Green Optical Network Design: 
Power Optimization of Wide Area and 
Metropolitan Area Networks 

by 
Katherine Xiaoyan Lin 
B.S., Massachusetts Institute of Technology (2010) 

Submitted to the Department of Electrical Engineering and Computer Science 
in partial fulfillment of the requirements for the degree of 
Master of Engineering in Electrical Engineering and Computer Science 
at the 
MASSACHUSETTS INSTITUTE OF TECHNOLOGY 
June 2011 
© Katherine Xiaoyan Lin, MMXI. All rights reserved. 
The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author ...................................................................... 
Department of Electrical Engineering and Computer Science 
May 18, 2011 

Certified by .................................................................. 
Vincent W.S. Chan 
Joan and Irwin Jacobs Professor of Electrical Engineering and 
Computer Science 
Thesis Supervisor 

Accepted by ......................................................... 
Dr. Christopher J. Terman 
Chair, Masters of Engineering Thesis Committee
Green Optical Network Design:
Power Optimization of Wide Area and Metropolitan Area Networks
by
Katherine Xiaoyan Lin

Submitted to the Department of Electrical Engineering and Computer Science on May 18, 2011, in partial fulfillment of the requirements for the degree of Master of Engineering in Electrical Engineering and Computer Science

Abstract
Advancements in technology are fueling huge growth in network traffic capacity. Demand for low cost, reliable, and high bitrate transmissions grows 40-110% internationally every year. To date, most research has focused on cost minimization of wide area and metropolitan area networks. In this thesis, we concentrate instead on finding scalable WAN designs with respect to power constraints and optimal MAN topologies with minimal capital and operating expenditures.

We find optical bypass networks to be most scalable with respect to power consumption, especially when quality of service and network flexibility, reliability, and protection are considered. The power consumption of the standard bypass network can be lowered further through a hybrid design in which whole wavelengths of core, stable traffic between node pairs are routed via direct, fixed lightpaths using patch panelling and unexpected, bursty traffic is switched on a standard optical bypass network. We analyze power distribution among components and find the OXC switch most scalable at each node and O/E/O switches and routers wasteful. Finally, we prove that shortest path and minimum hop routing is power optimal and traffic balanced routing should be avoided.

We approximate MAN topologies with regular graphs for tractable analysis. We augment a previous cost-based joint optimization formulation [13] with power expenditure modelling and obtain closed form solutions for optimal node degree and normalized network costs. We find that the optimal node connectivity increases 20-25% due to the added operating expenditures. Normalized network cost and normalized network cost per unit traffic also rise by approximately 25%. Our results show that the Generalized Moore graph with node degree between 0.05N and 0.08N is both power and cost minimal for a purely optical network.

Thesis Supervisor: Vincent W.S. Chan
Title: Joan and Irwin Jacobs Professor of Electrical Engineering and Computer Science
Acknowledgments

First and foremost, I thank my research advisor, Professor Vincent W.S. Chan, for his guidance, patience, and support. Professor Chan expresses unwavering faith in the ability of his students and mentors us to think creatively, to communicate effectively, and to lead with integrity. It has been my honor to work with him.

I am grateful to the Siebel Scholars Foundation, National Science Foundation, and MIT for the financial support to this research.

I thank my fellow graduate students Matthew Carey, David Cole, Shane Fink, Henna Huang, Rui Li, and Lei Zhang for their insightful discussions and enjoyable companionship.

Lastly, I thank my family. I am grateful to my parents, Zhiping Lin and Xiulan Ma, for their understanding and encouragement. To my love, Jieyun Fu, as a token of my appreciation for his unconditional friendship and support, I dedicate this thesis.
# Contents

1 Introduction 17
   1.1 Motivation of the Problem Statement 18
   1.2 Problem Approach 19
   1.3 Previous Work and Contribution for WAN 21
   1.4 Previous Work and Contribution for MAN 23
   1.5 Thesis Organization 24

2 Components of the Wide Area Network 27
   2.1 Transmitter/Receiver (Transceiver) 28
   2.2 Optical Amplifier and Regenerator 29
   2.3 Switch 31
   2.4 Router 33
   2.5 Summary of Chapter 2 35

3 Wide Area Network Power Consumption Model 37
   3.1 Physical Topology 38
   3.2 Traffic Model 39
   3.3 Network Parameters 44
   3.4 Bypass versus Non-bypass 46
   3.5 WAN Network Designs 47
   3.6 Summary of Chapter 3 49

4 WAN Power Analysis and Results 53
4.1 Analyzing Power Consumption with Baseline Traffic Model ........ 55
  4.1.1 Measuring Baseline Power Consumption ....................... 55
  4.1.2 Measuring the Effect of Switch Choices ....................... 61
  4.1.3 Measuring the Effect of Cooling Overhead .................... 64
  4.1.4 Measuring the Distribution of Power Consumption Among Com-
      ponents ....................................................... 65
4.2 Analyzing Power Consumption Under a Highly Loaded Traffic Model 73
4.3 Patched Bypass versus Bypass Network Designs .................... 77
4.4 Justification of Shortest Path Routing ............................ 79
4.5 Summary of Chapter 4 ............................................. 87

5 Components and Cost Model of the Metropolitan Area Network 89
  5.1 MAN Traffic Model ............................................... 90
  5.2 Parametric Network Cost Model ................................ 91
      5.2.1 Transceiver Cost ........................................ 91
      5.2.2 Fiber Connection Cost ................................... 92
      5.2.3 Switch Cost ............................................. 92
      5.2.4 Network Cost Model ...................................... 97
  5.3 Regular Topologies ............................................... 98
      5.3.1 Generalized Moore Graphs ................................ 99
      5.3.2 Δ-Nearest Neighbors Graphs .............................. 101
      5.3.3 Symmetric Hamilton Graphs ............................... 102
  5.4 Summary of Chapter 5 ............................................. 102

6 MAN Power Analysis and Results ..................................... 105
  6.1 Problem Formulation .............................................. 106
  6.2 Optimal Node Degree and Network Cost for Δ-Nearest Neighbors Graphs 107
  6.3 Optimal Node Degree and Network Cost for Symmetric Hamilton Graphs 108
  6.4 Optimal Node Degree and Network Cost for Generalized Moore Graphs 110
  6.5 Analyzing Optimal Node Degree .................................. 111
  6.6 Analyzing Minimal Network Cost Normalized for Network Size .... 114
6.7 Analyzing Minimal Network Cost Normalized for Network Size and Traffic ......................................................... 117
6.8 Comparing Power and Cost Minimal with Cost Minimal Network Topologies .................................................... 119
6.9 Summary of Chapter 6 ......................................................... 123

7 Conclusion .................................................................. 127
7.1 Summary of Contributions ................................................... 127
7.2 Future Work ................................................................. 131
List of Figures

1-1 United States Wide Area Network ........................................... 21
2-1 Functional Diagram of Transmitter/Receiver (Transceiver) .......... 28
2-2 Functional Diagram of Erbium-doped Fiber Amplifier (EDFA) .... 30
2-3 Functional Diagram of an OXC Switch ............................... 31
2-4 Functional Diagram of an O/E/O optical core switch ............. 32
2-5 Functional Diagram of an O/E/O electronic core Switch ......... 32
3-1 Comparison of Plant Topology and Fiber Topology ............... 39
3-2 United States Wide Area Network ........................................ 40
3-3 Sample Calculation of a Population Matrix (Upon Which the Traffic
    Demand Matrix Is Based) ........................................... 41
3-4 Schematic Diagram of WAN Designs Cases ......................... 51
4-1 Baseline Power Consumption ............................................. 56
4-2 Baseline Power Consumption with Cooling Overhead ............ 58
4-3 Effect of Cooling Overhead on Network Designs (on log-log scale) . 62
4-4 Effect of Cooling Overhead on Network Designs (on log-linear scale) . 63
4-5 Power Distribution Among Network Components for OXC Switched
    Networks .................................................................. 67
4-6 Power Distribution Among Components for OXC Switched Networks
    (regenLimit = 500km) ............................................. 69
4-7 Power Distribution Among Network Components for O/E/O_ELEC Switched
    Networks ................................................................. 70
6-14 Effect of Power Cost on Normalized Network Cost Per Unit Traffic under Increased Traffic Loading \( (t = 3) \) . . . . . . . . . . . . 126
List of Tables

2.1 Power Dissipation of XFP Tranceivers ......................... 29
2.2 Power Dissipation of OXC Switches .......................... 33
2.3 Power Dissipation of Routers at 40Gbps ................... 34

3.1 Sample Traffic Demand from Chicago to All Other Nodes .... 42
3.2 WAN Network Designs – Components at End Nodes .......... 47
3.3 WAN Network Designs – Components at Intermediate Nodes .. 47

4.1 Effect on Aggregate Capacity from Scaling Traffic Matrix ... 77
Chapter 1

Introduction

Advancements in technology over the past decade have led to new services and applications that require data transmission. For example, consumers have increased subscriptions to applications such as IP television, Voice over IP, video conferencing, interactive gaming, etc. that require low latency, high bitrate, and high reliability networks. The increase in underlying demand for network capacity leads to a need to build scalable networks with adequate capacity, throughput, and delays while maintaining acceptable capital and operating expenditures. In particular, network design needs to focus on the scalability with respect to power consumption. The optimal network topology needed is one in which power consumption grows sustainably as the number of users, capacity of network, and speed of network increase. Indeed, one of the current limiting factors to network growth is power consumption.

Ideal network architecture should minimize both cost and power consumption under any traffic profile (e.g. high user density, bursty traffic, sparse traffic, etc.). However, the dual optimization problem of minimizing both cost and energy is often intractable and finding an analytical solution is not always possible. Therefore, researchers usually tackle the cost and power consumption problems separately. Furthermore, the hierarchical organization of networks into Wide Area Networks (WAN), Metropolitan Area Networks (MAN), and Local Area Networks (LAN) leads to different optimization models due to differences in function, loading, and sometimes equipment. In other words, the precious resources to be shared is often different
among layers in the network hierarchy; in the WAN, the core routers consume enormous power whereas in the MAN, the grooming switches/aggregating routers are expensive. In addition, in WAN’s, the fiber path topology is often irregular and fixed whereas the graph structure of MAN’s can be approximated as regular, well-connected graphs. Therefore, researchers are forced to solve the cost and energy minimization problems differently for each network specification.

We identify four main areas of research:

1. Cost Optimization of Wide Area Networks
2. Power Optimization of Wide Area Networks
3. Cost Optimization of Metropolitan Area Networks
4. Power Optimization of Metropolitan Area Networks

We do not study the architecture design of LAN’s because LAN’s consist of mostly passive components and have minimal power consumption. In this thesis, we focus on topics (2), (3), and (4). We look to find power efficient WAN architectures and to minimize the combined capital and operating (power) expenditures of the MAN.

1.1 Motivation of the Problem Statement

Why is minimizing power consumption important? We first argue that energy limitations will be the single binding constraint of future large-scale networks and so power minimization is crucial to optical communications over the next few years. Developed countries have been increasing the energy dedicated to operating IT equipment. The energy consumption of the Telecom Italia network reached more than 2TWh in 2006, representing an 8% increase over 2005 and 12% over 2004 [7]. In the United Kingdom, operating IT equipment consumed about 10% of the country’s entire power consumption in 2007 [7]. When we consider the consequences of such increased traffic demand on individual equipment, the need to be power efficient is even more apparent. Currently, a rack of network equipment dissipates more than 15-20kW and
ambient air cooling solutions are limited in their ability to handle high heat loads [1] (the upper limit of air cooling is approximately 10kW per rack). Current air cooling technology cannot physically cool equipment racks and rooms as the heat dissipation density in W/m³ continues to rise. In addition, advanced cooling techniques such as liquid or mist cooling are either not mature enough technologies for deployment today or may not be feasible from a cost or maintenance standpoint. Therefore, all future large-scale networks must address power minimization as a primary constraint.

A second motivation for reducing power consumption is the observation that an energy-efficient architecture design is also a cost-effective design. The growth in routing needs leads to more and more switch and/or router ports at each node and consequently results in a larger demand for cooling. The cost versus heat dissipation curve can grow super-linearly as industry seeks (technically challenging) alternatives to air cooling, such as liquid cooling or mist cooling. Therefore, a energy-minimal network design should minimize operating expenditures, leading to cost savings. In fact, “if an energy-efficient network can be designed and energy-efficient operation strategies can be implemented to cut even 1% of the total energy consumption, then this will lead to a significant cost reduction to save about $5 billion per year given that the price of electricity is seventeen cents per kWh” [28]. Due to this significant factor, our thesis focuses on finding power minimizing architecture designs that will be green, economical, and scalable over the next few years.

1.2 Problem Approach

The power consumption optimization problem is complex with many interacting factors. In fact, “the topology design problem [in MAN’s] has a complexity of \( \Theta(2^n^2) \). For a design that involves 10 nodes, \( 3.518 \times 10^{13} \) scenarios need to be tested” [13]. In addition to the topology problem, the Routing and Wavelength Assignment (RWA) problem needs to be solved; via reduction to graph coloring, we can show that the RWA problem is NP-complete. Therefore, the only feasible solutions are numerical or analytical. The WAN and MAN have several inherently different characteristics,
so we can apply two different problem solving approaches.

In analyzing WAN’s, we must use a numerical-based model because the network nodes and fiber plant topology are fixed, often in strategic, but highly irregular graph locations, making it impossible to generalize analytical results into heuristics. Furthermore, the demographics of a WAN exert influence on the traffic model (i.e. we cannot assume every node in the United States is New York City and model the network using uniform all-to-all traffic as if every node were similar to New York City). Once we incorporate the demographics of a WAN, we lose perfect generalization and can only provide heuristic recommendations. As a result, the first half of our thesis focuses on finding an optimal architecture design with regard to energy consumption for the United States WAN, based upon the AT&T Next-Generation IP/MPLS Backbone network [2] (see Figure 1-1) using detailed simulations and linear programming.

When studying MAN’s, we can take advantage of the regular representation of network topologies to formulate general analytical solutions. In fact, optimal network designs based on cost minimization have been developed in detail [13]. In the second half of our thesis, we augment analytical models presented by Guan with power consumption costs to find the overall optimal MAN architecture with respect to both cost and power savings.

At all times, our research goal is to understand the relationships among competing factors in power consumption and to present simple and elegant architecture designs that minimize power consumption of simplified networks. What relationships do we mean? Take for example the number of nodes: as more users subscribe to the network, more transceivers and fibers must be installed, more ports must be added to routers, and more computing effort must be expended to handle the increased traffic. All these requirements translate into increased power consumption. Consider further the network connectivity: as we increase the connectivity among nodes in the network graph, we need to dedicate more resources to routing; however, higher connectivity also means a lower minimum maximum hop distance, which results in less switching — a source of possible power savings. A correct problem formulation for power minimization is complex and depends on critical understanding of the building
blocks of WDM networks and their interactions.

Figure 1-1: A 25-node, 56-link network representing the United States Wide Area Network, based upon the AT&T Next-Generation IP/MPLS Backbone network.

1.3 Previous Work and Contribution for WAN

Previous work in energy minimization in the WAN has two main branches. The first explores selectively turning off network components during off-peak usage. This idea is based upon the natural redundancy built into networks; as traffic load decreases (e.g. at night), nonessential elements can be put to “sleep” while a minimal number of components remain fully powered to handle residual traffic [23, 24]. However, because network traffic is dynamic, especially if we are forward-looking to the next generation of technology demands on IP-traffic capacity, few components in the core router can be dormant for any significant period of time. Countered with a potential decrease in end-user quality of service and the added cost and complexity of monitoring and managing hibernation patterns of each component, we feel a better solution is to
focus on energy-efficient network design.

In the second approach to power minimization, Shen [28] creates a mixed-integer linear programming problem to jointly minimize the power consumption of IP routers, erbium-doped fiber amplifiers (EDFA), transmitters, and receivers. His work shows that using optical bypass (in which wavelengths are able to directly connect source and destination nodes without being terminated and regenerated at each intermediate node) can reduce energy consumption from 25-40%. Shen also breaks down the power consumption in his optimized network by components and finds that routers consume much more energy than all the other elements combined. Our research falls into this branch and focuses on expanding the ideas that Shen presented with our own models, research results, and recommendations. Specifically, we analyze the traffic regime in which demand between any two node pairs (or between almost every node pair) increases hundredfold and can fill multiple wavelengths.

In our research, we first confirm the substantial energy savings that previous research suggests can be achieved with energy-efficient network designs, such as using full optical bypass. However, we acknowledge that current networks are far from being fully connected with dark fiber and so we compare the power consumption of bypass versus non-bypass network implementations. We also compare the different energy consumptions from using various network components (e.g. optical cross connect or OXC switches vs. optical-electronic-optical or O/E/O switches vs. routers). We use shortest path routing as a more power efficient routing method than some suggested in previous research and prove that a routing algorithm based upon both shortest path and minimum hop is power optimal. In our analysis, we use the US AT&T Next-Generation IP/MPLS Backbone based network map shown in Figure 1-1 as the WAN of interest and base our traffic matrices on metropolitan statistical area populations.

Our detailed models allow our results to be more realistic and our conclusions more directly applicable to the United States WAN, when compared to previous research. This is because we base the traffic modelling on network demographics instead of using either uniform all-to-all traffic or randomly generated traffic based upon a uni-
form distribution with identical mean for all nodes. Furthermore, because our model for the WAN includes only components in the core network (i.e. we exclude edge routers that aggregate traffic in a MAN or LAN), a comparison of power consumption by components is more relevant in identifying areas of potential energy savings. Finally, we frame our analysis uniquely as a function of link-reach/transmission-reach capabilities, since increasing the reach of long-haul fiber with respect to distance between regenerations is a current priority in optical network research [19]. This analysis presents comparisons of energy savings in the WAN for various stages of technological innovation over the next decade.

1.4 Previous Work and Contribution for MAN

The second half of our thesis focuses on the optimization of capital and operating expenditures in the MAN though analytical methods. Previous research has concentrated on only cost minimization of regular graph representations of the MAN [11, 12, 13]. Guan identified the cost minimal graph topology to be Generalized Moore graphs. Through performing normalized cost analysis on regular graphs, he identified the lower bound on capital expenditures to belong to the Generalized Moore graph and the upper cost bound to correspond to the $\Delta$-Nearest Neighbors graph. He further extended his results to provide heuristics to evaluate the cost feasibility of irregular networks.

We augment Guan’s cost optimization model with power consumption variables and parameters to analyze any impact in a MAN topology, optimized for capital expenditures, as a result of operating expenditure considerations. In particular, we seek the graph connectivity (node degree) that minimizes total expenditures. The scope of this thesis is limited to the analysis of the Generalized Moore, Symmetric Hamilton, and $\Delta$-Nearest Neighbor graphs. The main reason to focus on Moore graphs is the desire to conduct research on optimal structures instead of attempting to improve a recognized suboptimal structure. We analyze Hamilton graphs to provide a baseline comparison to the Moore graphs. Symmetric Hamilton graphs are popular in
present day networks (a degenerate Hamilton graph is the ring upon which previous
generation SONET networks were based) and so can give an indication of current
network capital and operating expenditures. We include $\Delta$-Nearest Neighbors graphs
as a potential upper bound on total expenditures since they have been shown to have
the worst cost (i.e. capital expenditure) performance.

1.5 Thesis Organization

The rest of the thesis is organized as follows.

In Chapter 2, we describe the basic components that make up a WAN: transmitter/receiver pairs (transceivers), erbium-doped fiber amplifiers (EDFA), switches
(optical and optical-electronic-optical), and routers. For each component, we describe
how it is used in the model and its representative power consumption value.

In Chapter 3, we build the WAN model. We discuss the physical topology, traffic
model, and network parameters. Furthermore, we describe the various network de-
signs that are compared. We look at four main network designs: patched bypass (in
which direct, non-switched lightpaths are allocated between any node pair), bypass
(in which lightpaths between any node pair are switched at intermediate nodes), non-
bypass (in which lightpaths between any node pair are terminated and re-initiated by
routers at intermediate nodes), and groomed non-bypass (in which whole lightpaths
between any node pair are switched while residual or fractional lightpaths are termi-
nated and groomed by routers at intermediate nodes). Each of the network designs
can be implemented with three switches: optical switch (OXC), O/E/O switch with
optical core, and O/E/O switch with electronic core.

In Chapter 4, we present our WAN power optimization results and recommenda-
tions. We first show the power consumption of each of the four network designs
when implemented with each of three types of switches. We then look at the power
breakdown via components and the effects of cooling costs on power consumption.
We compare the analysis from a static, expected traffic matrix with that of a high
variance, heavily-loaded one. We discuss how various network reliability safeguards
(such as 1-to-1 spare lightpath versus n-to-1 spare link) affect the network design choice. Finally, we justify the use of shortest path routing and prove it is power optimal in most topologies. Furthermore, we prove that the optimal routing algorithm uses an edge weighting function that depends on both shortest path and minimum hop.

In Chapter 5, we describe the basic components that make up a MAN: fiber connections, transceivers, and switches (optical and optical-electronic-optical). For each component, we discuss its use in the MAN and its average cost modelling, as originally presented by Guan [13]. We then build upon his cost model by incorporating operating expenditures, through converting the power consumption into cost figures, based on a five year component lifetime. Finally, we present an introduction to regular graph representations by summarizing the properties of the Generalized Moore, Δ-Nearest Neighbors, and Symmetric Hamilton graphs.

