

**Network Reliability and
Resiliency in Next
Generation Networks at
Physical, Datalink, and
Network Layers**

by

Alex Koudrin

A thesis

submitted to Victoria University of Wellington

in fulfilment of the

requirements for the degree of

Master of Science

in Computer Science.

Victoria University of Wellington

2007

Abstract

The “Next Generation Network”(NGN) will replace today’s separate networks for telephone and data. It is expected that the NGN will allow the introduction of new services, such as broadcast media transmission and video on demand. A major consideration with the NGN is that it provides an equivalent voice service quality and resiliency as the current telephone network. The analysis carried out for this thesis looked at the physical reliability of the network topology, network resiliency of the most popular candidate protocols at datalink and network layers, and the end-to-end voice performance over an NGN corresponding to the New Zealand network. This showed that the physical layer reliability of the network could surpass the reliability of the telephone network. The results also showed that in a representative ladder-based core network Multi Protocol Label Switching has a much better resiliency than Open Shortest Path First. In a representative tree-based access topology MPLS was found to be superior to Ethernet. The analysis of an end-to-end voice solution over the NGN yielded results which showed that the network complies with the most strict standards and the voice is of current telephone quality provided that packet losses in the NGN are below 1.34 %, which is relatively high to the expected NGN packet loss figures.

Acknowledgments

I would like to thank my supervisor Peter Komisarczuk for his feedback, patience, and support in helping me produce this thesis.

Contents

1	Introduction	1
1.1	Context and Motivation	1
1.2	Research Question	4
1.3	Contributions	4
1.4	Overview of Main results	5
1.5	Thesis Outline	6
2	Related Work	9
2.1	Public Switched Telephone Network	10
2.2	Next Generation Network Architecture	12
2.2.1	Packet Network Architecture	14
2.2.2	Migration Path	15
2.2.3	Security	16
2.3	Last Mile Access Technologies	17
2.3.1	Digital Subscriber Line	18
2.3.2	Cable	18
2.3.3	Passive Optical Network	19
2.4	Failure Modes	19
2.5	VoIP Transport Concerns	20
2.5.1	QoS Concerns of PSTN and VoIP Calls	21

2.6	VoIP Quality Assessment	24
2.6.1	E-model	24
2.6.2	Objective	26
2.7	Reliability	28
2.7.1	Physical Network Reliability	29
2.7.2	Metrics	30
2.8	Resiliency	31
2.9	Review of Key NGN Protocols	32
2.9.1	Legacy and Other Protocols	32
2.9.2	Open Shortest Path First	32
2.9.3	Multi Protocol Label Switching	33
2.9.4	Overview of Ethernet	35
2.9.5	Rapid Spanning Tree Protocol	37
2.10	NGN Transport and Signaling Protocols	38
2.11	Quality of Service	39
2.11.1	Integrated Services	40
2.11.2	Differentiated Services	40
2.11.3	Other Solutions	41
2.11.4	Metrics	42
2.12	Summary	44
3	Research Question	45
3.1	Research Question	45
3.2	Justification	46
3.2.1	Network Reliability	47
3.2.2	Resiliency	48
3.2.3	Experiment Analysis Steps	48

<i>CONTENTS</i>	vii
3.2.4 Simulation	49
3.2.5 Failure Modes	50
3.2.5.1 Other Metrics	50
3.2.6 End-to-end Quality Analysis	51
3.2.6.1 Metrics analysis	51
3.2.6.2 VoIP Analysis with E-model	51
3.2.7 Other Issues	52
3.3 Core of the Thesis	52
4 Next Generation Core Network	55
4.1 Introduction	55
4.2 Physical Topology	56
4.2.1 Ladder	56
4.2.2 Redundant Ladder	59
4.2.3 Physical Characteristics	60
4.3 Physical Reliability Analysis	62
4.4 Experiment Setup	65
4.4.1 Performance Metrics	65
4.4.2 Failure Modes	67
4.4.3 Protocols	67
4.4.3.1 Overview of OSPF Resiliency	68
4.4.3.2 Overview of MPLS Resiliency	72
4.4.4 Simulation Execution and Data Collection	74
4.5 Results	76
4.5.1 Uncertainties	76
4.5.2 Rerouting	78
4.5.3 Packet Loss	80

4.5.4	OSPF Hello Timer Variation	82
4.5.5	Delay	84
4.5.5.1	Small Link Capacity Study	84
4.5.5.2	Large Link Capacity Study	85
4.5.6	Jitter	86
4.6	Summary	87
5	Next Generation Access Network	89
5.1	Introduction	89
5.2	Physical Topology	91
5.2.1	Alternative Topologies	94
5.2.2	Physical Characteristics	94
5.3	Physical Reliability Analysis	95
5.4	Experiment Setup	97
5.4.1	Performance Metrics	97
5.4.2	Failure Modes	98
5.4.3	Protocols	99
5.4.3.1	MPLS in the Access Network	100
5.4.3.2	Ethernet in the Access Network	100
5.4.4	Simulation Execution and Data Collection	101
5.5	Results	102
5.5.1	Uncertainties	102
5.5.2	Rerouting	102
5.5.3	Packet Loss	104
5.5.4	Delay	105
5.5.5	Jitter	106
5.6	End-to-end Network Analysis	107

5.6.1	Conformance to ITU-T Standards	108
5.6.2	VoIP Call Quality Analysis	110
5.7	Summary	112
6	Conclusions	115
6.1	Contributions	116
6.2	Future Work	117
A	Miscellaneous Core Network Results	127
A.1	Rerouting	127
A.1.1	Delay	130
A.1.1.1	Small Link Capacity	130
A.1.1.2	Large Link Capacity	132
A.1.2	Jitter	134
A.1.2.1	Small Link Capacity	134
A.1.2.2	Large Link Capacity	136
B	Miscellaneous Access Network Results	139
B.1	Rerouting	139
B.2	Delay	141
B.3	Jitter	142
C	Reliability Concepts	143
C.1	Fundamental Concepts	143
C.2	Basic Network Configurations	146
C.2.1	Series	146
C.2.2	Parallel	146
C.2.3	r-out-of-n	148
C.3	Reliability in Graphs	149

C.3.1	N-modular Redundancy	149
C.4	Markov Modeling	150
C.5	Fault-Tree Analysis	152
C.6	Failure Mode and Effect Analysis	152
C.7	Reliability Optimisation	153
C.8	Network Reliability Implementation	154

List of Figures

1.1	Conceptual comparison of PSTN (top) and NGN (bottom).	2
1.2	Block structure of this thesis	7
2.1	SS7 signaling points and elements	11
2.2	NGN functional architecture	13
2.3	End-to-end packet network architecture	15
3.1	Network clouds and focus of experiments	45
4.1	Base ladder topology. Access network and edge router complexity are simplified.	57
4.2	Redundant ladder topology	59
4.3	Traffic received by a destination from one source versus simulation time. Illustration of measurements for a sample failure.	75
4.4	ETE Delay versus simulation time. Illustration of measurements.	76
4.5	OSPF <i>HelloInterval</i> variation	83
5.1	Base tree topology	91
5.2	Redundant tree topology	93

5.3	Traffic received by N 0 from all sources, N 0 – N 8 versus simulation time. Illustration of measurements for a sample failure.	102
5.4	Summary of the worst case delays in a VoIP path in milliseconds	108
5.5	E-model R-value versus error rate e	111
5.6	E-model R-value versus mouth-to-ear delay	112
C.1	Series configuration	146
C.2	Parallel configuration	148
C.3	Triple modular redundancy configuration with a non-redundant voter	150
C.4	Markov reliability model for two identical parallel elements	150
C.5	Inclusion-Exclusion Principle implementation	155

List of Tables

2.1	Categorisation of Unplanned Failures	20
2.2	Mapping between E-model R-value, MOS, and transmission speech quality	26
2.3	IP QoS class definitions from ITU-T Recommendation Y.1541 [5]	42
2.4	Recommended ETE mouth-to-ear delay for VoIP [1]	43
4.1	Effective reliability using two parallel paths, P_1 and P_2 , between E 0 and E 4, given router and link reliabilities	63
4.2	The ETE reliability, given the reliability of the access network and the core network, $ETERel = AccessRel^2 \cdot CoreRel$	64
4.3	Failure modes examined	67
4.4	OSPF configurable timers	69
4.5	Summary of uncertainties in milliseconds	77
	(a) Small link capacity	77
	(b) Large link capacity	77
4.6	OSPF effective rerouting times in milliseconds	78
	(a) Link failures. Uncertainty range is 145.42 ms.	78
	(b) Node failures. Uncertainty range is 993.73 ms.	78
4.7	MPLS effective rerouting times in milliseconds	79
	(a) Link failures. Uncertainty range is 0.41 ms.	79

(b)	Node failures. Uncertainty range is 0.17 ms.	79
4.8	OSPF packet loss to each destination A 0 – A 7 in packets per second	81
4.9	Small link capacity. Difference between delay after failure and delay before failure in milliseconds	84
(a)	OSPF. Uncertainty range is 0.33 ms.	84
(b)	MPLS. Uncertainty range is 0.13 ms.	84
4.10	Large link capacity. Difference between delay after failure and delay before failure in milliseconds	86
(a)	OSPF. Uncertainty range is 0.56 ms.	86
(b)	MPLS. Uncertainty range is 0.79 ms.	86
4.11	Small link capacity. Difference between jitter after failure and jitter before failure in milliseconds.	87
(a)	OSPF. Uncertainty range before failure is 0.10 ms and 0.09 ms after failure.	87
(b)	MPLS. Uncertainty range before failure is 0.07 ms and 0.06 ms after failure.	87
4.12	Large link capacity. Difference between jitter after failure and jitter before failure in milliseconds.	88
(a)	OSPF. Uncertainty range before failure is 0.002 ms and 0.002 ms after failure.	88
(b)	MPLS. Uncertainty range before failure is 0.001 ms and 0.001 ms after failure.	88
5.1	Effective reliability of two parallel paths, P_1 and P_2 , be- tween S 0 and S 7, given router and link reliabilities	95
5.2	Failure modes examined	98

5.3	Summary of uncertainties in milliseconds	103
5.4	Ethernet rerouting	103
	(a) Link failures. Uncertainty range is 133.05 ms.	103
	(b) Node failures. Uncertainty range is 180.67 ms.	103
5.5	MPLS effective rerouting times in milliseconds	104
	(a) Link failures. Uncertainty range is 0.42 ms.	104
	(b) Node failures. Uncertainty range is 0.21 ms.	104
5.6	Ethernet packet loss from source N 1 – N 8 to destination N 0 in pps. Without losses N 0 receives 800 pps. Uncer- tainty range is 0.	104
5.7	Difference between delay after failure and delay before fail- ure in milliseconds	106
	(a) Ethernet. Uncertainty range is 0.45 ms.	106
	(b) MPLS. Uncertainty range is 0.11 ms.	106
5.8	Difference between jitter after failure and jitter before fail- ure in milliseconds	107
	(a) Ethernet. Uncertainty range is 0.09 ms.	107
	(b) MPLS. Uncertainty range is 0.05 ms.	107
A.1	Uncertainty range for OSPF rerouting times in milliseconds	127
	(a) Link failures	127
	(b) Node failures	127
A.2	Uncertainty range for MPLS effective rerouting	128
	(a) Link failures	128
	(b) Node failures	128
A.3	MPLS average setup and rerouting times	129
	(a) Setup	129

(b) Rerouting	129
A.4 OSPF <i>HelloInterval</i> variation	129
A.5 Small link capacity. Uncertainty range for delay before failure in milliseconds	130
(a) OSPF	130
(b) MPLS	130
A.6 Small link capacity. Summary of uncertainty ranges for delay after failure in milliseconds.	131
(a) OSPF	131
(b) MPLS	131
A.7 Small link capacity. Delay after failure in milliseconds . . .	131
(a) OSPF. Uncertainty range is 0.38 ms.	131
(b) MPLS. Uncertainty range is 0.17 ms.	131
A.8 Large link capacity. Uncertainty range for delay before failure in milliseconds	132
(a) OSPF	132
(b) MPLS	132
A.9 Large link capacity. Uncertainty range for delay after failure in milliseconds.	132
(a) OSPF	132
(b) MPLS	132
A.10 Large link capacity. Delay after failure in milliseconds. . .	133
(a) OSPF. Uncertainty range is 0.77 ms.	133
(b) MPLS. Uncertainty range is 0.83 ms.	133
A.11 Small link capacity. Uncertainty range for jitter before failure in milliseconds	134
(a) OSPF	134

(b) MPLS	134
A.12 Small link capacity. Uncertainty range for jitter after failure in milliseconds.	134
(a) OSPF	134
(b) MPLS	134
A.13 Small link capacity. Jitter after failure in milliseconds.	135
(a) OSPF. Uncertainty range is 0.09 ms.	135
(b) MPLS. Uncertainty range is 0.06 ms.	135
A.14 Large link capacity. Uncertainty range for jitter before fail- ure in milliseconds	136
(a) OSPF	136
(b) MPLS	136
A.15 Large link capacity. Uncertainty range for jitter after fail- ure in milliseconds.	136
(a) OSPF	136
(b) MPLS	136
A.16 Large link capacity. Jitter after failure in milliseconds.	137
(a) OSPF. Uncertainty range is 0.002 ms.	137
(b) MPLS. Uncertainty range is 0.001 ms.	137
B.1 Uncertainty range for RSTP rerouting times in milliseconds	139
(a) Link failures	139
(b) Node failures	139
B.2 Uncertainty range for MPLS effective rerouting	140
(a) Link failures	140
(b) Node failures	140
B.3 MPLS average setup and rerouting times	140

(a)	Setup	140
(b)	Rerouting	140
B.4	Delay after failure in milliseconds	141
(a)	RSTP. Uncertainty range is 0.38 ms.	141
(b)	MPLS. Uncertainty range is 0.13 ms.	141
B.5	Summary of uncertainty ranges for delay after failure in milliseconds	141
(a)	RSTP	141
(b)	MPLS	141
B.6	Summary of uncertainty ranges for jitter after failure in milliseconds	142
(a)	RSTP	142
(b)	MPLS	142
B.7	Jitter after failure in milliseconds	142
(a)	RSTP. Uncertainty range is 0.38 ms.	142
(b)	MPLS. Uncertainty range is 0.13 ms.	142

Chapter 1

Introduction

1.1 Context and Motivation

Next Generation Network (NGN) is a broad concept of a converged network that would carry all services currently supplied by separate networks. Considering current residential services we would see best effort data and voice becoming the two primary services in the near future. Future services of the NGN include TV broadcast and high quality interactive data and voice services, called “triple play”. Voice traffic is currently transported using the Public Switched Telephone Network (PSTN), which has a dedicated transport and signaling network. Best effort data is transported using IP forwarding through a separate network, sometimes, using the same physical transport links as PSTN.

Figure 1.1 shows the conceptual comparison between PSTN and NGN. In PSTN an endpoint — the telephone, is connected to a local exchange (LE). LE is then connected to the core time-division multiplex (TDM) voice transport network, which consists of transit exchanges (TEs). Signaling System 7 (SS7) is a separate dedicated packet network that pro-

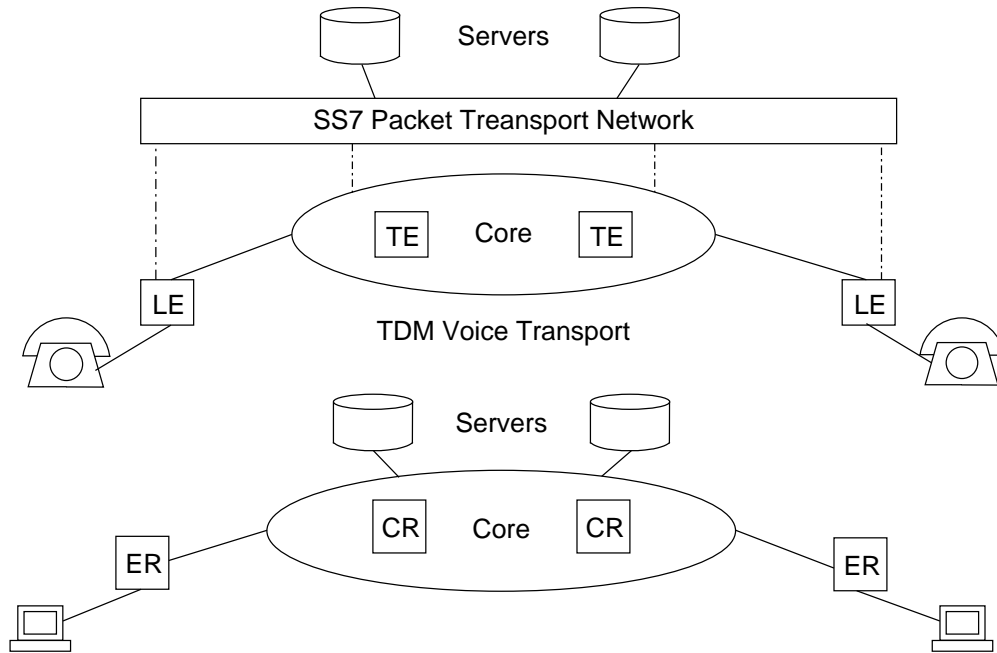


Figure 1.1: Conceptual comparison of PSTN (top) and NGN (bottom).

vides all signaling in the PSTN. Any additional services are provided by the servers attached to SS7 called the Intelligent Network (IN).

The second part of Figure 1.1 shows the NGN at the same level of abstraction as the PSTN. The local access parallels that of PSTN, but the access point aggregates many subscriber lines into a switch, which connects to an edge router (ER). An ER connects to the core network consisting of core routers (CRs). The major difference is that NGN accommodates signaling and transport using a single packet network, which provides financial and management advantages over the PSTN. Also, call control is a function of the residential gateway or the user equipment and servers in the core network. This is compared to a dumb phone with a local exchange on the end of a copper local loop. The LE and signaling network provide the intelligence to manipulate calls.

PSTN is by definition a carrier grade network, which provides highly available voice services. Historically, PSTN was built to have “five nines”

— 0.99999 element reliability, which translates into the often measured and quoted 0.9993 end-to-end reliability. For NGN to be a carrier grade network it must be at least as reliable as PSTN, provide high quality transmission guarantees, and be easy to manage and operate [41]. Most of today's deployed IP networks do not provide a network availability equivalent to the PSTN.

In the last decade there have been many advances in network technology [34]. Hardware switching elements are carrying higher than ever throughputs with high reliability. Market forces such as traffic and subscriber growth, equipment cost reduction, and new technology penetration, have deep impact on network buildouts. Some Quality of Service (QoS) can now be provided using such technologies as Multi-protocol Label Switching (MPLS).

The technology has matured to the point where it can be used to deliver NGN and in the last several years ITU-T has produced several standards addressing various aspects of the NGN. What remains is testing the technology in typical network configurations against the NGN standards. Once the network conforms to the standards it becomes a carrier grade solution that can safely replace the PSTN.

Migration from PSTN to NGN is happening throughout the world and New Zealand Telecom is planning to replace PSTN by NGN over the next decade. Thus, at this time it is vital to determine the effectiveness of the NGN technology prior to deployment.

1.2 Research Question

This thesis investigates the answer to the following question: *How well can current IP and Ethernet technology fulfill NGN network and service reliability and resiliency requirements?*

Many aspects of NGN technology need validation against the standards, however, the focus is narrowed down to the lower three network layer technologies. Three aspects of current technology are investigated in this thesis: physical topology; datalink layer protocols, which are MPLS and Ethernet; and a representative network layer protocols, which is Open Shortest Path First (OSPF).

Such an investigation needs to be undertaken prior to the NGN roll-out to make certain the new network will be at least equivalent to the PSTN voice network today. Physical topologies for various parts of the network need to be understood in terms of their reliabilities and whether reliability meets the set target. In particular, this thesis investigates a representative ladder topology for the core network and the tree topology for the access network. Resiliency of the most used protocols — OSPF, MPLS, and Ethernet — should be compared to select appropriate protocols for NGN deployment. Finally, the end-to-end quality of the voice service should satisfy the goal of matching or exceeding the PSTN using the best NGN technology.

1.3 Contributions

This thesis presents three main contributions.

First, is the application of reliability theory to a specific physical NGN topology, which is chosen to approximate a design for the NGN for New

Zealand. The core and the access network topologies are analysed separately and then combined. The analysis demonstrates the range of values of reliability for the New Zealand NGN and whether those figures conform to the NGN requirements.

Second, is the resiliency analysis of the main rival datalink and network layer protocols that are envisaged to be part of NGN. The analysis involves quality performance assessment by comparing various metrics to appropriate standards and using models to approximate voice quality. In the core network OSPF and MPLS are compared, whereas in the access network Ethernet and MPLS are compared.

Third, is the analysis of the complete end-to-end voice solution, using the best protocols determined through analysis of the second contribution.

1.4 Overview of Main results

Both the core and the access network each were found to require 0.9999 reliability to satisfy the end-to-end PSTN reliability of 0.9993. To satisfy 0.9999 reliability, thresholds for individual router and link reliabilities were calculated. For the core network the router and link reliabilities must be at least 0.999, whereas for the access network router reliability must be at least 0.999 and link reliability must be at least 0.9999 or vice versa.

Resiliency comparison of the datalink and network protocols used rerouting delay as the primary metric, with packet loss due to rerouting, delay and jitter used as secondary metrics. In the core network OSPF and MPLS were compared. MPLS was found a much better choice than

OSPF because the rerouting delay was at most 13 ms — much smaller than 6 – 40 seconds for OSPF, which was also supported by the packet loss endured during the rerouting. The difference between end-to-end delay after failure and delay before failure was approximately 1.5 ms, which is around 6 % of the delay before failure. In the access network Ethernet and MPLS were compared. Again, MPLS dominated the performance with 10 ms rerouting time versus 4 s for Ethernet when using the Rapid Spanning Tree Protocol to recover from failures. The difference between end-to-end delay after failure and delay before failure was approximately 1 ms, which is around 13 % of the delay before failure. The percentage is larger than that of the core network due to shorter propagation delay. In both cases jitter was seen to be insignificant within the end-to-end quality of service measurement.

The core and the access results were combined and the complete end-to-end network was analysed for the quality of voice delivered to users. Typical values for other parameters were used where the investigation was out of the scope. According to ITU-T recommendation the end-to-end solution conformed to the most strict requirements for VoIP service. The E-model also showed that voice quality is on the par with that of PSTN even during failures, provided that random packet loss remains below 1.34 %.

1.5 Thesis Outline

Figure 1.2 shows the overall structure of my thesis. Chapter 2 overviews the broad scope of related work on NGN. Chapter 3 develops the research question in some detail. The chapter also narrows down and analyses

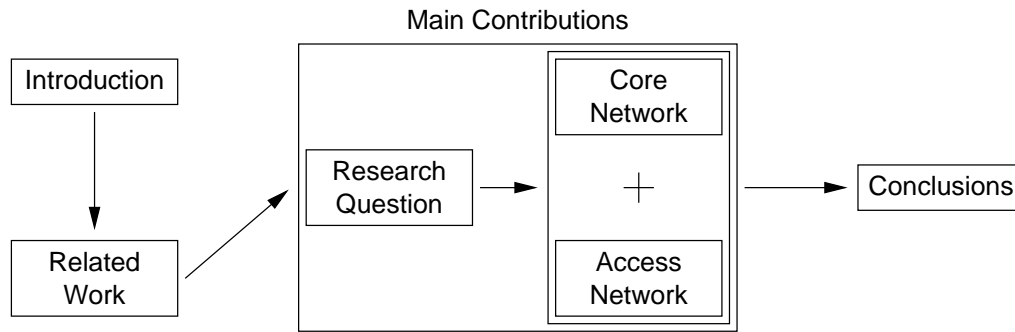


Figure 1.2: Block structure of this thesis

the information presented in Chapter 2 so it is directly relevant to the research question. Chapter 4 analyses reliability of the ladder topology representing the core network. It also compares resiliency and quality performance of OSPF and MPLS. Chapter 5 parallels Chapter 4 by analysing the physical reliability of a tree-based access topology and by comparing Ethernet and MPLS. At the end, the chapter combines the core and access results and analyses the end-to-end voice performance of the NGN. Chapter 6 draws conclusions, summarises the main contributions, and provides directions for future work.

Appendix A and B present more extensive results, including uncertainties, for the core network and the access network measurements respectively. Reliability concepts used for computations are described in Appendix C.

Chapter 2

Related Work

This chapter presents a broad review of literature available on NGN, which is narrowed down and analysed for the focus of this thesis in Chapter 3.

To establish the basis to which to compare the NGN, the chapter begins with a brief description of the current telephone network. The architecture of the NGN is discussed in Section 2.2, which also discusses the general packet network topology, migration path from PSTN to NGN, and security issues. Section 2.3 provides a brief overview of the last mile access technologies, including digital subscriber line and cable. The typical failure modes in a packet network are discussed in Section 2.4, which are used to decide what failures to introduce in the experimental work.

Section 2.5 discusses the voice over internet protocol (VoIP) issues, such as the technical factors that affect voice and general quality of service issues of a PSTN and VoIP call. To be able to assess the quality of a VoIP call Section 2.6 overviews the E-model, objective modeling, and combinations of the two approaches. Section 2.7 overviews reliability issues in more detail and discusses physical network reliability. The section

concludes by describing end-to-end downtime and defects per million as the two major metrics to measure physical network reliability.

Making the network reliable at the physical layer means failures still occur and additional failures may be introduced by higher layers in the protocol stack. The ability to react to a failure is called network resiliency, which is described in Section 2.8. Resiliency in this thesis refers to how a datalink or a network protocol reacts to failure. Section 2.9 overviews the protocols — OSPF, MPLS, Ethernet with RSTP — that are viewed by the NGN community as key for the new network. These protocols are used in the experimental work in this thesis.

Because the NGN combines transport and signaling, there is a set of signaling protocols identified in Section 2.10. However, these are of much less importance than transport protocols as they occupy a small fraction of the total transport bandwidth. Quality of service (QoS) is important for providing VoIP and there are some common methods for implementing it, such as differentiated and integrated service architectures. Section 2.11 gives a brief overview of QoS methods and more importantly of metrics that can be used to determine QoS level of a carrier grade service such as VoIP.

2.1 Public Switched Telephone Network

There are many books and other resources on Public Switched Telephone Network (PSTN). The majority of the material in this section is found in Davidson and Peters [31] and Modarressi and Mohan [58].

PSTN is based on time-division multiplexing (TDM) voice transport. The control is delegated to a physically separate signaling network, Sig-

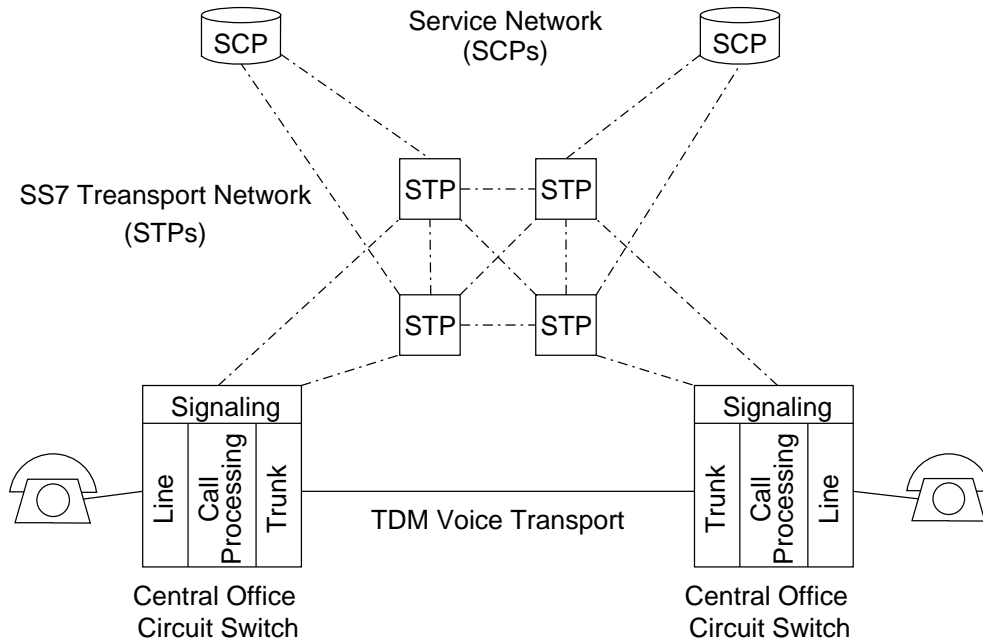


Figure 2.1: SS7 signaling points and elements. Based on Figure 4-2 from Davidson and Peters [31].

nalng System 7 (SS7), with its own distinct transport infrastructure and protocol suite. The SS7 network provides a real-time highly reliable means for switches to communicate and exchange call, connection, and service control information. PSTN also provides other services beyond call setup, such as call forwarding and call waiting provided by a service layer called the Intelligent Network (IN). PSTN is a network in which all intelligent components reside within the network, whereas the end points (telephones) have very little functionality.

Figure 2.1 shows the overall architecture of the PSTN, where:

- **Central Office Switch:** connects voice circuits and perform the necessary signaling functions to originate and terminate calls
- **Signal Transfer Point (STP):** routes all the signaling messages in the SS7 network. STPs are arranged in a hierarchy as are the switches. For redundancy, STPs are mated in pairs.

- **Service Control Point (SCP):** provides access to databases for additional routing information used in call processing. A SCP is the key element for delivering IN applications on the telephony network.

Using Figure 2.1 a simplified call setup that involves two local central office (CO) switches can be analysed. The line module on the incoming side is responsible for detecting the off-hook event of the connected telephone. The line module then provides dial tone and collects dialed digits. The processing module analyses the digits passed to it from the line module and initiates signaling over the SS7 network. The SS7 network routes the call setup message to the CO switch of the called party and initiates the establishment of the voice channel between the two parties over the TDM network.

2.2 Next Generation Network Architecture

The Next Generation Network (NGN) is a packet based network, which unifies voice, data, and other potential multimedia services that currently run over dedicated networks.

There are many organisations and forums that are involved in NGN development and producing numerous general NGN references, including the primary organisation ITU-T [7, 8, 48, 24, 26, 51] ATM Forum [25], and the Multiservice Switching Forum [33, 30]. There are also several comprehensive books [28, 57, 62] describe NGN background, issues, and detailed protocols. Numerous papers, for example, Lee et al. [58, 52] have also been written on the general NGN architecture and issues.

Figure 2.2 identifies the NGN architecture functional components. The

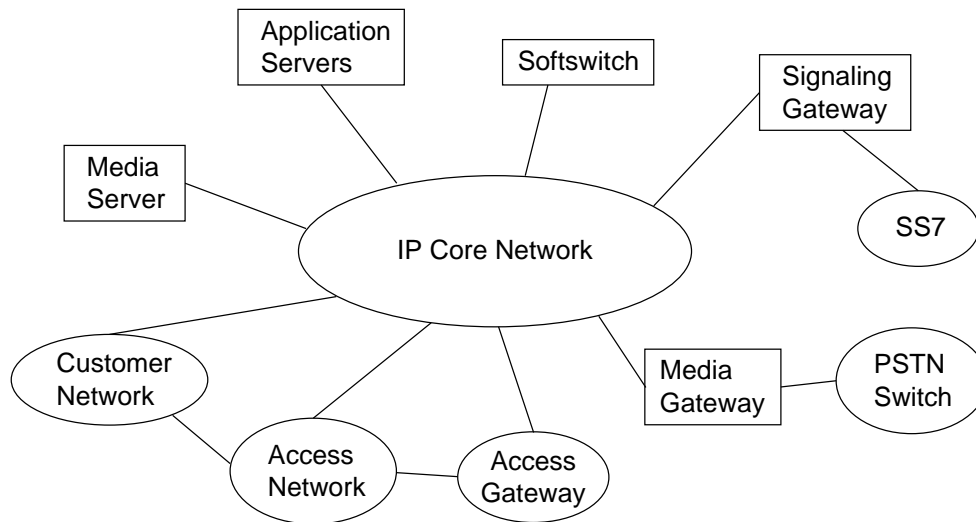


Figure 2.2: NGN functional architecture based on Figure 1 from Johnson et al. [46]

architecture has the intermediate elements for compatibility with the current PSTN. The short description of the functional elements follows:

Access Network provides connectivity between the customer premise equipment and the core network. Access method can be one of several including broadband and dial-up.

