Wavelength Routing for All-Optical Networks

by

Richard A. Barry

S.B., Massachusetts Institute of Technology 1989
S.M., Massachusetts Institute of Technology 1989

Submitted to the Department of
Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

September 1993

Signature of Author . ...
Department of
Electrical Engineering and Computer Science
September 1993

Certified by ....
Pierre Humblet
Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Frederic R. Morgenthaler
Chairman, Committee on Graduate Students

Wavelength Routing for All-Optical Networks

by

Richard A. Barry

Submitted to the Department of
Electrical Engineering and Computer Science
on September 1993, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We consider all-optical networks using a combination of wavelength and time division multiplexing where the path of a signal is determined by the network switches, the wavelength of the signal, and the signal origin. We present lower and upper bounds on the minimum number of wavelengths needed based on the connectivity requirements of the users and the number of switching states. The bounds hold for all networks with switches, wavelength routing, and wavelength changing devices. Passive and configurable devices in both non-blocking and blocking networks are considered. Switchless networks with near optimal wavelength re-use are presented.

Keywords: Wavelength Routing, All-Optical Networks, Connectors.

Thesis Supervisor: Pierre Humblet
Professor of Electrical Engineering

Acknowledgements

This thesis is dedicated to my friends, family, and loved ones.

I would especially like to thank my thesis advisor, Prof. Pierre Humblet, for guiding me through this endeavor.

Many thanks also go to my thesis readers Prof. Robert Gallager, Prof. Robert Kennedy, and Dr. Steve Finn for their interest, comments, and suggestions.

I thank my academic advisor, Prof. Dimitri Bertsekas, for his valuable academic and career advise.

Many students and former students of MIT offered technical support and advise. I would particularly like to acknowledge David Tse, Jane Simmons, Phil Lin, Peter Li, Rajesh Pankaj, Abhay Parekh, Emre Telatar and Sanjeev Kulkarni for their input. Your time and efforts are greatly appreciated.

Most of this research was performed at the Laboratory for Information and Decision Systems, MIT. Thanks to Sheila Hegarty, Margaret Flaherty, Kathleen O'Sullivan, Barbara Peacock-Coady, Giampiero Sciutto, and Nancy Young.

Part of this research was performed at the Institut Eurecom, Sophia Antipolis, France. To all the administration, staff, faculty, and students of Eurecom, thank you. In particular, I would like to thank Constantinos Papadias, Karim Maouche, and Revital Marom for helping to make my stay interesting and enjoyable.

This research was supported in part by the Army Research Office under contracts DAAL03-86-K-0171 and DAAL03-92-G-0115, The National Science Foundation Grant No. NCR-9206379, and the Defense Advanced Research Project Administration Grant #MDA972-92-J-1038.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 All-Optical Networks

We consider networks supporting high data rate *sessions*. Each session has an *origin* and a *destination*. At the origin, the data is converted from the electrical domain to the optical domain by modulating a laser. At the destination, the data is converted from the optical domain to the electrical domain via a photodiode and demodulation.

Optical networks consist of fiber optic links and nodes. They can be divided into two broad classes based on the function of the nodes: All-Optical Networks (AONs) and Electro-Optical Networks (EONs). The EON is the conventional approach. In an EON, also called Second Generation Optical Networks [Gre93] and Multi-Hop Networks [Aca93], the nodes perform optical to electronic conversion, electronic routing, e.g. virtual circuit or packet routing, followed by electronic to optical conversion. An advantage of EONs is that electronic routing can be performed. However, this advantage is also its potential downfall since the total throughput on the incoming links to a node must not exceed the electronic processing speed of the node. This has been called the *electronic bottleneck* [Gre93]. The name electronic bottleneck is somewhat of a misnomer; it is actually a cost/throughput trade-off since the throughput of a network is not necessarily limited by the throughput of a node. That is, high throughput networks can be made by simply increasing the number of nodes of the network. Note also that the number of links may also increase to interconnect the large number of nodes.

There is another way to describe the electronic bottleneck. In a network of high data rate sessions, where the data rate of a session is on the order of the processing speed of a node, then any given node can only process a small number of sessions. Therefore, networks with lots of sessions require lots of nodes. Again we see that the electronic bottleneck is a cost/throughput trade-off.

In All-Optical Networks (AONs), a.k.a. Third Generation Optical Networks [Gre93], a.k.a. Single-Hop Networks, the sessions remain in the optical domain from the origin to the destination. The nodes of the network are all-optical and eliminate the electronic bottleneck. Several laboratory and field experiment AONs have been demonstrated: Bellcore's LambdaNet [GKV$^+$90], AT&T Bell Laboratory's Wavelength Division Networks [WK$^+$90, KISS88, GS$^+$88], British Telecom's Wavelength Routing Network [W$^+$91], Rainbow-1 Network [Gre93], and NTT's 100 Wavelength Network [T$^+$90]. A description of these networks can be found in [Gre93]. Recently more ambitious AONs have been proposed for study: Linear Lightwave Networks (LLNs) [Ste90, Pan92], The MIT/DEC/AT&T AON [A$^+$93], and The Bellcore/Columbia/UPENN AON.

AONs eliminate the electronic bottleneck; however, since electronic routing is not possible, all-optical routing methods are required. One option is all-optical packet routing. Here the nodes of the network perform basic logic functions on the optical signals in order to route the messages from the inputs to the outputs of the node. There are severe technological problems with optical packet routing at this time. First of all, only the most basic of logic functions can be implemented. Second is the lack of optical buffers. If these technological limitations can be overcome or circumvented, optical packet routing may be the design of choice for future high speed networks.

Optical packet routing is probably the most technologically distant all-optical routing method. On the other hand, broadcast routing has already been demonstrated. In a broadcast network, each destination hears every signal from each origin. The signals are multiplexed using wavelength division multiplexing (WDM), time division multiplexing (TDM), polarization division multiplexing, code division multiplexing, or some combination of the above. Examples of broadcast networks that have been demonstrated are Bellcore's LambdaNet [GKV+90], IBM's Rainbow-1 [Gre93], and NTT's 100 Wavelength Network [T+90], achieving throughputs of 32 Gb/s, 9.6 Gb/s , and 62 Gb/s, respectively. Each of these networks use a combination of WDM and TDM to multiplex signals from different sources.

Broadcast networks have two inherent flaws which make them unscalable. First of all, the total throughput of a broadcast network can be no more than the total throughput of a fiber. Second, since each signal reaches each destination, there is an inherent power splitting loss which grows with the size of the network.

Another option is the topic of this thesis: wavelength routing [Goo89, Gre93]. Informally, the path a signal takes is a function of the wavelength of the signal. Most wavelength routing nodes are passive devices which work on interference effects and therefore do not suffer from the electronic bottleneck.

The next two sections describe wavelength routing networks in detail. Afterwards, important examples that we will return to many times are presented.

1.2 Wavelength Routing

In a wavelength routing all-optical network (λ-routing AON), the path a signal takes is solely a function of the state of the devices, the wavelength of the signal, and location of the signal transmitter. If the paths are under control of the network through the use of switches, configurable wavelength routing devices, or configurable wavelength changing devices, we say that the network is *configurable.* Otherwise, the network is *passive* or *fixed.*

A trivial example of a configurable wavelength routing node is an *elementary* 2×2 switch. The switch has 2 *states*, bar and cross. In the bar state, each signal on each wavelength from input i is routed to output i, for $i = 0, 1$. In the cross state, each signal on each wavelength from input i is routed to output $i + 1 \pmod 2$.

In a passive AON, the path is only a function of the wavelength and location of the signal transmitter. Here, the input/output connectivity of the network is fixed but routing through the network is possible since the connectivity of the devices, although also fixed, is a function of the optical wavelength. Connections are established by the tuning of the transmitters and/or receivers. Broadcast networks, discussed above, are the simplest example of a passive network where there exists a path from every input to every output on every wavelength.

A more complicated example of a passive network is shown in Fig. 1-1. The network has 2 origins and 3 destinations and uses passive λ-routing to establish sessions between origins and destinations. Each origin (destination) has one tunable laser (receiver) and **only one fiber** to access the network. The wavelength routing nodes (λ-nodes) selectively route the signals from the origins to the destinations based on wavelength only. Paths for three wavelengths Red, Green, and Blue are also shown in Fig. 1-1. Since the network is passive, the only freedom, after the network topology and wavelength paths have been determined, is in the tuning of the transmitters and receivers to different wavelengths.

Suppose that a session requires a full wavelength of bandwidth. Then the network in Fig. 1-1 can support sessions between any matching of the origins to the destinations (without multi-casting) except the matching $\phi = \{(1, Y), (2, Z)\}$. To see that this matching cannot be supported, notice that origins 1 and 2 must both be assigned wavelength Red to reach their intended destination. Since there is a Red-path from 2 to Y, two Red signals collide at destination Y. This need not happen for any other pair of sessions provided wavelengths are assigned properly.

However, if a session only requires half a wavelength of bandwidth so that sessions on the same wavelength can be time multiplexed, the network can support $\phi = \{(1,Y),(2,Z)\}$.

Since we are allowing the use of wavelength conversion within the network, a signal launched from a transmitter may arrive at a receiver on a different wavelength. In fact, a signal launched from a transmitter may arrive at a variety of receivers on many different wavelengths and/or arrive at a receiver on several different wavelengths.

Wavelength converters are represented by π-nodes. As for λ-nodes, there are two types of π-nodes: configurable and fixed. An example of a network with a fixed π-node is shown in Fig. 1-2. Here if a signal on wavelength Red is launched from source 2, it arrives at destination Y on wavelength Red **and** at destination Z on wavelength Green.

In the general situation, the π-nodes are configurable, e.g. depending on the state of the device, Red is converted to Green or to Blue. An AON containing configurable π-nodes is itself configurable. Also, if a passive AON contains π-nodes, they are, by definition, fixed.

1.3 Network Model

We consider networks with M_t transmitters, M_r receivers, and F wavelengths. When $M \stackrel{\text{def}}{=} M_t = M_r$, the network is *symmetric* and M is the number of users. Each transmitter (receiver) is connected to one outgoing (incoming) fiber. To model wavelength changing, we define an *origin-destination* channel, or OD channel, as an ordered pair of wavelengths and use the notation $f:f'$ to represent an OD channel. We say that transmitter n is connected to receiver m on OD channel $f:f'$ if a signal launched from n on wavelength f is received at m on wavelength f'. If a transmitter or receiver is tuned to wavelength f, we say it is *assigned* f. Note that there is no assumed relationship between the OD channels connecting transmitter n to receiver m and the OD channels connecting transmitter m to receiver n. Using the OD channel ter-

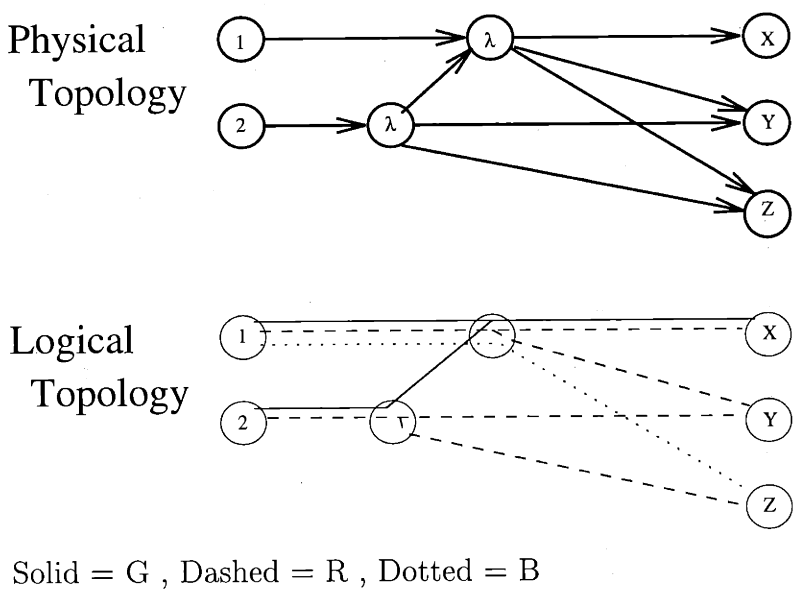


Figure 1-1: Example of a λ-routing network without wavelength changing

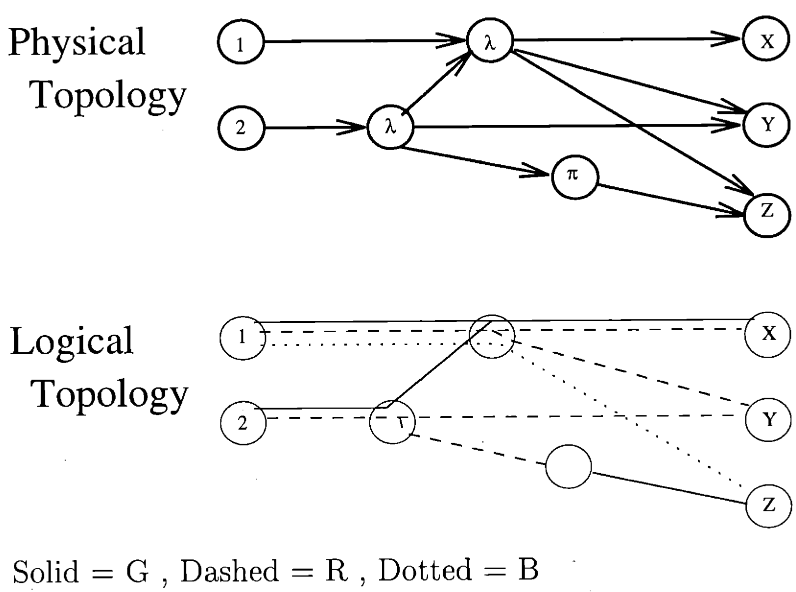


Figure 1-2: Example of a λ-routing network with wavelength changing

minology, the connectivity of a λ-routing network can be fully described by the set $\mathcal{H} = \{H_\psi | \psi \in \Psi\}$, where $H_\psi(n, m)$ is the set of OD channels connecting transmitter n to receiver m in state ψ, and Ψ are the *switching states* of the network. A switching state ψ represents the combined state of **all** configurable switches, λ-nodes, and π-nodes in the network.

In networks without wavelength changing, $f\!:\!f' \in H_\psi(n, m)$ implies that $f = f'$. In this case, we will use the obvious short hand notation of f for $f\!:\!f$.

If $|\Psi| = 1$, the network is passive. In this case, the connectivity of the AON is specified by a single connectivity matrix H. For example, the passive λ-routing network shown in Fig. 1-2 has the connection matrix,

$$H = \begin{bmatrix} \{G\!:\!G, R\!:\!R\} & \{R\!:\!R\} & \{B\!:\!B\} \\ \{G\!:\!G\} & \{R\!:\!R\} & \{R\!:\!G\} \end{bmatrix}$$

Note that our network model is a generalization of a conventional circuit switched network [Hui90]. In our terms, a conventional circuit switched network has $F = 1$ wavelength and $|\Psi| > 1$ states.

1.4 Light Tree AONs (LT-AON)

A symmetric *3-stage interconnection AON* is shown in Fig. 1-3. The first stage consists of M tunable transmitters, M wavelength demultiplexers, and up to M arbitrary fixed wavelength changers. The second stage consists of single wavelength fiber *trunks*. The third stage consists of up to M arbitrary fixed wavelength changers, M wavelength multiplexers and M tunable receivers. The first and third stages are called *peripheral stages* or *peripheral nodes*.

Each output (input) of the demultiplexers (multiplexers) are connected to at most one wavelength changer. In addition, each wavelength changer is connected to at most one trunk. Therefore, a signal from a transmitter travels on at most one

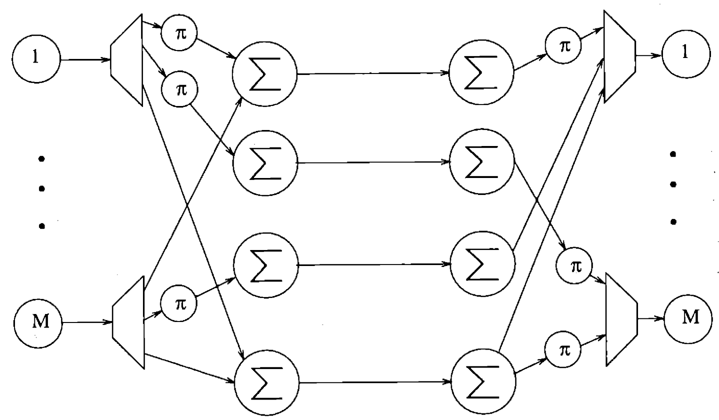


Figure 1-3: LT-AON

trunk. Similarly, a receiver can hear at most one trunk on each wavelength.

We call the trunks, along with the transmitters and receivers connected to a given trunk a *light tree*. The transmitters are the roots of the tree and the receivers as the leaves. The number of trees is simply the number of trunks and the wavelength of a tree is by definition the wavelength of the trunk. Note that the wavelength a transmitter uses to reach a trunk or a receiver uses to hear a trunk is not necessarily the wavelength of the trunk, due to the presence of wavelength changers. But since each output (input) of each wavelength changer in the first (third) stage is connected to at most one trunk, it does follow that each transmitter is connected to at most F trees and each receiver is connected to at most F trees. We now define a very important class of AONs.[1]

Definition 1 Light Tree AON (LT-AON)

A passive AON is a passive Light Tree AON iff there exists a 3-stage interconnection AON with the same connectivity matrix. A configurable AON is a configurable Light Tree AON iff the network is a passive Light Tree AON in each switching state.

[1]The definition is more general than that in [A+93].

Usually, this definition is little help in determining if an AON is a Light Tree. The following equivalent definitions generally remedy the situation.

Definition 2 Light Tree AON (LT-AON)

A passive Light Tree AON has a connectivity matrix H such that if three entries of H are of the form

	m	y
n	$\{f:f',...\}$	$\{f:g',...\}$
x	$H(x,m)$	$\{g:g',...\}$

then the fourth entry must contain $g:f'$,

	m	y
n	$\{f:f',...\}$	$\{f:g',...\}$
x	$\{g:f',...\}$	$\{g:g',...\}$

That is, if $n \neq x$ and $m \neq y$, $f:f' \in H(n,m)$, $g:g' \in H(x,y)$, and $f:g' \in H(n,y)$ then $g:f' \in H(x,m)$.

Note that there is nothing special about $H(x,m)$ being in the left hand corner and by permuting n, x and/or m, y, we can make the undetermined square in any of 4 possible positions. Also, the rows and columns need not be adjacent to each other.

When there is no wavelength changing, the connectivity matrix is a superposition of rectangles.

Definition 3 Light Tree AONs w/o Wavelength Changing

A passive Light Tree AON without wavelength changing has a connectivity matrix H

such that if three entries of H are of the form

	m	y
n	$\{f, ...\}$	$\{f, ...\}$
x	$H(x, m)$	$\{f, ...\}$

then the fourth entry must contain f*,*

	m	y
n	$\{f, ...\}$	$\{f, ...\}$
x	$\{f, ...\}$	$\{f, ...\}$

That is, if $n \neq x$ *and* $m \neq y$, $f \in H(n, m) \bigcap H(x, y) \bigcap H(n, y)$ *then* $f \in H(x, m)$.

Hypergraphs are another way to view the connectivity of a LT-AON.[2] To see this, form a bi-partite hypergraph with M_t input nodes and M_r output nodes. Then for each light tree in the network, form an edge containing the transmitters and receivers connected to the tree. Conversely, any AON whose connectivity is represented by a Hypergraph can be represented by a LT-AON.

1.5 Latin Routers

A special case of a passive wavelength routing device or network has a connection matrix which is a latin square. A latin square is an $N \times N$ matrix where each element (i, j) is one of N *symbols* such that no symbol appears in a row or column more than once. Two examples of 4×4 latin squares are shown in tables 1.1a and 1.1b. Notice that in table 1.1b the symbols are the four sets $\{0, 1\}, \{2, 4\}, \{5, 6\}, \{3, 7\}$.

We call a device or network whose connection matrix is a latin square a *Latin Router* (LR). We only consider Latin Routers without wavelength changing and where

[2] Hypergraphs are a generalization of undirected graphs where the arcs of the graphs are arbitrary sets of nodes. If each arc contains only 2 nodes, the hypergraph is a graph [Brg89].

each input is connected to each output on k wavelengths, some k. k is called the *coarseness* of the device. If $k = 1$, we say the device is *fine* and in this case the number of inputs and outputs of the device is equal to the number of wavelengths, i.e. $N = F$. In general, $F = Nk$.

It should be easy to see that the Latin Router, with any number of users connected to the inputs and outputs, is a LT-AON. It should also be easy to see that all F wavelengths can be simultaneously applied to each of the N inputs without any output contention accommodating a total of F^2/k simultaneous connections [Hil88, HCD89, WK89, Goo89, Bra90].

A well known physical implementation of a Latin Router with $k = 1$ is the WDM cross-connect [Bra90] shown in Fig. 1-4. The cross-connect is a 2 stage network where the first stage consists of F frequency demultiplexers of size $1 \times F$ each and the second stage consists of F multiplexers of size $F \times 1$ each. The demultiplexers separate the wavelengths on each input fiber onto a unique output fiber. The interconnection between the stages consists of F^2 fibers connecting each of the F demultiplexers to each of the F multiplexers in the final stage. By properly choosing the output and input ports of the demultiplexers and multiplexers any latin square can be implemented. Although the cross-connect uses an excessive $2F$ devices and F^2 interconnections, there are far more practical designs; for instance a 1000×1000 Latin Router can be implemented with 66 devices, each of size 33×33, and 1000 fiber interconnections [BH93a]. Each device used in this design is itself a Latin Router with

0	1	2	3
3	0	1	2
2	3	0	1
1	2	3	0

Table 1.1a
$k = 1$

{ 0,1 }	{ 2, 4 }	{ 5,6 }	{ 3,7 }
{ 3,7 }	{ 5,6 }	{ 2, 4 }	{ 0,1 }
{ 2, 4 }	{ 0,1 }	{ 3,7 }	{ 5,6 }
{ 5,6 }	{ 3,7 }	{ 0,1 }	{ 2, 4 }

Table 1.1b
$k = 2$

Table 1.1: Two Examples of 4×4 Latin Routers/Squares

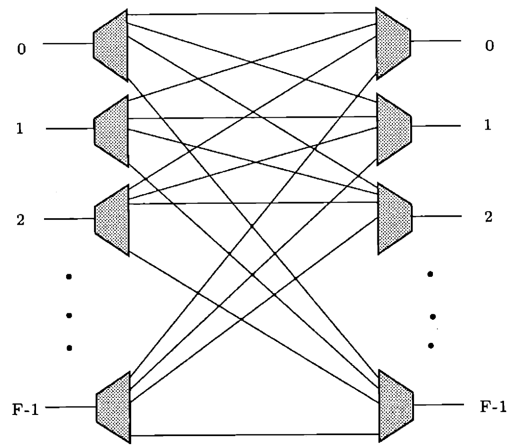


Figure 1-4: WDM Cross-Connect

a special periodic structure. These devices, similar to a generalized Mach-Zehnder interferometer, have been integrated onto silicon [Dra91].

1.6 Non-Light Tree AONs (NLT-AONs)

Not all networks are LT-AONs. An example of a non-Light Tree AON (NLT-AON) is shown in Fig. 1-5. The Σ-nodes represent star couplers which broadcast all input signals to all outputs. The connection matrix is

$$H = \begin{bmatrix} \{1,2,...F\} & \{1,2,...F\} & \emptyset \\ \emptyset & \{1,2,...F\} & \emptyset \\ \emptyset & \{1,2,...F\} & \{1,2,...F\} \end{bmatrix}$$

where $\emptyset$ is the empty set.

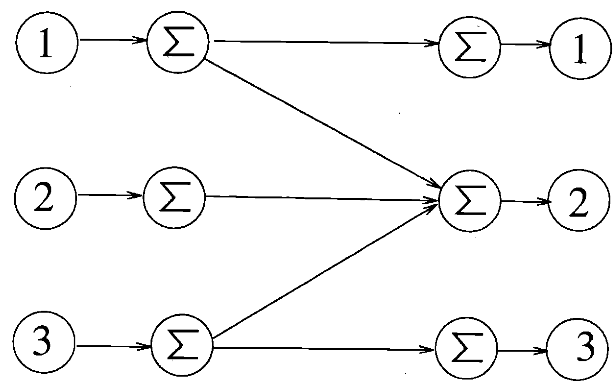


Figure 1-5: Non-Light Tree AON

Chapter 2

Problem Statement

Recall that in a broadcast AON each transmitter is connected to each receiver on all wavelengths and therefore as many wavelengths as there are simultaneous signals are required. In networks where the number of active users far exceeds the number of available wavelengths[1], it will be necessary to simultaneously assign many transmitters the same wavelength. In addition, it will be necessary for a signal originating at a source to reach many destinations. Since two sessions using the same wavelength cannot travel over the same fiber simultaneously, certain collisions within the network need to be prevented. In particular, we must insure that sessions do not collide at intended receivers. That is, if receiver m is listening to wavelength λ at time t, we must insure that only one signal on λ arrives at receiver m at time t. If two or more arrive, we say there is *contention.*

Contention is avoided by isolating signals of the same wavelength. This isolation can be done spatially and temporally. Spatial isolation is achieved by λ-routing. In this thesis, we show that there is a limit to the possible amount of isolation, or equivalently a limit on the wavelength re-use. This limit depends on the number of wavelengths, the number of devices, the functionality of the devices, and the requirements of the users. We will derive lower bounds on the required number of

[1]Bandwidth is a scarce commodity!

wavelengths for λ-routing networks under various constraints. We will also present constructions and existence proofs which show that some of these bounds can be tight.

We will outline the most important results in this chapter. Before that, we need a language to describe the connectivity requirements of the users. To that end, define a *session* (n, m) to be an ordered pairing of a transmitter, n, to a receiver, m. The session data rate is defined to be R_s b/s.

A *traffic* ϕ is a set of sessions. If $(n, m) \in \phi$, we say that (n, m),n, and m are *active* in ϕ. We assume that each transmitter and each receiver are active in at most one session in any traffic, that is we do not consider any multi-point connections. If two sessions are active in the same traffic ϕ they are said to be *concurrent* in ϕ. We say that a network *supports* ϕ if all the sessions in ϕ can be connected without contention.

The *traffic set* $\mathcal{T}$ is a set of traffics. A network *supports* $\mathcal{T}$ if the network supports all traffics in $\mathcal{T}$.

The *Permutation Traffic Set* for M users is defined to be the set of all traffics of size M and has $M!$ traffics. We say a network does *permutation routing* if it supports the Permutation Traffic Set. We also say it is a *permutation network* and a *connector.*

We will be interested in another important traffic set. Define the *ρ-Permutation Traffic Set* as the set of all traffics of size ρM. ρ is called the *utilization* and ρM is called the *load.* If a network supports the ρ-Permutation Traffic Set, then we say it does ρ-permutation routing. Also we say the network is a *ρ-permutation network* and a *partial-connector.* If $\rho = 1$, ρ-permutation routing is permutation routing and a ρ-permutation network is a permutation network.

The general question we are interested in is how many wavelengths are needed given a demand on the network, as measured by a traffic set, and given an acceptable level of performance, as measured by the blocking probability. We will mainly be interested in passive networks where the demand on the network is the ρ-Permutation

Traffic Set, but we will consider configurable networks and other demands as well.

Throughout this thesis we will use various combinatorial functions; these functions and some of their properties are described in Appendix A. We will also use the binomial and hypergeometric probability distributions; these distributions and their properties are described in Appendix B.

2.1 Non-Blocking WDM Networks

Throughout this thesis, M is the number of users, F is the number of wavelengths, R b/s is the maximum throughput of each wavelength, and R_s is the required bit rate of each session. When each session requires a full wavelength of bandwidth, i.e. $R_s = R$ b/s, we call the network a WDM network. When each session requires less than a full wavelength of bandwidth, the network is called a WDM/TDM network. WDM networks are the subject of Parts I and II. WDM/TDM networks are treated in Part III.

A non-blocking network is one which supports the traffic set which describes the demand on the network. For instance, a non-blocking network which supports permutation routing is a connector and a non-blocking network which supports ρ-permutation routing is a partial connector.

In Part I we consider building non-blocking WDM networks. We start in Chapter 3 with a special case. Specifically, we determine the minimum number of wavelengths required for a passive AON when pairs of users are connected by at most one OD channel, as in the WDM cross-connect. Such a network is called a *simple network.* We will see that there are essentially no good simple networks for ρ-permutation routing.

We therefore consider the general case where users are connected on more than one OD channel. Chapter 4 presents a lower bound on the number of wavelengths for any non-blocking λ-routing AON based on the traffic set and the number of switching states. The influence of wavelength changers on this bound is also discussed. We will see that wavelength changing plays a very insignificant role in this bound and that the bound can be tight without wavelength changing.

Chapter 5 is devoted to connectors. Four connectors are considered: rearrangeably non-blocking and wide-sense non-blocking connectors with and without wavelength changing. A rearrangeable network is allowed to reassign wavelengths and reconfigure devices in response to a session request or termination. A wide-sense network cannot disturb active sessions and therefore requires at least as many wavelengths as a rearrangeable network. As a consequence of the lower bound presented in Chapter 4, at least $\sqrt{M/e}$ wavelengths are required to build a passive connector. We will see that $\Theta(\sqrt{M \log M})$ wavelengths are sufficient for a wide-sense non-blocking network without wavelength changing.[2] This is currently the best asymptotic bound for all four types of passive connectors, but the constants are slightly smaller for rearrangeably non-blocking versus wide-sense non-blocking networks. There is currently no difference in constants with or without wavelength changing. We also consider configurable networks. Again there is no difference between the best known bounds with and without wavelength changing but unlike the passive case, the bounds presented here have a small asymptotic difference between rearrangeably and wide-sense non-blocking connectors. This difference has recently been closed [ABCR+93]. Unfortunately, we only prove the existence of wavelength routing networks with $\Theta\left(\sqrt{M \log M}\right)$ wavelengths; currently no constructions with this efficiency are known. This limitation is taken care of in Part II.

[2]We use $h(M) = \Theta(g(M))$ to mean that there exist positive bounded constants c_1, c_2, M_o such that for all $M \geq M_o$, $c_1 g(M) \leq h(M) \leq c_2 g(M)$. $h(M) = O(g(M))$ means $h(M) \leq \Theta(g(M))$ and $h(M) = \Omega(g(M))$ means $h(M) \geq \Theta(g(M))$.

2.2 Blocking WDM Networks

In Part II, we consider the problem of honoring as many requests as possible from a random list of requests. Let $s_1, ...s_{\rho M}$ be an ordered list of session requests where request i is made before request $i+1$ and where all lists without multi-point connections are equally likely.

We consider two types of networks: *non-sequential* and *sequential.* In a non-sequential network, analogous to a rearrangeable network, the network waits until the last request before deciding which requests to *honor* and which to *block.* In a sequential network, analogous to wide-sense non-blocking networks, the decision to honor or block the i^{th} request is made before the $(i+1)^{th}$ request. Also a sequential network is not allowed to reassign wavelengths nor reconfigure devices in order to honor the i^{th} request.

We show that for small blocking probability P_b, a non-sequential blocking network requires just about as many wavelengths as a rearrangeably non-blocking network. Then we construct sequential networks with very small blocking probabilities using only $c\sqrt{\rho M}$ wavelengths where c is a constant that depends weakly on the blocking probability and is between 6.5 and 9.2 for P_b between 10^{-3} and 10^{-6}. The constructed networks, which do not require wavelength changing, consist of broadcast local area networks (LANs) connected to a Latin Router. An optimization is done on the number of wavelengths connecting pairs of users. Therefore, unlike non-blocking connectors, we have essentially been able to meet the theoretical lower bound in terms of asymptotic growth of the required number of wavelengths. However the constants may be too large for applications where the number of wavelengths is limited.

Part II also treats the case of *almost-all connectors.* In an almost-all connector, the network must honor all of M randomly requested sessions with high probability. We explicitly construct almost-all connectors using $\Theta(\sqrt{M \log M})$ wavelengths without wavelength changing, meeting the best non-blocking existence proofs.

2.3 WDM/TDM Networks

Recall that R b/s is the maximum throughput of each wavelength and that R_s is the required bit rate of each session. When each session requires less than a full wavelength of bandwidth we call the network a WDM/TDM network. We assume a slotted system with T periodic time slots where each session is assumed to require one time slot of a wavelength, i.e. $R_s = R/T$ b/s.

