Robustness of Bus Overlays in Optical Networks

by

Ari Levon Libarikian

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2002

© Ari Levon Libarikian, MMII. All rights reserved.

Author . . . .
Department of Electrical Engineering and Computer Science
March 1, 2002

Certified by. . . . .
Muriel Médard
Assistant Professor
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Robustness of Bus Overlays in Optical Networks

by

Ari Levon Libarikian

Submitted to the Department of Electrical Engineering and Computer Science
on March 1, 2002, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

Abstract

Local area networks (LANs) nowadays use optical fiber as the medium of communication. This fiber is used to connect a collection of electro-optic nodes which form network clouds. A network cloud is a distribution network that connects several external nodes to the backbone, and often takes the form of a star or tree. Optical stars and trees have expensive and inefficient recovery schemes, and as a result, are not attractive options when designing networks. In order to solve this problem, we introduce a virtual topology that makes use of the robustness that is inherently present in a metropolitan area network (MAN) or wide area network (WAN) (long haul network). The virtual topology uses a folded bus scheme and includes some of the elements of the real topology (architecture). By optically bypassing some of the router/switch nodes in the physical architecture, the virtual topology yields better recovery performance and more efficient systems (with respect to cost related to bandwidth and recoverability). We present a bus overlay which uses simple access nodes and is robust to single failures. Our architecture allows the use of existing optical backbone infrastructure. We consider a linear folded bus architecture and introduce a T-shaped folded bus. Although buses are generally not able to recover from failures, we propose a loopback approach. Our approach allows optical bypass of some routers during normal operation, thus reducing the load on routers, but makes use of routers in case of failures. We analyze the behavior of our linear and T-shaped systems under average use and failure conditions. We show that certain simple characteristics of the traffic matrix give meaningful performance characterization. We show that our architecture provides solutions which limit loads on the router.

Thesis Supervisor: Muriel Médard
Title: Assistant Professor

Acknowledgments

I would like to express my gratitude to my thesis advisor, Muriel Médard, whose experience, knowledge, intuition, patience, and support helped me considerably in my journey towards completing my thesis. She continuously provided ideas, thoughts, and guidance, not only with regards to my thesis, but also with respect to my entire graduate experience at MIT.

I would also like to thank Professor Vincent Chan for his advice during my time at MIT.

The students at LIDS, particularly those of Room 35-316, made my time at MIT both educational and fun, and for that, I thank them.

A very special thanks goes out to my parents, Levon and Sylva, and brother, Vicken, for their encouragement during the writing of my thesis. I couldn't have gotten this far without their love and support.

Finally, I'd like to thank my dear friend Sabina, for her wonderful support, enduring patience and affection during my Masters degree.

Contents

List of Figures

Chapter 1

Introduction

Most current high speed local area networks (LANs) use a fiber optical communications medium to connect the nodes. This fiber is used to connect a collection of nodes which perform electro-optic processing. The LANs are connected to metropolitan area networks (MANs) which themselves access the core of a long-haul network.

A LAN often takes the form of an aggregation of connections to a router. Routers are then in turn aggregated until the point of access to the backbone. These aggregations of nodes form distribution networks that often take the form of stars or trees. These networks often lack efficient recoverability from failures. Either a large amount of bandwidth is used to recover from failures, or a very complex recovery scheme is used. Figure 1-1 shows such a system.

Nodes on LANs are traditionally connected through electro-optic routers (i.e. routers that receive optical signals, but perform data processing in the electronic domain). Routers receive data in the form of individual IP (Internet Protocol) packets, possibly encapsulated in MPLS (Multiprotocol Label Switching), or other formats of high granularity, and transmit them on another fiber, or wavelength. The difference between a router and a switch is that a switch connects a wavelength (or wavelengths) on a fiber to a wavelength (or wavelengths) on another fiber. Hence, the entire *stream* of data is connected (as opposed to individual packets or units of traffic as in the case of routers).

Routers usually implement their routing patterns by using table lookups. Once
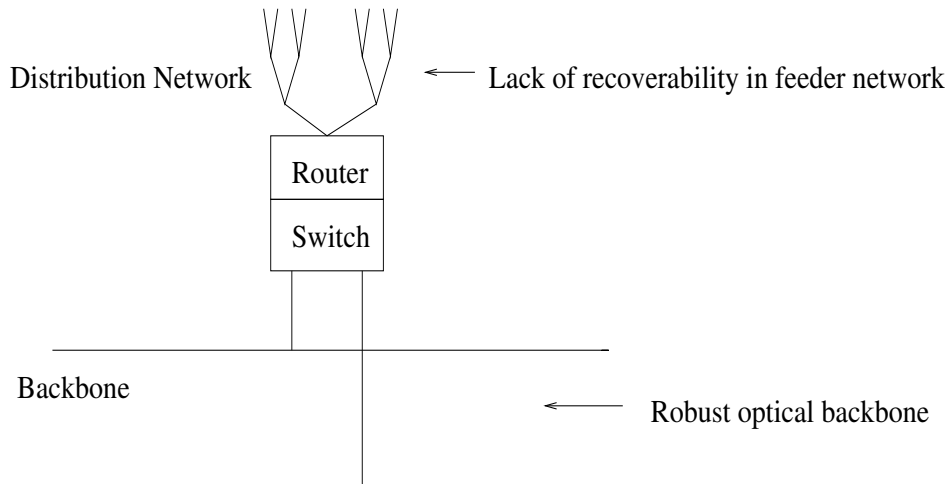
Figure 1-1: Diagram of a router/switch combination fed by a distribution network in the form of a tree.

a packet is received, the information in its header or associated label is read, and looked up in a table stored in the router's memory. The destination of the packet is determined by reading the element adjacent to the address in the table. Another function of routers is traffic grooming. Grooming corresponds to judicious assignment of low-granularity traffic onto discrete streams. It includes multiplexing different streams of traffic and effectively mixing different packets into one stream. This is often done when traffic from two (or more) fibers is to be merged and placed onto one.

Recovery is the ability of a network to resume proper operation during failure, whereas robustness is the ability of the network to maintain proper service. We do not consider the concept of time and protocols during failures, and, as a result, this thesis will look at recovery issues during failures. There are very few LAN topologies apart from rings that provide architectures with good recovery and robustness.

Optical star LANs are examples of architectures that require large amounts of bandwidth to provide robustness. The protection scheme used in a star is referred to as $1+1$ protection, in which there is a dedicated fiber link for protection of every working fiber link. A star LAN with protection usually involves two parallel stars (working and protection) that are connected by a fiber link. This fiber link often is the cause of bottleneck problems. Figure 1-2 shows this setup. In addition to
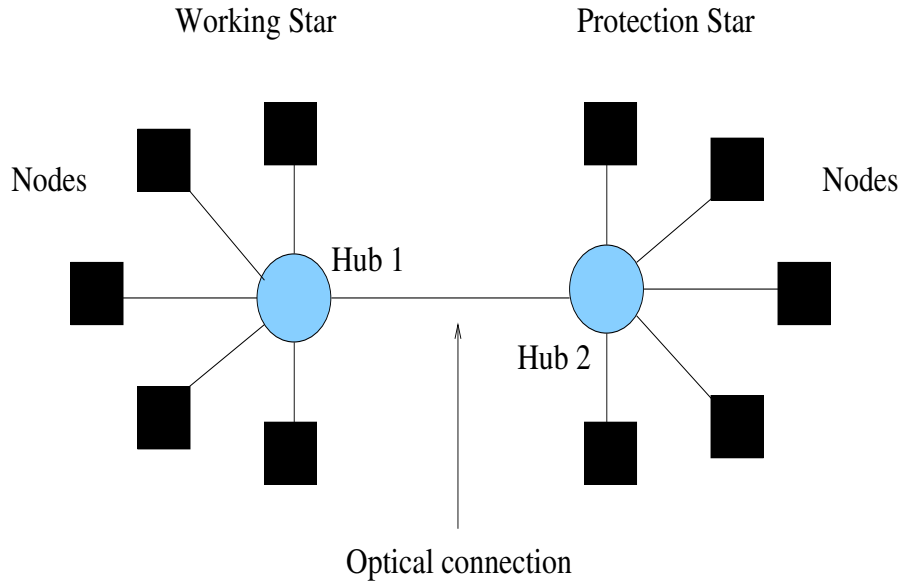
Figure 1-2: Diagram of an optical star LAN with both working and protection stars shown.

the bottleneck drawback, the architecture is not scaleable in that it is difficult to extend. For instance, if three stars were to be connected, there would be even more serious bottleneck issues with regards to the fibers interconnecting stars. Several architectures utilizing star topologies have been proposed, and have been discussed in depth ([2, 3, 4, 5, 9, 10, 11, 12, 14, 15, 17, 18, 20, 21, 22, 23, 24, 26, 27, 29, 34]).

The distribution networks (as shown in Figure 1-1) often suffer from a lack of recoverability (as explained in the previous paragraph). Because of this lack of recoverability and scaleability in distribution networks (such as stars or trees), nodes that are connected to LANs are susceptible to losing traffic in the event of a failure.

This thesis proposes the use of an overlay network to improve recoverability within a LAN/MAN while keeping cost and complexity low. An overlay is a virtual topology that is supported by a real (physical) topology. Certain nodes in the system are used, whereas certain router/switch combinations are bypassed. When a router is optically bypassed, the traffic on the fiber is not processed by the router. Figure 1-3 shows this setup. Traffic going to this switch/router node can either be processed by the router, or optically bypassed.

There are several reasons as to why optically bypassing a router is beneficial to

Figure 1-3: Diagram of a router/switch combination in the overlay. The diagram on the left represents traffic being processed by the router, whereas the diagram on the right represents the router being optically bypassed.
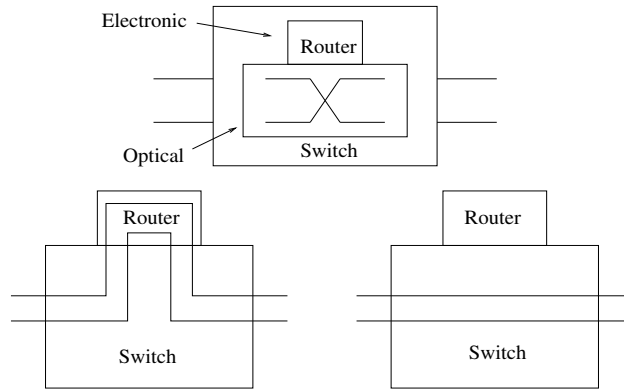
the system design. Firstly, because each router in the system is fed by a distribution network, often in the form of a tree (as shown in Figure 1-1), it is clear that the utilization of the router needs to be minimized. Since routers have a finite, and usually small, number of slots in which to insert blades, an expensive and difficult upgrade, requiring a physical replacement of the bay at the router (known as a forklift upgrade), is almost always necessary after the number of connections in the router reaches a certain threshold value. This threshold will undoubtedly be reached as the capacity of the optical switch to which the router is connected continues to rise at rates greater than the rate at which the electronic capacity of the router can grow. The necessary forklift upgrades are usually extremely expensive, and cause traffic delays and loss. As a result, it is desirable to reduce the amount of traffic that is actually processed by a given router. Secondly, since the same router can be, and almost always is, used in *several* virtual topologies, it is imperative to reduce the total traffic processed by a given router *per* virtual topology. This can be done by optically-bypassing the traffic that does not need to be handled by the router.

