Topological Design and Dimensioning of Agile All Photonic Networks

Ning Zhao

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

June 2005

A thesis submitted to the Faculty of Graduate Study and Research in partial fulfillment of the requirements for the degree of Master of Engineering.

Statement of Contributions of Authors

A co-authored paper is included in this thesis. The paper is entitled "Topological design and dimensioning of Agile All Photonic Networks". The co-authors are Lorne Mason, Anton Vinokurov, Ning Zhao, and David Plant. This paper has been admitted by the journal of Computer Networks of Elsevier B.V.

In this paper, Professor Lorne Mason presented the new mixed integer linear programming formulation and the network architecture concepts. Anton Vinokurov developed the JAVA based Topological Design Tool for visualization and analysis of the results. David Plant gives the analysis of optical switching system and the overview of an Agile All Photonic Network. Ning Zhao, the author of this thesis, developed a set of programs for two and three-layer topological designs and evaluated the results with various network models and traffic/cost assumptions.

The supervisor Professor Lorne Mason has attested to the accuracy of this statement.

Abstract

In this thesis, we present the identification of methods and tools for the design and analysis of an Agile All-Photonic Network (AAPN).

This thesis discusses the layered topology which comprises of a set of overlaid star/tree networks, with an optical core space switch at each of the star centers and hybrid photonic/electronic switches at the edges, and optionally, with Multiplexer/Selectors in between to concentrate traffic. Consequently, network cost is minimized while taking into consideration performance criteria such as delay and reliable traffic restoration upon network failure. A new mixed integer linear programming formulation is presented for core node placement and link connectivity to determine the near cost optimal designs. Both a Metropolitan Area Network (MAN) and a Canadian Wide Area Network (WAN) with two-layer or three-layer network topological implementations have been tested. Network models and their performance were evaluated with a set of software tools and methodologies to design and dimension our vision of an AAPN.

Sommaire

Dans cette thèse, nous présentons l'identification des méthodes et des outils de conception et d'analyse des réseaux agiles tout photoniques (RATP).

Cette thèse discute la topologie posée qui comporte un ensemble d'overlaid des réseaux en étoiles/arbres , avec un commutateur optique de l'espace de noyau a chacun des centres d 'étoile et des commutateurs hybrides de photonic/electronic aux bords, et sur option, avec Multiplexer/Selectors entre eux pour concentrer le trafic. En conséquence, le coût de réseau est réduit au minimum tout en prenant en compte des critères d'exécution comme le retard et la restauration fiable du trafic en cas de panne du réseau. Une nouvelle forme de programmation linéaire de nombres entiers mélangés est présentée pour le placement de noeud de noyau et la connectivité de lien pour déterminer les conceptions quasi-optimales en termes de coût pour les types de réseaux: “Metropolitan Area Network” (MAN) et “Wide Area Network” (WAN) avec des réalisations topologiques de réseaux composés de deux-couches ou de trois-couches. Des modèles de réseau et leur performances ont été évalués avec un ensemble d'outils et de méthodologies logicielles pour concevoir et dimensionner notre vision d'un RATP.

Acknowledgements

Firstly, I would like to express my deepest gratitude to my supervisor, Prof. Lorne Mason, for his indispensable guidance and invaluable advice throughout my graduate studies at McGill University. I am also grateful to Prof. Mason for providing financial assistance to complete this research. I further would like to thank Anton Vinokurov for the suggestions and great help in my simulations and programming. The Topological Design Tool (TD Tool) he developed greatly helped me in the visualization and analysis of my results.

I sincerely thank my fellow colleagues of the Telecommunication and Signal Processing Lab for their technical support and friendship. Especially, Fariba Heydari and Faker Moatamri's French translation of the abstract is much appreciated.

Finally, I am grateful to my parents and my sister for their unconditional love and support during my graduate studies at McGill. Also I need to thank my boyfriend Yan Chen for his continual encouragements and unwavering love to me. My sister Jing Zhao and my friend Xiang Zhou gave me great help in the grammar of this thesis. Last but not least, I need to thank all friends I met in McGill for being with me and helping me. Thank you.

Contents

List of Figures

Chapter 1 Introduction

1.1 Problem statement

The Agile All-Photonics Networks (AAPN) Research Network is aimed at developing an all-photonic network, where the core network will stretch as close as possible to the end-user, and the number of optical-electrical-optical conversions (OEOs) in network data paths will be reduced. This thesis is part of Task 1.1.4 in AAPN, network modeling and performance evaluation [1].

1.1.1 The topological design and dimensioning in an AAPN

The objective of this thesis is to design and dimension an all-photonic network to support the traffic requirements from end users. In the topological design, graph theory is commonly used for network modeling; and various optimization methods are employed to determine the best configuration of the network facilities while minimizing the overall costs.

In an AAPN, end users are connected to edge nodes, and their transmission data are exchanged in core node switches. Various services may be carried by an AAPN, while these services may have different requirements on QoS, bandwidth etc. The design of an AAPN needs to determine the allocation and connectivity of network elements. While the dimensioning of an AAPN means to scale up the capacity and the processing abilities of nodes and links in the network.

The goal of topological design of an AAPN is to achieve a specified performance at minimum cost. Because this complex problem is intractable, the optimization problem should be decomposed into separate components. A possible method is to separate the design of core node allocation (backbone networks) and Mux/Sel allocation (local access networks). Each of these problems is also difficult and complex, and extensive research in deriving solution techniques for them is needed.

The transmission of information always involves a transfer delay in the form of latency. An acceptable level of the traffic delay is one of the most important quality specifications. In the design of an AAPN, the transmission delay should be considered when designing the

network topology.

The AAPN is designed with service continuity and reliability even with heavy burden of traffic load. Thus it is inevitable for an AAPN to be designed to provide reliable services which is the ability to perform required functions when a set of specified components become unavailable.

1.1.2 Approaches adopted in this thesis

Based on our research, overlaid star/tree network architectures are applied, where core nodes are connected to Multiplexer/Selector switches (Mux/Sels) and then to edge nodes, or to edge nodes directly. There is no direct connection between any two core nodes. The network model includes the following elements:

- Electrical/Optical edge nodes, each equipped with wavelength-tunable or fixed lasers and receivers.
- Optical Mux/Sels, each concentrating closely-located edge nodes and connecting to core nodes using Dense Wavelength Division Multiplexing (DWDM).
- Optical core switches, each designed to deliver packets with low latency time.

Our research has been done in a systematic and comprehensive way. The network topological design and dimensioning for AAPN has been accomplished in several steps described as follows:

- Network architecture modeling: Design the building blocks (network elements) and feasible AAPN architectures to construct the hierarchical network.
- Network traffic demand and connectivity analysis: Calculate traffic demand, and select suitable point to point (edge-core, Mux/Sel-core and edge-Mux/Sel) circuit connectivity methods based on traffic demand and distance.
- Working network design: The design of a working network should be cost and performance oriented. This step can be accomplished in the following fashion:
  - Allocate edge nodes based on demographic data for the area being served.
  - Locate remote optical Mux/Sels and assign edge nodes to them in order to minimize access costs using *P-median* or Capacitated Plant Location Problem (CPLP) solutions.
  - Formulate the allocation of optical non-blocking core nodes and link connectivity between multi-layers as a cost optimization problem. Several methods including Lagrangian relaxation, Enumeration calculation, Simulated Annealing and CPLEX are employed to resolve these large combinatorial

optimization problems.

- Backup network design: For reliability and traffic load restoration, design a traffic restoration strategy upon network failure for the connectivity between Mux/Sels and core nodes, and between edge nodes and Mux/Sels.

1.1.3 Our Contributions

In this thesis, with the input of the population distribution data, traffic demand estimation, link capacity and facility costs, the problem of topological design and dimensioning of an AAPN is discussed. We can determine the number, size and location of edge nodes, remote optical Mux/Sels and optical core switches, and the homing patterns within them. We have developed several analytical models in MATLAB with the Lagrangian mechanism for the large-scale networks. Different parameter settings with different performance and efficiency of the program are tested. Also, we use CPLEX/Enumeration Calculation to solve the smaller-sized optimization problems and compare the results with heuristic algorithms. Customized JAVA software was developed for the visualization and analysis of results by Anton Vinokurov.

Specifically, a new Mixed Integer Linear Programming (MILP) formulation has been developed to resolve the location and assignment problem for core node allocation. In this formulation, the performance factors of delay and scheduling efficiency have been taken into account.

The test results show that the techniques employed are quite reasonable in both wide area and metro area networks. The approaches can also be extended to other similar network design problems.

1.2 Thesis structure

The remainder of this thesis is organized as follows.

- Chapter 2 gives a literature review for the network topological design, including the existing design problems and optimization approaches used.
- Chapter 3 proposes the topological design method for the Agile All Photonics Networks. It analyzes the models of network cost elements and solves them with optimization methods.
- Chapter 4 presents the results from our simulations and analysis of various designs.

- Chapter 5 summarizes the conclusions of the thesis and presents the applications of the proposed model in Agile All-Photonic Networks. Suggestions by Professor Lorne Mason for future research directions are also discussed.
- In the appendix, the cost data sets used for simulation and detailed programming procedures are provided.
- Lastly, the references used in thesis are provided.

Chapter 2 State of the art in network topological design

A network can be modulated as a series of points or nodes interconnected by communication paths or links. In communication networks, a topology is usually a schematic description of the arrangement of a network, including its nodes and connecting links. There are two ways of defining network geometry: the physical topology and the logical (or signal) topology. The physical topology of a network is the actual geometric layout of points and nodes. Logical (or signal) topology refers to the nature of paths the signals follow from node to node[1]. In the context of our study, the topological design is for the physical topology, which means the allocation and connectivity of network elements. This problem usually consists of the following component problems:

- The geographical placement of service points (hosts).
- The selection of links that satisfy traffic requirements, utilization thresholds, and other constraints.
- The optimization of network cost and performance.

However, most network design optimized problems are complex and computationally intractable. Thus heuristic algorithms are employed to approximately calculate the minimum cost within a reasonable computation time.

In this chapter, we will review several well-known problems of facility location and network topological design problems, and present some existing approaches to solve them.

[1] http://www.whatis.com/

2.1 Review of topological design problem sets

2.1.1 The *P-median* problem

The *P-median* problem is a classical location problem. Its purpose is to locate p "facilities" (medians) to serve a set of n "customers". The cost of every assignment is the sum of the distances from each customer to the facility that serves it. The distance can also be weighted by some factors such as the demand of a customer. If each candidate facility has a fixed capacity, i.e. a limited maximum number of customers that it can serve, the problem is called a capacitated *P-median* problem.

The *P-median* problem is well known to be NP-hard [2]. This means that as the problem's dimension becomes greater, heuristic methods are the only alternative to determine feasible solutions. Therefore, several heuristics are developed for *P-median* problems. The approaches that are widely used are the genetic algorithm, Tabu search, branch-and-price, and several variations of Lagrangian Relaxation heuristics etc.

The *P-median* problem can be formulated as follows:

$$\min_{y,z} \sum_{j=1}^{J} \sum_{i=1}^{N} y_{ij} \cdot d_{ij} \tag{2.1}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} y_{ij} = 1 \quad i = 1...N \tag{2.2}$$

$$\sum_{j=1}^{J} z_j = p \tag{2.3}$$

$$y_{ij} \le z_j \quad i = 1...N, j = 1...J \tag{2.4}$$

And integer constraints:

$$\underbrace{z_j \in \{0,1\}}_{\text{J variables}}, \underbrace{y_{ij} \in \{0,1\}}_{\text{N*J variables}} \quad i = 1...N, j = 1...J \tag{2.5}$$

Where,

i is the customer number in the network, and N is total number of customers.

j is the facility number in the network, and J is total number of candidate facilities.

$[d_{ij}]_{N \times J}$ is a distance matrix, and d_{ij} is the distance from customer i to facility j.

p is the number of facilities used as medians.

The variables are:

$$z_j = \begin{cases} 1 & \text{if candidate } j \text{ is a facility used as a median} \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

$$y_{ij} = \begin{cases} 1 & \text{if customer } i \text{ is served by facility } j \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

The objective is to minimize the representative function (2.1), the total distance between customers and facilities. The constraint (2.2) ensures that the demand of each customer *i* will be satisfied. Constraint (2.3) determines that the exact number of facilities to be allocated as medians is *p*. Constraint (2.4) ensures that only when a facility *j* is selected to be a median, can customer *i* be served by *j*. Constraint (2.5) gives conditions that all variables should be binary as shown in formula (2.6) and (2.7). In this case, one customer can only be served by one median, which is called "single source". If we set $y_{ij} \in [0,1]$ and $\sum_{j=1}^{J} y_{ij} = 1 \; \forall i = 1...N$, it means one customer can be served by multiple medians, which is called "multi sources".

From the literature, other researchers have put forward various optimization methods to solve the *P-median* problem. A successful approach to approximately solve this problem is the use of Lagrangian heuristics, based upon Lagrangian Relaxation (LR) and subgradient optimization. In [3], a modified Lagrangian Relaxation which generates an optimal integer solution was studied, which was called semi-LR. It is used to solve large-scale instances of the *P-median* problem. In [4], Lagrangian and surrogate relaxations were combined to relax the assignment constraints in the surrogate way in the *P-median* formulation. Then, the LR of the surrogate constraint was obtained and approximately optimized (one-dimensional dual). This method is tested to have same good result as Lagrangian alone but with more coding work and it is more suitable for very large problems. Lagrangian relaxations are proved to be very stable (low-oscillating) and reach good results in acceptable computational time.

Another algorithm often used is the genetic algorithm. In [5], a genetic algorithm (GA) was proposed to solve the capacitated *P-median* problem. The proposed GA used not only conventional genetic operators, but also a new heuristic "hyper-mutation" operator suggested in their work. Branch-and-price approach was also widely used for *P-median* problems. The reference [6] described a branch-and-price algorithm for the *P-median* location problem. A stabilized approach that combines the column generation and Lagrangian/surrogate relaxation was proposed in that paper. In [7], a multi-start hybrid heuristic that combines elements of several traditional meta-heuristics to find near-optimal solutions to *P-median* problem was presented.

Tests reported in these papers show that these methods can achieve acceptable results in

terms of both running time and solution quality. However, algorithms other than LR approaches tend to be more complicated and are more difficult to implement in programming software. As we have several kinds of optimization problems in our research of Mux/Sel allocation and core node allocation in AAPN, our work would be reduced significantly by choosing a universal optimization algorithm according to the problem size, efficiency requirement and complexity of implementation. This will be analyzed more by the end of Section 2.1.2.

2.1.2 Capacitated plant location problem

Quite similar to the *P-median* problem, another facility location problem formulation is the Capacitated Plant Location Problem (CPLP). The detailed clarified presentation and discussion of this problem can be found in [8]. In the CPLP, just as in the *P-median* problem, a set of potential facility locations and a set of customers, each with a known demand, are given. Each potential location has limited capacities, and each customer must be satisfied by one or more facilities. The objective of the CPLP is to assign customers to be served by facilities with minimized total cost. The distinct difference to distinguish CPLP from *P-median* problem is that there is no fixed number of facilities in CPLP.

The formulation can be presented as follows:

$$\min \sum_{j=1}^{J} c_j \cdot z_j + \sum_{j=1}^{J}\sum_{i=1}^{N} y_{ij} \cdot cost_{ij} \tag{2.8}$$

$$\sum_{j=1}^{J} y_{ij} = 1 \quad i = 1...N \tag{2.9}$$

$$\sum_{i=1}^{N} E_i \cdot y_{ij} \le Q_j \quad j = 1...J \tag{2.10}$$

$$y_{ij} \le z_j \quad i = 1...N, j = 1...J \tag{2.11}$$

And integer constraints:

$$\underbrace{z_j \in \{0,1\}}_{\text{J variables}}, \underbrace{y_{ij} \in \{0,1\}}_{\text{N*J variables}} \quad i = 1...N, j = 1...J \tag{2.12}$$

Where，

J is the set of potential facility sites, and N is the set of customers.

$[\cos t_{ij}]_{N \times J}$ is the cost matrix, and $cost_{ij}$ is the cost of supplying all demand of customer i from facility in j.

c_j is the fixed cost of opening a facility at j, and Q_j is its maximum capacity if it is

open.

E_i is the demand from customer *i*.

The binary variable z_j is equal to 1 if facility *j* is open and 0 otherwise.

The binary variable y_{ij} is equal to 1 if customer *i* is served by facility *j* and 0 otherwise.

Unlike general *P-median* problems, the total cost in CPLP normally includes start-up costs to open the facilities plus transportation costs required to satisfy customers' demand in (2.8).

The constraint (2.9) is the demand constraint that means all demands from a customer should be served. The constraint (2.10) is the capacity constraint, which means that facility *j* has maximum capacity Q_j if opened. Constraint (2.11) guarantees that a customer can only be supplied by an opened facility. For the side constraint(2.12), it ensures that a facility can only be either opened or closed, and one customer can only be served by one facility.

Because of the capacity threshold for each potential facility in (2.10), this problem is defined as "capacitated". Furthermore, because of (2.12), where any given customer can only be served by one facility, this is often referred as "single-source" CPLP (SSCPLP). And if we add one more constraint:

$$\sum_{j=1}^{J} z_j = p \tag{2.13}$$

This problem is turned into a capacitated *P-median* problem.

CPLP being a well-developed problem, the literature on it is very rich. In [8-16], the problem of CPLP location is addressed. Researchers have worked on both heuristic algorithms and exact methods to solve CPLP as reviewed in [14]. For the exact algorithms, there are also a lot of options such as Branch-and-Price algorithm. These algorithms are quite similar to those for *P-median* problems.

When the problem size is large and no explicit solutions can be obtained, heuristic algorithms are used. [9] gave the comparison of several heuristic schemes such as Evolutive Algorithms (EA), Greedy Randomized Adaptive Search Procedure (GRASP), Simulated Annealing (SA) and Tabu Search (TS) etc. In [14], the heuristics for CPLP was classified under three basic approaches: the greedy heuristics, interchange heuristics and Lagrangian heuristic.

In the greedy heuristics, neighborhood structures will usually involve the so-called "ADD/DROP" procedure. In the ADD procedure, facilities are opened and clients are connected to the nearest facilities. In the DROP procedure, opened facilities are closed and their customers are moved to other facilities to save cost.

In the category of greedy heuristics, once a decision is made, it will not be changed. However,

in interchange heuristics, improvement should be made on the greedy solution. Two different methods belong to the interchange heuristic category, (1) the Alternate Location Allocation (ALA), and (2) the Vertex Substitution Method (VSM).

In Lagrangian Heuristics, Lagrangian Relaxation is used for the optimization procedures. In [11-14,17], the LR approach was applied for CPLP problems. Several types of Lagrangian heuristics are presented in these papers in terms of the bounds and solution techniques. Results in these papers have shown that LR is an acceptable method for CPLP and can also be used in some other related problems such as *P-median*. When the upper bound is calculated in every Lagrangian iteration, the greedy and interchange heuristics can be applied. Reference [16] presented a combination of LR approach and restricted neighborhood search. Detailed discussion about various upper bound and lower bound calculation schemes can be found in these papers.

As we will discuss in Section 4.3.2, we can formulate the Mux/Sel allocation to a *P-median* or CPLP problem. As analyzed in [10], Lagrangian heuristics were tested for various kinds of location problems such as *P-median*, Uncapacitated Location problems and CPLP etc. Lagrangian Heuristics give quite good solutions in an acceptable time. Based on our investigation, the general method of Lagrangian Relaxation has been chosen for our problem sets since it is a widely used and efficient algorithm for both *P-median* and CPLP problems as shown in both Section 2.1.1 and 2.1.2.

2.1.3 Network topological design

Topological design, when compared with previous location problems, adds demand volume into the design problem together with the capacity-dependent costs, and both factors are considered in the total cost in network design. The objective of network topological design is to achieve a specified performance while minimizing the overall cost. When the network size gets large, the problem will become complex and virtually unsolvable and also the network operation and maintenance will be more complicated. Thus a hierarchical internetworking model is necessary for the overall design. A well-accepted network construction model to simplify the task of building a reliable, scalable, and less expensive hierarchical network is the three-functional-layer of a network:

Core layer: This layer is considered the backbone of the network and includes the high-end switches, such as the all-optical switches and high-speed cables, such as DWDM connections. High data transfer rate and high reliability are the most important performance criteria for this layer.

Distribution/Concentration layer: This layer ensures that packets are properly routed among the end users. Devices used in this layer include the Multiplexer/Selectors. Packets from different edge nodes are multiplexed and routed to different core nodes.

Access layer: This layer includes edge nodes such as hubs or switches to connect end users. In the access layer of an all-optical network, this layer is for electrical-optical conversion and will be connected with user networks such as LAN, ATM access network or servers as its downlink. Edge nodes should be allocated according to the distribution of network resources and demands.

In chapter 6 of [18], location and topological design problems are studied for different cases. Location problems are easier than topological design ones because the traffic demands between nodes are not considered. The location design problem is to connect customers to possible facility locations so that the total cost is minimized. Besides, in order to minimize the cost of the core network connection at the same time, the combined node location and link connectivity design problems are also discussed. Moreover, the topological design problem with the consideration of demand volumes is studied in several different cases. The reviews of [18] in literature give us a strong background on our formulation for AAPN topological design problems.

Numerous investigations have been conducted upon the topological design for layered networks. The topological design can be divided into two categories. One is the physical network design, where the physical interconnection of network elements is determined, as in [19-23]. Another is the logical topology where the routes/lightpaths are set up and nodes are configured, as discussed in [24,25]. Some previous work has combined the two designs together such as [26]. Because of the complexity of the large-scale AAPN, the physical and logical design has been separated. This thesis will only deal with the physical network design for the large-scale AAPN. The physical topological design problem has been investigated by different methods as following.

The general objective of this problem is to minimize the network cost including installation cost of nodes and links and capacity cost. In [22], the budget constraint is considered and the two-layer network with access nodes and transit nodes are assumed, while in [23], the design of an access network is studied with a binary linear programming model. In [19], similarly to [23], the problem of the design of telecommunication access networks with reliability constraints is studied. An optimization model is formulated and is solved with a simulated annealing algorithm, however, in both models of [19,23], the location of backbone switches are given, and the design is limited to the access network where customers' premises are connected to the network carrier's switches.

In [20,21], not only the network cost, but also the performance criteria are considered in the objective function when formulating the optimization problem. In [20], the objective is to balance the overall investment and delay imposed. The topological design problem is formed as a nonlinear programming problem and is solved with Lagrangian relaxation. Moreover, [21] takes the reliability and flow constraints into account while simultaneously minimizing

network delay and cost. The multi-objective problem is solved with a genetic algorithm. These approaches are quite useful for our AAPN topological design.

In this thesis, we present a systematic and step by step design methodology for the total network design of an AAPN, which will be discussed in detail in Section 4.3.

2.2 Network restoration strategies

Network optimization not only deals with normal network operation, but also considers a set of failure situations. Under different failure scenarios, the availability of the links and nodes, and demand volumes requested by some nodes may vary from one to another. Within an existing network, how to quickly restore the affected traffic upon various network failures is an important issue for network reliability performance. In an AAPN, the network is self-healing, which means that it should have the ability to perceive that it is not operating correctly and, without human intervention, to make the necessary adjustments to restore the network to the normal operation. Thus the network restoration strategies upon network failure should be considered for a robust design of all-optical networks.

In preplanned restoration schemes, the re-establishment mechanisms can be performed on the link or path basis shown in Figure 2-1.

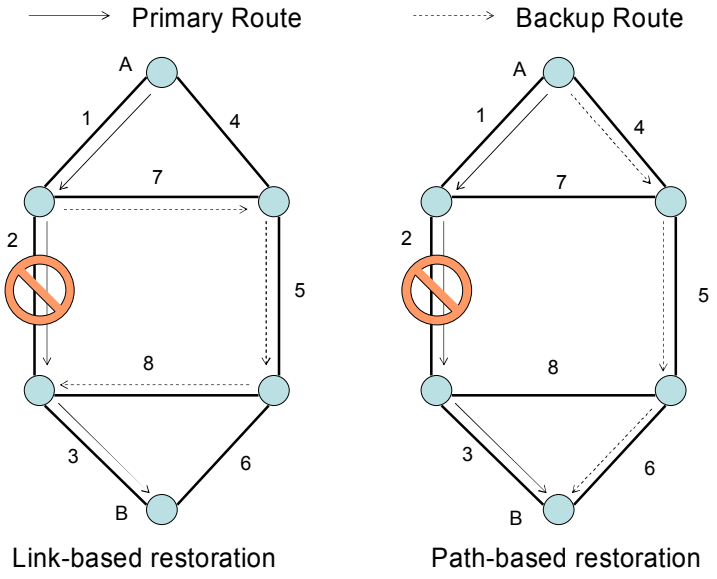


Figure 2-1 Re-establishment mechanisms

While path restoration individually re-establishes the end-to-end flows that use failed links, link restoration only re-establishes the failed links. For example, in Figure 2-1, when link 2 is broken, the route from A to B has to be re-established. In link-based restoration, the

re-established route is from link 1-7-5-8-3, and in path-based restoration, the re-established route is from link 4-5-6. Path restoration can be classified into failure-oriented reconfiguration and global reconfiguration. In failure-oriented reconfiguration, only the affected working paths are rerouted, while in global reconfiguration, the whole layout of working paths (affected and unaffected) may be rearranged to overcome a failed link or node. Global reconfiguration gives us the minimum spare capacity cost for the given restoration requirements but is more difficult to be implemented in practice.

In chapter 9 of [18], network resource failures and protection/restoration mechanisms in single-layer networks were studied, and thus corresponding to various problem categories, formulations were given for different protection mechanisms with their applicability to different technologies. The link protection strategies for optical networks have also been particularly studied in [27,28].