Access Gateway is located in the service provider's network. It supports the line interface to the core network for phones, PBXs, and devices. Functions such as packetisation and echo control are part of access gateway.

Media Gateway supports a trunk (voice) side interface between to the PSTN or IP router flows in the packet network. It compresses and packetises the voice data, and delivers compressed voice packets to the IP network and vice versa.

Softswitch (also called media gateway controller) handles the registration and management of resources at the media gateway and is

responsible for configuration and termination of calls, monitoring network resources, tracking billing, handling security and authentication, and performing a number of other critical administrative tasks.

Signaling Gateway provides a signaling interface between VoIP signaling and SS7 signaling.

IP Core Network serves as a high performance transport for IP packet transport

Media Server is under control of a call agent and provides announcements, tones, and collects user information

Application Server provides additional features not directly hosted on a call agent

NGN architecture is based around VoIP as its primary service. The key part of NGN is the softswitch architecture [62], which effectively replaces the PSTN switching functionality. The three main entities shown in Figure 2.2 are the signaling gateway interfacing with SS7, media gateway interfacing with PSTN voice transport, and the softswitch or the media gateway controller that provides call control. Section 2.10 identifies the different protocols used in the softswitch architecture to communicate various functions, such as call control and call state, between the different entities. Ohrtman [62] provides extensive detail on the softswitch architecture, which is out of scope of this thesis.

2.2.1 Packet Network Architecture

A packet network typically consists of several networks. Figure 2.3 shows

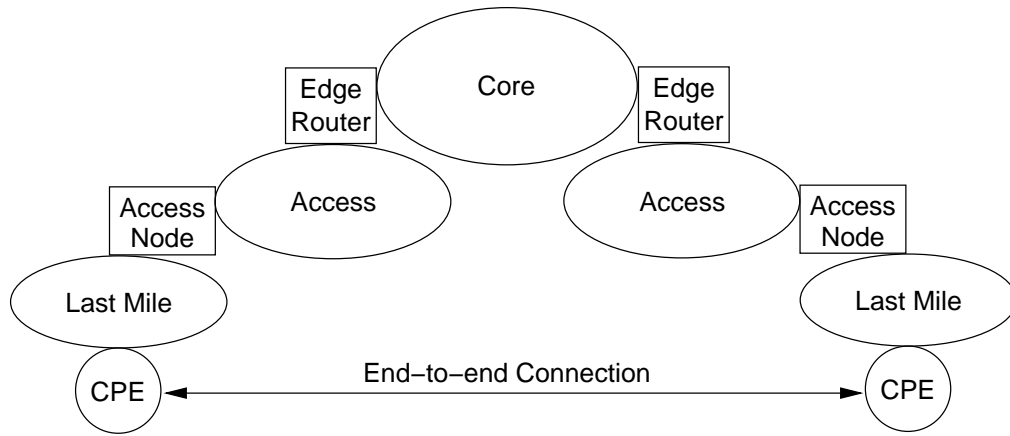


Figure 2.3: End-to-end packet network architecture

the typical architecture of an end-to-end packet network. It consists of a low latency and high bandwidth core (or backbone), which typically extends over a large geographic region such as a country, joined with an access network by an edge router. The access network aggregates the traffic from all the access nodes, which perform a lower level aggregation of individual users connected through a last mile access network. The last mile may use a variety of technologies, such as ADSL and Cable, which are briefly described in Section 2.3. Customer Premise Equipment (CPE) refers to all equipment located past the network interface, such as router, LAN, and host.

2.2.2 Migration Path

PSTN needs to be replaced by NGN, but it needs to be gradual and seamless. Migration from the PSTN to the full NGN is a big issue [28, 57, 25, 10, 61]. As the middle ground PSTN-IP-PSTN architecture is generally regarded as the immediate future of NGN. In this architecture the current PSTN and IP networks are inter-worked together with media gateways

and signaling gateways as the interfaces. Those gateways are key to reliability and availability of the entire system. For example, if there is only one gateway then it becomes a single point of failure seriously undermining end-to-end reliability.

Choi et al. [25] mentions some migration strategies for NGN. Migration path may involve replacement of class 4 and 5 switches with media gateways. Alternatively a media gateway may be placed along side of a switch handling some of the traffic. At the end of the switch lifetime the traffic can be seamlessly diverted to the media gateway.

2.2.3 Security

PSTN with its separate physical network is very secure [31] due to factors, such as network intelligence and simple end devices, the nature of billing makes it almost impossible to affect many network devices. It is virtually impossible to interfere with a voice conversation. At worst eavesdropping is possible, which still requires physical access to the line.

IP networks have more security vulnerabilities due to factors such as, intelligence is at the endpoints instead of the network and the cost of being connected is much lower than that of PSTN. As a result, NGN is likely to be susceptible to various security attacks, such as Denial of Service (DoS), theft of service, and invasion of privacy [33].

There are protective measures against insecurity of data networks [35]. One extreme option is a completely separate physical network for VoIP. This is currently the most common option to deliver VoIP capabilities. Virtual overlay networks are also used to achieve a similar goal. Less drastic measures, in terms of management and financial cost,

include firewalls, intrusion detection, encryption, and authorisation and authentication techniques.

Increase in security measures also increases the difficulty of NGN operation [33]. For instance, firewalls may make it more difficult to configure VoIP and may degrade its performance, because different and variable ports may be used by the VoIP signaling protocol. Adding more security means increasing the number of hardware and software components in the network, which increases financial costs and decreases network reliability. A significant piece of hardware that is specific to VoIP security is a Session Border Controller (SBC).

SBCs are typically located between two service provider networks in a peering environment, or between an access network and a backbone network to provide service to residential and/or enterprise customers. They provide a variety of functions to enable or enhance session-based multi-media services (e.g., VOIP). These functions include: a) perimeter defense (access control, topology hiding, DoS prevention, and detection); b) functionality not available in the endpoints (NAT traversal, protocol interworking or repair); and c) network management (traffic monitoring, shaping, and QoS). SBCs are envisaged to be a big part of NGN security. The IETF draft [23] describes more details of the function of SBCs.

2.3 Last Mile Access Technologies

This section gives a brief overview of some of the common last mile access technologies. A comprehensive up-to-date reference on last mile access technologies is Jayant [45]. Also Black [19] and Tan [79] were used as supporting references in this section.

2.3.1 Digital Subscriber Line

xDSL denotes Digital Subscriber Line, where x can be either A(symmetrical), H(igh-bit-rate), V(ery-high-bit-rate), or S(ymetrical). ADSL is the most common type of DSL technology used in last mile access. It runs over the existing PSTN twisted-pair copper cabling. Forward Error Correction (FEC) is used in ADSL due to PSTN lines being highly susceptible to noise, such as AM radio. FEC adds up to 20 ms to the end-to-end delay.

The ADSL modem sends and receives data at the customer premises. An ADSL filter at the customer premise combines voice from a PSTN phone with ADSL data on the same cable. At the central office exchange DSL Access Multiplexer (DSLAM) aggregates from 100s to 1000s subscriber lines onto an STM-1 or STM-4 backhaul link.

2.3.2 Cable

Bi-directional Hybrid Fibre-Coaxial (HFC) network is a broadcast technology created to broadcast video signals to all the endpoints. Data Over Cable System Interface Specification (DOCSIS) is a set of protocols for data transmission between a CM and a CMTS. Cable Modem (CM) interfaces the user PC on one side and a filter on the other. The filter, combines video and data channels. The setup parallels that of DSL. Cable Modem Termination System (CMTS), which is analogous to a DSLAM in ADSL. A cable network is contended by users in the upstream direction and broadcast technology is used to transmit downstream. The delay in a cable network is up to 8 ms, which is much lower than that of ADSL due to better transmission properties.

2.3.3 Passive Optical Network

Passive Optical Network (PON) consists of Optical Link Termination (OLT) located at the central office exchange; 1:N passive splitter/aggregator located near the access region connecting to a maximum of 64 Optical Network Units (ONUs). These entities are linked by fibre connections. For fibre to the curb deployment, ONUs are located in the road-side cabinets and users are connect to ONUs using DSL over copper connections.

The main fiber run on a PON network can operate at 155 Mps up to 2.5 Gbps. Downstream data is transported using a broadcast, whereas upstream data is transported using a time division, multiple access protocol to avoid collisions at the aggregation point — the splitter.

Ethernet over PON (EPON) is a very promising PON technology. It combines the ubiquitous low cost and efficient Ethernet technology with high capacity, low maintenance, up 20 km long reach PON technology. The delays introduced by any PON is within 1 ms, which is much less than DSL or cable.

2.4 Failure Modes

Typical failure modes of today's IP core network have recently been analysed by Markopoulou et al. [54]. As a result of analysis of Sprint's IP backbone network, Markopoulou et al. [54] categorises the failure modes in Table 2.1, which shows the breakdown of unplanned failures. Planned maintenance is not shown, because it cannot be avoided in present IP networks. In NGN maintenance failures may be eliminated as hot swappable IP and MPLS technology improves, which means maintenance can

Failure Class		% Unplanned
Unplanned		100.0
Shared	Router	16.5
	Optical	11.4
	Unspecified	2.9
Individual	Link	68.5

Table 2.1: Categorisation of Unplanned Failures. Adapted from Table 3 of Markopoulou et al. [54].

be performed without service disruption, except in cases of human error. My experiments deal with only unplanned failure scenarios.

Markopoulou et al. [54] defines shared and individual sub categories of unplanned failures. Shared failures are defined as originating from the same cause. For example, several link attached to a router will share the same failure if the router processor fails. Individual failures do not share a cause with any other failures. Such failures are typically link failures.

2.5 VoIP Transport Concerns

Voice over IP (VoIP) is a common term for carrying voice over a packet network. Much information on packet voice transport is contained in the general NGN references [7, 8, 48, 24, 26, 51, 25, 33, 30, 28, 57, 62].

The main concern over VoIP is quality [69, 44, 22, 77, 65, 55, 47, 17, 78], which must be maintained at or above the PSTN quality for NGN to be accepted by consumers. VoIP quality can be influenced in several different ways.

The coder/decoder (codec) used to encode voice packets greatly influences voice quality, encoding delay, and bandwidth. Many different codecs exist, but G.729a is the favorite choice for VoIP in NGN because

it satisfies PSTN quality and a good compromise between encoding time and bandwidth.

End-to-end (ETE) delay is another major factor that affects VoIP quality. Encoding (and decoding) delay is part of the total delay. There is also packet assembly, transmission buffer and receiver buffer delays. The other major source of delay is the general network delay experienced by all packets flowing between routers and switches to get from source to destination.

Section 2.6 discusses the E-model and Section 2.11.4 discusses metrics for packet transport.

2.5.1 QoS Concerns of PSTN and VoIP Calls

This section discusses QoS concerns relating to a typical call in a PSTN and VoIP networks [39]. The issues include network accessibility, routing speed, connection setup reliability, routing reliability, connection continuity, and disconnection continuity. This does not include voice quality, which is discussed in Section 2.6.

Accessibility is primarily a network issue and is the ability to initiate a call when desired. It can be measured by the probability that a user will be unable to use voice services for a given time period. Accessibility is a critical part of network reliability.

Accessibility of a packet network is similar to accessibility of the PSTN. However, accessibility may be lower in a hybrid solution where PSTN and packet network are interfaced together depending on the reliability of the interface.

Routing Speed is the rate at which the calls are set up. It can be measured by a post-dial delay, the time from the last entered piece of input information to the receipt of the disposition of the request, such as a ring back or a busy signal.

A packet network is likely to increase the post-dial delay for two reasons. First, translation between numbers and device addresses will require a registry look up. Second, negotiation between IP devices is more complex than that of PSTN, for example, Session Initiation Protocol (SIP) codec negotiation.

Connection Setup Reliability is the probability that a correctly executed request for a call setup will be extended to the required destination. If the connection cannot be configured, it is determined as a defect and clearly contributes to the measure of unreliability of the network.

Additional handling of packets in packet networks is likely to decrease connection setup reliability. Hardy [39] hypothesise that the call completion rate of 99.5 % for the PSTN may drop to 99.0 %, which is considered barely acceptable.

Routing Reliability is the accuracy of call setup in terms of reaching the specified destination. As for connection setup the measure is the misrouted number of call attempts to the total number of calls. This may be considered as a metric for network reliability as customer experience is affected by this issue.

This type of reliability is likely to fall due to the additional handling required as in the case of connection setup. There is also another source of potential misroutes originating from the possibility

of routing to multiple end-user devices. This is possible when using the same alias for different IP addresses of devices, such as a desk phone and a cell phone. If the call is routed to the wrong end device, this could lead to a misroute from the point of view of the customer. This aspect of routing reliability is a concern of the service logic.

Connection Continuity is the ability to maintain a connection of a satisfactory quality until call completion. Discontinuities may be the result of spontaneous disconnects, unacceptable degradation, and transfer error. A measure for connection continuity is the expected number of calls that result in disconnects or unacceptable quality degradation as a function of call duration. The issue is both call oriented and network oriented as both the user is concerned about call connection and thus reliability of network is affected and measured.

Additional handling of packets in IP networks will increase spontaneous disconnects and transfer errors by a marginal amount. The main concern with packet networks is quality degradation, which may result users deciding to disconnect voluntarily. The reason for this is the nature of IP networks where there is no hard reservation of the call path as in PSTN, where other traffic cannot interfere with the reserved path. Thus congestion can degrade voice quality in the middle of a call.

Disconnection Reliability is a measure of whether the disconnection instruction produces the desired result, such as correct call termination and billing. The number of failed disconnection attempts to the total number of calls is a metric of how good disconnection reliability is.

PSTN has very reliable hardware mechanisms for on-hook and off-hook events, which therefore can provide accurate billing. In packet networks a disconnection event is detected using software. There are also more devices and interactions between them to negotiate call termination. Thus the additional handling and processing may decrease the reliability of disconnection and billing by an estimation of 2 – 4 times that of the PSTN [39].

2.6 VoIP Quality Assessment

To compare VoIP call quality against the traditional PSTN call quality, a reliable measurement scheme is required. These schemes are outlined in this section.

2.6.1 E-model

For NGN to constitute an attractive alternative to the traditional PSTN, it must provide high-quality VoIP services. The problem of assessing the quality of voice communication over Internet backbones has been extensively studied in the literature. The most popular model is the E-model [3] devised by ITU-T, which is used in design of hybrid circuit-switched and packet-switched networks for carrying high quality voice applications. The model estimates the relative impairments to voice quality when comparing different network equipment and network designs. The E-model estimates the subjective Mean Opinion Score (MOS) rating of voice quality by using objectively measured quantities over these planned network environments.

In the E-model the terminal, network, and environmental quality factors are represented by 20 input parameters. A single output of the model is the R-value (also called the R-factor). Degradation of quality due to individual quality factors, such as loudness, echo, delay, and distortion, are calculated on the same psychological scale. Then these are separated from the reference value. Due to its nature, the E-Model applies only to telephone band (300 – 3400 Hz) handset communication and it is inapplicable to hands-free or wide band communication (150 – 7000 Hz) [3].

Cole and Rosenbluth [27] gives the simplified expression for the R-factor for various codecs under random packet loss:

$$R = \alpha - (\beta_1 \cdot d) - \beta_2(d - \beta_3) \cdot H(d - \beta_3) - \gamma_1 - \gamma_2 \cdot \ln(1 + \gamma_3 \cdot e),$$

where: $\alpha = 94.2$, $\beta_1 = 0.024 \text{ ms}^{-1}$, $\beta_2 = 0.11 \text{ ms}^{-1}$, $\beta_3 = 177.3 \text{ ms}$. For codec G.729a, $\gamma_1 = 11$, $\gamma_2 = 40$, $\gamma_3 = 10$, whereas for codec G.711 these values are: $\gamma_1 = 0$, $\gamma_2 = 70$, $\gamma_3 = 15$. The one way mouth-to-ear delay is given by following relation:

$$d = d_{\text{codec}} + d_{\text{de-jitter buffer}} + d_{\text{network}},$$

where the total loss probability e is given by:

$$e = e_{\text{network}} + (1 - e_{\text{network}}) \cdot e_{\text{jitter buffer}}$$

Typically $e_{\text{jitter buffer}}$ is assumed 0, so $e = e_{\text{network}}$. $H(x)$ is a Heavyside function:

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

R-value lower limit	MOS	Speech Quality
90	4.34	Best
80	4.03	High (lowest PSTN)
70	3.6	Medium
60	3.1	Low
50	2.58	Poor

Table 2.2: Mapping between E-model R-value, MOS, and transmission speech quality

The results for the G.729a codec assume a 20 ms packet size, while the G.711 results are for a 10 ms packet size. The results for both G.729a and G.711 listed above are limited to random packet loss. Other values need to be derived for other combinations of codecs; packet size and error mask distributions. For the G.729a codec:

$$R = 83.2 - 0.024d - 0.11(d - 177.3) \cdot H(d - 177.3) - 40\ln(1 + 10e)$$

Table 2.2 shows the R-values corresponding to MOS and speech quality. $R \geq 80$ corresponds to PSTN quality speech.

2.6.2 Objective

Objective modeling is a method to assess the VoIP quality by injecting sample speech segment across a voice transport path. Using psycho-acoustic principles, objective modeling compares the output speech with input speech to produce opinion without reference to underlying channel conditions. To capture the conversational impairments due to delay, the low level transport measurements of delay and echo must still be overlaid on top of this (using the E-model).

Cole and Rosenbluth [27] list the advantages of an objective model as follows:

- No assumptions are made about the underlying network (coder, de-jitter buffer, error mask, and packet size)
- Results may be more accurate as predicted opinion is based on fundamental psycho-acoustics, unlike the E-model, in which an interpolation of subjective testing results is used.

Speech layer and packet layer are the two main objective models currently under standardisation [78].

The disadvantages of objective models are as follows [27]:

- High cost and complexity
- Inaccurate under certain conditions, such as in the case of temporal clipping
- Intrusive by nature, whereas the E-model can be implemented as either intrusive or non-intrusive
- The cause of degradation of quality remains unknown

There are efforts to improve the E-model combining it with objective modeling. For example, Takahashi et al. [78] introduces a second order polynomial in terms of delay and changes the talker echo impairment factor. The new model has shown to be a better fit when it was validated against the data obtained from commercial VoIP products.

Another method [78] directly combines the packet level measurements, such as delay, jitter, and packet loss. This relies on using thresholds to define critical quality of voice conversation. Implementation of this method is very simple, however, defining arbitrarily chosen thresholds

is its weakness, which is overcome by the E-model. Also, it does not attempt to combine the transport metrics in a meaningful way with respect to voice quality and therefore we ignore it.

2.7 Reliability

PSTN has an ETE reliability of 99.93 % or 0.9993 [28, 46]. The goal of NGN is to have at least equivalent reliability for the same set of voice services.

Johnson et al. [46] is an excellent summary of VoIP reliability issues. It lists four techniques required to achieve VoIP reliability:

1. Fault-tolerant hardware is a traditional means of achieving reliability by seamlessly switching to a redundant element within a single platform
2. Fault-tolerant software relies on duplicating software processes
3. Physical network reliability relies on duplicating low reliability network elements to achieve higher network reliability
4. VoIP element interface redundancy employs multiple interfaces and real-time switch over between them

Out of the four techniques, physical network reliability is explored further. Reliability of networks is a vast and complex mathematical and statistical area. This has been a big research in early in 20th century as electricity grids and PSTN needed to be designed with very high reliability. For computer networking only a small part of the field is needed to estimate the physical layer reliability. There are many books written about reliability, a particularly good source is Shooman [74], which has been

used as an extensive reference to reliability and the underlying statistics concepts.

2.7.1 Physical Network Reliability

Appendix C defines basic reliability concepts delves into more technical details of reliability aspects used in my thesis. This section provides a non-technical overview.

Element reliability at the lowest level is reliability of a single network element such as an IP router or switch. A router is a device for switching packets and contains two processors: a route processor and a line card. Routers are usually designed with redundant processors, line cards, power supplies and so on. Current router availability is potentially better than 99.999 % with appropriate redundancy in place. In reality it is closer to 99.90 % – 99.99 % due to the following factors.

- Chassis remain a single point of failure with large mean time to repair
- To minimise time to market, development and testing is shortened
- Software and hardware upgrades can cause 10 – 60 min downtime per year [46]
- Denial of Service (DoS) attacks can cause outages [28, 57]

In 2004, effective router reliability was only 0.9990 to 0.9999 due to maintenance downtime, software and chassis low reliability and average link reliability was approximately the same [82, 46]. Currently, in the year 2006, vendors advertise routers and switches with 0.99999 [6] reliability

due to hitless switchover technology — eliminating maintenance downtime, improved software and hardware redundancy.

PSTN element reliability should not be used for all VoIP elements. VoIP networks vary in size, function, use fault tolerant protocols, and use more redundant components than PSTN. Thus the reliability requirements for a single element should be based on network design and complexity of the element.

2.7.2 Metrics

VoIP service is defined to be available only if the logical end-to-end connection can be completed with sufficient QoS and the calls can be maintained for sufficient time to complete the transaction [46].

Johnson et al. [46] proposes two major metrics for reliability: end-to-end downtime and Defects Per Million (DPM).

End-to-end downtime models a single path between the two end devices such as POTS phones. The path will involve a sequence of equipment such as switches and routers that make the call possible. Each element in the sequence has a reliability measurement and combined those reliabilities produce the total path reliability. This measure does not reflect the impact of outages on customers as it does not incorporate customer demand during outages. Downtime is usually calculated in minutes per year, thus yearly downtime $YD = (1 - A) \times 525600 \text{ min/yr}$ with $A = 99.999 \%$ corresponding to $YD = 5.26 \text{ min/yr}$.

DPM counts customer demands not served. More precisely DPM is the average number of blocked and cutoff calls per million attempted calls. $DPM = (1 - A) \times 10^6$. Thus $A = 99.999 \%$ corresponds to 10 DPM.

2.8 Resiliency

Resiliency is the ability to react to a failure, whereas reliability tries to minimise failures by using redundant network elements.

Failure resiliency is very important in NGN as it ensures the ETE service is highly available. Effectively, resiliency makes the reliability visible at a higher layer of the network protocol stack. For example, with two separate physical paths between two points if one of the paths fails, the a higher layer protocol needs to detect the failure and reroute the traffic using one of the redundant paths to take advantage of the high reliability. If on the contrary the protocol had no resiliency, only one redundant path could be used, which effectively decreases the reliability to a single path.

There has been some research in Europe in the Protection Across Network Layers (PANEL) project, into trying to build models to analyse the multilayer resilience mechanisms [32]. The models are high level, involving a lot of approximations. The difficulty is that resiliency depends on which protocols are used at each layer, how they interact with one another in the face of failure. Improvement in coordination between resilience mechanisms is the ultimate goal of the PANEL project. However, more research efforts are needed to get closer to that goal. The outcome is a set of guidelines for the type of failure recovery strategy needed in specific situations. In practice, miscoordination often occurs resulting in failure detection time being much longer than it could have been. From time to time network managers typically change protocol configuration in the network to try and improve failure resiliency.

2.9 Review of Key NGN Protocols

This section examines the most popular current protocols that are candidates for NGN. The three key protocols are OSPF, MPLS, and Ethernet. Some of these protocols can be used in the core network and/or the access (aggregation) network in Figure 2.3. OSPF is a routing protocol that is used only in the core network. Ethernet is a datalink protocol used only in the access network. MPLS is a datalink protocol and it can be used in both core and access networks.

2.9.1 Legacy and Other Protocols

Other technologies, such as Asynchronous Transfer Mode (ATM) and Frame Relay (FR), are not examined as they are recognised to be older technologies that are being gradually phased out of packet networks [57]. These protocols provide performance guarantees, but suffer from many disadvantages, such as being very complicated and difficult to scale.

Routing protocols such as RIP, ISIS, and BGP are either legacy protocols, much less popular, or inappropriate. RIP has been surpassed by OSPF, whereas a very similar Intermediate System to Intermediate System (ISIS) operates in parallel with OSPF but it is used much less in networks today. Border Gateway Protocol (BGP) [53] is not examined as it provides inter domain connectivity.

2.9.2 Open Shortest Path First

This section overviews OSPF [60] protocol, concentrating on failure recovery. A lot of the material is covered in Pasqualini et al. [64].

One of the most common intra-domain routing protocols in IP networks is OSPF. The Hello protocol is used for the detection of topology changes. Each router periodically emits *Hello* packets on all its outgoing interfaces. If a router has not received *Hello* packets from an adjacent router within the *RouterDeadInterval*, the link between the two routers is considered down. When a topology change is detected, the information is broadcasted to neighbours via Link State Advertisements (LSA).

Each router maintains a complete view of the OSPF area, stored as an LSA Database. Each LSA represents one link of the network, and adjacent routers exchange bundles of LSAs to synchronise their databases. When a new LSA is received the database is updated and the information is broadcasted on outgoing interfaces.

Routes calculation: configurable cost values are associated to each link. Each router then calculates a complete shortest path tree. Only the next hop is used for the forwarding process.

The Forwarding Information Base (FIB) of a router determines which interface has to be used to forward a packet. After each computation of routes, the FIB must be reconfigured.

Typically a router failure is detected within 30 to 40 seconds, while a link failure can be detected using hardware detection in under 10 seconds.

2.9.3 Multi Protocol Label Switching

There is a data plane and a control plane to Multi Protocol Label Switching (MPLS) [66].

The data plane is analogous to ATM virtual paths (VPs) and virtual circuits (VCs) as MPLS is a connection-oriented by means of Label Switch

Paths (LSPs). Instead of using ATM VP/VC identifier, MPLS uses 20-bit labels to identify a connection. This allows up to a theoretical maximum of 1 million LSPs per interface. This number can be much higher, since MPLS additionally supports label stacking, which is comparable to the use of ATM VP/VCS. However, contrary to ATM, label stacking is not limited to one level, but could be extended to an arbitrary depth.

In core networks, the endpoints and intermediate nodes of the LSP will typically be IP routers. The endpoints of the LSP are called label edge routers (LER), whereas the interior nodes are called label switching routers (LSRs).

MPLS control plane is responsible for establishing an MPLS connection, which can be done in several ways. The first method is to manually configure the LSP in each of the involved network elements using a network management tool. This is the equivalent of an ATM PVC and does not require the use of an MPLS control plane. Another model is to establish the LSP using MPLS signaling protocols, equivalent to ATM SVCs, or ATM soft-PVCs. Two commonly used MPLS signaling protocols are the Label Distribution Protocol (LDP) [12] and Resource Reservation Protocol with extensions for LSP establishment, or traffic engineering (RSVP-TE) [15]. Traffic engineering is about arranging traffic flows around the network to avoid congestion. RSVP-TE provides mechanisms for doing that. The control plane is based on network layer protocols and thus uses IP addressing.

MPLS is the best of both IntServ and DiffServ architectures. As in DiffServ traffic is aggregated into classes (forward equivalence class or FEC), but there is an unlimited amount of classes or labels in MPLS. A class can be introduced and all that needs to be done is to circulate the new label

and its level of QoS around the network. The labels are distributed using a Label Distribution Protocol (LDP).

Unlike DiffServ MPLS is not per hop, but similar to IntServ it sets up a path, called a label switch path (LSP) for an aggregated flow through the network. This way MPLS achieves the scalability of DiffServ and the path set up of IntServ.

MPLS is the most likely to be used to provide QoS. Specifically, providing voice QoS requires setting up a LSP for voice traffic only once. All voice traffic will be aggregated into that class and so all packets will follow the same path through the network.

The main benefits of MPLS are [63]:

- Similar to ATM, MPLS is connection oriented which provides aggregation and segregation of customer traffic
- Scalable technology as label stacking can be performed to arbitrary depth
- It can tunnel any IP or non-IP protocol
- Inherent support for CoS and QoS
- Many recovery options with performance close to SONET/SDH

2.9.4 Overview of Ethernet

Ethernet is a data link protocol, where the concept of connection is replaced by use of a single Ethernet broadcast domain. The network elements are self learning Ethernet switches that send frames to their final destination.

There is no signaling involved, nor is there a need for any higher layer protocol to be involved, such as a network protocol for label distribution in MPLS. This is because Ethernet uses the Rapid Spanning Tree Protocol (RSTP) to eliminate loops and recover in an event of a failure.

Ethernet has traditionally been a LAN technology, however, recently Meddeb [56] proposed Ethernet for WANs due to various advantages such as low cost and ease of use of the protocol. However, Ethernet needs significant alterations for it to become a carrier grade WAN protocol.

Ethernet is rapidly migrating from being a LAN technology to access network. It was described in an earlier section. The protocol is considered to be one of the two protocols for access networks.

There are advantages and disadvantages of using Ethernet in the access network [56]. The advantages are:

- Granular access speed ranges from 10 Mbps to 10 Gbps
- Efficient multipoint connections support
- No or little training and learning from customers
- No frame termination needed at the CPE
- Low deployment costs due to ubiquitous characteristics of Ethernet

The disadvantages include:

- Contention with proven and widely deployed WAN technologies
- Need for compromise between carrier grade and SLA support versus simplicity and flexibility of Ethernet
- Interoperability concerns between resilience at other layers and spanning tree protocol

- 802.1p frame prioritisation does not necessarily provide end-to-end QoS
- Potential scalability and latency issues due to large numbers of MAC addresses
- OAM protocols and tools need to be modified substantially

This thesis investigates resiliency, so for Ethernet, RSTP is the key focus.

2.9.5 Rapid Spanning Tree Protocol

Unlike routers, switches and bridges do not age-out old packets. This causes problems such as unicast frame duplication and multicast frame multiplication. When two or more switches are connected in a loop, they can multiply multicast frames, sending them round infinitely. The inability to tolerate active loops is a fundamental restriction of bridged networks.

Redundant connections are essential when designing high-availability systems, and that is where Rapid Spanning Tree Protocol (RSTP) IEEE 802.1D [9] is used. By blocking redundant connections, RSTP enables a single/primary data path between nodes, creating a tree topology, which is a graph with no loops with no disconnected nodes. If a device or a link failure causes this primary path to become unusable, RSTP will enable a secondary path. The predecessor of RSTP is STP [9], which has very slow re-convergence time of 30 to 50 seconds compared to RSTP with several seconds re-convergence.

2.10 NGN Transport and Signaling Protocols

There are many types of signaling involved in a VoIP network. It is important to discuss these protocols because they contribute to the reliability metrics discussed in Section 2.7.2. Three categories of signaling are distinguished, voice transport, transport of SS7 signaling, inter-gateway signaling within the softswitch model, and service layer signaling.

Voice Transport and Media Signaling Voice packets are carried over IP using Real Time Protocol (RTP) [71] which uses UDP. Real Time Control Protocol (RTCP) [72] is used to control the voice packets. RTCP mainly provides quality related feedback from the receivers to the transmitter about the state of the RTP flow. Network operator can use RTCP to monitor quality of voice calls and network problems.

Stream Control Transmission Protocol The function of Stream Control Transmission Protocol (SCTP) [76] is to carry SS7 signaling between a signaling gateway and a softswitch over the NGN. The protocol is designed to have the speed of UDP and reliability of TCP as provided by SS7 signaling in PSTN.

MGCP and MEGACO/H.248 Media Gateway Control Protocol (MGCP) [13] and the more popular MEGACO/H.248 [29] are the two master-slave protocols used between a softswitch and a media gateway.

H.323 ITU-T Recommendation H.323 [2] describes H.323, which is a signaling protocol for VoIP. The architecture consists of end terminals, gateways, gatekeepers to control functionality, and multipoint control units which are used for multi-terminal interaction.