Our implementation of the virtual topology is done through the use of a folded bus. The folded bus is an overlay requiring a fiber (or fibers), with a wavelength, or group of wavelengths, running through all the nodes (access ports and switches/routers) in each direction. The idea here is that the nodes should be able to both "place" their data onto the bus, as well as "collect" it. Figure 1-4 represents a $2N$-node, 3-router

14
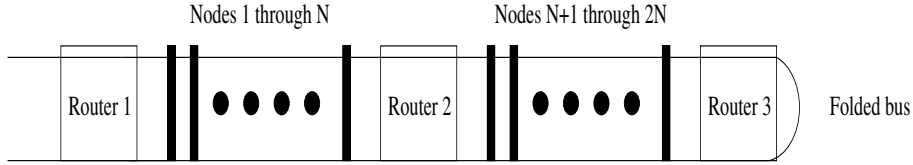
Figure 1-4: Diagram representing a system with $2N$ nodes and 3 active routers.
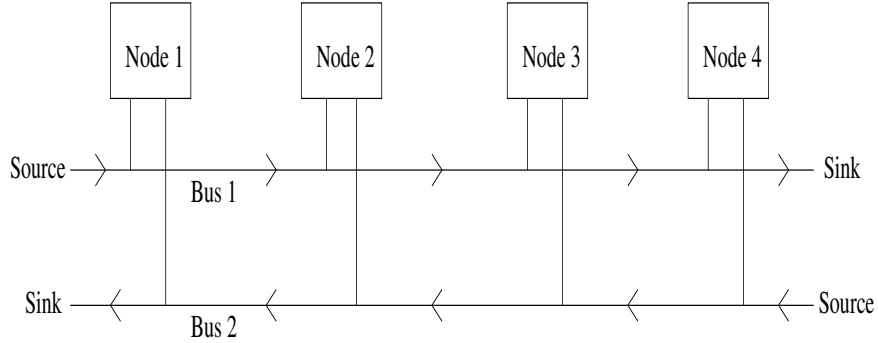
Figure 1-5: Diagram representing a 4-node dual bus system.

system employing a folded bus.

Folded buses have been used in other systems in previous architectures. The problem with bidirectional buses using optical fiber is that the use of bidirectional taps greatly reduces the distance that data can be transmitted without the need for regenerators/amplifiers. Hence, this problem is solved by the introduction of two parallel unidirectional buses, each connected to unidirectional taps, and carrying data in a single direction. Figure 1-5 shows a dual bus system.

Distributed Queue Dual Bus (DQDB) is an example of a folded bus system that is used to interchange data among nodes. DQDB has been analyzed extensively ([1, 7, 8, 13, 16, 25, 31, 32]). It is a system that uses a distributed queueing algorithm called *queued-packet distributed-switch*. It is implemented on two unidirectional folded buses, each carrying data in opposite directions. The nodes in the system tap the buses in order to place data on and collect data from the backbone. Each node has a buffer in which a distributed queue is set up.

HLAN (Helical Local Area Network) is another example of a bus system. HLAN may be constructed as a helical ring structure or a single folded bus and is used mainly to maintain fair utilization of bandwidth among all users. HLAN is implemented using
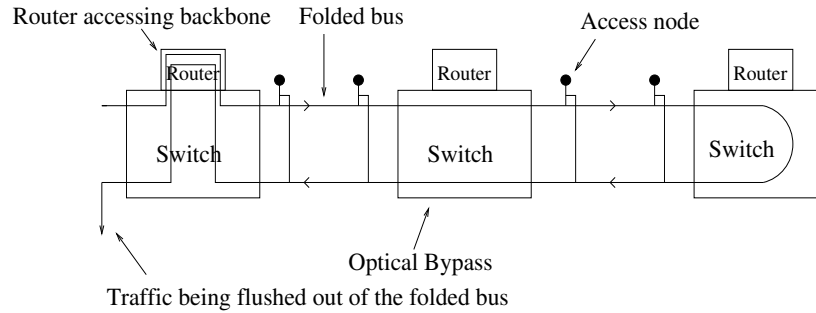
Figure 1-6: Diagram of a 3-router linear network with a folded bus, and 4 access ports.

a single folded unidirectional bus (as opposed to the two buses used in DQDB). Each node has three connections to the backbone. Two of the connections are bidirectional, whereas one of them is unidirectional. The unidirectional connection is called the RCV segment, and is the connection through which data is received. One of the bidirectional connections is the Guaranteed Bandwidth (GBW) segment, while the other bidirectional connection is the Bandwidth on Demand (BOD) segment.

ORMA (Optical Reservation Multiple Access) is another scheme that uses an optical folded bus as its backbone. ORMA is a highspeed LAN/MAN protocol that is fast and simple. Reservations are made by different users requiring bandwidth through simple hardware circuits. Results have shown that performance of ORMA is independent of the size and complexity of the network.

Besides HLAN, ORMA, and DQDB, other dual bus schemes have been studied ([6, 19, 28, 30, 33]).

A major drawback of these systems is the lack of recoverability. A virtual topology utilizing switches and routers in the system ensures recoverability and topology flexibility in the network. At the same time, these routers can access the backbone themselves.

Figure 1-6 shows an example of a linear system with a folded bus as its backbone. The middle router is optically bypassed, and data is flushed out of one end of the folded bus. Folded buses have the advantage that no routing is needed to transmit the data.

Access nodes are easily rented or leased by a user, and are desired because of their

simplicity and low cost in comparison to routers or optical switches.

The architecture presented here combines the advantages of folded buses, access ports, optical bypass in switches, and robustness of routing during failures. Although folded bus overlays have been limited to linear topologies, we introduce T-shaped systems as well (introduced later). The main idea behind the virtual topology is the flexibility introduced by using routers in combination with optical switches. The overlay networks provide increased flexibility to the local area network. A problem with using routers for recovery purposes, is that the load they support may exceed normal limits. Although networks often operate for long periods of time without failures, occasional failures do occur, and it is important to measure the performance of a system in these cases. Hence, we evaluate a system's recovery performance not only in the *time-average* sense, but also under worst-case *failure* conditions.

Routers process traffic in the electronic domain, and their performance degrades as the load supported increases towards a threshold value. As the amount of traffic processed by a router increases, more and more "packets" or units of traffic are dropped or lost. Hence, a good metric for performance is the load per router. The load per router gives us an indication of the average congestion in the network, as well as average local congestion around a router.

We compare the performance of various linear systems, as well as that of various T-shaped systems, which we describe in the next section. Linear systems are comprised of a set of colinear router/switch pairs and access nodes. The entire system lies on a line and folded buses are adjacent to each other. T-shaped systems are a special type of non-linear systems in which a central node may serve as a hub for up to three folded buses.

The thesis is organized as follows. Chapter 2 describes the system architectures and topologies examined, namely the structure and operating schemes of linear and T-shaped topologies. Chapter 3 presents the analysis of the performance of these systems as well as several relations and bounds. Chapter 4 introduces ratios that characterize the network traffic and present the simulation results for both the linear and T-shaped systems. Chapter 5 presents concluding remarks and directions for

further research.

Chapter 2

System Architecture

The access ports share resources such as a wavelength or a fiber. The nodes represent entities that "access" the fiber both for placing data to be transmitted to another node (distribution portion), and for collecting data transmitted from other nodes (collection portion). This structure is called a folded bus. The upper portion of the folded bus (shown in Figure 2-1), represents the distribution part, whereas the lower portion represents the collection part. As a result, any traffic meant to be transmitted from node i to node j is simply placed on the distribution portion at node i, and collected by node j on the collection portion.

Traffic is defined in terms of an integer number of units, such as packets or cells. For instance, if u units of traffic or data are to be transmitted from node i to node j in a given interval of time, then the element in the ith row and j^{th} column of the traffic matrix will be u. We define the load on a router as the statistical mean of the number of packets transmitted by it. Routers can handle less traffic than switches, but routers may not need to process all incoming traffic. Routers in the system support traffic emanating from or destined to nodes not included in the overlay (external traffic),

Nodes 1 through 2N

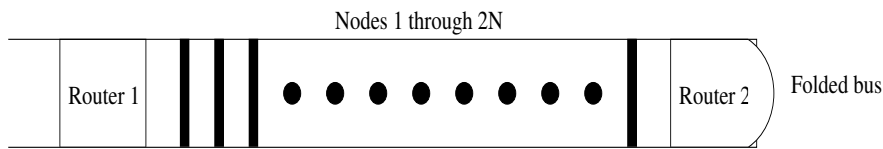Router 1   • • • • • • • • •   Router 2   Folded bus

Figure 2-1: Diagram of a 2-router folded bus overlay.

and as a result, may have to carry a large burden. Thus, it is desirable that the amount of load placed on them during recovery situations be minimized. The routers of this overlay can and do support all the external traffic and a portion of the internal traffic (during both failure and non-failure states). Figure 1-3 shows a typical system in which the router associated with a switch is not bypassed, as well as a system in which the router is bypassed. Once the decision has been made to make a router active (i.e. not optically bypassed; processing the traffic) in a local area network system, the decision of which traffic, and how much traffic it will handle also has to be made. These are the main issues of the analysis in this thesis. Figure 2-1 shows a system with the number of routers in the system set to $N_R = 2$. It is important to note here that N_R represents the number of routers that are *active* in the system. A router/switch combination may be physically present in the system, but if the switch is configured to bypass the traffic optically, then the router doesn't affect the traffic flow (i.e. the optical bypass would act as an extension of the folded bus). In the case of N_R active routers, there could be up to $N_R - 1$ folded buses (one between every pair of successive routers). It is this value, N_R, whose choice drastically affects system performance in both the *time-average* scenario and the *failure* scenario.

The proposed architecture has a recovery scheme on the folded bus. Let us assume that a failure occurs on link l (the link connecting nodes $l - 1$ and l). The system recovers as follows:

- Node $l - 1$ "connects" the two terminals of the folded bus so that all the data that was to be transferred across link l is now sent back, in the reverse direction, across link $l - 1$. This assures that the router immediately to the left of link l can access the data meant to be transferred across the link, and send it to routers on the other side of the link, via external connections.

- Node l "connects" the two terminals of the folded bus so that all the data meant to cross the link is now looped back, as in the previous case.

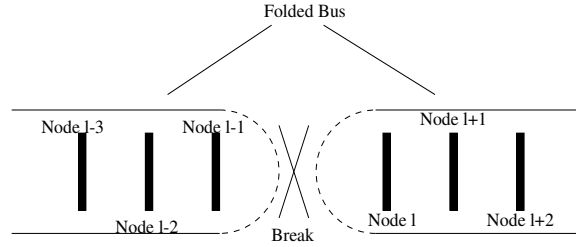Figure 2-2 shows this loopback effect.

Figure 2-2: Diagram representing the loopbacks performed on both sides of a folded bus during a failure (e.g. a break in the fiber).

Using these loopbacks, the routers now become active participants in the recovery-scheme, as all the data that was meant to cross the now-failed link will be sent through at least 2 routers.

Varying the number of routers present in a system causes different changes in system-performance. Firstly, in the *no failure* case, the routers have two purposes:

- To support traffic which does not have both a source and a destination on the same folded bus.

- To support traffic being sent between two nodes not located on the same folded bus.