In [29], the capacity and flow assignment problem from the design of self-healing ATM networks using the virtual path concept was studied. It was formulated as a linear programming problem with the objective to minimize the spare capacity cost for the given restoration requirement. A new heuristic algorithm based on the Minimum Cost Route concept was developed for the design of large self-healing ATM networks using path restoration.

In [30], a restoration scheme for IP over WDM networks was investigated. A simple integrated protection/restoration scheme was developed to coordinate both the IP and optical layers. In this joint two-layer recovery scheme for IP-centric WDM based optical networks, the optical layer will take the recovery actions first, and subsequently the upper IP layer initiates its own recovery mechanism, if the optical layer does not restore all affected services. The proposed two-layer recovery scheme led to better results than the traditional single-layer recovery scheme.

In [31], the dual-failure restorability and related availability considerations were analyzed in the Shared Backup Path Protection (SBPP). In their work, the network was designed to achieve the survival of services against all single failures first. Then analysis was done upon how the resulting network withstands the dual failure combinations. Following this, the network was changed to enhance the restorability for dual failures. SBPP capacity requirements were optimized with explicit limits on the number of primary service paths that are allowed to share the same backup link.

In our design, according to the spare capacity in the links after the working network design, the existing algorithm in [29] is borrowed for the reliable traffic reestablishment.

Chapter 3 Background knowledge

3.1 Traffic patterns for dimensioning

In this thesis, when the network topology is designed, traffic demands or traffic distribution patterns are necessary input. However, the traffic characteristics, or the traffic prediction and estimation in an all-photonic network are not discussed in this thesis.

A gravity model for traffic distribution and a flat community of interest factor are assumed in the study. Normally, the net traffic demand matrix is given by the gravity model[2]. In a gravity model, the traffic is modeled as following:

- Traffic between sites is proportional to traffic originated at each site, i.e. $\lambda_{ij} \propto I_i I_j$.
- There is no systematic difference between traffic in node *i* and node *j* and only the total volume matters.
- A distance term can be included, and the importance of locality of information varies depending on various services.

The general form of the gravity traffic model can be described as follows.

$$\lambda_{i,j} = \left[\frac{I_i \cdot I_j}{\left(d_{i,j}\right)^{\alpha}} \lambda_0 \right] \tag{3.1}$$

Where,

$\lambda_{i,j}$ – demand between nodes *i* and *j*,
I_i and I_j – "importance factor" assigned to nodes *i* and *j*, for example, population,
λ_0 – normalized demand unit,
$d_{i,j}$ – distance between nodes *i* and *j*,
α – power parameter related to traffic types

For example, in the telephone network $\alpha \to 2$ and for Internet traffic $\alpha \to 0$, where Internet

[2] Discussion about the gravity model can be found at: http://faculty.washington.edu/~krumme/systems/gravity.html.

traffic demand is almost independent of source and destination locations. The flat gravity based traffic model of Internet is selected initially to exercise our topological design tools in the absence of more accurate and service specific traffic models. This traffic demand matrix is only an input parameter of our design program. When better traffic models become available for the "services of the future", they can be easily incorporated in our AAPN topological design tools and procedures due to the modular structure we employ in the network and traffic modeling and the software implementation.

The net network traffic demand matrix is governed by the population of the originating area served by the edge node and the population of the terminating area served by the receiving edge node. By design we choose that the amount of customers served by each edge is approximately equal so that the traffic demand matrix is flat. This is quite realistic when edge nodes have the same size. Furthermore, even when edge nodes have different sizes, the same algorithm discussed in this thesis can still be employed for the design, only with some changes in the input traffic matrix. If there is a high population corresponding to one existing location (eg, a city), then one can place a certain number of edge switches at this location. By doing this we simplify the design considerably as all parts have equal demand requirements.

To dimension the AAPN, we need to compute the link capacity, which meets the Quality of Service (QoS) requirements of the supported services given the net traffic demand matrix and routing algorithm. For deterministic shortest path routing and the net traffic matrix computed above we can compute link traffic demands. Next we need a queuing model to relate traffic demand, link capacity and QoS. To put it in another way, net traffic demand is not sufficient to determine the bandwidth needed from one source to destination. The utilization ratio and blocking probability should also be considered. For this purpose the ErlangB and Farago (defined in [34]) models are used, as we are initially designing for a single high quality service class that can potentially handle all traffic types in a single unified manner by suitably over provisioning. The detailed examples to calculate the traffic will be given in Section 5.4 and 5.5 when we simulate the Wide Area Network of Canada and Metro Area Network of Gotham.

The famous ErlangB formula gives a simple way to calculate the blocking probability. Initially ErlangB model was used in telephone call center systems. In the ErlangB Model, the caller makes only one attempt to place the call. If all servers are busy (being blocked), the call is cleared from the system and will never return. It can be used for calculation for any one of these three factors if you know or predict the other two:

- Busy Hour Traffic (BHT), call traffic during the busiest hour of operation
- Blocking probability, or the percentage of calls that are blocked because not enough servers are available

- Number of servers

The ErlangB formula, though designed only for the classical case of single-rate Poisson traffic, is still often used even in today's complex networks. ErlangB can handle the common traffic engineering problems relatively easily.

In Farago's traffic model [34], traffic demand is modeled as a set of stochastic processes and the mean rate of the process is supposed to be F_t (the expected value of the offered load to a link). $B_{\max}$ is the largest component flow, and the line capacity is $C \cdot B_{\max}$, where C is similar as the number of servers in previous model. Thus the blocking probability estimated by ErlangB is:

$$P_{ErlangB} = \frac{F_t^C / C!}{\sum_{n=0}^{C} F_t^n / n!} \tag{3.2}$$

Moreover, in [34], Farago gives a simple upper bound for the blocking probability for general multi-rate traffic. If $C \cdot B_{\max} > F_t$, the following bound holds for the probability of the demand exceeding capacity:

$$P_{Farago} \le \left(\frac{F_t}{C}\right)^C e^{C-F_t} \tag{3.3}$$

Farago's bound is conservative in the sense that worst–scenario of traffic processes is assumed to be offered.

If we measure the flow bandwidth demand and link capacity in the unit $B_{\max}$=1, the following figure shows the relationship between probability of blocking and utilization as capacity varies.

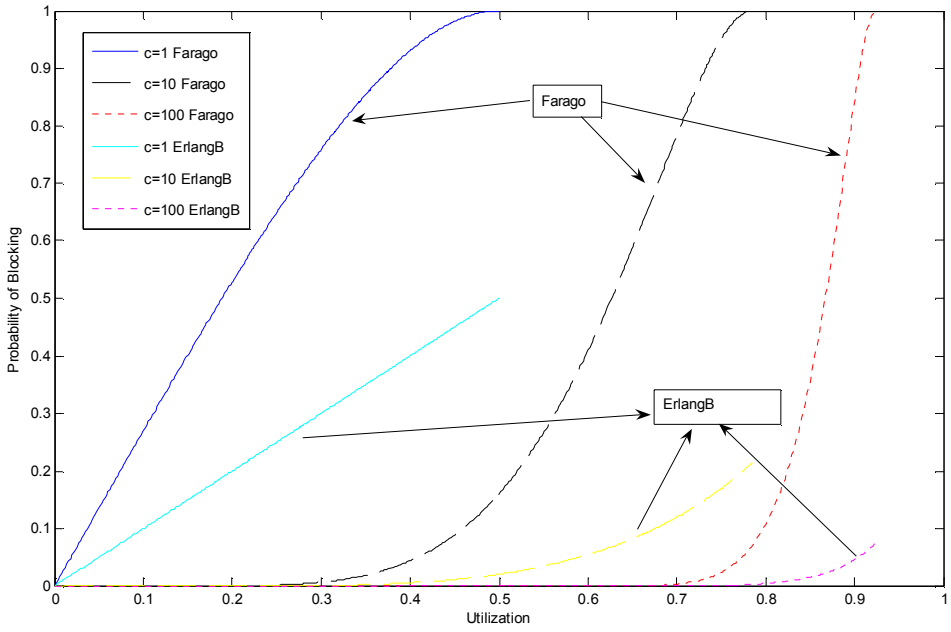


Figure 3-1 Farago and ErlangB blocking probability

This figure shows that:

When capacity is 1, 1% Farago loss probability occurs when utilization is 0.5%.

When capacity is 10, 1% Farago loss probability occurs when utilization is 32%.

When capacity is 100, 1% Farago loss probability occurs when utilization is 73%.

Farago's bound is derived for a pure blocking system. We can anticipate that adding "small" buffers should only reduce the blocking rate, so Fargo's bound will hold with strict inequality. As single best effort traffic class in our AAPN design is assumed, the loss probability should be sufficiently low.

Our design tool described in the following sections uses as input a link dependent utilization parameter. Hence alternative traffic models to Farago's bound could be used to generate the utilization parameter meeting the QoS requirements used in our design tool.

For example, in our national network design, as we have estimated the traffic from each Mux/Sel to core node would be several hundred Gb/s, while the flow traffic (from one Mux/Sel to another) is several Gb/s. So if we specify that the loss probability should be less than 1%, the utilization will be less than 73%. As Farago's bound is given under the worst scenario, normally the utilization should be much more than this percentage, which is an acceptable result for our design.

3.2 Cost models for circuit designs

In the topological design for AAPN, the connectivity methods between edge nodes, Mux/Sels and core nodes need to be defined. As we know, there is no wavelength direct connection in AAPN, so wavelength routing will be determined after we set up the physical topology, and logical topology will not be a concern in our research. However, this is not the end of our work. With given physical interconnection from node to node, we need to select a suitable connection method such as Dense Wavelength Division Multiplexing (DWDM), Coarse Wavelength Division Multiplexing (CWDM), direct fiber, amplified fiber etc, for the transmission systems to minimize the link cost. For local areas that direct fibers can support without amplifiers, we can use direct fiber or CWDM. While for longer distances, we need amplifiers and regenerators in certain intervals of the fiber link to support uninterrupted transmission. Similarly amplified fibers and DWDM are two choices for wide area connections. Usually the more traffic there is between source and destination; the better it is to employ CWDM/DWDM systems to save the connectivity cost. More information on the transmission systems can be found in [35,36].

3.2.1 Fiber connections

Direct fiber

Normally without any amplifying equipment, the data transmission distance that a direct fiber connection can support is less than 80 km with current technology. Different interface types on the equipment can support different non-distortion transmission distances. As in metro network, normally the distance range will be less than 80 km; we can use direct fibers without amplifier. Here direct means for one fiber, only one wavelength (color) is supported, and no multiplexer and amplifier is used. The connection is shown as follows in Figure 3-2:

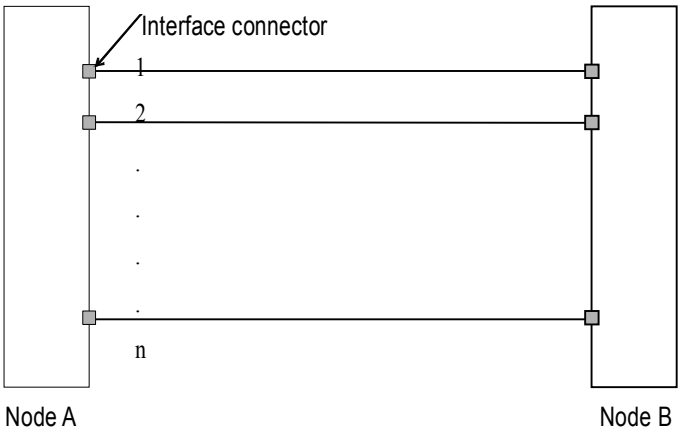


Figure 3-2 Direct fiber for local area

As shown in this figure, for the case without amplifiers, the total circuit cost has two parts. One is the interface cost. Normally interface connectors are necessary to support connections longer than dozens of kilometers on node switches. The other part is the fiber cost which is in direct proportion to distance.

$$COST\_direct = \underbrace{2 \cdot n \cdot C_{IF\_LH}}_{\text{node interface cost}} + \underbrace{n \cdot C_f \cdot d_{AB}}_{\text{fiber cost}} \tag{3.4}$$

Where,

C_f : Cost of fiber, \$/kilometer.

C_{IF_LH} : Node switch interface, cost for Long Haul interfaces in \$/each. This kind of optical interface normally can support transmission distance of 40, 60, 80 Km without amplifiers.

T_{AB} : Traffic between node A and node B in (Gb/s). Here we suppose that the traffic is symmetric, so $T_{AB} = T_{BA}$.

n: number of interfaces needed on node switch, which depends on the traffic between the two nodes. Then if one fiber can support 10Gb/s traffic, we have $n = ceil(T_{AB}/10)$.

d_{AB} : Distance between node A and node B.

Amplified fiber

For distance >80km where fiber transmission attenuation cannot be ignored, amplifiers should be used to amplify optical signals and regenerators may also be used to regenerate signals.

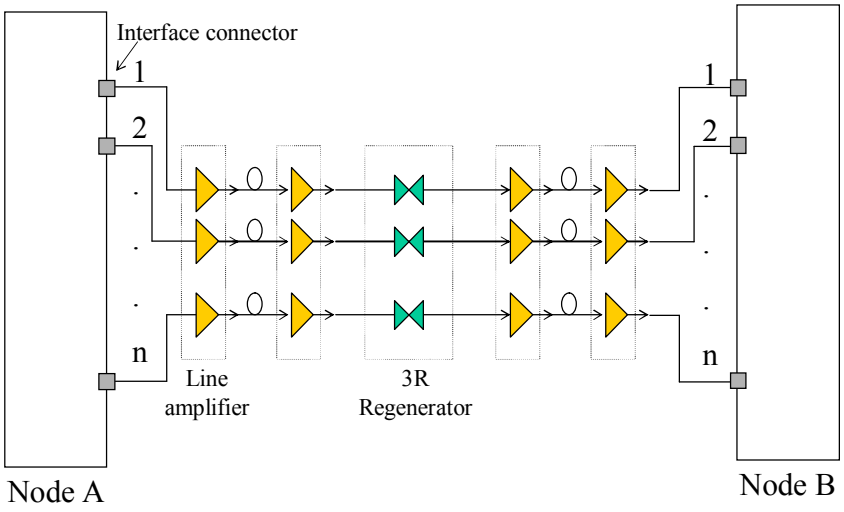
In the study and development of agile transparent optical transport networks, the electronic cross-connect and add/drop multiplexing switching systems are replaced with photonic counterparts; at the same time, the Optical-Electrical-Optical regenerators are also replaced by all optical ones. This enables the provision of light paths linking network ingress and egress nodes where the signal is transmitted entirely in the optical domain, thus eliminating the expensive OEO conversions associated with the SONET/SDH cross connect systems.

In an AAPN, optical amplification should be performed at certain intervals along a fiber span to compensate the inherent signal deterioration caused by propagation through the fiber. In the optical transmission systems, optical amplifiers boost not only the signal but also noise. Consequently, the original signal must be recovered and regenerated after a certain number of amplifications. Impairments in optical signals require periodical restoration of data. As presented in [37], an all optical 3R (Re-amplification, Reshaping, Re-timing) regenerator consists of:

- Clock recovery unit: extracts the repetition rate

- Optical pulse source: generates high-quality pulse stream at this rate
- Decision gate: imposes the data on this pulse stream

The amplified fiber connection is shown as in Figure 3-3.



Note:

1. This figure only shows equipment required for signals traveling in one direction.
2. Normally amplifiers and regenerators for different fibers are used separately. But in practice, they are deployed in the same hut for all fibers in the same way.

Figure 3-3 Amplified fiber for broad area

With amplifiers and regenerators, the total circuit cost has three parts: interface cost, fiber cost, and cost of amplifiers and regenerators.

$$COST\_fiber\_amp = \underbrace{2 \cdot n \cdot C_{IF\_LH}}_{\text{node interface cost}} + \underbrace{n \cdot C_f \cdot d_{AB}}_{\text{fiber cost}} + \underbrace{\left[ceil\left(\frac{d_{AB}}{max\_amp} \right) - 1 \right] \cdot C_{Amp} + \left[ceil\left(\frac{d_{AB}}{max\_reg} \right) - 1 \right] \cdot C_{Reg}}_{\text{cost of amplifiers and regenerators}} \quad (3.5)$$

$$\approx \underbrace{2 \cdot n \cdot C_{IF\_LH}}_{\text{node interface cost}} + \underbrace{n \cdot \left(C_f + \frac{C_{Amp}}{max\_amp} + \frac{C_{Reg}}{max\_reg} \right) \cdot d_{AB}}_{\text{line cost}}$$

Where,

C_{Amp} : cost of amplifier, \$/amplifier for one wavelength.

C_{Reg} : cost of regenerator, \$/regenerator for one wavelength.

max_amp : the maximum allowable distance (link budget value) that can be traversed between two amplifiers.

max_reg : the maximum allowable distance that can be traversed between two

regenerators.

We can see from the formula that with a given level of traffic, *n* is fixed, then node interface cost turns out to be a constant. The total cost will increase as the distance increases, but it is not linear with distance because of the discrete allocation of amplifiers and regenerators. However, in order to simplify the problem, it can be approximated as a linear function w.r.t. the distance as shown in (3.5).

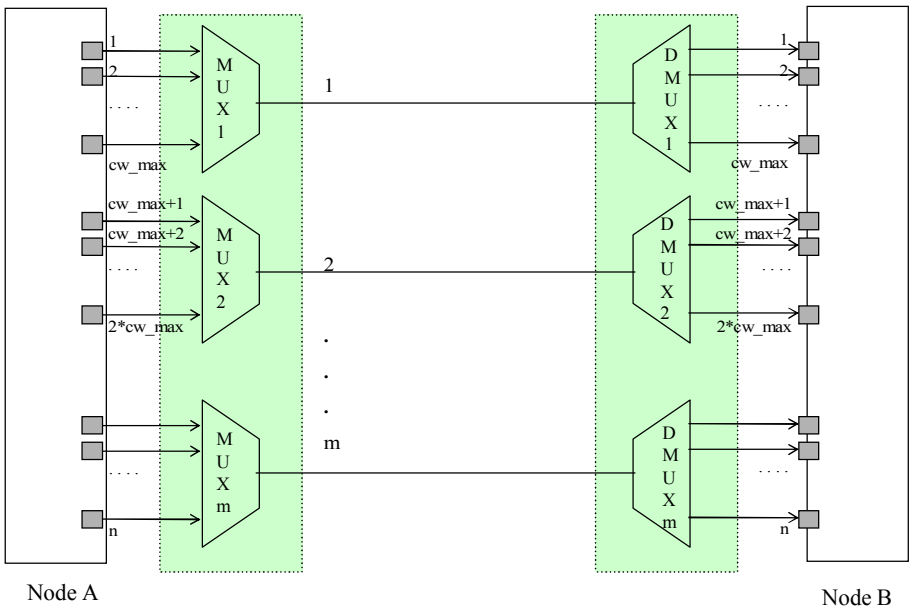
3.2.2 WDM connections

CWDM connection

Coarse Wavelength Division Multiplexing (CWDM) delivers multiple wavelengths over an optical fiber at a fraction of the cost of Dense Wavelength Division Multiplexing (DWDM). CWDM system is quite similar to DWDM system but it is simpler with the cost of about 40%~70% of DWDM system. It is a good fit for metro area network and practically supports 4~8 channels (wavelengths) in most widely distributed networks. CWDM is more suitable for networks with the following characteristics:

- Low channel count of 4 to 8 channels
- Transmission rates of <2.5 Gb/s per channel
- Short distances of <80 km, so no amplifier or regenerator is needed.

The connection of a CWDM system is shown in Figure 3-4.



Note: This figure only shows equipment required for signals traveling in one direction.

Figure 3-4 CWDM connection for local area

The total circuit cost is:

$$COST\_CWDM = \underbrace{2 \cdot n \cdot C_{IF\_SH}}_{\text{node interface cost}} + \underbrace{m \cdot d_{AB} \cdot C_f}_{\text{fiber cost}} + \underbrace{2 \cdot m \cdot C_{CWDM}(cw\_max) + 2 \cdot C_{CWDM}(cw\_last)}_{\text{CWDM equipment cost}} \quad (3.6)$$

Where,

cw_max: Maximum number of fibers that can be supported by CWDM equipment. Normally it is from 4~8.

m: number of CWDM equipment connected to one node. So we have: $m = floor(n / cw\_max)$, and the number of wavelengths that are supported by the last CWDM equipment is: $cw\_last = n - m \cdot cw\_max$.

C_{IF_SH}: Node switch interface cost in \$/each for Short Haul interfaces. This kind of interface can support very close transmission distance (within 0.5Km normally). It is much cheaper than the Long Haul fiber interface.

C_{CWDM}: an array of cost of CWDM equipment for different numbers of interfaces. Note that the cost here is for bidirectional transmission. For different interface number in one CWDM equipment from 1…cw_max, the cost would be $C_{CWDM}(1), C_{CWDM}(2), ..., C_{CWDM}(cw\_max)$.

Similarly, the total cost will increase as the distance increases but not linear because of the staircase increase of CWDM equipment cost. Suppose that CWDM equipment cost linearly increases with the number of interfaces from node switch and C_{CWDM_each} is CWDM cost for each wavelength, CWDM equipment cost would be:

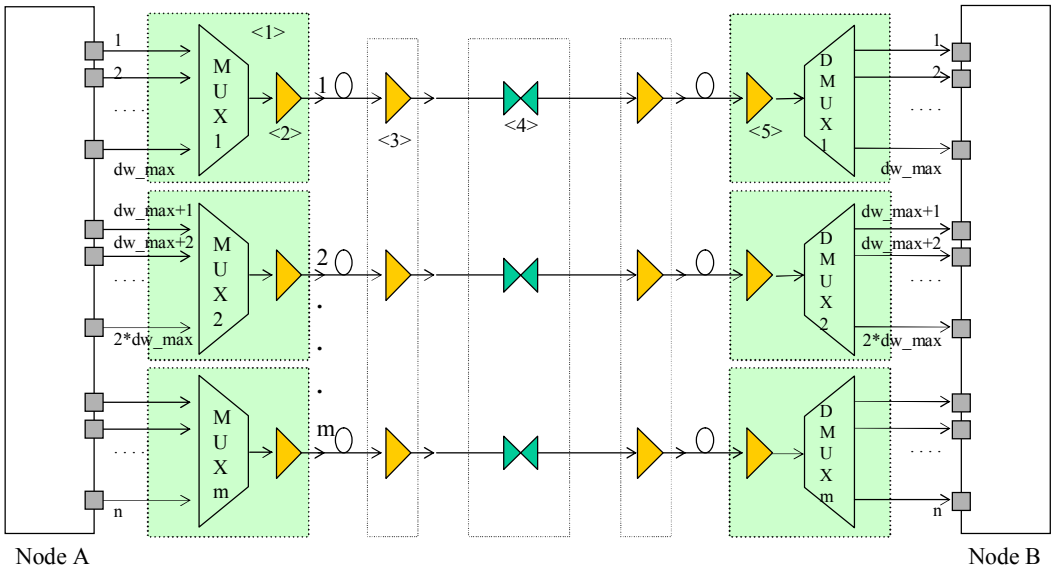
$$2 \cdot n \cdot C_{CWDM\_each} \quad (3.7)$$

We can approximate the total cost to a linear function w.r.t. distance as follows:

$$COST\_CWDM = \underbrace{2 \cdot n \cdot \left(C_{IF\_SH} + C_{CWDM\_each}\right)}_{\text{node interface cost}} + \underbrace{\left(n / cw\_max\right) \cdot d_{AB} \cdot C_f}_{\text{fiber cost}} \quad (3.8)$$

DWDM connection

With DWDM technology, we can transfer more wavelengths in one fiber over a longer distance. Currently the most commonly used technology is 8, 16, 32, 64 wavelength multiplexing. DWDM link architecture is shown in Figure 3-5.



<1> DWDM terminal equipment()
<2> BA (Booster Amplifier or Post Amplifier)
<3> LA (Line Amplifier)
<4> 3R Regenerator
<5> PA (Pre Amplifier)

Note: This figure only shows equipment required for signals traveling in one direction.

Figure 3-5 DWDM connection for broad area

With amplifiers and regenerators, the total cost of DWDM will be:

$$\begin{aligned} COST\_DWDM = & \underbrace{2 \cdot n \cdot C_{IF\_SH}}_{\text{node interface cost}} + \underbrace{m \cdot d_{AB} \cdot C_f}_{\text{fiber cost}} \\ & + \underbrace{2 \cdot m \cdot C_{DWDM}(dw\_max) + 2 \cdot C_{DWDM}(dw\_last)}_{\text{DWDM equipment cost}} \\ & + \underbrace{\left[ceil\left(\frac{d_{AB}}{max\_amp} \right) - 1 \right] \cdot C_{LA} + \left[ceil\left(\frac{d_{AB}}{max\_reg} \right) - 1 \right] \cdot C_{D\_Reg}}_{\text{cost of amplifiers and regenerators}} \end{aligned} \tag{3.9}$$

Where,

C_{LA}: cost of line amplifier, \$/amplifier for all wavelengths in one fiber.

C_{D_Reg}: cost of regenerator, \$/regenerator for all wavelengths in one fiber.

dw_max: Maximum number of fibers that can be support by DWDM equipment. Normally it is 8, 16, 32 or 64 etc.

m: number of DWDM equipment connected to one node. So we have: $m = floor(n / dw\_max)$, and the number of wavelengths that are supported by the last DWDM equipment is: $dw\_last = n - m \cdot dw\_max$.

C_{IF_SH} : Node switch interface cost in \$/each for Short Haul interfaces.

C_{DWDM} : An array of cost of DWDM equipment for different numbers of interfaces. Note that the cost here is for DWDM transceiver cost for bidirectional transmission on both ends. This cost includes the cost of booster amplifier and cost of pre amplifier. Interface number (bidirectional) in every DWDM equipment is from 1… dw_max, and the cost would be $C_{DWDM}(1), C_{DWDM}(2),..., C_{DWDM}(dw\_max)$.