H.323 is a complex protocol suit because it consists of many different protocols for doing various functions. Communication involves many message exchanges, which complicates and retards the speed of interaction considerably. H.323 is used extensively in today's enterprise VoIP in a variety of products and so forms part of NGN.

Session Initiation Protocol Session Initiation Protocol (SIP) [38], is as expressive a signaling protocol as H.323. SIP has a simpler architecture, message syntax, and it is easily extensible. SIP is based on a client-server model, in which a client sends SIP requests and a server responds to SIP requests.

SIP messages are plain text messages with a small number of headers. This makes it simple to communicate, around 5 messages accomplish call setup whereas it takes around 20 messages for H.323. Plain text messages are human readable, making it much easier to find errors than the Abstract Syntax Notation 1 (ANS.1) notation used in H.323.

Simplicity and extendibility of SIP make it attractive for VoIP use. However, it may need extensions to handle security and fault tolerance with distributed SIP servers. The extensions are considered far less formidable than with H.323.

2.11 Quality of Service

Quality of Service (QoS) is a very important aspect that separates NGN from the current best effort Internet. A particular type of traffic is classified into a service class (CoS) and is treated according to that classification

within the network. Prime examples are network management traffic is usually assigned to the best class and voice traffic is assigned to the second highest class. Typically the lower the CoS the worse its performance in the network, which may include higher end-to-end delay, packet drop, and so on.

There are three main methodologies for solving the QoS problem. First is reserving resources on the path to destination prior to the session. Second is classification of traffic into different classes that are treated with a different QoS in the network. Third,, bandwidth overprovisioning of network elements to avoid congestion of all traffic is the simplest but a more costly solution, since all sections of the network will be underutilized especially during the non-failure operation.

2.11.1 Integrated Services

Integrated Services (IntServ) [21] is a per flow architecture, where each individual traffic flow reserves resources on its path to the designation to satisfy required QoS. Resource ReserVation Protocol (RSVP) is used to reserve the path resources and admission control. The IntServ architecture is usually considered too complicated and demanding on the network and so it is scarcely used by itself.

2.11.2 Differentiated Services

Differentiated Services (DiffServ) [20] uses the concept of aggregating traffic flows into a set amount of classes. The marking and shaping of the traffic is achieved by using the DiffServ Code Point IP field. The Diff-Serv architecture is more scalable than IntServ as only little amount of

state needs to be maintained by the network. The main obstacle to DiffServ is it is on a per hop basis and so a voice packet may take a different route than the one in the same session thus changing the delay, jitter, and packet loss. Thus DiffServ is not suitable for sessions that require fixed QoS, such as VoIP or video conferencing.

A combination of IntServ and DiffServ is also possible, such as IntServ operation over DiffServ networks, defined in RFC 2998 [18].

2.11.3 Other Solutions

There are many QoS solutions based on QoS and constraint-based routing. Constraint-based routing is outside of the scope of this thesis, because it is a large separate field of research.

One interesting example is resilient differentiated QoS (RD-QoS) [14], which is a combination of DiffServ and routing mechanisms. The approach argues that resilience to failure of different traffic classes is as important as QoS classes. Four resilience classes are proposed, with the first most resilient class having 10 – 100 ms recovery time and a pre-established recovery path among other specifications. As the classes progress recovery time increases together with other parameters reflecting lower resilience needs. RD classes are proposed to be signaled together with QoS signaling by extending RSVP. The classes are mapped to MPLS recovery schemes and they can also be mapped to other QoS architectures such as DiffServ.

2.11.4 Metrics

The simple transport level metrics important to assess VoIP quality are ETE delay, jitter, packet loss, and packet error rate.

ITU-T Recommendation Y.1541 [5], summarised by Seitz [73], specifies numerical values to be achieved for each of the key performance parameters defined in Y.1540 [4], on international IP network paths. A path extends from the network interface of the source user to the network interface of the destination user. Customer premise equipment is not part of the path. The specified values are grouped in a number of distinct QoS classes to establish a practical basis for communication between end users and network providers, and among providers, on the quality to be supported in ETE IP paths.

Table 2.3 specifies Y.1541 performance objectives and QoS classes. Each QoS class creates a specific combination of bounds on a subset of the performance values. Classes 0 and 1 place upper bounds on packet transfer delay and packet loss. They also limit packet delay variation. Classes 2 and 3 place upper bounds on packet transfer delay and packet loss, but do not limit packet delay variation. Classes 0 and 2 differ from classes 1 and 3 in their packet transfer delay objectives. Class 4 limits packet loss and provides a very soft upper bound on delay. Y.1541 also defines an unspecified class (class 5) that provides no specific performance guar-

Parameter	QoS Classes					
	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5
ETE delay	100 ms	400 ms	100 ms	400 ms	1 s	-
ETE jitter	50 ms	50 ms	-	-	-	-
Packet loss	0.001	0.001	0.001	0.001	0.001	-
Packet error rate	0.0001	0.0001	0.0001	0.0001	0.0001	-

Table 2.3: IP QoS class definitions from ITU-T Recommendation Y.1541 [5]

antees. The value for the single packet error ratio objective was chosen to ensure that packet loss is the dominant cause of defects presented to upper layers.

The ETE delay objectives of Class 0 and 2 will not always be achievable on long paths [73]. If only propagation delay is taken into account the maximum length of a path satisfying the delay of 100 ms is $0.1 \text{ s} \times 2 \times 10^8 \text{ m/s} = 2 \times 10^7 \text{ m} = 20,000 \text{ km}$. Due to node delays this length is unachievable. However, for New Zealand this is not a problem as the maximum distance from the bottom of the South Island to the top of the North Island is less than 2000 km.

Table 2.4 identifies acceptable VoIP mouth-to-ear delays in terms of user experience, suggested by the ITU-T Recommendation G.114 [1], but has no relation to the E-model. The delay includes all possible delays incurred on the path between two users.

Delay will always be present in a VoIP path. The two major classes of delay are codec related delay and network delay [28, 57]. Codec related delay is a sum of codec processing, de-jitter buffer, serialisation, packetisation delays. Network delay applies to all packets in the network (provided they are in the same class of service). The major contributors to network delay are interface queuing, transmission and node processing delays [57].

In addition to the formal specifications there are two less formal criteria that a VoIP service may satisfy. One of these an interruption or a fail-

Delay (ms)	Acceptable Conditions
0 - 150	Intracountry calls
150 - 400	International calls
> 400	Calls with satellite hop

Table 2.4: Recommended ETE mouth-to-ear delay for VoIP [1]

ure below the network layer should not be longer than 200 ms, in order to prevent failure detection at the network layer following a timeout. This time is agreed upon in the network community [28] not only for VoIP but other real time carrier grade services. The second criteria concerns purely VoIP because it is a direct consequence of the way the call session is configured and maintained. According to Johnson [46], a VoIP call session is terminated if any activity between the calling parties is absent for longer than approximately 2 seconds. This can be regarded as a maximum time for failure recovery. If a failure cannot be detected and repaired in less than about 2 seconds all VoIP call sessions currently active (and any new attempts) will be terminated causing extremely adverse consequences.

2.12 Summary

A broad review of literature has been presented. Reliability, resiliency, routing protocols, and quality of service that are most significant for my research have been described in more detail. Chapter 3, analyses these aspects with respect to the aim of my work.

Chapter 3

Research Question

3.1 Research Question

The question this thesis tries to answer is:

How well can current IP and Ethernet technology fulfil NGN network and service reliability and resiliency requirements?

The three pieces of current technology investigated are the physical topology, datalink layer protocols, and network layer protocols. NGN is investigated in two sections, the core network and the access network. Figure 3.1 shows all the major parts of the ETE network topology. The focus of physical reliability and resiliency experimental work is the core and access network.

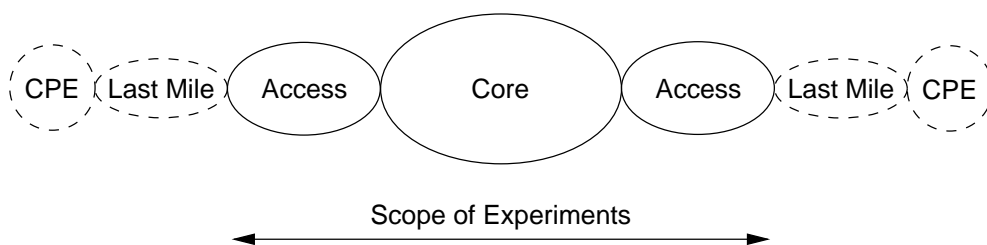


Figure 3.1: Network clouds and focus of experiments

Physical topology is the basis for other layers in the network protocol stack. The configuration of network elements and their types and quality determine physical layer reliability. My thesis investigates physical reliability of the core and access network separately, because they have different topologies and requirements. The results of these analyses are combined to develop an ETE reliability analysis.

Analysis in Chapter 2 shows the important aspects related to answering the research question. Section 2.7 identified physical network reliability as the most relevant issue and identified two measures in particular, DPM and end-to-end downtime. Section 2.5.1 showed that the most important aspect of resiliency is network accessibility. The best metrics to assess resilience were shown in Section 2.11.4 to be ETE delay, jitter, packet loss, and packet error rate. The selection of the protocols for the study is provided in Section 2.9. OSPF and MPLS are selected for comparison in the core network and Ethernet and MPLS in the access network. Section 2.11.4 describes the resulting end-to-end network solution assessment in terms of voice quality. Two ITU-T recommendations and some less formal measures are used to compare network metrics. The E-model in Section 2.6.1 is identified to be the best model to assess VoIP voice quality against that of the PSTN.

3.2 Justification

The focus of this research is determined as physical reliability, protocol resiliency, and assessing the ETE quality of NGN. These aspects are detailed in the following sections. In this section I determine the justification for this specific focus.

3.2.1 Network Reliability

Physical network reliability is selected in Section 2.7 out of a number of reliability issues, because it constitutes a large part of the reliability space.

Appendix C describes techniques for calculating network reliability. Using these techniques, the core and access topologies are analysed separately and combined for a complete reliability measure.

The metric used for reliability is the ETE downtime measure, which is identified in Section 2.7.2. For theoretical computation of reliability ETE downtime equivalent to the DPM measure. The two metrics are subtly different when using real network data because DPM counts user demands not served.

Reliability, which is regarded equivalent to availability in networks, is primarily a measure of network accessibility (Section 2.5.1), which constitutes the main part of quality concern. The secondary concern is connection continuity, because it is important that the call progresses without interruptions and is of consistent level of quality throughout the call. Connection continuity is indirectly measured by simulation rerouting time and other metrics indicating the level of ETE quality. The quality concern is investigated at the end of Chapter 5 as part of the ETE network analysis. The reliability issues identified in Section 2.5.1, but which are not part of the scope of this thesis are as follows. Routing speed is a customer related issue and does not have much bearing on network reliability. Connection setup reliability is not directly related to reliability. Routing Reliability is a concern of the service logic.

3.2.2 Resiliency

As Section 2.8 described, resiliency of higher layer protocols is very important because it effectively translates the physical layer reliability to higher layer. Examining and comparing resiliency of datalink and network layer protocols is therefore of great importance in answering the research question.

3.2.3 Experiment Analysis Steps

Hassan and Jain [40] in Chapter 4 describe a methodology for a systematic simulation study. The steps are as follows.

1. Define the objective of the study
2. Design reference network model and select fixed parameters
3. Select performance metrics
4. Select variable parameters
5. Construct the network model and set fixed parameters in simulation software
6. Configure simulation software to produce relevant performance data
7. Execute simulation program and collect performance data
8. Present and interpret results

For the purpose of the thesis these steps are modified and combined with Law and Kelton [50] general advice on performing simulation analysis.

3.2.4 Simulation

Law and Kelton [50] define Discrete Event Simulation (DES) as:

"Discrete event simulation concerns the modeling of a system as it evolves over time by a representation in which the state variables change only at a countable number of points in time. These points in time are the ones at which an event occurs, where an event is defined to be an instantaneous occurrence which may change the state of a system."

DES is a method used to model real world systems able to be decomposed into a set of logically separate processes autonomously progressing through time. Each event must take place on a specific process, and must be assigned a logical time (a timestamp). The result of this event can be a message passed to one or more other processes (including, if required, the process on which the event occurred). On arrival at this other process, the content of the message may result in the generation of new events, to be processed at some specified future logical time.

The principle restriction placed on DES is that an event cannot affect the outcome of a prior event, that is, logical time cannot run backwards.

DES is used because this is the only accurate method to investigate the performance of a network in response to a failure. An alternative approach to DES is analytical modeling, which cannot accurately predict the outcome when there is a complex mix of traffic from multiple traffic sources and failure consequences need to be analysed. Such models are only useful for calculating steady state of a network, when all processes have settled down.

For all my simulations I used Opnet, because it is one of the best DES simulators and it contains models of the majority of the NGN technology required. The other major alternative network simulator is NS 2, which was ruled out because it needed several non-trivial extensions to perform the required simulations.

3.2.5 Failure Modes

From Table 2.1 we see that more than two thirds of unplanned failures are individual link failures, whereas router-related failures contributes only 16.5 %. My thesis discards the optical and unspecified shared failures, because the optical layer is part of the physical layer, which is not investigated. Unspecified failures are difficult to simulate because their cause is not determined or classified.

According to Table 2.1, simultaneous failures of core routers is sufficiently low. The probability of a simultaneous failure of two independent routers is the probability of one router failure times the unavailability all squared. According to Miller [57], typical router availability is between 99.99 and 99.9999 %. Unavailability = $1 - \text{Availability}$, so the worst unavailability for a router is 0.01 %. Using the figure of single router failure probability of 16.5 % from Table 2.1, simultaneous router failure is $(16.5 \% \times 0.01 \%)^2 = 0.00027 \%$, provided router failures are independent events. Such an event is extremely negligible and therefore not examined.

3.2.5.1 Other Metrics

The main network metrics used for measurement and identified in Section 2.11.4 are ETE, because these are the most meaningful in all models and standards. The ETE metrics depend on delay, jitter, packet loss, and

packet error rate measurements. ETE delay is the most significant of the metrics in terms of influence on quality performance. Packet loss as a result of the network reconfiguration is a significant measure of protocol resiliency. Random packet loss is not simulated explicitly, but it is used in the E-model calculations to indicate the range of values of packet loss that can be tolerated.

Per network element metrics are not very useful for my experiments. These are processing delays, propagation delays, and queuing delays. These are analysed briefly to justify their simplification in Chapter 4.

3.2.6 End-to-end Quality Analysis

3.2.6.1 Metrics analysis

Network performance is compared to performance of ITU-T Recommendation Y.1541 [5] QoS classes. VoIP user experience is compared to ITU-T Recommendation G.114 [1] delay thresholds. Both of these recommendations were overviewed in Section 2.11.4. ITU-T's E-model [3] (Section 2.6.1 is used to compare VoIP quality more accurately.

3.2.6.2 VoIP Analysis with E-model

The E-model was introduced in Section 2.6.1 and it was shown that the E-model is the most useful tool to assess and compare the VoIP call quality to that of PSTN. This analysis is critical to perform because the ultimate goal of NGN is to provide VoIP service that is as good or better than PSTN service. If that goal cannot be achieved deploying such a network will be of little benefit to the end users.

3.2.7 Other Issues

There is a number of issues in Chapter 2 that are determined to be out of the scope of this thesis. Not every piece of NGN topology in Figure 3.1 is considered, CPE and last mile access are not explicitly examined. However, they do appear in the analysis using well known delay figures. Section 2.2.3 presents security as an important aspect of NGN, but it is a large separate area of research that is not considered.

QoS architectures and CoS issues in Section 2.11 are also not part of this work because it is a separate research field. My analysis assumes that voice is allocated the second highest CoS after network management traffic. Signaling protocols in Section 2.10 are important, but as Collins [28] shows, they utilise an insignificant amount of capacity relative to other traffic and therefore signaling traffic can be ignored.

PSTN aspects in Section 2.1 are not investigated further, except for a comparison with PSTN voice quality, which is used as a benchmark in the E-model for VoIP assessment. Migration scenarios and economic analysis identified in Section 2.2 are important issues during the transition from PSTN to NGN. However, these areas are separate areas of research as economic analysis is a non-technical area, while migration scenarios deal with interworking which need a different type of analysis than one used in this thesis.

3.3 Core of the Thesis

The research question is answered in Chapter 4 and Chapter 5. Chapter 4 analyses a representative ladder-based core network. The physical reliability of the network is analysed and the resiliency of OSPF and MPLS

is compared. In a similar manner Chapter 5 analyses a representative tree-based access network. The physical reliability of the access and of the total end-to-end network is analysed. The resiliency of Ethernet and MPLS was compared, followed by the analysis of the end-to-end voice solution.

Chapter 4

Next Generation Core Network

4.1 Introduction

A core or backbone network provides high bandwidth and low latency transport that carries aggregated traffic between all the top level access networks. A core network extends over a large region, typically a country. The network needs to be very reliable because consequences of a failure may be significant. However, core networks are not fully redundant due to prohibiting factors such as geographical and economic factors.

This chapter has two main goals. First, is to calculate reliability of the physical network topology. Second, is to compare resilience performance of the two most popular traffic routing protocols, OSPF and MPLS. The core network study does not include the access aggregation, which is the subject of Chapter 5.

A number of results were established in this chapter. To satisfy the physical reliability requirement of 0.9999, the core network was found to need individual router reliability of at least 0.999 and link reliability of 0.9999, or vice versa.

The comparison of the two protocols using simulations showed that OSPF experienced 6 to 40 seconds of rerouting delay resulting in significant packet loss. Rerouting delay for MPLS was found to be around 13 milliseconds, causing no interruption at network layer. The interruption can be measured by the length of the discontinuity of received traffic by the callee. The difference between the end-to-end delay after failure and before failure is around 1.5 milliseconds for both protocols, which is under 7 %. The jitter delay difference was even more negligible at around 0.002 milliseconds, which is under 5 % of total jitter. These delay and jitter degradations have no practical effect on voice service quality.

Section 4.2 presents the physical topology of the core network and its characteristics. Section 4.3 analyses the physical reliability of the topology. Configuration of the core experiments, including the objective, variables, failure modes, protocols used, and simulation execution, are covered in Section 4.4. Section 4.5 presents my simulation results. The summary in Section 4.6 completes the chapter.

4.2 Physical Topology

This section looks at the physical network topology that is used in the core network experiments. Redundancy and physical characteristics are described in detail.

4.2.1 Ladder

Figure 4.1 shows the ladder configuration of core routers as the base case topology for this study. This is a typical architecture for an elongated country such as New Zealand. The topology is a balance between an all

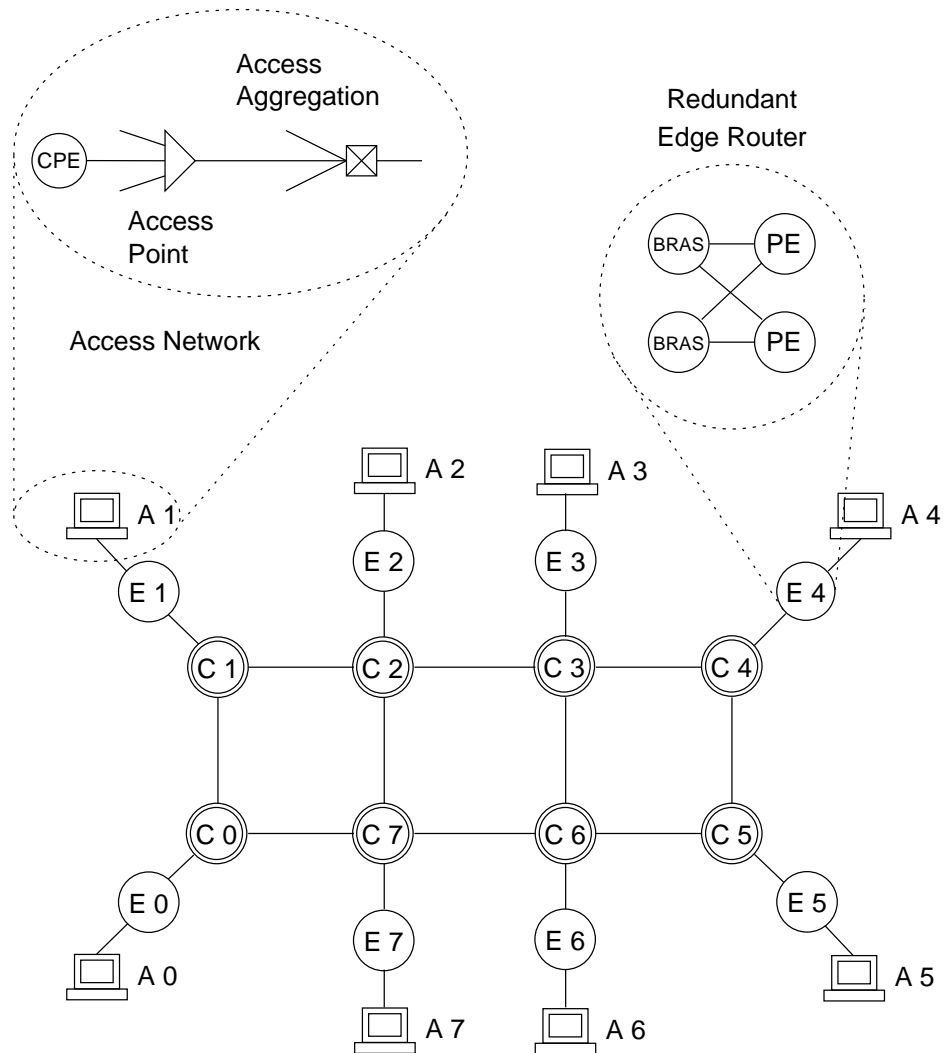


Figure 4.1: Base ladder topology. Access network and edge router complexity are simplified.

connected (also called a full mesh) topology and a topology with minimal connectivity, due to a tradeoff of two conflicting constraints, financial cost and network performance. The ladder topology is the middle ground that is the starting point of an NGN core for an elongated country. The network shows 8 core routers C 0 – C 7 and 8 access networks A 0 – A 7 connected by edge routers E 0 – E 7.

For the core network study all of the access network is replaced by a single VoIP endpoint. Figure 4.1 shows what the access network looks like in a real network. It consists of 100s or 1000s of users connected to an access point using some last mile technology such as ADSL. Aggregation of the access points occurs in an access aggregation network, which may be a tree of Ethernet switches with the root switch connecting to the edge router. Chapter 5 studies the access aggregation in detail.

The edge router on Figure 4.1 is also a simplified version of the configuration in a real network. A typical redundant node consists of a Broadband Remote Access Servers (BRAS) and Provider Edge (PE) routers. A BRAS aggregates user subscriber lines and provides functionality, such as billing and traffic shaping and monitoring. A BRAS may also incorporate PE function and SBC function. A PE router aggregates the traffic, which then goes to a core router (sometimes called Provider (P) router). For reliability purposes the PE router is often deployed in a redundant configuration and each BRAS connects to two PE routers. If one PE router fails, the redundant one can still be used. For the purposes of this study such complexity can be simplified and a single reliability figure used for the node. In a similar manner this thesis ignores the redundant configurations and other complexities of network equipments, such as routers and switches.

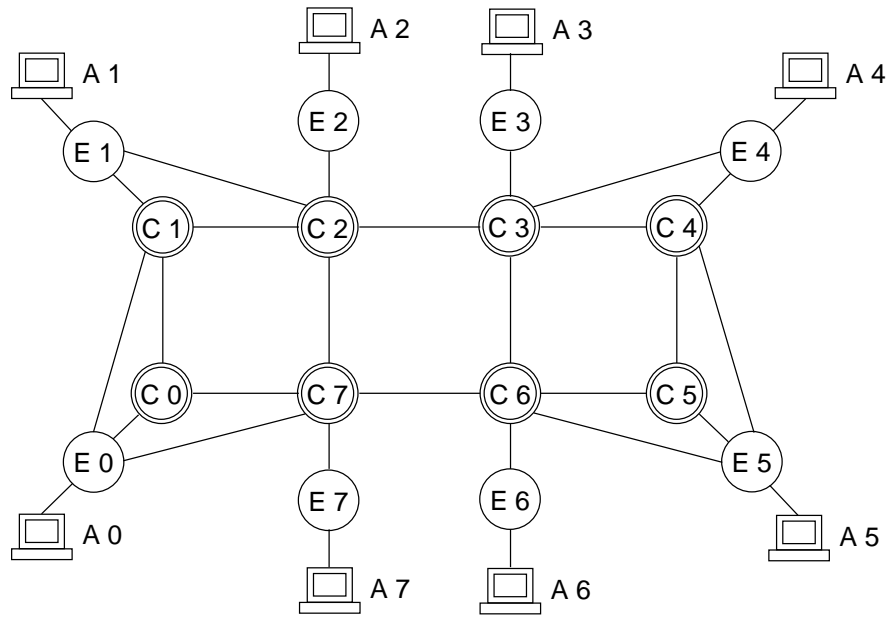


Figure 4.2: Redundant ladder topology

The ladder network contains resilient links between core routers, but no resilient links between edge routers and the core. A failure of a single core router will cause permanent failure of exactly one access network. For instance, if C 2 fails access network A 2 would be unreachable, but all the other access networks would remain reachable after route reconfiguration. For example, A 1 may need to change its route to A 3 if the original route went through core router C 2. To make the ladder more resilient to core router failures some redundant links need to be added between edge and core routers.

4.2.2 Redundant Ladder

My experiments use the redundant ladder topology shown in Figure 4.2. This is a more realistic scenario because it provides a level of resiliency to edge routers. There are many choices for the number and location of redundant links ranging from zero redundancy to a full connectivity be-

tween edge and core routers. A full mesh between core and edge routers is the best option from failure resiliency perspective. However, financial constraints prohibit such infrastructure. There is an equilibrium point between minimising costs and satisfying resiliency requirements. Figure 4.2 shows a more realistic compromise than a full mesh network. E 0 has two redundant links to C 1 and C 7, whereas E 1 has only one redundant link to C 2. E 2 and E 7 do not have any redundancy. The purpose of the asymmetry was to replicate the real world situation where the cost of redundant links is different depending on various factors such as geography, distance, technology and so on. Also this makes the results more interesting compared to a simpler completely symmetric topology. The second half of the network is horizontally symmetrical to the first. The redundant topology illustrates the benefits of redundancy for the corner edge routers E 0, E 1, E 4, E 5. The four middle edge routers E 2, E 3, E 6, E 7 demonstrate the effects of lack of redundancy.

4.2.3 Physical Characteristics

Link length between adjacent core routers is approximately 250 km. Link propagation speed is set to speed of light in fibre or copper, which is approximately 2.00×10^8 m/s. All routers have IP forwarding rate and MPLS datagram switching rate set to 500,000 packets per second (pps).

NGN is likely to use a Differentiated Services architecture, which allocates traffic to different classes of service depending on the type of traffic. My thesis investigates VoIP, which is the second highest class of service after management data. An interesting problem is what happens to all other classes during various failure conditions. For example, each class may have a different queue type, which would affect its performance in

different ways. My thesis leaves the problem of how different classes are affected for future work.

A full mesh of call session demands is configured between each access network. Thus, A 0 has seven connections to each of A 1 to A 7, A 1 has seven connections to each of A 0 and A 2 to A 7 and so on. A single call is a G.729a call with parameters of 100 pps or 30,400 bps.

Two link capacities are used in this analysis. During the analysis of failure recovery a link capacity of just 700,000 bps is used. This is the rate at which during any failure the maximum link utilisation of VoIP traffic is approximately 90 % to 95 %. The configuration of my experiments is such that there are no losses at any time due to insufficient link capacity through capacity planning. Since failures are easily simulated the best and simplest way to decide what capacity to use is to decide what is the maximum target utilisation of any link after a failure. A failure capacity of 95 % is used because the network provider seeks to minimise the cost of the network and therefore would like to minimise the link capacities.

A less accurate alternative strategy would be to dimension the network so that no more than a certain proportion of network capacity, such as 40 % link capacity, is used during non-failure operation. That way during failure there is 60 % spare capacity for rerouted traffic. However, this does not guarantee that during failure utilisation does not exceed 100 % and therefore packet losses would occur.

When approximating performance during failure conditions typical core network link capacity of 10 Gbps (9,510,912,000 bps), with a static background load of 9,510,212,000 bps in order to have the already calculated capacity (700,000 bps) for the seven calls from each access network to all others. Background network traffic is not rerouted during network

failure, it provides a constant load on the links and the routers.

4.3 Physical Reliability Analysis

Reliability at the physical layer is important to calculate because everything relies on physical connectivity. Each layer added above the physical layer potentially degrades the reliability by adding more failure modes. Physical layer reliability should ideally exceed the required reliability at the very top, service, layer. However, in practice this is not entirely true because failure recovery mechanisms at different layers can mask failures below that layer. Such behaviour is in accordance with Saltzer's end-to-end argument [68] that advocates that functions should be moved as close to the application in the network stack as possible. For example, as Section 2.11.4 mentioned, if the physical failure duration is under 200 ms at data link layer, network layer will not detect the failure experiencing very small increase in delay and jitter.

Appendix C, Section C.8 presents an implementation of an algorithm for calculating the theoretical reliability between two endpoints. The calculation suffers from the problem of overestimating the reliability, because all possible paths between a source and a destination are used in parallel and some of them may not be disjoint. Thus, a failure could affect more than one path. However, the theoretical reliability is useful as an upper bound on reliability.

A more practical approach is to consider only two paths in parallel, but choose maximally disjoint, longest paths. This gives a realistic lower bound on the reliability, whereas the absolute lower bound is given by

Link R	Router Reliability				
	0.99	0.999	0.9999	0.99999	0.999999
0.99	0.989046	0.997008	0.997542	0.997592	0.997597
0.999	0.996004	0.999880	0.999969	0.999975	0.999975
0.9999	0.996520	0.999958	0.999999	0.9999997	0.9999997

Table 4.1: Effective reliability using two parallel paths, P_1 and P_2 , between E 0 and E 4, given router and link reliabilities

a single path. Note that increasing the number of paths for the parallel reliability calculation results in improvement in the overall reliability.

From the topology in Figure 4.2, the two maximally disjoint and longest paths are $P_1 = \{E\ 0, C\ 0, C\ 1, C\ 2, C\ 3, E\ 4\}$ and $P_2 = \{E\ 0, C\ 7, C\ 6, C\ 5, C\ 4, E\ 4\}$. For simplicity, only two distinct reliability figures are used, router reliability r and link reliability l . Reliability of P_1 and P_2 happen to produce the same expression, $T_1 = T_2 = T = l^5 r^6$, using notation for path reliability from Section C.3. Therefore, assuming independent failures, the effective reliability of the two paths in parallel is

$$\begin{aligned}
 R &= P(T_1 + T_2) \\
 &= T_1 + T_2 - T_1 T_2 \\
 &= T(2 - T) \\
 &= l^5 r^6 (2 - l^5 r^6)
 \end{aligned}$$

Table 4.1 shows the effective reliability between E 0 and E 4, given router reliability r and link reliability l .

Depending on the requirement for the reliability of the core network, Table 4.1 can be used to check what values of r and l satisfy the total reliability. For example, if the core is required to have reliability of 0.9999 we can see that router and link reliability must be ≥ 0.9999 .

As a reference, for small reliability values of $r = 0.999$ and $l = 0.99$, the

Access R	Core Reliability				
	0.99	0.999	0.9999	0.99999	0.999999
0.99	0.970299	0.979120	0.980002	0.980090	0.980099
0.999	0.988021	0.997003	0.997901	0.997991	0.998000
0.9999	0.989802	0.998800	0.999700	0.999790	0.999799
0.99999	0.989980	0.998980	0.999880	0.999970	0.999979
0.999999	0.989998	0.998998	0.999898	0.999988	0.999997

Table 4.2: The ETE reliability, given the reliability of the access network and the core network, $ETERel = AccessRel^2 \cdot CoreRel$

theoretical reliability between E 0 and E 4 is 0.999946. For $r = 0.9999$ and $l = 0.99$, which is still a low reliability, the result exceeds 0.999999. If only the theoretical reliability is used, r and l would be greatly underestimated due to using 24 paths in parallel for the calculation.

