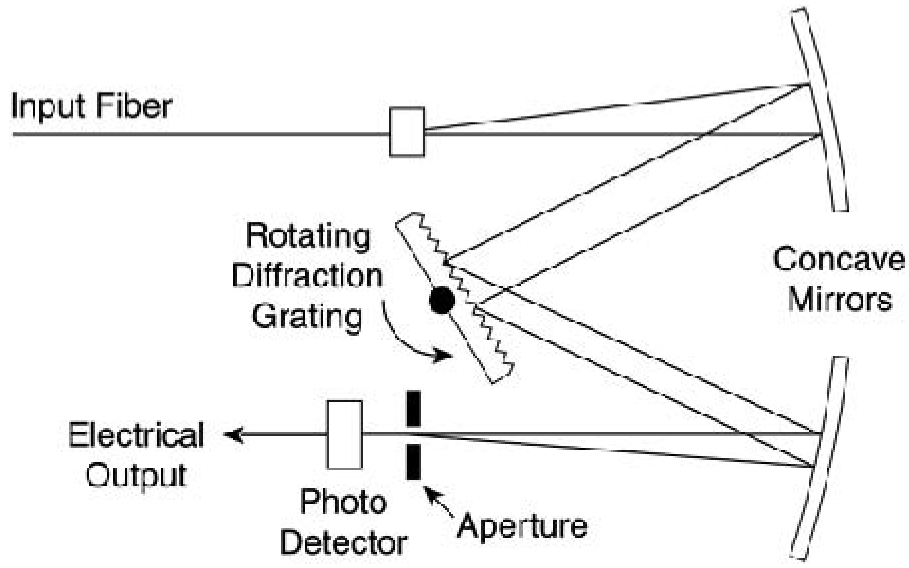
Figure 9-8. Diffraction Gratings



The diffraction grating separates the input signal into several output beams. Except for the zero-order beam, all the other beams are a separate wavelength. The zero-order beam (n = 0) is similar to the reflection course, if we replace the grating with a mirror. The first-order beam (n = 1) is formed as a result of the constructive interference of reflections from each grating. For constructive interference to occur, the path length difference between reflections from the adjacent gratings must be equal to a multiple of wavelengths. There-fore, for the first-order beam, the path length difference is equal to one wavelength; for the second-order beam, the path length difference is equal to two wavelengths, and so on. Each order beam is an individual wavelength that spectrum analyzers use to analyze the optical input signals.

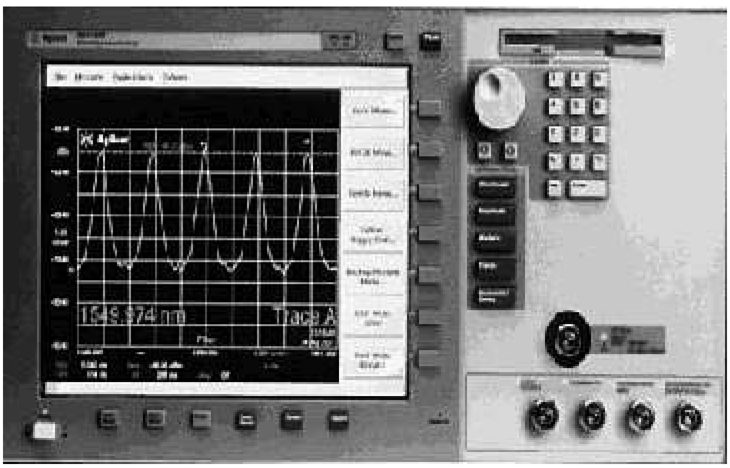
The operation of a diffraction grating-based optical spectrum analyzer is shown in Figure 9-9. As the diffraction grating rotates, the analyzer sweeps a range of wavelengths. The wavelength that falls on the photodiode depends on the position of the gratings. (Only one wavelength passes through the aperture at any given time.) This technique allows wide wavelength coverage. Concave mirrors are used to focus the diffracted signal onto a detector.

Figure 9-9. Diffraction Grating-Based Optical Spectrum (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



Even though the basic principles are the same, each vendor uses unique techniques to increase the dynamic range and sensitivity of the spectrum analyzers. For example, Agilent uses a unique wavelength selection scheme called double pass monochromator to increase the dynamic range and sensitivity. A typical spectrum analyzer from Agilent (model 8614xB) is shown in Figure 9-10.

Figure 9-10. Agilent 8614xB Family of Optical Spectrum Analyzers (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



Optical Time Domain Reflectometer (OTDR)

The Optical Time Domain Reflectometer (OTDR) is used to locate faults in the optical fiber links. With an OTDR, we can also determine the fiber loss (attenuation), fiber length, connector loss, splice loss, reflectance, and chromatic dispersion of the fiber under test. All measurements except chromatic dispersion require just one wavelength, chromatic dispersion needs multiple wavelengths. To determine loss and estimated length, an OTDR uses the technique of backscattering.

How OTDR Works

An *event* is a term that identifies any abnormalities in the fiber, including cracks, connectors, splices, bends, and so on. An event is anything that causes loss or reflection in the fiber medium other than the fiber itself. A high-power pulse is launched into a fiber cable. As light travels, some of the light is absorbed and some is scattered in all directions. The light that is scattered into the direction of an OTDR (back scattering) is used for measurement purposes. Distance is calculated using the formula shown in Equation 9-1.

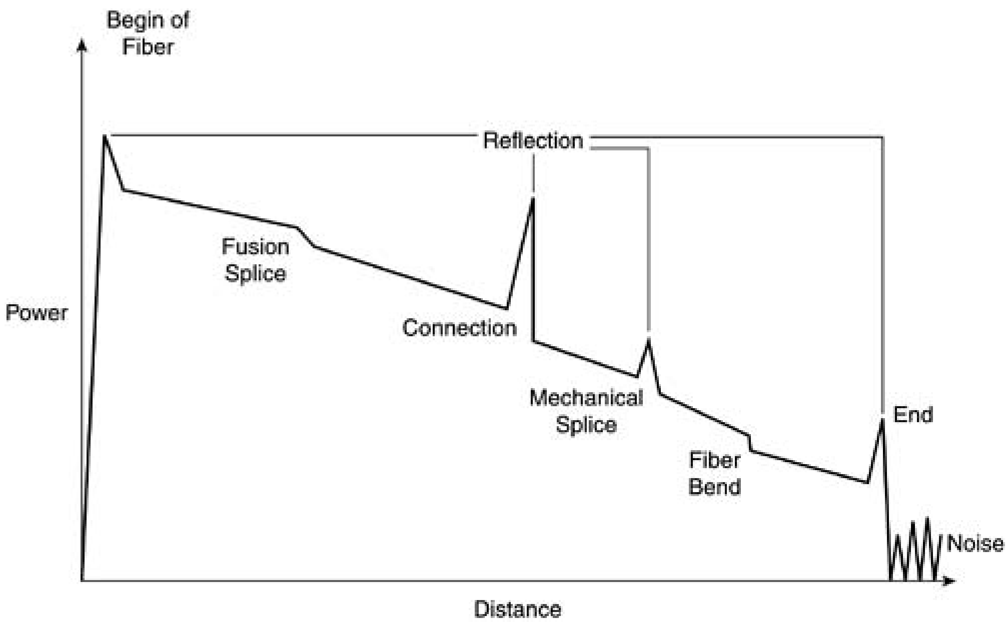
Equation 9-1

$$S = \frac{T_M \times v_0}{\mu}$$

In the equation, T_M is the time taken for the round trip (from the OTDR and back), v_0 is the velocity of light in a vacuum, and μ is the refractive index of the core. Attenuation is calculated by measuring the intensity of reflected light.

A typical OTDR trace with the most common events is shown in Figure 9-11. The vertical axis represents power (Y–axis), and the horizontal axis represents distance (X-axis). The OTDR displays the results graphically and can display the distance between multiple events.

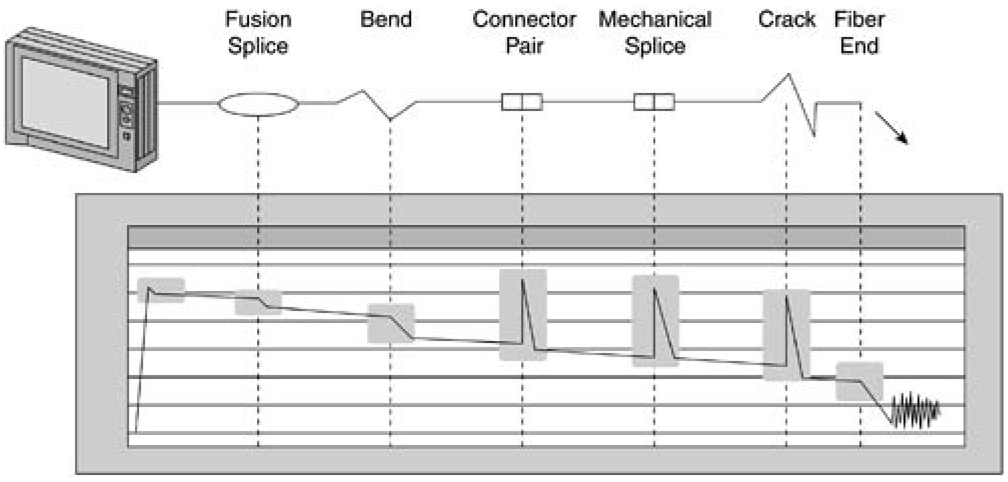
Figure 9-11. OTDR Trace



From the trace in Figure 9-11, we can see a small level of attenuation and strong reflection at the beginning and at the end of a fiber. A break in the fiber has a low reflectance; therefore, the trace data drops to the noise level. Connectors within a link cause attenuation and reflectance; a mechanical splice has a similar characteristic as that of a connector, but the loss and reflectivity are much less. In a fusion splice, only loss can be detected. Sometimes a splice exhibits a gain, and in that case, it is important to test the fiber from the other side. Bends can also be identified using OTDR, as explained in Chapter 1, "Introduction to Optical Networking." Macro-bend has higher attenuation at higher wavelengths. Using two different wavelengths, it is possible to isolate macro bends. Cracks produce reflection and loss. The loss due to a crack is high compared to the loss due to connectors, splices, and bends.

OTDR finds application during the manufacturing, installation, and maintenance phase. During installation, OTDR is typically used to perform an end-to-end quality check, fiber continuity check, and link loss measurement. During the maintenance phase, OTDR is used to isolate/locate fiber faults, faults due to stress location, test connectors, mismatches after repair, and troubleshooting. A typical OTDR from Agilent is shown in Figure 9-12.

Figure 9-12. OTDR (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



Polarization and Its Measurements

Polarization and its impact on fiber-optic transmission system are explained in Chapter 2. Here, we explain how polarization is calculated and what test equipments are used to analyze polarization. After chromatic dispersion is compensated, the system data rate in an optical system is limited by polarization mode dispersion (PMD). The polarization dependent loss affects the system by selectively degrading certain states of polarization (SOPs). Polarization dependent loss (PDL) can be measured in several different ways, out of which the *Jones method* (measures PDL from the Jones matrix) and the *Muller matrix method* (simulates a test path with four known SOPs) are most widely used. A polarimeter is used to calculate the SOP of input light. (The output of a polarimeter is a normalized Stokes' parameters.)

Jones Vector

A slowly varying field in a vectorized form depicted is Equation 9-2.

Equation 9-2

$$E(t) = E_x(t).E_y(t).E_z(t)$$

For general conditions, $E_z(t)$ is negligible, resulting in Equation 9-3.

Equation 9-3

$$E(t) \equiv E_x(t).E_y(t)$$

Therefore, the electric field vector is orthogonal to the direction of propagation, which in our case is the Z-axis. (Assume the fiber to be a cylinder and propagation to be in the center Z-axis.)

The vector $\{E_x(t), E_y(t)\}$ is then known as the Jones vector. It tells us the instantaneous state of polarization (SOP). The Jones vector is shown in Equation 9-4.

Equation 9-4

$$\{E_x, E_y\} = E\sqrt{1-k}.\sqrt{k}\exp(i\delta)$$

In the equation, k is the power-splitting ratio between E_x and E_y, and δ is the phase difference between E_x and E_y.

Stokes Vector

A further measurement of polarization is from the Stokes vector. The Stokes vector and the Jones vector are related. The Stokes vector is shown in Equation 9-5.

Equation 9-5

$$S(t) = \{S_0(t)S_1(t)S_2(t)S_3(t)\}$$

Upon simplification, we get Equation 9-6.

Equation 9-6

$$S_0(t) = S_1^2(t) + S_2^2(t) + S_3^2(t)$$

We get individual values in Equations 9-7–9-10.

Equation 9-7

$$So(t) = |E_x(t)|^2 + |E_y(t)|^2$$

Equation 9-8

$$S_1(t) = |E_x(t)|^2 - |E_y(t)|^2$$

Equation 9-9

$$S_2(t) = |E_x(t)E_y{}^*(t)| + |E_y(t)E_x{}^*(t)|$$

where, A^* denotes the complex conjugate of A.

Equation 9-10

$$S_3(t) = (|E_x(t)E_y{}^*(t)|) - |E_y(t)E_x{}^*(t)|$$

Following is an explanation of the subvectors S_0, S_1, S_2, and S_3:

- S_0 is the average optical power.
- S_1 is the power difference between horizontal and vertical components.
- S_2 gives the angle of polarization.
- S_3 denotes the preference for clockwise (right) or counterclockwise (left) polarization.

From the Stokes vector, we get Equation 9-11.

Equation 9-11

$$P = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}$$

where by 'P' is the degree of polarization. P = 1 denotes a fully polarized signal, such that its polarization state is time invariant.

The relationship to the Jones vector is shown in Equations 9-12–9-14.

Equation 9-12

$$S_1 = 1 - 2k$$

Equation 9-13

$$S_2 = 2\sqrt{k(1-k)}\cos\delta$$

Equation 9-14

$$S_3 = 2\sqrt{k(1-k)}\sin\delta$$

where

$$k = \frac{\langle E_x \rangle}{\langle E_y \rangle}$$

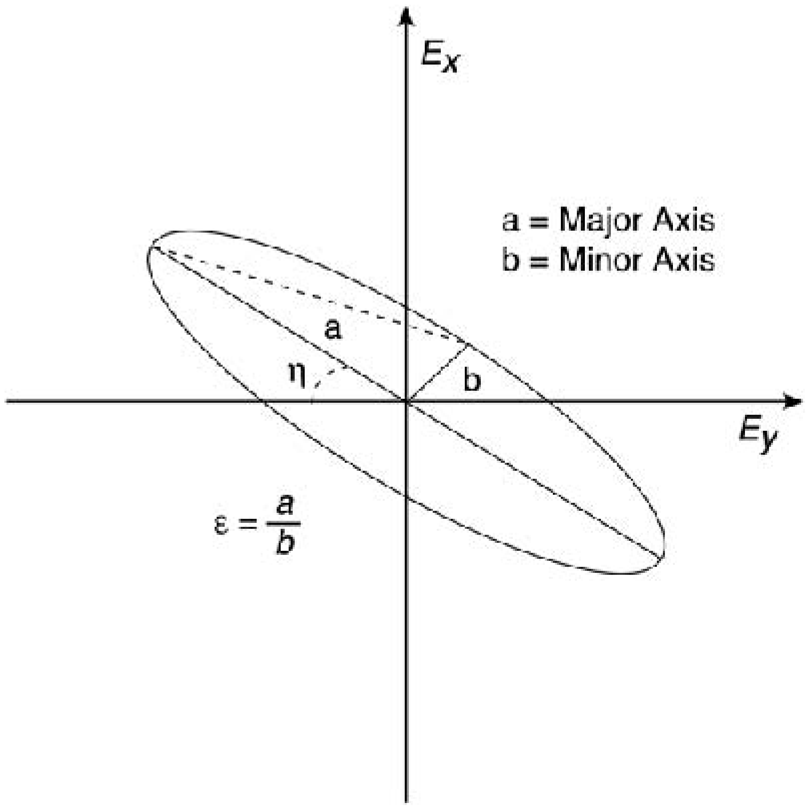
is the ratio of the average E_x and average E_y and δ is the phase difference (E_x, E_y).

Polarization Ellipse and Poincare' Sphere

A state of polarization (SOP) for a signal can be geometrically represented by a polarization ellipse, of ellipticity μ and azimuth η. The ellipticity depicts the ratio of minor axis to major axis. In other words, if μ is small, the ellipse is highly elongated; while if $\mu = 1$, the ellipse is a circle.

If μ is positive, the polarization is right-SOP, and if μ is negative, the polarization is left-SOP. The azimuth η describes the polarization state orientation with respect to the x-axis (Figure 9-13).

Figure 9-13. Polarization Ellipse



The ellipse is related to the Jones and Stokes vectors, as shown in Equations 9-15 and 9-16.

Equation 9-15

$$\tan 2\eta = \frac{2\sqrt{k(1-k)}\cos\delta}{1-2k}$$

Equation 9-16

$$\sin 2\varepsilon = 2\sqrt{k(1-k)}\sin\delta$$

The ellipse is related to the Stokes vector, as shown in Equations 9-17–9-19.

Equation 9-17

$$S_1 = \cos 2\eta \cos 2\varepsilon$$

Equation 9-18

$$S_2 = \sin 2\eta \sin 2\varepsilon$$

Equation 9-19

$$S_3 = \sin 2\varepsilon$$

Poincare' Sphere

From the theory of Stokes vector, we have Equation 9-20.