Thus, when $N_R = 2$, the only routers in the system are end-routers, and the second responsibility above becomes irrelevant. However, when $N_R > 2$, the routers work to carry traffic from one folded-bus to another. In the *failure* case, the routers have an additional responsibility:

- To provide new recovery paths and to support all re-routed traffic through these new paths.

It is this third purpose that can increase load on the routers in the case of failure.

As the number of routers present in the system increases, the average amount of traffic to be handled *per router* decreases. However, even when there is no link failure in the system, these additional routers are transferring data from one folded bus to another (something the $N_R = 2$ system doesn't suffer from), and as a result, the amount of average load per router is increased. We explicitly address this trade-off.
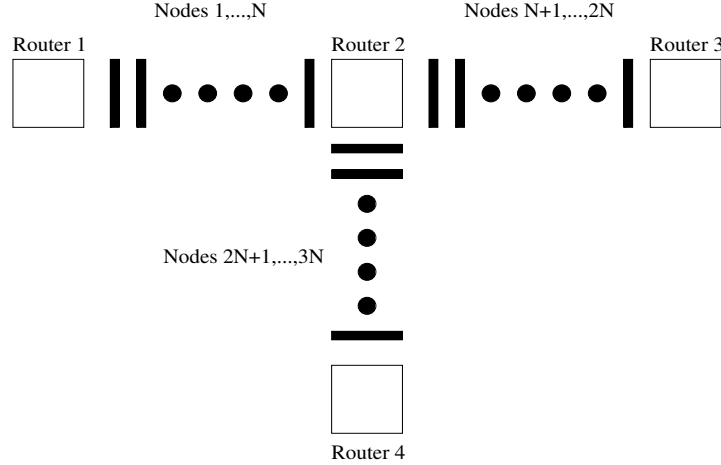
Figure 2-3: Diagram representing the general T-shaped system.

We may extend the concept of traditional linear architectures by introducing T-shaped systems. These systems are formed with four routers, one of which is a central router, with the other three each adjacent to it. N nodes separate each pair of routers. Nodes 1 to N in these systems represent the nodes in the first region, nodes $N+1$ to $2N$ represent the nodes in the second region, and nodes $3N$ to $3N+1$ represent the nodes in the third region. Figure 2-3 shows this general system. Figures 2-4, 2-5 and 2-6 show this system in different forms.

We consider four main variations, which we discuss below.

System A represents the system with 3 folded buses. This means that the central router is connected to every other router through a folded bus. Each folded bus is devoted entirely to the nodes between central router and the corresponding adjacent router. As a result, the central router is a junction where three folded buses meet. It is clear then that any traffic meant to go from one node to another node in another leg of the tree needs to be routed. This traffic "crosses over" the central router (in the physical topology), and thus is called the *cross-over* traffic. In System A, router 2, the central router, will carry all of the cross-over traffic. Figure 2-4 shows System A.

System B is another variation of the T-shaped system. In System B, there is one folded bus going through all four routers. However, as evident in Figure 2-5, this folded bus is not sufficient to transmit all the internal traffic. Specifically, all the
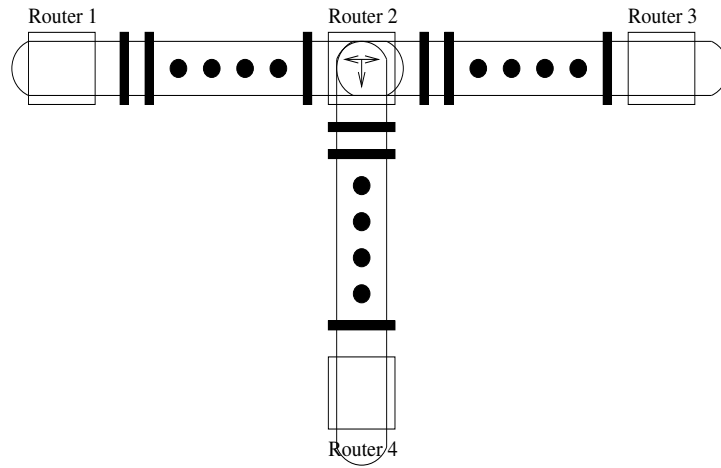
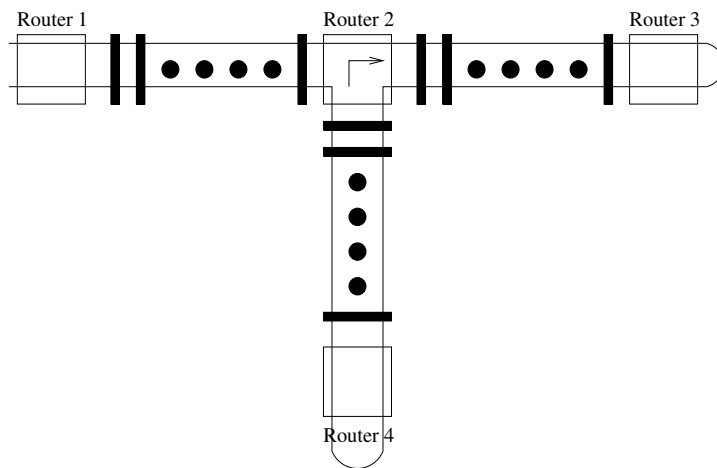Figure 2-4: Diagram representing the general T-shaped system with 3 folded buses.

Figure 2-5: Diagram representing the general T-shaped system with 1 folded bus.
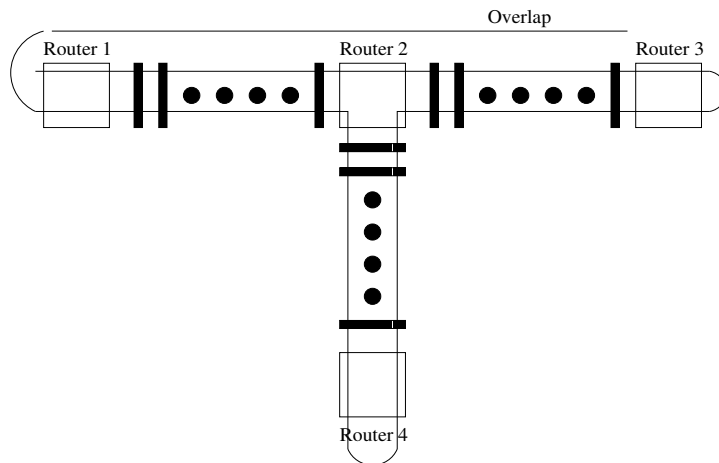
Figure 2-6: Diagram representing the general T-shaped system with an *extended* folded bus.

traffic meant to go from nodes in the third region to nodes in the second region cannot be transmitted through the single folded bus (unless the folded bus is extended, but this case will be considered in a later system). Hence, the central router is devoted solely to the transmission of data from region 3 to region 2 (but *not* vice-versa).

System C represents a variation of System B in which the central router is not used to transmit any traffic. The single folded bus is extended from the first to the third router, as shown in Figure 2-6. It is clear now that in the case of no failure, System C utilizes no routers for internal traffic.

System D represents a hybrid of Systems B and C. In effect, the central router is now used to transmit only a fraction of the total traffic meant to go from nodes in region 3 to nodes in region 2. The rest of this traffic will be sent over the extended folded bus. It is clear that in System D, less bandwidth is used on the "extension" portion of the folded bus, yet more load is applied to the central router.

System A routes all cross-over traffic, whereas System B only routes part of the cross-over traffic. All other traffic (non-cross-over and external) is routed the same way in both systems. Hence, we see that the average load per router is at least as high in System A as in System B. We may also see that Systems B, C, and D only differ in the proportion of cross-over traffic from the lower leg to the right leg routed through the central router (Note: This is related to the share-factor, to be defined

24

later). Since added traffic on the folded bus does not increase the average load per router, we see that the average load per router in System B is greater than that of System D which, in turn, is greater than that of System C.

When considering the performance of these systems, we will assume that the probability of link failure is uniform throughout the system (i.e. there is no bias towards one link failing more than another).

Chapter 3

Analysis

In this section, the load equations are outlined for two different linear systems ($N_R = 2$ and $N_R = 3$), as well as the T-shaped systems.

In the case of the linear systems, it is assumed that a $(K + 1)\text{x}(K + 1)$ traffic matrix (T) is available, where $K = N(N_R - 1)$, the total number of nodes in the system, and that there are N nodes per region. t_{ij} represents the traffic to be sent from node i to node j. The $(K + 1)^{st}$ row and column represent the external traffic (emanating from or destined to nodes outside the LAN).

Let S_i represent the total amount of traffic to be processed by router i (measured in units).

For $N_R = 2$:

In the case of no failure, we have

$$S_1 = S_2 = \tfrac{1}{2} \sum_{i=1}^{K} (t_{i,K+1} + t_{K+1,i}),$$

where the summation above represents all the traffic destined to, or arriving from the *external* node. It is important to note that, in the case of low or zero external traffic, the routers in this system have almost no load on them.

In the case of failure of link l:

$$S_1 = \sum_{i=1}^{l-1} (t_{i,K+1} + t_{K+1,i}) + \sum_{i=1}^{l-1} \sum_{j=l}^{K} (t_{ij} + t_{ji})$$

and

$$S_2 = \sum_{i=l}^{K}(t_{i,K+1} + t_{K+1,i}) + \sum_{i=1}^{l-1}\sum_{j=l}^{K}(t_{ij} + t_{ji}).$$

In the above equations, the first term represents traffic to be carried to and from the external node to the partition of the region accessible by the router. In other words, router 1 deals with the traffic meant for all the nodes to the *left* of link l, and router 2 deals with the traffic meant for all the nodes to the *right* of link l.

In addition to this traffic, the routers also have to process the traffic that is to be transmitted from one node on the folded bus on one side of the failure to another node on the folded bus on the other side of the failure. As discussed before, this traffic gets looped back to the router closest to it.

For $N_R = 3$:

In this scenario there are now 2 regions to deal with. In the case of no failure, we have:

$$S_1 = \sum_{i=1}^{N}(t_{i,K+1} + t_{K+1,i}),$$
$$S_2 = \sum_{i=1}^{N}\sum_{j=N+1}^{K}(t_{ij} + t_{ji}) + \sum_{i=N+1}^{K}(t_{i,K+1} + t_{K+1,i}) \text{ and}$$
$$S_3 = 0.$$

In the above equations, S_1 represents all the external traffic to the left-most region. S_2 represents all the external traffic to the right-most region plus the cross-over traffic (i.e. traffic meant to be transmitted from one region to the other). As a result, the right-most router has no traffic to process. It should be noted that this traffic can be balanced better, but, owing to linearity, this clearly will not affect the average load per router. In the case of a failure in link l:

For $l \leq N$:

$$S_1 = \sum_{i=1}^{l-1}\sum_{j=l}^{K}(t_{ij} + t_{ji}) + \sum_{i=1}^{l-1}(t_{i,K+1} + t_{K+1,i}),$$
$$S_2 = \sum_{i=l}^{N}\sum_{j=N+1}^{K}(t_{ij} + t_{ji}) + \sum_{i=l}^{N}(t_{i,K+1} + t_{K+1,i}) \text{ and}$$
$$S_3 = \sum_{i=N+1}^{K}\sum_{j=l}^{l-1}(t_{ij} + t_{ji}) + \sum_{i=N+1}^{K}(t_{i,K+1} + t_{K+1,i}).$$

Likewise, for $l \geq N$:

$$S_1 = \sum_{i=1}^{N}\sum_{j=l}^{K}(t_{ij} + t_{ji}) + \sum_{i=1}^{N}(t_{i,K+1} + t_{K+1,i}),$$

$$S_2 = \sum_{i=1}^{N} \sum_{j=N+1}^{l-1}(t_{ij} + t_{ji}) + \sum_{i=N+1}^{l-1}(t_{i,K+1} + t_{K+1,i}) \text{ and}$$

$$S_3 = \sum_{i=l}^{K} \sum_{j=1}^{l-1}(t_{ij} + t_{ji}) + \sum_{i=l}^{K}(t_{i,K+1} + t_{K+1,i}).$$

In the above equations, we see that, in the case of failure, the load is redistributed. The cross-over traffic used to be sent over 1 router (the middle one), but now is sent through 2 routers (the two end-routers). At the same time, depending on where the link failure occurs, different routers have to shoulder different proportions of the load.