To get an estimate of the target reliability of the core network, the NGN ETE reliability should be considered. The goal of NGN from the outset was to make it at least as reliable as the PSTN, which has an ETE reliability of 99.93 % or 0.9993 [28, 46]. However, there is an argument that users may be willing to accept lower reliability in return for more functionality. For example, mobile phone users accept lower voice quality because they are not bound to a land line. However, for an equivalent PSTN service, such as VoIP, NGN must provide the same or better reliability (and voice quality). Table 4.2 shows the ETE reliability, given core and access network reliabilities. The expression for ETE reliability is $ETERel = AccessRel^2 \cdot CoreRel$. Last mile access network is not explicitly included, but, as Section 5.3 shows, it can be incorporated as part of the access network. Despite its simplicity the expression provides a good estimation of the ETE reliability without using additional parameters. The bold region in Table 4.2 shows which pairs of core and access network reliabilities satisfy the target minimum ETE reliability requirement of 0.9993. The region indicates that the core and access reliabilities

must be at least 0.9999. The table will be further used in Section 5.3. Corresponding to that requirement, Table 4.1 shows that either $r \geq 0.999$ and $l \geq 0.9999$, or vice versa.

4.4 Experiment Setup

The core network study has two subparts, which will be referred to as the Small Link Capacity (SLC) and Large Link Capacity (LLC) studies or experiments. The SLC was analysed because simulations are faster and the results are much easier to analyse in terms of the changes in measured parameters, such as ETE delay, packet loss and so on. The only difference between the two cases is that the LLC estimates the ETE delay and jitter more accurately because the scenario is much closer to the real network. Rerouting times and packet loss are still the same as in the SLC experiments because the amount of dynamic calls have not changed, only that static background traffic has been introduced, which increases queuing delays.

4.4.1 Performance Metrics

All the experiments use several specific metrics: rerouting time, packet loss as a result of the rerouting, end-to-end (ETE) delay and jitter. All metrics are ETE. Queuing delay is a potential metrics, but in these experiments measurements show that it can be approximated by a constant time period.

The queue type is a default first-in-first-out with 16 MB buffer, which is a typical default size for the router of such size [57]. Queuing delay and buffer utilisation were measured for SLC and LLC configurations.

For the SLC scenario queuing delay between any two adjacent core routers is approximately a constant 1 ms (before and after failure) with uncertainty range of 0.06 ms, while queue buffer utilisation is around 1 KB with uncertainty range of 0.03 KB. Between access network and edge router (for example, A 0 – E 0) and between edge router and core router (for example, E 0 – C 0), the delay is approximately a constant of 0.010 ms with uncertainty range of 0.001 ms. Queue buffer utilisation is only 0.2 KB, with uncertainty range of 0.07 KB. In comparison with propagation delays, such as the delay between two adjacent core routers (250 km x speed of light in fibre or copper = 1.25 ms) the latter queuing delay is negligible.

For the LLC scenario a large static background utilisation is applied on all links. According to Opnet documentation and my results, the effect of this utilisation is solely to increase queuing delays, which simulates the delays in real networks without having multiple gigabytes of explicit data, which would make simulation time prohibitively large. The effect of the background traffic is that all queuing delays are approximately a constant of 3.8 ms, with a buffer utilisation of approximately 5 KB under steady state condition.

Buffer utilisation is under 0.031 %, which is extremely low. According to queuing theory if queue utilisation is less than 1 % then the delay is effectively constant and the exponential behaviour needs not be analysed. The utilisations and queuing delays are always constant and so my study does not investigate queue behaviour explicitly, but only as a constant component of the ETE delay. As a consequence of low queuing delay and utilisation, there are no random packet losses in my experiments.

4.4.2 Failure Modes

My experiments investigate single link, double link, and single node failures, which are justified in Section 2.4. Table 4.3 shows the link and node failure modes used in the core network experiments. Simultaneous failures of more than two links were not examined as those approximate node failure. Failure modes $\{C 2 - C 3, C 6 - C 7\}$ and $C 2, C 7$ are not shown but they were measured to confirm that they completely separate the core network into two disjoint entities. No traffic can reach either half for the duration of the failure. One improvement to the topology is to link the opposite sides of the ladder network by adding an extra link between either $C 0$ and $C 5$ or $C 1$ and $C 4$. The link would prevent the event of the network becoming disjoint with two router or link failures.

The focus of the core network study is core router and link failures. The effects of failure of edge routers or access networks are trivial as only the closest access network and all sessions with it are affected. For example, a failure of $E 0$ means $A 0$ will receive nothing from $A 1 - A 7$ and $A 1 - A 7$ will receive nothing from $A 0$.

4.4.3 Protocols

The main variables are associated with two routing protocols OSPF and MPLS. OSPF failure detection timer is varied to compare it with the stan-

Link Failures	Node Failures
C 2 - C 3	C 0
C 0 - C 1	C 1
C 1 - C 2, C 2 - C 7	C 2
C 0 - C 7, E 0 - C 7	C 3

Table 4.3: Failure modes examined

dard default of 10 s. MPLS is mainly investigated with its fast reroute (FRR) configuration, but some control simulations were created to compare this with the global recovery configuration.

For the network size in question, OSPF is the most popular routing protocol and serves as a good representation of the routing protocol family [35], whereas MPLS is a label switch technology, which is by many [57, 28, 62] believed to be the dominant protocol for the NGN core.

Section 2.9 overviews both OSPF and MPLS in general. This section concentrates on their failure resiliency analysis.

4.4.3.1 Overview of OSPF Resiliency

One of the most common intra-domain routing protocols in IP networks is OSPF. Pasqualini et al. [64] is the basis of much of the information on OSPF contained in this section.

The Hello protocol is used for the detection of topology changes. Each router periodically emits *Hello* packets on all its outgoing interfaces. If a router has not received *Hello* packets from an adjacent router within the *RouterDeadInterval*, the link between the two routers is considered down. When a topology change is detected, the information is broadcasted to neighbours via Link State Advertisements (LSA).

Each router maintains a complete view of the OSPF area, stored as an LSA Database. Each LSA represents one link of the network, and adjacent routers exchange bundles of LSAs to synchronise their databases. When a new LSA is received the database is updated and the information is broadcast on outgoing interfaces.

Name	Default value	Short description
<i>HelloInterval</i>	10 s	Time between two consecutive Hello packets
<i>RouterDeadInterval</i>	40 s	Assume neighbour is down if no Hello packets received after this time
<i>RetransmitInterval</i>	5 s	Time between LSA transmissions
<i>InterfaceTransmissionDelay</i>	1 s	Estimate of time to transmit LSA. Used to age LSA
<i>spfDelay</i>	5 s	Minimum time between LSA reception and start of SPF calculation
<i>spfHoldTime</i>	10 s	Minimum time between consecutive SPF calculations
<i>processLSA</i>	0.6 - 1.1 ms	Check that LSA is new and update LSA database
<i>lsaFlood</i>	33 ms	Process LSA, bundle LSAs and pacing timer
<i>updateFIB</i>	100 - 300 ms	From end of LSA processing to end of new routes installation

Table 4.4: OSPF configurable timers. Adapted from RFC 2328 [60] and Goyal et al. [36].

Route calculations are based on link costs, which are associated with each link. Each router then calculates a complete shortest path tree, however only the next hop is used for the forwarding process.

The Forwarding Information Base (FIB) of a router determines which interface has to be used to forward a packet. After each computation of routes, the FIB must be reconfigured. Table 4.4 shows the configurable OSPF parameters.

In OSPF, two adjacent routers in the same area periodically exchange *Hello* messages to maintain the link adjacency. If a router does not receive a *Hello* message from its neighbor within a *RouterDeadInterval* (typically set to 4 *HelloIntervals*), it assumes the link to its neighbor is down, and generates a new *Router LSA* to reflect the changed topology. All *Router LSAs*, generated by the routers affected by the failure, are flooded throughout the network and cause the routers in the network to recompute the shortest path first (SPF) calculation and update the next hop information in their respective forwarding tables. In summary, the time required to recover from a failure consists of:

- Failure detection time — *RouterDeadInterval*
- LSA processing and flooding time — *processLSA* and *lsaFlood*
- SPF calculation time — *spfDelay* and *spfHoldTime*
- Updating forwarding tables — *updateFIB*

Using Table 4.4 for the values of the timers, failure detection time is between 30 and 40 seconds. *lsaFlood* times consist of the propagation delays and any pacing delays resulting from the rate-limiting of (Link State Update) *LSUpdate* packets transmitted on an interface. Once a router receives a new LSA, it schedules an SPF calculation. LSA processing and flooding takes approximately 34 ms. Since the SPF calculation using Dijkstra's algorithm [60] constitutes a significant processing load, a router typically waits *spfDelay* time to let additional LSAs arrive, before doing the SPF calculation on a batch of LSAs. Routers also limit the frequency of SPF calculations introducing further delays. The frequency is set to *spfHoldTime*, typically 10 seconds, between successive SPF calculations. Thus, SPF calculation time is less than 10 seconds. Updating forwarding tables takes at most 300 ms.

Hence, failure detection is the main component of OSPF failure recovery, followed by the SPF calculation time. Most current efforts have been concentrated on reducing failure detection time. Goyal et al. [36] and Basu and Riecke [16] have found that reducing the *HelloInterval* can substantially reduce the failure detection time. However, there is a limit up to which the *HelloInterval* can be safely reduced. As the *HelloInterval* becomes smaller, there is an increased chance that network congestion will lead to loss of several consecutive *Hello* messages thus causing a false alarm that the adjacency between routers is lost, even though the routers

and the link between them are functioning perfectly well. The LSAs generated because of a false alarm will lead to new SPF calculations by all the routers in the network, to avoid the supposedly down link. This false alarm would soon be corrected by a successful *Hello* exchange between the affected routers, which then causes a new set of LSAs to be generated, and possibly new path calculations by the routers in the network. Thus, false alarms cause an unnecessary processing load on routers and sometimes lead to temporary changes in the path taken by network traffic. This can have a serious impact on the QoS achieved in the network. If false alarms are too frequent, routers have to spend considerable time doing unnecessary LSA processing and SPF calculations, which may significantly delay important tasks such as *Hello* processing, thereby leading to more false alarms. Persistent overloads on router processors may adversely affect the routing function in the whole network.

Pasqualini et al. [64] analysed failure detection by minimising the OSPF timers and comparing the performance with MPLS protection switching. The conclusion is that even with *HelloInterval* at 5 ms internal dependencies of OSPF specification restrict failure recovery. To overcome this the OSPF specification must be modified in order to have sub-second recovery time. The work [64] did not take account of false alarms however. MPLS global and local protection on the other hand proved to perform in a sub-second region without any modifications.

To improve failure recovery OSPF has an equal-cost multi-path mechanism [80, 42] where OSPF stores multiple paths to a single destination provided more than one such distinct path and the sum of link costs is the same for each path. In practice it is not straightforward to find link cost assignments yielding equal cost for several paths [75]. When a fail-

ure occurs on one path, the other one is used without recalculating the route. Failure detection still needs to be performed and equal paths need to be configured. There is a very recent attempt by Schollmeier et al. [70] at a new routing scheme, called O2 routing, which provides each node in the network with two or more outgoing links towards every destination. Two or more possible next hops are then used at each router towards any destination instead of OSPF's single next hop. This thesis does not attempt to study the newer routing schemes.

4.4.3.2 Overview of MPLS Resiliency

MPLS Recovery methods [64] provide alternative LSPs to which the traffic can be switched in case of a failure. There are two types of recovery mechanisms: protection switching and restoration. Protection switching includes recovery methods where a protection LSP is pre-calculated, just needing a switching of all traffic from the working LSP to the backup LSP after the failure detection. In restoration, the backup LSP is calculated dynamically after the detection. Another way to classify MPLS recovery mechanisms depends on which router along the LSP takes the rerouting decision. It can be done locally — called Fast Reroute (FRR), the node detecting a failure immediately switching the traffic from the working to the backup LSP, or globally when the failure is notified to upstream and downstream LSRs that reroute the traffic.

In global recovery secondary LSPs are set up manually (or using a constrained shortest path first algorithm), which means the network manager explicitly controls where the traffic will flow. The disadvantage of this method is a longer recovery time as the detection signal must come back to the router that originated the primary LSP. FRR configuration

performs local failure recovery automatically, which is faster, but there is less control of traffic flow after the failure.

In my thesis FRR was chosen as an MPLS recovery scheme for several reasons. Global recovery predetermines all traffic flows exactly and so there will be no unexpected congestion. This is good in general, but for the experimental purposes not very interesting and acts as the control or base case. Although, everything is predetermined with global recovery, there are important problems. One problem is that the backup LSPs must be configured manually, which is time consuming for the 56 LSPs used in the experiments. The other problem is that recovery takes much longer, because the failure signal must travel back to the ingress LSR to inform that there is a failure and only then will the new traffic (or the traffic rerouted back from the failure, depending on the exact scheme used) be rerouted onto the backup LSP.

FRR on the other hand, needs no manual configuration and it is much faster to recover from failure because the failure signal does not need to return all the way back to the ingress LSR. FRR is the best choice if minimum recovery time is required, as it is with NGN services such as VoIP. One potential problem is that there may be unexpected congestion if too many backup tunnels use the same link during a failure. My experiments showed that this is not a problem, because both the core and the access topologies provide enough multiple redundant paths for failure traffic to be distributed among them.

The FRR scheme used was a restoration category. A bypass tunnel is only created when a failure happens. This is much more efficient in terms of resources, than have every possibly bypass tunnel configured for every possible failure for every possible working LSP. As experiments

show the total recovery time is still much less than the guaranteed 50 ms for FRR, which is much less than the critical failure interval of 200 ms for real time NGN services. Results for FRR under protection scheme are easy to obtain by simply subtracting the bypass tunnel setup time from the effective rerouting time.

4.4.4 Simulation Execution and Data Collection

The experiments were performed using the following simulation configuration. For the topology and the protocols in the study, 160 seconds is allowed at the start of the simulation for the network to reach a steady state. In a steady state, on average, no changes occur in the network traffic and the route configuration is finished by the routing protocols. An example of a non-steady state would be if a failure was introduced in the middle of a queue build up, which increases as time advances. Extra time is added to protocol convergence to further ensure a steady state condition. A failure is introduced at 300 simulation seconds.

The failure persists for 100 seconds to let the failure propagate through the network. A short failure, such as 10 seconds, may not propagate completely if failure detection time is longer than the failure period. 100 seconds is more than enough for the purposes of my experiments. Simulation persists for further 200 seconds to ensure traffic returns to steady state after the failure is removed at 400 seconds.

Figure 4.3 shows traffic received by one of the destination access network from one source access network. The graph shows part of the raw data measured in the core network experiments. The graph shows that during protocol convergence no traffic is received. After around 160 seconds a VoIP call session is established with the mean rate of 100 packets

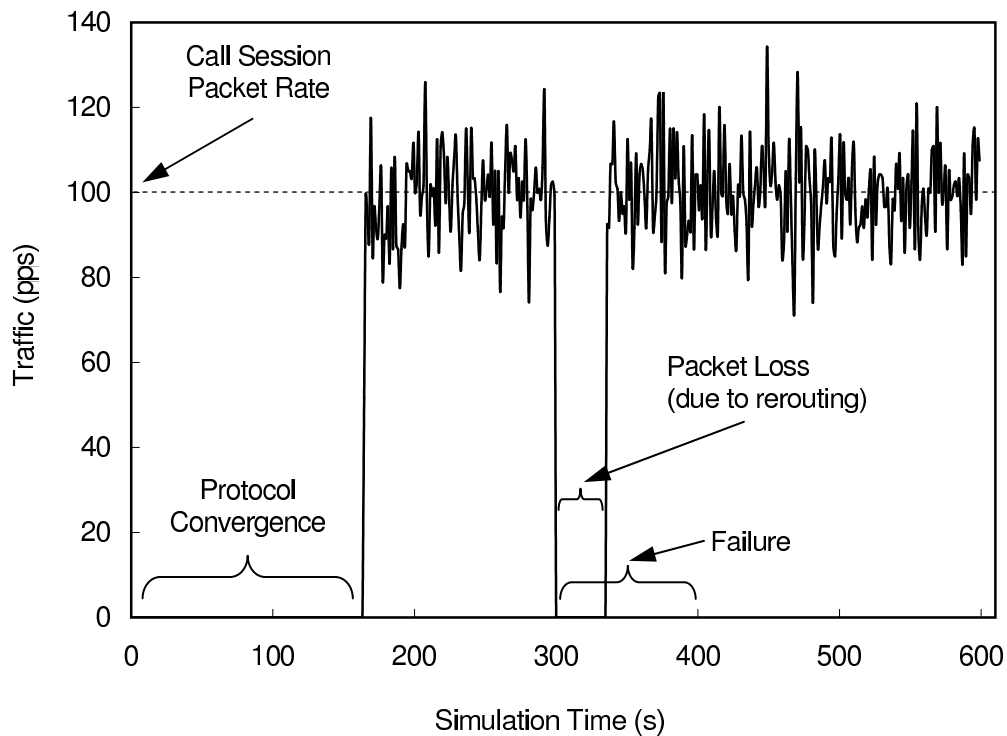


Figure 4.3: Traffic received by a destination from one source versus simulation time. Illustration of measurements for a sample failure.

per second (pps). To make sure steady state is achieved the failure is introduced over 100 seconds later, at 300 seconds. When the failure is introduced OSPF takes some time to react to failure by rerouting the session to a different physical route. The graph is used to measure the rerouting time. Also we can see the packet loss due to rerouting, which is simply all of the traffic. Once rerouting is finished the traffic is back to its normal level before the failure. The failure is repaired at simulation time of 400 seconds.

Figure 4.4 , which corresponds to Figure 4.3, shows the ETE delay experienced by one destination access network. We can see the ETE delay from around 160 simulation seconds. When the failure occurs no traffic is received for a short time, which accounts for the break in the delay data. The graph is used to measure the delay after failure, which could also be

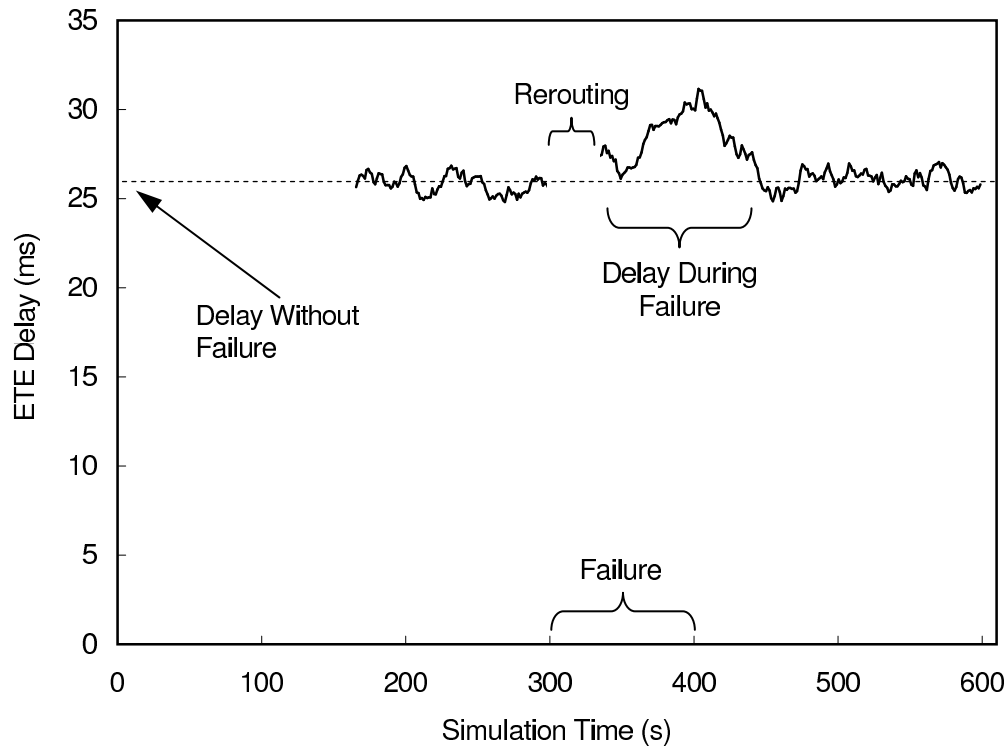


Figure 4.4: ETE Delay versus simulation time. Illustration of measurements.

thought of as delay during failure. The delay is significantly higher than delay elsewhere. For measurement purposes this portion of the graph is averaged over time to get the measure of delay after failure.

4.5 Results

4.5.1 Uncertainties

To be confident in accuracy of the results uncertainties were measured for each measurement category. The common method to measure uncertainty in discrete event simulation is to repeat the set of experiments for several random seed values that lie in a wide range. My simulations used five seed values ranging from 23 to 1000001.

Time taken	Metric	OSPF	MPLS
n/a	Loss	0	0
	Reroute	993.73	0.17
Before Failure	Delay	0.07	0.08
	Jitter	0.10	0.07
After Failure	Delay	0.38	0.17
	Jitter	0.09	0.06

(a) Small link capacity

Time taken	Metric	OSPF	MPLS
Before Failure	Delay	0.42	0.68
	Jitter	0.00	0.00
After Failure	Delay	0.77	0.83
	Jitter	0.002	0.001

(b) Large link capacity

Table 4.5: Summary of uncertainties in milliseconds

All experiments were repeated for each seed value and Appendix A shows the appropriate uncertainty ranges. Two exceptions to this are delay and jitter after failure. Unlike, for example, delay before failure, these measures are different for each failure mode. However, uncertainty range in all failure modes is very small and very close to each other. So that a sample of several points would serve as a good representation of the population uncertainties. Three failure modes were chosen to represent the uncertainty ranges for delay and jitter after failure. These failure modes are single link failure C 2 – C 3, double link failure C 0 – C 7, E 0 – C 7, and single node failure C 2. Link C 2 – C 3 was chosen from single link failures, {C 0 – C 7, E 0 – C 7} was chosen from double link failures, and C 2 was chosen from node failures. Each failure mode was chosen from a separate group of failure modes. The mechanism for choosing these failure modes from each group is so that, in terms of reaction to failure, each one is not the best or the worst case, but somewhere in the middle. Thus, the three chosen failure modes approximate the uncertainties for all failure modes well.

Table 4.5 summarises the uncertainties for all measurements taken. The summary only approximates the uncertainties. Strictly uncertainties should be stated with every figure. However, the uncertainties are close enough to make the average range a good representation of the uncer-

Failure Mode	Rerouting
C 2 - C 3	6341.27
C 1 - C 2, C 2 - C 7	7011.44
C 0 - C 7, E 0 - C 7	7102.86
Ave	6818.52
Rng	761.59

(a) Link failures. Uncertainty range is 145.42 ms.

Failure Mode	Rerouting
C 1	37227.09
C 2	37480.02
C 3	47991.71
Ave	40899.61
Rng	10764.62

(b) Node failures. Uncertainty range is 993.73 ms.

Table 4.6: OSPF effective rerouting times in milliseconds

tainty. For example, strictly uncertainty in Table 4.6 is different for each rerouting figure. However, uncertainties are close enough to simply state a single figure for uncertainty range of 145.42 ms. This means that each figure should be taken to be within $\pm\frac{1}{2}$ uncertainty range. More sophisticated statistical techniques such as standard deviations, sample means and so on are not used because the level of the analysis does not require this. The range can be used to state uncertainty because the distribution of uncertainties is normal (or Gaussian) according to Opnet documentation and measured results. Appendix A presents all the individual figures for uncertainties, which are used to derive the average uncertainties.

4.5.2 Rerouting

Rerouting and packet loss as its consequence are the main factors in determining how well the rerouting protocol reacts to failure. During the rerouting time the route is being recalculated by the routers in the network and packets cannot be routed to their destination. Table 4.6 presents the rerouting times due to each failure. Rerouting times are sufficiently different between link and node failures. Table 4.6a shows that the average rerouting time for a link failure using OSPF is approximately 7 s, whereas for a node failure the rerouting time is approximately 40 s, shown

Failure Mode	Reroute	Setup	Eff. Reroute
C 2 - C 3	0.91	14.21	15.12
C 0 - C 1	1.12	6.58	7.70
C 1 - C 2, C 2 - C 7	0.86	13.27	14.13
C 0 - C 7, E 0 - C 7	2.08	12.19	14.27
Ave	1.24	11.56	12.81
Rng	1.22	7.63	7.42

(a) Link failures. Uncertainty range is 0.41 ms.

Failure Mode	Reroute	Setup	Eff. Reroute
C 0	0.67	6.67	7.34
C 1	0.71	14.10	14.81
C 2	0.83	6.92	7.75
C 3	0.90	6.53	7.43
Ave	0.78	8.56	9.33
Rng	0.23	7.57	7.47

(b) Node failures. Uncertainty range is 0.17 ms.

Table 4.7: MPLS effective rerouting times in milliseconds

in Table 4.6b. The difference is explained by how OSPF detects failure [60]. Link failure is detected immediately through loss of carrier without waiting for a *RetransmissionInterval*. In this case, failure recovery time (or convergence time after the failure) depends on SPF calculation parameters. The default behaviour (and the one modeled in Opnet) is that SPF calculation parameters are set to periodic with a *spfHoldTime* of 10 s. This means that SPF calculations happen every 10 s. In case that a link fails, it detects that the neighbour is down immediately but the convergence happens in the next SPF calculation cycle, which is less than or equal to 10 seconds. During a router failure the neighbouring router waits until *RetransmissionInterval* elapses, which is between 30 to 40 seconds by default, then, calculate the SPF in the next SPC calculation cycle.

Table 4.7 shows that MPLS effective rerouting is on average approximately 13 ms for a link failure and 9 ms for a node failure. The effective rerouting of MPLS is made of two components, a setup time for a new LSP tunnel to reroute around the failure and a rerouting time, which is

the time for the traffic to be placed on the new LSP. Table A.3 shows the two components. The average for setup time is 10 ms, whereas the average for rerouting time is 1 ms.

A control experiment was undertaken to measure the effective rerouting of the global MPLS rerouting scheme. The global rerouting scheme is of no interest as part of my experiments mainly because the traffic before and after failure is explicitly mapped onto LSPs and traffic cannot be routed any other way. However, a comparison of the effective rerouting is useful to put the FRR figures in perspective. In the global scheme the secondary LSP is set up from the head end router (the same as the primary LSP). Thus even if the failure occurred close to destination the traffic must be routed back to the head end router. This suggests that the rerouting component of the effective rerouting delay is greater than that of FRR. The experiment showed that an average rerouting time for this topology is 29 ms, compared to the fast reroute rerouting time of 1 ms. The LSP setup time is around 17 ms, which is also greater than for the fast reroute case due to a longer path to set up. A new path must be signaled from the head end router to the destination LER in contrast to fast reroute which has to signal a shorter bypass tunnel around the failure. The effective rerouting time is 46 ms, which is still under 200 ms, so the network layer detects no interruption and thus there are no losses due to rerouting.

4.5.3 Packet Loss

In all my experiments packet loss refers only to losses created through failures resulting in rerouting. As Section 4.4.1 mentions, due to queuing utilisation being very small, there are no random losses in the network.

Failure Mode	A 0	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 2 - C 3	100	200	400	400	400	100	100	100	225	300
C 2 - C 3, C 6 - C 7	400	400	400	400	400	400	400	400	400	0
C 1 - C 2, C 2 - C 7	200	0	0	0	0	200	200	300	113	300
C 0 - C 7, E 0 - C 7	400	0	0	0	0	100	100	100	88	400
C 1	300	100	100	100	100	0	0	0	88	300
C 2	200	600	700	400	400	200	200	300	375	500
C 3	400	200	100	100	100	400	400	0	213	400

Table 4.8: OSPF packet loss to each destination A 0 – A 7 in packets per second. Without losses each destination receives 700 pps. Uncertainty range is 0.

Each access network experiences a different packet loss during the same failure. The duration of all losses is the same rerouting time for that failure. A negligible difference may exist due to propagation delays, as the failure is a different distance from each access network. When a failure happens and rerouting takes place as part of failure recovery, packets from some access network may not reach their destination. For example, if C 1 fails, A 2 will not receive packets from A 0 if the original route from A 0 to A 2 goes through C 1. Since each access network sends exactly 100 pps to each destination excluding itself (700 pps in total), A 2 would receive 100 pps less recording 100 pps loss.

Table 4.8 shows losses experienced by each end point during all failure mode. Packet loss is in multiples of 100 pps because that is the bandwidth of a single call between any pair of end points. Alternatively, the losses could be recorded as lost calls per destination (out of total 7 calls possible to receive).

The last two link failure modes, {C 1 – C 2, C 2 – C 7} and {C 0 – C 7, E 0 – C 7}, have lower losses due to OSPF using different links for the original routes. Table 4.8 does not include C 0 – C 1 and C 0 failures because they experienced 0 losses, due to OSPF choosing the shortest path for its routes. C 0 – C 1 link is never used by OSPF because E 0 – C 1

uses one less router (C 0) to get to C 1. Likewise C 0 is never used as E 0 – C 1 and E 0 – C 7 are always part of the shorter route to and from A 0.

In summary, the losses are very large, ranging from 88 to 400 pps out of total possible 700 pps. That is failures in the core network may create losses of over 50 % for 7 to 40 seconds, which is intolerable for a carrier grade service, such as VoIP.

MPLS experienced no losses at the network layer due to its extremely fast effective rerouting times ranging from only 9 to 13 ms, which conforms to the recommendation of the carrier grade real time service failure recovery of 200 ms, stated in Section 2.11.4.

4.5.4 OSPF Hello Timer Variation

The aim of this task was to examine the effect of *HelloInterval* on rerouting performance of OSPF. Adjusting *HelloInterval* is the best method of decreasing rerouting times. Two sets of experiments were simulated. The first has all other timers set to default and the second has all other timers set to the minimum allowed by Opnet. The latter was set up to assess if other timers have any significant effect on rerouting times.

Figure 4.5 shows an approximately linear positive relationship between *HelloInterval* and rerouting time. The default behaviour differs little from the one with all timers set to minimum. The biggest difference is at the smallest *HelloInterval* of 1 s, where the default gives over 8 s and the other gives just over 3 s.