C_{Reg} : cost of regenerator, \$/regenerator for all wavelengths.

A lot of researches have been done on the placement of amplifiers and regenerators to ensure error-free propagation and to minimize costs. This is also a complicated problem since the optimal placement would be influenced by a lot of factors such as fiber type, hut distribution, traffic demand etc. Here we just use an ideal scenario in which amplifiers and regenerators are placed at maximum distance span.

Similarly, with given traffic, we can approximate the total circuit cost to a linear function w.r.t. distance as follows:

$$COST\_DWDM = \underbrace{2 \cdot n \cdot \left(C_{IF\_SH} + C_{DWDM\_each}\right)}_{\text{node interface cost}} + \underbrace{\left(n / dw\_max\right) \cdot \left(C_f + \frac{C_{LA}}{max\_amp} + \frac{C_{D\_Reg}}{max\_reg}\right) \cdot d_{AB}}_{\text{line cost}} \tag{3.10}$$

Chapter 4 Network topological design for AAPN

Currently all photonic switches in development have a switching capacity dramatically larger than traditional electronic counterparts. Such a huge increase in capacity demands that a completely different philosophy and methodology be adopted for the design and dimensioning of networks.

As in [1], the Agile All-Photonics Networks (AAPN) Research Network (RN) is based on the observation that optical switching technologies will be introduced into optical networks. The practical paradigm in the near-term contains the following key ingredients: (1) rapidly reconfigurable all-optical space-switching in the core, (2) agility: the ability to perform time domain multiplexing to dynamically allocate bandwidth to traffic flows as the demand varies, and (3) control and routing functionality are concentrated at the edge switches that surround the photonic core. We refer to such networks as Agile All-Photonic Networks (AAPNs).

AAPN Theme 1, Networks and Architectures, is partitioned into two interrelated projects. The research under Project 1.1 on Network Architectures is oriented toward the overall evaluation of different network architectures and the consideration of different schemes for bandwidth sharing in the time domain. The research under Project 1.2 on Network Traffic Engineering deals with the detailed protocols that are required to control the different nodes of the photonic network (pure photonic core nodes and the electro-optical edge nodes) for the provisioning of the different classes of shared transmission services foreseen.

Our research is for Task 1.1.4 of AAPN to develop methods and tools for selecting network topologies including:

- Identification of existing methods and tools for selecting network node distributions and optimizing performance/cost parameters that can be suitable or adaptable to overlaid star/tree network architectures.
- Network models and their performance evaluation against the preliminary set of services.

This chapter will present solutions for this task. In order to compare the various design options, we need methods, models and computational tools to optimize and quantify equipment requirements and costs under different traffic demands and population distribution scenarios. Most of the content in this chapter has been previously reported in paper of [32] and the poster of [33].

4.1 Network architecture models

4.1.1 AAPN network infrastructure

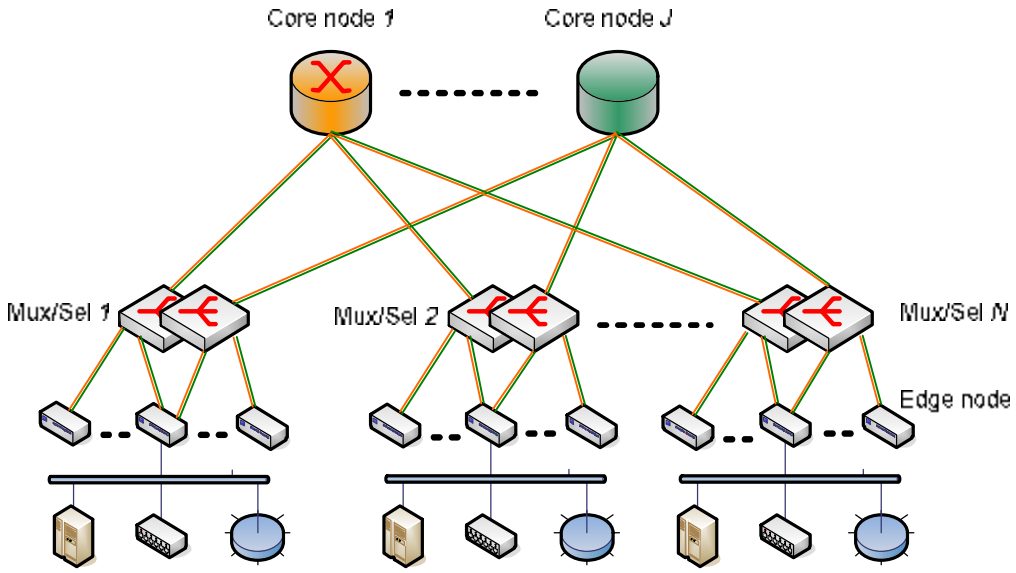


Figure 4-1 AAPN network model: three-layer overlaid tree topology

A review has been done to identify appropriate models for AAPN network topological design. Currently, many core networks have a mesh topology because it is robust and can distribute traffic load over switches. Initially in the definition for this task, we assume a two-layer design where core optical switch are connected directly to as many as 1000 or more edge nodes in an overlaid star configuration. However, this may result in more cost in the link connectivity and more optimal memory in core switches. Thus an investigation of a three or two-layer network topologies is required according to the number and traffic demands of edge nodes of the AAPN, which is referred as the overlaid star or tree topology. Figure 4-1 shows an example of the three-layer overlaid tree topology.

Stars and overlaid stars are robust to various traffic distributions. Dimensioning and performance is related to aggregate demand which is more easily to forecast. In tree topology, distributed core switches consist of a central stage with remote Mux/Sels located in the points of presence of the edge nodes. This configuration has the advantage of reducing the number of ports on the core switch as well as significantly reducing the quantity and cost of the transmission network linking the edge switches and the central core nodes. This requires additional synchronization of the Mux/Sels and the central core node switching function, relative to that of a single stage large optical core switch. Nevertheless the resulting tree topology obtained from the use of remote Mux/Sels can still be synchronized using a straightforward extension of the procedure originally developed for the star topology.

Traffic robustness is an important specification when designing a network. Here the definition for traffic robustness is that the network should be robust to variations in traffic distribution. The network should achieve high bandwidth efficiency with acceptably low blocking probability, as well as low delay. Accordingly for Metropolitan Area Network and Wide area Network, several scheduling alternatives may be applied such as slot by slot, frame by frame or call by call.

Due to the special structure resulting from the overlaid star/tree topology, the various network design algorithms reported in literature do not directly apply to the AAPN topological design problem, which is the initial motivation for task 1.1.4 in AAPN research.

Figure 4-1 only gives one of the general architectures for AAPN. In fact to realize an AAPN in Metro, Regional and National networks, the traffic demands, connectivity methods, device functionality and device inner fabric may be quite different. In [33], the alternatives architecture classes are described as symmetric and asymmetric circuit designs, and two and three layer networks designs with different device options.

4.1.2 Device options

As described in [33], the edge node and core node of AAPN are shown in Figure 4-2, Figure 4-3, Figure 4-4 and Figure 4-5 as below.

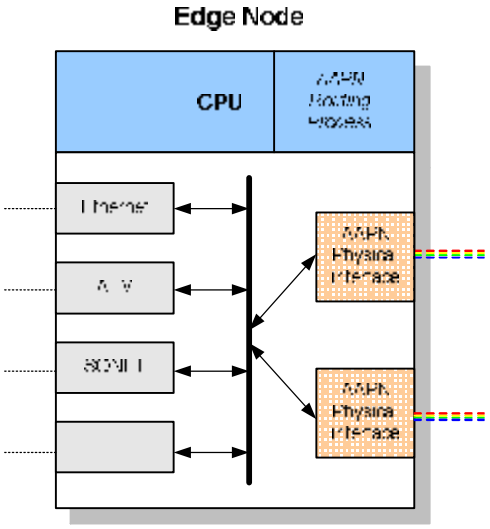


Figure 4-2 Edge node

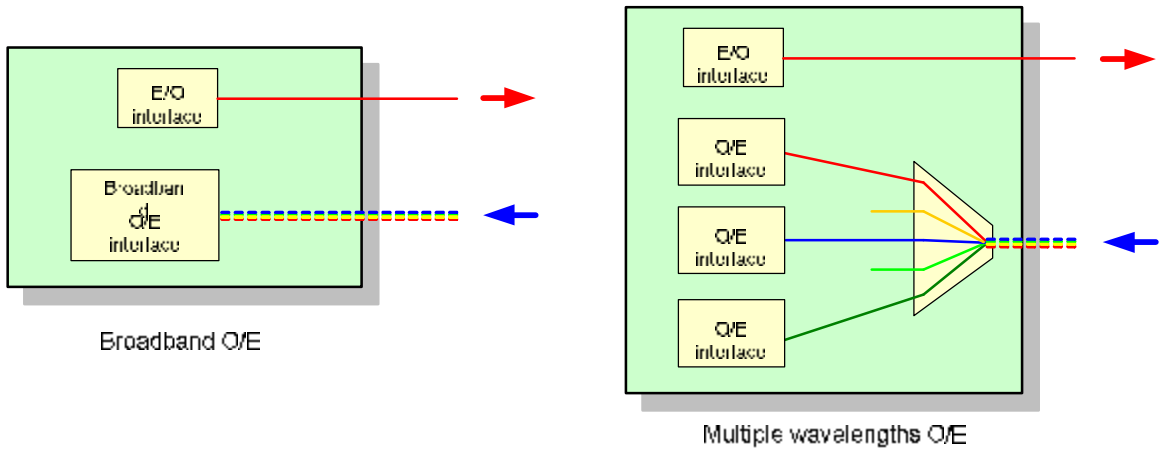


Figure 4-3 Edge node interface options

Figure 4-2 depicts the edge node fabric and Figure 4-3 depicts two kinds of interfaces of an Edge Node. The edge node incorporates the interface to legacy networks and to AAPN. In addition, the Edge Node will be equipped with transmitters (fixed or tunable) and receivers (broadband or multiband). Assume that one single color fiber has the capacity of 10Gb/s, single color (wavelength) in upstream direction is sufficient for the total traffic from one edge node to AAPN. In downstream direction, this element may be only one fiber with one color, or one fiber with multiplexed colors. For the latter case, de-multiplexing devices are necessary in edge nodes. In the following description, we refer to the upstream direction as the path taken from transmitter in edge node (Electrical to Optical) to core switch, and the downstream direction as the path taken from core switch to receiver in edge node (Optical to Electrical).

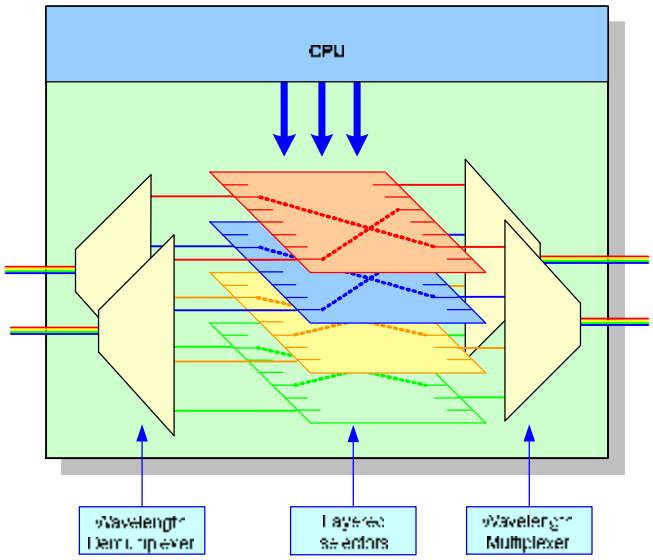


Figure 4-4 Fast switched core node

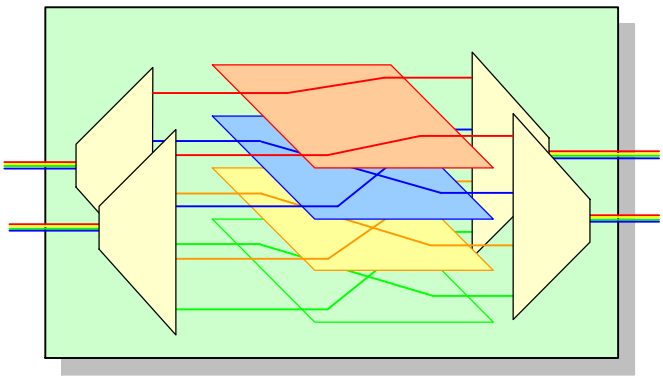


Figure 4-5 Passive switched core node

Figure 4-4 and Figure 4-5 depicts two kinds of core nodes. Core node is the Optical Switching equipment where the data switching is performed, which will be discussed in detail in section 4.1.4. One option switch is the symmetric architecture for both uplink and downlink. The picture shows that a core node includes a de-multiplexer, layered selectors and a multiplexer for both directions. Traffic of different colors is switched separately and there is no wavelength conversation. De-multiplexers and multiplexers enabled the combination of traffic from different colors. If edge node is using single color fiber to core node, it can also be connected to the switching plane of that specific color directly.

For core switch, according to the change frequency of switching methods, there are several options including fast-switched core, slow-switched core and passive core. According to the scheduling schemes, core nodes can use Optical Burst Switching (OBS) or synchronous Optical Time Division Multiplexing (OTDM) etc. The passive-optical switch is passive in that the traffic which is destined for another destination passes through the switch with predefined configurations. The switching methods can not be changed, regardless of the possible change of the traffic patterns. Correspondingly, in slow-switched optical core switch, switch methods are updated automatically after a certain period, and fast switched optical switch can change its switching operation momentarily. Thus, as shown in Figure 4-4 and Figure 4-5, fast switched core node needs CPU to calculate the switching connection while passive core does not.

Blocks of data are transmitted in the form of slots of 10 microseconds' duration. This is because the data switching time needed by current switch is 1 ms and another 9 ms are needed for the transmission.

The traffic transmission process from the Electrical to Optical to Electrical can be described as follows:

- Traffic inbound from existing networks is sorted by destination edge node and placed in a corresponding Virtual Output Queue (VOQ) in order of arrival after passing

through an adaptation layer that performs the necessary segmentation and framing functions.

- Data slots are read out from these electronic VOQ buffers. Under electronic control, at the appropriate time that is governed by the scheduling algorithm, data are converted to an appropriate wavelength in the optical domain.
- These optical data slots are launched into the photonic system where they are space switched at their time of arrival towards the destination edge node corresponding to the VOQ from which the data slots originated.
- The data slots remain on the same wavelength channel or light path throughout their journey until reaching the destination node where multiband or broadband receivers convert the received optical signal back to the electronic domain where slot reassembly is performed to reconstruct the data back to the form it had prior to entering the AAPN.
- The electronic data in its native form is then routed to its appropriate destination legacy network.

As the AAPN network infrastructure indicates, the AAPN architecture can be designed as three-layer networks, with groups of edge nodes homing on Mux/Sels via single fibers or DWDM. The Mux/Sels in turn home on a specific core node via DWDM. Figure 4-6 depicts the symmetric Mux/Sel, Figure 4-7 depicts the asymmetric Mux/Sel and Figure 4-8 depicts the asymmetric Mux/Sel with broadcast. Asymmetric circuit design means that upstream transmission path is not the inverse of the downstream transmission path with respect to the core switch, while the symmetric design has the same paths for both upstream and downstream directions. These three kinds of Mux/Sels lead to different circuit design options as will be discussed in the next section.

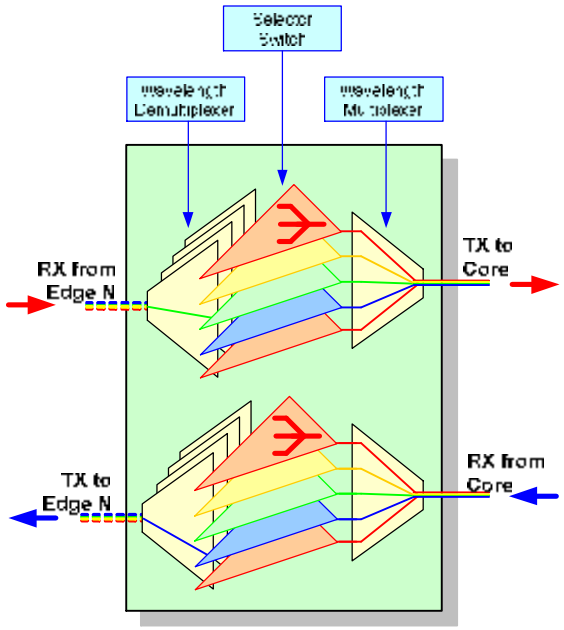


Figure 4-6 Mux/Sel: Symmetric

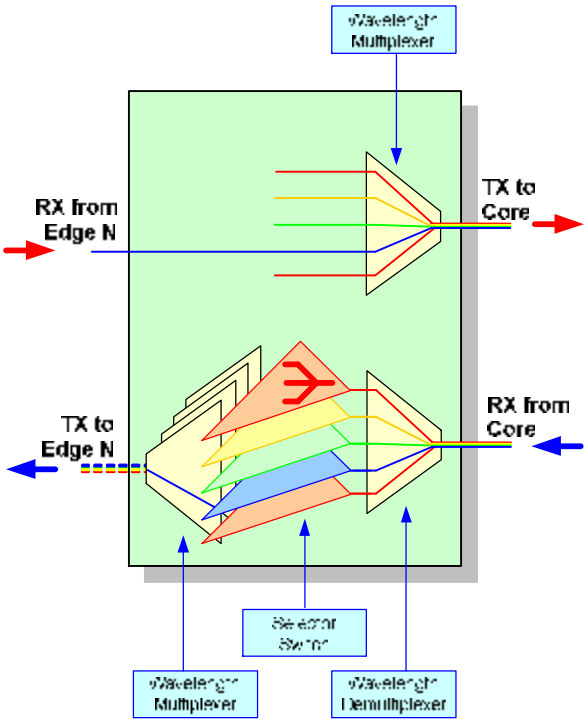


Figure 4-7 Mux/Sel: Asymmetric

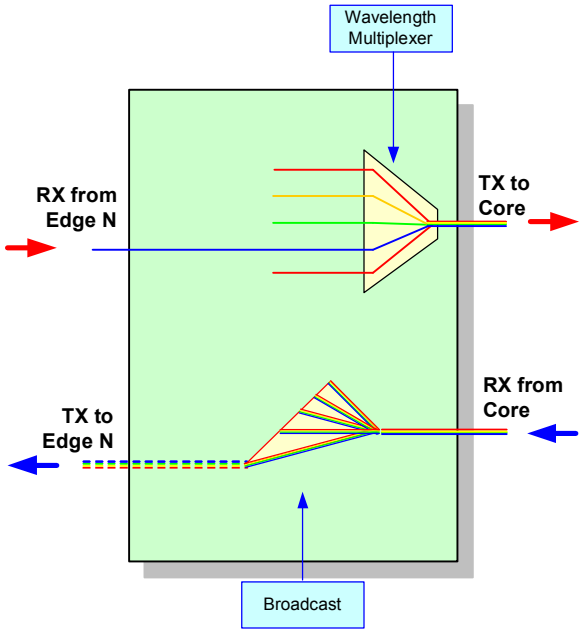


Figure 4-8 Mux/Sel: Asymmetric with broadcast

The Mux/Sel is also called concentrator, multiplexer or selector switch depending on its functionalities. As shown in Figure 4-6, the symmetric Mux/Sel has an active selector with the wavelength multiplexer and de-multiplexer in both upstream and downstream directions. In Figure 4-7, the asymmetric Mux/Sel has an active selector with the wavelength multiplexer and de-multiplexer in the downstream direction, and only a multiplexer in the upstream direction. In Figure 4-8, the asymmetric with broadcast Mux/Sel uses a multiplexer in the upstream direction and a passive broadcast Optical Star Coupler in the downstream direction. Albeit the symmetric Mux/Sel is the most expensive one among the three, it has the most flexible functionalities and can support symmetric traffic demands from edge nodes. The asymmetric Mux/Sel requires edge nodes to have single color fiber for upstream traffic. The asymmetric with broadcast Mux/Sel uses optical star coupler in downstream, where light from one incoming core node is combined and broadcast to N (N=8, 16, 32 etc) outgoing fibers with an intrinsic 1/N power loss. The asymmetric Mux/Sel is much cheaper than the symmetric one, and the broadcast Mux/Sel has the lowest price.

As figures in this section show, core switches and Mux/Sels have several model options for the device type when used in different situations. These options may influence the end to end traffic transmission and thus influence the layered design. One thing that should be noticed is that in our topological design scheme, no matter what kind of core switch or Mux/Sels is used, once the layered architecture has been determined, the topological design procedure will be the same for all of them. The only change is that we need to use different cost values for different devices.

4.1.3 End to end connection models

According to the layers of the network, the alternative architectures can be classified as two-layer and three-layer network design. While based on the circuit options, network is described as symmetric and asymmetric designs. In [33], the end to end connections of these architectures are shown in the following figures:

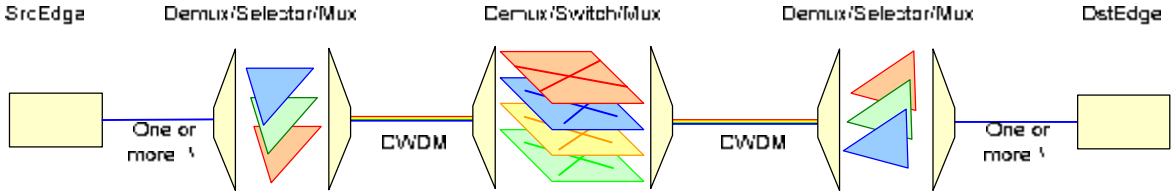


Figure 4-9 Symmetric Architecture

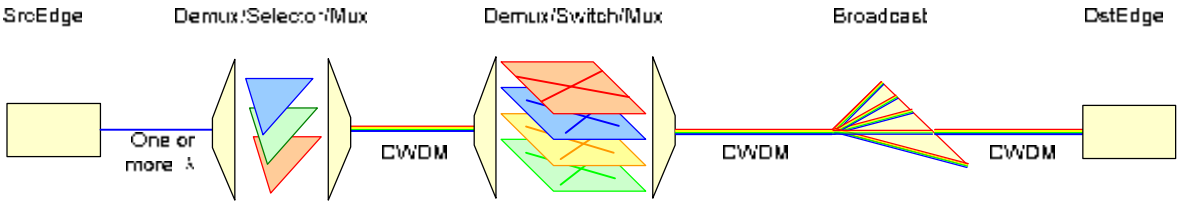


Figure 4-10 Active selector switch in upstream with broadcast in downstream

The symmetric designs are shown in Figure 4-9. The symmetric architecture is the most expensive one among all design options and it can realize the highest traffic robustness. Traffic routing is exactly the same for both upstream and downstream. The selector switch is an active device which requires color synchronization from edge to selector switch. In symmetric design, connections from edge to Mux/Sel use single color fiber while connections from Mux/Sel to core use DWDM links. In Figure 4-10, the downstream Mux/Sel is replaced with simple broadcast star coupler which is less expensive. At the same time, this architecture can still have high traffic robustness.

For the asymmetric case in Figure 4-11 and Figure 4-12, a circuit in the upstream direction is comprised of an: E/O edge node; fixed wavelength single fiber; Lambda multiplexer; DWDM Fiber; then core node with wavelength de-multiplexer and layered space switch. For the downstream direction there are also two options of Mux/Sel and broadcast star coupler. The former has single color fiber connected to edge nodes which is suitable for symmetric traffic demand while the latter uses DWDM connections to edges which provides more bandwidth for downstream traffic. Furthermore, the broadcast star coupler can realize broadcast easily.

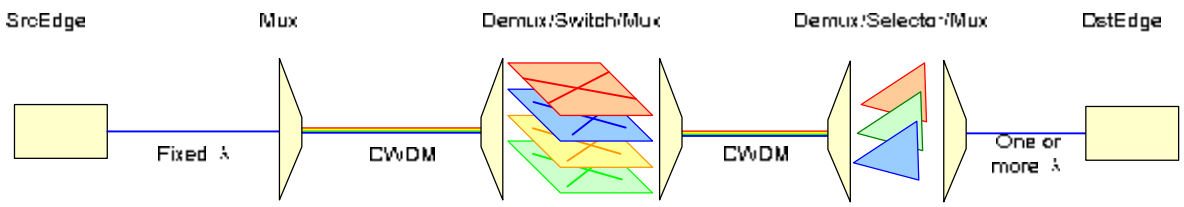


Figure 4-11 Asymmetric Architecture

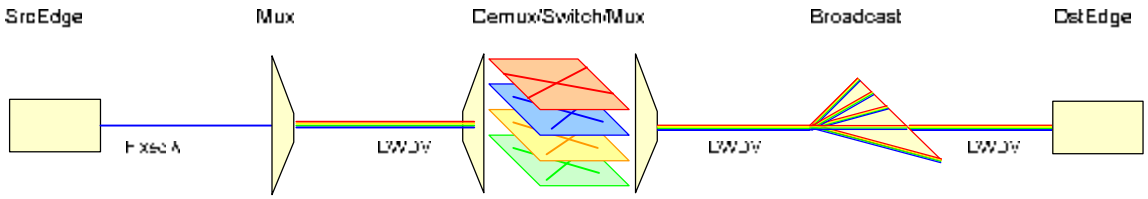


Figure 4-12 Asymmetric with broadcast in downstream

Compared with symmetric designs, several advantages of such an asymmetric design are as follows:

- There is no need to synchronize time slots in distinct layers (colors) in the core switches or selectors (potentially less complex and less costly).
- There is no need to co-ordinate scheduling strategy across multiple core layers in core switches and selectors (thereby reducing complexity for scheduling with some potential loss of traffic efficiency).
- Reduction in number of devices on end-to-end path (less transmission loss and network cost).
- Broadband O/E in edge node can be replaced with demultiplexer + separate O/E converters for each received color (will increase edge node's receiving capacity).

