DESIGN OF A SURVIVABLE METROPOLITAN AGILE ALL-PHOTONIC NETWORK

Ning Zhao, Anton Vinokurov, Lorne Mason

Department of Electrical and Computer Engineering, McGill University

ABSTRACT

We present models and methods for the design and dimensioning of a survivable metropolitan Agile All-Photonic Network (AAPN). This paper discusses the layered topology that is comprised of a set of overlaid star/star networks, with an optical core space switch at each of the star centres, hybrid photonic/electronic switches at the edges, and Multiplexer/Selectors in between. Network cost is minimized while taking into consideration performance criteria such as delay and reliable traffic restoration upon network failure. A mixed integer linear programming formulation is presented for core node placement and link connectivity to determine the near cost optimal designs. Network models and their performance were evaluated with a set of software tools and methodologies to design and dimension our vision of an AAPN.

Keywords: *Optical Networks, AAPN, MAN, Survivable Network Design*

1. INTRODUCTION

This paper focuses on survivable metropolitan area network design, where the network is resilient to any single link failure by preplanned spare capacity and restoration procedures. The AAPN architecture described in [1] and [2] has been proposed as a high speed transparent optical transport network utilizing sub microsecond optical switching elements. These elements are interconnected as illustrated in Figure 1. Alternative optical path designs are depicted in Figures 2 and 3 for the symmetric and asymmetric architectures respectively. The overlaid star/star topology of an AAPN facilitates network synchronization required for Optical Time Division Multiplexing (OTDM) of individual wavelengths or colors. A comparison of different OTDM methods [3] and Optical Burst Switching (OBS) [4] has been reported. OBS implemented without fiber delay lines or wavelength conversion, does not allow to achieve an acceptable loss rate even if traffic is low [5]. OTDM permits several connections to flexibly share a color's bandwidth, thereby extending the reach of the all optical transport network closer to the end users. In addition the statistical multiplexing made possible by sub micro second switching enables efficient handling of bursty traffic relative to that of the slower MEM's based optical cross connect equipment currently being deployed in transport networks.

Network design models and optimization approaches have been studied in [6] for a variety of technologies. The topological design and dimensioning methodology for the working AAPN in WAN and MAN applications has been described in detail in [2]. The topological design and dimensioning problem consists in determining the number, capacity and location of switching elements and their interconnection pattern so as to minimize costs while meeting performance and availability objectives. Because this complex problem is intractable, the optimization problem has been decomposed into separate components. The method applied is to separate the design of core node allocation (backbone networks) and Mux/Sel allocation (access networks). We also separate the design of the working network and the spare or backup network into a two stage process. The solution techniques for survivable AAPN MAN design has been carried out in this paper and design results are displayed for different traffic scenarios using the switch and transmission cost models reported in [2].

We have developed a set of modular software tools and methodologies in CPLEX, Matlab and custom Java applications. These tools are employed under various equipment cost assumptions, to evaluate a set of circuit design alternatives, and two-layer and three-layer designs in a metropolitan network, assuming a gravity model for traffic distribution with a flat community of interest factor.

The rest of the paper is organized as follows. In Section 2 we give the modeling of the network architecture for a MAN. In Section 3 we provide the solution procedure for designing and dimensioning of the network. In Section 4 we provide the computational results based on our solution procedure. In Section 5 we conclude and give some comments.

2. MODELING THE NETWORK

According to the number and traffic demands of edge nodes of the AAPN, an investigation of a three and two-layer network architectures of a MAN has been carried out. Figure 1 shows an example of the three-layer overlaid tree topology. Figure 2 and 3 give the end to end connectivity of two and three-layer architectures.

In Figure 2, the symmetric three-layer network design is given. Traffic routes for both upstream and downstream directions are the same. Edge nodes are connected to Mux/Sels with single wavelength fiber, and Mux/Sels are connected to core nodes with Dense Wavelength Division Multiplexing (DWDM). In the two-layer design of Figure 3, for upstream traffic, source edge node ports are connected directly to core node switching plane without any intermediate aggregation. However, in the downstream direction, if the number of edge nodes is greater than the number of DWDM links from one core node, the Mux/Sel should be used, 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.