Recall that at least $F \geq \Theta(\sqrt{\rho M})$ wavelengths are required for passive WDM networks. Inverting this formula, the maximum load ρM of such a network is $\Theta(F^2)$. It will be more convenient to express results on WDM/TDM networks in terms of the load.

First off, since a network with F wavelengths and T time slots cannot perform better than a network with FT wavelengths, the maximum load of a passive WDM/TDM network is $\Theta(F^2T^2)$. We show that $\Theta(F^2T)$ is achievable and that this is the best any passive LT-AON can do [Gal92b]. We also presents bounds for configurable LT-AONs.

Then, in Chapter 8 we show that if the Light Tree restriction is relaxed, $F^2T^{4/3}$ active sessions are possible as long as $F \geq T^{1/3}$. The difference between $F^2T^{4/3}$ and F^2T^2 is an open question.

Part I

Non-Blocking WDM Networks

Chapter 3

Simple Networks

A *simple* AON is defined to be a passive AON where each user pair is connected by at most one OD channel, i.e. $|H(n,m)| \leq 1$. Simple networks have the practical advantage that the OD channel used by a session is not a function of the other active sessions.

An important example of a simple network is the $F \times F$ WDM cross-connect, Fig. 1-4. Consider using a WDM cross-connect to build a partial connector with $\rho M \geq 2$ active sessions. Clearly $F = M$ wavelengths suffice by hooking one user to each input and each output. If $F < M$, at least one input port i and at least one output port j have more than one user attached. Since the cross-connect can support at most one session between any input and output port, the network cannot support ρ-permutation routing if $\rho M \geq 2$. Therefore a broadcast network, which requires ρM wavelengths, is more wavelength efficient than the cross-connect for any $\rho < 1$, and equally efficient for $\rho = 1$. Note that this conclusion may not hold if a session does not require a full wavelength of bandwidth.

Can we do better without relaxing the restriction $|H(n,m)| = 1$? The answer is yes, but only by a factor of two. Specifically, an $M \times M$ simple connector can be built with $\left\lceil \frac{M}{2} \right\rceil + 2$ wavelengths and this is the best we can do. The networks do not require wavelength changing and are LT-AONs. A 7 wavelength, 10 user

6	2	3	4	5	7	**1**	**1**	**1**	**1**
1	6	3	4	5	2	7	2	2	2
1	2	6	4	5	3	3	7	3	3
1	2	3	6	5	4	4	4	7	4
1	2	3	4	6	5	5	5	5	7
7	**1**	**1**	**1**	**1**	6	2	3	4	5
2	7	2	2	2	**1**	6	3	4	5
3	3	7	3	3	**1**	2	6	3	4
4	4	4	7	4	**1**	2	3	6	5
5	5	5	5	7	**1**	2	3	4	6

Table 3.1: 10 user AON with 7 wavelengths.

11	2	3	4	5	6	7	8	9	10	12	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
1	11	3	4	5	6	7	8	9	10	2	12	2	2	2	2	2	2	2	2
1	2	11	4	5	6	7	8	9	10	3	3	12	3	3	3	3	3	3	3
1	2	3	11	5	6	7	8	9	10	4	4	4	12	4	4	4	4	4	4
1	2	3	4	11	6	7	8	9	10	5	5	5	5	12	5	5	5	5	5
1	2	3	4	5	11	7	8	9	10	6	6	6	6	6	12	6	6	6	6
1	2	3	4	5	6	11	8	9	10	7	7	7	7	7	7	12	7	7	7
1	2	3	4	5	6	7	11	9	10	8	8	8	8	8	8	8	12	8	8
1	2	3	4	5	6	7	8	11	10	9	9	9	9	9	9	9	9	12	9
1	2	3	4	5	6	7	8	9	11	10	10	10	10	10	10	10	10	10	12
12	2	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	11	2	3	4	5	6	7	8	9	10
2	12	2	2	2	2	2	2	2	2	**1**	11	3	4	5	6	7	8	9	10
3	3	12	3	3	3	3	3	3	3	**1**	2	11	4	5	6	7	8	9	10
4	4	4	12	4	4	4	4	4	4	**1**	2	3	11	5	6	7	8	9	10
5	5	5	5	12	5	5	5	5	5	**1**	2	3	4	11	6	7	8	9	10
6	6	6	6	6	12	6	6	6	6	**1**	2	3	4	5	11	7	8	9	10
7	7	7	7	7	7	12	7	7	7	**1**	2	3	4	5	6	11	8	9	10
8	8	8	8	8	8	8	12	8	8	**1**	2	3	4	5	6	7	11	9	10
9	9	9	9	9	9	9	9	12	9	**1**	2	3	4	5	6	7	8	11	10
10	10	10	10	10	10	10	10	10	12	**1**	2	3	4	5	6	7	8	9	11

Table 3.2: A Non-blocking 20×20 AON with 12 wavelengths

example is shown in Table 3.1. Wavelength 1 is printed in boldface to help the reader discern the pattern. Another example is shown in table 3.2. Both these networks are non-blocking since if $H(n,m) = \{f\}$ and $H(x,y) = \{f\}$, then $f \notin H(n,y)$ and $f \notin H(x,m)$. Therefore when (n,m) and (x,y) are both active, neither signal is heard at the other's receiver. Note that $n \neq x$ and $m \neq y$ since multi-point connections are not allowed in permutation routing.

Surprisingly, this is the best we can do even if $\rho << 1$. Specifically, let F_s be the minimum number of wavelengths to do ρ-permutation routing with a simple AON.

Then,

$$\left\lceil \frac{M+1}{2} \right\rceil \leq F_s \leq \left\lceil \frac{M}{2} \right\rceil + 2$$

for all $\rho M \geq 2$. Notice that the required number of wavelengths is independent of ρ, for $\rho M \geq 2$. If $\rho M = 1$, a broadcast network supports the traffic set and exactly one wavelength is required. Also if $\rho \leq .5$, eqn. (3.1) shows that a broadcast network is more wavelength efficient than a simple wavelength routing network.

Theorem 1 Bounds for Simple AONs

Let F_s be the minimum number of wavelengths to do ρ-permutation routing with a simple AON. Then for $\rho M \geq 2$,

$$\left\lceil \frac{M+1}{2} \right\rceil \leq F_s \leq \left\lceil \frac{M}{2} \right\rceil + 2 \tag{3.1}$$

Proof: Let H be any simple AON that supports ρ-permutation routing, $\rho M \geq 2$. Then clearly H must support $\frac{2}{M}$-permutation routing. That is, H must support the following traffic set,

$$\mathcal{T}_2 \stackrel{\text{def}}{=} \{(n,m),(x,y) \mid n \neq x, m \neq y\} \tag{3.2}$$

We will show that at least $\left\lceil \frac{M+1}{2} \right\rceil$ wavelengths are required to avoid contention.

Let $f(n,m) : f'(n,m)$ be the OD channel in $H(n,m)$. We refer to $f(n,m)$ as the o-color and $f'(n,m)$ as the d-color of (n,m). Now define $d_r(n,m)$ to be the number of entries in row n with o-color $f(n,m)$. Note that $d_r(n,m) \geq 1$ since $H(n,m)$ contains o-color $f(n,m)$. Similarly define $d_c(n,m) \geq 1$ to be the number of entries in column m with d-color $f'(n,m)$. Note also that $d_r(n,m) \leq M$ and $d_c(n,m) \leq M$ since a color can be used at most M times in a row or column.

Notice that if $d_c(n,m) \geq 2$ and $d_r(n,m) \geq 2$, then there exists an m'

such that $f' = f'(n, m') = f'(n, m)$ and similarly there exists an n' such that $f = f(n, m) = f(n', m)$. It therefore follows that there is contention at receiver m' in traffic $\{(n', m), (n, m')\}$ as the reader can easily verify. So,

$$\frac{1}{d_r(n,m)} + \frac{1}{d_c(n,m)} \geq 1 + \frac{1}{M}$$

for all (n, m). Now summing over all (n, m) gives

$$\sum_{(n,m)} \frac{1}{d_r(n,m)} + \frac{1}{d_c(n,m)} \geq M^2 \left(1 + \frac{1}{M}\right). \tag{3.3}$$

To finish the proof, we need only count the left hand side of eqn. (3.3) in a different way. Notice that for all n,

$$\sum_{m=1}^{M} \frac{1}{d_r(n,m)} = \sum_{f=1}^{F} \left(\sum_{\substack{m, \\ H(n,m)=f}} \frac{1}{d_r(n,m)} \right) \leq F$$

since the inner sum is 1 if f is used in row n and 0 otherwise. Similarly, the number of d-colors used in column m is

$$\sum_{n=1}^{M} \frac{1}{d_c(n,m)} \leq F$$

Thus, the left hand side of eqn. (3.3) is at most $2FM$. The fact that F must be integer completes the proof of the lower bound.

The following construction, which uses $\left\lceil \frac{M}{2} + 2 \right\rceil$ wavelengths without wavelength changing is due to [E+86]. The construction supports ρ-permutation routing for all $0 \leq \rho \leq 1$. Specifically for a $2M \times 2M$ square, call the rows $x_1, x_2, ...x_M, y_1, y_2, ...y_M$ and the columns $u_1, u_2, ...u_M, v_1, v_2, ...v_M$. The coloring is as follows,

$$f(x_i, u_j) = f'(x_i, u_j) \quad = \quad j \quad \text{for } i = 1, 2, ...M, j \neq i$$

$$
\begin{aligned}
f(x_i, v_j) = f'(x_i, v_j) &= i \quad \text{for } i = 1, 2, ...M, j \neq i \\
f(y_i, u_j) = f'(y_i, u_j) &= i \quad \text{for } i = 1, 2, ...M, j \neq i \\
f(y_i, v_j) = f'(y_i, v_j) &= j \quad \text{for } i = 1, 2, ...M, j \neq i \\
f(x_i, u_i) = f'(x_i, u_i) &= M + 1 \quad \text{for } i = 1, 2, ...M \\
f(x_i, v_i) = f'(x_i, v_i) &= M + 2 \quad \text{for } i = 1, 2, ...M \\
f(y_i, u_i) = f'(y_i, u_i) &= M + 2 \quad \text{for } i = 1, 2, ...M \\
f(y_i, v_i) = f'(y_i, v_i) &= M + 1 \quad \text{for } i = 1, 2, ...M
\end{aligned}
$$

□

3.1 Conclusions and Extensions

In the remaining chapters, we will relax the restriction $|H(n,m)| = 1$. Before proceeding, we note 4 extensions of Theorem 1.

First, for simple AONs without wavelength changing, the minimum number of wavelengths is exactly $\lceil \frac{M}{2} + 2 \rceil$. The slight improvement to the lower bound is shown by relating the problem of supporting $\mathcal{T}_2$ to a previously solved bi-partite graph edge coloring problem [E+86]. The transmitters are one set of nodes, the receivers are the other set, and the wavelengths are the colors of the edges. A feasible H corresponds to a colored graph without monochromatic paths of length 3 edges. The added complexity is not worth the improved results for our purposes so we omit the details.

Second, the problem can be extended to the case where the number of transmitters M_t is not equal to the number of receivers M_r. Then by obvious modifications of the arguments,

$$F_s \geq \frac{M_t M_r}{M_t + M_r}\left(1 + \frac{1}{\max(M_t, M_r)}\right) \tag{3.4}$$

The case of no wavelength changing has been solved by [EU91]. For $M_t \geq M_r$, $\frac{M_t}{M_r}$

an integer, and $M_t \geq 6$,

$$F_s = \min\{M_r, \lceil M_t \frac{M_r + 2}{M_t + M_r} \rceil + 1\} \tag{3.5}$$

For $M_t < M_r$, reverse the rolls of M_t and M_r in the above expression. When $M_t = M_r = M$, this expression reduces to $F_s = \lceil \frac{M}{2} \rceil + 2$ as before. Notice that for large M_t , $F_s = M_r$ and the optimal network assigns one fixed wavelength to each receiver. Similarly, for M_r large, the optimal network requires one wavelength per transmitter.

Third, the results of this section hold for passive AONs that are not simple but instead assign a fixed OD channel to each session. In this case $f(n, m) : f'(n, m)$ represents the OD channel used by session (n, m). As is the case for simple networks, the session (n, m) is always routed through the same path in the network. This type of routing is called *oblivious* routing [Lei92].

Fourth, notice that the constructed networks are strict sense non-blocking.[1] We will see later that at least $\lceil M/2 \rceil$ wavelengths are needed for a strict sense non-blocking LT-connector [Chapter 5, p. 76]. Since the construction provided in this chapter is a LT-AON, it is optimal in another way.

[1] An AON is strict sense non-blocking (S.S.NB.) if given any set of active sessions without contention and a new session request between an inactive transmitter and an inactive receiver, there always exists an OD channel to feasibly assign the session.

Chapter 4

On the Number of Wavelengths and Switches Needed in AONs

4.1 Introduction

In this chapter, we only consider networks with M users where each user has one transmitter and one receiver, i.e. $M = M_t = M_r$. The techniques and results can be easily extended to the case where $M_t \neq M_r$.

We relax the restriction used in Chapter 3 and consider λ-routing networks with arbitrary topology, wavelength changers, and switches. A lower bound on the number of wavelengths is presented in section 4.2. Section 4.3 discusses the influence of wavelength changing on the bound. In section 4.4, we show that the bound can be very tight and that near optimal wavelength re-use is possible even without wavelength conversion. Partial connectors are discussed in section 4.5 followed by another example in section 4.6.

Conclusions on the scalability of non-blocking wavelength routing AONs are delayed until Chapter 6.

4.2 Lower Bound

Recall that $\mathcal{T}$ are the traffics the network must support. Each traffic $\phi \in \mathcal{T}$ is called an *allowable traffic*. Two sessions in the same traffic are said to be concurrent. Also, when considering a fixed traffic ϕ, we say that n, m, and (n, m) are all *active* in ϕ if $(n, m) \in \phi$. Also recall that two concurrent sessions *collide* if they arrive at the same receiver on the same wavelength. Such a collision may not be fatal; it is fatal only if the collision occurs at one of the intended receivers. If two sessions have a fatal collision, we say they *contend*.

Let $F(\mathcal{T}, S)$ be the minimum number of wavelengths needed to support all the traffics in $\mathcal{T}$ without contention for any λ-routing network with $|\Psi| = S$ states. Also let $d = \max_{\mathcal{T}} |\phi|$ be the maximum number of concurrent sessions. Our main result proved in this section is that

$$F(\mathcal{T}, S) \;\geq\; \left(\frac{|\mathcal{T}|}{S}\right)^{\frac{1}{2d}} \binom{M}{d}^{-1/d} - 1 \tag{4.1}$$

For example, the Permutation Traffic Set has $d = M$ and $|\mathcal{T}| = M!$ traffics. Therefore, applying the bound and using $M! \geq M^M e^{-M}$, at least $\sqrt{M/e} - 1$ wavelengths are needed to arbitrarily interconnect M users in any passive ($S = 1$) λ-routing network.

Define a *tuning state*, $\mathbf{v}$, as the $2 \times M$ matrix

$$\mathbf{v} \;=\; \begin{bmatrix} f_{o,1} & f_{o,2} & \cdots & f_{o,M} \\ f_{d,1} & f_{d,2} & \cdots & f_{d,M} \end{bmatrix}$$

where $f_{o,n}$ ($f_{d,m}$) is the wavelength assigned to transmitter n (receiver m) in this tuning state. If $f_{o,n} = 0$ ($f_{d,m} = 0$), we say that transmitter n (receiver m) is *off*. Let v_{in} and v_{out} be the first and second row of $\mathbf{v}$, respectively. v_{in} and v_{out} are known as the transmitting and receiving tuning states, respectively.

A tuning state $\mathbf{v}$, together with a switching state ψ, defines a *network state.* If $\mathcal{H}$ is a λ-routing network, we say that a network state is *feasible* for traffic ϕ if all active sessions (n, m) in ϕ can be assigned OD channel $f_{o,n} : f_{d,m}$ without contention when the network is in switching state ψ. Network state feasibility is formally defined below.

Definition 4 Network State Feasibility

Let $\mathcal{H} = \{H_\psi | \psi \in \Psi\}$ be a λ-routing network, $\mathbf{v}$ be a tuning state, and ψ a switching state. Then the network state defined by $\mathbf{v}$ and ψ is feasible for a traffic ϕ if

(C0) Inactive transceivers are off.

(C1) Active sessions are connected, i.e. $\forall (n, m) \in \phi \;,\; f_{o,n} : f_{d,m} \in H_\psi(n, m)$.

(C2) Concurrent sessions do not contend, i.e. if $\{(n, m), (x, y)\} \subseteq \phi$ *then* $f_{o,n} : f_{d,y} \notin H_\psi(n, y)$ **and** $f_{o,x} : f_{d,m} \notin H_\psi(x, m)$.

Condition (2) says that for all pairs of concurrent sessions (n, m) and (x, y), there is not a path from n to y on $f_{o,n} : f_{d,y}$ and similarly there is not a path from x to m on $f_{o,x} : f_{d,m}$.

A network $\mathcal{H}$ *supports* traffic set $\mathcal{T}$ without contention if for each allowable traffic, there is a feasible network state. We will show below in lemma 3, that a network state can be feasible for at most one traffic. Therefore, the number of network states must be at least the number of traffics. Bounding the number of network states will then prove eqn. (4.1). Following the proof of eqn. (4.1) and a few variations, we will discuss some further interpretations of this result with an emphasis on the role wavelength changers play in defining feasible network states. Afterwards, we will presents some examples.

We now formally prove our bound.

Theorem 2 Lower Bound

Let $F(\mathcal{T}, S)$ be the minimum number of wavelengths needed to support $\mathcal{T}$ without contention for any λ-routing network with S states. Then,

1)

$$F(\mathcal{T}, S) \geq \left(\frac{|\mathcal{T}|}{S}\right)^{\frac{1}{2M}} - 1 \tag{4.2}$$

where $|\mathcal{T}|$ is the number of traffics in $\mathcal{T}$.

2) Also, let $d = \max_{\mathcal{T}} |\phi|$ be the maximum number of active sessions, then

$$F(\mathcal{T}, S) \geq \left(\frac{|\mathcal{T}|}{S}\right)^{\frac{1}{2d}} \binom{M}{d}^{-1/d} - 1 \tag{4.3}$$

3) Also for all integer $1 \leq x \leq M$, let $\mathcal{T}_x = \{\phi \in \mathcal{T} \mid |\phi| = x\}$ be the traffics with x active sessions. Then,

$$F(\mathcal{T}, S) \geq \left(\frac{|\mathcal{T}_x|}{S}\right)^{\frac{1}{2x}} \binom{M}{x}^{-1/x} \tag{4.4}$$

4) which implies that for any λ-routing network with F wavelengths,

$$S \geq |\mathcal{T}_M| F^{-2M} \tag{4.5}$$

Proof: Part I: $S = 1$

First consider an arbitrary passive λ-routing network specified by a connection matrix H. Since H is assumed to be able to support $\mathcal{T}$, it must be true that there is a feasible tuning state for each $\phi \in \mathcal{T}$. We will show in the lemma below that for a passive network, conditions (C0), (C1), and (C2) imply that a tuning state can be feasible for at most one traffic. Therefore, the number of feasible tuning states must be greater than the number of allowable traffics. This proves eqn. (4.2) when $S = 1$ since the number of feasible tuning states is at most $(F+1)^{2M}$.

Equations 4.3 and 4.4 follow by using the added information to better bound the number of feasible tuning states. To see eqn. (4.3) note that at least $M-d$ transmitters and at least $M-d$ receivers are off. There are $\binom{M}{M-d}$ ways to pick the $M-d$ inactive transmitters and the same number of ways to pick the inactive receivers. The remaining d transmitters are either assigned a wavelength from 1 to F or off. Similarly for the remaining d receivers. Therefore, the number of feasible tuning states is no more than $(F+1)^{2d}\binom{M}{M-d}^2 = (F+1)^{2d}\binom{M}{d}^2$.

Eqn. 4.4 uses the same reasoning as above, except now there are exactly x active transmitters and x active receivers in each traffic of $\mathcal{T}_x$. There are $\binom{M}{x}^2$ ways to pick the active transmitters and receivers and each one is assigned a wavelength between 1 and F.

Part II: $S \geq 1$

Let $\mathcal{T}(\psi)$ be the traffics supported by the network in switching state ψ. Then since $\mathcal{T} = \bigcup_{\psi \in \Psi} \mathcal{T}(\psi)$, at least one switching state must support at least $|\mathcal{T}|/S$ traffics. Now apply Part I in this state. This proves eqns. (4.2),(4.3), (4.4) for arbitrary S. Solving eqn. (4.4) for S gives eqn. (4.5) when $x = M$.

□

Lemma 3 *In a passive network, a tuning state can be feasible for at most one traffic.*

Proof: Consider an arbitrary passive λ-routing network with connection matrix H. Let $V(\phi)$ be the set of feasible tuning vectors for traffic ϕ. That is $V(\phi)$ are those vectors $\mathbf{v}$ that satisfy conditions (C0), (C1) and (C2).

Suppose $\mathbf{v}$ were feasible for two traffics. Specifically, suppose $\mathbf{v} \in V(\phi)$ and $\mathbf{v} \in V(\phi')$. Then (C0) implies that ϕ and ϕ' have the same active transceivers. Since $\phi \neq \phi'$, there must be at least one receiver m matched to n in ϕ, i.e.

$(n,m) \in \phi$, and not matched to n in ϕ', i.e. $(n,m) \notin \phi'$. Since m is active in ϕ', there must be an $x \neq n$ such that $(x,m) \in \phi'$. From condition (C1) applied in ϕ, there must be a path from n to m on $f_{o,n}\!:\!f_{d,m}$, i.e. $f_{o,n}\!:\!f_{d,m} \in H(n,m)$. Also from condition (C1) applied in ϕ' there must be a path from x to m on $f_{o,x}\!:\!f_{d,m}$, i.e. $f_{o,x}\!:\!f_{d,m} \in H(x,m)$. Now since x is active in ϕ, there must be a y such that $(x,y) \in \phi$. Therefore condition (2) is violated since $\{(n,m),(x,y)\} \subseteq \phi$ and there is a path from x to m on OD channel $f_{o,x}\!:\!f_{d,m}$. This is a contradiction since we assume $\mathbf{v}$ was feasible for ϕ. □

4.3 Networks without Wavelength Changing

Consider the traffics where everybody is on. Let # network states = # switching states × # tuning states. According to Theorem 2, the total number of network states must be large enough to assign a unique network state to each traffic in the traffic set. That is,

$$\# \text{ switching states} \times \# \text{ tuning states} \geq \# \text{ traffic states}$$

is a necessary condition to avoid contention for any λ-routing network. Since there are no more than F^{2M} feasible tuning states,

$$\# \text{ switching states} \times F^{2M} \geq \# \text{ traffic states}$$

Now consider networks without wavelength changing. There is a temptation to apply the following erroneous reasoning. The number of traffic states is $|\mathcal{T}|$. There are M active sessions and each session is assigned one of the F wavelengths. Therefore, it must be true that $S \times F^M \geq |\mathcal{T}|$ for networks without wavelength changing. This is not true. We will show by example in section 4.4 that this argument can vastly overestimate the required number of wavelengths for networks without wavelength changing. The above reasoning fails because it counts the possible number of feasible tuning states for any traffic ϕ, not the total number of feasible tuning states. That is, it is true that for any ϕ, $|V(\phi)| \leq F^M$ for networks without wavelength changing and $|V(\phi)| \leq F^{2M}$ for networks with wavelength changing. However, it is not true that the total number of feasible tuning states for networks without wavelength changing is at most F^M.

In a network without wavelength changing, $f : f' \in H_\psi(n, m)$ implies $f = f'$. Combining this with condition (C1), we see that in order for $\mathbf{v}$ to be feasible, $f_{o,n} = f_{d,m}$ for all active sessions (n, m). This implies that (C3) *the number of receivers assigned f must equal the number of transmitters assigned f*. We use this condition in the following theorem to improve Theorem 2, but the improvement is negligible. Our reason for including it is to spare the ambitious reader the trouble of reproducing the argument.

Theorem 4 Lower Bound for Networks without Wavelength Changing

Let $F'(\mathcal{T}_M, S)$ be the minimum number of wavelengths needed to support $\mathcal{T}_M$ without contention for any λ-routing network without wavelength changing. Let S be the number of switching states. Then,

$$F'(\mathcal{T}_M, S) \geq (1 + \epsilon) \left(\frac{|\mathcal{T}_M|}{S} \right)^{\frac{1}{2M}} \tag{4.6}$$

where ϵ goes to 0 faster than $\frac{\ln \sqrt{M}}{\sqrt{M}}$.

Proof: Let V be the number of tuning states where v_{out} is a permutation of v_{in}. We first count V and then use this to bound F.

To count V, let $V(\mathbf{k})$, $\mathbf{k} = [k_1, k_2, ..., k_F]$, be the number of transmitting tuning states with k_i transmitters using wavelength i. Obviously, $V(\mathbf{k})$ is also the number of receiving tuning states with k_i receivers using wavelength i. In addition, each receiving state counted in $V(\mathbf{k})$ is a permutation of any v_{in} with

k_i transmitters using wavelength i. Therefore

$$\begin{aligned} V &= \sum_{\sum_{i=1}^{F} k_i = M} V^2([k_1, k_2, ..., k_F]) \\ &= \sum_{\sum_{i=1}^{F} k_i = M} \left(\frac{M!}{k_1! k_2! ... k_F!} \right)^2 \end{aligned} \tag{4.7}$$

Dividing both sides by F^{2M}, we see that the right hand side is a multinomial distribution. Using the Central Limit Theorem, we approximate the multinomial by an $F-1$ dimensional Gaussian density with mean $\bar{k} = \frac{M}{F}$ and variance $\sigma^2 = M\left(\frac{F-1}{F^2}\right)$ in each dimension.

$$\frac{V}{F^{2M}} \approx \sum_{k_1, k_2, ... k_{F-1}} \left(\frac{1}{(2\pi\sigma^2)^{\frac{F-1}{2}}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{F-1} (k_i - \bar{k})^2 \right\} \right)^2$$

where now the sum is over all integers in each dimension. Approximating the sum as an integral and letting $\mathcal{R}$ be the reals,

$$\begin{aligned} \frac{V}{F^{2M}} &\approx \int_{\mathcal{R}^{F-1}} \left(\frac{1}{(2\pi\sigma^2)^{\frac{F-1}{2}}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{F-1} (k_i - \bar{k})^2 \right\} \right)^2 \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{F-1}{2}}} \int_{\mathcal{R}^{F-1}} \frac{1}{(2\pi\sigma^2)^{\frac{F-1}{2}}} \exp\left\{ -\frac{1}{\sigma^2} \sum_{i=1}^{F-1} (k_i - \bar{k})^2 \right\} \end{aligned}$$

The integral is the weight under a scaled gaussian with variance $\sigma^2/2$. Therefore, V is approximated by

$$V \approx \frac{F^{2M}}{(4\pi\sigma^2)^{\frac{F-1}{2}}} = F^{2M} \left[\frac{\left(\sqrt{F}\right)^F}{\left(\sqrt{4\pi M}\right)^{F-1}} \right]$$

Now since the number of traffics can be no more than the number of network states, $V \times S \geq |\mathcal{T}|$. Let F^* be the smallest F such that $V \times S \geq |\mathcal{T}|$. Using the approximation for V and relaxing the integer constraints,

$$\ln F^* = \frac{\ln|\mathcal{T}| - \ln S}{2M} + \frac{F^*}{4M}\ln\frac{4\pi M}{F^*} - \frac{1}{4M}\ln 4\pi M \tag{4.8}$$

Now, define α such that $F^* = \alpha F_o$, where $\ln F_o = \frac{\ln|\mathcal{T}|-\ln S}{2M}$. Since F_o is the lower bound on F for networks with channel changing, $\alpha \geq 1$. Thus eqn. (4.8) reduces to

$$\ln\alpha = \frac{\alpha F_o}{4M}\ln\frac{4\pi M}{\alpha F_o} - \frac{1}{4M}\ln 4\pi M \tag{4.9}$$

Note that the right hand side of eqn. (4.9) is an increasing function of both α and F_o. Since F^* is no more than M (since $V \geq F^M$ and $|\mathcal{T}| \leq M!$), α is no more than M/F_o. Using $\alpha \leq \frac{M}{F_o}$ to bound the right hand side of eqn. (4.9) shows that $\ln\alpha \leq .25\ln 4\pi$. Using $\alpha \leq (4\pi)^{\frac{1}{4}}$, we get

$$\ln\alpha \leq \frac{(4\pi)^{\frac{1}{4}}F_o}{4M}\ln\frac{(4\pi)^{3/4}M}{F_o} - \frac{1}{4M}\ln 4\pi M \tag{4.10}$$

Since the right hand side of eqn. (4.10) is increasing with F_o, we can replace F_o by its largest possible value. F_o is increasing with $\frac{|\mathcal{T}|}{S}$ and $\frac{|\mathcal{T}|}{S} \leq M!$. Therefore, $F_o \leq (M!)^{1/2M}$. Using Sterling's approximation gives us our final bound on $\ln\alpha$,

$$\ln\alpha \leq \frac{(4\pi)^{\frac{1}{4}}(2\pi M)^{\frac{1}{4M}}}{4\sqrt{Me}}\ln\left((4\pi)^{\frac{3}{4}}(2\pi M)^{-\frac{1}{4M}}\sqrt{Me}\right) - \frac{1}{4M}\ln 4\pi M$$

Since $M^{1/M} = 1 + O(\frac{\ln M}{M})$, $\ln\alpha = O(\frac{\ln\sqrt{M}}{\sqrt{M}})$. Let $\epsilon = \alpha - 1$, so $\epsilon = O(\frac{\ln\sqrt{M}}{\sqrt{M}})$. This completes the proof.

As an aside, V can be strictly lower bounded by taking only the middle term, i.e. $k_1 = k_2 = ... = k_M = M/F$, in eqn. (4.7). The conclusion is the same except that the convergence of ϵ to 0 is slightly slower. □

The preceding theorem does not say that wavelength changing cannot help, only that the absence of wavelength changing will not significantly change the lower bound derived earlier. We included it here to emphasize the fact that the number of tuning states grows something like F^{2M} with and without wavelength changing.

4.4 Near Optimal Wavelength Reuse

We present an example to make two important points. First, $S \times F^M \geq |\mathcal{T}|$ is not a necessary condition to avoid contention in networks without wavelength changing. Second, the bound derived in Theorem 2 can be extremely tight, even in the absence of wavelength changing.

In order to describe the traffic set, group the transmitters into $\sqrt{M}$ disjoint *T-Groups* of size $\sqrt{M}$ each. Number the T-Groups from 1 to $\sqrt{M}$ and let $I(n)$ be the T-Group of transmitter n. Similarly group the receivers into *R-Groups* and let $1 \leq J(m) \leq \sqrt{M}$ be the R-Group of receiver m. Define the *Santa Barbara Traffic Set*, $\mathcal{T}_{SB}$, to be the set of all traffics without multi-point connections with exactly one active session between each [T-Group,R-Group] pair. A typical traffic is shown in Fig. 4-1. Notice that like permutation routing, the Santa Barbara traffics have M active sessions in each traffic.

To implement the Santa Barbara Traffics, we use a $\sqrt{M} \times \sqrt{M}$ WDM cross-connect with each T-Group connected to an input and each R-Group connected to an output; see Fig. 4-2. The network uses a total of $\sqrt{M}$ wavelengths. Numbering the wavelengths by the integers from 1 to $F = \sqrt{M}$, the connection matrix is defined by

$$H(n, m) = (I(n) - J(m)) \bmod \sqrt{M} + 1 \qquad (4.11)$$

To see that this is about the best possible, we count the number of traffics and use eqn. 4.4 with $x = M$ and $S = 1$. $|\mathcal{T}_{SB}|$ can be directly counted from its definition, however we will employ an easier trick. Consider only the case where every user is on and notice that the network of Fig. 4-2 supports exactly the Santa Barbara Traffics. Therefore, if we count the number of traffics supported by the network, we simultaneously count $|\mathcal{T}_{SB}|$. To count these traffics, notice that each feasible tuning state for the network produces a unique traffic since if a transmitter (receiver) retunes, it is talking (listening) to a different R-Group (T-Group). Therefore to count $|\mathcal{T}_{SB}|$ it suffices to count the number of feasible tuning states for Fig. 4-2 when everybody is on. So $|\mathcal{T}_{SB}| = (\sqrt{M})!^{2\sqrt{M}}$ since there are $2\sqrt{M}$ groups, each group can be independently tuned, and there are $(\sqrt{M})!$ ways to tune a group. Now using eqn. 4.4 with $x = M$,

$$F^{2M} \geq (\sqrt{M})!^{2\sqrt{M}} \geq \left(\frac{\sqrt{M}}{e}\right)^{2M} \tag{4.12}$$

or $F \geq \frac{\sqrt{M}}{e}$.

