The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.
Investigating Cost Effective Communication Alternatives for Geographically Hostile Regions

A dissertation submitted to the Department of Computer Science in partial fulfillment of the degree of Master of Science at the University of Cape Town

September 2001

Written by
Yakomba Yavwa

Supervised by
Professor P. S. Kritzinger

Data Network Architectures Laboratory
Department of Computer Science
University of Cape Town
Abstract

The lack of communication facilities in developing countries is a constraint to social, political and economic empowerment of the people. However, advances in technology promise to deliver voice, video and data communication services, that are urgently needed, under a common packet switched communication network.

Using Zambia as a case study, the cost effectiveness of three design scenarios which involve a hybrid of radio and microwave, radio and satellite, and radio and optic fibre were evaluated. The communication links in each scenario were modeled as E1 trunks. The cost of establishing these links was determined. Each scenario was then subjected to generic input traffic patterns using a multiclass queueing network analysis model to determine its effectiveness.

Our findings show that the scenario involving the microwave links as inter-regional links has the lowest cost. The second is the scenario involving satellite inter-regional links and the most expensive is the scenario involving optic fibre inter-regional links. Furthermore, the microwave-radio hybrid scenario was found to have the smallest cost-effectiveness ratio of 17.9 followed by the optic fibre-radio scenario with 20.3 and finally the satellite-radio scenario with 892.9.

These results effectively imply that the microwave-radio hybrid scenario is a better option for developing countries.
Acknowledgements

Professor PS Kritzinger, for being my Supervisor, his encouragement, patience, support, ever valued ideas, objective criticism and above all, his time in shaping this work is most treasured and appreciated.

Dr. Hutchison A., for his ideas, proof reading the thesis and his help in implementing the analytical model using Micro SNAPL.

The Department of Computer Science, for having made it possible for me to attend a conference on Telecommunication systems in Johannesburg.

My wife Felistah, my daughter Elizabeth and my son Jacob, for their understanding during difficulty times when I spent most of my time working on this project.

Mr. Chisenga (my Director), for having afforded me this rare opportunity to undertake these studies and his support during the course of the academic programme.

Mr. Oliver Chinganya (Deputy Director CSO), for his assistance in providing us with the statistics and information about Zambia and the regions in the case study.

Mr Cilliie Malan and various Suppliers, for their help in giving us costs of various components. For fear of competition, some suppliers opted not to be cited but all the same, we extend our sincere thanks.

My parents, Mr Muhombo Yavwa and Mrs Esnati Yavwa, for their sacrifice and dedication in shaping my life when I could not stand on my own.

Dr. Mulenga and Mr. Musongole, for having created a conducive social environment for me, their brotherly support and encouragement, I treasure most.

The DNA crew, Marc, Elton, Lourens, Micheal, Andy and Farrel for their ideas and help.
## Contents

1 Introduction 1

1.1 Motivation and Objectives 1 1

1.2 Type of Applications 3

1.3 Traffic model 4

1.3.1 Video traffic 5

1.3.2 Voice traffic 6

1.3.3 Data traffic 6

1.4 Topological Model 6

1.4.1 Ring Architecture 7

1.4.2 Mesh Architecture 7

1.4.3 Star Architecture 9

1.4.4 System overview 10

1.5 Related Work 11

1.6 Dissertation Roadmap 12

2 Evaluation of Candidate Technologies 14

2.1 Microwave Systems 14

2.1.1 Application 14

2.1.2 Cost 17

2.1.3 Advantages 17

2.1.4 Disadvantages 18
2.2 Optic Fibre Systems ........................................... 18
  2.2.1 Application .............................................. 18
  2.2.2 Cost ..................................................... 19
  2.2.3 Advantages ............................................ 20
  2.2.4 Disadvantages ......................................... 20
2.3 Satellite Systems .............................................. 21
  2.3.1 GEO Satellite systems .................................. 21
    2.3.1.1 Application ....................................... 21
    2.3.1.2 Cost ............................................... 23
    2.3.1.3 Advantages ....................................... 23
    2.3.1.4 Disadvantages .................................... 23
  2.3.2 MEO Satellite systems .................................. 24
    2.3.2.1 Application ....................................... 24
  2.3.3 LEO Satellite systems .................................. 24
    2.3.3.1 Application ....................................... 24
    2.3.3.2 Advantages ....................................... 26
    2.3.3.3 Disadvantages .................................... 26
  2.3.4 Selection of a suitable satellite system ............. 27
2.4 DSL Systems .................................................. 28
  2.4.1 Application .............................................. 28
  2.4.2 Cost ..................................................... 28
  2.4.3 Advantages ............................................ 28
  2.4.4 Disadvantages ......................................... 29
2.5 Radio Systems ................................................ 29
  2.5.1 Application .............................................. 29
  2.5.2 Cost ..................................................... 30
  2.5.3 Advantages ............................................ 30
  2.5.4 Disadvantages ......................................... 30
3 Analytic Modelling

3.1 Modelling techniques ........................................ 32
3.2 The Queueing Model ........................................... 33
3.3 Functional description of a customer .......................... 33
3.4 Arrival Process .................................................. 34
3.5 Class Switching .................................................. 34
3.6 Multiclass Networks ............................................ 35
3.7 Types of Servers ................................................ 36
3.8 The Service Discipline ......................................... 37
   3.8.1 PS .......................................................... 37
   3.8.2 FCFS ....................................................... 38
   3.8.3 Infinite Server (IS) ....................................... 38
3.9 Performance Measures ......................................... 38
   3.9.1 Throughput .............................................. 39
   3.9.2 The Mean Delay .......................................... 39
   3.9.3 Utilization ................................................ 41
3.10 Conclusion ...................................................... 41

4 Case Study ........................................................ 42

4.1 Existing Infrastructure ......................................... 43
4.2 The System Model .............................................. 43
   4.2.1 Regional Connectivity Pattern .......................... 43
   4.2.2 Possible Scenarios ....................................... 45
      4.2.2.1 Scenario 1 ........................................... 46
      4.2.2.2 Scenario 2 ........................................... 47
      4.2.2.3 Scenario 3 ........................................... 47
   4.2.3 The Workload Model ..................................... 48
CONTENTS

4.2.3.1 Customer behaviour on the inter-regional link .................. 50
4.2.3.2 The cost and performance models ................................. 51
4.2.3.3 Routing of customers .............................................. 52
4.2.4 The Machine Model .................................................. 53

5 Results 56

5.1 Scenario 1 .............................................................. 57
  5.1.1 Performance ....................................................... 57
  5.1.2 Cost .................................................................. 60
5.2 Scenario 2 .............................................................. 61
  5.2.1 Performance ....................................................... 61
  5.2.2 Cost .................................................................. 64
5.3 Scenario 3 .............................................................. 64
  5.3.1 Performance ....................................................... 65
  5.3.2 Cost .................................................................. 67
5.4 Summary of results .................................................... 68
5.5 Conclusion ............................................................... 68

6 Conclusion 70

6.1 Future Work ............................................................ 72
List of Figures

1.1 The Ring topology ................................................................. 7
1.2 The Mesh topology ............................................................... 8
1.3 The Star Architecture ............................................................ 9
1.4 The System Overview ........................................................... 10

2.1 Point-to-Point Microwave hop ............................................... 15
2.2 Intra-Region connectivity using LMDS technology .................. 16
2.3 A snapshot of a VSAT configuration ....................................... 22
2.4 The connectivity pattern for LEO satellites with onboard processing with ISL .... 25
2.5 Cell Configuration ................................................................. 26
2.6 An illustration of an interlocking web topology radio system .......... 30

3.1 The queueing model .............................................................. 33
3.2 Change of customer classes .................................................. 35
3.3 Types of classes ................................................................. 36
3.4 A system with multiple servers ............................................. 37
3.5 Average network delay as a function of utilization, $\rho$ ............... 40

4.1 Regions covered in the case study ......................................... 44
4.2 Microwave inter-regional links ............................................. 46
4.3 Satellite inter-regional links .................................................. 47
4.4 Optic Fibre inter-regional links ............................................ 47
4.5 Uni-directional traffic flow between the regions ....................... 49
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.6</td>
<td>The traffic flow between two adjacent regions</td>
<td>50</td>
</tr>
<tr>
<td>4.7</td>
<td>The M-QNA model</td>
<td>55</td>
</tr>
<tr>
<td>5.1</td>
<td>Mean Delay Vs Utilization of scenario 1</td>
<td>58</td>
</tr>
<tr>
<td>5.2</td>
<td>Mean Delay Vs Permissible throughput of scenario 1</td>
<td>58</td>
</tr>
<tr>
<td>5.3</td>
<td>The effect of an error on the Lundazi - Mbala link</td>
<td>60</td>
</tr>
<tr>
<td>5.4</td>
<td>Mean Delay Vs Utilization of scenario 2</td>
<td>62</td>
</tr>
<tr>
<td>5.5</td>
<td>Mean Delay Vs Permissible throughput of scenario 2</td>
<td>63</td>
</tr>
<tr>
<td>5.6</td>
<td>The effect of an error on the inter-regional link in Figure 4.3</td>
<td>63</td>
</tr>
<tr>
<td>5.7</td>
<td>Mean Delay Vs Utilization of scenario 3</td>
<td>65</td>
</tr>
<tr>
<td>5.8</td>
<td>Mean Delay Vs Permissible Throughput of scenario 3</td>
<td>66</td>
</tr>
<tr>
<td>5.9</td>
<td>The effect of an error on the inter-regional link in Figure 4.4</td>
<td>66</td>
</tr>
</tbody>
</table>
### List of Tables

1.1 Direct Services to End - Users. Source: 1999 IEEE Emerging technologies symposium [31] .................................................. 2

2.1 Unit costs for the microwave system ........................................... 17

2.2 Typical Fibre Characteristics Source: Sterling D., Technicians Guide to Fibre Optics, 1993 .................................................. 19

2.3 Inter-regional connectivity costs for configuration 3 - Source: University of Cape Town, ITS Department ......................................... 20

2.4 Inter-regional connectivity costs for configuration 2 using a leasing option .................................................. 23

2.5 Prominent Broadband Systems. Source: IEEE's GLOBECOM'99-MULTIMEDIA SATELLITE COMMUNICATIONS, pg 1167 .................................................. 27

4.1 Population distribution for the case study regions. Source: Census 2000 report, CSO Zambia .................................................. 43

4.2 Summary of System Parameters .................................................. 45

4.3 Hybrid Scenarios .................................................. 46

4.4 A loop to modify the traffic input patterns for each workload .................................................. 48

4.5 The Routing Probability Matrix .................................................. 53

5.1 Inter-regional connectivity costs for configuration 1 - Source: See Chapter 2 .................................................. 61

5.2 Inter-regional connectivity costs for scenario 2 using a leasing option: Source: See Chapter 2 .................................................. 64

5.3 Inter-regional connectivity costs for scenario 3 - Source: see Chapter 2 .................................................. 67

5.4 Summary of Results .................................................. 68
Notation

In order to ensure that the reader understands the abbreviations and notations used throughout this dissertation, we hereby present a list of such abbreviations and notations. The reader is cautioned that notation and terminology are not uniform among authors and therefore it is not uncommon for two authors to use the same symbol to mean two different things. For the purpose of this dissertation, the notation defined here is applicable.

*Customer* - represents a packet in a communication network

*N* - represents the maximum number of nodes in the network

*n* - represents a node in the network such that 1 ≤ *n* ≤ *N*

*l*<sub>*ij*</sub> - represents a communication link between regions *i* and *j*.

*K*<sub>*ij*</sub> - represents the capacity of the link *l*<sub>*ij*</sub>

*R*<sub>*i*</sub> - represents the number of the region, 1 ≤ *i* ≤ *n*

*λ*<sub>*c*</sub> - denotes the mean arrival rate of class *c* customers

*λ*<sub>*ic*</sub> - represents mean arrival rate of class *c* customers at node *i*.

*τ*<sub>*ij*</sub> - represents the throughput or average flow between regions *i* and *j*

*Site n*<sub>*i*</sub> - represents Site *n* in region *i*

*l*<sub>*ii*:in</sub> - represents a communication link between site 1 and Site *n* in region *i*

*p*<sub>*ic*:jc</sub> - denotes the probability of routing class *c* customers from region *i* to region *j*

*p*<sub>*ic*:jr</sub> - the probability that a customer of class *c* at region *i* is transferred to class *r* at region *j*

*T* - represents the total average network delay of a customer in the system

*T*<sub>*li*</sub> - represents the mean delay of a customer on a link *l*<sub>*ij*</sub>

*D* - represents the total cost of the system

*d* - represents the unit cost of a communication network resource

*d*<sub>*ij*</sub> - represents the cost of a communication link *l*<sub>*ij*</sub> including the cost of the terminal equipment

*ρ*<sub>*i*</sub> - represents the utilisation of the system at node *i*

*ρ*<sub>*ij*</sub> - represents the utilisation of the link *l*<sub>*ij*</sub>
\( \mu_i \) - represents the mean service rate at node \( i \)

\( \frac{1}{\mu_i} \) - represents the mean service time at node \( i \).

\( C \) - represents the maximum number of classes of customers in the network.

\( c \) - represents a class of customers such that \( 1 \leq c \leq C \)

\( e \) - represents the bit error rate of a communication link.

\textit{FCFS} - First Come First Serve

\textit{PS} - Processor Sharing

\textit{IF} - Infinite Server
Chapter 1

Introduction

1.1 Motivation and Objectives

The design of cost-effective communication systems for geographically hostile regions is an important feature to enable inexpensive end-user services. These geographically hostile regions are defined as lacking telecommunications infrastructure, being sparsely populated and have less economic activities. They are generally wide spread and are mainly located in Africa, South America, Eastern Europe and Asia [39, 32]. Guggenmons [12] has emphasized the need to provide inexpensive services to these regions by using cost-effective communication systems. The required services include remote data processing in financial institutions which is a fundamental factor in the economic and sociological development of any region. Due to lack of health facilities in most geographically hostile regions, the provision of tele-medical services would enhance medical care and keep these regions up to date with other developed regions. In the field of education, the emphasis is on interactive learning. The provision of interactive learning to these regions would facilitate the equal distribution of learning services. This would eventually close up the regional developmental gaps created by lack of adequate services for geographically hostile regions as a result of their geographical restrictions.