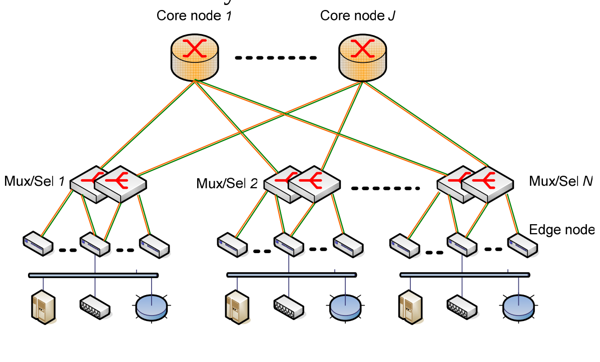

Figure 1 AAPN network model: three-layer overlaid tree topology

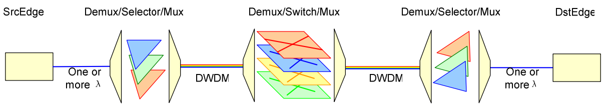

Figure 2 Three-layer network Architecture

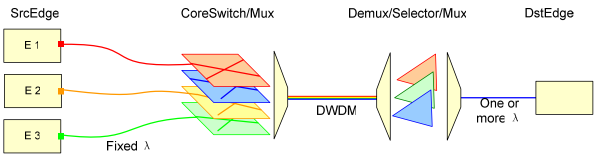

Figure 3 Two-layer network architecture

3. THE NETWORK DESIGN PROCEDURE

3.1 Mux/Sel 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 formulated as a Capacitated Plant Location Problem (CPLP) [7]. This problem is solved using a Lagrangian Relaxation (LR) approach along with CPLEX optimization and the Hamburger Heuristic [8]. The Hamburger method is 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, CPLEX program is applied to get the optimal connectivity design between edge nodes and Mux/Sels.
3) Each Mux/Sel and edge nodes that it connects are in one group. For each group, find the optimal Mux/Sel locations with least total distances. If any of them is different from current Mux/Sel locations, use the new locations as input and go to step 2). Otherwise, current solution is the desired result.

3.2 Core node allocation

3.2.1 Problem formulation

The core node location problem in the metro network is done with the enumeration method. Firstly a Mixed Integer Linear Programming (MILP) problem for the core node allocation is formulated as follows:

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

Subject to the capacity constraint and traffic demand constraints as below:

$$\sum\nolimits_{j=1}^{J} \alpha_{ik}^{j} = 1 \quad i,k = 1 \ldots N,\ j = 1 \ldots J \quad (2)$$

$$N \cdot y_{ij} \ge \sum\nolimits_{k=1}^{N} \alpha_{ik}^{j} \quad i,k = 1 \ldots N,\ j = 1 \ldots J \quad (3)$$

$$N^2 \cdot z_j \ge \sum\nolimits_{i=1}^{N} \sum\nolimits_{k=1}^{N} \alpha_{ik}^{j} \quad i,k = 1 \ldots N,\ j = 1 \ldots J \quad (4)$$

$$z_j \in \{0,1\},\ y_{ij} \in \{0,1\},\ \alpha_{ik}^{j} \in \{0,1\}\ i,k = 1 \ldots N,\ j = 1 \ldots J \quad (5)$$

The design formulation starts with the traffic sources (Mux/Sels) i, where $i=1 \ldots N$ is the Mux/Sel index in the set of N Mux/Sels. Suppose that a core node can only be located in place where there is already a Mux/Sel. $j=1 \ldots J$ is the core node index and $J<=N$.

λ_{ik} : Traffic demands between Mux/Sel i and k. The traffic is symmetric, which means $\lambda_{ik} = \lambda_{ki}$.

z_j, $j = 1 \ldots J$: The core node existence. $z_j = 1$ if there is a core node at location j and 0 otherwise.

C_{core} : The start-up cost for each core node.

C_{IF} : The core interface cost to connect the Mux/Sel.

d_{ij} : The distance between Mux/Sel i and core node j.

y_{ij} : The connection between Mux/Sel i and core node j. $y_{ij} = 1$ if there is a connection between core node j and Mux/Sel i and 0 otherwise.

α_{ik}^{j} : The proportion of traffic from Mux/Sel i to Mux/Sel k routed through core node j. If all traffic follows the shortest path routing, $\alpha_{ik}^{j} = 0$ 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.

b_{ij} : The number of DWDM links between Mux/Sel i and core node j. This coefficient is calculated from the traffic load and spare capacity requirement on each link.

c and w: Two constants which we will discuss later.

One more constraint for reliable network design is that for each source-destination Mux/Sel pair, at least two different routes passing two different core nodes should be available in case of the single link failure.

The solution methods are given with respect to two kinds of traffic assumptions as follows.

3.2.2 Light traffic load solution

The light traffic load means that the link capacity is sufficient for all normal traffic and traffic restoration requirement for any single link failure. Thus in this case b_{ij} is always 1 and does not change with different traffic load distributions. As in a metro network, there are limited edge nodes and thus the number of Mux/Sel is relatively small, the MILP problem of Equation (1) can be solved with enumeration calculation. 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).