Now consider the symmetric circuit design option. To exploit the additional flexibility in wavelength assignment that is made possible by the availability of the Mux/Sel in the upstream direction, the transmitter should be tunable. In this case, improvement in utilization over the asymmetric fixed laser design can be obtained by computing a coordinated schedule across both time slots and wavelengths. This however requires synchronizing all wavelength-switching planes in both clock rate and phase. To achieve the efficiency improvements given by the additional wavelength flexibility, the scheduling computation will be more complex as it must be performed in a unified way across the wavelength switching planes. On the positive side symmetric design enables time-sharing a wavelength across distinct edge nodes in the upstream direction, which may be desirable if there isn't sufficient traffic emanating from a single edge node to fully occupy a frame. The fixed allocation of wavelengths to ports of the passive multiplexer used in asymmetric design precludes such flexibility in bandwidth sharing of the link between the Mux/Sel and the core

node.

On the other hand, for the asymmetric case we can perform several independent scheduling calculations in parallel, one for each wavelength. For OTDM slot-by-slot scheduling, an attractive alternative to OBS, the global schedule must be computed in less than a slot time of 10 microseconds. We thus conclude that the asymmetric overlaid star design is well suited to the fast switching times associated with fast slot-by-slot OTDM scheduling.

For the asymmetric circuit design with fixed wavelength lasers and multiband receivers, each of these sub-networks can be synchronized independently, that is, the clock can differ in phase among component sub networks. As an alternative, if a bank of several separate narrow band receivers are employed then the receive capacity can be increased without having to synchronize across separate wavelength switching planes. Thus for asymmetric circuit design there are two receiver design options.

The above end-to-end connections are for three-layer network design. For two-layer network, there are two choices as shown in Figure 4-13 and Figure 4-14.

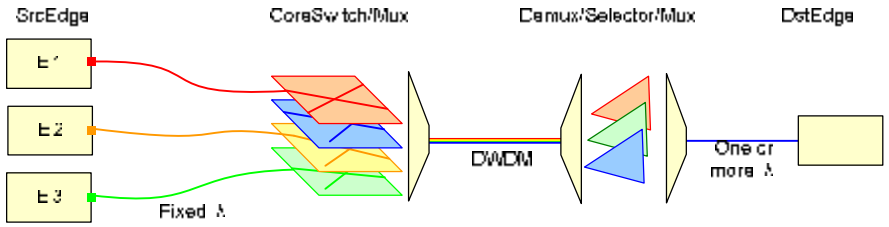


Figure 4-13 Two-layer network architecture

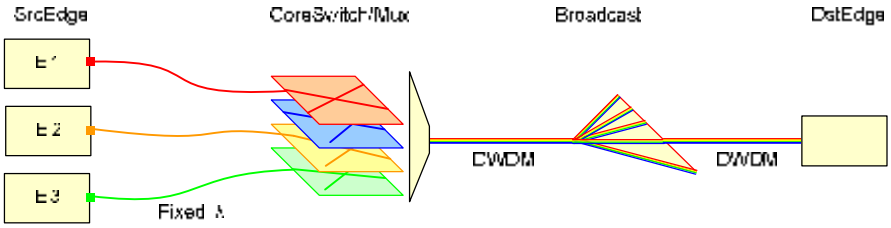


Figure 4-14 Two-layer network with broadcast in downstream

It is clearly shown in the two figures that in two-layer design for upstream traffic, source edge node ports are connected directly to core node switching plane without any intermediate aggregation. Thus there is no upstream Mux/Sel cost and cost for core switch will also be lower. However for downstream, the DWDM link from core nodes may not be able to connect to edge nodes directly. This depends on the edge number in the network. For example, if the selector switch plane in the core node is a 64×64 crossbar, there are at most 64 DWDM links for the downstream direction. Thus if there are more than 64 edges, Mux/Sel or broadcast star coupler is necessary for the downstream direction, which is different from the upstream direction. So in fact, the two-layer design mentioned here is a combination of two

and three-layer design, with two-layer in the upstream direction and three-layer in the downstream direction.

4.1.4 Optical switching options

There are many possible protocols for managing the bandwidth and buffering for such hardware structures, which are transparent optical paths. Recently considerable attention has been directed to time division multiplexing of light paths using asynchronous Optical Burst Switching (OBS), Optical Packet Switching (OPS) architectures, and synchronous Optical Time Division Multiplexing (OTDM) techniques as means of increasing network agility and reach. Introducing time-domain multiplexing is very challenging because switching requests come from multiple sources and the optical space-switches must be configured correctly before the arrival of the data to be switched. To efficiently transport bursty traffic such as found in the Internet, fast optical switching is required to time-share light paths. In addition the reduced granularity of the data volume carried in a time slot, packet or burst switching increases the potential reach of all optical networks to smaller aggregation points closer to the traffic demand sources.

In [32], the switching methods for an AAPN have been discussed. In the topological design and dimensioning of an AAPN, we use Optical Time Division Multiplexing (OTDM) and Overlaid Star and more generally, Overlaid Tree Topologies from both the cost and performance perspectives. This is because OTDM networks can be non-blocking while OBS networks cannot. Apart from the design simplicity and insensitivity to traffic distribution forecast errors, optical star networks are readily amenable to the synchronization required for OTDM. It can be argued that a variant of OTDM called synchronous slot-by-slot switching is a viable alternative to Optical Burst Switching (OBS) for the supporting bursty traffic in next generation all photonic transport networks. Synchronous slot switching in photonic networks by necessity requires global synchronization due to the lack of cost effective optical memory. This global synchronization requirement places restrictions on the class of network topologies that can easily support synchronous OTDM to composite star/tree networks and potentially ring topologies. Accordingly in order to evaluate and compare synchronous OTDM with other alternatives such as OBS one needs to design networks which are suitable for supporting OTDM. The topology design tool we have developed enables a planner to examine the impact of device and architectural design decisions and device costs on the resulting network designs, equipment quantities and costs by category for different traffic demand scenarios.

4.2 Discussion of integrated vs. tiered network design

Following Lorne Mason's proposal which is discussed in [32], the following two topological design approaches are investigated: Integrated Network Design (IND) and Tiered Network Design (TND).

In the IND approach, the entire network topology, including metropolitan and wide area, is determined in an integrated fashion, given the population distribution and traffic model. In this approach we employ a weighted objective function criterion consisting of network cost and network performance components. The appropriate value of the relative weights applied to the cost and performance terms in the scalar objective function is determined by a multi-criterion optimization methodology. This approach treats all traffic sources in a homogeneous manner attempting to minimize overall network cost while minimizing average point-to-point propagation delay, the principle source of performance impairment.

In the TND approach, we apply a tiered design methodology, whereby we design the local (metro) areas first, then regional networks and finally the national network. While this approach can result in somewhat larger network costs than the IND approach, it tends to provide a more uniform point-to-point propagation delay within the component networks. For example metro networks will have low propagation delay among all edge nodes within the metro area. This design approach would then facilitate the support of services requiring low propagation delay within a metro region, which could not be achieved in a national network due to the distances involved.

To explain the different design procedures and methods used in these two approaches, we will give an example of the Canadian network.

In the TND, we can define three kinds of core nodes as "local", "regional" and "national".

For local area of the city zone, core nodes can be allocated in densely populated cities to support edge nodes within the city. In Canadian network, Local Core Nodes are most likely placed in big cities like Toronto, Montreal and Vancouver.

Regional area is a province or state range area when a suitable distance range is selected. In Canada, as the population distribution shows, there should be one range in east Canada, and, maybe another region in west Canada. Edge nodes within one region will be connected to the Regional Core Nodes. If the region covers local areas, Local Core Nodes will transmit the traffic within edge nodes of the same city, and Regional Core Node will exchange the traffic to other edges in the region.

Similarly National Core Nodes can be allocated. Please be noted that in this procedure, there is no distance limit.

In this procedure, the number of core nodes can be fixed or be calculated in each step. When core node number is fixed, enumeration calculation can be used to find the optimal placement of core nodes. When core node number is not fixed, optimization methods will be employed to allocate core nodes and to find out connectivity between core nodes and edge nodes.

In this thesis, more emphasis has been put on the ITD approach, which is studied in details in the following chapters.

4.3 Working network design

The first step of our working network design is to define the elements in the circuit path in both directions of downstream and upstream. The original data we have is the demographic data of population distribution and the rough traffic demands per end user. In our research, two cases are studied. One is the Canadian network and the other is a metro network in a virtual city named Gotham. From previous internal work of Anton Vinokurov and Fariba Heydari, population data were collected and correlated with location information in the Canadian Network. And Anton Vinokurov has generated the virtual population distribution of Gotham city using the Gaussian distribution.

The proposed topological design methodology consists of a two-stage heuristic design process. First the insertion of remote Mux/Sels in the network architecture naturally leads to the question of their optimal placement, size and homing patterns. The Mux/Sel location problem is closely related to the classical plant and warehouse location problems, called Capacitated Plant Location Problem (CPLP) or the *P-median* problem considered in the Operations Research literature. Secondly, we determine the number and location of the central core switching stages and the multiplexer/selector to core switch homing pattern using the new ILP formulation developed for the overlaid star/tree architecture.

Our model for topological design can be used for various multiplexing methods such as various TDM based scheduling schemes and Optical Burst switching when suitable performance models are available.

4.3.1 Edge node location and cabling infrastructure

In this step, beginning with the population distribution for the area served, we first determine the location of the edge switches. Given the predicted originating and terminating traffic/user, at edge nodes we compute the total originating and terminating traffic per edge node.

Detailed discussion about traffic models is given in Section 3.1. We compute the mean traffic demand between edge nodes using a forecasted originating/terminating traffic per node along

with a gravity model and community of interest factors obtained from the literature, to estimate how that ingress node's traffic is distributed across the egress nodes of the AAPN network.

In the private comments, Anton Vinokurov and Fariba Heydari have collected demographic data for the country of Canada to represent the wide range of potential applications of the AAPN optical core transport network. All the population is supposed to be in cities as points in the topology. For the Canadian network we have obtained location data, and the population data are then correlated with (x,y) co-ordinates to enable computation of the lengths of transmission links, necessary for topological cost/performance optimization, as well as point to point traffic demand forecasts. After this step, we have 1024 edges that are distributed in 140 cities using the flat traffic demand model; each edge serves around 30,000 customers, and each customer in turn is generating a 30Mbits/s data stream in both directions.

One thing to point out is that in national network, we assume the edges in one city are collocated with each other. This is because the distance range of a city is normally less than 50 km, which is neglectable comparing with the average 2000 km distance of the national network nodes. This assumption will reduce complexity.

With the absence of fiber embedding information, our design starts with a country/city with conduits ready but without cables and fibers. The Floyd-Warshall algorithm is used to compute all pairs of city-to-city distances needed for the minimization calculations.

Furthermore, in order to compare network topology designs for global and metro networks, Anton Vinokurov "simulated" an artificial city in absence of real data of a city in his report for AAPN. In his model, the population density distribution is simulated with Gaussian random process, and edge nodes are distributed according to population-per-location in a Manhattan network layout. The properties of Manhattan network can be found in [38,39]. Cable infrastructure is simulated with a modified spanning tree algorithm. After the simulation, we have 300 edge nodes and the "edge-edge" distance matrix can be calculated.

With the input population distribution for the area being served (Canada and Gotham in our example), we can get the output of locations for edge nodes in the first step.

4.3.2 Multiplexer/Selector allocation

The problem of finding the location of Mux/Sels and connectivity between Mux/Sels and edge nodes, to serve a set of edge nodes at a minimum total cost is a Single Source Capacitated Plant Location Problem (SSCPLP). In this step we use an Integer Linear Programming (ILP) formulation as follows:

$$\min \underbrace{Cm \cdot \sum_{j=1}^{J} z_j}_{\text{MUX equipment cost}} + \underbrace{\sum_{j=1}^{J}\sum_{i=1}^{N} y_{ij} \cdot cost_{ij}}_{\text{link connectivity cost}} \tag{4.1}$$

$$\sum_{j=1}^{N} y_{ij} = 1 \quad i = 1...N, j = 1...J \tag{4.2}$$

$$K \cdot z_j \geq \sum_{i=1}^{N} Ei_i \cdot y_{ij} \quad i = 1...N, j = 1...J \tag{4.3}$$

$$2 \cdot z_j \leq \sum_{i=1}^{N} Ei_i \cdot y_{ij} \quad i = 1...N, j = 1...J \tag{4.4}$$

$$\underbrace{z_j \in \{0,1\}}_{\text{J variables}}, \underbrace{y_{ij} \in \{0,1\}}_{\text{N*J variables}} \quad i = 1...N, j = 1...J \tag{4.5}$$

Where,

J is the set of potential Mux/Sel sites and N is the set of edge nodes. In our model, we suppose that Mux/Sel can only be allocated in the place where there are already one or more edge nodes. So $N=J$. However, we can also set that potential Mux/Sel sites belong to a subset of edge node sites, where $J<N$.

$cost_{ij}$ is the cost of supplying all of edge demand in node i from Mux/Sel in node j. This can be calculated according to the distance and connection method (direct fiber or DWDM) between each edge-Mux/Sel pair.

Cm is the fixed cost of a Mux/Sel and K is its maximum capacity if it is open.

Ei_i is the demand from node i. (In national network, it is the remainder of edge node number in city i after divided by K. In Gotham network, we just set it to 1.)

The binary variable z_j is equal to 1 if Mux/Sel j is open and 0 otherwise.

The binary variable y_{ij} is equal to 1 if edges in node i are connected to Mux/Sel in j and 0 otherwise.

The constraint (4.2) is the demand constraint, which means all demands should be served. The constraints (4.3) and (4.4) are the capacity constraints, which mean that one Mux/Sel can serve at most K edges and at least 2 edges.

Analyzing the formula (4.1), we can expect that in wide area network, where distance is a large number and thus the cost of connectivity covers a bigger proportion of the total cost, the optimal design tends to have a lot of Mux/Sels to make them near the edge nodes. In contrast, in metro network, where equipment cost dominates the total cost, the optimal design may have less Mux/Sels and more links. Thus, the Mux/Sel allocation problem can be considered in two variations. One is the capacitated one, where we have a fixed number for total number

of Mux/Sels, which is a *P-median* problem. For example, when we use this formulation in Canadian national design with 1024 edges, we can fix the total number of Mux/Sels to 1024/16=64 if we use 16-interface Mux/Sels. And for metro network with 300 edges, the total number of Mux/Sels can be fixed as *ceil*(300/16)=19. Another one is the relaxed one, where there is no fixed total Mux/Sel number, which is a SSCPLP problem.

There is one thing that is not neglectable in the three-layer national network design. As we separated the two steps of Mux/Sel allocation (when it is a SSCPLP problem) and core allocation and our target being to minimize the overall cost, we should consider the best balance of costs from these two parts. As we analyzed, lots of Mux/Sels widely distributed can achieve less cost when no fixed Mux/Sel number in the SSCPLP. But this will result in more cost in connectivity and interface in core node allocation design. To solve this problem and find the least cost, we can try different cost values for Mux/Sel (Cm) and obtain different Mux/Sel allocations, then continue to the second step to get the corresponding core node allocation and then calculate the overall costs. CPLEX is used to solve this SSCPLP problem. However, because different Mux/Sel costs will result in different designs, they need different computation time in CPLEX. When we have a very big number for Mux/Sel cost, the CPLEX program may stop because of running out of memory. To get the allocation if Mux/Sel for *P-median* problem with 64 Mux/Sels (which is the design with least Mux/Sel number), Simulated Annealing (SA) algorithm is used.

While in metro network design, there are 300 edge nodes without collocating and it cannot be solved with CPLEX because of the limit of memory. So we can solve it with Lagrangian Relaxation. In pure three-layer metro network design, we have one more constraint on the total number of Mux/Sel, which is the least number to guarantee that all edges will be connected. As just discussed, this is because in metro network, the connectivity cost is much less compared with Mux/Sel equipment cost. So less Mux/Sel can reduce cost. In this case the previous constraint (4.4) is no longer necessary. We can rewrite the formula for the *P-median* problem as follows:

$$\min \underbrace{Cm \cdot \sum_{j=1}^{J} z_j}_{\text{MUX equipment cost}} + \underbrace{\sum_{j=1}^{J}\sum_{i=1}^{N} y_{ij} \cdot cost_{ij}}_{\text{link connectivity cost}} \tag{4.6}$$

$$\sum_{j=1}^{N} y_{ij} = 1 \quad i = 1...N, j = 1...J \tag{4.7}$$

$$K \cdot z_j \geq \sum_{i=1}^{N} Ei_i \cdot y_{ij} \quad i = 1...N, j = 1...J \tag{4.8}$$

$$\sum_{j=1}^{N} z_j = P \quad j = 1...J \tag{4.9}$$

$$\underbrace{z_j \in \{0,1\}}_{\text{J variables}}, \underbrace{y_{ij} \in \{0,1\}}_{\text{N*J variables}} \quad i = 1...N, j = 1...J \tag{4.10}$$

Where P is a constant and is defined as: $P = ceil\left(\sum_{i=1}^{N} Ei_i \Big/ K \right)$

The LR approach to solve this problem can be described in the following steps:

Step1: Lagrangian multiplexing

First we form the Lagrangian function by adjoining the constraints multiplied by the multipliers (dual variables). We use dual variables $\alpha_j, j = 1...J$ to relax constraints (4.8).

The relaxed function is:

$$\begin{aligned} W\left(z_j, y_{ij}, \alpha_j\right) &= Cm \cdot \sum_{j=1}^{J} z_j + \sum_{j=1}^{J}\sum_{i=1}^{N} y_{ij} \cdot cost_{ij} + \sum_{j=1}^{J} \alpha_j \cdot \left(\sum_{i=1}^{N} Ei_i \cdot y_{ij} - K \cdot z_j \right) \\ &= \sum_{j=1}^{J} z_j \cdot \left(Cm - K \cdot \alpha_j \right) + \sum_{j=1}^{J}\sum_{i=1}^{N} y_{ij} \cdot \left(\cos t_{ij} + \alpha_j \cdot Ei_i \right) \end{aligned} \tag{4.11}$$

Now, we can use iteration method to solve this optimization problem. After relaxation, we have separated minimization problems into two:

- Separated problem 1:

$$\min \sum_{j=1}^{J} z_j \cdot \left(Cm - K \cdot \alpha_j \right) \tag{4.12}$$

subject to:

$$\sum_{j=1}^{N} z_j = P \quad j = 1...J \tag{4.13}$$

- Separated problem 2:

$$\min \sum_{j=1}^{J}\sum_{i=1}^{N} y_{ij} \cdot \left(\cos t_{ij} + \alpha_j \cdot Ei_i \right) \tag{4.14}$$

subject to:

$$\sum_{j=1}^{N} y_{ij} = 1 \quad i = 1...N, j = 1...J \tag{4.15}$$

Step 2: Lower bound $\underline{W}$:

Since all the variables are binary, we can easily get the optimum value from the coefficients. From separated problem 1, we have:

$$z_j^{*(n)} = \begin{cases} 1 & (Cm - K \cdot \alpha_j) \quad \text{in those } j\text{'s with respect to the smallest } P \text{ values} \\ 0 & \text{otherwise} \end{cases} \tag{4.16}$$

For a specific edge node i, (there are N different cases), we will do the following:

We should minimize $\sum_{j=1}^{J}\sum_{i=1}^{N} y_{ij} \cdot \left(\cos t_{ij} + \alpha_j \cdot Ei_i\right)$ with respect to y_{ij}, subject to: $\sum_{j=1}^{J} y_{ij} = 1$.

Since y_{ij} can only be 0 or 1, we just need to calculate $\cos t_{ij} + \alpha_j \cdot Ei_i,\ j = 1...J$ and find the minimal of it for each j. Then the corresponding $y_{ij}^{*(n)}$=1, and all the others =0.

In this manner, we can get $y_{ij}^{*(n)}$.

Plugging these values of $z_j^{*(n)}$ and $y_{ij}^{*(n)}$ into the Lagrangian relaxed function gives the dual function value at the current iteration (n). Its value gives a current lower bound $\underline{W}$.

Step 3: Upper bound $\overline{W}$:

Find the feasible solution. We use the value of $z_j^{*(n)}$ to calculate the feasible y_{ij}.

Although we can not call for CPLEX in MATLAB directly, some commercial software has been developed for the optimization functions in MATLAB. Here we used the software of TOMLAB for calling CPLEX functions in MATLAB[3].

Step 4: Subgradient:

Taking partial derivatives with respect to the dual variables gives a subgradient.

$$\frac{\partial W^{(n)}}{\partial \alpha_j} = \sum_{i=1}^{N} Ei_i \cdot y_{ij}^{*(n)} - K \cdot z_j \quad j = 1...J \tag{4.17}$$

Thus we can update the multiplier values. Let π denote a vector of all dual variables, i.e. $\pi \overset{\Delta}{=} \{\alpha_j\} \quad \forall j = 1...J$. In this simplified notation we state the update rule as $\pi^{(n+1)} = \max\left\{\pi^{(n)} + t_n \frac{\partial W(\pi)}{\partial \pi^{(n)}}, 0\right\}$, where $(n+1)$ denotes $(n+1)$st iteration, t_n is the step

[3] Detailed discussion about this method can be found in http://tomlab.biz/products/cplex/.

size, which is a constant for all variables at iteration *n*. The step size t_n is computed from the equation:

$$t_n = \frac{\rho\left(\bar{W} - W\left(\pi^{(n)}\right)\right)}{\left\| \frac{\partial W^{(n)}(\pi)}{\partial \pi^{(n)}} \right\|^2} \tag{4.18}$$

Here $\bar{W}$ is the best upper bound which we calculated before (i.e. F_{best}), while $W\left(\pi^{(n)}\right)$ is the current lower bound.

$0 \le \rho \le 2$ is a number set by program, which is used to adjust the step size according to the progress in previous iterations. For example, we can set $\rho = 0.5\rho$ if there is no progress after three iterations.

$\left\| \frac{\partial W^{(n)}(\pi)}{\partial \pi^{(n)}} \right\|^2$ is the squared norm of the subgradient vector, computed at iteration (*n*).

As the LR approach is also used in core node allocation, the detailed discussion bout the LR approach such as the calculation performance, the selection of subgradient method etc will be discussed in section 4.3.3.

After enough runs of LR, the Mux/Sel allocation data will be put in CPLEX. Then we can use the Matlab-CPLEX method (which is the "Hamburger Heuristic" method as will discuss in Section 5.1) to get the optimal result. Results show that our near-optimal algorithm in getting upper bound is quite efficient.

4.3.3 Core node allocation

In this step, a mixed integer linear programming problem for the core switch node location has been formulated. New heuristics have been devised for large-scale problems. The problem is solved with a LR approach which is quite similar to the method for Mux/Sel allocation in section 4.3.2. Heuristic designs have been compared to the optimum solutions obtained from CPLEX or enumeration calculation for small-scale problems. The accuracy comparison will be discussed in Section 5.1.

Mixed Integer Linear Programming Formulation for core allocation is as follows:

$$\min\left(\sum_{j=1}^{J} z_j \cdot C_{core} + \sum_{j=1}^{J}\sum_{i=1}^{N}\left(c \cdot d_{ij} + C_{coreIF\_MUX}\right)\cdot y_{ij} + \sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=1}^{J}\left(d_{ij} + d_{jk}\right)\lambda_{ik}\alpha_{ik}^{j} w \right) \tag{4.19}$$

Subject to the capacity constraint and traffic demand matrix for shortest path routing in both inbound and outbound directions given as below:

$$\sum_{j=1}^{J} \alpha_{ik}^{j} = 1 \quad i,k = 1...N, j = 1...J \tag{4.20}$$

$$N \cdot y_{ij} \geq \sum_{k=1}^{N} \alpha_{ik}^{j} \quad i,k = 1...N, j = 1...J \tag{4.21}$$

$$N^2 \cdot z_j \geq \sum_{i=1}^{N} \sum_{k=1}^{N} \alpha_{ik}^{j} \quad i,k = 1...N, j = 1...J \tag{4.22}$$

$$\underbrace{z_j \in \{0,1\}}_{\text{J variables}}, \underbrace{y_{ij} \in \{0,1\}}_{\text{N*J variables}}, \underbrace{\alpha_{ik}^{j} \in \{0,1\}}_{\text{N}^2\text{*J variables}} \quad i,k = 1...N, j = 1...J \tag{4.23}$$

This formulation is applied when Mux/Sels are already allocated in three-layer design. The design formulation starts with the traffic sources (Mux/Sels) *i*, where *i*=1…*N* is the Mux/Sel index in the set of *N* Mux/Sels.

From results for Mux/Sel allocation, we have the factor of traffic generated per source terminal, i.e. Mux/Sel *i*. The traffic is symmetric, which means inbound and outbound bit rates of Mux/Sel i are equal. We use λ_{ik} to denote traffic demands between Mux/Sel *i* and Mux/Sel k, hence $\lambda_{ik} = \lambda_{ki}$.

In our design, we suppose that a core node can only be located in place where there is already a Mux/Sel. The core node existence is denoted by $z_j, \quad j = 1...J$.

$z_j = 1$ if there is a core node at location *j* and *0* otherwise.

We should notice that *J*<=*N*. *J* is set to equal to *N* in our current design.

The start-up cost for each core node is denoted by C_{core} and the interface cost to connect the Mux/Sel is denoted by C_{coreIF_MUX}. In another case, we can also suppose that core switch node has a fixed cost for fully equipped box. In this situation, C_{coreIF_MUX} would be *0* and C_{core} would be a bigger number.

The distance between Mux/Sel *i*, and core node *j* is given by d_{ij}. In practice, actual length of cable routes linking Mux/Sels and core nodes would be used instead, where a cable routes already exist.