Equation 9-20

$$P = \frac{(S_1^2 + S_2^2 + S_3^2)^{\frac{1}{2}}}{S0} = \left[\frac{S_1^2}{S_0} + \frac{S_2^2}{S_0} + \frac{S_3^2}{S_0}\right]^{\frac{1}{2}}$$

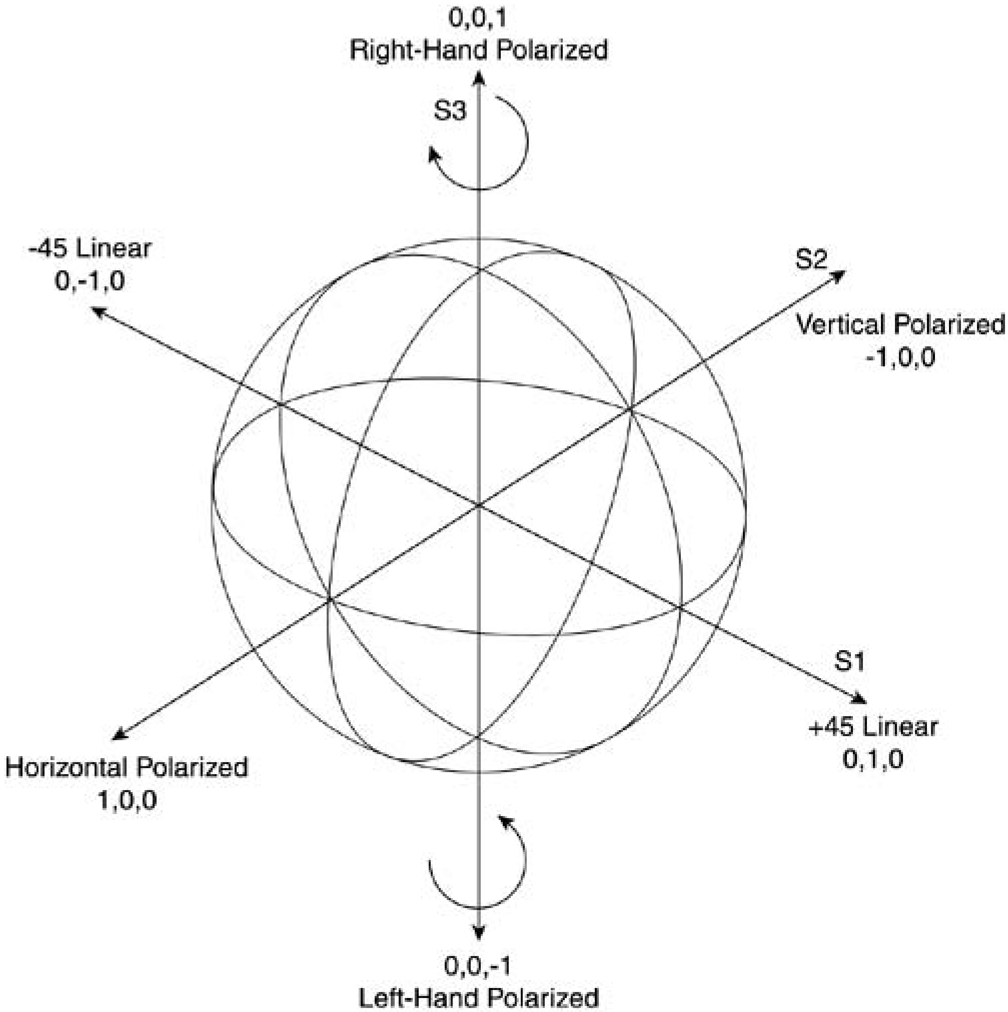
Polarization (P) is shown in Equation 9-21.

Equation 9-21

$$P = [s_1^2 + s_2^2 + s_3^2]^{\frac{1}{2}}$$

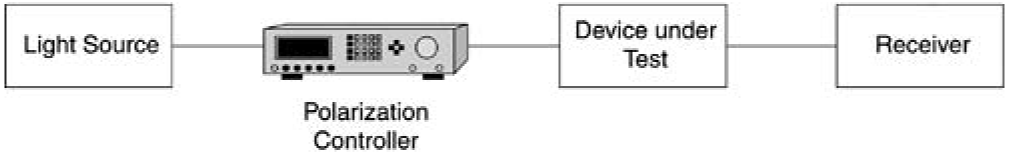
P can be plotted as a point in 3D space tracing out a sphere as the boundaries of P with radius S_0. This sphere is called the *Poincare' Sphere* and is shown in Figure 9-14. The equator represents linear polarization, whereas the North and South Poles represent right and left polarization. Points within the Northern hemisphere represent right-elliptical polarization, and points within the Southern hemisphere represent left-elliptical polarization.

Figure 9-14. Poincare' Sphere



A polarization controller can control the state of polarization of an optical signal. A polarization controller produces all states of polarization and covers the entire Poincare sphere in a pseudo-random manner. To measure losses dependent on states of polarization, it is essential to have polarization controllers, which synthesize the signal for the required output before it reaches the receiver.Figure 9-15 shows a conceptual setup to measure the polarization dependent loss of the device under test.

Figure 9-15. Polarization Controller (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



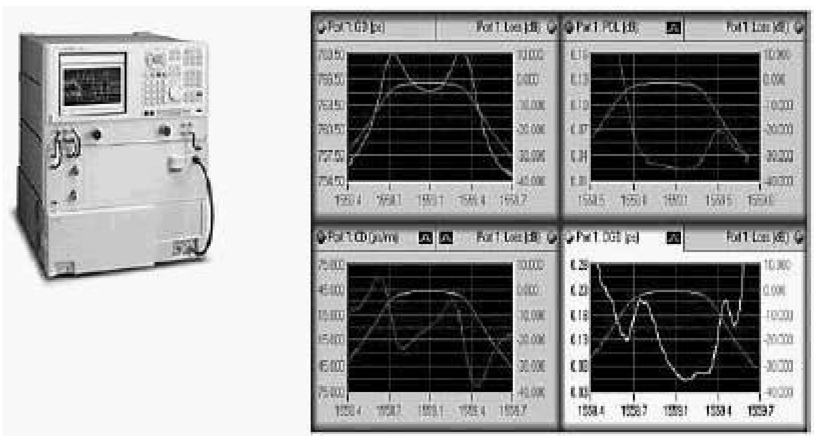
PDL and PMD Measurement

Polarization dependent loss can be defined as the difference in the maximum and minimum variation in loss (transmission/insertion) of an optical device over all states of polarization (SOP) and is expressed in dB.

A typical PDL for an optical connector is less than .05 dB and varies from component to component. The complete polarization characterization of optical signals and components can be accomplished using a polarization analyzer.

Chromatic dispersion can be measured using optical dispersion analyzers. Modern analyzers can simultaneously measure chromatic dispersion, polarization mode dispersion and polarization dependent loss. The Agilent 86038A Optical Dispersion Analyzer (ODA) (shown in Figure 9-16) can simultaneously measure CD, PMD, group delay, insertion loss, and PDL with a single connection.

Figure 9-16. Agilent 86038A Can Be Used to Measure CD, PMD, PDL, Group Delay, and Insertion Loss (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)

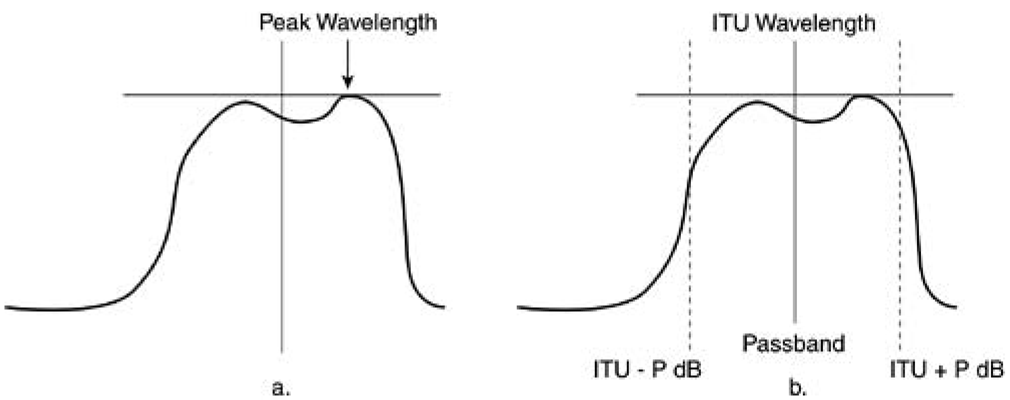


Characterization of a WDM Multiplexer/Demultiplexer and Optical Add/Drop Unit

The main terms that describe the performance of optical passive components are discussed here. The measurements of these parameters are important to evaluate total end-to-end system or network performance:

- Nominal wavelength— This is the wavelength for which the system is designed to operate. DWDM systems have more than one nominal operating wavelength. Channel spacing determines the wavelength spacing of the nominal wavelength(s). (The ITU proposed standard is followed.)
- Peak wavelength— This is the wavelength measure at the output of an optical device with the minimum loss between the input and output (refer to Figure 9-17).

Figure 9-17. (a) Peak Wavelength (b) Pass Band



- Mean center wavelength— This is the mean value of the wavelength measured at the edges of the band. (*Band edges* are the wavelengths below 1 dB peak.) In ideal cases, nominal wavelength and mean center wavelength should be the same.
- Bandwidth (BW)— This is the width of a filter at a specified level below the maximum peak.
- Passband— This is the reference point upon which measurements are taken. It refers to the useable operating bandwidth and is always in reference to ITU-defined wavelengths. If a passband is expressed as P, then the passband is mean central wavelength + P and mean central wavelength - P (refer to Figure 9-17b).
- Passband shape— This is the shape of the spectral response within the passband. Flat response, Gaussian response, and Lorentzian response are the three most commonly known response types.
- Insertion Loss— This is the difference in power at the input and the output of the device under test. It is usually expressed in decibels. (This is per channel loss on a DWDM system.)

- Cross-talk— This specifies how much power is received from the adjacent channels (both left and right channels). It is the worst-case difference between the minimum inband power of the channel under test into an adjacent channel passband and the maximum insertion loss within the passband of the channel under test.
- Return loss— This refers to the portion of the original input power that is reflected back. It is the ratio of the incident power and the reflected power and can also depend on the wavelength.
- PDL— This is the difference in the maximum and minimum variation in loss (transmission/insertion) of an optical device over all SOPs and is expressed in decibels. PDL adds to the insertion loss and cross-talk of the device.
- Best- and worst-case loss— If we consider all the previously defined parameters, there will be a combination of wavelength and state of polarization, where the loss of the device under test reaches its maximum and minimum. This maxima and minima represents the best and worst case losses.

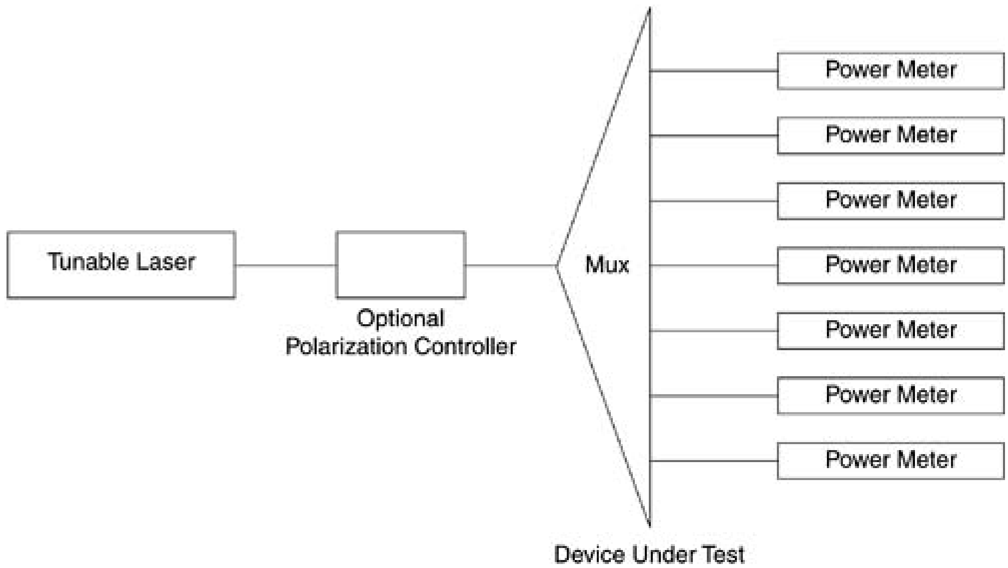
Passive Component Testing

A passive component serves in many different functions either as standalone or as building block for complex structures in an optical network. WDM networks require passive components to couple and decouple channels. The properties of interest in passive component testing are insertion loss, passthrough loss, flatness, rejection, and wavelength dependence. Following is the equipment that can be used for passive WDM testing:

- Tunable lasers
- Broadband light source/ASE source
- Power meter or optical multimeter
- Optical spectrum analyzer
- Polarization controller
- Polarization analyzer

Different combinations of equipment can be used to perform varieties of tests depending on the requirements.Figure 9-18 shows a typical solution for testing a DWDM multiplexer, which requires multiple channel characterization. Tunable lasers are capable of continuously tuning at various sweep speeds; power meters are used to record the power levels. Polarization controls are used along with polarization analyzers if PDL measurements are required. A broadband source with low spontaneous emission can be used instead of tunable lasers in this test case.

Figure 9-18. Testing WDM Mux


Tunable Laser
Optional
Polarization Controller
Mux
Power Meter
Power Meter
Power Meter
Power Meter
Power Meter
Power Meter
Power Meter
Device Under Test

EDFA Testing

To qualify an EDFA or any other optical amplifier, it is important to know how well it performs as a booster amplifier (high power) or as a preamplifier (low noise). Standard tests include the following measurements:

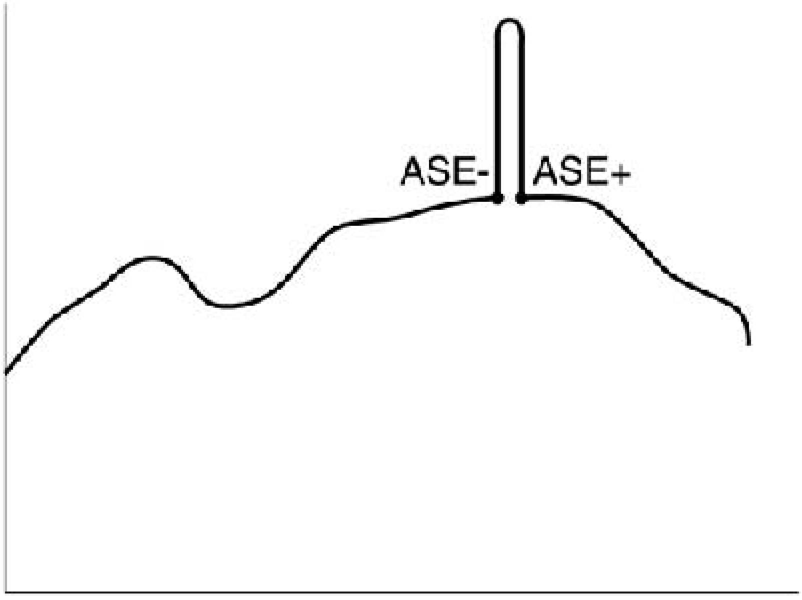
- Output verses input power
- NF verses input power
- Signal gain verses wavelength

Amplifier tests also show when the saturation begins and which wavelengths are best amplified using a particular EDFA. Apart from this, we need to measure gain flatness, power tilt, and OSNR values in multiwavelength applications. Different noise figure (NF) measurement techniques are introduced, each having its own advantage. The most common procedures used today are interpolated source subtraction (ISS) and time-domain extinction (TDE).

Interpolated Source Subtraction (ISS)

To accurately determine the NF, we measure the ASE level at each wavelength. It is impossible to measure the ASE level directly. In interpolation, we determine the ASE of the amplifier by measuring the total spontaneous emission at a wavelength just above and just below each signal (ASE- and ASE+, as in Figure 9-19), and then interpolating to determine the level at each wavelength. A two-point linear interpolation was used previously for measurement purposes. We can use four-point quadratic interpolation or curve-fitting algorithms for greater accuracy. The ISS method is used to test or characterize the amplifiers in different WDM applications (see Figure 9-19).

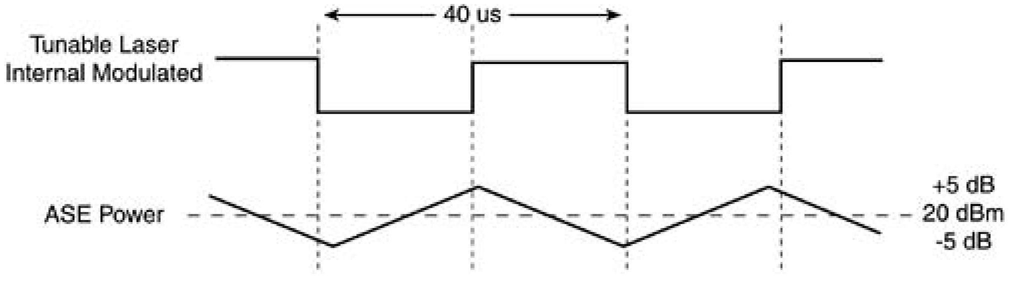
Figure 9-19. ASE Interpolation (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



Time Domain Extinction

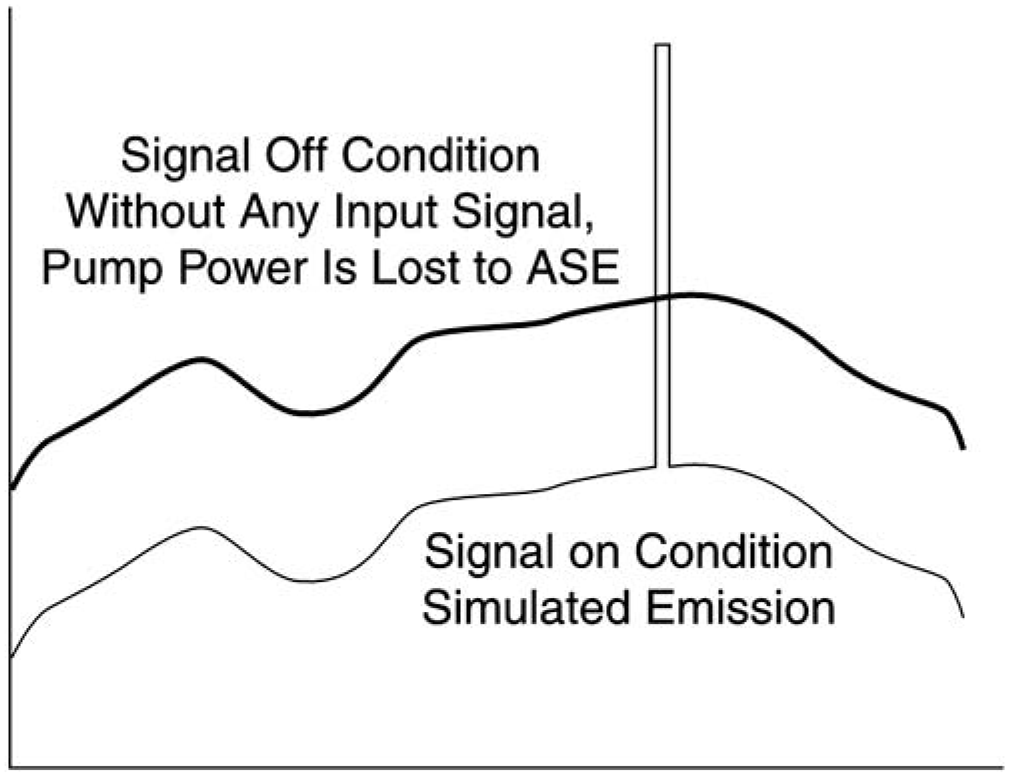
The ASE power level of EDFAs does not change immediately when the input signal is turned off. The time domain extinction method uses this characteristic to measure the ASE of any wavelength. During the measurements, the input signal is modulated on an ON/OFF (external or direct modulation) scheme at a frequency of 25 KHz or higher. The ASE change during the modulation is kept constant, as shown in Figure 9-20.