If we had used the incorrect bound $F^M \geq |\mathcal{T}_{SB}|$ for networks without wavelength changing, we would have predicted that at least $\frac{M}{e^2}$ wavelengths were needed. Note also that H is a simple network [Chapter 3] that uses $\sqrt{M}$ wavelengths. Theorem 1 does not apply in this case since $\mathcal{T}_2 \not\subseteq \mathcal{T}_{SB}$. Therefore even though H can support $|\mathcal{T}_{SB}|$, and $|\mathcal{T}_{SB}|$ is quite large, it cannot support many traffics with only two sessions. For instance, H cannot support two concurrent sessions between a T-Group and an R-Group, even if these two sessions are the **only** active sessions in the entire network. This is of course not true when each session requires less than a full wavelength of bandwidth.

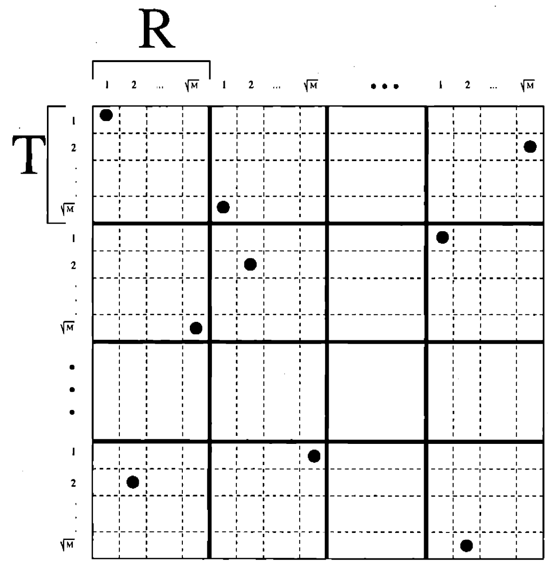


Figure 4-1: Typical Traffic of $\mathcal{T}_{SB}$

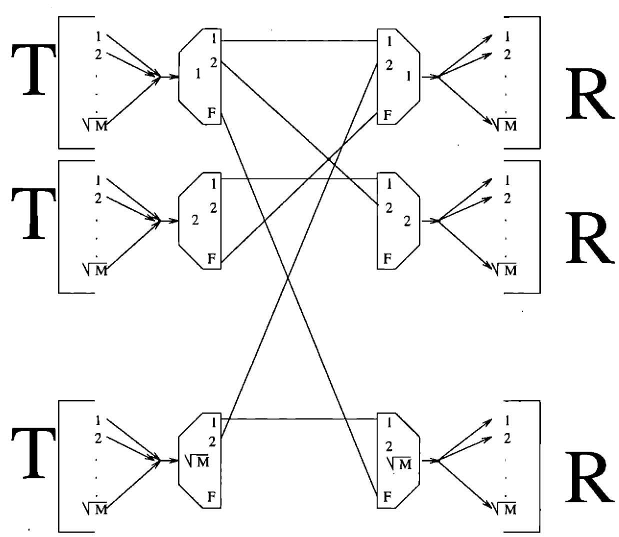


Figure 4-2: Implementation of $\mathcal{T}_{SB}$

4.5 Permutation Routing

As another example of how to apply Theorem 2, recall that the ρ-permutation traffic set is the set of all one-to-one matchings of transmitters to receivers with no more than ρM active sessions. Let $F(M, \rho, S)$ be the minimum number of wavelengths to do ρ-permutation routing over all λ-routing networks with S states. Also, let $S(M, \rho, F)$ be the minimum number of states to do ρ-permutation routing over all λ-routing networks with F wavelengths.

There are $\binom{M}{\rho M}$ ways to pick the number of active transmitters, $\binom{M}{\rho M}$ ways to pick the active receivers, and $(\rho M)!$ ways of matchings the transmitters to the receivers. Therefore, eqn. (4.4), with $x = \rho M$, gives the bounds

$$F(M, \rho, S) \geq S^{-\frac{1}{2\rho M}} \sqrt{\frac{\rho M}{e}} \tag{4.13}$$

$$\log_2 S(M, \rho, F) \geq \rho M \log_2 \rho M - 2\rho M \log_2 F - 1.44\rho M \tag{4.14}$$

For instance, passive AONs require $\sqrt{\rho M/e}$ wavelengths to do ρ-permutation routing.

4.6 Duplex Permutation Routing

In *duplex permutation routing* there are no multi-point connections and every user can call every other user, but if n is talking to m then m is talking to n. The duplex permutation traffics are therefore a subset of the permutation traffics and model typical phone service without conference calling.

Let's apply eqn. 4.4 to this case. There are M receivers transmitter 1 could request. Given that request, there are $M - 2$ receivers the next lowest available transmitter may request. Note if 1 requests 2, the next lowest available transmitter is 3, else it is 2. Given the first i requests, the next lowest available transmitter may request one of $M - 2i$ receivers. So the number of traffics is

$$M(M-2)(M-4)...2 = \frac{M!}{(M-1)(M-3)...1} \tag{4.15}$$

Since $M(M-2)...2 \approx (M-1)(M-3)...1$, the number of traffics is approximately $\sqrt{M!}$. Applying eqn. 4.4, we have $F \geq (M/e)^{1/4}$. This is much smaller than the bound for permutation routing!

However, we have picked this example to show that Theorem 2 need not be tight. Paradoxically, we show this by applying the bound to a subset of the traffics. Specifically consider only the duplex permutation traffics where all transmitters $n \leq M/2$ are talking to receivers $m > M/2$ and transmitters $n > M/2$ are talking to receivers $m \leq M/2$. Considering only the bottom half of the sessions, i.e. (n, m) where $n < m$, there are $M/2$ active transmitters, $M/2$ active receivers, and $(M/2)!$ traffics. Therefore applying eqn. (4.4), $F \geq \sqrt{M/2e} >> (M/e)^{1/4}$.

4.7 Conclusions

In this chapter we considered networks using a combination of wavelength routing, wavelength changing, and circuit switching. A general bound on the number of wavelengths was presented. The bound holds for all AONs and can be tight, even in the absence of wavelength converters.

Chapter 5

Connectors

A connector is a network which can support arbitrary one-to-one connectivity between its inputs and outputs. Connectors are traditionally built with one wavelength and many switches. Shannon proved that a single wavelength $M \times M$ connector requires at least $\Theta(M \log_2 M)$ 2×2 switches [Sha49] and Beneš showed that this number is sufficient [Ben65].

In this chapter, we consider building F wavelength all-optical $M \times M$ connectors. From Theorem (2) and the inequality $M! \geq M^M e^{-M}$, it follows that at least $\Theta(\sqrt{M})$ wavelengths are required to build a passive connector and at least $\Theta(S^{-\frac{1}{2M}}\sqrt{M})$ wavelengths are required to build a configurable connector with S states.

In section 5.1, we investigate connectors with wavelength changers. In this case, we are able to use known results in switching theory to get some quick answers. In particular, we show that no more than $\Theta(S^{-\frac{1}{2M}}\sqrt{M \log_2 M - \log_2 S})$ wavelengths are needed to build a rearrangeably non-blocking all-optical connector with wavelength changers. In addition, we show that no more than $\Theta(S^{-\frac{1}{2M}}\sqrt{M \log_2 M})$ wavelengths are needed to build a wide-sense non-blocking all-optical connector with wavelength changers.

Then in section 5.2, we consider building connectors without λ-changers. In this case, the results on switching networks are not as useful but the arguments can be modified with no penalty in the required number of wavelengths: no more than $\Theta(S^{-\frac{1}{2M}}\sqrt{M\log_2 M - \log_2 S})$ wavelengths for a rearrangeably non-blocking connector and no more than $\Theta(S^{-\frac{1}{2M}}\sqrt{M\log_2 M})$ for a wide-sense non-blocking connector.

We also consider the problem of building configurable AONs with 2×2 elementary switches. An elementary switch has two states: *bar* and *cross*. In the bar state, the switch passes signals on all wavelengths from input i to output i, $i = 1, 2$. In the cross state, the outputs are exchanged. The lower bound shows that at least $\Theta(M\log\frac{M}{F^2})$ switches are needed and the above result on rearrangeable connectors shows that $\Theta(M\log\frac{M}{F^2} - M\log\log F)$ switches suffice. [ABCR+93] has shown that for a wide-sense non-blocking connector, the same result hold.

For passive networks, there is still a factor of $\Theta(\sqrt{\log M})$ between the lower and upper bounds on the required number of wavelengths. Section 5.3 discusses some of the difficulties of closing this gap.

5.1 Connectors with Wavelength Changing

5.1.1 Rearrangeable Connectors

In this section, we prove that there exist rearrangeably non-blocking $M\times M$ connectors with no more than $\Theta(S^{-\frac{1}{2M}}\sqrt{M\log M - \log S})$ wavelengths.

Consider Fig. 5-1. Here a tunable transmitter is followed by a demultiplexer which separates each of the F wavelengths onto a unique fiber. Each demultiplexer output is then followed by a wavelength changer which converts the input signal to wavelength 1. All-to-one wavelength changing devices have been demonstrated, but currently their use is limited to signals using amplitude modulation [G+92b, G+92a, D+93]. The device in Fig. 5-1 is functionally equivalent to an $1\times F$ switch, the state of the switch being determined by the wavelength of the transmitter. Similarly, the device in Fig. 5-2 is functionally equivalent to a $F\times 1$ switch, where here the i^{th} wavelength changer converts wavelength 1 to wavelength i. Now consider a network with M

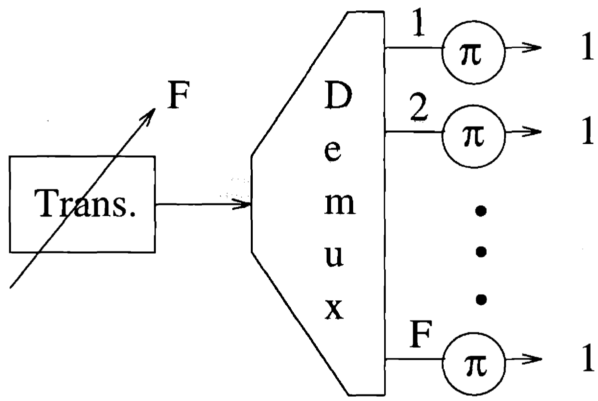


Figure 5-1: An all optical implementation of a $1 \times F$ switch.

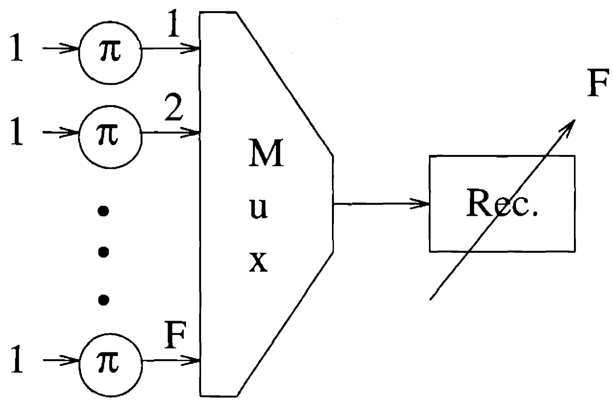


Figure 5-2: An all optical implementation of a $F \times 1$ switch.

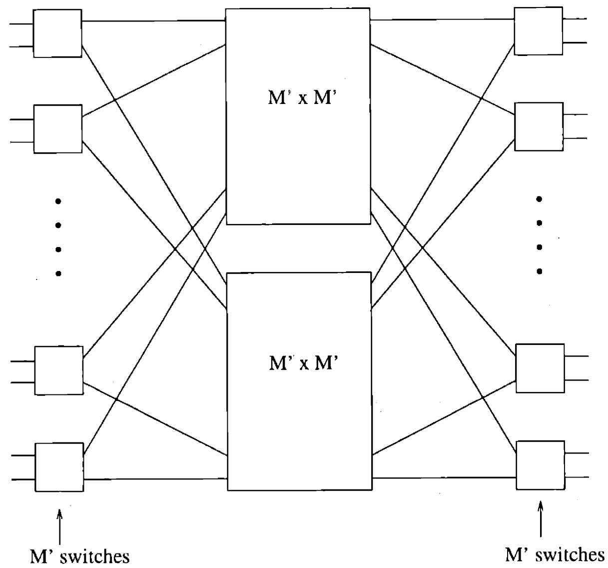


Figure 5-3: Beneš like network

transmitters, M receivers, and F wavelengths. Use the above constructions to build an $1 \times F$ switch after each transmitter and a $F \times 1$ switch before each receiver. By connecting outputs of the $1 \times F$ switches to one or more inputs of the $F \times 1$ switches, a network with two switching stages can be built. 2-stage wide-sense non-blocking, and hence also rearrangeably non-blocking connectors were considered in [FFP88]. The goal there was to minimize the number of switching cross-points; the goal here is to minimize the switch size. [FFP88, Theorem 3] proves the existence of 2-stage wide-sense non-blocking connectors using $\Theta(M^{3/2}\sqrt{\log_2 M})$ switching crosspoints and it is easy to verify that the proof uses switches of size no more than $8\left\lceil\sqrt{M\log_2 M}\,\right\rceil$.

Let $F(M,S)$ be the minimum number of wavelengths required to build an $M \times M$ connector with S states. We have just shown that $F(M,1) \leq 8\left\lceil\sqrt{M\log_2 M}\,\right\rceil$. For the configurable case, we use a technique of Sasaki [Sas93]. Let k be a positive integer. We will show below how to build a configurable $2^k M' \times 2^k M'$ connector with elementary switches and multiple copies of a passive $M' \times M'$ connector. We assume that the $M' \times M'$ connector uses the minimum number of wavelengths, $F(M',1)$.

The configurable connector uses the same number of wavelengths and $K_k = 2^k k M'$ switches so that $S = 2^{K_k}$. Therefore, we will show that

$$F(2^k M', 2^{K_k}) \leq F(M', 1) \tag{5.1}$$

for any $k \geq 1$. Making the change of variables, $M = 2^k M'$ and $k = \frac{1}{M} \log_2 S$,

$$F(M, S) \;\; \leq \;\; F(MS^{-\frac{1}{M}}, 1) \tag{5.2}$$

Using the above results on passive networks now shows that

$$F(M, S) \leq \Theta(S^{-\frac{1}{2M}} \sqrt{M \log M - \log S}). \tag{5.3}$$

We now iteratively build the desired connector. For $k = 1$, the $2M' \times 2M'$ connector is shown in Fig. 5-3. The first and last stage together consist of $K_1 = 2M'$, 2×2 switches and the middle stage consists of two $M' \times M'$ **passive** connectors with $F(M', 1)$ wavelengths. For $k = 2$, repeat the iteration using $4M'$ new switches and two of the $2M' \times 2M'$ connectors just constructed to build a $4M' \times 4M'$ connector with $K_2 = 4M' + (2M' + 2M') = 8M'$ switches and $F(M', 1)$ wavelengths. After k iterations, we have k stages of 2×2 switches, with $2^{k-1}M'$ switches per stage, 2^k $M' \times M'$ passive networks in the middle, followed by k stages of 2×2 switches with $2^{k-1}M'$ switches per stage. The final construction is a Beneš network with the middle $2 \log_2 M' - 1$ stages replaced by 2^k passive $M' \times M'$ connectors.

5.1.2 Wide-Sense Non-blocking Connectors

Let $F_w(M,S)$ be the minimum number of wavelengths to build a wide-sense non-blocking connector with S states. Since [FFP88, Theorem 3] proves the existence of 2-stage wide-sense non-blocking connectors with switch size no more than $8\left\lceil\sqrt{M\log_2 M}\right\rceil$,

$$F_w(M,1) \le \Theta(\sqrt{M\log_2 M}) \tag{5.4}$$

which is the same as the best known upper bound on $F(M,1)$. We cannot use the technique we used in the last section to build configurable connectors since the Beneš network is not wide-sense non-blocking. We can however get almost the same result using a different technique.

For $S>1$, replace the $1\times F$ switches and the $F\times 1$ switches by $1\times bF$ and $bF\times 1$ switches, where $b=\left\lfloor S^{\frac{1}{2M}}\right\rfloor$. The construction of the switches is straightforward. For instance, a $1\times bF$ switch is shown in Fig. 5-4. A network built from these switches has less than S states. Therefore since 2-stage connectors can be built with switch size $8\left\lceil\sqrt{M\log M}\right\rceil$, all-optical connectors can be built in this way with F wavelengths if

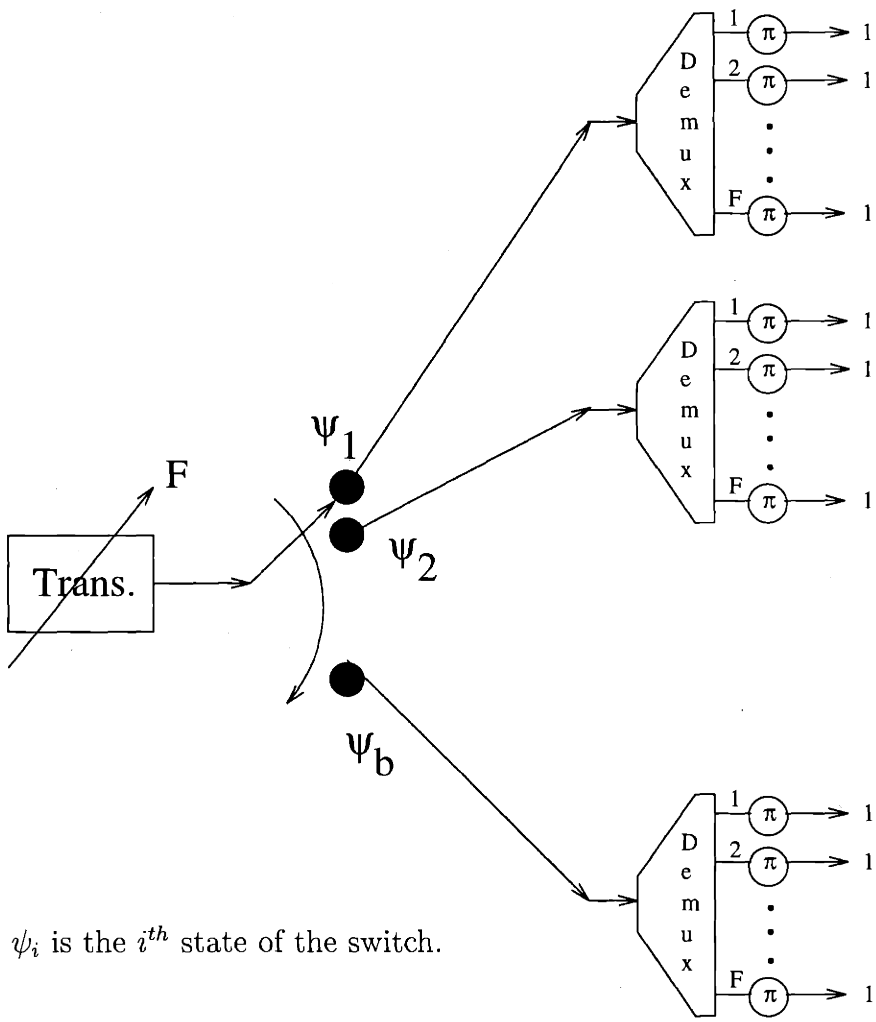
$$F\cdot b \le 8\left\lceil\sqrt{M\log_2 M}\right\rceil. \tag{5.5}$$

We have just proved the following theorem,

Theorem 5 WSNB Networks with Wavelength Changing

Let $F_w(M,S)$ be the minimum number of wavelengths to build a wide-sense non-blocking connector with S states. Then

$$S^{-\frac{1}{2M}}\sqrt{\frac{M}{e}} \le F_w(M,S) \le \Theta(S^{-\frac{1}{2M}}\sqrt{M\log_2 M}) \tag{5.6}$$



ψ_i is the i^{th} state of the switch.

Figure 5-4: An all-optical implementation of a $1 \times bF$ switch using a $1xb$ switch.

Unfortunately, [FFP88] does not construct these connectors; only their existence is proven. The proof randomly interconnects the inputs and outputs of the first and second stages. It is shown that if the interconnections are made in an appropriately random fashion, a connector will be constructed with probability greater than 0.

Up until very recently, the best 2-stage construction was also due to Feldman, Friedman, and Pippenger [FFP88] and required $\Theta(M^{5/3})$ cross-points. Since the switch size is $\Theta(M^{2/3})$, the construction is also a wide-sense non-blocking connector with $F = \Theta(M^{2/3})$ wavelengths.

Wigderson and Zuckerman [WZ93] have recently proposed a polynomial time algorithm for constructing 2-stage wide-sense non-blocking connectors with switch size $p(M)\sqrt{M}$, where

$$p(M) = 2^{(\log n)^{5/6+o(1)}} = M^{o(1)} \tag{5.7}$$

They also report that the 5/6 factor can be reduced to 4/5 using a more complicated algorithm. This leads to $p(M)\sqrt{M}$ wavelength connectors. Unfortunately, $p(M)$ decays so slowly that in order for $p(M)\sqrt{M}$ to be less than M, we need $M \geq 10^{28}$. Therefore for reasonable M, the $\Theta(M^{2/3})$ construction mentioned above is still the best.

5.2 Connectors Without Wavelength Changing

In the last section we showed how to implement any 2-stage switching network all-optically with wavelength changers. Without wavelength changers, a 2-stage switching network still has an all-optical implementation but may require many more wavelengths. We outline the difficulties below. Then we consider building all-optical connectors without wavelength changing. We consider rearrangeable connectors in section 5.2.1 and wide-sense non-blocking connectors in section 5.2.2.

Let G be a 2-stage switching network with switch size d. With wavelength changers, G can be implemented all-optically with d wavelengths using the constructions in Fig. 5-1 and Fig. 5-2. Without wavelength changers, many more wavelengths may be necessary. To understand why, let us say that G has *pin changing* if an output i of some first stage switch is connected to an input $j \neq i$ of some second stage switch. Clearly if G does not have pin changing, it can be implemented all-optically without wavelength changing with d wavelengths; proceed as before using the $1 \times F$ and the $F \times 1$ switches but removing the wavelength changers.

If G has pin changing, there may or may not be such an implementation. There will be an implementation if there is a consistent labeling of the inputs and outputs of the first and second stage switches with d labels such that if an output is connected to an input, they have the same label, and such that two ports (outputs or inputs) of a switch have different labels. If there were such a labeling, the ports could be renamed in such a way that the renamed network did not have pin-changing. Therefore we will assume that if G has pin-changing, there is no such labeling.

Now suppose G has pin changing and maximum switch size d. We can still implement this switch all-optically without wavelength changing, but the number of wavelengths may be much larger than d. We shall not try to determine the required number of wavelengths but simply show that no more than d^2 wavelengths are ever necessary. To see this, let us define the $M \times M$ switching connectivity matrix $[G(n,m)]$, where $G(n,m)$ is the set of paths connecting source n to source m. A path is represented by a pair $i:j$. $i:j \in G(n,m)$ means that the i^{th} output of source n's first stage switch is connected to the j^{th} input of sink m's switch. Now construct a new switching network G' with switch size d^2. Label the outputs (inputs) of each first (second) stage switch by the pairs (a,b), $a,b = 1,2,...,d$. Connect output (a,b) of the n^{th} first stage switch to input (a,b) of the m^{th} second stage switch iff $a\!:\!b \in G(n,m)$. This new network has a switch size of d^2, does not have pin changing, and is equivalent to G in terms of connectivity, i.e. it supports exactly the same traffics. Therefore it can be implemented all-optically without wavelength changers with d^2 wavelengths.

In section 5.1, we used results from [FFP88]. The proof randomly interconnects switches of size $d = \Theta(\sqrt{M \log M})$. Given the random nature of the argument, the connector may have pin changing and therefore require up to $d^2 = \Theta(M \log M)$ wavelengths if wavelength changers are not used.[1] This is why we need the proofs in sections 5.2.1 and section 5.2.2.

[FFP88] also presented a construction with switch size $\Theta(M^{2/3})$. Because of the difficulties discussed above, this design requires M wavelengths without wavelength changers. This construction will be discussed again in section 5.3.

Given that the best known 2-stage designs are not efficiently implementable all-optically without wavelength changers, it is surprising that the best bounds for connectors with and without λ-changing are currently the same for all posed connection problems.

5.2.1 Rearrangeable Connectors

We first show that $\Theta(\sqrt{M \log_2 M})$ wavelengths are sufficient to build passive rearrangeably non-blocking connectors **without** wavelength changing. The arguments used are an adaptation of those used in [Pip73] for rearrangeable networks taking into account the particular restrictions of networks without wavelength changers. Then using the Beneš construction we used in section 5.1.1, we show that $\Theta(S^{-\frac{1}{2M}}\sqrt{M \log_2 M - \log_2 S})$ wavelengths are sufficient for an S state network.

We are interested in $M \times M$ connectors, but it will help us later on if we construct an asymmetrical network. So consider a network with M_t inputs, M_r outputs, and M_r/a single-fiber trunks where a is a constant to be determined later. The outputs are divided in M_r/a groups, each group being served by one trunk. Input n is connected to trunk j through a fiber and a filter which passes the wavelengths in the set H_n^j

[1] Of course we would never need more than M wavelengths to build a connector; we are talking here about the all-optical implementation of a particular design.

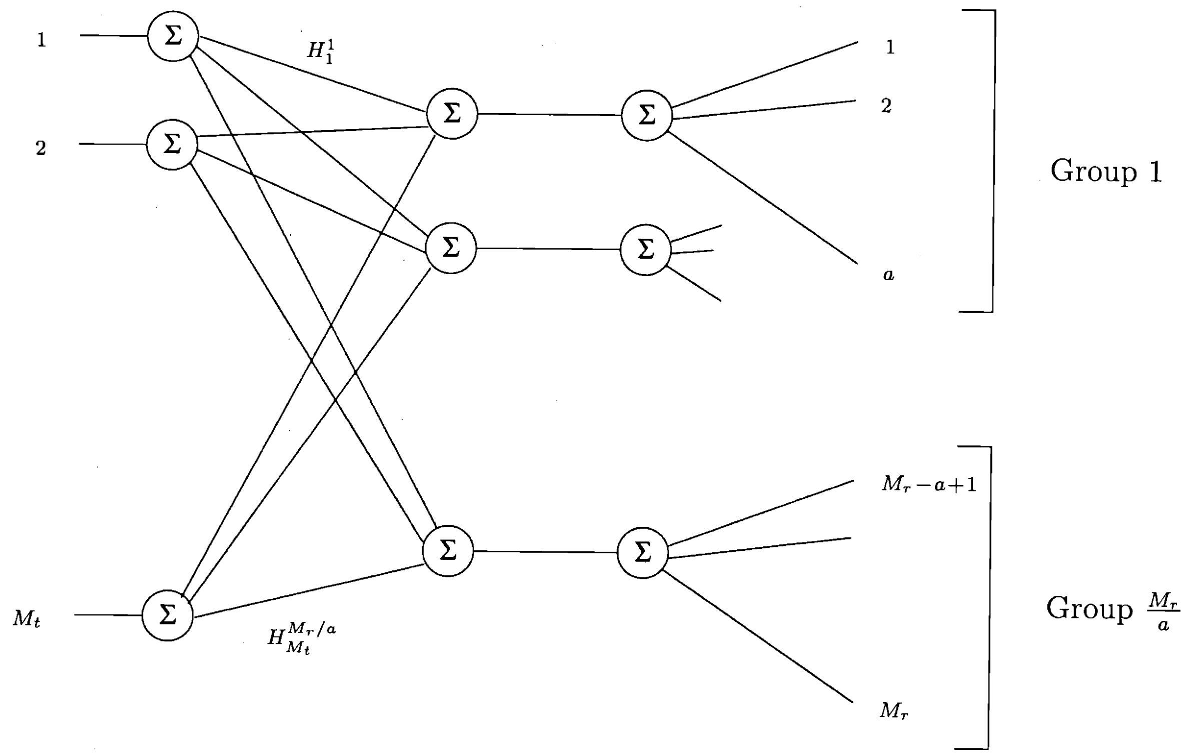


Figure 5-5: Rearrangeably non-blocking AON w/o wavelength changers.

and blocks the other wavelengths. This network is shown in Fig. 5-5. The filters H_n^j are picked such that H_n^j and $H_n^{j'}$ pass no common wavelengths for all choices of n and $j \neq j'$. This insures that the signal from input n passes through a maximum of one trunk. Call this the *trunk independence* property. Notice that this is a LT-AON.

Given a traffic ϕ, define ϕ^j to be the sessions in ϕ whose destinations are served by trunk j. Since each signal passes through at most one trunk, the traffic is supported by the network iff each ϕ^j is supported by the network. Furthermore, ϕ^j is supported iff each session in ϕ^j can be assigned a distinct wavelength such that the signals pass through the filters, i.e. if $(n, m) \in \phi^j$, then (n, m) must be assigned a wavelength in H_n^j.

Let $\phi^j = \{(n_1, m_1), (n_2, m_2), ..., (n_L, m_L)\}$. Then for the network to support ϕ^j, there must be a list of L distinct wavelengths, $w_1, w_2, ..., w_L$, such that $w_i \in H_{n_i}^j$.

Such a list is called a system of distinct representation (SDR) of the sets $H^j_{n_1}$, $H^j_{n_2}$,..., $H^j_{n_L}$. Hall's Theorem asserts that these sets will have an SDR iff the union of every $k \leq L$ of these sets has at least k elements [10]. We will prove that if $F \geq \Theta(\sqrt{M_r \log_2 M_t})$ and the sets H^j_n are picked in an appropriate random fashion satisfying the trunk independence property, Hall's Theorem will be satisfied for every possible ϕ^j with positive probability. Therefore, with positive probability we will have picked a connector. This establishes the existence of connectors with $F \leq \Theta(\sqrt{M \log_2 M})$ wavelengths [Bar93].

Theorem 6 Wavelength Efficient Passive Connectors w/o λ-Changing

For $M_t \geq M_r$, there exist passive $M_t \times M_r$ connectors with $\Theta(\sqrt{M_r \log_2 M_t})$ wavelengths.

Proof: Recall that there are M_r/a trunks. Let each filter pass d wavelengths. a and d are constants to be specified later but we require that $F = M_r d/a$. For each input n, pick H^1_n uniformly randomly from the set of all filters which pass d wavelengths, i.e. $p(H^1_n) = \binom{F}{d}^{-1}$. Having picked H^j_n for $j = 1, 2, ..., J$, pick H^{J+1}_n uniformly randomly from the set of all filters which pass d wavelengths not used in $\bigcup_{j=1}^{J} H^j_n$. Therefore, the trunk independence property is satisfied. Also, notice that H^j_n is statistically independent of $H^j_{n'}$ for each j and all choices of $n \neq n'$.