The set of experiments confirms the latest research on examining failure detection with OSPF in a real network by Goyal et al. [36]. An important factor considered by Goyal et al. is false alarms caused by network congestion. As the *HelloInterval* decreases more hello packets flow

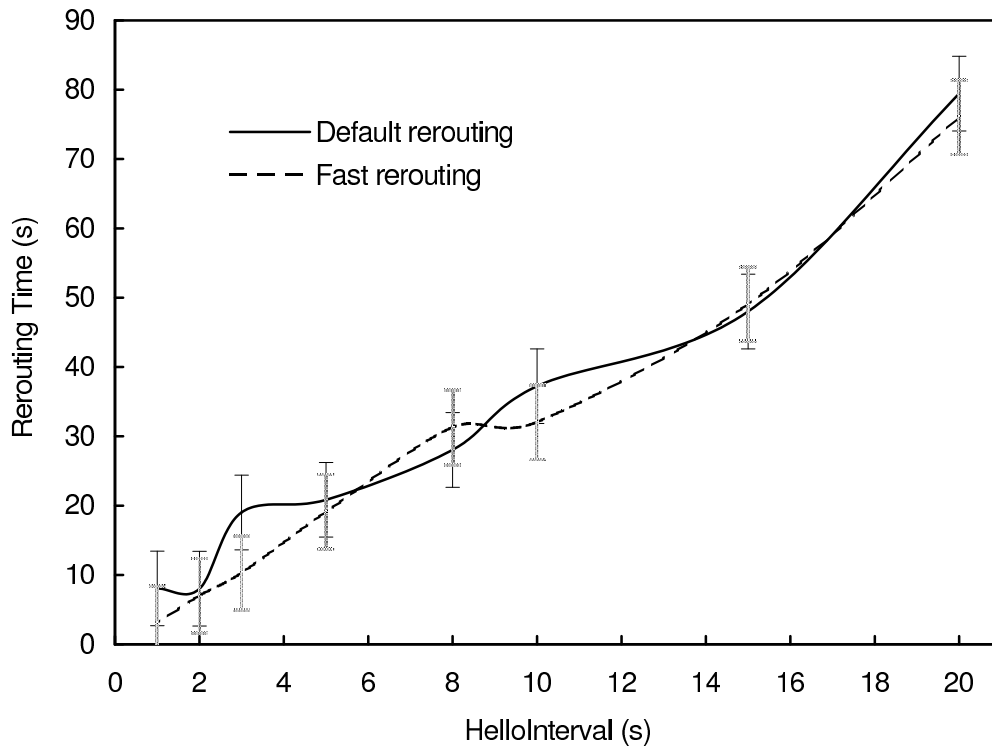


Figure 4.5: OSPF *HelloInterval* variation. Default rerouting has the default timers: *InterfaceTransmissionDelay* = 1 s, *RetransmissionInterval* = 5 s, *spfDelay* = 5 s, *spfHoldTime* = 10 s. For fast rerouting, all timers are set to minimum values in Opnet: *InterfaceTransmissionDelay* = 1 s, *RetransmissionInterval* = 1 s, *spfDelay* = 1 s, *spfHoldTime* = 1 s. Table A.4 shows the raw data. Uncertainty range is 5382.31 ms, represented as error bars.

through the network causing more congestion. Congestion can cause loss of consecutive *Hello* packets, which would cause the router sending *Hello* packets to believe the other router is down and begin a new SPF calculation sending new LSAs. The behaviour would be corrected by a successful *Hello* exchange between the affected routers. However, the effects are increased load on routers and temporary changes to the routes. Thus, *HelloInterval* can only be decreased to a level past which the amount of false alarms would be excessive. The optimal *HelloInterval* for the network studied was found by Goyal et al. [36] to be 250 ms. The rerouting time was found to be a little over 10 s, which slightly exceeds the results in my thesis based on Opnet models. The research also confirms that other

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	2.85	4.01	5.39	6.50	4.88	3.84	2.40	4.27	4.10
C 2 - C 3	-0.05	-0.07	4.14	6.87	7.06	5.27	0.06	3.33	7.14
C 1 - C 2, C 2 - C 7	-0.04	0.34	0.62	1.19	0.12	0.12	0.08	0.35	1.23
C 0 - C 7, E 0 - C 7	0.08	0.15	0.24	0.24	1.70	1.80	1.79	0.86	1.72
C 1	3.22	0.11	0.22	0.24	0.21	0.21	0.10	0.61	3.13
C 2	0.01	0.06	0.36	0.37	0.26	0.28	0.09	0.20	0.36
C 3	0.02	0.06	0.25	0.26	1.42	1.52	0.01	0.51	1.51

(a) OSPF. Uncertainty range is 0.33 ms.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	3.03	4.39	6.04	6.86	5.22	4.04	2.67	4.61	4.19
C 2 - C 3	-0.03	-0.14	2.16	1.36	0.03	0.00	-0.03	0.48	2.30
C 0 - C 1	-0.58	-0.77	-0.13	1.30	1.14	1.45	2.77	0.74	3.54
C 1 - C 2, C 2 - C 7	-2.82	-1.44	1.70	2.50	2.46	4.16	2.74	1.33	6.98
C 0 - C 7, E 0 - C 7	-3.89	-0.76	2.10	4.08	8.42	6.43	3.83	2.89	12.31
C 0	-2.29	-0.03	3.61	6.56	4.70	4.25	2.56	2.77	8.85
C 1	1.13	1.53	6.21	6.42	5.03	4.11	2.70	3.88	5.30
C 2	0.46	n/a	6.31	7.26	6.13	5.13	2.86	4.69	6.80
C 3	3.13	5.55	6.19	6.57	5.75	5.93	3.35	5.21	3.44

(b) MPLS. Uncertainty range is 0.13 ms.

Table 4.9: Small link capacity. Difference between delay after failure and delay before failure in milliseconds

timers do not have any significant affect as there are restricting internal dependencies of OSPF specification. Modifying the protocol to overcome these restrictions would lead to potentially a new protocol, which is out of the scope of my work.

4.5.5 Delay

My experiments analysed the end-to-end one way delay and jitter. The delay difference is extremely minor between OSPF and MPLS. Analysis of jitter follows a similar trend.

4.5.5.1 Small Link Capacity Study

Table 4.9 shows the delay before failure and the difference between delay after failure and delay before failure for OSPF and MPLS. Note the aver-

age delay before failure for OSPF is 4.27 ms and 4.61 ms for MPLS, which is slightly longer as primary LSPs were manually configured with a policy that redundant links from edge routers may not be used for traffic during non-failure situations. This way edge routers are not overloaded during non-failure operation and the redundant links remain free. The redundant links are only used during failure by FRR. Absolute delay values are not interesting because a real network will have much more traffic and links of much larger capacities. Section 4.5.5.2 undertakes a set of experiments to see the approximation to the actual ETE delay using the LLC study.

The most important factor in determining failure effects is the difference between delay before failure and delay after failure. The differences are very minor for both protocols, ranging up to 3 ms for OSPF and 5 ms for MPLS. Overall, these delay increases are insignificant when compared to the packet loss effects, which persists for several seconds.

4.5.5.2 Large Link Capacity Study

This set of experiments was performed to estimate the absolute ETE delay and jitter before and after a failure occurs. The experiments used link capacity of 10 Gbps (9,510,912,000 bps), with a static background load of 9,510,212,000 bps in order to have the already calculated capacity (700,000 bps) for the seven dynamic calls from each end point to all others. This is still only an approximation, because ideally the calculations that were done for the first case with link capacity 700,000 bps, should be done for the 10 Gbps links working out how many calls should be made from each end point. However, this is not feasible in Opnet as the number of calls

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	16.83	21.05	25.31	25.85	21.39	20.67	16.27	21.05	9.58
C 1 - C 2, C 2 - C 7	0.38	2.15	1.23	1.15	0.61	0.70	0.90	1.02	1.77
C 1	1.49	0.52	0.50	0.72	0.21	0.42	0.47	0.62	1.28

(a) OSPF. Uncertainty range is 0.56 ms.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	19.40	23.14	27.01	24.42	26.58	23.28	19.48	23.33	7.61
C 1 - C 2, C 2 - C 7	0.54	3.59	4.55	0.42	0.66	0.53	0.43	1.53	4.13
C 1	0.17	1.03	0.93	0.17	0.99	0.94	0.73	0.71	0.86

(b) MPLS. Uncertainty range is 0.79 ms.

Table 4.10: Large link capacity. Difference between delay after failure and delay before failure in milliseconds

would be in the order of tens of thousands, taking years to simulate. The approximation does provide a good idea of failure behaviour.

Table 4.10a shows that the delay before failure is on average 21 ms for OSPF and 23 ms for MPLS. As explained in Section 4.5.5.1, MPLS LSPs are not the shortest paths between access networks. LLC study investigated two failure modes that had the most significant failure recovery effects in SLC study. The difference in delay after and before failure is approximately 1.5 ms, which is very close to the SLC case. The figure is $< 7\%$ of the absolute delay, which is insignificant.

The longest delay is experienced by A 4 and A 5, which are the furthest from A 0 with respect to both hop count and physical distance. They experience delays of 25 to 27 ms at the longest, which is a very small delay for a network of such size.

4.5.6 Jitter

Jitter results in Tables 4.11 and 4.12 show relatively little absolute variation of less than 0.05 ms ($< 5\%$ of total jitter) for the SLC and less than 0.003 ms ($< 15\%$ of total jitter) for the LLC. The affect of such jitter is in-

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	0.48	0.86	1.36	1.70	1.36	0.95	0.95	1.09	1.22
C 2 - C 3	-0.03	-0.05	0.02	-0.03	0.31	0.39	0.03	0.09	0.44
C 1 - C 2, C 2 - C 7	-0.03	-0.05	0.02	0.12	0.04	0.09	0.05	0.03	0.17
C 0 - C 7, E 0 - C 7	-0.01	0.02	0.09	0.07	0.36	0.46	0.52	0.22	0.53
C 1	0.01	0.00	0.08	0.07	0.07	0.12	0.06	0.06	0.12
C 2	0.35	0.06	0.13	0.13	0.08	0.15	0.05	0.14	0.30
C 3	0.01	0.01	0.09	0.08	0.31	0.38	-0.01	0.12	0.39

(a) OSPF. Uncertainty range before failure is 0.10 ms and 0.09 ms after failure.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	0.58	1.03	1.54	1.83	1.33	1.00	0.53	1.12	1.30
C 2 - C 3	0.01	-0.02	0.02	0.02	0.00	-0.01	0.02	0.01	0.04
C 0 - C 1	0.04	0.02	0.05	0.05	-0.02	0.00	0.01	0.02	0.07
C 1 - C 2, C 2 - C 7	0.08	0.08	0.06	-0.13	0.03	0.06	0.07	0.04	0.21
C 0 - C 7, E 0 - C 7	0.10	0.07	0.04	-0.04	0.01	0.03	0.06	0.04	0.14
C 0	-0.02	-0.05	0.02	0.04	-0.05	-0.05	0.01	-0.01	0.09
C 1	-0.01	-0.04	0.02	0.03	-0.03	-0.02	0.04	0.00	0.08
C 2	0.03	-0.06	0.03	-0.10	-0.04	0.05	0.02	-0.01	0.15
C 3	0.05	0.01	0.07	0.06	0.02	0.07	0.03	0.04	0.06

(b) MPLS. Uncertainty range before failure is 0.07 ms and 0.06 ms after failure.

Table 4.11: Small link capacity. Difference between jitter after failure and jitter before failure in milliseconds.

significant to the end-to-end delay. For these failure conditions, the only difference between LLC and SLC experiments is that LLC absolute jitter of approximately 0.02 ms is much smaller than approximately 1.12 ms of SLC, which is due to larger capacity links. Appendix A Section A.1.2 presents more detailed jitter results.

4.6 Summary

This chapter obtained two major results. Reliability analysis of the ladder topology showed that to satisfy the ETE PSTN reliability requirement of 0.9993, the core network must have individual router and link reliabilities of three and four nines respectively or vice versa. The analysis was based

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	0.015	0.021	0.024	0.024	0.020	0.021	0.019	0.021	0.009
C 1 - C 2, C 2 - C 7	0.000	0.000	0.004	0.003	-0.001	0.000	-0.002	0.001	0.006
C 1	0.003	0.000	0.000	-0.001	0.001	0.000	0.000	0.000	0.004

(a) OSPF. Uncertainty range before failure is 0.002 ms and 0.002 ms after failure.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
Before Failure	0.018	0.021	0.024	0.022	0.021	0.022	0.020	0.021	0.006
C 1 - C 2, C 2 - C 7	0.001	0.002	0.002	0.001	0.000	0.001	0.001	0.001	0.002
C 1	0.002	0.000	-0.001	0.000	0.000	-0.001	-0.001	0.000	0.003

(b) MPLS. Uncertainty range before failure is 0.001 ms and 0.001 ms after failure.

Table 4.12: Large link capacity. Difference between jitter after failure and jitter before failure in milliseconds.

on a two parallel path calculation as opposed to the theoretical reliability, which grossly overestimated the realistic reliability.

The failure detection time of OSPF was found to be approximately linear as the *HelloInterval* increases with the range of 4 to 80 seconds rerouting time. It was also found that other OSPF timers had insignificant effect on the protocol's rerouting ability.

The comparison of OSPF and MPLS in the ladder network demonstrated the superiority of MPLS using the main metrics of rerouting delay and packet loss. OSPF rerouting was shown to be at best approximately 4 seconds with heavy packet losses as a result. On the other hand, MPLS took at worst 13 milliseconds to reroute around any of the failures used. The result also means that the network layer perceives no interruption, but merely a slight increase in delay of the VoIP session. For both protocols, the absolute delay was found to be over 20 milliseconds and delay variability was under 7 % with similar jitter results. These results were insignificant relative to the rerouting time and packet loss metrics.

Chapter 5

Next Generation Access Network

5.1 Introduction

An access network provides aggregation of individual users. Figure 3.1 shows the logical location of an access network within the physical NGN topology. Access networks aggregate from several hundred users in small communities to millions in large cities and can cover hundreds of kilometers. To provide efficient aggregation, access networks typically use a tree-like structure. The capacity and resiliency requirements are typically lower than those for the core network.

This chapter has three main goals. First, is to calculate the reliability of the physical network topology. Second, is to compare the resilience performance of Ethernet and MPLS, which are the main two rival datalink protocols expected in the next generation access network. Third, is to analyse the performance of the complete end-to-end VoIP solution, incorporating results from Chapter 4. The first two goals parallel those of Chapter 4.

Several results were established in this chapter. Physical reliability of

the access network is found to require at least 0.9999 reliability, which corresponds to individual router and link reliability of at least 0.999. The result was obtained using the analysis from Section 4.3.

Experimental findings were that Ethernet experienced rerouting delays of around 4 seconds resulting in significant packet loss, whereas MPLS reroutes in approximately 10 milliseconds, which is short enough to result in no losses at the network layer. The difference between the end-to-end delay after failure and before failure is approximately 1 ms or around 10 % of the ETE delay for each protocol. Jitter results were similar to delay, with the difference being 0.13 ms or 6 % for MPLS and 0.17 ms or 5 % for Ethernet. The value of delay and jitter differences are inconsequential to the performance of ETE voice service. The absolute values of these changes are insignificant to the ETE voice performance. MPLS showed its superiority in absolute figures by being more than 1 ms less than Ethernet.

Finally, the result of the access and core network were combined into the total end-to-end solution. The analysis showed that MPLS is the best protocol to use in both core and access networks. The end-to-end performance conforms to the most stringent requirements of the standards and thresholds consulted. The VoIP call quality was analysed using the E-model, which showed that PSTN quality is achieved provided random packet loss remains below 1.34 % during failure.

Section 5.2 describes the physical topology of the access network and its characteristics. Physical reliability is analysed in Section 5.3. Section 5.4 presents the configuration of my experiments, including the objectives, variables, failure modes, protocols used, and simulation execution. Results are described in Section 5.5 and Section 5.6 analyses the ETE

network solution. The summary in Section 5.7 concludes the chapter.

5.2 Physical Topology

A typical access network is a tree of switches or routers. Figure 5.1 shows a typical example of the access topology. Nodes S 0 – S 14 represent either Ethernet switches or MPLS enabled routers, which will be referred to as switches. The root node S 0 connects to an edge router of the core network. Each leaf node, S 7 – S 14, connects to multiple users through a last mile access, such as ADSL where a Digital Symmetric Line Access Multiplexer (DSLAM) is used to aggregate user connections.

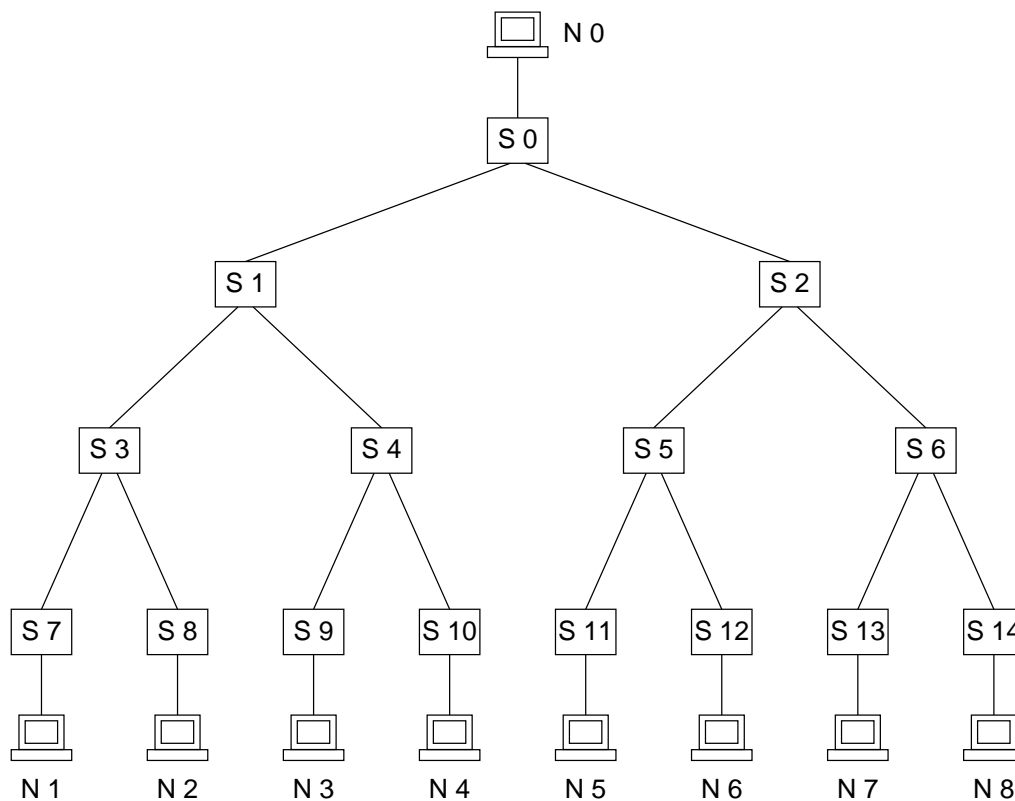


Figure 5.1: Base tree topology

According to Moerman et al. [59], which describes constructing the biggest Ethernet access network in the world, the number of levels in the tree typically would not exceed four. To accommodate a large number of users, the number of children on each node may be more than two. This is due to the leaf node aggregation, for example, an Ethernet switch may only be able to serve several hundred users at a time. Some access networks may be required to serve tens or even hundreds of thousands customers, which could mean over 50 leaf nodes. For the purpose of my thesis the number of nodes in Figure 5.1 is representative for the task of comparing protocol performance in terms of failure recovery. Also, Opnet simulation is prohibitively slow for a larger number of nodes. In my simulation study there are eight users, each setting up a call session with an endpoint N 0 attached to the root of the tree. N 0 is a user that only receives calls in the simulation, but in the total network it represent the edge router of the core network.

In Figure 5.1 there are no alternative routes between the switches, which makes it a highly non-resilient network. Figure 5.2 shows a redundant tree access network, which is much more resilient to failure. The redundant links are placed only between the switches, and not the users for financial reasons. In general, the number and type of redundant links is dictated by financial cost, distance between the switches, and performance requirements.

The network is highly symmetric so a description of one branch is enough to describe the whole tree. The redundant links put in place from S 7 and S 8 are both to S 4, because it is the closest located switch one level up. A redundant link from S 7 or S 8 to S 5 or S 6 is likely to create a much longer link, because the two halves of the tree are likely to separated by

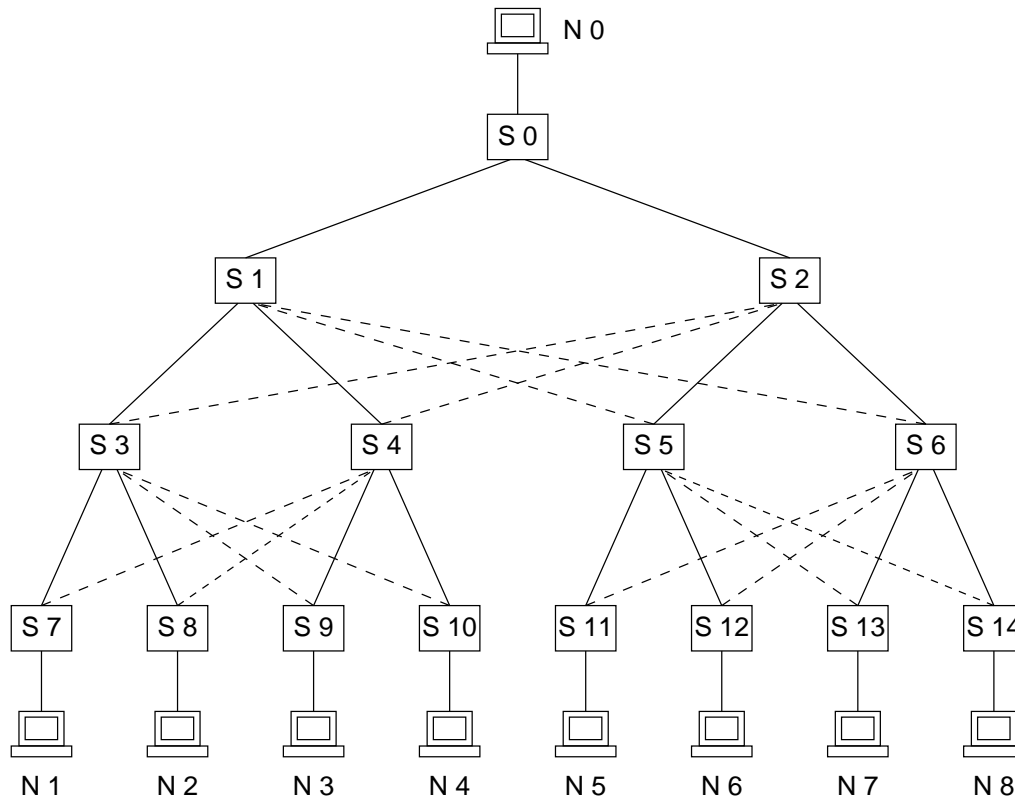


Figure 5.2: Redundant tree topology

a larger geographical distance than the nodes in the same half of the tree. A longer link means a larger ETE delay and higher costs. Another option is to create a redundant link to a switch that is two or three levels higher, for example, S 7 to S 1 or S 7 to S 0. Although the redundant link is longer, it bypasses at least one level of the tree. However, this breaks the structure of the tree topology, which is optimal for aggregation of many traffic sources.

The topology is configured so that the primary links are preferred during the non-failure operation, thus explicitly setting the other links to be redundant and carrying no traffic. One alternative is using load balancing over redundant links [37]. However, load balancing and load sharing schemes constitute a separate research area, which is not covered in this

thesis.

5.2.1 Alternative Topologies

Other possibilities for the access network topology include a ring topology, where the top switch is connected to all switches at the bottom in a chain or a ring. This is highly non-resilient so is not investigated here.

5.2.2 Physical Characteristics

The physical characteristics of the access network are very similar to those of the core network in Section 4.2.3. The length of the network from N 0 to the bottom N 1 – N 8 is approximately 150 km. Thus, each non-redundant link between two switches is 40 – 50 km. Link propagation speed is set to 2.00×10^8 m/s — the speed of light in fibre or copper. The packet processing rate for a switch and a router is 500,000 pps. The Ethernet processing rate of its control information or Bridge Protocol Data Unit (BPDU) for each switch is 100,000 pps. The number is may be considered infinite when compared to the actual control information exchange in my experiments.

Similar to the core network, each user N 1 – N 8 sets up a G.729a VoIP call session with the destination N 0. Each session is 100 pps or 30,400 bps, hence, the total amount of traffic received by N 0 with no failures is eight calls or 800 pps. Link capacity used is 10 Mbps (10,000,000 bps) to represent a relatively small access network. Contrary to the core network configuration, background utilisation is being used to utilise 80 % of the link capacity (8,000,000 bps). At this utilisation and the amount of dynamic traffic, during any failure there are no losses and the links are

Link R	Router Reliability				
	0.99	0.999	0.9999	0.99999	0.999999
0.99	0.994032	0.998130	0.998417	0.998444	0.998447
0.999	0.998130	0.999936	0.999981	0.999984	0.999984
0.9999	0.998417	0.999981	0.999999	0.9999998	0.9999998

Table 5.1: Effective reliability of two parallel paths, P_1 and P_2 , between S 0 and S 7, given router and link reliabilities

90 – 95 % utilised as for the core network. The reason the method in the core network cannot be used to configure link capacities is because Ethernet links have only several discrete capacity values. This configuration is similar to the Large Link Capacity case in Chapter 4.

5.3 Physical Reliability Analysis

This section analyses the physical reliability of the access network. It takes a similar approach to Section 4.3.

Two maximally disjoint paths used for realistic reliability calculations in the access network topology shown in Figure 5.2 are $P_1 = \{S\ 7, S\ 3, S\ 1, S\ 0\}$ and $P_2 = \{S\ 7, S\ 4, S\ 2, S\ 0\}$. Strictly, the last mile access, that is the link from S 7 to N 1 and N 1, should be included in the reliability calculation, but this will be explored later. As in the core network these two paths produce exactly the same reliability expression $T_1 = T_2 = T = l^4 r^4$, where l is the link reliability and r is the switch reliability. The effective reliability of the two paths in parallel is $R = l^4 r^4 (2 - l^4 r^4)$. Table 5.1 shows the effective reliability between S 0 and S 7, given router reliability r and link reliability l . The table is very similar to Table 4.1, but it contains slightly higher reliability figures for the same link and router reliabilities. Referring to the highlighted region of Table 4.2, to achieve PSTN reliability of 0.9993 the access network must be at least 0.9999 reliable. Table 5.1

then shows that $r \geq 0.999$ and $l \geq 0.999$, which are more lenient than r and l for the core network.

In summary of core and access reliability, both networks must be at least 0.9999 reliable according to Table 4.2. Referring to Table 4.1, the core network must have either router reliability of 0.999 and link reliability of 0.9999, or vice versa to satisfy 0.9999 core reliability. For the access network to have 0.9999 reliability, both router and link reliabilities must be at least 0.999.

There is a separate issue of including S 7 to N 1 link and N 1 end node into the reliability calculation. Strictly, for the PSTN ETE reliability calculation, the last mile access distribution to individual users should be included, not including reliability of the CPE. The last mile has a lower reliability, for financial reasons and due to greater number of potential failure modes, such as a cable cut. Reliability of the last mile access network usually does not exceed 0.9999 [82] and is typically much less.

The expression for the reliability of two parallel paths can be modified to include the common S 7 to N 1 link, *LastMile*. The new reliability R^* is the previous reliability R multiplied by a factor:

$$R^* = LastMile \cdot R$$

The expression for R^* can be rearranged, making R the subject.

$$R = \frac{R^*}{LastMile}$$

Because reliability is always less than 1, $R^* < R$. For example, if *LastMile* = 0.999 and $R = 0.9999$, the effective reliability $R^* = 0.9989$. The second expression tells us that $LastMile \geq R^*$ if the previous reliability $R \geq R^*$. In

other words, either a greater reliability is needed everywhere except for the last mile access, or the last mile access needs to be at least as reliable as the overall reliability of the rest of the components, R .

Therefore, to achieve the goal access reliability of 0.9999, $LastMile \geq 0.9999$, which is the maximum reliability figure for the last mile access stated earlier in this section. Such a requirement is too high to apply to last mile networks everywhere. There are efforts to bring redundancy to as near to the customer premise as possible, such as multiple links to the roadside cabinet so that only the last few meters are not redundant. Improving the local loop is problematic due to cost issues as each customer needs to be connected, potentially several kilometers away from the aggregation point to the access network. This is outside the scope of my thesis.

5.4 Experiment Setup

5.4.1 Performance Metrics

The performance metrics for the access network experiments are identical to those for the core network in Section 4.4.1. These are rerouting time, packet loss as a result of the rerouting, end-to-end (ETE) delay and jitter.

Queuing delay measurements were similar to the core network study. Queuing delay between MPLS routers has a mean of 1.0 ms with uncertainty range of 0.04 ms. Queuing delay between Ethernet switches is higher, 2.7 ms with uncertainty range of 0.5 ms, which makes the ETE delay slightly longer than that of MPLS due to the Opnet implementation of the Ethernet model.

Link Failures	Node Failures
S 1 - S 3	S 1
S 4 - S 7	S 3
S 0 - S 1	
S 0 - S 1, S 1 - S 3	
S 0 - S 1, S 2 - S 6	
S 0 - S 1, S 2 - S 3	

Table 5.2: Failure modes examined

Bandwidth wastage is one of potential problems with Ethernet as links may be blocked to prevent loops in the spanning tree. Some efforts have emerged that try to minimise wasted bandwidth of a single spanning tree, such as the the Multiple Spanning Trees Protocol (MSTP, IEEE 802.1Q [9]). However, improvements to the Ethernet and RSTP models, part of which is investigating the problem of queuing and congestion delays, is a separate area of research and is out of the scope of this thesis.

The explicit detailed model of Ethernet needs detailed investigation, part of which is investigating the problem of queuing and congestion delays. Such a study is a separate part of research concerning improvements to the Ethernet protocol and is not part of this thesis.

Queue buffer utilisation for MPLS is 0.067 % and for RSTP it is 0.346 %. Similar to Section 4.4.1, queuing delay can be treated as constant because buffer utilisation is less than 1 %.

5.4.2 Failure Modes

Table 4.3 shows the link and node failure modes (based upon those in Section 2.4) used in the core network experiments. The tree topology is highly symmetric so the failure modes represent most of the possibilities. The two node failures represent the distinct failures in level two and three of the tree, because all other node failures are identical to these due to

symmetry. A failure of a node that has no redundant path around it, such as the end nodes N 0 – N 8 and S 7 – S 14, is not interesting because the failure only affects one source. These failures will persist until the node recovers as there are no alternative paths. The same is true for the links, such as N 1 – S 7.

Link failures are chosen so that in the event of failure more traffic is affected rather than less. For example, although these represent the same kind of failure, failing the link S 3 – S 7 has less impact than S 1 – S 3 because the failure is one level higher up the tree. Double link failures are chosen to be a representative sample of all interesting possibilities. Some are on the same side of the tree, requiring for traffic to switch to the other side through redundant links early on. Some failures are a failure of the primary and redundant link to force a different reroute of traffic.

5.4.3 Protocols

The main simulation variables are the two protocols being used. The MPLS (and RSVP) and Ethernet RSTP parameters are kept at the default values because these are widely agreed to produce good failure resiliency. One RSTP parameter that could make a difference is the *BPDU Hello Time*, which is fixed to be 2 seconds in IEEE 802.1D-2004 [9]. However, it is impossible to change the *BPDU Hello Time* without producing a modified version of RSTP. Other RSTP parameters, although could be changed, do not impact the failure resiliency of the protocol.

Section 2.9 overviews both Ethernet and MPLS protocols in general. This section concentrates on their failure resiliency. Section 4.4.3.2 describes the resiliency of MPLS, so it is not mentioned here further.

5.4.3.1 MPLS in the Access Network

Section 2.9 and 4.4.3.2 overviewed MPLS in general. This section summarises the main reasons for using it in the access network. Hussain [43], Guichard et al. [37] are two good reference books on MPLS in both core and access networks. A very good reference on metropolitan MPLS networks is Tan [79].

Tan [79] advocates MPLS in access networks, because it satisfies scalability, failure resiliency, ease of network management, and service oriented perspective by using an ETE model and QoS classes. MPLS was designed from the outset to provide these features, which were problem areas of most prior protocols.

MPLS rivals Ethernet in the access network. While it is more costly financially and is not plug-and-play like Ethernet, it provides carrier grade guarantees, such as QoS, a number of resiliency choices and so on.

5.4.3.2 Ethernet in the Access Network

IEEE 802.1D-2004 [9] states that failure detection must be $\leq 3 \times BPDUHelloTime$, which is a fixed value of 2 s (also confirmed by Opnet technical support) according to page 153 of IEEE 802.1D-2004 [9]. The exact detection time is implementation specific. In Opnet the detection time is approximately 1 to 2 *BPDU Hello Time*. It is possible to get recovery of as low as 500 ms in some specific cases [63], however, this results in a modified protocol that does not conform to the 802.1D standard.

The requirement set out by the IEEE 802.1D-2004 [9] standard is that failure detection must happen in no more than 3 times hello time. In

most cases the time is the sync time, which is implementation specific. In Opnet this is equivalent to roughly 1 – 2 times hello time. In some rare cases the port state can change to a new forwarding state immediately, causing no interruption as will be evident later.

In the following Ethernet means Ethernet with RSTP.