In analyzing the T-shaped systems, we show here the equations pertinent to System A. Under non-failure conditions, the following equations hold:

$$S_1 = \sum_{i=1}^{N}(t_{i,3N+1} + t_{3N+1,i}),$$

$$S_2 = \sum_{i=1}^{N} \sum_{j=N+1}^{3N}(t_{ij} + t_{ji}) + \sum_{i=N+1}^{2N} \sum_{j=2N+1}^{3N}(t_{ij} + t_{ji}),$$

$$S_3 = \sum_{i=N+1}^{2N}(t_{i,3N+1} + t_{3N+1,i}) \text{ and}$$

$$S_4 = \sum_{i=2N+1}^{3N}(t_{i,3N+1} + t_{3N+1,i}).$$

Note that the system is symmetric, so a failure in any one of the three regions will yield the same equations (subject to a rotation).

Let us consider the case where the failure is between nodes l and $l + 1$ in region 1. We have:

$$S_1 = \sum_{i=1}^{l}(t_{i,3N+1} + t_{3N+1,i}) + \sum_{i=1}^{l} \sum_{j=l+1}^{3N}(t_{ij} + t_{ji}),$$

$$S_2 = \sum_{i=l+1}^{N} \sum_{j=N+1}^{3N}(t_{ij}+t_{ji}) + \sum_{i=N+1}^{2N} \sum_{j=2N+1}^{3N}(t_{ij}+t_{ji}) + \sum_{i=l+1}^{N}(t_{i,3N+1}+t_{3N+1,i}),$$

$$S_3 = \sum_{i=N+1}^{2N}(t_{i,3N+1} + t_{3N+1,i}) + \sum_{i=1}^{l} \sum_{j=N+1}^{2N}(t_{ij} + t_{ji}) \text{ and}$$

$$S_4 = \sum_{i=2N+1}^{3N}(t_{i,3N+1} + t_{3N+1,i}) + \sum_{i=1}^{l} \sum_{j=2N+1}^{3N}(t_{ij} + t_{ji}).$$

We may derive a break-even point between Systems A and B which represents the failure probability ("break-even" probability) at which the two systems exhibit the same average router load. Because System A routers are under non-zero load even under no failure, we would expect that under very small values of p_f, System B would exhibit better performance (i.e. lower load in System B due to the central router in System A). On the other hand, System A's load is distributed among 3 routers, whereas that of System B is distributed among 2 routers. Hence, under

certain conditions, the average load *per* router should be lower in System A than in System B.

Let F_C represent the amount of cross-over traffic in System A under failure conditions, let F represent the amount of traffic that has its transmission path changed due to a failure, let C represent the total cross-over traffic, under no failure conditions, and let p_f represent the total probability of failure in the system (i.e. the probability that at least one link is down in the system). We have:

$$A_{load} = \tfrac{1}{3}(C(1 - p) + pF + pF_C) \text{ and}$$
$$B_{load} = \tfrac{1}{2}pF.$$

For $A_{load} > B_{load}$,

$$p_f < \frac{C}{\frac{1}{2}F + C - F_C} = p^*.$$

Here, p^* represents the break-even point of p_f.

For $p_f \geq p^*$, $A_{load} \leq B_{load}$, whereas for $p_f \leq p^*$, $A_{load} \geq B_{load}$.

Let us assume that $t_{ij} = k \; \forall \; i \neq j$, and that System A has $N_R = 3$.
We find that

$$p^*(N) = \frac{3N^2}{13N^3 + 12N^2 - N}.$$

Hence, for $N = 2$ (i.e., 2 nodes per region),

$$p^*(2) = \tfrac{2}{25},$$

whereas for $N = 10$,

$$p^*(10) = \tfrac{10}{473}.$$

As the number of nodes in the system increases, p^* drops, and, as a result, System A becomes more and more attractive (for a given p_f). We see that for any reasonable value of N, the break-even probability is very high - orders of magnitude higher than a typical failure probability.

Chapter 4

Simulation Results

4.1   Simulation Rules

A $(K + 1)\text{x}(K + 1)$ traffic matrix T is assumed to represent all the traffic that is relevant to the system. The element t_{ij} represents the amount of traffic to be sent from node i to node j. The $(K + 1)^{st}$ row and column represent all the traffic to be sent to and received from any node outside of the system (i.e. the *external* traffic). The diagonal elements of this matrix represent all the traffic to be sent from a node to itself. We are going to assume that there is some type of internal processing at these nodes, and that the folded buses are not used to process this traffic. Figure 4-1 shows this structure.

Each nonzero element of T is assumed to have a probability distribution. For the purposes of the simulations in this thesis, all the distributions are assumed to be of the form:

$$t = \lceil x \rceil - 1, \text{ where } p(x) = \lambda \ e^{-\lambda x} \text{ for } 0 \leq x < \infty.$$

As a result, the entries in the matrix are integers that have distributions that are *quasi-exponential*. This distribution is chosen because of the need for a nonnegative, integer, memoryless random variable. The traffic incident for a given node-pair is often modelled to be a memoryless process. As a result, this quasi-exponential distribution most closely matches the requirements.
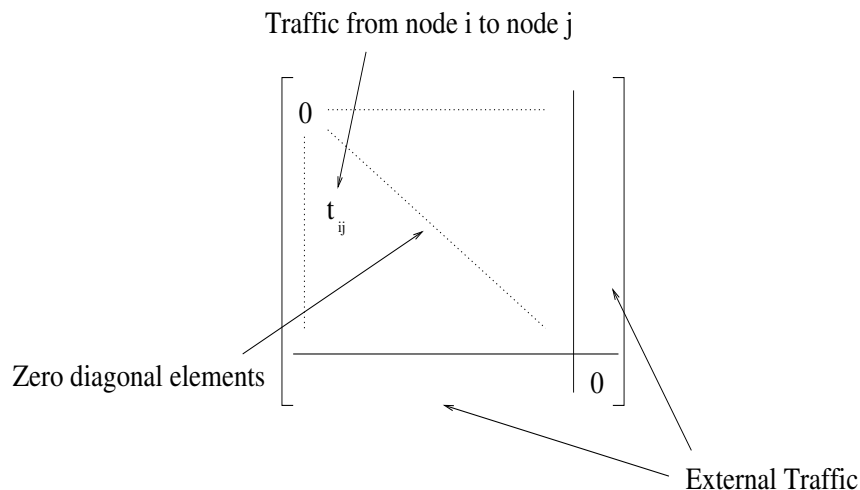
Figure 4-1: General structure of the traffic matrix T.

The parameter λ is chosen to vary across the matrix. We assume that it is more likely that two neighboring nodes send each other traffic than two nodes far from each other, hence the parameter is chosen so that nodes closer together will have a higher traffic rate than nodes farther apart. We therefore choose

$$\lambda_{ij} = a \mid i - j \mid,$$

where a represents positive a scale factor. The *external* traffic is also chosen to be exponentially distributed, but with a constant parameter λ across the nodes, that varies only with the total number of nodes in the system (and a positive scale factor):

$$\lambda_{ij}(K) = \frac{b}{K},$$

where b is another scale factor. $t_{K+1,K+1}$ is chosen to be zero, for obvious reasons.

Hence, the overall traffic matrix T is a $(K+1)\text{x}(K+1)$ matrix with zero diagonal elements, and *quasi-exponentially* distributed non-diagonal elements. For all the elements in the upper-left K x K matrix, the parameter λ_{ij} varies with i and j, whereas for all the elements in the last row and column (excluding $t_{K+1,K+1}$), λ is constant with respect to matrix coordinates.

4.2 Simulation Results of Linear Systems

During normal (*non-failure*) conditions, the system operates as follows:

- All nodes place the data they want to send onto the folded bus.

- All routers on the folded bus process data to be sent from one node in the LAN to another.

- All routers on the folded bus process data to be sent from external nodes to nodes in the LAN, and *vice-versa*.

Let us consider the case of a link failure. Let N now represent the number of nodes between any two successive routers on the folded bus. We will say that all nodes lying between the same pair of active routers represent a *region*. Hence, there are $N_R - 1$ regions, of N nodes each, yielding a total of $N(N_R - 1)$ nodes in the system. A *link* connects successive nodes, and nodes to neighboring routers. It is clear then, that there are $(N+1)(N_R-1)$ links in the system. We will assume that the probability of link failure is uniform throughout the system (i.e. there is no bias towards one link failing more than another).

Let p be the probability that a link between two successive nodes *not separated by a router* fails. Then, the probability that there is a failure between two nodes that *are* separated by a router is $2p$ because this section is comprised of two links and the links are dependent to the extent that only one link can fail at a time (multiple link failures are not considered here).

If p_f represents the probability of any link failure in the system, then

$$p = \frac{p_f}{(N+1)(N_R-1)}.$$

Thus, the value of p may be chosen such that the total probability of error equals p_f. When a link failure occurs, it is clear that no traffic can be sent across that link (at least until it recovers).

Let us assume that a failure occurs on link l (the link connecting nodes $l-1$ and l). The system recovers as follows:

- Node $l-1$ "connects" the two terminals of the folded bus so that all the data that was to be transferred across link l is now sent back, in the reverse direction, across link $l-1$. This assures that the router immediately to the left of link l can access the data meant to be transferred across the link, and send it to routers on the other side of the link, via external connections.

- Node l "connects" the two terminals of the folded bus so that all the data meant to cross the link is now looped-back, as in the previous case.

Figure 2-2 shows this loopback effect.

Using these loopbacks, the routers now become active participants in the recovery-scheme, as all the data that was meant to cross the now-failed link will be sent through at least 2 routers.

The first simulation tests the relative performance of two basic linear systems - the first of which is the $N_R = 3$ system (i.e. two end-routers and a router in the middle of the system), and the second of which is the $N_R = 2$ system (i.e. two end-routers). From now on, System A will represent the $N_R = 3$ system, and System B will represent the $N_R = 2$ system. The *average load* on the routers will serve as the measure of performance (i.e. the lower the average load on the routers, the more efficient and desirable the system). The average load will be defined as the statistical mean of the number of units (or packets) of traffic that is routed in a system per router. Depending on the incident traffic, the systems will perform differently according to this metric.

