The Conceptualisation, Design and Implementation of a National Research and Education Network

by

Sybrand Abraham Brink

Thesis presented in partial fulfilment of the requirements for the degree

Master of Science in Engineering

at the University of Stellenbosch

Supervisor:

Dr. R. Wolhuter

December 2009

DECLARATION

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2009

Abstract

This thesis discusses the steps that should be taken when conceptualising, designing and implementing a *South African national research and education network*. It looks at all aspects of specifying transmission mediums and technology, choosing hardware and interfaces and detailing node layout and optic fibre network architecture. It examines the design of a national transport backbone, as well as a metropolitan optic fibre network. It considers ways in which the network may be made future-proof, both in terms of capacity and technology, and suggests ways to integrate the incumbent tertiary and education network into a fully-fledged, world-class network. Finally it benchmarks the network against similar networks in other countries, most notably the United States of America.

Opsomming

Hierdie tesis bespreek die stappe wat geneem behoort te word wanneer 'n *Suid Afrikaanse nasionale navorsings en tersiêre opvoedings netwerk* gekonseptualiseer, ontwerp en geïmplementeer word. Dit beskou al die aspekte rondom die spesifiseering van transmissie mediums en -tegnologie, die keuse van hardeware en koppelvlakke en die detailering van nodus uitlegte en optiese vesel netwerk argitektuur. Dit ondersoek die ontwerp van 'n nasionale transmissie ruggraat netwerk, sowel as 'n metropolitaanse optiese vesel netwerk. Dit oorweeg maniere waarmee 'n netwerk toekomsproef gemaak kan word, beide in terme van kapasiteit sowel as tegnologie, en stel maniere voor waarmee die huidige tersiêre opvoeding en navorsingsnetwerk geïntegreer kan word na 'n volwaardige, wêreldklas netwerk. Uiteindelik word die netwerk ontwerp vergelyk met soortgelyke netwerke in ander lande, spesifiek in die Verenigde State van Amerika.

Acknowledgements

I would like to thank the following people for their help and support in completing this project and thesis:

- My wife Petro, for her patience and support through the long nights and trips away
- Duncan Martin, the CEO of TENET, and Andrew Alston, CTO of TENET, without whose vision and effort the SANReN would never have been built
- My supervisor Dr. R. Wolhuter, for guidance and insight

Table of Contents

List of Figures

Nomenclature

3G - Third generation mobile communications

AARNet - Australian Academic & Research Network

ADM – Add/Drop Multiplexer

ADSL - Asymmetric Digital Subscriber Loop

API - Access Point Identifier

AS - Autonomous System

ASE - Amplified Spontaneous Emission

ATM - Asynchronous Transfer Mode

ATOM - Any Transport Over MPLS

AU - Administrative Unit

AUP - Acceptable Use Policy

BER - Bit Error Rate

BGP - Border Gateway Protocol

BPDU - Bridge Protocol Data Unit

CAC - Call Admission Control

CANARY - Canadian Advanced Network and Research for Industry and Education

CAREN - Central Asian Research and Education Network

ccTV - Closed circuit television

CDVT - Cell Delay Variation Tolerance

CHPC - Centre for High Performance Computing

CINX - Cape Town Internet Exchange

CMOS - Complementary Metal-Oxide Semiconductor

CoCT - City of Cape Town

COLT - City Of London Telecom

CP - Customer Premises

CRC - Cyclic Redundancy Checking

CRS - Carrier Routing System

CSIR - Centre for Scientific and Industrial Research

CTO - Chief Technical Officer

DANTE - Delivery of Advanced Network Technology to Europe

dB - Decibel

DSL - Digital Subscriber Line

DST - Department of Science and Technology

DWDM - Dense Wavelength Division Multiplexing

EDFA - Erbium Doped Fibre Amplifier

EDGE - Enhanced Data rates for GSM Evolution

FDP - Fibre Distribution Point

FE - Fast Ethernet

FEC - Forward Error Correction or Forward Equivalence Class

FIB - Forwarding Information Base

FICON - Fibre Connectivity

FWM - Four-Wave Mixing

GbE - Gigabit Ethernet

GCC - General Communications Channel

GFP - Generic Framing Procedure

GRE - Generic Routing Encapsulation

HartRAO - Hartebeeshoek Radio Astronomy Observatory

HDPE - High Density Polyethylene

ICANN - Internet Corporation for Assigned Names and Numbers

I-ECNS - Individual Electronic Communications Network Services

IPoDWDM - Internet Protocol over Dense Wavelength Division Multiplexing

IPSEC - IP Security

IPv4 - Internet Protocol version 4

IPv6 - Internet Protocol version 6

IRU - Indefeasible Right of Use

IS-IS - Intermediate system to intermediate system

ITIL - Information Technology Infrastructure Library

ITU - International Telecommunication Union

JINX - Johannesburg Internet Exchange

LABS - Laboratory for Accelerator Based Sciences

LAN - Local Area Network

LCAS - Link Capacity Adjustment Scheme

LCR - Least-Cost Routing

LDP - Label Distribution Protocol

LER - Label Edge Router

LONI - Louisiana Optical Network Initiative

LSP - Label Switched Path

LSR - Label Switch Router

MAC - Media Access Control

MAREN - Malawi Research and Education Network

MCM - Marine & Coastal Management

MIT - Massachusetts Institute of Technology

MOA - Memorandum of Understanding

MoRENet - Mozambique Research and Education Network

MPLS - Multi-Protocol Label Switching

MRC - Medical Research Council

MSOH - Multiplexer Section Overhead

MTBF - Mean Time Between Failures

MTTR - Mean Time To Repair

KENET - Kenya Education Network

NGSDH - Next Generation Synchronous Digital Hierarchy

NLC - National Laser Centre

NLSA - National Library of SA

NMMU - Nelson Mandela Metropolitan University

NOC - Network Operations Centre

NRF - National Research Foundation

NSF - National Science Foundation

NWU - North West University

NZ-DSF - Non-Zero Dispersion Shifted Fibre

ODF - Optical fibre Distribution Frame

ODU - Optical Data Unit

ONS - Optical Network Switch

OPGW - Optical fibre composite overhead ground wire

OPU - Optical Path Unit

OSI - Open System Interconnection

OSNR - Optical Signal-to-Noise Ratio

OSP - Outside Plant

OSPF - Open Shortest Path First

OTU - Optical Transport Unit

PA-BA - Pre-Amp, Booster Amp

PABX - Private Automatic Branch Exchange

Petaflop – 10^{15} floating point operations per second

PER – Provider Edge Router

PMD - Polarisation Mode Dispersion

PNNI - Private Network-to-Network Interface

POTS - Plain Old Telephony Service

PR - Provider Router

PRI - Primary Rate Interface (Integrated Services Digital Network)

PSTN - Public Switched Telephony Network

PVC - Permanent Virtual Circuits

QinQ - IEEE 802.1q

QoS - Quality of Service

RIB - Routing Information Base

ROADM - Remotely reconfigurable Optical Add/Drop Multiplexer

RSOH - Regenerator Section Overhead

RSTP - Rapid Spanning Tree Protocol

RSVP-TE - Resource reservation protocol – traffic engineering

RU - Rhodes University

RwEDNet - Rwanda Education Network

SAAO - South African Astronomical Observatory

SAC - Satellite Applications Centre

SAN - Storage Area Networks

SANAP - SA National Antarctic Program

SANRAL - SA National Roads Agency

SBS - Stimulated Brillouin Scattering

SCE - Service Control Engine

SDR - Secure Domain Router

SIM - Subscriber Identity Module

SIP - Session Initiation Protocol

SLA - Service Level Agreement

SMF - Single Mode Fibre

SNR - Signal-to-Noise Ratio

SONET - Synchronous Optical Network

SPM - Self-Phase Modulation

SPVC - Soft Permanent Virtual Circuit

STM-1 - Synchronous Transport Module level 1

TEIN - Trans-Eurasia Information Network

Teraflop - 10^{12} floating point operations per second

TIA - Telecommunications Industry Association

TTL - Transistor-Transistor Logic

TUT - Tshwane University of Technology

UFH - University of Fort Hare

UJ - University of Johannesburg

UL - University of Limpopo

Unisa - University of South Africa

UP - University of Pretoria

UPS - Uninterruptable Power Supply

uPVC - unplasticised Polyvinyl Chloride

VANS - Value Added Network Services

vBNS - very-high-performance Backbone Network Service

VC - Virtual Container

VCAT - Virtual Concatenation

VLBI - Very Long Baseline Interferometry

VoIP - Voice over IP

VRF - Virtual private network Routing and Forwarding

VSAT - Very Small Aperture Terminal

VUT - Vaal University of Technology

WAN - Wide Area Network

WDM - Wavelength Division Multiplexing

Wits - University of the Witwatersrand

WSU - Walter Sisulu University

WSS - Wavelength Selective Switch

WSX - Wavelength Selective cross-connect

XPM - Cross-Phase Modulation

ZADNA - .ZA Domain Name Authority

Chapter 1

Introduction

Any country which takes research and tertiary education seriously needs a network that is dedicated to providing connectivity and information technology services to the research and education community. This network is known as a National Research and Education Network. The South African national government's Department of Science and Technology has embarked on a mission to build such a network, with assistance from the Tertiary Education and Research Network of South Africa (TENET), which has been the country's de facto NREN since 2000.

It was obvious from the start that there is much to be gained from all parties involved with these networks to work together to design and implement an integrated network which would be more than the sum of its parts, rather than to try and reinvent the wheel by building a new network from the ground up. The TENET staff has a great deal of experience in designing and managing networks and it was therefore natural that they would conceptualise and design the network.

1.1   Summary and Contributions

The author, in his position as network architecture consultant to TENET, was tasked with designing the optical backbone network for the SANReN. His responsibilities included the following:

- specifying all aspects of the optical network:
    - medium and optic fibre type
    - underlying transmission technology and equipment
- all aspects of the network design including
    - planning route diversity
    - planning card redundancy
    - router failover design
    - node layout and configuration
- installing and commissioning of the optical equipment in Johannesburg
- installing and configuring the optical equipment on the Mtunzini-Durban link

Finally, as TENET representative, he was asked to approach the City of Cape Town to align its dark fibre network routes with the needs of the NREN community. This led to the author being appointed as the City's principle telecommunications network architect, a position he currently holds.

As TENET's subject matter expert, the author was involved in every aspect of the design of the network, from planning routes for the national network and writing interface methodologies, to stipulating the backbone equipment and node layouts, down to card and port level.

1.2   Overview of the Thesis

This thesis documents the process that was followed to align the existing TENET network with the burgeoning SANReN network. It discusses all the stages of the design that had to be finalised, from choosing a transmission medium and technology, deciding on a Layer 3 technology and designing a network with it, as well as implementing the equipment and connecting it to the existing and new international circuits.

In this thesis the author follows the methodology of *action research* [1], in that he seeks to understand a complex, existing network of relationships as well as of equipment and data links, and then tries to develop a new and innovative way in which it may all be synthesised into a new and greater system, with the addition of new equipment, links and relationships.

To the best of the author's knowledge, no similar study about the design of an NREN has ever before been made in South Africa.

The thesis is structured as follows:

- Chapter 2 - A South African National Research and Education Network

  The need for a South African NREN is discussed and an overview is given of the attempt made to integrate the new government-sponsored network with the existing institution-owned network on a political and organisational level.

- Chapter 3 – Concerning NRENs and RRENs

  The role of NRENs in the wider global network is discussed and examples are given of what other NRENs do and how they integrate into larger RRENs. The specific function and history of TENET, as the de facto South African NREN, is examined and the integration of its network into the new SANReN is discussed.

- Chapter 4 – Choosing a transmission network medium

  An overview of the different transmission network mediums available to the NREN is given. Considering the amount of bandwidth required by the network, optic fibre is almost the only realistic choice of medium, but it comes with certain physical limitations that must be understood in order for the fibre to be used effectively. These limitations are documented and certain techniques discussed for limiting their impact on traffic over an optic fibre-based network.

- Chapter 5 – Choosing a transmission network technology

  Several technology options are available for putting traffic directly onto the chosen medium. In optical networks, this is known as "lighting the fibre". Each of the competing technologies has its

advantages, which makes the choice a difficult one. This chapter discusses each technology, lists its advantages and disadvantages, and finally discusses why certain technologies were chosen above others for implementing the SANReN.

- Chapter 6 – Choosing a Layer 3 technology

Once the underlying transmission technology has been chosen, a decision must be made as to which routing and switching technology should be implemented on top of it. The options available are discussed, once again with pros and cons, and a final choice of technology is made and motivated.

- Chapter 7 – Integrating the TENET GEN3 network

TENET has an existing network which covers more than 100 sites. It would be foolish for the network designers to disregard that, so this chapter examines ways in which a seamless integration of the existing network with the new network may be accomplished, with emphasis on both the national and international components.

- Chapter 8 – Designing the SANReN: principles

Before discussing the actual design of the SANReN, some fundamental principles behind the design should be understood. The design team had the luxury of designing a whole network from the ground up, which seldom happens in NRENs; and it was decided to set broad principles and follow them through. This chapter also documents the choice of equipment vendors and specific devices.

- Chapter 9 - Designing the SANReN: the national network

This chapter documents the design of the inter-city network which forms the national backbone of the SANReN. Methods of interconnecting small towns to larger cities are discussed, as well as the factors that were considered in deciding whether to purchase dark fibre- or managed connections.

- Chapter 10 - Designing the SANReN: metro network

The greater majority of the traffic on the SANReN will be generated by sites located within the borders of the four major cities: Pretoria, Johannesburg, Cape Town and Durban. Concurrently the most important part of the network is the multi-10Gbps metro networks to be implemented in these cities. The design of the metro networks is discussed in detail down to card and port level, as well as important considerations regarding redundancy and failover.

- Chapter 11 - Designing the SANReN: Layer 3 network

The Layer 3 network implemented on top of the optical network is discussed and details given about IP ranges defined, protocols used and network setups chosen.

- Chapter 12 – Benchmarking the SANReN

  It is important to know whether the design methodology used in the SANReN is truly according to world-class standards. This chapter discusses a similar network in the United States and draws parallels between it and the SANReN.

- Chapter 13 - Case study: designing a metropolitan fibre network

  One of the crucial impacts that the SANReN has already had on the real world of internetworking is in the design City of Cape Town's metro fibre network. This network was in its infancy when TENET approached the City and requested that its network be redesigned to include all the potential TENET sites. A detailed look is given to how such a network should be designed and what its impact may be on tertiary education and research, as well as the general community at large.

- Chapter 14 – Results and conclusions

  Clearly there is a difference between designing a network and building one, therefore this chapter documents what has already been done to implement the SANReN. It discusses some of the difficulties encountered and also suggests some ways in which the future rollout of the network may be facilitated.

- Appendix – Photos of the SANReN equipment

  A few photographs of the installed SANReN optical equipment is showcased.

- Bibliography

1.3   Technical Notes

Telecommunications is rife with acronyms and abbreviations, so footnotes are used to explain the nomenclature and abbreviations used in the text. The *ISO 690* system of citations is used, as well as some in-text hyperlinks to websites of institutions and the like. Extensive use has been made of figures and pictures, including some screen prints from *Google Earth*, to show relative distances and routes. All figures from external sources are cited in the bibliography, all others have been created by the author.

Chapter 2

A South African National Research and Education Network

If one takes a close look at what is happening in research and education networking in the developed world, one finds that more and more money and effort is being invested in creating high availability, very high bandwidth networks for the use of students and researchers. It is increasingly becoming possible for researchers of the same institution, or even of different institutions, to commission short-term links with very high bandwidths, on the order of 10 or even 40 gigabits per second between sites. On a network with a dedicated optical core, a group of researchers could set up a link like this for a day or even an hour, and tear down the link after they are done with it. This kind of service is very hard to come by in any commercial environment.

It is the nature of research traffic to be very bursty, i.e. an institution could easily consume an entire 10 Gbps circuit for a short time while an experiment is running, whereas the steady-state consumption of that site would be more on the order of 15% or so of the maximum burst speed. Clearly this kind of usage requires a very different sort of network to what is available commercially. If sufficient funding is available, the NREN would always strive to provide more bandwidth to each of its sites than *can* be used. A large site might be supplied with a dedicated 10 Gbps circuit, with the understanding that an additional circuit will be provided if the site ever manages to consume the first one on a regular basis. The philosophy behind the service provisioning is along the lines of: if you can consume what we give you, you don't have enough.

No profit-seeking service provider would ever countenance providing a connection that is usually only utilised at 15%, with occasional bursting to 70% of capacity. They only commission more bandwidth on the network when they absolutely have to, and will always try to maximise utilisation of all bandwidth available in the network. NRENs, on the other hand, are perfectly satisfied to *massively over-provision* bandwidth, as they take the view that bandwidth constraints limit research, whereas a bandwidth glut will tempt researchers into experimenting to try and utilise what they are given. With this approach, the NREN is actually stimulating research, rather than just supporting it.

In the era of easily available and affordable 10Gbps systems, even a large university or research institution may be hard-pressed to fully consume an entire 10Gbps circuit. However, if one were to add a supercomputer cluster into the network, the possibilities for high-bandwidth usage for research multiplies. It should come as no surprise therefore, that some of these high-bandwidth research and education networks are specifically being built around supporting and interconnecting supercomputer clusters. If several institutions which each have a supercomputer node could be interconnected on a high-speed backbone, a *supercomputer super-cluster* could be built. This configuration would also create new opportunities for research into distributed supercomputing.

One of the projects of the CSIR[1] is the *centre for high performance computing* in Cape Town. The centre has a satellite campus in Pretoria and several other sites are planned. On other fronts, South Africa has a large radio astronomy centre near Pretoria and will also most likely become the host country for the *square kilometre array*, which is set to become the largest radio astronomy project in the world (2). These and many other research projects that have very high bandwidth demands have placed the focus on the severe lack of bandwidth and the exorbitant costs thereof that South Africa in general and research specifically faces.

The Department of Science and Technology has embarked on a large and well-funded project to build a multi-10Gbps national backbone network to support these projects and research in general. This project, which will be implemented by the CSIR and its organ the *Meraka Institute*, will be known as the *South African National Research Network* (SANReN) [3].

South Africa has an existing formal tertiary education and research network which is owned by the tertiary education institutes themselves. The organisation which builds and manages this network is known as the *South African Tertiary Education and Research Network* (TENET). Since early 2008 the two organisations have had a signed memorandum of understanding which documents their firm intention to work together to build a single, integrated network to meet the bandwidth and interconnectivity needs of *all* the tertiary education institutions and research councils [4].

This thesis documents the growth of the relationship between the organisations as well as the conceptualisation, planning and realisation of this network, which will become the sole national research and education network of South Africa.

The first thing to do when designing such a national backbone network is to consider what underlying transmission medium should be used. The next chapter examines the different options and points out some of the limitations of these mediums that should be kept in mind when designing a network.

---

[1] Centre for Scientific and Industrial Research

Chapter 3

Concerning NRENs and RRENs

3.1 What is a NREN?

A NREN is a National Research and Education network. Its purpose is to connect many tertiary education and research institutions on a common backbone. This will allow each institution to have preferential access to its peers and also to shared infrastructure, such as mirror servers or computing facilities.

The NREN company or organisation (hereafter referred to as the NREN) is almost always a not-for-profit company and is usually owned by the institutions themselves, in the sense that these institutions are the shareholders. This status will often enable the NREN to negotiate very favourable rates from commercial carriers, who will in turn capitalise on their contribution as social responsibility.

3.1.1 Types of networks

The NREN network may take two forms, either as a real transmission network in the sense that the NREN company/organisation owns and operates its own equipment, or as a virtual network, i.e. a network built on top of a larger, existing network; typically a carrier's network.

3.1.1.1 Real NREN network

With contributions from its member institutions, donations from donor organisations and even government contributions, a NREN may be able to purchase its own transmission and routing equipment. This makes sense economically, as the non-profit NREN will always be able to operate its network more economically than a commercial operator could, or would be willing to. The only additional asset that the NREN would have to secure is long-term access to optical fibres, known as IRUs[1]. Few operators would be willing to donate so precious a resource as optic fibre pairs to the education and research cause, so the rights will also have to be purchased. Exceptions may be made when an operator has a "glut" of fibre, but in practice the fibre pairs required for the network are

---

[1] Indefeasible Right of Use

seldom available. In any case, the "last mile" of cable connecting the institutions' premises to the carrier's network will almost always have to be constructed on demand.

The NREN can then commission a transmission network to connect all of its sites to each other, as well to one or more core nodes. A core node would be situated where the NREN network is interconnected to other service providers and may also contain centralised equipment such as a mirror server or a computing node (even a supercomputer cluster). Operating its own network will give the NREN full control over the network and the ability to provide many ancillary services that would not be available to it otherwise.

3.1.1.2   Virtual NREN Network

A virtual NREN is a distinct and separate network in name only. It will be built and managed by a carrier/operator on top of their existing network. The role of the NREN in this kind of arena will be twofold. Primarily it will aggregate the buying power of the institutions and hence negotiate favourable services rates for them. The impact of a single provider supplying all the Internet/data services to a united community of research and education institutions will also be far greater than that same provider supplying institutions individually. The propaganda/advertising value that the provider may gain by being benevolent towards research and education may easily offset the loss of revenue it will experience by supplying services at below market value.

Secondly, the NREN may also leverage the virtual network to superimpose other services on it. This may extend to rate limiting and bursting (allowing a site to exceed its assigned bandwidth when there is spare bandwidth on the network), spam control and firewalling, as well as centralised services discussed below.

The advantage of running a virtual network is that the NREN is freed from the expense and complexity of owning/running its own equipment. It may still provide many of the ancillary services that would be expected of it when running a real network, such as billing and a central fault/service desk. However, the disadvantages of having little or no direct control of the network, as well as the inevitable higher costs of buying the services from a third party, rather than providing services itself, may force the NREN to move into the arena of becoming the owner/operator of a real network.

3.1.2   NREN services

3.1.2.1   Commodity Internet services

It may sound very useful for a NREN to provide interconnectivity between its member institutions for research and education traffic, but the reality is that up to 80% of the bandwidth required by an institution (especially a university) is so-called commodity traffic. This would mean regular Internet services such as file downloading, web browsing and email. It is unrealistic for a NREN to expect to provide certain network services to its clients, but not others. It therefore makes eminent sense for a NREN to provide regular, commodity Internet services alongside its "pure" research and education traffic.

3.1.2.2 Mirror server

A mirror server in the NREN context is a central data repository for the use of the NREN members. It would typically contain very large datasets that are of interest to several or all of the NREN members. It makes much more sense economically and practically to have one powerful, central server with a large storage capacity that holds datasets from several institutions, rather than for each university or research institution to try and buy and manage its own.

International (and even national) bandwidth is usually purchased in terms of download speed, not upload speed, even if the network is symmetrical (upload and download speed is equal). Given that most institutions (especially universities) will consume much more download than upload bandwidth, there will be unused upload capacity. A mirror server creates the opportunity to utilise the unused upload capacity. Academic/research data may be made available to the wider Internet at no extra cost, since the mirror has been established already and the upload bandwidth paid for.

It would also be possible to offer the mirror service to commercial companies, who would use it for software updates and the like. A typical example of this would be for web-browser or operating system software updates, which happen quite often and are downloaded by a very large number of users worldwide. The NREN could then negotiate favourable access to other databases and services in exchange for hosting these software updates.

3.1.2.3 Supercomputer cluster

Although many supercomputers are constructed for commercial purposes, few commercial entities can really justify the use of so much processing power. This places the typical usage of a supercomputer firmly within the realm of research and, by extension, tertiary education. The supercomputers available today can have data interfaces of as much as 10 Gbps[1], which places the usability bottleneck on the network links between the institution using the computing facility and the facility itself. A large research institution or university may indeed have its own supercomputer, but it is extremely unlikely that all possible users of such a facility will be onsite. Therefore, remote access to the facility must be provided. The NREN is ideally suited to build such a network, or else to provide extra capacity on its network for computing access. It is quite realistic for an NREN to build its entire network around such computing facilities, rather than to try and cater for it afterwards [5].

Another opportunity may be created by an NREN. If several institutions on the network have supercomputer clusters, these may all be linked on the NREN network into a super cluster. The network then becomes an integral part of a much larger and more powerful supercomputing grid, which would not have been possible without the existence of the network [6].

Such a grid would in its turn provide tertiary education and research opportunities for distributed computing. There is much focus on hard research into the implementation and application of distributed computing currently. As much as there will always be the necessity for research & education to do number crunching, the opportunities to do hard research in the field of distributed computing are much more promising [7].

---

[1] Giga (1,000,000,000) bits per second

3.1.2.4   Research datasets

Typical datasets would include a bio-informatics database containing data and images of diseases and viruses etc, and an astronomical database. The combined buying power of the NREN, as well as its access to donor and government funding may give it the means to afford expensive commercially available datasets which would be unavailable to smaller institutions individually [8].

The NREN may host databases locally that are originally hosted in other countries. This means that the NREN's members may have access to the datasets on the local (national) backbone, rather than download data from overseas via much more expensive international links. These datasets may then be synched at quiet times when there is little network traffic, such as over weekends or at night.

3.1.3   Examples of successful NRENs

3.1.3.1   Internet2

The NREN in the United States of America is an organisation called Internet2. It started in 1996 when researchers from 34 US universities met and decided to establish a dedicated network for research and education. Initially the network was built on the existing vBNS[1] of the NSF[2], but the needs of the education and research community diverged from those of the NSF, which led to the vBNS being used increasingly for government purposes [9].

The Internet2 community subsequently build the *Abilene* network and also invested heavily in the National Lambdarail network, which is an intercontinental DWDM network owned and operated by the education and research community itself [10].

In contrast to the Lambdarail, Internet2 is funded by university-corporate sponsorships. It does however, have a much larger reach, connecting over 60 000 institutions through the primary member institutions, which include 200-plus tertiary education- and 30 research institutions [5].

Internet2 tried to organise a merger with Lambdarail, but this was never realised. Since 2007 Internet2 dropped the *Abilene* moniker and renamed the network to *Internet2 Network* [11].

3.1.3.2   JANET

The NREN in the United Kingdom is called JANET, which was originally a contraction of *Joint Academic NETwork*. It traces its roots back to the 1970s when several small university networks started to grow and interconnect. By the 1980s it had transitioned to the X.25 technology, eventually becoming the fastest X.25 network in the world. JANET also led to member researchers advancing the X.25 protocol for their own uses, which resulted in the first complete X.25 standard [12].

JANET has gone through many incarnations and upgrades, including the move to a dedicated optic fibre backbone and the adoption of the *SuperJANET* moniker. The latest incarnation is called *SuperJANET5* and is based on 10 Gbps backbone. Plans are underway to upgrade it to 40 Gbps [12].

---

[1] very-high-performance Backbone Network Service
[2] National Science Foundation

3.2   Regional RENs

A Regional Research and Education Network (RREN) is a network which interconnects NRENs. The region in question usually spans several countries and in some cases even entire continents. The RREN organisation strives to provided a common backbone and interconnect standard which all its members can follow to ease connecting to each other [13].

3.2.1   Géant

Géant is the European RREN. It is a community comprised of 34 NRENs across Europe, all of which are interconnected on a dedicated optical network run by DANTE[1], which is itself a non-profit organisation dedicated to providing the underlying network that the NREN's interconnect on [14].

Gèant entered production at the end of 2001 and has gone through several project incarnations and name changes. It allows NREN connections at most common connection speeds, starting at 155 Mbps all the way up to 10 Gbps.

Gèant usually only interconnects its NREN community to other communities, i.e. allows peering connections only to other RRENs, such as TEIN3[2] in Asia Pacific and Australia, and CAREN[3] in central Asia. It is reluctant to interconnect its community to individual NRENs outside Europe, but has done so on occasion in the case of Internet2, TENET in South Africa and CANARIE[4] in Canada [14].

3.2.2   UbuntuNet

The *UbuntuNet Alliance* is the RREN for the African continent. It was established in 2005 by the five African NRENs then in existence, namely KENET[5], MAREN[6], MoRENet[7], RwEDNet[8] and TENET. It was formally incorporated in 2006 in Amsterdam [15].

UbuntuNet has tasked itself with fostering the creation of new NRENs in all the African countries which don't already have their own NREN. These NRENs will progressively be linked to each other on terrestrial backbones (preferably fibre) and also interconnected to the other RRENs of the world, most notably Gèant. In addition it will strive to develop IT resources in the tertiary education institutions comprising its member NRENs, both in terms of human capacity development as well as in actual equipment and infrastructure [16].

At the beginning the member NRENs were not directly interconnected, but only through low-earth

---

1 Delivery of Advanced Network Technology to Europe
2 Trans-Eurasia Information Network
3 Central Asian Research and Education Network
4 Canadian Advanced Network and Research for Industry and Education
5 Kenya Education NETwork
6 MAlawi Research and Education Network
7 Mozambique Research and Education Network
8 Rwanda Education Network

orbit VSAT[1] satellite uplinks, typically at speeds of 64kbps, although speeds of up to 2Mbps are possible. All the VSAT uplinks terminate at the Inmarsat ground station in Germany, but traffic between the NRENs is routed by the central UbuntuNet router, which is located in Telecity in London. From here the network has a 1 Gbps link to Gèant, as well as a 1 Gbps link to the commodity Internet from TeliaSonera [16].

On 23 July 2009 the 10 Gbps circuit provided on the Seacom cable was commissioned. It links the UbuntuNet central router in London to the TENET router in Durban. From there the bandwidth is transported to Johannesburg on the SANReN national backbone where it terminates on the UbuntuNet Southern Cluster router. The fledgling NRENs of Swaziland and Lesotho are also connected to this router (see 7.3).

3.3 TENET, the South African NREN

3.3.1 History

The original NREN in South Africa was called the Uninet, which was a project of the NRF[2]. In November 1998 Dr Khotso Mokhele, the president of the NRF, told the member institutions of the Uninet that it was not in the NRF's mandate to provide network services to universities and research institutions. He further stated that the day would come when the NRF "switched off" the Uninet and that the institutions would have to come up with a plan to provision a network themselves [17].

To this end, the Committee of Technikon Principals (CTP) and the SA Universities Vice-Chancellors Association (SAUVCA) formed a technical working group to negotiate a contract with Telkom SA for the provisioning of a network to replace the Uninet. During the same period, two American donor foundations, the Andrew W Mellon Foundation and the Atlantic Philanthropies, expressed their concern to Telkom about the excessively high bandwidth and service costs that the Uninet institutions had to pay [17].

These two stimuli led to the technical working group establishing a *section 21* non-profit organisation called the Tertiary Education Network (TENET). TENET has no shareholders, but instead belongs to its members, to whit the technikons, universities and other non-profit institutions such as the national research councils.

The two American donor foundations contributed around 2 million US dollars for the establishment of TENET, as well as to stimulate the growth of Internet usage by the institutions, and the expansion of technical skills in their IT departments [17].

TENET operates under an agency model, rather than a reseller/middle-man model. All participating institutions sign an agency agreement with TENET, who is thus empowered to negotiate contracts on the institutions' behalf. The institutions will also be bound by whatever agreements TENET signs. Telkom bills TENET directly and TENET bills the institutions in turn, adding a small agency fee which covers its running expenses [8].

---

1 Very Small Aperture Terminal
2 National Research Foundation

3.3.1.1   The HEIST Agreement

In December 2000 TENET succeeded in negotiating the *Higher Education Networking Solution with Telkom* (HEIST) agreement. Under this agreement Telkom would migrate all the existing Uninet services onto a new, integrated network to be built and managed by Telkom.

It took one year to the day for Telkom to migrate all the Uninet sites onto the HEIST network. All of the Uninet institutions, as well as a few other research institutions elected to join TENET and be included in the HEIST network. By April 2002 the number of sites had grown to 76 [18].

Under the agreement, an institution had to specify three different bandwidths for its sites: a backbone circuit connecting the site to the network, an amount of national traffic to South African sites outside the HEIST network, and an amount of international bandwidth. It could also ask for a portion of the backbone to be used to interconnect sites of the same institution. The international portion would include research traffic to overseas institutions as well as regular commodity Internet traffic [18].

For each portion the amount of bandwidth bought represents a committed information rate (CIR), which will always be available as well as always be non-over provisioned, meaning that the site will not share its bandwidth with any other site/institution. In addition, a site may burst its bandwidth usage by up to 100% when there is unused bandwidth on the network. This may happen during the evenings and over weekends and holidays; when some institutions are closed and therefore not consuming their share of the network bandwidth.

The agreement also states that the size of the backbone circuit may not be less than twice the size of the largest of the bandwidth types, i.e. inter-campus, national or international, nor smaller than the sum of the three. In practice a site will typically order considerably more capacity on its access circuit than the sum of its bandwidth partitions, to allow bursting. If a site only had as much capacity on the access circuit as the bandwidth it had ordered, it would not be able to burst.

Another salient point of the agreement is that the cost of bandwidth and access circuits is *distance-independent*. TENET insisted on this condition as many of the smaller and previously disadvantaged institutions are in geographically remote areas. In contrast, many of the larger sites that are able to afford higher prices are located in the major cities. Any agreement that was not distance-independent would therefore advantage larger, richer sites but disadvantage smaller, poorer sites [18].

3.3.1.2   The GEN2 Network

In December 2004 TENET signed a new agreement with Telkom called the "Second Generation" or GEN2 agreement. Aside from a much more attractive tariff structure, Telkom insisted on migrating the entire HEIST network which was based on ATM[1], to their new MPLS[2] backbone. This created several problems, not least of which was the fact that the cutover of the network took almost 10 months [19].

For the first time Telkom allowed TENET to request changes to the network routing schemes. TENET

---

[1] Asynchronous Transfer Mode
[2] Multi-Protocol Label Switching

took the opportunity to appoint a CTO[1] and introduce quite a few new services to the network. This included a mirror server, connections to Géant and Abilene in America, and a cache server. The servers were implemented in a new local node that TENET established in Telkom's Bellville exchange. This node was called the *Beachhead* and interfaced directly with the GEN2 network. In this way it was possible for the institutions on the network to access the data on the Beachhead servers without having to sacrifice bandwidth on either the national or international segments of their GEN2 service.

One of the advantages of the Telkom *IPNET MPLS* backbone network is that it offers different levels of QoS[2]. Thus a site's link may be segregated into different traffic streams, each at a different QoS. For example, commodity Internet traffic may be at the lowest available level (called "Bulk Business" in Telkom jargon), whereas a VoIP[3] service would require a higher level of service (called "Interactive Business"). Video conferencing would require the highest level of QoS available ("Real-Time"). Despite the obvious advantages the actual uptake of higher levels of service was very limited in the GEN2 network. Only one of the forty TENET institutions ever used the higher QoS levels, and even that was only for testing purposes. The cost of the lowest level of QoS was already so high that the added advantage of a higher QoS did not cost in [19].

The current TENET network is called the *GEN3 network* and is discussed in detail in section 7.1.

3.3.2   TENET Services

3.3.2.1   Cache server

During 2004 TENET introduced the cache server which kept track of pages and files that were frequently requested (downloaded) by any site in the GEN2 network and kept a local copy of that page/file on a server in the Beachhead. The next time that page/file was requested the cache server would serve it, saving the site using some of its international bandwidth to download it. After a year's service it was determined that the presence of the cache server made a difference of as much as 10% to the international bandwidth usage of a given site. Considering the cost of such bandwidth, the cache has more than justified its existence.

3.3.2.2   Mirror server

Many of the universities and research institutions on the overseas NREN's offer free online study courses to students, such as the *MIT[4] Open Courseware*. A local student would have to consume expensive international bandwidth to access such a course. However, if the courseware was kept on a local server, such as the one on the TENET network, it would not consume any international or national bandwidth that the student's institution would have to pay for, since the mirror server was (virtually) located on the network backbone. The TENET mirror server (*mirror.ac.za*) is such a server and keeps a local copy of the entire suite of courseware offered by MIT.

Secondly, the mirror server keeps a local copy of important research databases, such as the

---

[1] Chief Technical Officer
[2] Quality of Service
[3] Voice over IP
[4] Massachusetts Institute of Technology

bioinformatics database at *bio-mirror.net*, as well as the SAAO[1]'s astronomical database.

Thirdly, local copies of many software packages commonly used by students and researchers are kept on the mirror server. These range from Linux operating system software to antivirus and mapping tool software.

Lastly, the mirror serves as a distribution point for software updates such as *Mozilla Firefox*. It is the nature of the NREN in South Africa that the great majority of international bandwidth is used to download information from Europe and the rest of the world, rather than upload. This means that a lot of outgoing international bandwidth is paid for by the institutions, but cannot be utilised as the bandwidth cannot be bought asymmetrically. Companies like Mozilla are willing to pay third parties like TENET to host their software updates. TENET in effect resells the outgoing bandwidth that has already been paid for by the institutions and uses the income to pay for the common facilities on the network, such as the mirror server itself [20].

3.3.2.3   Other services

TENET offers its clients several other services [20]:

- It inherited several large blocks of IP space from the Uninet. It assigns these addresses to the member institutions on request and announces them to the other research networks of the world and the Internet in general.
- It also manages the academic domain name *AC.ZA*. Only academic institutions and other non-profit research institutions are assigned domain names in this range. TENET is registered with both *ICANN[2]* and *ZADNA[3]* as a local domain name authority.
- It has advocated the switch to IPv6[4] since the early days of the GEN2 network. It now offers native IPv6 to all its sites and insists that all parts of the NREN network must be IPv6 capable
- It runs several IT-related human capacity development programs. These programs are usually funded by philanthropic organisations and serves to expand the IT knowledge and skills in the tertiary education institutions. This includes periodically running training courses on IT management best-practice principles such as ITIL[5].

3.4   SANReN, the South African National Research Network

In 2005 the South African DST[6] obtained R365 million in funding from the National Treasury to build a high-speed national network to interconnect all the primary research institutions. Several RFPs were published in 2006 to investigate the options of building a network to connect all the research institution sites that housed A-rated scientists at speeds of up to 500 Mbps [21].

---

[1] South African Astronomical Observatory
[2] Internet Corporation for Assigned Names and Numbers
[3] .ZA Domain Name Authority
[4] Internet Protocol version 6
[5] Information Technology Infrastructure Library
[6] Department of Science and Technology

The Meraka institute of the CSIR[1] was given a three-year contract in 2007 to design and deploy the SANReN. It was obvious from the beginning that a close relationship with TENET could be very beneficial to all involved in the realisation of the SANReN. In 2008 Meraka and TENET signed a formal MoA[2] to the effect that TENET would assist with the planning and implementation of the network, as well as to manage it once it has been built [4].

It was soon apparent that the original list of sites with A-rated scientists was too exclusive, for it ignored many of the larger universities. The list was therefore expanded to include all 23 universities in the country. In addition, permission was granted for TENET to include any and all sites belonging to its member institutions, as long as additional funding was found for establishing connections to the SANReN. These connections would either be funded by the institution itself, or else by another organisation extending donations to TENET for the purpose. These donors could range from the *Department of Education* to international donor organisations such as the *Andrew W. Mellon Foundation* [4].

Under the MoA any site belonging to a TENET member institution could connect onto the SANReN backbone and would be considered an equal peer; i.e. funding aside, no distinction would be made between a SANReN site and a TENET site which is not on the SANReN deployment list. Traffic would be exchanged freely and all services available to a SANReN site would also be available to a TENET site. As TENET only includes not-for-profit research and education institutions to its membership, no questions of misuse of government-funded infrastructure would arise [4].

3.5   How SANReN integrates with TENET

Traditionally, TENET has been a virtual NREN, buying bandwidth and MPLS services from a national provider and building its value-added services on top of it. In the last few years TENET has expanded its range of services and moved closer to becoming a real NREN by buying routers and dark fibre IRUs and starting to manage a physical network itself.

SANReN, as a government-funded project, intends to also supply many of the services one would expect from a fully-fledged NREN. In addition, its entire client-institution list is a subset of the TENET client list. One may be excused for thinking that this creates an untenable position where the function of a NREN is being vied for by two completely separate organisations. In reality, however, there has been a remarkable level of cooperation between the CSIR and its organ Meraka, and TENET, to such an extent that it has been said by the management of TENET that it may be entirely possible that TENET as an organisation may cease to exist and all its functions subsumed into the SANReN organisation.

Practically speaking though, it will be many years before such a situation could occur, as the priorities of the two organisations still differ widely. TENET has always been focused on the needs of tertiary education, whereas SANReN was envisaged as playing a major role in promoting research. The two approaches also diverge considerably, SANReN following a top-down kind of approach, building a fast

---

[1] Council for Scientific and Industrial Research
[2] Memorandum of Understanding

backbone network with government funding, whereas TENET follows a bottom-up approach, concentrating on services and expanding the network as extra funding becomes available.

While neither approach can be faulted, it will take continued effort and cooperation between all concerned to be able to integrate the two networks into the world-class NREN that both organisations would like to see.

In the next chapter the first steps in building a real NREN, that of choosing a suitable medium on which to build a national network, are examined.

Chapter 4

Choosing a Transmission Network Medium

The first question to ask when designing a new network for a NREN is whether the NREN will own some or all of the physical infrastructure it uses. It has happened that the NREN is presented with the opportunity of buying a large section of an existing network, as in the case of AARNet[1]. However, it is much more likely that the NREN will have neither the means nor the operating license to own and run its own physical network. In such a case the NREN should try to obtain IRUs on either fibre pairs or else given amounts of bandwidth, for example a wavelength on a DWDM system or a leased line on copper or microwave.

In the event that IRUs are bought from an operator, the operator will usually not give the NREN a choice as to what medium is used in the core, as the carrier will use its existing network, regardless of what medium it uses. However, in the event that the last mile to the NREN's client site has not been constructed, the NREN may have the choice of which medium the carrier uses to deploy the last mile link.

In such a case the choice of medium will be dependent on two things: the physical circumstances and the bandwidth required.

4.1 Medium Choice

4.1.1 Wireless

There are several options available when considering installing a wireless link to a site. Microwave, free space optical or point-to-point wireless links are all good options when the bandwidth required is low and the cost of installation is must be minimised. Each technology has its own advantages and disadvantages, as well as having slightly different applications, but by and large it may be said that when the bandwidth required is less than the wireless system capacity, and no upgrades are foreseen, then wireless may be the best solution.

4.1.2 Optic fibre

Optic fibre is a superior transmission medium to microwave/wireless in many respects, however it

---

[1] Australian Academic & Research Network

may be much more expensive to deploy if the terrain is challenging. In a dense urban environment it is usually very expensive to trench and install pipes. The cost could easily exceed R600/m so it may be prohibitively expensive if the link from the site to the network is more than a kilometre or two.

If the site is in a rural area it may be feasible to install an overhead optic fibre route which is strung along on poles. In many places an existing pole run may be used, which will reduce the cost of building the route to the cable itself and the marginal cost of hanging it on the poles. The other advantage is that such a route may be constructed very quickly. A competent team may hang as much as 10km of cable a day.

If the bandwidth required is more than 155 Mbps, or an aggressive upgrade path is followed, optic fibre may be the only feasible choice. Its main advantage is that it is theoretically infinitely upgradeable. Even a very large campus or computing node may be satisfied with a single grey[1] 10 Gbps optical link.

It is advisable for the NREN to give very careful consideration to the upgrade path envisaged. It may be argued that any site that is able to consume an entire 10 Gbps link, or even fully utilise a 1 Gbps connection, way well expand to the point where a single 10 Gbps will not be enough. This will typically be the case with a supercomputer node or even a large university site.

An NREN is not like a commercial carrier that installs the minimum requirements initially and just upgrades capacity when the demand justifies it. Because of the non-profit arena in which an NREN plays it will most likely be forced to plan its network a lot more carefully than a carrier would. This could mean spending more money initially and installing a system that may be easily (and cheaply) upgraded in the future.

4.2   Optic Fibre

Although it is not common, it is not unheard-of for an NREN to build its own optic fibre network. Even if this network is not owned by the NREN, its carrier may be obliged to build some new cable routes for the use of the NREN. In such a case the NREN may have the opportunity to specify which fibre type it wants to use. Several types are available but the two types most commonly used are G.652 and G.655, according to the ITU[2] specifications. Careful consideration should be given to the choice of fibre type, depending on what use the NREN wishes to make of the network.

There are many characteristics of the optical fibre that impact on the transmission of optical signals. The most important of these are attenuation and chromatic dispersion. All of the optical effects should be considered when making the choice of fibre type.