5.4.4 Simulation Execution and Data Collection

Simulation execution and data collection for the access network are almost equivalent to that of the core network in Section 4.4.4. As in the core network the duration of the failure is 100 seconds, which occurs at the time period of 3 minutes. These figures are the same as in the core network experiments. The difference for the access scenario is that there is one destination and multiple sources, instead of multiple destination and a single source in the core network experiments. This effectively changes the way packet loss is measured. Rerouting time, ETE delay, and jitter are measured in the same way as in the core experiments.

Figure 5.3 shows traffic received from all sources by the single destination N 0. There are eight sources so N 0 receives 800 pps normally. During a failure some call sessions will not reach the destination during the rerouting period. The graph shows that the particular failure results in exactly half of the traffic being lost. It is possible to check packet loss on a per call session basis for each source and add the results together. The method confirms that exactly half of the traffic is lost.

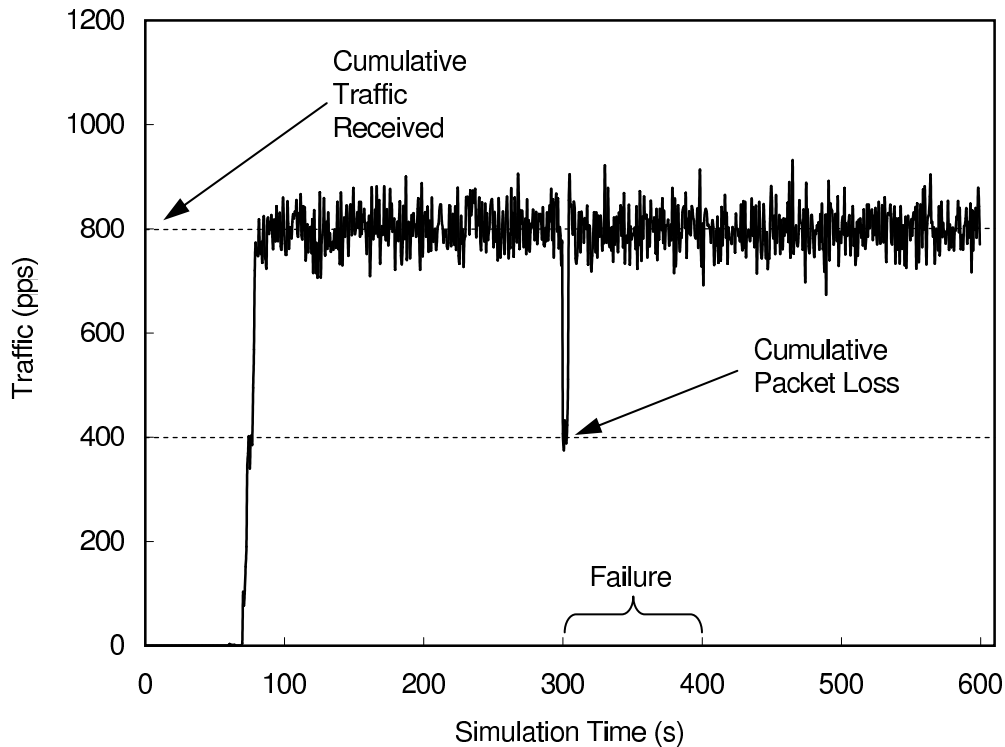


Figure 5.3: Traffic received by N 0 from all sources, N 0 – N 8 versus simulation time. Illustration of measurements for a sample failure.

5.5 Results

5.5.1 Uncertainties

Unlike the uncertainty calculations for the core network in Section 4.5.1, each experiment was simulated for each of the five different seed values. In the core network some of the experiments were not simulated multiple times due to the simulation duration and complexity. Table 5.3 summarises the uncertainties for the access network experiments.

5.5.2 Rerouting

Table 5.4 shows Ethernet rerouting time. The average rerouting time is close to 4 seconds for both link and node failures, which indicates that Ethernet treats both kinds of failures the same way. IEEE 802.1D [9] con-

Time taken	Metric	RSTP	MPLS
n/a	Loss	0	0
	Reroute	180.67	0.32
Before Failure	Delay	0.56	0.09
	Jitter	0.10	0.04
After Failure	Delay	0.38	0.13
	Jitter	0.07	0.05

Table 5.3: Summary of uncertainties in milliseconds

firmly that Ethernet RSTP treats node and link failures equally. The only exception is if switches can detect links down using hardware detection, which accelerates failure recovery. Such functionality is an optional part of the switch implementation. My experiments represent the worst case performance, so hardware detection is not implemented in the switches. The Ethernet rerouting time is close to constant with a range of only 40 ms.

Table 5.5 shows MPLS effective rerouting times. The figures are very similar to those of the core network shown in Table 4.7. For link failures the average effective rerouting time is 10.50 ms and for node failures it is a similar figure of 6.91 ms. To reroute traffic into a backup tunnel it takes at most 1.42 ms, whereas tunnel configuration time is the bulk of the effective rerouting time of at most 9.08 ms. If a hot standby technique [79] is used to preestablish the backup tunnels for all possible failures, the con-

Failure Mode	Rerouting
S 0 - S 1	3943.02
S 0 - S 1, S 1 - S 3	3954.21
S 0 - S 1, S 2 - S 6	3980.62
S 0 - S 1, S 2 - S 3	3960.97
Ave	3959.71
Rng	37.60

(a) Link failures. Uncertainty range is 133.05 ms.

Failure Mode	Rerouting
S 1	3960.35
S 3	3951.66
Ave	3956.01
Rng	8.69

(b) Node failures. Uncertainty range is 180.67 ms.

Table 5.4: Ethernet effective rerouting times in milliseconds

Failure Mode	Reroute	Setup	Eff. Reroute
S 0 - S 1	0.97	7.62	<i>8.59</i>
S 0 - S 1, S 1 - S 3	2.05	5.28	<i>7.33</i>
S 0 - S 1, S 2 - S 6	0.77	12.78	<i>13.55</i>
S 0 - S 1, S 2 - S 3	1.89	10.63	<i>12.52</i>
Average	1.42	9.08	10.50
Range	1.28	7.50	<i>6.22</i>

(a) Link failures. Uncertainty range is 0.42 ms.

Failure Mode	Reroute	Setup	Eff. Reroute
S 1	1.03	6.09	<i>7.12</i>
S 3	1.14	5.56	<i>6.70</i>
Average	1.09	5.83	6.91
Range	0.11	0.53	<i>0.42</i>

(b) Node failures. Uncertainty range is 0.21 ms.

Table 5.5: MPLS effective rerouting times in milliseconds

Failure Mode	N 0
S 1	400
S 3	200
S 1 - S 3	0
S 4 - S 7	0
S 0 - S 1	400
S 0 - S 1, S 1 - S 3	200
S 0 - S 1, S 2 - S 6	600
S 0 - S 1, S 2 - S 3	400

Table 5.6: Ethernet packet loss from source N 1 – N 8 to destination N 0 in pps. Without losses N 0 receives 800 pps. Uncertainty range is 0.

figuration time would become nearly zero. However, as Section 4.5.2 explains, that this could potentially bind too many resources because there are many possible failures and backup tunnels must be preconfigured for all of them if hot standby is to be used. My experiments do not investigate hot standby explicitly, but the result is predicted to be similar to the effective rerouting time, that is, in the order of a few milliseconds.

5.5.3 Packet Loss

Table 5.6 shows Ethernet packet loss. The length of the period of the packet loss corresponds to the values in Table 5.4. Failure modes S 1 –

S 3 and S 4 – S 7 experienced no losses because Ethernet was able react quickly to the failure and the switch was able to change to forwarding state immediately, causing zero rerouting time and loss as was mentioned in Section 5.4.3.2. In the case of S 1 – S 3 failure, the redundant link S 3 – S 2 becomes the primary link without any synchronisation or timer delay. In case of S 4 – S 7, the primary routing link S 7 – S 3 is not affected so the failure of the redundant link causes no interruption.

The other losses are explained by looking at the topology. For example, S 1 failure causes a temporary disruption to all traffic coming from the left side of the tree. Thus, half of the traffic ($\frac{800}{2} = 400$ pps) is lost during the rerouting period. The most serious packet loss of 600 pps occurs during {S 0 – S 1, S 2 – S 3} failure, which is the sum of the loss of the traffic from the left half of the tree and the right most quarter side of the tree.

5.5.4 Delay

Table 5.7 shows the difference between ETE delay after failure and delay before failure. Ethernet delay before failure was measured to be 15.3 ms with uncertainty range of 0.56 ms, whereas MPLS delay before failure was 8.08 ms with uncertainty range of 0.09 ms. Identical packet processing rates, link capacities, and background load were used so Ethernet delay is inherently slower than MPLS. Section 5.4.1 explains that the difference is caused by a longer queuing delay. The difference is approximately 1.7 ms per switch, which translates to 6.8 ms for a typical path with four switches corresponding to the difference between the ETE delays shown above. The delay difference potentially impacts the decision

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	0.84	0.56	0.23	0.21	-0.23	-0.20	0.16	-0.05	0.19	1.07
S 3	0.43	0.09	-0.07	0.03	-0.49	-0.24	-0.23	-0.41	-0.11	0.92
S 0 - S 1	0.53	0.41	0.26	0.23	0.22	0.03	0.09	0.24	0.25	0.50
S 0 - S 1, S 1 - S 3	0.71	0.70	0.62	0.63	0.18	0.13	0.38	0.28	0.45	0.58
S 0 - S 1, S 2 - S 6	0.69	0.49	0.31	0.17	0.27	0.00	0.09	0.31	0.29	0.69
S 0 - S 1, S 2 - S 3	5.02	4.85	0.37	0.54	0.49	0.44	0.30	0.39	1.55	4.72

(a) Ethernet. Uncertainty range is 0.45 ms.

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	0.35	0.38	0.33	0.22	0.11	0.25	0.19	0.14	0.25	0.27
S 3	0.18	0.07	0.02	0.01	0.02	0.12	0.07	0.05	0.07	0.17
S 0 - S 1	1.98	1.95	1.97	1.82	0.17	0.23	0.13	0.16	1.05	1.85
S 0 - S 1, S 1 - S 3	0.30	0.26	0.24	0.23	0.03	0.22	0.15	0.14	0.20	0.27
S 0 - S 1, S 2 - S 6	1.89	1.87	2.00	1.86	0.09	0.46	0.12	-0.10	1.02	2.10
S 0 - S 1, S 2 - S 3	1.96	1.93	1.92	1.79	0.15	0.20	0.13	0.11	1.02	1.85

(b) MPLS. Uncertainty range is 0.11 ms.

Table 5.7: Difference between delay after failure and delay before failure in milliseconds

between the protocols. However, the difference in the rerouting time is a more crucial metric, and clearly MPLS has a big advantage over Ethernet.

The difference between before and after failure delays is more important than the absolute delay for resiliency investigation. Table 5.7 shows that the delay difference range for Ethernet is 0.19 – 1.55 ms, which is 1 – 10 % of the total delay. For MPLS the difference range is 0.07 – 1.05 ms, translating into 1 – 13 % of the total delay. These differences are insignificant, caused by a longer propagation delay due to the redundant links being longer than the primary links. There is no indication of any significant congestion problems, which would cause a large increase in ETE delays.

5.5.5 Jitter

Jitter results parallel those of delay, showing the superiority of MPLS over Ethernet in absolute terms. Table 5.8 presents the difference between jitter after failure and jitter before failure. Jitter before failure was measured to

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	0.06	0.10	0.05	0.05	0.07	-0.03	0.08	0.02	0.05	0.13
S 3	0.06	0.03	0.09	0.12	0.01	0.00	0.07	0.03	0.05	0.12
S 0 - S 1	0.02	0.05	0.08	0.06	0.07	0.05	0.05	0.08	0.06	0.06
S 0 - S 1, S 1 - S 3	0.11	0.10	0.11	0.11	0.02	0.01	0.07	0.04	0.07	0.10
S 0 - S 1, S 2 - S 6	0.06	0.04	0.06	0.06	0.05	0.11	0.06	0.05	0.06	0.07
S 0 - S 1, S 2 - S 3	0.51	0.48	0.09	0.08	0.01	-0.03	0.09	0.09	0.17	0.54

(a) Ethernet. Uncertainty range is 0.09 ms.

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	0.02	0.01	0.04	0.02	-0.02	0.02	0.01	0.03	0.02	0.06
S 3	0.00	-0.02	0.05	0.01	0.01	-0.01	0.01	0.01	0.01	0.07
S 0 - S 1	0.21	0.24	0.26	0.22	0.02	0.02	0.03	0.00	0.13	0.26
S 0 - S 1, S 1 - S 3	-0.03	-0.02	0.04	0.01	-0.01	0.03	0.01	-0.01	0.00	0.07
S 0 - S 1, S 2 - S 6	0.25	0.24	0.25	0.23	-0.02	0.01	0.09	-0.04	0.13	0.29
S 0 - S 1, S 2 - S 3	0.21	0.23	0.24	0.23	0.06	0.05	0.03	0.02	0.13	0.22

(b) MPLS. Uncertainty range is 0.05 ms.

Table 5.8: Difference between jitter after failure and jitter before failure in milliseconds

be 3.70 ms with uncertainty range of 0.10 ms for Ethernet and 2.24 ms with uncertainty range of 0.04 ms for MPLS. We can see that MPLS is superior to Ethernet by more than 1 ms. The impact of a failure on jitter is 0.13 ms or 6 % for MPLS and 0.17 ms or 5 % for Ethernet. The absolute values of these changes are insignificant to the ETE voice performance.

5.6 End-to-end Network Analysis

Figure 5.4 shows the combination of various delays that make up the VoIP mouth-to-ear delay. The worst case behaviour must be analysed so the longest (as opposed to the average) delays for each network are chosen. Figure 5.4 indicates the delays after failure for the core and access network, which can be found in Table A.6 and B.4. These results show that the worst delay in core networks is experienced by MPLS with the highest value of 31.56 ms. Delay in access network is 20.28 ms for Ethernet and 10.06 ms for MPLS. When combined an all MPLS solution yields

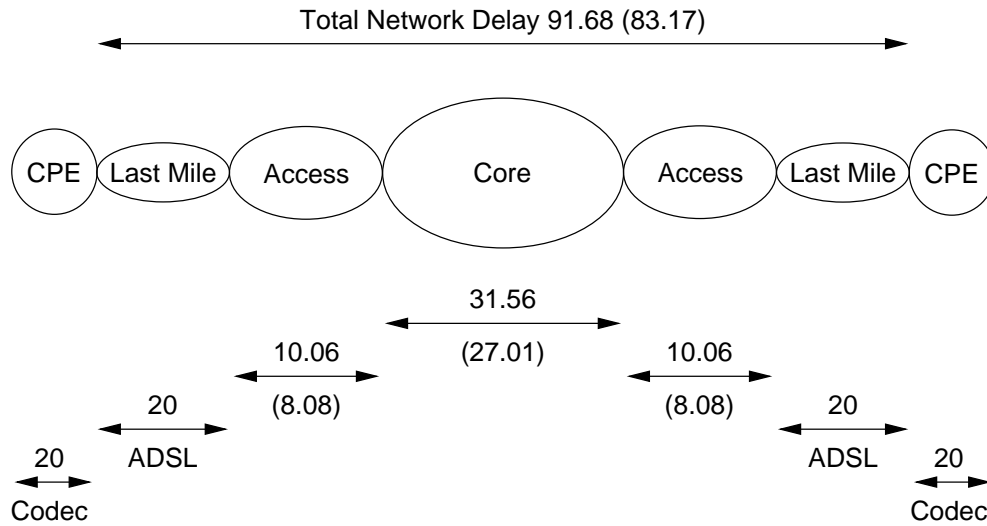


Figure 5.4: Summary of the worst case delays in a VoIP path in milliseconds. Figures in brackets show delays before failure. Network delay extends only to CPE interface. CPE delay represents codec related delays

an ETE delay of 51.68 ms. If Ethernet in access network is used, the ETE delay is 72.12 ms, however Ethernet performs significantly worse than MPLS rerouting and loss metrics so only the MPLS worst case ETE delay is analysed here.

The delay before failure for the all MPLS solution is shown in brackets in Figure 5.4. The figures come from Table 4.10b for the core network and from Section 5.5.4 for the access network.

Figure 5.4 indicates the worst case delay for the last mile network is 20 ms. Such delay is possible in ADSL due the use of forward error correction [81]. Other last mile technologies have much less delay, for example, data over cable has a maximum of 8 ms delay. With the last mile delays of ADSL the network delay becomes $d = 91.68$ ms.

5.6.1 Conformance to ITU-T Standards

ITU-T Recommendation Y.1541 [5] is applied to the total network delay in Figure 5.4 as described in Section 2.11.4. According to Table 2.3, the

total delay of 91.68 ms satisfies Class 0 delay threshold of 100 ms. The ETE jitter, adding two times the access jitter and core jitter, is 2.50 ms, which is below Class 0 jitter threshold of 50 ms.

To assess user experience of a VoIP call on this network ITU-T Recommendation G.114 [1] is used.

Section 2.6.1 discussed the E-model. For this thesis, the E-model in Section 2.6.1 can be modified further. My experiments contain no random loss, which means $e = 0$ and so the expression for the R-value simplifies to:

$$R = 83.2 - 0.024d - 0.11(d - 177.3) \cdot H(d - 177.3)$$

Without the Heavyside function the expression becomes:

$$R = \begin{cases} 83.2 - 0.024d - 0.11(d - 177.3) & \text{if } d > 177.3 \text{ ms} \\ 83.2 - 0.024d & \text{if } d \leq 177.3 \text{ ms} \end{cases}$$

Section 2.6.1 demonstrated that the total mouth-to-ear delay required to compute the R-value is,

$$d = d_{\text{codec}} + d_{\text{de-jitter buffer}} + d_{\text{network}}$$

For the examined G.729a codec $d_{\text{codec}} = 20$ ms, half of which is incurred on the sender side and half on the receiver side; $d_{\text{de-jitter buffer}} = 20$ ms, which is incurred only on the receiver side [28]. Thus, for G.729a, the total mouth-to-ear delay can be written simply as

$$d = d_{\text{network}} + 40$$

Because $d_{\text{network}} = 91.68$ ms, which is illustrated in Figure 5.4, the

total delay is $d = 131.68$ ms. The mouth-to-ear delay before failure in the access and core networks is 123.17 ms. According to Table 2.4, the delay is within 150 ms, which means it is acceptable for a national call.

Packet error rate is generally very good in current networks and the Y.1541 packet error rate of 0.0001 for all classes is easily satisfied [43]. Random packet loss, is not explicitly present in my simulations. However, Section 5.6.2 uses the E-model to compute the range of error rates that can be tolerated.

As stated in Section 2.11.4, there are two other benchmarks that can be used to assess resiliency performance. First, a carrier grade service requires failure repair of under 200 ms, which means the network layer does not detect the failure. MPLS achieves this easily with the worst effective rerouting time of approximately 13 ms in the core and 11 ms in the access network. OSPF rerouting ranges from 6 to 40 seconds, which is one or two orders of magnitude over the target 200 ms. The second benchmark relates to VoIP call continuity, which requires that call interruption should not be greater than 2 seconds or the call is terminated. Again MPLS satisfies the requirement with its short rerouting period and OSPF still fails even this accommodating criterion.

5.6.2 VoIP Call Quality Analysis

Substituting $d = 131.68$ ms into the simplified R-value expression,

$$R = \begin{cases} 83.2 - 0.024d - 0.11(d - 177.3) & \text{if } d > 177.3 \text{ ms} \\ 83.2 - 0.024d & \text{if } d \leq 177.3 \text{ ms} \end{cases}$$

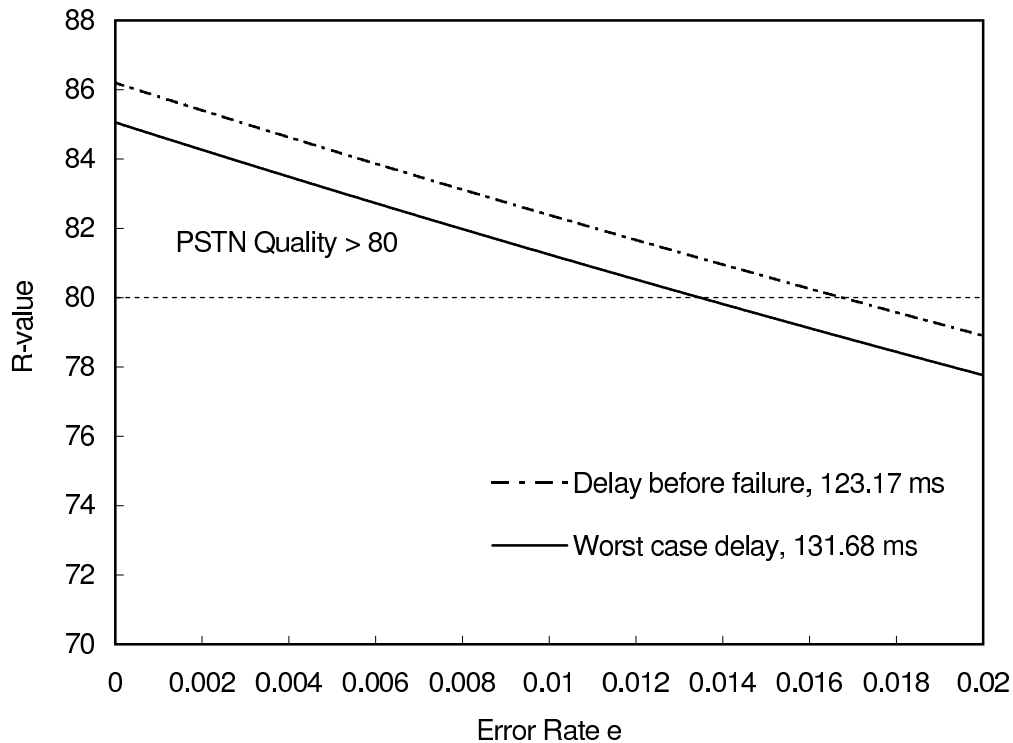


Figure 5.5: E-model R-value versus error rate e . The worst case delay and delay before failure are plotted.

yields $R = 85.06$, which is above the threshold of 80 in Table 2.2 and, hence, it achieves the PSTN call quality. As expected, the delay before failure give a slightly better result of $R = 86.20$.

Because my experiments assumed no random packet loss, Figure 5.5 shows the R-value as a function of error rate e (from Section 2.6.1):

$$R = 83.2 - 0.024d - 0.11(d - 177.3) \cdot H(d - 177.3) - 40 \ln(1 + 10e)$$

Using the worst case delay of 131.68 ms, PSTN quality ($R > 80$) is achieved provided $e < 0.0134$. With no failures ($d = 123.17$ ms), a higher error rate of 0.0167 may be tolerated.

Figure 5.6 is a different perspective on the R-value. The R-value is plotted as a function of delay to investigate what happens to the R-value

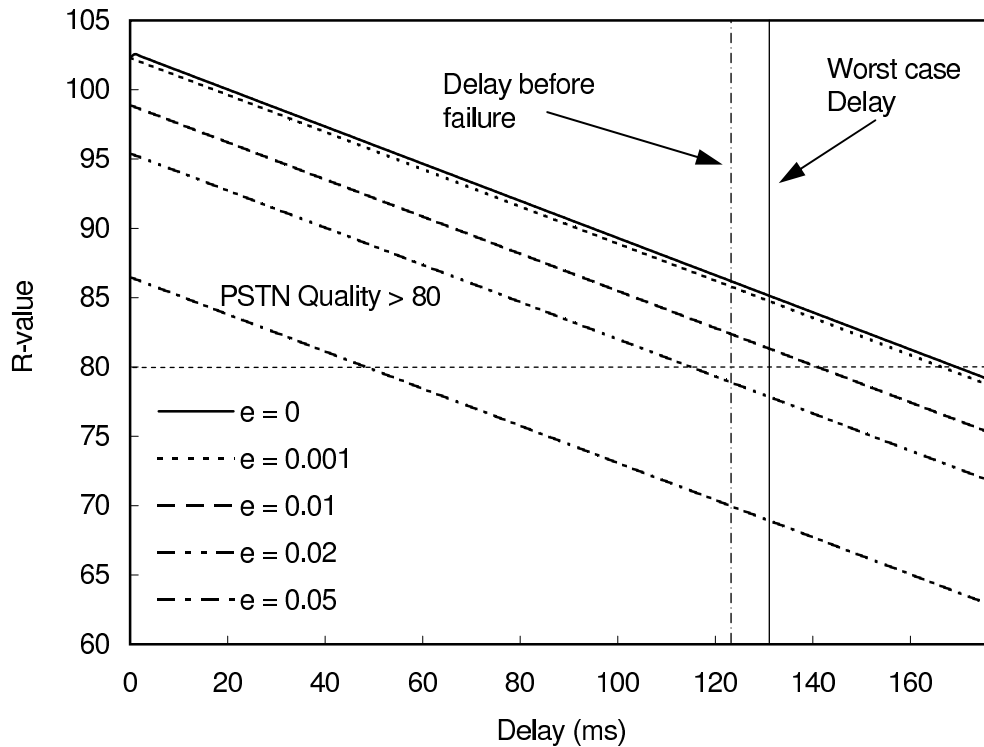


Figure 5.6: E-model R-value versus mouth-to-ear delay. Several error rates are plotted. The worst case delay and the delay before failure are shown.

when delay is varied. Delays for several representative error rates are plotted. When $e = 0$ PSTN quality can be achieved with $d < 169$ ms, whereas when $e = 0.05$ the delay must be less than 48 ms. Note that for the threshold error rate $e = 0.001$ in ITU-T Recommendation Y.1541 [5] shown in Table 2.3, the R-value is practically equivalent to $e = 0$. The main result from the graph is that the worst case mouth-to-ear delay satisfies PSTN quality, provided the total packet error rate is less than 1.34 %.

5.7 Summary

Three sets of results were obtained in this chapter. Reliability of the access network was found to need individual router and link reliabilities of at least three nines to satisfy the ETE PSTN reliability requirement of

99.93 %.

The comparison of Ethernet RSTP and MPLS yielded that MPLS is superior to Ethernet in its effective rerouting and absolute delays. The rerouting time of RSTP was approximately 4 seconds, compared to 10 milliseconds for MPLS. The absolute delay and jitter differences between before and after failure are insignificant to the ETE voice service.

The analysis of the total network solution showed that with MPLS in the core and access the performance of the whole network satisfies the most stringent ITU-T recommendations. The ETE delay of 91.68 ms and jitter of 2.50 ms are under the threshold of 100 ms delay and 50 ms jitter. The E-model shows that even during failure PSTN quality can be maintained, provided the random packet loss is below 1.34 %.

Chapter 6

Conclusions

The goal of my work was to answer the research question: *How well can current IP and Ethernet technology fulfil NGN network and service reliability and resiliency requirements?* This involved comparing the NGN physical reliability to that of the existing telephone network and measuring the resiliency of network layer and datalink layer protocols in the NGN. From the complete end-to-end network, core and access networks were investigated.

The ladder-type network topology was chosen for the core network because it is suitable for a country such as New Zealand. For the access network a redundant tree topology was investigated. Both networks were separately analysed from the physical reliability perspective and from failure resiliency perspective. The results were combined to estimate the end-to-end voice performance in the NGN using ITU-T standards and the E-model.

The conclusion of this work is that the NGN reliability requirement is fulfilled when using the redundant ladder topology in the core network and the redundant tree topology in the access network. The NGN re-

silience requirements can be optimally fulfilled by using MPLS in both the core network and the access network, which provides as good a quality for voice call as that of the telephone network today.

6.1 Contributions

In order to be equivalent to PSTN reliability, computations showed that both the core and the access network are required to have 0.9999 reliability. In turn this means that individual router and link reliabilities for the core must be 0.999, whereas for the access network router reliability must be 0.999 and link reliability must be 0.9999 or vice versa. These requirements are not very high by today's standards and therefore the physical reliability threshold for NGN should be easily satisfied.

The resiliency study of protocols in the core and the access network was primarily based on rerouting performance of the protocols. Resiliency was studied in a New Zealand like context of a representative ladder-based core network and a tree-based access network. The analysis of rerouting within OSPF and MPLS was extended beyond that of the current literature. The two protocols were compared and the analysis showed that an all-MPLS solution is most optimal.

The OSPF protocol performs much worse than MPLS in the core network, with network convergence 1000 times greater than with MPLS. The performance of Ethernet using the rapid spanning tree protocol is much closer to that of MPLS, but it still lags behind MPLS considerably. MPLS was designed from the beginning with resiliency performance in mind and it is considered that an all-MPLS solution is easier to manage, is more scalable and provides end-to-end quality of service.

The end-to-end performance of voice services was analysed in a New Zealand like NGN setting. The analysis of the complete end-to-end network performance demonstrated compliance with all ITU-T recommendations and informal requirements in terms of packet level measurements. The analysis using the E-model showed that the PSTN voice quality is achievable even during failures, provided that the end-to-end random packet loss is below 1.34 %.

6.2 Future Work

There are several directions in which to progress from this thesis.

One direction that needs to be investigated is QoS (or constraint-based) routing in protocols. QoS routing is based on more than one constraint, such as delay, jitter, and reliability. Most routing protocols used today are based on one metric, such as physical distance and hop count. However, choosing an optimal route in IP networks depends on more than just distance or hop count. This is even more vital for NGN because routes may be constrained by voice factors, such as delay, jitter, packet loss, and packet error rate. Also new NGN applications may bring extra constraints.

There has been a lot of research trying to find the best algorithm for QoS routing. Kuiper's PhD thesis [49] is a good overview of the state of the art in QoS routing. Constraint-based routing is an NP-hard problem, for which each algorithm is exponential in the number of network nodes. Most efforts come up with heuristics, most of which are inaccurate or can only deal with a small number of constraints. Kuiper [49] develops a new algorithm which achieves linear performance increase as the net-

work grows for most of the common Internet topologies. The algorithm uses multiple techniques to reduce the search space, which was never done in other previous algorithms.

The algorithm is a great breakthrough. However, it needs to be implemented with an existing algorithm, such as the Label Distribution Protocol for MPLS to find optimal paths given multiple constraints.

Another direction is investigating more topologies for both access and core networks. In this thesis topologies most suited to New Zealand NGN were chosen. Factors such as geographical location, financial costs, and other constraints may make the topology choices of my thesis inadequate.

A more extensive simulation should also be conducted, for example, with the use of GRID to distribute simulation processing to a cluster of machines. A full scale simulation can then be conducted, using a realistic number of access networks and number of calls on each access network. Such a simulation, would solidify and make more accurate the results of my work. Specifically congestion through queuing will be more accurately modeled. However, the results should not be far from results presented here as conservative estimates were used for the traffic utilisation and therefore queue utilisation.

Section 2.2.3 mentioned that to make the NGN secure a number of extra network elements, such as firewalls and SBCs, may need to be used. These elements may degrade the physical reliability as calculated in this work. This area needs to be analysed, to discover the effects of extra security and how it can be minimised. For example, a high reliability firewall may be the result of such work.

Bibliography

- [1] One way transmission time. ITU-T Recommendation G.114, May 2000.
- [2] Packet-based multimedia communications systems. ITU-T Recommendation H.323, 2000.
- [3] The E-Model, a computational model for use in transmission planning. ITU-T Recommendation G.107, May 2000.
- [4] IP Packet Transfer and Availability Performance Parameters. ITU-T Recommendation Y.1540, December 2002.
- [5] Network Performance Objectives for IP-Based Services. ITU-T Recommendation Y.1541, May 2002.
- [6] The avici terabit switch router. <http://www.avici.com>, 2004. Last accessed May 2006.
- [7] General Overview of NGN. ITU-T Recommendation Y.2001, 2004.
- [8] General principles and general reference model for next generation network. ITU-T Recommendation Y.2011, 2004.
- [9] 802.1D-2004, I. Ieee standards for local and metropolitan area networks. Higher Layer LAN Protocols Working Group, 2004.
- [10] A step-by-step migration scenario from PSTN to NGN. <http://www.alcatel.com>, 2001. Last accessed April 2006.
- [11] ALI, S. R. *Digital switching systems: system reliability and analysis*. McGraw-Hill Telecommunications, 1997.
- [12] ANDERSSON, L., DOOLAN, P., FELDMAN, N., FREDETTE, A., AND THOMAS, B. LDP Specification. IETF Request for Comments 3036, January 2001.
- [13] ARANGO, M., DUGAN, A., ELLIOTT, I., HUITEMA, C., AND PICKETT, S. Media gateway control protocol (mgcp) version 1.0. IETF Request for Comments 2705, October 1999.