The required services in these regions form sets of applications which are classified as real-time and non-real-time applications. These applications are described in detail in Section 1.2. Designing separate networks to support these different sets of applications would defeat the purpose of enabling inexpensive end-user services. Thus, the provision of integrated services under a common network, described in Section 1.4, which enables inexpensive end-user services is assumed in this study.
1.1 Motivation and Objectives

The type of connectivity required to provide communication services to geographically hostile regions is identified as comprising both intra-region and inter-region communication links illustrated in Section 1.4.

The intra-region communication links involve communication between the sites located within a region while the inter-regional communication links involve communication between different regions.

The intra-region links are further defined as short distance links because they provide the last mile communication links to the end-users. Communication systems considered to establish such links include DSL, Radio, the point-to-point Optic Fibre and Microwave systems. The sites in each region represent the various categories of end-users mentioned in Table 1.1, ranging from large organisations that require LANs to residential users that require dial-up services.

<table>
<thead>
<tr>
<th>User</th>
<th>User characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large Business</td>
<td>. generate constant symmetric traffic</td>
</tr>
<tr>
<td></td>
<td>. require guaranteed data rates &amp; QoS</td>
</tr>
<tr>
<td>Medium Business</td>
<td>. generate relatively constant symmetric traffic</td>
</tr>
<tr>
<td></td>
<td>. require guaranteed min/max data rates &amp; QoS</td>
</tr>
<tr>
<td>Small Business</td>
<td>. generate asymmetric traffic</td>
</tr>
<tr>
<td></td>
<td>. require guaranteed min data rates &amp; QoS</td>
</tr>
<tr>
<td>Residential</td>
<td>. generate sporadic symmetric and highly asymmetric traffic</td>
</tr>
<tr>
<td>(Households)</td>
<td>. require access to max available data rates</td>
</tr>
</tbody>
</table>

Table 1.1: Direct Services to End - Users. Source: 1999 IEEE Emerging technologies symposium [31]

These end-users communicate through regional routers which are connected using inter-region links. Thus, the inter-region links are used to provide long distance links between the regions. The communication systems considered to establish such links include Optic Fibre, Microwave and Satellite systems.

The major cause of the lack of telecommunications infrastructure in these regions is that Service Providers are unable to financially justify the cost of providing the required communication services. This is because of the low economic activity which would translate into inadequate return on investment to sustain the provision of the services. This presents an opportunity for investigating cost-effective communication technology alternatives to service these regions.

Based on this motivation, the communication problem posed in this chapter is tackled in three
ways. Firstly, different system scenarios comprising hybrid architectures and distinct backbone links are designed. For each scenario, the performance is modelled. The reasons for modelling the performance of scenarios with distinct links is made clear in Chapter 4. Secondly, the cost for each scenario is computed and thirdly, the cost-effectiveness of each scenario is evaluated. The results of the various scenarios are then compared to determine the most cost-effective alternative. At this point, it is important to note that the cost of communication equipment will usually differ with different suppliers. As a result, our focus is on the principle of investigating cost-effective communication alternatives rather than costing the communication systems.

An important point to note is that the problem being solved requires the design of communication systems yet to be installed. This means that operational data on which to base the performance evaluations of the systems does not exist. This presents a problem of uncertainty in the system parameters and the usage patterns which are not known exactly. This uncertainty is dealt with by doing parametric analysis in which the model is solved many times for different parameters. The usage of the system is represented as a workload imposed on the system. Uncertainty about the level of use, therefore, is also dealt with by doing parametric analysis for a given workload as recommended by R. Haerkort [13].

1.2 Type of Applications

Future applications such as voice, video and multimedia applications are expected to have stringent real-time constraints and will demand not only high-bandwidths but a predictable quality-of-service (QoS). The large amounts of bandwidth promised by future networks, such as the development of Gigabits per second (Gbps) Optic Fibre systems, also offer the possibility of integrating real-time applications together with non-real-time applications within a single common network. Although the scaling of bandwidth to more than a Gbps in next generation networks will certainly have a profound effect on all aspects of networking, the need to support a more diverse mix of services by accommodating the performance requirements of both real-time applications and non-real-time applications raises important design issues that go beyond bandwidth scaling.

The communication network applications such as file transfer, electronic mail and remote login are examples of non-real-time applications for which the performance metrics of interest are typically mean delay and throughput. These applications also have strict reliability requirements which resulted in the complexity of traditional network protocols because of the need for loss-
free communication between remote sites. For the purpose of this study, these non-real-time applications are classified as belonging to a data class.

The characteristics of real-time communication applications, which are classified as packetized voice and video, differ significantly from those of non-real-time applications. The distinguishing feature of real-time applications is the fact that the value or utility of the communication depends upon the time at which messages are successfully delivered to the destination. Typically, the desired delivery time for each message across the network is bounded by a specific maximum delay (or latency), resulting in a deadline being associated with each message. If a message arrives at the destination after its deadline has expired, its value to the end application may be greatly reduced or may completely become insignificant to perform the intended task. In such a situation, a message is considered perishable which means that it is useless to the application if delayed beyond the deadline. This message is then discarded or considered lost.

These applications can be further characterised by their bit error rates (BER) and sensitivity to delay. For data applications, although low latency is usually desirable and may be beneficial in some way, they are delay insensitive but are extremely error sensitive with an acceptable BER of $10^{-9}$. Voice applications, on the other hand, require low latency because they are delay sensitive with an acceptable BER of $10^{-3}$. Like voice, video applications are delay sensitive with an acceptable BER of $10^{-5}$ and thus require low latency too.

The three classes of traffic namely voice, video and data, all of which are packetized, are assumed to be integrated in a single network which is packet switched. Furthermore, the packetization of digitized user information provides a highly flexible and bandwidth efficient method of transferring integrated services such as voice, video and data in a single network infrastructure. The nature of this communication network is explained in Section 1.4.

### 1.3 Traffic model

There are usually two ways to model the traffic over a network. Either by using a statistical model or by using network traces (recorded on a real network). Since the network being modeled is yet to be implemented, there are no traces available for this purpose and hence a statistical model has been used in this study.

In this model, the messages from an infinite source population arrive at the gateway following a Poisson arrival process (random interarrival time with negative exponential distribution). At
1.3.1 Video traffic

the gateway, these messages are broken down into packets. The main advantage of this traffic model is that it is already widely used [10, 20, 2, 36] and it allows to explore the full range of network loads and packet sizes.

From Section 1.2, we know that these network loads are assumed to comprise three distinct customer classes that arrive to the communication system at rates $\lambda_{\text{voice}}$, $\lambda_{\text{video}}$ and $\lambda_{\text{data}}$ according to the Poisson arrival process. The overall external arrival rate $\lambda$ at the gateway is the sum of these class dependent arrivals. Each of these class dependent arrivals place different demands on the communication network. Ramakrishnan and Mitra [26] approximate this demand to be $0.4\lambda$, $0.1\lambda$, $0.25\lambda$ and $0.25\lambda$ for voice, video, data and best effort classes of customers respectively. However, this study does not consider the best effort class. Hence, based on the weightings in these values and the nature of the service regions considered in this study (i.e. with low economic activities), we heuristically approximate this demand to be $0.5\lambda$, $0.13\lambda$ and $0.37\lambda$ for voice, video and data respectively to the nearest 5 percent.

Furthermore, we know that voice, video and data customers (or packets) are generated by their respective sources differently before they are funneled through the common communication network. Therefore, we first seek to understand the nature of video traffic and then consider the voice traffic in order for us to correctly characterize them. Finally, we consider the data traffic.

1.3.1 Video traffic

Video is generally fragmented. This is because a video generated frame is quite large and hence, is not ideal for transmission. Fragmentation means that a video frame is divided into multiple fragments. Each fragment forms a video packet or customer that is sent across the communication network. There is normally an upper limit on the maximum fragment size used by the fragmentation process. The choice of this maximum fragment size should be made such that it does not penalize the voice customers. The voice customers are penalized if large video customers are generated that induce huge delays in the system. According to the work done by ElGebaly [7], we note that the video customer lengths are exponentially distributed with a mean of 384 bytes. Since video customers are delay sensitive, we assume UDP/IP sessions for their transmission. We assume UDP sessions because of its connectionless nature which does not permit the retransmission of any packets in error. As a result, we add 20 bytes of an IP header and 8 bytes of a UDP header to the mean customer length which gives a total of 412 bytes. We also know that video services, such as videoconferencing, generally require a higher image
quality which translates into a requirement for a relatively higher capacity. This knowledge is necessary for us to adequately analyse the scenarios designed in Chapter 4.

1.3.2 Voice traffic

Previous studies on the performance of voice traffic and VoIP assume that the voice customer lengths are exponentially distributed [3, 4, 37]. For the purpose of modelling this type of traffic in the network, we extend these assumptions to the work done in this dissertation. From the G.723.1 standard [18], we know that a voice customer is formed every 20 msec, which gives a mean customer length of 106 bits [18]. Like video customers, voice customers are real time and hence, they are delay sensitive. Therefore, we use the connectionless UDP/IP sessions for their transmission which gives 28 extra bytes that are added to the mean customer length resulting in a total of 394 bits.

1.3.3 Data traffic

Finally, we characterize the data traffic. P. Kritzinger et al. [30] has shown that the customer lengths for data customers are exponentially distributed with a mean customer length being 1000 bits. Thus, the mean customer lengths for data customers is assumed to be 1000 bits. Furthermore, the data traffic model considered in this dissertation does not implement any TCP flow control and congestion avoidance mechanisms, but only models packet distribution and retransmission mechanisms.

These values are necessary to model the service time distribution at the routers and the gateway.

1.4 Topological Model

The efficiency of any communication network depends on the design and the topology used. In this section, three different but commonly used network topologies for Wide Area Networks are presented. These include the ring, mesh and star topologies. The trade-offs between these topologies are analysed. These topologies are analysed for their application to geographically hostile regions. Finally, a system overview is presented.
1.4.1 Ring Architecture

The ring topology comprises nodes connected in a ring pattern. To effect communication in this architecture, all nodes have to communicate through their adjacent neighbours. This communication pattern is illustrated in Figure 1.1.

![Ring Architecture Diagram](image)

Figure 1.1: The Ring topology

As can be seen in Figure 1.1, the demand on the link increases as more nodes communicate. This increases the throughput on the link resulting in a requirement for more capacity. From our knowledge of the queueing theory, we know that as throughput increases, the network utilization also increases. This results in a corresponding increase in the network delay as the network load approaches the capacity of the link. As a result, this topology requires more capacity to be able to provide higher traffic demands. Although this topology has less links than a mesh architecture discussed in Section 1.2, the need for more capacity on the link poses a threat of higher link costs. However, due to the less economic activities and sparse population in geographically hostile regions, this threat may be unfounded.

Let $N$ represent the number of nodes in the topology and $d_i$ the unit cost of the link. The cost of the total links in this topology is given by

$$D = \sum_{i=1}^{N} d_i$$  \hspace{1cm} (1.1)

1.4.2 Mesh Architecture

The mesh topology comprises nodes connected in a mesh pattern. To effect communication in this architecture, a node can communicate with any other node without the services of intermediate nodes. This communication pattern is illustrated in Figure 1.2.
1.4.2 Mesh Architecture

This topology has the highest number of links in the network. The number of links to fully connect \( N \) regions is given by

\[
\binom{N}{2}.
\]

For the nine regions, this translates to a total of 36 possible links. Due to this high number of links, this topology incurs the highest link costs. However, due to its fully connected nature, this topology can enable both dynamic and fixed routing patterns. In the dynamic routing pattern, nodes have the option of sending traffic on the shortest or less congested routes. We know that the longer the route, the higher the average network delay. Presenting the network link as a FCFS queue, new packets arriving to the queue are forced to wait in a queue which eventually degrades the performance of the network, a feature that characterises the ring network discussed above. Since packets in this topology can choose to use the shortest routes, the performance of this topology is improved drastically. The throughput through the links is relatively lower than that of the ring topology. This results in less traffic on the links which consequently leads to lower link capacity demand. As a result the system costs are reduced. For technologies such as the optic fibre, the reduction in link capacity has no effect on the cost of the link. This is because the major cost of an optic fibre system is made up of the link and installation costs. For satellite and microwave systems, this reduction in link capacities has a major effect on the system costs.

The cost of the total links in this topology is given by

\[
D = \sum_{i=1}^{N(N-1)} d_i
\]

We noted earlier that the geographically hostile regions have less economic activities, are sparsely populated and are wide spread. This implies that a mesh topology would be too expensive to
provide a backbone network for a developing country with scarce financial resources.

### 1.4.3 Star Architecture

The star topology requires the services of a central node to establish communication as shown in Figure 1.3.

![Star Architecture Diagram](image)

Figure 1.3: The Star Architecture

This topology has the lowest number of links in the communication network. All nodes communicate through the central node and this places higher capacity demand on the central node to be able to handle the traffic in transit. Because it has the least number of links, the capacity demand on these links is higher for nodes that receive and send more traffic. Thus, capacity scaling can be done in this topology by considering each link separately.

The cost of the total links in this topology is given by

\[
D = \sum_{i=1}^{N-1} d_i
\]  

Although this topology has the lowest link cost, it has one major disadvantage. If the central node fails, the whole system is brought down as no communication can be established. Providing a backup node is expensive. The failure to provide resilience in the network presents it as a bad topology compared to the ring and mesh topology.

Using Equations 1.1, 1.2 and 1.3, we find that the mesh topology has the highest link costs followed by the ring topology. The star topology has the least link costs. However, due to the requirement of a central node, the star topology is not ideal. Based on this analysis and the analysis made in each section, the ring topology is adopted for providing connectivity to
1.4.4 System overview

geographically hostile regions. Thus, the design and evaluations carried out in this study are based on this topology. The system overview presented in Figure 1.4 is also based on this topology.

1.4.4 System overview

In Figure 1.4, Region \( i \) refers to a region at which messages arrive destined for the adjacent Region \( j \). As in Section 1.3, the composition of the total traffic in the network is as follows; 50% VoIP, 15% Video and 35% Data.

Figure 1.4 provides a system overview of the type of communication network under study. This network is assumed to be an IP based packet switched communication network in which the messages received at the gateway are broken down into individual packets. These packets are individually addressed, which allows the packet switches to deal with each independently to forward it along the most appropriate and available path. The integrated services provided by this network comprise voice, video and data services.