---

[1] A "grey" interface uses an uncooled laser operating at 1310nm or 1550nm. It is relatively cheap as little attention is paid by the manufacturer to spectral bandwidth or wavelength stability. For manufacturing reasons, silicon-based lasers usually cost less at 1310nm than at 1550nm

[2] International Telecommunication Union, a body that writes many specifications for IT and telecoms

4.2.1 Optical effects

4.2.1.1 Fibre attenuation

The attenuation that an optical signal experiences on a fibre is wavelength dependent. In pure, undoped silica the fibre has two attenuation troughs, i.e. areas of minimum attenuation. This occurs at 1310 nm and 1550 nm (Figure 1).
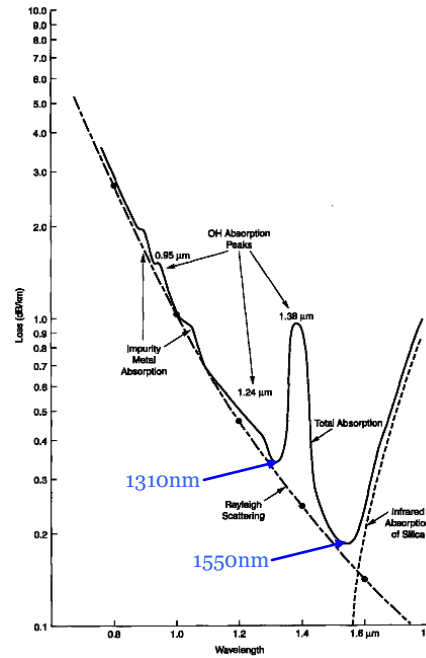
Figure 1: Signal attenuation vs. Wavelength [22]

It is clear from the illustration that the attenuation at the first trough (1310nm) is at approximately *0.35dB/km* and around *0.2dB/km* at 1550nm. Note the OH (water) absorption peak at 1380 nm. Modern manufacturing techniques have all but eliminated the absorption peak, significantly improving the usable wavelength range of the fibre.

The dotted line labelled *Rayleigh Scattering* indicates the theoretical minimum attenuation as a function of wavelength. The illustration shows that for modern single mode fibre types the real attenuation encountered by an optical signal is very close to the theoretical limit. Rayleigh scattering is the mechanism which causes attenuation in optical fibre. It is the elastic scattering of light due to collision with particles that are much smaller than the wavelength of the light (impurities in the optic fibre).

When short distances of fibre are to be lit, the fibre attenuation doesn't play such an important role, and either a 1310- or 1550nm laser may be chosen. The more important factor would be laser cost, as a 1550nm laser is somewhat more expensive than a 1310nm laser. Longer fibre distances are better approached with a 1550nm laser, as the attenuation per km plays a larger role. As a general rule of thumb, a fibre distance of a few kilometres may be lit by a 1310nm laser, but when the distance

approaches 10km or more, the 1550nm lasers should be used.

When several channels need to be commissioned, WDM[1] systems make more economic sense, as they can multiplex many channels onto the same fibre pair. Extending the reach of a WDM system to more than the system maximum requires amplification at an intermediate node. Such an amplifier will almost certainly be an EDFA[2]. Figure 4 shows that the EDFA functions well only in the 1550 nm band, rather than around 1310 nm. Therefore, choosing a WDM system that must be amplified necessitates the use of the 1550nm band, called the *C-band*.

4.2.1.2 Chromatic dispersion

The ideal optical signal source (laser) is monochromatic, meaning that it consists of a single wavelength of zero bandwidth. Such a signal is unfortunately impossible to produce with the silicon-based technology currently in use. A laser signal source hence has a non-zero bandwidth, which may be represented as having a spread of colours ranging from "blue" to "red" (Figure 2). This convention is used for illustration only as all wavelengths are outside the visible spectrum.
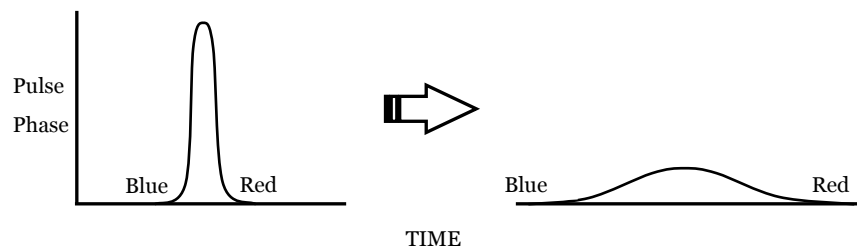
Figure 2: Pulse spreading due to chromatic dispersion [23]

As an optical pulse travels down a fibre the different wavelengths travel at different speeds, which means that the "red" wavelengths travel faster than the "blue" ones. This means that the length of time that the light pulse persists, increases as it travels down the fibre. This results in pulse spreading (Figure 2), which in turn leads to pulse aliasing (Figure 3) [24].

Note that the effect of the two adjacent pulses overlapping in the top part of Figure 3: Pulse aliasinghas the result that the signal amplitude at the overlap is the sum of the two signals. This means that the signal strength never returns to zero anymore. Because the signal conveys digital information aliasing can cause the signal to be corrupted.

As an example, the 0-5 V signal range is often used in digital systems like TTL[3] or CMOS[4]. A digital "zero" is defined as any voltage between 0 and 0.8 V. Likewise a "one" is defined as any voltage between 2 and 5 V. Any voltage between these two levels (0.8-2 V) would be an invalid signal level. As soon as the lowest signal voltage crosses this threshold, the receiving system can no longer distinguish

---

[1] Wavelength Division Multiplexing
[2] Erbium Doped Fibre Amplifier
[3] Transistor-Transistor Logic
[4] Complementary Metal–Oxide Semiconductor

between a "one" and a "zero" and the entire digital signal is corrupted and the information lost.
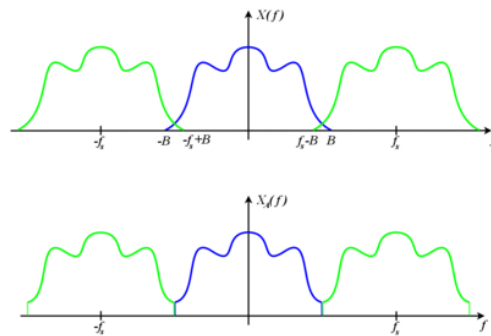
Figure 3: Pulse aliasing [25]

Chromatic dispersion in an optic fibre system causes pulse aliasing, which means that when the laser transmits a "zero", which represents the transmitting laser being "off", there is still some optical power being received. When the "zero" optical power threshold is crossed, the digital signal cannot be recovered.
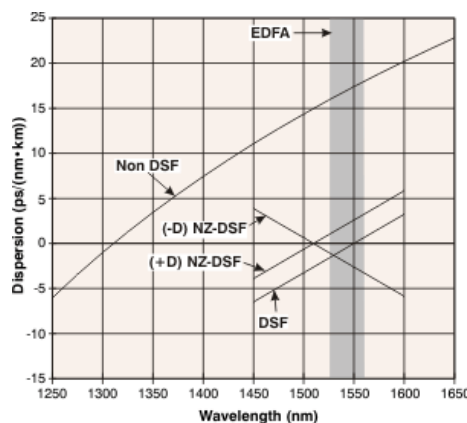
Figure 4: Dispersion curves for common fibre types [26]

Figure 4 illustrates chromatic dispersion vs. wavelength for the most common types of fibre commercially available.

4.2.1.3   Dispersion compensation

Chromatic dispersion may be compensated for by adding a length of fibre that has an opposite dispersion slope to the fibre in use on the network (Figure 4: *(-D) NZ-DSF*). To use the analogy of the "blue" and "red" wavelengths again, in this fibre type the "blue" wavelengths will travel faster than the "red" wavelengths and therefore will "catch up" to the "red" wavelengths. This will have the effect of reversing the pulse spreading illustrated in Figure 2. The disadvantage of this scheme is of course the extra length of fibre necessary. If the compensating fibre had the same (negative) slope as the network fibre, one would need an equal length of fibre to compensate. This would mean that a system with enough optical power for a 100km link, would now only be able to reach 50km, because the effective fibre length in the link would be twice the physical distance between nodes.

The solution to this problem is to use dispersion shifted fibre with a much sharper negative slope than the network fibre. Alternatively, on long-haul runs the route may be constructed using sections of fibre with alternating dispersion slopes. For example, if the first 10 km of the fibre cable had a positive dispersion slope, but the next 10 km had a equal (but negative) dispersion slope, the optical signal will experience dispersion all the time, but after 20km the dispersion will be back to zero (24).

4.2.1.4 Nonlinear effects

Several non-linear effects manifest themselves in optical systems, especially when the bit-rate of the optical signal becomes very high (in excess of 10 Gbps) or the channel spacing becomes very small in WDM systems (smaller than 50 GHz between channels)[1].

PMD[2] occurs when one polarisation axis (for example the vertical axis) of the optical signal experiences a different level of dispersion as the other (horizontal) axis. The effect of this is that the vertically polarised components (for example) of the signal travel faster than the horizontal components. In effect the optical wave front becomes inclined, rather than flat, with respect to the direction of travel. In high bit rate systems this will cause the horizontal component of the trailing end of one pulse to arrive at the same moment as the leading edge of the next pulse. The receiver cannot then distinguish between the end of one pulse and the start of the next.

SPM[3] occurs when the electric field aspect of ultra-short pulses of light induce a change in the refractive index of the fibre. This is known as the *optical Kerr effect* and increases the frequency spectrum of the light pulse, which leads to pulse spreading. It is only prominent at very high bit-rates and high power levels.

XPM[4] is also due to the Kerr effect and causes one channel of a WDM system to affect the phase of another channel. It is only prominent with very narrow channel spacing (smaller than 50 GHz).

FWM[5] occurs when three channels in a WDM system experience intermodulation distortion, which means that the interaction of the three signals gives rise to a fourth signal. This is caused when scattering of the incident photons of the three signals produces a fourth photon, at a slightly different wavelength. This produces a "ghost" signal, which may interfere with a channel at a different wavelength[6] (crosstalk). The original signals may interact in as many as 12 different ways, producing 12 "ghost" signals, of which three may be at the same wavelengths as the original signals. FWM is more prominent with very narrow channel spacing and high power levels and may be mitigated by using unequal channel spacing.

SBS[7] occurs only in very high power systems where the energy of an incident photon is so high it induces an acoustic wave (phonon), which travels in the opposite direction of the optical signal. SBS effectively limits the maximum launch power of a laser or amplifier in an optic fibre system.

---

[1] Some 25 GHz-spacing systems are coming onto the market now
[2] Polarisation Mode Dispersion
[3] Self-Phase Modulation
[4] Cross-Phase Modulation
[5] Four-Wave Mixing
[6] This is called crosstalk
[7] Stimulated Brillouin Scattering

Compensating for all the nonlinear effects mentioned is usually very difficult, but many of the effects may be avoided by choosing a channel spacing that is not too small (50 GHz or larger), limiting the optical launch power, restricting the channel bit-rate to 10 Gbps, and using fibre with enough chromatic dispersion (see section 4.2.2.2) [27].

4.2.2   Fibre types

4.2.2.1   G.652

The standard grade of fibre in most widespread use across the world is made in accordance with the ITU *G.652* recommendation, also known as SMF[1] or non-dispersion shifted fibre. It has a zero chromatic dispersion peak at 1310nm, as indicated by the *Non DSF* curve in Figure 4. The reason this fibre type is in such widespread use is twofold: firstly it was the first type of fibre that was widely deployed and therefore new networks still deploy it, to avoid having mismatches between different fibre types in the same network. Secondly, G.652 works very well with any single channel system employing 1310 nm laser sources, as these signals will encounter no chromatic dispersion, even at 40 Gbps or higher [28].

Two problems arise when one tries to use DWDM[2] on this fibre: as is evident from Figure 4, the amount of dispersion introduced in the 1550 nm band is quite large. Secondly, the use of DWDM technology and EDFAs dictate the use of the 1550nm band, as EDFAs do not function in the lower wavelength bands. If G.652 is the choice of fibre type, some effort will have to be made to compensate for the dispersion that the optical signals experience.

4.2.2.2   G.655

The most common NZ-DSF[3] in common use is made according to the G.655 recommendation. It has a zero dispersion point at 1530nm, which is just outside the C-band.

The reason for not having the zero dispersion point at 1550 nm is that the wavelength grid recommended for the C-band by the ITU is spread out on either side of the 1550nm centre frequency, ranging from 1530 to 1565nm. Therefore, all channels in the C-band would experience a positive amount of dispersion. If the zero dispersion point was exactly at 1550 nm, all channels at a shorter wavelength will experience negative dispersion, whereas longer wavelengths will experience positive dispersion. Such a situation will make dispersion compensation on the whole system more difficult.

It can be seen from Figure 4 that the dispersion in the C-band is considerably less using the G.655 fibre (*+D NZ-DSF*) than with G.652 fibre (*Non-DSF*).

There are several grades of G.655 fibre. After the recommendation was initially released, it was noticed that several of the nonlinear optical effect are inversely proportional to the amount of chromatic dispersion. Therefore, while it is preferable to have a low chromatic dispersion in the C-band, it should not be too low; otherwise the nonlinear effects have a larger impact on the signal

---

[1] Single Mode Fibre
[2] Dense Wavelength Division Multiplexing
[3] Non-Zero Dispersion Shifted Fibre

transmission. Generally accepted values of C-band dispersion are between 6 and 10 ps/nm^2*km [29].

4.2.2.3   Other fibre types

Several other types of fibre are commercially available. If one wants to build a single-channel system and to make use of the lower attenuation at 1550 nm, one would still be concerned with the high dispersion at this wavelength. The solution would be to use a fibre type that has a zero dispersion point at 1550 nm. *G.653* is a fibre type that would suit this application (*DSF* in Figure 4).

In long-haul DWDM systems, especially where the channel spacing is very small (≤50 GHz between channels) or where the channel bit speed is very high (≥10 Gbps), the factors influencing optical transmission most are the nonlinear effects, rather than chromatic dispersion. *G.656* fibre is optimised to minimise these nonlinear effects and should be used for long-haul DWDM systems.

4.2.3   Cable types

Many kinds of fibre cable are in widespread use and may be deployed when building core and access networks. The application will determine the cable type.

4.2.3.1   OPGW

Optical fibre composite overhead ground wire is a hollow steel cable which contains one or several tubes of optic fibres. The outer jacket consists of (usually) six bundles of braided aluminium cable. The aluminium is not only very strong, which makes for a robust cable, but is also a very good conductor of electricity. OPGW cable is usually used as a replacement for the topmost ground wire on high voltage power transmission lines. In this way the cable is simultaneously the transmission line's protection grounding wire and an optic fibre cable. The most popular size of OPGW contains four tubes of 12 fibre strands (48 strands).
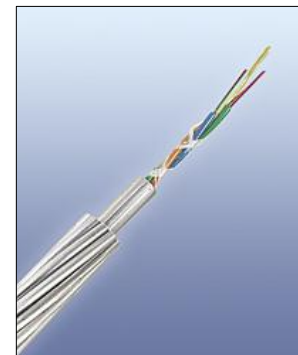
Figure 5: OPGW cable [52]

4.2.3.2   CST

Corrugated steel tape outside plant cable is the most commonly used underground fibre cable in the world. It consists of a central strengthening member (usually aramid yarn), several tubes of fibre strands (usually 6 or 12 strands/tube), and a layer of corrugated steel tape, sandwiched between two layers of HDPE[1]. CST cable is usually considered to be both water- and rodent proof, although some rodents have been known to gnaw through standard CST cable.
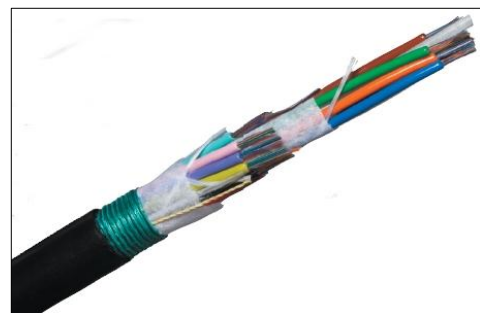
Figure 6: CST cable [51]

---

[1] High Density Polyethylene

4.2.3.3   Aerial cable (ADSS)

All-dielectric self-supporting cable is the preferred cable for installing on overhead pole routes. It is similar to CST cable except that the CST layer is replaced with a layer of aramid yarn. This increases the tensile strength of the cable considerably and enables it to be installed between poles, without any other supporting guy wire. The outer HDPE layer is ultraviolet resistant and the aramid layer will allow installation of unsupported lengths of 50m or more.
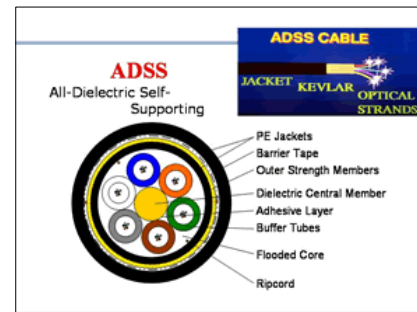
Figure 7: Aerial cable [53]

4.2.3.4   Blown fibre mini cables

Blown fibre mini cables are very similar to CST cables except they do not have a CST layer. This makes the cable diameter much smaller, but of course less robust. Popular fibre counts include 12, 24, 48, 72 and even 96. The cable is air-blown under pressure (typically 15 bar) into a mini-tube of 10mm inner diameter. The tube will be part of a tube assembly (mini-duct) consisting of between one and seven mini-tubes. The mini-duct may be installed into a larger duct (110mm) or be directly buried in the ground. The direct buried version usually has an additional HDPE outer jacket.
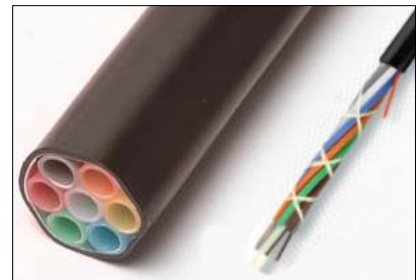
Figure 8: Mini cables [54]

4.2.3.5   Blown Fibre micro cables

The last mile of a fibre network, especially in an urban area, will typically be constructed from blown fibre micro-cables, which are much smaller in diameter than mini-cables (between 1 and 1.5mm) and are air blown into micro-ducts with a 3.5mm inner diameter. Building entries are easier to construct and manage with micro-cables than any other kind of cable. It is also very suited to cabling inside buildings, even high-rises. Micro cables never exceed 12 strands, to maximise blowing distance.
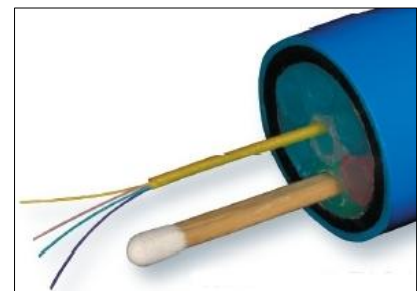
Figure 9: Micro cables [55]

4.3   Wireless

4.3.1   Microwave

The term "microwave" usually refers to electromagnetic radiation with a frequency of between 300 MHz and 300 GHz. In the telecoms industry it is often used to describe a specific type of technology. Microwave devices are synchronous[1], symmetric[2] and function in the band mentioned. They are almost exclusively used as backhaul connections in the Core network, or as leased-line connections in

---

[1] Circuit-based, rather than packet-based
[2] The data bit-rate in the downstream direction is equal to that in the upstream direction

the Access network and are best suited for permanent, fixed bandwidth, always-on, symmetric applications. A microwave connection will typically be available in all the standard PDH bit-rates; up to 155 Mbps SDH (see 5.1 and 5.2). Microwave circuits are always point-to-point, never point-to-multipoint and may have a range of as much as 60km or more.

The main disadvantage of traditional microwave systems is that they pollute the frequency spectrum. The power output of a microwave is usually only controllable in very coarse steps, which means that the much more than the required amount of power is transmitted. This is also done to mitigate the effects of bad weather. The problem is that the microwave links are often used in densely populated areas, so when the microwave beam extends further than necessary for operation, it "jams" the spectrum for a considerable distance beyond the receiver, making that portion of the spectrum unusable for other users.

4.3.2   Point-to-point

Many other devices may be found that operate in the microwave frequency band that are also point-to-point, but do not qualify as microwave devices. The systems use links that are usually not permanent, meaning that they only transmit when in use, and are often not symmetric either. Such systems will be employed to connect sites that have much smaller upload than download requirements, or that need a connection which is not in use all the time (and therefore only generates cost when in use).

4.3.3   Point-to-multipoint

Point-to-multipoint systems are similar to point-to-point systems, but with the added limitation that the available bandwidth is shared between the number of users connected to the fixed access point (base station). This kind of connection is becoming more prevalent as more end-users are prepared to accept the trade-off between cost and connection quality. 3G[1] and EDGE[2] are generally available examples of this connection type.

4.3.4   Free-space optical

If the connection required is short-distance and there is line-of-sight between the customer site and a provider node, a free-space optical link may be used. This is a short range laser beam, often in the visible or near-infrared spectrum. The FSO will most often serve as a leased-line type of service, and may be either a circuit-based (PDH) or packet-based (Ethernet), and is usually symmetric.

The advantages of FSO links are that no spectrum licence is required, the link is very secure, the equipment is usually inexpensive and easy to install, and that the system is immune to interference. On the other hand, the FSO beam is usually very sensitive to atmospheric conditions and may be disrupted completely by rain, fog or pollution. FSO is most suited to links that are short-distance (typically less than 5km) and sites that are cost-sensitive, but not averse to possible signal degradation/loss due to weather conditions.

---

[1] Third generation mobile communications
[2] Enhanced Data rates for GSM (Global System for Mobile telecoms) Evolution

This chapter examined some of the medium options available to the NREN designers. The choice of medium may sometimes be difficult to make, especially where the budget is very limited or two options offer similar benefits. Generally, however, the NREN will always use optic fibre as its default choice, since no other medium offers the same bandwidth and upgrade path. Microwave should only be used when optic fibre is prohibitively expensive and the site in question is within one "hop" of an established network point. The other wireless technologies are really only intended for very low usage sites, when it may be possible to establish a low-bandwidth connection via the NREN virtual private network.

In the next chapter, the underlying transmission technologies that the NREN may choose from are examined in detail, to enable the design team to choose a technology that is future-proof but still satisfies all the clients' needs.

Chapter 5

Choosing a Transmission Network Technology

Many different choices are open to the NREN when it comes to choosing a type of technology to build the network with. It is obvious that this choice should be made very carefully as well, to avoid choosing a technology that may be defunct in the near future.

5.1   Synchronous Digital Hierarchy

SDH is the transmission technology of choice for most carriers in the world outside the USA. A very similar system called SONET[1] is the standard in America.

SDH is a circuit-based transmission technology, which gives rise to its synchronous nature. It was developed specifically for optic fibre networks and works very well in ring and mesh architectures. It is specifically well adapted to networks with high QoS requirements and can be used to build networks with 99.999% availability.

SDH is eminently suited to circuit-based networks such as voice where QoS is the paramount consideration and the price-per-bit is of lesser importance. In modern terms, however, its principle strength is also its major weakness. Data networks are steadily evolving towards packet-based architectures, which bring the advantages of statistical multiplexing to the fray. Now that switched voice in the PSTN[2] is slowly giving way to VoIP[3], the market for SDH is steadily shrinking and within a few years the technology will be defunct.

In SDH networks the network manager has to commission each and every circuit, whether it is assigned to voice or to a leased-line data service. This used to be seen as an advantage over packet based networks because the network may be provisioned according to a plan and this can lead to very good bandwidth utilisation of the network with a minimum of unused/dead bandwidth. Very strong protection schemes may be also be implemented whereby a circuit, group of circuits or even an entire fibre link may have alternate routes provisioned with restoration times in the order of 50ms.

However, once again the advantages have become disadvantages. In comparison with the new packet-

---

[1] Synchronous Optical Network
[2] Public Switched Telephony Network
[3] Voice over Internet Protocol

based technologies SDH is very rigid and inflexible. Each circuit has to be provisioned individually, either by a human operator or by the network management system itself. Such a system may easily cost millions of Rands and is usually based on megalithic operating cores which are very difficult to modify. To make matters worse, the paths provisioned for protection are usually reserved, which means that these links may not be used for live traffic. In the extreme this could mean that half the network is idle, waiting for something to go wrong.

5.1.1 The SDH frame

The basic frame of SDH is called an STM-1[1] frame (Figure 10). In the simplest terms the basic frame has four parts: the payload (data) section; two data communications channels (the RSOH[2] & MSOH[3]), which are used for communicating with respectively, a regenerator node and a terminal multiplexer node; and the AU[4], which contains the pointers [30].
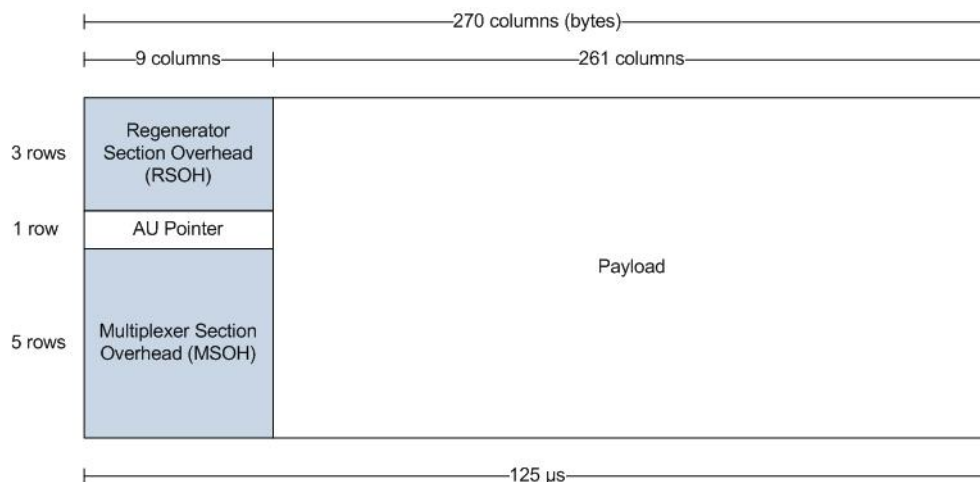
Figure 10: The SDH STM-1 Frame [31]

One frame, containing 2430 bytes, takes 125 µs to transmit, which results in a line speed of (19440 bits[5]/ 125 µs) = 155.52 Mbps. It is now simple to calculate that the maximum data speed is 150.336 Mbps, by using only the payload bytes.

The data section may be subdivided into several standard size "virtual containers", designed to transport standard legacy data rates, like PDH 140 Mbps (VC[6]-4), 34/45 Mbps (VC-3) and 2.048 Mbps (VC-12) (see 0). 4x STM-1 data containers may also be multiplexed into a STM-4 with a line speed of 622.08 Mbps and so on, up to a STM-16 (2.488 Gbps), STM-64 (9.95 Gbps) and even STM-256 (39.8 Gbps) [30].

---

[1] Synchronous Transport Module level 1
[2] Regenerator Section Overhead
[3] Multiplexer Section Overhead
[4] Administrative Unit
[5] (270*9*8 bits/byte)
[6] Virtual Container

5.1.2 Protection switching

SDH has several protection switching schemes that may be used to provide backup connectivity if the main (worker) circuit fails. The main reasons for circuit failure would be path disruption (fibre cable breaks) and equipment failure (most often laser failure).

5.1.2.1 Automatic Protection Switching

APS, also known as 1+1 switching, is used to protect against card/laser failure between two pieces of equipment in the same facility. Two paths are set up between the switching fabrics of the two sets of equipment (a worker- and a protector path) and switching occurs when the worker path degrades or fails. APS is sometimes implemented on submarine fibre networks, where route diversity is not available [32].

This added level of protection is required for very high availability networks, for example to improve the statistical availability of a network from 99.9% to 99.99%[1] or higher. As the paths are only between two sets of equipment, protection switching may be as fast as 10ms or better.

5.1.2.2 Sub-Network Connection Protection

SNCP is the main type of circuit protection employed in SDH networks. It is usually used to establish a ring between two points, for example between two sites of the same customer. The provider would establish two physically diverse circuits between the two points, which may both pass through many sets of equipment on the provider's network. One may use the analogy of the transmitting system "talking" down both paths at the same time, and the receiver "hearing" both paths, but ignoring the protector circuit. During normal operation all the data on the protector path is ignored (lost) and the only the worker path is used. However, when the QoS on the worker path degrades beyond a certain point, or fails altogether, the receiver switches over to the protector path. Almost always there is a difference in distance between the two paths, and the receiver therefore needs some time to synchronise with the new path, before service can be restored. The switching always takes less than 50ms [32].

5.1.2.3 Multiplex Section-Shared Protection Ring

MS-SPRing is used in more complex networks to mitigate the effects of deploying dedicated protections circuits. In simple terms, a section of the network (usually a fibre span) is configured with MS-SPRing, which will then reroute all the traffic on that section along other paths over the network to reconnect the two points that experienced service disruption. In this way the protection sections are not dedicated to specific circuits, but are used for all circuits that are on the disrupted path [32].

This scheme adds a lot of complexity to the network, as each section must have a protection section configured. One the positive side, if one assumes that only one section of the network will be down at any given time, several production sections may be protected by the same protection section.

---

[1] The availability percentage indicates the statistical probability of a circuit failing. It is often translated into a number of minutes per year of permitted downtime, but this is a misleading interpretation.

Once the entire network has been configured in this way, new non-protected circuits may be added that will automatically have the benefit of the protection scheme. In addition, dedicated protection sections can be used to carry low-priority traffic during normal working conditions. When the protection circuit is activated, the low-priority traffic is dropped.

5.1.3   Routing circuits

The SDH network management system consists of two components: an element manager and a network manager. The element manager configures all aspects of the individual SDH elements, including configuring cards and ports, and updating the firmware. The network manager uses the information from the element manager, as well as its own information about the state of the network, and will plan and commission new circuits across the entire network, given only the start and end points, protection scheme and route weightings that the network designer specifies. It will interrogate individual elements and seize available ports for the new circuit. It also serves as the network database and keeps track of where all circuits are, including the path taken across the network.

The network manager may use any of a number of routing and weighting schemes, but the straight-forward shortest-path-first algorithm[1] is employed most often.

5.1.4   Synchronisation in SDH networks

SDH is a fully synchronous network, which means that the clocks of all elements must be synchronised. It is not enough to just specify a given clock frequency, as two individual clock crystals can never be perfectly identical. Even the smallest difference may cause a synchronisation error given enough time.

When a SDH network is designed, a synchronisation plan must also be created. One element on the whole network is designated to be the master clock, and all other "downstream" elements must be synchronised to that clock signal. An element may be configured to run on its own internal clock, or to get its clock from one of the incoming lines. With careful planning and implementation a network may run error-free for years. Without a proper plan implemented, an entire network may suddenly fall out of synchronisation and shut down completely, as has happened before. In such a case the only thing that can be done is to reinitialise every single element [30].

5.1.5   Next Generation SDH

In comparison with packet-based systems SDH seems wasteful and inflexible, as well as very expensive. New mapping schemes like GFP[2], LCAS[3], and VCAT[4] have helped to extend the lifetime of the exiting SDH carrier networks. GFP allows arbitrary sizes of SDH streams to be specified. One could for example assign a 15Mbps stream to an Ethernet input. Previously one only had the option of assigning a VC-12 (2.048 Mbps) or an entire VC-3 (45 Mbps) to the Ethernet input [32].

---

[1] Also known as Dijkstra's algoritm
[2] Generic Framing Procedure
[3] Link Capacity Adjustment Scheme
[4] Virtual Concatenation

VCAT allows the combining of several VCs into a single, virtually contiguous stream. LCAS allows the system to dynamically redefine the size of the assigned stream, based on the changing bandwidth demands of a typical Ethernet system.

These new framing and assignment procedures have considerably eased the transition of circuit-based devices and services into packet-based networks, however it is clear that any transmission system that remains firmly rooted in circuit based- rather than packet-based technology is doomed to become obsolete. In the final reckoning the ubiquitousness of Ethernet simply means that it is cheaper to deploy than SDH. It can be shown that in a carrier-class network SDH can be as much as four to five times as expensive to deploy as Ethernet.

5.2   Plesiochronous Digital Hierarchy

PDH is a legacy transmission technology that has almost entirely disappeared from transmission networks in the developed world. Plesiochronous means "semi-synchronous" in the sense that all the elements of the network work at the same clock speed, but that there is no attempt to synchronise the different clocks. The underlying assumption is that there is always a slight difference in the clock speed of the different elements in the network, but that the elements can compensate for it.

The one aspect of PDH that has outlived the technology as a whole is the standard data rates. From standard telephony a voice signal is band limited to 3.1 kHz and sampled at 8 kHz and 8 bits, to produce a 64kbps[1] digital signal. 30 Data channels and 2 signalling/synchronisation channels are multiplexed together to form a 2Meg[2] circuit, with a bit rate of 2 048 000 bps. Some stuffing bits are added to aid synchronisation when four 2Megs are multiplexed into an 8Meg (8.448 Mbps). Four 8Megs make a 34Meg[3] (34.368Mbps), four 34Megs make a 140Meg (139.264 Mbps) and four 140Megs make a 565 Mbps circuit [33].

The American version of PDH multiplexes 24x 64kbps circuits into a DS-1 circuit (1.544 Mbps) and 8x DS-1s into a DS-3 (44.736 Mbps).

Even in modern SDH networks one still encounters the E1, E3 and DS-3 circuits very frequently, especially in mainframe computing and SANs[4], attesting to the influence that PDH still has in the world.

5.3   Asynchronous Transfer Mode

ATM is a transmission technology that is still in very wide-spread use. It was the first serious attempt to marry the advantages of circuit-based switching and packet-based switching into a single one-size-fits-all transmission network. ATM uses fixed-length packets called "cells" that always consist of 48 bytes of data and 5 bytes of overheads, making a 53-byte cell [34].

ATM networks establish virtual circuits (VCs) between the origin and destination nodes that ensure

---

[1] Also called an E0 service
[2] Also known as an E1
[3] E3 service
[4] Storage Area Networks

that all the cells of a given data stream (or call in voice and video applications) are transmitted over the same path through the network, eliminating the packet network problem of packets arriving out of turn.

The fixed-length cells eliminate the problem of jitter (variance in latency) that is present in packet networks. When a time-sensitive service such as voice or video is introduced onto a packet network, the packets may encounter other packets of near-maximum length (e.g. 1500 bytes on Ethernet networks), and be forced to wait until that packet has been transmitted. Such a delay may be unacceptable for voice or video, since the user will notice the delay. When one uses a technology where *all* services are broken down into small, equally sized packets, the problem of jitter is eliminated.

5.3.1   The ATM cell

There are two versions of the ATM header; one is used for user to network interface (UNI) and the other for network to network interface (NNI). The UNI header (Figure 11), which is used in most links) consists of only five bytes, which include (30):

- **Generic flow control** (4 bits, all 0 by default) – for identifying multiple stations that use the same ATM interface, seldom used
- **Virtual path identifier** (8 bits) – helps identify the path through the network
- **Virtual channel identifier** (16 bits) – used in conjunction with VPI to identify the next element in the ATM virtual circuit
- **Payload type** (3 bits) – 1st bit indicates whether user data or control data is in the payload, 2nd bit indicates congestion, 3rd bit specifies whether the cell is the last of an AAL5 frame
- **Congestion loss priority** (1 bit) – indicates whether the cell may be dropped if the network experiences congestion
- **Header error control** (8 bits) – the header checksum
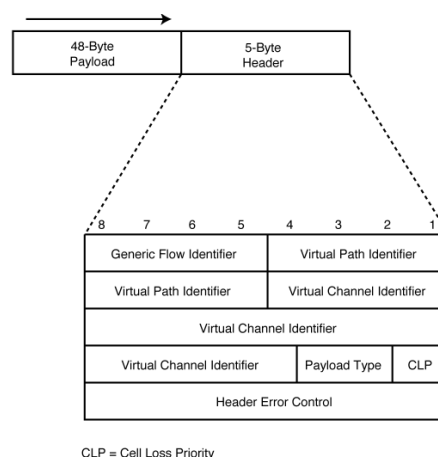
Figure 11: ATM UNI header [35]

The only difference between the UNI and NNI headers is that the four GFI bits in the UNI header

expands the VPI field to 12 bits in the NNI header.

5.3.2   ATM adaptation layer

The AAL converts the user input protocols (Ethernet, SDH, etc.) into ATM cells that may be switched by the ATM switch. The AAL manages the timing and flow control, handling of lost cells and transmission errors, as well as the segmentation and reassembly of the client protocols. The AAL type is not specified in the header, but is negotiated between the origin and destination nodes. Figure 12 illustrates the role of the AAL in integrating other protocols into ATM.
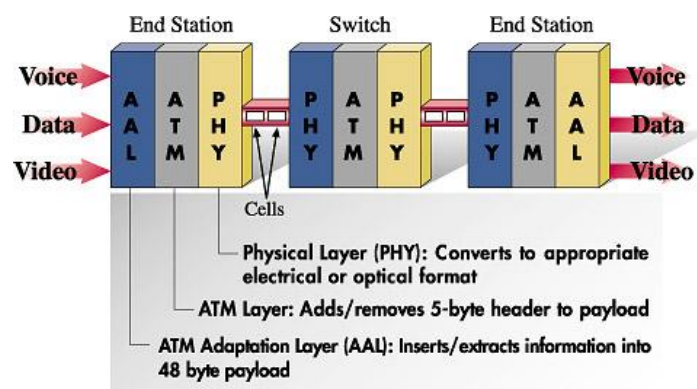
Figure 12: Integrating other protocols into ATM [36]

There are five different types of AAL, each specified for a specific type of data/service:

AAL1 – Constant bit rate, connection-oriented and synchronous traffic. Typically used for E1/DS-1 emulation

AAL2 - Variable bit rate, connection-oriented and synchronous traffic. Used for video and voice

AAL3/4 – Asynchronous or connectionless packet data. Used for Frame Relay and X.25

AAL5 – Simplified version of AAL3/4, with reduced transmission and protocol overheads. Widely used for packet data, especially IP- and Ethernet over ATM

5.3.3   Traffic engineering

One of the principle advantages of ATM over traditional packet-based networks is its traffic engineering feature (traffic contracts). Four types of QoS may be specified: the *Constant Bit-Rate* type specifies a constant peak cell rate, used for real-time services such as video and voice. *Variable Bit-Rate* specifies an average bit rate, which may peak to a specified level for a maximum length of time. *Available Bit-Rate* specifies only a minimum guaranteed rate and *Unspecified Bit-Rate* only allocates whatever bandwidth is available at the time to the input signal. Most types also allow CDVT[1], which

---

[1] Cell Delay Variation Tolerance

allows the system to clump together certain cells [37].

ATM networks use *shaping* (marking and queuing of cells) to ensure a cell VC will meet its traffic contract. *Policing* is the action of enforcing the traffic contracts. If a VC attempts to exceed its traffic contract, or if network congestion occurs, the ATM switches can discard cells with a lower QoS. Of course, it does not make sense to discard only part of a larger frame which has been broken into cells, so the network uses *early- and partial packet discard* techniques to discard all of the packets of the frame. This saves on defunct cells taking up bandwidth.

5.3.4   Routing virtual circuits

ATM networks can establish static, permanent virtual circuits (PVCs) or build a link dynamically[1] (on demand). This is done by configuring each hop between elements as a segment, with the origin and destination nodes negotiating the AAL type and QoS level. The PNNI[2] protocol now routes traffic over the intermediate segments using the Dijkstra algorithm, just as the OSPF[3] protocol does in IP networks. The CAC[4] algorithm calculates whether there is enough bandwidth left on the chosen path to support the virtual connection; if there is then the VC is routed and the "call" established [30].

ATM is a technology that has never managed to fulfil its potential. One of its main original advantages, that of jitter reduction, is rendered obsolete by the speed of modern transmission networks. At 1 or 10 Gbps the longest frames admissible on the network are transmitted in so little time they do not add a significant amount of jitter. The fact that the cell payload (48 bytes) was chosen arbitrarily, rather than as the result of what works well, has also slowed its adoption in large networks. To make matters worse, the cell header is more than 10% the size of the data packet, which makes for an unacceptably large "cell tax".

In the final reckoning, ATM is too complicated (and therefore too expensive) to deploy on the scale that the designers originally intended. The attempt to make one technology that is all things to all people is what is ultimately responsible for ATM's downfall. While it is still widely deployed in carrier networks and it is gaining ground for multiplexing low-speed non-symmetric services such as DSL[5], ATM has been superseded by less complex technologies such as MPLS (see 6.1.2) that offer similar advantages at lower cost.

5.4   Ethernet

Ethernet and its successors FE[6] and GbE[7] have become the default choice for data networks because of its ease of use and adaptability to new services. It is also considerably cheaper than circuit-based technologies like SDH and ATM. It can be shown [38] that a SDH network can be as much as five times as expensive to deploy as carrier-class Ethernet. The main reasons for this are the relative simplicity of Ethernet, as well as the fact that Ethernet equipment has become almost ubiquitous in

---

[1] Soft Permanent Virtual Circuit (SPVC)
[2] Private Network-to-Network Interface
[3] Open Shortest Path First
[4] Call Admission Control
[5] Digital Subscriber Line
[6] Fast Ethernet
[7] Gigabit Ethernet

small and medium sized data networks. The availability of both low-end (Enterprise) and high-end (Carrier-Class) equipment has seen a proliferation of the technology and a subsequent consistent drop in the price-per-bit of an Ethernet system. ATM and SDH have never had low-end solutions in the Enterprise sphere and thus have remained very expensive.

The traditional problem with Ethernet has always been that it does not really qualify as a transmission technology because it was not designed to be one. This means that the likes of SDH have far higher levels of availability in both equipment and network resilience. Ethernet also has very limited Layer 1 (transmission) functionality, even when using high quality optic fibre and laser transmitters. FE and GbE systems may have links in the region of 10km, but if one compares this to the 100km or more per link possible in SDH, it becomes clear than Ethernet was not designed for transmission. In addition, Ethernet only uses CRC[1], which only tells the receiver whether or not the Ethernet frame has any errors or not. If it is found that the frame has errors, it is considered corrupt and the frame discarded, necessitating retransmission of the frame. SDH and ATM use FEC[2], which enables the receiver to "fix" transmission errors, saving on retransmission bandwidth and latency.

5.4.1.1 The Ethernet frame

From the view of the OSI[3] protocol stack model, Ethernet is both a Layer1 and a Layer2 protocol. The Layer1 part of Ethernet manages everything concerned with the physical layer, meaning the actual transmission medium (copper, fibre, etc). Many different bit encoding and transmission schemes are used by the various versions of Ethernet, but this is transparent to the network designer. It is enough to know what data is actually transmitted on the PHY[4].

The datagram that is transmitted over the physical layer is known as the MAC[5] Frame[6] (Figure 13). It contains 7 octets[7] (called a preamble) with a specific data pattern (10101010), which signals the receiver that a frame is inbound. It also helps with clock recovery and synchronisation. The next octet also has a fixed bit-pattern (10101011) and indicates to the receiver that the Ethernet frame is about to follow. After the Ethernet frame a "space" is left called the *interframe gap*, which ensures a low rate of frame collisions. The gap is usually 12 octets long, but may be shortened with some varieties of Ethernet [39].

Preamble	Start-of-frame delimiter	Ethernet Frame	Interframe Gap
7 octets of 10101010	1 octet of 10101011	64-1518 octets	12 octets

72-1526 octets
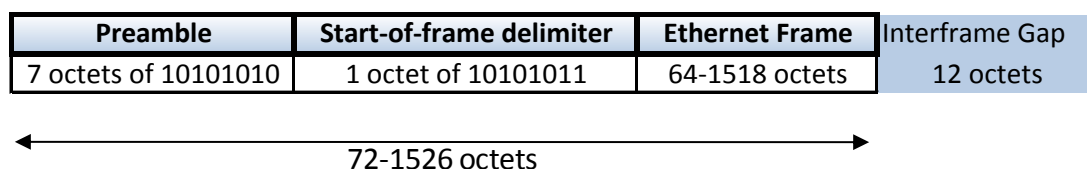
Figure 13: The 802.3 MAC frame

---

[1] Cyclic Redundancy Checking
[2] Forward Error Correction
[3] Open System Interconnection
[4] physical layer
[5] Media Access Control
[6] According to the IEEE 802.3 standard
[7] Unlike a byte, an octet is transmitted least-significant-bit first

The Ethernet frame (Figure 14) contains the MAC address of the destination node, followed by the address of the source node (both 6 octets long). The MAC addresses are always universally unique, but may be overridden. The next two octets indicate the length of the frame. They also indicate the type of Ethernet frame, as the commonly used Ethernet II always has frames of between 64 and 1518 octets. If the length field has a value of 1536 or greater it means that a different frame type is in use. Then the value of the field specifies which protocol is encapsulated (IP, IPX, etc) [40].