In Chapter 6, we present our MAN capital and operating expenditure minimization results based upon a uniform all-to-all traffic model. We first show the optimal node degree for each topology and discuss the effects of increased data rate and traffic loading on the connectivity. We analyze network cost normalized by number of nodes in the network and then by both network size and traffic load. We find that the optimal topology for cost optimized MAN (i.e. Generalized Moore graphs) is also optimal for a power and cost minimized network. The main difference is an approximate 25% increase in network connectivity and cost.

Finally, in Chapter 7, we conclude the thesis with a summary of our contributions and recommendations. We also discuss areas of promising future research.
Chapter 2

Components of the Wide Area Network

In this chapter, we build the foundation for our thesis by describing each component used in our network and providing its representative power consumption value. In selecting the representative components from available models, we choose the most power efficient ones for two reasons. First, our research is forward-looking. Since technological innovation emphasizes making components more “green”, we believe future models will only be more power efficient than even the most efficient components today. Second, by using the most power efficient elements, we isolate the power savings so that all savings arising from our optimization exercises will be the power efficiency gained by using the optimal architecture. In other words, we eliminate the low-hanging fruit that is the obvious power savings from transforming a suboptimal network topology composed of inefficient components to the same suboptimal network topology with more power efficient components.

In the following sections, we present the components used in modelling our networks and provide representative efficient power consumption values for each component. Our model consists of four main components: transceivers/receivers (transceivers), amplifiers (erbium-doped fiber amplifiers in long-haul fibers), switches (OXC and O/E/O switches), and routers.
2.1 Transmitter/Receiver (Transceiver)

The main function of a transmitter is to “send out a modulated optical signal complying with a set of specifications, such as bit error rate or signal-to-noise ratio” and the main function of a receiver is to “detect a modulated photonic signal with a predetermined level of accuracy, which is measured in bit error rate” [13]. Figure 2-1 is a schematic diagram of transceiver application and placement within a network.

Traditionally, 850nm transmitter/receiver pairs were used in O/E/O switches to convert optical signals to electronic and back to optical at short distances; 1550nm transceivers were used to link the optical ends to long-haul fiber and to regenerate signals when the signal-to-noise ratio was no longer tolerable. The 850nm wavelength transceivers were used because the shorter wavelength was easier to focus and the component was easier to build in a small size. However, now that manufacturing technology has advanced to produce small silicon lenses capable of focusing 1550nm wavelengths, the reasons for using the 850nm transceivers are no longer relevant. As a result, our forward-looking models use 1550nm transceivers exclusively.

![Figure 2-1: A functional diagram of transceiver component use. Transmitters on one end of an optical fiber send out modulated optical signals which are detected at the destination end by receivers.](image-url)
In Table 2.1, we list the power dissipation data of popular models [16, 15, 14, 5]. We consider the XFP (10 Gb/sec small form factor pluggable) optical transceiver as our transmitter/receiver element. We define $P_{1550\text{nm}}$ to be the power consumption of the 1550nm wavelength transceiver for use in connecting the network. Representative power dissipation per transceiver is 2.3W under typical operation from the Bookham IGF-17511J or IGF-32511 models; maximum power dissipation is 3.5W from any of the Bookham or JDS Uniphase 1550nm wavelength transceivers.

Table 2.1: Power Dissipation of XFP Transceivers

<table>
<thead>
<tr>
<th>XFP Transceiver Model</th>
<th>Transceiver Wavelength</th>
<th>Typical Power Dissipation</th>
<th>Maximum Power Dissipation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bookham IGF-17511J</td>
<td>1550nm</td>
<td>2.3W</td>
<td>3.5W</td>
</tr>
<tr>
<td>Bookham IGF-32511</td>
<td>1550nm</td>
<td>2.3W</td>
<td>3.5W</td>
</tr>
<tr>
<td>Bookham IGF-42311J</td>
<td>1310nm</td>
<td>2.2W</td>
<td>2.5W</td>
</tr>
<tr>
<td>Bookham IGF-42312J</td>
<td>1310nm</td>
<td>2.2W</td>
<td>2.7W</td>
</tr>
<tr>
<td>JDSU JXP-01EMAB1</td>
<td>1550nm</td>
<td></td>
<td>3.5W</td>
</tr>
<tr>
<td>JDSU JXP-01EMAC1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDSU JXP-01EEAB1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDSU JXP-01EGAB1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDSU JXP-01LMAB1</td>
<td>1310nm</td>
<td></td>
<td>2.5W</td>
</tr>
<tr>
<td>JDSU JXP-01LMAC1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDSU JXP-01LEAB1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDSU JXP-01LGAB1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDSU PLRXXL-SC-S43-C1</td>
<td>850nm</td>
<td>1.2W</td>
<td>1.5W</td>
</tr>
</tbody>
</table>

2.2 Optical Amplifier and Regenerator

An optical erbium-doped fiber amplifier (EDFA) placed onto a fiber simultaneously adds gain to all wavelengths on the fiber. Signals on the fiber are not read, converted, or processed. Amplification is achieved when erbium in the fiber is excited by optical pumping. Figure 2-2 is a schematic of an EDFA placed onto an optical fiber.

A regenerator enhances signals carried on wavelengths by “converting the optical signal to an electronic signal of the same bitrate, amplifying it, and then converting the electronic signal back to the optical domain” [13]. This $O/E/O$ conversion ensures that the output signal has a high signal-to-noise ratio.
Figure 2-2: A functional diagram of an erbium-doped fiber amplifier (EDFA). An EDFA adds gain to all wavelengths carried on the fiber.

In order to accurately model the power consumption of fiber amplifiers in a network, we must incorporate the distance of each fiber edge traversed by our optical signals because optical signals attenuate with distance and must be amplified every so often. We define linkReach [km] to be the distance a signal can travel in the fiber before needing amplification. Furthermore, we cannot ignore that signals become distorted as they travel through the fiber; coupled with multiple rounds of amplification, these distortions result in high bit error rates unless the signal is regenerated every so often. We define regenLimit [km] to be the distance a signal can travel with gain added at intermediate amplification points before needing to be completely regenerated. We refer to large regenLimit’s as good quality optical transmission or long optical reach.

We use the JDSU WaveReady WRA-217 Multichannel EDFA as our amplifying component. We define $P_{amp}$ to be the power consumption of the EDFA. Representative power dissipation per amplifier is 18W under typical operation; maximum power dissipation is 24W [27, 21, 20, 25]. We note that an amplifier adds gain to an entire fiber, which can support up to 200 wavelengths, so the power per wavelength is actually $\frac{1}{200}P_{amp}$. Our representative regenerator will be the 1550nm transceiver (described previously) with $P_{regen} = P_{1550nm}$, as the $O/E/O$ conversion between a transmitter/receiver pair successfully converts the input signal into an output one with a high signal-to-noise ratio and minimal distortion.
2.3 Switch

There are two general cross connection fabrics in optical networks:

1. **Optical Cross Connect switches (OXC)** are pure optical switches in which the optical data streams are cross connected within the optical domain. Figure 2-3 is a schematic diagram showing the functionality of a pure optical switch.

![Figure 2-3: A functional diagram of an optical cross connect (OXC) switch, in which optical data streams are cross connected within the optical domain.]

2. **Optical-Electrical-Optical switches (O/E/O)** convert optical data streams into electronic data streams and back. There are two types of O/E/O switches.

   The first type of O/E/O switch has an optical core (hereafter referred to as O/E/O_{OXC}). A few variations of the O/E/O_{OXC} exist. The one we use first converts the optical data streams into electronic data streams and back to optical (equivalent to a regeneration of the signal), then cross connects within the optical domain, and finally converts the optical streams to electronic streams and back to the optical domain so that the final output signal is optical (equivalent to a second regeneration of the signal). Figure 2-4 is a schematic diagram showing the functionality of an O/E/O optical core switch.

   The second type of O/E/O switch has an electronic core (hereafter referred to as O/E/O_{ELEC}). The O/E/O_{ELEC} first converts the optical data streams into electronic data streams, cross connects within the electronic domain, and then converts the streams back to the optical domain. Figure 2-5 is a schematic diagram showing the functionality of an O/E/O electronic core switch.
In Table 2.2, we list power dissipation data of popular OXC switch models [8, 9, 10, 17, 4, 3, 18]. We approximate the power dissipation per connection pair, $P_{OXC_{conn}}$, with the maximum power dissipation of each model divided by its respective maximum port configuration. The entries in Table 2.2 suggest that the relation between port count and power consumption is sublinear; however, to be conservative in our analysis, we assume that the power dissipation per connection pair scales linearly with the number of ports needed in our network model.

The most power efficient OXC switches are the Calient DiamondWave FiberConnect with 470mW per connection, Glimmerglass Intelligent Optical System 500 or 600 with 443mW per connection, and JDS Uniphase 401-04-O-S 64x64 configuration with 391mW per connection. We estimate $P_{OXC_{conn}}$ to be the average power dissipation per connection of these three models, i.e. $P_{OXC_{conn}} = 430mW$. 

32
Table 2.2: Power Dissipation of OXC Switches

<table>
<thead>
<tr>
<th>OXC Switch Model</th>
<th>Maximum Power Dissipation</th>
<th>Maximum Port Configuration</th>
<th>Power Dissipation per Connection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calient DiamondWave FiberConnect</td>
<td>470mW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glimmerglass Intelligent Optical System 100</td>
<td>50W</td>
<td>96×96</td>
<td>520mW</td>
</tr>
<tr>
<td>Glimmerglass Intelligent Optical System 500, 600</td>
<td>85W</td>
<td>192×192</td>
<td>443mW</td>
</tr>
<tr>
<td>JDSU 401-04-0-S</td>
<td>15W, 25W</td>
<td>32×32, 64×64</td>
<td>469mW, 391mW</td>
</tr>
<tr>
<td>JDSU 402-04-O-S Single mode</td>
<td>25W, 45W</td>
<td>32×32, 64×64</td>
<td>780mW, 730mW</td>
</tr>
<tr>
<td>JDSU 405-04-0-S VST</td>
<td>25W, 45W</td>
<td>32×32, 64×64</td>
<td>780mW, 703mW</td>
</tr>
<tr>
<td>JDSU 414-02-0-S Single mode</td>
<td>30W</td>
<td>32×32</td>
<td>938mW</td>
</tr>
</tbody>
</table>

To model an O/E/O\textsubscript{OXC}, we use an optical switch as the core switch, and note that each connection pair also requires two transceivers (see Figure 2-4). We define the power of an O/E/O\textsubscript{OXC} switch to be \( P_{O/E/O\textsubscript{OXC}} = P_{OXC\textsubscript{conn}} + 2 \cdot P_{1550\textsubscript{t/r}} \).

To model an O/E/O\textsubscript{ELEC}, we use an electronic switch as the core switch, and note that each connection also requires a transceiver (see Figure 2-5). We approximate the power dissipation per connection, \( P_{ELEC\textsubscript{conn}} \), with the maximum power dissipation of the electronic core divided by its maximum port configuration. Our representative electronic switch is the Vitesse VSC3144 10.709Gbps crosspoint switch [6] with typical power dissipation of 21W for a 144×144 switch, yielding \( P_{ELEC\textsubscript{conn}} = 146mW \). We define the power of an O/E/O\textsubscript{ELEC} switch to be \( P_{O/E/O\textsubscript{ELEC}} = P_{ELEC\textsubscript{conn}} + P_{1550\textsubscript{t/r}} \).

### 2.4 Router

A router is a network device that forwards incoming packets from one network to another based on internal routing tables. The line or output port that outgoing packets should be directed to is determined by the destination address in the packets. We focus on two main types of routers.

1. **Core routers** forward packets to computer hosts within a network (but not between networks). Core routers are usually found in the nodes of IP backbone
Table 2.3: Power Dissipation of Routers at 40Gbps

<table>
<thead>
<tr>
<th>Router Model</th>
<th>Power Dissipation</th>
<th>Port Configuration</th>
<th>Power Dissipation per Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco CRS-1 16-Slot Single-Shelf System</td>
<td>9630</td>
<td>16</td>
<td>602W</td>
</tr>
<tr>
<td>Cisco CSR-1 8-Slot Single-Shelf System</td>
<td>4834</td>
<td>8</td>
<td>604W</td>
</tr>
<tr>
<td>Juniper T1600</td>
<td>8352</td>
<td>16</td>
<td>522W</td>
</tr>
<tr>
<td>Juniper T640</td>
<td>7296</td>
<td>8</td>
<td>912W</td>
</tr>
</tbody>
</table>

networks (i.e. in the WAN). They can be configured to help optimize end-user performance, network cost, or, in our case, network power consumption because the line a router forwards to is determined algorithmically by the current traffic load, congestion, line costs, and cooling costs, among other factors.

2. Edge routers route packets between a self-contained access network (MAN/LAN) and other access networks via a network backbone. The edge router can sit on the boundary between a MAN and WAN, or connect a LAN to a MAN or even to a WAN.

In Table 2.3, we list the power dissipation data of popular core routers [29, 26]. We approximate the power dissipation per connection pair $P_{router}$ with the maximum power dissipation of each model divided by its respective port configuration. The power consumption data of Cisco routers in Table 2.3 imply that the relation between port count and power consumption is linear; this assumes that the cooling cost of routers grows linearly as the number of ports grows. In other words, two CRS-1 16-Slot Systems consume twice the power as one CRS-1 16-Slot system on the same rack just as 200 CRS-1 16-Slot Systems consume twice the power as 100 CRS-1 16-Slot Systems in the same room. This most likely is not the case.

First, the connection network among routers at a node grows as the number of ports per node increases. We consider this network to be passive although there may be negligible power needs for equalization of copper cable or vixel laser transmitters driving fiber interconnects.

Secondly, as the density of routers in a closed space increases dramatically, air cooling becomes more power-intensive and expensive until it is no longer a feasible (or even physically possible) cooling option. Liquid cooling and mist cooling are capable
of withdrawing more heat from a densely packed space, but are energy expensive and not mature enough technologically to be deployed into current networks. However, to be conservative in our analysis, we will ignore the super-linear cooling curve and assume that the power dissipation per connection scales linearly with the number of ports in our network model.

2.5 Summary of Chapter 2

In this chapter, we laid the foundation for our thesis by introducing the basic components that make up a network: transmitter/receiver pairs (transceivers), erbium-doped fiber amplifiers (EDFA), switches (optical and optical-electronic-optical), and routers. For each component, we described how it is used and modelled in a network and provided its representative power consumption value.
Chapter 3

Wide Area Network Power Consumption Model

In this chapter, we build the WAN model in detail, beginning with the physical topology of the United States IP-backbone network. Then, we derive the traffic matrix used in dimensional analysis of the network (i.e. how many wavelengths, fibers, and ports are needed to serve all end-users according to projected usage patterns). Finally, we define network parameters based on technological capabilities (e.g. wavelengths per fiber, optical fiber reach, regeneration reach, etc).

At the end of the chapter, we discuss the various network designs and implementations that are compared. We first divide WAN designs into two broad cases, bypass (using switches at intermediate nodes) and non-bypass (using routers at intermediate nodes). We then present the four network infrastructure designs we analyze: patched bypass (in which direct, non-switched lightpaths are allocated between any node pair), bypass (in which lightpaths between any node pair are switched at intermediate nodes), non-bypass (in which lightpaths between any node pair are terminated and re-initiated by routers at intermediate nodes), and groomed non-bypass (in which whole lightpaths between any node pair are switched while residual or fractional lightpaths are terminated and groomed by routers at intermediate nodes).
3.1 Physical Topology

The physical architecture of an optical network consists of cable plants (which house fibers) that connect network nodes (OXC switches, O/E/O switches, or routers). We refer to this fixed, often irregular, physical equipment as the plant topology. A fiber topology sits on top of the plant topology. Both the plant and fiber topologies consist of a set of $N$ nodes corresponding to the switches and routers, usually placed at strategic cities, and a set of $E$ edges corresponding to the set of optical fiber links connecting the nodes. We are interested mainly in designing power efficient fiber topologies for the WAN.

We show in Figure 3-1 that the difference between the fiber and the plant topologies is the connection of the fibers in the plant topology. That is, the geographic layout of nodes A, B, C, and D and the optical fibers that connect them are fixed (shown on the left). However, the way in which those fibers actually connect nodes can be very different. For example, although nodes C and D are not directly connected in the physical layout, a patch panel at node B allows a direct lightpath from node C to D as if there were a direct path (albeit a longer distance path). Furthermore, we see that node A can communicate to node D via two hops $(A \rightarrow B$ followed by $B \rightarrow C$, shown in blue) but also via a direct patched bypass path $(A \rightarrow C$, shown in red). This is because the patch panelling at node B allows architecture designers to connect the fibers in creative ways to minimize network cost, or in our case, power consumption.

The physical topology we analyze is a 25-node, 56-link model of the United States WAN, based upon the AT&T Next-Generation IP/MPLS Backbone network, completed as of October 2008 [2]. The only direct connections in the physical topology are the 56 links between nodes, as shown in Figure 3-2. One goal of our research is to design a power efficient fiber topology for the graph (i.e. adding patch panels as needed for patched optical bypass). We will also explore the energy consumption of individual nodes and highly loaded links to argue that shortest path routing is more energy efficient than load balancing (which aims to equalize traffic across nodes and
Figure 3-1: Comparison of plant topology (shown in A) and fiber topology (shown in B). The plant topology shows physical nodes and geographic fiber layout. The fiber topology shows connections among the fibers. For example, although nodes C and D are not directly connected in the physical layout, a patch panel at node B allows a direct lightpath from node C to D as if there were a direct path.

3.2 Traffic Model

In this section, we build a detailed static, all-to-all optical flow traffic model based on projected user traffic demand and usage patterns, parametrized to each node pair. The traffic model is developed in two steps. We first create a population matrix $M_P = [P(i,j)]$ in which $P(i,j)$ indicates the expected number of users in node $i$ who wish to communicate with a user in node $j$. In the second step, we convert the population matrix into a traffic matrix $M_T = [T(i,j)]$ in which $T(i,j)$ indicates the estimated number of wavelengths needed to serve the expected traffic load between nodes $i$ and $j$, where $T(i,j) = F(P(i,j))$.

Population matrix $M_P$ is generated using the population of the area (served by a node) as a proxy for the traffic that node is responsible for originating and terminating. The traffic model is based on population distribution because we reason that a densely populated region generates more traffic than a sparsely populated region. Furthermore, of all the traffic requests generated by a node $i$, we assume they will be distributed proportionally among all the other nodes in the network so that a densely populated region served by node $j$ receives more requests from $i$ than sparsely
Figure 3-2: A 25-node, 56-link network representing the United States Wide Area Network, based upon the AT&T Next-Generation IP/MPLS Backbone network.

populated region served by node $k$.

We construct our population matrix using the following algorithm. For each node, take the metropolitan statistical index (MSI) to be a proxy for the population served by that node (define as $MSI(i)$). For each node $j$ connected to node $i$, we calculate its share of $MSI(i)$ by scaling $MSI(i)$ by $MSI(j)$ divided by the sum of the MSI of all nodes connected to $i$. We provide a sample population matrix calculation in Figure 3-3.

Given the above population matrix $M_P = [P(i, j)]$, we can then apply function $F$ to transform $M_P$ into traffic matrix $M_T = [T(i, j)]$. We first scale $P(i, j)$ so that the WAN serves the US population. This is because while the nodes in the WAN are the twenty-five largest, strategically chosen cities in the continental United States, the sum of their metropolitan statistical indices is less than 116 million residents. The population of other areas are not included in the MSI of these cities. As of 2009, the US population was estimated at over 307 million. Therefore, we must convert
Figure 3-3: Sample calculation of population demand $P(A, \cdot)$ for a four-node network. The traffic from any node to itself (i.e. $P(A, A)$) is zero; local user activity does not need to be transported across a WAN. The requests from node $A$ to any other node to which $A$ is connected is the population of $A$ multiplied by the destination node's population, relative to the total population of all possible destinations (e.g. $P(A, B) = MSI(A) \cdot \frac{MSI(B)}{MSI(B) + MSI(C) + MSI(D)}$). This calculation can be carried out for all node pairs so that the outgoing demand is $MSI(A)$ and the incoming traffic is $P(B, A) + P(C, A) + P(D, A)$.

$T(A, A) = 0$

$T(A, B) = \frac{P(A)P(B)}{P(B) + P(C) + P(D)}$

$T(A, C) = \frac{P(A)P(C)}{P(B) + P(C) + P(D)}$

$T(A, D) = \frac{P(A)P(D)}{P(B) + P(C) + P(D)}$

$2.65 \cdot P(i, j)$ to $T(i, j)$, not simply $P(i, j)$, to avoid under-provisioning in the network. We can then convert the scaled number of users into fractional wavelengths based on their expected network activity. We define $q$ to be the probability a user is active on the network at any time, $p$ to be the probability an active user is sending or receiving traffic, and $r$ to be the bitrate. Then, $F = 2.65 \cdot p \cdot q \cdot r$ and $T(i, j) = 2.65 \cdot P(i, j) \cdot p \cdot q \cdot r$.