Packet switched networks were originally developed for communication involving the data class of traffic only which is tolerant to delay and sensitive to errors. The switches that comprise these
1.5 Related Work

Investigating cost-effective communication alternatives for geographically hostile regions is an important area of research which needs consented efforts. In this regard, recent work done by Nkambule and Mneney [27] shows that Radio systems are more cost-effective than traditional wired systems for interconnecting remote sites that are geographically close to each other. However, they do not investigate the most cost-effective means of providing integrated services to regions that are separated by long distances, a feature considered in this dissertation.

To work towards this goal, we build on the work done by Leonard Kleinrock [20] who modelled the ARPANET with the objective of seeking the most cost-effective design to provide communication to the connected regions while at the same time providing guaranteed quality of services. Kleinrock's model was based on a single customer class and different link capacities other than different communication technology alternatives. The model also includes various networks and underlying protocols implemented in those switches were designed based on these assumed characteristics. Voice, video and other stream-oriented applications do not fit this assumption well. However, these networks can be optimized to provide future integrated services through various mechanisms such as compression algorithms, packet prioritization techniques, bandwidth reservation mechanisms, provision of high speed switches, gateways and routers. Other techniques that can be applied to optimize such a network are documented by Darren L. Spohn [36].

Real-time applications, such as voice over IP, would involve the formation of packets comprising multiple pulse code modulation voice bytes, with a compression algorithm being applied in order to reduce the packet size for efficiency reasons as recommended by James F. Kurose et al. [25]. This is done in a gateway which accomplishes the overall process of protocol conversion. The packets are presented to the network in sequence and they traverse the network independently. As these real-time packets are received at the destination gateway, they are re-assembled into original messages. Any error and out of sequence packets are discarded since there is no time for re-transmitting or re-sequencing these delay sensitive packets. Hagsand, Hanson and Marsh [29] have shown, through their experiment consisting of four sites, that voice over IP is feasible in a packet switched communication network. Information on the transmission of non-real-time data packets is found in many good network books, such as Data Network Design by Darren L. Spohn [36], and thus will not be repeated here.
design parameters and variables such as the mean delay, cost constraint, the link capacity and the system throughput. Although this model was not designed for geographically hostile regions, the approach taken by Kleinrock is sufficient to provide a basis for investigating cost-effective communication alternatives for these regions.

We thus, extend this approach to determine the most cost-effective communication technology alternative for geographically hostile regions that require diverse services which include packetized voice, packetized video and data under a common IP based communication network. We also apply the techniques used by P. Kritzinger et al. [30] in modelling the performance of a South African Post Office Network in which the regional nodes and links were subjected to different levels of input traffic to ensure system reliability.

1.6 Dissertation Roadmap

This dissertation is arranged as follows.

Chapter 1

This chapter emphasizes the need to design systems that provide coverage for geographically hostile regions while minimizing the cost of communication and highlights the motivation and objectives on which the work in this study is based. The term geographically hostile region has been defined and examples of such regions are given. This chapter also highlights the approach taken in this dissertation to tackle this problem. The nature of applications targeted for these regions and the type of network are also described. The work done by Kleinrock in modelling the ARPANET and that done by P. Kritzinger et al. [30] has been used as the foundation for the design in this study. To ensure consistency, the notation used in this dissertation has been described and finally the dissertation roadmap presented herein.

Chapter 2

This chapter presents an evaluation of the candidate communication technologies used for intra- and inter-region connectivity respectively. The objective of this chapter is to study the application of these candidate technologies to geographically hostile regions. We study these technologies in two parts. Firstly, we present the long distance technologies and secondly the short distance technologies. The long distance technologies include Microwave, Optic Fibre and
Satellite systems. Since there are three main types of satellites, each of them is discussed with the objective of selecting a type that best suits the needs of geographically hostile regions. The short distance technologies, on the other hand, include DSL and Radio systems.

Chapter 3

This chapter provides the analytical modelling techniques required to understand the work in this study. The analytical model is based on Open Multiclass Queueing Networks. The key features in this chapter are the functionality of a customer, the queueing model, the types of servers, the arrival process, open and closed classes, change of classes, the service disciplines to which the customer classes are subjected and the performance techniques used.

Chapter 4

This chapter deals with the design issues which include the system model, the workload model and the multiclass queueing network analysis (machine) model. The system model describes the topological design and the different design configurations based on the technologies being examined. The workload model describes the usage patterns that the network is subjected to and defines the cost and performance models of these patterns. The multiclass queueing network analysis model serves as the machine model that represents the communication network being modelled.

Chapter 5

In this chapter, the experimental results of the scenarios designed in Chapter 4 are analysed. For each scenario, the average network delay is mapped against the utilisation of the system, and then against the throughput. The aim of doing so is to determine the level of utilization and throughput at which the maximum permissible mean delay for real-time customers is attained. Based on these measures and the cost of each scenario, we determine the cost-performance profiles for each scenario. These cost-performance profiles provide the basis for choosing the most cost-effective scenario that serves as a desirable solution for the problem defined in this chapter. Furthermore, the effect of link errors on the performance of the network is measured.

Chapter 6 presents the conclusions.
Chapter 2

Evaluation of Candidate Technologies

The objective of this chapter is to evaluate the candidate technologies, considered in Chapter 1, which include Microwave, Optic Fibre, Satellite, DSL and Radio technologies for their suitability to serve geographically hostile regions. These technologies are evaluated in terms of their application, cost, advantages and disadvantages. Based on their suitability, cost and the performance measurements presented in Chapter 5, the most cost-effective communication technology for geographically hostile regions is determined.

We group these candidate technologies into two categories which include long distance technologies and short distance technologies. The long distance technologies comprise Microwave, Satellite and Optic Fibre while the short distance technologies comprise DSL and Radio. Since Microwave and Optic Fibre can also be used for short distance communication, they will be considered.

The long distance technologies mentioned above have been considered to provide connectivity between regions that are several kilometers apart.

2.1 Microwave Systems

2.1.1 Application

A typical microwave system consists of three basic components: a digital modem for interfacing with digital terminal equipment, a radio frequency (RF) unit for converting a carrier signal from the modem to a microwave signal, and an antenna to transmit and receive the signal. The combination of these three components is referred to as a radio terminal. A minimum of two
terminals are required to establish a microwave communications link, commonly referred to as a microwave hop as shown in Figure 2.1.

![Digital Microwave Radio Link](image)

**Figure 2.1: Point-to-Point Microwave hop**

The antennas are installed in such a way that a line-of-sight (LoS) is established. The digital modem performs the modulation and demodulation functions. Over long distances, such as 100 Km, repeaters are installed to boost the quality of the signal. For point-to-point communications, such a repeater would consist of two directional antennas facing in each direction of propagation. The repeater unit in a microwave relay system picks up the signal sent to it, amplifies it and re-transmits it to the next LoS antenna on the way to the destination point.

Since this technology depends on the existence of a clear LoS, Valkenburg [40] has shown that the maximum distance between any pair of adjacent antennas is given by

$$X = 7.14\sqrt{\omega h}$$

where $X$ is the distance measured in Kilometers, $h$ is the antenna height in metres and $\omega$ is the adjustment factor. The adjustment factor accounts for the refractions of the microwaves with the curvature of the earth and is approximately $\frac{4}{3}$.

The main components of a Microwave system considered in our costs include a tower, modem, antenna, RF unit, transceiver, multiplexer and a power unit. This equipment is useful for both inter-regional and intra-regional links. However, according to a survey carried out by Okunev [28], and Dong Weiner et al. [35] a variation of Microwave technology called Local Multipoint Distribution System (LMDS) offers high bandwidth for the provision of voice, data and video communication services within a region and hence, is most suitable for intra-regional links.
This technology operates at 27.5-28.35 GHz for two way communication services and 29.1-29.25 GHz for broadcast communication services.

The architecture of LMDS consists of, primarily, four parts which are the networks operations centre (NOC), the fibre base infrastructure, base station and the customer premises equipment (CPE). The NOC contains the network management system equipment which manages the network. The fibre base infrastructure consists of synchronous optical network (SONET) optical carrier, central office equipment, switching system and interconnects with the internet and the public switched telephone network (PSTN). The base station equipment includes the network interface for fibre termination and microwave transceiver equipment. The conversion from fibre to wireless infrastructure is performed at the base station. Figure 2.2 illustrates the LMDS connectivity within a single region.

---

**Figure 2.2: Intra-Region connectivity using LMDS technology**
2.1.2 Cost

As indicated in Chapter 2, the major components that make up the cost of a Microwave system are the digital modem, the RF unit and the antenna. Since the design of the network requires that hops be established in two directions from each region, two directional antennas will be required at each region. Furthermore, the long distances involved dictate the use of repeaters between a pair of regions. Since, in general, Microwave systems can transmit up to 50 km, several repeaters need to be installed. Applying this to Figure 4.1, we find that 128 repeaters are required to ensure connectivity. This, therefore, means that the highest cost of this system is made up of the cost of the repeaters.

Table 2.1 shows the cost of the standard radio unit, repeaters, installation and maintenance.

<table>
<thead>
<tr>
<th>Description</th>
<th>Unit cost USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site equipment$^1$</td>
<td>28,820</td>
</tr>
<tr>
<td>Repeater tower</td>
<td>25,000</td>
</tr>
<tr>
<td>Multiplexer</td>
<td>2,320</td>
</tr>
<tr>
<td>Installation</td>
<td>1,599</td>
</tr>
<tr>
<td>Maintenance</td>
<td>3,000</td>
</tr>
</tbody>
</table>

Table 2.1: Unit costs for the microwave system

2.1.3 Advantages

- microwave systems are quicker to deploy
- microwave systems permit the transmission of full motion video
- microwave antennas are directional and hence, there is control over who receives the signal
- voice, video and data signals can be transmitted on the same link
- no right-of-ways are required
- LMDS provides high capacity for regional connectivity

$^1$Protected transceiver, RF unit, digital modem and antennas
2.1.4 Disadvantages

- Microwave systems require a LoS from the transmitter to the receiver, as a result several antennas have to be installed as repeaters.

- The LMDS waves suffer interference caused by changes in the weather. These waves at high frequencies are absorbed by water in tree leaves and in rainfall which cause the signals to deteriorate. However, depending on the severity of the rainfall (tropical regions) and distance travelled, LMDS systems compensate by providing a large signal-to-noise ratio margin. Furthermore, rain detection devices included in the transmitter can automatically boost the transmitter power when it rains.

- The links in a microwave system are not very secure. However, for services such as tele-education, the need for security is not a critical requirement. For other services such as tele-medicine and electronic commerce that require secure sessions, the security of the links can be enhanced by the use of appropriate cryptographic algorithms.

- Microwave systems experience losses with a bit error rate of $10^{-6}$ due to multipath dispersion and Rayleigh fading.

- LMDS are very distance limited, about 8 Km maximum.

- A license is required to operate a Microwave system.

2.2 Optic Fibre Systems

2.2.1 Application

This technology provides the highest capacity, in access of a Tbps [34], to the end users. To achieve this increase in capacity, a technology referred to as Wavelength Division Multiplexing (WDM) [17] is used. This technology allows the sending of many modulated light waves at different wavelengths down each fibre strand. These light waves are separated at the terminal equipment. A study by Riezenman [33] has reviewed that the advent of the dense WDM has increased the wavelength multiplexing capabilities to 320 wavelengths on a single fibre strand from the 160 wavelengths on a single fibre strand using WDM.

The main components that comprise a digital fibre system include the fibre cable, multiplexer, codec, optical transmitter, optical receiver or photodetector and repeater.
2.2.2 Cost

The Optic Fibre cable has three forms namely, the Single-mode, Graded-index and the Step-index. These forms have varying characteristics in terms of core diameter, cladding diameter, attenuation and bandwidth, which have a bearing on the performance of a particular form. Table 2.2 highlights the varying characteristics.

<table>
<thead>
<tr>
<th>Fibre type</th>
<th>core (μm) diameter</th>
<th>cladding diameter (μm)</th>
<th>850nm (dB/km) (Max)</th>
<th>1300nm</th>
<th>1500nm</th>
<th>Bandwidth (MHz/km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-mode</td>
<td>8.1</td>
<td>125</td>
<td>2.3</td>
<td>0.5</td>
<td>0.25</td>
<td>5000 @ 850nm</td>
</tr>
<tr>
<td>Graded-index</td>
<td>50</td>
<td>125</td>
<td>2.4</td>
<td>0.6</td>
<td>0.5</td>
<td>600 @ 850nm 1500 @ 1300nm</td>
</tr>
<tr>
<td></td>
<td>62.5</td>
<td>125</td>
<td>3.0</td>
<td>0.7</td>
<td>0.3</td>
<td>200 @ 850nm 1000 @ 1300nm</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>140</td>
<td>3.5</td>
<td>1.5</td>
<td>0.9</td>
<td>300 @ 850nm 500 @ 1300nm</td>
</tr>
<tr>
<td>Step-index</td>
<td>200 or 300</td>
<td>380 or 440</td>
<td>6.0</td>
<td></td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

Table 2.2: Typical Fibre Characteristics Source: Sterling D., Technicians Guide to Fibre Optics, 1993

Table 2.2 clearly shows that single-mode fibre has the greatest bandwidth and the lowest attenuation. Based on this fact, the Optic Fibre link depicted in Figure 4.2 is assumed to be a single-mode fibre link.

The multiplexer converts the signal to and from an electrical signal. The codec changes the signal to digital and the optical transmitter converts the signal to an optical signal. At the receiving end, the receiver reconverts the optical signal to digital. The transmitters are either lasers or light emitting diodes. The receivers are either positive-intrinsic-negative or avalanche photodiode (APD). Since the design to which this technology is subjected involves distances greater than 100 Km, see Figure 4.1, APD is a more preferrable option. The repeater is required to reconstruct the signal to extend the transmission distance.

Although improvements in Optic Fibre technology has made them increasingly attractive, their suitability for geographically hostile regions requires re-examining and comparing the results against competing technologies. This is done in Chapter 5.

2.2.2 Cost

The costs presented in Table 5.3 are merely unit costs of the major components in a region. The
2.2.3 Advantages

link costs and subsequently the system costs are discussed in a greater detail in Chapter 5 where each scenario developed in Chapter 4 is analysed.