The MAC header is followed by the actual data which the frame must deliver from source to destination. The data must always be 46 bytes or more, otherwise the frame is called a *runt*. It must also be shorter than 1501 bytes, otherwise it will be a *jumbo* frame. Most networks don't allow the transmission of runts and jumbo frames, as they are usually caused by errors in the software or hardware and will simply be discarded.
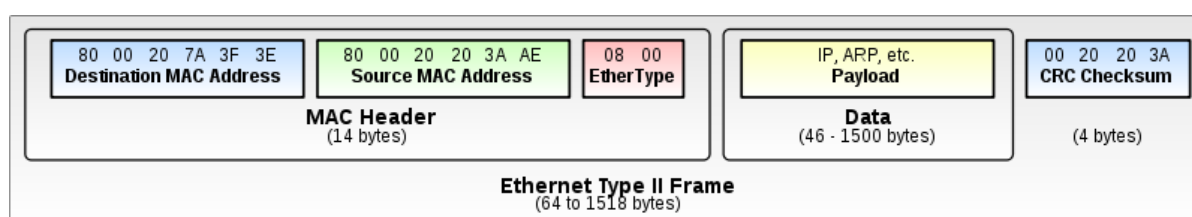
Figure 14: The Ethernet II frame [41]

The last section of the frame is called the CRC and contains the checksum of the entire frame. If a recalculation of the checksum at the receiving node is different from the value stated in the CRC field, the frame is assumed to be corrupt and is discarded.

5.4.1.2 Building networks with Ethernet

Ethernet is eminently suited to building small networks such as LANs[1] and other enterprise networks for several reasons: the equipment is inexpensive; the network does automatic speed negotiation and will to some extent configure itself.

When an Ethernet network is first switched on, none of the switches in the network know anything at all about the network so each switch must learn its place in the network. When a MAC frame from an unknown source is received on a switch interface, the switch stores the source MAC address in its MAC lookup table, as well as the port number the frame was received on. In this way the switch builds up a picture of what other switches and devices are connected to which of its ports. To forward a MAC frame the switch looks up the destination MAC address in its lookup table and then forwards the frame out the corresponding port. When a frame is received with an unknown destination address, the switch floods the frame out every port that connected to the network, except for the incoming port.

If redundant links are established between switches, loops could occur when frames are endlessly circulated around the ring formed by the redundant links. The Spanning Tree Protocol was invented

---

[1] Local Area Network

to reliably stop loops from occurring. STP will elect a root bridge, which will usually be the switch with the lowest MAC address. The lowest cost paths to the root bridge will then be calculated and the redundant ports in a switch, i.e. the ports with higher cost paths to the root, will be disabled. Only when a failure on one of the primary paths occurs will the redundant links be activated. To ensure that all of the elements on the network know where the root bridge is and how the network is connected, BPDUs[1] (special control frames) are sent to all elements, which contain information on bridge IDs and root path costs. The RSTP[2], which was introduced in 1998, can converge a network after a topology change within a second, as opposed to 30 to 50 seconds with the original STP [40].

5.4.1.3   *Virtual local area networks*

VLANs can be used to subdivide the network into segments. A university may choose to divide its network so that each faculty is assigned a specific VLAN. This would mean that no device on one segment of the network can access resources on another segment. It also ensures security, as no device on one VLAN can access data from another VLAN, even though frames from both VLANs could be switched on the same switch and even on the same cable (where shared infrastructure is used).

Aside from the ability to segregate traffic, VLANs also reduce broadcast traffic on the network, as a switch will only send out a broadcast to all the ports connected to a specific VLAN, rather than all the switch ports. In addition, network administration becomes easier, as a user may move his computer to a different part of the network, but still remain connected to the same VLAN, thus obviating any manual configuration changes.

5.5   Wavelength Division Multiplexing

WDM is a system that allows many independent signals, each at a bit-rate of 10Gbps or more to be optically multiplexed onto a single fibre. Optic fibre pairs are often the most expensive parts of a large optical transmission system because of the high cost of deployment. When one moves into the arena of having several 10 Gbps signals that are transmitted over a long distance (more than 100km), then the cost of regenerating each signal individually becomes prohibitive. A multi-channel optical amplifier is much more expensive than a single-channel regenerator, however its cost is independent of the number of channels employed, therefore it costs in quite quickly when many channels are used.

The steady growth of the bandwidth needs of NREN clients could mean that the high cost of deploying WDM will be offset by the cost of deploying several channel regenerators and more than one fibre pair on a route.

5.5.1   CWDM

Coarse WDM employs a wide channel spacing of 20 nm which means that only 20 channels are available for use on most optic fibre types. The older speciation fibre (G.652) has a water absorption peak at around 1380nm, which effectively means that several CWDM channels are unusable on this fibre type (Figure 1). Newer types of fibre (G.652D, G.655) have much lower attenuation peaks and

---

[1] Bridge Protocol Data Units
[2] Rapid Spanning Tree Protocol

may be used over the entire CWDM band. The 20 channels are spaced between 1270 and 1610 nm.

CWDM is much cheaper than DWDM because of the wider channel spacing. This means that low-cost optical filters and uncooled lasers may be used [42].

5.5.2   DWDM

Dense WDM uses a very close channel spacing of either 100 GHz (40 channels) or 50 GHz[1] (80 channels) in the C-band (1530 – 1565 nm). This requires laser sources that are very stable and have a very narrow spectral bandwidth. This is only possible with cooled laser units. Recovery of a multiplexed signal also requires very narrow-band optical filters, all of which make DWDM very expensive. This is offset by the very large channel capacity.

The C-band is also the wavelength band which has the lowest attenuation, usually below 0.25 dB/km, depending on the type of fibre used.

5.5.2.1   *Transponders vs. muxponders*

A transponder is a line card on a WDM element that converts the "grey" client signal (uncooled laser at 1310 or 1550nm) into a "coloured" wavelength (cooled laser at a specific wavelength). The client signal maybe anything from Ethernet to SDH to FICON[2], depending on the model of the transponder.

Some transponders have two line laser sources, which enable the one card to establish both sides of a ring on a ringed network. Alternatively it can be used to light two different wavelengths on a point-to-point link.

A muxponder[3] is a tributary card which combines the transponder functionality of wavelength translation with an Ethernet Layer2 switch. This enables the network designer to establish a Layer2/3 network on the transmission medium without having to depend on the underlying transmission technology to provide protection switching. In this way MPLS *fast reroute* or Ethernet *rapid spanning tree protocol* may be used rather than the G.709 protection. Other Layer2 functionalities such as VLANs will also be available, making the transition from packet-based Ethernet to the synchronous G.709 OTN easier (see 5.6).

5.5.2.2   *C1 vs. C2 band*

The ITU have standardised the specific wavelengths that may be used in a DWDM system. The C1 band is specified as 40 channels in the band between 1530 and 1565nm, at a spacing of 100 GHz (1.6nm). The C2 band is very similar, 40 channels at 100 GHz spacing with a *0.8nm offset*. This allows the *interleaving* of the two bands. The full C-band channel count would then look as follows: 1, 41, 2, 42, 3, 43, etc. to produce 80 channels with 50 GHz spacing. Not all manufacturers count from the same side or use the same first channel wavelength, but all obey the grid wavelength plan, to ease interworking [43].

---

[1] 0.8 nm
[2] Fibre Connectivity, an IBM proprietary interconnect standard
[3] Also known as a Xponder

New super-high density systems are being developed with 25 GHz channel spacing, but as discussed in 4.2.1.4, such systems encounter extreme nonlinear effects, which are difficult to compensate for. Buying a second fibre pair with another 80 channels at 50GHz spacing may prove to be more cost-effective than buying a 160 channel system at 25GHz spacing.

Alternatively, one could employ a similar system in the next wavelength band, called the L-band (1575-1610nm). Commercial systems exist that offer 80 channels in the C-band as well as 80 channels in the L-band, however the L-band amplifiers (EDFAs) tend to be quite a bit more expensive than their C-band counterparts. Even so, expanding an 80-channel system to 160 channels may still be cheaper in some instances than buying another fibre pair.

5.5.2.3   Amplification vs. regeneration

A single-channel regenerator is less expensive than a multi-channel optical amplifier by as much as an order of magnitude. However, one must calculate all system costs to get a real picture of the cost of amplifier deployment.

A regenerator performs what is often referred to as 3R: resizing, reshaping and retiming. The incoming signal is resampled and the information "copied" to the output. The signal is now transmitted at the same optical intensity level as it started with at the originating terminal (resizing), thus "resetting" the OSNR[1]. It is also transmitted at the optimal pulse shape (as close to a sign-wave shape as possible to limit the bandwidth of the signal: reshaping). In addition the signal is synchronised with the incoming signal and retransmitted at the original period to eliminate any dispersive effects (retiming) [42].

The optical input signal may also be amplified by using and EDFA. The EDFA consists of a length of fibre doped with the trivalent rare-earth element Erbium. A pump laser with a wavelength of 1480nm[2] is multiplexed with the incoming signal, shining continuously against the signal's direction of travel (Figure 15). The outer electrons of the Erbium atoms in the fibre are raised to a higher energy level and kept there by the pump laser. When a signal photon strikes an Erbium atom, the outer electron releases two photons in the direction of travel through a quantum leap to a lower energy level. Both photons are at the wavelength of the incoming one, thus achieving a doubling in optical power of the incident photon. With a long enough doped fibre section and a high power
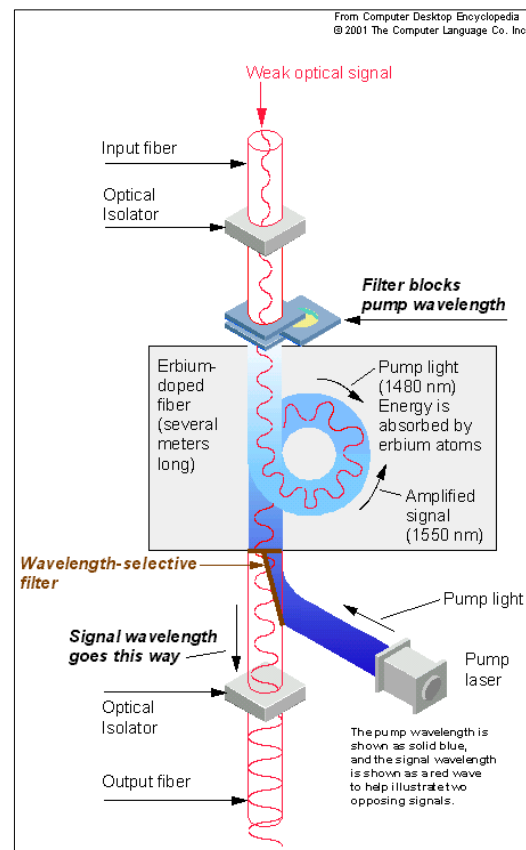
Figure 15: Optical amplification in an EDFA [56]

---

[1] Optical Signal-to-Noise Ratio
[2] Sometimes 980nm is also used

pump laser amplification in the order of 15dB or more may be achieved [43].

There is a noise penalty to pay, as with all active components. ASE[1] occurs when the Erbium electrons decay spontaneously, rather than because of stimulation from incident photons. The photons produced by ASE may also be released in the signal band, becoming optical noise and reducing the signal OSNR.

Regeneration should only be used when the maximum possible channel speed (usually 10 Gbps) is fast enough in the medium term (5 years). In other words, if there is any chance that more than 10Gbps will ever be needed during this time frame, careful consideration is necessary to determine what method should be used. If the decision taken to implement a single-channel, non-DWDM system with regeneration turns out to be short-sighted, it will become a costly mistake, as all the regenerators will have to be replaced with DWDM amplifiers.

Two single-channel systems would require two complete terminals at either end, two regenerators at each span point (roughly every 100km), as well as two fibre pairs. The other option would be the deployment of a WDM system. The additional system cost is then the two WDM terminals and a multi-channel amplifier at each span point.

Depending on the cost of the fibre, a WDM system may cost in with as few as 2 channels, but research suggests that with very inexpensive fibre, the WDM will still cost in at around eight wavelengths.

5.5.2.4   Raman amplification

A Raman amplifier works on a similar basis, except that the underlying principle uses the nonlinear Raman Effect. A pump laser (980nm) working against the signal direction of travel excites silica atoms in the line fibre (no length of doped fibre is required). During the process of stimulated emission two photons are produced for each incident photon, both at the wavelength of the incident signal. Thus the signal is amplified with the amount of amplification increasing the closer the signal gets to the receiving element (43).

5.5.2.5   CWDM vs. DWDM

CWDM offers 20 channels at 10 Gbps using uncooled lasers, as discussed in (5.5.1). This makes the coloured interfaces much less expensive than the cooled lasers used in DWDM (as little as a quarter of the cost). The problem with CWDM is it is only useful on the *reduced water peak* fibre types such as G.652D, which is not yet widely deployed. Much of the available optic fibre is still the older G.652A type, which significantly reduces the usable spectrum of the fibre.

This consideration makes the decision to opt for DWDM a foregone conclusion. In addition, the maximum capacity of the C-band (96 channels) far exceeds the potential capacity of a CWDM system. Also, EDFA optical amplifiers only work in the C-band, and not across the entire wavelength range used by CWDM. EDFA amplification will be required at some parts of the network, so this becomes an important consideration.

---

[1] Amplified Spontaneous Emission

5.5.2.6   Fixed vs. reconfigurable wavelength filters

The Cisco 15454 can be configured either as a ROADM or as a regular ADM with fixed wavelength filters. With the ADM option the designer can choose the wavelength range to be dropped at a site (say lambda 1-4) and drop a different set at the next site (say lambda 5-8). Both sites would then have as much as 40 Gbps which could be dropped there. The only problem with such a scenario would be that the same wavelengths could not be dropped at both sites. It would be possible to retrofit both sites with the optical filters they don't have, but aside from the equipment lead times, two sets of fixed filters cost more or less the same as one fully reconfigurable add-drop unit.

SANReN, true to its mission of building a network that may be easily upgraded without changing the optical design, opted to deploy the full ROADM version of the 15454 at each site.

5.5.2.7   Shared vs. dedicated transponders

After the decision has been made to connect each major site with a 10Gbps wavelength, the question arises of whether to roll each site out on a dedicated wavelength, or to share a wavelength between sites, so as to deploy bandwidth more slowly. In this way, when the point is reached where a single wavelength no longer has sufficient capacity for all the traffic, the larger sites may be upgraded to their own dedicated lambda.

Consider a scenario where the bandwidth was slowly deployed one lambda at a time, on a ring with three client sites as well as a core node, and where all sites shared every lambda (Figure 16). The bandwidth deployed would be 30 Gbps (one lambda to each site). To obey the requirements for card redundancy, each site would contain two transponders per wavelength. Thus 30 Gbps shared over four sites would require 24 transponders.

■ 10 Gbps Transponder
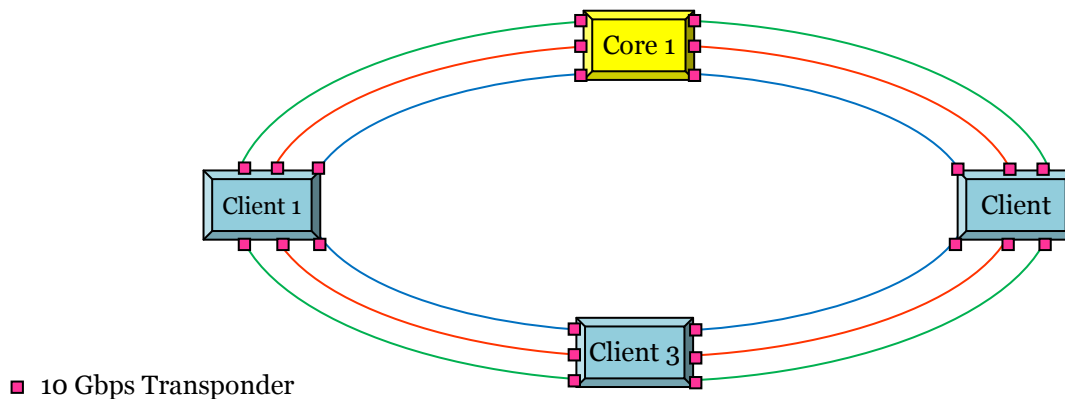
Figure 16: Shared wavelengths

Compare this to the scenario where each site was deployed from the beginning with a dedicated lambda, which was only dropped at the particular site, as well as the core node (Figure 17). One can see that these dedicated lambdas to each site would require only one transponder pair per client site, necessitating only 12 transponders for the entire network.
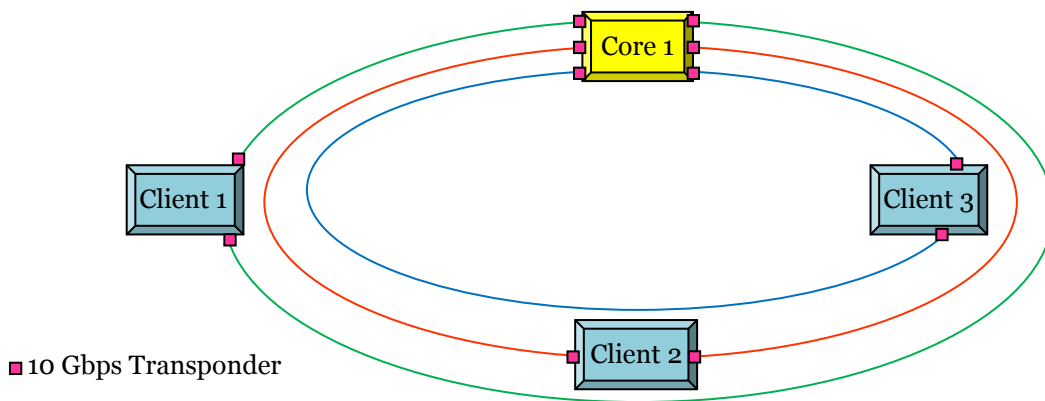
Figure 17: Dedicated wavelengths

The obvious conclusion is that it is better to install a dedicated lambda to each client site during the initial deployment, rather than trying to save first-in cost by sharing the bandwidth around the ring.

5.5.2.8   Coloured vs. grey interfaces

Careful consideration should be given to the use of coloured router interfaces. A "coloured" interface uses a cooled, narrowband laser which functions at a wavelength specified by the ITU wavelength grid (not at 1310 or 1550nm as with grey lasers). It is usually several times more expensive than a grey one. See ITU specification G.694.1 [44].

When mating a DWDM port to a router port, one can use one of two methods:

1. A grey transponder port on the DWDM ADM connected to a grey router interface
2. A coloured interface on the DWDM add/drop module connected to a coloured router interface

Case 1 requires a transponder card with two interfaces (one grey and one coloured), as well as a router card with a grey interface. Case 2 only requires one card with one interface (coloured). Using a coloured router interface obviates the need for a transponder card on the WDM node. This will most often compensate for the added cost of using a coloured router interface instead of a grey one, as the same type of coloured interface is usually used by both the router and the ADM.

5.6   Optical Transport Network

The OTN standard, commonly referred to as G.709 or *digital wrapper*, is an optical encapsulation procedure for adding long-distance transmission capabilities to a variety of common data streams, including SONET OC192, SDH STM-64 and 10 gigabit Ethernet (45). SDH/SONET already has a very strong transmission structure which allows it to be used natively on WDM systems. However, the advanced FEC[1] implemented in OTN permits the use of much longer links and low-power amplifiers.

---

[1] Forward Error Correction

FEC can easily add an effective 5dB of optical budget to a link. It also has its own GCC[1] for out-of-band communications and control. This allows client SDH streams to be encapsulated and transported natively, rather than having to strip the SDH frame and retranslating the client data into a new frame.

Four line rates have been defined [32]:

- **OTU[2]1** runs at around 2.7 Gbps and was designed for to transport SONET OC-48 and SDH STM-16, both of which have a bit speed of approx. 2.5 Gbps
- **OTU2** has a line rate of 10.7 Gbps and is intended for OC-192 and STM-64 signals (which run at just less than 10 Gbps). There is also an over clocked version (OTU2e) for transporting 10 GbE WAN PHY signals which run at 10.3 Gbps.
- **OTU3** was designed for 40 Gbps STM-256 and OC-768 and has a line speed of 43 Gbps.
- **OTU4** is still under development with the view of future deployment of 100Gbps signals.

5.6.1   The OTN frame

The OTN frame encapsulates the native client signal and adds a header and footer (digital wrapper). The payload area is preceded by a 64-byte header and followed by a footer which is 1024 bytes long and contains the FEC. The header consists of a frame alignment section, an OTUk[3] overhead section, a ODUk[4] overhead section and a mapping and concatenation section called a OPUk[5] (Figure 18).
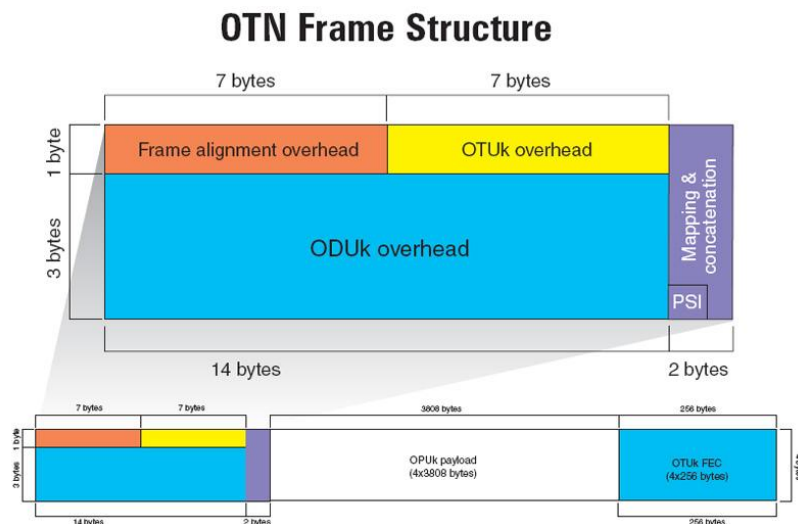
Figure 18: The OTN frame structure [46]

The frame alignment overhead has a fixed binary value of six bytes[6] as well as a multi-frame alignment signal of one byte, which increments once per frame so that overhead signals which span

---

[1] General Communications Channel
[2] Optical Transport Unit
[3] Optical Channel Transport Unit - the *k* indicates the OTU type
[4] Optical Data Unit
[5] Optical Path Unit
[6] F6 F6 F6 28 28 28

more than one frame may be synchronised.

The OTUk overhead is similar to the SDH regenerator section overhead and contains source and destination APIs[1], four error-detection indicators and a GCC.

The 42-byte ODUk overhead is analogous to the multiplex-section overhead in SDH and contains similar functions to the OTUk overhead, as well as six layers of tandem connection monitoring, two GCCs, automatic protection switching bits, a fault location reporting channel and some bytes reserved for future use and experimentation [32].

The mapping and concatenation section supports the adaptation of client signals and contains a mapping specific overhead and a payload structure identifier, which indicates the payload type and structure.

5.6.2   Forward error correction

FEC is a method where some redundant data contained in the data payload is encoded according to a certain algorithm[2] that enables the receiving station to detect transmission errors in the data and also to correct errors, up to a point. Technologies such as Ethernet only employ error detection (cyclic redundancy check). Any frame containing an error is discarded and therefore that frame has to retransmitted.

The encoded data in the FEC portion of the frame enables the receiving node to correct some errors. This means that an error-frame is not discarded, and would not have to be retransmitted. This saves on bandwidth as well as data latency.

Another reason why it is beneficial to keep close track of all transmission errors is that it provides advance warning of link failures. The BER[3] is the measure of how many errors per unit of time are recorded by the system, and is calculated as a ratio of errored to normal bits. A BER of 10^{-6} indicates that one out of every million bits received contained an error.

Some link failures occur suddenly and without warning, such as a fibre break. Others, like deteriorating weather conditions on a microwave link, or a laser reaching the end of its useful life, display a slowly increasing BER, which indicates that link failure is imminent. This advance warning may be used to activate protection schemes, such as switching the traffic to a redundant link. Such switchovers may be done seamlessly, without any loss of data. Even the fastest protection switching involves some link interruption, resulting in data loss [47].

Some BER values are commonly used to determine link quality. A link with a BER of 10^{-12} or better is considered fault-free, whereas a BER of 10^{-7} indicates a severely degraded link. By the time the BER drops to 10^{-5}, the link is considered to have failed completely and all incoming data on that link is ignored [33].

---

[1] Access Point Identifier
[2] Usually the Reed-Solomon algorithm
[3] Bit Error Rate

5.7 Comparing Technologies

Deciding which transmission technology to use may be a difficult decision, so it is important to keep in mind what the network will be used for. This will dictate to a large extent what evaluation criteria should be used.

The SANReN design team decided from the beginning to build a world-class network with full redundancy, high capacity, flexibility and expandability. The available technologies should be considered and compared from several different aspects to determine which one to choose.

5.7.1 Equipment resilience

When building a high-availability transmission network one should consider the innate reliability of the equipment in the network. Ethernet has traditionally been manufactured in two distinct classes, Enterprise and Carrier class. While the proliferation of low-cost Enterprise devices has made Ethernet almost ubiquitous, it is to the carrier-class devices that a NREN must turn to when building its own network. The new generations of carrier-class Ethernet have some levels of built-in equipment redundancy, and although it may be argued that carrier-class Ethernet equipment is still not as reliable as carrier-class SDH, the difference is shrinking and one can now seriously consider building a network with a carrier-class Ethernet core, rather than a traditional SDH core. On the other hand, both ATM and OTN were conceived as carrier-class technologies from the start and offer several levels of equipment redundancy. There are three aspect of equipment resilience that should be kept in mind: the control plane, the switching core and the transmission components.

When the control plane fails in an Ethernet device, the lookup table is lost and the switch can no longer perform its function. This risk may be mitigated by having a duplicated control plane. SDH, ATM and OTN on the other hand, do not require their control cards to carry traffic, so when the control plane fails, only the configuration, management and network visibility of the device is lost, it will not drop traffic.

As many of the other critical and traffic-affecting components as possible should be duplicated, such as laser transceivers, power supplies and switching fabrics. All carrier-class devices should be able to satisfy these requirements.

5.7.2 Network resilience

The main advantage of SDH and OTN has always been its very fast rerouting, which is always less than 50ms. In comparison, traditional Ethernet rerouting/convergence times take about one second in modern networks with rapid spanning tree protocol, and could be in the order of several minutes for very large networks with older protocols. When one considers that at 10 Gbps line speed, even a 50ms downtime translates to 500 Mb. This amount of data loss can be unacceptable to high-availability systems like SANs and supercomputer grids. It is therefore crucial that a transmission technology is chosen that can detect slow failures and automatically reroute traffic to compensate, so that no traffic is lost, except in disastrous circumstances. Even then, the traffic loss must be kept to an absolute minimum. SDA, OTN and ATM are all well-suited to satisfy these requirements.

5.7.3 Technology for the SANReN network

It may be argued that a non-commercial network like a NREN does not have very high availability requirements, but when one considers that the SANReN will provide access to the CSIR's supercomputer grid, as well as be the backbone for a diverse set of research projects, one starts to appreciate the importance of building the best network that can be afforded.

The SANReN design team considered all of these factors and decided that for implementing a high-availability, high bandwidth metro backbone network, the only transmission technology that could satisfy its needs over the long term was an optic fibre-based DWDM system with a G.709 transport core.

Neither ATM nor SDH have the capacity and expandability that SANReN looked for and both are very expensive. To make matters worse, both technologies have innate disadvantages when one is trying to build a next generation, future-proof, packet-based network. This makes the technology choice much simpler. OTN/DWDM has a clear advantage when it comes to resilience, robustness and expandability, and proved to be the obvious choice for the network backbone. It meets the requirements of very high capacity, expandability, flexibility and redundancy that the NREN will have to deliver to its client institutions. In addition, it features FEC, which none of the other technologies can offer.

At the lower end of the scale, to implement access rings, carrier-class Ethernet will be deployed. Full redundancy will be used wherever possible, including dual laser transceivers, power supplies etc.

This two-tier approach should satisfy all the transmission needs of the SANReN network, and also facilitate an easy upgrade path between low-end and high-end sites.

The next chapter will focus on the routing technology needed to implement a fully-fledged Layer 3 network on top of a backbone transmission network.

Chapter 6

Choosing a Layer 3 Technology

Several choices of network layer (Layer 3) technology are available once the underlying transmission technology has been decided on. Since many of the services that the network should support are data services, and more specifically *packet* data services, it is inevitable that there will be a foray into the OSI protocol stack. The most important question to ask is how complex a level of service will be offered by the network. The NREN could decide to offer nothing more complicated than simple Layer2 bandwidth pipes to connect the institutions together, leaving all higher functionality up to the institutions themselves. This will probably not be a satisfactory solution, as it would leave all routing and interconnecting issues to its users, defying the very purpose of the NREN, which is to create a common network that all parties can connect to and providing access to the same services for all users.

Further, the NREN could decide to build an IP network, and only manage the routing of all traffic over the network. In such a case all the NREN would have to do is establish some large routers, and interconnect them to each other and the client sites. This would once again leave the institutions themselves to establish more advanced services between themselves, since a straight-forward IP network does not support services like VRF[1]s, VoIP and QoS. Again, it could be argued that the purpose of the NREN is to build a network that is as easy to use as possible and that supports every advanced service that the institutions could ever require. TENET/SANReN has taken this view, and has undertaken to build a network that is future-proof, meaning it will support any service that can be envisaged at the moment. If any new services are developed in future, and demanded by the client institutions, TENET/SANReN will try and upgrade the network to support those services.

This narrows the choosing of a data network technology down to two options: ATM and MPLS. Even though ATM is primarily a transmission technology, it is much more advanced and versatile than SDH or similar technologies, and may be considered to be a Layer 3 protocol as well. ATM will not be discussed further, please refer to *section 5.3* for details.

---

[1] (Virtual private network) Routing and Forwarding

6.1   Advanced Network Services

6.1.1   Quality of service

QoS in a network may be hard to define and even harder to implement. Specifying very high levels of QoS may also make the network very expensive to build, so the QoS levels must be chosen with care. In data networks, QoS usually refers to the resource reservation control mechanisms, and not to the actual quality of service achieved on the network. This means that having a specified QoS guarantee mechanism is really only an issue when there is congestion on the network.

The easiest way to ensure a high level of service quality is simply to have more bandwidth available than is strictly required. On commercial networks it costs money to provide more bandwidth than is needed, so there is seldom if ever a surplus of bandwidth. On a research and education network, however, it would be possible to deploy more bandwidth than is needed. However, these kinds of institutions usually find a way of using up however much bandwidth is supplied to them. It may be argued that in a complex network the flow of data is so deterministic and unpredictable that it can never be assumed that there will *always* be enough bandwidth available to obviate the need for a QoS mechanism.

Some kinds of traffic are naturally more resilient to QoS problems than others. If limited bandwidth forces a file download to a very slow rate of packet transfer, then the user has to wait longer, but the transfer is unaffected otherwise. On the other hand, if a VoIP or video service is subjected to the same bandwidth constraints, the signal may become so distorted as to be rendered useless.

The only effective way of ensuring QoS is to assign priorities to certain types of traffic. Video traffic has higher demands on QoS because a video signal is more easily distorted by poor quality than most other forms of signal. For example, in analogue systems the minimum SNR[1] for video is 40dB[2] (a factor of 10 000), but for voice telephony it is only 30dB (a factor of 1 000). Therefore, priority is given to video signals, then to voice and after that to other kinds of data. Intra-campus file transfer will typically also be given a higher priority than, say, Internet browsing [47].

The only conclusion one can draw is that regardless of the amount of bandwidth planned for the network, the underlying technology should at least be capable of establishing some level of QoS, i.e. one cannot just ignore prioritisation issues on the assumption that there will always be enough bandwidth.

6.1.2   Voice over IP

Packet data networks have for a long time offered a significantly lower price-per-bit than traditional, circuit switched networks, primarily because of the proliferation of low-cost Layer2 and Layer 3 devices. Voice telephony, on the other hand, is almost singlehandedly responsible for the entrenchment of circuit-based networks. Telephony requires very high performance standards and telcos have traditionally exploited the situation by justifying the deployment of very expensive

---

[1] Signal-to-Noise Ratio
[2] decibel

equipment (such as ATM and SDH) and recovering the costs by charging exorbitant fees for phone calls. This has lead to the present-day phenomenon where telephony represents less than 10% of all traffic on global networks, yet earns between 70 and 80% of the revenue [49].

Clearly this situation is unsustainable economically and sooner or later a technology would be developed that used the packet data networks' cheap price-per-bit but offered the same service as POTS[1]. VoIP is a concept that has been talked about for many years but has only really become mature enough to be used commercially in the last half-decade or so.

Once a QoS mechanism has been established on a network that ensures low latency[2] and guarantees that packets comprising a voice conversation will always be delivered, and in the right order, the scene has been set for translating an analogue voice conversation into a packet stream.

It goes without saying that all the packets in the conversation data stream must be routed along the same path through the network, to avoid packets being delayed or delivered out-of-turn. For this to be implemented successfully, a path or virtual circuit must be commissioned between the caller and the receiving party. SIP[3] is the most widely used way of initiating a call over an IP network.

In practise, running a VoIP network and integrating it with existing POTS networks requires several components. Firstly, the analogue voice call must be digitised and packetized. This may either be done by a VoIP phone handset[4] or by a VoIP-capable PABX[5] in office/campus environments. Legacy, non-VoIP PABXs may be integrated by using a VoIP gateway device, which intercepts the circuit-switched lines to the PSTN[6], usually ISDN PRIs[7], and routes the traffic onto the packet network [48].

Secondly, a device must set up the call between the calling and receiving parties. This is usually performed by the VoIP switch, which uses SIP signalling to notify all the intervening network devices to establish the call. On a VoIP-only network (such as campus network) this is all that is required, other that an underlying Layer 3 network which supports QoS and SIP.

To connect a call to a POTS phone on the worldwide PSTN, however, one also requires a way to interconnect the VoIP/packet network with the PSTN/circuit network. The high-end VoIP switches are capable of this as well, which is why they are situated on the border between the two worlds. This interconnect point is usually also where the billing takes place. Clearly, some arrangement must be made between the VoIP and PSTN providers to reduce costs, otherwise there would be no cost-saving in using VoIP. Typically the PSTN providers offer reduced interconnect charges for national and international calls, but not for local (same-city) calls. This forces the VoIP provider to keep the PRIs between the customer PABX and the PSTN in place, which protects the PSTN provider's access network investment [49].

In South Africa, the unique situation has developed where the cell phone interconnect charges are so

---

[1] Plain Old Telephony Service
[2] Data packets get delivered with a minimum of delay
[3] Session Initiation Protocol
[4] The standard way of translation for residential use
[5] Private Automatic Branch Exchange
[6] Public Switched Telephony Network
[7] Integrated Services Digital Network, Primary Rate Interface

high that it has spawned a whole new industry where LCR[1] providers intercept all PABX calls to cell phone numbers and reroute them to a so-called *SIM[2]-farm,* where the call goes directly from the PABX onto the relevant cell phone network instead of onto the PSTN or VoIP networks. The only reason this makes sense economically is because the SIMs in the SIM-farms are all purchased with free call-minutes, which the cell phone providers offer as an incentive for subscribers. Using the free minutes, the average call charge to the institution drops to below a level that can be offered by either the PSTN or VoIP providers.

This is clearly an artificial situation which is unsustainable in the long run. Future drops in VoIP-to-cell phone interconnect charges will make VoIP more affordable and render the entire LCR industry obsolete.

TENET was involved in a lengthy RFP process to examine the soundness of converting all its client institutions' PSTN traffic over to VoIP on the GEN3/SANReN network. The outcome was inconclusive, but firm plans remain in place for TENET to become a VoIP service provider in its own right. Once the interconnect charge situation has normalised a VoIP switch may be added to the SANReN network and TENET can then manage its own interconnects with the other providers, quite likely at rates much more favourable than is available to other commercial entities.

6.1.3   Virtual private network routing and forwarding

VPNs are used when establishing sub-networks on a common backbone infrastructure. A NREN institution may have several campuses connected to the common backbone which is also used by other institutions. The institution would not want users from other organisations to have access to its data, and would therefore request the NREN to establish a VPN tunnel between the campuses. The dataflow could then not be intercepted by anyone else and the institution would not have to worry about security and firewalling issues. Different departments inside the same institution could also establish VPNs between themselves. This is not unlike VLANs in Layer2 networks.

When a VPN is built across a network, an extra tunnel header/label is added to the Layer 3 packet for all packets that must traverse the VPN. Those packets are then routed according to the tunnel label first, before the IP header itself is read.

While there are several IP protocols like GRE[3] and IPSec[4] which perform VPN functions in IP networks, this kind of safe virtual circuit is particularly well managed by MPLS [50].

6.1.4   IPv6

Individual network devices will always need an addressing scheme that assigns each device a unique name. The problem with the current IP addressing scheme IPv4[5], is that it has far too few unique

---

[1] Least-Cost Routing
[2] Subscriber Identity Module
[3] Generic Routing Encapsulation
[4] IP Security
[5] Internet Protocol version 4

numbers[1]. It is already the norm in first-world countries for each person to own several networked devices: be it a PC, a laptop and a cell phone, or even home appliances such as fridges and dish washers. When one considers that there are fewer IPv4 addresses than there are people on Earth, one sees that the world is fast approaching the situation where every single IPv4 address is assigned.

The only real solution to this problem is a numbering scheme which has a far greater address space. IPv6 has a total of 3.4E38 addresses. To see this in perspective, if one imagines the IPv4 space as being the size of a cigarette box, IPv6 is the size of the solar system. All that remains to do is to upgrade the existing IPv4 Internet into a network that can support the IPv6 addressing scheme. Networks that only support IPv4 can be used to "tunnel" IPv6 over by introducing IPv6 as a layer4 protocol. If the *Protocol* flag in the regular IP header is set to a value of "41"[2] the routers at the edge of the network will recognise that IPv6 addresses are being used, which can be used to expand a small number of IPv4 addresses assigned to an institution into a much larger address space. Ideally though, the NREN should support IPv6 natively, which means that all the NREN routers will recognise the source and destination addresses of all packets as IPv6. This will allow the NREN to assign as large a block of IP address to a given institution as they will ever require [51].

6.2 Multi-Protocol Label Switching

MPLS is a switching scheme that superimposes another layer of control on top of the Layer2 (Ethernet) backbone in what is sometimes referred to as Layer2.5. Although it is also a packet-based technology, it has many of the advantages of connection-oriented technologies, like traffic engineering, QoS and out-of-band control. On WAN[3] networks it combines the best characteristics of packet-based and circuit-based (connection-oriented) networks, and it is increasingly being deployed in carrier networks to improve the resilience and performance of Ethernet/IP networks. The MPLS protection scheme called Fast Reroute can achieve SDH-like rerouting times, in the order of 50ms.

The other great advantage of MPLS is that is largely protocol agnostic, meaning that it may be deployed in almost any kind of network. This is called ATOM[4] and is possible due to MPLS' ability of accepting any length of packet, unlike ATM [52].

6.2.1 Teletraffic engineering

Teletraffic engineering is the science of improving the performance of networks by examining traffic patterns and flows, and using the resulting statistics to predict future traffic trends, with the view of avoiding congestion and generally improving the throughput and service levels offered by the network. It may be used in real-time to switch traffic from a congested or failed path to another, in order to minimise traffic downtime and data loss.

MPLS network use a scheme called RSVP-TE[5]. If a path through the network becomes congested, the network control layer can dynamically assign new paths for the traffic to avoid the congested areas.

---

[1] Less than 4.3 million
[2] The Transfer Control Protocol (TCP) has a value of "6"
[3] Wide Area Network
[4] Any Transport Over MPLS
[5] Resource reservation protocol – traffic engineering

This form of traffic engineering is not available in simpler networks like IP networks, which use routing protocol based on fixed criteria, such as shortest path [53].

RSVP is a network *control* protocol that allows elements of the network to reserve resources for a specific purpose, along a specified path. It is used to set up LSPs[1] between nodes, which function as virtual circuits, not unlike PVCs[2] in ATM networks.

When traffic engineering is not used to establish paths through the network, labels are distributed across the network using LDP[3]. This is only used for best-effort LSPs.

6.2.2 Tunnelling over a MPLS network

Figure 19 illustrates how a MPLS Layer may be established over a transport infrastructure consisting of edge routers, core routers and DWDM ROADMs.

Once an integrated MPLS network has been established it may be used to encapsulate and transport all manner of traffic, from IP packets to Ethernet frames and ATM cells.

Referring to the diagram in Figure 19, the edge router[4] accepts the packet (frame, cell) from the external network (in this case the CPE switch/router). It calculates the FEC[5], which may be based on a IP address, and designates the proper LSP[6].The router then adds (pushes) a MPLS label onto the packet and forwards it to the core router[7], which either just switches the packet onto the correct LSP, or adds a second MPLS label. In this way two layers of tunnelling may take place on the Layer 3 network. The LSR creates a small number of high capacity paths between its peers (other LSRs), each of which contains many lower capacity paths between the LERs.

The MPLS traffic is now itself encapsulated into an OTN frame for transport over the DWDM network. Each wavelength will now represent a virtual circuit which transports the MPLS traffic over the optical network.

At the destination of the path the DWDM node strips off the G.709 frame and delivers the MPLS path to the LSR, which pops the second MPLS label (*decapsulation*) and forwards it to the LER, which in turn pops the first label and routes it to the destination router/switch.

Each element in the MPLS network only looks at the topmost label of the "stack" of MPLS labels and makes the routing decision based on it, rather than looking at the lower level protocol labels such as the IP header. In effect the MPLS network thus consists of a small number of paths across the network and the routing decisions that the routers have to make are therefore quite simple, which means that traffic may be routed very fast [53].

---

[1] Label Switched Paths
[2] Permanent Virtual Circuits
[3] Label Distribution Protocol
[4] Also called a LER – Label Edge Router
[5] Forward Equivalence Class
[6] Label Switched Path
[7] Also called a LSR – Label Switch Router
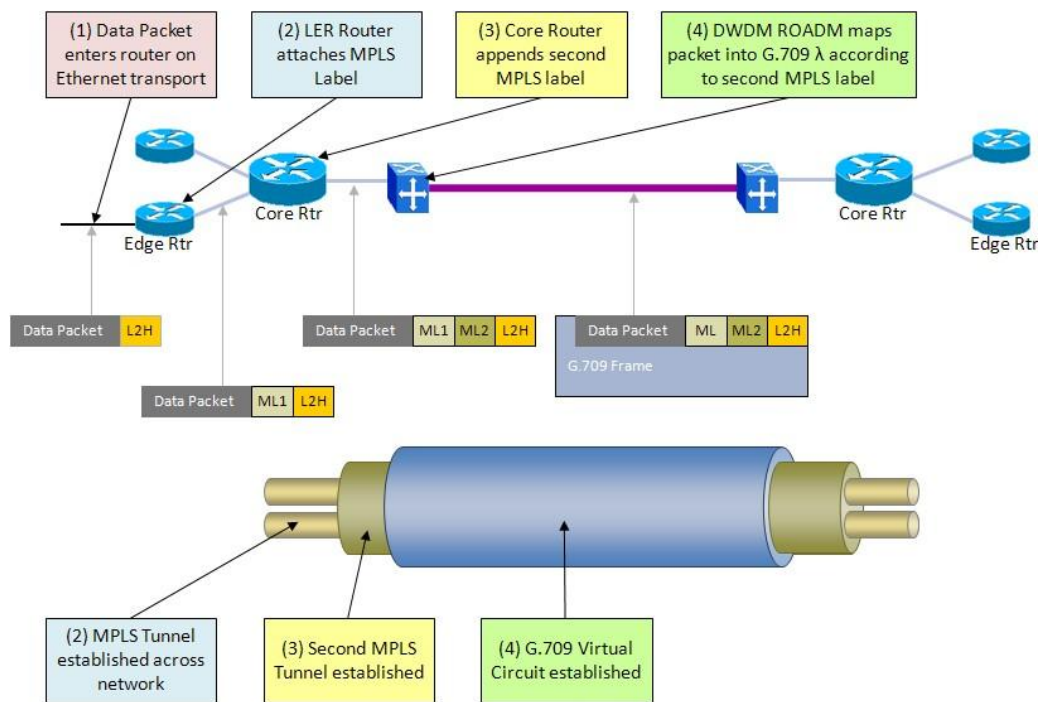
Figure 19: MPLS tunnelling across the network

6.2.3 MPLS labels

The MPLS label that gets appended to the data frame consists of a 20-bit *label* field followed by a 3-bit *traffic class* field, which indicates the QoS level, as well as *explicit congestion notification.* The next bit indicates whether the current label is the last one in the stack. Lastly there is an 8-bit *time-to-live* field, which is decremented at every hop (Figure 20). Once the value reaches zero, the label (and corresponding frame of data) will be discarded. Note that the label is inserted between the Layer2 (Ethernet) and Layer 3 (IP) headers. This explains why MPLS is referred to as Layer2.5.
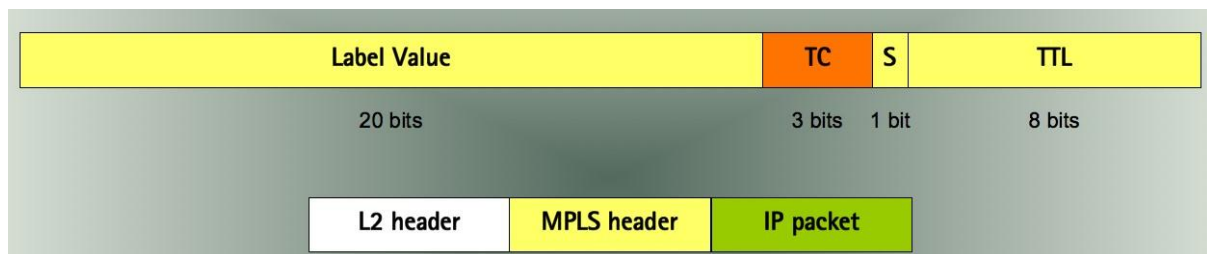
Figure 20: The MPLS label header [63]

6.2.4 Local protection (Fast Reroute)

One feature of RSVP-TE is the ability to configure backup routes to cater for link or element failure. Each LSP through the network may have primary path which carries traffic, as well as an alternate

route assigned. If the LSP crosses many elements, a single link failure between adjacent nodes will trigger a local reroute, where the element before the failure establishes a new path to the element after the failure, as soon as it notices that a failure has occurred. This usually takes less than 50ms, providing SDH-like protection. Notifying all the nodes before the failure to establish the backup path may take too long. RSVP-TE may now activate the *planned* backup path when all the nodes between the origin and failure have been notified of the failure. The local reroute path will then be torn down and normal traffic flow resumes, until the failure is repaired, at which time a new update circulates through the network and the network goes back to the state it was in before the failure [50].