Almost all users of a network currently own, or will come to own over the next few years, at least one device (smartphone, laptop, etc.) that converses with the network, regardless of whether the user is actively using the device. For our purposes, all such devices are considered to be “active”; thus, we assume $q = 0.1$. We further assume $p = 0.1$ and $r = 100$Mb/sec. Assuming 10Gb/sec optical fibers, this specification corresponds to a network with expected static traffic flow of approximately 300-Terabit/sec (without network protection), compared to current AT&T network capacity of 30-Terabit/sec (with network protection). This corresponds to approximately a hundredfold increase in expected traffic demand. In Table 3.1, we provide an excerpt of actual calculations for a traffic matrix entry where node $i$ (the originating sender node) is Chicago.

Our resulting traffic matrix is based on the following two assumptions:
Table 3.1: Sample Traffic Demand from Chicago to All Other Nodes

<table>
<thead>
<tr>
<th>Node j</th>
<th>MSI(i,j)</th>
<th>P(i,j)</th>
<th>T(i,j)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cleveland</td>
<td>2,250,871</td>
<td>207,782</td>
<td>55.063</td>
</tr>
<tr>
<td>Orlando</td>
<td>2,082,628</td>
<td>192,254</td>
<td>50.947</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>5,838,471</td>
<td>538,967</td>
<td>142.826</td>
</tr>
<tr>
<td>San Antonio</td>
<td>2,031,445</td>
<td>187,529</td>
<td>49.695</td>
</tr>
<tr>
<td>Phoenix</td>
<td>4,364,034</td>
<td>402,858</td>
<td>106.757</td>
</tr>
<tr>
<td>Washington, DC</td>
<td>5,400,000</td>
<td>498,491</td>
<td>132.100</td>
</tr>
<tr>
<td>St. Louis</td>
<td>2,828,990</td>
<td>266,153</td>
<td>69.206</td>
</tr>
<tr>
<td>Boston</td>
<td>4,522,858</td>
<td>417,519</td>
<td>110.643</td>
</tr>
<tr>
<td>Sacramento</td>
<td>2,136,604</td>
<td>197,237</td>
<td>52.268</td>
</tr>
<tr>
<td>Dallas</td>
<td>6,477,315</td>
<td>597,941</td>
<td>158.454</td>
</tr>
<tr>
<td>Nashville</td>
<td>1,666,566</td>
<td>153,846</td>
<td>40.769</td>
</tr>
<tr>
<td>Chicago</td>
<td>9,785,747</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>15,250,000</td>
<td>147,775</td>
<td>373.060</td>
</tr>
<tr>
<td>San Francisco</td>
<td>4,203,898</td>
<td>388,075</td>
<td>102.840</td>
</tr>
<tr>
<td>Raleigh</td>
<td>1,125,827</td>
<td>103,926</td>
<td>27.541</td>
</tr>
<tr>
<td>Denver</td>
<td>2,552,195</td>
<td>235,601</td>
<td>62.434</td>
</tr>
<tr>
<td>Seattle</td>
<td>3,407,848</td>
<td>314,589</td>
<td>83.366</td>
</tr>
<tr>
<td>San Diego</td>
<td>2,880,000</td>
<td>265,862</td>
<td>70.453</td>
</tr>
<tr>
<td>Atlanta</td>
<td>5,475,000</td>
<td>505,414</td>
<td>133.935</td>
</tr>
<tr>
<td>Kansas City</td>
<td>2,053,928</td>
<td>189,604</td>
<td>50.245</td>
</tr>
<tr>
<td>Houston</td>
<td>5,867,489</td>
<td>541,646</td>
<td>143.536</td>
</tr>
<tr>
<td>Portland</td>
<td>2,217,325</td>
<td>204,688</td>
<td>54.242</td>
</tr>
<tr>
<td>New Orleans</td>
<td>1,235,650</td>
<td>114,067</td>
<td>30.228</td>
</tr>
<tr>
<td>Albany</td>
<td>19,006,798</td>
<td>1,754,577</td>
<td>464.963</td>
</tr>
<tr>
<td>Salt Lake City</td>
<td>1,130,293</td>
<td>104,341</td>
<td>27.650</td>
</tr>
</tbody>
</table>

1. Traffic from a node to itself is zero. This assumption is motivated by the observation that traffic within a closed network such as a MAN or LAN do not need to be transported across the WAN. In fact, if $T(A, A)$ did indeed consume resources of a core router at $A$, we can immediately improve the power consumption of the network by withdrawing this route.

2. The MSI is a valid proxy for traffic demand. Furthermore, the MSI of the twenty-five metro areas covered by our network can be uniformly scaled to serve the US population. This assumption forces us to ignore potential differences in network users (e.g., we assume a farmer in the Midwest generates
statistically identical network activity as an MIT student would). We also ignore traffic hotspots that are uncorrelated with region population (e.g. locations of data centers). We acknowledge these two weaknesses in using the MSI to generate traffic matrices, but argue that this model is still one of the most accurate approximations for test networks among published studies. By using the population to determine traffic demand, we have controlled for the biggest influence on demand. Furthermore, by introducing variance on that demand that is a function of the population, we can scale the entire network in a meaningful manner. We compare the MSI-based traffic model to those used in previous research in which traffic is generated either using uniform all-to-all traffic calculations or using a random traffic generator based upon a uniform distribution “centered at an identical average” with identical variance [13, 28]. The first method is valid only for MAN’s in which the region served by the entire network is homogeneous. The second method ignores the inherently irregular, heterogeneous nature of WAN’s completely. That is, even though the traffic is generated randomly, the generating function is independent of the node characteristics and can be described as adding noise to an otherwise uniform all-to-all traffic model.

Finally, we note that the traffic demand of the network, based on the assumptions we have made in this section, is much larger than that served by current networks. Looking at the number of wavelengths that need to be provisioned between Chicago and the rest of the US network in Table 3.1, we argue that, intuitively, any power savings from using traffic grooming are minimal. For example, Chicago traffic towards Los Angeles requires 373.06 wavelengths. The 0.06 fraction of a wavelength that would have been able to be groomed to maximize wavelength utilization is almost negligible compared to the 373 full wavelengths and the number of components (amplifiers, transceivers, and switch/router ports) needed to support those wavelengths. That the research we conduct is in a new, more voluminous regime of traffic than had been previously considered is one differentiating aspect of our research approach, results, and conclusions when compared to previous research.
3.3 Network Parameters

In Chapter 2, we introduced the parametric modelling of network components and gave representative power consumption values for each parameter. In this section, we define more network parameters and assign values to those parameters we had introduced previously but had not discussed in detail.

Regeneration Limit

We introduced \textit{regenLimit} in Chapter 2 as the maximum distance an optical signal can travel, with amplification, before needing to be regenerated due to a poor signal-to-noise ratio.

In our research, \textit{regenLimit} is perhaps the most interesting parameter because increasing the \textit{regenLimit} is one of the most active areas of network infrastructure research today. A longer distance before regeneration (or equivalently, a longer signal transmission or optical reach) can reduce the number of transceivers used in the network tremendously and therefore reduce power consumption.

Because the path lengths (node-to-node) range from 121km for the shortest edge to 4820km for the longest path, we allow \textit{regenLimit} to float between 100km and 5000km. The power consumption of networks is presented as a function of \textit{regenLimit} so that we can study how improved optical reach can influence infrastructure design decisions. We note that most of the current network infrastructure operates at \textit{regenLimit} \approx 500km for 10Gb/sec lines. However, research has shown that networks with \textit{regenLimit} \geq 1500km can easily be deployed, with experimental data showing that \textit{regenLimit} can reach up to 2500km for 100Gb/sec lines \cite{19}.

Wavelengths per Fiber

We define \textit{lambdaFiber} as the number of wavelengths per optical fiber. We set \textit{lambdaFiber} = 200 and ignore quantization effects in our modeling (i.e. we fully prorate the amplification power per wavelength regardless of how many wavelengths are used per fiber). We acknowledge that the decision to ignore quantization effects
can bias our power consumption results downwards. However, we argue that the effects for a 25-node, 300-Terabit/sec network are small (i.e. well within one order of magnitude error).

If the number of wavelengths on a fiber is close to \( \lambda_{Fiber} \) (e.g. 150 or more fibers), the fiber can be considered to be well utilized; if the number of wavelengths is much smaller than \( \lambda_{Fiber} \) (e.g. 10 or less), the traffic on these wavelengths can be routed along alternate paths (to avoid using a new fiber) or be given their own fiber at additional amplification cost and power. If the excess number of wavelengths (modulo \( \lambda_{Fiber} \)) is distributed uniformly from 0 to 200 (which is plausible since traffic demand among several nodes may have their own wavelengths but share fiber connections, making the utilization of fiber cables relatively independent of population demographics), the expected number of wavelengths per fiber is 100, leading to a factor of two underestimation in the power consumption of amplifiers.

**Amplifier Reach**

We introduced \( linkReach \) in Chapter 2 as the maximum distance an optical signal can travel before needing amplification to boost the signal-to-noise ratio. We set \( linkReach = 50 \text{km} \).

We further define \( lastLinkReach \) as the maximum distance an optical signal can travel without amplification, right before regeneration. That is, if the distance between the last amplifier and a regenerator on a fiber is greater than \( linkReach \) but less than \( lastLinkReach \), we allow the signal to skip the last amplification to save power. We set \( lastLinkReach = 100 \text{km} \).

**OXC Switch Loss**

We define \( OXCLoss \) as the equivalent extra distance an optical signal travels for each \( OXC \) it traverses due to the loss across the \( OXC \). We assume a 4dB loss per \( OXC \). Assuming 20dB gain per EDFA and \( linkReach = 50 \text{km} \), each kilometer of fiber is equivalent to a 0.4dB loss. Therefore, \( OXCLoss = \frac{4dB}{20dB/\text{linkReach}} = 10 \text{km} \).
3.4 Bypass versus Non-bypass

In this section, we discuss the two broad cases of WAN network designs: optical bypass and non-bypass.

Currently, most traffic is routed in a non-bypass fashion in which signals from node $i$ to node $j$ are terminated and re-initiated at routers in all intermediate nodes $k$. The main benefit from using non-bypass is that traffic demands between various node pairs can be shared on a common lightpath via multiplexing. In other words, if traffic between nodes $i$ and $j$ is not sufficient to fill a lightpath and traffic between nodes $m$ and $n$ is also not sufficient to fill its own lightpath, but the traffic shares a common subpath $e$ that can be filled with their combined traffic, non-bypass allows network designers to allocate a single common subpath for shared traffic. This can save up to half the number of components (amplifiers and regenerators) when compared with allocating the two lightpaths separately. The more lightpaths that share a common edge in this manner, the greater the number of components saved.

However, non-bypass requires the use of routers, which consume far more power per port when compared to the per port power consumption of a switch. Therefore, when the traffic is sufficiently large so that wavelength utilization is high, optical (or direct) bypass is preferred. In this design, signals are switched at intermediate nodes until they reach the destination node. Optical bypass in large traffic demand cases is a highly flexible, easily implemented, and power efficient design. Furthermore, in the event that traffic demand is large and static, direct bypass can be improved upon by hard-wiring all lightpaths. In other words, patch panelling can be installed at all nodes so that switching costs can be minimized. This approach minimizes power consumption when the traffic demand is high and a large (integer) portion of that traffic is static and can be statically provisioned. This “patched” bypass state is the lower bound on direct bypass energy consumption.
3.5 WAN Network Designs

In this section, we analyze the four network designs we consider and point out any distinctive characteristics the designs have that may affect power consumption.

Our basic research set-up consists of six bypass implementations and six non-bypass implementations. In Table 3.2, we show the equipment at originating and receiving nodes for each design. In Table 3.3, we show the equipment at each intermediate node for each design. We explain the functionality of each design by discussing a specific implementation from each of the four cases. Figure 3-4 is a schematic of how the various designs function. Each design case can further be implemented using OXC, OEO\textsubscript{OXC}, or OEO\textsubscript{ELEC} switches for a total of twelve separate design implementations that can be tested.

Table 3.2: WAN Network Designs – Components at End Nodes

<table>
<thead>
<tr>
<th>Patched Bypass</th>
<th>Bypass</th>
<th>Non-bypass</th>
<th>Groomed Non-bypass</th>
</tr>
</thead>
<tbody>
<tr>
<td>OXC Switch</td>
<td>OXC Switch</td>
<td>OXC Switch</td>
<td>OXC Switch and Router</td>
</tr>
<tr>
<td>OEO\textsubscript{OXC} Switch</td>
<td>OEO\textsubscript{OXC} Switch</td>
<td>OEO\textsubscript{OXC} Switch</td>
<td>OEO\textsubscript{OXC} Switch and Router</td>
</tr>
<tr>
<td>OEO\textsubscript{ELEC} Switch</td>
<td>OEO\textsubscript{ELEC} Switch</td>
<td>OEO\textsubscript{ELEC} Switch</td>
<td>OEO\textsubscript{ELEC} Switch and Router</td>
</tr>
</tbody>
</table>

Table 3.3: WAN Network Designs – Components at Intermediate Nodes

<table>
<thead>
<tr>
<th>Patched Bypass</th>
<th>Bypass</th>
<th>Non-bypass</th>
<th>Groomed Non-bypass</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>OXC Switch</td>
<td>OXC Switch</td>
<td>OXC Switch and Router</td>
</tr>
<tr>
<td>-</td>
<td>OEO\textsubscript{OXC} Switch</td>
<td>OEO\textsubscript{OXC} Switch</td>
<td>OEO\textsubscript{OXC} Switch and Router</td>
</tr>
<tr>
<td>-</td>
<td>OEO\textsubscript{ELEC} Switch</td>
<td>OEO\textsubscript{ELEC} Switch</td>
<td>OEO\textsubscript{ELEC} Switch and Router</td>
</tr>
</tbody>
</table>

Patched Bypass Network Design

We explain the patched bypass design by looking at the patched bypass design implemented with OXC switches. This all optical network design is the absolute lowest bound on power consumption. Originating traffic from node $i$ to node $j$ is inserted into the WAN network by an OXC switch. Once in the network, the flow traffic has dedicated lightpaths. At each intermediate node the signal passes, the signal is left intact by patch panelling at the node. The signal is amplified and regenerated as needed, taking the entire path length from $i$ to $j$ as a single “virtual” link. Using
OEO_{OXC} or OEO_{ELEC} switch implementations do not impact any component or network functions. This static design may present difficulties in protection switching and a pragmatic network using patch panelling will be a hybrid of patched bypass and standard bypass designs.

**Bypass Network Design**

We explain the bypass design by looking at the bypass design implemented with OEO_{OXC} switches. Originating traffic from node $i$ to node $j$ is inserted into the WAN network by an OEO_{OXC} switch. At each intermediate node in the network, the OEO_{OXC} switch can select the next edge the signal should be sent out on. Because OEO_{OXC} switches have two transceivers (see Chapter 2, Switch Power Consumption), regeneration is pegged to occur at each intermediate node (what we refer to as minimum, mandatory regeneration instances). Additional regeneration and any amplification occurs on the network edges as needed. Note that using the OXC switch implementation results in an OXC_{Loss} incurred with each switch traversed; however, we are able to float regenerations along the entire length of the path. As a result, the OXC switch implementation may save slightly on transceiver usage while expending more power on amplifiers when compared with the O/E/O switch implementations.

**Non-bypass Network Design**

We explain the non-bypass design by looking at the non-bypass design implemented with OEO_{ELEC} switches. Under a non-bypass design, the optical signal is terminated and re-initiated by the router at each intermediate node. All traffic enters the WAN through the router at source node $i$ and exits through the router at the destination node $j$. Minimum regenerations are then pegged to occur at each intermediate node. Using OXC or OEO_{OXC} switch implementations do not impact any component or network functions.
Groomed Non-bypass Network Design

Groomed non-bypass is a bit of a misnomer. The groomed non-bypass design actually is a hybrid between the bypass and non-bypass designs. That is, we attempt to merge the power savings from optical bypass with the flexibility of traffic grooming from non-bypass. This allows us to account for reduced component usage and power savings through traffic grooming.

We explain the groomed non-bypass design by looking at the groomed non-bypass design implemented with $OEO_{OXC}$ switches. The traffic from node $i$ is separated into full lightpaths and partial lightpaths (e.g. if 11.11 wavelengths are needed for traffic from node $i$ to node $j$, there are eleven full lightpaths and 0.11 partial lightpaths). The full lightpaths are switched in a bypass model so that eleven wavelengths enter and exit the switches at each intermediate node. The partial lightpaths are routed in a non-bypass model through the router at each intermediate node. The power savings compared to using a standard non-bypass design come from provisioning fewer router ports since the routers are used only for residual traffic grooming purposes.

3.6 Summary of Chapter 3

In this chapter, we built the WAN model by discussing the physical topology, traffic model, network parameters, and potential network designs. Our traffic model is based upon network demographics so that the expected traffic between any two nodes is a function of the population of each node’s metropolitan area (MSI). Then, nodes serving densely populated regions both originate and terminate a larger proportion of overall network traffic compared with nodes serving sparsely populated regions. Variance in the traffic demand between any node pair reflects a likely quote of network fluctuation. We also described four main network designs. Under the patched bypass design, direct, non-switched lightpaths are allocated between any node pair. With a standard bypass network, lightpaths between any node pair are switched at intermediate nodes. Under a non-bypass design, lightpaths between any node pair are terminated and re-initiated by routers at intermediate nodes. Finally, under a
groomed non-bypass design, whole lightpaths between any node pair are switched while residual or fractional lightpaths are terminated and groomed by routers at intermediate nodes.
Figure 3-4: A schematic diagram of the four WAN designs. Under the patched bypass design, direct lightpaths are set up between any two nodes with patch panelling at intermediate nodes. Under the bypass design, signals between any two nodes are switched at all intermediate nodes. Under the non-bypass design, signals are terminated at routers in intermediate nodes and re-initiated at each intermediate node. Under the groomed bypass design, traffic between two nodes that fills whole lightpaths are switched at intermediate nodes (like it would be under a bypass design), but residual traffic that fills only a partial lightpath is terminated at intermediate nodes (like it would be under a non-bypass design).
Chapter 4

WAN Power Analysis and Results

In this chapter, we present the results and recommendations based upon our research on WAN's. We first focus on analyzing the network with the traffic model presented in Chapter 3. We discuss our findings for the base network and for the network with cooling overhead (in which we allocate 100% power overhead to allow for cooling of components). We then break the power consumption of each network design into its component constituents to analyze the power distribution in each architecture design.

That analysis can be repeated for stressed (or highly loaded) traffic models. We present the power consumption analysis (with cooling overhead) and power distribution analysis for a network in which the traffic matrix is loaded with $k$ standard deviations of traffic between all node pairs. We compare the power consumption of the designs under the expected value of a static traffic load to the designs under a more dynamic traffic load to show that our recommendations are also valid when the network needs to be over-provisioned.

In this chapter, we also delve into the effect of reliability protection on network power consumption. We compare the 1-plus-1 spare lightpath design for patched bypass networks (in which, when a signal fails to transmit along any intermediate edge along the virtual lightpath, the entire path is switched onto a second mirror path at end nodes) with the reliability methods available for optical switching (e.g. $N$-plus-1 spare edge design in which a spare link can cover all the lightpaths that traverse the edge).
Finally, we motivate that the popular networking goal of equally distributing traffic load among all paths (so as to balance power consumptions at each node) is not power minimizing [28]. That is, under a bypass model, moving traffic off highly loaded shortest paths to a path that has more available capacity but is longer in length consumes strictly greater power than using shortest path or shortest hop routing for all traffic. Under a non-bypass model, traffic balanced routing consumes at least as much power as shortest path routing.

As a final note, we would like to clarify a few points about the traffic model and network graph. The traffic load is static and all requested lightpaths are available. In other words, allocation and provisioning of wavelengths are performed off-line, and network capacity is increased as needed by growth in traffic demand. Therefore, to minimize blocking when network demand may exceed capacity, over-provisioning must be incorporated into the network infrastructure. However, such over-provisioning can also be considered inefficient for two reasons: first, the links will be under-utilized any time the traffic demand is less than the maximum traffic capacity and second, all network components must remain fully powered regardless if they are actively being used so that the fixed power consumption at any time for a heavily over-provisioned network is high.

In the network graph, the length of links between nodes in the physical topology is determined by geographic shortest distance ("as the crow flies") using the latitude and longitude of node locations. These link distances may underestimate the actual length of the fiber, since optical fiber is often laid along major highways (i.e. actual link lengths may correspond more to driving distance). However, since common network research practice is to use the geographic shortest distance, we model our WAN using the same standard. Then, to find path distances between any two nodes $A$ and $B$ via shortest path routing, we run Dijkstra on the network graph. The shortest path between $A$ and $B$ must equal the shortest path from $B$ to $A$, despite possible differences in traffic demand from $A$ to $B$ and $B$ to $A$. 

54
4.1 Analyzing Power Consumption with Baseline Traffic Model

In this section, we analyze the network using the traffic model presented in Section 3.2. We will refer to this set of analysis as the “Baseline Cases” because we use the static traffic demand matrix formed by taking the expectation of the population served by each node and the populace’s usage patterns. The baseline traffic demands correspond to a 300-Terabit/sec aggregate network capacity, compared to today’s 30-Terabit/sec IP Backbone network.