Let P_f be the probability of not having a connector. We will show that $P_f < 1$. If we do not have a connector, then there is some trunk j such that there exists a traffic ϕ^j that cannot be supported. By the union bound and by symmetry,

$$P_f \leq \frac{M_r}{a} P'_f$$

where P'_f is the probability that there exists a traffic ϕ^1 that cannot be supported by trunk 1. If there exists a traffic which cannot be supported by trunk

1, then by Hall's theorem, there exists L sets, $L \leq a$, whose union contains less than L elements. Therefore, P'_f is the probability that there exists an L, $d+1 \leq L \leq a$, and $i_1, i_2, \ldots i_L$ such that

$$|\bigcup_{l=1}^{L} H^1_{i_l}| < L$$

L can be any number between $d+1$ and a; there are $\binom{M_t}{L}$ ways to pick the L sets and $\binom{F}{L-1}$ to choose the $L-1$ wavelengths the union of these sets must fall in. Since the sets are picked independently, P_f is no more than

$$P_f \leq \frac{M_r}{a} \sum_{L=d+1}^{a} \left[\frac{\binom{L-1}{d}}{\binom{F}{d}} \right]^L \binom{F}{L-1} \binom{M_t}{L}$$

Now using $\binom{M_t}{L} \leq M_t^L$, $\binom{F}{L-1} \leq F^L$, and $\binom{L-1}{d} / \binom{F}{d} \leq (L/F)^d$, we see that

$$\begin{aligned} P_f &\leq \frac{M_r}{a} \sum_{L=d+1}^{a} \left[\left(\frac{L}{F}\right)^d M_t F \right]^L \\ &\leq \frac{M_r}{a} \sum_{L=d}^{\infty} \left[\left(\frac{a}{F}\right)^d M_t^2 \right]^L \end{aligned}$$

since F need never be more than $\max\{M_t, M_r\} = M_t$. Evaluating the sum,

$$P_f \leq \frac{M_r}{a} \left[\left(\frac{a}{F}\right)^d M_t^2 \right]^d \frac{1}{1 - \left(\frac{a}{F}\right)^d M_t^2}$$

Substituting $a = M_r d/F$ and $F = \sqrt{2M_r d}$, we see that

$$P_f \leq \sqrt{2M_r/d} \left[2^{-d} M_t^2 \right]^d \frac{1}{1 - M_t^2 2^{-d}}$$

which is less than 1 if $d \geq \log_2(2M_t^2)$. Therefore if $F \geq \sqrt{2M_r \log_2(2M_t^2)}$, $P_f < 1$. □

Using the analogy to 2-stage switching networks, we now have the following corollary.

Corollary 7 *There exists 2-stage switching networks without pin changing with only $\Theta(M^{3/2}\sqrt{\log_2 M})$ cross-points.*

Building the Beneš network as before, we have the following theorem.

Theorem 8 Wavelength Efficient Connectors w/o λ-Changing

Let $F'(M, S)$ be the minimum number of wavelengths to do permutation routing without wavelength changing over all λ-routing networks with S states. Then

$$S^{-\frac{1}{2M}}\sqrt{\frac{M}{e}} \leq F'(M, S) \leq \Theta(S^{-\frac{1}{2M}}\sqrt{M \log M - \log S})$$

5.2.2 Wide-Sense Non-blocking Connectors

We show that $\Theta(S^{-\frac{1}{2M}}\sqrt{M \log_2 M})$ wavelengths are sufficient to build wide-sense non-blocking connectors without wavelength changing. The arguments are an adaptation of those used in [FFP88] and are almost the same as the rearrangeable case. The key difference is that instead of Hall's Theorem, an expander theorem will be used.

First consider the passive case. The network is the same as the last section, but we'll find that d must be slightly larger. Let $\phi^j = \{(n_1, m_1), (n_2, m_2), ..., (n_L, m_L)\}$ be a traffic with destinations being served by trunk j. A necessary condition for the network to support ϕ^j is that $H^j_{n_1}$, $H^j_{n_2}$,..., $H^j_{n_L}$ has an SDR. For the network to be wide-sense non-blocking, there must exist an algorithm for determining this SDR as sessions come on and off without re-assigning wavelengths. The following expander theorem guarantees the existence of such an algorithm. We omit the proof which can be found in [FFP88].

Theorem 9 Expander Matching Theorem

Given sets $H_1, H_2, ..., H_M$, there exist an algorithm for dynamically assigning wavelengths for up to b sessions if the union of any k of the sets has at least $2k$ elements for each $k \leq 2b$.

We now have the following theorem.

Theorem 10 Efficient Passive WSNB Connectors w/o λ-Changing

For $M_t \geq M_r$, there exists passive wide-sense non-blocking $M_t \times M_r$ connectors with $\Theta(\sqrt{M_r \log_2 M_t})$ wavelengths.

Proof: There are M_r/a trunks. Let each filter pass d wavelengths. a and d are constants to be specified later but we require that $F = M_r d/a$. Randomly pick the filters in the same way as before.

Let P_f be the probability of not having a wide-sense non-blocking connector. We will show that $P_f < 1$. If we do not have such a connector, then there is some trunk j such that $H_1^j, H_2^j, ..., H_M^j$ do not satisfy the conditions in the Expander Matching Theorem. By the union bound and by symmetry, $P_f \leq \frac{M_r}{a} P_f'$, where P_f' is the probability that trunk 1 fails. By the expanding theorem, if trunk 1 fails to be wide-sense non-blocking, then there exists L sets, $L \leq 2a$, whose union contains less than $2L$ elements. Therefore, P_f' is the probability that there exists an L, $\frac{d}{2} + 1 \leq L \leq 2a$, and $i_1, i_2, ... i_L$ such that

$$|\bigcup_{l=1}^{L} H_{i_l}^1| < 2L$$

L can be any number between $\frac{d}{2} + 1$ and $2a$; there are $\binom{M_t}{L}$ ways to pick the L sets and $\binom{F}{2L-1}$ to choose the $2L-1$ wavelengths the union of these sets must fall in. Since the sets are picked independently, P_f is no more than

$$P_f \leq \frac{M_r}{a} \sum_{L=\frac{d}{2}+1}^{2a} \left[\frac{\binom{2L-1}{d}}{\binom{F}{d}} \right]^L \binom{F}{2L-1} \binom{M_t}{L}$$

Now using $\binom{M_t}{L} \leq M_t^L$, $\binom{F}{2L-1} \leq F^{2L}$, and $\binom{2L-1}{d} / \binom{F}{d} \leq (2L/F)^d$, we see that

$$P_f \leq \frac{M_r}{a} \sum_{L=\frac{d}{2}+1}^{2a} \left[\left(\frac{2L}{F} \right)^d M_t F^2 \right]^L$$

$$\leq \frac{M_r}{a} \sum_{L=\frac{d}{2}}^{\infty} \left[\left(\frac{4a}{F}\right)^d M_t^3 \right]^L$$

Evaluating the sum,

$$P_f \leq \frac{M_r}{a} \left[\left(\frac{4a}{F}\right)^d M_t^3 \right]^{\frac{d}{2}} \frac{1}{1 - \left(\frac{4a}{F}\right)^d M_t^3}$$

Substituting $a = M_r d/F$ and $F = \sqrt{8M_r d}$, we see that

$$P_f \leq \sqrt{8M_r/d} \left[2^{-d} M_t^3\right]^{d/2} \frac{1}{1 - M_t^3 2^{-d}}$$

which is less than 1 if $d \geq \log_2(2M_t^3)$. Therefore if $F \geq \sqrt{8M_r \log_2(2M_t^3)}$, $P_f < 1$. $\square$

For a configurable $M \times M$ connector, we use the construction shown in Fig. 5-6. Let $b = \left\lfloor S^{\frac{1}{M}} \right\rfloor$. The network consists of M, $1 \times b$ elementary switches followed by b, $M \times \frac{M}{b}$ passive wide-sense non-blocking networks. Now use the above theorem with $M_t = M$ and $M_r = \frac{M}{b}$. This shows that $\Theta(S^{-\frac{1}{2M}} \sqrt{M \log_2 M})$ wavelengths are sufficient.

Aggarwal et. al. have improved this bound to

$$F'_w(M, S) \leq \Theta(S^{-\frac{1}{2M}} \sqrt{M \log_2 M - \log_2 S}) \tag{5.8}$$

matching the rearrangeable case [ABCR+93]. They have also figured out how to use the results of Widgerson and Zuckerman to construct all-optical wide-sense non-blocking connectors with $p(M)\sqrt{M}$ wavelengths **without** wavelength changers. $p(M)$ is given by eqn. (5.7) as before. This is now the best known construction for large M. However, $p(M)$ decays so slowly that in order for $p(M)\sqrt{M}$ to be less than M, we need $M \geq 10^{28}$. Therefore, for $M \leq 10^{28}$, the best known connector without wavelength changers is the $\left\lceil \frac{M}{2} + 2 \right\rceil$ connector presented in Chapter 3. The

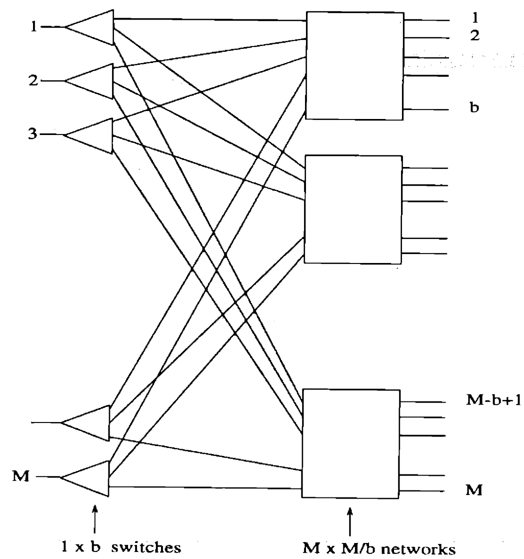


Figure 5-6: Wide-sense non-blocking AON w/o wavelength changers

best known construction with wavelength changers for $M \leq 10^{28}$ is the all-optical implementation of the 2-stage construction in [FFP88] and requires $\Theta(M^{2/3})$ wavelengths.

5.3 Light Tree AONs

Throughout this chapter, we have used results and techniques from the study of 2-stage switching networks. Actually, only a subset of these networks has been considered in the literature for connectors: *depth 2 interconnection networks.*[2] A depth 2 interconnection network is a directed graph with 3 stages of nodes: one stage of input nodes, one stage of middle nodes, and one stage of output nodes. The nodes correspond to busses and the edges correspond to switching cross-points. Each cross-point can be open or closed; a closed cross-point allows a signal to flow across the edge. At

[2] 2-stage switching networks which are not interconnection networks were considered in [BLM93], but not for building connectors. We will discuss these results in Part III.

most one signal can tranverse any bus or cross-point so each input and output can serve at most one user if the network is to be a connector. In general, the depth of an interconnection network is the maximum number of cross-points (edges) a signal can pass through from an input to an output. Unfortunately, depth 2 interconnection networks are sometimes referred to as 3-stage interconnection networks which should not be confused with 3-stage switching networks; the former has 3 stages of nodes and the latter 3 stages of switches. A 3-stage switching network may or may not have an interconnection network representation. If it does, it would be depth 3 and have 4 stages of nodes in the representation.

All the AONs considered in this chapter, in fact in this part of the thesis, are LT-AONs. LT-AONs are in some sense equivalent to depth 2 interconnection networks. This section discusses that equivalence and contains various results which should provide the reader with a deeper insight into these networks and the difficulty of constructing wavelength efficient passive connectors.

First we need a couple of definitions. Referring to Fig. 1-3, consider an arbitrary passive LT-AON with L light trees. Form a 3-stage directed graph where the first stage nodes are transmitters, the second stage nodes are trunks, and the third stage nodes are receivers. Put a directed edge from a periphery node to a middle stage node if the transceiver is part of the light tree. Now label each of the middle nodes with the wavelength of the light tree. We call this a *labeled interconnection network of depth 2.*

For an arbitrary LT-AON H[3], let $\mathcal{G}_H$ be its labeled interconnection network. A session in H corresponds to a length two path in $\mathcal{G}_H$ (path length measured in edges). Similarly, a length two path in $\mathcal{G}_H$ corresponds to a session in H. Two non-contending sessions in H cannot share a light tree, so two non-contending paths in $\mathcal{G}_H$ cannot share a middle vertex. Likewise, any set of middle vertex disjoint paths

[3]Since whether a network is a LT-AON is completely determined by its connectivity matrix H, we will typically refer to H as the network itself.

in $\mathcal{G}_H$ represents a set of non-contending sessions in H.

Interconnection networks for permutation routing have been well studied. The goal is generally to minimizing the number of edges as a function of the depth of the network. In 1949, Shannon [Sha49] proved that a non-blocking interconnection network of any depth must have at least $M \log_2(M/e)$ edges. The argument is as follows. Let $|E|$ be the number of edges. Then the network has at most $2^{|E|}$ states since an edge can be used or not used in a path. No two permutations share a state so it must be that $2^{|E|} \geq M!$ which gives $|E| \geq M \log_2(M/e)$. Notice that if $\mathcal{G}_H$ is a labeled interconnection network of a LT-AON with F wavelengths, then $\mathcal{G}_H$ has less than $2FM$ edges. This is true since each transmitter or receiver can be connected to at most one light tree of a given wavelength. Therefore, for a LT-AON

$$2FM \geq |E| \geq M \log_2 M - 1.44M \tag{5.9}$$

which gives $F \geq .5 \log_2 M$. Since we have already shown that at least $\sqrt{M/e}$ wavelengths are required, the bound is not useful here. Shannon's bound fails because it does not take into account the fact that the network has depth 2.

A better bound was derived in 1982 by Pippenger and Yao [PY82]. Here, they showed that at least $2M^{3/2}$ edges are required for a rearrangeably non-blocking depth 2 interconnection network. Therefore, a Light Tree Connector requires

$$F \geq \sqrt{M} \tag{5.10}$$

wavelengths, a factor of $\sqrt{e}$ better than the bound of Theorem 2. We will derive eqn. (5.10) from a more general bound, the Light Tree Bound, in Chapter 13. The relationship between the Light Tree Bound and Pippenger and Yao's bound will be discussed there but note for now that the proof does not generalize to all AONs.

It is interesting to note that Pippenger and Yao proved a much stronger result than we have just stated. They proved that at least $2M^{3/2}$ edges are needed for a depth 2 interconnection network to be an M-shifter. An M-shifter is a network that supports the M traffics of the form

$$\{(1,i),(2,i+1),...,(M,i+M-1)\} \tag{5.11}$$

for $i = 0, 1, ...M-1$, where addition is done modulo M. Therefore a Light Tree M-shifter requires $\sqrt{M}$ wavelengths. In fact this bound is tight since a $\sqrt{M} \times \sqrt{M}$ WDM cross-connect with $\sqrt{M}$ users connected to each input and each output can function as an M-shifter. To see this, number the users connected to input n of the cross-connect (n,i), for $i = 0, 1, ..., \sqrt{M} - 1$. Similarly label the users connected to output m of the cross-connect (j,m), for $j = 0, 1, ..., \sqrt{M} - 1$. Each shift i is representable by a pair (a_i, b_i) where user (n,j) is shifted to $(n + a_i, j + b_i)$, addition modulo $\sqrt{M}$. Since two users connected to the same input are never shifted to the same output, the LR is an M-shifter.

We say an AON is *strict sense non-blocking* (S.S.NB.) if given any set of active sessions without contention and a new session request between an inactive transmitter and an inactive receiver, there always exists an OD channel to feasibly assign the session. At least $\left\lceil \frac{M}{2} \right\rceil$ wavelengths are required for a strict sense non-blocking Light Tree AON since Friedman [Fri88] showed that at least $2M^2$ edges are required for a strict sense non-blocking connector. We present this proof in terms of LT-AONs below. Since the $\left\lceil \frac{M}{2} + 2 \right\rceil$ connector presented in Chapter 3 is a LT-AON and is strict sense non-blocking, this bound is almost achievable even without wavelength changers. Currently, the best lower bound for strict sense non-blocking NLT-connectors is $\sqrt{M/e}$.

Theorem 11 Strict Sense Non-blocking LT-AONs

A S.S.NB. LT-connector requires $F \geq \left\lceil \frac{M}{2} \right\rceil$ *wavelengths.*

Proof: Number the light trees from 1 to L and let $V(n, m)$ be the light trees connecting n to m. We begin with an observation. Suppose that $V(1, 1) = \{c_1, c_2, ..., c_K\}$, $K < M$, and that for each $k = 1, ..., K$, channel c_k is used at least k more times in row 1. Then it is always possible to select K receivers $r_1, r_2, ..., r_K$ other than 1, such that $c_k \in H(1, r_k)$. To see this, notice that wavelength c_1 is used at least one other time in row 1. So pick r_1 such that $c_1 \in H(1, r_1)$ and $r_1 \neq 1$. Now since wavelength c_2 is used at least two other time other than in $(1, 1)$, pick r_2 such that $c_2 \in H(1, r_2)$ and $r_2 \neq 1, r_1$. Continue this procedure until $r_1, r_2, ...r_K$ have been determined. The following lemma is based on this observation.

Lemma 12 *Let* h_i *be the number of transmitters connected to light tree* i, $i = 1, 2, ...L$. *Similarly, let* w_i *be the number of receivers connected to light tree* i, $i = 1, 2, ..., L$. *Then if* H *is S.S.NB.,*

$$\sum_{i \in V(n,m)} \frac{1}{h_i} + \frac{1}{w_i} \geq 1 \tag{5.12}$$

for all (n, m).

Proof: We'll prove eqn. (5.12) for any arbitrary (n, m). Let

$$V(n, m) = \{c_1, c_2, ..., c_K\} = \{c'_1, c'_2, ..., c'_K\}$$

where c'_k is just another ordering of the elements of $V(n, m)$. Let $a_k = h_{c_k}$ and $b_k = w_{c'_k}$ and assume without loss of generality that the ordering has been chosen so that a_k and b_k are non-decreasing with k. Then we want to show that

$$\sum_{k=1}^{K} \frac{1}{a_k} + \frac{1}{b_k} \geq 1$$

if H is S.S.NB. There are two cases. First suppose that for some i, $a_i \leq i$

or $b_i \leq i$. Then the claim is true since

$$\begin{aligned}\sum_{k=1}^{K} \frac{1}{a_k} + \frac{1}{b_k} &\geq \sum_{k=1}^{i} \frac{1}{a_k} + \frac{1}{b_k} \\ &\geq i * \left(\frac{1}{a_i} + \frac{1}{b_i}\right) \quad \text{since } a_k \text{ and } b_k \text{ are non-decreasing} \\ &\geq 1\end{aligned}$$

Now suppose that $a_k \geq k+1$ and $b_k \geq k+1$ for $k = 1, 2, ...K$. We will show that H is not S.S.NB. using the observation made earlier. Since $a_1 \geq 2$ there is another receiver connected to light tree c_1. Call this rec_1. Now since $a_2 \geq 3$, there is a receiver different from m and rec_1 connected to c_2. Proceeding in the same way, we can select K distinct receivers $rec_1, rec_2, ..., rec_K$ such that $c_k \in V(n, rec_k)$. Similarly, we can select K transmitters, $tran_1, tran_2, ..., tran_K$ such that $c'_k \in V(tran_k, m)$. Let $tran'_1, tran'_2, ..., tran'_K$ be a re-ordering of the transmitters $tran_1, ..., tran_K$ such that $c_k \in H(tran'_k, m)$.

Now, pick the traffic $\{(tran'_1, rec_1), (tran'_2, rec_2), ..., (tran'_K, rec_K)\}$ and assign session $(tran'_k, rec_k)$ light tree c_k. Since H is a LT AON, $c_k \in V(tran'_k, rec_k)$. There is no way to route the session (n, m) since all its channels have been blocked. Therefore H is not S.S.NB. □

This lemma and the fact that a transmitter (receiver) can be connected to at most F LTs will give us our desired result.

By the lemma,

$$\sum_{(n,m)} \left[\sum_{i \in V(n,m)} \left(\frac{1}{h_i} + \frac{1}{w_i}\right) \right] \geq M^2$$

But this sum is identical to

$$\sum_{i=1}^{L} \left(\frac{1}{h_i} + \frac{1}{w_i}\right) h_i w_i = \sum_{i=1}^{L} (h_i + w_i)$$

Since each transmitter n and each receiver m can be connected to at most F LTs, this sum is no more than $2FM$. □

Note that a depth 2 interconnection network with $|E|$ edges is not in general a LT-AON with $F = .5|E|/M$ wavelengths. There are two ways the interconnection network can fail to represent a LT-AON. First, the degree of a peripheral node may be greater than F in an interconnection network but not in a labeled interconnection network. Second and more troublesome, is that without wavelength changers, there may not be a way to consistently label the middle stage nodes with only F labels (wavelengths) since in this case each input or output can be connected to at most one light tree of the same wavelength. In fact, to transform the $O(M^{5/3})$ edge construction mentioned above to a LT-AON without wavelength changing requires M wavelengths. The reason being that any two middle nodes of this construction have a common input. Therefore, those two nodes must have distinct wavelengths. There are M middle nodes and therefore M wavelengths are required.

We are now in a better position to understand the difficulties of constructing a passive LT-connector without wavelength changing with only $\Theta(\sqrt{M})$ wavelengths. This problem is equivalent to constructing a depth 2 interconnection network with $\Theta(M^{3/2})$ edges, with bounded degrees on the first and third stages, and which can be labeled with $\Theta(\sqrt{M})$ wavelengths. If we were able to construct such a network, we could construct a depth 2 interconnection network with $\Theta(M^{3/2})$ edges beating the best known construction by a factor of $p(M)$ and the construction in [FFP88] by a factor of $M^{1/6}$. With λ-changers, labeling the middle nodes is no longer a problem. We still have the bounded degree property so strictly speaking constructing a wavelength efficient passive LT-connector with wavelength changing is a harder problem than constructing an edge efficient depth 2 interconnection network. However it would not be surprising if for symmetric problems, like a connector or some other homogeneous traffic, the number of edges is minimized by graphs with constant degree. There is no general proof of this though.

5.4 Determining the Number of Switches

Since for $F < \sqrt{M/e}$, it is impossible to build a connector without switches, we now consider the issue of whether using a combination of switching and wavelength routing can significantly reduce the number of switches required when $F < \sqrt{M/e}$. Let $S(M, F)$ be the minimum number of states to do permutation routing over all λ-routing networks with F wavelengths. Then from eqn. (4.14) with $\rho = 1$,

$$\log_2 S(M, F) \geq M \log_2 M - 2M \log_2 F - 1.44M \qquad (5.13)$$

When $F = 1$, eqn. (5.13) agrees with Shannon's bound.

As a consequence of eqn. (5.13), if the number of users is much more than the square of the number of wavelengths available, the number of switches needed in a permutation AON with λ-routing cannot be significantly reduced over the number of switches needed in a conventional circuit switched network with 1 wavelength. Specifically, define the *λ-routing gain* to be the possible reduction in elementary switches due to λ-routing. That is, define G to be

$$G = \frac{\log_2 S(M, 1)}{\log_2 S(M, F)} \qquad (5.14)$$

It is easy to see that a gain of 2 may be achieved if the number of wavelength grows at a rate of $M^{1/4}$. However to achieve a gain of 10, F must grow like $M^{9/20}$. Larger gains require F to grow at a rate rapidly approaching $\sqrt{M}$.

Therefore, for configurable networks WDM combined with wavelength routing and wavelength changing cannot change the order of growth of the number of switches. However it may be possible to reduce the number of switches by a factor. For instance, with $\approx 10^8$ users and 1000 wavelengths, wavelength routing could possibly reduce the number of switches by about a factor of 5. Therefore, even in very large networks, wavelength routing may reduce hardware cost and switching control complexity.

Chapter 6

Conclusions on Scalability

In this part of the thesis we considered both configurable and passive non-blocking networks where each session requires a full wavelength of bandwidth. The emphasis has been on passive connectors but configurable networks and other traffic demands have been introduced.

The most important conclusion is that if the number of active sessions far exceeds the square of the number of wavelengths, passive networks cannot provide reasonable user connectivity without contention. We will next see in Part II that the same conclusion holds for blocking networks with small blocking probability. Therefore, passive WDM networks are not scalable. Furthermore, we will see in Part III that this conclusion continues to hold even if each session requires only a fraction of the wavelength.

We also showed that there are no good simple networks, i.e. networks where pairs of users are connected on only one wavelength [Chapter 3]. The same conclusion continues to hold when a small amount of blocking is tolerated [Part II]. However in WDM/TDM networks the conclusion may not hold when $T >> 1$, or equivalently $R_s << R$. The reason is clear, pairs of users are connected on T frequency time slots even though they are connected on only one wavelength. Simple networks will play an essential role in building channel efficient WDM/TDM networks [Part III].

6.1 Open Problems

We have shown that $\Theta(\sqrt{M \log M})$ wavelengths are sufficient to build a wide-sense non-blocking connector without wavelength changing. This is the best known bound. An interesting open question[1] is whether wavelength changing or the ability to rearrange sessions can reduce this bound. We also showed that $\lceil \frac{M}{2} \rceil$ wavelengths are required for a LT-connector to be strict sense non-blocking. It is not known if relaxing the LT restriction is beneficial.

In addition, the most wavelength efficient networks are constructed probabilistically. An important open problem is to explicitly construct a wavelength efficient network.

All the networks considered in this part of the thesis were Light Tree and there is a factor of $\Theta(\sqrt{\log M})$ wavelengths between the lower and upper bounds for passive networks. This raises the question of whether NLT-connectors are more wavelength efficient than LT-connectors. This is an open question, but we will see in Part III that NLT-AONs are more time slot efficient than LT-AONs. Note that to prove that passive connectors can be built in $\Theta(\sqrt{M \log M})$ wavelengths, we made use of Hall's Matching Theorem and the Expander Matching Theorem. It is unclear how to use these theorems, or some variant, on networks which are not Light Tree. Since a LT-AON is equivalent to a depth 2 interconnection network, finding NLT-AONs which are more efficient than LT-AONs would imply that the it may be beneficial to relax the self-imposed restrictions of interconnection networks and consider more general switching networks.

[1]Not a very practical one given the results to come in Part II.

Part II

Blocking WDM Networks

Chapter 7

Introduction

Since the construction of wavelength efficient non-blocking connectors has been so elusive and since a small amount of blocking in a network is usually tolerable, we turn our attention and effort to networks with blocking. In this case, progress is much more encouraging but less complete: specifically, we only consider passive networks.

Let $s_1, ..., s_{\rho M}$ be a random list of session *requests* where session s_i is requested before session s_{i+1} and where $\phi = \{s_1, ..., s_{\rho M}\}$ is a traffic without multi-point connections. The random requests are such that given ρM, all traffics and all orderings of the traffics are equally likely.

Each session is *honored* or *blocked* according to some *routing strategy.* We consider two types of strategies: *sequential* or *non-sequential.* In a sequential strategy, the decision to honor or block request i is made without knowledge of requests $i+1, i+2, ...\rho M$. If i is honored, it must be assigned an OD channel that does not contend with the honored requests proceeding i. In addition, the previously honored requests cannot be reassigned OD channels in a sequential strategy. In a non-sequential strategy, the network waits until the last request before deciding which requests to honor or block. These two modes are not all inclusive and we could imagine networks operating in a partially sequential strategy where the network waits a few requests after request i before deciding whether to route or block s_i.

Two different but related performance measures are considered. The first is the *failure probability* P_f, defined to be the probability that one or more requests cannot be honored. The second is the *blocking probability* P_b, defined to be the expected fraction of requests that are blocked. Note that a rearrangeably non-blocking connector is a network with $P_b = P_f = 0$ when $\rho = 1$ and when operated with an appropriate non-sequential strategy and that a wide-sense non-blocking connector is a network with $P_b = P_f = 0$ when $\rho = 1$ and when operated with an appropriate sequential strategy.

For both these performance measures, we present lower bounds on the number of required wavelengths for a network operating under any strategy and then construct networks which operate with a particularly simple sequential strategy. The constructions do not require wavelength changing.

First in Chapter 8, we focus on *almost-all connectors*, i.e. networks with a small failure probability P_f under heavy load $\rho = 1$. We'll see that except for $P_f \approx 1$, at least $\sqrt{M/e}$ wavelengths are required for any passive network; the same as in the non-blocking case. We then explicitly construct sequential networks which have very small P_f using $\Theta(\sqrt{M \log M})$ wavelengths; the same number as the "probabilistically designed" networks in Chapter 5. The constructions are called LAN-LRs since they consist of broadcast LANs connected to a Latin Router backbone.

We then turn our attention to networks with small blocking probability P_b. The lower bounds are essentially the same as the non-blocking and almost-all cases. In Chapter 9 we will show that at least

$$F \geq \left(\sqrt{\frac{\rho M}{e}}\right)^{1-P_b} \left(1 + O\left(\frac{\ln \rho M}{\rho M}\right)\right) - 1. \tag{7.1}$$

wavelength are needed.

Then in Chapter 10, we show that the LAN-LR network only needs $c\sqrt{\rho M}$ wavelengths where c is a constant which depends on P_b and is between 6.5 and 9.2 for P_b

between 10^{-3} and 10^{-6}. Therefore unlike the non-blocking and almost-all case, we have essentially met the theoretical lower bound.

Chapter 8

Almost-All Permutation Routing

Let $s_1, ..., s_M$ be a random list of requests where all lists without multi-point connections are equally likely. We will investigate the number of wavelengths required to keep the failure probability P_f small. Clearly at least $\sqrt{M/e}$ wavelengths are required and $\Theta(\sqrt{M \log M})$ are sufficient when $P_f = 0$ [Chapter 5].

First we show that except for $P_f \approx 1$, the ability to block the list of requests does not reduce the lower bound [Section 8.1]. Then we describe the *LAN-LR* networks [Section 8.2]. The LAN-LR is a two level hierarchical network. The bottom layer consists of N broadcast local area networks with b users per LAN; $M = Nb$ is the total number of users as always. The upper layer interconnects the LANs through an $N \times N$ Latin Router (LR) [Section 1.5]. Each pair of LANs are connected on k wavelengths, where recall that k is called the coarseness of the LR. The WDM cross-connect ($k = 1$) is shown to be very wavelength inefficient in Section 8.3; however with proper choice of the number of wavelengths connecting any input to any output, the LAN-LR is very wavelength efficient [Section 8.4].

8.1 Lower Bound

To derive the lower bound, we first need a definition.

Definition 5 Combinatorial Power

Given a passive AON H, let $\mathcal{T}_M$ be the permutation traffics that H can support. Following Beneš [Bén65], we define the combinatorial power r *of a network as the fraction of the number of permutations supported. In particular,*

$$r \stackrel{\text{def}}{=} \frac{|\mathcal{T}_M|}{M!} \tag{8.1}$$

so that r is the probability that a random permutation is supported by the network when all permutations are equally likely.

Consider an arbitrary passive AON and a list of requests $s_1, ..., s_M$ for the network. If $\{s_1, ..., s_M\}$ is not supported by the network then any routing strategy must fail. Therefore the failure probability is at least $P_f \geq 1 - r$. The following two bounds are now trivial.

Theorem 13 Almost-All Lower Bounds

1) A passive AON with failure probability P_f requires at least

$$F \geq \left[(1 - P_f)\sqrt{2\pi M}\right]^{\frac{1}{2M}} \sqrt{\frac{M}{e}} \tag{8.2}$$

wavelengths.

2) If $P_f \leq 1 - \frac{1}{\sqrt{2\pi M}}$ then at least $\sqrt{M/e}$ wavelengths are needed.

Proof: Since the number of feasible tuning states must be no less than the number of traffics and since there are $rM! \geq (1 - P_f)M!$ permutation traffics,

$$F^{2M} \geq (1 - P_f)M! \geq r\sqrt{2\pi M}(M/e)^M \tag{8.3}$$

which proves statement 1). Now if $P_f \leq 1 - \frac{1}{\sqrt{2\pi M}}$, then

$$(1 - P_f)\sqrt{2\pi M} \geq 1 \tag{8.4}$$

This combined with eqn. (8.2) shows that in this case at least $\sqrt{M/e}$ wavelengths are required. □

8.2 LAN-Latin Router Network

The networks considered have N LANs, with b users per LAN. We think of each LAN consisting of two parts: a T-LAN containing only the transmitters and the R-LAN consisting of the receivers. Number the LANs from 0 to $N-1$ and use the notation $[x, y]$ to stand for a T-LAN, R-LAN pair. Each $[x, y]$ is called a *block* for reasons that will become apparent. Each transmitter in T-LAN x is connected to each receiver in R-LAN y on the wavelengths $J[x, y]$, where [1]

$$J[x,y] = \{f \mid \left\lfloor \frac{f}{k} \right\rfloor \equiv y - x \pmod{N}\} \tag{8.5}$$

The connection matrix of the network is given by $H(n, m) = J[x_n, y_m]$ where x_n is the T-LAN of transmitter n and y_m is the R-LAN of receiver m. The connection matrix H is shown in Fig. 8-1 and a possible implementation is shown in Fig. 8-2.