Variables y_{ij} denote the connections between Mux/Sels and cores. $y_{ij} = 1$ if there is a connection between core node *j* and Mux/Sel *i* and *0* otherwise.

Variable $\alpha_{ik}^{j} \in [0,1]$ means the proportion of traffic from Mux/Sel *i* to Mux/Sel k routed through core node *j*. As we are assuming here that traffic between a given Mux/Sel pair (*i*,*k*)

all follows the shortest path between the Mux/Sels (*i*,*k*), we will have $\alpha_{ik}^{j} = 0 \text{ or } 1$ only. $\alpha_{ik}^{j} = 1$ if traffic from Mux/Sel *i* to Mux/Sel *k* is routed through core node *j* and *0* otherwise.

In the case of relaxing constraint (4.21) and (4.22), the relaxed function is:

$$L\left(z_j,\ y_{ij},\ \alpha_{ik}^{j}, \mu_{ij}, \sigma_j\right) = \left[\begin{array}{l}\left(\sum_{j=1}^{J} z_j C_{core} + \sum_{j=1}^{J}\sum_{i=1}^{N}\left(c\cdot d_{ij} + C_{coreIF\_MUX}\right)\cdot y_{ij} + \sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=1}^{J}\left(d_{ij}+d_{jk}\right)\lambda_{ik}\alpha_{ik}^{j} w\right) \\ +\sum_{i=1}^{N}\sum_{j=1}^{J}\mu_{ij}\cdot\left(\sum_{k=1}^{N}\alpha_{ik}^{j} - N\cdot y_{ij}\right) + \sum_{j=1}^{J}\sigma_j\cdot\left(\sum_{i=1}^{N}\sum_{k=1}^{N}\alpha_{ik}^{j} - N^2\cdot z_j\right)\end{array}\right] \tag{4.24}$$

Subject to:

$$\sum_{j=1}^{J}\alpha_{ik}^{j} = 1 \tag{4.25}$$

$$\underbrace{z_j \in \{0,1\}}_{J\ \text{variables}},\ \underbrace{y_{ij} \in \{0,1\}}_{N*J\ \text{variables}},\ \underbrace{\alpha_{ik}^{j} \in \{0,1\}}_{N^2*J\ \text{variables}} \quad i,k = 1...N,\ j = 1...J \tag{4.26}$$

Rearranging the above Lagrangian and leading the minimization problem of the dual function:

$$\begin{aligned} W\left(\mu_{ij},\sigma_j\right) &= \min_{z_j}\sum_{j=1}^{J}\left(C_{core} - N^2\cdot\sigma_j\right)\cdot z_j + \min_{y_{ij}}\sum_{j=1}^{J}\sum_{i=1}^{N}\left(c\cdot d_{ij} + C_{coreIF\_MUX} - N\cdot\mu_{ij}\right)\cdot y_{ij} \\ &+\min_{\alpha_{ik}^{j}}\left\{\sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=1}^{J}\left[\left(d_{ij}+d_{jk}\right)\cdot\lambda_{ik}w + \mu_{ij} + \sigma_j\right]\cdot\alpha_{ik}^{j}\right\} \end{aligned} \tag{4.27}$$

Subject to:

$$\sum_{j=1}^{J}\alpha_{ik}^{j} = 1. \tag{4.28}$$

Notice that there is no longer dependency between α_{ik}^{j} and z_j, y_{ij}, in the constraints. In this case, we decoupled the problem into three independent optimization problems, corresponding to three terms in the dual function.

Iterations for the optimization problem:

In the current iteration, we know the correct values of the multipliers μ_{ij},σ_j. Using these multipliers, we then perform the minimizations one by one for each multiplier.

First considering the variables $z_j (j = 1...J)$, let $z_j^{*(n)}$ denote the optimum value of z_j in iteration (n).

Then,

$$z_j^{*(n)} = \begin{cases} 1 & \left(C_{core} - N^2 \cdot \sigma_j\right) < 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.29}$$

There are *J* separate minimizations to do for each value of z_j. Similarly,

$$y_{ij}^{*(n)} = \begin{cases} 1 & \left(c \cdot d_{ij} + C_{coreIF\_MUX} - N \cdot \mu_{ij}\right) < 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.30}$$

Notice that the last term is a much simple optimization problem.

$$\min_{\alpha_{ik}^j} \sum_{i=1}^{N} \sum_{k=1}^{N} \sum_{j=1}^{J} \left[\left(d_{ij} + d_{jk}\right) \cdot \lambda_{ik} w + \mu_{ij} + \sigma_j \right] \cdot \alpha_{ik}^j \tag{4.31}$$

Subject to: $\sum_{j=1}^{J} \alpha_{ik}^j = 1$.

We can solve it by a simple method of optimization as follows:

For a specific multiplexer pair (*i,k*), we should minimize $\sum_{j=1}^{J} \left[\left(d_{ij} + d_{jk}\right) \cdot \lambda_{ik} w + \mu_{ij} + \sigma_j \right] \cdot \alpha_{ik}^j$ with respect to α_{ik}^j, subject to: $\sum_{j=1}^{J} \alpha_{ik}^j = 1$. Since α_{ik}^j can only be 0 or 1, we just need to calculate $\left[\left(d_{ij} + d_{jk}\right) \cdot \lambda_{ik} w + \mu_{ij} + \sigma_j \right], j = 1...J$ and find the minimal of it for each *j*. Then the corresponding $\alpha_{ik}^{j^*}$=1, and all the others =0.

In this manner, we can get $\alpha_{ik}^{j^{*(n)}}$.

Plugging these values $\left(z_j^{*(n)}, y_{ij}^{*(n)}, \alpha_{ik}^{j^{*(n)}}\right)$ into the Lagrangian function gives the dual function value $W\left(\mu_{ij}^{(n)}, \sigma_j^{(n)}\right)$ at the current iteration. Its value gives a current lower bound.

Because the variables $z_j \in \{0,1\}, y_{ij} \in \{0,1\}, \alpha_{ik}^j \in \{0,1\} \quad i,k = 1...N, j = 1...J$ are binary variables, the objective optimization function in equation (4.19) is a discrete function and it is not differentiable. The application of classical methods that use gradients based on finite differences may no be applicable for finding an optimal solution. For this nondifferentiable optimization, a subgradient optimization method is used.

Let π denote a vector of all dual variables, i.e. $\pi \overset{\Delta}{=} \{\mu_{ij}, \sigma_j\}$ $\forall i, j$. Taking partial derivatives with respect to the dual variables σ_j and μ_{ij} gives a subgradient $g_n = \frac{\partial W^{(n)}}{\partial \pi}$, which is calculated as:

$$\frac{\partial W^{(n)}}{\partial \sigma_j} = -N^2 \cdot z_j^{*(n)} + \sum_{i=1}^{N}\sum_{k=1}^{N} \alpha_{ik}^{j*(n)} \quad j = 1...J \quad \text{(J equations altogether)} \tag{4.32}$$

$$\frac{\partial W^{(n)}}{\partial \mu_{ij}} = -N \cdot y_{ij}^{*(n)} + \sum_{k=1}^{N} \alpha_{ik}^{j*(n)} \quad i = 1...N \; j = 1...J \quad \text{(N×J equations altogether)} \tag{4.33}$$

We state the update rule as $\pi^{(n+1)} = \max\left\{\pi^{(n)} + t_n \frac{\partial W(\pi)}{\partial \pi^{(n)}}, 0\right\}$, where (n+1) denotes (n+1)st iteration, t_n is the step size, which is a constant for all variables at iteration n. The step size t_n is computed from the equation:

$$t_n = \frac{\rho\left(\bar{W} - W\left(\pi^{(n)}\right)\right)}{\left\| \frac{\partial W^{(n)}(\pi)}{\partial \pi^{(n)}} \right\|^2} \tag{4.34}$$

Here $\bar{W}$ is the best upper bound that we calculated before (i.e. F_{best}), while $W\left(\pi^{(n)}\right)$ is the current lower bound. ρ is the coefficient defined to adjust step size. The initial value of ρ is set to be 2. If there is no progress in certain times of iterations, it may be because the step size is too big. Then we will halve the step size value.

Lots of simulations have been done with different cost settings and initial values. To illustrate the working process of the LR algorithm, an example of the iterations in LR is given as follows. The following Figure 4-15 and Figure 4-16 show the change of dual and primal values and the step size when iteration number is increasing.

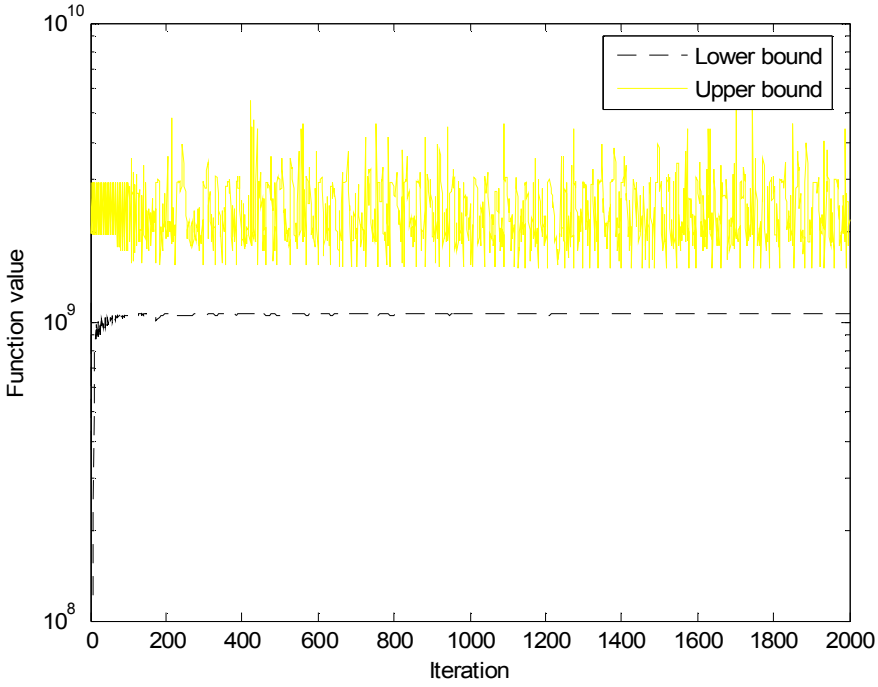


Figure 4-15 The lower and upper bounds of LR algorithm

Figure 4-15 shows the convergence behaviour of the dual and primal functions. The abrupt change in lower and upper bounds in the first several iterations may be introduced by large step-size. And the values tend to be stable as more iterations are applied. The "duality gap" is defined as the difference between the minimum of the upper bound and the maximum of lower bound. In this example, the gap is 16.85%, which shows that the algorithm works well and can get a relatively converged state with acceptable results in the iterations shown above (2000 in this example).

However, it also shows that the pure subgradient method has some limits on its computational performance, due to the direction of motion used. The subgradient direction for this non-differentiable case results in the zig-zagging phenomenon in the upper bound that might cause the procedure to crawl toward optimality. To overcome this problem, the concept of conjugate subgradient method has been proposed which is similar to the conjugate gradient method for differentiable cases. To deflect the subgradient, the conjugate subgradient is obtained by combining the current subgradient with the previous direction. Several subgradient deflection algorithms have been introduced. In these methods, the direction of motion d_n at n-th iteration is not the subgradient g_n only as in pure computed as follows:

$$d_n = g_n + \psi_n d_{n-1} \tag{4.35}$$

Where ψ_n is the deflection parameter. In different subgradient deflection algorithms, ψ_n is computed in several methods. In the Modified Gradient Technique in [43],

$$\psi_n = \begin{cases} \dfrac{-1.5(g_n' d_{n-1})}{\|d_{n-1}\|^2} & if\ g_n' d_{n-1} < 0 \\ 0 & otherwise \end{cases} \tag{4.36}$$

And in the Average Direction Strategy in [44],

$$\psi_n = \frac{\|g_k\|}{\|d_{n-1}\|} \tag{4.37}$$

However, when these algorithms are applied in our particular problem in the core node allocation problem, we found that although the zig-zagging phenomenon is less severe than the pure subgradient algorithm, the computation time for convergence does not improve so much, and the best solutions found by the algorithms are the same in most simulations.

Another problem of the subgradient method is the discontinuity near the optimal solution point. This may result in the oscillatory when reaching the optimal solution. Thus, an alternative "subgradient" method should be proposed to solve the problems mentioned above. This will be part of our future research.

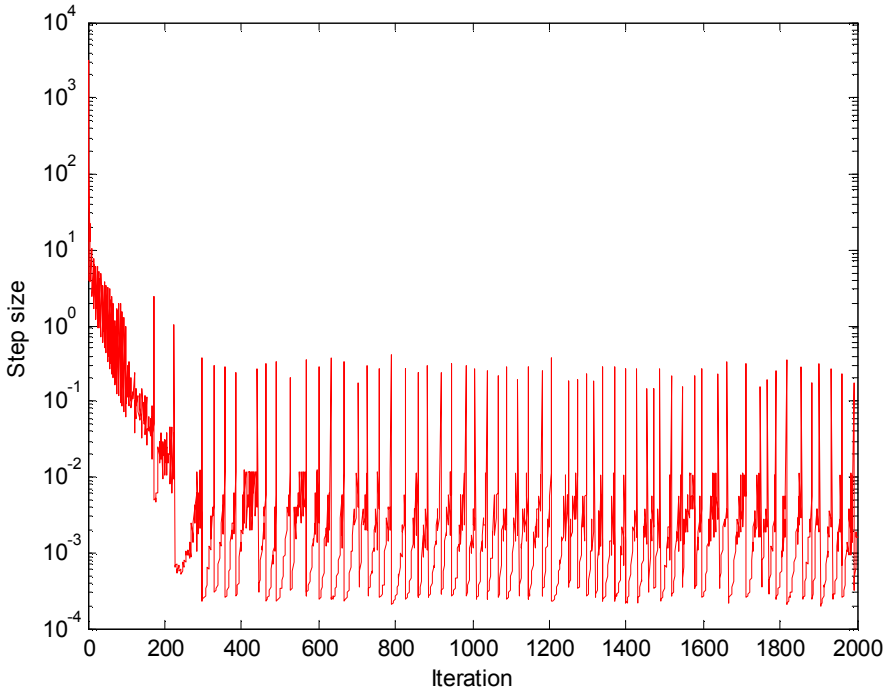


Figure 4-16 The step size of LR algorithm

Figure 4-16 shows the change of the step size. In the several iterations at first, the step size changes sharply, which results in the jump of lower and upper bounds. Step size region is roughly from 0.0001~0.5 when it is in the relatively stable state. At the same time, the resulting primal and dual curves are smoother.

From the simulations we have done, the LR algorithm has the following properties:

1. The setting of the initial values of coefficients is very important to the performance of the program. It may influence the running time for the algorithm to converge to an acceptable sub-optimal result.
2. Different initializations also may yield different core node allocations. This is because the program may not converge to the optimal solution in limited running time.
3. Obviously, more iterations may provide better result. However, we can not have infinite iterations.

4.3.4 Weighted objective function for cost and delay

The IND approach requires determining a weighting factor w (as shown in equation (4.19)), which accounts for the relative importance of the two criteria, cost and delay, when selecting the optimal network topology.

Below is an example to elaborate how to determine the factor w in national network design for core node allocation. Assume the DWDM link between Mux/Sel and core has 16 colors/wavelengths per fiber with a capacity of 160 Gb/s. The number of fibers required is computed from the traffic demand from the Mux/Sel to the core node. Note that not every Mux/Sel and every core need to be connected together and in fact they generally will not since the shortest path routing is used. No matter how they are connected, the $\text{cost}/(\text{unit distance}\cdot\text{single DWDM link})$ is the fiber cost part of w in the formula. The cost of the link from Mux/Sel i to core node j is then $w\sum_{k=1}^{N}\lambda_{ik}\cdot\alpha_{ik}^{j}\cdot d_{ij}/160$. Here we just formulate the cost into a linear function w.r.t. distance, which we will discuss more in this section. Similarly the cost for the link from core node j to egress Mux/Sel k is given by $w\sum_{i=1}^{N}\lambda_{ik}\cdot\alpha_{ik}^{j}\cdot d_{jk}/160$. The cost of DWDM links from core node j to ALL egress Mux/Sels is $w\sum_{i=1}^{N}\sum_{k=1}^{N}\lambda_{ik}\cdot\alpha_{ik}^{j}\cdot d_{jk}/160$, while the cost from all ingress Mux/Sels to core node j is given by $w\sum_{i=1}^{N}\sum_{k=1}^{N}\lambda_{ik}\cdot\alpha_{ik}^{j}\cdot d_{ij}/160$. Summing the two terms gives the cost of DWDM links connected

to core node j as $w\sum_{i=1}^{N}\sum_{k=1}^{N}\lambda_{ik}\cdot\alpha_{ik}^{j}\cdot(d_{ij}+d_{jk})/160$. Finally summing over all core nodes j where α_{ik}^{j}=1, gives the total cost of DWDM fibers linking Mux/Sels to all core nodes. This is just $w\sum_{j=1}^{J}\sum_{i=1}^{N}\sum_{k=1}^{N}\lambda_{ik}\cdot\alpha_{ik}^{j}\cdot(d_{ij}+d_{jk})/160$. Noting that it is proportional to the traffic weighted delay since distance and delay are linearly related by propagation speed, we do not need to add another term if we wish to emphasize the delay aspect.

Suppose the price for DWDM is 3000$/km (average price including amplifiers, 3R regenerators etc). We can find the value of delay and cost for a range of w higher than 3000$/km, where the average delay of whole network is defined as:

$$delay=\frac{\sum_{i}\sum_{k}\sum_{j}\alpha_{ik}^{j}\cdot\left(d_{ij}+d_{jk}\right)}{N^{2}\cdot c\cdot 0.75} \tag{4.38}$$

The higher value of the parameter w will cause more emphasis on the delay term than those involving smaller w. Here w is a parameter that allows us to use the optimization routine to generate a range of values of configurations with corresponding true costs and true delay values. And the true costs and the true delay values are always calculated with w=3000, regardless of the value of w used in the objective function to get that topology. Upon these results, we can get the maximum of delays, D_max, and the maximum of costs, C_max.

Viewed as a multicriterion problem, we can plot the Pareto Boundary[4] by varying w over a wide range of values, which will impact the number of core switches in the optimized design in Figure 4-17. Pareto optimality is widely used in economic analysis for efficient and economic allocation of resources. [40,41] give detailed introduction of this scheme. In our figure, the point of interest is obtained when w is set at the value where the hyperbola is tangent to the curve with the polygonal line with stars.

The following figure is obtained from the Lagrangian relaxation model (has been discussed in details in Section 4.3.3) by varying the weighting factor w. Various numbers of core nodes and their locations in national network design have been obtained with 64 fully loaded 16-port multiplexers.

[4] http://www.ime.auc.dk/people/employees/no/notes/OPT8.pdf.

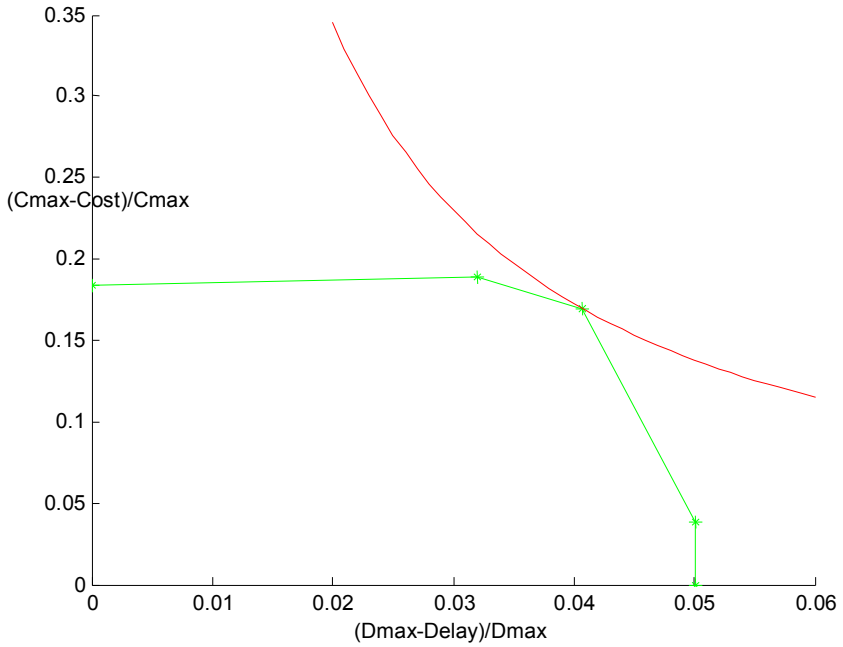


Figure 4-17 Pareto Boundary of delay vs. cost

Data used for the nodes in this picture (from left to right) are as follows:

Core node num. in each case:	4	5	8	27	30
w used for designing network:	3000	6000	10^4	10^6	10^8
Delay (in millisecond):	7.8965	7.6442	7.5751	7.5009	7.5004
Cost:	1.5832×10^9	1.5729×10^9	1.6109×10^9	1.8646×10^9	1.94×10^9

The picture shows a non-monotonic behavior of the curve, where the cost when w=3000 is more than the cost when w=6000 while when w=3000 the delay is less. This is because the heuristic solution is closer to optimum when w=6000 than it is when w=3000. The actual optimum at w=6000 cannot be cheaper than the optimum at w=3000.

Furthermore, we notice that there is very little difference in delay between the case with w=10^8 which has many more core nodes than the case with w=3000 which has 3 or 4 core nodes only. This is mainly because of Canada's geographic character. In Canada, all nodes are roughly in two groups, the west and the east. Traffic can be divided into three groups: west<->east, within west and within east. If we have core nodes in both west and east groups, the actual routing distance doesn't save too much even if we have many more core nodes. But adding more core nodes doesn't decrease the delay that much because the absolute distance from west to east is significant and delay in traffic between west and east is always much longer.

The real costs we get from the formula include:

Core node cost: including start up equipment cost + interface cost

$$\sum_{j=1}^{J} z_j \cdot C_{core} + \sum_{j=1}^{J}\sum_{i=1}^{N} C_{coreIF\_MUX} \cdot y_{ij} \tag{4.39}$$

Fiber cost: When w=3000 is used, the estimated fiber cost is

$$\sum_{j=1}^{J}\sum_{i=1}^{N} c \cdot d_{ij} \cdot y_{ij} + \sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=1}^{J}\left(d_{ij} + d_{jk}\right)\lambda_{ik}\alpha_{ik}^{j} w \tag{4.40}$$

For fiber cost, we may notice that the real cost model is modular, while the MILP used here is linear in the number of fibers and cost per unit distance respectively. The modular cost is of course more accurate than the linear cost model but does not result in an MILP. Thus a linear or more generally an affine cost model is applied to approximate it to a linear function.

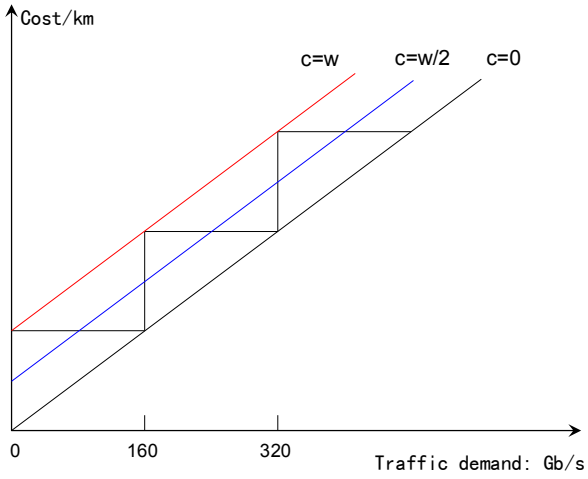


Figure 4-18 An affine cost model for fiber cost

When we set c=0 we are using the linear model as the bottom line in Figure 4-18. If the point-to-point traffic demands are much less than 160 Gb/s, then the linear model will greatly underestimate the true cost given by the modular cost model as well as the number of fibers. In fact the linear model implies that it is possible to have a fractional number of fibers, for example, 0.06 fiber of a single Mux/Sel to Mux/Sel demand through some single core node. Indeed if the demand is >0 then at least one fiber must be present, not 0.06 fiber as given by the last term in the MILP.

For large traffic demands >> 160 Gb/s the error is less severe, however for the case with lots of core nodes it will be very significant and we should take this fact into account. One way to mitigate this problem while still retaining the MILP formulation would be to use a non-zero value for the c parameter in the second term involving y_{ij}. If we set c=w then we have an

affine cost function shown in the upper line which upper bounds the true modular cost function, as shown in Figure 4-18. A compromise would be an affine function shown in the middle, which corresponds to a *c* value of *w/2*.

4.4 Backup network design

In this context agility is a required network capability because it provides a degree of service quality robustness in the face of traffic forecasting errors and facility failures.

For the connectivity between Mux/Sel and core nodes, traffic restoration upon network failure is necessary. We apply the MCR algorithm from [29] in the core node allocation problem. Results show that this algorithm also works well in our network.

The procedure is as follows:

1) Based on the results from Lagrangian Relaxation, check all the possible routes between each source-destination pair. If there is only one route, then another route is added for backup usage. The algorithm is described as follows:

   a) Check a source-destination pair. If there are at least two routes available for them, then go to step b. If there is only one route, then find another second-shortest path route and add a tag on the link that needs to be added.

   b) Repeat step a, until all source-destination pairs are considered. If there is no tagged link, then stop here, otherwise go to step c.

   c) Add the connectivity link which is with most tags, then go to step a.
   After this step, the connectivity matrix y_{ij} may be changed.

2) Calculate the working link capacity and traffic on each link in the non-failure state.