4.1.1 Measuring Baseline Power Consumption

We subject each of the four network designs (patched bypass, bypass, non-bypass, and groomed non-bypass) to the baseline traffic demand on the U.S. Next-Generation IP/MPLS Backbone network. We implement each design using each of three switches (using all OXC switches, all O/E/O\textit{OXC} switches, or all O/E/O\textit{ELEC} switches). In Figure 4-1, we graph a comparison of the power consumption of each network implementation as a function of \textit{regenLimit}.

In the top graph, we plot the power consumption of each network design implementation on a log-log scale. We find that the patched bypass design, regardless of the switch type used, is the most power efficient architecture. In fact, the patched bypass design consumes between 20 to 25 times less power than the bypass design, under any switch implementation and at any \textit{regenLimit}. The patched bypass design is between 2000 to 7000 more efficient than the non-bypass design using OXC switches (consuming 2500 times less power at \textit{regenLimit} = 500km), between 1000 to 2500 times more efficient using O/E/O\textit{OXC} switches (consuming 1500 times less power at \textit{regenLimit} = 500km), and between 1500 to 4000 times more efficient using O/E/O\textit{ELEC} switches (consuming 2000 times less power at \textit{regenLimit} = 500km).

The power consumed by router ports in the non-bypass design raises serious doubts on the scalability of this architecture. Even when compared to the standard opti-
Figure 4-1: A comparison of the network power consumption for four network designs (Patched Bypass, Bypass, Groomed Non-bypass, and Non-bypass) implemented with each of three possible switches ($OXC, O/E/O_{OXC}, O/E/O_{ELEC}$). Power consumption is based on a 300-Terabit/sec network. The top graph plots power consumption on a log-log scale. The lower graph plots power consumption on a log-linear scale to underscore the large power gap between non-bypass and every other design.
cal bypass architecture design, the non-bypass design consumes between 100 to 300 times more power under using \( OXC \) switches (consuming 120 times more power at \( \text{regenLimit} = 500\text{km} \)), between 50 to 100 times more power using \( O/E/O_{OXC} \) switches (consuming 64 times more power at \( \text{regenLimit} = 500\text{km} \)), and between 50 to 150 times more power using \( O/E/O_{ELEC} \) switches (consuming 85 times more power at \( \text{regenLimit} = 500\text{km} \)). In the bottom graph of Figure 4-1, we plot the power consumption of each network design implementation on a log-linear scale. We emphasize the large gap between the non-bypass design and any other architecture designs. In fact, the differences in power consumed among the patched bypass, bypass, and groomed non-bypass designs are barely distinguishable compared to the differences between these designs and the non-bypass design.

In Figure 4-2, we graph the comparison of the power consumption (with additional power consumption due to cooling overhead in transceivers, switches, and routers factored into total power requirements) of each network implementation as a function of \( \text{regenLimit} \). Koomey estimates that routers require an estimated 100% of operational energy intake for cooling purposes [22]. His estimations are taken from air-cooled server farms. As the router density increases with hundredfold increase in traffic load, air cooling will no longer be sufficient. Liquid or mist cooling are expected to be extremely complex and expensive in both power and cost; however, because the technology is relatively new, we have no reliable data on the power demands of these cooling technologies. As a result, we continue to assume a 100% cooling overhead for all network components except for amplifiers. Because transceivers in fibers and switches in nodes can be cooled relatively efficiently and are numerous in the network, we feel the sub-100% cooling overhead for those components can contribute to balancing the possibly super-linear cooling curve of the router ports.

In the top graph, we plot the power consumption (accounting for cooling overhead) of each network design implementation on a log-log scale. We find that the patched bypass design, regardless of the switch type used, is still the most power efficient architecture. The patched bypass design again consumes between 20 to 25 times less power than the bypass design, under any switch implementation and at any \( \text{regenLimit} \).
Figure 4-2: A comparison of the network power consumption for four network designs (Patched Bypass, Bypass, Groomed Non-bypass, and Non-bypass) implemented with each of three possible switches \((OXC, O/E/O_{OXC}, O/E/O_{ELEC})\), when a 100% power overhead is assumed for transceivers, switches, and routers. Power consumption is based on a 300-Terabit/sec network. The top graph plots power consumption on a log-log scale. The lower graph plots power consumption on a log-linear scale to underscore the large power gap between non-bypass and every other design.
When compared to the non-bypass design, the power savings achieved using patched bypass is even greater. The patched bypass design is between 2000 to 12000 more efficient than the non-bypass design using \textit{OXC} switches (consuming 3000 times less power at \textit{regenLimit} = 500km), between 1000 to 3000 times more efficient using \textit{O/E/O\textit{OXC}} switches (consuming 1600 times less power at \textit{regenLimit} = 500km), and between 1500 to 5000 times more efficient using \textit{O/E/O_{ELEC}} switches (consuming 2200 times less power at \textit{regenLimit} = 500km).

In the bottom graph of Figure 4-2, we plot the power consumption (accounting for cooling overhead) of each network design implementation on a log-linear scale. We emphasize the large gap between the non-bypass design and any other architecture design. The differences in power consumed among the patched bypass, bypass, and groomed non-bypass designs are again barely distinguishable compared to the differences between these designs and the non-bypass design.

Figures 4-1 and 4-2 suggest several interesting recommendations. First, when traffic demand is high and stable, network architecture should be dominated by patched bypass. In other words, when a substantial but fixed fraction of the traffic matrix is well-defined and predictable (e.g. comprises mostly of circuit-based connections), this fraction should be routed using a patched bypass architecture. The routing of the remaining bursty, high variance traffic and network protection can be implemented on a bypass or groomed bypass network. The two networks in hybrid can deliver significant power savings without sacrificing quality of service or network flexibility.

However, we must point out that using patched bypass is only feasible for relatively static traffic. The number of wavelengths that are patch panelled between any pair of nodes are fixed and costly to change. When traffic has high variance, we must either over-provision the number of fixed lightpaths leading to low link utilization, or depend heavily on a hybrid network (i.e. the majority of the traffic matrix is routed on an optical bypass or groomed non-bypass network) which decreases the potential power savings. Even if the aggregate network traffic does not fluctuate (i.e. overall traffic is 300-Terabit/sec) but the internal traffic demand shifts among nodes (e.g. on certain days, north to south traffic is prevalent but on other days east to west traffic
forms the bulk of the overall traffic demand), patched bypass is difficult to deploy and can only be used for the minimum “core” static traffic.

The second observation is that the optical bypass and groomed non-bypass designs have very similar power consumption characteristics because almost all traffic in the groomed non-bypass design are routed using optical bypass. Only residual fractions of wavelengths are routed using traffic grooming. As the traffic demand matrix scales upward, the difference in power consumption between the two designs becomes even more negligible. However, at all values of \( \text{regenLimit} \), we find that the groomed non-bypass design consumes more power than the optical bypass design. Note that in our model of groomed non-bypass, the residual traffic is in fact fully traffic groomed; each fractional wavelength is given the prorated fraction of fiber, switch, and/or router power consumption so that the power calculations fully mimic the operation of the groomed network described by Shen [28]. We find that even the most sparing use of routers (so as to only provision for fractional wavelength traffic capacity on each lightpath) is more power intensive than a full optical network, due the high power needs of routers. Therefore, our findings suggest that bypass designs dominate groomed non-bypass designs.

We have pointed out earlier how power intensive the non-bypass architecture design is. We note that the most striking characteristic of non-bypass networks is the high fixed power consumption of network ports. At over 600 Watts per pair of ports, the router far overshadows any other component’s power demands. As a result, any potential power savings from increasing optical capacity (i.e. reducing number of transceivers used by increasing \( \text{regenLimit} \)) is negligible compared to the routers’ power needs. Therefore, when the traffic demand is high (as with optical flow switched traffic), the non-bypass design is dominated by all the other designs, under any switch implementation.

Finally, we comment on the importance of increased optical reach. Using coherent detection, Renaudier has shown that long-haul optical signals can travel for 1500 or 2500km before needing regeneration [19]. When compared to current network technology that regenerates signals approximately every 500km, we consume
between 30-60% less power under all network designs except under the non-bypass (which we argued is not scalable with respect to power due to the high fixed power consumption of router ports). Such power savings comes with another added benefit; fewer regenerations are synonymous with faster end-to-end transmission for users since intermediate signal processing at accessory transceivers can be eliminated.

4.1.2 Measuring the Effect of Switch Choices

We compare the power efficiency of using each of the three switches available (OXC, O/E/O\textsubscript{OXC}, and O/E/O\textsubscript{ELEC}). In Figure 4-3 and 4-4, we plot the effects of switch choice and cooling overhead for each design choice. Figure 4-3 is graphed on a log-log scale while Figure 4-4 is graphed on a log-linear scale. Each graph shows the power consumption of a network design using only OXC switches, then using only O/E/O\textsubscript{OXC}, and finally O/E/O\textsubscript{ELEC} switches in solid lines; the dashed lines show the power consumption of that same network implementation when cooling overhead is factored in.

We observe that the three power consumption curves for each switch choice are parallel and that OXC switches consistently outperform O/E/O\textsubscript{ELEC} switches, which consume much less power than O/E/O\textsubscript{OXC} switches. The difference in power efficiency grows as \texttt{regenLimit} increases. This phenomena is especially prevalent in the patched bypass and bypass designs. This effect is most likely due to the decreased use of transceivers (which consume equivalent power under any switch implementation) that allows the power difference of the switch choices to be highlighted. Note that the transceivers used in a switch (an O/E/O\textsubscript{OXC} switch contains two transceivers and an O/E/O\textsubscript{ELEC} switch contains one) are not accounted for in the transceiver cost, but their power consumption are considered part of the switch power, since the switch is not fully functional without the transceiver(s) and we do not wish to double count the transceiver power.

This transceiver power assignment can explain most of the power difference we observe. The O/E/O\textsubscript{OXC} switches are the most power intensive because they operate two extra transceivers compared to OXC switches. The O/E/O\textsubscript{ELEC} switches
Figure 4-3: A comparison of the network power consumption (for four network designs implemented with each of three possible switches) to the power consumption when a 100% power overhead is assumed for transceivers, switches, and routers. Power consumption is based on a 300-Terabit/sec network. Solid lines denote baseline power consumption without cooling; dashed lines denote baseline power consumption with cooling overhead. The top left, top right, bottom left, and bottom right graph shows the effects on a patched bypass, bypass, groomed non-bypass, and non-bypass design, respectively.
Figure 4-4: A comparison of the network power consumption (for four network designs implemented with each of three possible switches) to the power consumption when a 100% power overhead is assumed for transceivers, switches, and routers. Power consumption is based on a 300-Terabit/sec network and plotted on a linear scale. Solid lines denote baseline power consumption without cooling; dashed lines denote baseline power consumption with cooling overhead. The top left, top right, bottom left, and bottom right graph shows the effects on a patched bypass, bypass, groomed non-bypass, and non-bypass design, respectively.
consume only slightly more power than OXC switches despite containing an extra transceiver because electronic switches are very power efficient (e.g. each electronic switch port consumes only a third of the power of an OXC switch port). Since O/E/O_{ELEC} switches can perform the same function as O/E/O_{OXC} with lower power needs, O/E/O_{ELEC} switches dominate O/E/O_{OXC} switches and O/E/O_{OXC} switches should be avoided altogether in future networks. When OXC switches may be used, we recommend that they be deployed because they offer a 25-50% power savings over O/E/O_{ELEC} switches in the patched bypass network design and 30-60% power savings in the bypass design, for \textit{regenLimit} ranging from 500-2000km.

4.1.3 Measuring the Effect of Cooling Overhead

We compare the effect of adding cooling overhead to the network designs by interposing the cooled and non-cooled power consumption curves. The effects of switch choice and cooling overhead for each design choice can be seen in Figure 4-3 and 4-4.

At low values of \textit{regenLimit} (i.e. \textit{regenLimit} \leq 500km), the power consumption curves are all parallel, as might be expected. However, in the patched bypass and bypass designs, as \textit{regenLimit} increases further, we find that the power demand of an O/E/O_{ELEC} implementation with cooling overhead actually approaches the power demand of an O/E/O_{OXC} implementation without cooling considerations! The power demand of the OXC switched implementations with cooling overhead outperforms both O/E/O_{OXC} and O/E/O_{ELEC} switched implementations without cooling. This phenomena can be explained by counting the number of transceivers used in each design implementation. As the optical reach increases, the number of transceivers required for regeneration decreases so that at \textit{regenLimit} = 5000km, the difference between solid and dashed curves is due solely to switch choice and indicates how much of the network power demand is consumed by switches.

In the groomed non-bypass and non-bypass designs, we see that the power consumption curves are parallel, regardless of the value of \textit{regenLimit}. The differences in power consumption among the implementations still grow as \textit{regenLimit} increases, but the level of increase is constant between the power consumption regardless of
cooling overhead. In other words, under the groomed non-bypass and non-bypass designs, power consumption is dominated by transceivers, switches, and routers; then, when we include 100% power consumption for cooling, the power consumption curves are scaled by approximately two, with relationships among the curves perfectly maintained.

Finally, we state that all the analysis hereafter is run on network design implementations with cooling overhead. It can be misleading to discuss results and make recommendations based on incomplete network models. All network infrastructure must consider the challenge of cooling the equipment and so to discuss networks that exclude such cooling concerns is inaccurate. Hereafter, networks with cooling overhead will be simply referred to as networks. Furthermore, we will focus our analysis on OXC and O/E/OELEC switched patched bypass, bypass, and (as comparison) non-bypass implementations since we have argued that O/E/OOXC switches are dominated in function and power efficiency by these switches just as the bypass design dominates groomed non-bypass designs.

4.1.4 Measuring the Distribution of Power Consumption Among Components

We have mentioned previously that routers ports are very power intensive, that O/E/OOXC switches are less power efficient than other switches, etc. In this section, we take a close look at the breakdown of total power into the fractions used by amplifiers, transceivers, switches, and routers across network designs and as functions of regenLimit. The goal is to better understand the relationship among components in a WAN so that we can appropriately target areas for power reduction.

If we look at the raw power consumption of each component, the router dominates over any other component. Based upon the components we selected in Chapter 2, each amplifier is 0.09W per wavelength. The power consumption of an OXC, O/E/OELEC, and O/E/OOXC switch port equals the energy needs of 5, 27, and 56 amplifiers, respectively. A transceiver consumes 25 times the power of an amplifier
while a router port consumes 6688 times! Based upon these raw comparisons, the biggest power savings come from reducing the number of router ports, using efficient switches, and minimizing the number of transceivers. We hypothesis that these three goals can be reconciled in a patched bypass, bypass, or hybrid patched bypass/bypass network design using OXC or O/E/OELEC switches while simultaneously increasing the \( \text{regenLimit} \geq 1000 \text{km} \). To test the hypothesis, we look at the actual distribution of power by wattage among all components because the relative number of components in each network design implementation is different. For example, the patched bypass network with OXC switches has a minimal number of overall components whereas the non-bypass network has a maximal number of router ports.

In Figure 4-5, we plot the ratio of component power to total network power across four network designs implemented with OXC switches, as a function of \( \text{regenLimit} \). The solid black line shows how much the total network power consumption falls (when compared to \( \text{regenLimit} = 100 \)) as \( \text{regenLimit} \) increases, indicating the power savings due to improvements in optical reach. The shape of the curve is essential in properly understanding the evolution of power distribution among components, so that relative increases in power used can be read in the context of overall power savings.

In the top two graphs, we plot the power distribution for OXC switched patched bypass and bypass designs. We find that transceivers consume the bulk of total power. In fact, at \( \text{regenLimit} = 500 \text{km} \), over 75% of the power is consumed by transceivers for signal regeneration. We find that transceiver use decreases and switch use increases steeply for \( \text{regenLimit} \) between 500 to 2000km in the bypass design. This reflects that most long-haul links are between 500 to 1000km so pegged regeneration at intermediate nodes (regardless of actual distance signal has travelled since the last regeneration) becomes more prevalent over \( \text{regenLimit} \)-based regenerations.

As transceiver usage drops at very large values of \( \text{regenLimit} \), we find that amplifiers become the dominant network component, consuming over half the network power for \( \text{regenLimit} \geq 2000 \text{km} \). Switch and router power seem to increase with regeneration improvements; however, this is an illusion. Switch usage is constant and
Figure 4-5: Network power distribution for four network designs implemented with OXC switches. Power consumption is based on a 300-Terabit/sec network and plotted as a ratio of component power to total network power for each value of \textit{regenLimit}. The solid black line shows how much the total network power consumption falls (as a ratio to aggregate power at \textit{regenLimit} = 100) as \textit{regenLimit} increases. The top left, top right, bottom left, and bottom right graphs show the power breakdown for patched bypass, bypass, groomed non-bypass and non-bypass designs, respectively.
amplifier usage increases very slightly. At each original regeneration point, we save
one amplifier through \textit{lastLinkReach} = 100km so as regeneration points disappear
by increasing \textit{regenLimit}, we can no longer claim this amplifier savings. Switch and
router power usage seems to increase relative to other components in the network
because the overall power usage decreases substantially. As we increase \textit{regenLimit}
from 500km to 3000km, we save 70\% in aggregate network power consumption.

In the bottom two graphs, we plot the power distribution for \textit{OXC} switched
groomed non-bypass and non-bypass designs. In both implementations, we find the
power consumption due to amplification and switching to be minimal. At \textit{regenLimit}
= 500km, approximately 60\% of the groomed non-bypass network power is used by
routers and 30\% are used by transceivers. Power savings in the non-bypass state is
insignificant.

In Figure 4-6, we summarize our power distribution analysis by presenting a snap-
shot slice of Figure 4-5 in a bar chart at \textit{regenLimit} = 500km. This chart is useful
when considering our hypothesis that the patched bypass and/or bypass design is the
most scalable design. We see that under both designs, the power consumption used
by transceivers is over four times that used by amplifiers or \textit{OXC} switches (both of
which are already very power efficient and technologically mature components). Re-
search to increase the optical reach of signals in fibers can greatly reduce transceiver
usage, resulting in significant power savings. Such savings cannot be found in the
non-bypass design. Even if the power consumption of routers is improved by 50\%, on
an absolute scale, the total power consumed by the non-bypass design is still orders of
magnitude larger than the energy required in the patched bypass or bypass designs.

In Figure 4-7, we plot the ratio of component power to total network power
across four network designs implemented with \textit{O/E/O_{ELEC}} switches, as a function
of \textit{regenLimit}. In Figure 4-8, we present a snapshot slice of Figure 4-7 in a bar chart
at \textit{regenLimit} = 500km. Our previous observations remain applicable.

For patched bypass and bypass networks, transceivers dominate power consump-
tion with transceiver use falling steeply when \textit{regenLimit} is between 500 and 1000km.
\textit{O/E/O_{ELEC}} switches consume more power than amplifiers (as expected). However,
Figure 4-6: Network power distribution for four network designs implemented with OXC switches at \textit{regenLimit} = 500km. Power consumption is based on a 300-Terabit/sec network and plotted as a ratio of component power to total network power for each value of \textit{regenLimit}. 
Figure 4-7: Network power distribution for four network designs implemented with O/E/O<sub>ELEC</sub> switches. Power consumption is based on a 300-Terabit/sec network and plotted as a ratio of component power to total network power for each value of regenLimit. The solid black line shows how much the total network power consumption falls (as a ratio to aggregate power at regenLimit = 100) as regenLimit increases. The top left, top right, bottom left, and bottom right graphs show the power breakdown for patched bypass, bypass, groomed non-bypass and non-bypass designs, respectively.
Figure 4-8: Network power distribution for four network designs implemented with \( O/E/O_{ELEC} \) switches at \( \text{regenLimit} = 500\text{km} \). Power consumption is based on a 300-Terabit/sec network and plotted as a ratio of component power to total network power for each value of \( \text{regenLimit} \).
that switches consume more power than transceivers when \( \text{regenLimit} \geq 800 \text{km} \) for the patched bypass network is surprising because the use of switches is so minimal in the design. Switches are only used at endpoints to select network lightpaths to be used. Similarly, that switches and transceivers each consume similar percentage of the total network power at \( \text{regenLimit} = 500 \text{km} \) and that switches dominate aggregate power consumption for optical reaches greater than 600km is notable. These observations support our hypothesis that \( \text{OXC} \) switches should be used whenever possible because they are significantly more power efficient than \( \text{O/E/OELEC} \) switches.

Finally, we note that under the \( \text{O/E/OOXC} \) switched non-bypass design, the power consumption is still completely dominated by routers (see Figure 4-8). The effect of regeneration improvements is negligible because transceiver power usage is negligible compared to router usage. This concurs with the observation that under a non-bypass design, the fixed power consumption of routers (both in terms of port count and power per port) is so high that no other technological advances can result in any meaningful energy savings. Even if industry can reduce router power consumption per port by 50%, a very ambitious amount, non-bypass total power consumption will still be 25 to 50 times as high as bypass or groomed non-bypass designs, and 500 to 1500 times as high as under the patched bypass designs.