<table>
<thead>
<tr>
<th>Description</th>
<th>Unit cost (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single mode Fibre (4 strands)</td>
<td>1.36/meter</td>
</tr>
<tr>
<td>Multiplexer</td>
<td>3,758</td>
</tr>
<tr>
<td>Codec</td>
<td>13,000</td>
</tr>
<tr>
<td>E1 Fibre Modem</td>
<td>925</td>
</tr>
<tr>
<td>Optical Transmitter</td>
<td>1,400</td>
</tr>
<tr>
<td>Optical receiver</td>
<td>1,450</td>
</tr>
<tr>
<td>Repeater</td>
<td>213.97</td>
</tr>
<tr>
<td>Copper to Fibre converter</td>
<td>173.00</td>
</tr>
<tr>
<td>Installation</td>
<td>0.49/meter</td>
</tr>
<tr>
<td>Maintenance</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.3: Inter-regional connectivity costs for configuration 3 - Source: University of Cape Town, ITS Department

2.2.3 Advantages

- offers the highest capacity to the end users
- suffers the least signal attenuation with a bit error rate as low as $10^{-12}$
- difficult to tap and hence highly secure
- The high durability of fibre results in low maintenance costs
- The large bandwidth of fibre allows voice, video and data to be combined on one line
- Permits full motion video transmissions
- Fibre is unaffected by electromagnetic currents, static interference, radiation and electric motors

2.2.4 Disadvantages

- The installation of the cable is labour intensive
- Fibre systems require high start-up costs, see Chapter 5, Scenario 3
- Right-of-way costs for placing cables in the ground are required
- Difficult to install additional links on existing infrastructure
2.3 Satellite Systems

This technology comprises four types of satellite systems namely the Geostationery (GEO) satellite system, the Middle Earth Orbit (MEO) satellite system, the Low Earth Orbit (LEO) satellite system and the High Elliptical Orbit (HEO) satellite system. HEO satellite systems were designed (by the Russians) to serve the polar regions and therefore are not applicable to the problem being solved. Thus only the remaining three systems are examined and the most suitable system that meets the requirements of geographically hostile regions is further evaluated against other competing technologies.

2.3.1 GEO Satellite systems

2.3.1.1 Application

Located at an altitude of 36,000 Km above the equator, a geostationery satellite system provides different kinds of services which include voice, data and video. This system operates using the bent pipe architecture in which the satellite operates as a repeater that translates the radio frequency RF energy from the uplink onto the downlink beam. These systems are increasingly being used to provide coverage to underserved regions.

There are three main applications that can be implemented in geographically hostile regions using these systems. These include Single Carrier Per Channel (SCPC), Very Small Aperture Terminals (VSATs) and SCPC Demand Assigned Multiple Access (SCPC DAMA).

SCPC applications have the simplest configuration to implement and operate. The configuration involves a collection of terminals, 64Kbps link and the hub. Each end user is assigned a separate single channel and the system does not permit the sharing of the channel by multiple users. However, this application is rigid and does not adapt to changes in the workload on the network. It also does not scale up when the need to increase the number of regions to be connected arises.

The VSAT applications are a collection of an intelligent earth station(Hub) and terminals that support a variety of communication services such as voice, data and video. To set-up communication using VSAT, a terminal initially requests the hub for a channel to be assigned for transmission. Only after the hub has acknowledged the electronic request, which is in the form of a short signal burst, does the terminal start transmitting. The end users in a region share capacity using the time division multiple access in which a single 128, 256, 512 or 1024 Kbps
pipe is time shared as IP sessions to transmit traffic to other regions as shown in Figure 2.3. The VSAT Hub transmits to these regions at capacities ranging from 64kbps to 2Mbps. By design, VSAT applications are hub centric, which means that all the traffic passes through the central hub which is referred to as the gateway. The implementation of this application enables the networked regions to access the multimedia services from a central location where the hub is installed. Unlike the SCPC, this application is scalable. A snapshot of a VSAT configuration is presented in Figure 2.3.

[Diagram of VSAT configuration]

SCPC DAMA applications are a hybrid of SCPC and VSAT applications. Instead of using the central hub as the case is with VSATs, these applications have the ability to establish peer-to-peer communications without the necessity to go through the hub. This is because the electronics associated with its central hub are partly incorporated into the individual terminals for direct transmission and reception from the satellite without the physical involvement of the hub. The incorporation of part of the electronics associated with the central hub into the individual terminals makes the cost of these terminals more than that of the VSAT terminals.
2.3.1 GEO Satellite systems

2.3.1.2 Cost

The VSAT equipment comprises a 4.8 C-Band Rx/Tx antenna for the Hub, 3.8m C-Band Rx/Tx antenna for the rest of the geographically hostile regions, 10W C-Band transciever, and a satellite modem. To determine the cost of this equipment and that of leasing bandwidth and maintenance, we contacted several companies to request for quotations. Only two companies responded, Plessey Solutions (Pty) Ltd and Global (Pty) Ltd. Their costs were identical and hence we picked only one of the two quotations.

<table>
<thead>
<tr>
<th>Description</th>
<th>Unit Cost (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.8m C-Band Rx/Tx antenna (gateway)</td>
<td>32,000</td>
</tr>
<tr>
<td>3.8m C-Band Rx/Tx antenna (others)</td>
<td>15,800</td>
</tr>
<tr>
<td>10W C-Band Transciever</td>
<td>12,300</td>
</tr>
<tr>
<td>Multiplexer</td>
<td>2,320</td>
</tr>
<tr>
<td>Satellite Modem with V.35 interface</td>
<td>1,900</td>
</tr>
<tr>
<td>Installation</td>
<td>3,500</td>
</tr>
<tr>
<td>Maintenance</td>
<td>2,000/year</td>
</tr>
<tr>
<td>Link lease (E1)- 1 year</td>
<td>404,628/year</td>
</tr>
</tbody>
</table>

Table 2.4: Inter-regional connectivity costs for configuration 2 using a leasing option

2.3.1.3 Advantages

- GEO systems provide a wider coverage area
- The satellite antennas can be focused on regions that require the communication services and thus optimising the system usage
- GEO systems, VSAT in particular, are scalable. They can provide services at data rates ranging from 64kbps to 2Mbps
- The distance between the regions has no cost or performance implications over the network

2.3.1.4 Disadvantages

- The links in a GEO system have high propagation delay which is calculated to be 120 msec on either the uplink or downlink
2.3.2 MEO Satellite systems

2.3.2.1 Application

These systems comprise a constellation of satellites in order to provide global coverage of the service regions. Jian, Hu, and Yeung [19] have shown that a two-tier network architecture in which a layer of MEO satellites hovers above a layer of LEO satellites to provide reduced re-routing probability for end-to-end connections is feasible. However, this architecture attracts high system costs because of the use of both MEO and LEO systems.

The usage of MEO satellite systems is not widespread as it only currently services regions in the northern hemisphere. Since most of the geographically hostile regions defined in Chapter 1 are located in the Southern hemisphere, this system would not be applicable.

2.3.3 LEO Satellite systems

2.3.3.1 Application

In a LEO satellite system, satellites move in circular or elliptic polar orbits. Elliptic polar orbits are designed to increase the visibility period of the satellite over the highly populated regions such as the northern hemisphere [14]. The objective of the evaluation in this chapter is to determine the suitability of the communication technologies to geographically hostile regions. Thus, we consider the LEO satellites in circular polar orbits.

According to the study done by Akyildiz [15], LEO satellite systems in circular polar orbits can either be implemented as systems that have on-board processing capabilities with inter-satellite links (ISL), like Teledesic [38], or merely as bent-pipe systems, like Globalstar [8]. On-board processing satellites with ISLs perform dynamic routing of traffic and achieve load balancing [43] without the intervention of earth stations. The inter-satellite links enable the establishment of hand-offs between these satellites. This technology is also designed for regions
that are densely populated and not the geographically hostile regions defined in Chapter 1 as can be seen from the illustration of Figure 2.4.

Figure 2.4: The connectivity pattern for LEO satellites with onboard processing with ISL

Figure 2.4 shows the connectivity pattern of a LEO system with ISLs. The coverage regions are divided into several footprints. As the satellite traverses its orbit, it provides antenna beams to these footprints to enable the end-users transmit their traffic to the satellite. These footprints are designed in such a way that any adjacent pair overlap to enable effective hand-off. Since the geographically hostile regions are defined as being sparsely populated and wide spread, a number of these footprints would be idle. This means that the system would have to perform frequent load balancing and hand-off functions, which would increase the system's operational costs as these functions require more transmitter power. Therefore, in order to recover these
2.3.3 LEO Satellite systems

operational costs, the service charges would have to be high. This consequently would limit the affordability of these systems by the end-users in geographically hostile regions as a result of less economic activity in these regions.

A bent-pipe system involves satellites that do not have ISLs and depends on the earth stations to perform the routing and load balancing functions. This implies that this system depends on terrestrial network backbone. Since geographically hostile regions lack telecommunications infrastructure, this system would also not be ideal.

LEO systems have cellular configuration which means that the footprints have to be partitioned into cells. A study by Pratt et al. [32] shows a specific satellite system with a footprint diameter of 4,021 Km which has 48 cells in each footprint. Since each cell represents a spot beam, this translates to 48 spot beams as illustrated in Figure 2.5.

Since the distribution of users in geographically hostile regions is sporadic and wide spread, the configurations in Figures 2.4 and 2.5 entail that some of the spot beams will not be used resulting in wastage of transmitter energy which is caused by constant signaling.

2.3.3.2 Advantages

- The system provides global coverage.
- Has lower propagation delay compared to GEO systems.

2.3.3.3 Disadvantages

- The number of satellites in the constellation always remains fixed, it is not possible to increase the bandwidth capacity by launching new satellites.
2.3.4 Selection of a suitable satellite system

- The footprint covered by each transmitter is very large and available bandwidth needs to be shared by people within the footprint to provide lower rates. Geographically hostile regions are sparsely populated and wide spread and thus cannot benefit from the lower rates which result from shared bandwidth.

### Table 2.5: Prominent Broadband Systems. Source: IEEE’s GLOBECOM’99-MULTIMEDIA SATELLITE COMMUNICATIONS, pg 1167

<table>
<thead>
<tr>
<th>Program</th>
<th>Orbit</th>
<th>Band</th>
<th>Sats</th>
<th>US$M</th>
<th>Gbps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astrolink</td>
<td>GEO</td>
<td>Ka</td>
<td>9</td>
<td>3,994</td>
<td>61</td>
</tr>
<tr>
<td>EuroSkyWay</td>
<td>GEO</td>
<td>Ka</td>
<td>5</td>
<td>2,500</td>
<td>45</td>
</tr>
<tr>
<td>INTELSAT</td>
<td>GEO</td>
<td>Ka</td>
<td>6</td>
<td>2,500</td>
<td>48</td>
</tr>
<tr>
<td>OXYGEN</td>
<td>None</td>
<td>Fiber</td>
<td>N/A</td>
<td>8,000</td>
<td>2,560</td>
</tr>
<tr>
<td>SkyBridge 2</td>
<td>LEO</td>
<td>Ku</td>
<td>80</td>
<td>9,600</td>
<td>67</td>
</tr>
<tr>
<td>Spaceway</td>
<td>GEO</td>
<td>Ka</td>
<td>8</td>
<td>3,200</td>
<td>40</td>
</tr>
<tr>
<td>Teledesic</td>
<td>LEO</td>
<td>Ka</td>
<td>288</td>
<td>8,900</td>
<td>64</td>
</tr>
<tr>
<td>WEST</td>
<td>GEO</td>
<td>Ka</td>
<td>6</td>
<td>2,600</td>
<td>55</td>
</tr>
<tr>
<td>ICO Global Communications [5]</td>
<td>MEO</td>
<td>Unknown</td>
<td>12</td>
<td>2,600</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Although GEO systems have high latencies, they have the ability to provide focused services to regions that require them. As a result, these systems are a better choice for the geographically hostile regions.

From the evaluations made in this section, Microwave systems, Optic Fibre systems and GEO satellite systems have been identified as technologies that provide long distance solutions. Each of these technologies is then subjected to the configuration of Figure 4.1 and the M-QNA model of Figure 4.2.4, in Chapter 4 for the purpose of determining their performance and cost. The results are discussed in Chapter 5.

The short distance technologies are used to interconnect sites located within one region. For
2.4 DSL Systems

2.4.1 Application

DSL is a technology that provides high bandwidth communication services to users over ordinary copper telephone lines [6]. The use of this technology enables the transmission of traffic directly to a computer and eliminates the need to convert from analogy to digital and vice versa. As a result, a much wider bandwidth for the transmission of traffic is enabled. Depending on the quality of installation, the maximum repeaterless range that can be covered is about 5.5 Km.

DSL has three variations which include Asymmetric DSL (ADSL), High bit rate DSL (HDSL) and Very high bit rate DSL (VHDSL) [41]. ADSL offers asymmetric services because it provides different bandwidth from the exchange to the subscriber (downstream) and the subscriber to the exchange (upstream). The recommended upstream and downstream bandwidths are 640 Kbps and 6.144Mbps respectively, and covers a distance of approximately 3.5 km [1]. Both HDSL and VHDSL offer symmetric services whereby the downstream and the upstream bandwidth are the same. Since these geographically hostile regions are described as having less economic activities, more traffic is expected on the downstream than the upstream [17], and hence ADSL becomes a better choice for the intra-region connectivity.

2.4.2 Cost

The Subscriber set-up cost for ADSL is **USD 179.53** and a monthly running cost of **USD 47.69** [1].

2.4.3 Advantages

- The asymmetrical nature of the ADSL variation of DSL enables effective utilisation of the transmission channel. This is because the larger portion of the traffic is found on the downstream than on the upstream
- It allows the Service Providers to use a much wider bandwidth for the transmission of traffic within intra-regions
2.4.4 Disadvantages

- DSL is a local loop technology and therefore cannot be used for inter-region connectivity.

2.5 Radio Systems

2.5.1 Application

The nature of the required network for which these systems are being evaluated is a packet switched network discussed in Chapter 1. Therefore, in evaluating this system, we assume a packet radio network. A packet radio network is a collection of radio transmitters and receivers (transceivers) located in a geographical region. Associated with each transceiver is a transmission range, which depends on its transmission power.

Among the Radio systems currently in use, a Rooftop Community Radio System (RCRS) [16] is best suited for communication within a geographically hostile region. It uses a different approach to that of wireless systems, aiming to create fast and robust networks with an interlocking web topology that is constructed entirely by end-users located in a rural setting. The RCRS allows many users to share a link by joining a local Rooftop network. To join the network, users purchase a Spirit 2000 internet radio and antenna at a cost of USD 500.00, consisting of a high speed (115Kbps to 10 Mbps) digital radio that uses ISM bands and an embedded microprocessor to run Internet Radio Operating System (IROS) software.