When one compares traditional IP routing, where the global IP routing table may contain close to a quarter of a million entries, it is easy to see why MPLS is fast becoming the defacto choice for building complex networks [50].

6.2.5 Routing equipment

A router is a network device that routes traffic according to the Layer 3 (IP) header. It operates on two planes, the forwarding plane and the control plane. The control plane is what maps the network, based on network information received on all its incoming interfaces. The forwarding plane is where actual traffic is received on one interface and forwarded (transmitted) onto another interface, according to routing information supplied by the control plane.

The control plane constructs a RIB[1], also called a routing table, from information received from incoming data packets from other routers, as well as the status of its own interfaces and pre-configured routes called *static* routes. Some routers keep separate FIB[2]s, which are optimised to minimise the address lookup time, increasing the speed at which a data packet may be forwarded, whereas RIBs are optimised for updating by routing protocols.

The data (forwarding) plane must also handle congestion, which occurs when the router receives packets faster than it can forward them. Several schemes are employed to alleviate congestion, but all involve dropping packets, incurring data loss and necessitating packet retransmission.

6.2.5.1 Label edge routers

The LER handles the task of labelling the incoming data packets and turning an IP packet or ATM cell into a MPLS datagram. It must also perform the task of establishing the LSPs across the network and distributing the labels to all routers downstream. It has a relatively complex function but as it is usually used in an aggregation function, throughput speed is not necessarily a primary requirement.

6.2.5.2 Label switch routers

The LSR is on the highest level of routing in the network and as such its primary function is to switch packets as fast as possible. It has a relatively simple function since all inbound data packets already have MPLS labels; the most complex function it ever has to perform is aggregate LSPs into other LSPs, appending new labels as it does so. Since it does not have to break the data packets down to extract

---

[1] Routing Information Base
[2] Forwarding Information Base

the IP addresses or route on the IP address, it can be optimised to switch very fast. This is important as it usually collects data streams from several LERs and consequently has to be able to match the throughput of all its subtended LERs simultaneously [53].

6.3 Comparing Technologies

ATM was designed to be the ultimate combination of the advantages of circuit switching, such as fast rerouting and error detection, and the advantages of packet switching, such as low-cost hardware and statistical multiplexing. However, it was ultimately more expensive and less flexible than the existing packet-based technologies that it was meant to replace. Even though it is still widely deployed throughout the world, it is generally recognised that ATM is becoming obsolete. It would therefore be very short-sighted if SANReN were to choose it as a Layer 3 technology for the national NREN network.

IP is without a doubt the most widely deployed Layer 3 technology in the world. The success of the entire global Internet can be ascribed to IP's simplicity, flexibility and ease of use. However, it has some fundamental flaws. It can never hope to provide the kind of QoS, resilience and protection switching speed that present and future users of the NREN will require of it. The low-cost, best-effort kind of approach that it has always followed just does not stand up to scrutiny when a high-performance research and computing network is built.

This leaves MPLS the clear winner. It is a relatively simple method of switching, which implies reduced cost, but still offers all the advanced features and QoS that an NREN needs. It could be argued that it is the ultimate compromise between the advantages of packet and circuit switching, without any of the drawbacks that doomed more complex technologies such as ATM.

SANReN can rest assured that in deciding on MPLS as a Layer 3 technology it can build a flexible and future-proof network.

The next chapter will discuss the integration of the existing TENET network into the burgeoning SANReN network.

Chapter 7

Integrating the TENET GEN3 Network

TENET has been the ISP of all the so-called *TENET institutions*[1] since 2000. As such it has been responsible for requesting site connections, ordering bandwidth upgrades and monitoring all traffic usage on the network. Until the end of 2007 the entire network was built on the Telkom MPLS network and TENET had very little direct influence over the structure and operations of the network. This virtual network was called the GEN2 network.

In January 2008 TENET started to migrate its sites off the Telkom GEN2 network and onto the new GEN3 network. This network is built mainly on the Neotel MPLS core and is markedly different from the old GEN2 in several respects.

7.1 The GEN3 Network

Firstly, there are several kinds of services, as opposed to the GEN2, which only had one kind of service. Neotel has a very new access network and was therefore forced to extend their network a great deal to cater for all 106 sites on the TENET network. All sites that Neotel managed to connect before the termination of the GEN2 (31 March 2008), were connected straight onto the Neotel MPLS cloud, either with NGSDH[2], MetroEthernet or else a wireless solution such as WiMax. These services are all traditional "leased lines"[3].

Sites that could not be reached in time had to accept a slightly degraded non-MPLS service from Telkom. Small sites that either cannot afford the minimum Neotel connection speed of 2Mbps or are not satisfied with either of the other options, are supplied with an ADSL[4] connection by IS.

7.2 International Bandwidth

Secondly, International bandwidth is supplied by IS as a Layer2 service. This was never available in the GEN2 era because Telkom insisted on supplying the bandwidth as a layer 3 service. This bandwidth is delivered as one STM-1 circuit between London and Cape Town, as well as one STM-1 between London and Johannesburg (Figure 21). IS will deliver this service to TENET only until the

---

[1] All the universities in SA, as well as all the research councils except two (Agricultural- and Nuclear Energy research councils)
[2] Next Generation Synchronous Digital Hierarchy
[3] A connection with fixed, uncontended bandwidth
[4] Asymmetric Digital Subscriber Loop

Seacom submarine cable is ready for service (July 2009).
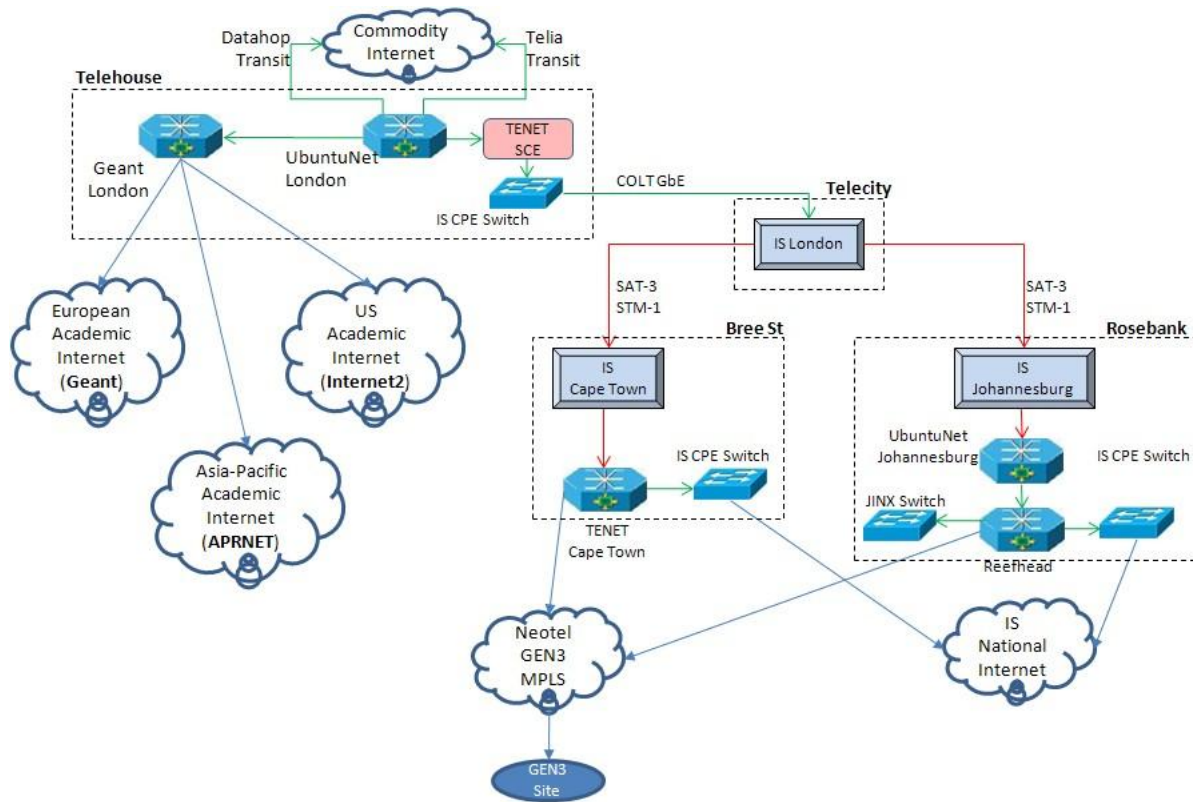
Figure 21: The GEN3-era international network

In the GEN2 network the international traffic was distributed locally to the TENET institutions at the Beachhead by Telkom. In THE GEN3 network the splitting of the international bandwidth stream into two destinations (Johannesburg and Cape Town), necessitated situating the distribution functionality at the London end of the international link. This also gives TENET far more Layer 3 control over the GEN3 network than before. A Cisco SCE[1] splits the bandwidth into many streams so that each institution can only consume the amount of bandwidth it has paid for. The SCE may be configured in such a way that an institution may "burst" into the traffic of other institutions after hours. Many institutions are only active during working hours, so this allows other institutions who have students in hostels etc. to make use of the unutilised bandwidth after hours. Clearly limits should be set on bursting, but as a rule of thumb a large institution with many students who work during the evenings may be allowed to burst to 150% of its purchased capacity.

---

[1] Service Control Engine

7.3   UbuntuNet

As explained before, UbuntuNet functions as the Regional Research and Education Network for Africa, of which TENET is a member NREN. Therefore, in keeping with international standards for NRENs, TENET is supplied with international bandwidth by UbuntuNet. What this means in practice is that all international peering and transit traffic purchased in London for use in Southern Africa is routed via the London UbuntuNet router (Figure 21). This router is also connected to the other RRENs of the world like Gèant and Internet2 for peering research and education traffic. All these academic Internet connections are managed by Gèant, and the traffic flows between the Gèant and UbuntuNet routers situated in Telecity. Regular commodity Internet peering and transit traffic is purchased from commercial suppliers like TeliaSonera and Datahop, in London.

Since IS supplies the international connectivity between the London and South African UbuntuNet nodes, a connection must be established between the UbuntuNet router in Telehouse and the IS router in Telecity. TENET rents a 1 GbE service from COLT[1] which connects the IS router to an IS CPE switch in Telehouse (*green line marked COLT GbE*).

In Johannesburg and Cape Town, TENET manages the South African Cluster on behalf of UbuntuNet. This takes the form of another router in Johannesburg, which receives the International circuits from the transit provider (*red lines marked SAT-3 STM-1*). Since the GEN3 international links would be short-lived, UbuntuNet could not afford to provide a second router in Cape Town. Instead, the local TENET router is used.

From both the Reefhead and Breehead routers there are high capacity (1 GbE) links to both the Neotel MPLS cloud as well as the IS National MPLS cloud (*blue lines*). The Neotel cloud interconnects all sites on the GEN3 network to the TENET/SANReN routers and the IS national network delivers South African commodity Internet service to the TENET/SANReN network.

Peering to national Internet sites that are not on the IS network are established by peering with the other large local ISPs at the JINX[2], to which the SANReN router has a direct fibre connection.

7.4   Seacom

TENET, with financial backing from most of the larger institutions, has negotiated the purchase of a STM-64 circuit (10 Gbps) on the new Seacom East Coast submarine cable. As soon as Seacom indicates that the circuit is commissioned (July 2009), SANReN will accept it at the landing station in Mtunzini and redistribute it among all the TENET institutions (see 9.1).

The circuit is on a 20 year IRU and the capital amount will be paid over six years. The service will be provided at such a competitive price that the contributing institutions will be able to continue paying the same amount for Seacom service as they are currently paying for international bandwidth on the GEN3 network, which is sourced via the SAT-3 cable. After the first six years of service, the institutions will collectively have paid off the IRU. The institutions are all satisfied to continue paying

---

[1] City Of London Telecom
[2] Johannesburg Internet Exchange

the same amount for international bandwidth year on year, so as the bandwidth becomes cheaper, they will simply buy more. Therefore, after the capital amount for the IRU has been recovered by Seacom, the institutions will not stop paying TENET for bandwidth, but instead TENET will negotiate the purchase of more bandwidth, either one more IRU on Seacom, or else a similar circuit on another submarine cable.

Some of the larger institutions have purchased as much as 1 Gbps of bandwidth. Such an institution would then be upgrading its international connection by (typically) *a factor of 50*. The total international bandwidth consumed by the TENET family will jump from approximately 200 Mbps to 10 Gbps.

Institutions that have not been able to contribute to the purchase of the IRU will still get the benefit of it, as the circuit is undersubscribed but overfunded. In other words, the combined funding from the TENET members that did contribute only represents about 85% of the available bandwidth. The other 15% may now be sold by TENET to non-funding institutions. TENET, as a non-profit organisation, will hold the additional income in escrow for future bandwidth purchases on Seacom or other submarine cables.

Figure 22 illustrates how the international network will change with the introduction of the Seacom circuit. Both STM-1 circuits from IS will be decommissioned and the Seacom circuit will become the only source of international bandwidth. The SANReN national backbone will be used to distribute the bandwidth nationally among the various SANReN sub networks, as will be discussed in detail later (see Chapter 9).
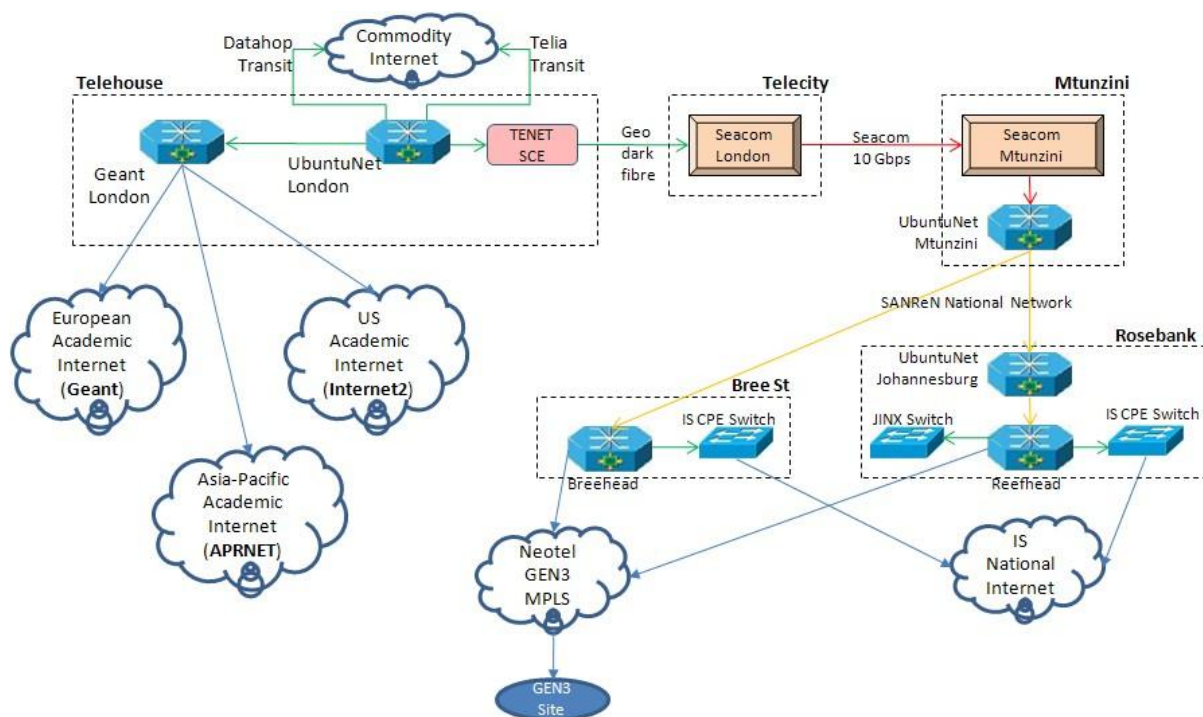
Figure 22: The SANReN-era international network

The connections to the GEN3 MPLS network will remain until all sites have been migrated to SANReN. All the peering links for national Internet traffic will also remain.

7.5   Regional Connections

The UbuntuNet core sites, as the *Southern Cluster* of African RREN nodes, also form the local points of interconnect between the Southern African NRENs. As such, networks from three different countries are currently connected to the Johannesburg UbuntuNet node. Lesotho and Swaziland have not yet established their own fully-fledged NRENs, but in anticipation TENET has already segregated the traffic from these countries' institutions into sub networks.

The [University of Lesotho](#) (at Roma, near Maseru) and the [University of Swaziland](#) (at Kwasuleni near Manzini) have been members of the TENET network for many years, under both HEIST and GEN2). When the GEN3 network was constructed, it was decided to actualise the UbuntuNet network locally, to interconnect networks from different countries in the region. The GEN2 cross-border links were replaced by non-GEN3 leased line circuits which terminate on the UbuntuNet router in the Johannesburg node.

In the future connections will also be established to other NREN's further north, such as KENET (Kenya) and MAREN (Malawi).

7.6   The TENET Core Nodes

In the GEN2 era TENET maintained a single core node in Bellville, called the Beachhead. This node contained the TENET mirror server ([mirror.ac.za](#)) as well as several smaller servers which performed network enhancement functions for the TENET GEN2 network, such as a cache server and an instant messaging server.

After the new GEN3-era core node was commissioned at IS in Bree St, Cape Town (called the Breehead), the equipment at the Beachhead was removed and the node decommissioned. All the functions that the old Beachhead performed were moved to the two new core nodes, the Breehead and the Reefhead in Johannesburg. The Reefhead is the first primary core node on the SANReN network, and will be followed by primary and secondary core nodes in each of the four major cities.

Integrating the existing TENET network into the new SANReN network is a very important step in building a complete NREN. So much work and effort has gone into the TENET network that it would be foolish to ignore the existing network and do everything again.

The next chapter looks at the basic principles that were employed in designing and building the SANReN.

Chapter 8

Designing the SANReN: Principles

8.1　Design Principles

The SANReN was designed primarily as a research network. In the light of the fact that research institutions traditionally have a large demand for bandwidth but a limited budget to buy it with, it was decided to make the fundamental design principle of the network the provisioning of more bandwidth than the institution can use. If, for example, an institution was given a 1 Gbps connection and managed to use it to capacity regularly, the connection would be upgraded to 2 Gbps. Similarly, if an institution managed to fully utilise an entire 10Gbps wavelength, it would be given a second one.

This meant from the outset that the SANReN would be a WDM-based multi-10 Gbps network, as several of the target institutions had the potential of making full use of an entire lambda. Smaller institutions as well as smaller sites of large institutions would be provisioned on a shared 10 Gbps grey ring.

A shared ring could provide 10 sites with 1 Gbps each, or 20 sites at 500Mbps, or any similar combination where the available 10 Gbps of bandwidth was divided between the sites on the ring. If ever a smaller site that was connected on a grey ring required more bandwidth than it was assigned, but there was no more unassigned bandwidth on the ring, and the other sites could not release some bandwidth to cater for this site, then the entire local network would be upgraded. If there was already a DWDM system running on that geographic part of the network, then the largest (or fastest growing) site on the grey ring would be upgraded to a DWDM site. It would then be assigned its own wavelength and 10 Gbps service. The bandwidth released on the grey ring would then be available for reassignment to the sites left on the ring.

 In the case where no DWDM system was yet commissioned on that part of the network, a new system would be designed and installed. If more than one site warranted a future upgrade to 10 Gbps, all the sites would be upgraded at the same time, to justify the expense and effort involved in commissioning a whole new system.

8.2   Technology Choice

The SANReN design team decided on a minimum specification list for deciding on a vendor to provide the equipment for the network.

8.2.1   Core router specification

The primary purpose of the core label switch router is to route packets at a very high speed. Since it forms the backbone of the national network, high availability is crucial, but the complexity of functions it performs is not as high as that of the PER, since all traffic in and out is labelled MPLS frames.

Required features:

- Very high throughput
- 10 GbE wire speed interfaces
- Coloured (C-band) interfaces with WAN PHY to G.709 conversion
- Carrier-class availability
- Full MPLS support

Nice to have features:

- Secondary router core for building test and research networks

8.2.2   Provider edge router specification

The edge router is really the heart and soul of the network. It routes all traffic between sites on the same ring but different VLANs; traffic between sites on different rings in the same city and also traffic to the core router. In MPLS terms it appends the MPLS labels to the incoming packets, establishes label switched paths over the network and makes all routing decisions based on cost as topology and congestion changes take place.

Required features:

- 10 GbE wire speed interfaces
- Support for other standard interfaces such as STM-64
- QinQ support (see 11.1.6)
- Very fast routing backplane
- Carrier-class availability
- Full MPLS support
- Detachable laser sources (SFP type)

Nice to have features:

- Coloured (C-band) interfaces with WAN PHY to G.709 conversion

8.2.3   Switch specification

The grey rings serving the smaller sites that do not justify DWDM equipment will be on simple Layer2

switches. These switches, however, would need to satisfy some stringent requirements:

Required features:

- 10 Gbps line and switch fabric
- Dual, redundant line inputs for ring operation
- Support for RSTP and QinQ
- At least 12x 1 Gbps tributary ports
- All tributary ports must support either copper of fibre interfaces
- Carrier-class availability
- Fully managed

Nice to have features:

- Support for FastEthernet tributary inputs (100 Mbps)
- Automatic speed negotiation

8.2.4   WDM Node specification

The WDM layer underpins the MPLS network and expands the capacity of the network tenfold. In keeping with the design philosophy of providing up to 10 Gbps per site, the WDM layer is crucial to ensure expandability of each large node.

Required features:

- Large channel capacity (30 or more) at 10 Gbps
- G.709 support with FEC
- ROADM[1] configuration
- Fully redundant ring support
- Full network and element management
- 10 GbE LAN/WAN PHY support
- Coloured tributary inputs

Nice to have features:

- Layer2 switching functionality

8.3   Equipment Decisions

8.3.1   Core routers

Each of the four major cities (Johannesburg, Pretoria, Cape Town and Durban) would have two core SANReN nodes, one primary and one secondary. The primary core node would have a large aggregation router as well as a core router for routing inter-city traffic. In a carrier network the aggregation routers would be called Provider Edge Routers (PERs), as they would be placed on the edge of the core network. Only Access routers and switches would be connected on the Access side of

---

[1] Remotely reconfigurable Optical Add/Drop Multiplexer

these routers. On the network side the core router would be called the PR[1].The national network could then be represented by the diagram in Figure 23.
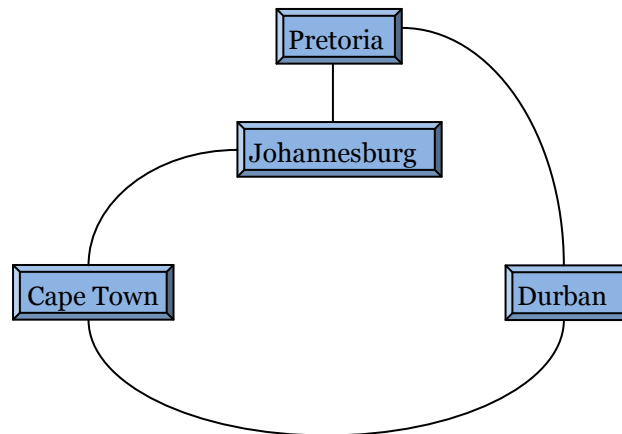
Figure 23: The Core router network

The local network in each city may presented by the diagram in Figure 24.

All local rings would be sub-tended from the PER so that any local site may have traffic routed to any other site on the same local network. Each wavelength on the DWDM network would be directly connected to the PER via the DWDM ADM[2]s. The local network would in effect consist of several 10 Gbps rings, one ring for each lambda and one for each grey ring. The sole point of interconnection would be the PER. In MPLS terms the PER would fulfil the function of the *label edge router*.
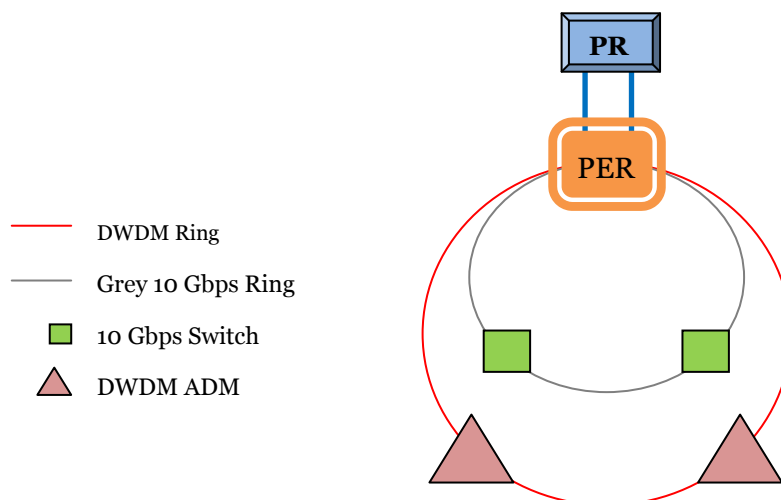
Figure 24: The Local network

The core router (*label switch router*) routes traffic between the different networks on the core layer, in this case the four major cities. Thus traffic between different rings in the same city (or different nodes on the same ring), would be routed only by the PER, whereas traffic between nodes in different cities would be routed by the PR, after passing through the respective local PERs.

---

[1] Provider Router
[2] Add Drop Multiplexer

SANReN chose the Cisco 7609s as the provider edge router. It has 7 line card slots, each of which has a traffic capacity of 40 Gbps, with a total backplane throughput of 720 Gbps. In addition it supports a wide range of interface types including 10 Gbps Ethernet and SDH up to STM-64. It supports full MPLS functionality and QinQ, and has hot-swappable 10 Gbps grey lasers. It does not yet support coloured line interfaces, but these are on the roadmap for the next year.

The Cisco CRS-1[1] was chosen as the PE core router. In the 16-slot configuration deployed by SANReN it has a switch fabric throughput capacity of 1.2 Tbps. It supports integrated IPoDWDM[2] with coloured lasers, fully tuneable throughout the C-band, enabling integration with a DWDM platform without the use of transponders. In addition is has an extra router core called a SDR[3], which may be used in conjunction with the production routing fabric for testing and research. It is completely separate and no problem encountered during operation would affect the production fabric.

8.3.2 Switches

It was decided early on that all of the grey rings would be built using 10 Gbps Ethernet switches. This would ease the interfacing of client LANs, since all the institutions' LANs use Fast Ethernet or Gigabit Ethernet. The 10 Gbps backbone could then be partitioned into bandwidth chunks, depending on the speed required by the site. Larger sites would typically consume an entire 1 Gbps of bandwidth or even more, whereas smaller sites would be assigned a 100 Mbps line.

SANReN chose to use the Cisco 4924 switch, which offers up to 24x 1 Gbps ports, either optical or electrical (copper). In addition to being carrier class (managed) and fully non-blocking, it also supports QinQ and RSTP. It has 2x 10 Gbps line interfaces which support hot-swappable laser interfaces, either at 1310nm (short range) or 1550nm (long range). It does not support any interfaces at 100 Mbps or anything slower than 1 Gbps, but the design team decided that this was acceptable. It would simply state in the AUP[4] that any client site would be obliged to connect at 1 Gbps, as Ethernet equipment at this speed is quite inexpensive.

8.3.3 WDM nodes

In a market where dark fibre is still very hard to come by and consequently very expensive, WDM is really the only sensible option when one is trying to build a very high capacity, expandable network.

The Cisco 15454 platform offers up to 40 channels at 10Gbps each with full remote reconfigurable optical add-drop multiplexing functionality. All channels support standard G.709 encoding with enhanced Reed-Solomon FEC. In addition, the platform features 10 Gbps Ethernet muxponders with built-in Layer2 switching functionality. It will also accept coloured tributary interfaces and is fully managed. Lastly it supports WAN/LAN PHY and STM-64 interfaces.

The SANReN design team decided to deploy full 32-channel ROADMs at each of the large client sites. Using the built-in Layer2 functionality of the muxponder card, no external Layer2 switch would be

---

[1] Carrier Routing System
[2] Internet Protocol over Dense Wavelength Division Multiplexing
[3] Secure Domain Router
[4] Acceptable Use Policy

needed at these sites. All the client lines would simply be interfaced with 10GbE. Since the muxponders do not operate at any other speed than 10 Gbps, the onus is on the site to ensure it can interface at this speed. Client sites will get a single-shelf, two-degree ONS[1] node, whereas core nodes that have three degrees or more will get full WSX[2] nodes. To save on card space, each degree will get a full shelf.

8.4   Network Decisions

8.4.1   Equipment redundancy

On the grey rings, each site switch has to have redundant optics. It can be shown that the component most likely to fail[3] in an optical system such as a switch or ADM is the optical transceiver. For this reason many manufacturers now make optical cards such as transponders with pluggable optical transceivers. In the case of a transceiver failure only the transceiver needs to be replaced, not the entire card. The grey ring switches would have redundant power supplies as well, but it was decided not to attempt a duplication of the switch fabrics, as this additional redundancy did not justify the added cost. If the entire switch failed, it would have to be replaced with the standard MTTR[4] in effect, in accordance with the SLA[5] between SANReN and the equipment provider.

The larger sites that were assigned a DWDM ADM would benefit from several layers of redundancy. In the first instance they would have redundant transponders as well as redundant optics. Each side of the ring (East and West) would therefore have a transponder/optic pair feeding into the client's equipment. This would be sufficient to survive any single equipment or fibre failure, short of a catastrophic router failure (Figure 25).
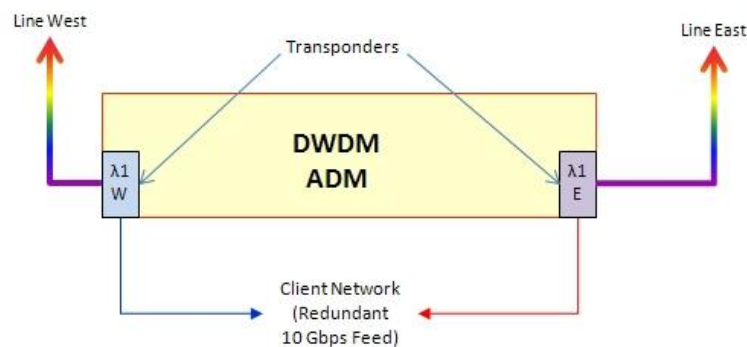
Figure 25: DWDM client site

8.4.2   Router failover

So much reliance is placed on the PER that it becomes the single point of failure in the local network. If the primary core PER fails, the entire local network would be down, as it is the sole point of

---

[1] Optical Network Switch
[2] Wavelength Selective cross-connect
[3] The likelihood of any component failing is expressed in its published MTBF figure (Mean Time Between Failures). The shorter the MTBF time (expressed in years), the more likely it is to fail.
[4] Mean Time To Repair
[5] Service Level Agreement

interconnect between the various rings. It is true that sites connected to the same grey ring would still be interconnected, however, each DWDM site would be completely cut off and the local network as a whole would be isolated from the national and international network.

The solution to this single point of failure is to have a second PER, which plays the role of backup to the primary PER. In the event of primary PER failure, the backup would step in to keep the network running. The only problem would then be to re-establish connections between the backup PER and the Core router, which provides connectivity to the national and international network.

It would be wise to establish the backup router in a physically separate location to the primary one, so it makes eminent sense to locate it at a major client site with a DWDM node. This also provides the opportunity of using the DWDM network to backhaul the traffic on the backup PER to the Core router. Figure 26 illustrates how to establish this second level of redundancy.
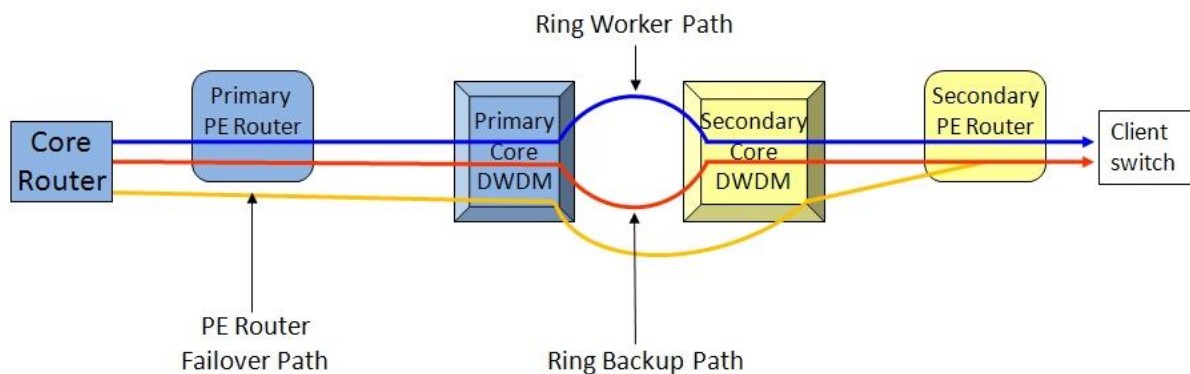
Figure 26: PER failover

Both traffic paths (worker and backup) run through a PER at either end of the ring. A third path has to be commissioned that runs through the secondary PER and the entire WDM system, but not through the primary PER. Instead it is directly connected between the primary core WDM node and the core router. This path will only be used if the primary PER fails and will be implemented with a Layer 3 protection scheme, such as *MPLS fast reroute*. Both the ring backup path as well as the router failover path will be inactive and will only be activated by the MPLS control plane in case of a failure. On the secondary core site the failover path will require a port on both the PER and WDM node; however at the primary core it would require only a WDM port and a core router port.

All other rings will also be fed through the secondary PER and have backup routes configured in this way. All WDM lambdas will have to add/dropped into the secondary router and have an additional failover path configured. The grey rings will also have to be "dual-homed" around both PERs.

It is very important to establish firm ground rules when embarking on the design of a large network, and just as important that the principles be adhered to. Once compromises are made on the principles, it may be very hard indeed to design the network in such a way that it still satisfies the design criteria and user needs.

The next chapter will discuss the design of the national backbone network.

Chapter 9

Designing the SANReN: National Network

Most of the links required on the national SANReN backbone will be in the form of managed wavelengths between the main Cities, but there are some notable exceptions to the rule.

The SANReN network consists of four main sub networks (one in each of the cities of Johannesburg, Pretoria, Cape Town and Durban); several smaller sub networks in other cities, and a national backbone interconnecting all the sub networks. Figure 27 illustrates the network at its highest level.
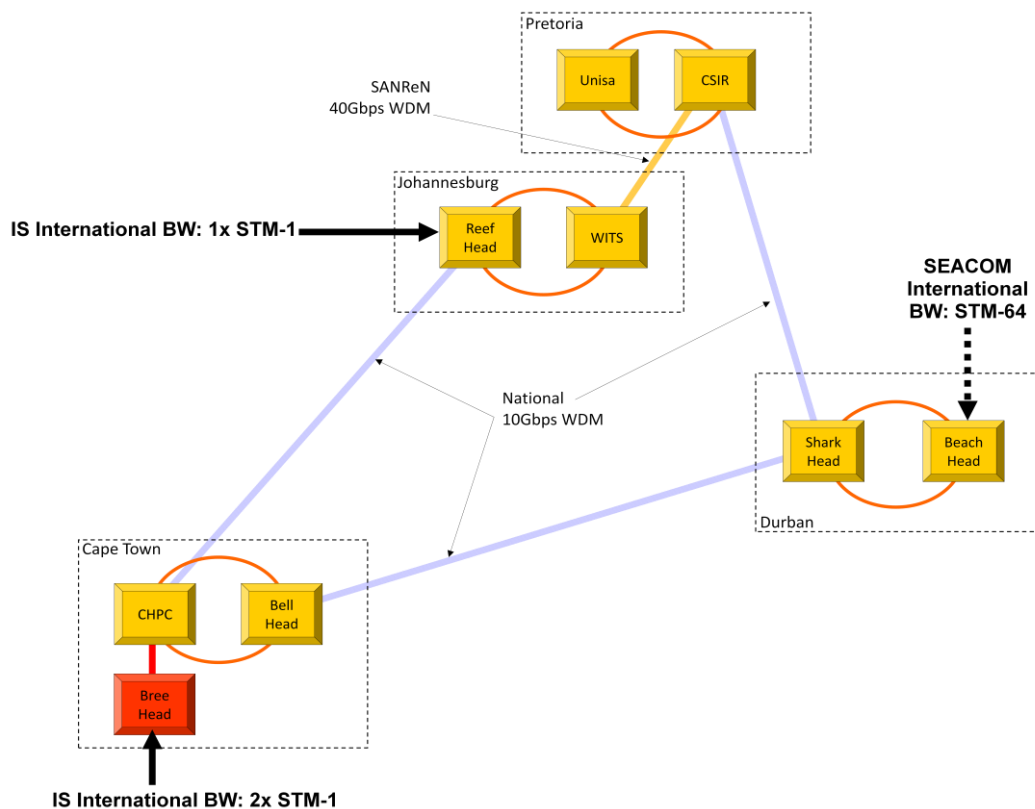
Figure 27: The SANReN National Network

9.1 Durban - Pretoria

9.1.1 Mtunzini on the backbone

The Neotel national backbone is built on top of the Infraco national network, and thus there are many places where the two networks are, for all practical purposes, indistinguishable.

One such place is the Athena carrier room to the north of Mtunzini, near Empangeni. If the national network link between Pretoria and KwaZulu-Natal could be diverted to pass through Athena and then Mtunzini, then the Beachhead node at Mtunzini could become a node directly on the national backbone, rather than on a separate ring out of Durban (Figure 28). This, however, presupposes that Infraco would be willing to create a new node at Mtunzini.
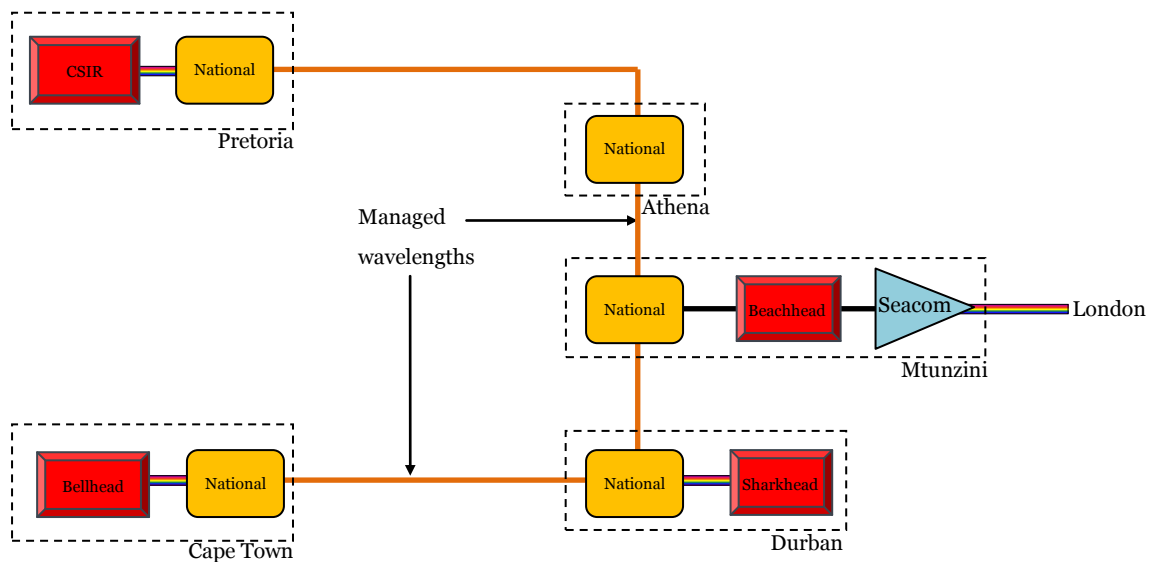
Figure 28: Managed wavelengths via Beachhead

There is much to recommend this architecture, as it places the access point of the international bandwidth on the national ring, which means that the traffic no longer has to be backhauled to Durban before it can be distributed on the national network. Considering the amount of new traffic suddenly available in Mtunzini, Infraco would quite likely be interested in establishing a new node there.

The main function of the national backbone to KwaZulu-Natal is not to interconnect the institutions, but rather to provide a backhaul/distribution path for the international traffic. Consider the default network design again (Figure 27). The 10 Gbps international can only be distributed into the network at four places: Durban, Johannesburg, Pretoria and Cape Town. From studying the traffic patterns on the GEN3 TENET network, one can deduce that the international bandwidth demands may be subdivided roughly as follows:

- Durban: 10%
- Pretoria: 30%
- Johannesburg: 20%
- Cape Town: 30%
- Other sites (nationally): 10%

Geographically speaking then, only 10% of the bandwidth is dropped in Durban, 35% is transported to sites south of Durban on the backbone and 55% is transported to sites north of Durban. If another 10 Gbps circuit became available, one could expect the ratios to remain the same even when the amount of bandwidth to each city doubled.

Consider also, that the amount of non-international traffic between Gauteng and Durban, as well as between Durban and sites south of Durban is minimal (in the region of 30Mbps), so less than 1% of Seacom).
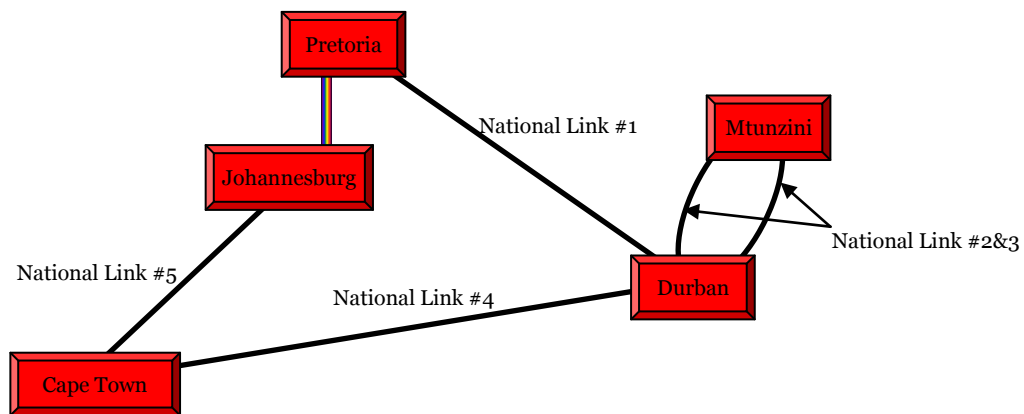
Figure 29: Default national backbone

So, if the default network design is followed where Mtunzini is connected to Durban on a redundant link (Figure 29), a second circuit on Seacom would require a *redundant* 20 Gbps of traffic between Mtunzini and (Durban), which would push both link 2 & 3 up to 20 Gbps.
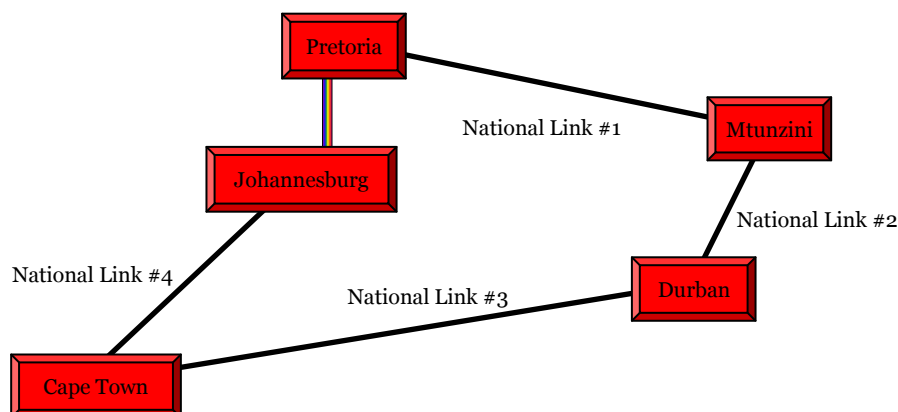
Figure 30: National backbone via Mtunzini

However, if Mtunzini were on the national backbone (Figure 30), then only a single 10 Gbps circuit would be required to the north (Gauteng), as well as one to the south (Durban), which would save SANReN upgrading the link at all.