Finally, in Figure 4-8, at the current regeneration limit of 500km, we show exactly how the power is distributed across components. In the top graph, we normalize the power consumption. The bottom graph, in which the power consumption is not normalized to show the magnitude of power consumed by routers in the non-bypass design, is interesting. We see that the magnitude of power consumed by routers in a non-bypass design is thousands of times more than any other component in any other design. This again confirms our recommendation that non-bypass designs should be avoided in future networks. Even if router technology becomes more efficient, the order of magnitude savings that must be achieved in router power reduction for overall network power to be comparable with other designs may not be physically feasible.
4.2 Analyzing Power Consumption Under a Highly Loaded Traffic Model

The analysis we conduct in Section 4.1 can be augmented to study the effects of rare but large transmissions on network power consumption. In this section, we look at how the network infrastructure must scale as traffic demand increases by $k$ standard deviations between all node pairs.

In the past decade, traffic was mostly circuit based and as a result, traffic demands on the network were easy to calculate. Link utilization was high because growth in traffic was steady and additional lightpaths can be lit as needed. In other words, the communications network was very much static and predictable. However, the shift from circuit-based connections to packet-based, unscheduled transmissions has forced network designers to grapple with the complexity of handling unpredictable and potentially huge transactions. These “large elephants” are highly unpredictable and account for over 90% of network traffic.

As a result, our network research in Section 4.1 on the expected traffic load may prove to widely underestimate the network capacity in the future, especially when we consider static infrastructure and inflexible designs such as patched bypass. In fact, at any time, any node pair has a low, but non-zero, probability of needing to handle a large elephant transaction. As a result, the network must be over-provisioned to allow for such large transactions on any lightpath. This leads to low utilization most of the time (since if we over-provision by three standard deviations, we can expect to be under-utilized by that much traffic 99% of the time). However, since the network infrastructure is static, to prevent denial of service, a trade-off between link utilization and quality of service must be made by the network designer.

The arrival of large and bursty transactions can be modelled as a Poisson process. For the most part, these large transactions can be assumed to occur independently but continuously. In other words, we assume that a user’s desire to download a high definition movie from Netflix is constant with time and is not correlated with his neighbor’s movie activity. Of course, during evenings, most of the population is
asleep and so a Poisson model is not entirely accurate. Nevertheless, we study the basic case in which the bursty traffic can be modelled as Poisson. Then, the variance of the traffic between any two nodes is equal to the mean of the traffic demand. We previously used the expected value of traffic loading in Section 4.1 and discuss our modelling of the expectation of traffic in Section 3.2. With these parts, we can easily add $k$-standard deviations to each entry in the traffic matrix to signify the large elephant events.

The main contribution of this approach is that the proportion of large transactions is modelled characteristically to the demographics of the network. In other words, regions with more people generate more traffic and has higher variance so a uniform scaling of capacity on each edge is a laboratory idealization that is unrealistic. We scale the traffic of each lightpath via population proxy.

In Table 4.1, we show the increase in aggregate traffic demand as a result of increasing traffic between each node pair by $k$ standard deviations. In Figure 4-9, we present a comparison of the aggregate power consumption for patched bypass and bypass network designs under $k = 0$ (corresponding to baseline), 3, 5, and 10 standard deviation traffic loading models. The top graphs plot power consumption on a log-log scale while the bottom graphs present a log-linear view. In Figure 4-10, we present a comparison of the power distribution for patched bypass and bypass networks under $k = 0, 3, 5,$ and 10 standard deviation traffic loading models. The top two graphs plot power consumption for $OXC$ switched patched bypass and bypass designs. The bottom two graphs show power consumption for $O/E/O_{ELEC}$ switched designs.

We notice that the curves for all values of $k$ lie very closely together. In fact, power distribution curves are indistinguishable among the various values of $k$. This is because the aggregate increase in traffic demand is small (approximately 2% increase at $k = 10$ standard deviations increase in traffic). As a result, the change in network power consumption and component power distribution is negligible. We make no additional comments beyond the analysis and recommendations made in Section 4.1.
Figure 4-9: A comparison of the network power consumption for patched bypass and bypass network designs under various traffic loading models. Solid lines denote OXC switched network; dashed lines denote O/E/O_{ELEC} switched network. The top left and right graphs plot the power comparison on a log-log scale and the bottom left and right graphs plot on a log-linear scale for patched bypass and bypass network, respectively.
Figure 4-10: Network power distribution for patched bypass and bypass network designs under various traffic loading models. Power consumption is plotted as a ratio of component power to total network power for each value of \textit{regenLimit}. The top left and right graphs plot the power comparison for \textit{OXC} switched and the bottom graphs plot power comparison for \textit{O/E/O_{ELEC}} switched patched bypass and bypass network, respectively.
Table 4.1: Effect on Aggregate Capacity from Scaling Traffic Matrix

<table>
<thead>
<tr>
<th>$k$</th>
<th>Standard Deviation Increase</th>
<th>Network Size [Terabits]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 0$ (Baseline)</td>
<td></td>
<td>306.85</td>
</tr>
<tr>
<td>$k = 3$</td>
<td></td>
<td>308.65</td>
</tr>
<tr>
<td>$k = 5$</td>
<td></td>
<td>309.85</td>
</tr>
<tr>
<td>$k = 10$</td>
<td></td>
<td>312.84</td>
</tr>
</tbody>
</table>

4.3 Patched Bypass versus Bypass Network Designs

In the previous sections, we have identified the patched bypass network design as the power minimal design. The power savings is considerable (approximately twenty-five times more efficient compared to the bypass designs). A simplistic recommendation would be to immediately deploy patched bypass networks across the country. This action may not be wise. The relationship between patched bypass and bypass infrastructure is complex and not fully discussed in previous research. In this section, we discuss the two main considerations that influence which design may be better under normal circumstances.

We have noted in Section 4.1.1 that a large downside to using patched bypass is network inflexibility. That is, because each lightpath under the patched bypass design must be statically provisioned and physically patched at each intermediate node, the network capacity between any node pair is fixed. When the traffic demand has a high variance, network designers must over-provision the number of lightpaths to handle fluctuations in traffic patterns. This results in lowered link utilization during normal operation. However, we notice that the potential power savings (25 times more power efficient than using bypass networks that can switch lightpaths at each intermediate node) suggest that this downside can be overcome.

What would happen if we uniformly over-provision the network by twenty times? Then, we allow for total scalability to future capacity since the aggregate traffic capacity increases by a factor of twenty. Simultaneously, we may handle the temporary fluctuations in traffic patterns. In other words, we should not discount patched by-
pass networks as infeasible simply because they are targeted for static traffic matrices; the large power consumption difference between the patched bypass and standard bypass designs offer a potential compromise in which we can increase network capacity while still reducing the aggregate network power consumption. We feel this is a very interesting area for future research.

The second network consideration regards network reliability and protection. Optical switching at intermediate nodes allow designers to implement $N$-plus-1 protection. Under $N$-plus-1 protection, for any edge, there is one extra set of lightpaths (usually larger than the maximal set of lightpaths connecting any node pair whose routing traverses the edge) that serves to protect the entire set of lightpaths that use the edge. In other words, if the lightpaths from node $i$ to node $j$ fail on edge $e$, switching at nodes $i$ and node $j$ allow the network to recover quickly and efficiently. The additional power consumption for $N$-plus-1 protection can be approximated as the amplifier power and switching power corresponding to extra set of lightpaths (and transceiver power if intermediate regeneration is needed). This “protection overhead” is in addition to the baseline bypass power consumption.

The alternative protection technique is 1-plus-1 lightpath protection, in which a spare lightpath (end-to-end) stands ready to be switched into service at end nodes should any part of the lightpath fail to transmit signal effectively. The protection power overhead is a factor of two. Considering that the patched bypass network is over 25 times more power efficient than a bypass network, the alternative protection technique seems to be able to protect a patched bypass network at substantial power savings. One potential downside is that 1-plus-1 lightpath protection is inefficient. The entire network capacity is doubled but not used (standing by as a spare), which results in link utilization lowered by a factor of two. In addition, because errors on a lightpath can occur at any of the links connecting intermediate nodes, additional research needs to be conducted to see if one set of spare lightpaths is enough, or if a larger set is needed due to a higher end-to-end of failure rate. In other words, lightpath protection is less hardy than edge protection because under $N$-plus-1 edge protection, should the edge and the spare edge both fail, a switched bypass network
can reroute traffic along a completely independent geographic route at the cost of lower performance. This reroute is not possible under the patched bypass network design.

To determine whether patched bypass is a strong contender for best network architecture, future research must measure network fluctuation accurately. In other words, we must first determine the multiple by which the patched bypass network needs to be over-provisioned by, so that once we factor in 1-plus-1 lightpath protection power overhead, we can compare the power consumption of the patched bypass design with that of the optical bypass network architecture. For stable networks with static traffic matrices, our research suggests the answer is yes. However, for what level of unpredictability in the network is this observation still valid? This question is an important future research area.

4.4 Justification of Shortest Path Routing

In our research, we have used shortest distance path routing, motivated by the desire to minimize the number of components on any independent path. That is, we route the traffic of node $i$ without concern about the capacity utilization of nearby nodes. However, is this absolutely the minimum power routing algorithm?

Previous research suggests that a traffic balancing routing technique may be optimal. The optimal routing algorithm presented by Shen “allows traffic demands between different node pairs to share capacity on common virtual (lightpath) links in order to improve capacity utilization. Although such an effort may elongate the traversing lengths of some (but not many) traffic flows...the overall improvement of network capacity utilization...[corresponds to] less energy consumption” [28]. In other words, allowing traffic to be routed along more circuitous routes, if those routes have excess network capacity, consumes less energy than along a shortest path without traffic grooming in an optical bypass design. He uses a greedy algorithm for the routing assignment in which node pairs with large network demand define which virtual lightpaths are allocated and exist. Node pairs with smaller traffic demands are
expected to fill the residual traffic capacity on these large flow lightpaths.

We respectfully disagree with Shen’s routing recommendation for power savings. First, we believe that a network should be statistically fair. In other words, there may be instances in which traffic from node \(i\) to node \(j\) is routed along a path resulting in lower quality of service (e.g. in the case of edge outage, or temporary traffic overload on an edge), but this should not be the default behavior for a subset of customers residing in a less active network region compared with another subset of customers. The exception to this argument would be if all subsets of customers whose traffic were routed as “filler” were provided an appropriate discount for this service. However, we do not believe this “lightpath stuffing” is even necessary. Under our traffic model assumptions (incorporating international growth estimates of 40-110% annually), the traffic demand is large enough that even the nodes serving more rural regions have traffic that fills entire lightpaths. For example, the smallest traffic demand in our matrix is 3 wavelengths between Raleigh and Salt Lake City, in each direction.

We do not agree that traffic grooming is power efficient. In order to combine traffic from multiple source streams into a single wavelength, routers are needed. In Section 4.1.1, we have argued that optical bypass is dominant over groomed non-bypass, both of which are significantly more power efficient than the non-bypass network design recommended by Shen. Under the groomed non-bypass design, we model a more efficient version of Shen’s “multi-hop bypass” algorithm. We route whole lightpaths of traffic on an optical bypass network with no router ports (aside from aggregating routers at source and destination nodes) and groom only residual fractional lightpaths in routers at intermediate nodes. We find that for all values of \(\text{regenLimit}\), the groomed non-bypass design consumes more power than the optical bypass design because each router port pair consumes over 600W of power, compared with 0.09W per amplifier, 0.430W per pair of \(OXC\) switch ports, and 2.3W per transceiver. Despite our sparing use of the router ports, the difference in aggregate power consumption between the groomed non-bypass and bypass designs is obvious.

We further find Shen’s claim that “direct bypass” (corresponding to our bypass
design) and “multi-hop bypass” can “help equalize the power consumption at each network node” [28] to be based on inconclusive data. First, his models place routers at all intermediate nodes so that his network power distribution among components is over 90% dominated by routers, as our power distribution for the non-bypass design was. Therefore, we suggest that while Shen did not observe large variances in traffic processed at each node in his test networks, there are in fact still large variances in power consumption among these nodes. Because the router power is order(s) of magnitude larger than these differences, variances in traffic processed at each node are not obvious. In an optical bypass network without router ports, these differences between nodes can become highly significant.

Another reason why previous research was able to equalize power consumption at each node is because the traffic model used in previous research is based on a uniform distribution with identical mean for all nodes. Then, the goal of equalizing traffic capacity across nodes is similar to smoothing out noise across several data points and is almost trivial if the variance in the sampling is small. However, in our traffic model based on actual population demographics, we find balancing traffic loads across all edges to be extremely difficult. In other words, if no one communicates with Santa at the North Pole during the summer, it does not make sense to route cross-continental traffic up to the North Pole simply because we can (i.e. there is excess capacity). In this example, there is no reason why the North Pole traffic should be equal to Chicago traffic and so the power consumed by the two nodes cannot be comparable.

In Figure 4-11, we plot the aggregate number of lightpaths that traverse any edge in our network graph. We color the edges to indicate density of traffic loading. For example, the green and yellow edges denote well traversed edges while orange and red colors indicate very highly loaded edges or that the edges connect “hotspot nodes”. Not surprisingly, the most heavily used link connects Chicago and Albany (Albany serves the tri-state area, including New York City). This edge is the lifeline of the country’s financial capitals, serving companies who run dedicated lightpaths (without regard to link utilization and with concern only for transmission performance and zero blocking probability). In that case, we actually underestimate the number
Figure 4-11: A 25-node, 56-link network representing the United States Wide Area Network, in which edges are colored according to how many lightpaths traverse the edge. Path routing is via the shortest path algorithm.
of wavelengths and switch/router ports used on the Chicago ↔ Albany edge. Over 10% of the traffic carried on that link are direct unidirectional traffic from Chicago to Albany or vice versa. If the link utilization on those lines are contractually obligated to be low (far less than 1%), then that estimate of single hop traffic can jump. We point out that one cannot minimize the lightpaths on a edge if the traffic originates or ends at either of the nodes served by that edge. Therefore, while using another routing algorithm may be able to shift a few wavelengths from heavily loaded to lightly loaded edges, the majority of heavy traffic must still originate or terminate in a hotspot node.

Finally, ignoring the contractual reasons (i.e. for commercial customers) and customer satisfaction (i.e. quality of service) arguments to use shortest path routing, we provide a semi-formal proof by cases why shortest path routing is more power efficient under the majority of circumstances to traffic balanced routing. In fact, we argue that a routing algorithm that uses an edge weighting that is a function of both path distance and hop count is power optimal.

**Case I: Routing in a Patched Bypass Design**

Take any lightpath composed of edges that are patch panelled together. Suppose a shorter path exists by replacing a subset of the edges. By selecting the shorter path, the number of amplifiers (and possibly transceivers depending on `regenLimit` and amount of distance reduced) decreases leading to reduced power consumption. This reduction terminates once we route all traffic on the shortest path. No grooming is available so low link utilization is irrelevant. Therefore, under the patched bypass network design, shortest path routing is power optimal.

**Case II: Routing in Bypass Design**

**Part A: Same Number of Intermediate Nodes on Both Paths**

If the paths have the same number of intermediate nodes, then when the network is OXC switched, the shorter path is more power efficient. The number of amplifiers
and transceivers are calculated the same way as in the patched bypass scenario, so shortest path routing is optimal.

If the network is $O/E/O_{ELEC}$ switched, minimum regeneration points are pegged at intermediate nodes. In that case, the location of the nodes is important. In most cases, lightpaths on the shortest path consume the least amount of power because the edge(s) that are different between the longer and shorter path are usually longer and so will need at least the same number, if not more, of amplifiers (and of transceivers if needed). We acknowledge that cases can exist in which shortest path routing is not optimal. This is because whether shortest path routing is optimal depends on the actual switch placement and the resulting number of switches and transceivers.

Take for example the extreme case illustrated in Figure 4-12. Assume that $\text{linkReach} = 50\text{km}$, $\text{lastLinkReach} = 100\text{km}$, and $\text{regenLimit} = 500$. Then, for this graph configuration, the shortest path routing is not power minimal. Even though the path through the upper node is shorter by 37km, it uses two amplifiers (after the first 50km on each edge) while the bottom path only uses one amplifier (after 50km on the second edge). The number of switch ports at the intermediate nodes is the same. Then, the longer path is more power efficient by deploying fewer components.

Of course in real network deployment, the longer path would not actually be more power efficient because while amplifiers are recommended to be used every $\text{linkReach}$ times, the signal is not lost if each amplifier’s reach is extended by a few kilometers. Then, the network designer can space the amplifiers and transceivers on that path so that a minimal number of amplifiers and transceivers are used on the shorter path. In other words, each edge on the shorter path can only be less efficient by at most one amplifier. The network designer can distribute the single $\text{linkReach}$ distance in question over the other amplifiers so that the edges on the shorter path consist of the same number of amplifiers as the edges on the longer path in such edge counterexamples.

Therefore, under the $OXC$ switched bypass design, shortest path routing is power optimal. Under the $O/E/O_{ELEC}$ switch design, shortest path routing is usually optimal, but may also be considered suboptimal in certain circumstances where the
Figure 4-12: A routing example for which the shortest path routing is not power optimal due to the location of O/E/O\textsubscript{ELEC} switch placement. Assume that linkReach = 50km, lastLinkReach = 100km, and regenLimit = 500. Then, the path through the upper node is shorter by 37km, but uses two amplifiers (after the first 50km on each edge) while the bottom path only uses one amplifier (after 50km on the second edge). The number of switch ports at the intermediate nodes is the same.

placement of intermediate nodes on the longer path leads to fewer components. We note that network designers can intervene by adjusting placement of amplifiers and transceivers to maximum reach so that shortest path routing is still optimal.

Part B: More Intermediate Nodes on the Longer Path Than on the Shorter Path

If a longer path has more hops than a shorter path does, then shortest path routing is equivalent to shortest hop routing and is power optimal. Let m be the number of extra intermediate nodes and n be the number of wavelengths traversing the path.

If the network is OXC switched, we approximate the power consumed by the m extra switches used on the longer path as $m \cdot n \cdot P_{OXC,\text{conn}}$. Note that because OXC switches suffer from OXC\text{Loss}, in addition to the extra length in the longer path, we should add $m \cdot OXC\text{Loss}$ to the total path length (or more accurately, OXC\text{Loss} after each of the n extra switches). We ignore this effect in this approximation.

The most inefficient node placement on the shortest path (based on Figure 4-12) results in one extra amplification needed per edge except for the last edge. (Note, it is possible to find numbers such that one extra regeneration per edge is needed, however, these values are not realistic in real life as it requires $\text{regenLimit} \approx \text{linkReach}$ or $\text{lastLinkReach}$, such as $\text{linkReach} = 50\text{km}$, $\text{lastLinkReach} = 100\text{km}$, and $\text{regenLimit} = 100\text{km}$. Therefore, we discard this possibility.) Then, the differ-
ence in amplification power is approximately $m \cdot n \cdot \frac{P_{amp}}{200}$.

Because $P_{OXC} < \frac{P_{amp}}{200}$, the longer path is never more power efficient than the shorter path using OXC switches. The same argument can be made to compare $P_{O/E/O_{ELEC}} < \frac{P_{amp}}{200}$ to argue that longer path is never more power efficient using O/E/O_{ELEC} switches either.

Therefore, under either OXC or O/E/O_{ELEC} switched bypass design in which the longer path has more hops than a shorter path, shortest path or shortest hop routing is power optimal.

**Part C: More Intermediate Nodes on the Shorter Path Than on the Longer Path**

If a longer path has fewer hops than a shorter path does, then shortest hop routing may be more power efficient than shortest path routing. Let $m$ be the number of fewer intermediate nodes and $n$ be the number of wavelengths traversing the path.

If the network is OXC switched, we approximate the power consumed by the $m$ extra switches used on the shorter path as $m \cdot n \cdot P_{OXC_{conn}}$. Note that because OXC switches suffer from OXCLoss, in addition to the extra length in the longer path, we should add $m \cdot OXCLoss$ to the total path length (or more accurately, OXCLoss after each of the $n$ extra switches). We ignore this effect in this approximation.

Let $\Delta_{length}$ denote the difference in path distance between the longer and shorter paths and $t = \frac{\Delta_{length}}{regenLimit}$ denote the approximate number of regenerations needed over the extra distance. Again, we ignore the effect of OXCLoss and node placement since network designers can reposition amplifiers and transceivers within a link as needed.

We approximate the power consumed by extra transceivers as $n \cdot t \cdot P_{1550/r}$. We approximate the power consumed by extra amplifiers as

$$n \cdot \frac{P_{amp}}{200} \left\{ \frac{t \mod regenLimit}{linkReach} + \text{floor}(t) \cdot \frac{regenLimit - \text{lastLinkReach}}{linkReach} \right\} \quad (4.1)$$
\[
\begin{align*}
    m \cdot P_{\text{OXC}} - t \cdot P_{\text{1350/1}} + n \cdot P_{\text{amp}} \left\{ \frac{t \mod \text{regenLimit}}{\text{linkReach}} \right. \\
    \left. + \; \text{floor}(t) \cdot \frac{\text{regenLimit} - \text{lastLinkReach}}{\text{linkReach}} \right\} 
\end{align*}
\]

then shortest hop routing is power optimal for OXC switched bypass networks. Otherwise, shortest path routing is power optimal. The same argument can be made to compare routing algorithms when \(O/E/O_{\text{ELEC}}\) switches are used.