The RCRS is a self managing web of peers in which each radio serves not only as a user’s connection to the network, but also as an automatic repeater to forward other users’ traffic as required. This makes the system unique in that the traffic is automatically routed via the nearest repeater to its final destination. If a potential user is not within range of the internet gateway, then a closer user must install a system to provide the route to the gateway.

The IROS software in the radios automatically control the routing of packets across multiple links between their sources and destinations. This software requires no configuration or intervention by users beyond that which is required to connect a computer to the internet over a typical dial-up line. A gateway hub unit to provide an interface to the LAN is also necessary.

The system covers a range varying from 8 Km, in multipoint-to-multipoint mode, 18 Km, in point-to-multipoint mode, and 30-50 Km in point-to-point mode. The configuration of a RCRS is depicted in Figure 2.6.
Figure 2.6: An illustration of an interlocking web topology radio system

2.5.2 Cost

It costs USD 500.00 for a Subscriber to join a RCR system. This cost covers the antenna and the internet radio.

2.5.3 Advantages

- It works out cheaper to connect to an existing link
- The system offers much higher capacity than that of ADSL within intra-regions
- The system can be implemented in different modes

2.5.4 Disadvantages

- For a potential user not within range of the internet gateway, it becomes difficulty to get service if the closer person or organisation does not need the service
- Systems using this configuration are at high risk of being infected by computer worms if appropriate measures are not put in place
2.6 Conclusion

In this chapter, two types of communication systems are evaluated, the short and long distance communication systems. Although the long distance communication systems such as Microwave and Optic Fibre can be used within short distances, the main short distance communication systems are the DSL and Radio systems. A comparison of these two systems indicates that a Packet Radio communication system is more suited for intra-regional communication within geographically hostile regions due to its lower setup costs and higher capacity.

In the same vein, the long distance communication systems such as Microwave and Optic Fibre have higher installation costs (to be discussed in Chapter 5) than that of a Packet Radio system. This leaves the Radio system as the only cheaper communication system for intra-region connectivity.

At this point, we conclude that the three long distance technologies are all capable of providing communication services to geographically hostile regions.
Chapter 3

Analytic Modelling

In this chapter, we present the relevant details of an analytic technique called a Multiclass Queueing Network Analysis (M-QNA) used to model the scenarios presented in Chapter 4. M-QNA model involves the analysis of queueing networks which consist of a collection of (connected) single server stations and multiple customer classes. The customers arrive to the communication network as independent Poisson streams with exponentially distributed inter-arrival times. Upon arrival at the single server station, these customers are served according to the prescribed service discipline and immediately leave the service facility.

This chapter is organised as follows; we first present the modelling techniques, then the queueing model followed by the description of the functionality of a customer, the arrival process, class switching, types of classes in multiclass networks, types of servers, the service disciplines and the performance measures. We finally summarize the major points in the conclusion.

3.1 Modelling techniques

The two commonly used modelling approaches are simulation and analysis. A simulation model is normally used when equations that relate the desired performance quantities to the parameters are insoluble or when it is not even obvious what the equations of the system are. Simulation models are very useful in investigating the detailed operation of a system, which can lead to key insight into the system [36]. Thus, by using this model, a possible sequence of events in the network is generated and the desired performance quantities for that sequence are derived.

Since there is an infinite number of possible sequences of events, a large number of such events has to be simulated before any confidence can be placed in the model. As a result, simulation
models are more costly [30] in terms of computer processing time.

Analytic models seek equations relating the desired performance quantities to the parameters. Although these models represent less detail, they are generally much cheaper to use than simulation models of the same system [30]. Based on this, analytical models have been used to evaluate the performance of the network configurations presented in Chapter 4.

The principal output of an analytical model are mainly measures of performance and cost. The measures of performance being the average network delay that a customer might be forced to endure, the number of customers that are served per unit of time (throughput) and the utilization of the system. These are discussed further in Section 3.9. The cost constitutes the sum of all the costs required to set-up each link in the communication network. The cost and performance are then used to derive a cost-performance ratio for each scenario to assess its cost-effectiveness.

3.2 The Queueing Model

In this section, we present the basics of the queueing model that describe a queueing system.

![Queueing Model Diagram]

Figure 3.1: The queueing model

In this model, customers arrive to the queue at the mean arrival rate $\lambda$ and leave the system at a mean departure rate equal to the mean arrival rate after waiting in the queue and receiving service at the server. In this system, customers arrive to the queue as independent Poisson streams and are served at an arbitrarily average service rate $\mu$ which is almost general.

3.3 Functional description of a customer

In this dissertation, we use the term “customer” to represent packets that receive service in a communication network. The format of these customers comprise header and payload informa-
We thus define a customer as a packet comprising payload information and protocol overheads. The behaviour of the customer is discussed in Sections 3.2 and 3.5.

### 3.4 Arrival Process

The arrival process describes the average rate, denoted by $\lambda$, at which customers enter the system. Some models comprising voice, video and data [3, 4, 37, 7] assume this arrival process to be Poisson. We therefore, describe the Poisson arrival process in detail.

The Poisson process is often used to model random arrivals of customers. In applying this arrival process, the following assumptions are made:

1. In a small time interval $\Delta t$, the probability of exactly one arrival is proportional to the size of the interval. i.e. $\text{Prob}[1 \text{ arrival in } \Delta t] = \lambda \Delta t + o(\Delta t)$.

2. In $\Delta t$, the probability of more than one arrival is negligible. $\text{Prob}[1 \text{ arrival in } \Delta t] = o(\Delta t)$.

3. Arrivals are independent of other arrivals and also independent of the time since the last arrival.

Furthermore, if we denote the number of arrivals in interval $(0, t)$ by $N(t)$, then the probability of having $n$ arrivals is given by

$$
\text{Prob}[N(t) = n] = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.
$$

### 3.5 Class Switching

This section presents the theory required to understand how customers may change their classes. The key objective is to model this change as it affects the behaviour of the communication network being investigated. Customers are forced to change their class largely due to errors on the communication link. The customers that change their class, from class $c$ to class $z$, are put back on the output link and retransmitted as illustrated in Figure 3.2.
3.6 Multiclass Networks

The net effect is that a retransmission adds the delay of contending for the medium and retransmitting the customer. If the rate of retransmission becomes high, the network gets congested causing the average network delay $T$ to tend to infinity. In Figure 3.2, we see that customers arrive to the system as class $c$ customers. They receive service at the gateway and are subsequently put on the communication link. While on the link, a fraction of the customers switch classes from $c$ to $z$. The customers that change class are kept in the timeout server for a period of time equal to their timeout duration and then put back on the link to model the retransmission process. For example, a system involving satellite links would have a timeout duration equal to twice the propagation delay plus the processing time at the satellite of approximately 250 msec. It is these retransmitted customers that begin to load the network and cause link failures even before the theoretical threshold is reached. We define the theoretical threshold as the throughput limit at which the queueing system experiences an infinite queue in the absence of link errors. Since customers are conserved, the blocking probability at the timeout server is zero.

We know, from [22], that in the absence of error clustering, the probability that a block of $n$ bits is in error is given by $1 - (1 - e)^n$, where $e$ is the bit error rate of the link. This gives us the fraction of timeouts in Figure 3.2.

Figure 3.2: Change of customer classes

Multiclass networks include multiple customer classes within a single communication network. The state probability of such a network is represented by $\pi(r_1, \ldots, r_N)$. The normalization condition that the sum of the probabilities of all possible states ($r_1, \ldots, r_N$) is 1 must also be
satisfied. A multiclass network can have both open classes and closed classes.

An open class is one that has an infinite source population of customers where customers arrive, receive service, and leave the network. The number of open class customers in the system varies dynamically and the number of possible system states is infinite.

A closed class is one that has a fixed finite number of customers. These customers circulate among the system resources where they receive service and do not leave the system. Hence, a closed class does not admit new customers. The number of customers in the system at any time is fixed. Also, the number of possible system states is finite.

Figure 3.3 shows that an open class has a source and a sink where as a closed class has neither source nor sink. An example of an open class is a class of voice signals in a telecommunication system. The voice signals from source enter the system through a gateway and after receiving service, they leave the system to the destination. On the other hand, an example of a closed class is a class of data on disk that moves between the main memory and the disk. It circulates between the disk and the main memory and thus does not leave the system.

3.7 Types of Servers

A queueing system can either have a single server, presented in Figure 3.1, or multiple servers. In a system with multiple servers, customers are allocated to the servers where they are served according to some service discipline. The allocation of these customers is also done according to some scheduling discipline. An example of a system with multiple servers is illustrated in Figure 3.4.

If we assume the same amount of network load, we find that systems with multiple servers
are less utilized than those with single servers. As a result, single server systems offer an efficient utilization of system resources than multiple server systems. Logically, its cheaper to implement a single server system than a multiple server system. Based on this knowledge, the model considered in this study is a single server multiclass queueing network model.

3.8 The Service Discipline

This theory defines the order of service and determines which customer is served at any given time. There are essentially many service disciplines that are used in queueing systems. This section, however, examines only those service disciplines that are relevant to this study [11]. The M-QNA model allows for 3 different service disciplines; Processor Sharing (PS), First Come First Serve (FCFS) and Infinite Server.

3.8.1 PS

This service discipline allows customers from the three classes to receive service. It is as if all customers are served simultaneously and the service time is increased correspondingly. This mechanism can be used to provide some kind of priority service based on time. Different classes can be allocated different service times. Processor sharing, in this case, becomes the limiting behaviour of a timesliced server when the timeslices become very small. The various customer classes may have different service time distributions which maybe almost general. Because of its ability to provide time priority to different classes, this service discipline is used to model the service time distribution at the routers, gateway and the switch at the satellite.
3.8.2 FCFS

This service discipline allows customers to receive service in the order of entry into the queue. The rate at which customers receive service depends on the propagation time of the link thus, this service discipline is normally used to model the service on the output links in a communication network.

3.8.3 Infinite Server (IS)

This service discipline is such that there is always a server available for a customer entering the queue. At the server, a customer is delayed for a length of time equal to its service demand at the centre and no longer. This is applied in modelling the time-out of the customers that experience errors on the link. For customers that belong to the voice and video classes, an infinite service is offered to ensure that they are not delayed any further. Thus, no queue ever forms at the IS.

3.9 Performance Measures

The measures of performance considered in this study include the average network delay, throughput and the utilization. The average network delay is the average time the customer spends in the system and is referred to as 'residence'. For the network under consideration, described in Chapter 1, Section 1.4, residence is a very critical measure for customer classes such as voice and video because of their inherent sensitivity to delay. For example, a customer belonging to the voice class is associated with a time stamp within which it should reach its destination otherwise it would be dropped or simply rendered useless. Stephen et al. [32], in his survey of the Iridium system, and Krunz [23] indicate that this time stamp which is referred to as the maximum tolerable mean delay for voice and video is 400 msec. This measure is used as an upper bound for the mean delay of real-time customers through the system to guarantee the acceptable QoS levels.

The inherent delay in the system is caused mainly by two random variables, the throughput and the mean service time. When the throughput is increased, the mean delay also increases basically because there are more customers to be served and therefore these customers are forced to wait in a queue. Likewise, increasing the mean service time implies that each customer stays in the service facility for relatively a longer time which would result in more customers waiting in the queue. This gives rise to an increase in the mean delay. As this increase becomes higher,
the number of customers in the communication system also increases, which eventually leads to the network being in a congestion state and its utilization is seen to tend to unity. These performance measurements are further discussed in Sections 3.9.1, 3.9.2 and 3.9.3.

3.9.1 Throughput

The throughput $\tau$ of an elementary queueing system is defined as the mean number of jobs whose processing is completed in a single unit of time, i.e. the departure rate. It is measured in packets per second (in packet switched networks), frames per second (in Frame Relay networks) and cells per second (in ATM networks). For the purpose of this study, a packet switched network is assumed and thus throughput will be measured in packets per second.

The amount of throughput through a system depends on factors such as the packet size measured in bytes or Kb, the service time distribution at the service centres and the propagation delay of the link. These factors determine how much throughput goes through the communication system. Although the packet sizes and service time distribution at the gateway and routers maybe varied to increase or reduce the amount of throughput, the fact that the link propagation delay is fixed limits the amount of throughput in the network.

Since we have assumed that the communication network reaches an equilibrium state, the throughput at each regional router $i$ is given by

$$\tau_i = \lambda_i + \sum_{j=1}^{N-1} \lambda_j p_{ji}; \quad 1 \leq j \leq N - 1 \quad (3.2)$$

where $p_{ji}$ is the probability that a customer upon receiving service at regional router $j$ will proceed to regional router $i$ and $\sum_{j=1}^{N-1} \lambda_j p_{ji}$ is the sum of the rate of customers that arrive at regional router $j$ destined for regional router $i$.

3.9.2 The Mean Delay

This is a measure of network performance used to characterize the average delay a customer suffers through the system. It is best described using John D. C. Little's Law [24] which states that the average number of customers in a queue is equal to the average rate at which they arrive to the queue multiplied by the average time they spend in the queue.

We illustrate this measure by considering customers originating from router $i$ destined for router
3.9.2 The Mean Delay

Let \( \frac{T_{ij}}{\lambda} \) represent the fraction of the total customers in the system that suffer the average delay \( T_{ij} \). \( T_{ij} \) represents the average delay on the link plus the average delay suffered in the service facility. Applying Little's Law, the average network delay \( T \) is given by the sum of the product of the number of customers in the system and the mean delay \( T_{ij} \) expressed by Equation 3.3

\[
T = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{T_{ij}}{\lambda}.
\]

Equation 3.3 represents the decomposition of the communication network on the basis of the origin-destination pairs. The measure \( T \) is very important in determining the most suitable network configuration to provide communication services to geographically hostile regions. Its importance lies in its adequacy to sufficiently characterize the average delay suffered by customers with similar statistical characteristics when subjected to different network configurations. This measure is obtained using the keyword “RES” in the MicroSNAP modeling language. Using Figure 3.5, we show the behaviour of the average network delay as a function of utilization.

![Figure 3.5: Average network delay as a function of utilization, \( \rho \)](image)

The average network delay reaches its threshold at the point \( T^* \), after which, it is seen to grow without bound causing the system to be unstable. Thus, in designing telecommunication systems, it is important to ensure that the average network delay is kept below its threshold.