Figure 9-20. Timing Diagram (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



The OSA is then gated to produce measurement in both signal-off and signal-on conditions. The signal-off spectrum allows the direct determination of ASE. Figure 9-21 shows the signal-off and signal-on conditions. This is an accurate method for measuring the NF of EDFA amplifiers. This technique will not work with SOA and Raman amplifiers.

Figure 9-21. EDFA Spectrum (Signal-On and Signal-Off Conditions)

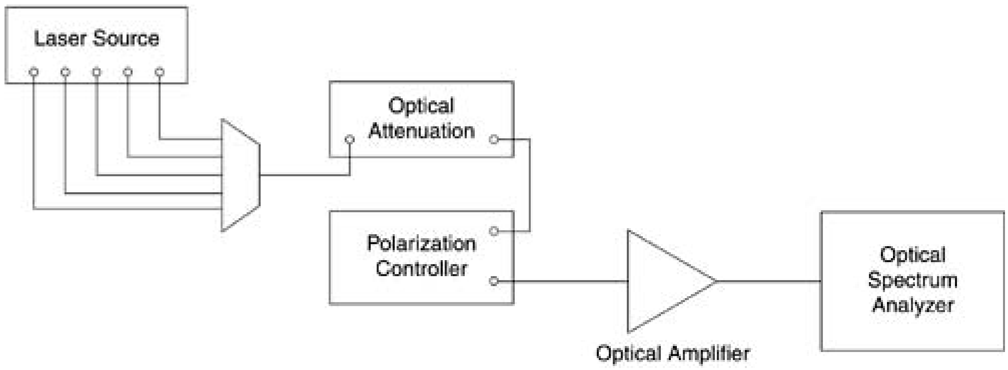


To test an optical amplifier used in a WDM system, we need the following test equipment:

- A tunable laser (should have a wide tunability) or multiple laser sources
- A broadband light source and ASE source
- A polarization controller
- An optical spectrum analyzer
- An optical attenuator

A typical test setup is shown in Figure 9-22.

Figure 9-22. Test Setup to Test Optical Amplifier in a WDM Application (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



The laser sources are used as input channel to the optical amplifier under test. The laser sources can be DFB lasers or tunable lasers (DBR). The variable attenuators vary the power for different input power selections, and the polarization controllers randomize the state of the polarization. The OSA displays the result in a graphical form.

Noise Gain Profile

A faster and alternative method to characterize the gain shape is to use noise source in combination with the TDE technique, as shown in Figure 9-23. When the signal from the tunable laser reaches saturation, the EELED probes the amplifier with a broadband signal. Figure 9-23 shows Trace B and the LED lead simulation of the EDFA output spectrum. When the LED remains off, the OSA measures only the ASE (see Trace C). NF can be calculated as the difference between Trace B and Trace C divided by Trace A, which is the LED spectrum when the laser is off and the LED is on.

Figure 9-23. Noise Gain Measured at Amplifier Output

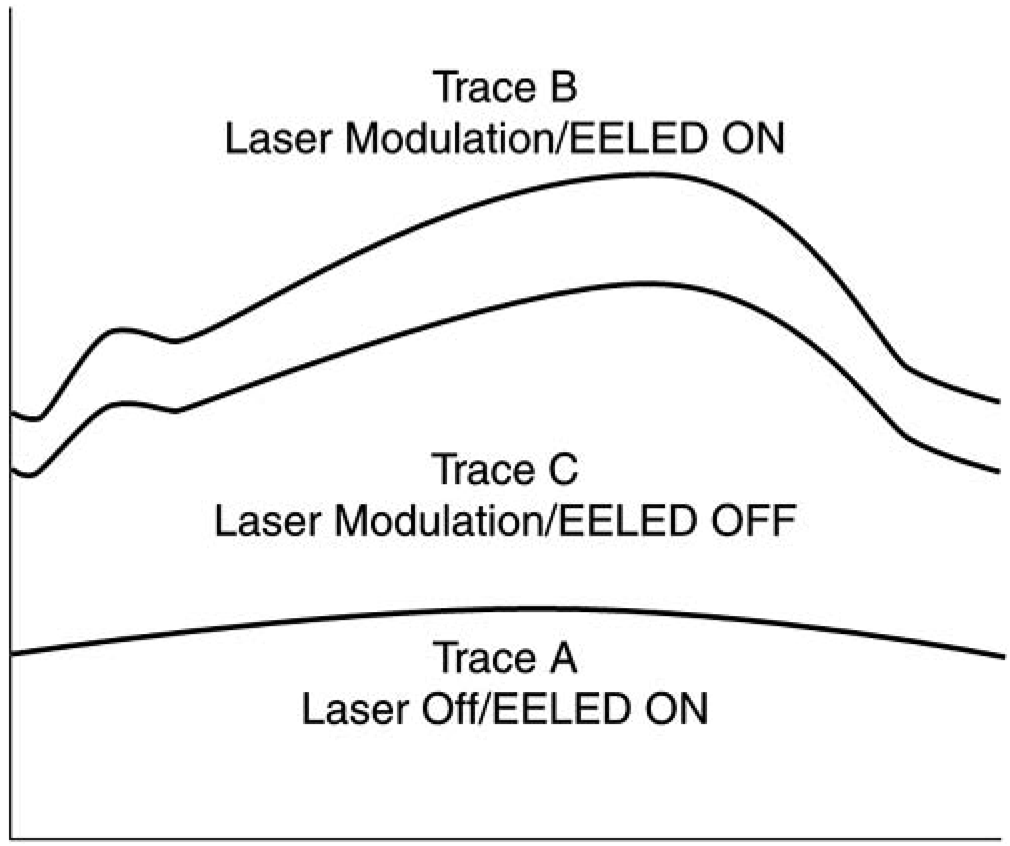
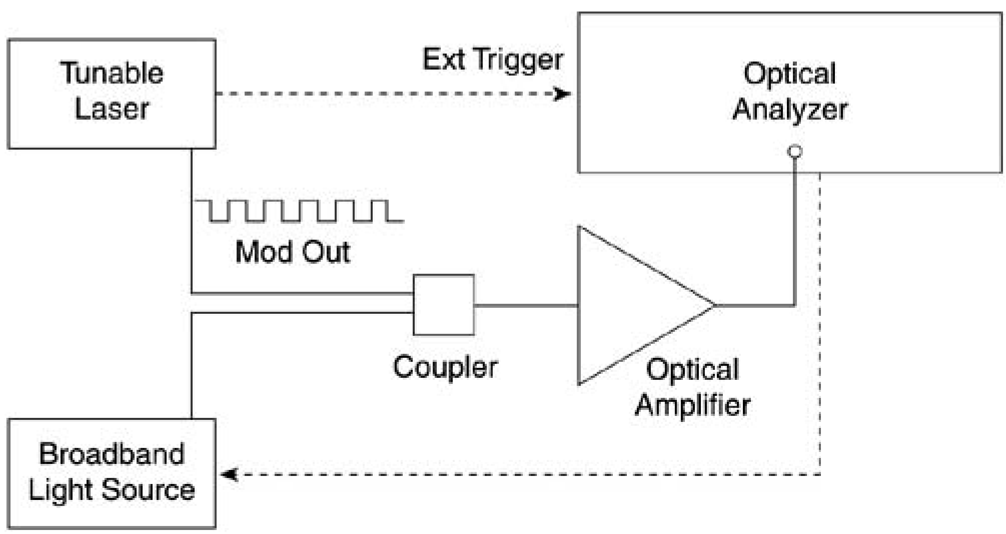


Figure 9-24. Setup for Noise Gain Profile Measurements (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)

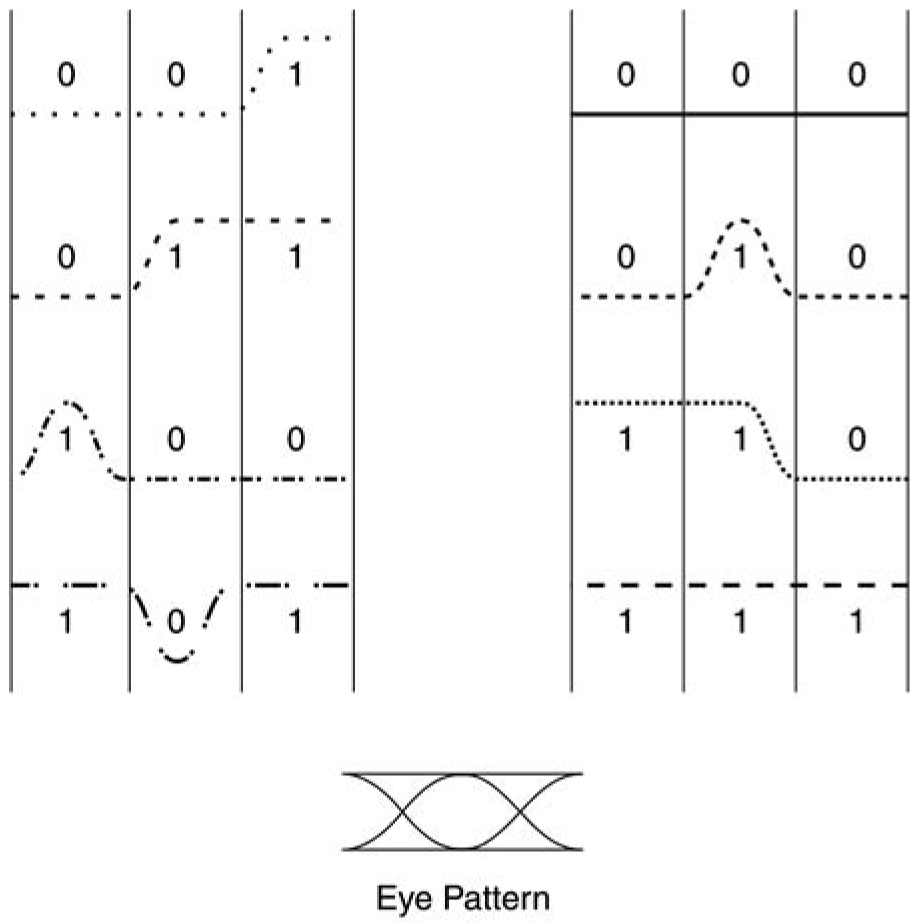


Optical Waveform Analysis

An eye pattern technique is used to analyze signals and verify that the signals have enough quality to support the communication system. This method has long been used in digital copper transmission systems to evaluate performance. The same technique is used now in optical systems to analyze the effects of waveform in fiber-optic transmission systems.

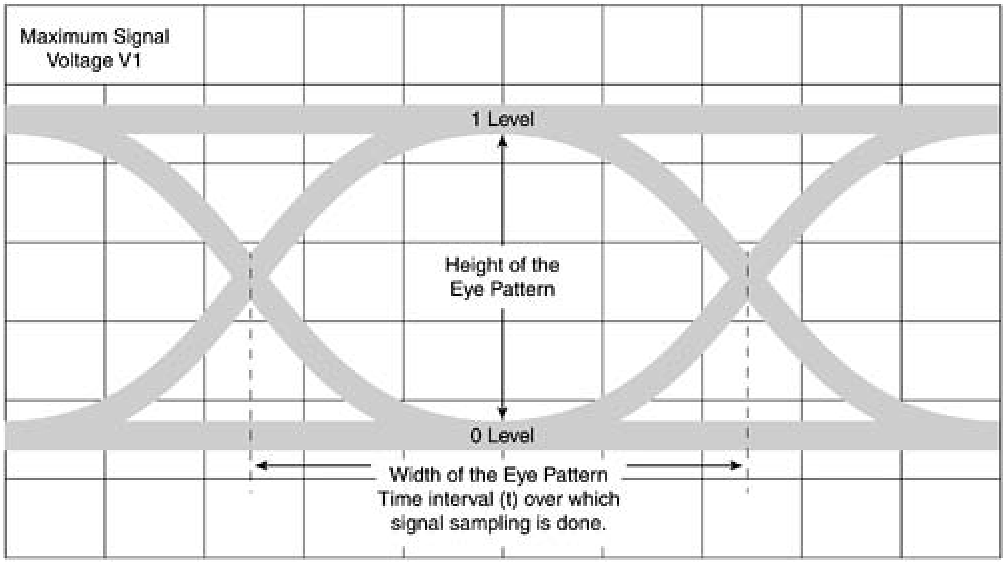
The eye diagram measurements are done in a time domain and displayed on an oscilloscope (sampling oscilloscope). The eye pattern is constructed by superimposing the various bit sequences on top of each other. Consider all eight possible combinations of 3-bit sequences: 000, 001, 010, 011, 100, 101, 110, and 111 (see Figure 9-25). If we superimpose all the combinations and display them on an oscilloscope, we get an eye diagram or eye pattern, as shown in Figure 9-25. An eye pattern gives us information about the signal distortion, pulse rise time, and jitter, and also helps us calculate extinction ratio, Q factor, and laser chirp.

Figure 9-25. Formation of Eye Diagram



An eye diagram shows a relative performance of the signal. The horizontal and vertical openings of the eye provide valuable information about the ability of the receiver to detect a 1 bit and a 0 bit correctly. Figure 9-26 shows the characteristics of the eye diagram. The vertical *eye opening* (height) shows the ability to distinguish between a 1 and 0 bit, and the width or the horizontal opening gives the time period over which the signal can be sampled without any errors.

Figure 9-26. Eye Pattern of Optical Signal as Viewed Using Sampling Oscilloscope

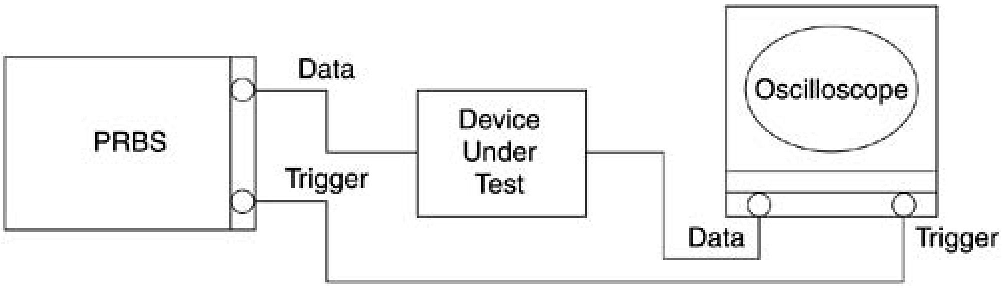


For a good transmission system, the eye opening should be as wide and open as possible. The eye diagram also displays information such as maximum signal voltage, rise and decay time of pulse, and so on.

The *extinction ratio* (ratio of a 1 signal level to a 0 signal level) and *Q factor* (ratio of peak-to-peak signal strength to total noise in electrical domain) are also calculated from eye diagrams. To determine whether a transmitter is conforming to the standards and is capable of error-free transmissions, the eye diagram is tested for correct shape (eye mask test) and proper extinction ratio.

A basic setup for measuring an eye pattern is shown in Figure 9-27. Because data is random in nature, the best approach is to generate a random sequence of data. A pseudo-random binary sequence (PRBS) generator is used to generate data for test purposes. A PRBS can generate data patterns of 2^N x 1, where N can be any integer. Typical values of N can be 7, 9, 13, and 31. After the limit is reached, the data sequence repeats.

Figure 9-27. Typical Setup for Viewing Eye Pattern



A typical setup for measuring eye pattern is shown in Figure 9-27. In this setup, the method of triggering an analyzer has an impact on how the data will be displayed. The ideal approach is to use the clock signal that is synchronized with data. Data can be used as a trigger if this clock signal is unavailable. During the jitter measurements, triggering using the clock source is more accurate than using data. Jitter is the variation in the location in the time of the rising and falling edge of the signal. In the eye diagram, this is shown as the thickness (spread) of the crossing point. (In the photonic layer, jitter accumulation is not a primary verification parameter as compared to the SONET/SDH and Gigabit [Layer 2] verifications.)

To measure optical signals and to display the eye diagram, an optical-to-electrical conversion is needed. Such a device is used to emulate the effect of an optical oscilloscope. Because the analysis is done in a wide range of frequencies, the frequency response of the instrument used to analyze this is important. A reasonable combination of frequency, bandwidth, and noise characteristics are essential for accurate display of optical signals in an oscilloscope.

Extinction Ratio

The extinction ratio is derived from eye pattern measurement. *Extinction ratio* is the ratio of a 1 signal level (power) to a 0 signal level (see Equation 9-22).

Equation 9-22

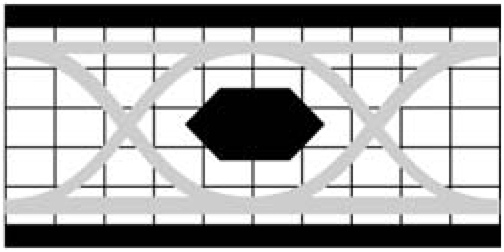
$$\text{Extinction ratio} = 10\log\left(\frac{v_1}{v_2}\right)dB$$

In the equation, v_1 is the mean of a 1 signal level, and v_2 is the mean of a 0 signal level.