If
\[
\begin{align*}
    m \cdot P_{\text{OXC}_{\text{ELEC}}} - t \cdot P_{\text{1350/1}} + n \cdot P_{\text{amp}} \left\{ \frac{t \mod \text{regenLimit}}{\text{linkReach}} \\
    \left. + \; \text{floor}(t) \cdot \frac{\text{regenLimit} - \text{lastLinkReach}}{\text{linkReach}} \right\} 
\end{align*}
\]

then shortest hop routing is power optimal for \(O/E/O_{\text{ELEC}}\) switched bypass networks. Otherwise, shortest path routing is power optimal.

Therefore, under either OXC or \(O/E/O_{\text{ELEC}}\) switched bypass designs in which the longer path has fewer hops than a shorter path, either the shortest path or shortest hop routing is power optimal. If the longer path has a very large \(\Delta_{\text{length}}\), shortest path routing will be more power efficient. Otherwise, shortest hop routing will be optimal.

### 4.5 Summary of Chapter 4

In this chapter, we provided detailed power optimization results and design heuristics for the WAN. We found that optical bypass networks are the most scalable architecture design with respect to power consumption, especially when quality of service, network flexibility, reliability, and protection are considered. The power consumption of the standard bypass design can be further improved through a hybrid patched bypass/bypass network. Under the hybrid scheme, whole wavelengths of core, stable
traffic between node pairs is routed via direct, fixed lightpaths using patch panelling to avoid lightpath switching at intermediate nodes and all other unexpected, fluctuating, and/or bursty traffic is switched on a standard optical bypass network. We found the hybrid network to be power minimal with exceptional latency, transmission, and reliability performance. We also analyzed power distribution among components and found the $OXC$ switch to be most scalable and the $O/E/O_{OXC}$ switch to be most wasteful (signals traversing an $O/E/O_{OXC}$ switch are regenerated twice due to its use of two transceivers, compared to one regeneration using an $O/E/O_{ELEC}$ switch and no regenerations using an $OXC$ switch). Finally, we proved that shortest path and minimum hop routing is power optimal and traffic balanced routing should be avoided when a majority of the traffic demand between any pair of nodes is sufficiently large to warrant whole wavelengths, as is the case when network capacity grows from the current 30-Terabit/sec to a projected 300-Terabit/sec capacity.
Chapter 5

Components and Cost Model of the Metropolitan Area Network

As capacity in the wide area network grows, infrastructure in the metropolitan and local area networks must also scale appropriately to handle network traffic efficiently. Traffic added into the WAN at each core node must be aggregated by the routers of the metropolitan feeder network and all traffic terminated at a node in the WAN must be processed by the MAN. Then, finding a scalable MAN architecture is as important as implementing a power efficient WAN design.

In the previous two chapters, we presented the assumptions and reasoning for the traffic demand in the WAN to grow by a hundredfold over the next five to ten years. We also presented heuristics for designing a scalable WAN with respect to power. In this half of the thesis, we look at developing a scalable MAN architecture via joint optimization over physical topology, routing and wavelength assignment, and dimensioning of network resources, all with respect to both power and cost. Much of the analysis in this chapter is based upon the work of Guan [13, 11, 12], especially the capital expenditure models of network components and analytical descriptions of graph structures. Our goal is to develop bi-scalable MAN designs and compare them with the cost-scalable architecture developed by Guan to characterize power and cost efficient network optimization heuristics and recommendations.

What is a MAN? The main functionalities of a MAN include aggregating traffic,
delivering traffic to a hub, transferring traffic from one LAN to another, and performing network management and control functions to maintain high reliability and quality of service for end-users. The MAN can be thought of as a small WAN with a regular topology. Where the WAN connects many MAN nodes, the MAN is a network of many LAN nodes and interfaces between the LAN and WAN. The MAN nodes in a region serve to create a stream of traffic for their WAN node. MAN topology can be described as a set of hub nodes and access nodes to the LAN. These nodes are connected in a ring or mesh, which is further connected to a WAN node.

In this chapter, we present the parametric cost model of each component with representative capital and operating expenditure values for each component. Then, we combine the components into networks and formulate analytical expressions for network and normalized network cost. Finally, because the normalized cost functions depend on the type of topologies used, we present a summary of the properties of the Generalized Moore, Δ-Nearest Neighbors, and Symmetric Hamilton graphs (taken from [13], Chapter 4).

5.1 MAN Traffic Model

We choose a deterministic, off-line, non-blocking traffic model. That is, the entire traffic demand is provided *a priori*. We assume uniform all-to-all traffic in which each node sends exactly $t$ wavelengths of traffic to each of the other nodes in the network. This results in a fully connected logical topology in which $T(i,j) = t$ for $i \neq j$ in traffic matrix $M_T = [T(i,j)]$. This type of traffic matrix best models a dense metropolitan area in which transactions among all nodes are uniform and well-balanced. In other words, our model describes a network in which communications from any node are as likely to go to one node in the network as to any other node. In Figure 5-1, we show the fully connected logical topology that results from a uniform all-to-all traffic model.
Figure 5-1: Schematic diagram showing the fully connected topological nature of a network with uniform all-to-all traffic among nodes. Each edge carries $t$ wavelengths of traffic.

5.2 Parametric Network Cost Model

In this section, we set up a parametric network cost model. The model must be parametric because the cost and power estimates for key network components can vary significantly depending on the configuration and operating environment. With a parametric model, the network dependencies can be analyzed easily over a reasonable range of parameter values. We note that the cost parameters provided by Guan are dated as of 2007 and may be inflated by today’s technological standards. Nevertheless, we expect the relative switch to fiber cost ratio to be accurate. We translate power consumption into cost figures so that capital and operating expenditures can be incorporated into a single cost function. In our model, the network consists of three components: fiber connections, transceivers, and switches. The costs are modelled as functions of the number of nodes ($N$) and the node degree ($\Delta$).

5.2.1 Transceiver Cost

To support uniform all-to-all traffic demand, each node must have the ability to send/receive traffic to/from any other node. This assumption is necessary so that in the worst case, if all $N - 1$ other nodes wish to send traffic to the same single node at the same time, all the traffic can be received by the node. Similarly, if a
single node wishes to broadcast traffic to all other $N - 1$ nodes, the node does not experience blocking due to a lack of transmitters. Therefore, the network contains $N(N - 1)t$ transceivers, independent of the node degree $\Delta$. We can consider both the capital and operating expenditures as a constant offset to the network cost function and disregard them in our optimizations.

5.2.2 Fiber Connection Cost

In the MAN, a fiber connection spans a much shorter distance than it would in a WAN. Therefore, amplifiers (which are needed only every 50km or 30 miles) are rarely used in the MAN. For example, the most populous city in the United States, New York City, is only about 20 square miles, a size which obviates the need for amplifiers. Then, we can assume that all fiber connections have approximately the same cost, independent of the length of each connection. The only cost is from capital expenditures. There are no operating expenditures to model because the fibers contain no active components. We define the fiber cost function $C_f(N, \Delta)$ as a linear function of $N$ and $\Delta$

$$C_f(N, \Delta) = \alpha N \Delta$$  \hspace{1cm} (5.1)

where coefficient $\alpha$ is the marginal cost of a new fiber connection. If cable plants pre-exist, the marginal cost of a new fiber connection is approximately $2K/km$. However, if the cable plants do not pre-exist, then network operators need to break ground and lay new ducts and cables, driving up the marginal cost significantly to $25K/km$. Because a typical fiber is 5 to 20km, Guan estimates the value of $\alpha$ to lie between $10K/fiber$ and $500K/fiber$.

5.2.3 Switch Cost

We assume that network traffic is uniform all-to-all so that each node sends exactly $t$ wavelengths of traffic to every other node with each wavelength modulated at a data rate of $r$ (Gb/sec). We further assume that networks are regular to facilitate the analysis of pass-through traffic. We consider two types of switches: the optical
cross connect, $OXC$, switch and the optical-electronic-optical with electronic core, $O/E/O_{ELEC}$ switch.

**Amount of Add-Drop and Pass-Through Traffic**

Traffic is added and removed from the network at switches in each node. For traffic paths of more than one hop, lightpaths must be switched at intermediate nodes (referred to as pass-through traffic). In Figure 5-2, we provide a schematic diagram illustrating the relationship between local traffic added and dropped at each node and the traffic that is switched at each node.

Each logical connection between nodes carries $t$ wavelengths of traffic so that each node sends exactly $t$ wavelengths of traffic to every other node. Each node can originate and add $t(N - 1)$ lightpaths to the network; similarly, each node can terminate and drop $t(N - 1)$ lightpaths received from the network. The add-drop traffic at each node is $2t(N - 1)$.

The pass-through traffic at each node is $t(N - 1)[H_{min}(N, \Delta)] - 1]$ where $H_{min}$ is the average minimum hop distance of the network. Each node can pass the traffic of $N - 1$ other nodes in the network. The average number of nodes that pass traffic is $H_{min} - 1$ because the node at the last hop will terminate and drop the traffic instead of passing it.

**Cost of an $OXC$ Switch**

We define the network cost due to $OXC$ switching as $C_{OXC}(N, \Delta, t)$. To find an analytical expression of the cost, we first find the number of ports needed by a node to switch traffic. Then, we model the capital expenditure per port and convert power consumption per switch port into an operating expenditure per port. Then, $C_{OXC}(N, \Delta, t)$ is simply the product of the number of ports, the total cost per port, and the number of switches in the network.

First, we dimension the size of an $OXC$ switch. We note that because the traffic matrix is uniform all-to-all, the size of each $OXC$ switch is identical. The capacity of an $OXC$ switch is independent of the actual data rate of each wavelength so that the
Figure 5-2: Schematic diagram serving as a guideline to account for the amount of local traffic added or dropped at each switch and the traffic that passes through the switch ports. Assumes $t = 1$ wavelengths of traffic between any node pair. Reproduced from [13], Figure 5-3.

The size of the switch is determined entirely by the number of ports required. Each added, dropped, and passed-through lightpath requires one port pair. Therefore, the size of an OXC switch $K_{OXC}$ is equal to the sum of number of add-drop and pass-through lightpaths.

$$K_{OXC}(N, \Delta, t) = 2t(N - 1) + t(N - 1)[H_{\text{min}}(N, \Delta) - 1]$$

$$= t(N - 1)[H_{\text{min}}(N, \Delta) + 1]$$  \hspace{1cm} (5.2)

We assume that optical switches allow full wavelength conversion so that any wavelength on an input port can be switched to any wavelength on an output port. Then, $K_{OXC}$ ports are sufficient to switch all the lightpaths in our traffic model. We further assume that the switching architecture is 3-D so that the capital expenditure function is

$$F_{\text{cap}}(K_{OXC}(N, \Delta, t)) = \beta_{OXC,1} K_{OXC}(N, \Delta, t)$$  \hspace{1cm} (5.3)

where $\beta_{OXC,1}$ is $10K$/port pair for an 8X8 switch.

We assume that the operating life of an OXC switch is five years. To maintain consistency between our operating cost parameters and the capital expenditure parameters used by Guan, we disregard inflation, net present value discounting, and
cost of capital. Then, the operating expenditure function is

$$F_{op}(K_{OXC}(N, \Delta, t)) = \beta_{OXC2} K_{OXC}(N, \Delta, t)$$ (5.4)

where $\beta_{OXC2}$ is the operating expenditure converted from expected power consumption over the lifetime of the switch. $\beta_{OXC2} = 2 \cdot P_{OXC_{conn}} \cdot 8766 \text{ hours/year} \cdot 0.098 \text{ kW/h} \cdot 5 \text{ years}$, where the factor of two accounts for estimated air cooling power overhead. $\beta_{OXC2}$ is approximately $3700$/port pair for an 8X8 switch, using the representative power consumption values from Section 2.3 (see Table 2.2).

The total cost of OXC switches in a network with $N$ nodes, $C_{OXC}(N, \Delta, t)$, is

$$C_{OXC}(N, \Delta, t) = N[F_{cap}(K_{OXC}(N, \Delta, t)) + F_{op}(K_{OXC}(N, \Delta, t))]$$

$$= N(\beta_{OXC1} + \beta_{OXC2}) K_{OXC}(N, \Delta, t)$$

$$= t\beta_{OXC} N(N - 1)[H_{min}(N, \Delta) + 1]$$ (5.5)

where $\beta_{OXC} = \beta_{OXC1} + \beta_{OXC2}$.

Cost of an O/E/O Switch

We define the network cost attributable to O/E/O_{ELEC} switching as $C_{O/E/O}(N, \Delta, t)$. There are technically two available types of O/E/O switches: O/E/O_{OXC} which uses an optical switching core and the O/E/O_{ELEC} which uses an electronic switching core (see Section 2.3 for detailed distinctions). However, during power consumption analysis on the WAN, we found that O/E/O_{ELEC} switches dominate over the O/E/O_{OXC} switches in power while providing the same functionality. O/E/O_{OXC} switches were left out of Guan’s network cost modelling, possibly for the same performance versus power and cost trade-off. Therefore, we do not conduct analysis on the MAN using O/E/O_{OXC} switching. Hereafter, all references to O/E/O switches refer to the O/E/O_{ELEC} configuration, as diagrammed in Figure 2-5.

To find an analytical expression of the O/E/O switching cost, we first find the number of ports needed by a node to switch traffic. We model the capital expenditure per port and convert power consumption per switch port into an operating expendi-
ture per port. Then, $C_{O/E/O}(N, \Delta, t, r, R, \eta)$ is simply the product of the number of ports, the total cost per port, and the number of switches in the network.

The total traffic switched at an $O/E/O$ switch is dependent on the data rate per wavelength $r$, so that the traffic is actually $r \cdot K_{OXC}$. The size of an $O/E/O$ switch, $K_{O/E/O}$ is

$$K_{O/E/O}(N, \Delta, t, r, R, \eta) = \frac{rt(N - 1)[H_{\text{min}}(N, \Delta) + 1]}{R\eta}$$  \hspace{1cm} (5.6)

where $r$ is the Gb/s data rate, $R$ is the port (interface) rate $R$, and $\eta$ is the port utilization.

We model the capital expenditure cost of an $O/E/O$ switch as a linear function of the number of switching ports $K_{O/E/O}$ such that

$$F_{\text{cap}}(K_{O/E/O}(N, \Delta, t, r, R, \eta)) = \beta_{O/E/O,1} \cdot K_{O/E/O}(N, \Delta, t)$$  \hspace{1cm} (5.7)

where $\beta_{O/E/O,1}$ is $\$80K$/port pair for an 10Gb/s interface switch.

We assume that the operating life of an $O/E/O$ switch is five years. To maintain consistency between our operating cost parameters and the capital expenditure parameters used by Guan, we disregard inflation, net present value discounting, and cost of capital. As a result, the operating expenditure function is

$$F_{\text{op}}(K_{O/E/O}(N, \Delta, t, r, R, \eta)) = \beta_{O/E/O,2} \cdot K_{O/E/O}(N, \Delta, t, r, R, \eta)$$  \hspace{1cm} (5.8)

where $\beta_{O/E/O,2}$ is the operating expenditure converted from the expected power consumption over the lifetime of the switch. $\beta_{O/E/O,2} = 2 \cdot P_{O/E/O_{\text{ELEC}}} \cdot 8766 \text{ hours/year} \cdot 0.098 \text{ kW/h} \cdot 5 \text{ years}$, where the factor of two accounts for estimated air cooling power overhead. $\beta_{O/E/O,2}$ is approximately $\$20K$/port pair for an 10Gb/s interface switch, using the representative power consumption values from Section 2.3.

The total cost of $O/E/O$ switches in a network with $N$ nodes, $C_{O/E/O}(N, \Delta, t, r, R, \eta)$,
is
\[
C_{O/E/O}(N, \Delta, t, r, R, \eta) = N[F_{\text{cap}}(K_{O/E/O}(N, \Delta, t, r, R, \eta)) + F_{\text{sp}}(K_{O/E/O}(N, \Delta, t, r, R, \eta))]
\]
\[
= N(\beta_{O/E/O,1} + \beta_{O/E/O,2}) K_{O/E/O}(N, \Delta, t, r, R, \eta)
\]
\[
= \frac{rt\beta_{O/E/O}N(N - 1)[H_{\min}(N, \Delta) + 1]}{R\eta}
\]
(5.9)

where \(\beta_{O/E/O} = \beta_{O/E/O,1} + \beta_{O/E/O,2}\).

### 5.2.4 Network Cost Model

In this section, we model the total parametric cost model of the optical MAN, based on the cost functions of individual components presented earlier. The total network cost is the sum of the fiber connection cost and the switching cost. We have assumed the transceiver cost (for regeneration in fiber connections) is static and independent of network design. Given the two types of switches, optical cross connect (OXC) and \(O/E/O\) electronic core (\(O/E/O\)), we have two network cost models.

For a network equipped with OXC switches, the total network cost function, \(C_{\text{OXC}_\text{total}}\), is

\[
C_{\text{OXC}_\text{total}}(N, \Delta, t) = C_f(N, \Delta) + C_{\text{OXC}}(N, \Delta, t)
\]
\[
= \alpha N\Delta + t\beta_{\text{OXC}}N(N - 1)[H_{\min}(N, \Delta) + 1]
\]
(5.10)
\[
= N(\alpha\Delta + t\beta_{\text{OXC}}(N - 1)[H_{\min}(N, \Delta) + 1])
\]

The total power consumption per node, \(C_{\text{OXC}_\text{norm}}\), is

\[
C_{\text{OXC}_\text{norm}}(N, \Delta, t) = \frac{C_{\text{OXC}_\text{total}}(N, \Delta, t)}{N}
\]
\[
= \alpha\Delta + t\beta_{\text{OXC}}(N - 1)[H_{\min}(N, \Delta) + 1]
\]
(5.11)

Similarly, for a network equipped with \(O/E/O\) switches, the total power consump-
tion for $C_{O/E/O_{total}}$ is

$$C_{O/E/O_{total}}(N, \Delta, t, r, R, \eta) = C_f(N, \Delta) + C_{O/E/O}(N, \Delta, t, r, R, \eta)$$

$$= \alpha N \Delta + \frac{rt \beta_{O/E/O} N (N - 1)[H_{min}(N, \Delta) + 1]}{R \eta}$$

$$= N \left( \alpha \Delta + \frac{rt \beta_{O/E/O} (N - 1)[H_{min}(N, \Delta) + 1]}{R \eta} \right) \quad (5.12)$$

The total power consumption per node, $C_{O/E/O_{norm}}$, is

$$C_{O/E/O_{norm}}(N, \Delta, t, r, R, \eta) = \frac{C_{O/E/O_{total}}(N, \Delta, t, r, R, \eta)}{N}$$

$$= \alpha \Delta + \frac{rt \beta_{O/E/O} (N - 1)[H_{min}(N, \Delta) + 1]}{R \eta} \quad (5.13)$$

### 5.3 Regular Topologies

In this section, we first provide a working definition of regular topology to cover a broad class of topologies with symmetric and well-defined structures. Then, we summarize the topological properties of Generalized Moore, $\Delta$-Nearest Neighbors, and Symmetric Hamilton graphs.

We say that a topology is regular with node degree $\Delta$ when it satisfies the following conditions:

- $\Delta$ outgoing edges from and $\Delta$ incoming edges to each node
- Nodal symmetry such that each node links to $\Delta$ other nodes following predefined connectivity rules
- $\Delta$-connectedness such that the number of nodes that are $i$ hops away from a node (define as $n(i)$) via minimum hop routing for $i$ less than the diameter of the network is at least $\Delta$
- Average minimum hop distance $H_{min}$ can be expressed as

$$H_{min} = \frac{1}{N - 1} \sum_{i=1}^{D} in(i) \quad (5.14)$$

98
Figure 5-3: An example of a Moore graph. In (a), we show the Petersen graph with \( N = 10, \Delta = 3, \) and \( D = 2. \) In (b), we show the routing spanning tree as seen from Node 1. Reproduced from [13], Figure 4-3.

5.3.1 Generalized Moore Graphs

A Moore graph is an ideal (not necessarily realizable) regular topology that satisfies the Moore bound as established by E. F. Moore. Under the Moore bound, a directed graph can support up to \( N_{\text{max}}(\Delta_{\text{max}}, D) \leq 1 + \sum_{i=1}^{D}(\Delta_{\text{max}})^i \) nodes and an undirected graph can support up to \( N_{\text{max}}(\Delta_{\text{max}}, D) \leq 1 + \Delta_{\text{max}} \sum_{i=0}^{D-1}(\Delta_{\text{max}} - 1)^i \) where \( D \) is the diameter of the graph and \( \Delta_{\text{max}} \) is the maximum node degree. An important characteristic of the Moore graph is that each node can reach every other node in a fully populated \( \Delta \)-ary minimum hop routing spanning tree in which each path is unique. Furthermore, among regular topologies with the same number of nodes and same node degree, the Moore graph is the lower bound on average minimum hop distance. In Figure 5-3, we present one of the existing Moore graphs, the Petersen graph, as an example.

A Generalized Moore graph is a regular graph which does not achieve upper bound on the number of nodes, but does achieve the lower bound on the average minimum hop distance. The routing spanning tree of a Generalized Moore graph has all the...
levels of a Moore graph, except possibly a semi-filled last level. In Figure 5-4, we present an example of a Generalized Moore graph, the Heawood graph.