For the best selection of c and w in Equation (1), we will do the following. The $w\sum_{j=1}^{J}\sum_{i=1}^{N}\sum_{k=1}^{N}\lambda_{ik}\cdot\alpha_{ik}^{j}\cdot(d_{ij}+d_{jk})$ gives the cost of DWDM fibers linking Mux/Sels to all core nodes. And at the same time, it is also proportional to the traffic weighted delay since distance and delay are linearly related by propagation speed. Thus we do not need to add another term if we wish to emphasize the delay aspect. The average delay of whole network is defined as:

$$delay = \frac{\sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=1}^{J}\alpha_{ik}^{j}\cdot\left(d_{ij}+d_{jk}\right)}{N^2\cdot c_l} \qquad (6)$$

$c_l = 0.75\cdot 3\cdot 10^8$ is the speed of light in glass.

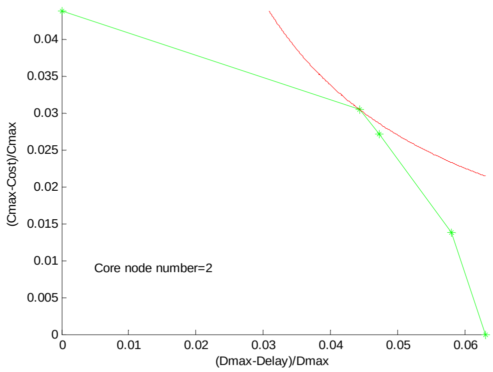

Figure 4 Pareto Boundary of delay vs. cost (2 cores)

In light traffic case, if we set *c=fiber cost/distance* and *w*=0, Equation (1) represents the real cost. By varying the weighting factor *w* from 0 to some large number, Equation (1) will account for different relative importance of the two criteria, cost and delay, when selecting the optimal network topology, then we can get the maximum of delays, D_{max}, and the maximum of costs, C_{max}. Viewed as a multicriterion problem, the Pareto Boundary is plotted in Figure 4. In this figure, the star point in the left top corner represents the case with minimum cost and maximum delay, while the star point in the right bottom corner represents the case with maximum cost and minimum delay. The point of interest is obtained where the hyperbola is tangent to the curve with the polygonal line with stars.

3.2.3 Heavy traffic load solution

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 b_{ij} of the links, we apply the Minimum Cost Routing (MCR) algorithm [9] in all iterations of the enumeration calculation. Results show that this algorithm also works well in our network.

The design procedure is similar to the light traffic load case. 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.

3.2.4 Two-layer network design

For a two-layer network, data will be transmitted via different routes for upstream and downstream directions. With the selector switch node allocation resulting from three-layer design, the enumeration calculation is also applied to get the optimal allocation of core nodes. The selector switch node will be much cheaper than the three-layer design because it can be a passive device and is used for only one direction. Also the line and interface costs are calculated for uni directional transmission

4. COMPUTATIONAL STUDY

In absence of real data for a city such as fiber embedding information, we have created an artificial city named Gotham. In this model, the population density distribution is simulated with a Gaussian random process, and edge nodes are distributed according to equalize population-per-location in a Manhattan network layout. Cable infrastructure is simulated with a modified spanning tree algorithm. Gotham city has a size of 60km*80km and a population of 4.5 million connected to 300 edge nodes. A detailed list of data sets used is given in [10].

Two traffic load models are considered in the design. One is the light traffic model with 10 Mb/s inbound/outbound from each edge node to another edge, and another is the heavy traffic model with 30Mb/s. Assuming that each Mux/Sel serves 16 edge nodes and the scheduler performance results is 80% utilization, in the light traffic load case, traffic from one Mux/Sel destined for another Mux/Sel is 10Mb/s×16×16=2.56Gb/s. The total traffic from one Edge is 10M×300/0.8=3.75Gb/s, which is less than the single fiber’s capacity of 10Gb/s.

For any Mux/Sel to all core nodes connected, the total traffic is: 2.56x19/0.8=60.8 Gb/s. If the capacity of one DWDM link is 16×10=160Gb/s, the minimum spare capacity in a DWDM fiber is 160-60.8=99.2Gb/s. Thus in all failure states, the spare capacity is always sufficient. Similarly, we can calculate the traffic parameters when the heavy traffic model is used. Figure 5 shows the 16-interface Mux/Sel locations.