Note that each transmitter is connected to each receiver on k wavelengths and recall that k is called the *coarseness* of the network. The *backbone* network $J[x, y]$ is an example of a Latin Router (LR) [Section 1.5]. In a Latin Router, all $F = kN$ wavelengths can be simultaneously applied to each of the N inputs without any output contention providing a throughput of $\frac{F^2}{k}$ wavebands, i.e. simultaneous connections. When $k = 1$, a possible implementation of a LR is the well known

[1] Any LR with coarseness k will perform equivalently to J. We use eqn. (8.5) only as a concrete example.

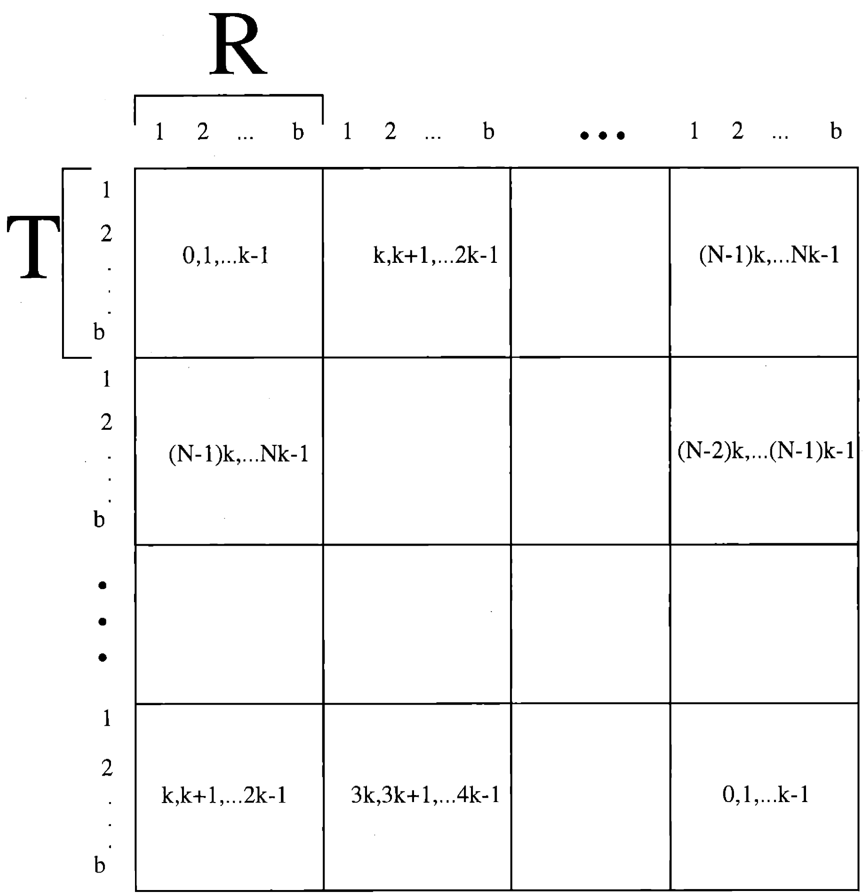


Figure 8-1: LAN-LR connection matrix

WDM cross-connect.

We call the network a *LAN-LR*. Note that the connectivity of a LAN-LR is completely determined by N, b and k. The total number of users is $M = Nb$ and the total number of wavelengths is $F = Nk$. Since the maximum throughput of a LAN-LR is $\frac{F^2}{k}$ wavebands, it seems reasonable at first thought to only consider the WDM cross-connect ($k = 1$) since this maximizes the maximum throughput. However, the logic of maximizing the maximum throughput starts to break down as F gets large since for $F > \sqrt{M}$ and $k = 1$, the maximum throughput is larger than the achievable

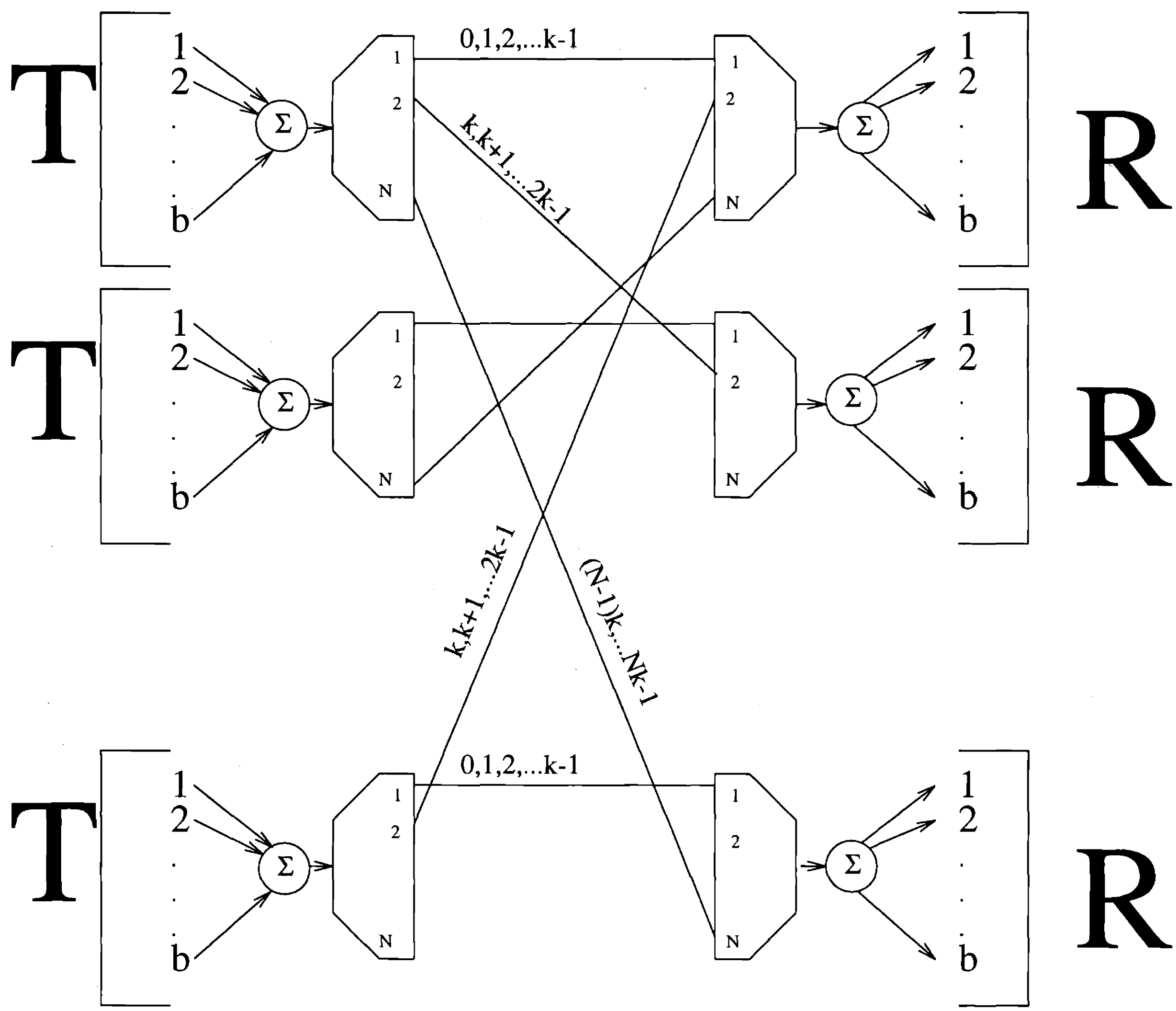


Figure 8-2: The LAN-LR network

throughput of M. Since $F > \sqrt{M/e}$ is a necessary condition, this case cannot be discounted. We will show below that the cross-connect is a poor choice if the goal is to minimize P_f unless $F \approx \sqrt{M}$. However, in this case $P_f \approx 1$. In Chapter 10, we show that the same conclusion holds if the goal is to minimize P_b, the blocking probability, and that this conclusion continues to hold for $\rho < 1$.

Before proceeding, note that the LAN-LR requires M wavelengths to be non-blocking since all b transmitters in a T-LAN could request all b receivers in an R-LAN and $F = Nk$. Therefore, the LAN-LR is a very wavelength inefficient **inefficient** non-blocking connector. However, we will see below that for proper choice of N, b, and k, the LAN-LR is an **efficient** almost-all connector. Specifically, for any $P_f > 0$, no more than $\Theta\left(\sqrt{M \log M}\right)$ wavelengths are required.

8.3 WDM Cross-Connect

We'll show that for P_f not to approach 1 rapidly, the cross-connect (i.e. $k = 1$) requires at least $\Theta(M^{2/3})$ wavelengths. Since there are networks with $P_f = 0$ and $\Theta\left(\sqrt{M \log M}\right)$ wavelengths, the cross-connect is a very wavelength inefficient design.

Notice that $P_s \stackrel{\text{def}}{=} 1 - P_f$ is no more than the probability that the first b sessions can be routed. Let P_s' be that probability. The total number of ways of matching the first b receivers to b of the M transmitters is $(M)_b$ since there are M ways to match the first transmitter, $M - 1$ ways to match the second, etc.[2] The first b sessions can be honored iff each falls in a different block. It follows that the total number of successful ways to match the first b transmitters to b of the M receivers is $M(M-b)(M-2b)...(M-(b-1)b)$ since there are M ways to successfully match transmitter 1, $M - b$ ways of successfully matching transmitter 2, etc. Therefore, P_s

[2] $(M)_b$ is read M lower factorial b. For a description of the lower factorial function, see Appendix A.

is upper bounded by

$$\begin{aligned} P_s &\leq P_s' = \frac{M(M-b)...(M-(b-1)b)}{(M)_b} = \frac{(M-b)...(M-(b-1)b)}{(M-1)_{b-1}} \\ &\leq \prod_{i=1}^{b-1}\left(\frac{M-ib}{M-b}\right) = \prod_{i=1}^{b-1}\left(1-\frac{(i-1)b}{M-b}\right) \\ &\leq \prod_{i=1}^{b-1}\exp\left\{-\frac{(i-1)b}{M-b}\right\} \end{aligned} \tag{8.6}$$

where the last inequality follows from $1-x \leq e^{-x}$. Therefore

$$P_s \leq \exp\left\{-\sum_{i=1}^{b-1}\frac{(i-1)b}{M-b}\right\} = \exp\left\{-\frac{b(b-1)(b-2)}{M-b}\right\} \tag{8.7}$$

So in order for P_s not to go to 0, $b^3 = \frac{M^3}{F^3}$ must grow slower than M, or equivalently F must grow at least as fast as $M^{2/3}$.

8.4 Performance Analysis

Relax the cross-connect restriction and consider LAN-LRs with k wavelengths per block. We operate the network in a sequential mode using a greedy strategy: if s_i falls in block $[x, y]$, then the i^{th} request is honored iff less than k previous requests have been honored in $[x, y]$. The failure probability P_f is the probability that more than k requests fall in some block. Since there are N LANs,

$$P_f \leq N^2 \cdot P(i \geq k+1) \tag{8.8}$$

where $P(i)$ is the probability of i requests in an arbitrary block $[x, y]$ and where we have used the union bound.

Now let's derive $P(i)$. Since there are b transmitters in T-LAN x requesting a session from one of M receivers, there are $\binom{M}{b}$ ways to pick the b requested receivers. Out of these ways there are $\binom{b}{i}\binom{M-b}{b-i}$ ways that i receivers in R-LAN y will be requested from T-LAN x. Therefore, the *request distribution* is given by

$$P(i) = \frac{\binom{M}{b-i}\binom{M-b}{b-i}}{\binom{M}{b}} \tag{8.9}$$

for $0 \leq i \leq b$. The distribution is known as a hypergeometric distribution and in order to bound the failure probability we need to bound the tail of $P(i)$. For a description of the hypergeometric random variable and its relationship to the binomial random variable, see Appendix B [Think of the b receiver in R-LAN y as "blue balls", the remaining $M - b$ receivers as "red balls", and the b requests as "picks"].

Let $\lambda \stackrel{\text{def}}{=} M/N^2$ be the expected number of requests in a block. The weight in the tail of a hypergeometric is bounded above by [Appendix B, eqn. (B.40)]

$$\ln P(i \geq k+1) \leq \ln P(i \geq k) \leq (k - \lambda) + k \ln \frac{\lambda}{k} + \frac{b}{M-b} \tag{8.10}$$

for all $k \geq \delta\lambda$, where $\delta = \exp(\frac{1}{N-1}) = 1 + \Theta\left(\frac{1}{N}\right)$. Now using eqn. (8.8), $N = kF$, and $b = M/N = Mk/F$

$$P_f \leq \frac{F^2}{k^2} \cdot e^{\left(k - \frac{Mk^2}{F^2}\right) + k \ln\left(\frac{Mk}{F^2}\right) + \frac{k}{F-k}} \tag{8.11}$$

which is valid for integer F,k,M such that $N = F/k$ and $b = Mk/F$ are integer. Since we are interested in asymptotic results in this chapter, we will relax these integer constraints.

In order to get the best asymptotic results, we optimize eqn. (8.11) over k for a fixed M and F. To that end, let $p \stackrel{\text{def}}{=} \frac{F}{\sqrt{M}}$ be a measure of the wavelength efficiency of the network. Re-writing eqn (8.11) in a more convenient form,

$$P_f \leq \frac{Mp^2}{k^2} \cdot e^{g(k) + \frac{k/p}{\sqrt{M} - k/p}} \tag{8.12}$$

$$g(k) = \left(k - \frac{k^2}{p^2}\right) + k \ln \frac{k}{p^2} \tag{8.13}$$

which is valid for $k \geq \delta\lambda = \delta k^2/p^2$.

Since the bound on P_f is strongly dependent on $g(k)$ and weakly dependent on the other terms, we will minimize $g(k)$ over k to approximate the optimal value of k. We are interested in k in the range $k < p^2$ since in order to keep the failure probability low, $k > \lambda = k^2/p^2$. $g(k)$ is strictly negative in the range $0 < k < p^2$ and has a unique minimum at $k \approx .2p^2$. At this value $k = .2p^2 >> \lambda = k^2/p^2 = .04p^2$ so that the bound in eqn. (8.12) is valid. Therefore since $g(.2p^2) \approx -.16p^2$

$$P_f \;\leq\; \frac{25M}{p^2} \cdot e^{-.16p^2 + \frac{.2p}{\sqrt{M}-.2p}}. \tag{8.14}$$

From eqn. (8.14) it is easy to see that for fixed $P_f > 0$, p need not grow faster than $\Theta(\sqrt{\log M})$. Specifically, let $p = \sqrt{\frac{\ln M}{.16}} = 2.5\sqrt{\ln M}$ so that $k = .2p^2 = 1.25 \ln M$. Eqn. (8.14) becomes

$$\begin{aligned} P_f \;&\leq\; \frac{.16 \cdot 25M}{\ln M} \cdot e^{-\ln M + \frac{\sqrt{\ln M}}{2\sqrt{M}-\sqrt{\ln M}}} \\ &= \frac{4}{\ln M} \cdot \left(1 + O\left(\frac{\ln M}{\sqrt{M}}\right)\right) \end{aligned} \tag{8.15}$$

Therefore since $F = p\sqrt{M}$, an almost-all connector with any $P_f > 0$ asymptotically requires no more than $2.5\sqrt{M \log M}$ wavelengths.

Chapter 9

Lower Bounds for Networks with Blocking

In the previous chapter we considered networks which with high probability routed **all** of the requests. We now consider networks which with high probability route **most** of the ρM requests from a random list of requests.

First we give lower bounds on the number of wavelengths needed in a passive AON that has some blocking. The first section shows that for sufficiently small P_b at least $\sqrt{\rho M/e}$ wavelengths are required. The next section shows that for any P_b,

$$F \geq \left(\sqrt{\frac{\rho M}{e}}\right)^{1-P_b} \left(1 + O\left(\frac{\ln \rho M}{\rho M}\right)\right) - 1 \approx \sqrt{\frac{\rho M}{e}} \tag{9.1}$$

wavelengths are needed. In the next chapter, we will show that the LAN-LR requires about this many wavelengths.

A few preliminaries will simplify the coming proofs. Consider any AON with connectivity matrix H and let $\mathcal{T}(H)$ be the set of traffics H can support. Also let $s_1, s_2, ..., s_{\rho M}$ be a list of session requests where session s_i is requested before session s_{i+1}. Recall that each session is either honored or blocked according to some strategy. If the decision whether to block the i^{th} request s_i, is always made before the

$(i+1)^{th}$ request, we say the strategy is sequential. Also by definition, OD channel rearranging is not permitted in a sequential strategy. If the strategy makes use of all the information, it is non-sequential. An *optimal* strategy is one which always honors the maximal number of requests possible. Formally, an optimal strategy is one which always honors

$$q_\phi = \max\{|\phi'| \ : \phi' \subseteq \phi \quad \text{and} \quad \phi' \in \mathcal{T}(H)\} \tag{9.2}$$

sessions where $\phi = \{s_1, ..., s_{\rho M}\}$. Note that the optimal strategy depends on the network.

The optimal blocking probability is the probability that a request picked at random from a traffic picked at random is blocked under an optimal strategy. The expected number of honored sessions under an optimal strategy is

$$E[q] = \frac{1}{\binom{M}{\rho M}^2 (\rho M)!} \sum_\phi q_\phi \tag{9.3}$$

where the sum is taken over all ρ-permutations. The optimal blocking probability is

$$P_b = \frac{\rho M - E[q]}{\rho M} \tag{9.4}$$

The lower bounds of the next two sections will be proved by assuming an optimal strategy and then proving the lower bound when $\rho = 1$. Then using the following lemma, we can simply substitute ρM for M when $\rho < 1$.

Lemma 14 *Let $F(M, \rho, P_b)$ be the minimum number of wavelengths for a network with M users, ρM requests, and optimal blocking probability P_b. Then $F(M, \rho, P_b) \geq F(\rho M, 1, P_b)$.*

Proof: Consider an arbitrary passive AON operating with an optimal strategy and a blocking probability P_b. Conditioning on the set of transmitters and set of receivers requesting sessions, the optimal blocking probability can be written as the expected blocking probability given a set of transmitters and a set of receivers, i.e.

$$P_b = \frac{1}{\binom{M}{\rho M}^2} \sum_{Tran,Rec} P_b(Tran, Rec) \tag{9.5}$$

where the sum is taken over all sets of ρM transmitters and all sets of ρM receivers and where $P_b(Tran, Rec)$ is the optimal blocking probability given $Tran$ and Rec. Pick a $(Tran, Rec)$ with $P_b(Tran, Rec) \leq P_b$; there must be at least one. Now form a new network with the ρM transmitters in $Tran$ and the ρM receivers in Rec with the same wavelength connectivity between these users as the original network. Then at least $F(\rho M, 1, P_b(Tran, Rec))$ wavelengths are required for this new network. Since $P_b(Tran, Rec) \leq P_b$, the result follows. □

9.1 Very Small P_b

The case when P_b is very small is easily shown. This is done in the following theorem.

Theorem 15 Lower Bound for Small P_b

1) A passive AON with blocking probability P_b requires at least

$$F \geq \left[(1 - \rho M P_b)\sqrt{2\pi\rho M}\right]^{\frac{1}{2\rho M}} \sqrt{\frac{\rho M}{e}} \tag{9.6}$$

wavelengths.

2) If $P_b \leq \frac{1}{\rho M} - \frac{1}{\rho M\sqrt{2\pi\rho M}}$, then at least $\sqrt{\rho M/e}$ wavelengths are needed.

Proof: We first do the $\rho = 1$ case. Let H be any network operating with an optimal strategy and let $\mathcal{T}_M(H)$ be the set of permutation traffics that H can support. By the union bound, the failure probability $P_f \leq MP_b$ and therefore $\mathcal{T}_M(H)$ contains at least $M!(1-P_f) \geq M!(1-MP_b)$ permutation traffics. Since the number of feasible tuning states must be no less than the number of traffics,

$$F^{2M} \geq (1 - MP_b)M! \geq (1 - MP_b)\sqrt{2\pi M}(M/e)^M \tag{9.7}$$

where the second inequality follows from Sterling's formula [Appendix A]. This proves statement 1) when $\rho = 1$. Now if $MP_b \leq 1 - \frac{1}{\sqrt{2\pi M}}$, then

$$(1 - MP_b)\sqrt{2\pi M} \geq 1 \tag{9.8}$$

This combined with eqn. (9.7) shows that in this case at least $\sqrt{M/e}$ wavelengths are required for $\rho = 1$.

For $\rho < 1$, use the lemma to prove 9.6. Statement 2) follows exactly as before replacing M by ρM. □

9.2 Moderate P_b

If $P_b \leq \frac{1}{\rho M} - \frac{1}{\rho M\sqrt{2\pi\rho M}}$, then the expected number of blocked requests is less than 1. Therefore, the fraction P_f of times not all requests can be supported must be small. This fact proved the last theorem. Here we treat the case of larger P_b. We are particularly interested in the case when $E[q] = (1 - P_b)\rho M >> 1$.

The next theorem is based on the intuition that if we want $E[q]$ to be large, the network must support a large number of traffics. We proceed as before: assume an optimal strategy, bound F when $\rho = 1$, and then replace M by ρM in the bound.

Theorem 16 Lower Bound for Moderate P_b

$$F \geq \left(\sqrt{\frac{\rho M}{e}}\right)^{1-P_b}\left(1 + O\left(\frac{\ln \rho M}{\rho M}\right)\right) - 1 \tag{9.9}$$

wavelengths are needed for a passive AON with blocking probability P_b.

Proof: Let H be any passive AON and $\mathcal{T} = \mathcal{T}(H)$ be the traffic set the network supports. We use a slight abuse of notation and set $|H| = |\mathcal{T}(H)|$ to be the number of traffics supported by the network. Assume an optimal strategy. By the lemma, it is sufficient to prove the theorem for $\rho = 1$.

The first step of the proof is to show that if $E[q]$ is large, then $|H|$ must also be large. Specifically, we first show that for any integer $0 \le k \le M-1$,

$$E[q] \le k + \left(\frac{M}{M-1}\right)\frac{(k+1)|H|}{(M)_{k+1}} \tag{9.10}$$

where recall that $(n)_i = \binom{n}{i} i! = n(n-1)...(n-i+1)$ is the lower factorial function.

The expected number of sessions routed under the optimal strategy is

$$E[q] = \sum_{q=1}^{M} qP(q) \tag{9.11}$$

where $P(q)$ is defined to be the probability that the maximum number of sessions that can be routed is q. Define $P'(q) \ge P(q)$ to be the probability of being able to route q sessions, i.e. the probability that there exists a subset of ϕ of size q that can be supported by the network.

Now, for any k between 1 and $M-1$,

$$\begin{aligned} E[q] &= \sum_{q=1}^{M} qP(q) && (9.12)\\ &\le k + \sum_{q=k+1}^{M} qP(q) && (9.13)\\ &\le k + \sum_{q=k+1}^{M} qP'(q) && (9.14) \end{aligned}$$

Now let $\mathcal{T}_k$ be the set of traffics in $\mathcal{T}$ with exactly k sessions. There are $\binom{M}{q}$ q-subsets of ϕ, i.e. subset of ϕ with size q. If any of these subsets is in $\mathcal{T}_q$, we can route q sessions. The probability that one of these subsets picked at random is in $\mathcal{T}_q$ is

$$\frac{|\mathcal{T}_q|}{\binom{M}{q}^2 q!} \tag{9.15}$$

since there are a total of $\binom{M}{q}^2 q!$ possible q-traffics, i.e. traffics of size q, and if we pick a q-subset of ϕ any q-traffic is equally likely. Using the union bound, the probability of being able to route q sessions is no more than

$$P'(q) \;\leq\; \frac{\binom{M}{q}|\mathcal{T}_q|}{\binom{M}{q}^2 q!} = \frac{|\mathcal{T}_q|}{\binom{M}{q}q!}. \tag{9.16}$$

Therefore, the expected number of sessions that can be routed is upper bounded by

$$E[q] \;\leq\; k + \sum_{q=k+1}^{M} q\frac{\alpha_q|H|}{\binom{M}{q}q!} \tag{9.17}$$

where $\alpha_q = |\mathcal{T}_q|/|H|$ and $\sum_{q=k+1}^{M}\alpha_q \leq 1$. Using lemma 17 in the appendix to this chapter $E[q]$ is no more than

$$E[q] \;\leq\; k + |H|\left(\frac{M}{M-1}\right)\frac{k+1}{(M)_{k+1}} \tag{9.18}$$

which proves eqn. (9.10).

For the second part of the proof, pick $k = \lfloor E[q]\rfloor - 1 \leq M-1$. The second term in eqn. (9.10) above must be at least 1, and solving for $|H|$ this gives

$$|H| \;\geq\; \left(\frac{M-1}{M}\right)\left(\frac{1}{\lfloor E[q]\rfloor}\right)(M)_{\lfloor E[q]\rfloor} \tag{9.19}$$

$$\tag{9.20}$$

Now, using the lower bound on $(M)_x$ presented in the second appendix to this chapter,

$$|H| \geq \left(\frac{M-1}{M}\right)\left(\frac{1}{\lfloor E[q] \rfloor}\right)\left(\frac{\sqrt{2\pi}}{e}\right)\left(\frac{M}{e}\right)^{\lfloor E[q] \rfloor} \tag{9.21}$$

and since $E[q] \geq \lfloor E[q] \rfloor \geq E[q] - 1$,

$$|H| \geq \left(\frac{M-1}{M}\right)\left(\frac{1}{E[q]}\right)\left(\frac{\sqrt{2\pi}}{e}\right)\left(\frac{M}{e}\right)^{E[q]-1} \tag{9.22}$$

Finally, since $E[q] = M(1 - P_b)$, since the number of feasible tuning states can be no less than the number of supported traffics, and since the number of feasible tuning states cannot be more than $(F+1)^{2M}$,

$$(F+1) \geq \epsilon \left(\sqrt{\frac{M}{e}}\right)^{1-P_b} \tag{9.23}$$

where

$$\epsilon \stackrel{\text{def}}{=} \left[\left(1 - \frac{1}{M}\right) \cdot \frac{1}{M(1-P_b)} \cdot \frac{\sqrt{2\pi}}{e} \cdot \frac{e}{M}\right]^{\frac{1}{2M}} \tag{9.24}$$

Now $\ln \epsilon = O\left(\frac{\ln M}{M}\right)$ so that $\epsilon = 1 + O\left(\frac{\ln M}{M}\right)$. □

9.3 Appendix

Here we prove the upper bound referred to above.

Lemma 17 *Let $\{\alpha_i\}_k^n$ be a sequence of non-negative real numbers such that $\sum_{i=k}^n \alpha_i \leq 1$. Then for $n \geq 3$,*

$$\sum_{i=k}^{n} \alpha_i \frac{i}{(n)_i} \leq \left(\frac{n}{n-1}\right) \frac{k}{(n)_k} \tag{9.25}$$

Proof: Let $b_i = \frac{i}{(n)_i}$. The sum is upper bounded by $\bar{b} \stackrel{\text{def}}{=} \max_{k\leq i\leq n} b_i$. We now show that $\bar{b}$ is never more than $n/(n-1)$ times the first term. To see this first notice that $b_k \geq b_{k+1} \geq \ldots \geq b_{n-1}$ since

$$\frac{b_i}{b_{i+1}} = \frac{i}{i+1} \frac{(n)_{i+1}}{(n)_i} \tag{9.26}$$

$$= \frac{i}{i+1}(n-i) \tag{9.27}$$

and $i(n-i)/(i+1)$ is at least 1 for i no more than $n-2$. Also, $\frac{n-2}{(n)_{n-2}} \geq \frac{n}{(n)_n} \geq \frac{n-1}{(n)_{n-1}}$ for all $n \geq 3$. So

$$\max_{k\leq i\leq n} \frac{i}{(n)_i} = \begin{cases} \frac{k}{(n)_k} & k \neq n-1 \\ \frac{n}{(n)_n} & k = n-1 \end{cases} \tag{9.28}$$

and therefore the $\bar{b} = b_k$ unless $k = n-1$. If $k = n-1$ then the maximum is b_n but

$$b_n = \frac{n}{(n)_n} = \left(\frac{n}{n-1}\right) b_{n-1} \tag{9.29}$$

which completes the proof. □

9.4 Appendix

Here we bound $(M)_x$. Using [Gal92a]

$$\sqrt{2\pi x}\left(\frac{x}{e}\right)^x \leq x! \leq e\sqrt{x}\left(\frac{x}{e}\right)^x \tag{9.30}$$

we get

$$\begin{aligned}
(M)_x = \frac{M!}{(M-x)!} \quad &\geq \left[\left(\frac{M}{e}\right)^M \sqrt{2\pi M}\right]\left[\left(\frac{M-x}{e}\right)^{M-x} e\sqrt{M-x}\right]^{-1} && (9.31)\\
&= \left(\frac{M}{M-x}\right)^{M-x+.5}\left(\frac{M}{e}\right)^x \frac{\sqrt{2\pi}}{e} && (9.32)\\
&\geq \left(\frac{M}{e}\right)^x \frac{\sqrt{2\pi}}{e} && (9.33)
\end{aligned}$$

Chapter 10

Latin Routers

In the previous chapter, we essentially showed that at least $\sqrt{\rho M/e}$ wavelengths are required for any network with a small blocking probability. Here we will construct a network with a very small blocking probability using less than $(2.5\sqrt{.8 - \ln P_b})\sqrt{\rho M}$ wavelengths. For P_b between 10^{-3} and 10^{-6}, between $6.5\sqrt{\rho M}$ and $9.2\sqrt{\rho M}$ wavelengths suffice.

The networks are the LAN-LR described in section 8.2. In section 10.1 we bound the blocking probability of a LAN-LR and optimize the number of wavelengths connecting each pair of LANs. The optimization is very similar to the optimization in Section 8.4. In 10.2, we analyze the WDM cross-connect and show that this design performs quite poorly compared with the first design. Conclusions appear in section 10.3.

10.1 Performance

We analyze the blocking probability of a LAN-LR with ρM requests where all lists of requests $s_1, ..., s_{\rho M}$ equally likely. Recall that the LAN-LR network has N local area networks, each with b users. The LANs are interconnected by an $N \times N$ Latin Router with coarseness k [Section 1.5].

Before deriving an expression for P_b, note that for all $\rho \geq \frac{b}{M} = \frac{1}{N}$, the LAN-LR requires M wavelengths to be non-blocking. To see this, note that for $P_b = 0$, k must be at least b since all b transmitters in a T-LAN could request all b receivers in an R-LAN. Therefore since $F = Nk$, non-blocking operation requires $F \geq Nb = M$ wavelengths which makes a LAN-LR a very wavelength **inefficient** non-blocking connector even for very small ρ.[1] However, we will see below that for proper choice of N, b, and k, the LAN-LR is a very wavelength **efficient** blocking connector.

Now let's derive P_b, the expected number of blocked requests. Since there are N^2 blocks and since at most k requests can be honored in any block,

$$P_b = \frac{1}{\lambda} \sum_{i=k+1}^{b} (i-k)P(i) \tag{10.1}$$

where $P(i)$ is the probability of i requests in a block and $\lambda \stackrel{\text{def}}{=} \rho M/N^2$ is the expected number of requests per block. P_b can be rewritten and bounded as

$$\begin{aligned} P_b &= \frac{1}{\lambda}\left\{k \sum_{i=k+1}^{2k} \frac{i-k}{k} P(i) + \sum_{i=2k+1}^{b} (i-k)P(i)\right\} \\ &\leq \frac{1}{\lambda}\left\{k \cdot P(i \geq k+1) + b \cdot P(i \geq 2k+1)\right\} \quad . \end{aligned} \tag{10.2}$$

In the appendix to this chapter, we use eqn. (10.2) to bound P_b when $\rho = 1$, i.e. there are M requests. In this section, we do an approximate analysis for any ρ.

[1] Recall that $F = \Theta(\sqrt{M \log M})$ wavelength suffice for a non-blocking connector; see Chapter 5.

When $\rho = 1$, the results here agree with the exact analysis given in the appendix.

Consider an arbitrary block $[x, y]$ and pick a transmitter n uniformly from the b transmitters in T-LAN x. The probability that n is requesting a session is ρ since the total number of requests is ρM and there are a total of M transmitters. If n requests, the probability that the requested receiver is in R-LAN y is $1/N$ since there are N R-LANs. We make the approximation that the number of requests in each block is binomially distributed with mean λ; specifically let

$$\tilde{P}(i) \;=\; \binom{b}{i}(\rho/N)^i \left(1 - \frac{\rho}{N}\right)^{b-i}. \qquad (10.3)$$

Note that $\rho b/N = \frac{\rho M}{N^2} = \lambda$.