Since System A has more routers than System B, it may appear that the average traffic handled per router is less in System A, than in System B (because the total traffic is the same in both systems). However, the middle router in System A has non-zero load even in the case of *no failure*. As a result, the traffic that is simply transported down the folded bus from one node in the first region to another node in the second region will be processed by the middle-router. This load is not present in System B. Hence, we see a tradeoff arising between Systems A and B. As a result, we need to measure how the relative performance of the systems varies with respect

to a variable measuring the cross-over traffic. Let

$$R = \frac{Intra-region\ Traffic}{Inter-region\ Traffic}.$$

R measures the ratio of non-cross-over traffic to cross-over traffic. One would expect that, as this ratio increases, the performance of System A relative to that of System B improves. It should be made clear that R represents a statistic that extracts most of the pertinent information from the traffic matrix T. R can be viewed as a simplification of considering the entire traffic matrix. Let

$$\rho = \frac{A_{load}}{B_{load}},$$

where A_{load} represents the average load per router in System A, and B_{load} represents the average load per router in System B. We can use ρ as a measure of the relative performance of the two systems, and see how it varies with R. The main results of the simulation are outlined and described below.

- Because R does not capture all the information in the matrix, it does not fully determine the value of ρ. In other words, the simulations may, and in fact do produce multiple values of ρ for the same value of R. This is because different matrices with the same value of R may have different distributions of traffic across the matrix, resulting in different recovery mechanisms, and overall load on the routers.

  However, for a very special case of $K = 4$ and cross-over traffic restricted to 1 (2 nodes per region, and a total of 1 packet crossing over), the randomness is lost because we have $R = t_{12} + t_{21} + t_{34} + t_{43}$. In this case, two matrices with the same value of R have the same traffic distribution across the matrix.

- In general, as R increases, ρ decreases. Altough R is a simplification of the entire matrix T, it is still a good metric.

  Regardless of the value of p_f, the value of ρ (ratio of average load of System A to that of System B) drops as R increases. Since the middle router in System A is always processing cross-over traffic, as the relative amount of this traffic

is decreased (with respect to the local traffic), the average load in System A with respect to that of System B drops, as expected. Figure 4-2 shows the ρ vs. R plot for a given system (whose details are outlined on the plot itself) with $p_f = 10^{-3}$. In typical local area networks, normal values of p_f are much smaller than 10^{-3}. This is because a typical optical network transmitting information using the SONET (Synchronous Optical Network) standard only fails once or twice every 2 or 3 years, and of those failures, only a very small percentage cannot be corrected within the 50ms SONET recovery time. Although more reasonable values of p_f are orders of magnitude smaller, the qualitative results are unaffected. It is interesting to note that the ρ reaches values on the order of 10^2 or 10^3. This can be explained as follows. Since the probability of link-failure is so low, and the external traffic flow has a lower rate than the internal traffic flow, System B's end routers are rarely used. However, the middle-router in System A is always being used to process cross-over traffic. Hence, System B is only processing significant traffic 0.1% of the time, whereas System A is always processing significant traffic. This accounts for the extremely high values of ρ. As R increases, most of the traffic remains local (in the same region as it was generated), and this deficiency of System A is not exploited as often.

- If the number of nodes, and external traffic rate are chosen properly, the graph of ρ vs. R becomes decomposed into two curves, one representing the cases with zero external traffic, and the other representing the cases with positive external traffic. Figure 4-3 shows this effect. We see in Figure 4-3 that there is a general monotonically decreasing characteristic for $\rho > 10^5$, as well as a characteristic at the bottom of the graph around where $\rho << 10^5$. Since the external traffic flows to each node are represented by I.I.D. random variables, as the number of nodes are increased, the probability that there is no external traffic decreases. Alternatively, for a fixed number of nodes, as the external traffic rate increases, the probability of no external traffic decreases. The following equation summarizes these statements.
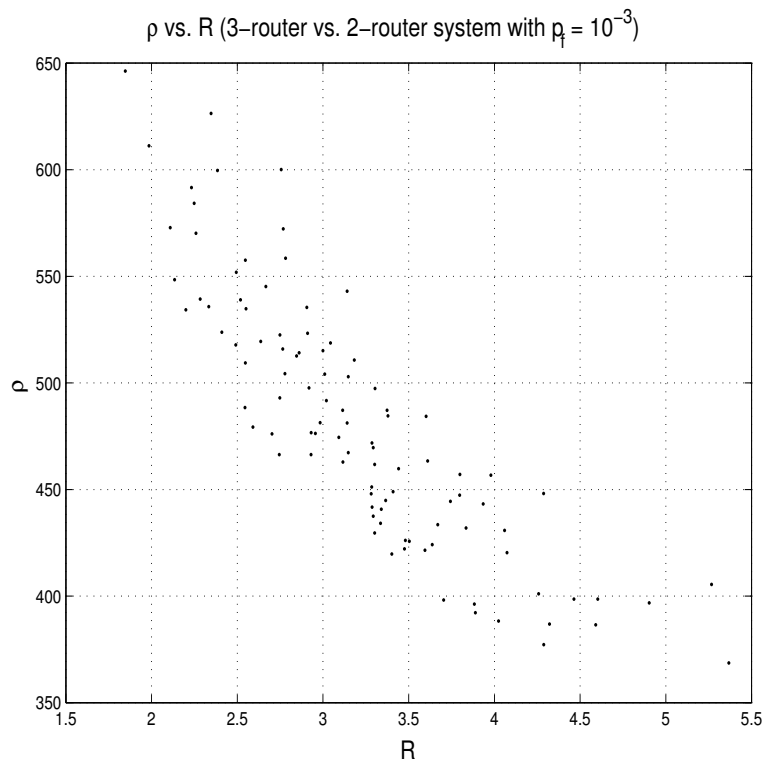
Figure 4-2: Graph representing the general monotonically-decreasing relationship between ρ and R.
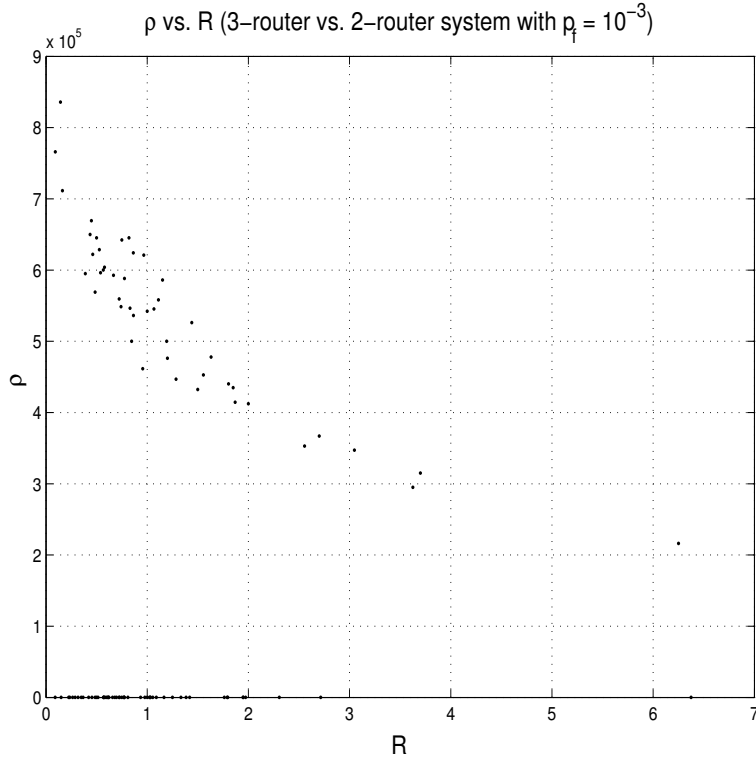
Figure 4-3: Graph representing the "splitting" of the data points into two categories.

Let p_e represent the probability that there is no external traffic sent to or from the system, and let p_{e1} represent the probability that there is no external traffic sent to a given node in the system. Then,

$$p_e = (p_{e1})^K,$$

where K here is the total number of nodes in the LAN.

If the values of K and p_{e1} are chosen properly (p_{e1} can be adjusted by adjusting the incoming traffic rate), situations can occur where different iterations within the same simulation cause traffic matrices with both positive and zero total external traffic.

When the external traffic is zero, the end-routers in System B have no load in the case of no failure. As a result, the only load incurred on those routers is during the time when a failure has occured. When p_f is low, it is clear that B_{load} will be correspondingly low. As a result, two separate curves are present,
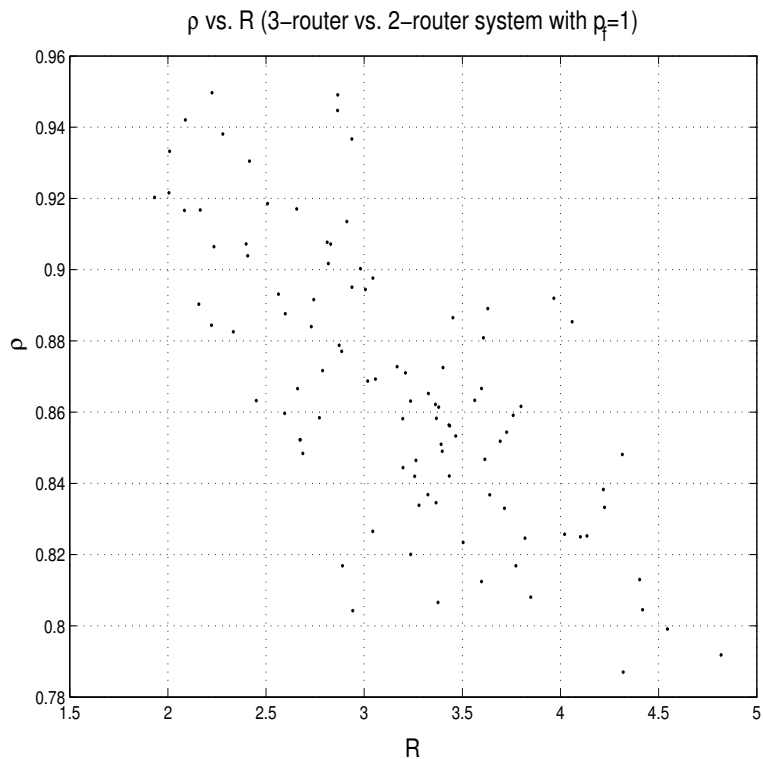
Figure 4-4: Graph showing the comparable performance between Systems A and B in the failure case.

the lower of which represents those traffic matrices with positive external traffic, and the higher of which represents those traffic matrices with zero total external traffic.

- Conditioning on the *failure-state* ($p_f = 1$, i.e. *worst-case scenario*), we see that System A and System B have comparable performance. Figure 4-4 shows this result.

This is because of the fact that there are two contending factors. Firstly, the cross-over traffic causes System A's performance to degrade with respect to that of System B. However, because 3 routers are shouldering the load in System A, as opposed to 2 routers in System B, we can expect the performances to be comparable. Figure 4-5 shows the histogram of the proportion of iterations that result in an average value of load within an interval (for both Systems A and B). It is clear that the histograms for both systems are similar and represent

data of the same order of magnitude.

When the amount of traffic crossing over is regulated (i.e. the cross-over traffic is still random, but is now modelled as a bounded integer random variable with a finite maximum) and bounded from above to a low value, the performance of System A actually is better than that of System B. It is clear from Figure 4-5 that System A can possibly outperform System B. Since a large portion of System A's load is due to the cross-over traffic during non-failure states, we see that if we limit this to a very low value, the performance of System A dramatically improves, and eventually surpasses the performance of System B.