   The definitions are as follows:

   a) Link: The direct connection between two nodes. There may be more than one fiber in the link depending on the traffic.

   b) Route: Especially in our design, route means the connection of the source-core node-destination.

   c) Link traffic: The link traffic in non-failure state:

$$traffic_{ij} = \sum_{k=1}^{N} \alpha_{ik}^{j} \cdot \lambda_{ik} \quad \forall i = 1...N,\ j = 1...J \tag{4.41}$$

d) Link capacity: As 16-interface DWDM connection is used on the link, the link capacity in non-failure state is:

$$capacity_{ij} = 160 \cdot ceil(traffic_{ij}/160) \quad \forall i = 1...N,\ j = 1...J \tag{4.42}$$

e) Spare capacity:

$$spare_{ij} = capacity_{ij} - traffic_{ij} \quad \forall i = 1...N,\ j = 1...J \tag{4.43}$$

Link capacity should be bigger than or equal to link traffic. For example, if the total traffic from Mux/Sel i to core node j is 300G, then *ceil*(300/160)=2 fibers are needed, thus the initial link capacity in non-failure state is 2×160=320G. The spare capacity is 320-300=20G.

3) Calculate the link capacity in failure state. Here the case of failure-oriented, state dependent reconfiguration scheme is considered. In the single link failure scenario, when the network is in each state, only a portion of working connections is affected.

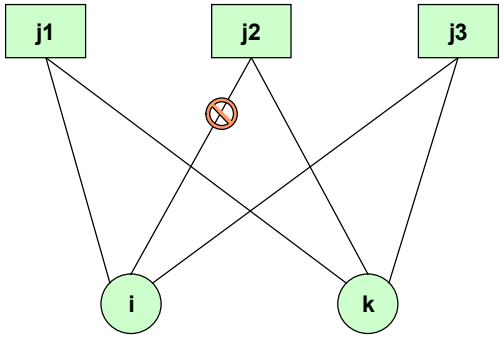


Figure 4-19 A single link failure example

The new connections and traffic distribution in each failure state are calculated using MCR concept.

a) In a link failure state, find all the affected source-destination traffic flows. For example, in Figure 4-19, for source-destination pair $i-k$, if the working route is $i-j_2-k$, when the link $i-j_2$ of working route is broken, all traffic from i to other Mux/Sels passing j_2 is affected. This traffic should be transferred to another backup route $i-j_1-k$ or $i-j_3-k$.

Minimize the spare capacity cost using MCR. Suppose there are R candidate routes for the source-destination pair $i-k$, with the core nodes $j_1, j_2 \cdots j_R$. We can see that only when the working route of the pair is broken, will the traffic be affected. In state s, when the working route of $i-k$, i.e. j_m, is broken, a route with minimum cost will be chosen as backup route with the following formula:

$$cost_{ik}^{s} = \min_{j=j_1,j_2\cdots j_R \text{ except } j_m} \left(d_{ij} \cdot (B_{ik}^{j_m} - spare_{ij})^{+} + d_{kj} \cdot (B_{ik}^{j_m} - spare_{kj})^{+} \right) \quad \forall i,k,s \quad (4.44)$$

The $B_{ik}^{j_m}$ here is the bandwidth of the working route for i and k, passing j_m, which should be equal to the traffic λ_{ik}. We should notice that when the cost is equal to 0, it means the spare capacity in link $i-j$ and $j-k$ are both bigger than traffic and $i-j-k$ is a valid route. In the case where the costs of all routes are bigger than 0, it means the traffic is bigger than the initial spare capacity. On the route with minimal cost, one fiber with 160G capacity should be added to the link $i-j$ or $j-k$ or both as needed. In this case, the available spare capacity and the bandwidth of link should be updated.

b) The required spare capacity and bandwidth of each link in state *s* will be saved.

c) Repeat step a), until all failure states are considered.

Find the maximum bandwidth needed on each link. This will result in the initial assignment of link capacity for single-failure restoration.

For example, if the traffic from Mux/Sel i to core node j in non-failure state is 300G, and maximum traffic in some failure state is 500G, then ceil(500/160)=4 fibers are needed, the maximum required spare capacity is 500-300=200G. The spare capacity in non-failure state is 4×160-300=340G.

4) The heuristic algorithm is to modify the connectivity capacity. This part is quite similar to what is described in the paper. Follow the heuristic modification phases until the total spare cost cannot be reduced.

Step A:

a) For a given (tagged) link, find the worst failure state and the corresponding broken link. Then find all the affected traffic routes in this state.

b) Restore failed working routes if their backup routes do not pass the tagged link;

c) If the backup route passes the tagged link, try to find a new backup route not passing the arc but with zero additional spare cost with MCR;

d) If such a new backup route is not available, then keep the original backup; if the new backup route is available, use the new one. Go to step b for next route until all routes are considered.

e) Recalculate the spare capacity on all links.

if (total spare cost decreased)

go back to a);

elseif (total spare cost unchanged)

if (at least one new backup route was found in c)

go back to a);

else

go to step B;

end

elseif (cost increase)

set back modified backup routes; go to step B;

end

Step B: Repeat step A until all links are considered.

Step C: Repeat steps A and B until the total spare cost cannot be further reduced.

Chapter 5 Data analysis

Before analyzing the results, we hasten to point out that equipment costs used in the calculations are extrapolated estimates based on the similar devices currently available, and the costs will likely decrease significantly with time as larger equipment volumes are deployed. Nevertheless we believe that the results can provide some guidance and insight into what equipment items will dominate network capital cost. These cost figures reported are for the working network design and reliable network design for networks protected against single link fiber failures.

The tools and design methods mentioned in Chapter 1 are exercised under a wide variety of equipment cost assumptions for demographic data collected for a pan Canadian wide area network and an artificial metropolitan network called Gotham. For Canadian network, as an integrated design case, we have done the working network and survivable network designs in sequence. The survivability and physical diversity and capacity requirements for survivability are studied for single link failure case. The metro network design is a part of the tiered network design approach.

5.1 Design problem categories and algorithms

As discussed in Section 4.3, we have designed the network topologies for a WAN for Canada and a MAN of Gotham. Two-layer and three-layer topologies are considered.

Given the initial location of edge nodes as well as the possible fiber path/conduit infrastructure, the design process consists of determining the optimum number, placement and interconnection pattern for core nodes and Multiplexer/Selector switches to minimize the overall cost subject to a set of given constraints.

WAN design for Canada:

In Canadian national network, there are 1024 edge nodes distributed in 140 cities. As analyzed in 4.1.1, for such a large network, we can use an overlaid tree topology with three layers. For the Mux/Sel as the middle layer, different sizes of 8, 16 or 32-interfaces have been tested.

In particular, the placement and interconnection of fixed port count, with fixed number of Mux/Sels referred as a *P-median* problem and if Mux/Sel number is not fixed, it is referred as

a Single Source Capacitated Plant Location Problem (SSCPLP). Both problems can be solved by the MILP method (CPLEX), Lagrangian Relaxation (LR) and Simulated Annealing (SA) (MATLAB). In our work, for the *P-median* problem, SA is used to get the 26 Mux/Sel locations. And LR or CPLEX (in some cases) can be used for the SSCPLP problem.

For core node allocation, as discussed in 4.3.3, a new Integer Linear Program is applied to formulate the problem. It is solved with the LR algorithm.

MAN design for Gotham:

In Gotham network, there are 300 edge nodes. As the distance range in metro network is not too wide, fiber and cable cost is no longer as significant as it is in national network. So we have tried the two-layer and three-layer design to compare the total cost.

In three-layer design, for the Mux/Sel allocation, an approach borrowed from [42] has been used (LR for location + CPLEX for connectivity) to solve the *P-median* problem. For the core node allocation problem in three-layer design, as there are not many Mux/Sels, explicit enumeration calculation is used.

For two-layer network of MAN, as we analyzed in Section 4.1.3, the network layout is different for upstream and downstream traffic. With the Mux/Sel allocation in our three-layer design, enumeration calculation is employed to find the optimal allocation of core nodes.

The mathematical problems we have in our design, and algorithms to solve them have been listed in Table 5-1.

Table 5-1 Solution methodology (Examples with 16-interface Mux/Sels)

Common part	
Edge node allocation based on population information	
Cabling infrastructure simulation based on modified spanning tree	
two-layer network problem	
Connect edges to core nodes	
MAN: 300 edges to 1...6 cores	LR, Enum
three-layer network problem	
Connect edges to switches	
WAN: 1024 edges to 64 (or more) Mux/Sels	SA, LR, CPLEX
MAN: 300 edges to 10,19,38 Mux/Sels	LR+CPLEX
Connect Mux/Sels to core nodes	
WAN: 64 (or more) Mux/Sels to 1…6 core nodes	LR, Enum.
MAN: 19 Mux/Sels to 1…6 core nodes	Enum.

In the design process, a lot of assumptions were made re traffic demand and some cost constants are assumed in the design for both WAN and MAN cases.

5.2 Accuracy validation of optimization methods

Because CPLEX has limitation on problem size and cannot get results for large problems, we have used LR and SA algorithm for the optimization procedure. For the MILP case, there is an imposed limit on problem size due to the computational complexity of direct algorithm. Concerning the accuracy measured in percentage, the approximate methods like LR, SA are compared to the exact solution from CPLEX or Explicit Enumeration. Undoubtedly we cannot do comparisons on the large problems because if we had CPLEX solutions or Explicit Enumeration solutions we would not really need the LR/SA methods. The accuracy of our programming implementation of heuristic methods is verified with MILP on smaller separate cases.

For example, CPLEX program is unable to solve the 26 Mux/Sels to 464 edge nodes *P-median* problem for three-layer WAN. A reduced dataset of 128 edge nodes is successfully solved with CPLEX and the result is less than 1% different from results from LR and SA algorithms. Then, 464-edge problem is solved by SA and solution is imported to CPLEX as a starting point. No further improvement is obtained.

Similarly, in the Mux/Sel allocation of three-layer design for MAN, because CPLEX cannot handle this big problem, LR is used instead. A reduced data set with 96 edges is tested for the accuracy of optimization algorithms comparing with CPLEX.

Some statistics are provided in the following Table 5-2. We found that in most cases LR and SA are able to obtain an exact solution for a given problem.

Table 5-2 Computation complexity and cost values

	Algorithm	Result	Calc. time
Edge connectivity			
WAN 1024x64 WAN 1024x67	SA CPLEX	41681 34548	12 hours 1.5 hours
WAN 128x8	LR +CPLEX SA CPLEX	22172 to 20245 20083 20020	5 min. 7 min. 25 min.
MAN 300x19	LR +CPLEX	167300	3 min.
MAN 96x6	LR +CPLEX SA CPLEX	71100 to 68500 68300 68300	2 min 12 min. 3 min.
Core connectivity			

MAN 19x1…6	LR	1673138	1 min.
	Enum.	1673138	30 sec.
WAN 67x1…6	LR	1.51×10^{9}	6 hours

5.3 Circuit design

In order to evaluate the circuit cost of direct fiber vs. CWDM in the local area and amplified fiber vs. DWDM in the national area, a program in MATLAB has been developed with results shown in Figure 5-1 and Figure 5-2.

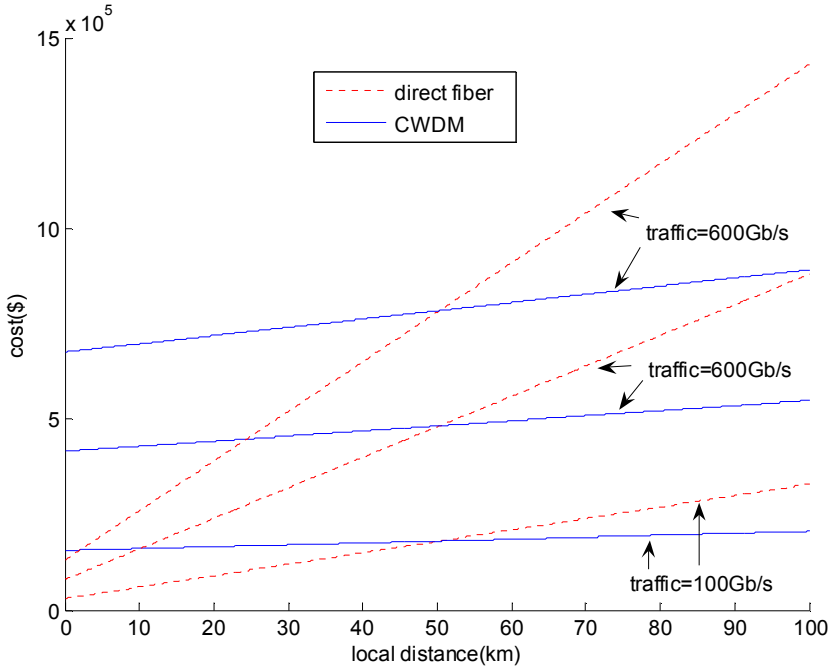


Figure 5-1 Cost of direct fiber vs. CWDM

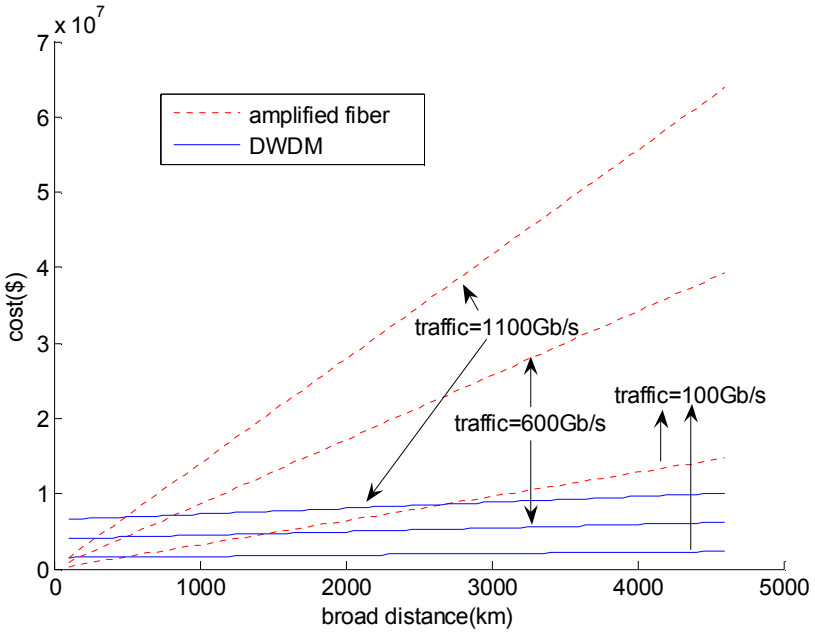


Figure 5-2 Cost of amplified fiber vs. DWDM

As the results shown above, within distance range of 50 kilometers, which is the normal range for a metro network area, direct fiber is a better choice than CWDM in most cases. And within distance range of 500 kilometers, which is the normal distance for nodes in a national network, DWDM is much economical than amplified fibers.

Another noticeable trend from the figures is that the costs increase a lot as the traffic increase. As we will discuss in Section 5.4 and 5.5, traffic can be calculated and thus a suitable circuit connection method can be chosen for both MAN and WAN.

5.4 National network

Actual population information as well as geographical coordinates for 140 Canadian cites were obtained from a Census database. In absence of existing cabling infrastructure data we have generated a representative infrastructure layout with the help of the TD Tool using a modified Minimum Spanning Tree algorithm as in Figure 5-3. The resulting wide-area network then consists of 1024 edge nodes distributed over 140 locations using the flat traffic demand model; each edge serves around 30,000 customers where each in turn is generating a 30Mbits/s data stream in both directions.

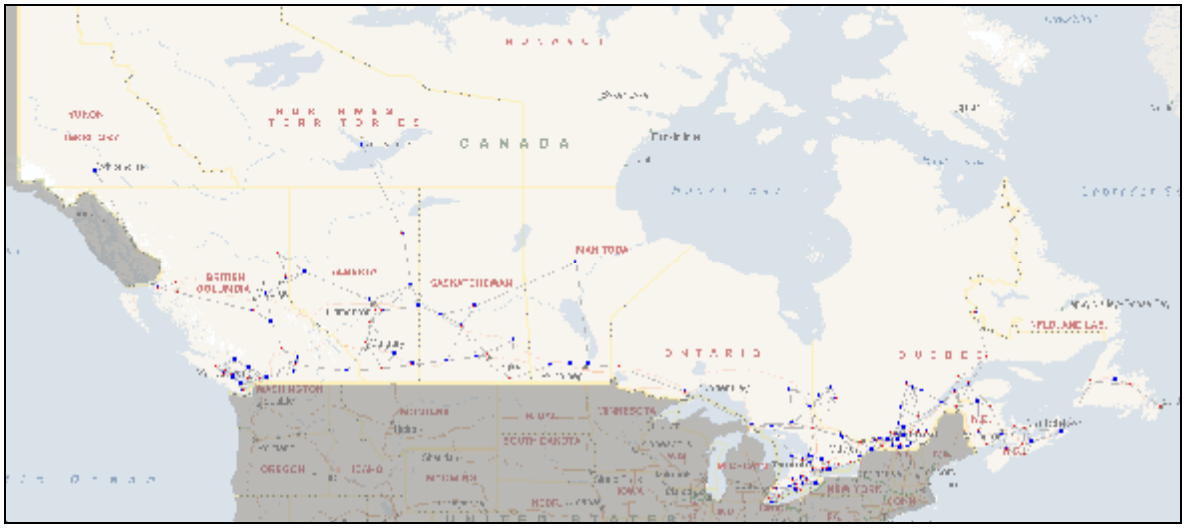


Figure 5-3 National network, cabling infrastructure

As what we assumed in our previous traffic demand estimation in Section 3.1, the Canadian network consists of about 30,000,000 people. If every user generates a peak traffic demand outgoing=incoming of 10 Mb/s, assuming 10% active users during busy hour, the outbound traffic is 30 Tb/s. Assuming 1024 edge nodes and a flat traffic demand used, there is about 30 Gb/s outbound from each edge node. As there are 1024 edge nodes, there will be roughly 30 Mb/s outbound from each edge node to another.

Assuming that each Mux/Sel serves 16 edge nodes both incoming and outgoing, there will be 30 Mb/s×16×16= 7.68 Gb/s outbound from each Mux/Sel destined for each other Mux/Sel. However, for Mux/Sel that is not fully loaded, the traffic will be certainly less. Suppose the link utilization is 80% (this number is an estimated number, which has been discussed in Section 3.1), we have the capacity needed as the following:

- Edge to Mux/Sel connection capacity needed: 30M×1024/0.8=38.4Gb/s (outbound capacity for each edge)
- Any Mux/Sel to any other Mux/Sel capacity needed: 7.68/0.8=9.8 Gb/s
- Thus the total traffic from one Mux/Sel is: 9.8×64=627.2 Gb/s

5.4.1 Multiplexer/Selector allocation

We will explain the design schemes using the instance of 16-interface Mux/Sel. It is quite straightforward that to reduce the complexity of the problem, the required Mux/Sels can be explicitly positioned in cities where the number of installed edge nodes is high. In other words, if one city has more than 16 or 16's multiple (32, 48, 64…) edges, 1 (or 2, 3 or 4) Mux/Sel should be allocated in the same city. This will minimize their fiber connectivity cost to 0 because we assume the edges in the same city have distance=0 from each other.

Assume the 16-interface Mux/Sels are used. Mux/Sels are first allocated in the city where there are more than 16 edge nodes. For example, in Toronto there are 201 edge nodes, so (201)mod(16)=12 Mux/Sels are allocated first, with 201-12×16=9 edges left. Altogether in Canadian network, 35 Mux/Sels are allocated with edges in the same city with 464 edge nodes left. Then these 464 edges should be connected to 464/16=29 or more Mux/Sels, which is the problem we need to solve as discussed in Section 4.3.2. In our Canadian network case, 35 Mux/Sels with 560 edges are first allocated in some cities. For the remaining 464 edges, the allocation and number of Mux/Sels can be formulated as an SSCPLP (relaxed Mux/Sel number) or a *P-median* problem (fixed Mux/Sel number) as discussed in Section 4.3.2. In fact the Mux/Sel number used here is not precise because the traffic demand and link capacity issues are mot considered here. More details about this will be illustrated by the end of this section.

The capacitated *P-median* problem where the Mux/Sel number is set to be 1024/16=64, is solved with SA algorithm. The MILP methods for solving the SSCPLP problem are applied to the edge node data set in CPLEX. Different settings for Mux/Sel start-up cost are tried and the results are as follows:

If all Mux/Sels are fully occupied, there are 35+29=64 Mux/Sels (solved by SA).

Case 1: $Cm_{16} = 10^5$: 124+35=159 Mux/Sels (This Mux/Sel cost is the reasonable one we used in our design, solved by CPLEX)

Case 2: $Cm_{16} = 10^6$: 46+35=81 Mux/Sels (solved by CPLEX)

Case 3: $Cm_{16} = 4 \times 10^6$: 32+35=67 Mux/Sels (solved by CPLEX)

(When Mux/Sel cost is higher, the CPLEX program runs out of memory.)

To choose a proper design to minimize the overall cost, all the results are substituted into the core node allocation problem. Results show that the case with 67 Mux/Sels achieves least overall cost. The results were imported back to the TD Tool (Figure 5-4). This gives the locations for the 67 Mux/Sels in the relaxed case.

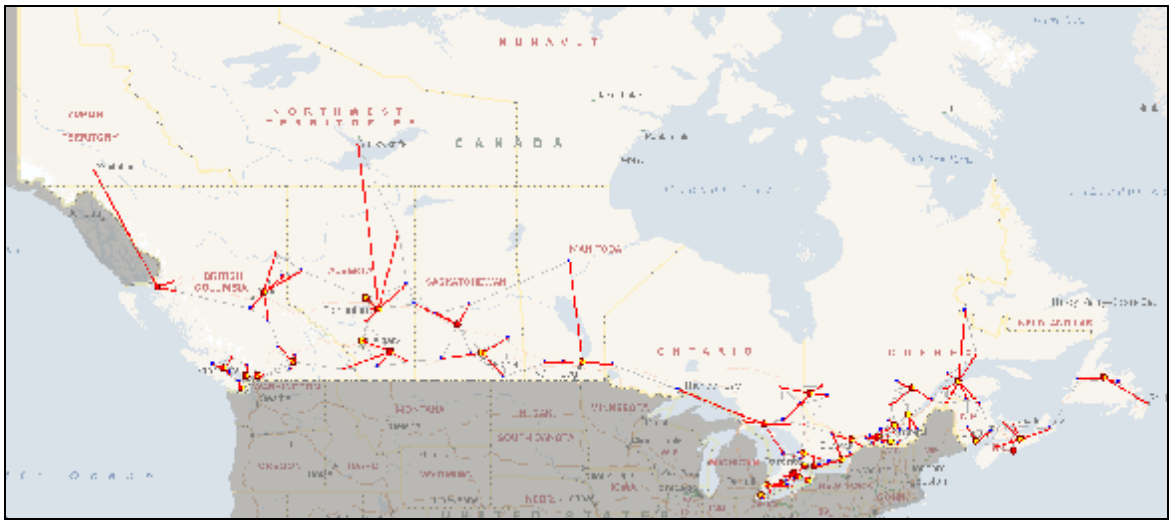
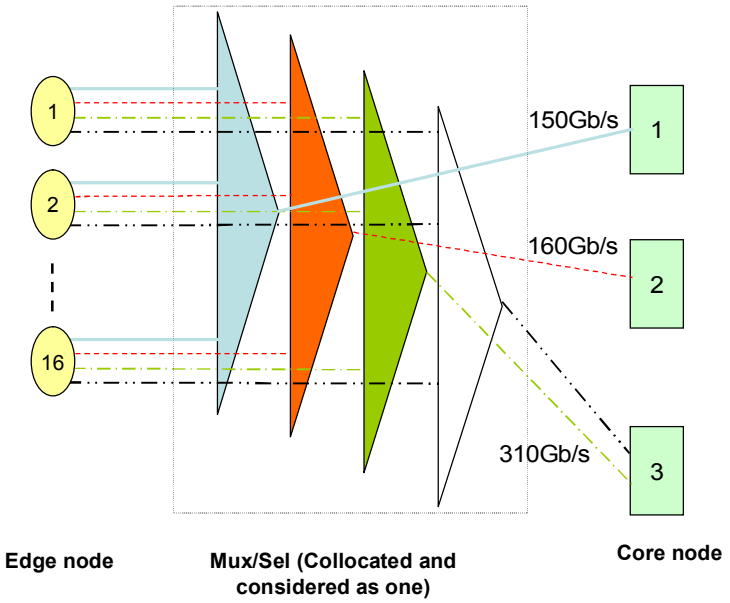


Figure 5-4 National network, 67 Mux/Sels in 38 cities.

There is one thing we have to clarify about the Mux/Sel number, like “67” here. First of all, each Mux/Sel has only one uplink to one core node as we indicated in Section 4.1.1. Thus if one Mux/Sel needs to connect to more than one core nodes, we actually need more than one Mux/Sels collocated. Another consideration is about the DWDM capacity. The actual number of Mux/Sel devices and fibers will be bigger because the calculation algorithm assumes infinite switch-to-core capacity while the real demand may exceed the actual capacity that the DWDM interface/fiber can support. The connection has been shown in Figure 4-9 and Figure 4-10 in Section 4.1.3. Suppose the maximum link capacity is 10Gb/s, if the Mux/Sel has16 interfaces to edges, the maximum capacity for one DWDM link from Mux/Sel to core node is 16×10=160Gb/s. So, as shown in Figure 5-5, if the capacity needed from a certain Mux/Sel to a core node is more than 160G, we will need more Mux/Sels collocated and more interfaces in the core switch. Consequently, from each edge node to its connected Mux/Sels, more connectivity fibers are needed. An example has been given in Figure 5-5, where four Mux/Sels are needed to be collocated instead of one.