With such a configuration the two links out of Mtunzini would carry 20 Gbps of traffic International traffic (10 Gbps each). Using the same ratio as above for traffic dropped in Durban (10% of 20 Gbps = 2 Gbps) Link #1 would carry a full 10 Gbps, as would Link#2. Link #3 to Cape Town would then carry 7 Gbps of international traffic, leaving more than enough capacity for national traffic to and from Durban.

In such a case all the national traffic into Durban would come from the south, but if a break ever occurred, Links #1&2 could be throttled back to cater for the national traffic, which could then be rerouted via Mtunzini.

Ideally, one would upgrade the capacity of the national network at the same time as upgrading the international link, but with this topology an international upgrade would not *force* a national upgrade to take place concurrently.

9.1.2   Managed wavelengths vs. dark fibre

When one considers such a topology change to the network, other options present themselves. For example, SANReN can consider the possibility of purchasing a dark fibre link between Mtunzini and Durban and lighting it with a DWDM system. This would be in line with the future upgrade paths, since the International bandwidth may then easily be upgraded to a second 10 Gbps circuit without a simultaneous upgrade in the national backbone capacity.

Consider once more the default design (Figure 29), which with dark fibre would call for redundant links on the route between Mtunzini and Durban: a total of 320 km of dark fibre. If the traffic was routed through Athena instead, and lit on a dark fibre on the route Athena-Mtunzini-Durban, only 210km of dark fibre would be required.

Moreover, in comparing managed wavelengths to dark fibre one should keep in mind that wavelengths are usually not priced according to distance, but rather according to the number of transponder hops. For example, if the provider has a system of optical amplifiers, the entire distance between Pretoria and Durban would count as a single transponder hop. There will be an additional charge for each amplifier hop of course, but even so a single transponder hop with several amplifier hops would be less expensive than two short transponder hops (see 5.5.2.3). Knowing this, a systems designer would try and avoid buying short-distance managed links. After all, if the price for a 500 km hop is more or less the same as for a 50 km hop, it would make no sense to purchase any short hops. Rather, the designer would always attempt to design a DWDM dark fibre link for any distance that may be built without amplifier stations. This could be as much as 120km with a standard DWDM system, and as much as 200km with Raman amplification (see 5.5.2.4).

Consequently, it would be far more attractive to consider commissioning a DWDM link on dark fibre between Athena, Mtunzini and Durban than to purchase a managed wavelength which could be put to better use on a much longer route.

If Infraco is the service provider, it would make eminent sense to extend the dark fibre link all the way up to Athena (Figure 31), saving Infraco the trouble and expense of putting a DWDM node at Mtunzini landing station and picking up the SANReN link there. Such a cost saving would be reflected in the reduced cost of a new managed service from Athena to Pretoria, rather than from Mtunzini to Pretoria.
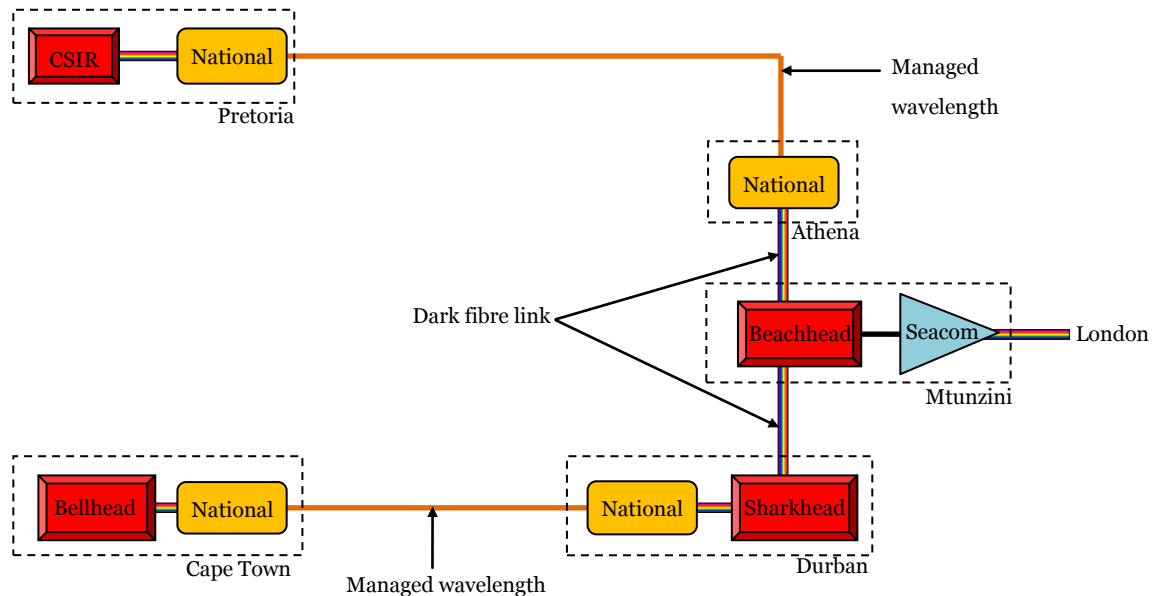
Figure 31: Dark fibre link Athena-Durban

9.1.3 Athena - Pretoria

SANReN has been unable to negotiate a satisfactory arrangement with either of the two potential long-distance network operators (Neotel and Infraco) for the purchase of 10 Gbps wavelength service between Durban and Pretoria, and has therefore begun to consider various alternatives.

DFA[1] extended SANReN a reduced price for a long distance dark fibre connection on their new Gauteng to Durban route. There is a question about diversity with the Pretoria-Johannesburg fibre link, but assuming another route may be found in DFA's Gauteng network, it creates the possibility of SANReN constructing a 700km+ DWDM link between the Pretoria CSIR core node and the Sharkhead in Durban, via the Beachhead (Figure 32).

Clearly, no single-hop DWDM link may be established over such a distance outside a laboratory environment, which means that intermediate amplification/regeneration is required. If SANReN saw no possibility of ever needing more than a single 10 Gbps link, a STM-64 SDH link might have been considered, which would only need a regenerator every 100 km or so. However, since even the initial national requirement of 10 Gbps would soon be insufficient, seen in the light of the imminent 10 Gbps

---

[1] Dark Fibre Africa

international Seacom link, SANReN must assume that it will soon need several wavelengths on this route.
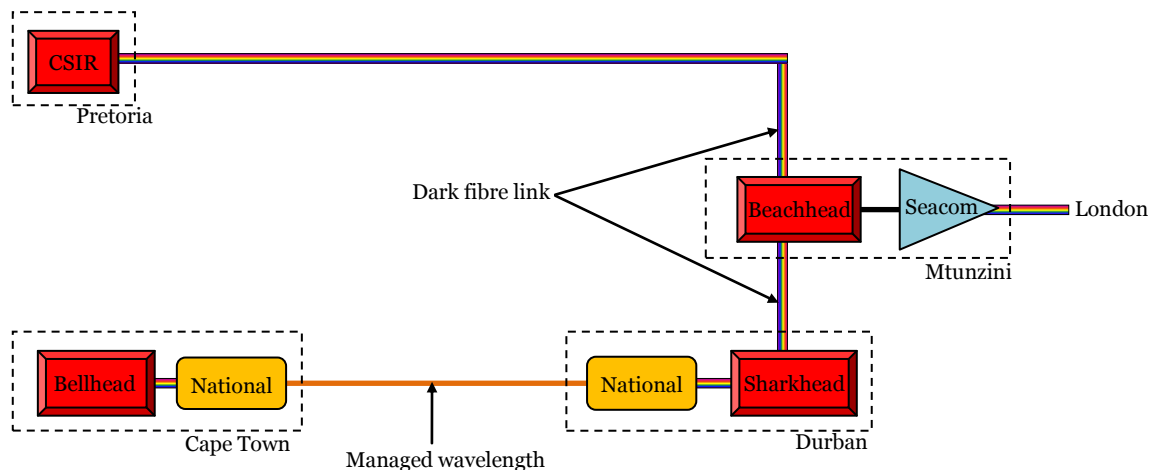
Figure 32: Dark fibre link Durban-Pretoria

However, if a dark fibre DWDM solution is considered, provision must be made for intermediate amplifiers along the fibre route. DFA will have to provide facilities for amplifiers for the use of their other carrier clients, which means placing an "amplifier hotel" with power, air-conditioning and security every 100km or so, according to standard DWDM design principles. This will satisfy the needs of most clients.

SANReN now has the option to either use a standard PA-BA[1] amplifier arrangement to span the 100 km fibre hops, or else to use extended-reach Raman amplification to stretch the amplifier spacing to 200 km. Careful consideration should be given to the decision, as the 20 year operational cost of operating more, cheaper amplifiers may outweigh the upfront capital expense of buying fewer, more expensive amplifiers. One must consider that the expected lifetime of an active optical component such as an amplifier card is only five years, so the designers must expect to replace all the active components at each site at least *three times* during the lifetime of the IRU.

The main motivation of considering the greater upfront cost of a dark fibre IRU instead of one or two managed wavelengths is in the long-term capacity of the network. While buying a fibre IRU and all the equipment to run a DWDM link over it may seem prohibitively expensive, it must be considered that such a system may be upgraded by 10 Gbps at the cost of two cards. The amplifiers are not dependent on the number of channels employed, which means that in the 20-year lifespan of the network the link may grow to as much as 100 Gbps or more for very little added expense. Compared to the incremental cost of 100%+ for adding another wavelength in a managed system, the dark fibre approach may cost in with only a few wavelengths.

---

[1] Pre-Amp, Booster Amp

9.1.4   Other sites along the route: Witbank

Witbank is on the main fibre route between Kwazulu-Natal (Mtunzini) and Pretoria. Both the dark fibre route of DFA and the national backbone network of Infraco pass through (and are terminated) in the Witbank CBD, which is also home to an important satellite campus of TUT[1]. TENET can construct a 4.5 km fibre cable either by itself or by co-build with DFA or another provider (Figure 33).
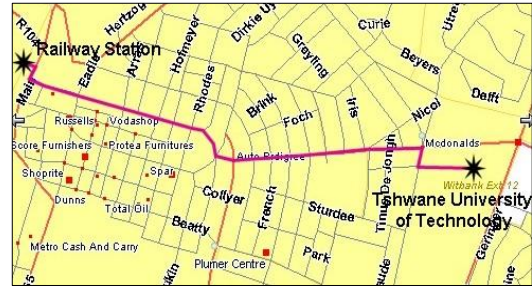
Figure 33: Witbank access

9.2   Pretoria – Johannesburg

At present the only SANReN link between Pretoria and Johannesburg consists of two 10Gbps links running from the Pretoria CSIR node to SAC[2] at Hartebeeshoek and then from SAC to Wits Main in Johannesburg (Figure 34).
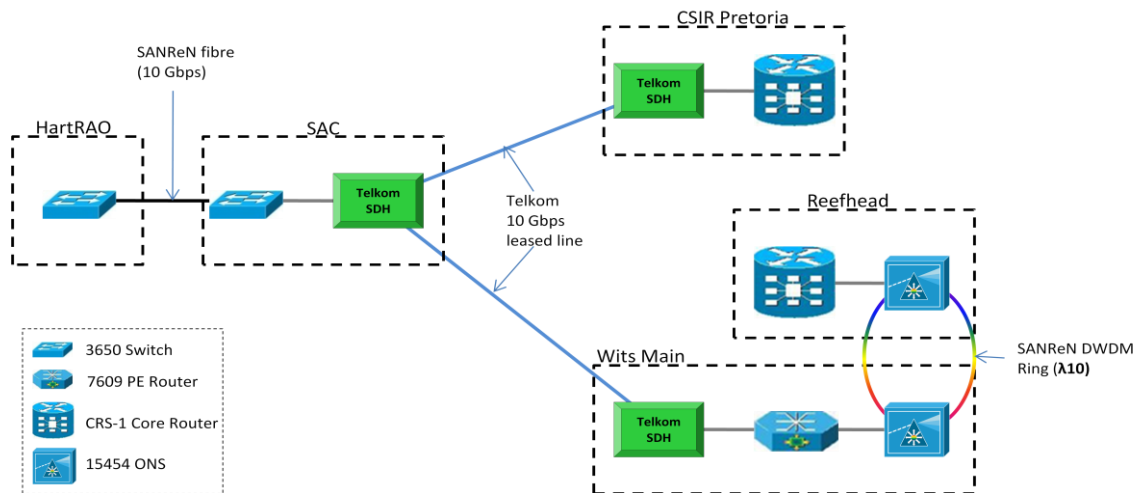
Figure 34: Telkom Connections to SAC

SANReN's top priorities were the two installations at Hartebeeshoek. The SAC is a project of the CSIR which provides tracking, telemetry and command (TT&C) services for geo-synchronous and polar orbiting spacecraft. HartRAO[3] is a project of the NRF[4] and is the only major radio astronomy observatory in Africa. Both installations need very fast connections to the Internet, more specifically to other research networks (NRENs) in Europe and elsewhere. HartRAO is involved in several multi-national VLBI[5] projects and needs a high bandwidth connection (in excess of 1 Gbps) for these projects alone.

---

[1] Tshwane University of Technology
[2] Satellite Applications Center
[3] Hartebeeshoek Radio Astronomy Observatory
[4] National Research Foundation
[5] Very Long Baseline Interferometry

SANReN was unable to interest any of the providers in building a new fibre route out to the sites because they all said it would be prohibitively expensive. Only Telkom had any infrastructure in the area and they refused to relinquish a fibre pair to SANReN under any circumstances. The only service they were prepared to offer was a 10Gbps leased line between a Johannesburg or Pretoria location and SAC. Since SANReN was struggling to establish any connection at all between Pretoria and Johannesburg, it was decided to purchase a 10-year IRU at 10Gbps for a connection between Wits Main campus and SAC, as well as one between the CSIR in Pretoria and SAC. This established a single 10 Gbps connection between CSIR Pretoria and Wits Main (Figure 34).

HartRAO is about 2 km from SAC so SANReN commissioned a new fibre cable between the two sites. Both installations would now have access to the redundant 10Gbps connection to the rest of the SANReN network.

As soon as SANReN can close a deal with a dark fibre provider it will purchase a link between the Reefhead and the CSIR core node in Pretoria and proceed to light the fibre with a DWDM system at multi-10 Gbps speed. The initial deployment plan calls for 40 Gbps of bandwidth between Johannesburg and Pretoria.

9.3   Johannesburg – Cape Town

The long-term plan for this route is to pass it through several cities on the way down from Johannesburg. Initially it will be a single 10 Gbps managed wavelength, to be provided by one of the national operators on an IRU basis. Expansion to 20 Gbps will only occur in some year's time, when the demand justifies the expense.

9.3.1   Vanderbijlpark

Just to the south of Johannesburg, the town of Vanderbijlpark hosts two large TENET campuses: North West University[1] and Vaal University of Technology[2]. Both these sites may be reached by a small access network if SANReN establishes a node in nearby Vereeniging. If the national link is routed through the Vereeniging station, then the two SANReN sites may easily be connected to the backbone with the use of a short, dedicated fibre link (Figure 35).
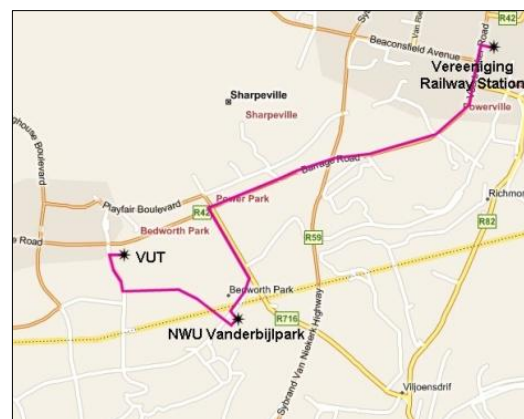
Figure 35: Vanderbijlpark access

[1] NWU
[2] VUT

9.3.2 Bloemfontein

Two large universities (University of the Free State & Central University of Technology) have their main campuses in Bloemfontein. SANReN will construct an access network between the campuses and the national node at the Bloemfontein station to connect these sites to the national network (Figure 36).
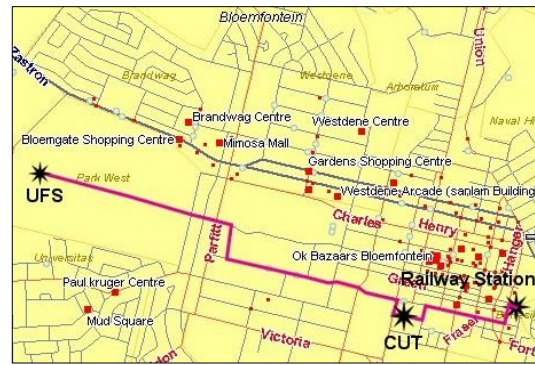
Figure 36: Bloemfontein access

9.4 Cape Town – Durban

Several large cities which are on this route should also be included in the SANReN network. This connection will consist of a 10 Gbps managed wavelength, as with the Johannesburg - Cape Town link.

9.4.1 Port Elizabeth

The Nelson Mandela Metropolitan University[1] is situated on the southern edge of the city and has four campuses of which three are contiguous, which means they may be easily connected by a small fibre network (Figure 37). PE also hosts a small branch of the CSIR, which is very close to NMMU and may be connected to the main site by a short fibre link. These campuses then only need to be connected to the national carrier's node, which is in downtown Port Elizabeth, at the North station.

Figure 37: Port Elizabeth access

9.4.2 Grahamstown

Rhodes University[2] has its main campus in Grahamstown, which is on the N2 national road between PE and East London. A single point-to-point fibre connection could easily be constructed to connect the campus to the national provider node on the outskirts of the town (Figure 38). A node on the national backbone would then connect it onto the 10 Gbps national network.
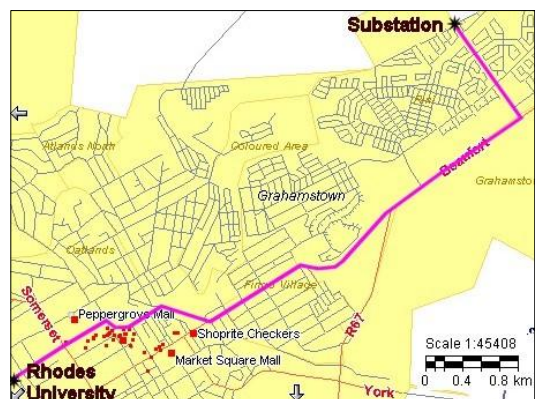
Figure 38: Grahamstown access

---

[1] NMMU
[2] RU

9.4.3 King William's Town

The Bisho campus of the University of Fort Hare is only 6.5 km from the King William's Town station (Figure 39). Therefore, if the SANReN national backbone is provided by Infraco and the SANReN 10 Gbps circuit is regenerated at King William's Town, TENET may request that the circuit be dropped into a local router at the node, since Infraco will have to provide back-to-back transponders for the circuit regeneration anyway (see 9.6).
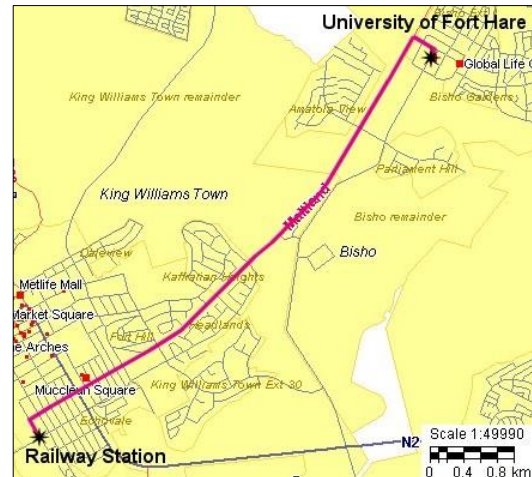
Figure 39: King William's Town access

9.4.4 East London

Both the University of Fort Hare[1] and Walter Sisulu University[2] have satellite campuses in East London. UFH has a single campus but WSU has four separate campuses, which has prompted SANReN to consider designing a full access network in this city. This may take the form of managed Metro Ethernet circuits, but would ideally be a dark fibre ring running through the city, connecting all the sites in a grey 10 Gbps shared ring. Initial deployment calls for a simple fibre spur through the CBD, connecting the *UFH* campus, *WSU Chiselhurst* and *WSU College St.* to the backbone node at the station (Figure 40).
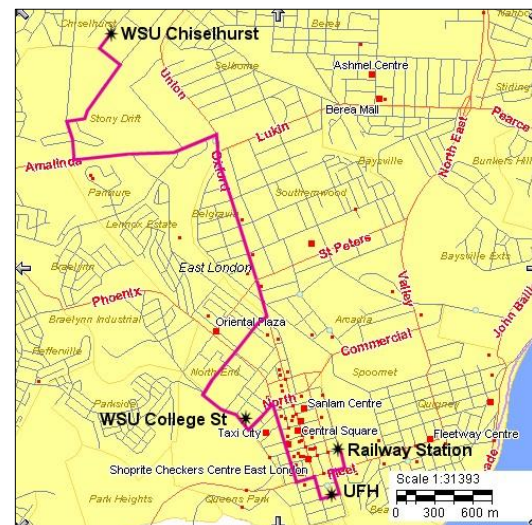
Figure 40: East London access

9.4.5 Mthatha

The main campus of WSU is located just outside Mthatha, and may be the site of a node on the national network. Failing that, a dark fibre may be commissioned from the national provider's node in the centre of the city to connect WSU to the national backbone.
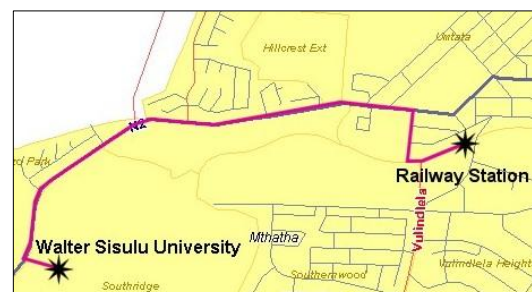
Figure 41: Mthatha access

---

[1] UFH
[2] WSU

79

9.5   Other National Routes

9.5.1   Polokwane

The main campus of the University of Limpopo[1] is in Sovenga, just to the east of the city of Polokwane. A long-distance leased line between this site and the CSIR node in Pretoria would connect it to the national backbone. If the national provider cannot deliver the circuit to the campus, SANReN may have to commission a fibre connection between Sovenga and Polokwane and light it. A connection can then be established between the SANReN node in Polokwane and the national provider node.

9.5.2   Potchefstroom

NWU's main campus is in the town of Potchefstroom. It will also be connected to the local provider's node on SANReN dark fibre, for transporting at 1-2 Gbps to the nearest SANReN national node, most likely in Vereeniging.

9.6   Connecting Nodes to the National Backbone

In many of the small cities and towns along the route followed by the Infraco backbone the SANReN circuit will be regenerated. At such a Point-of-Presence (PoP) the regeneration of the 10 Gbps optical signal is accomplished by using back-to-back transponders on the OADM[2], which are then simply connected by a short patch cable. If TENET has a site in that town/city the circuit may be dropped into a local TENET router. From the router a dark fibre connection may be established to the site, terminating on a small Ethernet switch, typically at 10 Gigabit speed (Figure 42).

TENET will therefore have to provide the following in order to connect the local campus to the national backbone:

- 1x 10 Gbps router (typically a Cisco 7609s)
- 1x 10 Gbps line card (Cisco 6708)
- 3x 10 GbE router ports (XFP interfaces)
- 1x 10 Gbps switch (Cisco 4924)
- 1x 10 GbE switch port (XFP)
- 1x 1 Gbps switch port (copper SFP)
- Dark fibre connection between the site and the national PoP

---

[1] UL
[2] Optical Add-Drop Multiplexer
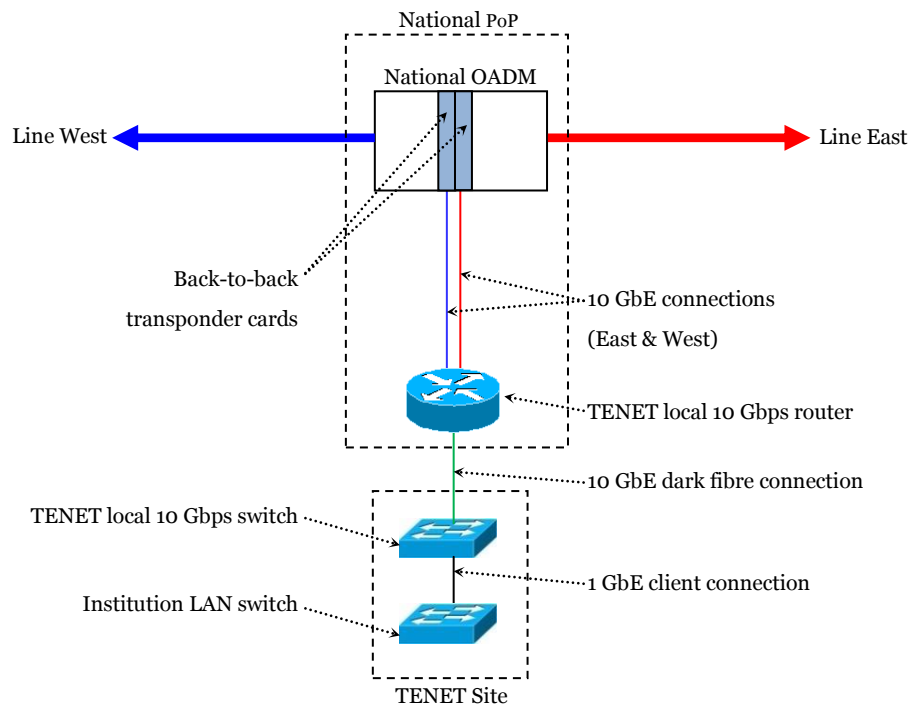
Figure 42: Connecting a node to the backbone

9.7 The Complete National Network

 Figure 43 illustrates what the entire national backbone network looks like, with all the secondary sites included.
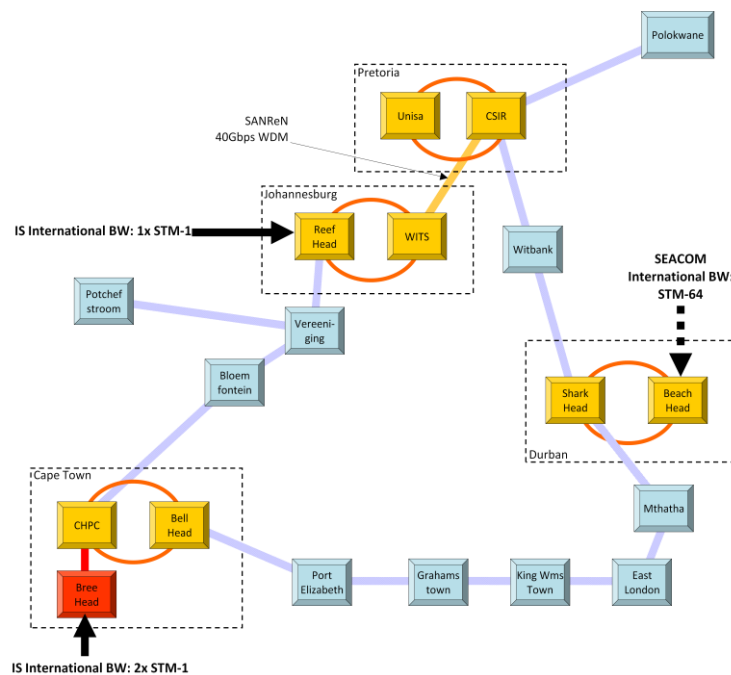
Figure 43: Complete national network

Chapter 10

Designing the SANReN: Metro Network

Each of the four major cities' networks consists of the same basic elements: a primary- and secondary core node, a DWDM ring containing several major client sites and one or more grey rings containing the smaller sites.

10.1 The Johannesburg Network

The three most important institutions in Johannesburg are Wits[1], UJ[2] and CSIR. It was quickly established that a single fibre ring could connect all of the most important sites of these institutions to a core site.

10.1.1 Participating institutions

10.1.1.1 University of the Witwatersrand

Wits has four main campuses, which are located more or less in a straight line. Working from the outside towards the main campus, Wits Education is across the road from Wits Medical, the Business School campus is just a few blocks away and from there to the main campus is just a few kilometres (Figure 44).

Wits had constructed its own WAN to interconnect its campuses, but the links were all too small and some were unstable wireless links. Only the main campus had an existing TENET GEN2 connection, so it was decided to provide a fibre connection to all of the campuses.

The main campus, which has by far the largest bandwidth demand, would be fed by a DWDM node, which will allow for future expansion as the students and researchers start using more bandwidth-intensive applications. The other three sites would all be supplied with service on a grey 10 Gbps ring.

---

[1] University of the Witwatersrand
[2] University of Johannesburg
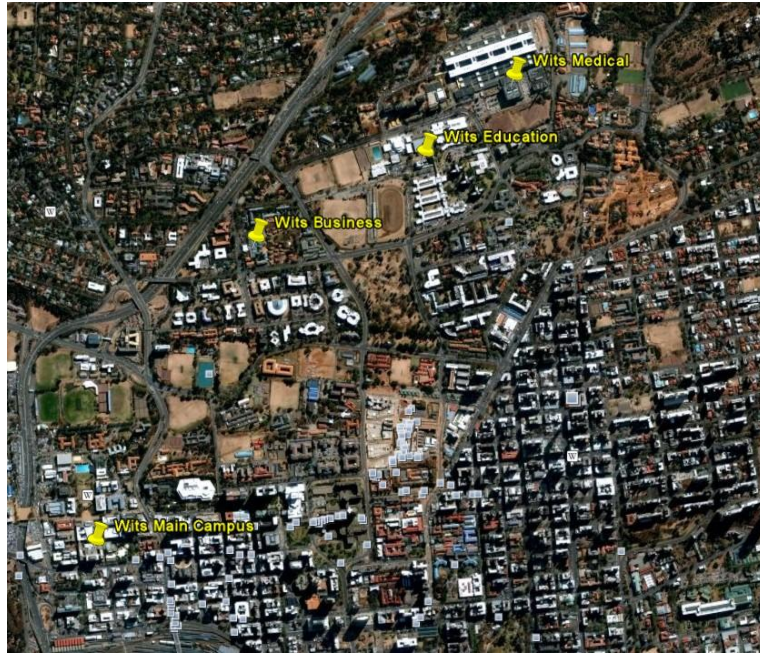
Figure 44: University of the Witwatersrand campuses

10.1.1.2   *University of Johannesburg*

UJ has two large campuses (Kingsway and Bunting Rd), as well as a smaller campus called Doornfontein. The smaller campus is quite a distance from the two larger ones, however they are situated on either side of the Wits campuses (Figure 45).

Figure 45: Johannesburg sites

UJ Kingsway is the main campus and is also responsible for the bulk of the bandwidth demand; therefore it is a logical DWDM site. At first, UJ Bunting Rd was also earmarked as a DWDM site, but

after consultation with the institution it was decided to put both of the smaller sites on the same grey ring as the small Wits campuses.

All three sites were existing TENET GEN2 sites, so during the cutover the traffic would be migrated onto the fibre ring.

10.1.1.3   Centre for Scientific and Industrial Research

The CSIR has five small campuses in the city, but only two are in the general vicinity of the university campuses. Since the CSIR is not only a primary TENET client but also the owner of the SANReN network, it was decided to include both of these sites onto the ring, even though neither have a large demand for bandwidth.

The CHPC[1] is also a project of the CSIR Meraka Institute, and the management at the CHPC main site in Cape Town had mentioned their long term plans of establishing a supercomputer node in Johannesburg. After being told of the network design in Johannesburg, they expressed great interest in the two small CSIR sites and mentioned that access to this ring would make a strong case for situating the new supercomputer node at one of these two sites.

10.1.2   The Johannesburg Ring

SANReN purchased a 20-year IRU on the Neotel fibre network for a ring of three fibre pairs connecting the SANReN sites in a ring (Figure 46). This ring would be maintained by Neotel for the duration of the IRU contract. Each site would be connected to the ring on a Neotel patch panel equipped with LC-APC connectors.

10.1.2.1   The DWDM ring

Each of the primary customer sites (UJ Main and Wits Main) is equipped with a DWDM ROADM node to supply a dedicated 10 Gbps connection to each site. The bandwidth is easily upgradeable to multiple wavelengths per site should the need arise.

Further wavelengths are commissioned on the ring to supply the other connections required for PER failover etc.

10.1.2.2   The grey ring

In addition to the DWDM ring that feeds the primary customer sites, a grey 10 Gbps ring is required to connect the smaller sites to the Reefhead redundantly. Each site is connected on fibre pair 2 and equipped with a Cisco 4924 10 Gbps switch. Initially each site will be given a dedicated 1 Gbps connection to the ring, with further spare capacity available should a site require more bandwidth.

---

[1] Centre for High Performance Computing
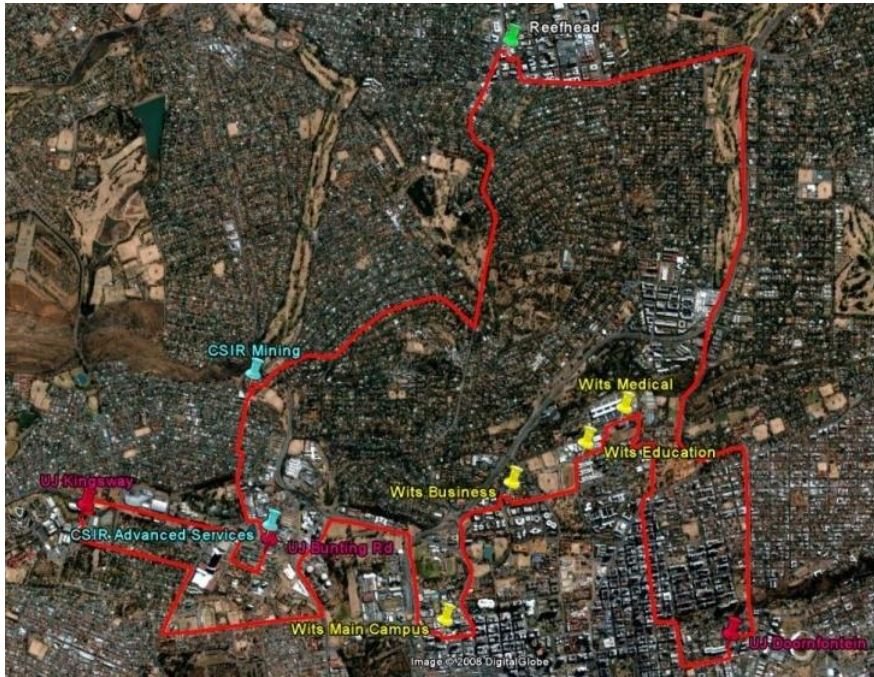
Figure 46: The Johannesburg Ring (physical view)

10.1.3   Other connections

10.1.3.1   The Pretoria connection

One dilemma faced by the design team was concerning the connection between the Johannesburg network and the one in Pretoria. It was decided that the connection definitely had to be a dark fibre, which could then carry a multi-10 Gbps DWDM system. The problem was that none of the providers approached for this connection were prepared to commit to a solution, and therefore the network design team could not finalise the fibre network design.

Two possible routes were considered. If an agreement could be reached with the National Roads Agency[1], the dark fibre pair would follow the N1 south from the Pretoria SANReN core node at the CSIR, up to the Buccleuch interchange (pink line, Figure 47). From there it would follow the N3 until it reached the N17, after which it would follow the N17 until it reached the M1 in the centre of Johannesburg (red line).

The implication of this would be that the connection to Pretoria could not be at the Reefhead, as the SANRAL network does not extend that far north of the N12, but rather at the main campus of Wits, which is located quite close to the M1. The impact that it will have on the network design is that the DWDM node at Wits Main would have to become a three-degree node (see Figure 48). The node would therefore have to become a WSX site, rather than a simple two-degree, one shelf ONS.

---

[1] SANRAL
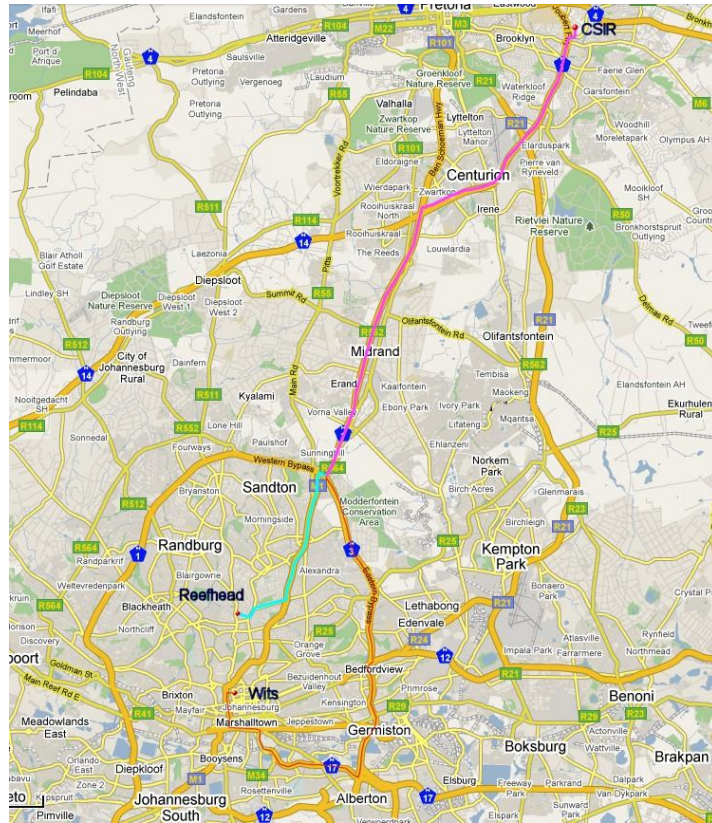
Figure 47: The dark fibre connection to Pretoria
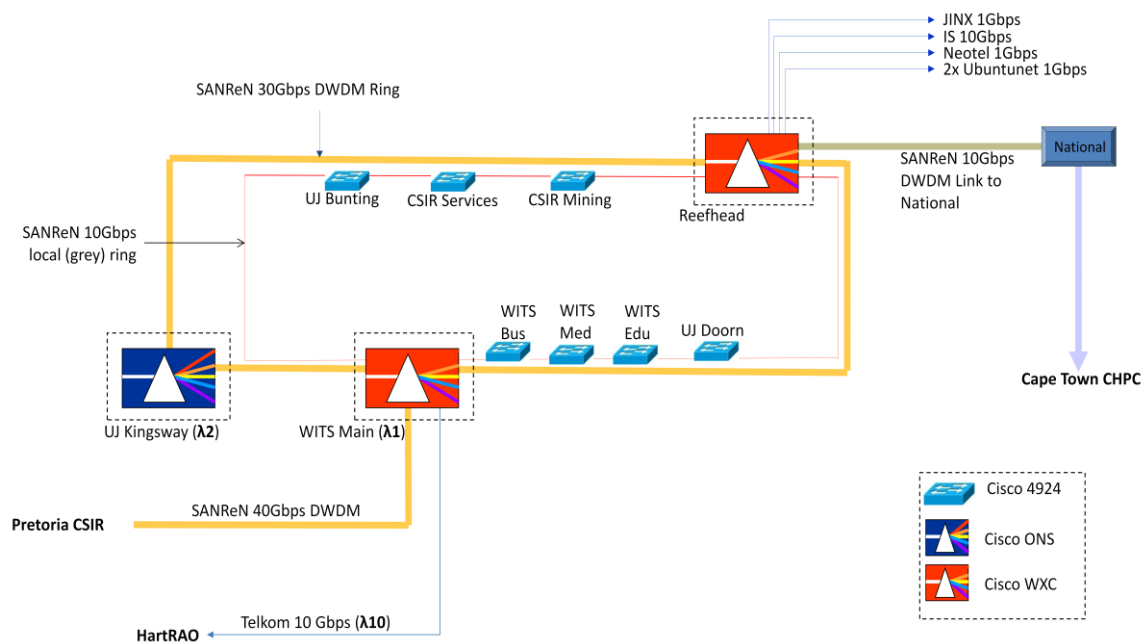
Figure 48: The Johannesburg Ring (logical view with connection at Wits)

The second possibility is obtaining a fibre IRU from either Neotel or DFA. In both cases there would be a direct dark fibre connection from the Reefhead all the way along the M1, N1 and into the CSIR core node in Pretoria (blue line Figure 47). This would make the Reefhead node a four-degree node (see Figure 49), whereas the Wits node would only be a two-degree node.

The problem was that the equipment had to be installed and commissioned before the question around the fibre route was settled. It was therefore decided to commission Wits as a full WSX site, but with only two shelves (and therefore only two degrees). This would cover either eventuality, as the node could easily be upgraded to a three-degree node should it be necessary.

10.1.3.2 The HartRAO/SAC connections

The connection to the CSIR core node via SAC at Hartebeeshoek would remain in any event, as its main function is to supply a redundant connection for the sites there, rather than to connect the Johannesburg and Pretoria networks.
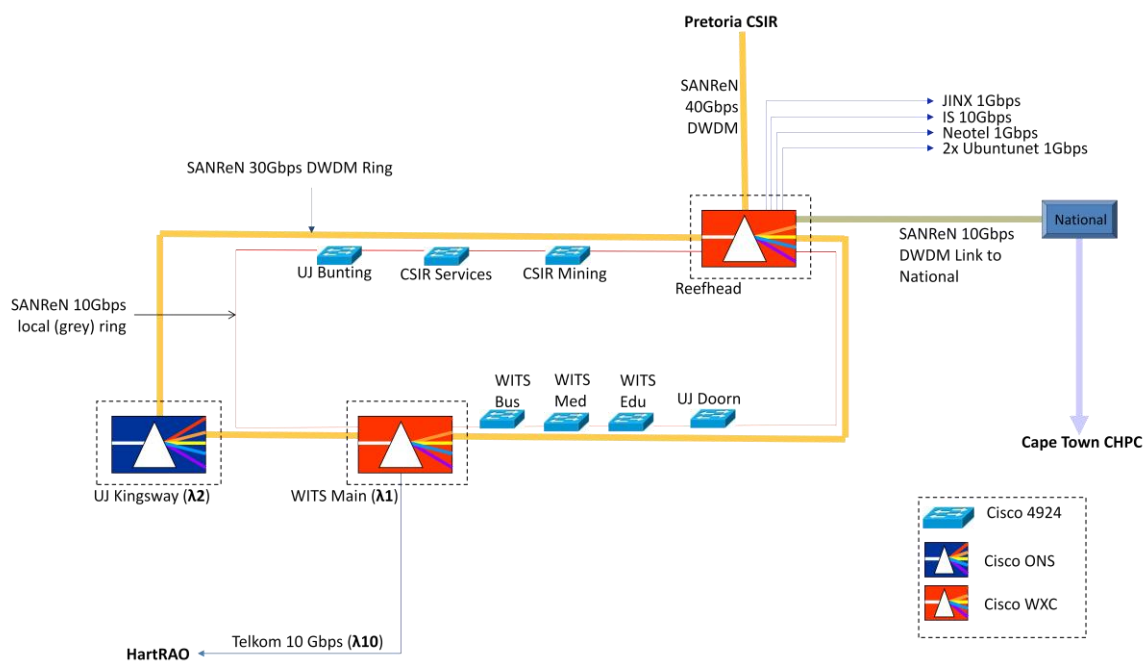
Figure 49: The Johannesburg Ring (logical view with connection at Reefhead)

10.1.3.3 The Cape Town connection

In anticipation of SANReN building an inter-city national backbone, the Reefhead core node had to have a connection to the backbone network provider for transporting traffic to the core node in Cape Town (Figure 50). Negotiations have not been concluded with any provider but it is foreseen that in the future the backbone will exceed 10 Gbps, therefore it was decided to plan for a dark fibre/DWDM connection between the provider's node and the Reefhead primary core node. Initially this connection will only have a single 10Gbps wavelength; however it is wise to plan for DWDM equipment at both ends of the link (Figure 50). This connection represents the third degree at the Reefhead, in addition

to the East and West connections to the DWDM ring. In accordance with the stated SANReN design principle, each degree in a WSX node will have a shelf dedicated to it.
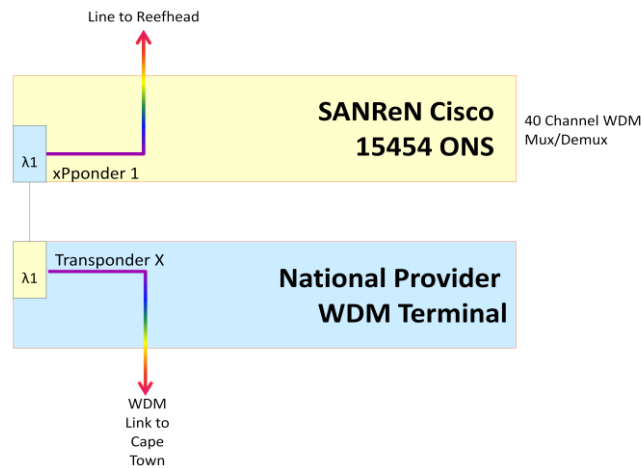
Figure 50: The SANReN/national provider node

10.1.3.4   Other external connections

The SANReN network has to be connected to the outside world; therefore the following connections are also required at the Reefhead:

- JINX (1 Gbps, upgradeable to 10 Gbps), for peering with other ISPs
- IS (10 Gbps), for peering with Internet Solutions
- Neotel (1 Gbps), for connecting the network to the TENET GEN3 MPLS network
- UbuntuNet (redundant 1 Gbps), for receiving the IS international bandwidth, and eventually the Seacom circuit

10.1.4   The primary core node (Reefhead)

Internet Solutions, which is the largest corporate ISP in South Africa, is the hosting company for the JINX. JINX is the virtual place where all the South African ISPs interconnect, so the Rosebank node of IS, where the JINX router is located, is a logical place to host the SANReN core node.