Another measure relevant to this study is the utilization of the system described in Section 3.9.3.
3.9.3 Utilization

For any queueing system that consists of a single server, such as those designed in Chapter 4, the utilization $\rho$ is defined as the fraction of the time in which the server is busy or occupied. This utilization is given by

$$\rho = \frac{\text{mean service time}}{\text{mean interarrival time}} = \frac{\text{arrival rate}}{\text{service rate}} = \frac{\lambda}{\mu}.$$  \hspace{1cm} (3.4)

The most important point about this measure which directly applies to the work done in Chapter 5 (Experimentation) is the condition for stability given by the inequality $\rho < 1$. This condition puts a limit on the arrival rate $\lambda$ such that $\lambda < \mu$. In other words, customers must not arrive to the system at a rate higher than that at which they are being served. This knowledge allows us to effectively determine appropriate random values for the service time distribution such that the arrival rate is kept within permissible values.

3.10 Conclusion

In solving any problem that requires modelling, either a simulation or analytical approach is normally used. The use of either approach depends on the amount of information required from the model. In this study, our interest is in the cost and performance of the scenarios designed in Chapter 4. Since the principle output of an analytical model are measures of performance and cost, we adopt the analytic technique in this study.
Chapter 4

Case Study

This chapter presents the system, workload and the M-QNA models used to determine the most cost-effective communication system required to provide communication services to geographically hostile regions defined in Chapter 1 and evaluated in Chapter 2. The design of these models is based on the theory of Open Multiclass Queueing Network Analysis described in Chapter 2.

For the purpose of being able to design and characterize these models, Zambia, being a country within the geographically hostile regions with a population of 10.3 million people, was chosen as a case study. Any other country within the geographically hostile regions could have been chosen. The choice of Zambia was motivated by the experience that the author has with this country and the current need for communication infrastructure in the regions indicated in Figure 4.1.

To enable sufficient data for parametric analysis and for deriving probability distributions, nine regions which include Lusaka, Livingstone, Zambezi, Mwinilunga, Mbala, Mansa, Mkushi, Lundazi and Siavonga were selected. These regions, except Lusaka, were selected mainly because they have inadequate communication infrastructure and in some cases it is non-existent. Lusaka was included in this design for the purpose of linking the geographically hostile regions to external networks. These regions are sparsely populated with the population distribution as shown in Table 4.1.
4.1 Existing Infrastructure

<table>
<thead>
<tr>
<th>Region</th>
<th>Total Population</th>
<th>No. of Households</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lusaka</td>
<td>1,103,413</td>
<td>198,290</td>
</tr>
<tr>
<td>Siavonga</td>
<td>58,932</td>
<td>10,624</td>
</tr>
<tr>
<td>Lundazi</td>
<td>236,732</td>
<td>46,738</td>
</tr>
<tr>
<td>Mbala</td>
<td>161,533</td>
<td>30,907</td>
</tr>
<tr>
<td>Mansa</td>
<td>186,506</td>
<td>15,537</td>
</tr>
<tr>
<td>Mkushi</td>
<td>109,546</td>
<td>16,143</td>
</tr>
<tr>
<td>Mwinilunga</td>
<td>131,515</td>
<td>23,138</td>
</tr>
<tr>
<td>Zambezi</td>
<td>66,694</td>
<td>12,483</td>
</tr>
<tr>
<td>Livingstone</td>
<td>158,149</td>
<td>32,856</td>
</tr>
</tbody>
</table>

Table 4.1: Population distribution for the case study regions. Source: Census 2000 report, CSO Zambia.

This chapter proceeds as follows; we first discuss our approach towards any existing infrastructure, we then describe the system model and characterize the workload to which the system model is subjected and finally, design the M-QNA model which describes the communication system being modeled in detail.

4.1 Existing Infrastructure

There is no infrastructure that provides services such as video teleconferencing in these regions. However, there is limited infrastructure for voice and data services. Hence, for the purpose of determining the most cost-effective way of providing communication with integrated services to geographically hostile regions, we neglect any existing infrastructure.

4.2 The System Model

4.2.1 Regional Connectivity Pattern

The regions selected for this study can be interconnected in many ways. These ways vary according to the number of links required in the network. From Chapter 1, we learned that the number of links to fully connect $N$ regions is given by

$$C = \binom{N}{2}$$
4.2.1 Regional Connectivity Pattern

For the nine regions, this translates to a total of 36 possible links. We note that these regions are defined as being sparsely populated, widespread with less economic activity and therefore establishing a fully connected topological design with 36 links would be expensive. Thus, we consider a least connected topological design which gives rise to the connectivity pattern presented in Figure 4.1.

Figure 4.1: Regions covered in the case study

Each region connects to the adjacent region through its regional router. The router in Lusaka was modelled as a gateway to link other regions to the external networks. The links \( l_{ij} \) between these routers are assumed to be reliable with capacities \( K_{ij} \) which depend on the maximum load that the link is required to carry. The inter-regional distance indicated in the figure is relevant for the purpose of calculating the cost of terrestrial systems and the link propagation delay for Microwave systems in particular.

The propagation delay is defined as the average time a customer spends on the communication link and does not include the time it spends in the node. This delay is very much dependent on the type of technology used and is analysed in Chapter 2 for each technology considered for inter-regional connectivity. The link capacities also vary with different technologies. However, these
capacities may still differ even with the same technology and hence depended on the maximum load required on a particular link.

The system parameters are tabulated in Table 4.2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Denotes the source population from which external arrivals of customers originate. It is assumed to be infinite.</td>
</tr>
<tr>
<td>Maximum number of regions $N$</td>
<td>$N$ represents the maximum such that $1 \leq n \leq N$ where $n$ is the scaling factor for the regions and each region is assumed to have a finite number of sites.</td>
</tr>
<tr>
<td>Processing time of a router</td>
<td>This is the average time a receiver takes to process or offer service to an arriving customer. This time is assumed to be independent identically distributed.</td>
</tr>
<tr>
<td>Propagation delay</td>
<td>This parameter gives the average delay of a customer from ENTRY to its EXIT from the system. It is technology dependent.</td>
</tr>
<tr>
<td>Link cost</td>
<td>Represents the cost of the system components required to establish communication and is dependent on the technology used.</td>
</tr>
<tr>
<td>Link capacity</td>
<td>Represents the maximum bit rate on a communication link and is dependent on the maximum required load and the technology used.</td>
</tr>
</tbody>
</table>

Table 4.2: Summary of System Parameters

The possible number of scenarios that can be used to investigate the cost of the system models developed in this section are presented in Section 4.2.2.

4.2.2 Possible Scenarios

In order to determine the effectiveness of the communication systems considered in Chapter 1, we consider three scenarios to interconnect the regions in Figure 4.1 based on the recommendations made in Chapter 2. We use these scenarios to investigate the cost and performance of the communication systems with the objective of determining the most cost-effective scenario that can serve as a solution for geographically hostile regions. These scenarios are indicated in Table 4.3.

As indicated in Table 4.3, only three distinct inter-regional links are possible. Since the performance of the communication system depends on the load that these inter-regional links are able to carry, the three scenarios shown in the table are sufficient to assess the performance of
Table 4.3: Hybrid Scenarios

<table>
<thead>
<tr>
<th>Inter-regional link technologies</th>
<th>Intra-regional link technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Microwave</td>
<td>Radio</td>
</tr>
<tr>
<td>2. Satellite</td>
<td>Radio</td>
</tr>
<tr>
<td>3. Optic Fibre</td>
<td>Radio</td>
</tr>
</tbody>
</table>

each scenario. In each scenario, the intra-regional link technologies are only relevant for costing purposes. These three scenarios are illustrated in Figures 4.2, 4.3 and 4.4.

4.2.2.1 Scenario 1

This scenario considers a pattern in which microwave technology is used for the network backbone.
4.2.2 Possible Scenarios

4.2.2.2 Scenario 2

In this scenario, satellite technology is used for the network backbone.

4.2.2.3 Scenario 3

This scenario uses optic fibre technology for the network backbone.

We now describe the workload model that is subjected to the scenarios designed in Section 4.2.2.
4.2.3 The Workload Model

This section is aimed at characterizing the workload that is subjected to the scenarios designed in Section 4.2.2. This workload is defined as the inputs that the system receives from its environment during any given period of time. These inputs, which we shall refer to as external arrivals, enter the communication system through the gateway. They are assumed to originate from an infinite source population and follow the Poisson arrival process described in Chapter 3. In the analytic model, the workload is defined by means of the statement below:

\[
\text{wrkld stmt} \rightarrow \text{‘workload’ label [class clause] (arrival clause)[11]};
\]

In the statement, ‘workload’ is a keyword and label defines the name(s) of the workload. The class clause identifies and names the classes which belong to the particular workload. Each name must be unique within a workload, but classes in different workloads may have the same name. In this latter case, the class label must be qualified by the relevant workload name whenever referred to.

The arrival clause specifies the average arrival rate at which customers arrive from an infinite external source to a single entry point in the communication network. For the purpose of parametric analysis, an index is used so that the arrival rate for each workload is modified for every value of the index within a loop. This is done using the ‘modify workload’ [11] statement as in Table 4.4. Note that the initialisation statements for the ‘index’ and ‘randomnumber’ variables are not included.

```
LOOP loop = 1 TO X
MODIFY CENTRE centrename Service Discipline
    time one:voice \mu_1^{-1}
time two:video \mu_2^{-1}
time three:data \mu_3^{-1};
MODIFY WORKLOAD one Arrival ar_voice*index;
MODIFY WORKLOAD two Arrival ar_video*index;
MODIFY WORKLOAD three Arrival ar_data*index;
LET index = index + randomnumber;
ENDLOOP;
```

Table 4.4: A loop to modify the traffic input patterns for each workload

In the loop, `centrename` defines the centre at which arrivals take place. At this centre, a specific service discipline is assigned. As discussed in Chapter 3, this is either PS, FCFS or IS. The time
allocated for each workload class, i.e. one: voice, two: video and three: data, is almost general.

These arrivals are then routed through the network as customers destined for Region $j$. At Region $j$, there are several sites to which the customers may be destined. The number of sites in Regions $i$ and $j$ may differ according to the economic activity in each region. For the purpose of traffic generation in each scenario, the number of sites in each region are assumed to be the same. The uni-directional traffic flow describing the connectivity of Figure 4.1 without considering the links at this point is illustrated in Figure 4.5.

$$\lambda_c = \lambda_{\text{voice}} + \lambda_{\text{video}} + \lambda_{\text{data}}$$

![Diagram of network connectivity](image)

1 $\rightarrow$ From Hub (Earth Station)  
2 $\rightarrow$ To Hub (Earth Station)  
3 $\rightarrow$ Departures

Figure 4.5: Uni-directional traffic flow between the regions

In Figure 4.5, the external arrivals are seen to enter the system through the gateway in Lusaka at the rate $\lambda_c$, where $c$ is a positive number representing a customer class. The maximum number
of customer classes \( C \) in this design is 3. In most network models such as those studied by Kleinrock, Woodward and Gorney [20, 42, 9], the networks considered are those where the maximum number of customer classes \( C \) is equal to 1.

The lines labeled 1 and 2 in the figure are used to model the links to and from the gateway other than the adjacent regions. These links are considered when modelling the satellite communication system. They are not considered when modelling terrestrial systems for the reasons given in Section 4.2.1.

We note that Figure 4.5 does not explicitly model the behaviour of the customers on the link. To model this part of the communication system, we consider the gateway and the adjacent router as shown in Section 4.2.3.1.

### 4.2.3.1 Customer behaviour on the inter-regional link

This section models the behaviour of the customer on the communication link between two adjacent regions. The reason for modelling the behaviour of the customers on the link is to enable us characterize the behaviour of the workload classes explicitly. As shown in Figure 4.6, external arrivals enter the communication system through the gateway \( i \) at the mean rate equal to \( \lambda_c \). A fraction of these arrivals leave the system to the LANs within the region. The remaining customers are then routed on the link \( l_{ij} \) to the adjacent regional router \( j \).

![Diagram of traffic flow between two adjacent regions](image)

**Figure 4.6:** The traffic flow between two adjacent regions

As explained in Chapter 3, the time-out server on link \( l_{ij} \) models the mechanism for retransmission of customers that are in error. The probability of errors on such a link is denoted by \( e_{ij} \).
4.2.3 The Workload Model

The value of $e_{ij}$ depends on the communication technology used. In Chapter 2, we learned that this value is $10^{-6}$ for microwave links, $10^{-7}$ for satellite links and $10^{-12}$ for optic fibre links. These values have been used in the model presented in Section 4.2.4 to enable the computation of the probability of retransmission for customers that belong to the data class. This is because the recipients of the voice and video messages are assumed to be intelligent enough to correctly interpret the meaning of messages even when errors occurred. Furthermore, voice and video are delay sensitive and thus cannot be retransmitted. Hence, these customers are not delayed at this server. Modelling each interconnection between regions as illustrated in Figure 4.6, a full M-QNA model shown in Section 4.2.4 is derived.

4.2.3.2 The cost and performance models

The cost variable denoted by $D$, represents the cost of establishing all the links in the communication network. For each link $l_{ij}$, its cost is denoted by $d_{ij}$. In this study, $d_{ij}$ comprises the cost of the physical link, repeaters, and the terminal equipment at the source. Therefore, summing up these costs for all the links in the network, Equation 4.1 is derived.

$$D = \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij}$$  (4.1)

To compute this equation, the following model is used:

Given: Network topology

Determine: $D = \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij}$

With respect to: capacity $K$ and flow $\{r_{ij}\}$

This model is applied to each of the scenarios designed in Section 4.2.2.

The second model determines the throughput at which the average network delay is equal to the maximum permissible average network delay. In this model, a three-dimensional performance criterion which includes the mean network delay $T$, the throughput and utilization are used. Using these variables, we attempt to determine the maximum throughput in each scenario such that $T = T_{max}$, where $T_{max}$ has been defined to be 400 msec. This model is structured as follows:

Given: $\{K_{ij}\}$ and Network topology
4.2.3 The Workload Model

Determine: \( \{\tau_{ij}\} \).
such that: \( T \leq T_{MAX} \) or \( T = T_{\text{threshold}} \).