The second simulation produced two comparisons:

$$N_R > 3 \text{ vs. } 2 < N_{R2} < N_R$$

and

$$N_R > 2 \text{ vs. } N_{R2} = 2.$$

The main results of this simulation were:

- In general, when $p_f << 1$, the higher the value of N_R, the poorer the performance. As N_R is decreased to 2, the performance improves. Figures 4-6 and 4-7 shows the ρ vs. R plots, and histograms.

- A ρ vs. R plot is bounded from below by $\frac{N_{R2}}{N_{R1}}$. In other words, when comparing the performance of the $N_R = 5$ system to the $N_R = 3$ system, the curve will be lower-bounded by $\frac{3}{5}$, and will approach this limit as $R \to \infty$. The reason for this is as follows. Assuming no external traffic, as $R \to \infty$, most of the traffic becomes intra-region traffic, meaning the routers are not used in the case of no failure. Hence, when a failure *does* occur, and each unit passes through exactly 2 routers (because of the recovery scheme), the average load in the system is simply equal to the ratio of the total traffic to be sent through routers to the number of routers, N_R. We see that the ratio of average load values from system to system approaches $\frac{N_{R2}}{N_{R1}}$.
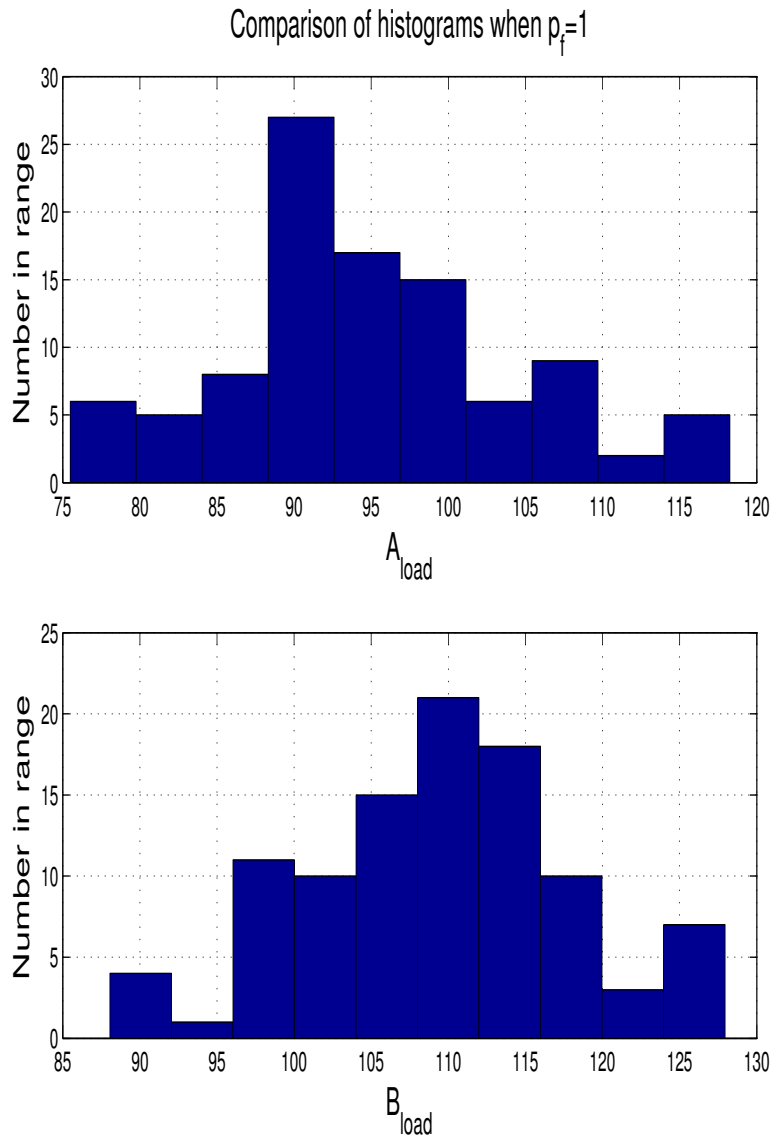
Figure 4-5: Comparison of histograms of Systems A and B in the failure case. This Figure shows that System A can actually outperform System B.

Figure 4-6: Comparison of load-ratios for different cases of N_R. System A represents a 5-router linear system, System A_2 represents a 3-router linear system, and System B represents a basic 2-router linear system.
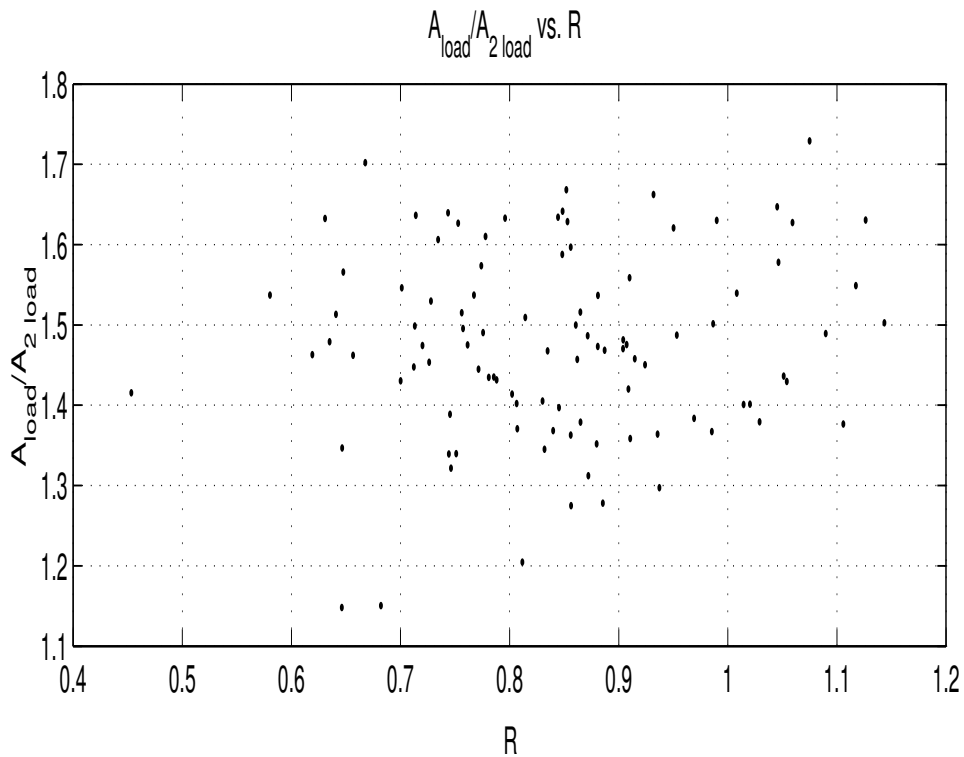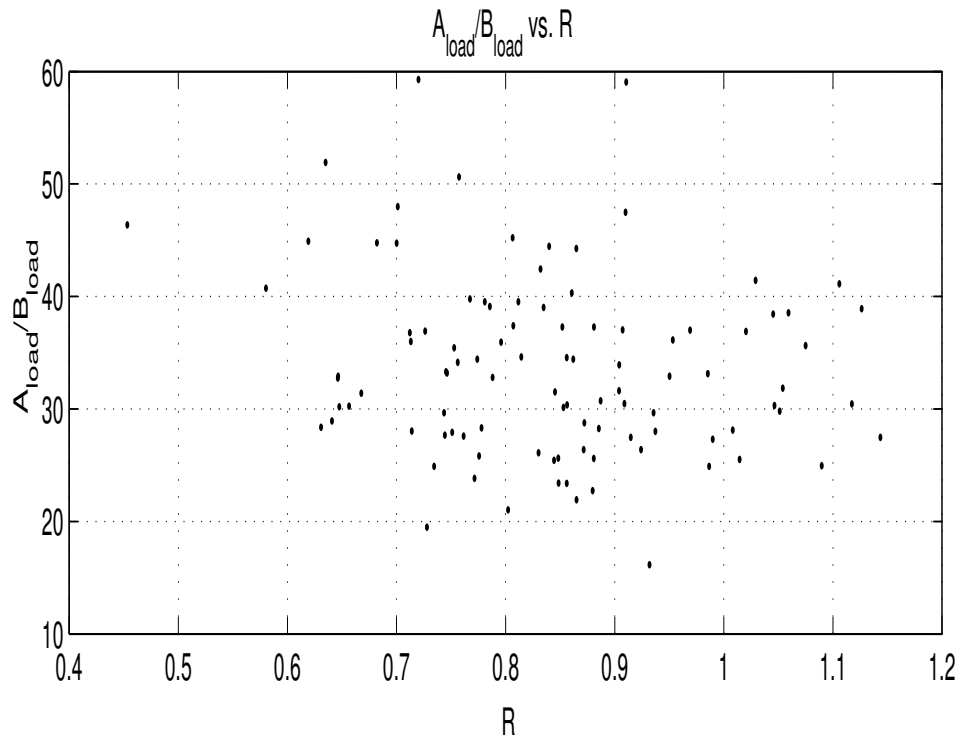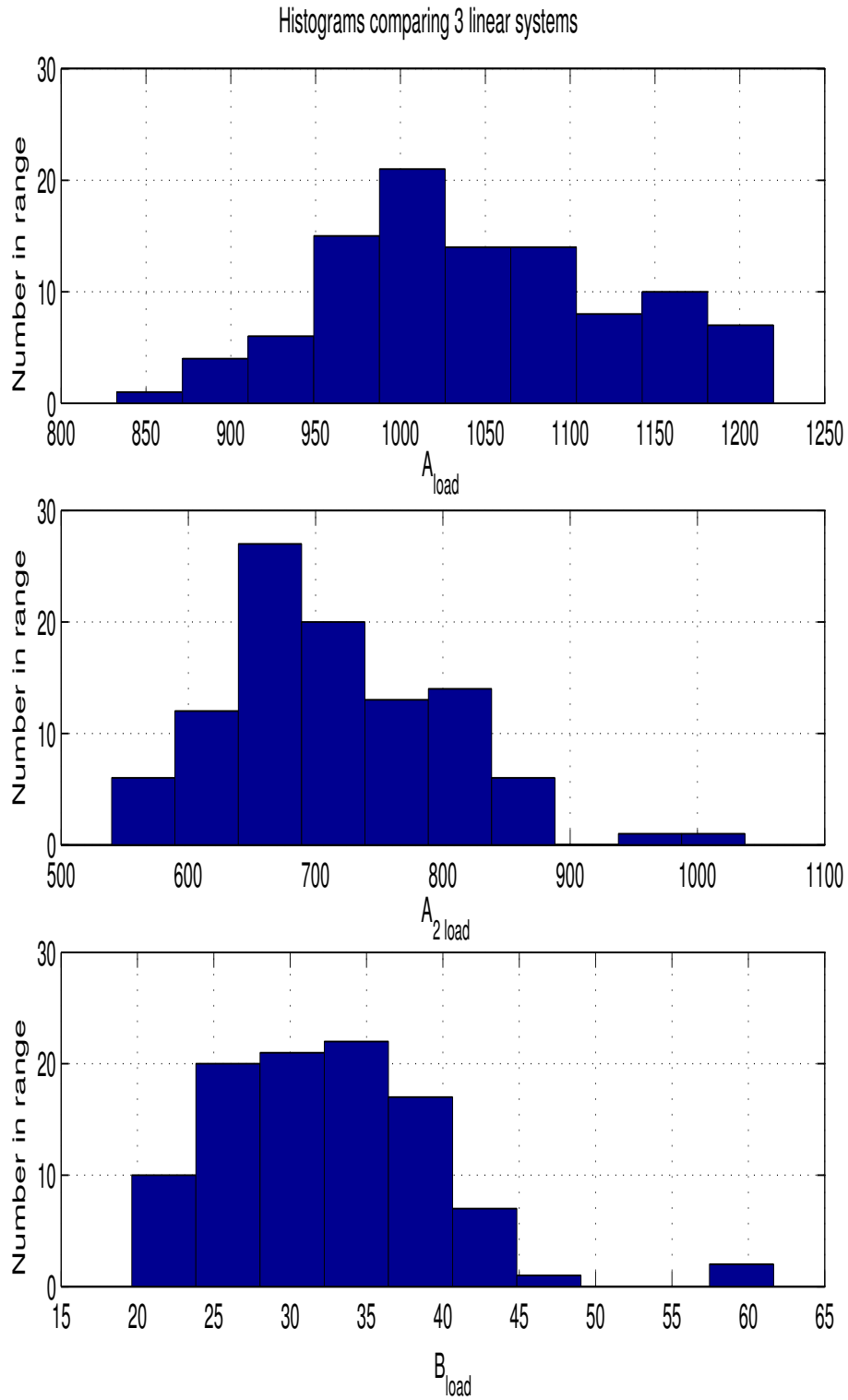
Figure 4-7: Comparison of histograms for different cases of N_R. System A represents a 5-router linear system, System A_2 represents a 3-router linear system, and System B represents a basic 2-router linear system.