- [14] AUTENRIETH, A., AND KIRSTDTER, A. Engineering End-to-End IP Resilience Using Resilience-Differentiated QoS. *IEEE Communications Magazine* 40, 1 (January 2002), 50–57.
- [15] AWDUCHE, D., BERGER, L., GAN, D., LI, T., SRINIVASAN, V., AND SWALLOW, G. RSVP-TE: Extensions to RSVP for LSP Tunnels. IETF Request for Comments 3209, December 2001.
- [16] BASU, A., AND RIECKE, J. G. Stability issues in OSPF routing. In *ACM SIGCOMM* (August 2000).
- [17] BEARDEN, M., DENBY, L., KARACALI, B., MELOCHE, J., AND STOTT, D. T. Assessing network readiness for IP telephony. In *ICC 2002 - IEEE International Conference on Communications* (April 2002), vol. 25, pp. 2568–2572.
- [18] BERNET, Y., FORD, P., YAVATKAR, R., BAKER, F., ZHANG, L., SPEER, M., BRADEN, R., DAVIE, B., WROCLAWSKI, J., AND FELSTAINÉ, E. A framework for integrated services operation over diff-serv networks. IETF Request for Comments 2998, November 2000.
- [19] BLACK, U. *Residential Broadband Networks*. Prentice Hall, 1998.
- [20] BLAKE, S., BLACK, D., CARLSON, M., DAVIES, E., WANG, Z., AND WEISS, W. An architecture for differentiated services. IETF Request for Comments 2475, December 1998.
- [21] BRADEN, R., CLARK, D., AND SHENKER, S. Integrated services in the internet architecture: an overview. IETF Request for Comments 1633, June 1994.
- [22] BRUNNER, S. Voice over IP 101. <http://www.juniper.net>, August 2004. Last accessed December 2005.
- [23] CAMARILLO, G., HAUTAKORPI, J., PENFIELD, R., AND HAWRYLYSHEN, A. Functionality of existing session border controller. Tech. rep., IETF, February 2005.
- [24] CARUGI, M., HIRSCHMAN, B., AND NARITA, A. Introduction to the ITU-T NGN Focus Group Release 1: Target Environment, Services, and Capabilities. *IEEE Communications Magazine* 43, 10 (October 2005).
- [25] CHOI, D., AND ET AL. The Voice of the Future: Next Generation Networks. <http://www.atmforum.com/aboutatm/ngn.html>, April 2002. ATM Forum. Last accessed February 2006.

- [26] COCHENNEC, J.-Y. Activities on next-generation networks under global information infrastructure in ITU-T. *IEEE Communications Magazine* 40, 7 (July 2002), 98–101.
- [27] COLE, R. G., AND ROSENBLUTH, J. Voice over ip performance monitoring. *Computer Communications Review* 31, 2 (April 2001), 9 – 24.
- [28] COLLINS, D. *Carrier grade voice over IP*, 2 ed. McGraw-Hill, 2003.
- [29] CUERVO, F., GREENE, N., RAYHAN, A., HUITEMA, C., ROSEN, B., AND SEGERS, J. Megaco protocol version 1.0. IETF Request for Comments 3015, November 2000.
- [30] DANIEL, C., AND WALKER, S. Service Solutions in Next-Generation Networks. <http://www.msforum.org/techinfo/reports/MSF-TR-ARCH-003-FINAL.pdf>, April 2003. Multiservice Switching Forum.
- [31] DAVIDSON, J., AND PETERS, J. *Voice over IP fundamentals*. Cisco Press, 2000.
- [32] DEMEESTER, P., GRYSEELS, M., AUTENRIETH, A., BRIANZA, C., CASTAGNA, L., SIGNORELLI, G., CLEMENTE, R., RAVERA, M., JAJSZCZYK, A., JANUKOWICZ, D., DOORSELAERE, K. V., AND HARADA, Y. Resilience in multilayer networks. *IEEE Communications Magazine* 37, 8 (August 1999), 70–76.
- [33] DREW, P., AND GALLON, C. Next-Generation VoIP Network Architecture. <http://www.msforum.org/techinfo/reports/MSF-TR-ARCH-001-FINAL.pdf>, March 2003. Multiservice Switching Forum.
- [34] EL-SAYED, M., AND JAFFE, J. A view of telecommunications network evolution. *IEEE Communications Magazine* (December 2002), 74–81.
- [35] FOROUZAN, B. A. *Data communications and networking*, 3 ed. McGraw-Hill Companies, Inc, 2003.
- [36] GOYAL, M., RAMAKRISHNAN, K. K., AND CHI FENG, W. Achieving faster failure detection in OSPF networks. In *ICC 2003 - IEEE International Conference on Communications* (May 2003), pp. 296 – 300.
- [37] GUICHARD, J., FAUCHEUR, F. L., AND VASSEUR, J.-P. *Definitive MPLS Network Designs*. Cisco Press, 2005.
- [38] HANDLEY, M., SCHULZRINNE, H., SCHOOLER, E., AND ROSENBERG, J. Sip: Session initiation protocol. IETF Request for Comments 2543, March 1999.

- [39] HARDY, W. C. *VoIP service quality: measuring and evaluating packet-switched voice*. McGraw-Hill Companies, Inc, 2003.
- [40] HASSAN, M., AND JAIN, R. *High performance TCP/IP networking*. Pearson Prentice Hall, 2004.
- [41] HOOGENDOORN, C., SCHRODI, K., HUBER, M., WINKLER, C., AND CHARZINSKI, J. Towards Carrier-Grade Next Generation Networks. In *ICCT* (Beijing, China, April 2003).
- [42] HOPPS, C. Analysis of an equal-cost multi-path algorithm. IETF Request for Comments 2992, November 2000.
- [43] HUSSAIN, I. *Fault-Tolerant IP and MPLS Networks*. Cisco Press, 2005.
- [44] JAMES, J. H., CHEN, B., AND GARRISON, L. Implementing VoIP: A voice transmission performance progress report. *IEEE Communications Magazine* 42, 7 (July 2002), 36–41.
- [45] JAYANT, N., Ed. *Broadband Last Mile*. CRC, 2005.
- [46] JOHNSON, C. R., KOGAN, Y., LEVY, Y., SAHEBAN, F., AND TARAPORE, P. VoIP reliability: A service provider’s perspective. *IEEE Communications Magazine* 42, 7 (July 2004), 48–54.
- [47] KARAM, M. J., AND TOBAGI, F. A. Analysis of the delay and jitter of voice traffic over the internet. In *IEEE INFOCOM 2001 - The Conference on Computer Communications* (April 2001), no. 1, pp. 824–833.
- [48] KNIGHTSON, K., MORITA, N., AND TOWLE, T. NGN Architecture: Generic Principles, Functional Architecture, and Implementation. *IEEE Communications Magazine* 43, 10 (October 2005).
- [49] KUIPERS, F. *Quality of Service Routing in the Internet: Theory, Complexity and Algorithms*. PhD thesis, Delft University Press, The Netherlands, September 2004.
- [50] LAW, A. M., AND KELTON, W. D. *Simulation modeling and analysis*. McGraw-Hill, 1982.
- [51] LEE, C.-S., AND KNIGHT, D. Realization of the Next-Generation Network. *IEEE Communications Magazine* 43, 10 (October 2005).
- [52] LEE, K.-H., LEE, K.-O., PARK, K.-C., LEE, J.-O., AND BANG, Y.-H. Architecture to be deployed on strategies of next-generation networks . In *ICC 2003 — IEEE International Conference on Communications* (May 2003), vol. 26, pp. 819–822.

- [53] LI, Y. R. T. A border gateway protocol 4. IETF Request for Comments 1142, March 1995.
- [54] MARKOPOULOU, A., IANNACCONE, G., BHATTACHARYYA, S., CHUAH, C.-N., AND DIOT, C. Characterization of failures in an IP backbone. In *IEEE INFOCOM 2004 - The Conference on Computer Communications* (March 2004), vol. 23, pp. 2307 – 2317.
- [55] MARKOPOULOU, A. P., TOBAGI, F. A., AND KARAM, M. J. Assessment of VoIP Quality over internet backbones. In *IEEE INFOCOM 2002 - The Conference on Computer Communications* (June 2002), vol. 21, p. 21.
- [56] MEDDEB, A. Why ethernet wan transport? *IEEE Communications Magazine* 43, 11 (2005), 136–141.
- [57] MILLER, M. *Voice over IP technologies: Building the converged network*. M & T Books, 2002.
- [58] MODARRESSI, A. R., AND MOHAN, S. Control and management in next-generation networks: Challenges and opportunities. *IEEE Communications Magazine* 38, 10 (October 2000), 94–102.
- [59] MOERMAN, K., FISHBURN, J., LASSERRE, M., AND GINSBURG, D. Utah's utopia: An ethernet-based mpls/vpls triple play deployment. *IEEE Communications Magazine* 43, 11 (2005), 142–150.
- [60] MOY, J. Ospf version 2. IETF Request for Comments 2328, April 1998.
- [61] NASSAR, A. Strategies for Next-Generation Networks: Migration and Deployment. http://www.iec.org/pubs/print/browse_white_archive.html, 2003.
- [62] OHRTMAN, F. D. *Softswitch: Architecture for VoIP*. McGraw-Hill, 2003.
- [63] OOGHE, S., CLERCQ, J. D., DE VOORDE, I. V., T'JOENS, Y., AND JAEGHER, J. D. Impact of the evolution of the metropolitan network on the DSL access architecture. *IEEE Communications Magazine* 41, 2 (February 2003), 140–145.
- [64] PASQUALINI, S., KIRSTADTER, A. I. A., AND FROT, A. Mpls protection switching vs. ospf rerouting - a simulative comparison. In *Fifth International Workshop on Quality of future Internet Services (QofIS'04)* (Barcelona, Spain, September 2004).
- [65] RAM, A. *Assessment of Voice over IP as a Solution for Voice over ADSL*. PhD thesis, Virginia Polytechnic Institute and State University, May 2002.

- [66] ROSEN, E. Multiprotocol label switching architecture. IETF Request for Comments 3031, January 2001.
- [67] ROSEN, K. H. *Discrete Mathematics and Its Applications*. McGraw-Hill, 1991, ch. pp. 284–286.
- [68] SALTZER, J. H., REED, D. P., AND CLARK, D. D. End-to-end arguments in system design. *ACM Trans. Comput. Syst.* 2, 4 (1984), 277–288.
- [69] SCHEETS, G., PARPERIS, M., AND SINGH, R. Voice over the Internet: A Tutorial Discussing Problems and Solutions Associated with Alternative Transport. *IEEE Communications Surveys and Tutorials* 6, 2 (Second Quarter 2004).
- [70] SCHOLLMEIER, G., CHARZINSKI, J., KIRSTADTER, A., REICHERT, C., SCHRODI, K. J., GLICKMAN, Y., AND WINKLER, C. Improving the Resilience in IP Networks. In *IEEE HPSR* (June 2003).
- [71] SCHULZRINNE, H., CASNER, S., FREDERICK, R., AND JACOBSON, V. Rtp: A transport protocol for real-time applications. IETF Request for Comments 1889, January 1996.
- [72] SCHULZRINNE, H., CASNER, S., FREDERICK, R., AND JACOBSON, V. Rtp: A transport protocol for real-time applications. IETF Request for Comments 3550, February 2003.
- [73] SEITZ, N. ITU-T QoS Standards for IP-Based Networks. *IEEE Communications Magazine* 41, 6 (June 2003), 82–89.
- [74] SHOOMAN, M. L. *Reliability of computer systems and networks: Fault tolerance, analysis, and design*. John Wiley & Sons, Inc., 2002.
- [75] SRIDHARAN, A., GUERIN, R., AND DIOT, C. Achieving near optimal traffic engineering solutions in current ospf/isis networks. In *INFOCOM* (San Fransisco, USA, 2003).
- [76] STEWART, R., XIE, Q., MORNEAULT, K., SHARP, M. C., SCHWARZBAUER, H., TAYLOR, T., RYTINA, I., KALLA, M., ZHANG, L., AND PAXSON, V. Stream control transmission protocol. IETF Request for Comments 2960, October 2000.
- [77] SUN, L. *Speech quality prediction for voice over internet protocol networks*. PhD thesis, University of Plymouth, January 2004.
- [78] TAKAHASHI, A., YOSHINO, H., AND KITAWAKI, N. Perceptual QoS assessment technologies for VoIP. *IEEE Communications Magazine* 42, 7 (July 2004), 28–34.

- [79] TAN, N.-K. *MPLS for Metropolitan Area Networks*. CRC Press, 2005.
- [80] THALER, D., AND HOPPS, C. Multipath issues in unicast and multicast next-hop selection. IETF Request for Comments 2991, November 2000.
- [81] TOUMPAKARIS, D., CIOFFI, J. M., GARDAN, D., AND OUZZIF, M. A square distance-based byte-erasure method for reduced-delay protection of dsl systems from non-stationary interference. In *GLOBECOM 2003 - IEEE Global Telecommunications Conference* (December 2003), no. 1, pp. 2114–2119.
- [82] VOGT, M., AND ANDVAAG, R. M. T. Availability modeling of services in ip networks. In *Fourth International Workshop on Design of Reliable Communication Networks* (October 2003), pp. 167–17.
- [83] WOOD, A. P. Reliability-metric varieties and their relationships. In *Annual Reliability and Maintainability Symposium* (2001), pp. 110–115.

Appendix A

Miscellaneous Core Network Results

A.1 Rerouting

Failure Mode	Min	Ave	Max	<i>Rng</i>	% Ave
C 2 - C 3	6288.22	6320.23	6345.10	<i>56.88</i>	0.90
C 1 - C 2, C 2 - C 7	6916.99	7003.89	7078.08	<i>161.09</i>	2.30
C 0 - C 7, E 0 - C 7	6933.37	7041.24	7151.65	<i>218.28</i>	3.10
Min	6288.22	6320.23	6345.10	<i>56.88</i>	0.90
Ave	6712.86	6788.45	6858.28	145.42	2.10
Max	6933.37	7041.24	7151.65	<i>218.28</i>	3.10
Rng	645.15	721.01	806.55	<i>161.40</i>	2.20

(a) Link failures

Failure Mode	Min	Ave	Max	<i>Rng</i>	% Ave
C 1	37044.18	37302.76	37603.72	<i>559.54</i>	1.50
C 2	37412.45	38021.49	38439.03	<i>1026.58</i>	2.70
C 3	47613.31	48105.91	49008.38	<i>1395.07</i>	2.90
Min	37044.18	37302.76	37603.72	559.54	1.50
Ave	40689.98	41143.39	41683.71	993.73	2.37
Max	47613.31	48105.91	49008.38	1395.07	2.90
Rng	10569.13	10803.15	11404.66	<i>835.53</i>	1.40

(b) Node failures

Table A.1: Uncertainty range for OSPF rerouting times in milliseconds

Failure Mode	Min	Ave	Max	<i>Rng</i>	% Ave
C 2 - C 3	14.79	14.95	15.19	0.40	2.70
C 0 - C 1	7.59	7.63	7.85	0.26	3.44
C 1 - C 2, C 2 - C 7	13.76	14.09	14.32	0.56	3.99
C 0 - C 7, E 0 - C 7	14.19	14.22	14.62	0.43	3.02
Min	7.59	7.63	7.85	0.26	2.70
Ave	12.58	12.72	13.00	0.41	3.29
Max	14.79	14.95	15.19	0.56	3.99
Rng	7.20	7.32	7.34	0.30	1.29

(a) Link failures

Failure Mode	Min	Ave	Max	<i>Rng</i>	% Ave
C 0	7.09	7.13	7.22	0.13	1.88
C 1	14.61	14.73	14.88	0.27	1.81
C 2	7.60	7.64	7.67	0.07	0.96
C 3	7.34	7.58	7.56	0.22	2.87
Min	7.09	7.13	7.22	0.07	0.96
Ave	9.16	9.27	9.33	0.17	1.88
Max	14.61	14.73	14.88	0.27	2.87
Rng	7.52	7.60	7.65	0.19	1.90

(b) Node failures

Table A.2: Uncertainty range for MPLS effective rerouting in milliseconds, which is a sum of LSP rerouting and LSP setup

Failure Mode	Min	Ave	Max	Rng	% Ave
C 2 - C 3	13.90	14.02	14.26	0.36	2.60
C 0 - C 1	6.39	6.41	6.63	0.24	3.70
C 1 - C 2, C 2 - C 7	12.92	13.24	13.46	0.54	4.10
C 0 - C 7, E 0 - C 7	12.18	12.20	12.58	0.40	3.30
C 0	6.56	6.60	6.69	0.13	2.00
C 1	13.87	13.98	14.12	0.25	1.80
C 2	6.88	6.90	6.94	0.06	0.80
C 3	6.45	6.61	6.65	0.20	3.10
Min	6.39	6.41	6.63	0.06	0.80
Ave	9.89	10.00	10.17	0.27	2.68
Max	13.90	14.02	14.26	0.54	4.10
Rng	7.51	7.61	7.64	0.49	3.30

(a) Setup

Failure Mode	Min	Ave	Max	Rng	% Ave
C 2 - C 3	0.89	0.93	0.93	0.04	4.20
C 0 - C 1	1.20	1.22	1.23	0.03	2.10
C 1 - C 2, C 2 - C 7	0.84	0.85	0.86	0.02	2.20
C 0 - C 7, E 0 - C 7	2.01	2.02	2.04	0.03	1.30
C 0	0.53	0.53	0.53	0.00	0.40
C 1	0.74	0.75	0.75	0.01	1.90
C 2	0.72	0.74	0.74	0.02	2.50
C 3	0.89	0.97	0.90	0.01	1.30
Min	0.53	0.53	0.53	0.00	0.40
Ave	0.98	1.00	1.00	0.02	1.99
Max	2.01	2.02	2.04	0.04	4.20
Rng	1.48	1.49	1.50	0.04	3.80

(b) Rerouting

Table A.3: MPLS average setup and rerouting times

Hello	Rtr. Dead	Reroute	Reroute F.
20	80	79442.13	76038.02
15	60	48002.47	49073.9
10	40	37227.09	32051.82
8	32	28034.55	31309.03
5	20	20845.01	19083.41
3	12	19009.98	10291.06
2	8	8020.19	7011.72
1	4	8061.64	3057.92

Table A.4: OSPF *HelloInterval* variation. Rerouting has the default timers: *InterfaceTransmissionDelay* = 1 s, *RetransmissionInterval* = 5 s, *spfDelay* = 5 s, *spfHoldTime* = 10 s. Fast reroute has all timers set to minimum amounts in Opnet: *InterfaceTransmissionDelay* = 1 s, *RetransmissionInterval* = 1 s, *spfDelay* = 1 s, *spfHoldTime* = 1 s. Uncertainty range is 5382.31 ms.

A.1.1 Delay

A.1.1.1 Small Link Capacity

Dest	Min	Ave	Max	Rng	% Ave
A 1	2.77	2.81	2.85	0.08	2.92
A 2	3.87	3.94	4.01	0.14	3.48
A 3	5.34	5.37	5.40	0.06	1.10
A 4	6.51	6.53	6.56	0.05	0.83
A 5	4.83	4.86	4.89	0.06	1.25
A 6	3.82	3.84	3.86	0.04	1.09
A 7	2.41	2.45	2.49	0.08	3.14
Min	2.41	2.45	2.49	0.04	0.83
Ave	4.22	4.26	4.29	0.07	1.97
Max	6.51	6.53	6.56	0.14	3.48
Rng	4.09	4.08	4.07	0.10	2.65

(a) OSPF

Dest	Min	Ave	Max	Rng	% Ave
A 1	3.02	3.03	3.03	0.00	0.13
A 2	4.31	4.35	4.39	0.09	1.98
A 3	5.93	5.99	6.04	0.11	1.87
A 4	6.63	6.75	6.87	0.24	3.56
A 5	5.15	5.16	5.18	0.04	0.72
A 6	4.02	4.04	4.06	0.04	1.09
A 7	2.67	2.67	2.68	0.01	0.49
Min	2.67	2.67	2.68	0.00	0.15
Ave	4.53	4.57	4.61	0.08	1.68
Max	6.63	6.75	6.87	0.24	3.56
Rng	3.96	4.08	4.19	0.24	5.79

(b) MPLS

Table A.5: Small link capacity. Uncertainty range for delay before failure in milliseconds

Dest	Min	Ave	Max	Rng
A 1	0.04	<i>0.06</i>	0.08	0.04
A 2	0.06	<i>0.09</i>	0.13	0.07
A 3	0.16	<i>0.27</i>	0.38	0.23
A 4	0.16	<i>0.31</i>	0.48	0.32
A 5	0.04	<i>1.00</i>	1.98	1.94
A 6	0.04	<i>0.66</i>	1.45	1.41
A 7	0.04	<i>0.26</i>	0.68	0.64
Min	0.04	<i>0.06</i>	0.08	0.04
Ave	0.08	0.38	0.74	0.66
Max	0.16	<i>1.00</i>	1.98	1.94

(a) OSPF

Dest	Min	Ave	Max	Rng
A 1	0.01	<i>0.02</i>	0.05	0.04
A 2	0.06	<i>0.25</i>	0.59	0.53
A 3	0.18	<i>0.24</i>	0.32	0.14
A 4	0.20	<i>0.55</i>	0.78	0.58
A 5	0.04	<i>0.06</i>	0.08	0.04
A 6	0.06	<i>0.07</i>	0.09	0.03
A 7	0.01	<i>0.02</i>	0.02	0.01
Min	0.01	<i>0.02</i>	0.02	0.01
Ave	0.08	0.17	0.28	0.20
Max	0.20	<i>0.55</i>	0.78	0.58

(b) MPLS

Table A.6: Small link capacity. Summary of uncertainty ranges for delay after failure in milliseconds. For each destination min, ave, max, range of the delay range for failure modes: C 2 – C 3, C 0 – C 7, E 0 – C 7, C 2 is shown.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 2 - C 3	2.80	3.93	9.53	13.37	11.94	9.11	2.46	7.59	10.92
C 0 - C 1	2.85	4.02	5.39	6.52	4.88	3.85	2.41	4.27	4.11
C 1 - C 2, C 2 - C 7	2.81	4.35	6.02	7.70	5.00	3.96	2.48	4.62	5.22
C 0 - C 7, E 0 - C 7	2.93	4.15	5.63	6.75	6.58	5.64	4.19	5.12	3.82
C 0	2.85	4.03	5.39	6.53	4.88	3.84	2.41	4.27	4.12
C 1	6.07	4.11	5.61	6.74	5.09	4.05	2.50	4.88	4.24
C 2	2.86	4.06	5.76	6.87	5.14	4.12	2.49	4.47	4.38
C 3	2.87	4.06	5.65	6.76	6.30	5.36	2.41	4.77	4.36

(a) OSPF. Uncertainty range is 0.38 ms.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 2 - C 3	3.00	4.25	8.20	8.21	5.25	4.04	2.64	5.09	5.57
C 0 - C 1	3.81	5.28	6.73	6.52	5.18	4.12	2.77	4.92	3.96
C 1 - C 2, C 2 - C 7	3.22	5.42	6.92	6.53	5.13	4.16	2.74	4.87	4.19
C 0 - C 7, E 0 - C 7	2.97	4.46	6.13	6.75	8.42	6.43	3.83	5.57	5.45
C 0	2.93	4.00	6.28	6.56	4.70	4.25	2.56	4.47	4.01
C 1	5.16	4.20	6.21	6.42	5.03	4.11	2.70	4.83	3.72
C 2	3.13	n/a	6.31	7.26	6.13	5.13	2.86	5.13	4.40
C 3	3.13	5.55	6.19	6.57	5.75	5.93	3.35	5.21	3.44

(b) MPLS. Uncertainty range is 0.17 ms.

Table A.7: Small link capacity. Delay after failure in milliseconds

A.1.1.2 Large Link Capacity

Dest	Min	Ave	Max	Rng	% Ave
A 1	16.36	16.75	17.14	0.78	4.66
A 2	20.89	21.16	21.42	0.53	2.51
A 3	25.14	25.36	25.57	0.43	1.70
A 4	25.65	25.80	25.94	0.29	1.12
A 5	20.91	21.17	21.43	0.52	2.46
A 6	20.45	20.58	20.71	0.26	1.26
A 7	16.27	16.33	16.39	0.12	0.73
Min	16.27	16.33	16.39	0.12	0.73
Ave	20.81	21.02	21.23	0.42	2.06
Max	25.65	25.80	25.94	0.78	4.66
Rng	9.38	9.47	9.55	0.66	3.92

(a) OSPF

Dest	Min	Ave	Max	Rng	% Ave
A 1	19.29	19.48	19.66	0.37	1.90
A 2	23.12	23.55	23.97	0.85	3.61
A 3	27.01	27.35	27.69	0.68	2.49
A 4	23.71	24.14	24.57	0.86	3.56
A 5	26.58	26.83	27.07	0.49	1.83
A 6	23.25	23.74	24.22	0.97	4.09
A 7	19.08	19.34	19.60	0.52	2.69
Min	19.08	19.34	19.60	0.37	1.91
Ave	23.15	23.49	23.83	0.68	2.88
Max	27.01	27.35	27.69	0.97	3.55
Rng	7.93	8.01	8.09	0.60	7.49

(b) MPLS

Table A.8: Large link capacity. Uncertainty range for delay before failure in milliseconds

Dest	Min	Ave	Max	Rng
A 1	0.52	0.71	0.90	0.38
A 2	0.70	0.84	0.97	0.27
A 3	0.75	1.05	1.34	0.59
A 4	0.92	1.05	1.17	0.25
A 5	0.46	0.61	0.75	0.29
A 6	0.43	0.65	0.87	0.44
A 7	0.21	0.51	0.80	0.59
Min	0.21	0.51	0.75	0.25
Ave	0.57	0.77	0.97	0.40
Max	0.92	1.05	1.34	0.59

(a) OSPF

Dest	Min	Ave	Max	Rng
A 1	0.35	0.77	1.18	0.83
A 2	0.09	0.39	0.69	0.60
A 3	1.03	1.38	1.72	0.69
A 4	0.36	0.93	1.49	1.13
A 5	0.30	0.83	1.36	1.06
A 6	0.31	0.74	1.16	0.85
A 7	0.49	0.82	1.15	0.66
Min	0.09	0.39	0.69	0.60
Ave	0.42	0.83	1.25	0.83
Max	1.03	1.38	1.72	1.13

(b) MPLS

Table A.9: Large link capacity. Uncertainty range for delay after failure in milliseconds. For each destination min, ave, max, range of the delay range for failure modes: C 2 – C 3, C 0 – C 7, E 0 – C 7, C 2 is shown.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 1 - C 2, C 2 - C 7	17.21	23.20	26.54	27.00	22.00	21.37	17.17	22.07	9.83
C 1	18.32	21.57	25.81	26.57	21.60	21.09	16.74	21.67	9.83

(a) OSPF. Uncertainty range is 0.77 ms.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 1 - C 2, C 2 - C 7	19.94	26.73	31.56	24.84	27.24	23.81	19.91	24.86	11.65
C 1	19.57	24.17	27.94	24.59	27.57	24.22	20.21	24.04	8.37

(b) MPLS. Uncertainty range is 0.83 ms.

Table A.10: Large link capacity. Delay after failure in milliseconds. Uncertainty range is 0.77 ms.

A.1.2 Jitter

A.1.2.1 Small Link Capacity

Dest	Min	Ave	Max	Rng	% Ave
A 1	0.41	0.45	0.49	<i>0.07</i>	16.52
A 2	0.75	0.86	0.97	<i>0.22</i>	25.73
A 3	1.31	1.33	1.36	<i>0.05</i>	3.53
A 4	1.70	1.79	1.88	<i>0.18</i>	10.00
A 5	1.33	1.35	1.36	<i>0.03</i>	2.23
A 6	0.96	0.98	1.00	<i>0.04</i>	4.20
A 7	0.46	0.50	0.53	<i>0.08</i>	15.32
Min	0.49	0.41	0.45	<i>0.03</i>	2.23
Ave	1.08	0.99	1.03	0.10	11.08
Max	1.88	1.70	1.79	<i>0.22</i>	25.73
Rng	1.40	1.29	1.34	<i>0.19</i>	23.50

(a) OSPF

Dest	Min	Ave	Max	Rng	% Ave
A 1	0.57	0.58	0.59	<i>0.02</i>	3.45
A 2	0.97	1.00	1.03	<i>0.06</i>	6.00
A 3	1.52	1.56	1.59	<i>0.07</i>	4.50
A 4	1.62	1.73	1.84	<i>0.22</i>	12.72
A 5	1.30	1.32	1.33	<i>0.03</i>	2.28
A 6	0.97	0.99	1.01	<i>0.04</i>	4.04
A 7	0.53	0.54	0.55	<i>0.02</i>	3.70
Min	0.55	0.53	0.54	<i>0.02</i>	2.28
Ave	1.13	1.07	1.10	0.07	5.24
Max	1.84	1.62	1.73	<i>0.22</i>	12.72
Rng	1.29	1.09	1.19	<i>0.20</i>	10.44

(b) MPLS

Table A.11: Small link capacity. Uncertainty range for jitter before failure in milliseconds

Dest	Min	Ave	Max	Rng
A 1	0.04	<i>0.05</i>	0.06	0.02
A 2	0.04	<i>0.09</i>	0.12	0.08
A 3	0.03	<i>0.19</i>	0.45	0.42
A 4	0.03	<i>0.13</i>	0.29	0.26
A 5	0.01	<i>0.05</i>	0.10	0.09
A 6	0.03	<i>0.07</i>	0.14	0.11
A 7	0.04	<i>0.05</i>	0.06	0.02
Min	0.01	<i>0.05</i>	0.06	0.02
Ave	0.03	0.09	0.17	0.14
Max	0.04	<i>0.19</i>	0.45	0.42

(a) OSPF

Dest	Min	Ave	Max	Rng
A 1	0.01	<i>0.02</i>	0.04	0.03
A 2	0.02	<i>0.04</i>	0.07	0.05
A 3	0.04	<i>0.07</i>	0.10	0.06
A 4	0.07	<i>0.19</i>	0.28	0.21
A 5	0.01	<i>0.02</i>	0.05	0.04
A 6	0.03	<i>0.04</i>	0.06	0.03
A 7	0.01	<i>0.02</i>	0.03	0.02
Min	0.01	<i>0.02</i>	0.03	0.02
Ave	0.03	0.06	0.09	0.06
Max	0.07	<i>0.19</i>	0.28	0.21

(b) MPLS

Table A.12: Small link capacity. Uncertainty range for jitter after failure in milliseconds. For each destination min, ave, max, range of the delay range for failure modes: C 2 – C 3, C 0 – C 7, E 0 – C 7, C 2 is shown.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 2 - C 3	0.45	0.81	1.38	1.67	1.67	1.34	0.49	1.12	1.22
C 0 - C 1	0.49	0.86	1.37	1.73	1.37	0.98	0.45	1.04	1.28
C 1 - C 2, C 2 - C 7	0.45	0.81	1.38	1.82	1.40	1.04	0.51	1.06	1.37
C 0 - C 7, E 0 - C 7	0.47	0.88	1.45	1.77	1.72	1.41	0.98	1.24	1.30
C 0	0.54	0.94	1.36	1.71	1.35	0.97	0.45	1.05	1.26
C 1	0.49	0.86	1.44	1.77	1.43	1.07	0.52	1.08	1.28
C 2	0.83	0.92	1.49	1.83	1.44	1.10	0.51	1.16	1.32
C 3	0.49	0.87	1.45	1.78	1.67	1.33	0.45	1.15	1.33

(a) OSPF. Uncertainty range is 0.09 ms.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 2 - C 3	0.59	1.01	1.56	1.85	1.33	0.99	0.55	1.13	1.30
C 0 - C 1	0.62	1.05	1.59	1.88	1.31	1.00	0.54	1.14	1.34
C 1 - C 2, C 2 - C 7	0.66	1.11	1.60	1.70	1.36	1.06	0.60	1.16	1.10
C 0 - C 7, E 0 - C 7	0.68	1.10	1.58	1.79	1.34	1.03	0.59	1.16	1.20
C 0	0.56	0.98	1.56	1.87	1.28	0.95	0.54	1.11	1.33
C 1	0.57	0.99	1.56	1.86	1.30	0.98	0.57	1.12	1.29
C 2	0.61	0.97	1.57	1.73	1.29	1.05	0.55	1.11	1.18
C 3	0.63	1.04	1.61	1.89	1.35	1.07	0.56	1.16	1.33

(b) MPLS. Uncertainty range is 0.06 ms.