Eye Mask Testing

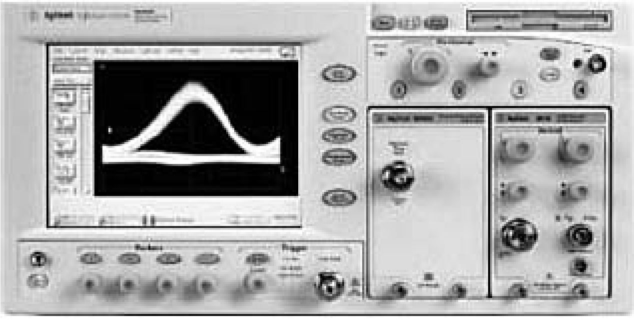
It is difficult to measure a numerical value that can describe the openness of the eye diagram. A process called *eye-mask testing* is used instead to test the openness of the eye. An *eye mask* is a collection of polygons that represents where the eye waveform cannot exist (see Figure 9-28). An eye mask consists of one polygon in the center of the eye, one polygon above, and one polygon below. The eye diagram is not allowed to cross the polygons or intersect them. A minimum acceptable opening is defined by the shape and size of the polygon in the middle.

Figure 9-28. Polygons in Eye Masking



To ensure standard results, in addition to defining the shape of the eye mask, the measurement system also specifies the frequency response of the measurement channel. This is done by defining a reference receiver that consists of a photodetector followed by a low-pass filter. The output of the combined element is fourth-order Bessel-Thomason frequency response, which is a close approximation of the response of a Gaussian filter. (Gaussian response yields minimum distortion of the waveform.) A communication analyzer from Agilent is show in Figure 9-29. It can be used for eye mask analysis and eye pattern testing.

Figure 9-29. Digital Communication Analyzer (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



BER Measurement

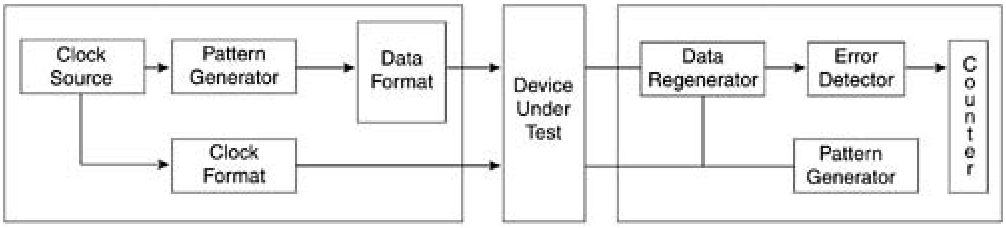
In a digital transmission system, the most fundamental performance measure is to analyze the probability of a bit that is not transmitted correctly (error). The probability is measured using BER tester or BERT. BER is defined as the number of bits received in error divided by the number of bits transferred (see Equation 9-23).

Equation 9-23

$$BER = \frac{No\ of\ bits\ received\ in\ Error}{No\ of\ bits\ transmitted} = \frac{Error\ count\ in\ Measured\ Period}{(Bit\ rate) \times (Measured\ Period)}$$

A *BERT* consists of a pattern generator and an error detector circuit (see Figure 9-30). The pattern generator generates random patterns by using a PRBS generator. It is important to use specific known patterns so that the pattern generator can easily verify whether the device under test correctly transports the data. The pattern generator also provides control over test signal characteristics such as waveform amplitude and shape.

Figure 9-30. BER Test System (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)



The error-detection circuit verifies that each bit sent through the device under test has been correctly received. The error detector and the signal pattern are synchronized and compared bit for bit. Whenever there is a discrepancy, an error is counted. Therefore, BER is determined from the number of bits found in error and the total number of bits verified.

BER results indicate the performance of the system, but they do not provide reasons for the errors or for poor performance. We can use several techniques for root cause analysis. For example, if many errors are found, we can check whether the errors are all because of 0 bit or 1. We can also observe whether the errors repeat in an interval, such as a single error after every 100 bits.

Today, optical communication specifies BER to be 10^{-12} or lower. To be confident, it is necessary to collect 1000 or more errors to see if the bit error ratio performance is achieved (epoch condition). At an error ratio of 10^{-12} and a data rate of 1 Gbps, it will take 11 and a half days to collect 1000 errors. Even if we increase the data rate and reduce the number of errors collected, it will take hours to complete the measurements.

There are other techniques that are used to determine the BER in a reduced time. One technique is to stress the signal being tested, and another technique is to adjust the threshold of the error detector to nonideal levels. One way to stress a signal is to attenuate prior to reaching the system. Here, the BER is artificially increased by attenuating the signal and diminishing the ability to distinguish a 1 bit from a 0 bit. We can plot a graph by measuring the BER at various levels of attenuation.

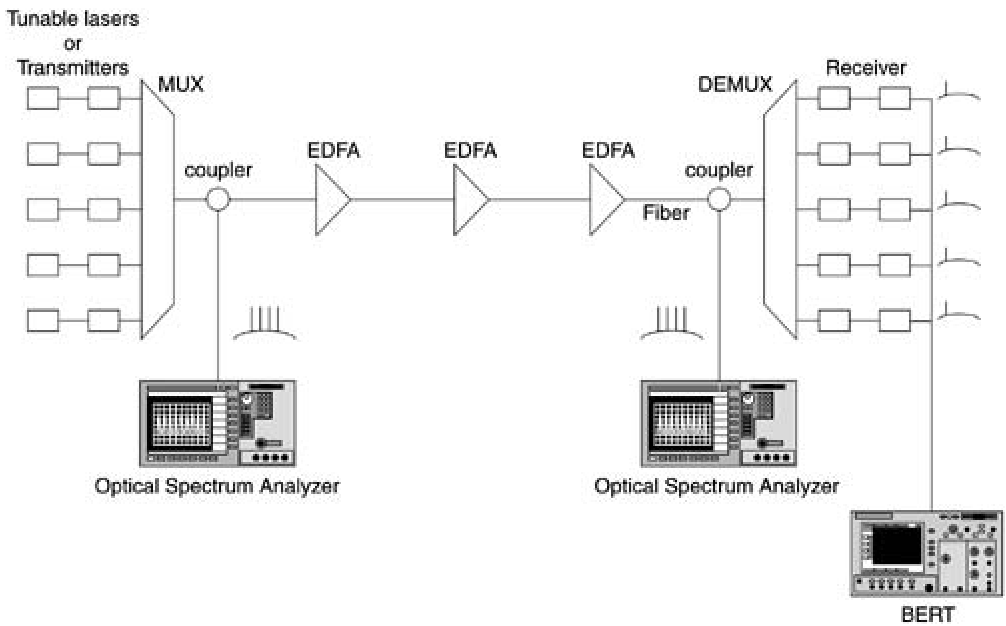
The BER increases artificially if the error detector-sampling threshold is changed (increased or decreased) from the ideal value. We perform this type of BER analysis at the test receiver input. This is the valuation of the signal quality and is called the Q factor measurement. The BER measurements are made above and below the ideal sampling threshold. For each BER result, the Q value is calculated and plotted (BER vs. Q). A typical BERT is shown in Figure 9-31 with courtesy from Agilent technologies, Inc.

Figure 9-31. Error Performance Analyzer (Reproduced with Permission, Courtesy of Agilent Technologies, Inc.)

Complete WDM End-to-End Testing

A complete point-to-point DWDM system that includes the transport units, multiplexer/demultiplexer, EDFA, and cables is shown in Figure 9-32. The testing of this system ensures that the whole system works and conforms with margin requirements (based on system design parameters) before it can carry live traffic. The deployment of the system is usually stage-by-stage; in other words, the WDM system is never fully loaded initially; rather, channels are added according to the requirement.

Figure 9-32. WDM End-to-End Testing



WDM Link Acceptance Testing

Optical spectrum analyzers are employed to analyze each wavelength that is used in the system. A power meter can also be used in this case, depending on the level of analysis. A bit analyzer is used to analyze errors in the system.

WDM Provisioning Test

The following procedure simulates the provisioning exercise to verify whether the system is behaving according to the specifications:

- Add a channel.
- Check to ensure that all existing channels indicated zero errors when the channel was added.

- Measure the tilt, power level, and OSNR using an OSA as each channel is configured.

Repeat this process until all wavelengths are operational.

WDM Stress Test

The following procedure simulates the WDM stress test to verify whether the system is behaving according to the specifications:

- Start the BER test.
- Check the power, tilt, and OSNR using the OSA.
- Check the BER performance

Repeat until the system margins have been verified (BER vs. OSNR).

Summary

In this chapter, we outlined the need for testing and the methods to measure different parameters in a generic optical network. The major issues such as attenuation, power measurements, dispersion, PMD measurements, and BER were extensively discussed in this chapter.

This chapter also showcased WDM end-to-end testing. We discussed WDM component testing for both passive and active components. We also explained testing methodologies, the underlying theories, and the technologies involved. Special attention was paid to dispersion measurement using a Poincare' sphere because dispersion is a big hurdle in deployment of high bit rate (10 Gbps and beyond) long-haul (1000 km and beyond) WDM transmission links.

The basis of all considerations is the standard from ITU-G.692, which defines testing and measuring benchmark parameters. Much of the equipment used for tests and measurements comes from Agilent, who helped us prepare for the test scenarios. Most of the instruments are generic, but for better understanding, we worked closely with Agilent to provide more of the details. Emphasis was given to optical spectrum analyzers (OSAs) and power meters, which are essential for measurements in WDM networks.

References

[1]Poole, C.D., J. H. Winters, and J.A. Nagel. "Dynamical Equation for Polarization Dispersion." *Optics Letters*. Vol. 16, No 6. March 1991.

[2]Agilent Technologies. "Measuring Extinction Ratio of Optical Transmitters." Application note 1550-8.

[3]Agilent Technologies. "Lightwave Test and Measurement." 2002 Catalog.

[4]Agilent Technologies. "Optical Spectrum Analyzer." Application note 1550-8.

[5]www.agilent.com/comms/lightwave.

[6]Agilent Technologies. "Frequency Agile Jitter Measurements System." Application note 1267.

[7]Agilent Technologies. "EDFA Testing with the Interpolation Technique." Product note 71452-1.

[8]Agilent Technologies. "8614xB Optical Spectrum Analyzer Family." Technical specifications.

[9]Agilent Technologies. "Measuring Extinction Ratio of Optical Transmitters." Application note 1550-8.

[10]Agilent Technologies. "Optical Spectrum Analysis." Application note 1550-4.

[11]Agilent Technologies. "86100B Wide-Bandwidth Oscilloscope." Technical specifications.

Chapter 10. Simulations of WDM Systems

This chapter discusses the basic aspects of simulation in WDM networks. It is intended to add practical impetus to the theories explained in the rest of the text. Much of this chapter focuses on running simulation setups and exercises based on VPI software. Many of the design cases in Chapter 4, "WDM Network Design -1" and Chapter 5, "WDM Network Design -2," are generically simulated in this chapter. The authors would like to thank Elizabeth P. Morgan (VPIphotonics) for coordinating this project. Special thanks to Arthur Lowery (for authoring the body of this chapter) and Peter Wildhagen (for creating the simulation demos), both from VPIphotonics.

Need for Simulation

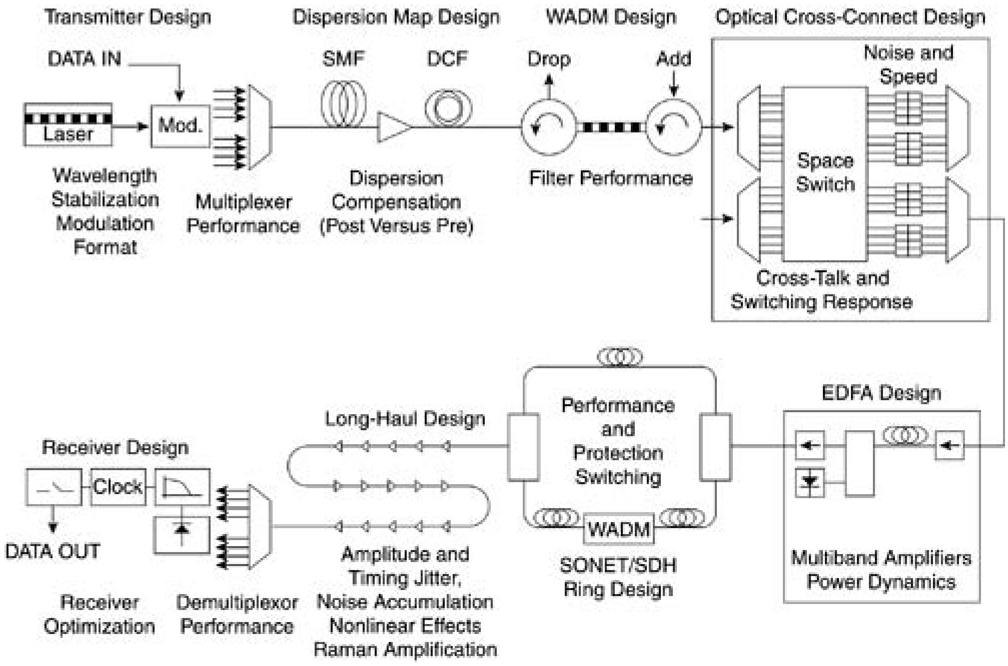
When a network design is sketched-out for the first time, many issues might require further optimization. One of the most important tools for designing and optimizing networks is simulation. Simulation uses interconnected mathematical models of components to predict the performance of a system as a whole. By incorporating a simulation within an optimization loop, the system's performance can be optimized for a particular scenario. This may mean reducing CapEx, or designing to meet future upgrades without major reconfiguration of the network.

The simulation of optical fiber links has only become commonplace over the past five years[1]. This is partly because researchers, who calculated conservative design rules that guaranteed the operation of a system in a well-defined set of circumstances, initially developed many commercial systems. In addition, a single manufacturer supplied the whole link—from transmitter, through fiber, to receiver—to guarantee performance through thorough testing. Of course, simulations of performance have been performed to design specialized links, such as submarine systems, for far longer. These systems were costly to prototype and operated as state of the art; saving design effort and lowering risk by simulations was well worth the effort.

With the advent of WDM systems, manufacturers began using commercial simulation tools to develop new products, to exchange details of components and systems electronically, and to make proposals to service providers. This boost in the commercial tool market was initially driven by the nonlinearity of optical fibers[2], which, although small, causes significant cross-talk along a long-haul link, where optical amplifiers keep the power levels high.

However, as the performance of systems increased and a need arose to reduce costs particularly for access systems, many new problems could be identified and solved using simulation. For example, increased channel densities (closer spacing of WDM channels) required optical demultiplexers with well-defined amplitude and phase characteristics, advanced modulation techniques, and optical channel-monitoring schemes. Figure 10-1 illustrates some of the technologies that can be optimized by simulation with a system.

Figure 10-1. Simulation Technologies



This chapter shows how practical WDM systems can be designed using computer simulation. It also introduces the design software from VPIphotonics provided with this book. The version provided is a standalone tool—that is, it does not include the project and data management features of the VPI's Design and Deployment Center. The commercial release can be obtained by contactingsales@VPIphotonics.com.

Inside the CD-ROM in This Book

VPIphotonics has created applications demos for its WDM systems simulator, VPItransmissionMakerWDM, to go with this book. These cover metro ring and long-haul design issues. A brief description of each demo is provided in the Case Studies section of this chapter. However, the best way to understand these applications is to run them and explore the results and parameter settings.

The demos can be run by first installing VPItransmissionMakerWDM. First, complete the supplied End User License Agreement and fax this to the U.S. headquarters of VPIphotonics. After you receive confirmation from VPI, install the software and e-mail your license application file to CISCOwdmDemo@VPIphotonics.com. The VPI team will send you a 30-day demo license file.

Note that the software requires the following: Windows XP, Microsoft NT 4.0 with Service Pack 6a, or Microsoft Windows 2000 with Service Pack 2; at least 300 MB of hard disk space and 256 MB of RAM and an installed and configured TCP/IP Protocol. On computers without a network card, you must install Microsoft Loop Back Adaptor or Dialup Networking first.

After VPItransmissionMaker WDM has been installed and started, you can import the specific demos described in the Case Studies section. These reside in a CISCO WDM Demos folder on the CD-ROM. You need to use File > Import... to select one of these demos. After the demo's schematic appears, press the Green-man and then OK in the Submit Simulation job window to run the demo.

VPItransmissionMaker also includes more than 200 standard demos showing how to simulate and characterize devices, amplifiers, and systems. These are available from the Demos tab pane at the top left of the GUI. You can open a demo by double-clicking on it. The GUI *User's Manual* provides detailed instructions on how to create your own simulations in its "Getting Started" chapter.

NOTE

The enclosed software, VPItransmissionMaker, shows just some of VPIphotonics' capabilities in designing for the physical layer. VPI provides tools covering all aspects of optimization and economic design of networks, and specific tools for active components and optical amplifiers.

Standalone Versus Collaborative Design Tools

Standalone photonic computer-aided design tools cannot support the management of design data among the many complementary-skilled engineers necessary to design a complex system. Therefore, regardless of the best efforts of the individual who is using the tool, the overall design will be suboptimal. For this reason, a design environment needs to be capable of the following tasks:

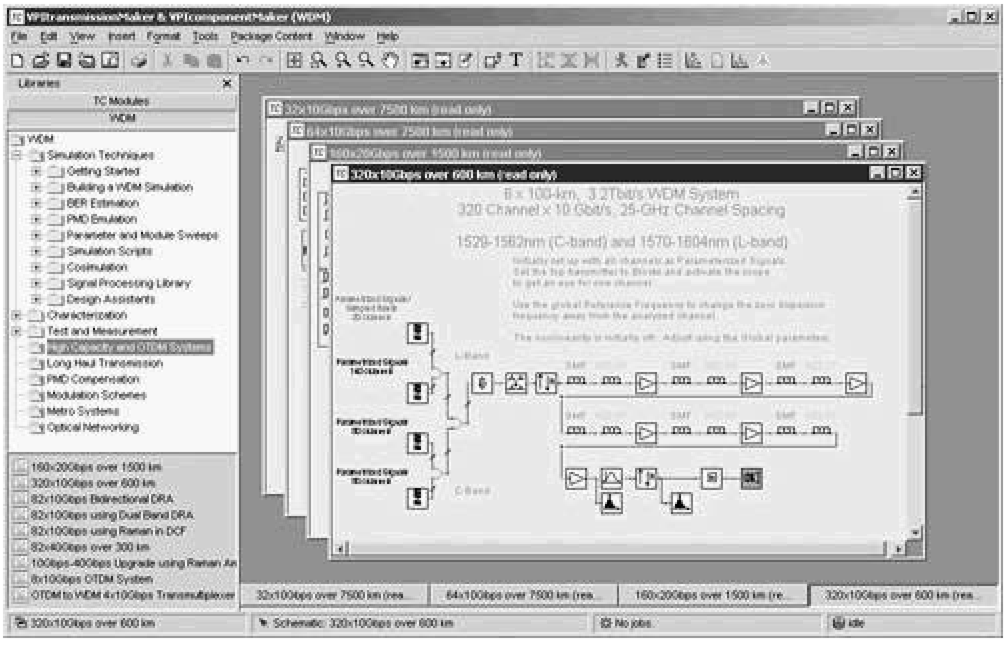
- Organizing project tasks into work packages
- Assigning teams to the tasks and allowing team members to check out and check in work packages (with notification to other team members)
- Linking component designs to systems designs for automatic updating of a systems design (always state-of-the-art, if required)
- Providing a wide range of templates for component characterization, systems design, and systems testing
- Allowing every stage of a design to be revisited and allowing important decisions to be documented along with the design stage
- Capturing all results, documentation, and specifications that are associated with a design

VPI provides a set of software and lifecycle management tools that meet the above requirements. One tool, VPItransmissionMaker is provided with this book to illustrate the concepts of earlier chapters.