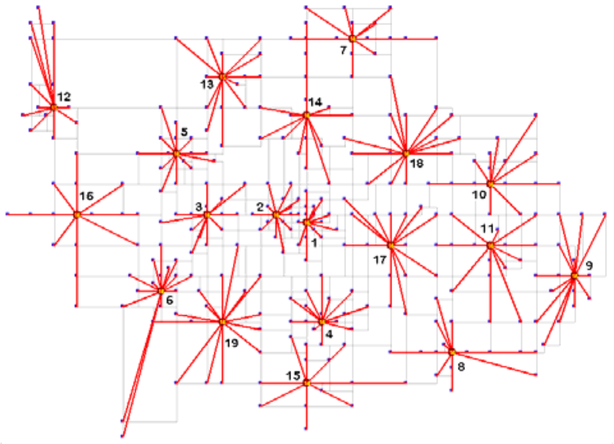

Figure 5 Three-layer network: Mux/Sel Allocation

The relatively small-scale core node location problem is solved with Enumeration Calculation. The test results are given in the following Table 1 and 2:

Layer	3-layer				2-layer	
Core locations	2, 17		2, 13, 17		1, 2	
Item	Number	Cost %	Number	Cost %	Number	Cost %
Core Nodes	2	0.21%	3	0.30%	2	0.21%
Core SC Ports	0	0.00%	0	0.00%	600	2.25%
Core DWDM Ports	38	1.19%	49	1.47%	38	0.61%
Mux/Sel Nodes	38	3.97%	49	4.89%	38	2.04%
Mux/Sel Ports	600	4.39%	774	5.40%	600	2.25%
Edge Nodes	300	62.75%	300	59.82%	300	64.42%
Edge Node Ports	600	4.39%	774	5.40%	600	2.25%
Location Startup	300	15.69%	300	14.95%	300	16.11%
Cable installation (km)	1069	5.59%	1107	5.52%	1045	5.61%
Fiber (km)	4343	1.82%	5662	2.26%	19705	4.23%
Total price	95,622,200		100,305,800		93,136,000	

Table 1 Light traffic load results

Layer	3-layer				2-layer	
Core locations	2, 17		2, 14, 17		1, 2	
Item	Number	Cost %	Number	Cost %	Number	Cost %
Core Nodes	2	0.18%	3	0.28%	2	0.19%
Core SC Ports	0	0.00%	0	0.00%	1200	3.97%
Core DWDM Ports	76	2.06%	68	1.89%	76	1.08%
Mux/Sel Nodes	76	6.87%	68	6.31%	76	3.59%
Mux/Sel Ports	1200	7.59%	1074	6.97%	1200	3.97%
Edge Nodes	300	54.20%	300	55.63%	300	56.69%
Edge Node Ports	1200	7.59%	1074	6.97%	1200	3.97%
Location Startup	300	13.55%	300	13.91%	300	14.17%
Cable installation (km)	1069	4.83%	1117	5.18%	1045	4.94%
Fiber (km)	8686	3.14%	7713	2.86%	39410	7.45%
Total price	110,699,400		107,846,200		105,847,000	

Table 2 Heavy traffic load results

5. CONCLUSIONS AND COMMENTS

The AAPN survivable network design problem for metropolitan area networks has been presented. To reduce complexity of the design procedure we have decomposed the design problem into a set of simpler problems for edge node placement, Mux/Sel placement and core node placement respectively and link connectivity for the working AAPN network architecture. To provide resilience to single link failures a procedure has been developed to augment capacity where needed. For the artificial city, Gotham, alternative designs have been created and compared under light and heavy traffic assumptions. 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.

6. REFERENCES

[1] G.V. Bochmann, T. Hall, O. Yang, M.J. Coates, L. Mason, R. Vickers, "The Agile All-Photonic Network: an architectural outline", *Queen's Biennial Conference on Communications*, 2004.

[2] L. Mason, A. Vinokurov, N. Zhao, and D. Plant, "Topological design and dimensioning of agile all photonic networks", *Journal of Elsevier Computer Networks* (in press, web published), 2005.

[3] X. Liu, N. Saberi, M. Coates and L. Mason, "A Comparison between time-slot scheduling approaches for All-Photonic Networks", *in Proc. IEEE ICICS*, December 2005.

[4] X. Liu, A. Vinokurov, and L. Mason, "Performance comparison of OTDM and OBS scheduling for agile all-photonic network", *in Proc. IFIP Metropolitan Area Network Conference*, April 2005.

[5] A. Vinokurov, X. Liu, and L. Mason, "Resource sharing for QoS in Agile All Photonic Networks", *OPNNETWORK 2005*, August 2005.

[6] M. Pioro and D. Medhi, *"Routing, Flow, and Capacity Design in Communication and Computer Networks"*, Morgan Kaufmann Publishers, 2004.

[7] P. Mirchandani and R. L.Francis, *"Discrete Location Theory"*, John Wiley and Sons, July 1990.

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

[9] Y. Xiong and L. 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.

[10] N. Zhao, *"Topological design and dimensioning of agile all photonic networks"*, Master thesis, McGill University, June 2005.