For the directed Generalized Moore graph, the average minimum hop distance, $H_{\text{min, dir. Moore}}$, is

$$H_{\text{min, dir. Moore}}(\Delta, D) = \frac{\Delta - \Delta^{D+1} + ND(\Delta - 1)^2 + D(\Delta - 1)}{(N - 1)(\Delta - 1)^2}$$  \hspace{1cm} (5.15)

For large values of $N$, $H_{\text{min, dir. Moore}}$ approaches the asymptotic limit $\log_\Delta N$.

For the undirected Generalized Moore graph, the average minimum hop distance, $H_{\text{min, undir. Moore}}$, is

$$H_{\text{min, undir. Moore}}(\Delta, D) = \frac{\Delta[1 - (\Delta - 1)^D] + ND(\Delta - 2)^2 + 2D(\Delta - 2)}{(N - 1)(\Delta - 2)^2}$$  \hspace{1cm} (5.16)

For large values of $N$, $H_{\text{min, undir. Moore}}$ approaches the asymptotic limit $\log_{\Delta-1} N$. 

Figure 5-4: An example of a Generalized Moore graph. In (a), we show the Heawood graph with $N = 14$, $\Delta = 3$, and $D = 3$. In (b), we show the routing spanning tree as seen from Node 1. Reproduced from [13], Figure 4-6.
Figure 5-5: An example of a $\Delta$-Nearest Neighbors graph. In (a), $N = 10$, $\Delta = 3$, and $D = 3$. In (b), we show the routing spanning tree as seen from Node 1. Reproduced from [13], Figure 4-9.

5.3.2 $\Delta$-Nearest Neighbors Graphs

In a $\Delta$-Nearest Neighbors graph, for each node $i$, there are $\Delta$ directed connections from node $i$ to node $(i+1) \mod N$, $(i+2) \mod N$, ..., $(i+\Delta) \mod N$. Compared with a Generalized Moore graph, the $\Delta$-Nearest Neighbors graph provides the upper bound on the average minimum hop distance among all regular topologies with the same number of nodes and node degree. This type of topology is very flexible, allowing us to construct a $N$-node network with $\Delta$ ranging from 2 to $N - 1$, all with only slight modifications to existing infrastructure. In Figure 5-5, we present an example of a $\Delta$-Nearest Neighbors graph.

For a $\Delta$-Nearest Neighbors graph in which $N$ is not divisible by $\Delta$, the average minimum hop distances, $H_{\text{min},\Delta}$, is

$$H_{\text{min},\Delta}(\Delta) = \left(1 + \frac{\Delta}{2(N-1)}\right)\left[\frac{N-1}{\Delta}\right] - \frac{\Delta}{2(N-1)}\left(\left\lfloor\frac{N-1}{\Delta}\right\rfloor\right)^2$$

(5.17)

For a $\Delta$-Nearest Neighbors graph in which $N$ is divisible by $\Delta$, the average minimum hop distances, $H_{\text{min},\Delta}$, is

$$H_{\text{min},\Delta}(\Delta) = \frac{1}{2} + \frac{N-1}{2\Delta}$$

(5.18)
5.3.3 Symmetric Hamilton Graphs

A Hamilton graph contains a cycle that connects all the nodes but traverses each node exactly once. In a Symmetric Hamilton graph, each node connects to $\Delta$ other nodes with an even spacing parameter $s$ between the nodes in a cyclical fashion. In Figure 5-6, we present examples of Symmetric Hamilton graphs.

For a Symmetric Hamilton graph with $s$ odd, the average minimum hop distances, $H_{\text{min}_{\text{odd}} \text{ Ham}}$, is

$$H_{\text{min}_{\text{odd}} \text{ Ham}}(\Delta) = \frac{\Delta}{N - 1} + \frac{(N - \Delta - 1)(N + 5\Delta - 7)}{4(N - 1)(\Delta - 1)}$$  (5.19)

For a Symmetric Hamilton graph with $s$ even, the average minimum hop distances, $H_{\text{min}_{\text{even}} \text{ Ham}}$, is

$$H_{\text{min}_{\text{even}} \text{ Ham}}(\Delta) = \frac{\Delta}{N - 1} + \frac{(N - 2)^2 + 4(N - 2)(\Delta - 1) - 4(\Delta - 1)}{4(N - 1)(\Delta - 1)}$$  (5.20)

These values can be approximated as

$$H_{\text{min}\text{Ham}}(\Delta) \approx \frac{3}{4} + \frac{N - 2}{4(\Delta - 1)}$$  (5.21)

5.4 Summary of Chapter 5

In this chapter, we modelled the capital and operating expenditures of basic components that make up a MAN: fiber connections, transceivers, and switches (optical and optical-electronic-optical). We base power costs on a five year operating lifetime. We also presented a introduction to regular graph representations by summarizing the properties of Generalized Moore, $\Delta$-nearest Neighbors, and Symmetric Hamilton graphs.
Figure 5-6: Examples of Symmetric Hamilton graphs. In (a), $N = 8$, $\Delta = 4$, $s = 2$, $D = 2$. In (b), $N = 8$, $\Delta = 3$, $s = 3$, $D = 2$. In (c), $N = 10$, $\Delta = 3$, $s = 4$, $D = 2$. In (d), $N = 11$, $\Delta = 4$, $s = 3$, $D = 2$. Reproduced from [13], Figure 4-12.
Chapter 6

MAN Power Analysis and Results

In this chapter, we formulate the physical topology design problem as an optimization over the type of graph topology used to represent each MAN, the routing algorithm, and the network node degree $\Delta$. By solving the problem analytically, we obtain upper and lower bounds on the network cost that can provide insight into optimal and scalable network architectures.

We focus on solving for optimal $\Delta^*$ because there exists a fundamental cost trade-off between increasing fiber connections and increasing switching capability. In other words, if fiber plants dominate the cost function (e.g. extreme case of breaking new ground to lay all cable plants), we would prefer a sparsely connected network topology to minimize cost. On the other hand, if switching costs completely dominate the cost function, then we would prefer a fully connected network topology to utilize the relatively cheaper fiber connections. At the optimal $\Delta^*$, the cost trade-off is balanced, resulting in a minimal cost MAN design.

In this chapter, we present the optimal $\Delta^*$ normalized for network size for each topology (e.g. $\Delta$-Nearest Neighbors, Symmetric Hamilton, and Generalized Moore graphs). We then compare minimal normalized network cost and minimal normalized network cost per unit traffic across the optimal topologies. We further study the effects increasing network data rate or traffic loading have on the optimal normalized network connectivity, cost, and cost per unit traffic. Finally, we compare our results to the ones Guan found using a capital expenditure based cost model (i.e. excluding
6.1 Problem Formulation

The general form of the convex optimization problem is

\[
\min_{\{tpl\}, \{r.a.\}, \Delta} \quad C_n(N, \Delta, t) \text{ or } C_n(N, \Delta, t, r, R, \eta)
\]

s.t. \(2 \leq \Delta \leq N - 1\)

\[
\Delta \in Z^+
\]

\[N \text{ and } t \text{ are given}\]

where \(tpl\) denotes the type of regular topology, \(r.a.\) denotes the routing algorithm used (i.e. shortest path routing), \(N\) is the number of nodes in the network, and \(t\) is the traffic demand between any node pair (i.e. uniform all-to-all traffic).

The constraint \(\Delta \leq N - 1\) imposes an upper limit (i.e. full connectivity) to the possible values of optimal node degree \(\Delta^*\) and the constraint \(\Delta \geq 2\) ensures that the topology is more than one-connected for reliability. We ignore quantization effects for fiber connections so the cost per wavelength is fully prorated. We also omit the details of wavelength assignment since the costs and efficiencies of an architecture design do not depend on channel assignment. Guan details a feasible wavelength assignment methodology in Chapter 7 of his thesis [13].

The number of components (e.g. \(K_{OXC}\) and \(K_{O/E/O}\) for switch size), component cost functions, and average minimum hop distance, \(H_{\min}\), of each regular topology we study are all convex in \(\Delta\). Then, the aggregate network cost functions are also convex in \(\Delta\). Therefore, a local optimal \(\Delta^*\) must exist which is a global optimum. In the following sections, we solve for the optimal node connectivity, \(\Delta^*\), and the resulting global minimal normalized network cost.
6.2 Optimal Node Degree and Network Cost for \(\Delta\)-Nearest Neighbors Graphs

We start with analysis of the \(\Delta\)-Nearest Neighbors graph because this topology provides an upper bound on network cost under uniform all-to-all traffic. The generic normalized cost functions from Chapter 5 are

\[
C_{OXC_{\text{norm}}}(N, \Delta, t) = \alpha \Delta + t \beta_{OXC}(N-1)[H_{\text{min}}(N, \Delta) + 1]
\]

(6.2)

where \(\beta_{OXC} = \beta_{OXC,1} + \beta_{OXC,2}\), as defined in Section 5.2.3.

\[
C_{O/E/O_{\text{total}}}(N, \Delta, t, r, R, \eta) = \alpha \Delta + \beta_{O/E/O} \frac{rt(N-1)[H_{\text{min}}(N, \Delta) + 1]}{R\eta}
\]

(6.3)

We approximate the average minimum hop distance of a \(\Delta\)-Nearest Neighbors graph with the \(H_{\text{min}\Delta}\) expression for when \(N\) is divisible by \(\Delta\)

\[
H_{\text{min}\Delta}(\Delta) = \frac{1}{2} + \frac{N-1}{2\Delta}
\]

(6.4)

The normalized network cost functions to be minimized by the \(\Delta\)-Nearest Neighbors graph are

\[
C_{OXC_{\text{norm}}}(N, \Delta, t) = \alpha \Delta + t \beta_{OXC}(N-1) \left( \frac{3}{2} + \frac{N-1}{2\Delta} \right)
\]

(6.5)

\[
C_{O/E/O_{\text{total}}}(N, \Delta, t, r, R, \eta) = \alpha \Delta + \frac{r t \beta_{O/E/O}(N-1)}{R \eta} \left( \frac{3}{2} + \frac{N-1}{2\Delta} \right)
\]

(6.6)

Using Lagrange multipliers, we find the first order condition expressions with respect to \(\Delta\) to be

\[
\alpha - \frac{t \beta_{OXC}(N-1)^2}{2\Delta^2} = 0
\]

(6.7)

\[
\alpha - \frac{r t \beta_{O/E/O}(N-1)^2}{2R \eta \Delta^2} = 0
\]

(6.8)

The resulting optimal network connectivity for \(OXC\) switched \(\Delta\)-Nearest Neigh-
bors graphs are
\[ \Delta^* = (N - 1) \sqrt{\frac{\beta_{OXC}}{2\alpha}} \]  
(6.9)
and the optimal network connectivity for \( O/E/O \) switched \( \Delta \)-Nearest Neighbors graphs are
\[ \Delta^* = (N - 1) \sqrt{\frac{rt\beta_{O/E/O}}{2\alpha R\eta}} \]  
(6.10)
We note that these \( \Delta^* \) are still constrained by \( 2 \leq \Delta^* \leq N - 1 \).

To find the optimal normalized network cost, we substitute \( \Delta^* \) into the appropriate normalized network cost functions (6.5) or (6.6). The minimal network cost function for an \( OXC \) switched \( \Delta \)-Nearest Neighbors graph is
\[ C^{*}_{OXC_{norm}}(N, \Delta, t) = (N - 1) \left( \sqrt{2\alpha t\beta_{OXC}} + \frac{3t\beta_{OXC}}{2} \right) \]  
(6.11)
The minimal network cost function for an \( O/E/O \) switched \( \Delta \)-Nearest Neighbors graph is
\[ C^{*}_{O/E/O_{norm}}(N, \Delta, t, r, R, \eta) = (N - 1) \left( \sqrt{\frac{2rt\beta_{O/E/O}}{R\eta}} + \frac{3rt\beta_{O/E/O}}{2R\eta} \right) \]  
(6.12)

### 6.3 Optimal Node Degree and Network Cost for Symmetric Hamilton Graphs

We approximate the average minimum hop distance of a Symmetric Hamilton graph with
\[ H_{\text{min,Ham}}(\Delta) = \frac{3}{4} + \frac{N - 2}{4(\Delta - 1)} \]  
(6.13)
The normalized network cost functions to be minimized by the Symmetric Hamilton graph are
\[ C_{OXC_{norm}}(N, \Delta, t) = \alpha\Delta + t\beta_{OXC}(N - 1) \left( \frac{7}{4} + \frac{N - 2}{4(\Delta - 1)} \right) \]  
(6.14)
\[ C_{O/E/O_{\text{norm}}}(N, \Delta, t, r, R, \eta) = \alpha \Delta + \frac{rt\beta_{O/E/O}(N-1)(\frac{7}{4} + \frac{N-2}{4(\Delta-1)})}{R\eta} \] (6.15)

Using Lagrange multipliers, we find the first order condition expressions with respect to \( \Delta \) to be

\[ \alpha - \frac{t\beta_{OXC}(N-1)(N-2)}{4(\Delta^*-1)^2} = 0 \] (6.16)

\[ \alpha - \frac{rt\beta_{O/E/O}(N-1)(N-2)}{4R\eta(\Delta^*-1)^2} = 0 \] (6.17)

The resulting optimal network connectivity for \( OXC \) switched Symmetric Hamilton graphs are

\[ \Delta^* = 1 + \frac{1}{2} \sqrt{\frac{t\beta_{OXC}(N-1)(N-2)}{\alpha}} \] (6.18)

and the optimal network connectivity for \( O/E/O \) switched Symmetric Hamilton graphs to be

\[ \Delta^* = 1 + \frac{1}{2} \sqrt{\frac{rt\beta_{O/E/O}(N-1)(N-2)}{\alpha R\eta}} \] (6.19)

We note that these \( \Delta^* \) are still constrained by \( 2 \leq \Delta^* \leq N-1 \).

To find the optimal normalized network cost, we substitute \( \Delta^* \) into the appropriate normalized network cost functions (6.18) or (6.19). The minimal network cost function for an \( OXC \) switched Symmetric Hamilton graph is

\[ C^*_{OXC_{\text{norm}}}(N, \Delta, t) = \alpha + \left( \sqrt{\alpha t\beta_{OXC}} + \frac{7t\beta_{OXC}}{4} \right) \sqrt{(N-1)(N-2)} \] (6.20)

The minimal network cost function for an \( O/E/O \) switched Symmetric Hamilton graph is

\[ C^*_{O/E/O_{\text{norm}}}(N, \Delta, t, r, R, \eta) = \alpha + \left( \sqrt{\frac{\alpha rt\beta_{O/E/O}}{R\eta}} + \frac{7rt\beta_{O/E/O}}{4R\eta} \right) \sqrt{(N-1)(N-2)} \] (6.21)
6.4 Optimal Node Degree and Network Cost for Generalized Moore Graphs

The Generalized Moore graph provides a lower bound on network cost under uniform all-to-all traffic and minimum hop routing. We approximate the average minimum hop distance of a Generalized Moore graph by assuming $\Delta \approx \Delta - 1$ and $N \gg \Delta$ so that

$$H_{\text{min,undir, Moore}}(N, \Delta) = H_{\text{min, dir, Moore}}(N, \Delta) = \log_\Delta N$$  \hspace{1cm} (6.22)

The normalized network cost functions to be minimized by the Generalized graph are

$$C_{OXC_{\text{norm}}}(N, \Delta, t) = \alpha \Delta + t \beta_{OXC}(N - 1)(\log_\Delta N + 1)$$  \hspace{1cm} (6.23)

$$C_{O/E/O_{\text{norm}}}(N, \Delta, t, r, R, \eta) = \alpha \Delta + \frac{rt \beta_{O/E/O}(N - 1)(\log_\Delta N + 1)}{R \eta}$$  \hspace{1cm} (6.24)

Using Lagrange multipliers, we find the first order condition expressions with respect to $\Delta$ to be

$$\Delta(\ln \Delta)^2 = \frac{t \beta_{OXC}(N - 1) \ln N}{\alpha}$$  \hspace{1cm} (6.25)

$$\Delta(\ln \Delta)^2 = \frac{rt \beta_{O/E/O}(N - 1) \ln N}{\alpha R \eta}$$  \hspace{1cm} (6.26)

The resulting optimal network connectivity for $OXC$ switched Generalized Moore graphs is

$$\Delta^* = \frac{t \beta_{OXC}(N - 1) \ln N}{4\alpha} \left\{ W\left( \frac{t \beta_{OXC}(N - 1) \ln N}{2\sqrt{\alpha}} \right) \right\}^{-2}$$  \hspace{1cm} (6.27)

and the optimal network connectivity for $O/E/O$ switched Generalized Moore graphs is

$$\Delta^* = \frac{rt \beta_{O/E/O}(N - 1) \ln N}{4\alpha R \eta} \left\{ W\left( \frac{rt \beta_{O/E/O}(N - 1) \ln N}{2\sqrt{\alpha R \eta}} \right) \right\}^{-2}$$  \hspace{1cm} (6.28)

where $W(\bullet)$ denotes the Lambert function. The Lambert function is the inverse of the function $f(W) = W e^W$.

To find the optimal normalized network cost, we substitute $\Delta^*$ into the appro-
appropriate normalized network cost functions (6.23) or (6.24). The minimal network cost function for an OXC switched Generalized Moore graph is

$$C^*_{OXC_{norm}}(N, \Delta, t) = \alpha \Delta^* + t \beta_{OXC}(N - 1) \left( \frac{\ln N}{\ln \Delta^*} + 1 \right)$$

(6.29)

The minimal network cost function for an O/E/O switched Generalized Moore graph is

$$C^*_{O/E/O_{norm}}(N, \Delta, t, r, R, \eta) = \alpha \Delta^* + \frac{rt \beta_{O/E/O}(N - 1)}{R \eta} \left( \frac{\ln N}{\ln \Delta^*} + 1 \right)$$

(6.30)

### 6.5 Analyzing Optimal Node Degree

In this section, we evaluate the expressions for optimal network connectivity with representative values for parameters. Our goal is to describe how the node connectivity of three topologies — Δ-Nearest Neighbors, Symmetric Hamilton, and Generalized Moore graphs — scale with respect to the size of the network.

In Figure 6-1, we plot the optimal node degree $\Delta^*$ normalized by the maximum node degree in a fully connected network. A densely connected network has normalized optimal node degree $\frac{\Delta^*}{\Delta_{max}}$ closer to 1 while a sparsely connected network has $\frac{\Delta^*}{\Delta_{max}}$ closer to 0. We note that the optimal node degree node degree is bounded by $2 \leq \Delta^* \leq N - 1$ so that the normalized optimal node degree must lie between $\frac{2}{N-1} \leq \frac{\Delta^*}{\Delta_{max}} \leq 1$, where the lower bound corresponds to a ring topology and the upper bound is a fully connected mesh topology.

The first observation we make is that the connectivity of the Δ-Nearest Neighbor graph is independent of the size of the network. The normalized optimal node degree is $\frac{t \beta_{OXC}}{2\alpha}$ for OXC switched networks and is $\frac{t \beta_{O/E/O}}{2\alpha R \eta}$ for O/E/O switched networks, where all parameters are constants. The normalized optimal node degree of Symmetric Hamilton and Generalized Moore graphs decreases as the network size increases. This indicates that switching in those topologies is more cost efficient than fiber connections, which holds when the topologies have small values of $H_{min}$. We note that in fact, Symmetric Hamilton and Generalized Moore graphs do have smaller $H_{min}$.
values than $\Delta$-Nearest Neighbors graphs do, with the difference growing as the number of nodes in the network grows. As a result, for moderate to large sized networks of twenty or more nodes, we find that the Generalized Moore graph has the lowest network connectivity and is therefore the most efficient switching architecture.

We also observe the differences between $OXC$ and $O/E/O$ switched topologies and find that $OXC$ switched topologies are always more sparsely connected than $O/E/O$ switched ones. In Figures 6-1 and Figure 6-2, we plot the normalized optimal node degree for $O/E/O$ switched topologies in dashes. We find the dashed lines to always lie above their respective solid lines. Furthermore, as the data rate increases, the difference between the connectivity of $OXC$ and $O/E/O$ switched topologies grows. In Figure 6-1, we assume $r = 1\text{Gb/sec}$ and in Figure 6-2, $r = 10\text{Gb/sec}$. The optimal connectivity of $OXC$ switched topologies is independent of the data rate $r$. However, the magnitude of $r$ has a huge effect on the optimal connectivity of $O/E/O$ switched
topologies.

Figure 6-2: Normalized Optimal Node Degree of Network Topologies. The solid lines indicate OXC switched topologies while the dashed lines indicate O/E/O switched topologies. Parameter assumptions: $\alpha = \$400K$, $\beta_{OXC} = \$13.7K$, $\beta_{O/E/O} = \$100K$, $r = 10\text{Gb/sec}$, $t = 1$.

When $r = 10\text{Gb/sec}$, we find the networks to be fully connected under all three topologies for small and moderate sized networks, mainly due to the network connectivity constraint $\Delta^* \leq N - 1$. Even for large networks (more than 50 nodes), the Symmetric Hamilton and Generalized Moore graphs rely heavily on fiber connections ($\Delta^*$ lies between $0.5N$ and $0.8N$). As the data rate increases, the electronic router’s capability to process the signals efficiently and economically creates a bottleneck in the network. This effect results in fiber connections being more cost efficient than switching. Even in a network with 100 nodes, at $r = 10\text{Gb/sec}$, the three O/E/O switched network designs are dominated by direct fiber connections due to cost effectiveness.