VPItransmissionMaker Simulation Technology

VPItransmissionMaker is designed for almost all physical layer simulation tasks. This flexibility is possible because of the graphical user interface, shown in Figure 10-2. This software has libraries of modules that you can place on a schematic and link together to fully describe an optical system or network. Each module has a set of parameters to describe its physical characteristics, although you can also use behavioral parameters. The hierarchy is supported by having subsystem modules that each have a schematic to describe their contents.

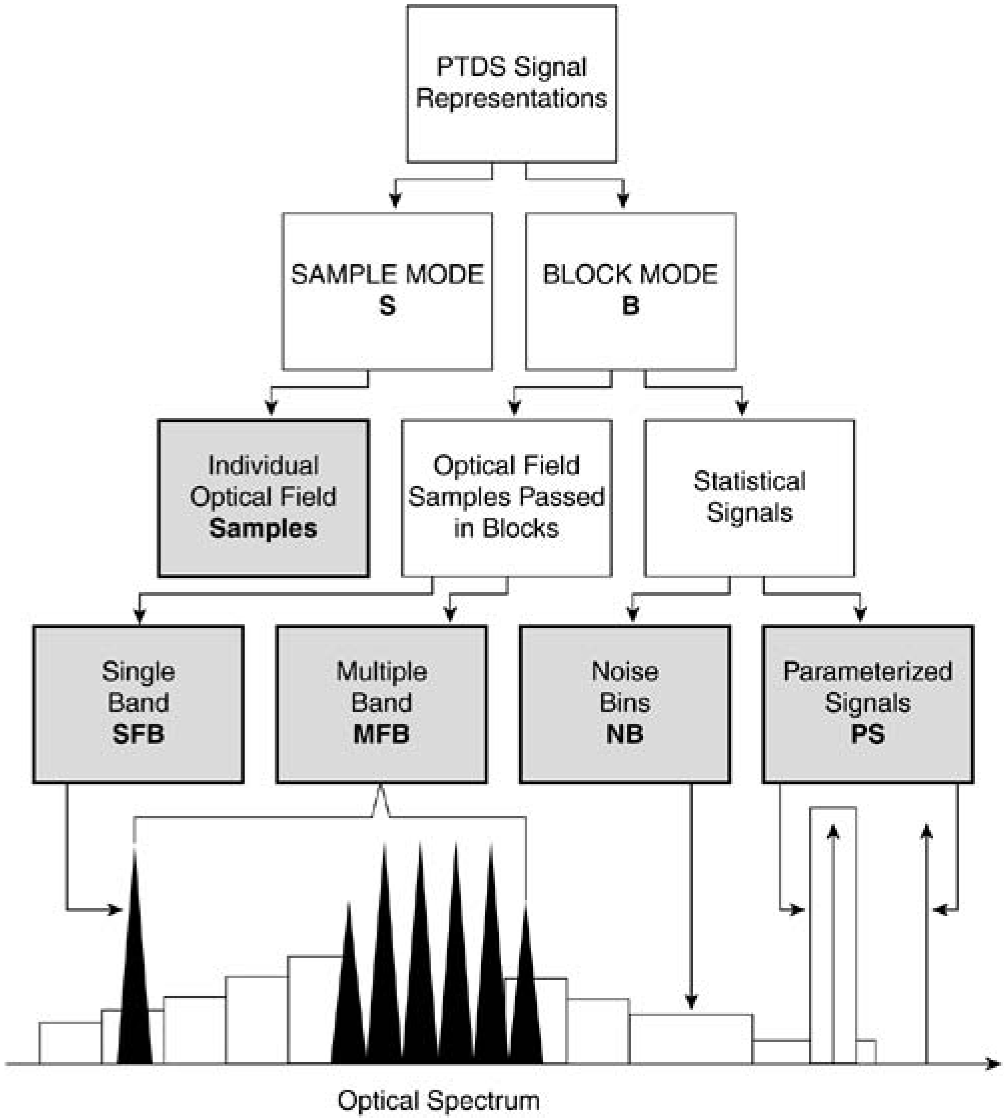
Figure 10-2. Graphical Interface of VPItransmissionMaker Showing a Typical Simulation Schematic



Photonic simulation covers many levels of scale, from subwavelength active regions and gratings in active devices to global networks. Every level of scale is important because a wavelength-scale phase shift in an optical device can have as much effect on a system as a hundred kilometers of fiber.

VPItransmissionMaker has been designed with the ability to cover many orders of scale by using multiple signal representations[3], shown in Figure 10-3. These go from detailed (representing time-sampled optical fields in two polarizations) to individual sampled bands that only cover areas of interest (multiple frequency bands) to statistical signals for noise (noise bins) and signals (parameterized signals). These signal representations can be mapped onto a design problem spectrally (each part of an optical spectrum is represented by the most appropriate signal representation) and spatially (each section of a system is represented by the most appropriate signal representation). Details of how to use signal representations can be found on the CD-ROM in the *WDMUser's Manual*. However, all demos' signal representations are already set up.

Figure 10-3. Spectral Mapping of Signal Representations



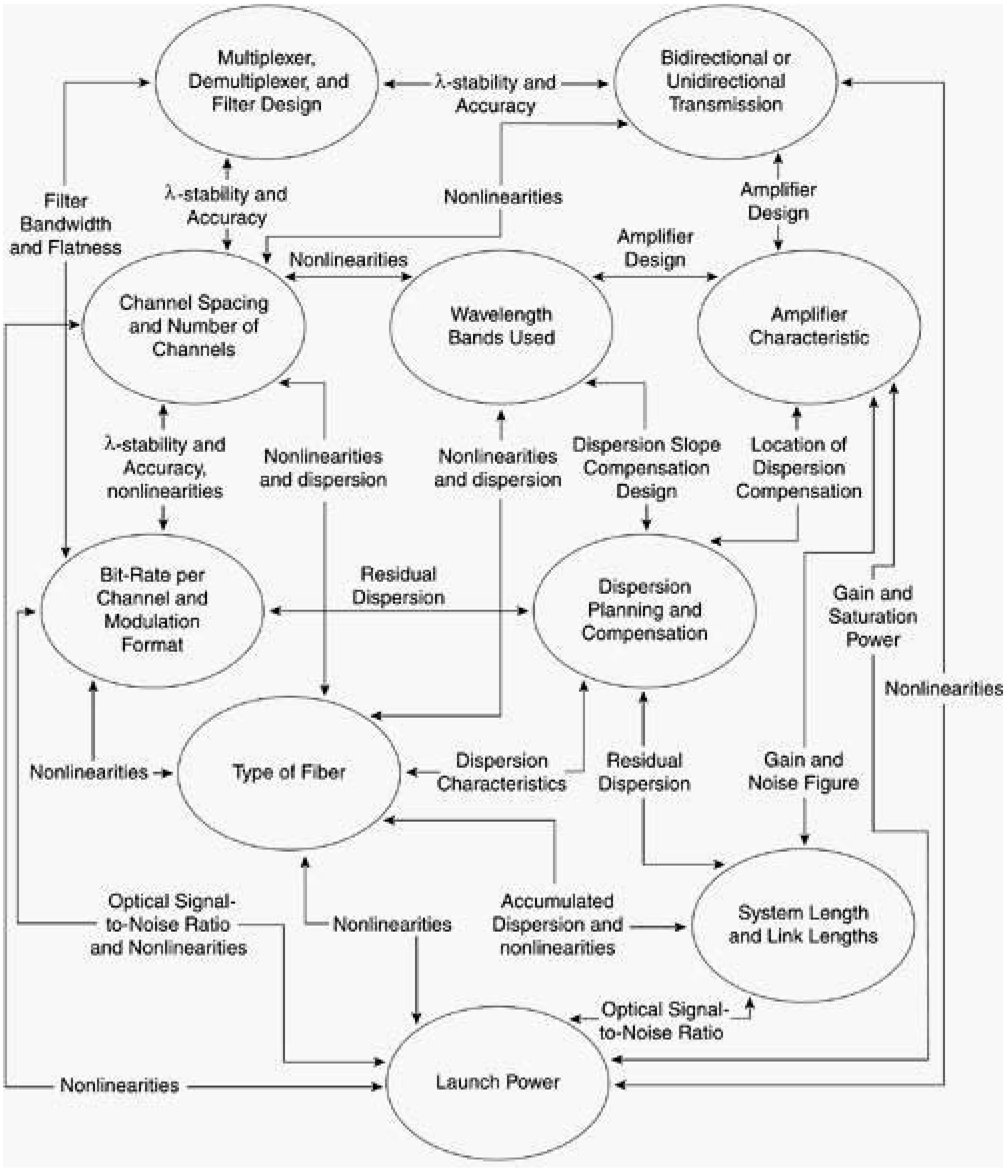
Typical Physical-Layer Design Processes Using VPItransmissionMaker

Current-generation optical point-to-point systems are extremely complex. A wide range of options is available to the product design team at all layers of the network hierarchy, and the introduction of new component technologies brings large changes in systems design philosophy. For example, multiple band (L, C, S) optical amplifiers can multiply the capacity of systems[4]. Similarly, high-efficiency modulation schemes and well-designed demultiplexers can pack more bit(s) into each hertz of optical bandwidth.

Many interrelations exist among system design variables. The difficulty of managing the complex interactions among all aspects of the design has in the past led to over-engineering of many optical transmission systems. This, in turn, has meant exceeding the minimum required specifications in all subsystems to provide sufficient margin to guarantee the performance of the overall system over its lifetime. Such over-engineering necessarily incurs costs, including the initial increased cost of deployment compared to the optimum system and the loss of revenue that could be realized if the system were used to its full potential.

Figure 10-4 shows some of the interconnected decisions that need to be made when designing a link. Obviously, many complex multivalued optimizations must be performed. The following sections show some examples of design refinement that can be performed using VPItransmissionMaker. The CD-ROM provides more than 200 examples, and the possibilities are infinite, given the permutations of hundreds of modules and parameters that can be created. You can edit the design examples and create different scenarios by varying network parameters.

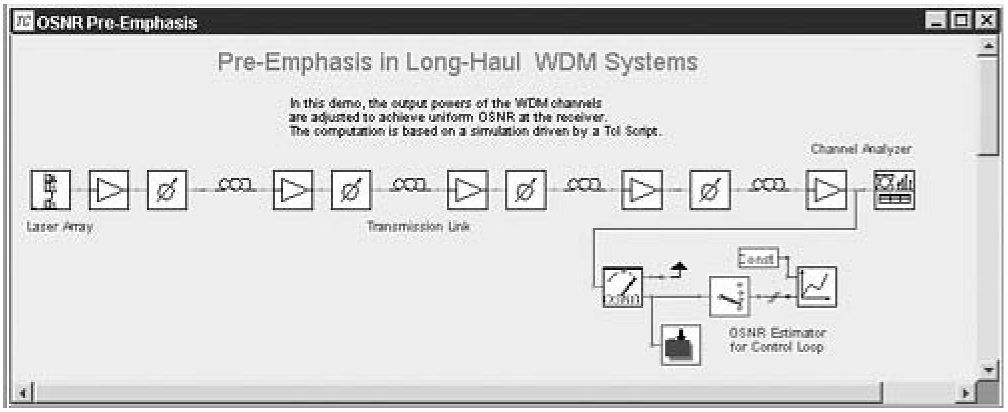
Figure 10-4. Some Interactions (Lines) Between Design Variables (Ovals) in a Systems Design Illustrating the Complexity of a Systems Design



Signal-to-Noise Simulation

The optimum information-carrying performance of a long-haul saturated amplifier link is obtained when the signal-to-noise ratios are equalized over all the channels. Figure 10-5 shows the schematic of a 16-channel WDM system with a chain of six amplifiers, three sections of dispersion-compensating fibers (DCFs), and two spans of single-mode fiber (SMF). The amplifiers have no gain equalization, so they suffer from a large spectral ripple. The object of this simulation is to optimize the input spectrum of the chain so that each WDM channel has the same optical signal-to-noise ratio (OSNR) at the output of the chain. This gives the maximum information capacity for the link, but it is difficult to calculate because the gain spectrum of the amplifiers depends on their input powers. Therefore, a self-consistent solution must be found iteratively.

Figure 10-5. Schematic of a Link Optimization for OSNR Equalization



For efficiency, parameterized signals are used to represent the mean power in a WDM channel over a data sequence and noise bins can be used to represent the noise in and around each channel. The converged output spectrum after five iterations is shown in Figure 10-6. This demonstrates a constant signal-to-noise ratio (SNR) for all the channels. (The parameterized signals are equal ratios above the noise bins.) Note that the widths of the noise bins have been automatically reduced around the ASE peak to maintain amplitude accuracy. This feature is designed to increase efficiency by optimizing the number of noise bins that cover the spectrum. The convergence process is shown in Figure 10-7.

Figure 10-6. Spectrum After Equalization Showing Signals as Arrows and Noise as Bins of Equal Power Spectral Density (Note Bin Width Reduction Around ASE Peak)

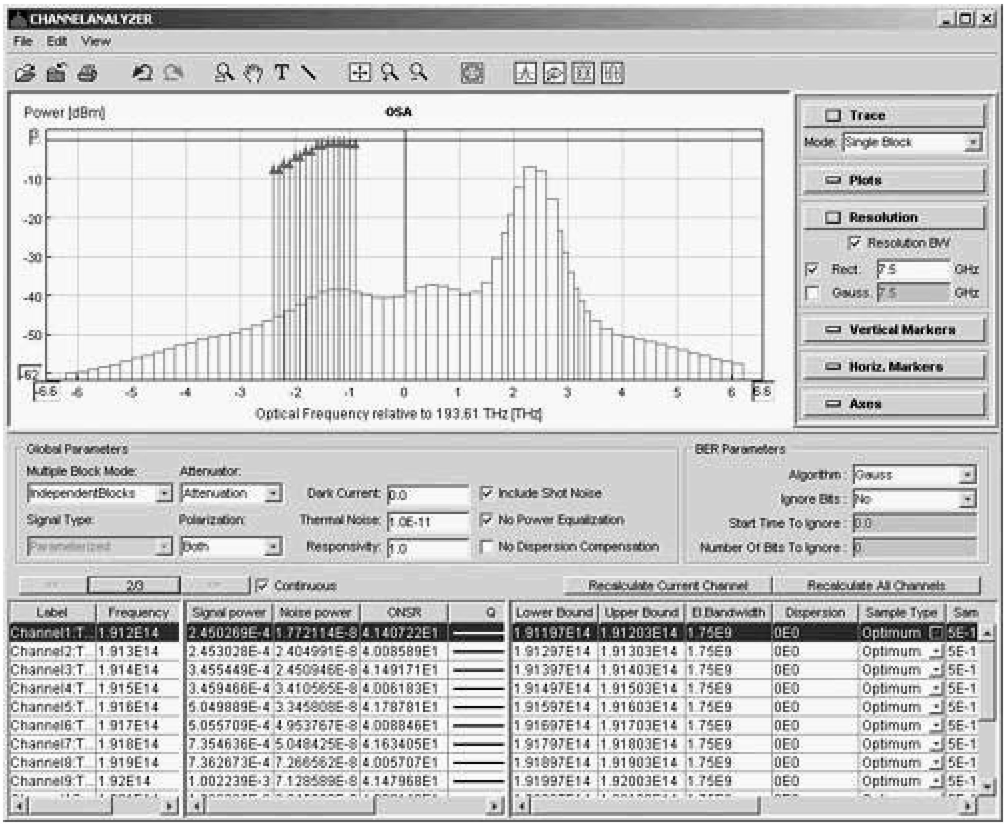
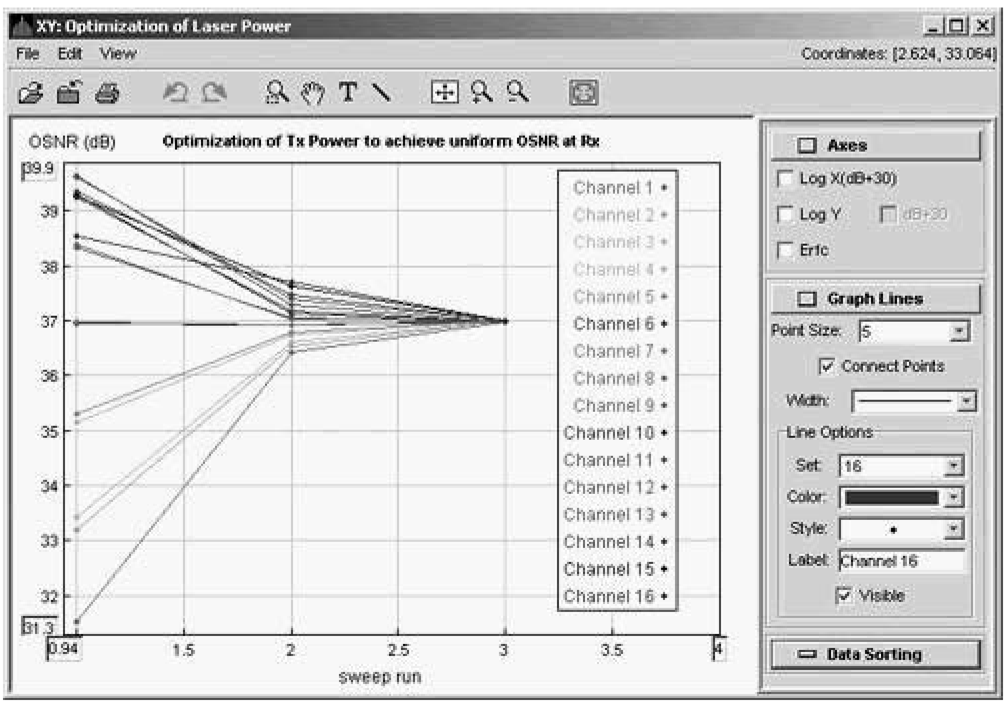


Figure 10-7. Iteration of Input Powers to Achieve Equal SNRs. At Iteration 2, the SNRs have crossed over because the Spectral gain of an EDFA is a function of input power.