To proceed, we need the following bound on the tail of the binomial

Lemma 18 Binomial Bound

Let u be a binomial random variable with mean $\bar{u} = \zeta n$ and distribution

$$p(u) \;=\; \binom{n}{u}\zeta^u(1-\zeta)^{n-u}.$$

Then for $k \geq \bar{u}$,

$$\ln Pr(u \geq k) \;\leq\; (k - \bar{u}) + k \ln \frac{\bar{u}}{k} \;. \qquad (10.4)$$

Proof: We start with the Chernoff bound [Gal68],

$$\ln Pr(u \geq k) \;\leq\; n\left[\frac{k}{n}\ln\frac{\zeta n}{k} + \left(1 - \frac{k}{n}\right)\ln\frac{1-\zeta}{1-\frac{k}{n}}\right] \;.$$

Note that since $(1-\zeta)/(1-k/n) = 1 + \frac{-\zeta + k/n}{1-k/n}$ and since $\ln(1+x) \leq x$,

$$\ln Pr(u \geq k) \;\leq\; k \ln\frac{\zeta n}{k} + (k - n\zeta) \;.$$

Substituting $\bar{u} = \zeta n$ finishes the proof. □

Using the bound,

$$\tilde{P}(i \geq k+1) \;\leq\; \left(\frac{\lambda}{k+1}\right)^{k+1} \cdot e^{k+1-\lambda} \;. \tag{10.5}$$

Regrouping,

$$\tilde{P}(i \geq k+1) \;\leq\; \frac{\lambda}{k} \cdot \left(\frac{\lambda}{k}\right)^{k} e^{k-\lambda} \cdot \left[e \left(\frac{k}{k+1}\right)^{k+1} \right] \;. \tag{10.6}$$

Now since

$$e \left(\frac{k}{k+1}\right)^{k+1} \;<\; 1 \tag{10.7}$$

we have

$$\tilde{P}(i \geq k+1) \;<\; \frac{\lambda}{k} \cdot \left(\frac{\lambda}{k}\right)^{k} e^{k-\lambda} \;, \tag{10.8}$$

and similarly

$$\tilde{P}(i \geq 2k+1) \;<\; \frac{\lambda}{2k} \cdot \left(\frac{\lambda}{2k}\right)^{2k} e^{2k-\lambda} \;. \tag{10.9}$$

Therefore the blocking probability $\tilde{P}_b$ is no more than

$$\begin{aligned} \tilde{P}_b \;&<\; \left(\frac{\lambda}{k}\right)^{k} e^{k-\lambda} + \frac{b}{2k} \cdot \left(\frac{\lambda}{2k}\right)^{2k} e^{2k-\lambda} \;. \\ &= \left\{ \left(\frac{\lambda}{k}\right)^{k} \cdot e^{k-\lambda} + \frac{b}{2k} \cdot \left(\frac{\lambda}{2k}\right)^{2k} \cdot e^{2k-\lambda} \right\} \\ &= \left(\frac{\lambda}{k}\right)^{k} \cdot e^{k-\lambda} \cdot \left\{ 1 + \frac{b}{2k} \cdot \left(\frac{e\lambda}{4k}\right)^{k} \right\} \end{aligned} \tag{10.10}$$

where we have made use of eqn. (10.2).

This bound can be used to find the minimum k for a given N, b, and desired blocking probability. Here we take a different approach. We use the bound on $\tilde{P}_b$ to determine the best value of k holding the number of wavelengths F and the total number of users $M = Nb$ fixed but allowing the number of LANs N and the number of users per LAN b to vary with k. We will see that for proper choice of k, the blocking probability can be made very small with $c\sqrt{\rho M}$, where c is between 7 and 10. On the other hand, section 10.2 shows that if care is not taken in choosing k, many more wavelengths may be required.

Since the LAN-LR has $M = Nb$ users, and $F = Nk$ wavelengths, we can re-write eqn. (10.10) in terms of M, k and F. Noting that $\lambda = \rho M k^2/F^2$ and $b = Mk/F$,

$$\lambda = k^2/c^2 \tag{10.11}$$

$$\tilde{P}_b \leq e^{g(k)} \cdot \left\{ 1 + \frac{\sqrt{M/\rho}}{2c} \cdot \left(\frac{ek}{4c^2}\right)^k \right\} \tag{10.12}$$

$$g(k) = \left(k - \frac{k^2}{c^2}\right) + k \ln \frac{k}{c^2} \tag{10.13}$$

where $c \stackrel{\text{def}}{=} \frac{F}{\sqrt{\rho M}}$ is a measure of the wavelength efficiency of the network and from the last chapter we know that $c \geq 1/\sqrt{e}$. Note that by definition, $k \geq 1$ and in order to keep the blocking probability low, $k \geq \lambda = k^2/c^2$. Therefore, we are interested in k in the range

$$1 \leq k \leq c^2 \tag{10.14}$$

and this implies that

$$k - \frac{k^2}{c^2} > 0 \tag{10.15}$$

$$k \ln \frac{k}{c^2} < 0 \tag{10.16}$$

Since the bound on $\tilde{P}_b$ is strongly dependent on $e^{g(k)}$ and only weekly dependent on the other terms, we will optimize $g(k)$ over $1 \leq k \leq c^2$ and use this value of k

$\tilde{P}_b$	F	$k = \lceil .2c^2 \rceil$
10^{-3}	$6.5\sqrt{\rho M}$	9
10^{-4}	$7.5\sqrt{\rho M}$	12
10^{-5}	$8.4\sqrt{\rho M}$	15
10^{-6}	$9.2\sqrt{\rho M}$	17

Table 10.1: Upper bounds on the required number of wavelengths.

to bound the blocking probability as a function of F and ρM. Note that this is an integer optimization over k and since k will be relatively small, we will not relax the integer constraints on k. As k varies, integers $b = \frac{Mk}{F}$ and N also vary. We ignore the integer constraints on b and N for now; the significance of this will be discussed at the end of this section.

As a first step only, we relax the integer constraints on k and find the value of k which minimizes $g(k)$. The minimum occurs at $k \approx .2c^2$ and at this value $g(.2c^2) \approx -.16c^2$. Since k must be an integer this solution is not strictly valid. Therefore we set $k = \lceil .2c^2 \rceil$ for which

$$\tilde{P}_b \leq e^{g(\lceil .2c^2 \rceil)} \cdot \left\{ 1 + \frac{\sqrt{M/\rho}}{2c} \left(\frac{e}{20} \right)^{.2c^2} \right\} \qquad (10.17)$$

where for simplicity we have relaxed the integer constraints on the error term. Using this expression, we can find the minimum c and hence the minimum $F = c\sqrt{\rho M}$ for any desired blocking probability. The results for various $\tilde{P}_b$ are presented Table 10.1. Note that P_b and c are expressed to 10% accuracy. The third column lists the approximate optimal k in each instance. It is easy to verify that the error terms are negligible and therefore will be dropped from now on.

For most values of interest, we need between $6\sqrt{\rho M}$ and $10\sqrt{\rho M}$ wavelengths. Put another way, for a given F and small P_b, the maximum expected number of sessions the LAN-LR can handle is between $F^2/100$ and $F^2/36$. The results are both encouraging and discouraging. First of all, we have established an order of growth of

$\Theta(\sqrt{\rho M})$ for the number of wavelengths, basically equal to the lower bound. On the other hand, the constants are fairly large for many applications. To see that this is so, imagine a system with 100 wavelengths and a blocking probability of 10^{-3}. Then only $F^2/6.5^2 \approx 236$ expected sessions can be accommodated. However if $F = 1000$, then for the same blocking probability more than $23,600$ expected sessions can be accommodated.

It is difficult to see the trade-off between blocking probability and number of wavelengths in eqn. (10.17) because of the ceiling functions. However the bound can be simplified with only slight increase in the number of required wavelengths for a given $\tilde{P}_b$. To that end, first notice that since $\ln \tilde{P}_b \geq g(.2c^2) \approx -.16c^2$, the bound is more than $1/e$ unless $c \geq 1/\sqrt{.16} = 2.5$. Since we are interested in small $\tilde{P}_b$, we will assume $c \geq 2.5$. The next lemma shows that under this condition, setting $k = \lceil .2c^2 \rceil$ is almost as good as $k = .2c^2$.

Lemma 19

1) If $.2c^2$ is an integer then,

$$\tilde{P}_b \;\leq\; e^{-.16c^2} \tag{10.18}$$

Furthermore, for a given $\tilde{P}_b$, c need not be larger than

$$c \;\leq\; 2.5\sqrt{-\ln \tilde{P}_b} \tag{10.19}$$

2) Now for any $c \geq 2.5$,

$$\tilde{P}_b \;\leq\; e^{-.16c^2+.8} \tag{10.20}$$

Furthermore, for a given $\tilde{P}_b$, c need not be larger than

$$c \;\leq\; 2.5\sqrt{.8 - \ln \tilde{P}_b} \tag{10.21}$$

Proof: To prove the first statement of 1), simply note that $g(.2c^2) \leq -.16c^2$. For the second statement, solve for c.

To prove the first statement of 2), first notice that

$$\begin{aligned} g(\lceil .2c^2 \rceil) &= \lceil .2c^2 \rceil \ln \frac{\lceil .2c^2 \rceil}{c^2} + \lceil .2c^2 \rceil \left(1 - \frac{\lceil .2c^2 \rceil}{c^2}\right) \\ &\leq \lceil .2c^2 \rceil \ln \frac{\lceil .2c^2 \rceil}{c^2} + .8\lceil .2c^2 \rceil \end{aligned}$$

Now since $\lceil .2c^2 \rceil \leq .2c^2 + 1$,

$$g(\lceil .2c^2 \rceil) \leq \lceil .2c^2 \rceil \ln \frac{\lceil .2c^2 \rceil}{c^2} + .16c^2 + .8$$

Finally note that $x \ln \frac{x}{c^2}$ is decreasing with x for $x < c^2/e$. So if $\lceil .2c^2 \rceil \leq c^2/e$,

$$\lceil .2c^2 \rceil \ln \frac{\lceil .2c^2 \rceil}{c^2} \leq .2c^2 \ln .2$$

Lastly, $\lceil .2c^2 \rceil \leq c^2/e$ for all $c \geq \sqrt{e} = 1.64$ which is true by assumption. Therefore for all $c \geq \sqrt{e}$,

$$\begin{aligned} g(\lceil .2c^2 \rceil) &\leq (.16 + .2 \ln .2)c^2 + .8 \\ &\leq -.16c^2 + .8 \end{aligned}$$

which proves the first statement of item 2).

Now for the second statement of 2), solve for c. □

To derive the above results, we relaxed the integer constraints on N and b; let's now address the validity of these approximations. Let c^* be the minimum c for a given $\tilde{P}_b$, e.g. $c^* = 7$ for $\tilde{P}_b = 10^{-3}$. Under the above approximations, $F = c^* \sqrt{\rho M}$ wavelengths suffice for a blocking probability of $\tilde{P}_b$ in a network with M users and utilization ρ.

First since $N = F/k$, we should have restricted ourselves to k which divide F.

But since k divides $c^*\sqrt{\rho M} + a$ for some $a < k$, the number of wavelengths can be increased by at most k, i.e. $F' = c^*\sqrt{\rho M} + k$ wavelengths suffice.

Now imagine that we did not relax the integer constraints on b, but wished to keep the number of users M fixed. We could still use a network with a LR backbone and N LANs, but some LANs will have $\lfloor \frac{Mk}{F} \rfloor$ users and some $\lceil \frac{Mk}{F} \rceil$ users. However, since

$$\left\lfloor \frac{Mk}{F} \right\rfloor = \frac{Mk}{F}\left(1 + O\left(\frac{1}{\sqrt{M}}\right)\right) \tag{10.22}$$

$$\left\lceil \frac{Mk}{F} \right\rceil = \frac{Mk}{F}\left(1 + O\left(\frac{1}{\sqrt{M}}\right)\right) \tag{10.23}$$

the approximations on b are sufficient for our purposes.

10.2 WDM Cross-Connect

Recall that the familiar implementation of a LR with $k = 1$ is the WDM cross-connect. Here we show that cross-connect performs quite poorly compared to the optimal case. The approximate blocking probability is

$$\begin{aligned} \tilde{P}_b &= \frac{1}{\lambda} \cdot \sum_{i=2}^{b} (i-1)\,\tilde{P}(i) \\ &= \frac{1}{\lambda} \cdot \left[\sum_{i=2}^{b} i\tilde{P}(i) - \sum_{i=2}^{b} \tilde{P}(i)\right] \\ &= \frac{1}{\lambda} \cdot \left[\lambda - 1 + \tilde{P}(0)\right] \end{aligned} \tag{10.24}$$

Recall that $\tilde{P}(i) = \binom{b}{i}\left(\frac{\rho}{N}\right)^i\left(1-\frac{\rho}{N}\right)^{b-i}$. Therefore $\tilde{P}(0)$ is

$$\begin{aligned} \tilde{P}(0) &= \left(1 - \frac{\rho}{N}\right)^b \\ &\approx 1 - \frac{\rho b}{N} + \frac{1}{2}\left(\frac{\rho b}{N}\right)^2. \end{aligned} \tag{10.25}$$

Now since $\lambda = \rho b/N = 1/p^2$, $\tilde{P}_b$ is approximately,

$$\tilde{P}_b \approx \frac{\lambda}{2} = \frac{1}{c^2} \tag{10.26}$$

So for a given blocking probability, about $\sqrt{\tilde{P}_b^{-1}} \cdot \sqrt{\rho M}$ wavelengths are required. For instance, $\tilde{P}_b = 10^{-3}$ requires about $31.6\sqrt{\rho M}$ wavelengths, a factor of 5 more than the optimal design.

10.3 Conclusions

The results are both encouraging and discouraging. First of all we have established an order of growth of $\sqrt{\rho M}$ for the number of wavelengths, basically equal to the lower bound. On the other hand, the constants are fairly large for many applications. We also saw that the WDM Cross-Connect is only optimal if the number of wavelength is small, i.e. $F \leq \sqrt{5\rho M}$. However, in this case the blocking probability is large.

10.4 Appendix

Here, we bound the blocking probability for the LAN-LR when $\rho = 1$ without making the binomial approximation. In particular, we show that

$$\tilde{P}_b \leq e^{g(k)} \cdot \left\{1 + \frac{\sqrt{M}}{2c} \cdot \left(\frac{ek}{4c^2}\right)^k\right\} \cdot \left(1 + O\left(\frac{1}{N}\right)\right) \tag{10.27}$$

where $g(k)$ is as before. Therefore for $\rho = 1$, the binomial approximation gives the same answer as the hypergeometric.

In Chapter 8, the request distribution was derived. Recall that

$$P(i) = \frac{\binom{M}{b-i}\binom{M-b}{b-i}}{\binom{M}{b}} \tag{10.28}$$

and that

$$\ln P(i \geq q) \;\leq\; (q-\lambda) + k\ln\frac{\lambda}{k} + \frac{b}{M-b} \qquad (10.29)$$

Since $M = Nb$, $\ln P(i \geq q) \leq (q-\lambda) + k\ln\frac{\lambda}{k} + \frac{1}{N-1}$. Now the claim follows by comparing this bound to eqn. (10.4) and using eqn. (10.2) as before.

Chapter 11

Conclusions

In this part of the thesis we considered passive networks where each session requires a full wavelength of bandwidth and where a small amount of blocking is tolerated.

The most important conclusion is that these networks are not scalable. Specifically, if the expected number of session requests far exceed the square of the number of wavelengths, passive WDM networks cannot provide reasonable user connectivity without contention or without a high probability of blocking.

Unlike the case of non-blocking connectors, we have been able to construct wavelength efficient blocking networks. Specifically, for a blocking probability of P_b, the minimum number of wavelengths is somewhere between

$$\left(1 + O\left(\frac{\ln^2}{M}\right)\right)\left(\sqrt{\frac{M}{e}}\right)^{1-P_b} \le F \le 2.5\sqrt{.8 - \ln P_b} \cdot \sqrt{M} \qquad (11.1)$$

The upper bound is achieved by the LAN-LR network. This network has recently been proposed as part of a larger Wide Area AON [A+93].

We also saw that the LAN-LR with $O(\sqrt{M} \log M)$ wavelengths supports almost-all permutations. This is asymptotically the same number of wavelengths as the best existence proofs for non-blocking connectors. The advantage here is that we have been able to explicitly construct the networks without wavelength changing. In addition, the algorithm used to assign wavelengths to sessions in a LAN-LR is trivial whereas the existence proofs of Chapter 5 do not address this important control issue.

11.1 Open Problems

There are two significant open questions. First, are there better topologies than the LAN-LR when $F < 7\sqrt{\rho M}$. Most likely the answer is yes. The problem with the LAN-LR is that if P_b is small, the k^{th} wavelength in each block is rarely used. This implies an inefficiency in wavelength re-use. However, it may turn out that more efficient networks require non-sequential operation; this also is an open question.

Second, what are good configurable topologies? One possibility is to use constructions like Fig. 5-4 of Chapter 5 on the input side and the equivalent constructions on the output side to build a configurable network with F wavelengths and b^{2M} switching states which "looks" like a passive network with Fb wavelengths. However this design is impractical and would throw away one of the important advantages of AONs, the ability to organize users into broadcast LANs. Another approach has been taken by Karol [Kar92]; however configurable wavelength changers are required. It is clear that this topic deserves more attention.

Part III

WDM/TDM Networks

Chapter 12

Time Slots and Wavebands

We now consider WDM/TDM networks which support sessions that do not require a full wavelength of bandwidth. As always, the networks have M users, F wavelengths, and S switching states. A wavelength supports $= R$ b/s and a session requires $R_s = R/T$ b/s. For simplicity we assume that T is an integer and that each wavelength is divided up into T periodic time-slots. Each frequency time slot is called a *channel.* A session therefore requires one channel of bandwidth.

Of course TDM is only one way to share bandwidth; consider instead that each wavelength is further divided into T wavelengths for a total of $F \cdot T$ wavelengths. In this way, a WDM/TDM network with F wavelengths and T time slots is analogous to a WDM networks with $F \cdot T$ wavelengths where the λ-nodes and π-nodes are restricted to operate solely on *wavebands.* In particular, the T wavelengths of a waveband are always routed together through a λ-node and a π-node changes all the signals in a waveband to another waveband but cannot operate on individual wavelengths. Similar analogies can be drawn to networks using CDMA or subcarrier multiplexing to share the wavebands.

It should be clear that there is more freedom in designing a network with $F \cdot T$ wavelengths than in designing a network with F wavelengths and T time slots per wavelength. Therefore, if F wavelengths are required in an AON when each session requires a wavelength, at least F/T wavelengths are required when each session only requires $1/T$ wavelengths. From eqn. (4.4), at least

$$F \cdot T \;\geq \left(\frac{|\mathcal{T}_d|}{S}\right)^{\frac{1}{2d}} \binom{M}{d}^{-\frac{1}{d}} \tag{12.1}$$

channels are required to support $\mathcal{T}$ where $\mathcal{T}_d$ are the traffics in $\mathcal{T}$ with d active sessions. All the lower bounds on wavelengths developed in Parts I and II immediately generalize to lower bounds on channels. For instance, a non-blocking partial connector requires $F \cdot T \geq \sqrt{\rho M/e}$ channels and a blocking network requires at least $\left(\sqrt{\rho M/e}\right)^{1-P_b} (1 + O\left(\frac{\ln \rho M}{\rho M}\right)) - 1$ channels. However, since λ-routing AONs only have to ability to route on the basis of wavelengths but not on the basis of channels, eqn. (12.1) seems optimistic.[1]

Now consider the following straightforward way of building a WDM/TDM network from a WDM network. Let $\mathcal{H}$ be an $M \times M$ connector with F wavelengths, S states, and 1 time slot. Connect T users as a *group* to each input and each output of $\mathcal{H}$. Then by time sharing, we have an $MT \times MT$ connector with F wavelengths, S states, and T time slots. To see this, use Hall's Theorem to decompose each traffic of the MT users into T traffics of the M groups. [2] Each group traffic is supported by the $M \times M$ connector so the new network is non-blocking in T time slots. Since passive WDM connectors can be built with $\Theta(\sqrt{M \log M})$ wavelengths, passive WDM/TDM connectors can be built with $\Theta(\sqrt{\frac{M}{T} \log(M/T)})$ wavelengths.

Using the same technique and the results in Chapter 10, a WDM/TDM network can be built with $\Theta(\sqrt{\rho M/T})$ wavelengths if a small probability of blocking P_b is

[1]We could imagine channel routing networks built from channel routing and channel changing nodes. For instance a channel Latin Router can be built from a cascade of wavelength Latin Routers and time slot Latin Routers [BH93a]. Then using the results in Chapter 10, eqn. (12.1) would in fact be tight.

[2]For each traffic on MT users form an $M \times M$ matrix where the $(n,m)^{th}$ entry is the number of requests between input group n and output group m. Each row and column of the matrix sums to T so by Hall's Theorem the matrix can be written as a sum of T permutation matrices [Ber92]. Each matrix represents a traffic on the groups.

tolerated. Therefore, the maximum load on a passive network lies somewhere between

$$\Omega\left(F^2T\right) \leq \rho M \leq O\left((FT)^{\frac{2}{1-P_b}}\right). \tag{12.2}$$

Note that since we are interested in P_b very small, we will approximate the upper bound by $O(F^2T^2)$ from now on. Closing this gap is the topic of the next two chapters. LT-AONs are covered in Chapter 13. NLT-AONs are treated in Chapter 14.

Chapter 13

Light Tree AONs

In this Chapter, we improve the bound in eqn. (12.1) by a factor of $\sqrt{T}$ for LT-AONs. Equivalently, we show that at most $O(F^2T)$ active sessions can be supported by a passive LT-AON. Two different proofs will be given: the Tuning State Bound and the Light Tree Bound. As the name implies, the Tuning State bound is analogous to Theorem 2. It holds for any traffic set and can be used for configurable networks. The Light Tree Bound holds for passive networks under an important class of traffics called *homogeneous* traffics. For these networks and traffics, the Light Tree Bound is tighter than the Tuning State Bound in terms of the constant in front of the F^2T term.

13.1 Equivalent Tuning States

Recall that $\mathcal{T}_d$ are the traffics in $\mathcal{T}$ with exactly d sessions and that at least

$$F \cdot T \geq \left(\frac{|\mathcal{T}_d|}{S}\right)^{\frac{1}{2d}} \binom{M}{d}^{-\frac{1}{d}} \tag{13.1}$$

channels are required to avoid contention. Therefore a passive partial connector requires $F \geq \frac{1}{T}\sqrt{\rho M/e}$ wavelengths. However, since λ-routing AONs only have to ability to route on the basis of wavelengths but not on the basis of time slots, this bound is optimistic. Specifically we show:

Theorem 20 *LT-AON Lower Bound A necessary condition for a λ-routing network to support a traffic set $\mathcal{T}$ without contention is*

$$F\sqrt{T} \;\geq\; \frac{1}{\sqrt{e}}\left(\frac{\sqrt{2\pi}}{e}\cdot\frac{|\mathcal{T}|}{S}\right)^{\frac{1}{2d}}\binom{M}{d}^{\frac{1}{d}} \tag{13.2}$$

where F is the number of wavelengths, T is the number of time slots, and S is the number of switching states.

Since $(\sqrt{2\pi}/e)^{1/2d} \approx 1$, a LT-partial connector requires $F \geq \frac{1}{e}\sqrt{\frac{\rho M}{T}}$ wavelengths and the Santa Barbara Traffics requires $F \geq \frac{1}{e^2}\sqrt{\frac{M}{T}}$ wavelengths. These bounds are a factor of $\sqrt{T/e}$ more pessimistic than the bound in eqn. (12.1).

Let's first update a few definitions from Chapter 4. A tuning state $\mathbf{v}$ for a WDM/TDM network is represented by the $2 \times M$ vector

$$\mathbf{v} = \begin{bmatrix} c_{o,1} & c_{o,2} & \cdots & c_{o,M} \\ c'_{d,1} & c'_{d,2} & \cdots & c'_{d,M} \end{bmatrix} \tag{13.3}$$

where $c_{o,n} = (f_n, t_n)$ is the channel assigned to transmitter n in $\mathbf{v}$ and $c_{d,m} = (f'_m, t'_m)$ is the channel assigned to receiver m in $\mathbf{v}$. A network state $(\mathbf{v}, \phi)$ is a tuning state v and a switching state ψ. A network state is feasible for a traffic if C0) inactive transceivers are off, C1) sessions are connected, and C2) there are no fatal collisions. Notice that since we can think of the time slots as wavelengths within a wavelength, these definitions are exactly the same as in Chapter 4. Therefore no network state can be feasible for more than one traffic and

$$S \cdot \binom{M}{d}^2 (FT)^{2d} \;\geq |\mathcal{T}_d| \tag{13.4}$$

is a necessary condition to avoid contention. Solving for $F \cdot T$ gives eqn. (13.1) which is the WDM/TDM equivalent to eqn. (4.4).

To see that eqn. (13.4) is optimistic when $T > 1$, observe that if $(\mathbf{v}, \psi)$ is a feasible network state for a traffic ϕ, we could exchange the sessions in time t with those in time slot t' and still maintain feasibility. Therefore rescheduling produces a second feasible network state for ϕ.[1] Since a network state cannot be feasible for two traffics, ϕ uses up at least 2 network states and therefore the left hand side of eqn. (13.4) can be reduced by a factor of 2. This is the idea behind the following definition and lemma.

Definition 6 Equivalent Tuning States

For a given AON, two network states are equivalent *if they support the same traffic or neither supports a traffic.*

Notice that the definition depends only on the network and not on the traffic set we wish the network to support. The following lemma formally shows that the idea of equivalent tuning states can be used to reduce the left hand side of eqn. (13.4). The lemma holds for all networks, not just LT-AONs.

Lemma 21 *An AON supports at most V traffics, where V is the number of non-equivalent network states.*

> **Proof:** If the statement is not true, there exist two distinct traffics supported by equivalent network states. But by definition of equivalent, this means that each of the two network states is feasible for the two traffics. Since no network state can be feasible for more than one traffic, this is a contradiction. □

We now show that Light Tree Networks must have a large number of equivalent tuning states. That limits the number of non-equivalent tuning states which will prove eqn. (13.2).

[1] Note that exchanging sessions on wavelength f with those on wavelength f' does not necessarily produce equivalent tuning states because signals are routed on the basis of their wavelength. As a counter example, consider the simple network in Chapter 3.

Lemma 22 *For a given LT-AON, let $\mathcal{V}(\phi)$ be the network states which support traffic ϕ and let $d = |\phi|$. Then if the network supports ϕ, at least*

$$|\mathcal{V}(\phi)| \geq (T!)^{\lfloor d/T \rfloor}(T)_{d \bmod T} \tag{13.5}$$

network states support ϕ.

Proof: Pick $(\mathbf{v}, \psi) \in \mathcal{V}(\phi)$ which is always possible since the network supports ϕ. $\mathbf{v}$ can be represented as

$$\mathbf{v} = \begin{bmatrix} (f_1, t_1) & (f_2, t_2) & ... & (f_M, t_M) \\ (f'_1, t'_1) & (f'_2, t'_2) & ... & (f'_M, t'_M) \end{bmatrix} \tag{13.6}$$

where (f_n, t_n) is the channel assigned to transmitter n in $\mathbf{v}$ and (f'_m, t'_m) is the channel assigned to receiver m in $\mathbf{v}$. Let $[L] = \{1, 2, ..., L\}$ be the light trees used in $\mathbf{v}$. Then $\mathbf{v}$ can also be written as

$$\mathbf{v} = \begin{bmatrix} (c_1, t_1) & (c_2, t_2) & ... & (c_M, t_M) \\ (c'_1, t'_1) & (c'_2, t'_2) & ... & (c'_M, t'_M) \end{bmatrix} \tag{13.7}$$

where c_n is the light tree assigned to transmitter n in $\mathbf{v}$, c'_m is the light tree assigned to receiver m in $\mathbf{v}$, and t_n, t'_m are as before.

Now let h_l be the number of sessions assigned light tree l in $\mathbf{v}$, for $l \in [L]$. Since there are T time slots and since a light tree can support at most one session per time slot, $h_l \leq T$ for all $l \in [L]$. Also since there are d active sessions,

$$\sum_{l \in [L]} h_l = d \tag{13.8}$$

There are $(T)_{h_l}$ ways of scheduling the h_l sessions assigned light tree j since there are T ways of scheduling the first session, $(T-1)$ ways of scheduling the second session, etc. Since no signal travels on more than one light tree, the sessions on each light tree can be scheduled independently without interfering with sessions on other light trees. Therefore, the number of feasible network states for ϕ is at least

$$|V(\phi)| \geq \prod_{l=1}^{L}(T)_{h_l} \tag{13.9}$$

To complete the proof, we show that

$$\min_{\{h_l\}}\left\{\left(\prod_{l=1}^{L}(T)_{h_l}\right) \ : \ s.t. \ \sum_{l\in[L]} h_l = d\right\} \geq (T!)^{\lfloor d/T \rfloor}(T)_{d \bmod T} \tag{13.10}$$

The key to proving eqn. (13.10) is the observation that for integers a and b such that $0 < a \leq b < T$,

$$(T)_{b+1}(T)_{a-1} < (T)_b(T)_a \tag{13.11}$$

To see this, divide both sides by $(T)_{a-1}(T)_b$. This gives

$$\frac{(T)_{b+1}}{(T)_b} < \frac{(T)_a}{(T)_{a-1}} \tag{13.12}$$

The left hand side is $T - b$ and the right hand side is $T - a + 1$ and since $b \geq a$, the inequality is true.

Now let h_l^*, $l = 1, 2, ..., L$ minimize the left hand side of eqn. (13.10). Then exactly $\left\lfloor \frac{d}{T} \right\rfloor$ of the h_l^*'s have value T, one h_l^* has value $d \bmod T$, and the remaining $d - \left\lceil \frac{d}{T} \right\rceil$ variables have value 0. If this were not the case then there would exist and i and j such that $0 < h_i^* \leq h_j^* < T$. But if this were so,

$$(T)_{h_j^*+1}(T)_{h_i^*-1} < (T)_{h_j^*}(T)_{h_i^*} \tag{13.13}$$

and h_l^*, $l = 1, 2, ..., L$ would not minimize the function. □

By the two lemmas,

$$S \cdot \frac{\binom{M}{d}^2 F^{2d} T^{2d}}{(T!)^{\lfloor d/T \rfloor} (T)_{d\bmod T}} \geq V \geq |\mathcal{T}_d| \tag{13.14}$$

since there are at most $S \cdot \binom{M}{d}^2 (FT)^{2d}$ feasible network states for $\mathcal{T}_d$. We simplify eqn. (13.14) by noting that $T! \geq (T/e)^T$ and $(T)_{d\bmod T} \geq \frac{\sqrt{2\pi}}{e}(T/e)^{d\bmod T}$; the latter was proved Section 9.4. Therefore

$$(T!)^{\lfloor d/T \rfloor} (T)_{d\bmod T} \geq \left(\frac{T}{e}\right)^{d - d\bmod T} \left(\frac{T}{e}\right)^{d\bmod T} \frac{\sqrt{2\pi}}{e} \tag{13.15}$$

Use this to upper bound the left hand side of eqn. (13.14) and then raise both sides by $1/2d$ to prove eqn. (13.2).

Notice for $T = 1$, eqn. (13.2) is about a factor of $1/\sqrt{e}$ worse than the bound of Chapter 4. This is only due to the inequality used in eqn. (13.15). For a small number of time slots eqn. (13.14) can be used for a better bound.

13.2 Light Tree Bound

In this section, we will present and discuss a bound due to Gallager [Gal92b]. The bound is valid for passive Light Tree Networks under certain traffic restrictions. The traffic is modeled using a more conventional flow model instead of the traffic sets we have been using. The bound, hereafter referred to as the Light Tree Bound, will be formally presented below. First, let us informally present the definitions and results.