Table A.13: Small link capacity. Jitter after failure in milliseconds.

A.1.2.2 Large Link Capacity

Dest	Min	Ave	Max	Rng	% Ave
A 1	0.015	0.016	0.016	<i>0.001</i>	6.452
A 2	0.018	0.020	0.021	<i>0.003</i>	15.385
A 3	0.022	0.023	0.024	<i>0.002</i>	8.696
A 4	0.020	0.022	0.024	<i>0.004</i>	18.182
A 5	0.019	0.020	0.021	<i>0.002</i>	10.000
A 6	0.020	0.021	0.021	<i>0.001</i>	4.878
A 7	0.018	0.019	0.019	<i>0.001</i>	5.405
Min	0.016	0.015	0.016	<i>0.001</i>	4.878
Ave	0.021	0.019	0.020	0.002	9.857
Max	0.024	0.022	0.023	<i>0.004</i>	18.182
Rng	0.008	0.007	0.008	<i>0.003</i>	13.304

(a) OSPF

Dest	Min	Ave	Max	Rng	% Ave
A 1	0.017	0.018	0.018	<i>0.001</i>	5.71
A 2	0.020	0.021	0.021	<i>0.001</i>	4.88
A 3	0.023	0.024	0.024	<i>0.001</i>	4.26
A 4	0.021	0.022	0.023	<i>0.002</i>	9.09
A 5	0.021	0.022	0.022	<i>0.001</i>	4.65
A 6	0.022	0.023	0.024	<i>0.002</i>	8.70
A 7	0.019	0.020	0.020	<i>0.001</i>	5.13
Min	0.018	0.017	0.018	<i>0.001</i>	4.26
Ave	0.022	0.020	0.021	0.001	6.06
Max	0.024	0.023	0.024	<i>0.002</i>	9.09
Rng	0.006	0.006	0.006	<i>0.001</i>	4.84

(b) MPLS

Table A.14: Large link capacity. Uncertainty range for jitter before failure in milliseconds

Dest	Min	Ave	Max	Rng
A 1	0.001	<i>0.002</i>	0.003	0.002
A 2	0.003	<i>0.003</i>	0.003	0.000
A 3	0.001	<i>0.002</i>	0.002	0.001
A 4	0.001	<i>0.002</i>	0.003	0.002
A 5	0.001	<i>0.001</i>	0.001	0.000
A 6	0.000	<i>0.000</i>	0.001	0.001
A 7	0.001	<i>0.001</i>	0.001	0.000
Min	0.000	<i>0.000</i>	0.001	0.000
Ave	0.001	0.002	0.002	0.001
Max	0.003	<i>0.003</i>	0.003	0.002

(a) OSPF

Dest	Min	Ave	Max	Rng
A 1	0.001	<i>0.001</i>	0.001	0.000
A 2	0.000	<i>0.000</i>	0.001	0.001
A 3	0.000	<i>0.001</i>	0.001	0.001
A 4	0.001	<i>0.002</i>	0.002	0.001
A 5	0.001	<i>0.001</i>	0.001	0.000
A 6	0.000	<i>0.001</i>	0.001	0.001
A 7	0.000	<i>0.000</i>	0.000	0.000
Min	0.000	<i>0.000</i>	0.000	0.000
Ave	0.000	0.001	0.001	0.001
Max	0.001	<i>0.002</i>	0.002	0.001

(b) MPLS

Table A.15: Large link capacity. Uncertainty range for jitter after failure in milliseconds. For each destination min, ave, max, range of the delay range for failure modes: C 2 – C 3, C 0 – C 7, E 0 – C 7, C 2 is shown.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 1 - C 2, C 2 - C 7	0.015	0.021	0.025	0.024	0.020	0.021	0.019	0.021	0.010
C 1	0.018	0.021	0.024	0.023	0.021	0.021	0.019	0.021	0.006

(a) OSPF. Uncertainty range is 0.002 ms.

Failure Mode	A 1	A 2	A 3	A 4	A 5	A 6	A 7	Ave	Rng
C 1 - C 2, C 2 - C 7	0.019	0.023	0.026	0.023	0.021	0.023	0.021	0.022	0.007
C 1	0.020	0.021	0.023	0.022	0.021	0.021	0.019	0.021	0.004

(b) MPLS. Uncertainty range is 0.001 ms.

Table A.16: Large link capacity. Jitter after failure in milliseconds.

Appendix B

Miscellaneous Access Network Results

B.1 Rerouting

Failure Mode	Min	Ave	Max	<i>Rng</i>	% Ave
S 0 - S 1	3931.20	3992.78	4054.36	<i>123.16</i>	3.08
S 0 - S 1, S 1 - S 3	3922.61	4011.85	4101.09	<i>178.48</i>	4.45
S 0 - S 1, S 2 - S 6	3943.56	3965.60	3987.64	<i>44.08</i>	1.11
S 0 - S 1, S 2 - S 3	3923.71	4016.95	4110.19	<i>186.48</i>	4.64
Min	3922.61	3965.60	3987.64	<i>44.08</i>	1.11
Ave	3930.27	3996.80	4063.32	133.05	3.32
Max	3943.56	4016.95	4110.19	<i>186.48</i>	4.64
Rng	20.95	51.35	122.55	<i>142.40</i>	3.53

(a) Link failures

Failure Mode	Min	Ave	Max	<i>Rng</i>	% Ave
S 1	3912.14	4004.74	4097.33	<i>185.19</i>	4.62
S 3	3951.14	4039.21	4127.28	<i>176.14</i>	4.36
Min	3912.14	4004.74	4097.33	<i>176.14</i>	4.36
Ave	3931.64	4021.97	4112.31	180.67	4.49
Max	3951.14	4039.21	4127.28	<i>185.19</i>	4.62
Rng	39.00	34.48	29.95	<i>9.05</i>	0.26

(b) Node failures

Table B.1: Uncertainty range for RSTP rerouting times in milliseconds

Failure Mode	Min	Ave	Max	Rng	% Ave
S 0 - S 1	3.39	15.04	7.14	0.41	2.71
S 0 - S 1, S 1 - S 3	4.64	7.62	11.14	0.26	3.45
S 0 - S 1, S 2 - S 6	3.02	14.16	8.59	0.56	3.98
S 0 - S 1, S 2 - S 3	5.19	14.26	8.79	0.43	3.01
Min	3.02	7.62	7.14	0.26	2.71
Ave	4.06	12.77	8.92	0.42	3.29
Max	5.19	15.04	11.14	0.56	3.98
Rng	2.17	7.42	4.00	0.30	1.27

(a) Link failures

Failure Mode	Min	Ave	Max	Rng	% Ave
S 1	5.09	7.35	13.63	0.14	1.84
S 3	8.35	15.72	18.13	0.28	1.81
Min	5.09	7.35	13.63	0.14	1.81
Ave	6.72	11.54	15.88	0.21	1.82
Max	8.35	15.72	18.13	0.28	1.84
Range	3.26	8.37	4.50	0.15	0.03

(b) Node failures

Table B.2: Uncertainty range for MPLS effective rerouting in milliseconds, which is a sum of LSP rerouting and LSP setup

Failure Mode	Min	Ave	Max	Rng	% Ave
S 1	2.41	14.02	6.09	0.36	2.60
S 3	3.53	6.41	9.84	0.24	3.70
S 0 - S 1	2.16	13.24	7.62	0.54	4.10
S 0 - S 1, S 1 - S 3	3.16	12.20	6.71	0.40	3.30
S 0 - S 1, S 2 - S 6	4.44	6.60	12.78	0.13	2.00
S 0 - S 1, S 2 - S 3	6.80	13.98	16.20	0.25	1.80
Min	2.16	6.41	6.09	0.13	1.80
Ave	3.75	11.08	9.87	0.32	2.92
Max	6.80	14.02	16.20	0.54	4.10
Rng	4.64	7.61	10.11	0.41	2.30

(a) Setup

Failure Mode	Min	Ave	Max	Rng	% Ave
S 1	0.98	1.02	1.05	0.04	4.20
S 3	1.11	1.21	1.30	0.03	2.10
S 0 - S 1	0.86	0.92	0.97	0.02	2.20
S 0 - S 1, S 1 - S 3	2.03	2.06	2.08	0.03	1.30
S 0 - S 1, S 2 - S 6	0.65	0.75	0.85	0.00	0.40
S 0 - S 1, S 2 - S 3	1.55	1.74	1.93	0.03	1.90
Min	0.65	0.75	0.85	0.00	0.40
Ave	1.20	1.28	1.36	0.03	2.02
Max	2.03	2.06	2.08	0.04	4.20
Rng	1.38	1.31	1.23	0.04	3.80

(b) Rerouting

Table B.3: MPLS average setup and rerouting times

B.2 Delay

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	16.10	15.82	15.49	15.47	15.03	15.06	15.42	15.21	15.45	1.07
S 3	15.69	15.35	15.19	15.29	14.77	15.02	15.03	14.85	15.15	0.92
S 0 - S 1	15.79	15.67	15.52	15.49	15.48	15.29	15.35	15.50	15.51	0.50
S 0 - S 1, S 1 - S 3	15.97	15.96	15.88	15.89	15.44	15.39	15.64	15.54	15.71	0.58
S 0 - S 1, S 2 - S 6	15.95	15.75	15.57	15.43	15.53	15.26	15.35	15.57	15.55	0.69
S 0 - S 1, S 2 - S 3	20.28	20.11	15.63	15.80	15.75	15.70	15.56	15.65	16.81	4.72

(a) RSTP. Uncertainty range is 0.38 ms.

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	8.43	8.46	8.41	8.30	8.19	8.33	8.27	8.22	8.33	0.27
S 3	8.26	8.15	8.10	8.09	8.10	8.20	8.15	8.13	8.15	0.17
S 0 - S 1	10.06	10.03	10.05	9.90	8.25	8.31	8.21	8.24	9.13	1.85
S 0 - S 1, S 1 - S 3	8.38	8.34	8.32	8.31	8.11	8.30	8.23	8.22	8.28	0.27
S 0 - S 1, S 2 - S 6	9.97	9.95	10.08	9.94	8.17	8.54	8.20	7.98	9.10	2.10
S 0 - S 1, S 2 - S 3	10.04	10.01	10.00	9.87	8.23	8.28	8.21	8.19	9.10	1.85

(b) MPLS. Uncertainty range is 0.13 ms.

Table B.4: Delay after failure in milliseconds

Source	Min	Ave	Max	Rng
N 1	0.55	0.69	0.82	<i>0.27</i>
N 2	0.24	0.34	0.44	<i>0.20</i>
N 3	0.41	0.56	0.71	<i>0.30</i>
N 4	0.52	0.65	0.79	<i>0.27</i>
N 5	0.25	0.53	0.80	<i>0.55</i>
N 6	0.17	0.35	0.53	<i>0.36</i>
N 7	0.41	0.65	0.89	<i>0.48</i>
N 8	0.24	0.54	0.83	<i>0.59</i>
Min	0.17	0.34	0.44	<i>0.20</i>
Ave	0.35	0.54	0.73	0.38
Max	0.55	0.69	0.89	<i>0.59</i>

(a) RSTP

Source	Min	Ave	Max	Rng
N 1	0.07	0.13	0.18	<i>0.11</i>
N 2	0.11	0.16	0.21	<i>0.10</i>
N 3	0.12	0.18	0.24	<i>0.12</i>
N 4	0.12	0.20	0.29	<i>0.17</i>
N 5	0.07	0.13	0.19	<i>0.12</i>
N 6	0.14	0.23	0.33	<i>0.19</i>
N 7	0.09	0.17	0.24	<i>0.15</i>
N 8	0.05	0.09	0.14	<i>0.09</i>
Min	0.05	0.09	0.14	<i>0.09</i>
Ave	0.10	0.16	0.23	0.13
Max	0.14	0.23	0.33	<i>0.19</i>

(b) MPLS

Table B.5: Summary of uncertainty ranges for delay after failure in milliseconds

B.3 Jitter

Source	Min	Ave	Max	Rng
N 1	0.07	0.11	0.14	<i>0.07</i>
N 2	0.05	0.10	0.14	<i>0.09</i>
N 3	0.05	0.09	0.13	<i>0.08</i>
N 4	0.07	0.10	0.13	<i>0.06</i>
N 5	0.10	0.12	0.14	<i>0.04</i>
N 6	0.09	0.12	0.15	<i>0.06</i>
N 7	0.03	0.07	0.11	<i>0.08</i>
N 8	0.05	0.08	0.10	<i>0.05</i>
Min	0.03	0.07	0.10	<i>0.04</i>
Ave	0.06	0.10	0.13	0.07
Max	0.10	0.12	0.15	<i>0.09</i>

(a) RSTP

Source	Min	Ave	Max	Rng
N 1	0.05	0.07	0.09	<i>0.04</i>
N 2	0.01	0.03	0.06	<i>0.05</i>
N 3	0.04	0.07	0.10	<i>0.06</i>
N 4	0.04	0.05	0.06	<i>0.02</i>
N 5	0.02	0.04	0.06	<i>0.04</i>
N 6	0.02	0.04	0.07	<i>0.05</i>
N 7	0.04	0.08	0.11	<i>0.07</i>
N 8	0.04	0.06	0.08	<i>0.04</i>
Min	0.01	0.03	0.06	<i>0.02</i>
Ave	0.03	0.06	0.08	0.05
Max	0.05	0.08	0.11	<i>0.07</i>

(b) MPLS

Table B.6: Summary of uncertainty ranges for jitter after failure in milliseconds

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	3.76	3.80	3.75	3.75	3.77	3.67	3.78	3.72	3.75	0.13
S 3	3.76	3.73	3.79	3.82	3.71	3.70	3.77	3.73	3.75	0.12
S 0 - S 1	3.72	3.75	3.78	3.76	3.77	3.75	3.75	3.78	3.76	0.06
S 0 - S 1, S 1 - S 3	3.81	3.80	3.81	3.81	3.72	3.71	3.77	3.74	3.77	0.10
S 0 - S 1, S 2 - S 6	3.76	3.74	3.76	3.76	3.75	3.81	3.76	3.75	3.76	0.07
S 0 - S 1, S 2 - S 3	4.21	4.18	3.79	3.78	3.71	3.67	3.79	3.79	3.87	0.54

(a) RSTP. Uncertainty range is 0.38 ms.

Failure Mode	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	Ave	Rng
S 1	2.26	2.25	2.28	2.26	2.22	2.26	2.25	2.27	2.26	0.06
S 3	2.24	2.22	2.29	2.25	2.25	2.23	2.25	2.25	2.25	0.07
S 0 - S 1	2.45	2.48	2.50	2.46	2.26	2.26	2.27	2.24	2.37	0.26
S 0 - S 1, S 1 - S 3	2.21	2.22	2.28	2.25	2.23	2.27	2.25	2.23	2.24	0.07
S 0 - S 1, S 2 - S 6	2.49	2.48	2.49	2.47	2.22	2.25	2.33	2.20	2.37	0.29
S 0 - S 1, S 2 - S 3	2.45	2.47	2.48	2.47	2.30	2.29	2.27	2.26	2.37	0.22

(b) MPLS. Uncertainty range is 0.13 ms.

Table B.7: Jitter after failure in milliseconds

Appendix C

Reliability Concepts

Much of the information presented in this chapter comes from Shooman [74], supported by other references [11, 46, 82, 83].

C.1 Fundamental Concepts

Availability function $A(t)$ is the probability that the system is operational at time t . With $A(250) = 0.95$, if 100 units operated for 250 hours, on average 95 will be operating at the end of that time and 5 units will be undergoing repair.

Reliability function $R(t)$ is the probability that the system has operated without failure over the interval $[0, t]$. With $R(250) = 0.95$, if 100 units operated during $[0, 250]$ hours time period 95 units will have no failure and 5 units will have failed at some time.

From the two definitions $R(t) \leq A(t)$, as reliability is not concerned with when a unit fails whereas for availability the unit may be fixed before the time t . Essentially $R(t)$ does not allow for repair and $A(t)$ does. The special case $R(t) = A(t)$ is when a single unit is irreparable. Re-

pairable systems with redundant components do improve $R(t)$. For example, a simple two unit parallel system will benefit from repairs as if unit one fails and gets repaired before the second unit fails the system will not fail in that time interval.

For networks concepts of availability and reliability are often considered the same. Steady state availability

$$A_{ss} = \lim_{t \rightarrow \infty} A(t)$$

is one common measure. $A(t)$ is typically a decreasing function, with the highest availability of 1 at the start with all units operational. $A(t)$ decreases with time to some steady value (very close to one in a good system) called a steady state availability A_{ss} . The reason A_{ss} is used so frequently is because it gives a lower bound on availability and it is simpler to calculate.

$A_{ss} = \frac{\mu}{\lambda + \mu}$ is a steady state availability of single component, where λ is the failure or hazard rate and μ is the repair rate. A_{ss} is often written as

$$Availability = \frac{Uptime}{Downtime + Uptime}$$

Markov modeling is needed in more complicated systems to determine A_{ss} correctly.

More mathematically,

$$R(t) = e^{\int_0^t z(\xi) d\xi}, \text{ where } z(t) \text{ is the hazard function}$$

In most cases the hazard function is assumed to be constant $z(t) = \lambda$. This makes calculations less complicated and it fits most circumstances

extremely well. Generally, $z(t)$ follows a bathtub shape. In the early stages of deployment of an element more failures occur due to new parts defects, incompatibility, and so on. As time progresses those early failures disappear and we enter the main lifespan of the element. In this middle stage relatively few random failures occur at approximately the same level throughout the time period. In the last stage of the element's lifespan wearout failures start to happen and increase at some variable rate. Thus, if we assume an element always operates in its middle stage, the constant failure rate assumption should hold.

There are more complex models, such as linear and logarithmic models, but those are not necessary for the analysis in here. Using a constant failure rate in the definition of reliability we get $R(t) = e^{-\lambda t}$.

Hazard rate λ is usually measured in failures in time (FITs) in telecommunication industry. One FIT is one failure per a billion hours of operation. For the earlier example $\lambda = 0.05/250 = 2 \times 10^{-4} = 200,000 \times 10^{-9}$ failures per billion hours. Now $R(250) = e^{-2 \times 10^{-4} \times 250} = 0.95$, which agrees with the reliability figure.

Besides reliability Mean Time To Failure (MTTF) is another measure which can be compared when deciding between different systems. MTTF can be easier to compute by using Laplace theory. MTTF is defined as

$$MTTF = \int_0^{\infty} R(t) dt$$

The MTTF for a single element with reliability expression $R(t) = e^{-\lambda t}$ is

$$MTTF = \int_0^{\infty} e^{-\lambda t} dt = \frac{1}{\lambda}$$

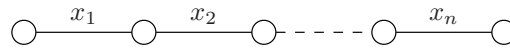


Figure C.1: Series configuration

For comparison purposes MTTF of a two element parallel system with repair is

$$MTTF = \frac{3\lambda + \mu}{2\lambda^2}$$

C.2 Basic Network Configurations

C.2.1 Series

A series configuration is one in which its elements are distributed in a chain, one after another as in Figure C.1. Failure of any one of the elements will cause system failure. Thus reliability of a series system is $R = P(x_1, x_2, \dots, x_n)$ and provided all units are identical $R = P(x_1)P(x_2)\dots P(x_n)$. It follows that a series configuration is worse than a single element.

Series configuration is useful when analysing or designing a large system. Such system is usually decomposed into a number of subsystems linked by couplers. In effect the subsystem form a series configuration.

C.2.2 Parallel

A parallel configuration is one in which elements are distributed such that the only one of the elements is needed to be operational for the system to function correctly. Figure C.2 shows a parallel configuration. Reliability of a parallel system is the probability of the sum of probabilities must be expanded using Inclusion-Exclusion Principle. Inclusion-

Exclusion Principle is very well known from set theory, it is as follows:

$$\begin{aligned}
 R = P(x_1 + x_2 + \dots + x_j) = & \sum_{1 \leq i \leq j} x_i \\
 & - \sum_{1 \leq l < m \leq j} (x_l \cdot x_m) \\
 & + \sum_{1 \leq n < o < p \leq j} (x_n \cdot x_o \cdot x_p) \\
 & \vdots \\
 & - (-1)^{j-1} (x_1 \cdot x_2 \cdot \dots \cdot x_j)
 \end{aligned}$$

For example:

$$P(x_1 + x_2 + x_3) = x_1 + x_2 + x_3 - x_1x_2 - x_1x_3 - x_2x_3 + x_1x_2x_3$$

The indices for each term are given by the number of combinations possible $\binom{n}{k}$. For example, the second term in the general expression contains $\binom{j}{2}$ expressions.

There are two categories of parallel systems. A parallel (or online or hot standby) system has each of the redundant units operating at all times. A (cold) standby system is superior since the redundant units are not powered and so cannot fail. A standby system is superior as time to system failure of a parallel system is the longest time to failure of one of n components. In a standby configuration the system failure occurs after the sum of time to failure of each redundant component.

Both parallel and standby systems have disadvantages. It is rare that a big system can be constructed using a parallel configuration. Usually it is several redundant systems joint together by couplers. Even if all those subsystems employ parallel redundancy, the reliability of the system is

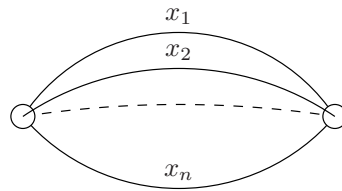


Figure C.2: Parallel configuration

limited to the reliability of the couplers. Calculations show that coupler reliability must be greater than the reliability of individual components.

Standby systems must switch to use the redundant component in the presence of failure. Such a switch is never perfect and must detect the failure of the active element, switch power and I/O to the standby element efficiently. The effect of the switch is to multiply the existing reliability by its own.

C.2.3 r-out-of-n

A configuration of units is called r-out-of-n when r units of the total n units must be functional for the system to continue its operation. Whenever there are less than r operational units left the system fails. This setup is a version of parallel redundancy which is more restricted. This type of redundancy is often employed as it provides big cost savings when compared to a fully redundant parallel setup. Examples of use of the r-out-of-n system can be found in optical networks with the (N:M) fault-tolerance scheme, network cabling, and many others.

The reliability expression is governed by the binomial coefficient as any combination of r out of n units may be operational. Therefore $R = P_s = \sum_{k=r}^n \binom{n}{k} p^k (1-p)^{n-k}$ as $B(r; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$ is the probability that r units succeed with probability p.

C.3 Reliability in Graphs

Reliability of a large network is not easy to assess due to the multiple number of paths between any two endpoints. A brute force method is to enumerate each path and do the reliability calculations using all possible combinations of these paths. For very small network this may be possible, but for larger structures a method from graph theory is used.

The method uses the notion of tie and cut sets. A tie-set is a set of branches in the a graph or network connecting two nodes. A cut-set is a set of branches, which when removed interrupts all connections between two nodes. Depending on the network topology it may be easier to find its tie-sets or its cut-sets. Reliability is calculated as follows

$$R = 1 - P_f = 1 - P(C_1 + C_2 + \dots + C_i) = P(T_1 + T_2 + \dots + T_j),$$

where C_k is the probability of a cut-set and T_l is the probability of a tie-set. Both probabilities are computed based on the probability of failure of each network branch and node that is part of the cut-set or tie-set.

It is difficult enough to calculate the reliability between only two nodes. But in general k-node reliability is required for full assessment of reliability. Approximation techniques based on heuristics and graph transformations exist. The problem is still too extensive to do by hand and a reliability computer package is likely to be used.

C.3.1 N-modular Redundancy

Practical parallel systems present the difficulties of either couplers in hot standby system or switches in cold standby system. N-modular redun-

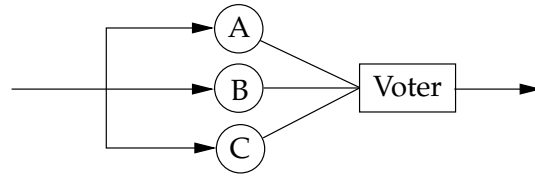


Figure C.3: Triple modular redundancy configuration with a non-redundant voter

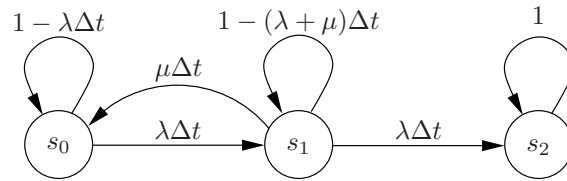


Figure C.4: Markov reliability model for two identical parallel elements

dancy is a simpler and more reliable mechanism to deploy a parallel configuration by taking advantage of digital nature of system inputs. The basic idea is that the same input is used for at least N units, each one produces and output which is put into a voter. The voter uses majority voting of its inputs to produce a correct output. Figure C.3 gives the most common N -modular redundancy scheme, Triple Modular Redundancy (TMR).

Redundant voters may also be used to improve reliability. Redundant voters are easier to implement than redundant couplers.

C.4 Markov Modeling

Markov modeling for reliability and availability is described in Chapter 3 and Appendix B of Shooman [74].

The essence of Markov modeling for reliability and availability is to build a model of states the system can enter and specify transitional probabilities of entering each state.

Figure C.4 illustrates the Markov model for a simple two element parallel system consists of four states: state s_0 with no elements failed; state s_1 with either one elements failed (counts as two); and state s_2 with both elements failed. Transitions from each state are specified by failure rates λ and repair rates μ . The sum of transitional probabilities at each state must equal to 1. Once all states and transitions are specified, a system of first order differential equations with respect to time are written down. The rules for formulating the differential equations are on pages 113 and 450 of Shooman [74]. The rule can be summarised as equating the derivative of the probability at any node to the sum of the transmissions coming into the node (or equivalently, transmissions coming out, with an opposite sign in front). The initial conditions for the model is that we start with a good system so probability of the first state is 1. The probability of all other states is set to 0. Solution to such a system of first order differential equations can be arrived at using a technique of Laplace transforms, which convert the differential system (in time domain) into algebraic set of equations (referred to as frequency domain). The new set of equations is solved and then transferred back to time domain. Mathematics computer packages such as Maple, Matlab, or Mathematica are likely to be used for systems with more than four states.

A slightly different Markov model is needed to calculate availability. The difference is that a repair transition from the system down state is allowed. The change to Figure C.4 is the addition of $s_2 \rightarrow s_1$ transition with a new repair rate. The rest of the procedure remains the same as for reliability calculation, with the outcome being an availability function.

To calculate A_{ss} several methods exist. One of the simpler methods is setting each time derivative to 0 in the system of differential equations

formed from the Markov graph.

Markov models can deal with different failures such as coverage (detection) failures and recovery failures. This is done by adding more states representing different failures and adding appropriate transitions. Chapters 3 and 4 of Ali [11] give examples of such models in the context of switching systems.

C.5 Fault-Tree Analysis

To analyse a system it needs to be decomposed into either a reliability block diagram or a fault tree [74].

Reliability block diagram diagrams represent elements and their connections. Configurations such as series or parallel can be recognised from such a diagram and appropriate calculations can be made.

Fault tree diagrams represent different failure modes of the system and its parts. Failure modes that lead to specific effects are organised in a hierarchical tree diagram. A failure at one level of the tree may trigger another failure one level above it. In such manner chains of events may be carefully traced with probabilities assigned to effects from failure modes and hazards. fault trees provide a view of how the system may fail from the failure modes point of view, rather than physical layout of elements shown by a reliability block diagram.

C.6 Failure Mode and Effect Analysis

Failure mode and effect analysis is a tabular procedure that identifies single-event chains that trigger distinct failures. Such analysis is done

early in the design process and serves as an input for many further procedures such as constructing fault trees, reliability block diagrams, and testing.

C.7 Reliability Optimisation

In this context optimisation means to design the system with the best possible reliability or availability given some constraints. Often only the financial cost constraint is used, but there are other often implicit constraints including network topology, time, and complexity.

To optimise reliability or availability in a complex system the system must be decomposed or apportioned into subsystem, which can in turn be optimised. Chapter 7 of Shooman [74] discusses reliability optimisation and states five main approaches to apportionment.

Equal weighting apportionment states that each subsystem must have the same reliability. This is an easy first attempt to apportionment.

Relative difficulty method improves on equal weighting by factoring in relative cost and difficulty of each subsystem. Usually rough heuristics are used to estimate costs and difficulty.

Relative failure rates apportionment requires knowledge of failure rates of each subsystem. The method is easier than relative difficulty and is likely to be more accurate.

Albert's method is an algorithm for apportionment. First an estimate of each subsystem reliability is needed, with the system reliability lower than the target. The subsystems that need to have a higher reliability than estimated to meet the target have their reliability increased so that they are equal. The algorithm is a good way to calculate the minimum

reliability of the subsystems required to meet the target system reliability (rather than trial and error).

Finally apportionment can be multi-layered in large systems. In such cases stratified optimisation is used, which reiterates optimisation in required subsystems. Engineering process of building a system is usually a trial-and-error process. This is due to many parameter being unknown prior to the design, difficult constraints, estimation of maintenance costs and knock-on effects, and so on. Thus often upper and lower bounds on the design are calculated using the techniques for apportionment. Several designs can then be modeled to give reliability within the bounds. The best design in terms of reliability and all the constraints is selected.

C.8 Network Reliability Implementation

This section shows how the theoretical network reliability computation was implemented in this thesis.

According to Section C.2.2, the general method for working out reliability between two points in a network is to compute the probability of the sum of all paths probabilities. All paths between source and destination must be found, where a path is a sequence of links and nodes from source to destination without visiting the same node more than once.

Practically this posed two problems. First, all the correct paths must be found. Second, the probability of the sum of the path individual reliabilities must be calculated. The first problem was solved recursively as most standard graph algorithms, with graph construction such that no loops were allowed in a path. The second problem is more difficult

than appears, because the probability of the sum of probabilities must be expanded using Inclusion-Exclusion Principle, stated in Section C.2.2.

To implement the Inclusion-Exclusion expansion an algorithm from Rosen [67] was used to systematically generate all combinations of n elements, taken k at a time, that is $\binom{n}{k}$. This way each term in the expression can be generated. The pseudocode is shown below:

```
1: list of individual probabilities
2: for all  $i$  such that  $1 \leq i \leq listSize$  do
3:   Compute all combinations of size  $i$  out of  $listSize$ ,  $\binom{i}{listSize}$ 
   {Using implementation of Rosen [67] algorithm}
4:   for all combinations of size  $i$  do
5:     for all indecies in one combination do
6:       Multiply appropriate values in list
7:     end for
8:     Add the combination to the totalProb with appropriate sign
9:   end for
10: end for
11: return totalProb
```

Figure C.5: Inclusion-Exclusion Principle implementation

Figure C.5 is exact, but it is solving an NP-hard problem, for which any exact solution is exponential in the number of input elements. This means that for a medium to large network the method takes too long to compute the answer. In complex networks, such as electricity grids and PSTN, complicated (unpublished) heuristic algorithms are used. The execution time of such heuristics depends on the required precision of the answer. Investigation of such heuristics and their efficiencies is not part of my thesis.