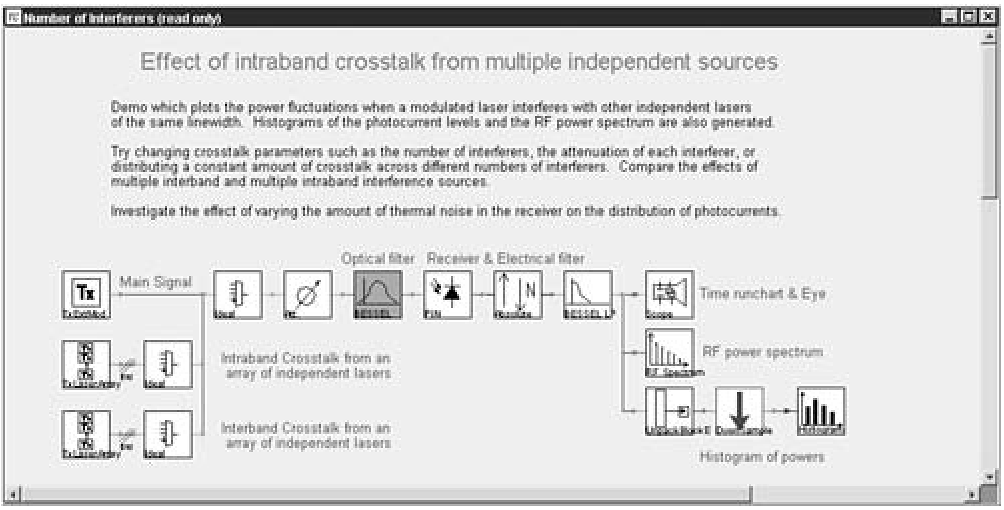


Optical Cross-Talk

You can use photonic simulation to assess the performance of optical cross-connects (switches) within systems. Of particular interest is optical crosstalk, which can severely limit the number of optical interconnects in a system[5]. Many different technologies can be compared, including blocking, nonblocking, and wavelength converting (using cross-gain, cross-phase, four-wave mixing, and opto-electronic technologies). VPI provides templates for many forms of optical switches, including wavelength-converting switches. However, the hierarchical nature of the user interface allows almost any form of switch to be defined and investigated.

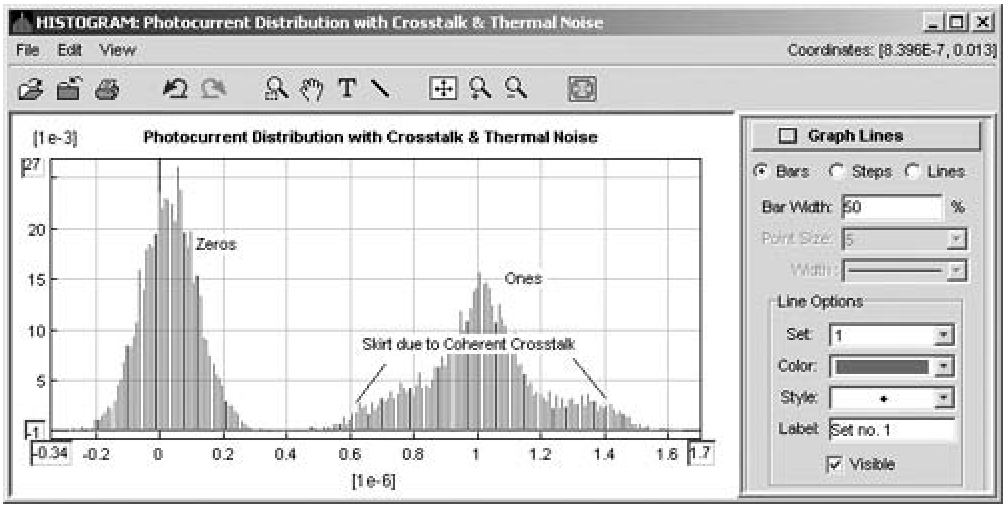
Figure 10-8 shows a simulation that illustrates and quantifies the effects of optical cross-talk. The severity of the crosstalk depends on whether the interfering carrier lies within the signal bandwidth of the desired signal. If it does, the cross-talk produces a strong beat signal that lies within the electrical bandwidth of the receiver. The interference is often called *coherent* because its severity depends on the addition of the optical fields rather than the optical powers.

Figure 10-8. Simulation to Assess Coherent and Incoherent Cross-Talk Effects



The histogram of Figure 10-9 shows the electrical signal levels at the sample time of the receiver. For a low BER, the distributions of the 1s and 0s should be well separated. The coherent cross-talk causes a large spread in the 1-bit signal levels because the fields for 1 bits are higher than those for 0 bits. Thermal noise and incoherent cross-talk cause the spread in the 1 bits.

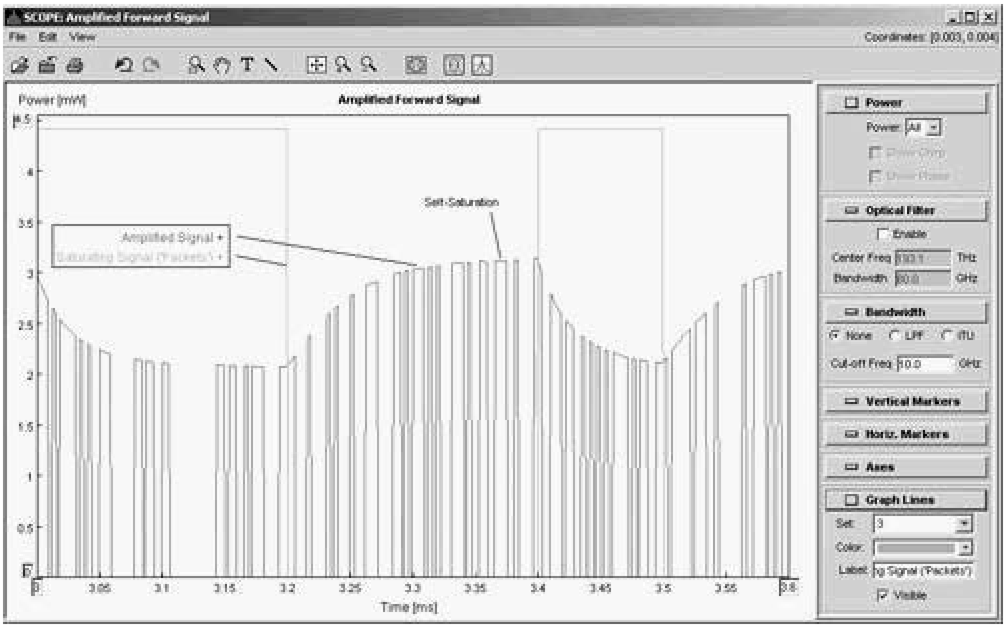
Figure 10-9. Probability Distribution of Powers in the 1 and 0 Bits. The skirts on the 1s distribution indicate coherent (or homodyne) optical cross-talk.



Millisecond Power Transient Analysis

In a system that has optical amplifiers, you must take care to minimize interactions between channels due to the gain of an amplifier changing with input power and the input power changing during channel switching or fault conditions[6]. With Erbium doped fiber amplifiers (EDFAs), the transients will be on a millisecond time scale (the lifetime of the excited Erbium ions). This scale would be much shorter if we considered chains of amplifiers. Furthermore, control circuits around the amplifiers also affect their dynamics as well as the dynamics of chains of amplifiers. Automatic gain control (AGC) and automatic balancing control (ABC) are two examples of circuits that improve the performance of an amplified system. Figure 10-10 shows the cross-talk between two channels. The saturating signal is modulated on a millisecond timescale. (This could be carrying an optical packet.) The saturating signal reduces the amplifier's gain and modulates the second channel (amplified signal). The second channel also modulates the gain of the amplifier, producing small gain reductions within each bit.

Figure 10-10. Cross-Gain Modulation in an Erbium-Doped Fiber Amplifier Causes Cross-Talk Between Channels on a Millisecond Timescale



Polarization Mode Dispersion

Polarization mode dispersion (PMD) affects the following systems (a) older fiber that has high PMD coefficients running at 10 Gbps (b) 40 Gbps systems running on low PMD fiber[7]. PMD is a random phenomenon. The severity of PMD varies with time due to the polarization states along the fiber changing due to temperature and stress. Certain alignments of the polarization states can cause severe eye closure over short intervals. Figure 10-11 shows a histogram of BERs obtained by simulating a fiber 250 times. (Each simulation could represent a given time of day.) A typical eye diagram showing distortion due to PMD is shown in Figure 10-12. Simulation can be used to assess the effects of fiber PMD on system performance and to identify solutions. The solutions might include using novel modulation formats, such as RZ pulses, multilevel coding running at a lower symbol rate, or PMD compensators.

Figure 10-11. Distribution of BERs in a 40 Gbps Fiber System Showing the Random Nature of PMD

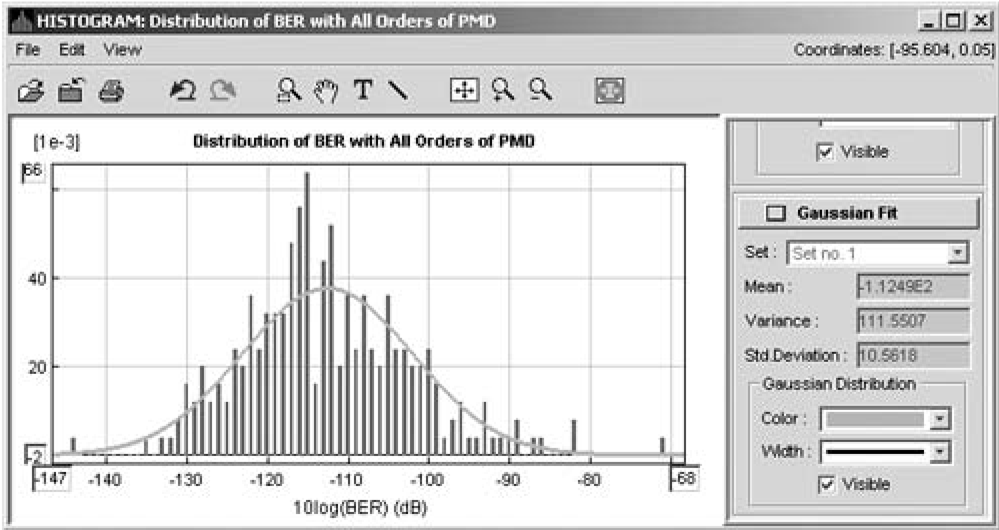
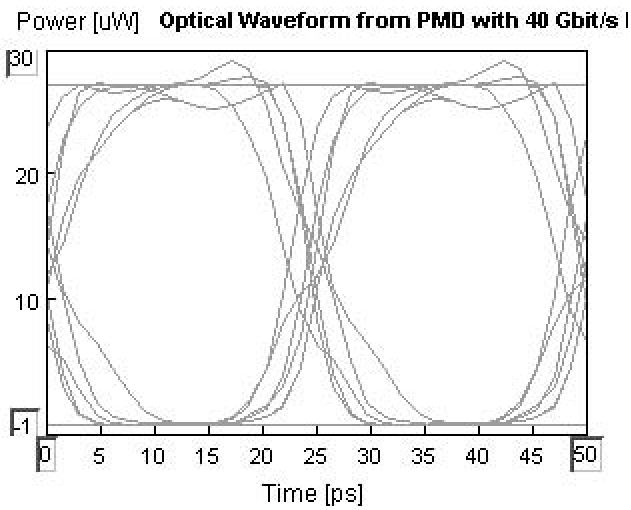


Figure 10-12. Eye Diagram Distorted by PMD. The spreading and timing Jitter is associated with first-order PMD, whereas the Peaking is associated with higher-order PMD.

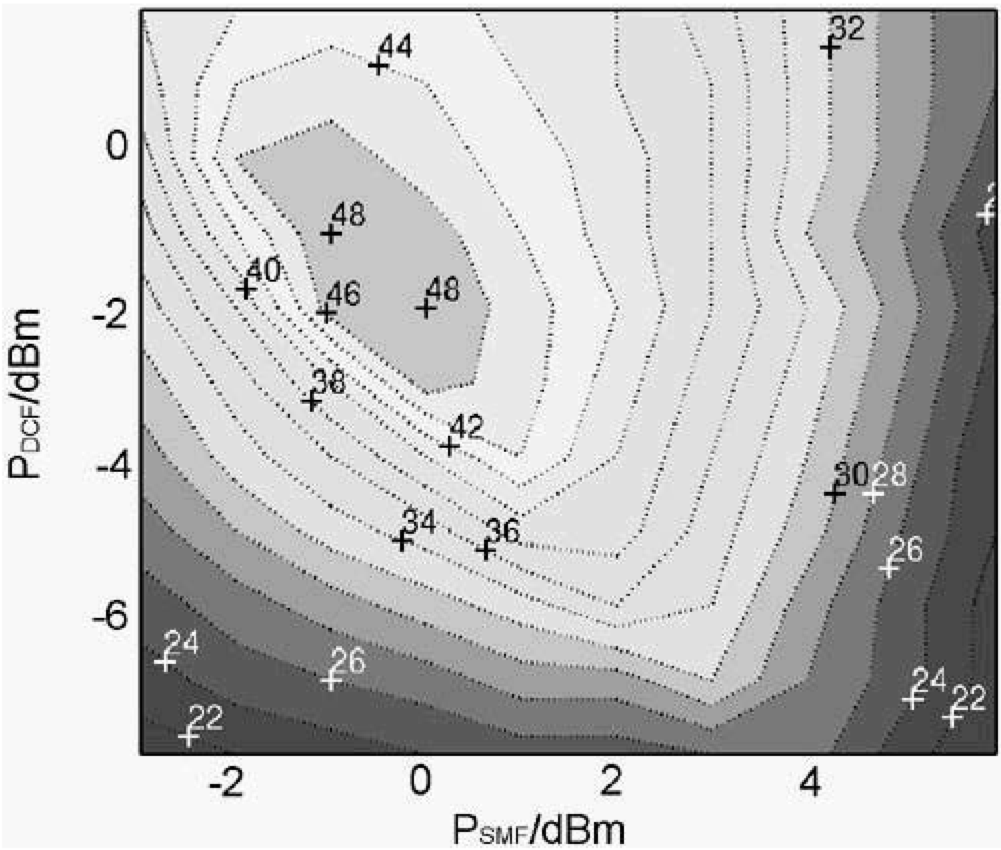


Maximizing Transmission Length

VPItransmissionMaker uses sampled optical signals (sampled optical fields) for nonlinearities, dispersion (fiber and filter), mixing of noise and signal, and polarization effects. Typical simulation output includes optical spectra, eye diagrams, and BER estimation. Photonic subsystems such as wavelength conversion and optical regeneration can be included in such simulations, and these can show many possible impairments.

As an example, Figure 10-13 shows a typical result of a link design in which the task is to optimize the amplifier powers to obtain the maximum number of spans before the signal degrades to unacceptable levels. The task is complicated because of the nonlinearities in the fiber coupled to the dispersion in the fiber[6]. This means that the nonlinearity induces a power-dependent phase shift on the pulse, following the pulse envelope, which is translated to a change in pulse shape due to the effect of fiber dispersion. In other words, a certain amount of nonlinearity is useful in sharpening pulses. However, in link designs that use lengths of opposite-dispersion fiber (to gain an overall dispersion close to 0), the powers into the fiber lengths are critical. This presents a complex optimization problem.

Figure 10-13. Maximizing the Number of Spans in a Long-Haul System by Optimizing Optical Powers into Nonlinear Fibers with Standard (SMF) and Dispersion-Compensating (DCF) Characteristics



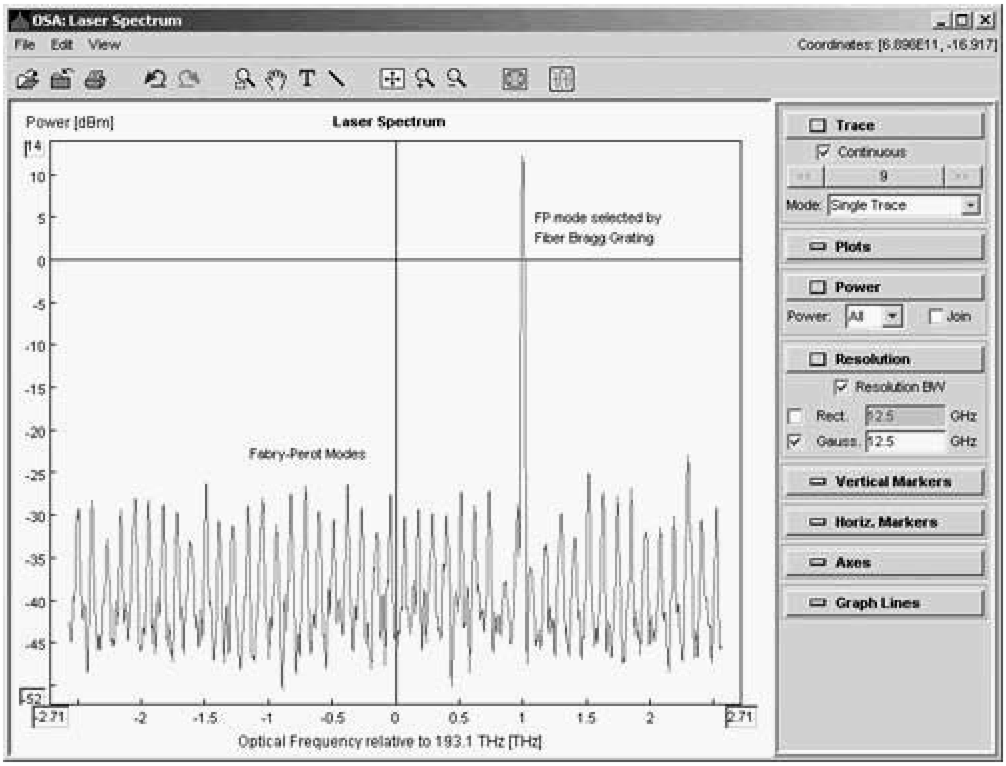
Photonic Component Design

Active photonic devices will perform many complex functions in future photonic networks, from providing stable as well as tunable single-mode wavelength references, to photonic switching and signal processing, to regeneration and wavelength conversion. VPIcomponentMaker Active Photonics provides advanced, wide-spectrum dynamic models of active devices, which are connected to model sources and photonic circuits bidirectionally. In this way, multisection and externally stabilized lasers can be constructed from elementary modules of active regions and passive resonators[9]. Both the optical and electrical interfaces are bidirectional, allowing sophisticated functionality to be designed, such as laser amplifier-detectors, injection-locked lasers, all-optical regenerators, clock recovery and optical logic circuits, and active and passive mode-locking of integrated and external cavity devices.