Consider any passive LT-AON H with L light trees, each of which can support at maximum of R b/s. Let $\gamma(n,m)$ be the throughput between user n and user m measured in b/s and let γ be the average throughput, where the average is taken over pairs of users. The $M \times M$ matrix $[\gamma(n,m)]$ is called the *flow.* The relationship between flows and traffics will be discussed later.

The Light Tree Bound says that if $\gamma(n,m) = \gamma$ for all (n,m), the total throughput

$Z \stackrel{\text{def}}{=} M^2\gamma$ is no more than F^2R. In addition if $\gamma(n,m) \approx \gamma$ for all (n,m), the total throughput cannot be much greater than F^2R. When $\gamma(n,m) = \gamma$ for all (n,m), we say the flow is *homogeneous*. When $|\gamma(n,m) - \gamma| \leq \epsilon$, we say that flow is *$\epsilon$-homogneous.*

We now prove the bound for ϵ-homogneous traffic. Then we relate this bound and the flow model to the traffic set model. Conclusions on previously studied traffic sets are presented there.

Theorem 23 Light Tree Bound

A passive Light Tree AON supporting ϵ-homogeneous flow has a throughput no more than

$$Z \leq F^2R\left(\frac{\gamma+\epsilon}{\gamma-\epsilon}\right). \tag{13.16}$$

Therefore if the flow is homogeneous, $Z \leq F^2R$ and also $F \geq M\sqrt{\gamma/R}$.

Proof: Let z_l be the throughput of light tree l, $l = 1, 2, ...L$. Also, let $h(l)$ and $w(l)$ be the number of transmitters and receivers connected to light tree l, respectively. A light tree cannot support more b/s on average than the users connected to it can give it. Since there are $h(l)w(l)$ pairs of users connected to light tree l and since the flow between two users is no more than $\gamma + \epsilon$,

$$z_l \leq (\gamma+\epsilon)h(l)w(l). \tag{13.17}$$

Also, since a light tree can support no more than R b/s.

$$z_l \leq \min\{(\gamma+\epsilon)h(l)w(l), R\} \tag{13.18}$$

and therefore,

$$Z = \sum_{l=1}^{L} z_l \leq \sum_{l=1}^{L} \min\{(\gamma+\epsilon)h(l)w(l), R\} \tag{13.19}$$

$$= R\sum_{l=1}^{L} \min\{h(l)w(l)\frac{\gamma+\epsilon}{R}, 1\}. \tag{13.20}$$

Since $\min\{x, 1\} \leq \sqrt{x}$, we have

$$Z \leq R\sum_{l=1}^{L} \sqrt{h(l)w(l)\frac{\gamma+\epsilon}{R}} \tag{13.21}$$

$$= \sqrt{R(\gamma+\epsilon)}\sum_{l=1}^{L} \sqrt{h(l)w(l)}. \tag{13.22}$$

Now since $\sqrt{xy} \leq .5(x+y)$, which can be seen by squaring both sides of the inequality,

$$Z \leq .5\sqrt{R(\gamma+\epsilon)}\sum_{l=1}^{L} h(l) + w(l). \tag{13.23}$$

Each transmitter is connected to at most F light trees and similarly true for each receiver, so

$$Z \leq .5\sqrt{R(\gamma+\epsilon)}(2FM) = FM\sqrt{R(\gamma+\epsilon)}. \tag{13.24}$$

Now squaring both sides,

$$Z^2 \leq F^2RM^2(\gamma+\epsilon). \tag{13.25}$$

By definition of throughput and ϵ-homogeneous flow,

$$Z = \sum_{(n,m)} \gamma(n,m) \geq M^2(\gamma-\epsilon). \tag{13.26}$$

Therefore, $M^2 \leq \frac{Z}{\gamma-\epsilon}$ and

$$Z^2 \leq \ F^2 R \left(\frac{\gamma+\epsilon}{\gamma-\epsilon}\right). \tag{13.27}$$

This completes the proof of the first statement. Now if $\epsilon = 0$, the second statement follows. Also, if $\epsilon = 0$, $Z = M^2\gamma \leq F^2 R$ and therefore, $F \geq M\sqrt{\gamma/R}$. □

Care needs to be taken not to overstate the results of the Light Tree Bound. In particular, the Light Tree Bound does not say that F^2 is the maximum number of light trees. Consider the following generalization of the ring network. Label MF trunks by the doublet (i, f) for $i = 0, ..., M-1$ and $f = 0, ..., F-1$. Connect input n to trunk (n, f) on wavelength f for each $f = 0, 1, ..., F-1$. Also connect the output of trunk (n, f) to the receiver $n+f \bmod M$. Then this is a light tree network with FM trees; however this network cannot support homogeneous traffic. In fact, the network is not even connected. To make a connected network, form the ring network described above with $F-1$ wavelengths and hook all transmitters and all receivers to a light tree on wavelength F. Now there are $M(F-1)+1$ light trees and the network is connected.

In the above examples, either the network is not connected or the light trees are overlapping. The quick reader may have hypothesized that at most F^2 light trees are possible if the light trees do not overlap and the network is connected. This hypothesis is also incorrect; an example is given in [Gal92b].

13.2.1 Traffic Set Light Tree Bound

The Light Tree Bound can be used to show that $F\sqrt{T} \geq \sqrt{\rho M}$ is a necessary condition to do ρ-permutation routing. This is a factor of e better than the tuning state bound in section 13.1. Similarly, the Light Tree Bound can be used to show that $F\sqrt{T} \geq \sqrt{M}$ for the Santa Barbara Traffics, a factor of e^2

better. To see how to apply the Light Tree Bound, we first need the following definition.

Definition 7 Homogeneous Traffic Set

A traffic set is homogeneous if each session appears in the same number of traffics.

Consider any traffic set $\mathcal{T}$ and let $p(n,m)$ be the fraction of traffics in which session (n,m) appears. Then $p(n,m)$ is given by

$$p(n,m) = \frac{1}{|\mathcal{T}|} \sum_{\phi \in \mathcal{T}} 1_{(n,m)}(\phi) \tag{13.28}$$

where $1_{(n,m)}(\phi) = 1$ if $(n,m) \in \phi$ and 0 otherwise. So $p(n,m)$ is the probability that session (n,m) is active in a traffic picked uniformly at random from the traffics in $\mathcal{T}$ and $\mathcal{T}$ is homogeneous iff all sessions are equally likely.

All of the traffic sets used in this thesis are homogeneous: $p(n,m) = 1/M$ for permutation routing, $p(n,m) = \rho/M$ for ρ-permutation routing, and the Santa Barbara Traffics have $p(n,m) = 1/M$. Note that when $p(n,m) = p$ for all (n,m), the expected number of active sessions in a traffic picked at random is pM^2. This provides an easy way to compute p for symmetric traffics like ρ-permutation routing.

As before, we will only need a lesser assumption, that of δ-homogeneous traffics.

Definition 8 δ-homogeneous Traffic Set

A traffic set, $\mathcal{T}$, is δ-homogeneous if for all (n,m),

$$|p(n,m) - \delta| \leq \frac{1}{M^2} \sum_{(i,j)} p(i,j) \stackrel{\text{def}}{=} \bar{p} \tag{13.29}$$

Consider an AON supporting $\mathcal{T}$ in F wavelengths and T time slots. When the network is supporting $\phi \in \mathcal{T}$, the total throughput is $Z_\phi = R_s|\phi|$ b/s where recall that $R_s = R/T$ is the session bit rate. Now by routing each traffic in $\mathcal{T}$ exactly once, we have a network which has a throughput of

$$\begin{aligned} Z &= \frac{1}{|\mathcal{T}|}\sum_{\phi\in\mathcal{T}} Z_\phi = \frac{R_s}{|\mathcal{T}|}\sum_{\phi\in\mathcal{T}} |\phi| \\ &= R_s\bar{p}M^2 \end{aligned} \tag{13.30}$$

Also, the throughput between n and m is

$$\gamma(n,m) = R_s \sum_{\phi\in\mathcal{T}} 1_{(n,m)}(\phi) = R_s p(n,m) \tag{13.31}$$

By the Light Tree Bound, $Z \leq F^2 R\left(\frac{\gamma+\epsilon}{\gamma-\epsilon}\right)$, where $\epsilon = R_s\delta$. Therefore,

$$F \geq M\sqrt{\frac{\bar{p}}{T}\left(\frac{\bar{p}-\delta}{\bar{p}+\delta}\right)}. \tag{13.32}$$

13.3 Previous Results

In Chapter 5 we discussed a bound derived by Pippenger and Yao [PY82]. They showed that at least $2M^{3/2}$ edges are required for a depth 2 interconnection network to be an M-shifter and therefore a Light Tree Connector without time slots requires $F \geq \sqrt{M}$ wavelengths. Since an M-shifter is a homogeneous traffic set, the same result follows from the Light Tree Bound. In fact, Pippenger and Yao's bound holds for all homogeneous traffic sets and the proof is very similar to the proof of the Light Tree Bound. There are a couple differences. First, their bound is not as general since it uses the traffic set model instead of the flow model and also does not consider time slots. Second, Pippenger and Yao were interested in the total number of edges required, not the maximum degree of a peripheral node, the latter being the required number of wavelengths in the light tree representation. However it would not be surprising if for homogeneous traffic sets, the number of edges would be minimized by graphs with constant degree.

Birk, et. al. [BLM93] and Liew [Lie88] have independently derived bounds similar to the Light Tree Bound. Birk, et. al. bound the maximum throughput of a multiple bus network whereas Liew bounds the sum of the number of bus connections for a given throughput. Note that the number of bus connections is analogous to the number of edges in Pippenger and Yao's bound.

Chapter 14

The All-to-All Problem

In the *All-to-All* problem each of N LANs has exactly one R_s b/s session for each other LAN. In addition, each LAN has one outgoing fiber, one incoming fiber, and a large but unspecified number of users. The outgoing (incoming) fiber of a LAN is connected by a broadcast star to all the transmitters (receivers) of that LAN. Since there is exactly one active session between each pair of LANs, the network supports N^2 sessions. Define the *capacity*, C, as the largest value of N^2 possible as a function of F and T. Also define the *maximum throughput* to be $Z = R_s C$ b/s.

We study the capacity and throughput of passive AONs. Before proceeding, let's discuss two special cases: broadcast AONs and LT-AONs. Recall that in a broadcast AON, each receiver hears the signals from each transmitter on each wavelength. Since there is no wavelength re-use, the class of broadcast networks has capacity $C_B = F \cdot T$.

By the Light Tree Bound $Z_{LT} \leq F^2 R$ or equivalently $C_{LT} \leq F^2 T$. To see that this capacity is achievable, fix F and T and connect $N = \sqrt{T}$ LANs to each input and each output of an $F \times F$ WDM cross-connect so that $N^2 = (F \times \sqrt{T})^2 = F^2 T$. Since each LAN has one session for each other LAN, there are a total of T sessions between any input and any output of the LR. This is exactly the number of sessions supportable by the one wavelength connecting any input to any output. If T is not a perfect square, a slight inefficiency is incurred.

Most of this chapter investigates NLT-AONs for the All-to-All problem. We will see that by relaxing the Light Tree assumption, capacity can be greatly increased.

First in section 14.1, we precisely formulate the problem and lay out the rules of contention. Then in section 14.2, we discuss the relationship between the All-to-All problem and the Santa Barbara Traffic Set introduced in Chapter 4. A bound on capacity is presented in section 14.3. We then turn our attention to constructions.

14.1 Modeling

As in Chapter 10, we think of each LAN consisting of two parts, a T-LAN containing only the transmitters and an R-LAN consisting of the receivers. Number the LANs from 0 to $N-1$ and use the notation $[x, y]$ to stand for a T-LAN, R-LAN pair. Each $[x, y]$ is called a block. Since each block must support exactly one session in All-to-All traffic, we will also refer to $[x, y]$ as a session. It should be kept in mind that "session $[x, y]$" is a short hand way of saying a session between a transmitter in T-LAN x and a receiver in R-LAN y.

Each transmitter in T-LAN x is connected to each receiver in R-LAN y on the wavelengths $J[x, y]$, where unlike Chapter 10, J is unspecified at this time. The connection matrix of the network is given by $H(n, m) = J[x_n, y_m]$ where x_n is the T-LAN of transmitter n and y_m is the R-LAN of receiver m. However we will not have any need of H since all relevant information is in J.

Let $W[x, y] = f[x, y] : f'[x, y]$ be the OD channel used by session $[x, y]$. The

$N \times N$ matrix, W is called the *assignment matrix.*[1] Since for connectivity we need $f[x,y]\!:\!f'[x,y] \in J[x,y]$, we will simply assume that $f[x,y]\!:\!f'[x,y] \in J[x,y]$. However since $|J[x,y]| > 1$ in the general situation, the assignment matrix, W, should not be confused with the connectivity matrix, J.

Having said that, we will adopt the convention that $J = \{W[x,y]\}$ and therefore drop the distinction between J and W from now on. A cautionary word to the reader is in order. This approach is not valid if there are other restrictions like J is a LT-AON. For instance, suppose that $W[0,0] = W[0,1] = W[1,1] \neq W[1,0]$, then $J[n,m] = \{W[n,m]\}$ is not a LT-AON.

Let $X[n,m]$ be the time slot used by session $[x,y]$. We call X the *scheduling matrix.* Two sessions $[n,m]$ and $[x,y]$ using the same time slot are said to be *simultaneous.* It should be kept in mind that having $[n,m]$ and $[n,y]$ simultaneously active is not a multi-point connection between two users. Similarly for $[n,m]$ and $[x,m]$ simultaneous.

Let the symbol $\star$ stand for any wavelength or time slot depending on the context. For instance, $f\!:\!\star \in J[n,m]$ means that there exists a wavelength f' such that $f\!:\!f' \in J[n,m]$. Using this notation, we say that session $[n,m]$ *kills* session $[x,y]$ if they are simultaneous and the relevant entries in the assignment matrix look like

	R-LAN m	R-LAN y
T-LAN n	$f\!:\!\star$	$f\!:\!f'$
T-LAN x	$\star\!:\!\star$	$\star\!:\!f'$

If $[n,m]$ kills $[x,y]$ then there is a fatal collision at y since when $[n,m]$ launches on f the signal is received at y on f' preventing the reception of $[x,y]$. If $n = x$ then the above definition is equivalent to: $[n,m]$ kills $[n,y]$ when

[1] In [BLM93], W is called the *wiring* matrix and indicates which transmitters are connected to which receivers.

$f[n,m] = f[n,y]$. Clearly there is contention in this case since two sessions emanating from the same T-LAN n must start on different wavelengths. Similarly, $[n,m]$ kills $[x,m]$ if $f'[n,m] = f'[x,m]$ since two sessions entering an R-LAN must be on different wavelengths. Note that in the general situation, $[n,m]$ killing $[x,y]$ does not mean that $[x,y]$ kills $[n,m]$. We say that there's a killing if $[n,m]$ kills $[x,y]$ or $[x,y]$ kills $[n,m]$.

An assignment matrix W along with a scheduling matrix X is said to be *feasible* if there is no killing. The goal in this chapter is to find feasible (W,X) in the fewest number of wavelengths and time slots possible.

14.2 Santa Barbara Traffics

In Chapter 4, the Santa Barbara Traffics were introduced to show that the tuning state bound could be tight. We repeat the definition below. It turns out that the Santa Barbara Traffics and the All-to-All problem are very related. We illustrate that relationship here and use it to bound the capacity of an AON.

Given M users, group the transmitters into groups of size $\sqrt{M}$. Call each group a T-Group. Similarly group the receivers into $\sqrt{M}$ R-Groups. The Santa Barbara Traffics are those which contain exactly one session between each T-GROUP and each R-GROUP. The traffic set is denoted by $\mathcal{T}_{SB}$ and we showed in Chapter 4 that $|\mathcal{T}_{SB}| = (\sqrt{M}!)^{2\sqrt{M}} \approx e^{-M}M!$.

To see the relationship between $\mathcal{T}_{SB}$ and the All-to-All problem, suppose that J is an $N \times N$ AON which supports All-to-All routing in F wavelengths and T time slots. Connect a T-LAN with N transmitters to each input of J and an R-LAN with N receivers to each output. The network supports $\mathcal{T}_{SB}$ and has $M = N^2$ users. The converse is not true since we have not assumed that a T-Group (R-Group) is a T-LAN (R-LAN). Therefore the All-to-All problem is a harder problem than the Santa Barbara problem. Since

$|\mathcal{T}_{SB}| = (\sqrt{M}!)^{2\sqrt{M}} \approx e^{-M}M!$, where $M = N^2$ here, and since the number of feasible tuning state is no more than $(FT)^{2M}$,

$$F \cdot T \geq \frac{\sqrt{M}}{e} = \frac{N}{e}. \tag{14.1}$$

is a necessary condition for any AON to support All-to-All Traffic. Thus C, the largest possible value of N^2, satisfies

$$C \;\leq\; e^2F^2T^2. \tag{14.2}$$

However, observe that since each LAN has N sessions leaving and entering it, and since each session requires one frequency time slot of bandwidth, it must be true that $FT \geq N$ and therefore, $C \leq F^2T^2$ for any AON. Therefore the tuning state bound is not useful for us here. For a LT-AON, we showed in Chapter 13 that a necessary condition for feasibility was $F^{2M}(eT)^M \geq \frac{\sqrt{2\pi}}{e}|\mathcal{T}|$. Therefore

$$C_{LT} \;\leq e^3F^2T. \tag{14.3}$$

Since we have already seen that $C_{LT} \leq F^2T$ using the Light Tree Bound, this result doesn't provide any new information for passive networks; it is just interesting that counting tuning states gives almost as good a bound. Also, this technique can be used to provide a bound on C_{LT} for configurable networks as well.

14.3 A Bound on Capacity

In this section, we show that

$$C \leq \frac{F^2(T+1)^2}{4} \tag{14.4}$$

Notice that if $T = 1$, $C \leq F^2$ and that this capacity is achievable by using an $F \times F$ WDM cross-connect as discussed in the introduction. Therefore, $C = C_{LT} = F^2$ for any F and $T = 1$. However for $T >> 1$, we'll see that NLT-AONs have much larger capacity than LT-AONs.

Theorem 24 All-to-All Capacity Bound

$$C \leq \frac{F^2(T+1)^2}{4}. \tag{14.5}$$

Proof: Let $d_r[n,m]$ be the number of times $f[n,m]$ is used in row n. Also let $d_c[n,m]$ be the number of times $f'[n,m]$ is used in column m. Then for any n,

$$\sum_{m=1}^{N} \frac{1}{d_r[n,m]} \tag{14.6}$$

is the number of origin wavelengths used in row n. Similarly, for all m,

$$\sum_{n=1}^{N} \frac{1}{d_c[n,m]} \tag{14.7}$$

is the number of destination wavelengths used column m. It follows that

$$\sum_{(n,m)} \frac{1}{d_r[n,m]} + \frac{1}{d_c[n,m]} \leq 2FN \tag{14.8}$$

Consider any $[n,m]$ and let $d_r = d_r[n,m]$, $d_c = d_c[n,m]$, $f = f[n,m]$ and $f' = f'[n,m]$. This situation is represented in eqn.(14.9). There are d_r sessions in the top row with the same origin wavelength and d_c sessions in the right column with the same destination wavelength. The row and column meet at $[n, m]$. It's easy to see that if any of the sessions are simultaneous, there's a killing. Since there are $d_r + d_c - 1$ sessions, $T \geq d_r + d_c - 1$.

$$\begin{array}{c|c|c|c|c} f:\star & f:\star & \dots & f:\star & f:f' \\ \hline \star:\star & \star:\star & \dots & \star:\star & \star:f' \\ \hline \star:\star & \star:\star & \dots & \star:\star & \star:f' \\ \hline \star:\star & \star:\star & \dots & \star:\star & \vdots \\ \hline \star:\star & \star:\star & \dots & \star:\star & \star:f' \end{array} \tag{14.9}$$

Therefore,

$$\frac{1}{d_r} + \frac{1}{d_c} \geq \frac{1}{d_r} + \frac{1}{T + 1 - d_r} \tag{14.10}$$

Relaxing the integer constraints, the right hand side is minimized at $d_r = (T + 1)/2$, so

$$\frac{1}{d_r} + \frac{1}{d_c} \geq \frac{4}{T + 1} \tag{14.11}$$

Summing over all $[n, m]$ and using eqn. 14.8,

$$2FN \geq \frac{4N^2}{T + 1} \tag{14.12}$$

which gives

$$N^2 \leq \frac{F^2(T+1)^2}{4} \tag{14.13}$$

□

14.4 Constructions

We present various constructions. The constructions are ordered from worst to best performance in terms of time slot efficiency. In particular, we start with the Latin Router design which achieves a throughput of F^2T wavelengths. Then we discuss several designs which support $N^2 = \Theta(F^2T \log T)$ sessions. Finally in section 14.4.5 we present the best known design which supports $N^2 = F^2T^{4/3}$ sessions.

14.4.1 Latin Router

This design was discussed in the introduction. We include it here for completeness and because we will use it later as a building block.

Hook $\lfloor\sqrt{T}\rfloor$ LANs to each input and each output of an $F \times F$ WDM cross-connect for $N^2 = F^2 \lfloor\sqrt{T}\rfloor^2 \approx F^2T$ possible concurrent connections. An example is shown in Fig. 14-1.

We formally define the *LR design* (W_{LR}, X_{LR}) with F wavelengths and T time slots to be the following $F\lfloor\sqrt{T}\rfloor \times F\lfloor\sqrt{T}\rfloor$ matrices

$$
\begin{aligned}
W_{LR}[n,m] &= \left\lfloor \frac{m}{\lfloor\sqrt{T}\rfloor} \right\rfloor - \left\lfloor \frac{m}{\lfloor\sqrt{T}\rfloor} \right\rfloor \pmod F \\
X_{LR}[n,m] &= (n \bmod \lfloor\sqrt{T}\rfloor) * \lfloor\sqrt{T}\rfloor + (m \bmod \lfloor\sqrt{T}\rfloor) \qquad (14.14)
\end{aligned}
$$

In Fig. 14-1, wavelength 0 is shown as R, 1 is shown as G, 2 is shown as B, and 3 is shown as Y. The scheduling matrix is not shown.

14.4.2 Birk's Design

We only present the most relevant design from [BLM93]. This design uses 2 wavelengths which we call Red and Blue.

The key part of the design is an asymmetric network with k transmitters

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	R	R	R	R	G	G	G	G	B	B	B	B	Y	Y	Y	Y
1	R	R	R	R	G	G	G	G	B	B	B	B	Y	Y	Y	Y
2	R	R	R	R	G	G	G	G	B	B	B	B	Y	Y	Y	Y
3	R	R	R	R	G	G	G	G	B	B	B	B	Y	Y	Y	Y
4	Y	Y	Y	Y	R	R	R	R	G	G	G	G	B	B	B	B
5	Y	Y	Y	Y	R	R	R	R	G	G	G	G	B	B	B	B
6	Y	Y	Y	Y	R	R	R	R	G	G	G	G	B	B	B	B
7	Y	Y	Y	Y	R	R	R	R	G	G	G	G	B	B	B	B
8	B	B	B	B	Y	Y	Y	Y	R	R	R	R	G	G	G	G
9	B	B	B	B	Y	Y	Y	Y	R	R	R	R	G	G	G	G
10	B	B	B	B	Y	Y	Y	Y	R	R	R	R	G	G	G	G
11	B	B	B	B	Y	Y	Y	Y	R	R	R	R	G	G	G	G
12	G	G	G	G	B	B	B	B	Y	Y	Y	Y	R	R	R	R
13	G	G	G	G	B	B	B	B	Y	Y	Y	Y	R	R	R	R
14	G	G	G	G	B	B	B	B	Y	Y	Y	Y	R	R	R	R
15	G	G	G	G	B	B	B	B	Y	Y	Y	Y	R	R	R	R

Figure 14-1: The 4 wavelength, 16 time slot Latin Router Assignment Matrix W_{LR}.

and $N = 2^k$ receivers. For this part of the design k can be any positive integer; however, we will further restrict k below. The first row of the assignment matrix W contains 2^{k-1} Reds followed by 2^{k-1} Blues. See Fig. 14-2 for a $k = 4$ example (Only the Red wavelengths are shown for clarity). The next row contains 2^{k-2} Reds followed by 2^{k-2} Blues followed by 2^{k-2} Reds, and ended with 2^{k-2} Blues. In each subsequent row, the number of adjacent Reds and Blues is cut in half until the last row which contains the pattern RBRB...RB. To be more precise, let $s_n = 2^{k-n-1}$, for $n = 0, 2, ...k-1$. Then $f[n, m] =$ Red if $\left\lfloor \frac{m}{s_n} \right\rfloor$ is even and $f[n, m] =$ Blue if $\left\lfloor \frac{m}{s_n} \right\rfloor$ is odd.

The $k \times 2^k$ assignment can be scheduled with 2^k time slots by setting $t[n, m] = m + s_n \pmod{2^k}$ if $f[n, m] =$ Red and $t[n, m] = m - s_n \pmod{2^k}$ if $f[n, m] =$ Blue. For this reason we call s_n the *shift* of row n. So the first Red of row n uses time slot s_n and the first Blue of any row uses time slot 0. The

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	R	R	R	R	R	R	R	R	·	·	·	·	·	·	·	·
1	R	R	R	R	·	·	·	·	R	R	R	R	·	·	·	·
2	R	R	·	·	R	R	·	·	R	R	·	·	R	R	·	·
3	R	·	R	·	R	·	R	·	R	·	R	·	R	·	R	·

Figure 14-2: Birk's 2 Wavelength Design

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	8	9	10	11	12	13	14	15	·	·	·	·	·	·	·	·
1	4	5	6	7	·	·	·	·	12	13	14	15	·	·	·	·
2	2	3	·	·	6	7	·	·	10	11	·	·	14	15	·	·
3	1	·	3	·	5	·	7	·	9	·	11	·	13	·	15	·

Figure 14-3: Schedule for Design

feasibility of (W, X) is shown in the next lemma.

Lemma 25 *The design is feasible.*

Proof: It suffices to prove that R-LAN y only hears T-LAN x in time slot $t[x, y]$, for all $[x, y]$. Note that this is a stronger condition than feasibility, i.e. feasibility is R-LAN y only hears T-LAN x in $t[x, y]$ on $f[x, y]$, for all $[x, y]$.

First we need a few definitions. Let $\hat{s}_n$ be the binary representation of s_n. $\hat{s}_n$ is all 0's except for a 1 in the $(k-n)^{th}$ bit. Also let $\hat{m}$ be the binary representation of m, for $m = 0, 1..., 2^k$. Then $f[n, m]$ is Red iff the $(k-n)^{th}$ bit of $\hat{m}$ is 0. Associating Red with 0 and Blue with 1,

$$f[n, m] = \hat{s}_n \odot \hat{m} \qquad (14.15)$$

where $\odot$ is the dot product. The time slot of $[n, m]$ in binary is

$$\hat{t}[n, m] = \hat{s}_n \oplus \hat{m} \qquad (14.16)$$

where addition is done bitwise modulo 2.

Now suppose $[n, m]$'s signal is heard at T-LAN y in the same time slot that $[x, y]$ is using, $[n, m] \neq [x, y]$. Then it must be true that $f[n, m] = f[n, y]$ and $t[n, m] = t[x, y]$. Since $f[n, m] = f[n, y]$,

$$\hat{s}_n \odot \hat{m} = \hat{s}_n \odot \hat{y} \tag{14.17}$$

or equivalently $\hat{m}$ and $\hat{y}$ agree in the $(k-n)^{th}$ bit. Also, they must be simultaneous so

$$\hat{s}_n \oplus \hat{m} = \hat{s}_x \oplus \hat{y} \tag{14.18}$$

The $(k-n)^{th}$ bit on the left hand side is the negative of the $(k-n)^{th}$ bit of $\hat{m}$. Since $\hat{m}$ and $\hat{y}$ agree in the $(k-n)^{th}$ bit, the equality can only be true if $x = n$. But if $x = n$, then $s_n \oplus \hat{m} = s_n \oplus \hat{y}$ and since s_n is all 0's except for the $(k-n)^{th}$ bit, $\hat{m} = \hat{y}$. Therefore $[n, m] = [x, y]$ which is a contradiction. □

Now to build a symmetrical $2^k \times 2^k$ assignment from the $k \times 2^k$ assignment, we repeat the above design $\frac{2^k}{k}$ times. Fig. 14-4 shows how this is done for $k = 4$. If k does not divide 2^k, then we must repeat the design $\left\lceil \frac{2^k}{k} \right\rceil$ times and discard the extra $2^k - k \cdot \left\lceil \frac{2^k}{k} \right\rceil$ rows. Since this will introduce inefficiency, we will restrict our attention to k's which divide 2^k, i.e. k a power of 2.

Each vertical repetition uses 2^k distinct time slots for a total of $T = \frac{4^k}{k}$. Therefore in T time slots and 2 wavelengths, we can support $N = 2^k$ LANs. For instance, if $k = 4$, then with 2 wavelengths and 64 time slots, a total of 16 LANs are supportable. This is a throughput of 256 wavelengths, exactly the same as the LR design. However for larger k, the design outperforms the LR. In general for any k a power of 2, the design achieves a throughput of $N^2 = 4^k$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	R	R	R	R	R	R	R	R	·	·	·	·	·	·	·	·
1	R	R	R	R	·	·	·	·	R	R	R	R	·	·	·	·
2	R	R	·	·	R	R	·	·	R	R	·	·	R	R	·	·
3	R	·	R	·	R	·	R	·	R	·	R	·	R	·	R	·
4	R	R	R	R	R	R	R	R	·	·	·	·	·	·	·	·
5	R	R	R	R	·	·	·	·	R	R	R	R	·	·	·	·
6	R	R	·	·	R	R	·	·	R	R	·	·	R	R	·	·
7	R	·	R	·	R	·	R	·	R	·	R	·	R	·	R	·
8	R	R	R	R	R	R	R	R	·	·	·	·	·	·	·	·
9	R	R	R	R	·	·	·	·	R	R	R	R	·	·	·	·
10	R	R	·	·	R	R	·	·	R	R	·	·	R	R	·	·
11	R	·	R	·	R	·	R	·	R	·	R	·	R	·	R	·
12	R	R	R	R	R	R	R	R	·	·	·	·	·	·	·	·
13	R	R	R	R	·	·	·	·	R	R	R	R	·	·	·	·
14	R	R	·	·	R	R	·	·	R	R	·	·	R	R	·	·
15	R	·	R	·	R	·	R	·	R	·	R	·	R	·	R	·

Figure 14-4: Vertical Repetition of Fig. 14-2

which is approximately $\frac{1}{2}T\log T$. To see the approximation, notice that

$$\frac{4^k}{\frac{1}{2}T\log T} = \frac{4^k}{\frac{1}{2}\frac{4^k}{k}(2k-\log k)} = 1 + O\left(\frac{\log k}{k}\right) \tag{14.19}$$

Since the error term is decreasing with k for $k \geq e$, we can upper bound it by lower bounding k. Now note that $k \geq .5\log T$ since $\log T = 2k - \log k$ so

$$N^2 = \left(\frac{1}{2}T\log T\right)\cdot\left[1 + O\left(\frac{\log\log T}{\log T}\right)\right] \tag{14.20}$$

In [BuHa90], it was shown how to improve the $k \times 2^k$ design to $k+1 \times 2^k$ using the same number of time slots. Repeating this design vertically $\frac{2^k}{k+1}$ times, assuming $k+1$ divides 2^k, we end up with a network with $N = 2^k$ LANs,

$T = \frac{4^k}{k+1}$ time slots and 2 wavelengths. Now note that

$$\frac{4^k}{\frac{1}{2}T \log 4T} = \frac{4^k}{\frac{1}{2}\frac{4^k}{k+1}(2k+2-\log k)} = 1 + O\left(\frac{\log k}{k}\right) \qquad (14.21)$$

and that $k \geq .5 \log T$ since $\log T = 2k - \log k + 1$ so

$$N^2 = \left(\frac{1}{2}T \log 4T\right) \cdot \left[1 + O\left(\frac{\log \log T}{\log T}\right)\right] \qquad (14.22)$$

14.4.3 Birk*2 Design

In the proof of lemma 25, we showed that if R-LAN y is listening to B in time slot t, then no other signals are received at y in time slot t, not even on wavelength R.[2] Therefore, it seems likely that we should be able to supplement the design by adding rows so that y also hears a signal on R in time slot t. In fact this is possible.