Suppose the traffic from the Mux/Sel to core node 1, 2 and 3 is 150Gb/s, 160Gb/s and 310Gb/s. In fact, four Mux/Sels are collocated to support all the traffic.

Figure 5-5 Collocated Mux/Sels

5.4.2 Core node allocation

Further LR calculations to solve the core node location problem resulting in the final topology are carried out. Figure 5-6 shows the logical core-Mux/Sel connectivity for node 1 of 5. The accuracy of calculations depends upon the actual cost parameters used as well as the maximum allowed run time.

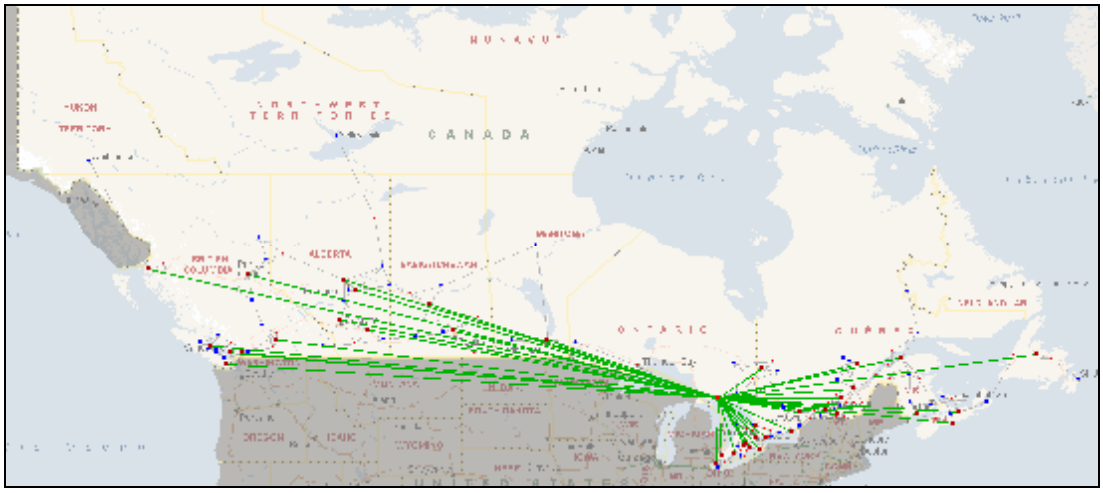


Figure 5-6 National network core node no. 1 (of 5)

Quantities of network equipment and cost breakdown for the WAN models are summarized for two Mux/Sel sizes in Table 5-3.

Table 5-3 Core node allocation with 16 and 32-interface Mux/Sels

Name	16-port Mux/Sels, 5 core nodes				32-port Mux/Sels, 3 core nodes			
	Number	Price	Total	Fraction,%	Number	Price	Total	Fraction,%
Core Nodes	5	100000	500000	0.029	3	100000	300000	0.016
Core Node DWDM Ports	179	30000	5370000	0.316	83	30000	2490000	0.132
Selector Switch Nodes	179	100000	17900000	1.055	83	100000	8300000	0.439
Selector Switch Ports	2740	7000	19180000	1.131	2487	7000	17409000	0.922
Edge Nodes	1024	200000	204800000	12.07	1024	200000	204800000	10.84
Edge Node Ports	2740	7000	19180000	1.131	2487	7000	17409000	0.922
Location Start-up	140	50000	7000000	0.413	140	50000	7000000	0.371
Fiber and amplifiers	474209	3000	1422627000	83.85	543827	3000	1631481000	86.36
Total			1696557000	100			1889189000	100

It was found that the major components of the total installed cost of a WAN are the fiber and amplifiers. The results shown are for the preferred WAN solution namely a three-layer network which includes Mux/Sels and DWDM transmission links. Designs with 3 core nodes and 5 core nodes indicate that the latter design is better in both cost and mean network propagation delay. We note that due to the large fraction of the total costs attributed to fiber and amplifiers, this motivates efficient traffic handling as gains in utilization and then efficiency will be translated more directly into cost savings than is the case the MAN designs.

5.4.3 Network reliability design

The working network design results from Lagrangian Relaxation where there are 67 Mux/Sels and 5 core nodes are used as input for reliability design. First some links are added to guarantee that the traffic from any source to any destination has at least two alternative routes. Afterwards, the Minimum Cost Route (MCR) algorithm is used to minimize the spare capacity cost for the given restoration requirements upon single-link failure.

As mentioned in Section 4.4, firstly some links may need to be added where necessary to guarantee there are at least two candidate routes from any source to any destination. In our national network case of Canada, after this step, 210-179=31 links are added into the network for reliability purpose (calculated from $\sum_{i=1}^{N}\sum_{j=1}^{J} y_{ij}$).

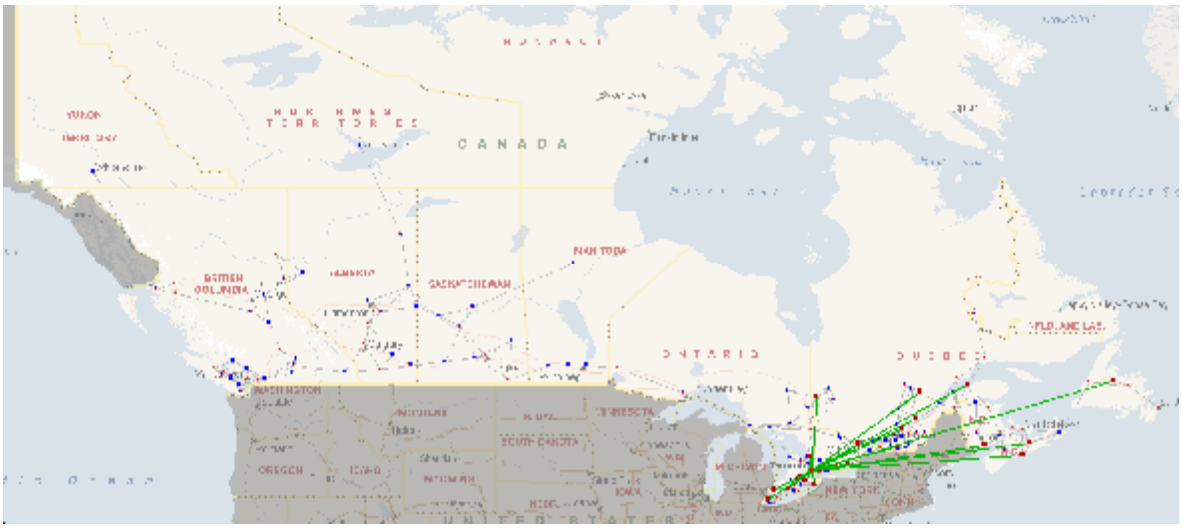


Figure 5-7 National network core node no. 4 (of 5): for working network design

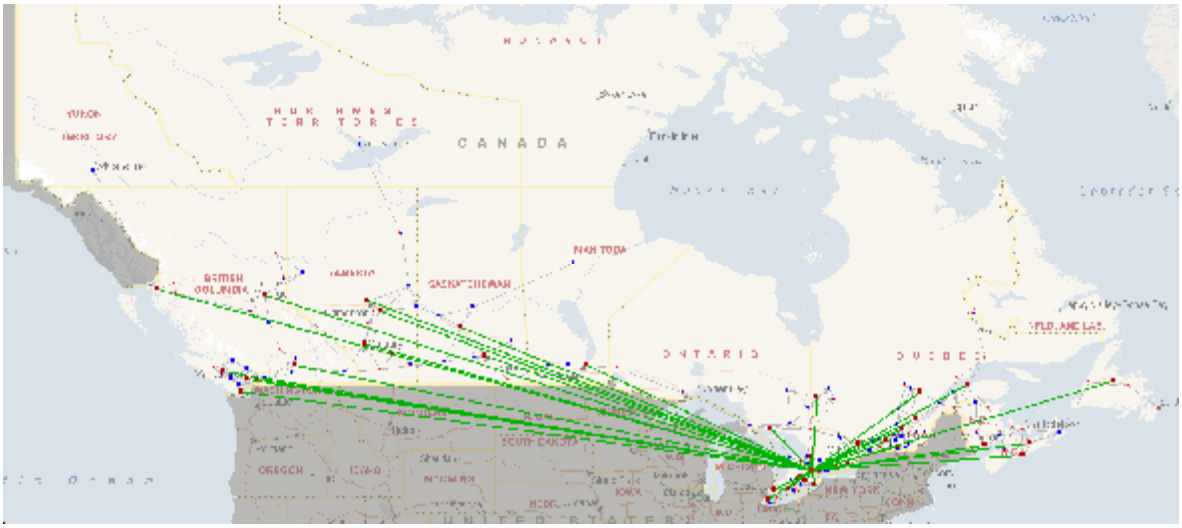


Figure 5-8 National network core node no. 4 (of 5): for reliable network design

We use the total fiber distance $=\sum_{i=1}^{N}\sum_{j=1}^{J} d_{ij} \cdot y_{ij} \cdot f_{ij}$ as the capacity cost measurement of the reliable design, where f_{ij} is the number of fibers on link $i-j$.

Following the step by step procedures in Section 4.4, we can see that the total fiber distance has changed as follows:

- The total fiber distance without failure-oriented backup is: 290859
- The initial total fiber distance for single-link failure restoration is: 595865
- The modified total fiber distance after the heuristic MCR algorithm is: 565968

5.5 Metro network

For the metro network, the 4.5-million, 300-edge node artificial city "Gotham" is simulated. This network covers an area of about 80*60 square kilometers. A custom MATLAB program has been used to distribute population with multiple Gaussian-like functions to model the downtown and suburban areas. Cabling infrastructure is simulated by the custom MST-based algorithm; Manhattan distances are used in the distance matrix preparation.

Similarly as the national network design case, we can get the traffic patterns for metro network. In our Gotham network, there are 300 edge nodes. Two traffic load models are considered in the design. One is the light traffic model where each edge node can be supposed to have less traffic demand than in the national network, which is 10 Mb/s inbound/outbound from each edge node to any other edge node. And another is the heavy traffic model with 30Mb/s.

First we will calculate the traffic for light traffic model. Assuming that each Mux/Sel serves 16 edge nodes both incoming and outgoing, there will be 10 Mb/s×16×16=2.56Gb/s traffic from each Mux/Sel destined for any other Mux/Sel. Suppose the scheduler performance is 80%, we have the capacity needed from Edge to Mux/Sel =10M×300/0.8=3.75Gb/s (traffic capacity for each edge). So from edge to Mux/Sel, the traffic is less than 10G, and single fiber is enough for the capacity.

For any Mux/Sel to all core nodes connected, the total traffic in both directions would be: 2.56×19/0.8=60.8 Gb/s, which is the upper bound of the traffic transmitted in one DWDM fiber. As we know, the capacity of one DWDM link is 16×10=160Gb/s. Thus the minimum spare capacity in a DWDM fiber should be 160-60.8=99.2Gb/s. As single link failure is assumed in reliable design, the minimum spare capacity 99.2Gb/s>60.8Gb/s, which is the maximum rerouted traffic to this fiber of all failure states. Thus in all failure states, once there is reroute path for the affected traffic, the spare capacity is always sufficient. It is not necessary to apply the MCR algorithm as described in Section 4.4 for metro network design with light traffic load.

Similarly we can calculate the traffic parameters when the heavy traffic model is used.

5.5.1 Three-layer metro network

For the three-layer network design, both MILP and LR algorithms are used to locate Mux/Sels (Figure 5-9). We have used an approach from [42] to solve the *P-median* problem. Similarly to WAN design, different sizes of 8, 16 or 32-interface Mux/Sels are tested. Our

heuristic method for light traffic model can be described as follows:

1) First LR is used to get the initial solution for the *P-median* problem.

2) With the input of Mux/Sel allocation from LR, CPLEX is applied to get the optimal connectivity design between edge nodes and Mux/Sels.

3) Suppose each Mux/Sel and edge nodes connected to it are in one group. For each group, check which edge node location can minimize the total group distance. It can be easily done as follows:

   a) Set one of the edge locations as the location for Mux/Sel in this group;

   b) Calculate the total distance to connect all edges to this Mux/Sel in this group;

   c) Find the optimal Mux/Sel position with least distance in each group. If there is any change of the optimal Mux/Sel locations from previous results, use these Mux/Sel locations as input and go to step 2). Otherwise, we can consider this as the last solution.

Figure 5-9 shows the Mux/Sel locations in Gotham city.

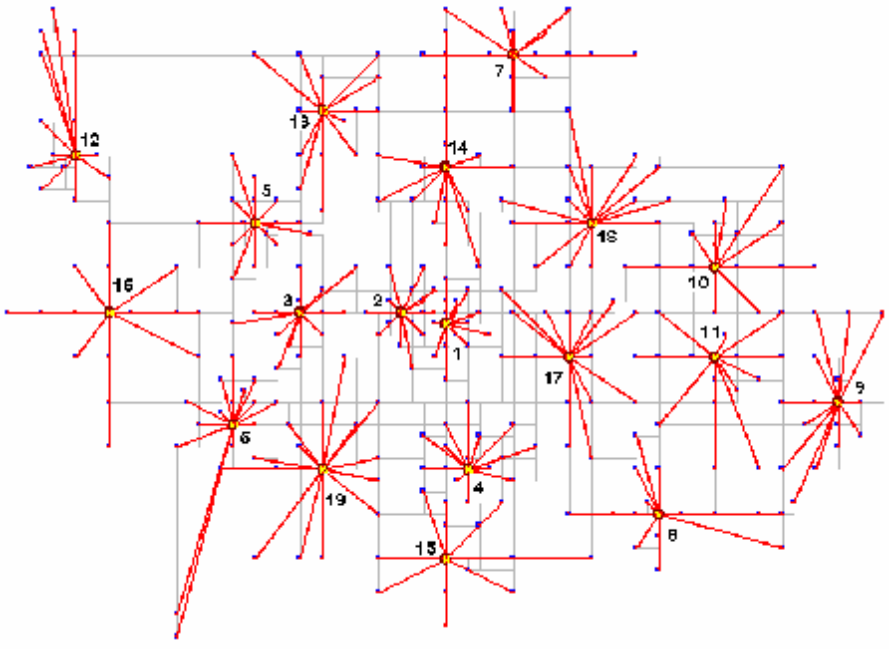


Figure 5-9 Three-layer metro network: Mux/Sels & edge node connectivity

The relatively small-scale core node location problem (MILP in Equation (4.19)) is solved with Enumeration Calculation when we assign two or three core nodes connected to 19 Mux/Sels.

In each iteration, the number and allocation of core nodes are assigned. The calculation

procedure is as follows.

First, for working network design, the connection between core nodes and Selector Switches are calculated using shortest path routing.

Second, for the reliable network design, all the possible routes between each source-destination pair are checked. If there is only one route, then another route is added for backup usage. It is clear that with the reliability requirement the network should have at least two core nodes. The algorithm is described as follows:

1) Check a source-destination pair. If there are at least two routes available for them, then go to step b. If there is only one route, then find another second-shortest path route and add a tag on the link that needs to be added.

2) Repeat step 1), until all source-destination pairs are considered. If there is no tagged link, then finish here, otherwise go to step 3).

3) Add the connectivity link which is with most tags, then go to step 1).

The Pareto boundary method in Section 4.3.4 is applied here. In light traffic case, if we set *c=fiber cost/distance* and w=0, Equation (4.19) represents the real cost. By varying the weighting factor w from 0 to some large number, corresponding designs put different relative importance on the two criteria, cost and delay.

When the capacity of one fiber is not sufficient for the traffic load, the allocation of normal and restored traffic to different links should be considered when dimensioning the network. In order to update the connectivity y_{ij} and capacity of the links, we apply the Minimum Cost Routing (MCR) algorithm [29] in all iterations of the enumeration calculation. Results show that this algorithm also works well in the metro network.

The design procedure for heavy traffic load case is similar to the light traffic load one. First some additional links are added after shortest path routing algorithm to guarantee that any source-destination pair has at least two different routes. In the MCR algorithm, for each source-destination pair, multiple failed working routes are restored by multiple restoration routes. The worst failure state for a given link is found. Then a new backup route is used to minimize the spare capacity on that link. The algorithm can yield near-optimal solution for spare optimization.

In the metro area, because the distance range is relatively small, the device cost (core node, Mux/Sel, interfaces) tends to dominate the core node allocation cost. However, for the purpose of reliability, at least two core nodes are needed to ensure that at least two routes are available for any source-destination Mux/Sel pair. Two and three core nodes are assigned in the metro designs for easier comparison of the results. Below we only give the core node distribution pictures for the case with three core nodes with heavy traffic load (Figure 5-10,

Figure 5-11, Figure 5-12 and Figure 5-13).

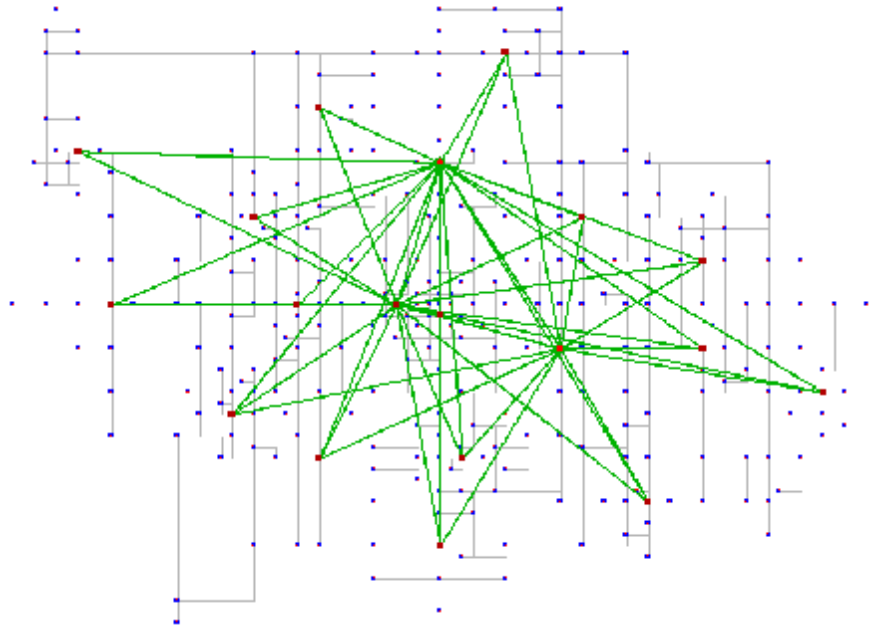


Figure 5-10 Three-layer metro network (heavy traffic): core nodes

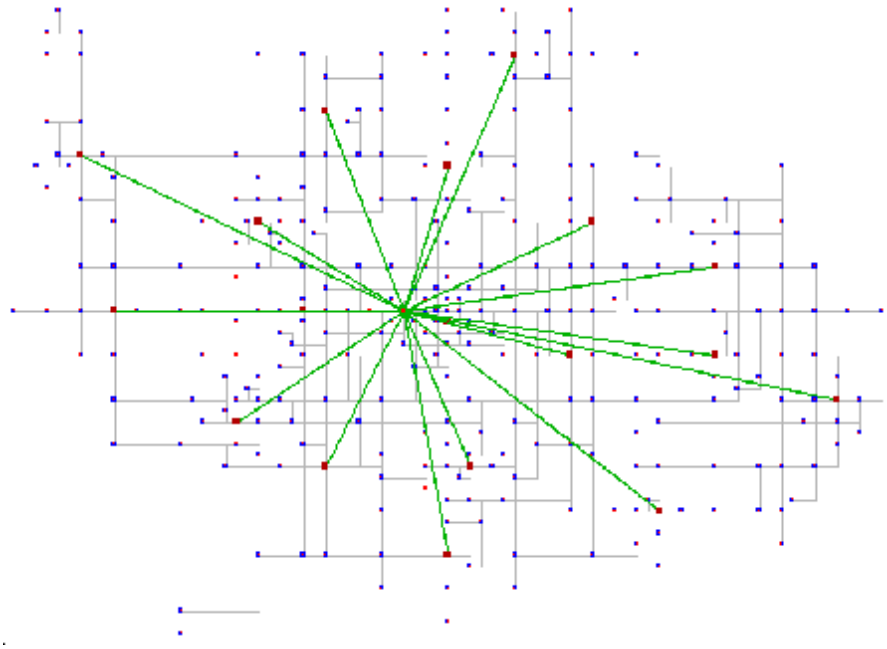


Figure 5-11 Three-layer metro network (heavy traffic): core node 1

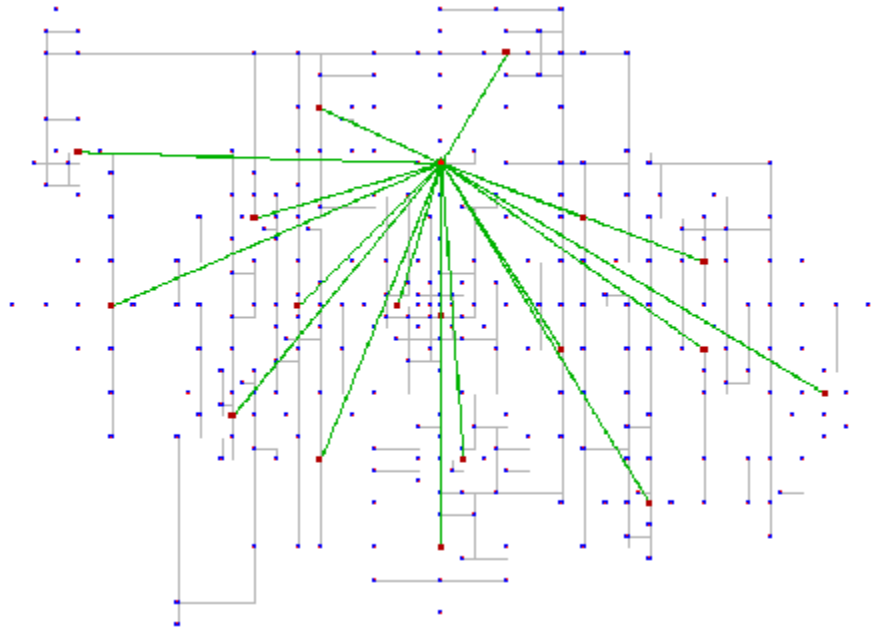


Figure 5-12 Three-layer metro network (heavy traffic): core node 2

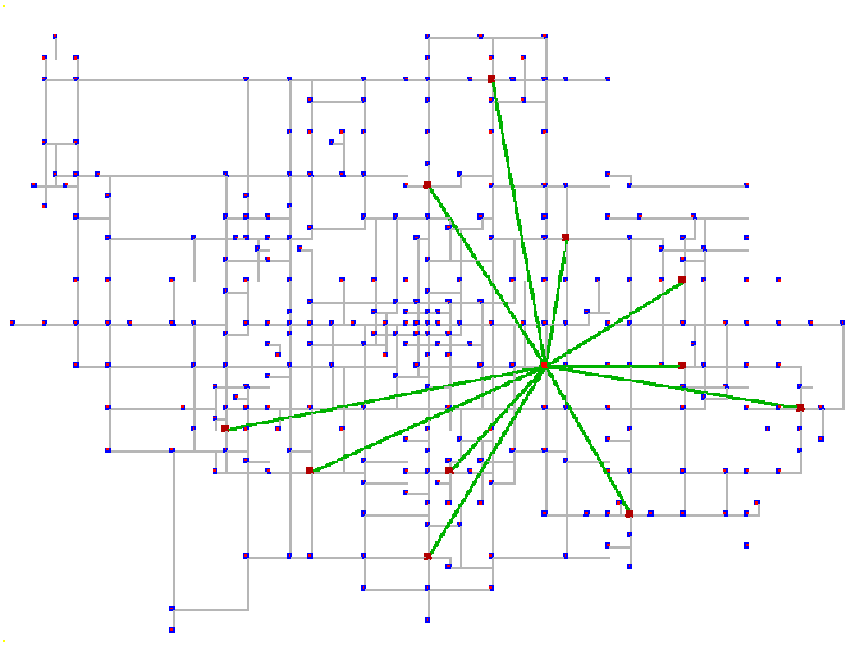


Figure 5-13 Three-layer metro network (heavy traffic): core node 3

The shortest path routing table is then calculated for both primary and backup connections (Figure 5-14). The links in this figure only indicates the logic links, and the real physical link connections are shown in Figure 5-15. The actual single-color and DWDM infrastructure is represented with different colors and are allocated along the Manhattan cabling models. In this figure, the grey lines represent Single Color Direct fibers while black lines are DWDM. Thickness of lines is proportional to the number of separate fibers per each cable section.