The only constraint considered here is the maximum permissible mean network delay and the threshold value of \( T \) beyond which the system becomes unstable. Thus, the most cost-effective scenario is obtained by varying the rate of flow \( \{\tau_{ij}\} \) such that the mean network delay is either equal to \( T_{MAX} \) or \( T_{\text{threshold}} \) if \( T_{\text{threshold}} \leq T_{MAX} \). Using the value of \( \tau_{ij} \) and \( D \), the cost-effectiveness ratio is calculated. The lowest ratio provides the best cost-effective scenario for geographically hostile regions.

In this second model, the mean network delay \( T \) is obtained by varying the rate of customer flow \( \{\tau_{ij}\} \) on the link \( l_{ij} \). The experimental results for these models are analysed in Chapter 5.

Another key feature in characterizing the workload is the routing pattern subjected to the customers. In this study, a non-adaptive routing pattern is adopted and is discussed in Section 4.2.3.3.

4.2.3.3 Routing of customers

The distribution of customers in the network is based on the routing of class \( c \) customers that arrive at the router and is specified by a transition probability matrix,

\[
P = \{p_{ic,jc} \mid 1 \leq i \leq N, \ 1 \leq j \leq N, \ 1 \leq c \leq C \}.
\]

A class \( c \) customer currently at router \( i \) is next routed to router \( j \) with probability \( p_{ic,jc} \). Therefore,

\[
\sum_{j=1}^{N} p_{ic,jc} = 1. \tag{4.2}
\]

for a particular router \( i \) and customer class \( c \). Since the customers on the network do not change their initial routes to accommodate dynamic load balancing needs, the routing of these customers in this model is thus non-adaptive and is specified by a set of matrices

\[
P = \{p_{ic,jc} \}, \quad 1 \leq c \leq C. \tag{4.3}
\]

The routing probability matrix which is shown in in Table 4.5 was estimated based on the knowledge of potential subscribers in each region as shown in Table 4.1. We defined a potential subscriber as a household in a region that requires the services defined in Chapter 1. Since
all households in a region have equal opportunity to use these services, we considered all the households in each region.

<table>
<thead>
<tr>
<th>From</th>
<th>LUS</th>
<th>SIA</th>
<th>LUN</th>
<th>MBA</th>
<th>MAN</th>
<th>MKU</th>
<th>MWI</th>
<th>ZAM</th>
<th>LIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUS</td>
<td>-</td>
<td>0.24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.76</td>
</tr>
<tr>
<td>SIA</td>
<td>0.81</td>
<td>-</td>
<td>0.19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LUN</td>
<td>0</td>
<td>0.26</td>
<td>-</td>
<td>0.74</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MBA</td>
<td>0</td>
<td>0</td>
<td>0.89</td>
<td>-</td>
<td>0.11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MAN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.66</td>
<td>-</td>
<td>0.34</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MKU</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.19</td>
<td>-</td>
<td>0.81</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MWI</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.56</td>
<td>-</td>
<td>0.44</td>
</tr>
<tr>
<td>ZAM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.41</td>
<td>-</td>
</tr>
<tr>
<td>LIV</td>
<td>0.94</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Table 4.5: The Routing Probability Matrix

Table 4.5 shows the probability routing matrix of the regions shown in Figure 4.1. In this table, a region can only route its traffic and the traffic from other regions to its adjacent regional routers.

4.2.4 The Machine Model

Figure 4.7 models the connectivity and customer flow patterns in the terrestrial systems identified in Section 4.2.2. For the system which involves the satellite links, the model is modified in such a manner that there is only one link from the gateway to the satellite and eight links from the satellite to the regional routers. Each router is modelled as a processor sharing server and all the customer classes \( c \) have different mean service times as shown in Section 1.3. The output links are modelled by FCFS queues with service rate \( \mu \) where \( \frac{1}{\mu} \) is the average service time.

Since the voice and video classes are delay sensitive but not error sensitive and the recipient of the messages is assumed to be intelligent enough to correctly interpret any errors in the message,
they are modelled without acknowledgements and time-outs. However, the customers that belong to the data class experience time-outs and re-transmissions as a result of their sensitivity to errors as earlier discussed in Chapter 1. Therefore, as a data class customer moves through the network from its source router to its destination router, each intermediate router stores a copy until it is positively acknowledged by the succeeding router. If no positive acknowledgement is received, the copy of the customer is re-transmitted. This is modelled by the time-out server which is more elaborately shown in Figure 4.6.

To model the re-transmissions, we make reference to Figure 4.6. A fraction $e_{ij}$ of data class customers leaving on link $l_{ij}$ are assumed to change class to class $z$ (where $z \neq c$ for all $c \in \{c\}$) and are delayed in the time-out server for the duration equal to the time-out specified period. Beyond that period, they are put back on the queue of the outgoing link. The remaining $1 - e_{ij}$ customers experience no errors on the link and are not delayed. As stated in Section 4.2.3, this mechanism is modelled by an Infinite Server queue. The service rate for class $z$ customers is inversely proportional to the time-out period and a service rate for class $c$ customers is infinite to ensure that these customers are not delayed any further.

The assumption taken into account to solve the M-QNA model of Figure 4.2.4 include:

- External arrivals follow a Poisson distribution as earlier mentioned in Section 4.2.3.
- The central processor of the router, gateway and the switch at the satellite are processor sharing with class dependent independent identically distributed service times.
- The outgoing link $l_{ij}$ is a FCFS queue with a service time that is equal to the link propagation delay.
- The blocking probability was assumed to be zero as they are no loses in M-QN system.

The service time distributions for each customer class, their routing patterns (defined in Section 4.2.3.3) and the arrival pattern are used as input to the M-QNA model presented in Figure 4.7.

To solve the analytic model in Figure 4.2.4, we used a programming package called Stochastic Network Analysis Programming Language (Micro SNAPL) [11]. Thus, the use of Micro SNAPL enables the analysis of different characteristics of a communication network which are relevant for the purpose of network planning and decision making. The model results are discussed in Chapter 5.
Figure 4.7: The M-QNA model
Chapter 5

Results

The objective of this chapter is to determine the most cost-effective communication system for geographically hostile regions based on the experimental results and the scenario costs.

For each experiment, the input traffic pattern at the gateway depended on how much load the communication network could carry. This input traffic pattern was chosen to fit the Poisson arrival process and differed for each scenario designed in Section 4.2.2.

In each of these experiments, our interest is in the trading relations among the mean delay \( T \) defined in Equation 3.3, the throughput \( \tau \) defined in Equation 3.2, the utilization \( \rho \) defined in Equation 3.4 and the system capacity \( K \). The relationship among these performance variables is significantly affected by the system design. In Chapter 4, three different system scenarios are presented as Scenario 1, 2 and 3, all with different parameter values. Furthermore, the configuration of Scenario 2 differs from that of the other scenarios in that its inter-regional links are non-terrestrial and form a star topology as shown in Figure 2.3. These differences between the scenarios are fundamental in determining the trading relations.

Using the trading relations among the performance variables mentioned above, we seek to minimize the mean delay \( T \) and determine the throughput at which \( T \leq T_{\text{MAX}} \) or \( T = T_{\text{threshold}} \) if \( T_{\text{threshold}} \leq T_{\text{MAX}} \), where \( T_{\text{MAX}} \) is taken to be 400 msec as recommended by ITU [32] for real time applications. To do this, we have an option to change the system parameters, particularly the service time distribution at the regional routers since they are assumed to be arbitrary.

The mean delay \( T \) can also be reduced if we reduce the utilization \( \rho \) by either increasing the system capacity \( K \) or reducing the system throughput \( \tau \). This approach is not satisfactory since we are paying the price of increased system cost (more capacity \( K \)) or of reduced throughput in order to reduce the mean delay \( T \).
Another trade-off is that of attempting to increase the throughput of the system. If we simply scale up the throughput $r$, $T$ will also increase. However, we can attempt to reduce the rate at which $T$ grows by increasing both $r$ and $\rho$ at a slow rate. We now investigate this trade-off with an objective of determining the most cost-effective scenario. In addition, we also investigate the effect of scaling up throughput on the link errors.

### 5.1 Scenario 1

This experiment was based on scenario 1 (see Figure 4.2) which depicts the communication system as a collection of Microwave inter-regional links and Radio intra-regional communication links.

#### 5.1.1 Performance

The propagation delay for the inter-regional links used in this experiment was calculated to be $3.3 \times 10^{-3}$ msec per Km by applying the basic formula relating time, distance and the speed of light. In this experiment, the M-QNA model took 1300 msec to solve on a Sun SPARCstation 5 with a cpu speed of 170 MHz.

The mean delay $T$ for the three customer classes was determined as a function of the link utilization. The objective for doing so was to determine whether the system can guarantee the stringent QoS requirements for real-time customers under a given load.

- The link from Lundazi to Mbala could not cope with traffic at 98% utilization under the assumed input traffic pattern
- The link capacities were assumed to be E1 full duplex
5.1.1 Performance

Figure 5.1: Mean Delay Vs Utilization of scenario 1

Figure 5.1 shows that Scenario 1, depicted in Figure 4.2, achieves the objective of minimizing the mean delay $T$ while at the same time ensuring high utilization of the system. This is clearly seen in the figure where the mean delay is below 104 msec at 89% utilization where the system reaches its threshold. Since the maximum tolerable mean delay for the real time customers is 400 msec, these results show that a microwave - radio hybrid system is efficient and able to provide guaranteed QoS levels for the end Users described in Table 1.1.

Figure 5.2: Mean Delay Vs Permissible throughput of scenario 1

Using the stated mean customer length and propagation delay, the system throughput was seen
to reach a maximum of 225 packets per second. Furthermore, we note that the mean delay $T$ grows at a fairly slow rate as the throughput $\tau$ increases. However, at the rate slightly above 200 packets per second, the mean delay $T$ is seen to exhibit a threshold behaviour. Beyond the threshold, $T$ is seen to suddenly grow without bound. This was shown by an infinite queue on the link from Lundazi to Mbala in the analytic model. Apparently, this link is the longest of all the links in the network. Hence, we can further deduce that $T$ is a strictly increasing function of the average link length. The longer the link, the quicker the value of $T$ is seen to grow and this puts a limit on the amount of throughput permitted. On the contrary, shorter links were seen to permit higher throughputs with the value of $T$ growing fairly slowly. These results show that the distribution of capacity on the links in the network should not be uniform, longer links should be given less capacity than shorter links. This suggests that the topological structure of the network should perhaps be chosen to yield a minimum average link length. Although this is achievable with the use of a highly connected network topology, the cost involved is prohibitive as explained in Chapter 1.

From Chapter 3, we know that $T$ exhibits a threshold behaviour at a point where the system throughput approaches the capacity of the system. In this experiment, this is exhibited at a throughput $\tau^*$ which is slightly above 200 packets per second. The throughput $\tau^*$ is therefore, the network saturation point. Beyond this saturation point, the link mentioned above could not cope with the amount of throughput as $T \to \infty$ (becomes unbounded) and thus creating a network bottleneck. This bottleneck would, however, be overcome by increasing the capacity $K$ on this link. Based on these results, we can further say that the most effective design in a scenario such as Scenario 1 is one which considers different capacities $K$ on each link, a feature not considered in this study.

We now investigate the effect of link errors which result in retransmissions for the customers that belong to the data class on the performance of the network. As discussed in Chapter 2, microwave systems have a bit error rate $e$ of $10^{-6}$. From Chapter 3, we know that in every block of $n$ bits, the probability that a bit is in error in the absence of error clustering is given by $1 - (1 - e)^n$. Applying this to the M-QNA model, the results in Figure 5.3 are obtained.
Figure 5.3 indicates that as the rate of retransmission on the link increases, the performance of the link deteriorates causing the average network delay to increase. This means that the customers being retransmitted will continue to congest the network. As a result, new customers are forced to wait in a queue. As the rate of retransmission increases, the queue length also increases tending to infinity. At a retry rate slightly above 3.5%, the Lundazi - Mbala link is seen to deteriorate at a faster rate causing the link to be put out of service. Hence, the traffic to Mbala would have to be routed through Mansa.

5.1.2 Cost

As indicated in Chapter 2, the major components that make up the cost of a microwave system are the digital modem, the RF unit and the antenna. Since the design of the network requires that hops be established in two directions from each region, two directional antennas will be required at each region. Furthermore, the long distances involved dictate the use of repeaters between a pair of regions. Since, in general, microwave systems can transmit up to 50 km, several repeaters need to be installed. Applying this to Figure 4.1, we find that 128 repeaters are required to ensure full connectivity. This, therefore, means that the highest cost of this system is made up of the cost of the repeaters. Table 5.1 shows the cost of the standard terminals (Radio unit), repeaters, installation and maintenance.
5.2 Scenario 2

<table>
<thead>
<tr>
<th>Description</th>
<th>Unit cost USD</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microwave Radio unit</td>
<td>28,820</td>
<td>259,380</td>
</tr>
<tr>
<td>Microwave repeater</td>
<td>25,000</td>
<td>3,200,000</td>
</tr>
<tr>
<td>Multiplexer</td>
<td>2,320</td>
<td>20,880</td>
</tr>
<tr>
<td>Installation (inclusive of repeaters)</td>
<td>1,599</td>
<td>219,063</td>
</tr>
<tr>
<td>Maintenance</td>
<td>3,000</td>
<td>27,000</td>
</tr>
<tr>
<td><strong>Total cost for 9 regions</strong></td>
<td></td>
<td><strong>3,726,323</strong></td>
</tr>
</tbody>
</table>

Table 5.1: Inter-regional connectivity costs for configuration 1 - Source: See Chapter 2

While the cost of the equipment, repeater and installation is fixed, the cost of maintenance is incremental and was calculated over a period of 1 year. This was done simply because budgets are reviewed annually, after which, changes in the cost structure may be possible.

Applying the cost equation in Chapter 4, we note that the average cost is merely the total system cost divided by the number of regions. Thus, we formulate the average cost equation shown in Equation 5.1.

\[
Av.\text{Cost} = \frac{1}{N} \sum_{ij} d_{ij}
\]  

For this particular scenario, the average cost is calculated to be USD **414,035.89**

5.2 Scenario 2

The second experiment was based on scenario 2 (see Figure 4.3) and depicts the communication system as a collection of satellite inter-regional links and Radio intra-regional communication links as shown in Figure 4.3.

5.2.1 Performance

For the satellite links, as discussed in Chapter 2, the distance between the regions has no effect on the performance of the communication system and thus was not considered. The roundtrip propagation delay for the satellite link was calculated to be 240 msec. The service time distribution was kept the same as in Experiment 1.