Another requirement for the core site is that it be the place of interconnection between the Neotel GEN3 MPLS network and the International circuits supplied by IS. Again the Rosebank site is ideal because it is both a core node of IS as well as a Neotel interconnect site, because of JINX.

10.1.4.1   Equipment at the Reefhead

The Reefhead core node hosts the following equipment:

- Core router (Cisco CRS-1)
- SANReN Provider Edge router (Cisco 7609s)
- UbuntuNet PER (Cisco 7609s)
- Wavelength cross-connect (Cisco 15454 WSX)
- ADSL aggregation router (Cisco 7200)

- Mirror server (Dell PowerEdge R900)

In order to fully understand the way the Reefhead node functions, it is necessary to construct a detailed wiring diagram, both logical and physical. The logical diagram is most important because it enables one to determine which ports are required and hence, how many cards are required (*Figure 51*). Only by constructing this diagram can the design team be sure that all the necessary cards and optics are ordered.
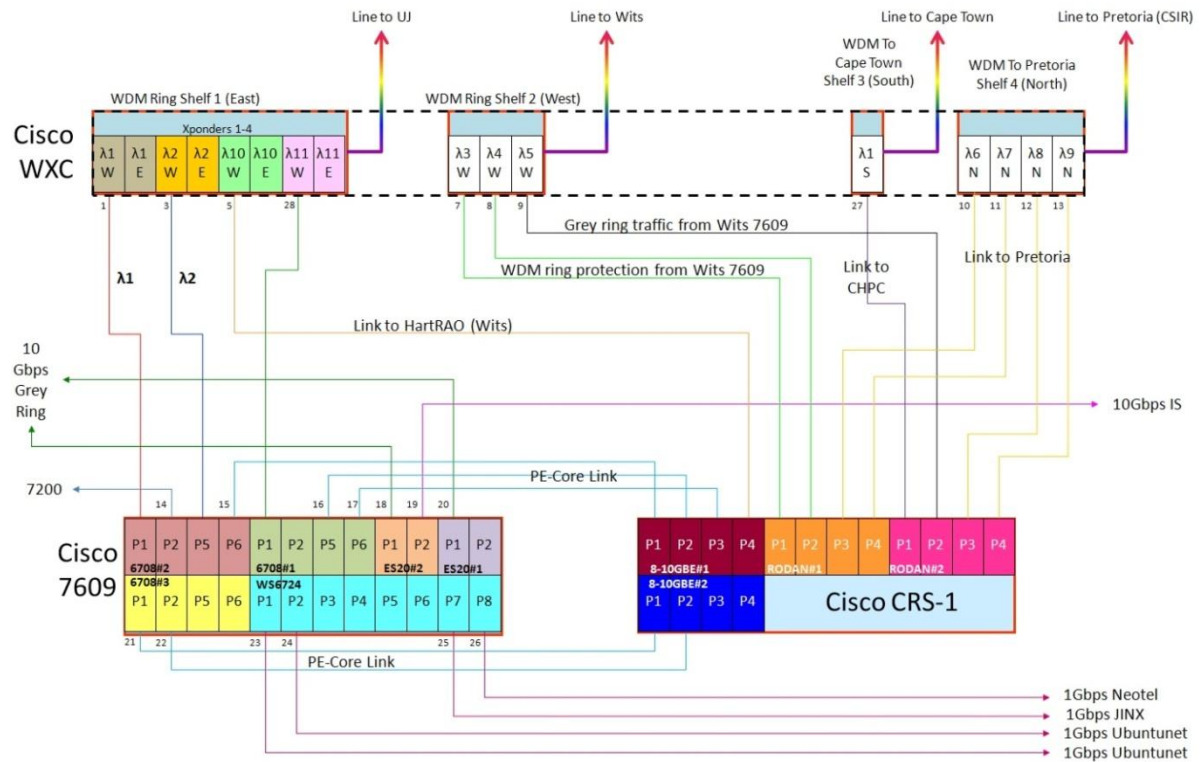
Figure 51: Reefhead logical wiring diagram

The connections may be summarised thus:

- 1: λ1 ring for the Wits traffic. Even though the muxponder has a connection to the West as well as East side of the ring, only a single connection is required between the Xponder and the PER card (6708). The potential single-point-of-failure in the Xponder is compensated for by the additional router failover path (Figure 26). Since the Layer2 functionality of the muxponder is employed, the ring connection to the card is seen as a Layer2 ring and therefore requires the spanning tree algorithm
- 3: λ2 ring for the UJ traffic
- 5: λ10 ring for the HartRAO traffic from Wits
- 7-8: λ3 and λ4 for single PER failover paths from the Wits PER, for the ring traffic (UJ and Wits). These ports are connected directly to the coloured optical inputs on the CRS-1 Rodan cards, and therefore do not require muxponder ports, only ports on the wavelength cross-connect
- 9: λ5 for PER failover path for the grey ring

89

- 10-13: λ6- λ9 for the dark fibre connection to Pretoria CSIR. The wavelengths were chosen in the event that the traffic must be carried over the ring to Wits
- 14: Connection to the 7200 router for access to the mirror server and ADSL aggregation router
- 15,16,17,21,22: 50 Gbps non-contended connection between PER and Core routers – 10 Gbps each for UJ and Wits traffic, Grey ring and IS peering link. 10 Gbps for other small connections (JINX, UbuntuNet, Neotel GEN3)
- 18,20: 10 Gbps grey ring connections onto ES20 QinQ cards
- 19: 10 Gbps IS peering connection
- 23,24: Redundant connections to UbuntuNet router on 12-port 1 Gbps WS6724 card
- 25: JINX peering connection (1 Gbps)
- 26: 1 Gbps Neotel connection for TENET MPLS network
- 27: 10 Gbps for DWDM link to national provider (Cape Town)
- 28: PER-PER link for closing the Layer 3 connections on the grey ring

Figure 52: Reefhead physical wiring diagram

Figure 52 illustrates the physical layout of the Reefhead node. Note the two 7609S PE routers (SANReN and UbuntuNet), the four WSX shelves and the ancillary equipment. The CRS-1 core router is unable to fit into the same data centre as the rest of the node and will therefore be hosted in an adjacent building. Clearly marked are the external fibre connections: a 12-strand cable (blue) to

Internet Solutions, a 12-strand cable (orange) to Neotel and a 48-strand cable to the core router (red).

10.1.5 The secondary core node (Wits)

10.1.5.1 *Equipment*

Wits is the secondary core site of the Johannesburg network, in addition to being a primary client (DWDM) site. It contains a two-degree wavelength cross-connect and a 7609s PER. The router's main function is to act as standby for the PER at the Reefhead.
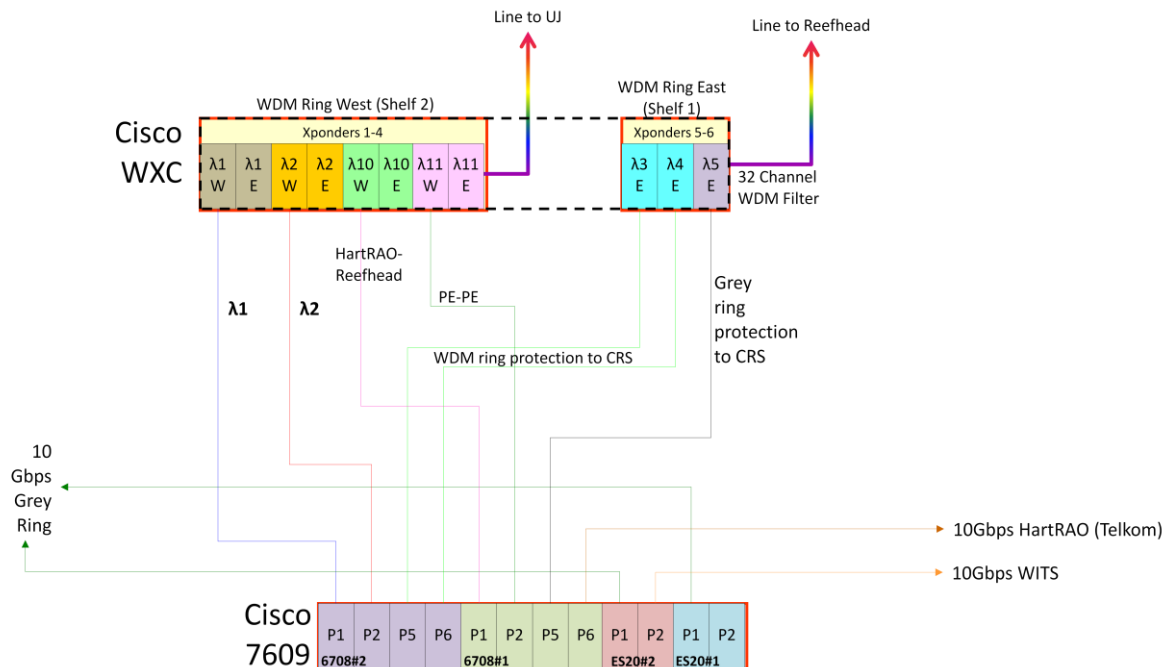
Figure 53: Wits Main logical wiring diagram

As can be seen from the network diagram (Figure 49), the grey ring also terminates here; however, the traffic from the smaller Wits sites would not normally be routed from the grey ring into the Wits Main switch. Instead, to keep the routing scheme simpler, the traffic is hauled back to the Reefhead and routed between the grey and DWDM ring by the primary PER. In the event of primary PER failure all the traffic from the Wits sites would be hauled back to the Core router and routed there. All sites' international traffic is received from the UbuntuNet router and distributed among the sites by the primary PER, according to the institutions' instructions. The actual subdivision of the international traffic among institutions is accomplished by the SCE in London, as explained before.

Referring to Figure 53 one can see that both wavelengths of the WDM ring are dropped at Wits, in case of primary PER failure. The grey ring traffic also flows through the secondary PER, as does the traffic from HartRAO, delivered to Wits on a Telkom SDH multiplexer. From here all four wavelength's traffic is routed to the core router via the ONS, respectively on lambdas 3, 4, 5 and 10. Lambda 11 is used for the Layer 3 grey ring closure.

Each wavelength is transported redundantly around the ring, and therefore requires both East and West outputs of the muxponders, except for Lambda 3-5, which are already redundant paths to the core router, and therefore do not require an additional level of redundancy. This enables the network to use a single muxponder for two wavelengths, so muxponder #5 lights both lambda 3 and 4, both in the anti-clockwise direction around the ring (East from Wits).

All ports are on the quad-port 6708 line cards, except for the grey ring and input from Wits campus itself. The inputs must be on ES20 line cards, to preserve the QinQ functionality.

10.2 The Pretoria Network

Pretoria has four large institutions to connect, as well as large number of smaller sites; both smaller sites of the large institutions, as well as the main sites of small institutions. The large universities are UP[1], TUT[2] and Unisa[3]. The main site of the CSIR is also located in Pretoria, and contains quite a number of smaller institutions.

10.2.1 Participating institutions

10.2.1.1 University of Pretoria

UP has a large main campus in the east of Pretoria, as well as several smaller but important sites scattered across the city, such as the Onderstepoort Veterinary School, the Pretoria Academic Hospital and the Groenkloof education campus. The main campus is by far the largest consumer of bandwidth and was therefore identified as a DWDM node. The smaller sites would all be interconnected on a shared 10 Gbps grey ring.

10.2.1.2 Tshwane University of Technology

TUT has a large main campus in the West of Pretoria, as well as 8 smaller sites spread across the city. All sites are interconnected via a high-bandwidth microwave network, with most of the links running at STM-1 wire speed[4]. In discussion with the IT department it was proposed to make the main site a DWDM node and to keep the microwave network in place to interconnect the smaller sites. As the SANReN fibre network expands over time, smaller sites would be picked up and their microwave links decommissioned, to decrease the institution's cost.

10.2.1.3 University of South Africa

Unisa has a very large main campus in the south of Pretoria, as well as a large satellite campus in Sunnyside, which is only a few blocks away. It was decided to make the main campus a DWDM node and pick up the smaller site on the grey ring. The size of the node, as well as a historical hosting relationship with TENET makes Unisa Main a good place for situating the SANReN secondary core node.

---

[1] University of Pretoria
[2] Tshwane University of Technology
[3] University of South Africa
[4] 155 Mbps

10.2.1.4   Centre for Scientific and Industrial Research

The CSIR's main campus in situated in the East of the city and is host to a number of smaller but important institutions, including the Meraka Institute (the owner of SANReN), the NRF[1] main campus, the NLC[2] and the DST[3] (the government body funding SANReN). It also has a large building dedicated for IT purposes, which makes it an ideal place to host the primary core node of SANReN in Pretoria. In addition to the CSIR as a "client" of the network, the IT building (building 9), also hosts the Pretoria node of the CHPC, called the C4 supercomputer.

The primary core node would contain a wavelength cross-connect, a PER and a CRS-1 core router. One 10 Gbps would be dedicated to the CSIR as a client, one to the C4 and one to a small intra-campus grey ring that would interconnect the smaller sites on the CSIR campus.

10.2.1.5   Other Institutions

Quite a few smaller institutions are located in Pretoria and, as existing TENET clients, will become fully-fledged SANReN nodes as soon as the network reaches them. These include the National Library, Water Research Commission, National Biodiversity Institute and the Human Sciences Research Council. All of these sites will be picked up on a shared 10 Gbps grey ring.

10.2.2   The Pretoria rings

The City of Tshwane has a very extensive optic fibre network that consists mainly of OPGW[4] cables strung along the high voltage (132 KV) power routes belonging to the City. The network was constructed originally for controlling the power grid, and as such connects most of the City's power substations on a 440 km network. A few years ago the City approached some of the TENET institutions such as UP and offered them access to the fibre network, which the City was prepared to see as its contribution to higher education. The offer could not be accepted at the time because the institutions were all committed to TENET as service provider, which did not at the time have a license to provide services on its own fibre systems.

In the SANReN era, and in the light of the famous Altech court case ruling, TENET (as the SANReN operator) has had its VANS[5] license upgraded to a full I-ECNS licence, which permits it to light dark fibres. The City was approached to supply fibre pairs for the network and proved to be very keen to help SANReN in its role as service provider to higher education. Negotiations for access to the network took a long time for various reasons, not least of which was the City's inability to provide SLAs[6] for the maintenance of the network. SANReN concluded negotiations and agreed that where there were "missing" pieces of the network, SANReN would undertake to construct the fibre cable links itself, after which it would turn ownership of the infrastructure over to the City. In exchange, the City would grant SANReN access to the network at only a nominal charge, instead of selling fibre IRUs at market value, which was the case in Johannesburg.

---

[1] National Research Foundation, owner of HartRAO
[2] National Laser Centre
[3] Department of Science and Technology
[4] Optical fibre composite overhead ground wire
[5] Value Added Network Services
[6] Service Level Agreements

Using the City's existing network and extending it in places as described, it is possible to interconnect the four primary sites in a single ring. A second fibre ring, which coincides with the primary ring only in certain places, will be used to commission the grey ring to interconnect the smaller sites.
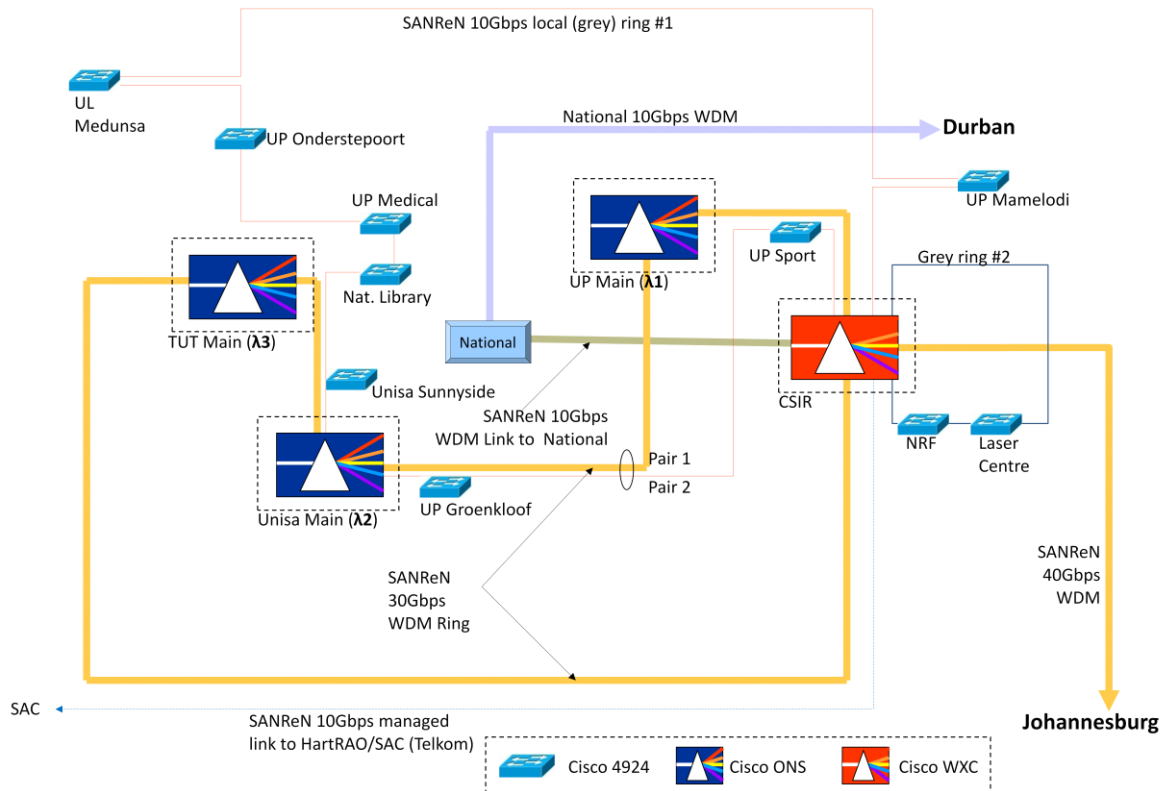
Figure 54: The Pretoria ring (logical view)

10.2.2.1 The DWDM ring

The DWDM ring has a single WSX at the CSIR primary core node, as well as three   2-degree ONS nodes, as illustrated in

Figure 54. TUT Main and UP Main have only the single shelf ONS, whereas Unisa Main has a single shelf ONS as well as a router rack containing a 7609s PER.

As in Johannesburg, the grey ring runs through the secondary PER at Unisa. Each primary site has a dedicated 10 Gbps wavelength backhauling its traffic to the primary core node. There are also failover paths configured for each lambda between the secondary PER and the core router for PER failover. Lastly there is the Layer 3 PE-PE link for closing the grey ring.

10.2.2.2 The grey rings

A single fibre pair runs throughout most of the City, picking up all the SANReN secondary client sites outside the CSIR campus (15 sites in total). Each site is equipped with a 4924 10 Gbps switch with QinQ functionality. By default each site would be connected by a 1Gbps connection, on either fibre or copper, as the site prefers.

The 10 Gbps bandwidth would be shared among the sites, with each site getting a fair share of it. Larger sites may require as much as 2 Gbps, whereas some of the smaller sites would only consume 50 Mbps or less.

The second grey ring connects all the other sites on the CSIR site to the core node. This includes the DST, NRF, National Laser Centre and Meraka itself.

10.2.3   The primary core node (CSIR)

The node consists of a CRS-1 core router, 7609s PE router and WSX with four shelves, one for each of the connections to Unisa and UP (DWDM ring), Johannesburg and Durban (Figure 55). As at the Reefhead, two kinds of cards are used on each of the routers. Grey rings and local traffic terminate on ES20 cards in the PER, all other connections are on quad-port 6708 cards. The CRS-1 has coloured-optic RODAN cards for ll direct interfaces with the WSX, all other connections (80 Gbps) are to the PER on quad-port 8-10GBE cards.
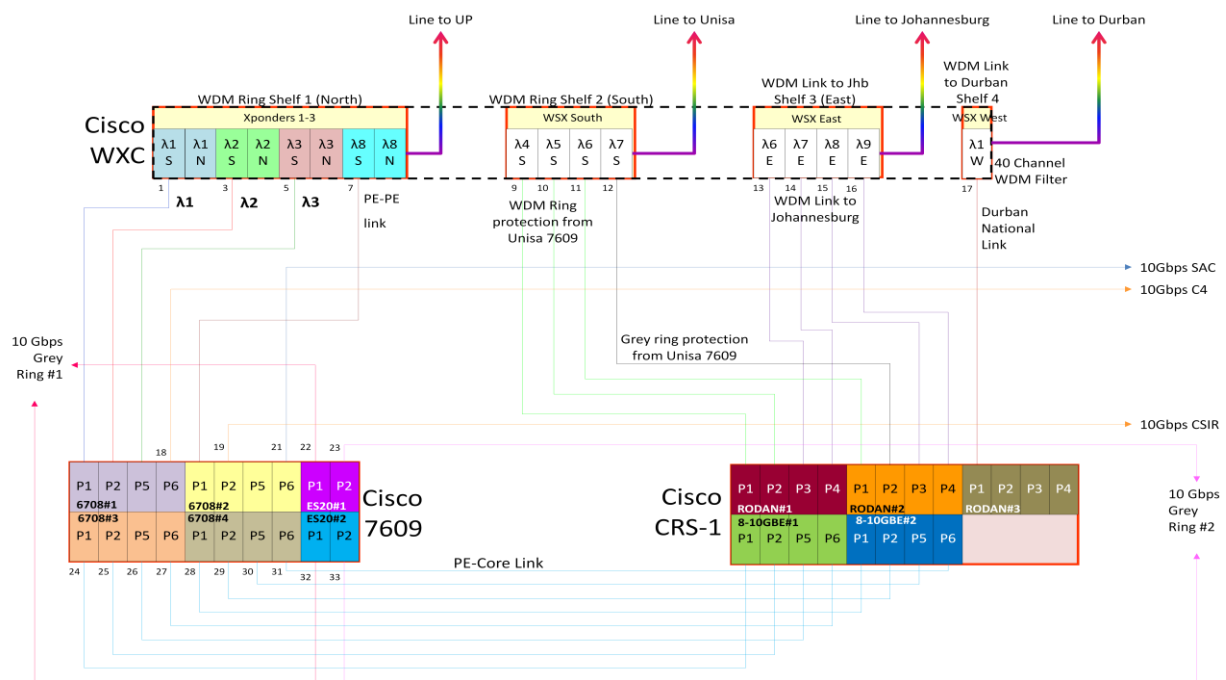
Figure 55: The CSIR Main logical wiring diagram

The following connections are made:

1. λ1 (UP traffic)
3. λ2 (Unisa traffic)
5. λ3 (TUT traffic)
7. PE-PE Link
9-11. WDM ring traffic from backup 7609 at Unisa in case CSIR 7609 fails
12. Grey ring#1 traffic from backup 7609 at Unisa in case CSIR 7609 fails
13-16. WDM link to Johannesburg (Reefhead)

17. WDM link to Durban (National)
18. C4 link
19. CSIR link
21. SAC link (Telkom)
22. Grey ring #1 West
23. Grey ring #2 West
24. Grey ring #1 traffic PER-Core (7609-CRS)
25. Grey ring #2 traffic PER-Core
26. λ1 (UP) traffic PER-Core
27. λ2 (Unisa) traffic PER-Core
28. λ3 (TUT) traffic PER-Core
29. C4 traffic PER-Core
30. CSIR traffic PER-Core
31. SAC traffic PER-Core
32. Grey ring #1 East
33. Grey ring #2 East

10.2.4   The secondary core node (Unisa)

The node at Unisa is very similar to the one at Wits, except for the fact that its DWDM node is a single-shelf ONS. It also contains a 7609s PER.
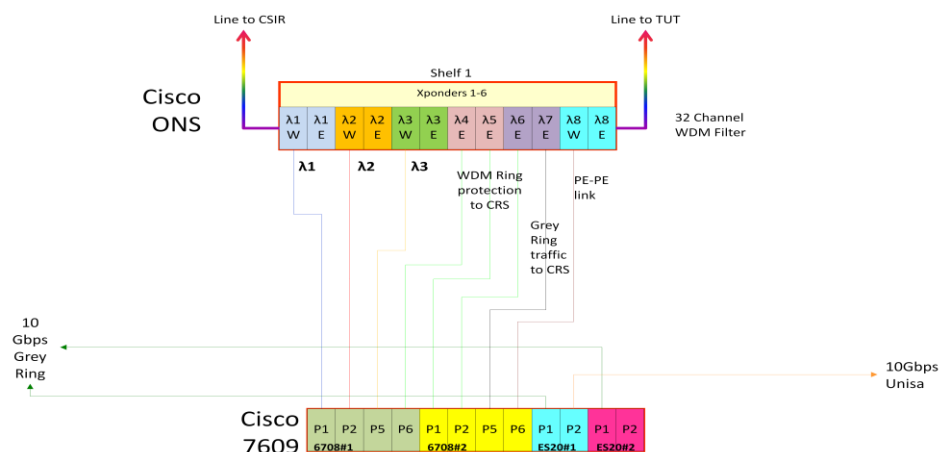
Figure 56: The Unisa logical wiring diagram

As before, the first wavelengths (1-3) are for the DWDM ring, including the traffic of Unisa itself (λ2). The grey ring also terminates here on ES20 cards, along with the local Unisa traffic, whereas all the DWDM traffic terminates on 6708 cards. Both the DWDM ring and grey ring traffic have failover paths to the CRS-1. Lastly there is the PER-PER Layer 3 ring for maintaining the routing failover system (λ8).

10.3 The Cape Town Network

Cape Town has four large universities, as well as several smaller institutions. The SANReN design team was involved in talks with the City's network design team at an early stage of the City's network design, and was therefore able to influence the City's network design (see section 13.1   ). The City was keen to help the tertiary education institutions and was amenable to the idea of adapting the design in order to build the network past all the sites of the universities while pursuing its own network requirements.

10.3.1 Participating institutions

10.3.1.1 University of Stellenbosch

US has three large campuses, two of which fall within the boundaries of the Cape Town Metro. Its main campus is in the town of Stellenbosch which, although a separate municipality not incorporated into the Metro, is adjacent to the Metro and also relatively close. The US Business School and US Medical School at Tygerberg Hospital are not far from the Bellville CBD, which is the Metro's second major business centre, after the Cape Town CBD itself. The SANReN design calls for a grey 10 Gbps ring dedicated to this institution, which will include all three campuses (Figure 57).

10.3.1.2 University of Cape Town

UCT has a large main campus on the edge of the Cape Town CBD, a medical school at Grootte Schuur Hospital, a business school in the waterfront and an art school in the CBD. It also has a small campus at the Red Cross Hospital in Athlone. It maintains a substantial fibre network of its own on the main campus and has also extended this network to include the Medical School, which is not far from the main campus. The main campus will be supplied with a DWDM node, whereas the two smaller campuses in the city will be part of a shared 10 Gbps grey ring.

10.3.1.3 University of Western Cape

UWC has a single large campus just to the south of Bellville CBD and will be serviced by a DWDM node.

10.3.1.4 Cape Peninsula University of Technology

CPUT has two equally large main campuses, one in the Cape Town CBD and one south of Bellville CBD, next to UWC. In addition, it has five other small campuses in the Metro area. The two large campuses will both be equipped with a DWDM node, whereas all the other campuses will be connected on shared 10 Gbps grey rings.

10.3.1.5 Other institutions

Other TENET/SANReN institutions in the Cape Town Metro include the MRC[1], iThemba LABS[2], the SAAO[3], MCM[1], NLSA[2], SANAP[3] and TENET itself. All these institutions will be connected on the

---

[1] Medical Research Council
[2] Laboratory for Accelerator Based Sciences
[3] South African Astronomical Observatory

shared grey 10 Gbps rings.
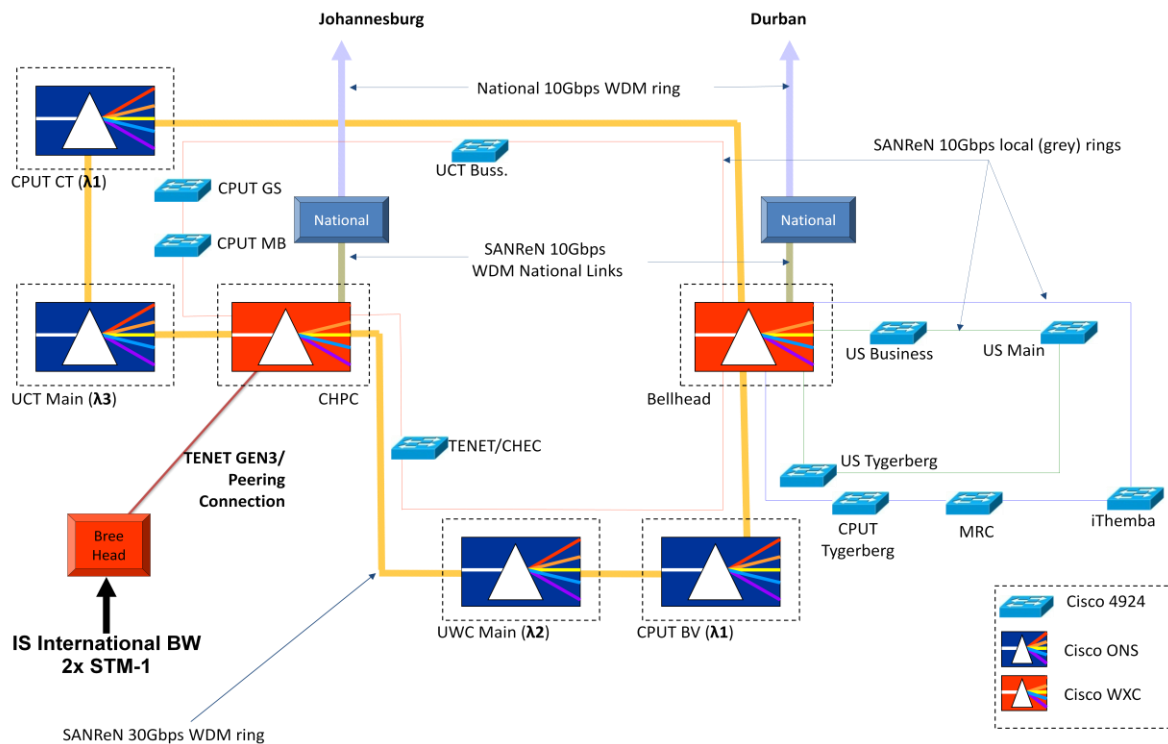
Figure 57: The Cape Town ring (logical view)

10.3.2   The DWDM ring

One can see from Figure 57 that there are the main campuses of UCT, UWC and CPUT (both) are connected on the DWDM ring to the primary (CHPC) and secondary (Bellhead) core nodes. Each of the institutions has a dedicated 10 Gbps wavelength assigned to it. The two campuses of CPUT share a lambda, but the bandwidth is of course, upgradeable on demand.

10.3.3   The grey rings

10.3.3.1   Ring #1

The main grey ring follows the same route as the DWDM ring and picks up all the smaller sites that do not justify the expense of a DWDM node. A single fibre pair is connected in turn to each of the sites and delivers service at 1 Gbps connection speed to every site. The ring covers about ¾ of the SANReN sites and includes 1 UCT campus, 3 CPUT sites and TENET. This ring is dual homed around both core nodes.

---

[1] Marine & Coastal Management
[2] National Library
[3] SA National Antarctic Program

10.3.3.2   Ring #2

This 10 Gbps ring is dedicated to the US and includes their main campus and both satellite campuses (Business and Medical School). This ring is single homed from the Bellhead.

10.3.3.3   Ring #3

This ring follows the same route as Ring #2 and picks up all the non-US sites along the route, such as iThemba LABS and the MRC, as well as the Stellenbosch CSIR and CPUT's Tygerberg School of Dentistry.

10.3.4   The national connections

All three potential national backbone providers (Telkom, Neotel & Infraco) have major nodes in Cape Town CBD as well as Bellville CBD. This creates the opportunity to ask for full diversity from Cape Town to the other major cities. Therefore, the Bellhead will have a connection to Durban (Sharkhead) and the CHPC will have a diverse connection to Johannesburg (Reefhead).

In each case SANReN will maintain a dark fibre DWDM connection between the relevant core node and the provider's node. Even though the initial capacity of 10 Gbps may not seem to justify the expense of a DWDM link, the ease of upgrading capacity more than compensates for it.

10.3.5   Other connections

The future plans to degrade the Breehead's functions means that it does not make sense to include it in the DWDM ring. Therefore, it has to fed on a separate fibre pair running on the City's network between the CHPC and the Breehead. This link will be lit at 10 Gbps to cater for gigabit-level peering between the SANReN network and the present IS network, as well as the future CINX[1].

For the lifetime of the TENET GEN3 network, all TENET sites in the Western Cape that have not been cut over onto SANReN will still be connected to the Neotel MPLS cloud and therefore have to receive their international bandwidth via the Breehead. As mentioned before, these sites will continue to receive this bandwidth directly from IS until the Seacom circuit is live. After the Seacom cutover, the GEN3 sites will receive their international bandwidth from Seacom, via the Sanren national backbone into the CHPC, and from there into the Neotel MPLS cloud via the CHPC-Breehead connection.

10.3.6   The primary core node (CHPC)

The Centre for High Performance Computing is another Meraka/CSIR project and is based at the CSIR site not far from UCT main campus. It is the main national node of the CHPC and as such is a logical location for a core SANReN node. The fact that it also has a fully equipped data room with air-conditioning, backup power and security makes it ideal.

The CHPC itself has had many problems in the past with bandwidth capacity. There are many institutions all over the country that would like to make use of its enormous computing capacity, but

---

[1] Cape Town Internet Exchange

are unable to do so because of the very limited bandwidth that the centre can afford. Even UCT, which is less than 4 km from the centre, has been unable to make good use of the centres' facilities.

The CHPC data room will host the primary core node for SANReN. This will include a CRS-1 core router, a 7609 PER and a full 3-degree DWDM ROADM. In addition to the two sides of the DWDM ring connecting the institutions; it will also be connected to the national backbone provider's node in Cape Town CBD to receive the connection to Johannesburg.
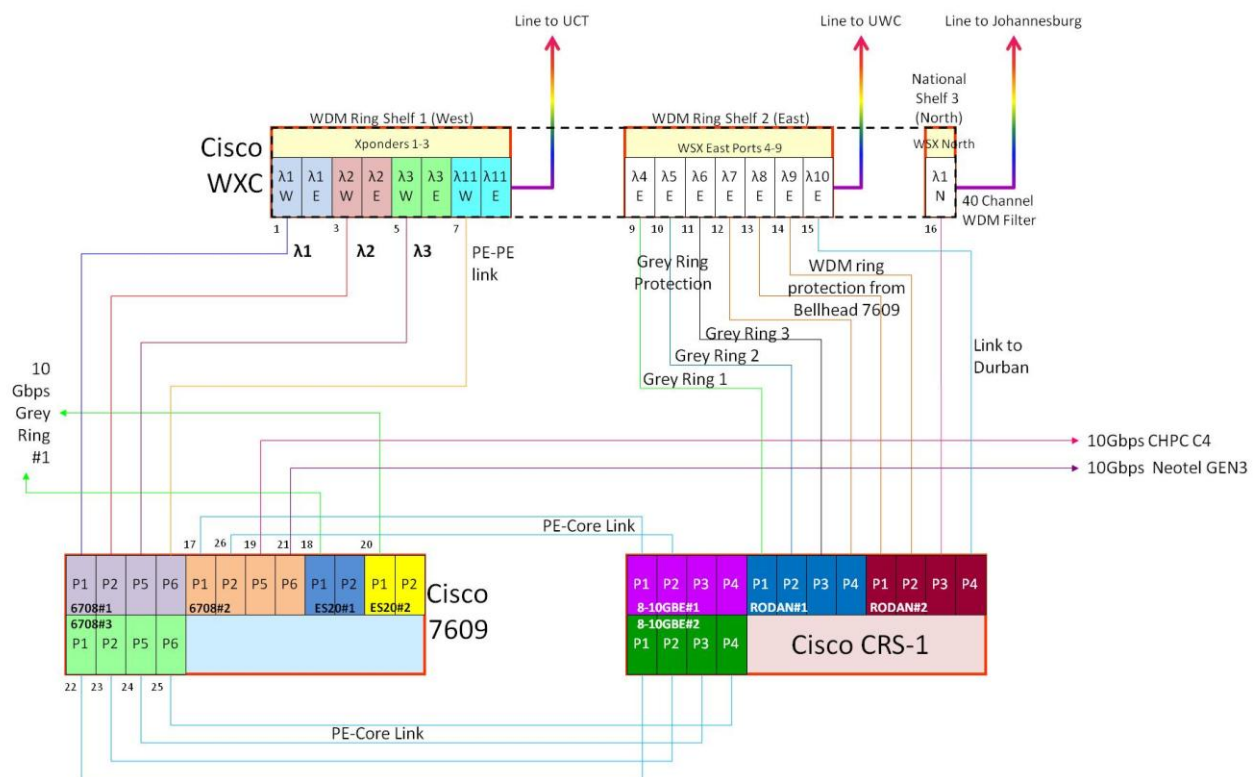
Figure 58: CHPC logical wiring diagram

Referring to Figure 58, the following patch connections exist:

> 1: Lambda 1 (CPUT traffic) link between WSX and PER. Only one connection is required for redundancy as the Layer2 functionality of the Xponder maintains both the East and West connections around the ring. To protect against card failure, as well as primary PER failure, router failover paths exist from the secondary PER
>
> 3: Lambda 2 (UWC traffic)
>
> 5: Lambda 3 (UCT traffic)
>
> 7: Lambda 11: PER-PER link to close the grey ring at Layer 3
>
> 9-11: router failover paths for grey rings 1-3 respectively
>
> 12-14: router failover paths for DWDM traffic
>
> 15: Durban link, via DWDM ring to Bellhead, then to national provider
>
> 16: Johannesburg link to national provider

17, 26: PER-Core router link for grey rings 1&2

19: Link to CHPC supercomputer

21: grey 10 Gbps link to Breehead for GEN3/peering

18, 20: East and West of grey ring 1

22: PER-Core router link for grey ring 3

23-25: PER-Core router link for DWDM traffic

10.3.7 The secondary core node (Bellhead)

The City of Cape Town is constructing 7 new switching centres as part of its core broadband network (Figure 73).

The Bellville site will host the SANReN secondary core node called the Bellhead, which will contain a 7609 PER and a 3-degree DWDM node. It will be a termination point for all three grey rings, the DWDM ring and the link to the Sharkhead in Durban, via the national backbone providers' Bellville node (Figure 59).
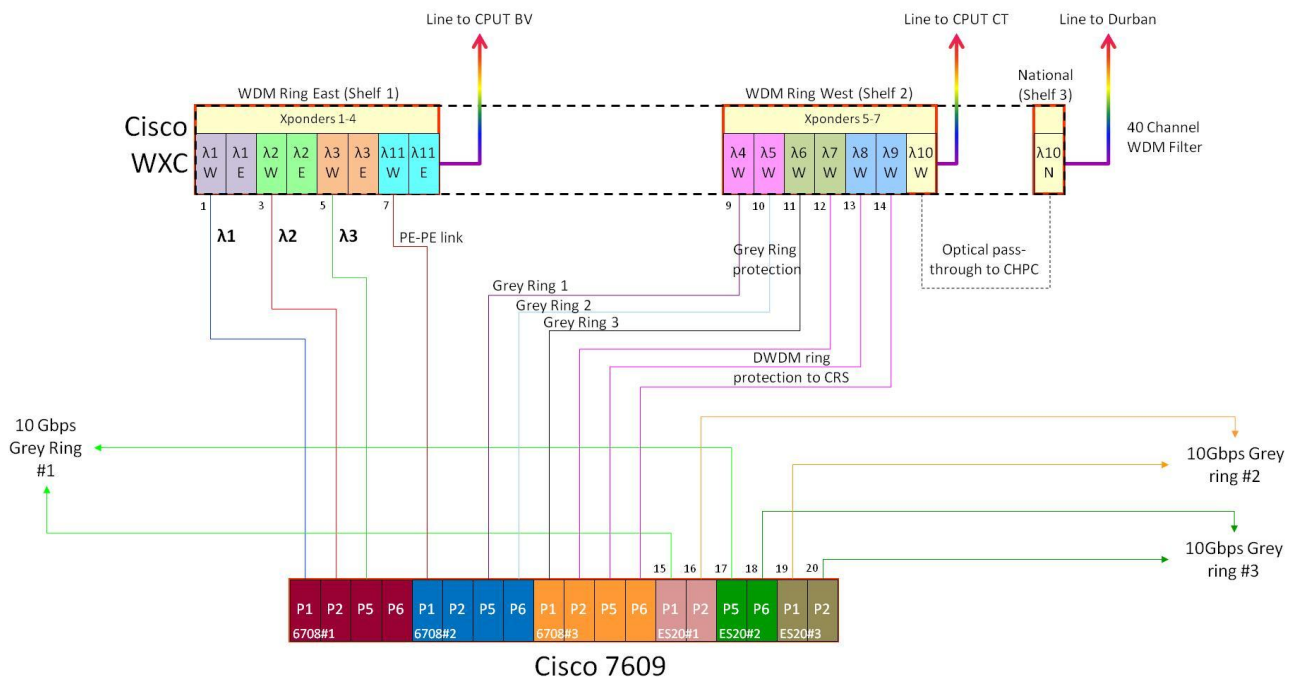
Figure 59: Bellhead logical wiring diagram

The Bellhead requires the following connections:

1: λ1 link for CPUT traffic from WSX into secondary (backup) PER

3: λ2 link for UWC traffic WSX -PER

5: λ3 link for UCT traffic WSX -PER

7: Layer 3 PER-PER link for grey ring #1 closure

9-11: PER failover paths between secondary PER and Core router for all three grey rings

12-14: PER failover paths for DWDM ring traffic

15,17: grey ring #1 **dual** homed to secondary PER

16,19: grey ring #2 **single** homed to secondary PER

18,20: grey ring #3 **single** homed to secondary PER

Note also the optical pass-though wavelength (λ10) from the CHPC to the national provider's Bellville node for the Durban national backbone connection.

10.3.8   The interconnect node (Breehead)

The Breehead is a TENET site that has to be incorporated into the SANReN network for several reasons. Until July 2009 when the Seacom 10 Gbps circuit is ready for service, all the southern sites on the TENET network will still be fed international bandwidth from the IS STM-1 circuit which is delivered at the Bree St facility of IS. Therefore, all SANReN sites in Cape Town that go live before July 2009, will also be fed international bandwidth from the Breehead. In addition, Bree St is also where TENET peers with IS and its customers (like Akamai and Google), as well as being the edge of the TENET/Neotel GEN3 MPLS cloud.