The design is as follows. For $k \geq 1$ an integer, the top k rows are the $k \times 2^k$ rows described above. The second k rows are the negative of the first k rows, i.e. the first k rows with R's and B's replaced. The schedule for each half is the same as the original design. To be more precise, let $f[n,m]$ and $t[n,m]$ be the assignment and schedule for the $k \times 2^k$ design and let $f'[n,m]$ and $t'[n,m]$ be the assignment and schedule for the supplemented design. Then letting $\bar{f}$ stand for the inverse of f,

$$\begin{aligned} f'[n,m] &= \begin{cases} f[n,m] & \text{if } n < k \\ \bar{f}[n-k,m] & \text{if } n \geq k \end{cases} \\ t'[n,m] &= t[n \bmod k, m] \end{aligned} \qquad (14.23)$$

for $n = 0, 1, ..., 2k-1$, $m = 0, 1, ...2^k - 1$. An example is shown in Fig. 14-5.

[2]However it is not true that an R-LAN only ever hears one signal. Consider Fig. 14-2 and time slot 8. R-LANs 12,13,14, and 15 all hear a Red signal from $[0,0]$ and a Blue signal from $[1,12]$.

To see that the design is feasible, suppose that $[n, m]$ kills $[x, y]$. There are two possibilities: $[n, m]$ and $[x, y]$ in the same half or in opposite halves. The first case is impossible by lemma 25. To see that the second case is impossible, assume without loss in generality that $[n, m]$ is in the bottom half and refer to the following picture.

		$\dots$	y	$\dots$	m	$\dots$
	x		B, t			
Top Half	$\vdots$					
	$n-k$		R		R, t	
	$\vdots$					
	$\vdots$					
Bottom Half	n		B		B, t	
	$\vdots$					

(14.24)

If $[n, m]$ kills $[x, y]$ then $f[n, m] = f[x, y] = f[n, y]$ which we have assumed is B in the picture. Therefore $f[n-k, y] = f[n-k, m] = R$ by the definition of the design. Since $t[n, m] = t[n-k, m] = t$, the session $[n-k, m]$ is heard at R-LAN y during time slot t. Now if $[n, m]$ kills $[x, y]$ then they must also be simultaneous, so $t[x, y] = t$. Therefore considering only the top half during time slot t, R-LAN y hears two signals, one of which is intended for R-LAN y. By the proof of lemma 25, this is impossible so $[n, m]$ does not kill $[x, y]$ and the design is feasible.

To build a symmetrical $2^k \times 2^k$ design from the $2k \times 2^k$ design, we repeat the above design $\frac{2^k}{2k}$ times. If k does not divide 2^{k-1}, then we must repeat the design $\left\lceil \frac{2^k}{2k} \right\rceil$ times and discard the extra $2^k - 2k \cdot \left\lceil \frac{2^k}{2k} \right\rceil$ rows. Since this will introduce inefficiency, we will restrict our attention to k's which divide 2^{k-1}, i.e. k a power of 2.

Each vertical repetition uses 2^k distinct time slots for a total of $T = \frac{4^k}{2k}$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	R	R	R	R	R	R	R	R	·	·	·	·	·	·	·	·
1	R	R	R	R	·	·	·	·	R	R	R	R	·	·	·	·
2	R	R	·	·	R	R	·	·	R	R	·	·	R	R	·	·
3	R	·	R	·	R	·	R	·	R	·	R	·	R	·	R	·
4	·	·	·	·	·	·	·	·	R	R	R	R	R	R	R	R
5	·	·	·	·	R	R	R	R	·	·	·	·	R	R	R	R
6	·	·	R	R	·	·	R	R	·	·	R	R	·	·	R	R
7	·	R	·	R	·	R	·	R	·	R	·	R	·	R	·	R

Figure 14-5: Birk*2 Design

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	8	9	10	11	12	13	14	15	·	·	·	·	·	·	·	·
1	4	5	6	7	·	·	·	·	12	13	14	15	·	·	·	·
2	2	3	·	·	6	7	·	·	10	11	·	·	14	15	·	·
3	1	·	3	·	5	·	7	·	9	·	11	·	13	·	15	·
4	8	9	10	11	12	13	14	15	·	·	·	·	·	·	·	·
5	4	5	6	7	·	·	·	·	12	13	14	15	·	·	·	·
6	2	3	·	·	6	7	·	·	10	11	·	·	14	15	·	·
7	1	·	3	·	5	·	7	·	9	·	11	·	13	·	15	·

Figure 14-6: Schedule for Design

Therefore, in T time slot and 2 wavelengths, we can support $N = 2^k$ LANs. For instance, if $k = 4$, then with 2 wavelengths and 32 time slots, a total of 16 LANs are supportable. This is a throughput of 256 wavelengths, which is twice the throughput of the LR. For larger k, this design does even better. In general, the design achieves a throughput of $N^2 = 4^k$ which is approximately $T \log 2T$. To see the approximation, notice that

$$\frac{4^k}{T \log 2T} = \frac{4^k}{\frac{4^k}{2k}(2k - \log k)} = 1 + O\left(\frac{\log k}{k}\right) \qquad (14.25)$$

and that $k \geq .5 \log T$ since $\log T = 2k - 1 - \log k$ so

$$N^2 = (T \log 2T) \cdot \left[1 + O\left(\frac{\log \log T}{\log T}\right)\right] \qquad (14.26)$$

14.4.4 An 8 by 8 Example

An important design that we'll return to later is the 2 wavelength 8×8 design shown in Fig. 14-7. It is based on the theory of the last section with $k = 3$ and using the extension laid out in [BuHa90]. The design uses 8 time slots. Only the Red sessions are shown for clarity.

The example is important for two reasons. Most importantly, we will use this design to build larger very efficient networks. This is done below. Secondly, the design shows that the bound of the last section can be tight. In particular, only 8 time slots are used and from Theorem 24 at least 7 are needed.

14.4.5 Composition of Designs

Let (W_1, X_1) be a feasible $N_1 \times N_1$ design with F_1 wavelengths and T_1 time slots. Also assume for simplicity that W_1 does not use wavelength changing. Suppose also that (W_2, X_2) is a feasible $N_2 \times N_2$ design with F_2 wavelengths and T_2 time slots that also does not use wavelength changing. In [BLM93], it

R	R	R	R	·	·	·	·
R	R	·	·	R	R	·	·
R	·	R	·	R	·	R	·
R	·	·	R	·	R	R	·
·	·	·	·	R	R	R	R
·	·	R	R	·	·	R	R
·	R	·	R	·	R	·	R
·	R	R	·	R	·	·	R

Figure 14-7: Assignment for 8 × 8 Design

4	5	6	7				
2	3			6	7		
1		3		5		7	
0			3		5	6	
				0	1	2	3
		0	1			4	5
	0		2		4		6
	1	2		4			7

Figure 14-8: Schedule for Design

was shown how to construct an $N_1 \cdot N_2 \times N_1 \cdot N_2$ feasible design with $F_1 \cdot F_2$ wavelengths and $T_1 \cdot T_2$ time slots using (W_1, X_1) and (W_2, X_2) as building blocks. Here we lay out the theory of this composition as well as using it to construct new designs.

Denote the *composition* of two matrices as $W = W_1 \otimes W_2$. An example is shown in Fig. 14-9 where W_1 is the 4×4 two wavelength design and W_2 is the two wavelength 2×2 design. The entries of W are wavelengths, written in the form, (f, w). We call w the *waveband* of the wavelength (f, w). Informally, the composition operator replaces each entry of W_2 with a copy of W_1 where the wavelengths of W_1 fall in the *wavebands* of W_2.

Definition 9 Composition

Let W_i be an $N_i \times N_i$ matrix for $i = 1, 2$. Then the composition of W_1 and W_2, denoted $W = W_1 \otimes W_2$, is the $N_1 \cdot N_2 \times N_1 \cdot N_2$ matrix with entries

$$W[(n_1, n_2), (m_1, m_2)] = (W_1(n_1, m_1), W_2(n_2, m_2)) \tag{14.27}$$

where we have labeled the rows and columns of W by the pairs (n_1, n_2), $0 \leq n_1 < N_1$, $0 \leq n_2 \leq N_2$.

The entries $W[(\star, n_2), (\star, m_2)]$ are called the *children* of $[n_2, m_2]$. Also, $[n_2, m_2]$ is their unique parent. Notice that children of $[n_2, m_2]$ all use wavelengths in band $W_2[n_2, m_2]$ so that entries in W having different parents use different wavelengths.

Now denote the composition of two designs as

$$(W, X) = (W_1, X_1) \otimes (W_2, X_2) = (W_1 \otimes W_2, X_1 \otimes X_2). \tag{14.28}$$

Note that the entries of the composition $X = X_1 \otimes X_2$ are time slots and are written in the form (t, b). We call b the *time band* of time slot (t, b). From

Fig. 14-9 it is fairly obvious that the composition of two feasible designs is also feasible; this is shown in the next theorem, adapted from [BLM93].

Theorem 26 Feasible of Compositions

If (W_1, X_1) and (W_2, X_2) are feasible, $(W, X) = (W_1, X_1) \otimes (W_2, X_2)$ is feasible.

Proof: By contradiction.

Suppose $s = [(n_1, m_1), (n_2, m_2)]$ kills $s' = [(x_1, y_1), (x_2, y_2)]$. A necessary condition is that s and s' use the same time slot and wavelength.

Now there are two possibilities: either s and s' have the same parent or they have different parents. If they have the same parent, i.e. $[n_2, m_2] = [x_2, y_2]$, then the session $[n_1, m_1]$ would kill $[x_1, y_1]$ in (W_1, X_1), contradicting the assumption that (W_1, X_1) was feasible.

If they have different parents it must be true that $W_2[n_2, m_2] = W_2[x_2, y_2] = W_2[n_2, y_2]$ else they would have different wavebands and therefore could not contend. But then s and s' use different timebands so they do not contend. □

Using this composition, we now have the following interesting design. Let (W_1, X_1) be any feasible $N_1 \times N_1$ design with F_1 wavelengths and T_1 time slots. Let (W_{LR}, X_{LR}) be the $F_2 \times F_2$ Latin Router Design with F_2 wavelengths and 1 time slot. Then $(W_1, X_1) \otimes (W_{LR}, X_{LR})$ is a feasible $F_2N_1 \times F_2N_2$ design with F_1F_2 wavelengths and T_1 time slot. Call this the *LR extension*. Fig. 14-9 is a LR extension since the second matrix is a 2×2 Latin Router.

For instance consider the following design which is valid for F a multiple of 2 and $T = 4^k/2k$ for some k a power of 2. Set $F_2 = F/2$ and $F_1 = 2$. Use the design laid out in section 14.4.3 how to form a feasible (W_1, X_1) with 2 wavelengths and T time slots. It was shown there that $N_1^2 = (T \log 2T)\left[1 + O\left(\frac{\log\log T}{\log T}\right)\right]$. Therefore, the LR extension supports $N^2 = (.25F^2T\log T)\left[1 + O\left(\frac{\log\log T}{\log T}\right)\right]$.

R	R	R	R
R	R	b	b
R	b	R	b
R	b	b	R

$\otimes$

1	0
0	1

$=$

(R,1)	(R,1)	(R,1)	(R,1)	(R,0)	(R,0)	(R,0)	(R,0)
(R,1)	(R,1)	(b,1)	(b,1)	(R,0)	(R,0)	(b,0)	(b,0)
(R,1)	(b,1)	(R,1)	(b,1)	(R,0)	(b,0)	(R,0)	(b,0)
(R,1)	(b,1)	(b,1)	(R,1)	(R,0)	(b,0)	(b,0)	(R,0)
(R,0)	(R,0)	(R,0)	(R,0)	(R,1)	(R,1)	(R,1)	(R,1)
(R,0)	(R,0)	(b,0)	(b,0)	(R,1)	(R,1)	(b,1)	(b,1)
(R,0)	(b,0)	(R,0)	(b,0)	(R,1)	(b,1)	(R,1)	(b,1)
(R,0)	(b,0)	(b,0)	(R,0)	(R,1)	(b,1)	(b,1)	(R,1)

Figure 14-9: Composition of Two Designs

We are now in a position to present the best known design for large F and T; the construction is trivial given what we have already done. First, compose the 8×8 design on itself n times. Then compose the final answer with a $k \times k$ LR. In the end, we have an $k8^n \times k8^n$ network using $F = k2^n$ wavelengths and $T = 8^n$ time slots. Notice that $N^2 = F^2T^{4/3}$ and that $F = kT^{1/3}$. Therefore $C \geq \Theta\left(F^2T^{4/3}\right)$ for $F \geq T^{1/3}$. Also notice that this design greatly outperforms the LR for large T since then $F^2T^{4/3} >> F^2T$.

14.5 Conclusions

Using NLT-AONs for the All-to-All problem was first studied in a different context [BLM93]. There it was shown that for $F = 2$, $C \geq \Theta(T \log T)$, beating the F^2T light tree limitation; however no upper bound on C was presented.

We showed that for any F and T, $C \leq .24F^2(T+1)^2$. By combining Birk's design and the LR, we also showed that $C \geq \Theta(F^2T \log T)$ for all $F \geq 2$. In addition, for $F \geq T^{1/3}$, $C \geq \Theta(F^2T^{4/3})$. Note that surprisingly, both results are achievable even if $F << T$ as long as $F \geq T^{1/3}$. None of the AONs require wavelength changing.

The capacity of an AON can be increased by increasing T at the cost of decreasing the session bit rate $R_s = R/T$. Since $C_{LT} \leq F^2T$, increasing T increases the capacity of a LT-AON but does not increase the maximum total throughput $Z_{LT} = C_{LT}R/T = F^2R$ b/s. However Z_{NLT} increases as the session rate decreases! This is a fundamental design trade-off that does not exist in traditional multi-access networks and is an area for future study.

Appendix A

Identities, Bounds, and Approximations

The factorial function $n!$ is defined to be

$$n! = n \times n - 1 \times n - 2 \times ... \times 1 \tag{A.1}$$

The famous Sterling's formula is

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{\frac{\theta}{12n}} \tag{A.2}$$

where $0 \leq \theta \leq 1$ and therefore

$$n! \geq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \geq \left(\frac{n}{e}\right)^n \tag{A.3}$$

The last inequality will generally suffice for our purposes.

A related function is $(n)_i$, read n lower factorial i. By definition $(n)_0 = 1$, and for any other positive integer i,

$$(n)_i = n(n-1)...(n-i+1). \tag{A.4}$$

Notice that $(n)_n = n!$. Two other useful identifies are

$$\binom{n}{i} = \frac{(n)_i}{(i)_i} \tag{A.5}$$

$$(n)_b = (n)_i(n-i)_{b-i} \tag{A.6}$$

where the last one is valid for $b \geq i$. We will generally encounter ratios of the lower factorial function. In that case the following bound is useful

$$\frac{(n)_i}{(m)_i} \leq \left(\frac{n}{m}\right)^i \tag{A.7}$$

for $n \leq m$ since in this case $(n-j)/(m-j)$ is decreasing with j. If $n > m$ reverse the inequality.

Appendix B

Hypergeometric Distribution

A probability distribution used in this thesis is the hypergeometric distribution. Consider a vat with b blue balls and $M-b$ red balls. Define the random variable i to be the total number of blue balls picked out of the vat in n trials, where the balls are not replaced after each pick. Then i is given by the hypergeometric distribution,

$$p(i) = \frac{\binom{b}{i}\binom{M-b}{n-i}}{\binom{M}{n}} = \frac{\binom{n}{i}\binom{M-n}{b-i}}{\binom{M}{b}} \tag{B.1}$$

for $0 \leq i \leq \min(n, b)$ and $n \leq M - b$, where M is the total number of balls. The mean and variance of the hypergeometric are [Fel68]

$$E[i] = n\epsilon \tag{B.2}$$

$$var[i] = n\epsilon(1-\epsilon) * \left\{\frac{M-n}{M-1}\right\} \tag{B.3}$$

where $\epsilon = \frac{b}{M}$ is the probability of picking a blue ball on the first pick.

Now define the random variable j to be the total number of blue balls picked out of the vat in n trials, where the balls are replaced after each pick. Then j

is given by the familiar binomial distribution,

$$b(j) = \binom{n}{j} \epsilon^j (1-\epsilon)^{n-j} \tag{B.4}$$

where ϵ is as before. The mean and variance of the binomial are

$$E[j] = n\epsilon \tag{B.5}$$

$$var[j] = n\epsilon(1-\epsilon) \tag{B.6}$$

If the expected number of red balls chosen without replacement is small compared to the total number of red balls, the hypergeometric is well approximated by the binomial [Fel68]. That is if

$$n\frac{b}{M} << b \leftrightarrow n << M \tag{B.7}$$

then $p(i) \approx b(j)$. Below we will bound the tail of the hypergeometric. Before doing that, recall that the tail of the binomial can be bound using the Chernoff bound. That is if

$$b(j \geq k) = \sum_{j=k}^{n} \binom{n}{j} \epsilon^j (1-\epsilon)^{n-j} \tag{B.8}$$

then using the Chernoff bound,

$$\ln b(j \geq k) \leq n\left[\frac{k}{n}\ln\frac{\epsilon n}{k} + \left(1-\frac{k}{n}\right)\ln\frac{1-\epsilon}{1-\frac{k}{n}}\right] \tag{B.9}$$

for $k \geq E[j] = n\epsilon$. Also note that since $(1-\epsilon)/(1-k/n) = 1 + \frac{-\epsilon+k/n}{1-k/n}$ and since $\ln(1+x) \leq x$,

$$\ln b(j \geq k) \leq k\ln\frac{\epsilon n}{k} + (k - n\epsilon) \tag{B.10}$$

Now in order to bound the tail of the hypergeometric, we first show that the hypergeometric can be bounded above by a binomial times a correction factor. The derivation is taken from [Bol85].

$$\begin{aligned}
p(i) &= \frac{\binom{n}{i}\binom{M-n}{b-i}}{\binom{M}{b}} && \text{(B.11)}\\
&= \binom{n}{i}\frac{(M-n)_{b-i}b!}{(b-i)!(M)_b} && \text{(B.12)}\\
&= \binom{n}{i}\frac{(M-n)_{b-i}(b)_i}{(M)_b} && \text{(B.13)}
\end{aligned}$$

Separating out the first $b-n$ terms of $(M-n)_{b-i}$ using eqn. (A.6), we get

$$(M-n)_{b-i} = (M-n)_{b-n} \times (M-b)_{n-i} \qquad \text{(B.14)}$$

Also using eqn. (A.6) twice on $(M)_b$,

$$(M)_b = (M)_n(M-n)_{b-n} = (M)_i(M-i)_{n-i}(M-n)_{b-n} \qquad \text{(B.15)}$$

Plugging these in,

$$\begin{aligned}
p(i) &= \binom{n}{i}\frac{(M-n)_{b-n}(M-b)_{n-i}(b)_i}{(M)_i(M-i)_{n-i}(M-n)_{b-n}} && \text{(B.16)}\\
&= \binom{n}{i}\frac{(M-b)_{n-i}(b)_i}{(M)_i(M-i)_{n-i}} && \text{(B.17)}\\
&\leq \binom{n}{i}\left(\frac{b}{M}\right)^i\left(\frac{M-b}{M-i}\right)^{n-i} && \text{(B.18)}
\end{aligned}$$

where the inequality follows from eqn. (A.7). Now since

$$\frac{M-b}{M-i} = \frac{M-b}{M} \times \frac{M}{M-i} = \left(1-\frac{b}{M}\right)\left(1+\frac{i}{M-i}\right) \qquad \text{(B.19)}$$

we have

$$
\begin{aligned}
p(i) &\leq \binom{n}{i}\epsilon^i(1-\epsilon)^{n-i} \times \left(1+\frac{i}{M-i}\right)^{n-i} && \text{(B.20)}\\
&\leq \binom{n}{i}\epsilon^i(1-\epsilon)^{n-i} \times exp\left\{\frac{i(n-i)}{M-i}\right\} && \text{(B.21)}\\
&\leq \binom{n}{i}\epsilon^i(1-\epsilon)^{n-i} \times exp\left\{\frac{i(n-i)}{M-b}\right\} && \text{(B.22)}
\end{aligned}
$$

where the last inequality is true since $i \leq b$. Now since $i(n-i)$ is maximized at $i = n/2$,

$$
p(i) \leq \binom{n}{i}\epsilon^i(1-\epsilon)^{n-i} \times exp\left\{\frac{n^2}{4(M-b)}\right\} \qquad \text{(B.23)}
$$

Eqn. (B.23) will be useful to us at times. It can also be used to bound the tail of the hypergeometric in terms of the tail of the binomial. Specifically

$$
\begin{aligned}
\ln p(i \geq k) &\leq \ln b(i \geq k) + \frac{n^2}{4(M-b)} && \text{(B.24)}\\
&\leq k \ln\frac{\epsilon n}{k} + (k - n\epsilon) + \frac{n^2}{4(M-b)} && \text{(B.25)}
\end{aligned}
$$

where the last inequality follows from eqn. (B.10) and is valid for $k \geq E[i] = n\epsilon$. This formula can be useful when $n^2 << M$. However, when n is on the order of the $\sqrt{M}$, the bound we will now derive is better. Starting again from eqn. (B.22),

$$
\begin{aligned}
p(i) &\leq \binom{n}{i}\epsilon^i(1-\epsilon)^{n-i} \times exp\left\{\frac{i(n-i)}{M-b}\right\} && \text{(B.26)}\\
&\leq \binom{n}{i}\epsilon^i(1-\epsilon)^{n-i} \times exp\left\{\frac{in}{M-b}\right\} && \text{(B.27)}\\
&= \binom{n}{i}(\epsilon\gamma)^i(1-\epsilon)^{n-i} && \text{(B.28)}
\end{aligned}
$$

where $\gamma = exp(\frac{n}{M-b})$. Note that the random variable e^{si} takes on only positive values. Therefore by Markov's inequality and the fact that $i \geq k$ is equivalent to $e^{si} \geq e^{sk}$ for all $s \geq 0$,

$$p(i \geq k) = p(e^{si} \geq e^{sk}) \tag{B.29}$$

$$\leq \frac{\sum_{i=0}^{n} p(i) e^{si}}{e^{sk}} \tag{B.30}$$

Now using eqn. B.28, we get that for all $s \geq 0$,

$$p(i \geq k) \leq e^{-sk} \sum_{i=0}^{n} \binom{n}{i} (e^s \epsilon \gamma)^i (1-\epsilon)^{n-i} \tag{B.31}$$

$$= e^{-sk} \left(e^s \epsilon \gamma + (1-\epsilon)\right)^n \tag{B.32}$$

Let $g_i(s) \stackrel{\text{def}}{=} (e^s \epsilon \gamma + (1-\epsilon))^n$. Then the minimum of the right hand side occurs at the point

$$k = \frac{\frac{\partial}{\partial s} g_i(s)}{g_i(s)} \tag{B.33}$$

$$= \frac{\partial}{\partial s} \ln g_i(s) \tag{B.34}$$

Taking the derivative of $g_i(s)$ and solving the above equation for s gives

$$s = \ln \frac{(1-\epsilon)k}{\epsilon\gamma(n-k)} \tag{B.35}$$

which is non-negative if $n \geq k \geq \frac{n\epsilon\gamma}{1+(\gamma-1)\epsilon}$. Using this value of s,

$$p(i \geq k) \leq e^{-sk} \left(e^s \epsilon \gamma + (1-\epsilon)\right)^n \tag{B.36}$$

$$\ln p(i \geq k) \leq n \left[\frac{k}{n} \ln \frac{\epsilon \gamma n}{k} + \left(1 - \frac{k}{n}\right) \ln \frac{1-\epsilon}{1-\frac{k}{n}} \right] \tag{B.37}$$

Notice that if $\gamma = 1$, this is the Chernoff bound on the tail of the binomial.

One more step and we're done,

$$\ln \frac{1-\epsilon}{1-\frac{k}{n}} \leq \frac{k/n-\epsilon}{1-k/n} \tag{B.38}$$

and therefore

$$\ln p(i \geq k) \leq (k-n\epsilon) + k \ln \frac{n\epsilon\gamma}{k} \tag{B.39}$$

$$= (k-n\epsilon) + k \ln \frac{n\epsilon}{k} + \frac{n}{M-b} \tag{B.40}$$

which is valid for all $n \geq k \geq \frac{n\epsilon\gamma}{1+(\gamma-1)\epsilon}$. Since $\gamma = e^{\frac{n}{M-b}} \geq 1$, eqn. (B.40) is valid for all $n \geq k \geq n\epsilon\gamma$.

Bibliography

[A$^+$93] Steve Alexander et al. A precompetitive consortium on wideband all-optical networks. *Journal of Lightwave Technology*, to appear in special issue on broadband optical networks 1993.

[ABCR$^+$93] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing and scheduling algorithms for optical networks. Technical report, IBM Research Report RC 18967 (82799), June 1993.

[Aca93] Anthony S. Acampora. Intelligent optical networks: Research, education, and industrial programs at the center for telecommunications research. *Proceedings of the IEEE*, 81(1):111–131, January 1993.

[Bar93] Richard Barry. An all-optical non-blocking $M \times M$ switchless connector with $O(\sqrt{M \log M})$ wavelengths and without wavelength changers. *IEE Electronics Letters*, 29(14):1252–1254, July 1993.

[Ben65] V.E. Beneš. *Mathematical Theory of Connecting Networks and Telephone Traffic*. Academic Press, New York, 1965.

[Brg89] C. Berge. *Hypergraphs*. Elsevier Science Publishing Company, Inc., 1989.

[Ber92] Dimitri P. Bertsekas. *Linear Network Optimization, Algorithms and Codes.* MIT Press, Cambridge, MA, 1992.

[BuHa90] M.T. Busche and B. Hajek. On optical interconnections of stations having multiple transmitters and receivers per station. In *Proc. 1990 International Symposium on Information Theory and its Applications (ISITA)*, pages 967–970, November 1990.

[BH92] Richard Barry and Pierre Humblet. On the number of wavelengths needed in WDM networks. In *LEOS Summer Topical Meeting*, pages 21–22, Santa Barbara, CA, Aug. 3-5 1992. LEOS '92.

[BH93a] Richard Barry and Pierre Humblet. Latin routers, design and implementation. *Journal of Lightwave Technology*, May 1993.

[BH93b] Richard Barry and Pierre Humblet. On the number of wavelengths and switches needed in all optical networks. *To appear in IEEE Trans. on Comm.*, 1993.

[BLM93] Yitzak Birk, N. Linial, and R. Meshulam. On the uniform-traffic capacity of single-hop interconnections employing shared directional multichannels. *IEEE Trans. on Information Theory*, 39(1), January 1993.

[Bol85] Bela Bollobas. *Random Graphs.* Academic Press, London UK, 1985.

[Bra90] C.A. Brackett. Dense wavelength division multiplexing networks: Principles and applications. *Journal on Selected Areas in Communications*, 8(6):948–964, Aug 1990.

[D+93] T. Durhuus et al. Optical wavelength conversion over 18 nm at 2.5 Gb/s by DBR-laser. *IEEE Photonics Technology Letters*, 5(1):86–88, Jan 1993.

[Dra91] C. Dragone. Integrated optics NxN multiplexer on silicon. *IEEE Photonics Technology Letters*, 3(10):896–899, Oct 1991.

[E+86] Yoshima Egawa et al. A decomposition of complete bipartite graphs into edge-disjoint subgraphs with star components. *Discrete Mathematics*, 58:93–95, 1986.

[EU91] H. Enomoto and Y. Usami. The star arboricity of complete bipartite graphs. In *Graph Theory, Combinatorics, and Applications, Vol. 1*, pages 389–396. Proceedings of the Sixth Quadrennial International Conference on the Theory and Applications of Graphs, John Wiley & Sons, 1991.

[Fel68] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. John Wiley & Sons, 3 edition, 1968.

[FFP88] P. Feldman, J. Friedman, and N. Pippenger. Wide-sense nonblocking networks. *SIAM Journal of Discrete Mathematics*, 1(2):158–173, May 1988.

[Fri88] Friedman. A lower bound on strictly non-blocking networks. *Combinatorica*, pages 185–188, 1988.

[G+92a] Bernard Glance et al. Broadband optical wavelength shifter. In *CLEO '92*, page PD10, 1992.

[G+92b] Bernard Glance et al. High performance optical wavelength shifter. *IEE Electronics Letters*, 28(18):1714–1715, 1992.

[Gal68] Robert G. Gallager. *Information Theory and Reliable Communications.* John Wiley & Sons, 1968.

[Gal92a] Robert G. Gallager. Fixed composition arguments and lower bounds on error probability. Supplementary Lecture Notes for Course 6.441, Massachusetts Institute of Technology, April 1992.

[Gal92b] Robert G. Gallager. Spatial scalability of B service. Technical report, Laboratory for Information and Decision Systems, MIT, July 1992.

[GKV+90] M. Goodman, H. Kobrinski, M. Vecchi, R. Bulley, and J. Gimlett. The lambdanet multiwavelength network: Architecture, applications, and demonstrations. *IEEE Journal on Selected Areas in Communications*, 8(6):995–1004, Aug 1990.

[Goo89] M. Goodman. Multiwavelength networks and new approaches to packet switching. *IEEE Communications Magazine*, 27(10):27–35, Oct 1989.

[Gre93] Paul E. Jr. Green. *Fiber Optic Networks.* Prentice Hall, New Jersey, 1993.

[GS+88] B. S. Glance, J. Stone, et al. Densely spaced FDM coherent star network with optical signals confined to equally spaced frequencies. *Journal of Lightwave Technology*, 6(11):1770–1781, 1988.

[HCD89] G.R. Hill, P.J. Chidgey, and J. Davidson. Wavelength routing for long haul networks. In *Proceedings of ICC*, pages 734–738, June 1989.

[Hil88] G.R. Hill. A wavelength routing approach to optical communications networks. In *Proceedings of the IEEE INFOCOM*, pages 354–362, March 1988.

[Hui90] Joseph Y. Hui. *Switching and Traffic Theory for Integrated Broadband Networks.* Kluwer Academic Publishers, Massachusetts, 1990.

[Kar92] Mark Karol, April 1992. personal discussion.

[KISS88] I.P. Kaminow, P.P. Iannone, J. Stone, and L.W. Stulz. FDMA-FSK star network with a tunable optical filter demultiplexer. *Journal of Lightwave Technology*, 6(9):1406–1414, 1988.

[Lei92] Thomas F. Leighton. *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes.* Morgan Kaufmann Publishers, 1992.

[Lie88] Soung Chang Liew. *Capacity Assignment in Non-Switching Multichannel Networks.* PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1988.

[Pan92] Rajesh Pankaj. *Architectures for Linear Lightwave Networks.* PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1992.

[Pip73] Nicholas Pippenger. *The Complexity Theory of Switching Networks.* PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1973.

[PY82] Nicholas Pippenger and Andres C.-C. Yao. Rearrangeable networks with limited depth. *SIAM Journal of Algorithms and Discrete Methods*, 3(4):411–417, Dec. 1982.

[Sas93] Galen Sasaki. A linear lightwave Beneš network. *to appear in IEEE/ACM Transactions on Networks*, 1993.

[Sha49] Claude E. Shannon. Memory requirements in a telephone exchange. *Bell System Technical Journal*, pages 343–349, 1949.

[Ste90] Thomas Stern. Linear lightwave networks. Technical report, Center for Telecommunications Research, Columbia University, 1990.

[T+90] H. Toba et al. 100-channel optical FDM transmission/distribution at 622 Mb/s over 50 km using a waveguide frequency selection switch. *IEE Electronic Letters*, 26(6):376–377, March 1990.

[W+91] H. J. Westlake et al. Reconfigurable wavelength routed optical networks: a field demonstration. In *European Conference on Optical Communications*, pages 753–756, September 1991.

[WK89] S. Wagner and H. Kobrinski. WDM applications in broadband telecommunication networks. *IEEE Communications Magazine*, 29(3):22–30, March 1989.

[WK+90] A. E. Willner, I. P. Kaminow, et al. 1.2 Gb/s closely-spaced FDMA-FSK direct detection star network. *IEEE Photonics Technology Letters*, 2(3):223–226, 1990.

[WZ93] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. In *25th ACM STOC 1993*, pages 245–251. Association for Computing Machinery, May 1993.