The active models are based on an enhanced transmission line laser model, which simulates the full longitudinal and spectral dynamics of advanced lasers and semiconductor optical amplifiers. The enhancements allow lasing to be tuned over the entire simulation bandwidth with a single parameter without loss of accuracy or spectral artifacts.

A typical design problem is shown in Figure 10-14. This is a fiber Bragg grating stabilized laser that is designed to produce temperature-stable wavelength due to the low temperature coefficient of the Bragg grating compared to the semiconductor laser. The spectrum shows a single-mode output, but with close-in side modes due to the long lasing cavity. The design is sensitive to the optical phase between the grating and the front facet of the laser. A typical design question is how good the antireflection coating on the laser facet has to be to achieve a high side-mode suppression ratio. In Chapter 2, "Networking with DWDM -1," we discussed the creation of single-mode lasers and using a filter (in this case, FBG) to ensure the presence of only the dominant mode at a given wavelength.

Figure 10-14. VPIcomponentMaker Model of a Bragg Grating Stabilized Laser and Its Predicted Spectrum


OSA: Laser Spectrum
File Edit View
Coordinates: [6.898E11 , -16.917]
Power [dBm]
Laser Spectrum
FP mode selected by
Fiber Bragg Grating
Fabry-Perot Modes
Optical Frequency relative to 193.1 THz [THz]
Trace
Continuous
9
Mode: Single Trace
Plots
Power
Power: All
Join
Resolution
Resolution BW
Rect. 12.5 GHz
Gauss. 12.5 GHz
Vertical Markers
Horiz. Markers
Axes
Graph Lines

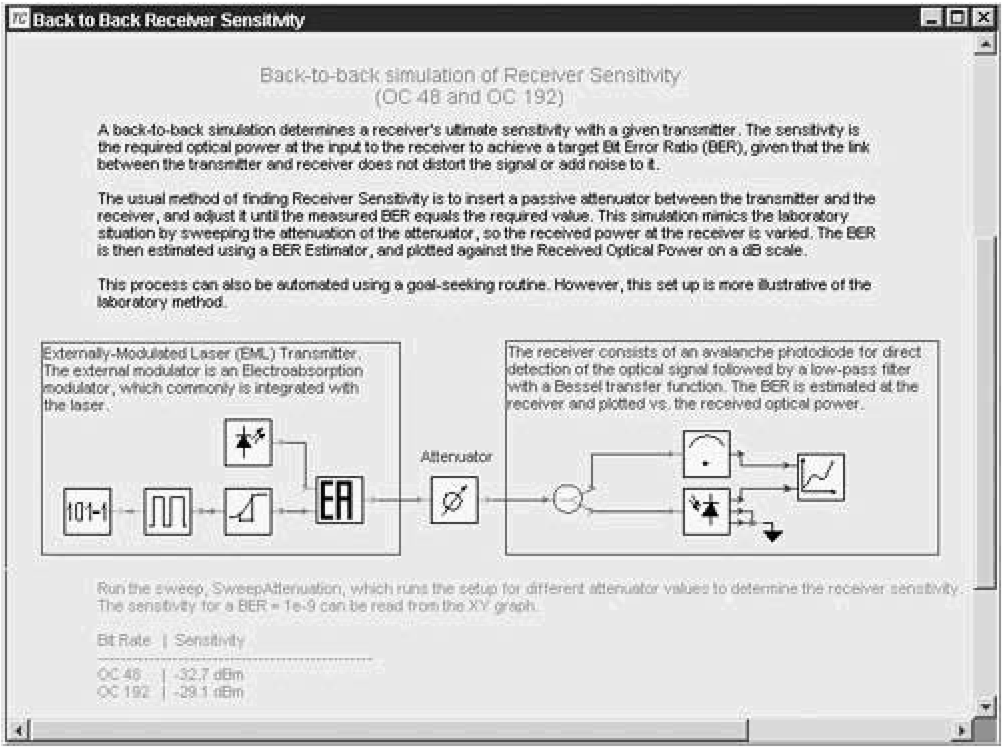
Case Studies

The following five demos have been especially created for this book. They illustrate physical-layer design issues peculiar to optical access rings. These demos reside in a CISCO WDM Demos folder on the CD-ROM. You need to use File > Import... to bring up the schematic for the demo.

Back-to-Back Receiver Sensitivity

A back-to-back simulation determines a receiver's ultimate sensitivity with a given transmitter. The sensitivity is the required optical power at the input to the receiver to achieve a target bit error ratio (BER), given that the link between the transmitter and receiver does not distort the signal or add noise to it. This demo (see Figure 10-15) is used to set the baseline for the performance of the metro core ring demos that follow.

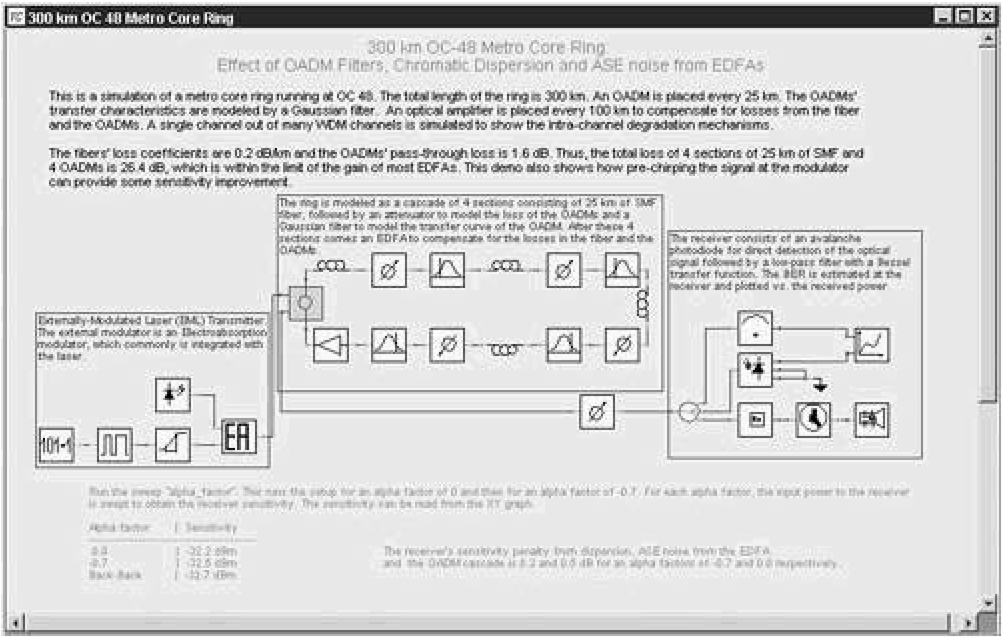
Figure 10-15. Schematic of the Simulation to Determine Receiver Sensitivity



300 km OC-48 Metro Core Ring

This is a simulation of a metro core ring running at OC 48 (see Figure 10-16). The total length of the ring is 300 km, and an OADM is placed every 25 km. The OADMs' transfer characteristics are modeled by a Gaussian filter. An optical amplifier is placed every 100 km to compensate for losses from the fiber and the OADMs. A single channel out of many WDM channels is simulated to show the intrachannel degradation mechanisms.

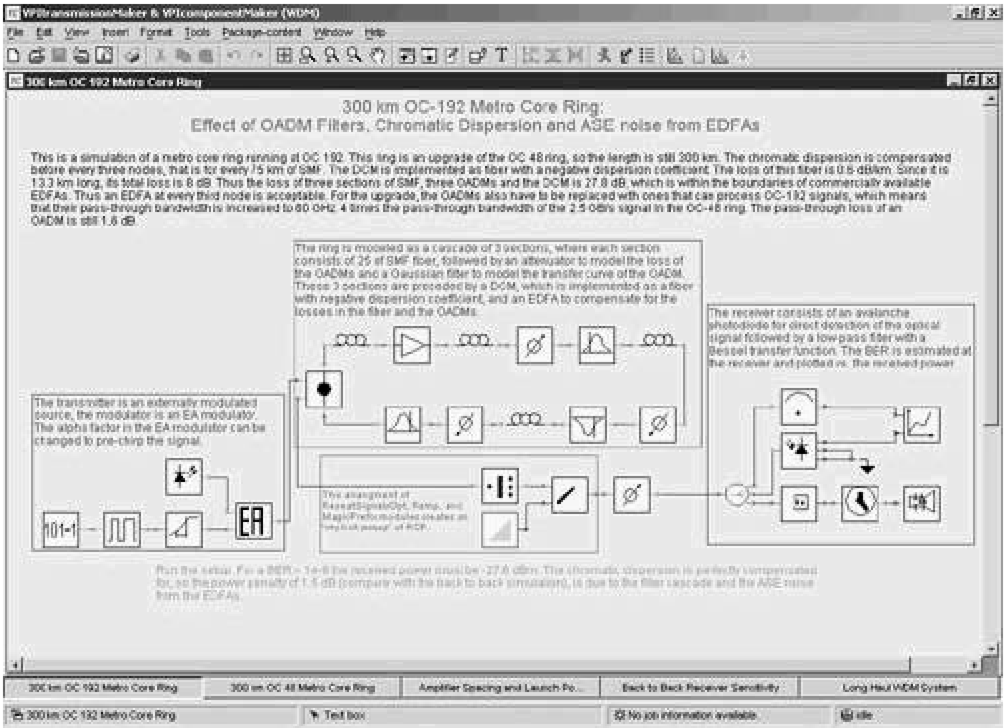
Figure 10-16. Simulation of a Metro Ring Running at 2.5 Gbps



300 km OC-192 Metro Core Ring

This is a simulation of a metro core ring running at OC 192 (see Figure 10-17). This ring is an upgrade of the OC 48 ring to keep the length at 300 km. The chromatic dispersion is precompensated for every three nodes (that is, before every 75 km of SMF). The dispersion compensation module (DCM) is implemented as a spool of fiber with a negative dispersion coefficient. The loss of this fiber is 0.6 dB/km and because it is 13.3 km long, its total loss is 8 dB. Therefore, the loss of three sections of SMF, three OADMs, and the DCM is 27.8 dB, which is within the boundaries of commercially available EDFAs. In other words, an EDFA at every third node is acceptable. For the upgrade, the OADMs have to be replaced with ones that can process OC-192 signals, which means that their passthrough bandwidth is increased to 80 GHz—four times the passthrough bandwidth of the 2.5 Gbps signal in the OC-48 ring.

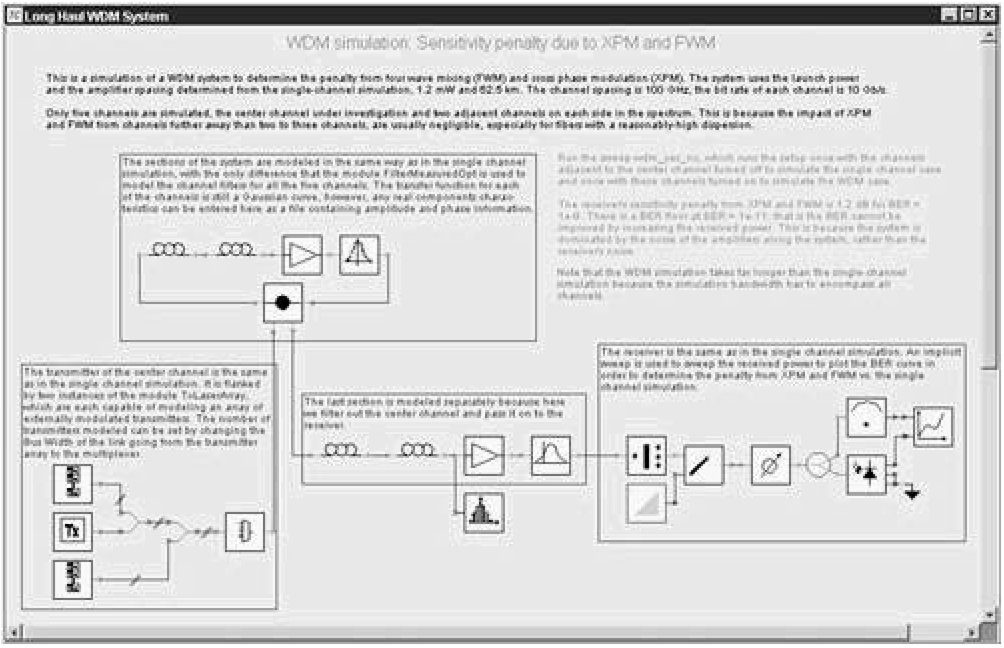
Figure 10-17. Simulation of a Metro Ring Running at OC-192 Rate



Long-Haul WDM System

This is a simulation of a WDM system to determine the penalty from four-wave mixing (FWM) and cross-phase modulation (XPM) by comparing a single-channel system to a WDM system (see Figure 10-18). The system uses a mean launch power of 1.2 mW and the amplifier spacing determined from the single-channel simulation, that is 62.5 km. The channel spacing is 100 GHz, and the bit rate of each channel is 10 Gbps. Only five channels are simulated: the center channel under investigation and two adjacent channels on each side in the spectrum. FWM and XPM effects for channels further than three-channel spacing are negligible, depending on the fiber plant[10].

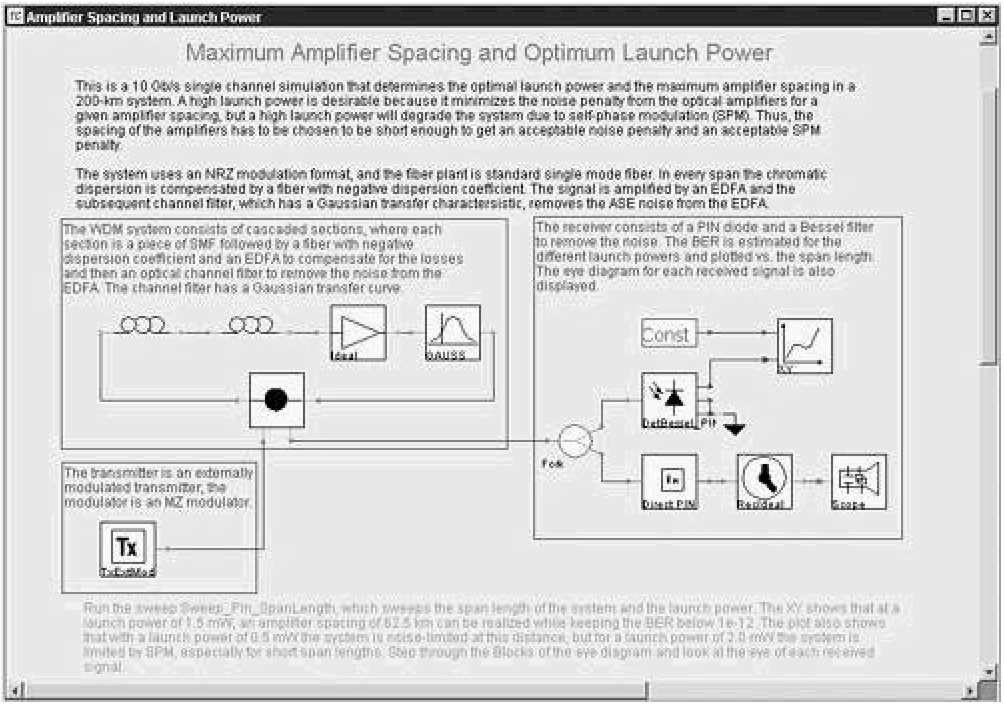
Figure 10-18. Simulation of a Long-Haul WDM System to Investigate FWM and XPM Effects



Amplifier Spacing and Launch Power

This is a 10 Gbps single-channel simulation that determines the optimal launch power and the maximum amplifier spacing in a 200 km system (see Figure 10-19). A high launch power is desirable because it minimizes the noise penalty from the optical amplifiers for a given amplifier spacing, but it does degrade the system due to self-phase modulation (SPM). Therefore, you must choose the spacing of the amplifiers to be short enough to get an acceptable noise penalty and an acceptable SPM penalty.

Figure 10-19. Simulation of a Long-Haul WDM System to Determine Optimum Launch Power and Amplifier Spacing



Amplifier Spacing and Launch Power
Maximum Amplifier Spacing and Optimum Launch Power
This is a 10 Gb/s single channel simulation that determines the optimal launch power and the maximum amplifier spacing in a 200-km system. A high launch power is desirable because it minimizes the noise penalty from the optical amplifiers for a given amplifier spacing, but a high launch power will degrade the system due to self-phase modulation (SPM). Thus, the spacing of the amplifiers has to be chosen to be short enough to get an acceptable noise penalty and an acceptable SPM penalty.
The system uses an NRZ modulation format, and the fiber plant is standard single mode fiber. In every span the chromatic dispersion is compensated by a fiber with negative dispersion coefficient. The signal is amplified by an EDFA and the subsequent channel filter, which has a Gaussian transfer charactersistic, removes the ASE noise from the EDFA.
The WDM system consists of cascaded sections, where each section is a piece of SMF followed by a fiber with negative dispersion coefficient and an EDFA to compensate for the losses and then an optical channel filter to remove the noise from the EDFA. The channel filter has a Gaussian transfer curve.
Ideal
GAUSS
The receiver consists of a PIN diode and a Bessel filter to remove the noise. The BER is estimated for the different launch powers and plotted vs. the span length. The eye diagram for each received signal is also displayed.
Const
XY
DetBessel
Plt
Fork
Direct PIN
RecIdeal
Scope
The transmitter is an externally modulated transmitter, the modulator is an MZ modulator.
Tx
TxExtMod
Run the sweep Sweep_Pin_SpanLength, which sweeps the span length of the system and the launch power. The XY shows that at a launch power of 1.5 mW, an amplifier spacing of 62.5 km can be realized while keeping the BER below 1e-12. The plot also shows that with a launch power of 0.5 mW the system is noise-limited at this distance, but for a launch power of 2.0 mW the system is limited by SPM, especially for short span lengths. Step through the Blocks of the eye diagram and look at the eye of each received signal.