After the Seacom circuit goes live, the Breehead will in turn be fed from Mtunzini, via the SANReN national backbone. The site will not be decommissioned before all TENET sites are cut over to SANReN, at which point only the peering function will remain. The proposed CINX distributed network will almost certainly have a node here, which means that the Breehead will remain the peering point for the SANReN network. As sites are cut over from the TENET network and onto the SANReN network, there will be a decrease in traffic on the Breehead-CHPC link. However, the link will be maintained at 10 Gbps speed as the amount of peering traffic is expected to increase over the next few years. It is conceivable that the demand for peering bandwidth will exceed 10 Gbps as more and more commodity Internet traffic is moved onto local clusters such as Google, Akamai and Limelight. When this happens traffic that was previously considered as international (and therefore Seacom traffic) in effect becomes local (national) traffic.

10.4   The Durban Network

The City Council of Durban (eThekwini) has constructed an extensive municipal fibre network that is an extension of its legacy network used for controlling the city power grid. The council have said in the media that it intended to have all higher education campuses connected to the network by end the of 2008. However, the costs involved in making new fibre connections and buying 10 Gbps services from the City's operator (MetroConnect), proved to be prohibitive. SANReN thus decided to purchase dark fibre pairs form DFA, who have also built an extensive metro fibre network. None of the sites in Durban justify the use of DWDM yet, therefore the Durban network consists of a DWDM ring to connect the Beachhead in Mtunzini with the Sharkhead core node in the Durban CBD and the national provider node at the Durban station, as well as two grey rings to connect the client sites to the Sharkhead (Figure 60).
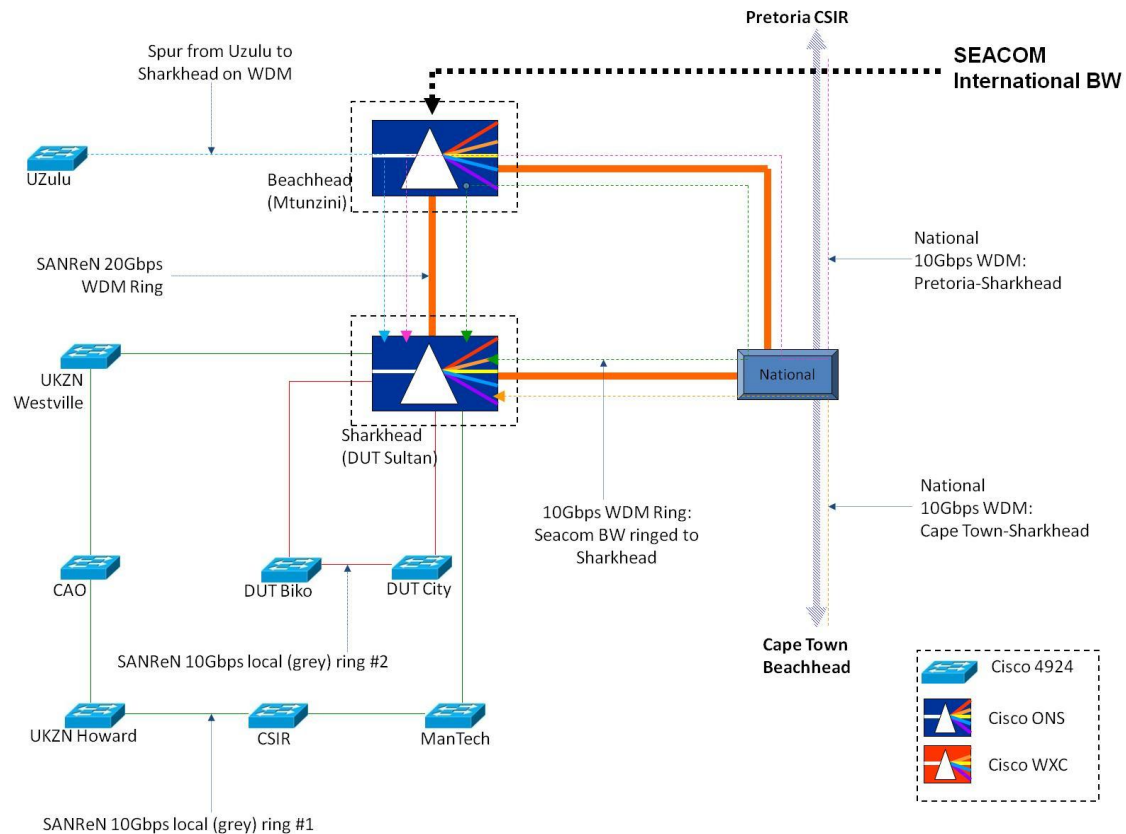
Figure 60: The Durban network (logical view)

10.4.1 Participating institutions

10.4.1.1 University of KwaZulu-Natal

The UKZN has two large campuses and two smaller ones. The Howard College site used to be the main site of the University of Natal and houses among other things, the IT department and engineering faculty. The Westville campus used to be the main campus of the University of Durban Westville and is the administrative centre for the institution. Both sites will be connected on a 10 Gbps grey ring to the SANReN core site. Depending on bandwidth demands, these sites may be upgraded to DWDM in future, leaving their 4924 switches spare for redeployment. The Edgewood and Medical School campuses will be connected at a later stage, as the SANReN network grows.

10.4.1.2 Durban University of Technology

The DUT has five campuses in Durban, of which four are clustered together in the CBD, including the main campus (ML Sultan). The IT department have already connected all the Durban sites together on a WAN, comprising some dark fibre and some wireless links. Therefore, all the institution requires is for the main site to be connected to the Durban SANReN network. The ML Sultan site will also host the Sharkhead core node of SANReN, and therefore have a direct 10 Gbps link into the Sharkhead PER.

10.4.1.3   Mangosuthu Technikon

ManTec has a single large site to the south of Durban in Umlazi, not far from the airport. The site will be connected on a shared 10 Gbps grey ring by "tromboning" a fibre link from the grey ring #1 to the site.

10.4.1.4   The Central Applications Office

The CAO is a small site not far from the Howard College site of UKZN, which handles student applications for the KwaZulu-Natal universities.

10.4.1.5   CSIR

The CSIR has an important site in Durban which is very close to the Howard College branch of UKZN and may easily be connected on the grey ring serving UKZN.

10.4.2   The DWDM ring

Durban has a single core node (the Sharkhead) around which all the customer sites are single homed. However, there must also be a PoP in Mtunzini to accept the 10 Gbps Seacom circuit, so it makes sense to establish a node at the Seacom landing station. This node will be known as the Beachhead.

In order to establish a secure backhaul for the Seacom circuit to the rest of the SANReN network, redundant dark fibre links must be established between:

1. The Beachhead and Sharkhead
2. The national provider's network and the Sharkhead

Assuming that (a) the national links to Pretoria and Cape Town terminate at the same node (as seems likely) and (b) that the Sharkhead is not co-hosted at the same site as the national provider's node: the simplest configuration is three dark fibre links connecting the three nodes on the SANReN network.

Referring to Figure 60, the Seacom circuit is transported on TENET DWDM equipment between the Sharkhead and the Beachhead. The direct (worker) path goes from Beachhead to Sharkhead via the national provider's node; whereas the secondary (backup) path goes directly from Beachhead to Sharkhead on the redundant dark fibre link.

Once this DWDM ring has been established, it is a simple matter to make the national connections redundant as well. If both the Pretoria and Cape Town circuits were on the shortest path between the national provider's node and the Sharkhead, a single fibre break would be able to drop the entire Durban network off the SANReN national network. Therefore, the southern circuit (Cape Town) is taken directly to the Sharkhead on the shortest path (Figure 61). The Pretoria circuit, however, is diverted to the Beachhead, where it is just patched across from East to West.
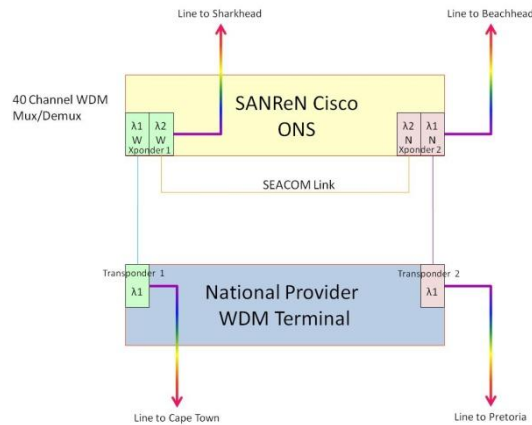
Figure 61: The national provider node in Durban

10.4.3   The grey rings

10.4.3.1   Ring #1

This ring is used to connect all of the non-DUT in Durban to the Sharkhead core node. These include Mantech, CSIR, CAO and both sites of UKZN (Howard College and Westville). In due course, as the Durban fibre network expands and more fibre is available, this ring will be diverted to include the two other UKZN sites (Medical School and Edgewood).

As the bandwidth needs of the institution increases, UKZN may opt for upgrading its grey ring to a DWDM ring, in which case each site will have as much as an entire wavelength (or even several) dedicated to it. This will mean that at least the two large sites be upgraded to have an ONS node each.

10.4.3.2   Ring #2

The second grey ring will pick up all the sites of DUT in the CBD (ML Sultan-, Steve Biko-, Brickfield-, Ritson- and City campus) using a combination of existing cables belonging to DUT and new dark fibre pairs bought from DFA. As usual each campus will be fed at GbE speed from a Cisco 4924.

10.4.4   Other connections

The network architecture explained above is the preferred one for it allows easy capacity expansion in the future. It is more than likely that the SANReN network will outgrow its single 10 Gbps circuit and the option of upgrading it must be considered. As explained before, most of the TENET institutions (26 of 40) have contributed to the financing of the Seacom circuit. However, it will only take 6 years from commissioning to pay off the capital amount for the purchase of the IRU. After this time the price per Mbps will collapse. As the institutions do not necessarily want to reduce their bandwidth spending, but would rather buy more bandwidth with the same monthly expenditure, one must assume that after the IRU has been paid off the institutions will be interested in purchasing another 10 Gbps international circuit.

It is also quite likely that other organisations such as the Department of Education will wish to purchase IRU's on the Seacom cable to add to the bandwidth available to the SANReN network. It would therefore be foolish indeed to assume that only one circuit will ever be needed between the Mtunzini site and the Sharkhead.

10.4.4.1   University of Zululand connection

The campus of Uzulu is in a small town called KwaDlangezwa, which about 25km north of Mtunzini. As may be expected from its remote location, it has never had any satisfactory service from TENET or anyone else. TENET has made arrangements with DFA to extend the Durban-Mtunzini fibre link all the way up to Uzulu. Since the Beachhead site does not have a router the only way to connect the site to the rest of the SANReN network is to provide it with a dedicated lambda on the DWDM core which is backhauled to the Sharkhead.

Figure 62: Uzulu connection to Beachhead

10.4.5   The Durban core node (Sharkhead)

In keeping with the SANReN principle of having a core router and PE router in each major city the Sharkhead will get both, as well as the DWDM ADM to transport the national and Seacom traffic on. The CRS-1 uses the standard 8-10GBE cards for all ports, rather than the coloured-optic RODAN cards used in the other cities as it would require at least two RODANs for redundancy, which would be much more expensive than simply adding two muxponders to the DWDM ADM.
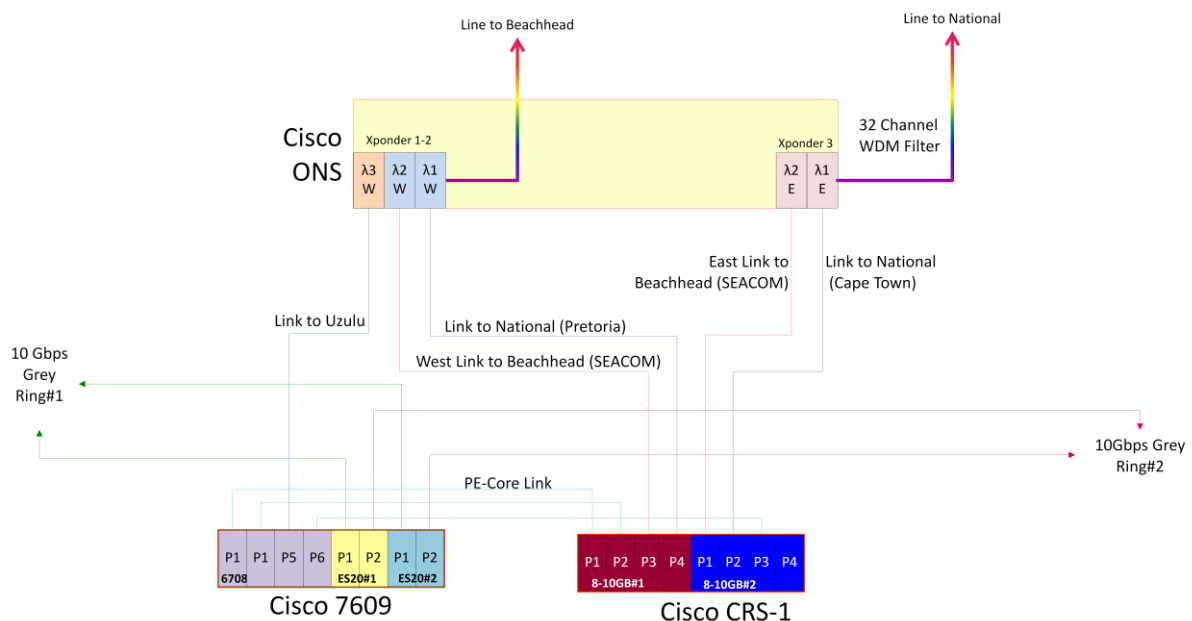
Figure 63: The Sharkhead logical wiring diagram

As usual each port on the tributary side of the PER has a corresponding connection to the core router to ensure full non-blocking traffic flow. The local grey rings are also terminated on the usual ES20 line cards for QinQ integrity (Figure 63).

10.4.6   The Mtunzini core node (Beachhead)

Space constraints in the Seacom landing station meant that SANReN did not have the option of placing a PE router there. The only piece of equipment that could be placed here was a single-shelf ONS which picks up the Seacom STM-64 circuit, as well as the Uzulu circuit.

The approximate fibre distance from Uzulu to Mtunzini is 25km, so a short range 1550nm optical interface is required, rather than a short range 1310nm interface, which only has a reach of about 10km. Unexpectedly, Seacom specified a 1550nm interface to their ADM, which necessitated a swop-out with the standard 1310nm interface which was originally ordered.
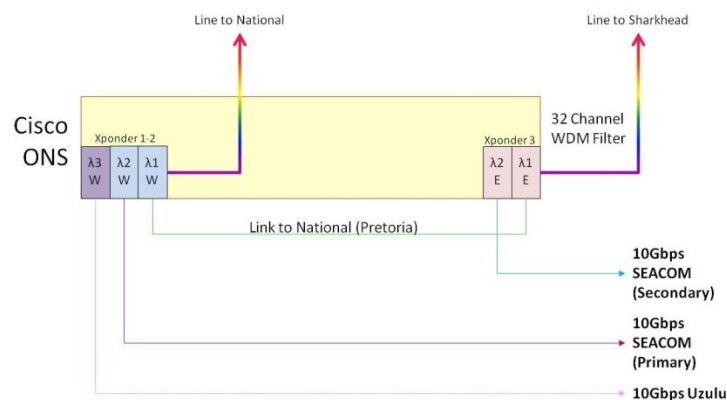
Figure 64: The Beachhead logical wiring diagram

Ultimately, SANReN decided not to get involved in the Mtunzini-Durban link, and it was therefore up to TENET to buy the dark fibre link from DFA, as well as install the ONS equipment at either end.

Chapter 11

Designing the SANReN: Layer 3 Network

In addition to building a transmission network interconnecting all the SANReN nodes (what may be called a Layer 0/1/2 network), SANReN also has to construct a Layer 3 services network. The transmission tiers do not fit neatly into the standard OSI[1] model, but for the purposes of illustration, and with reference to the SANReN network, it may be said that Layer 0 is the dark fibre, Layer 1 is the DWDM layer and Layer 2 is the gigabit and 10 gigabit Ethernet layer.

TENET has never before managed its own Layer 3 network. In the days of the Heist and GEN2 networks, TENET was forced to use the Telkom national network, which was ATM-based in the HEIST era and MPLS-based in the GEN2 era. TENET never had any say in any network configurations and was forced to take whatever Telkom dictated. TENET has always wanted to build a Layer2 network, which would enable it to build its own Layer 3 network.

It was only in the GEN3 era that TENET was able to influence the network design. The international connectivity that it buys from IS is a Layer 2 service only, which means that TENET may build its own Layer 3 network over the international links and distribute the bandwidth as it sees fit. It is true that the GEN3 network itself is still a third-party MPLS network, but TENET has read-only access to all the CPE routers on the network, and may ask for configuration changes as long as they do not impact the wider Neotel national MPLS network.

With the commissioning of the SANReN network, TENET is able to realise its dreams of managing a full Layer 3 network, which is not dependent on any third party for configuration and bandwidth management. TENET, as the operator of the SANReN, will have full control of the entire network, down to CPE level.

11.1 Layer 3 Considerations

Once the underlying transmission network has been designed, the Layer 3 MPLS network can be superimposed over it. Several issues should be addressed in the design, such as IP space, network routing protocols and virtual LANS.

11.1.1 IP address space

It was agreed between SANReN and TENET that the IP address space needed for the SANReN network would be assigned from the TENET block of addresses, as TENET is a registered local domain name authority. Also, all of the SANReN client institutions are already TENET clients and as such have blocks of TENET IP space assigned to them. It makes sense to keep the status quo intact

---

[1] Open Systems Interconnection

during the migration from TENET to SANReN, as this will keep the institutions part of the TENET AS[1] (AS 2018), which means that the outside world would not notice the move. If SANReN insisted on any other solution, it would have to register its own AS and apply for a new block of IP numbers.

Two /24 IPv4 blocks and four /64 IPv6 blocks have been allocated to SANReN:

- IPv4: 155.232.152.0 to 155.232.155.0
- IPv6: 2002:155.232.152::/64, 2002:155.232.153::/64, 2002:155.232.154::/64 and 2002:155.232.155::/64

11.1.2   Rerouting: STP vs. G709

Since the muxponder cards on the ONS multiplexers are to be configured with Layer2 functionality, the opportunity exists to see the entire optical network as a Layer2 (Ethernet) network. This would mean that the redundant route switching would have to be done using *spanning tree protocol*. However, since the G.709 transmission system has a much faster rerouting/convergence time (approximately 50ms), it was decided to use G.709 wherever possible. On the grey rings, which do not run on the G.709 platform, rapid spanning tree protocol[2] will be used.

11.1.3   Interior Gateway Protocol

IS-IS[3] is a commonly used IGP and was chosen to route traffic across the SANReN network. IS-IS is a link-state protocol which means that it floods topology information through the network using *hello* packets, not unlike what happens in an Ethernet network. Each router on the network then independently forms a picture of the network topology and then uses the *Dijkstra shortest path* algorithm to determine the best route to forward packets through the network.

SANReN chose IS-IS over the more popular OSPF[4] because it scales more easily into large networks, and also because it can easily be used for IPv6, whereas OSPF was designed for IPv4 networks.

11.1.4   Exterior Gateway Protocol

BGP[5] is by far the most widely used EGP used to route traffic between autonomous systems, and it may be said that it is the core routing protocol of the entire Internet. TENET already uses it to interconnect its AS to the outside world, and it is thus inevitable that SANReN should use it too.

11.1.5   Label Distribution Protocol

The SANReN is built as a standard MPLS network with Cisco 7609s provider edge routers and Cisco CRS-1 core routers. The standard MPLS LDP is used to distribute labels between the routers in the network. The routers in use on the SANReN feature LDP session protection which automatically creates a targeted LDP session between directly connected routers and also keeps the labels alive in

---

[1] Autonomous System
[2] IEEE 802.1w
[3] Intermediate system to intermediate system
[4] Open Shortest Path First
[5] Border Gateway Protocol

the case of link failure. Once the link has been re-established the LSP is reconnected automatically, reducing convergence times by addressing the LDP/IGP synchronisation problem, which occurs when the link is re-established and the IGP has converged, but labels have not been exchanged. During this period VPN traffic will be black-holed (packets discarded).

11.1.6   QinQ

QinQ is a variation of the IEEE 802.1Q standard, which allows for Ethernet switches to "nest" VLAN tags. This enables the network to have VLANs within VLANs and is implemented by adding two sets of VLAN tags to the Ethernet frame.

The reason SANReN decided to employ this technology is to enable user VLANing. If one considers a grey ring on the SANReN network, each site will be assigned a VLAN which will keep that sites' traffic separate and secure from every other site on the ring. If two sites of the same institution were on the same ring, and hence on different VLANs, traffic would be routed between the two sites by the core router, via the PER.

QinQ enables each site to set up several VLANs of its own within the VLAN assigned by SANReN, considerably easing the transition from the local LAN onto the wider SANReN network.

The main issues regarding the design of the SANReN Layer 3 network design have been dealt with and individual router configuration will set up the rest of the network.

The next chapter draws parallels between the TENET/SANReN network and other established NRENs and seeks to determine whether the SANReN network design is truly comparable with world-standard networks.

Chapter 12

Benchmarking the SANREN

12.1   Louisiana Optical Network Initiative

Strictly speaking, LONI is not a NREN, but a member network of both the *National Lambdarail* and *Internet2 Network*; however for purposes of illustration one may liken Internet2 to a continent-wide RREN, of which LONI would form one of the NREN members as the "national" network for the state of Louisiana.

LONI's first and foremost purpose is to serve as the backbone network interconnecting the supercomputer clusters scattered around the state of Louisiana and also in neighbouring Mississippi. The primary cluster contains a supercomputer called the "Queen Bee", which has a processing capacity of 50 Teraflops[1], making it one of the top 150 supercomputers in the world. Together with the other clusters on the LONI backbone, the total amount of processing available to researchers connected to LONI is 85 Teraflops.

In addition to interconnecting supercomputer clusters themselves, the LONI also connects user institutions to the supercomputer grid, as one would expect the user researchers to be located at sites other than ones containing the computers themselves. The institutions connected to LONI are the Louisiana State University (LSU), LSU Health Sciences Centre in New Orleans, LSU Health Sciences Centre in Shreveport, Louisiana Tech University, Southern University, Tulane University, University of Louisiana at Lafayette (ULL) and University of New Orleans (54).

LONI also forms a part of the wider national *Teragrid*, which is the largest and most powerful distributed supercomputer super grid in the world, with over one petaflop[2] of processing power (6).

The reason LONI makes a good example to benchmark the SANReN against is that it is built along very similar needs and principles. Although supercomputing is not the primary reason for building the SANReN, the CHPC nodes do form cornerstones of the SANReN network.

Examining the LONI network diagram (Figure 76), one can see that it forms three rings which are interconnected with a four-degree Cisco WSX cross connect at ULL and a six-degree WSX at the LSU main campus. All other nodes on the network are served by Cisco ONS nodes.

---

[1] 10^{12} floating point operations per second
[2] 10^{15} floating point operations per second
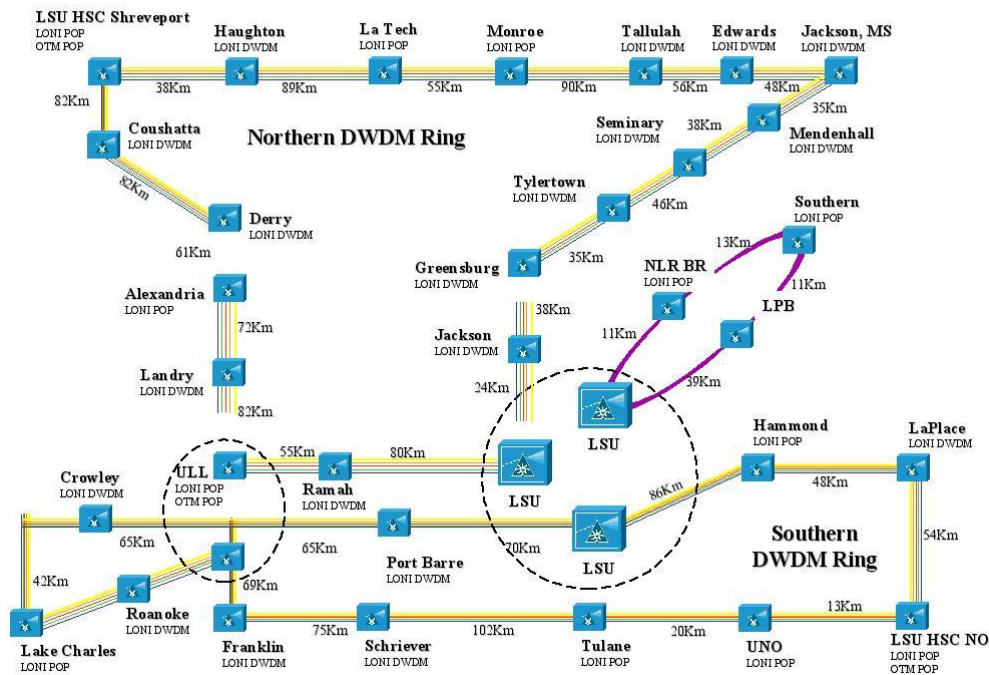
Figure 65: The LONI network (64)

Each node contains at least one PE router (Cisco 6500) in addition to the ONS node. To date LONI have installed 52 ONS nodes, 28 PERs and 31 x 10 Gbps links between nodes (54). The LONI is interconnected to the national Lambdarail at three sites, enabling its member researchers to commission almost any amount of bandwidth, up to and including a full 10 Gbps wavelength between their site and any other site on the Lambdarail, anywhere in the US(55). Also, in keeping with its secondary function of providing commodity Internet services to its member sites, LONI has recently acquired a 600 Mbps GbE connection to the commercial Internet, courtesy of *tw telecom inc*. (56)

Comparison between LONI and SANReN makes it clear that the SANReN network design measures up well to one of the most sophisticated research and tertiary education networks in the world. Since the SANReN network has not been completed, it is difficult to analyse traffic patterns to see if it does indeed exhibit the same kind of usage and impact that an established network such as LONI does. However, the design team from TENET and SANReN may feel confident that the network has been designed according to established principles and will someday rival the top REN networks of the world.

In the next chapter a closer look is taken at what is involved in designing a metropolitan fibre network, such as the one that is being constructed for the City of Cape Town.

Chapter 13

Case Study:

Designing a Metropolitan Fibre Network

Universities have a very long history of being involved with the development of new technologies, especially as far as networking and the Internet in general is concerned. Consider the valuable contributions that institutions like Berkeley have made to the development of the Internet. It stands to reason that the NRENs may have a lot to add to the development of the actual underlying infrastructure of the global networks as well.

Take as an example a large city like Cape Town. If the city builds a fibre network to benefit its citizens, some of the most important sites to be served by the network would have to be the large universities and research institutions, alongside major city and provincial sites such as hospitals, clinics, civic buildings and libraries.

When the TENET staff first heard of the City's intentions to construct a large fibre network, it met with the consulting company tasked with designing and implementing the network to make sure that all the major TENET sites would be well serviced by the City's fibre network. The TENET staff was informed that only a broad plan was in place, and that no actual network design of any kind had been completed. The TENET designer then offered to assist the consultation company in designing the network. This would ensure that all tertiary education and research institutions would be included on the City network, and ultimately, that SANReN would not have to purchase fibre IRUs from yet another party anywhere in Cape Town. The City in its turn was willing to allow TENET to co-design the network as servicing the university sites was always going to be one of its top priorities.

The design criteria for building the SANReN network were always going to be stringent, so it makes for a good case study to examine how an entire metropolitan fibre network should be designed in such a way that all major city and provincial sites are serviced sufficiently. It is a very good idea for a city to include the NREN in its planning process, as the NREN transmission network may well be responsible for the majority of traffic on the fibre network, which is very likely to be true in SANReN's case.

13.1 The City of Cape Town Optic Fibre Network

The City initially decided on building a city-wide optic fibre network to try and address the huge telecommunications costs it had to bear, which were in excess of R100 million per year by 2008. The CoCT[1] is an amalgamation of seven municipalities, in effect seven different cities of varying sizes. This means that the City was faced with a very complex legacy network of telecommunications services which ranged from simple direct telephone lines and dial-up Internet connections to PABXs with thousands of lines and multi-Mbps dedicated data lines, as well as any number of non-standard connections, including microwave. The upshot of it all was that the City had very little idea of which connections were required and which weren't. There were even some questions regarding the City being billed for services that were no longer in use or that had been decommissioned.

The City telecommunications department decided that the only long-term solution to this problem was to build a municipal fibre network, to which all City buildings would eventually be cut over. This would provide the opportunity of eliminating dead connections and dramatically decrease the cost of service and dependency on Telkom. On a City-owned fibre network there would be no recurring monthly cost to connect each City building. In fact, the operating costs could be reduced to a small fraction if the City only had to pay some usage-independent costs to operate the network, as well as usage of Internet bandwidth and voice calls.

The City is also primarily a service provider to its citizens, and such could play a very important role in economic development by creating an optic fibre network that would be accessible to any entrepreneur at a nominal monthly fee. The City adopted the viewpoint that the capital cost of the network would be recovered from the savings in telecommunications costs the network would bring, which means that it could make the fibre network available to end users at a very low cost indeed, much lower than any profit-seeking commercial company could afford.

13.1.1 A fibre network for the people

The problem with designing any city-wide fibre network is that it is very difficult to estimate the needs of the end customers, since they may cover the entire range of potential clients, all the way from very small local ISPs who only serve a very small number of customers to the very large ISPs and even telcos.

Broadly speaking, companies who require fibre connections may be sub-divided into two categories: small companies who are not very concerned with the service levels and may be content to lose service for a few days, as long as the fibre price is very low; and large companies, who don't really care about the price of the fibre, just as long as the service is first-rate and guaranteed.

The City took the view that the very large companies, such as the corporate ISPs and telcos, are not the target market of the City network, as they have enough resources to either build their own fibre networks or lease them from other telcos who are in the market of supplying fibre and services.

---

[1] City of Cape Town

This means that the City's chosen target market would be the small operators, companies who could never afford to build their own fibre networks, but who could see the advantages of operating their own fibre links, especially if it was offered at a very reasonable price.

13.1.2   Dark fibre vs. services

When building a fibre network, the first question must always be what kind of fibre network is to be built. Should the City become a full telco, providing ISP and telecoms services to all and sundry, or should it just build the fibre infrastructure and let the commercial companies provide the services? It could even choose to become a carrier-of-carriers, supplying only simple (typically Layer2) but fast services to other telcos and ISP, connections of the order of 1 Gbps and above.

Many other cities across the world have tried to follow the services model, with varying degrees of success. A closer look at one of these cities will be useful to understand the issues.

13.1.2.1   Industry example: Provo, Utah

The City of Provo in Utah, USA, set out to build a world-class metropolitan fibre network in July 2004. 18 months later is was the most extensive city-owned metropolitan fibre network in the entire USA. The municipal entity that the city created to build and manage the network is called iProvo. It makes a good case study because iProvo was very careful in designing and building their network, and yet it turned out to be something just short of a disaster.

The City allocated $39.5 million for the construction of the network, which was scheduled to take place over 24 months. At the end of the build period (July 2006) the entire network was constructed as planned and was built within the budget. As early as December 2005 fibre connections were available to more than half of the city's 27 000 residences and more than 4000 small businesses.

The network was intended to attract a commercial cable company who would manage the last mile connections, as well as the customer relations. It was anticipated that having a municipal fibre network would significantly lower the costs of broadband, as well as speed up the adoption rate and penetration.

iProvo started out well, having a successful build program and a customer base of almost 8000 by mid-2006. However, when one takes a closer look, the picture is less rosy. The network did not manage to improve the penetration of broadband much, as the vast majority of its clients were already broadband customers that simply migrated to iProvo. The price at which it offered its services was also very comparable to its commercial competitor's prices, belying its goal of providing a cheaper price than was already available.

To make matters worse, the network was unable to pay for itself, necessitating the City providing ongoing cash influx to enable iProvo to meet its costs. Not only does this defy the object of creating a network to lower broadband costs, the operating costs must eventually be borne by the taxpayer.

Even the plan of getting a commercial entity to run the access network went awry, as the single company iProvo managed to attract (HomeNet) went bankrupt after only one year, losing iProvo almost half the customers they started with in the process [57].

To add insult to injury, it became clear that iProvo sold its assets to HomeNet at 40% of their market value, which the taxpayer once again has to make up for. Eventually, the city decided to sell the entire network to a commercial provider called Broadweave Networks (May 2008).

The lesson to be learned from the mistakes iProvo made is that a municipality must be very careful when it tries to compete with established, commercial service providers. A city is usually geared for providing services which change very slowly, such as electricity, water or sewage. After all, once one of these services is installed, one would not expect it to change for years.

IT services is a very different story. The type of service demanded from the customer changes very fast and even the connection between the customer and provider will have to be upgraded frequently, and at very short notice. Commercial IT companies are very used to playing this kind of game, so when a city municipality announces its intention of competing with these companies, the kid gloves come off. As long as the city is expected to recoup its expenses it can never hope to compete with such a company, as the bureaucracy inherent in a municipality will always prevent it from providing services as quickly as a profit seeking entity.

Therefore, the only IT service that a city can safely and realistically offer is dark fibre. The "last mile" dark fibre connections may be considered to be the same kind of service as connecting that building to the electricity grid. Different service providers may provide service to the customer on the connected infrastructure, but the connection always remains the property of the city and will be leased to the service provider, not the end customer. In this manner the city will avoid competing with its own customers and competition among service providers may be encouraged, which benefits everyone.

13.1.3   An open access, dark fibre network

Ultimately, the examples offered by other cities convinced the CoCT to avoid the minefield that was the services market and opt for a pure open access, dark fibre network. The network will be designed and built as robustly as possible, be open to all companies with a valid telecoms I-ECNS[1] license, and be offered on an equal opportunity, regulated tariff model which would offer the same low price to all comers, regardless of size or importance.

Each company may request dedicated fibre links to be built, linking its end-user premises to the rest of the fibre network, from where it will pay the standard price per meter that all users pay. The capital cost of the last-mile connection will be recovered in part by an upfront installation cost, typically 30-40% of the total cost; the rest will be recovered from the monthly per-meter-per-pair-per-month price charged to the customer. If there are several customers in the same building, the installation costs may be shared among them.

The advantage that the City has is that the revenues from the network do not have to cover the deployment costs, which would be several hundred million Rands. As long as the revenues cover the last-mile installation costs and operating expenditures, the City management is content to rent out its "spare" network capacity to all comers, as its contribution to economic development.

---

[1] Individual Electronic Communications Network Services

13.1.4   Developing the local economy

The City commissioned an economic impact study from the University of Cape Town to examine the possible benefits of a large fibre network open to the public. Although the document has not been released into the public domain, it may be stated that the study found that the network could contribute as much as 3.5% to the City's GDP after 20 years, which may be in excess of *R200 billion*. It may also create in excess of 200 000 new job opportunities.

As much as these figures seem staggering, one should remember that a dark fibre network servicing previously under-serviced areas may be likened to the first deployment of an electricity grid. Before electricity, people in under-privileged areas did not have many job prospects in the area, but with the advent of an electricity grid, cottage industries such as welding shops and very light manufacturing appear everywhere, subject only to the local peoples' innovation and the laws of capitalism.

It may be argued that deploying a low-cost optic fibre network into such an area will have a similar effect. In today's market where fast, reliable switching and routing hardware is very cheap and easily available, all that stands in the way of the entrepreneur establishing a small, local fixed or wireless network is the access to a low-cost fibre network which may connect his network to the rest of the world.

Spurred on by the very optimistic results of the economic impact study, the City management felt confident that it could, with careful planning and management, create a fibre network that would make a considerable contribution to the economic welfare of its citizens, especially those in areas that have traditionally benefitted less from advances in telecoms and services.

13.2   Design Principles

Given the opportunity (and sufficient funding) any good network designer would attempt to build the best network possible. This means that the network should have ample spare capacity and that all possible levels of redundancy must be available at all points in the network. Only in extreme circumstances should any spurs be allowed. The default topology must always remain ring-based, if possible dual-homed (connected to two central network nodes).

Few service providers can afford dedicated links to all their clients, especially in a city as geographically large as Cape Town. Therefore, strategically placed interconnect points should be created where the service providers may consolidate and multiplex many short access links onto a backbone consisting of a small number of long, high capacity links which are shared by many end-users. To facilitate this requirement, the City has designed seven "switching centres" scattered across the city, which afford the opportunity to build multi-tiered services networks, geographically separated but interconnected to other networks as well as the Internet at large.

13.2.1 Levels of redundancy on a ringed network

Given that all access points are located on a fully redundant, dual homed ring, there are three levels of redundancy that may be extended onto the customer premises. Firstly one may build a simple, single connection from the FDP[1] into the CP[2] (Figure 65).
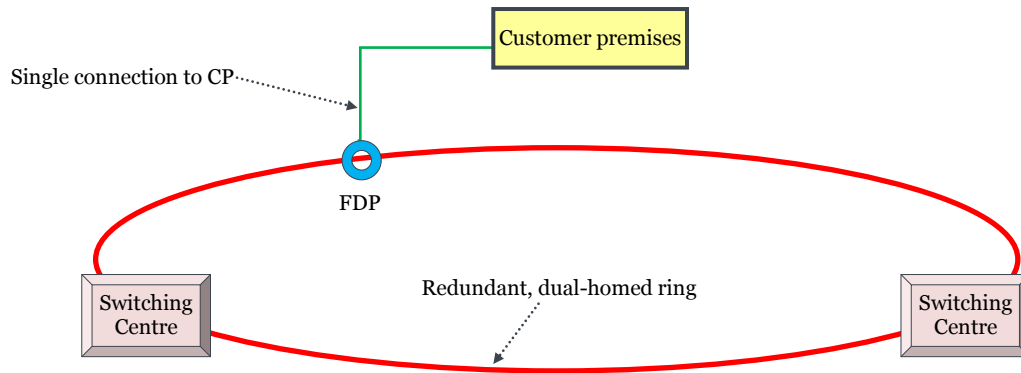
Figure 66: Single connection to CP

There is full redundancy available from the FDP to the rest of the network via the SCs, but no route redundancy between the FDP and the CP.

Secondly, one may build two diverse connections between the FDP and the CP (Figure 66).
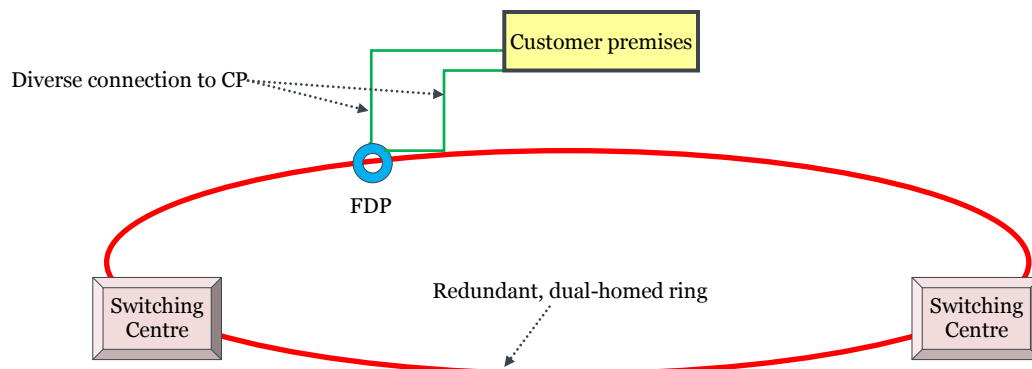
Figure 67: Redundant connections from FDP to CP

Such a design may seem to be fully redundant but it is clear from the diagram that the FDP is the single point of failure. It is unlikely that the FDP will be damaged or cut off, however it does expose the client to the risk of a fibre splicer making a mistake and cutting off both sides of the redundancy at the same time while installing other services.

---

[1] Fibre Distribution Point
[2] Customer Premises

Thirdly, one may install two fully redundant links from different FDPs to the CP (Figure 67).

The difficulty in designing a network that allows for all three types of FDP-CP redundancy is that one must allocate the fibre pairs in the FDP carefully. First of all, one must limit the risk of damaging pairs that are in use. This is done by only making certain fibre pairs available in any given FDP. This means that when a site has the potential of requesting a fully redundant solution, the same pairs must be available in both the FDPs illustrated in Figure 67. Creating this type of allocation throughout an entire network is not a trivial exercise.
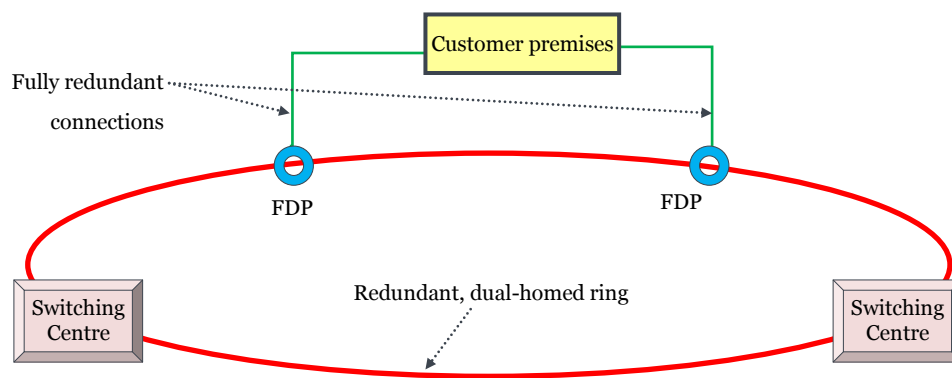

Figure 68: Fully redundant connections to CP

The only way of ensuring that non-allocated fibre pairs cannot accidentally be damaged during new installations is to divert the allocated fibre pairs into a second dome splice enclosure.
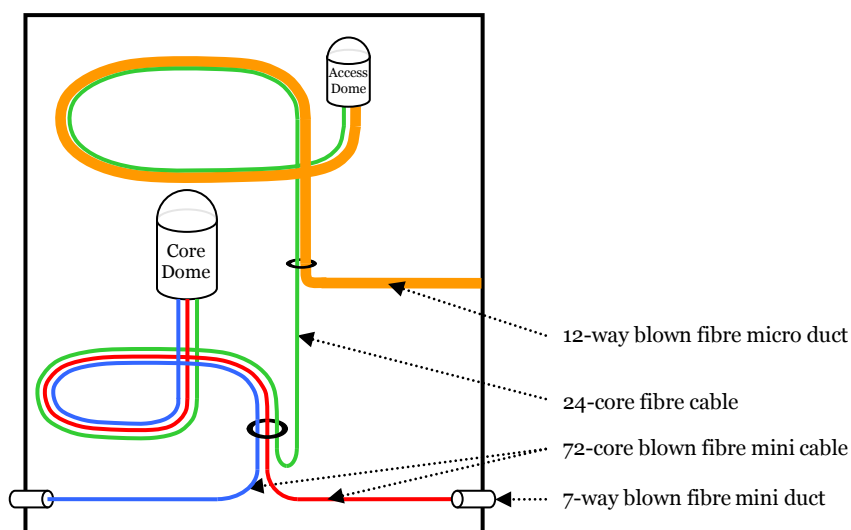

Figure 69: The dual-dome installation

As can be seen in Figure 68, the FDP consists of two dome splice enclosures. The first one (core dome) contains only 24 splices, as a single fibre tube (12 fibre strands) of the main (72) cable is diverted by cutting the tube and splicing the 24 resulting strands (12 East, 12 West) onto a new 24-strand cable. During the initial construction, when there are no customer connections yet, the two tubes in the second dome are simply spliced to each other: fibre 1 is spliced to fibre 13, 2 to 14 etc. The original tube of 12 fibres in the core dome has now been diverted into the second (access) dome, but in such a way that the fibre integrity has been restored, and each fibre is still continuous from the originating SC to the terminating SC.

Each core dome would only contain the 24 splices necessary to connect the one tube to the 24-cable leading to the access dome. When the manhole containing the FDP corresponds to the end of the cable drum, in which case the entire cable must be spliced to a new drum, an extra set of five splice cassettes is added into the core dome to accommodate the additional splices (with a 72-strand cable this equates to 60 splices).

Now, when the first case of redundancy (single connection) is to be employed to connect a new CP to the FDP (Figure 65), a small cable (4 fibre strands) may be installed between the FDP and CPE. One of the splice pairs in the access dome is now cut off, producing two fibre pairs. Pair 1 is now connected to pair 1 of the CP cable, pair 2 connected to pair 2. The CP now has full use of the network redundancy between the FDP and the SC, even though the cable between the FDP and the CP is not redundant. In the case where the access cable is very short, but the distance to the SC is very long, this still results in a very high level of redundancy.

The second case may also easily be realised, with a cable containing as little as 1 pair of fibres. Each end of the cable ring to the CP is now simply spliced onto the two pairs available in the access dome and full cable redundancy is achieved.