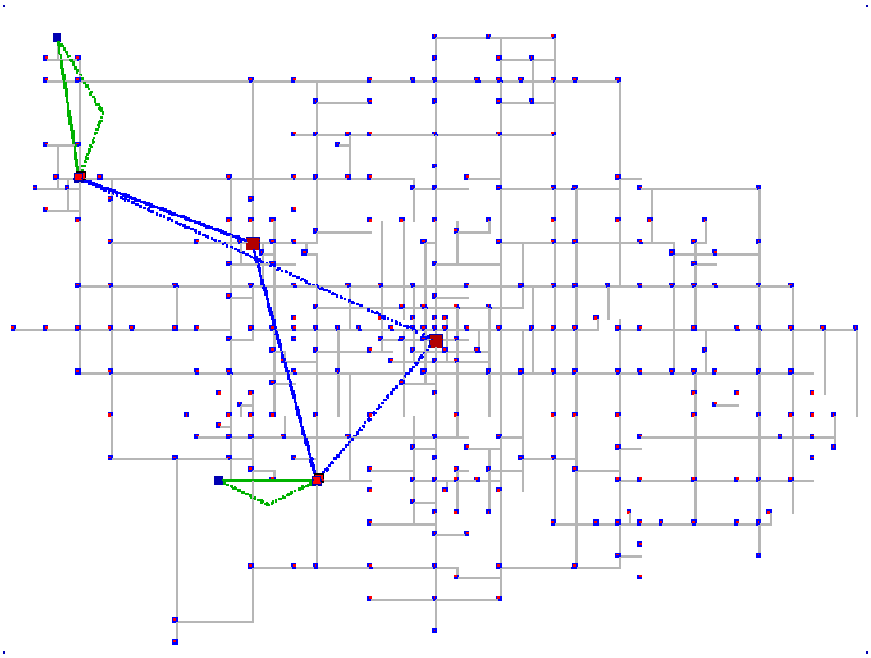


Figure 5-14 Three-layer metro network: primary and backup route

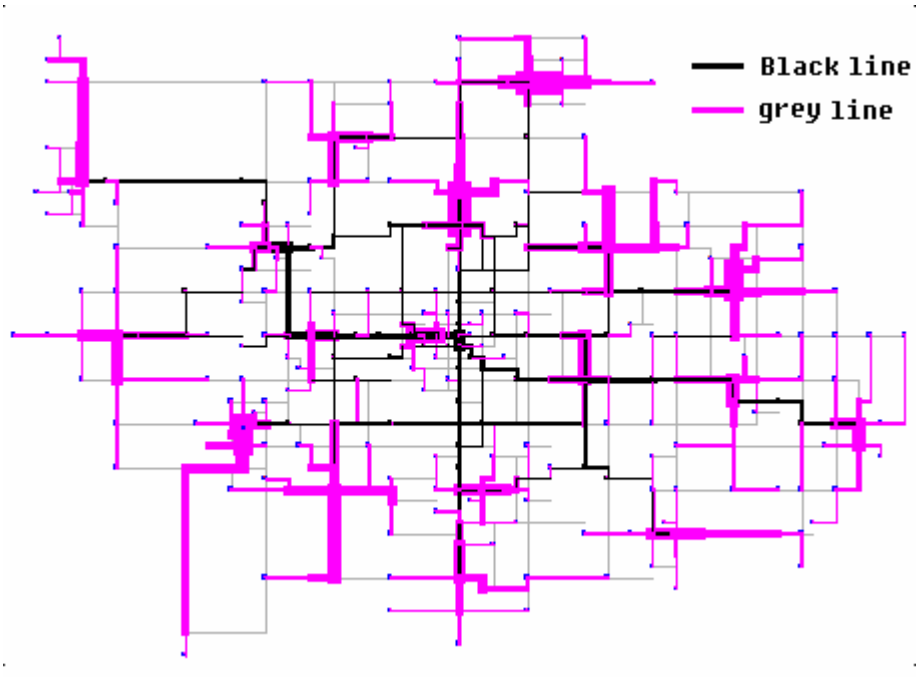


Figure 5-15 Three-layer metro network: fiber infrastructure

Table 5-4 Metro network design results with heavy traffic load

Layer	3-layer							2-layer			
Core locations	Unit price	2, 17			2, 13, 17			2, 17			
Item		Number	Sub Total	price %	Number	Sub Total	price %	Unit price	Number	Sub Total	price %
Core Nodes	100,000	2	200,000	0.21%	3	300,000	0.30%	100,000	2	200,000	0.21%
Core SC Ports	7,000	-	-	0.00%	-	-	0.00%	3,500	600	2,100,000	2.25%
Core DWDM Ports	30,000	38	1,140,000	1.19%	49	1,470,000	1.47%	15,000	38	570,000	0.61%
Mux/Sel Nodes	100,000	38	3,800,000	3.97%	49	4,900,000	4.89%	50,000	38	1,900,000	2.04%
Mux/Sel Ports	7,000	600	4,200,000	4.39%	774	5,418,000	5.40%	3,500	600	2,100,000	2.25%
Edge Nodes	200,000	300	60,000,000	62.75%	300	60,000,000	59.82%	200,000	300	60,000,000	64.42%
Edge Node Ports	7,000	600	4,200,000	4.39%	774	5,418,000	5.40%	3,500	600	2,100,000	2.25%
Location Startup	50,000	300	15,000,000	15.69%	300	15,000,000	14.95%	50,000	300	15,000,000	16.11%
Cable (m)	5	1,069,000	5,345,000	5.59%	1,107,000	5,535,000	5.52%	5	1,045,000	5,225,000	5.61%
Fiber (m)	0.4	4,343,000	1,737,200	1.82%	5,662,000	2,264,800	2.26%	0.2	19,705,000	3,941,000	4.23%
Total price				95,622,200			100,305,800				93,136,000

Table 5-5 Metro network design results with light traffic load

Layer	3-layer							2-layer			
Core locations	Unit price	2, 17			2, 14, 17			2, 17			
Item		Number	Sub Total	price %	Number	Sub Total	price %	Unit price	Number	Sub Total	price %
Core Nodes	100,000	2	200,000	0.18%	3	300,000	0.28%	100,000	2	200,000	0.19%
Core SC Ports	7,000	-	-	0.00%	-	-	0.00%	3,500	1,200	4,200,000	3.97%
Core DWDM Ports	30,000	76	2,280,000	2.06%	68	2,040,000	1.89%	15,000	76	1,140,000	1.08%
Mux/Sel Nodes	100,000	76	7,600,000	6.87%	68	6,800,000	6.31%	50,000	76	3,800,000	3.59%
Mux/Sel Ports	7,000	1,200	8,400,000	7.59%	1,074	7,518,000	6.97%	3,500	1,200	4,200,000	3.97%
Edge Nodes	200,000	300	60,000,000	54.20%	300	60,000,000	55.63%	200,000	300	60,000,000	56.69%
Edge Node Ports	7,000	1,200	8,400,000	7.59%	1,074	7,518,000	6.97%	3,500	1,200	4,200,000	3.97%
Location Startup	50,000	300	15,000,000	13.55%	300	15,000,000	13.91%	50,000	300	15,000,000	14.17%
Cable (m)	5	1,069,000	5,345,000	4.83%	1,117,000	5,585,000	5.18%	5	1,045,000	5,225,000	4.94%
Fiber (m)	0.4	8,686,000	3,474,400	3.14%	7,713,000	3,085,200	2.86%	0.2	39,410,000	7,882,000	7.45%
Total price				110,699,400			107,846,200				105,847,000

Table 5-6 Metro network designs with 8 and 32-interface Mux/Sels

Mux/Selportcount	8							32						
CoreNodemunber	UnitPrice	Twocorenodes			Threecorenodes			UnitPrice	Twocorenodes			Threecorenodes		
Item		Number	SubTotal	price%	Number	SubTotal	price%		Number	SubTotal	price%	Number	SubTotal	price%
CoreNodes	100,000	2	200,000	0.20%	3	300,000	0.30%	100,000	2	200,000	0.21%	3	300,000	0.31%
CoreSCPorts	7,000	-	-	0.00%	-	-	0.00%	7,000	-	-	0.00%	-	-	0.00%
CoreDWDMPorts	18,000	76	1,368,000	1.40%	79	1,422,000	1.44%	50,000	20	1,000,000	1.05%	22	1,100,000	1.13%
Mux/SelNodes	70,000	76	5,320,000	5.44%	79	5,530,000	5.58%	150,000	20	3,000,000	3.16%	22	3,300,000	3.40%
Mux/SelPorts	7,000	600	4,200,000	4.30%	624	4,368,000	4.41%	7,000	600	4,200,000	4.42%	655	4,585,000	4.72%
EdgeNodes	200,000	300	60,000,000	61.37%	300	60,000,000	60.57%	200,000	300	60,000,000	63.11%	300	60,000,000	61.81%
EdgeNodePorts	7,000	600	4,200,000	4.30%	624	4,368,000	4.41%	7,000	600	4,200,000	4.42%	655	4,585,000	4.72%
LocationStartup	50,000	300	15,000,000	15.34%	300	15,000,000	15.14%	50,000	300	15,000,000	15.78%	300	15,000,000	15.45%
Cable(m)	5	1,126,000	5,630,000	5.76%	1,184,000	5,920,000	5.98%	5	1,045,000	5,225,000	5.50%	1,134,000	5,670,000	5.84%
Fiber(m)	0.4	4,626,000	1,850,400	1.89%	5,359,000	2,143,600	2.16%	0.4	5,620,000	2,248,000	2.36%	6,313,000	2,525,200	2.60%
TotalPrice				97,768,400			99,051,600				95,073,000			97,065,200

5.5.2 Two-layer metro network

For a two-layer network, data will be transmitted via different routes for upstream and downstream. Upstream traffic from edge node goes directly to core node in single color fiber. For downstream traffic, selector switch or broadcast coupler is indispensable. With the selector switch node allocation resulted from three-layer design, enumeration calculation is applied, resulting in three core nodes. The selector switch node will be less expensive than the three-layer design because it can be a passive one and is used for only one direction. Also line and interface are calculated with one direction cost.

For the artificial city, Gotham, Two-layer design is marginally less expensive than the three layer design. Using the cost models for switching elements and transmission facilities, total switching costs strongly dominate transmission in metropolitan area applications of AAPN.

Table 5-4 shows the quantities and costs for metro network designs with light traffic load. And Table 5-5 gives the results with heavy traffic load. In both cases, 16-interface Mux/Sels have been used. The optimal designs in these two tables have taken into account both the cost and delay, which is calculated from the Pareto Boundary method in Section 4.3.4. Results show that although the two-layer design is cheaper than the three-layer design in our test cases, the difference is less than 10% and is sensitive to the initial settings of the cost assumptions. Another interesting thing is that in three-layer design, with the light traffic load, the two-core case is cheaper than the three-core one. But with the heavy traffic load, the results are opposite. The reason is that with light traffic load, traffic load does not influence the number of fibers. With more core nodes, more fibers are used only for the connectivity. So the three-core case is more expensive than the two-core one. But with heavy traffic load, with more core nodes, it is easier to adjust the traffic restoration to different links, and fewer fibers are used for the traffic supported than the two-core one. So the three-core case is less expensive.

Table 5-6 gives the optimal metro network designs using 8 and 32-interface Mux/Sels with light traffic load. The results in Table 5-6 are calculated only to minimize the total cost without considering the delay influence. The results show that designs with 8-interface Mux/Sels are cheaper than those with 32-interface ones. This is because the dominant costs in the metro network designs are attributed to the equipment cost. Again, the designs and costs are sensitive to the initial settings of cost parameters.

We can conclude from these results that two-layer and three-layer AAPN designs have similar costs which are sensitive to the cost of equipments, fiber and cables. According to our cost assumptions, two-layer AAPN designs will likely be the least costly option for metro networks. Mega cites with large population would likely benefit from optimized hybrid 2 and 3-layer designs and the use of DWDM equipment. And over provisioning for a

single high quality best effort class is a reasonable approach to design for metro networks.

Comparing WAN and MAN scenarios, it is clear that:

- The most expensive part of the network is cable&fiber in WAN, while equipment (especially edge nodes) dominates in the MAN design.
- DWDM and three-layer network is preferable for WAN while two-layer network better for MAN.

Chapter 6 Conclusions and future work

6.1 Thesis summary

This thesis has focused on the topological design and dimensioning of an AAPN. We have proposed a new mathematical programming model and a design methodology. A set of integrated planning tools has been developed which is intended to display and analyze the results.

In order to minimize network costs while satisfying performance requirements of the supported traffic, we are now facing the topological design challenges to determine the optimal number, size and placement of edge nodes, Multiplexer/Selectors and core switches, and to allocate direct fiber/DWDM links between them. Two networking models of WAN and MAN are tested, representing different geographic distributions, population densities, traffic flows and communities of interest. We have developed a set of modular software tools and methodologies for AAPN topological design and visualization in MATLAB and JAVA software. These tools are evaluated under various equipment cost assumptions, a set of circuit design alternatives, and two-layer and three-layer designs for both a metropolitan network and a long-haul network, assuming a gravity model for traffic distribution with a flat community of interest factor. We also show the topology sensitivity to the equipment cost and the number of ports, and show the preferred size of selector switches.

The design schemes proposed enable network planners to design near optimal network topologies given the facility and equipment cost capacity information and the traffic demand. The graphical display of the TD Tool allows the planner to view the resulting network designs. This optical network planning tool should be useful to both equipment vendors and network operators as it provides an effective means of evaluating the network cost and the performance impact of various equipment design options (capacities and costs) as well as alternative network architectures under different hypothesized traffic demand scenarios.

6.2 Future work

From the internal proposal for AAPN by Lorne Mason, the emphases for our future work would be enriching the functionality and performance of our TD Tool. The main objectives include:

- Develop queuing models for network dimensioning purposes. Currently we use a utilization factor obtained by OPNET simulation for specific scheduling schemes that is very time consuming. The availability of queuing models will accelerate the network design process by permitting rapid assessment of a broader range of design options than it is at present. These queuing models will also facilitate dimensioning for load shared routing.
- Propose alternative "subgradient" that do not have the discontinuity associated with the dual's gradient.
- Extend the economic analyses to include Net Present Value (NPV) of the revenues, capital and expense cash flows over a finite planning horizon taking into account the dynamic growth in demand and provisioned capacity. This will facilitate planning for alternative service offerings and tariff requirements for profitability. In particular it should be useful in assisting operators in setting tariffs for services in order to assure profitability over a planning horizon.
- Design the prototype of AAPN TD Tool for commercial use. In order to make further enhancements on the functionality and richness of output of this AAPN topological design and visualization tool, a prototype suitable for the commercial use should be developed. Additional coding and documentation are necessary. For example, Matlab and java coded programs can be integrated into the program as background applications and can be called by prompts or some input windows.
- Extend the TD Tool to include the traffic demand by service categories and provisioned capacity over a planning horizon. The existing tool produces the aggregate traffic demand workload at a snapshot in time. The enhancement involves the development of software for the forecasting traffic demand and the distribution by service category as well as aggregate demand. Based on the forecasted traffic demand by service category per user, we will build up workload traffic matrices by service category based on demographic data and forecasted penetration levels. This bottom up approach to traffic forecasting will be compared with top down estimates obtained from aggregate traffic forecasts to reconcile the two approaches.
- Extend the TD Tool outputs to include network transmission and traffic performance metrics. Based on estimated impairments obtained from other researchers for various

circuit elements we will use the design tool to compose these into end-to-end and network impairments such as S/N ratios, attenuation etc at a network as opposed to a device or span level. The traffic impairments such as the blocking delay and jitter will be computed at a network and end-to-end level based on the queuing models of individual network elements for the specified traffic workload. This proposed extension to the TD Tool will provide a richer description of the designed networks to assist network planners in the decision making process of selecting alternative network designs and evolution scenarios.

- Modeling an AAPN implementation with an existing equipment and facility infrastructure, to provide a migration path compatible with the evolution of existing and planned transport networks.

Appendix: Data sets used in simulations

Cost of core node:

$C_{core} = 10^5$: Start up cost for core switch, no interface cost is included

$C_{core\_full} = 2 \times 10^6$: fully loaded 64×64 optical switches cost

$C_{coreIF\_mux16} = 3 \times 10^4$: Cost of core switch interface for 16-interface Mux/Sel connection (160Gb/s)

$C_{coreIF\_mux32} = 5 \times 10^4$: Cost of core switch interface for 32-interface Mux/Sel connection (320Gb/s)

$C_{coreIF\_edge} = 7000$: Cost of core switch interface for edge connection (10Gb/s)

Cost for Mux/Sel

$Cm_{16} = 10^5$: Start up cost for 16-interface Mux/Sel, no interface cost is included

$Cm_{32} = 1.5 \times 10^5$: Start up cost for 32-interface Mux/Sel, no interface cost is included

$C_{mux\_IF\_LH} = 7000$: cost/Mux/Sel downlink interface for Long Haul connection (used when edge is connected to Mux/Sel or core with single fiber.)

$Cm_{16\_full} = 2 \times 10^5$; (More expensive than a normal DWDM equipment)

$Cm_{32\_full} = 3.7 \times 10^5$.

Note: In the connection between Mux/Sel and core node, cost for DWDM equipment is included in the price of Mux/Sel equipment and core Interface. But in the connection from edge to Mux/Sel, if the total traffic from one city to another need DWDM connection, the DWDM equipment cost is a separate part.

Cost of Edge node

$C_{edge} = 2 \times 10^5$: Start up cost for edge node.

Uplink interface cost: 7000

Link cost

Cable cost:

$C_{cable} = 5000$: Cable cost per kilometre including cable installation, right of way cost etc. One cable has infinite capacity of fiber.

Single fiber:

$Cf = 300$: single fiber cost/kilometre (10Gb/s, without amplifiers, regenerators etc)

$Cf_{amp} = 400$: single fiber cost/kilometre (10Gb/s, with amplifiers, regenerators etc)

$C_{amp} = 4000$: Amplifier cost for single fiber

$C_{reg} = 10^4$: Regenerator cost for single fiber

$max_{amp} = 80$: Distance span for amplifier

$max_{reg} = 600$: Distance span for regenerator

(Price source: From Nortel transmission cost model and information from industry.)

DWDM:

$C_{DWDM} = 5 \times 10^5$: DWDM terminal equipment: (16 wavelengths, full loaded)

$C_{D\_amp} = 4 \times 10^4$: DWDM amplifier cost, for 16 wavelengths altogether

$C_{D\_amp} = 7 \times 10^4$: DWDM amplifier cost, for 32 wavelengths altogether

$C_{D\_reg} = 3 \times 10^4$: DWDM regenerator cost, this is price for each wavelength in it

So DWDM fiber cost with everything included roughly is:

$Cf_{DWDM} = 3000\$ / \text{km}$

References

[1] Agile All Photonic Networks (AAPN), "Research Planning Document".

[2] O.Kariv and SL Hakimi, "An algorithmic approach to network location problems. ii: the p-medians", *SIAM Journal of Applied Mathematics*, vol. 37, no. 3, pp. 539-560, December 1979.

[3] C.Beltran, C.Tadonki, and J.-Ph.Vial, "Solving the *p-median* problem with a semi-Lagrangian relaxation", University of Geneva, 2004.

[4] E.L.F.Senne and L.A.N.Lorena, "A Lagrangian/ surrogate approach to *p-median* problems", *Computers and Operations Research*, March 1998.

[5] ES Correa, MTA Steiner, AA Freitas, and C Carnieri, "A genetic algorithm for solving a capacitated *p-median* problem", *Numerical Algorithms*, vol. 35, no. 2-4, pp. 373-388, April 2004.

[6] E.L.F.Senne, L.A.N.Lorena, and M.A.Pereira, "A branch-and-price approach to *p-median* location problems", *Computers & Operations Research*, vol. 32, no. 6, pp. 1655-1664, 2005.

[7] Resende MGC and Werneck RF, "A hybrid heuristic for the *p-median* problem", *Journal of Heuristics*, vol. 10, no. 1, pp. 59-88, January 2004.

[8] Pitu B.Mirchandani and Richard L.Francis, "Discrete Location Theory", Wiley, July 1990.

[9] Hugues Delmaire, Juan A.Diaz, Elena Fernandez, and Maruja Ortega, "Comparing new heuristics for the pure integer capacitated plant location problem", Investigacion Operativa, pp. 217-242, 1999.

[10] J.E.Beasley, "Lagrangean heuristics for location problems", *European Journal of Operational Research*, vol. 65, pp. 383-399, 1993.

[11] J.E.Beasley, "An algorithm for solving large capacitated warehouse location problems", *European Journal of Operational Research*, vol. 33, pp. 314-325, 1998.

[12] JA Diaz and Fernandez, "A branch-and-price algorithm for the single source capacitated plant location problem", *Journal of the Operational Research Society*, vol. 53, pp. 728-740, 2002.

[13] R.Sridharan, "A Lagrangian heuristic for the capacitated plant location problem with single source constraints", *European Journal of Operational Research*, vol. 66, pp. 305-312, 1993.

[14] R.Sridharan, "The capacitated plant location problem", *European Journal of*

Operational Research, vol. 87, pp. 203-213, January 1995.

[15] Qian Wang, Rajan Batta, Joyendu Bhadury, and Christopher M.Rump, "Budget constrained location problem with opening and closing of facilities", *Computers & Operations Research*, vol. 30, no. 13, pp. 2047-2069, November 2003.

[16] KS Hindi and K. Pienkosz, "Efficient solution of large scale, single-source,capacitated plant location problems", *Journal of the Operational Research Society*, vol. 50, pp. 268-274, 1999.

[17] Suda Tragantalerngsak, John Holt, and Mikael Ronnqvist, "Lagrangian heuristics for the two-echelon, single-source, capacitated facility location problem", *European Journal of Operational Research*, vol. 102, no. 3, pp. 611-625, 1997.

[18] Michal Pioro and Deepankar Medhi, "Routing, flow, and capacity design in communication and computer networks", Morgan Kaufman Publishers, June 2004.

[19] Samit Soni, Sridhar Narasimhan, and Larry J.LeBlanc, "Telecommunication access network design with reliability constraints", *IEEE Transactions on Reliability*, vol. 53, no. 4, pp. 532-541, December 2004.

[20] Bezalel Gavish, "Topological design of computer communication networks", Proceedings of the 22-nd HICSS, vol. 3, pp. 770-779, 1989.

[21] Rajeev Kumar, Prajna P.Parida, and Mohit Gupta, "Topological design of communication networks using multiobjective genetic optimization", Evolutionary Computation, 2002. CEC '02, vol. 1, pp. 425-430, May 2002.

[22] M.Pioro, A. J ttner, J. Harmatos, A. Szentesi, P. Gajowniczek and A. Myslek, "Topological design of telecommunication networks - nodes and links localization under demand constraints", 17th ITC, September 2001.

[23] Steven Chanberland, Brunilde Sanso, and Odile Marcotte, "Topological design of two-level telecommunication networks with modular switches", *Operations Research*, vol. 48, no. 5, pp. 745-760, September 2000.

[24] Rajesh M.Krishnaswamy and Kumar N.Sivarajan, "Design of logical topologies: A linear formulation for wavelength-routed optical networks with no wavelength changers", *IEEE/ACM Transaction on networking*, vol. 9, no. 2, pp. 186-198, April 2001.

[25] J.A. Bannister, L. Fratta, and M.Gerla, "Topological design of the wavelength-Division Optical Network", Infocom '90 of IEEE, vol. 3, pp. 1005-1013, June 1990.

[26] Yufeng Xin, George N.Rouskas, and Harry G.Perros, "On the physical and logical topology design of large-scale optical networks", *Journal of Lightwave Technology*, vol. 21, no. 4, pp. 904-915, April 2003.

[27] Bezalel Gavish, Pierre Trudeau, Moshe Dror, Michel Gendreau, and Lorne Mason, "Fiberoptic circuit network design under reliability constraints", *IEEE Journal on Selected Areas in Communication*, vol. 7, no. 8, pp. 1181-1187, October 1989.

[28] Anthony Sack and Wayne D.Grover, "Hamiltonian p-cycles for fiber-level protection in homogeneous and semi-homogeneous optical networks", *IEEE Network*, vol. 18, no. 2,

pp. 49-56, March 2004.

[29] Yijun Xiong and Lorne Mason, "Restoration strategies and spare capacity requirements in self-healing ATM networks", *IEEE/ACM Transactions on Networking*, vol. 7, no. 1, pp. 98-110, February 1999.

[30] Yang Qin, Lorne Mason, and Ke Jia, "Study on a joint multiple layer restoration scheme for IP over WDM networks", *IEEE Network*, vol. 17, no. 2, pp. 43-48, March 2003.

[31] John Doucette, Matthieu Clouqueur, and Wayne D.Grover, "On the availability and capacity requirements of shared backup path-protected mesh networks", *Optical Networks Magazine*, November 2003.

[32] Lorne Mason, Anton Vinokurov, Ning Zhao, and David Plant, "Topological design and dimensioning of agile all photonic networks", *Elsevier Computer Networks Journal*, (To be published).

[33] Anton L.Vinokurov and Lorne Mason, "AAPN Architecture Options (poster)", AAPN Annual Research Review, June 2005.

[34] A.Farago, "Blocking probability estimation for general traffic under incomplete information", Proc. IEEE ICC, June 2000.

[35] Giray Birkan, Jeffery Kennington, Eli Olinick, Augustyn Ortynski, and Gheorghe Spiride, "Optimization-based design strategies for DWDM networks: Opaque versus All-Optical networks", Southern Methodist University, May 2003.

[36] Mansoor Alicherry, HarshaNagesh, and Vishy Poosala, "Constraint-based design of optical transmission systems", *Journal of Lightwave Technology*, vol. 21, no. 11, pp. 2499-2510, November 2003.

[37] Christopher Joshua Ito, "All-Optical 3R regeneration for Agile All-Photonic networks", Queen's University, July 2003.

[38] B.Khasnabish, "Topological properties of Manhattan street networks", *Electronics letters 28th*, vol. 25, no. 20, September 1989.

[39] JT Brassil and RL Cruz, "Nonuniform traffic in the Manhattan Street Network", Proceedings of ICC'91, vol. 3, pp. 1647-1651, June 1991.

[40] "http://cepa.newschool.edu/het/essays/paretian/paretoptimal.htm".

[41] "http://members.aol.com/prfeubanks/pareto.PDF".

[42] AA Kuehn and MJ Hamburger, "A heuristic program for locating warehouses", *Management Science*, vol. 9, no. 4, pp. 643-665, July 1963.

[43] P.M.Camerini, L. Fratta, F. Maffioli, "On Improving Relaxation Methods by Modified Gradient Techniques", *Mathematical Programming Study*, vol. 3, pp. 26-34, 1975.

[44] H.D. Sherali, and O. Ulular, "A Primal-Dual Conjugate Subgradient Algorithm for Specially Structured Linear and Convex Programming Problems", *Applied Mathematics and Optimization*, Vol.20, pp. 193-221, 1989.