Summary

This chapter introduced the world of simulation that is being employed to develop new technologies and ensure that they are effectively used in telecommunications networks. To ensure that the latest advances in components are used when designing networks and to speed innovation by connecting the component and systems design processes, VPIphotonics developed the VPI Design and Deployment Center, which connects simulation tools to design teams and ensures that both are using the most current data.

The enclosed CD-ROM contains a full version of VPItransmissionMaker, which includes a wealth of simulation examples. Together with the design examples illustrated in this chapter, these demos show some of the wide range of applications that VPItransmissionMaker can be put to. We hope you enjoy using this software.

References

[1]Lowery, A.J. "Photonic Simulation Tools." *Optical Fiber Telecommunications IVB: Systems and Impairments*. Ed. Ivan Kaminow and Tingye Li. San Diego, CA: Academic Press; 2002. 564–610.

[2]Feature Issue on "Fundamental Challenges in Ultrahigh-Capacity Optical Fiber Communications Systems."*IEEE J. Quantum Electron*. Vol. 34 (11). Nov. 1998. 2053–2103.

[3]Lowery, Arthur, et al. "Multiple Signal Representation Simulation of Photonic Devices, Systems, and Networks." *IEEE J. Selected Areas in Quantum Electronics*. Vol. 6 (2). March/April 2000. 282–296.

[4]Srivastava, Atul K., and Yan Sun. "Advances in Erbium-Doped Fiber Amplifiers." Ed. Ivan Kaminow and Tingye Li. *Optical Fiber Telecommunications IVA: Components*. San Diego: Academic Press; 2002. 174–212.

[5]Ramaswami, Rajiv, and Kumar N. Sivarajan. *Optical Networks: A Practical Perspective*, Second Edition. San Francisco, CA: Morgan Kaufmann Publishers, Academic Press, San Diego; 2002. 299–308.

[6]Bononi, A. and L. A. Rush. "Doped-Fiber Amplifier Dynamics: A System Perspective." *J. Lightwave Technology*. Vol. 16 (5). 1998. 945–956.

[7]Kogelnik, Herwig, Robert M. Jopson, and Lynn E. Nelson. "Polarization Mode Dispersion." Ed. Ivan Kaminow and Tingye Li. *Optical Fiber Telecommunications IVB: Systems and Impairments*. San Diego, CA: Academic Press; 2002. 725–861.

[8]Essiambre, René-Jean, Gregory Raybon, and Benny Mikkelsen. "Pseudo-Linear Transmission of High-Speed TDM Signals: 40 Gbit/s and 160 Gbit/s." Ed. Ivan Kaminow and Tingye Li. *Optical Fiber Telecommunications IVB: Systems and Impairments*. San Diego, CA: Academic Press; 2002. 232–304.

[9]Lowery, A. J. "Semiconductor device and lightwave system performance modeling" in *Tech. Digest of Optical Fiber Communications 1997 (OFC'97)*, paper ThG5, OSA Tech. Digest Series, vol. 6, 267-268 (OSA, Washington D.C., 1997).

[10]Yu, T., et al. "A Mean-Field Approach for Simulating Wavelength Division Multiplexed systems." *IEEE Photonics Technology Letters*. Vol. 12 (4). April; 2000. 443–445.

Bibliography

Much of this book was generated from the experience of the authors in the optical networking industry. However, much of the text was inspired from the following works. Some of them are an excellent source for further information if you are interested in developing your understanding of the subject. Most of the works are authored either by industry leaders in technology or by leading academicians.

Agrawal, G.P. *Fiber-Optic Communication Systems*. Wiley Interscience, 1992.

______.*Non-Linear Fiber Optics*, Second Edition. Academic Press, 1995.

______.*Fiber-Optic Communication Systems*. Wiley Interscience, 1992.

Bala, K. et al. *Benefits of Minimum Wavelength Interchange in WDM Rings*. Proceedings of OFC, 1997.

Barry, R. and P. Humblet. "Models of Blocking Probability in All-Optical Networks With and Without Wavelength Changers."*IEEE Journal on Selected Areas in Communications*. Vol. 14 (5), June 1996.

Bertsekas, D. and R. Gallagher. *Data Communication Networks*. McGraw-Hill, 1996.

Black, U. *MPLS and Label Switching Networks*. Prentice Hall, 2000.

Bononi, A. *Optical Networking*. London, England: Springer Publishers, 1999.

Bononi, A. and L.A. Rush. "Doped-Fiber Amplifier Dynamics: A System Perspective."*J. Lightwave Technology*. Vol. 16 (5). 1998. 945–956.

Bouda, M. et al. *Tunable Lasers*. OFC 2001. Anaheim, CA.

Brackett, C. "Dense Wavelength Division Multiplexing Networks: Principles and Applications." *IEEE Journal On Selected Areas in Communications*. Vol. 8 (6). August 1990. 948–964.

______.*Transaction on Communications*. May 1990.

Chlamtac, I., A. Ganz, and G. Karmi. "Lightpath Communications: An Approach to High Bandwidth Optical WANs."*IEEE Transactions on Communications*. Vol. 40 (7). July 1992. 1171–1182.

Chlamtac, I. et al. "Lightnets: Topologies for High-Speed Optical Networks."*Journal of Lightwave*

Technology. Vol. 11 (5/6). May/June 1993.

______."CORD: Contention Resolution Using Delay Lines."*Journal on Selected Areas in Communication*. June 1996. Special Issue on Optical Networks.

______."Scalable WDM Network Architecture Based on Photonic Slot Routing and Switched Delay Lines." IEEE Proc of INFOCOM. 1997.

Chlamtac, I., A. Ganz, and G. Karmi. *Lightpath Communications: An Approach to High Bandwidth Optical WANs. IEEE Transactions on Communications* Vol. 40 (7). July 1992. 1171–1182.

Chow, Ming-Chwan. *Understanding SONET/SDH Standards and Applications*. Holmdel, NJ: Andan Publisher; 1996.

Cisco Systems, Inc. Official Web Site. www.cisco.com.

Crovella, M. and A. Bestavros. *IEEE/ACM Transactions on Networking*. Vol. 5. 1997. 836–846.

Essiambre, René-Jean, Gregory Raybon, and Benny Mikkelsen. "Pseudo-Linear Transmission of High-Speed TDM Signals: 40 Gbit/s and 160 Gbit/s."*Optical Fiber Telecommunications IVB: Systems and Impairments*. Ed. I. Kaminow and Tingye Li. San Diego, CA: Academic Press; 2002. 232–304.

Feature Issue on "Fundamental Challenges in Ultrahigh-Capacity Optical Fiber Communications Systems."*IEEE J. Quantum Electron*. Vol. 34 (11). Nov. 1998. 2053–2103.

Fukuda, I. *Optical Semiconductor Devices*. New York, NY: Wiley; 1990.

Gibbons, A. *Algorithmic Graph Theory*. Cambridge, England: Cambridge University Press; 1985.

Goodman, J.W. *Physics of Optoelectronic Devices*. Wiley; 1995.

Green, P. *Fiber Optic Networks*. Prentice Hall, 1993.

Gumaste, A. and J. Jue. "A Scheduling Procedure for Control Signaling in an Optical Burst Switched Network." First International Conference on Optical Communication and Networks. Singapore, 2002.

Gumaste, A. et al. "BITCA Bifurcated Interconnection to Traffic and Channel Assignment in WDM Rings." Proceedings of OFC. Anaheim, CA. 2002.

Gumaste, A. "TIONETS: Triangular Interconnection of Optical Networks."*IEEE Communications Letters*. Vol. 5 (2). 83–85.

IEEE 802.3ae Standard. http://grouper.ieee.org/groups/802/3/ae/criteria.pdf.

IEEE Communications Magazine. Special Issue on Optical Networks. April 2002.

IEEE Journal of Lightwave Technology. April 1987.

IEEE Quantum Photonics Letters issue. March 1984.

Internetworking Technologies Handbook, Second Edition. Indianapolis, IN: Cisco Press; 1998.

ITU Standard Document GR 652.

ITU Standard Document GR 653.

Journal of Lightwave Technology. Special Issue on Optical Technology and Multiwavelength Networks. June 1993.

Kazovsky, L. et al. Hornet Architectures OCRL Laboratory Web Site. University of Stanford. www.wdm.stanford.edu.

Keiser, G. *Optical Fiber Communications*. New York, NY: McGraw-Hill; 1996.

Kinoshita, S. Proceedings of OAA Torino, Italy, 2001.

Klienrock, L. *Queuing Systems*, Vol. 1. Canada: John Wiley and Sons; 1975.

Kogelnik, Herwig, Robert M. Jopson, and Lynn E. Nelson. "Polarization Mode Dispersion."*Optical Fiber Telecommunications IVB: Systems and Impairments*. Ed. Ivan Kaminow and Tingye Li. San Diego, CA: Academic Press; 2002. 725–861.

Kovacevic, M. and A. Acampora. "Electronic Wavelength Translation in Optical Networks." *Journal of Lightwave Technology*. Vol. 14 (6). June 1996.

Kramer, Glen and Gerry Pesavento. "Ethernet Passive Optical Network (EPON): Building a Next-Generation Optical Access."*IEEE Communications Magazine*. Vol. 40 (2). February 2002. 58.

Kressel, H. and J.K. Butler. *Semiconductor Lasers and Heterojunction LEDs*. San Diego, CA: Academic Press; 1977.

Krishnaswamy, R. and K. Sivarajan. "Design of Logical Topologies: A Linear Formulation for Wavelength Routed Optical Networks with No Wavelength Changers."*IEEE/ACM Transactions on Networking*. Vol. 9 (2). April 2001.

Lee, T., ed. *Current Trends in Optical Amplifiers and Their Applications*. World Scientific Publishers, 1999.

Leland, W.E. et al. "On the Self-Similar Nature of Ethernet Traffic" (Extended Version). *IEEE/ACM Transactions on Networking*. Vol. 2. 1994. 1–15.

Liu, M. *Principles and Applications of Optical Communications*. Kluwer Publications, 1996.

Lowery, A.J. "Photonic Simulation Tools."*Optical Fiber Telecommunications IVB: Systems and Impairments*. Ed. Ivan Kaminow and Tingye Li. San Diego, CA: Academic Press, 2002. 564–610.

Lowery, A.J. et al. "Multiple Signal Representation Simulation of Photonic Devices, Systems, and Networks."*IEEE Journal on Selected Areas in Quantum Electronics*. Vol. 6 (2). March/April 2000. 282–296.

______."Time-Domain Simulation of Photonic Devices, Circuits, and Systems." Lasers and Integrated Devices Symposium, Photonics West '96. San Jose, CA: Jan./Feb. 1996. 624–635.

Lu, K., G. Xiao, and I. Chlamtac. "Blocking Analysis of Dynamic Lightpath Establishment of Wavelength Routed Networks." Proceedings of the IEEE ICC. 2002.

Mukherjee, B. *Optical Communications Networks*. New York, NY: McGraw-Hill; 1997.

Mukherjee, B. et. al. "Some Principles for Designing a Wide-Area WDM Optical Network." *IEEE/ACM Transactions on Networking*. Vol. 4 (5). Oct. 1996.

Narula, A., P. Lin, and E. Modiano. "Efficient Routing and Wavelength Assignment for Reconfigurable WDM Networks."*Journal on Selected Areas in Communication*. Jan. 2002. 75–88.

Park, K. and W. Willinger, ed. *Self Similar Network Traffic and Performance Evaluation*. Wiley Interscience Series. 2000.

Proceedings of Symposium on Polarization Mode Dispersion. Venice, March 2002.

Raghavan, P. and E. Upfal. "Efficient Routing in All Optical Networks." Proceedings of 26th ACM Symposium on Theory of Computing. May 1994.

Ramamurthy, B. *Design of Optical WDM Networks-LAN, MAN and WAN Architectures*. Kluwer Publications, 2001.

Ramaswami, R. and G. Sasaki. "Multiwavelength Optical Networks with Limited Wavelength Conversion."*IEEE/ACM Transactions on Networking*. Vol. 6 (6). Dec. 1998. 744–754.

Ramaswami, R. and K. Sivarajan. *Optical Networks: A Practical Perspective*, Second Edition. Morgan Kauffman; 2001. Chapter 3–4.

______."Routing and Wavelength Assignment in All-Optical Networks."*IEEE/ACM Transactions on Networking*. Vol. 3 (5). 1995.

______.*Optical Networks: A Practical Perspective*. San Francisco, CA: Morgan Kauffman; 1999.

______."Routing and Wavelength Assignment."*IEEE/ACM Transactions on Networking*. April 1995.

Robinson, F.N.H. *Noise and Fluctuations*. Oxford, England: Oxford University Press; 1977.

Sahasrabuddhe, L., S. Ramamurthy, and B. Mukherjee. "Fault Management in IP-Over-WDM Networks: WDM Protection Versus IP Restoration."*Journal of Selected Areas in Communication*. Vol. 20 (1). 21–33.

Saleh, B. and M. Teich. *Fundamentals of Photonics*. Wiley, 1991.

Senior, J., C. Qiao, and S. Dixit. "All Optical Networking 1999: Architecture, Control and Management Issues." Proceedings SPIE. 2000.

Shimada, S. and H. Ishio. *Optical Amplifiers and Their Applications*. John Wiley and Sons, 2000.

Siller, Curtis A. Jr. and Mansoor Shafi. *A Sourcebook of Synchronous Networking*,*SONET/SDH*. IEEE Press, 1996.

Srivastava, Atul K. and Yan Sun, "Advances in Erbium-Doped Fiber Amplifiers."*Optical Fiber Telecommunications IVA: Components*. Ed. I. Kaminow and Tingye Li. San Diego: Academic Press, 2002. 174–212.

Stamotios, V. *Introduction to DWDM Technologies*. Wiley and Sons, 1998.

Steeleman, G.E. and C.M. Wolfe. *Avalanche Photodiodes*. Academic Press, 1977.

Su, Chao, Lian-Kuan Chen, and Kwok-Wai Cheung. "Theory of Burst-Mode Receiver and Its Applications in Optical Multiaccess Networks."*Journal of Lightwave Technology*. Vol. 15 (4). April 1997. 590–606.

Tam, A. and P. Lin. "Wavelength Requirements for Virtual Topology Reconfiguration in WDM Ring Networks." Proceedings of Infocom. 2001.

Taub, M. and S. Shilling. *Principles of Communication Systems*. New York, NY: McGraw-Hill; 1992.

Wartikar, J.N. *Applied Mathematics*. S. Chand and Co., June 1976.

www.10gea.org/Tech-whitepapers.htm.

Yoo, M., C. Qiao, and S. Dixit. "Optical Burst Switching for Service Differentiation in the Next-Generation Optical Internet."*IEEE Communications Magazine*. Feb. 2001. 98–104.

Yoo, S.J. "Wavelength Conversion Technologies for WDM Network Applications."*Journal of Lightwave Technologies*. Vol. 14 (6). June 1996. 955–966.

Yu, T. et al. "A Mean-Field Approach for Simulating Wavelength Division Multiplexed Systems." *IEEE Photonics Technology Letters*. Vol. 12 (4). April 2000. 443–445.

Zhang, Lixia et al. *IEEE Communications Magazine*. May 2002. 116–127.

Zhang, Z. and A. Acampora. "A Heuristic Wavelength Assignment Algorithm for Multihop WDM Networks with Wavelength Routing and Wavelength Reuse."*IEEE/ACM Transactions on Networking*. Vol. 3 (3). 1995. 281.

Zirngibl, M. et al. IEEE Photonics Technology Letters. Aug. 1986.

System Requirements and Recommendations

Hardware Requirements

VPItransmissionMaker™/VPIcomponentMaker™ requires the following minimum hardware configuration:

- Installed TCP/IP protocol: The Simulator will not run without this protocol! On stand-alone computers (e.g., without a network card), you must install Microsoft® Loop Back Adapter or Dialup Networking first.
- Personal computer with a Pentium III 800 MHz or higher processor
- 512 megabytes (MB) of RAM
- 150 MB of free hard disk space for software installation, 1 BG recommended for file storage
- XGA monitor (1024x768 pixels minimum display resolution)
- Video adapter that supports High Color (15-bit/65,536 colors)
- CD-ROM drive

Software Requirements

VPItransmissionMaker™/VPIcomponentMaker™ requires the following minimum software configuration:

- Microsoft® Windows NT® 4.0 with Service Pack (SP) 6a or Windows® 2000 with SP2
- TCP/IP protocol installed and configured
- Microsoft® Internet Explorer 5.5 or higher
- Microsoft® Office 2000 (Service Release 1) or higher

Additional Software Provided in the Package

The following software is needed for proper operation and is included on the CD:

- Adobe® Acrobat® Reader 5.05

[SYMBOL] [A] [B] [C] [D] [E] [F] [G] [H] [I] [J] [L] [M] [N] [O] [P] [Q] [R] [S] [T] [U] [V] [W] [X] [Z]

[SYMBOL] [A] [B] [C] [D] [E] [F] [G] [H] [I] [J] [L] [M] [N] [O] [P] [Q] [R] [S] [T] [U] [V] [W] [X] [Z]

[SYMBOL] [A] [B] [C] [D] [E] [F] [G] [H] [I] [J] [L] [M] [N] [O] [P] [Q] [R] [S] [T] [U] [V] [W] [X] [Z]

[SYMBOL] [A] [B] [C] [D] [E] [F] [G] [H] [I] [J] [L] [M] [N] [O] [P] [Q] [R] [S] [T] [U] [V] [W] [X] [Z]

[SYMBOL] [A] [B] [C] [D] [E] [F] [G] [H] [I] [J] [L] [M] [N] [O] [P] [Q] [R] [S] [T] [U] [V] [W] [X] [Z]

[SYMBOL] [A] [B] [C] [D] [E] [F] [G] [H] [I] [J] [L] [M] [N] [O] [P] [Q] [R] [S] [T] [U] [V] [W] [X] [Z]