When full network redundancy is needed (Figure 67), the East pair of the four fibres available is spliced to the East access cable in FDP1. The same is done in the FDP2, except that this time the West pair is used. This would mean that on the main cable between FDPs (red ring in Figure 67) now has one dead pair between FDP1 and FDP2. This pair has now been diverted via the CP and the total cable capacity remains the same.

13.2.2   Switching centres

The central interconnect nodes are full co-hosting facilities designed to host the multiplexing equipment of the client ISP/carriers. In keeping with the City philosophy of building a network for those companies that cannot afford their own nodes, or even the commercial charges of regular co-hosting facilities, the City SC services will be offered at the lowest price possible. However, the City will not compromise on service, so the SCs will be built and managed according to the standard Tier2[1] data centre standard.

They will therefore have false floors, redundant air-conditioning, fully redundant power with dual

---

[1] TIA (Telecommunications Industry Association) - 942

(A+B) power feeds, as well as a full load UPS[1] per feed and a standby diesel generator. The centres will also have biometric entry control, ccTV[2] security and a triple-knock fire detection system with gas fire suppression. All aspects of the SCs will be monitored in real time by a building management system, which in turn is monitored by the NOC[3].

To ensure full optic fibre redundancy, each SC will have two separate ODFs[4], one on each end of the room, as well as redundant patch frame routes (fibre raceways) between each rack and the two ODFs. This will enable a fibre network redundancy of up to Four 9s (99.99% availability)[5].

In addition, the City will provide the customer racks, according to a standard specification, which will be available in three flavours: a half-rack at 1kW, a full rack at 1.5kW and a full rack at 6kW. The City will rent the rack space to users of the fibre network at a flat monthly fee, electricity included, at a much lower rate than is available commercially.

13.2.3   Blown fibre vs. CST cable

A fundamental question that must be answered during the design phase is what kind of cable must be used in the outside plant. Two main options are available in a metro environment: rodent proofed cable (see 4.2.3.2) or blown fibre (4.2.3.4). The issue hinges on a trade off between risk and capacity. The CST cable is installed in an empty 40mm sub-duct and has a maximum capacity of 144 strands[6], but has the advantage of being rodent proof. Mini cables are installed in a 40mm, seven-way mini duct, which has a maximum capacity of 672 fibre strands[7], but is not rodent proof. Clearly, in cases where there is a significant risk of rodents damaging the cables the CST cables should be used. On the other hand, where the risk of rodent damage is low, it makes sense to opt for the higher capacity mini duct option.

The CoCT decided to standardise on the non-rodent proof mini cable system, with the proviso of retrofitting the network with rodent-proofed mini-cables where necessary.

13.2.3.1   Blown fibre in the Access network

The issue arises again when the Access network is considered. The traditional way of building an Access fibre network is to build a new pipe-and-chamber system into each and every building that has to be connected to the network. This is both very expensive and very disruptive to the road and pavement infrastructure outside the buildings in question. A much cheaper and less disruptive method is available in blown micro fibre (4.2.3.5).

---

[1] Uninterruptable Power Supply
[2] Closed circuit television
[3] Network Operations Centre
[4] Optical fibre Distribution Frame
[5] 99.9 is achievable with simple redundancy, whereas 99.99 requires full redundancy everywhere [42]
[6] without resorting to fibre ribbons, which are much more difficult to deploy and maintain
[7] 7 cables with 96 strands each
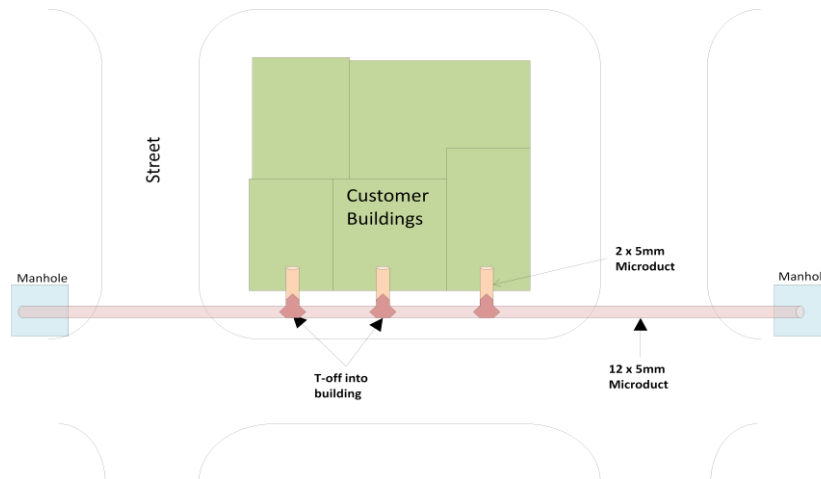
Figure 70: Accessing buildings with micro ducts

Figure 69 illustrates the basic method adopted by the CoCT. A 32mm, 12-way micro duct is directly buried in the pavement outside the building using a slot-cutter (Figure 70).

Figure 71: Constructing a micro duct system

Each building now has a single, 2-way microduct installed into its basement from the T-off junction on the main 12-way micro duct. Inside each T-shaped branching unit a single micro tube is cut and diverted into the adjacent building by connecting the 2-way branch tubes with air-tight connectors (Figure 71).

Figure 72: Detail of branching units

122

In this way, each building gets a dedicated, redundant connection to the rest of the fibre network. During the initial build, only city blocks that contain at least one building that must be connected have micro ducts installed, but once a given city block has micro duct installed, each and every building that the micro duct passes, has a 2-way connection installed, whether anyone ha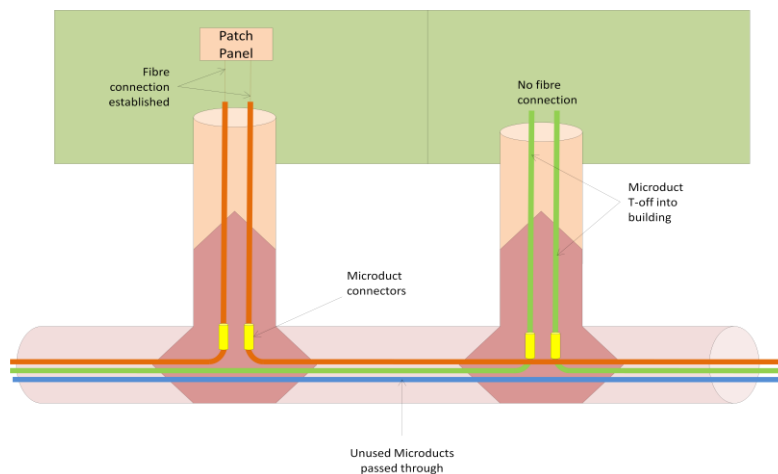s requested for the building to be connected or not. This ensures ease of connection at any time in the future with no further civil work required.

13.2.4   A three-tiered network

A complex fibre network should be partitioned into tiers so that it is always clear what purpose a given cable has and there is no possibility of misusing a cable, whether by accident or design.

It was decided that the CoCT network would consist of three tiers: Core, Local and Access. The Core layer would contain cables whose purpose is to carry traffic between SCs, and would therefore have no FDPs along the route. When it becomes necessary to splice the entire cable (when the end of the cable drum is reached), a single core dome housing all 72 of the splices on the cable, will be installed in a manhole. Each core cable will be dual homed, i.e. will originate and terminate at different SCs.

The Local layer will also be dual homed but each cable will contain many FDPs. Each Core cable will be accompanied by a Local cable, so that no segment that the Core layer passes through is without a means of access.

Access cables are any cables that do not conform to the definition of either the Core layer or the Local layer. Therefore, an access cable is one which: 1) is in a ring but originates and terminates at the same SC; 2) originates at a SC but terminates at any point/node which is not also a SC; 3) originates on a local cable and terminates either on another cable or in a CP.

Building a network strictly according to these definitions will mean that a stringent method is enforced to connect all points on the network. Without such a stringent enforcement it may be easy to cut the network up into bits and pieces which can no longer be interconnected in rings. This will invalidate the entire redundancy philosophy which has been implemented.

Tiering a network also introduces a system of priorities which is applied to the network maintenance. For example, one may assume that any given live fibre on a Core cable carries much more traffic than in any other layer, and must therefore receive priority in any emergency maintenance event. A similar principle holds for Local cables compared to Access cables.

Figure 73: The CoCT core network

Figure 74: The CoCT local network

Figure 75: Detail of the CoCT Local network: CBD

13.3   Co-Building of Infrastructure

The most expensive aspect of building an optic fibre network especially inside a city, is the civil infrastructure, known as the OSP[1], or the pipe-and-chamber network. Under certain conditions the OSP infrastructure may cost as much as R1000/m. When one compares the price of a 72-strand blown fibre cable at R25/m, it becomes clear that the cost of the cable is but a small fraction of the cost of trenching and re-instating. Certain areas will inflict a premium cost, in particular a CBD area where the construction contractor has to take out insurance against possible damage caused to existing infrastructure during construction, as well as pay an indemnity to the city to cover any damages to vehicles etc. that may be affected. When presented with such costs, it is wise to look at the existing OSP networks and try and purchase parts of them, rather than building everything from scratch.

Two commercial companies operating in Cape Town (Neotel and DFA) have infrastructure for sale that the City was interested in buying on a co-build basis. Most of what DFA has for sale is in the area outside the Cape Town CBD, whereas the Neotel infrastructure is mostly inside the CBD.

13.3.1   Dark Fibre Africa

DFA approached the CoCT in early 2008 to see how their proposed new fibre network in the city could be integrated with the City's, with the view to share civil construction costs. When it became clear that there was a greater than 60% overlap in desired routes, DFA suggested a combined plan that would

---

[1] Outside Plant

serve both parties' needs. Both network plans were adjusted accordingly with the result that the overlap was increased to around 80%.

DFA would now take on the role of primary route construction contractor and in effect, build the City network along with its own. The standard that DFA follows is to use mechanical trenching in the road surface, as opposed to building a duct route in the sidewalk or road reserve using manual labour. The mechanical trenching can produce as much as 1km of trenching a day, depending on the terrain and number of services to be crossed. This is up to five times as fast as hand trenching.

DFA constructs routes by cutting a 100mm wide trench in the road surface, 450mm deep. Into the trench is inserted a bundle of 6 ducts in 2x3 configuration. The top two layers contain one empty 50mm sub duct as well as one 40mm 7-way mini duct for blown fibre. The bottom layer consists of two 7-way mini ducts which will be reserved for the use of the City (Figure 75).

After the installation of the ducts the trench is filled up to 100mm below the road surface with cement slurry, after which it is filled in with bitumen, once the slurry has set. Thus the six ducts are safely embedded in the road surface, with very little risk of duct damage.

DFA does not build manholes for the CoCT, which means that the City has to install its own new manholes on top of the installed duct infrastructure.
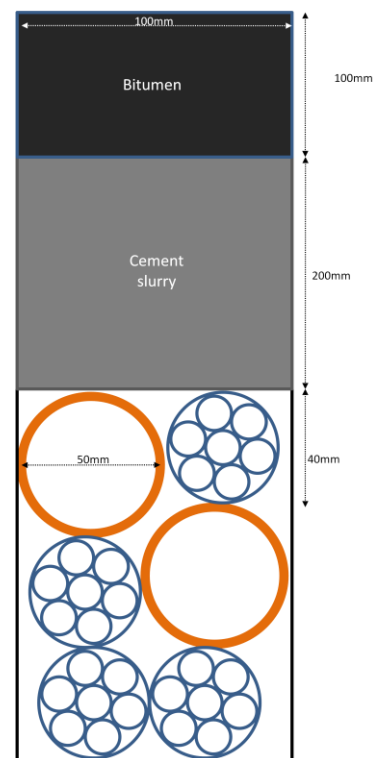
Figure 76: Trench for mini ducts

13.3.1.1   *Competing fibre networks*

It may sound like a bad idea for the City to allow a for-profit company like DFA to co-build its fibre network. After all, once both networks are operational, they will be competing for the same market with the same product, i.e. dark fibre. None of the other service providers offer dark fibre as a product, which means that the entire market segment for dark fibre is now divided between two providers who have cables along exactly the same routes. There is a lot of evidence to suggest that a government entity such as a city can never successfully compete with a profit-seeking company such as a telco.

The solution to this dilemma is for the City to target a different market segment with its dark fibre product than the commercial company would be interested in. End customers who may be interested in renting dark fibre fall into two broad companies: at the high end there are companies who service large corporate companies and other telcos or ISP. These customers are prepared to pay a premium for good service and will insist on redundancy and very high service levels. This makes them ideal targets for the profit-seeking dark fibre provider. At the other end of the scale there are the small ISPs

and other small companies who desire a dark fibre connection, but are prepared to accept a much lower level of service and longer MTTR[1]s, as long as the price is very low. When one considers that one of the main purposes of the City network is to service under-serviced areas and foster economic development, it becomes clear that the City will target the low-end users, offering very low prices and reasonable service levels. This sharp contrast between the two types of user should keep cut-throat competition for the same target market down to a minimum.

13.3.2   Neotel

Neotel has created a very extensive duct infrastructure inside the Cape Town CBD. Most of the routes consist of four 110mm uPVC[2] ducts buried in the sidewalks along most of the major streets in the CBD. Many of the minor streets only have two ducts. Neotel has expressed its willingness to sell some of these ducts to the CoCT, according to the principle of selling one 110m duct along each route where four are available, or one 40mm sub duct in each route where only two ducts are available.

Once again, the City has to build its own manholes on top of the existing infrastructure, after which sub ducts, mini ducts and optic fibre cables may be installed. This will once again result in two logically and physically separate fibre networks which only share the same trench.

13.4   Aligning the CoCT Network with SANReN

SANReN has received very good cooperation from the CoCT, to the effect that nearly all of the 20 or so TENET sites are within a city block of the fibre network. Full diversity is available to each and every site and SANReN will now be able to build the complete network that it has planned (Figure 57). A contract of purchase was drawn up to sell SANReN a total 265km of dark fibre pairs on a 20-year IRU basis. The six sites in Cape Town that are on the SANReN priority list[3] were also included in the build list. Links to all the other non-SANReN sites would be permitted under the same contract, i.e. any other TENET sites could ask SANReN to request a link from the City. This link would be built by the City under the same contract, but the institution would have to refund the City for the expense. As mentioned before, SANReN would make no distinction between its sites and such TENET sites, so all sites would be considered equal peers.

---

[1] Mean Time To Repair
[2] unplasticised Polyvinyl Chloride
[3] UCT main, UWC main, MRC, MCM, SAAO and iThemba

Chapter 14

Results and Conclusions

The actual build of the SANReN has been fraught with difficulties. Between contractual complications, recalcitrant municipalities and infighting in the ranks of the CSIR, the deployment of the network has been set back by 18-24 months, perhaps even longer. Aside from finalising the planning and establishing good relationships with municipalities and service providers, only a few of the most crucial elements of the build plan have been realised to date.

14.1 The National Backbone

As mentioned in section 9.2, the CSIR has established a pair of Telkom 10Gbps managed services between the CSIR main campus and SAC at Hartebeespoort, as well as between SAC and Wits main campus. The CSIR has been unable to negotiate a dark fibre on any of the provider's networks that could connect the SANReN section in Johannesburg with the one in Pretoria.

After much speculation that Infraco, as another project of the DST, would be able to provide SANReN with a 10Gbps wavelength in a ring around the country connecting all the large cities, it has turned out to be a false hope. Infraco has no ECNS license and can therefore not deliver services to any other company or organisation except Neotel, for which it provides the entire national backbone.

SANReN was forced to release a RFQ to ask for pricing to build a national ring and in the end Telkom was awarded a contract for supplying a 10-year IRU on a national ring for one 10 Gbps managed wavelength.
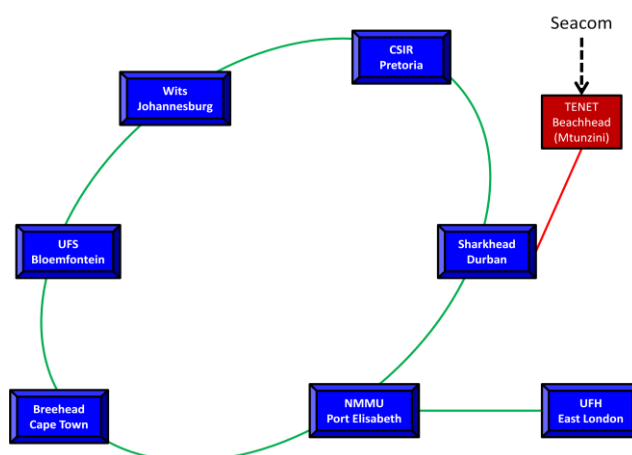

Figure 77: The Telkom-SANReN national backbone

Figure 77 illustrates the backbone that Telkom offered in response to the RFQ. Note that none of the smaller cities like Grahamstown were included on the ring and that even East London is set to be connected on a spur. The proposed date for the completion of the ring is December 2009, with the East London spur only being completed in July 2010.

As the diagram also shows, Telkom has elected to deliver the services directly to the seven SANReN core nodes, which should make a future upgrade to another wavelength somewhat easier.

14.2   The Seacom Link

TENET organised the purchase of the 10Gbps Seacom circuit, so the CSIR was unwilling to get involved in connecting the site at Mtunzini to the SANReN node in Durban. TENET managed to strike a deal with Dark Fibre Africa for the purchase of 12 fibre pairs on a new cable between Durban and Mtunzini, on a 20-year IRU basis. This cable was funded on a co-build project with other interested parties, so TENET was forced to buy a portion of the cable, rather than a single fibre pair or two.

According to the purchase agreement with Seacom, TENET was entitled to establish a node in the Seacom landing station. As the nearest national node is in Durban, it was decided not to deploy a router at the Mtunzini node (Beachhead) but only a DWDM terminal. For standardisation reasons, TENET decided to buy the same equipment as SANReN is using (Cisco ONS).

It became apparent during planning that the actual fibre distance between the Sharkhead node at DUT and the Beachhead is marginally too far (180km) to reach with a single hop, so it was decided to insert an amplifier node in Umhlanga, which is about 30km north of the Sharkhead. The main node of IS in Durban is also there, so it made good sense to co-locate the node in the IS data centre. There is a possibility of an Internet exchange developing in Durban and it is likely that the IS node will be one of the locations in a distributed DINX[1], so in future the TENET amplifier may be upgraded to a full add-drop multiplexer.

The Seacom circuit is not Ethernet based, as TENET would have preferred, but a SDH STM-64. This necessitated TENET installing a very expensive SDH POS[2] blade into the Sharkhead router to translate the SDH into Ethernet.

The Seacom circuit was finally declared ready for service on 23 July 2009, and TENET had managed to have both the fibre link and the ONS equipment installed, so the circuit was delivered to the Sharkhead node a few days later. This gave DUT full access to a 10Gbps international link. Only when the SANReN national network is completed will the Seacom bandwidth be delivered to the institutions in the quantities they ordered.

In the interim, TENET has made arrangements to deliver cheaper bandwidth to the rest of its sites. IS has purchased two STM-16 circuits on Seacom and was prepared to deliver some bandwidth to all the non-GEN3 sites of the TENET family at around 36% of the cost per Mbps that TENET had previously paid for bandwidth on the SAT-3 cable. These sites include all the small ADSL sites on the IS cloud,

---

[1] Durban Internet Exchange
[2] Packet Over SDH

the two foreign universities (Lesotho and Swaziland), as well as all the sites already on the SANReN network in Johannesburg (UJ, Wits and CSIR).

TENET also asked Neotel to commission a temporary 400 Mbps link between the Neotel GEN3 MPLS cloud and the Sharkhead, which would enable TENET to feed at least some of the bandwidth on the Seacom circuit to the GEN3 sites. On average, those sites could now afford three times as much international bandwidth as before, for the same price.

Once the SANReN backbone connects to the Sharkhead and to a TENET site, that site could expect to pay between 3.8 and 5% of the cost of SAT-3 bandwidth per Mbps, depending on whether that institution co-funded the Seacom circuit or not. This price would only stand for the first six years, after which it would collapse, once the loan has been repaid.

14.3   The Johannesburg Network

In January 2008 TENET started to migrate its sites from the Telkom GEN2 network onto the Neotel GEN3 network. SANReN had an aggressive rollout plan at the time and undertook to commission the first metropolitan network in Johannesburg in time for the GEN2 cut-off date, which was on 31 March 2008. Accordingly, the CSIR obtained a 3-pair dark fibre ring around Johannesburg that included all the sites on the original design, as discussed in section 10.1.2. These included one site of the CSIR, three sites of UJ and four sites of Wits. The ring was procured from Neotel on a 10-year IRU, the first such deal that Neotel ever made.

The first sections of the ring were completed almost two weeks before the deadline; allowing SANReN to get some equipment installed and configured some days before final cut-off of the site from the GEN2 network.

Unfortunately, the DWDM equipment had not yet been delivered to SANReN, so an interim plan was made. The large sites (UJ Doornfortein and Kingsway, as well as Wits Main) had SANReN 7609s installed to connect them to the Reefhead, and the other sites were kept on the existing institution LANs, as none of them were TENET sites and as such were not in danger of being cut off.

Once the new batch of equipment arrived a second round of installation was performed. During this phase (end of June 2008) DWDM equipment was installed at the main sites of UJ and Wits, as well as the full four-degree node at the Reefhead. All the other sites mentioned were connected to the fibre ring on 4924 switches. The initial ring was configured on the first pair of the ring, leaving the second pair open for the DWDM system and the third pair for the grey ring. Once all the new nodes were configured and synched, the interim ring could be torn down and the equipment removed.

The two 7609s at Doornfontein and UJ Main were removed for redeployment, but of course the one at Wits stayed, as it now became the backup PE router of the Johannesburg ring. This decommissioning of the temporary ring should have been very simple, but in fact all manner of problems became apparent on the physical fibre ring, with the result that the DWDM ring was only taken into production more than six months later, and the equipment installation has not been signed off by SANReN to this day, more than 14 months later. The current configuration of the Reefhead node is exactly as illustrated in Figure 52, except that the CRS-1 has not been taken into service yet.

14.4   The Pretoria Network

Despite assurances from the CSIR that they were on a very good footing with the Tshwane Municipality, a contract for the use of their dark fibre network has not materialised. As no other dark fibre network is available in the city, SANReN will have to wait until the contract has been finalised. Only then can it start to install equipment and build the network.

To make matters worse, serious infighting has taken place between the Meraka institute and the IT department of its parent organisation (the CSIR) to the effect that the CSIR refused to be involved in the SANReN network at all, much less host its Pretoria core node. This struggle has gone on for more than a year, but fortunately it seems to have been settled now, and the TENET community can look forward to the building of the Pretoria ring as planned.

Only the link to Johannesburg via SAC has been commissioned, as well as a short on-campus fibre link connecting Meraka and the NRF to the SANReN node.

14.5   The Cape Town Network

The City of Cape Town has had innumerable delays in starting to build their metro fibre network, but by all accounts the relevant tenders have been finalised and construction will start on 1 October 2009. The entire phase 1 of the build plan will be ready for service by the end of April 2010, which will include the construction of five switching centres and 171km of duct runs, which will be used to build a fibre network containing 485km of cable.

SANReN and the City could not come to a satisfactory arrangement for a long time, and even though both parties have now agreed to a final version of the IRU purchase contract, it has not yet been signed. This does not affect the City build plan though, so confidence remains high that SANReN will take possession of its designated fibre pairs by April 2010.

The network design remains largely the same, except that the SANReN core node will now be located at the TENET Breehead, which is in the Cape Town CBD.

14.6   The Durban Network

Much has been said about who should do what between TENET and SANReN in Durban, with the metro plan being redesigned several times. To date, no SANReN equipment has been installed in Durban and no fibre IRUs have been purchased. The only sign of the network is the node that TENET has commissioned to receive the DWDM circuit from Mtunzini (Sharkhead), which consists of an ONS terminal with two wavelengths (one for Uzulu and one for Seacom) and the receiving TENET router with a SDH blade. As soon as the national backbone is dropped at the Sharkhead node, it will be interconnected to the TENET router so that the Seacom bandwidth may be distributed nationally.

SANReN has made a commitment to commission a dark fibre network (probably on the DFA network) which will connect all the other sites mentioned in section 10.4, but this has yet to transpire. In all likeliness, the CRS-1 core router will be moved to the CSIR site and be connected to the Sharkhead on a dark fibre ring.

14.7 The Equipment

Mention should be made that, unlike what is usually done during network deployment, SANReN opted for upfront buying of almost all the equipment scheduled on the build plan. The list includes 23 complete ONS terminals, 30 Ethernet switches (Cisco 4924), eight 7609s PE routers and 4 CRS-1 core routers. Ordering all the equipment in one go sounded like a good idea initially, especially when one considers the lead times on the manufacture and delivery of some of the devices, but consider the fact that one always takes a risk storing advanced equipment like this for a long time (it has been more than a year). Added to that is the risk that new versions of both the hardware and software may be updated over such a timeframe, which means that if the order was placed one year later, new and better features may be available that would now be missed out on.

On the other hand, it may be argued that the bureaucratic delays and infighting among the various departments of the CSIR may have delayed the ordering of equipment so long that SANReN would still have had nothing to show, two and a half years after the project was initiated. So perhaps having more than eight tons of routing and transmission equipment in long-term storage is not the worst that could happen.

14.8 Summary of Contributions

The core design team of the SANReN consisted of only two people: the author and the CTO of TENET, Andrew Alston. The author was responsible for all aspects of the optical network design, whereas Mr. Alston was responsible for the Layer 3 design, as well as all facets of the integration between the SANReN and GEN3 network, which he also designed in its entirety. Of course, many technical experts from Cisco were involved in the configuration and commissioning of the network, and their contributions are gratefully acknowledged.

14.9 Conclusions

It cannot be doubted that the benefits of having a high-speed, high availability network for the exclusive use of the tertiary education and research community will be of incalculable value. South Africa is on the verge of becoming a serious contender in the world-wide research community, especially in the fields of radio- and optical astronomy, but these projects will be very dependent on the local research community having access to very high bandwidth, both locally and internationally.

TENET, as the de facto South African NREN, has done a remarkable job of lowering the cost of bandwidth and services to the community, and at the same time increasing the number of services available. SANReN, on the other hand, has made very big promises and have so far delivered very little in terms of either services or connectivity. Many reasons may be given for why this is so but one has to wonder what would have happened if the DST had simply given the task of deploying the SANReN to the community itself, i.e. TENET, rather than trying to magnanimously build a new NREN with little regard for the existing one's priorities and relationships with the community.

It is the opinion of the author that, as much as it is commendable for the national government to try and boost research networking, it may have been a mistake to task a single research entity that has its own agenda, to build a nationwide production network. TENET has for the best part of a decade

shown that is more than willing and able to build such a network. Perhaps in the near future this may become obvious to everyone.

14.10   Future Work

Much still has to be done where the deployment of the SANReN is concerned. No hardware has been commissioned in any of the cities except Johannesburg, and even there the optical network is still short of some components, not least of which is the core router. The CSIR keeps insisting that the signing of dark fibre IRUs with the Cities of Tshwane and Cape Town are imminent, but they have been saying that for at least 18 months.

The contract that has been signed between the CSIR and Telkom for the national backbone is a major step in the right direction, but it leaves many sites out of the original design and some plan will still have to be made to connect them. TENET has revised its position in the relationship with SANReN/CSIR to take a much more active role. In association with some of the more forward-thinking institutions, plans are being drawn up for many additions and extensions to the national network that fall outside the purview of SANReN itself.

It has become clear that it will be many years before the TENET network is fully converged with the SANReN, but perhaps all the role players can set aside their differences and allow the construction of a new unified network; one which will be greater than either TENET or SANReN can envisage at present.

Appendix – Photos of the SANReN equipment

## One of the four WSX shelves in the Reefhead
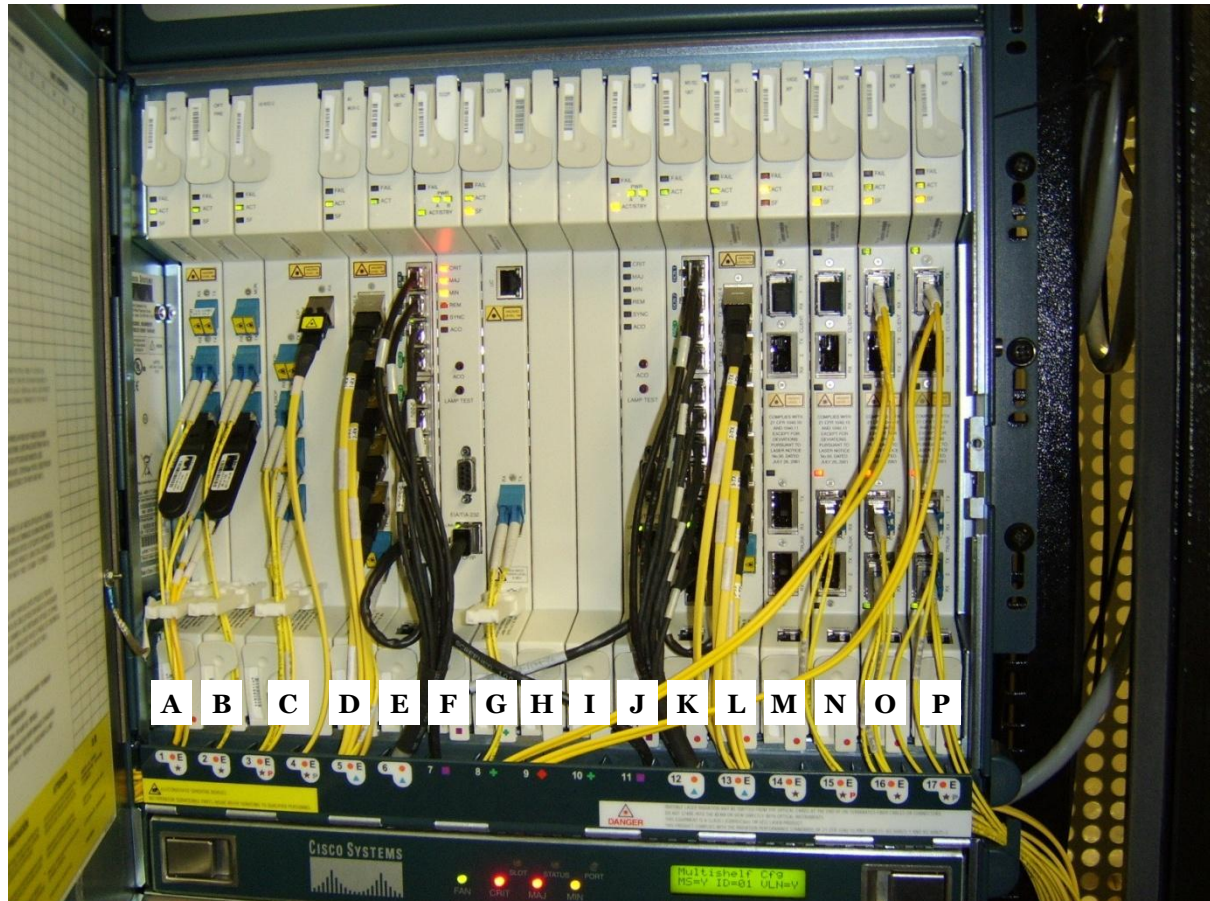
- A. Booster amp (line to network)
- B. Pre-amp (amplifies incoming signal)
- C. Wavelength selective switch (passes or drops channels)
- D. Demultiplexer (demultiplexes signal into 32 channels)
- E. Multishelf management (connects to other shelves)
- F. Timing and control (primary)
- G. Optical supervisory channel (adds comms for remote node)
- H. Blank slot for transponder
- I. Blank slot for transponder
- J. Timing and control (secondary)
- K. Multishelf management
- L. Multiplexer
- M. – P. Muxponder (grey ports at the top and coloured at the bottom)

The Reefhead fibre shelves

The provider edge routers at Reefhead and Wits
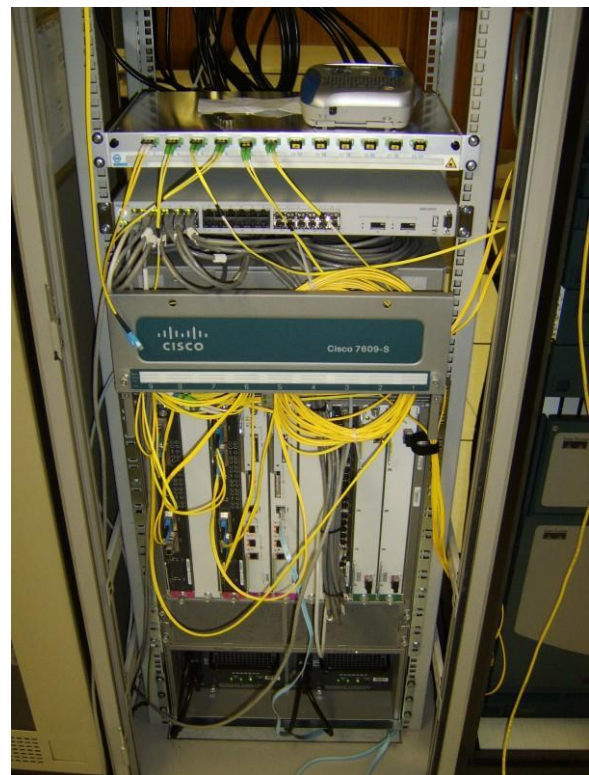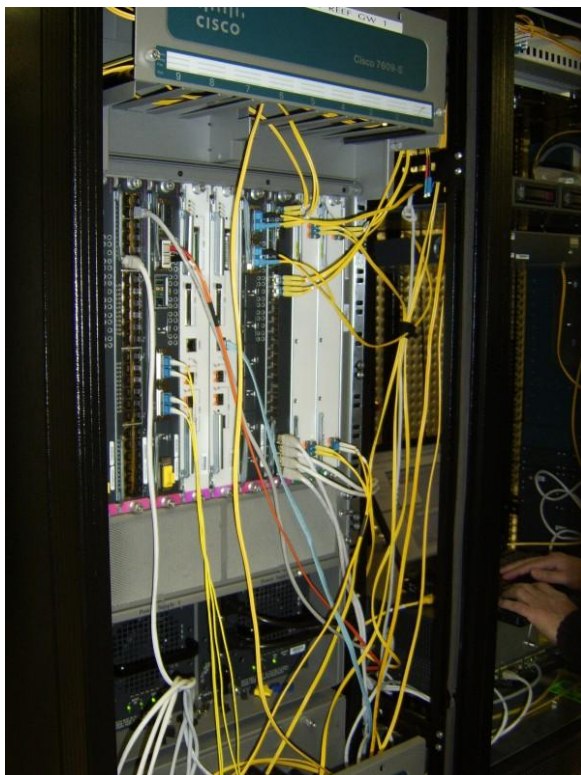
Bibliography

1. **Coghlan, David.** *Doing action research in your own organisation.* London : Sage Publications, 2005.

2. **Square Kilometer Array SA.** The SKA project - global perspective. *http://www.ska.ac.za.* [Online] [Cited: September 6, 2009.] http://www.ska.ac.za/bid/inafrica.php.

3. **CSIR.** South African Research Network. *http://www.meraka.org.za.* [Online] [Cited: September 6, 2009.] http://www.meraka.org.za/sanren.htm.

4. **Martin, Duncan.** *TENET Annual Report 2008.* Cape Town : TENET, 2008.

5. **Louisiana Optical Network Initiative.** *LONI Delivering on Promise of Attracting Federal Research Funding.* New Orleans : Louisiana Optical Network Initiative, 2005.

6. **Board of Regents.** *Louisiana to Construct One of Nation's Most Powerful Supercomputing.* New Orleans : BOARD OF REGENTS, 2006.

7. **Erlanger, Leon.** Distributed Computing: An Introduction. *www.extremetech.com.* [Online] April 4, 2002. [Cited: August 14, 2008.] http://www.extremetech.com/article2/0%2C2845%2C11769%2C00.asp.

8. **Martin, Duncan.** *TENET Annual Report 2005.* Cape Town : TENET, 2005.

9. *Internet2 Plans an Improved Version of Its Academic High-Speed Network.* **Chronicle of Higher Education.** 35, Chicago : Chronicle of Higher Education, 2006, Vol. 52.

10. *RESEARCH ON THE RAILS.* **Hohlhepp, Robert J.** 7, New York : Network Computing, 2006, Vol. 17.

11. *The big Internet merger is back on.* **Read, Brock.** 30, Washington DC : Chronicle of Higher Education, 2007, Vol. 53.

12. *Interview with Peter Kent.* **Allen, Keri.** 15, New York : Engineering & Technology, 2008, Vol. 3.

13. **Pale, Predrag.** Nation-Wide ICT Infrastructure Introduction and its Leverage for Overall Development. [book auth.] Mehdi Khosrow-Pour. *Cases on Telecommunications and Networking.* New Delhi : IGI Global, 2006.

14. *DANTE: Global collaboration demonstrated.* **Szomoru, Arpad.** London : M2 Presswire, 2008. TERENA Networking Conference.

15. **UbuntuNet Alliance.** *What is UbuntuNet?* Kigali : UbuntuNet Alliance, 2008.

16. **Martindale, Helen.** *UbuntuNet Alliance: Global Research Community Bridges Digital Divide Between Africa And Europe Through High Speed Network Link.* Mombasa : M2 Presswire, 2009.

17. **Martin, Duncan.** *TENET Annual Report 2001.* Cape Town : TENET, 2001.

18. —. *TENET Annual Report 2002.* Cape Town : TENET, 2002.

19. —. *TENET Annual Report 2004.* Cape Town : TENET, 2004.

20. —. *TENET Annual Report 2007.* Cape Town : TENET, 2007.

21. **CSIR.** National research cyberinfrastructure goes live in Gauteng. *CSIR eNews.* Monthly, 2008, Vol. 8, May 2008.

22. **Hecht, Jeff.** Fibre Attenuation - Developer Zone - National Instruments. *National Instruments.* [Online] [Cited: July 15, 2008.] http://zone.ni.com/devzone/cda/ph/p/id/88.

23. **Cord.** MODULE 8 LASER/FIBER-OPTIC COMMUNICATION SYSTEMS. *cord.org.* [Online] [Cited: June 15, 2008.] http://cord.org/cm/leot/Module8/module8.htm.

24. **Okamoto, Katsunari.** *Fundamentals of Optical Waveguides.* New York : Academic Press, 2006.

25. **Wikipedia.** Nyquist–Shannon sampling theorem. *www.Wikipedia.org.* [Online] [Cited: June 15, 2008.] http://en.wikipedia.org/wiki/Nyquist%E2%80%93Shannon_sampling_theorem.

26. **Mr Fiber.** Dispersion Compensation. *www.mrfiber.com.* [Online] 2005. [Cited: June 16, 2008.] www.mrfiber.com/Dispersion_Compensation.htm.

27. **Bennett, Charles A.** *Principles of Physical Optics.* Hoboken : John Wiley & Sons , 2008.

28. **ITU-T.** *ITU-T Recommendation G.652.* [Document] Geneva : ITU-T, 2005. G.652.

29. —. *ITU-T Recommendation G.655.* [Document] Geneva : ITU-T, 2006. G.655.

30. **Perros, Harry G.** *Connection-Oriented Networks: SONET/SDH, ATM, MPLS and Optical Networks.* New York : John Wiley & Sons, 2005.

31. **Wikipedia.** Synchronous Digital Hierarchy. *www.wikipedia.org.* [Online] [Cited: February 2, 2009.] http://en.wikipedia.org/wiki/Synchronous_Digital_Hierarchy.

32. **Simmons, Jane M.** *Optical Network Design and Planning.* New York : Springer, 2008.

33. **Kazi, Khurram.** *Optical Networking Standards: A Comprehensive Guide for Professionals.* New York : Springer, 2006.

34. **Coover, Edwin R.** *ATM Switches.* Boston : Artech House, 1997.

35. **Microsoft.** Asynchronous Transfer Mode. *http://technet.microsoft.com*. [Online] [Cited: February 2, 2009.] http://technet.microsoft.com/en-us/library/bb726929.aspx.

36. **UIC.** Building the Data Highway. *http://www.uic.edu*. [Online] [Cited: February 2, 2009.] http://www.uic.edu/depts/accc/newsletter/adn14/datahigh.html.

37. **Ahmad, Khalid.** *Sourcebook of ATM and IP Internetworking*. New York : IEEE Press, 2002.

38. **Brink, Sybrand.** *Comparing SDH to Carrier-Class Ethernet*. Backnang : Marconi Communications, 2003.

39. **Miller, Philip.** *LAN Technologies Explained*. New York : Digital Press, 2000.

40. **Seifert, Rich.** *The Switch Book: The Complete Guide to LAN Switching Technology.* Hoboken : John Wiley & Son, 2000.

41. **Answers.com.** Ethernet. *http://www.answers.com*. [Online] [Cited: February 2, 2009.] http://www.answers.com/topic/ethernet.

42. **Goralski, Walter.** *Optical Networking & WDM*. New York : McGraw-Hill/Osborne, 2001.

43. **Laude, Jean-Pierre.** *DWDM Fundamentals, Components, and Applications*. Boston : Artech House, 2002.

44. **ITU-T.** *ITU-T Recommendation G.694.1*. Geneva : ITU-T, 2003. G.694.1.

45. —. *ITU-T Recommendation G.709*. Geneva : ITU-T, 2005. G.709.

46. **xchange.** Getting on the Same Wavelength. *http://www.xchangemag.com*. [Online] [Cited: March 5, 2009.] http://www.xchangemag.com/articles/526/69h1515402888778.html.

47. **Valdar, Andy.** *Understanding Telecommunications Networks*. London : IET, 2006.

48. **Ohrtman, Frank.** *Voice over 802.11*. New York : Artech House, 2004.

49. **(ed), Richard Swale.** *Voice over IP: Systems and Solutions*. Stevenage : IET, 2001.

50. **Gallagher, Rick.** *Rick Gallagher's MPLS Training Guide: Building Multi-Protocol Label Switching Networks*. New York : Syngress Publishing, 2003.

51. **Davies, Joseph.** *Understanding IPv6*. Redmont : Microsoft Press, 2008.

52. **Serrat, Joan.** *Deploying and Managing IP over WDM Networks*. Norwood : Artech House, 2003.

53. **Gray, Eric W.** *MPLS - Implementing the Technology*. Upper Saddle River : Addison-Wesley, 2001.

54. **Leger, Lonnie.** *Board of Regents Technology Committee Meeting.* Louisiana : LONI, 2007.

55. **Coveney, P.V.** Large scale computational science on federated international grids: The role of switched optical networks . *Future Generation Computer Systems.* Monthly, 2008, Vol. 9, 10.

56. **Telecomworldwire.** tw telecom's high speed Internet service deployed by Louisiana universities. *Telecomworldwire.* Monthly, 2008, Vol. 18, 10.

57. **Titch, Steven.** *Spinning its Wheels:An Analysis of Lessons Learned from iProvo's First 18 Months of Municipal Broadband.* Los Angeles : Reason Foundation, 2006.

58. **OFS Optics.** http://www.ofsoptics.com/cable/category.php?txtCategoryID=20. *www.ofsoptics.com.* [Online] [Cited: January 21, 2009.] http://www.ofsoptics.com/resources/osp-134-0803(press).pdf.

59. **Net tech DI.** Optical Ground Wire (OPGW) Cable. *www.nettechdi.com.* [Online] [Cited: January 21, 2009.] https://www.nettechdi.com/categories/Fiber-Optics/Fiber-Optic-Cable/Optical-Ground-Wire-(OPGW)-Cable/.

60. **Tyco Electronics.** Fiber optic. *bestenergys.com.* [Online] [Cited: January 21, 2009.] http://bestenergys.com/cable.html#fibre.

61. **Emtelle.** Products. *www.blownfibre.co.za.* [Online] [Cited: January 21, 2009.] http://www.blownfibre.co.za/home/products/index.php.

62. **OFS Optics.** http://www.ofsoptics.com/cable/category.php?txtCategoryID=18. *http://www.ofsoptics.com.* [Online] [Cited: January 22, 2009.] http://www.ofsoptics.com/resources/accubreezefxosp133web.pdf.

63. **The Computer Language Co. Inc.** ZDNet Definition for: EDFA. *http://dictionary.zdnet.com.* [Online] [Cited: March 5, 2009.] http://dictionary.zdnet.com/definition/EDFA.html.

64. Geert's blog – Casu Consulto. *http://www.hauwaerts.be.* [Online] [Cited: May 17, 2009.] http://www.hauwaerts.be/2009/03/06/mpls-experimental-field-renamed-to-traffic-class-field/.

65. **LONI.** Network architecture. *http://www.loni.org.* [Online] [Cited: June 3, 2009.] http://www.loni.org/network/.