---

1Digital modem, 2*Radio Frequency unit, and 2* Antennas
5.2.1 Performance

Figure 5.4: Mean Delay Vs Utilization of scenario 2

- The link from Lusaka to the satellite could not cope with the traffic above 90% load under the assumed input traffic pattern
- The link capacities were assumed to be E1 Kbps full duplex

In this experiment, the system was seen to exhibit a threshold behaviour at 80% utilization giving a throughput of 6.7 packets per second. Although the system is stable at this utilization, the inherent sensitivity to delay by voice and video customers prohibits its operations. The experimental results show that the mean delay in this scenario is much higher than that experienced in Scenario 1 typically because of the high propagation delay suffered by customers as they traverse the link. The results show that this scenario is able to support real-time customers at a utilization with an upper bound of 50% which corresponds to 4.3 packets per second. If the system is allowed to operate at any higher load, above 50%, the mean delay becomes too high to permit the usefulness of the real time messages.
5.2.1 Performance

Figure 5.5: Mean Delay Vs Permissible throughput of scenario 2

It can also be seen that the throughput permitted by this scenario is also much lower than that permitted by scenario 1.

Like in Scenario 1, we investigate the effect of an error on the link in this scenario. Figure 5.6 shows that the probability of retransmission in this scenario is much lower than that of Scenario 1. The mean delay shown in the figure represents the average network delay suffered by customers belonging to the data class.

Figure 5.6: The effect of an error on the inter-regional link in Figure 4.3
At a retry rate of $10 \times 10^{-5}$, the mean delay $T$ is seen to reach its threshold. Although the number of class $z$ customers at this retry rate is marginally small, the net effect on the average network delay is substantial.

### 5.2.2 Cost

Table 5.2 presents the costs of the inter-regional link presented in Scenario 2 Chapter 4, Section 4.2.2. It is modelled as a link in a VSAT configuration described in Chapter 2, Section 2.3.1. We opted to calculate these costs over a period of 1 year simply because most budgets are reviewed annually, after which, changes in the cost structure may be administered. Within that period, we do not anticipate major changes in the cost structure.

<table>
<thead>
<tr>
<th>Description</th>
<th>Unit Cost (USD)</th>
<th>Sub total (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.8m C-Band Rx/Tx antenna (gateway)</td>
<td>32,000</td>
<td>32,000</td>
</tr>
<tr>
<td>3.8m C-Band Rx/Tx antenna</td>
<td>15,800</td>
<td>126,400</td>
</tr>
<tr>
<td>10W C-Band Transciever</td>
<td>12,300</td>
<td>110,700</td>
</tr>
<tr>
<td>Satellite Modem</td>
<td>1,900</td>
<td>17,100</td>
</tr>
<tr>
<td>Multiplexer</td>
<td>2,320</td>
<td>20,880</td>
</tr>
<tr>
<td>Installation</td>
<td>3,500</td>
<td>31,500.00</td>
</tr>
<tr>
<td>Maintenance</td>
<td>2,000</td>
<td>18,000.00</td>
</tr>
<tr>
<td>Link lease (E1)-1 year</td>
<td>404,628</td>
<td>3,641,652.00</td>
</tr>
<tr>
<td><strong>Total Cost for 9 regions</strong></td>
<td></td>
<td><strong>3,998,232.00</strong></td>
</tr>
</tbody>
</table>

Table 5.2: Inter-regional connectivity costs for scenario 2 using a leasing option: Source: See Chapter 2

Using the cost equation in Chapter 4, we note that the average cost is merely the total system cost divided by the number of regions. Thus, we use the average cost equation shown in Equation 5.1 above.

For this scenario, the average cost is calculated to be **USD 444,248**.

### 5.3 Scenario 3

The third experiment was based on Scenario 3 (see Figure 4.4) which depicts the communication system as a collection of Optic Fibre inter-regional links and Radio intra-regional communication links shown in Figure 4.4.
5.3.1 Performance

The propagation delay for the inter-regional link, in the absence of repeaters, was assumed to be negligible (about $4.7 \times 10^{-6}$ msec per Km [36]) and hence, the mean delay in the system merely results from the switching and queueing delay at the routers. Therefore, the utilization of the system, Equation 3.4, is only limited by the utilization of the gateway and the routers and not by the link as seen from the experimental results.

![Figure 5.7: Mean Delay Vs Utilization of scenario 3](image)

- The gateway at Lusaka could not cope with the traffic above 70% load under the assumed input traffic pattern
- The link capacities were assumed to be E1 Kbps full duplex

We know that the fibre link is capable of providing very high bit rates which are only limited by the switching equipment. However, since VSAT terminals can only transmit up to E1 speeds, we opted to use E1 speeds for this scenario in order to perform comparisons at the same level. In this experiment, the M-QNA model took 1100 msec to solve on a Sun SPARCstation 5 with a cpu speed of 170 MHz. The experimental results shown in Figure 5.7 represent the performance of the gateway and not that of the link. This is because, as the level of input traffic was increased, an infinite queue was experienced at the gateway in Lusaka implying that
5.3.1 Performance

the maximum throughput $\tau$ the gateway can permit had been exceeded. Therefore, the major bottleneck in this scenario is the processor at the gateway.

![Graph showing mean delay vs permissible throughput of scenario 3](image)

Figure 5.8: Mean Delay Vs Permissible Throughput of scenario 3

Using the mean customer length and service time distribution at the service centers, the system throughput was seen to reach a maximum of 570 packets per second. Even after the maximum permissible throughput had been reached, we note that the network is still able to meet its design goal of less than 400 msec mean delay. Thus, as desired, the communication network is essentially transparent to the user, as far as delay is concerned. We now investigate the effect of an error on the link.

![Graph showing the effect of an error on the inter-regional link](image)

Figure 5.9: The effect of an error on the inter-regional link in Figure 4.4
Figure 5.9 shows the effect of an error on the link in this scenario. Based on this figure, we can see that this scenario has the lowest probability of retransmission which agrees with the theoretical notion that an Optic Fibre link has the lowest bit error rates.

## 5.3.2 Cost

Table 5.3 presents the costs of the inter-regional link presented in Scenario 3 Chapter 4, Section 4.2.2. It is modelled as a single mode Optic Fibre link described in Chapter 2, Section 2.2.

<table>
<thead>
<tr>
<th>Description</th>
<th>Unit cost (USD)</th>
<th>Sub total (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single mode Fibre (4 strands)</td>
<td>1.36/meter</td>
<td>8,684,960</td>
</tr>
<tr>
<td>Multiplexer</td>
<td>3,758</td>
<td>33,822</td>
</tr>
<tr>
<td>Codec</td>
<td>13,000</td>
<td>117,000</td>
</tr>
<tr>
<td>E1 Fibre Modem</td>
<td>925</td>
<td>8,325</td>
</tr>
<tr>
<td>Optical Transmitter (E1)</td>
<td>1,400</td>
<td>12,600</td>
</tr>
<tr>
<td>Optical receiver</td>
<td>1,450</td>
<td>13,050</td>
</tr>
<tr>
<td>Repeater</td>
<td>213.97</td>
<td>27,388</td>
</tr>
<tr>
<td>Copper to Fibre converter</td>
<td>173.00</td>
<td>1,557</td>
</tr>
<tr>
<td>Installation</td>
<td>0.49/meter</td>
<td>3,129,140</td>
</tr>
<tr>
<td>Maintenance</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total cost for 9 regions</strong></td>
<td></td>
<td><strong>12,027,842</strong></td>
</tr>
</tbody>
</table>

Table 5.3: Inter-regional connectivity costs for scenario 3 - Source: see Chapter 2

We can clearly see from Table 5.3 that the biggest cost of a fibre system results from the installation and cost of the fibre itself. Like in the case for Microwave systems, signals in Optic Fibre systems have to be boosted every after 50 Km which also translates to 128 repeaters in the system. Unlike Microwave and Satellite systems that require frequent maintenance, the only maintenance work that can be done on an Optic Fibre system is to replace the obsolete parts or if an upgrade is required. Such costs are difficult to determine since it takes a long time before an upgrade is made or parts become obsolete and hence, we assume that the maintenance costs are negligible over short periods of time, such as one year.

We use the average cost equation shown in Equation 5.1 above. Subjecting the costs in this scenario to the average cost equation, the average cost is calculated to be **USD 1,336,426.89**
5.4 Summary of results

We now seek to determine the most cost effective scenario in terms of its cost-effectiveness (CE) ratio. Comparing the costs in each scenario, we note that the scenario with optic fibre backbone links has the highest system costs, followed by satellite and microwave being the least. On the other hand, at $T^*$, the optic fibre backbone link is seen to offer the highest throughput of 593,598 bps, followed by 208,280 bps for a Microwave backbone link and 4,478.02 bps for the satellite backbone link. The bps values were calculated by taking into account the average packet sizes and the proportions of each class of customer in the network as discussed in Chapter 1 Section 1.3.

Based on the cost and the throughput as a measure of performance, the cost-effectiveness ratios, derived by dividing cost by throughput, were calculated to be 17.9 for Scenario 1, 892.9 for Scenario 2 and 20.3 for Scenario 3. These results are summarised in Table 5.4.

<table>
<thead>
<tr>
<th>Performance Variable</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost USD</td>
<td>3,726,323</td>
<td>3,998,232</td>
<td>12,027,842</td>
</tr>
<tr>
<td>Throughput $T^*$</td>
<td>208,280 bps</td>
<td>4,478.02 bps</td>
<td>593,598 bps</td>
</tr>
<tr>
<td>CE Ratio</td>
<td>17.9</td>
<td>892.9</td>
<td>20.3</td>
</tr>
</tbody>
</table>

Table 5.4: Summary of Results

Comparing the costs in each scenario, we note that Scenario 1, comprising microwave inter-regional links, has the lowest system costs followed by Scenario 2, which comprises satellite inter-regional links. Scenario 3, which comprises optic fibre inter-regional links has the highest initial costs for a period of 1 year. However, the inter-regional link costs in Scenario 2 are incremental and hence, as the period exceeds one year, this system becomes more expensive.

5.5 Conclusion

From these results, we can therefore conclude that Scenario 1, comprising microwave inter-regional links and radio intra-regional links, offers the best cost-effectiveness ratio rendering it a more cost-effective solution for geographically hostile regions. However, we wish to state that this scenario has the highest probability of retransmission. Since the amount of traffic in geographically hostile regions is low, the impact of these retransmissions on the network
performance is expected to be low.
Chapter 6

Conclusion

The key issues in the design of any communication network that require much attention include the network topology, the required capacity on the links, the average network delay and the cost of the communication network.

The decision to use the least connected topology in this study was motivated by two things; firstly, the regions considered for the case study are wide spread and characterized by less traffic due to sparse population and less economic activity. Secondly, this study involves three different inter-regional technologies which are implemented differently. Although, a least connected topology may require more capacity than a relatively highly connected topology, the cost of a media, such as a single mode fibre, is irrespective of the capacity. From Chapter 5, we found that the optic fibre media has the highest cost and hence, a relatively highly connected topology would be costly for a network whose demand is in the E1 range. Furthermore, for a system involving microwave links, a highly connected topology would imply a corresponding increase in the repeaters which would effectively increase the system costs. We also know, from Chapter 2, that a VSAT system has a fixed topology and therefore, it is not realistic to consider a highly connected topology in this case. Based on these findings, we thus, conclude that a least connected topology is more cost-effective for geographically hostile regions when the backbone inter-regional technology involved is either microwave or optic fibre.

The results in Chapter 5 clearly show that the most efficient design in any communication network is one that considers different capacities on each communication link. In the scenario with a microwave inter-regional link, we saw that the utilization approached unit on the longest link more quickly than it did on the other links. This is basically because the average network delay \( T \) is a strictly increasing function of the average link length. Thus, as the delay increases,
more customers are forced to queue which leads to an increase in the network load. For this reason, we saw that the satellite link had the least amount of throughput and the highest average network delay. Therefore, for microwave systems, shorter links are capable of carrying more network load than longer links and as such more capacity should be allocated.

As stated in Chapter 1, the key objective of this study is to determine the most cost-effective way of providing connectivity to geographically hostile regions. From Chapter 5, we note that Scenario 1 involving microwave inter-regional links is the cheapest (USD 3,726,323.00) followed by Scenario 2 (USD 3,998,232) involving satellite inter-regional links (VSATs). Scenario 3, which involves optic fibre links, is the most expensive (USD 12,008,123) for the period considered. However, we wish to note that the major cost of a system involving satellite inter-regional links is the cost of leasing the satellite space segment which is charged on a monthly or yearly basis. This cost is incremental and thus, is likely to exceed the cost of Scenarios 3 which is characterised by higher initial costs.

On the other hand, Scenario 3 is seen to exhibit the highest amount of throughput (593,598 bps) followed by Scenario 1 (208,280 bps) and Scenario 2 (4,478.02 bps) being the least. Based on these results, the cost-effectiveness ratio for each scenario was calculated with Scenario 1 achieving 17.9, Scenario 2 achieving 892.9 and Scenario 3 achieving 20.3. These results, summarised in Table 5.4, clearly show that a communication network that involves microwave backbone links and radio intra-regional links is cost-effective for geographically hostile regions. Thus, this solution can be used to provide communication services countrywide at minimum costs. Due to this minimum cost, the ultimate goal of ensuring lower end-user service rates, a very important feature for developing regions, can be achieved. However, we wish to state that this scenario has the highest probability of retransmission. Since the amount of traffic in developing regions is low, the impact of these retransmissions on the network performance is expected to be low.

This study set out to investigate cost effective communication alternatives, and the models which have been developed have provided insight into the costs and capabilities of three scenarios in a way which allows comparison. It is contended that the results and methodology presented offer insight and benefit to telecommunications operators who are considering establishing services in geographically hostile regions.
6.1 Future Work

The results of this study, presented in Chapter 5, were derived using analytic techniques. It would, therefore, be useful to tackle the problem presented in Chapter 1 using simulation techniques and compare how close the results are. These results would provide the relevant insight details to telecommunications operators who are in the process of implementing communication systems for geographically hostile regions.

Furthermore, this study assumed uniform link capacities across the network. It would be useful to perform a further study that determines the capacity of each link independently. As noted in Chapter 5, the mean network delay is a function of link length and thus, the link capacity depends on this function. Therefore, the study could consider the different link capacities for each technology and optimize them for the given topology to determine a scenario that provides the least network cost.
Bibliography