Finally, we look at the effect of traffic loading on the optimal node connectivity among topologies. In Figure 6-3, we plot the normalized optimal node degree $\Delta^*$ of
OXC switched topologies with uniform all-to-all traffic $t = 1$ (solid lines) and $t = 3$ (dashed lines), with $r = 1$Gb/sec. We find that increasing $t$ raises the normalized optimal node degree for all topologies.

Figure 6-3: Normalized Optimal Node Degree of Network Topologies. The solid lines indicate OXC switched topologies under uniform all-to-all traffic $t = 1$ while the dashed lines indicate $t = 3$. Parameter assumptions: $\alpha = \$400K$, $\beta_{OXC} = \$13.7K$, $\beta_{O/E/O} = \$100K$, $r = 1$Gb/sec.

6.6 Analyzing Minimal Network Cost Normalized for Network Size

In this section, we address the essential question: which topology is cost optimal? We evaluate the expressions for minimal network cost with representative values for parameters. Our goal is to look at how the cost of three topologies – $\Delta$-Nearest Neighbors, Symmetric Hamilton, and Generalized Moore graphs – scale with respect to the size of and the traffic carried on the network.
In Figure 6-4, we plot the network cost normalized by the number of nodes in the network. As expected, the normalized cost per node grows as the network size increases, leading to scalability concerns. In small OXC switched networks (less than 10 nodes), we find that the normalized network cost of the three topologies are similar. As the network grows, Generalized Moore graphs are more economical. For example, for networks with \( N = 100 \) nodes, Symmetric Hamilton graphs cost over 40% more and \( \Delta \)-Nearest Neighbors graphs cost approximately 50% more to install and operate than Generalized Moore graphs. In O/E/O switched networks at \( r = 1 \) Gb/sec, the cost relationship among topologies also holds. In moderate to large sized networks, Generalized Moore graphs are the least expensive topology, although the minimal normalized cost for each topology is very similar.

![Figure 6-4: Minimal Normalized Network Cost of Network Topologies. The solid lines indicate OXC switched topologies while the dashed lines indicate O/E/O switched topologies. Parameter assumptions: \( \alpha = 400 \) K, \( \beta_{OXC} = 13.7 \) K, \( \beta_{O/E/O} = 100 \) K, \( r = 1 \) Gb/sec, \( t = 1 \).]

How does the normalized network cost change as the network data rate rises? Again, the cost of deploying and operating OXC switched networks is independent
of the data rate. However, as shown in Figure 6-4, the cost curves for $O/E/O$ switched topologies rise across the board. Because the optimal network connectivity is constrained by $\Delta^* \leq N - 1$, at $r = 10\text{Gb/sec}$, all three topologies have similar (close to fully connected, if not fully connected) node degrees. Then, for high data rates, as the network size grows, the corresponding network cost for $O/E/O$ switched topologies are identically high per node.

![Figure 6-5: Minimal Normalized Network Cost of Network Topologies. The solid lines indicate $OXC$ switched topologies while the dashed lines indicate $O/E/O$ switched topologies. Parameter assumptions: $\alpha = \$400K$, $\beta_{OXC} = \$13.7K$, $\beta_{O/E/O} = \$100K$, $r = 10\text{Gb/sec}$, $t = 1$.](image)

Finally, we look at the effect of traffic loading on normalized network cost among topologies. In Figure 6-6, we plot the minimal normalized network cost $C^*_n$ of $OXC$ switched topologies with uniform all-to-all traffic $t = 1$ (solid lines) and $t = 3$ (dashed lines), with $r = 1\text{Gb/sec}$. We find that increasing $t$ raises the normalized cost for all topologies. For moderate to large sized networks, the Generalized Moore graph is still the lower bound on total network cost while the $\Delta$-Nearest Neighbors is the upper bound.
6.7 Analyzing Minimal Network Cost Normalized for Network Size and Traffic

In this section, we take a closer look at scalability of three network topologies — $\Delta$-Nearest Neighbors, Symmetric Hamilton, and Generalized Moore graphs — by evaluating the minimal network cost normalized by both network size and traffic demand. Under uniform all-to-all traffic, the add-drop traffic at each node grows linearly with the network size and both add-drop and pass-through traffic grow linearly with the traffic demand between node pairs. Therefore, the best indicator of network scalability with respect to capital and operating expenditures is the minimal normalized network cost per unit traffic, $\frac{C^*}{t(N-1)}$.

In Figure 6-7, we plot the minimal normalized network cost per unit traffic. As expected, $\frac{C^*}{t(N-1)}$ decreases as the network size increases. The normalized cost per
unit traffic reaches a constant asymptote for network sizes of $N \geq 10$. However, the normalized network cost per unit traffic continue to fall for Generalized Moore topologies. This observation is an important key to scalability. In future networks, traffic load is certain to increase (i.e. industry estimates WAN traffic growth of 30-40% each year and we assume a hundredfold growth over the next five to ten years), so that even if no new nodes are deployed, the traffic carried by each node will grow explosively. The Generalized Moore graph is the only topology that has a decreasing marginal normalized network cost. This relationship holds for both $OXC$ and $O/E/O$ switched networks.

![Figure 6-7: Minimal Normalized Network Cost of Network Topologies. The solid lines indicate $OXC$ switched topologies while the dashed lines indicate $O/E/O$ switched topologies. Parameter assumptions: $\alpha = $400K, $\beta_{OXC} = $13.7K, $\beta_{O/E/O} = $100K, $r = 1$Gb/sec, $t = 1$.](image)

In Figure 6-8, we show the effect of traffic loading on the minimal normalized network cost per unit traffic. We plot $\frac{C^*}{i(N-1)}$ for $OXC$ switched topologies with $r = 1$Gb/sec. We find that as traffic loading increases, the network cost per node per unit traffic falls. At all times, the Generalized Moore graph topology is the most cost-effective.
efficient.

Figure 6-8: Minimal Normalized Network Cost of Network Topologies. The solid lines indicate OXC switched topologies under uniform all-to-all traffic $t = 1$ while the dashed lines indicate $t = 3$. Parameter assumptions: $\alpha = \$400K$, $\beta_{OXC} = \$13.7K$, $\beta_{O/E/O} = \$100K$, $r = 1$Gb/sec.

### 6.8 Comparing Power and Cost Minimal with Cost Minimal Network Topologies

In this section, we compare the dual power and cost optimized network topologies based upon our expanded joint optimization problem formulation with the original solution based on Guan’s cost minimization formulation [13]. We search specifically for differences in optimal connectivity among the three network topologies — $\Delta$-Nearest Neighbors, Symmetric Hamilton, and Generalized Moore graphs — resulting from enhanced cost models incorporating operating expenditures.

We present comparisons only for OXC switched topologies. Any conclusions
extend readily to O/E/O switched MAN designs due to the similarities we found in Sections 6.5 through 6.7 between the two switching architectures.

In Figure 6-9, we plot the optimal node degree $\Delta^*$ normalized by the maximum node degree in OXC switched network topologies. The dashed lines trace $\Delta^*$ under power-enhanced cost modelling and solid lines graph $\Delta^*$ under the original cost modelling. We find that the Generalized Moore graph has the lowest network connectivity for moderate to large sized networks and that the $\Delta$-Nearest Neighbor graph has the highest network connectivity, under either cost model. We see that the optimal normalized node degree increases by approximately $0.01N$ in Generalized Moore graphs with more than 50 nodes when power costs are incorporated. This translates into an increase of almost 20% over Guan’s proposed optimal connectivity.

![Figure 6-9: Effect of Power Cost on Optimal Node Degree. The solid lines indicate OXC switched topologies in which the cost function is modelled with only capital expenditures while dashed lines indicate OXC switched topologies in which the cost function consists of both capital and operating expenditures. Parameter assumptions: $\alpha = \$400K, \beta_{OXC} = \$10K$ for solid lines, $\$13.7K$ for dashed lines, $t = 1$.](image-url)
We also look at the effect of traffic loading on the optimal node connectivity among OXC switched topologies. In Figure 6-10, we plot the normalized optimal node degree $\Delta^*$ to the capital expenditure minimization problem in solid lines next to the capital and operating expenditure minimal solutions in dashed lines. We find that tripling $t$ raises the connectivity significantly under either cost model (e.g. at $N = 100$, $\frac{\Delta^*}{N-1}$ is almost double its value under $t = 1$). The difference between optimal connectivities increases slightly. For example, at $N = 100$, the $\frac{\Delta^*}{N-1}$ of the cost optimized Generalized Moore graph is $0.08N$ while the $\frac{\Delta^*}{N-1}$ of the power and cost optimized graph is just less than $0.1N$, representing an increase of approximately 25% over Guan’s proposed optimal connectivity.

![Figure 6-10: Effect of Power Cost on Optimal Node Degree under Increased Traffic Loading. The solid lines indicate OXC switched topologies in which the cost function is modelled with only capital expenditures while dashed lines indicate OXC switched topologies in which the cost function consists of both capital and operating expenditures. Parameter assumptions: $\alpha = \$400K$, $\beta_{OXC} = \$10K$ for solid lines, $\$13.7K$ for dashed lines, $t = 3$.](image)
Next, we consider the effect of power expenses on the minimal normalized network cost. In Figure 6-11, we plot the network cost $C^*_n$ normalized by network size for $OXC$ switched network topologies. The dashed lines trace $C^*_n$ under power-enhanced cost modelling and the solid lines graph $C^*_n$ under the original cost modelling. We find that the Generalized Moore graph is the most economical topology for moderate to large sized networks and that the $\Delta$-Nearest Neighbor graph is the most expensive to install and operate, under either cost model. In Figure 6-12, we replicate the plot with increased traffic loading ($t$ wavelengths between any node pair is tripled). Similar cost savings can be achieved by using Generalized Moore graphs. Under either traffic loading scenario, we find that the network cost increases slightly due to power considerations, as is expected.

Next, we discuss the effect of including operating expenditures on minimal normalized network cost per unit traffic. In Figure 6-13, we plot the network cost $C^*_n$ normalized by both network size and traffic demand for $OXC$ switched topologies. We find that the Generalized Moore graph is still the only scalable network topology for moderate to large sized networks under either the power enhanced or capital expenditure only cost model, as it is the only topology with a decreasing marginal cost function. For a network with 100 nodes, we find that incorporating power costs increases the minimal normalized network cost for the Generalized Moore graph by $15$ per unit traffic, an increase of 25% over the original minimal network cost achieved by minimizing capital expenditures.

In Figure 6-14, we regraph the minimal normalized network cost per unit traffic with increased traffic loading ($t$ wavelengths between any node pair is tripled). Under heavier traffic demand, the difference between normalized network costs per unit traffic is even larger. For example, at $N = 100$, the cost per node per unit traffic of the Generalized Moore graph optimized for power and cost is over 30% more expensive than the cost of Guan’s original solution.
Figure 6-11: Effect of Power Cost on Normalized Network Cost. The solid lines indicate OXC switched topologies in which the cost function is modelled with only capital expenditures while dashed lines indicate OXC switched topologies in which the cost function consists of both capital and operating expenditures. Parameter assumptions: $\alpha = 400K$, $\beta_{OXC} = 10K$ for solid lines, $13.7K$ for dashed lines, $t = 1$.

6.9 Summary of Chapter 6

In this chapter, we provided detailed capital and operating expenditure minimization results for the MAN, based upon a uniform all-to-all traffic model. We found that the Generalized Moore graph with node degree between $0.05N$ and $0.08N$ is both power and cost minimal for a purely optical network. In fact, for moderate to large sized networks, the Generalized Moore graph forms the lower bound on connectivity and cost regardless of switching architecture, data rate, and traffic loading while the $\Delta$-Nearest Neighbor topology provides the upper bound. When network cost is normalized for both number of nodes and units of traffic, we found that the Generalized
Figure 6-12: Effect of Power Cost on Normalized Network Cost under Increased Traffic Loading. The solid lines indicate OXC switched topologies in which the cost function is modelled with only capital expenditures while dashed lines indicate OXC switched topologies in which the cost function consists of both capital and operating expenditures. Parameter assumptions: $\alpha = 400K$, $\beta_{OXC} = 10K$ for solid lines, $13.7K$ for dashed lines, $t = 3$.

Moore graph is the only topology with a decreasing cost curve as the network size grows. Compared with a capital expenditure minimized network, our results show the overall power and cost minimal network architecture is also the Generalized Moore graph topology, albeit one with an approximately 20-25% higher network connectivity and cost as the relative switching cost increases due to power considerations.
Figure 6-13: Effect of Power Cost on Normalized Network Cost Per Unit Traffic. The solid lines indicate OXC switched topologies in which the cost function is modelled with only capital expenditures while dashed lines indicate OXC switched topologies in which the cost function consists of both capital and operating expenditures. Parameter assumptions: $\alpha = $400K, $\beta_{OXC} = $10K for solid lines, $13.7K$ for dashed lines, $t = 1$. 
Figure 6-14: Effect of Power Cost on Normalized Network Cost Per Unit Traffic under Increased Traffic Loading. The solid lines indicate $OXC$ switched topologies in which the cost function is modelled with only capital expenditures while dashed lines indicate $OXC$ switched topologies in which the cost function consists of both capital and operating expenditures. Parameter assumptions: $\alpha = \$400K$, $\beta_{OXC} = \$10K$ for solid lines, $\$13.7K$ for dashed lines, $t = 3$. 
Chapter 7

Conclusion

Advancements in technology over the past decade have led to new services and applications that require data transmission. For example, consumers have increased subscriptions to services such as streaming HD television, Voice over IP, video conferencing, interactive gaming, etc. that require low latency, high bitrate, and high reliability networks. The increase in underlying demand for network capacity leads to a need to build scalable networks with adequate capacity, throughput, and delays while maintaining acceptable capital and operating expenditures. To date, most previous research has focused on cost minimization of wide area and metropolitan area networks. However, the optimal network topology needed is one in which power consumption grows sustainably as the number of users, capacity of network, and speed of network increase. Indeed, one of the most pressing limiting factors to network growth today is power consumption.

In this thesis, we concentrated on finding scalable WAN designs with respect to power constraints and optimal MAN topologies with minimal capital and operating expenditures.

7.1 Summary of Contributions

In this thesis, we presented detailed analysis of the United States IP backbone to provide power minimizing design heuristics for WAN architecture. We also expanded
a cost-based joint optimization formulation of the MAN to include power expenditure estimates over the network's operating lifetime and provided closed form solutions for cost and power minimal MAN topologies.

In the introductory chapter of this thesis, we presented the motivations for developing scalable network architectures with respect to power and outlined our approach to analyze and optimize both the WAN and MAN. We briefly described related research and distinguished our problem solving approach from previously conducted work. In particular, our detailed modelling on the WAN is based on population demographics, resulting in one of the most realistic and accurate traffic matrices to be studied. The incorporation of operating expenditures in cost optimization of the MAN completes and confirms the analytical work started by Guan [13].

In Chapter 2, we laid the foundation for our thesis by introducing the basic components that make up a network: transmitter/receiver pairs (transceivers), erbium-doped fiber amplifiers (EDFA), switches (optical and optical-electronic-optical), and routers. For each component, we described how it is used and modelled in a network and provided its representative power consumption value.

In Chapter 3, we built the WAN model by discussing the physical topology, traffic model, network parameters, and potential network designs. Our traffic model is based upon network demographics so that the expected traffic between any two nodes is a function of the population of each node's metropolitan area (MSI). Then, nodes serving densely populated regions both originate and terminate a larger proportion of overall network traffic compared with nodes serving sparsely populated regions. Variance in the traffic demand between any node pair reflects a likely quote of network fluctuation. We also described four main network designs. Under the patched bypass design, direct, non-switched, lightpaths are allocated between any node pair. With a standard bypass network, lightpaths between any node pair are switched at intermediate nodes. Under a non-bypass design, lightpaths between any node pair are terminated and re-initiated by routers at intermediate nodes. Finally, under a groomed non-bypass, whole lightpaths between any node pair are switched while residual or fractional lightpaths are terminated and groomed by routers at intermediate
nodes.

In Chapter 4, we provided detailed power optimization results and design heuristics for the WAN. We found that optical bypass networks are the most scalable architecture design with respect to power consumption, especially when quality of service, network flexibility, reliability, and protection are considered. The power consumption of the standard bypass design can be further improved through a hybrid patched bypass/bypass network. Under the hybrid scheme, whole wavelengths of core, stable traffic between node pairs is routed via direct, fixed lightpaths using patch panelling to avoid lightpath switching at intermediate nodes and all other unexpected, fluctuating, and/or bursty traffic is switched on a standard optical bypass network. We found the hybrid network to be power minimal with exceptional latency, transmission, and reliability performance. We also analyzed power distribution among components and found the OXC switch to be most scalable and the O/E/O_{OXC} switch to be most wasteful (signals traversing an O/E/O_{OXC} switch are regenerated twice due to its use of two transceivers, compared to one regeneration using an O/E/O_{ELEC} switch and no regenerations using an OXC switch). Finally, we proved that shortest path and minimum hop routing is power optimal and traffic balanced routing should be avoided when a majority of the traffic demand between any pair of nodes is sufficiently large to warrant whole wavelengths, as is the case when network capacity grows from the current 30-Terabit/sec to a projected 300-Terabit/sec capacity.

In Chapter 5, we modelled the capital and operating expenditures of basic components that make up a MAN: fiber connections, transceivers, and switches (optical and optical-electronic-optical). We base power costs on a five year operating lifetime. We also presented an introduction to regular graph representations by summarizing the properties of Generalized Moore, $\Delta$-nearest Neighbors, and Symmetric Hamilton graphs.

Finally, in Chapter 6, we provided detailed capital and operating expenditure minimization results for the MAN, based upon a uniform all-to-all traffic model. We found that the Generalized Moore graph with node degree between $0.05N$ and $0.08N$ is both power and cost minimal for a purely optical network. In fact, for
moderate to large sized networks, the Generalized Moore graph forms the lower bound on connectivity and cost regardless of switching architecture, data rate, and traffic loading while the \( \Delta \)-Nearest Neighbor topology provides the upper bound. When network cost is normalized for both number of nodes and units of traffic, we found that the Generalized Moore graph is the only topology with a decreasing cost curve as the network size grows. Compared with a capital expenditure minimized network, our results show the overall power and cost minimal network architecture is also the Generalized Moore graph topology, albeit one with an approximately 20-25\% higher network connectivity and cost as the relative switching cost increases due to power considerations.

Our main findings for the wide area network can be summarized:

- Optical bypass is the single most scalable network design. It is power efficient, flexible, and reliable. Furthermore, it can be combined with patched panelling to create a hybrid patched bypass-bypass network design that handles a static subset of traffic in direct, non-switched, power minimal lightpaths and remaining bursty traffic in a power efficient OXC switched manner.

- \( O/E/O \) switches and routers are power intensive and should not be used in large scale backbone networks. In fact, router ports consume so much more power compared to OXC switches that achieving high link utilization though traffic grooming consumes more power than a pure optical bypass network with lower link utilization.

- Shortest path routing is more power efficient than traffic balanced routing approaches. In fact, the power minimized routing algorithm is a function of shortest path and shortest hop routing.

Our main findings for the metropolitan area network can be summarized:

- As the ratio of switch cost to fiber connection cost increases, the optimal node connectivity increases. For example, we found that when power considerations raised the switch cost by 30-40\%, \( \Delta^* \) increased by 25\%. This increased network
connectivity does not require new fiber to be pulled. Instead, patch panels can be installed to create the desired fiber topology from the given cable plant topology.

- Any subset of the traffic demand that is stable and fixed can be hard-wired using patched panelling to improve further upon the provided closed form cost minimal solution. This approach is similar to the patched bypass design in WAN in which lightpaths are statically provisioned and the stable traffic served on this fixed network, reducing the total traffic demand to be switched in the flexible part of the MAN.

7.2 Future Work

One of the biggest challenges network designers face is accurate estimation of traffic fluctuation in a network. In our power analysis of the WAN, we had assumed that the optical flow traffic can be modelled using a constant Poisson arrival process. We did so mainly because the variance in the mean of the true arrival rate is a proprietary industry secret for each network provider and unknown to us. The size of fluctuations in network traffic from day to night and from week to week has a profound effect on the viability of a patched bypass network design and influences the allocation of network resources in a hybrid patched bypass/bypass network dedicated to direct patched lightpaths between nodes. We identify more accurate modelling of traffic variance in the WAN as a critical area in future research.

Another direction for future work lies in more accurate forecasting of network cooling costs as the number of ports at each node scales to process increased traffic demand. Previous research indicated that air cooling consumes approximately the same amount of power required for component operation (i.e. a 100% overhead). However, in our research, we had argued that the actual cooling curve should be piecewise-linear or super-linear as a function of port count so that as the heat density on a rack exceeds the limits of air cooling, more technically advanced, power intensive, and expensive liquid or mist cooling is required. Depending on the shape of the
actual cooling curve, a traffic balanced routing algorithm may have merits over a shortest path/minimum hop routing algorithm because it would shift lightpaths off of extremely highly loaded network edges with power consumption in the steepest region of the cooling curve to lightly loaded edges that still operate in the linear air-cooled region of the curve.
Bibliography


[16] Bookham Inc. Xfp optical transceivers for 40km 10g datacom and storage applications, Feb 2007.