4.3   Simulation Results of T-shaped Systems

As before, random traffic matrices were generated, and the systems were simulated over a large number of iterations. In this case, since there are N nodes in each region, and 3 regions in each system, the T-matrices are $(3N + 1)\mathrm{x}(3N + 1)$ matrices, with diagonal entries being zero. Systems A, B, C, and D were all compared.

The ratios of the average load in each system were plotted against new values of R. Let C represent the total cross-over traffic, and L represent all the traffic to be sent from the lower leg (region 3) to the right leg (region 2) in the network. Specifically, when comparing Systems A and B, R_2 was selected to be

$$R_2 = \tfrac{C}{L}.$$

R_2 now replaces R as the metric because the difference in the traffic routed through the central router in Systems A and B is highly dependent on the ratio of the cross-over traffic to the traffic meant to go from region 3 to region 2. This appears to be a good choice because the dominant factors in determining the average load of the system when $p_e << 1$ are the traffic going from region 3 to region 2 (in System B), and all the cross-over traffic (in System A).

When comparing Systems B and C, we notice that the only difference in load among the routers when $p_e << 1$ is the amount of traffic to be transferred from region 3 to region 2. Hence, we choose the following value for R_3:

$$R_3 = \tfrac{C}{L} - 1.$$

Again, the choice of this value can be attributed to the structures of the two systems. Since System D is a slight variation of Systems B and C, we see that R_3 can be used to measure the relative performance of System D and either System B or System C, and therefore replaces R as the best metric. Once again, all the plots of load-ratios versus values of R are created, and the histograms of the amount of average load per router of each system were created.

In comparing Systems A and B, we see that System A always performs worse than System B. The reason for this is simple. A certain fraction f of all the cross-over
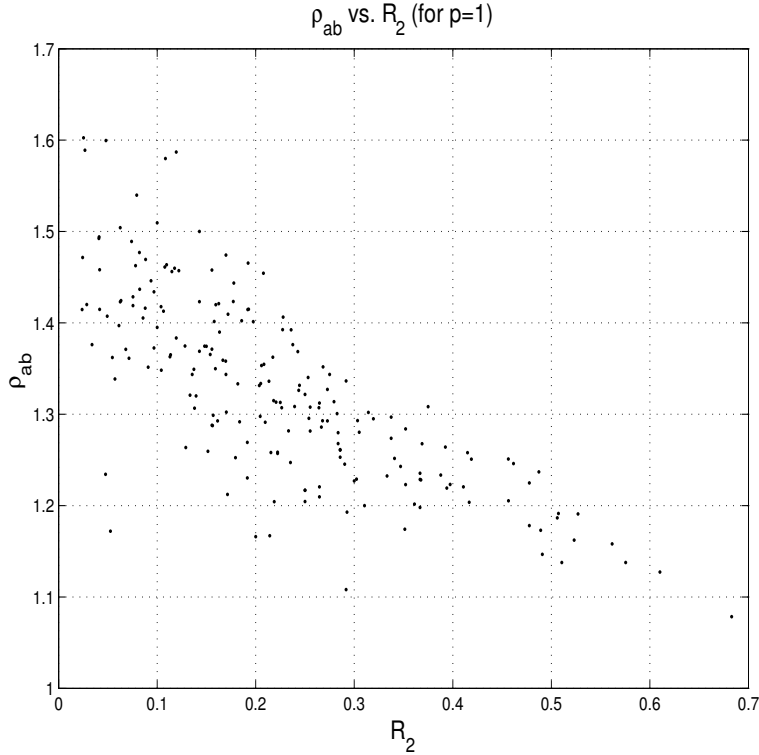
Figure 4-8: Graph comparing the performance of Systems A and B (different T-shaped systems).

traffic is not routed in System B, whereas all of the cross-over traffic is routed in System A. Moreover, this fraction f is usually greater than $\frac{1}{2}$. Recovery of all other traffic is performed the same way. It is interesting to note that, as R_3 increases, the two systems become more and more comparable in performance. This is because as R_3 increases, the ratio $\frac{L}{C}$ decreases, and more of the cross-over traffic is not in the region included in L. Hence, the ratio of the average load per router in System A to that of System B approaches 1 (the system performances become more and more similar). As R_2 decreases, System A's performance suffers relatively to that of System B. Figure 4-8 shows these results.

In comparing Systems B and C, we see that System C always performs better in terms of average load per router. This can be explained by the fact that System C utilizes more bandwidth in order to save router load. The traffic meant to go from region 3 to region 2 is passed through the folded bus (in C), thus resulting in savings on load. The obvious trade-off is that bandwidth on two links is needed for System C.
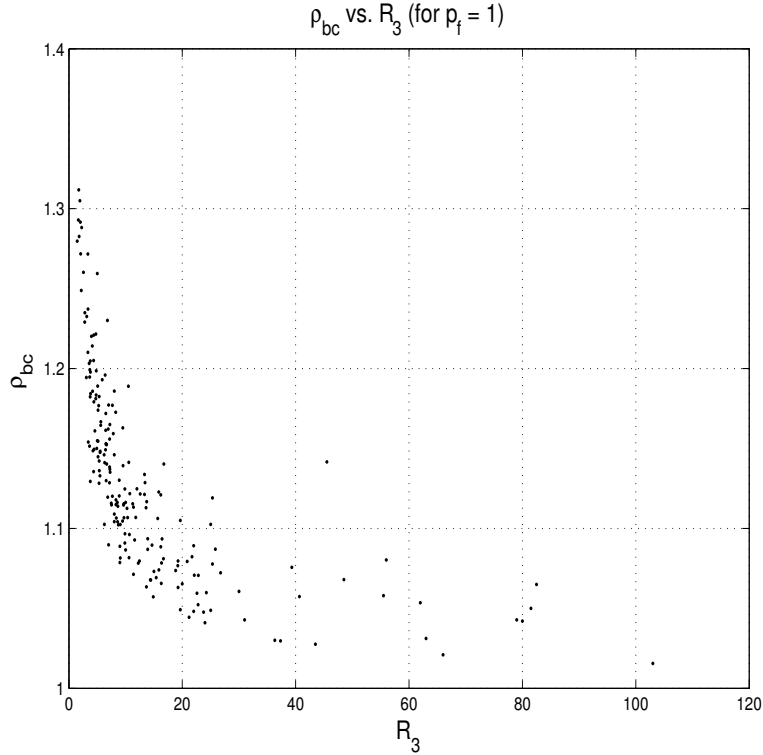
Figure 4-9: Graph comparing the performance of Systems B and C (different T-shaped systems).

As expected, as R_3 increases, the systems become more and more comparable. This is because most of the cross-over traffic is not included in L, and so neither system carries much internal traffic in the case of no failure. Figure 4-9 shows these results.

Finally, in comparing Systems B and D, a share-factor (α) needs to be selected. This value represents the proportion of the traffic included in L that is to be routed through the central router (in System D), and, together with R_3 provides for a good metric in comparing Systems B and D. For example, if $\alpha = \frac{3}{4}$, then 75% of the traffic to be sent from the lower leg (region 3) to the right leg (region 2) is processed by the central router. The remaining 25% of the traffic will be will be sent down the folded bus. Hence, the traffic is being "shared" by the bus and the central router.

Figure 4-10 shows the results for $\alpha = \frac{1}{2}$. As expected, System B's performance is worse than System D's performance (this is because System D is a hybrid of System B and System C), and the ratio of the average load is bounded from above by $\frac{1}{\alpha}$. As $R_3 \to 0$, almost all the traffic is included in L, and so the share-factor comes through

46
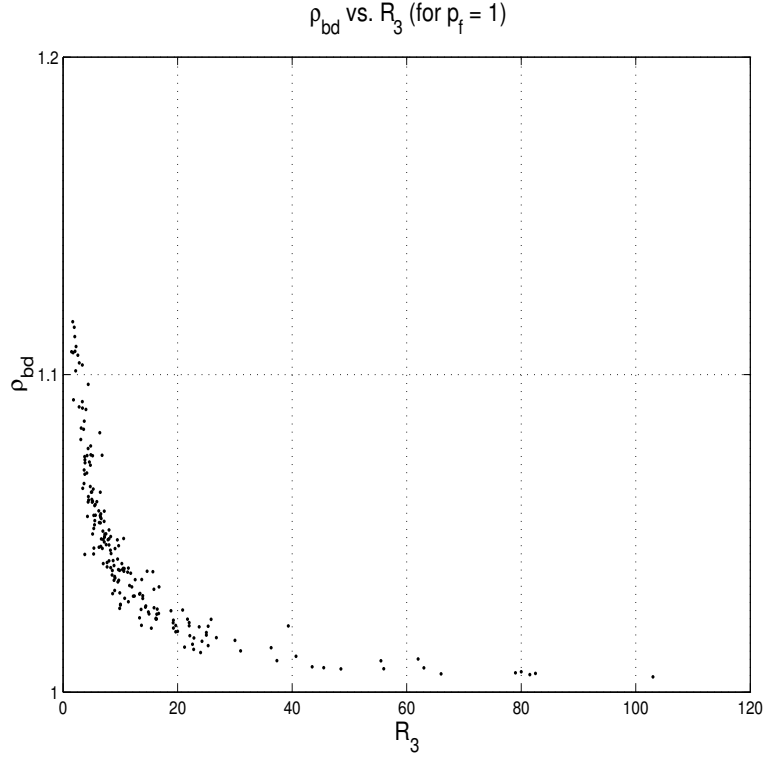
Figure 4-10: Graph comparing the performance of Systems B and D (different T-shaped systems).

in the average load. On the other hand, as R_3 gets large, the system performances become closer to each other, as expected.

In conclusion, we must note that the simulation results showed that the following inequalities always hold, independently of the traffic matrix, as predicted.

$$A_{load} \geq B_{load} \geq D_{load} \geq C_{load}.$$

It should be noted that Systems C and D use extra bandwidth, the amount of which depends on the traffic matrix. The ratio of the extra bandwidth needed in System C to that needed in System D is

$$r = \frac{1}{1-\alpha}, \text{ for } 0 \leq \alpha < 1.$$

47

The main results of this section are:

In the linear systems:

- Due to the fact that the value of R for a given matrix does not capture all the information in that matrix, the value of ρ is not a deterministic function of R. Hence, ρ vs. R plots generally have more than one point per value of R.

- In general, as R increases, ρ decreases. This is because of the fact that, as R increases, more and more of the total traffic is intra-region traffic.

- If the number of nodes and failure probability is chosen properly, the ρ vs. R graph decomposes into two graphs.

- In the failure state, the two systems (Systems A and B) have comparable performance.

- In general, when $p_f << 1$, the higher the value of N_R in a given system, the poorer the performance of the system (with respect to average load per router).

- A general ρ vs. R plot is bounded from below by $\frac{N_{R2}}{N_{R1}}$.

In the T-shaped systems:

- ρ vs. R plots behave the same way, in general, as for the linear systems.

- The following inequalities always hold: $A_{load} \geq B_{load} \geq D_{load} \geq C_{load}$.

Chapter 5

Conclusion

In this thesis, it was shown that a simple ratio R, extracted from the elements of the traffic matrix, characterizes the performance of the routers within the system with respect to certain metrics (i.e. average load per router, bandwidth needed,...etc.).

In particular, two systems were examined: the linear system, and the T-shaped system.

Analysis and simulations done on the linear system showed that under time average scenarios (very low probabilities of link failure), the performance of a system with more routers improves with respect to the performance of a system with fewer routers as the traffic becomes more and more local (i.e. more and more of the traffic has source and destination nodes that lie on the same folded bus). In other words, as R increases, the performance of the system with many routers gradually improves with respect to the performance of the system with fewer routers. Another interesting result that was obtained is that in the failure-state, two systems with a different number of routers perform comparably. The reason for this is that, although there are more routers in one system, forcing more traffic to be processed per router, the load is distributed among more routers, hence reducing the average load per router. Finally, when comparing the performance of two systems (System A and System B) with different numbers of routers (N_{R1} and N_{R2}), it was determined that the curve of the ratio of the average load in System A to that of System B is bounded from below by $\frac{N_{R2}}{N_{R1}}$.

The analysis and simulations performed on the four different T-shaped systems showed that the results from linear systems could be extended to T-shaped systems, when considering the ratio R. In addition to this, the average load per router in all four systems satisfied inequalities that indicated that a network designer can reduce router usage by increasing appropriately placed bandwidth (wavelengths or fibers in the system). Alternatively, one can decrease the amount of bandwidth needed in the system by increasing the usage of certain routers.

A direction for further research is the performance of these different systems using different traffic models. The elements of the traffic matrix are independent, identically-distributed (I.I.D.) exponential random variables with different parameters. Interesting paths to follow would be testing the system under uniform, and Pareto (or other heavy-tailed) traffic. Using uniform traffic would be similar to modelling a system in which all node-pairs have the same amount of demand to be exchanged. Analysis and simulations using heavy-tailed traffic would help us understand how the system performs under more realistic traffic in which the probability of a very large deviation is small, but not negligible.

Another avenue to follow would be to consider non-linear bus architectures beyond the T-shaped systems. Examining combinations of linear and T-shaped systems may yield results about general mesh networks. It would be interesting to see if the results pertaining to the simpler linear and T-shaped systems can be extended to account for cascading.

A more interesting result would be measuring the dropoff in performance of a particular system with respect to changing the distribution of the input traffic. For instance, how would the average load of a system change under exponential traffic with a smaller parameter? Given, the performance of a system under a certain traffic model, how would the performance degrade with respect to another distribution? This question is particularly important when considering the fact that virtual topologies lie on top of pre-existing physical architectures. Hence, not all virtual topologies are possible (i.e. they are contrained to a subset of topologies due to the nature of the physical topology), and as a result, the "best case" topology needs to be

used. This "best case" topology is generally not optimized to the proper traffic distribution, but rather, to another distribution with nonzero distance from the proper distribution. The "distance" between the distributions of traffic (i.e. determined using the Kullback-Liebler distance or some other metric) can be measured, and in a sense, the relationship between the amount a system degrades from one traffic distribution to another, and the distance between the two distributions can be developed.

Bibliography

[1] C. Bisdikian. Waiting time analysis in a single buffer DQDB (802.6) network. In *Proceedings IEEE INFOCOM*, pages 610–616, 1990.

[2] K. Bogineni, K.M. Sivalingam, and P.W. Dowd. Low-complexity multiple access protocols for wavelength-division multiplexed photonic networks. In *Jsac*, volume 11, pages 509–604, 1993.

[3] D. Callahan and G. Grimes. An intelligent hub protocol for local area lightwave networks. In *Conference on Local Computer Networks*, pages 260–261, 1999.

[4] M.-S. Chen, N.R. Dono, and R. Ramaswami. A media-access protocol for packet-switched wavelength division multiaccess metropolitan area networks. In *IEEE Journal on Selected Areas in Communications*, volume 8, pages 1048 –1057, 1990.

[5] I. Chlamtac and A. Ganz. Design alternatices of asynchronous WDM star networks. In *IEEE International Conference on Communications*, pages 23.4.1–23.4.5, 1989.

[6] I. Cidon and Y. Ofek. Metaring - a full duplex ring with fairness and spatial reuse. In *Proceedings IEEE INFOCOM*, pages 969–981, 1990.

[7] M. Conti, E. Gregori, and L. Lenzini. DQDB under heavy load: performance evaluation and fairness analysis. In *Proceedings IEEE INFOCOM*, pages 133–145, 1990.

[8] M. Conti, E. Gregori, and L. Lenzini. A methodological approach to an extensive analysis of DQDB performance and fairness. In *IEEE Journal on Selected Areas in Communications*, volume 9, pages 76–87, January 1991.

[9] P.W. Dowd. Random access protocols for high-speed interprocessor communicationbased on an optical passive star topology. In *Journal of Lightwave Technology*, volume 9, pages 799 –808, June 1991.

[10] A. Ganz and Y. Gao. Time-wavelength assignment algorithms for high performance WDM star based systems. In *IEEE Transactions on Communications*, pages 1827 –1836, May 1994.

[11] A. Ganz and Z. Koren. WDM passive star - protocols and performance analysis. In *Global Telecommunications Conference*, pages 9A.2.1–9A.2.10, 1991.

[12] I.M.I. Habbab, M. Kavehrad, and C.E.W. Sundberg. Protocols for very high speed optical fiber local area networks using a passive star topology. volume LT-5, pages 1782–1794, December 1987.

[13] E.Y. Huang and L.F. Merakos. On the access fairness of the DQDB MAN protocol. In *Proceedings of IPCC*, pages 325–329, 1990.

[14] P.A. Humblet, R. Ramaswami, and K.N. Sivarajan. An efficient communication protocol for high-speed packet-switched multichannel networks. In *IEEE Journal on Selected Areas in Communications*, volume 11, pages 568 –578, May 1993.

[15] H.B. Jeon and C.K. Un. Contention based reservation protocols in multiwavelength protocols in multiwavelength protocols with passive star topology. In *Proceedings of the IEEE International Conference on Communications*, June 1992.

[16] A.E. Kamal. Efficient multi-segment message transmission with slot reuse on DQDB. In *Proceedings IEEE INFOCOM*, pages 869–878, 1991.

[17] J.H. Laarhuis and A.M.J. Koonen. An efficient medium access control strategy for high-speed WDMmultiaccess networks. In *Journal of Lightwave Technology*, page 1078, May 1993.

[18] D.A. Levine and I.F. Akyildiz. PROTON:a media access control protocol for optical networks with star topology. In *Proceedings of the 20th Annual Computer Science Conference*, volume 3, pages 158 –168, April 1995.

[19] J. Limb and C. Flores. Description of Fasnet - a unidirectional local area communication network. In *Bell System Technical Journal*, volume 61, pages 1413–1440, September 1982.

[20] M. Maode, B. Hamidzadeh, and M. Hamdi. A receiver-oriented message scheduling algorithm for WDM lightwave networks. In *Global Telecommunications Conference*, volume 4, pages 2333 –2338, May 1998.

[21] N. Mehravari. Performance and protocol improvements for very high speed optical fiber local area networks using a passive star topology. volume 8, pages 520–530, April 1990.

[22] M. Mishra, E.L. Johnson, and K.L. Sivalingam. Scheduling in optical WDMnetworks using hidden markov chain-based traffic predictors. In *IEEE International Conference on Networks*, pages 380 –384, 2000.

[23] E. Modiano and R.A. Barry. Design and analysis of an asynchronous WDMlocal area network using a master/slave scheduler. In *INFOCOM '99*, volume 2, pages 900 –907, May 1999.

[24] K. Nosu and H. Toba. An optical multiaccess network with optical collision detection and optical frequency addressing. In *IEEE International Conference on Communications*, volume 3, pages 968 –975, March 1990.

[25] M.A. Rodrigues. Erasure nodes: performance improvements for the IEEE 802.6 MAN. In *Proceedings IEEE INFOCOM*, pages 636–643, 1990.

[26] T. Strosslin and M. Gagnaire. A flexible MACprotocol for all-optical WDM metropolitan area networks. In *Proceeding of the IEEE International Performance, Computing, and Communications Conference*, pages 567 –573, 2000.

[27] G.N.M. Sudhakar, N.D. Georganas, and M. Kavehrad. A multichannel optical star LAN and its application as a broadband switch. volume 5, December 1987.

[28] F.A. Tobagi, F. Borgonovo, and L. Fratta. Expressnet: a high performance integrated-services local area network. In *IEEE Journal on Selected Areas in Communications*, volume SAC-1, pages 898–913, November 1983.

[29] L. Wang and M. Hamdi. Efficient protocols for multimedia streams on WDMAnetworks. In *Twelfth International Conference on Information Networking*, volume 1, pages 241 –246, 1998.

[30] G. Watson and S. Ooi. What *should* a Gbits/s network interface lool like. In North-Holland, editor, *Protocols for High-Speed Networks*, volume II, pages 237–250, Amsterdam, 1990. M.J. Johnson, editor.

[31] G.C. Watson and S. Tohme. S++-anew MAC protocol for Gb/s local area networks. In *IEEE Journal on Selected Areas in Communications*, volume 11, pages 531–539, may 1993.

[32] J.W. Wong. Throughput of DQDB networks under heavy load. In *Proceedings of EFOC/LAN Conference*, pages 146–151, 1989.

[33] H.-T. Wu, Y. Ofek, and K. Sohraby. Integration of synchronous and asynchronous traffic on the MetaRing architecture and its analysis. In *Proceedings of the IEEE International Conference on Communications*, 1992.

[34] A. Yan, A. Ganz, and C.M. Krishna. A distributed adaptive protocol providing real-time services on WDM-based LANs. In *Global Telecommunications Conference*, volume 14, pages 1245 –1254, June 1996.