Simple Layout Algorithms To Maintain Network Connectivity Under Faults

Galen H. Sasaki[1]
Department of Electrical, University of Hawaii
2540 Dole Street, Honolulu, HI 96822 USA

Ching-Fong Su and David Blight
Fujitsu Laboratories of America
595 Lawrence Expressway, Sunnyvale, CA 94086-3922 USA

Abstract: A problem of survivable layout of an IP network topology on a wavelength division multiplexed network is considered. Simple layout algorithms are given that have low time complexities. The algorithms are variations of a basic algorithm, that lays out IP links as lightpaths sequentially according to some order and with respect to link costs. Simulations show that the ordering and link costs can significantly affect the performance of the algorithm.

1. Introduction

Wavelength division multiplexing (WDM) is the current technology of choice for broadband backbone networks. WDM long-haul, transport systems are being deployed due to their low cost per bit, good quality of service, and very wide bandwidths. WDM is essentially frequency division multiplexing of optical signals onto optical fibers, and *wavelengths* are equivalent to carrier frequencies (actually, the inverse of carrier frequencies). The technology's capability of providing very wide bandwidths using the existing optical fiber resources make it a good match to meet the exponential growth of Internet traffic.

In this paper we consider the problem of supporting IP networks over WDM networks. Figure 1 shows this two-layer network hierarchy, where the IP links are realized by optical end-to-end connections, referred to as *lightpaths*, at the WDM layer. At the WDM layer are fiber-pairs, each carrying multiple optical signals at different wavelengths. A fiber-pair forms a full-duplex physical link, with the fibers carrying signals in opposite directions. The fiber-pair carries full-duplex optical channels at each wavelength. In this paper, the fiber-pairs will be referred to as *fiber-links*. The fiber-links are terminated by transport equipment that multiplexes/demultiplexes the optical signals so that the fiber-link operates as a collection of parallel, full-duplex optical channels at different wavelengths. Today, lightpaths and WDM channels are 2.5 Gbps (OC-48) and 10 Gbps (OC-192), while 40 Gbps (OC-768) is the emerging technology.

The WDM optical channels from the transport equipment may be cross-connected at *optical cross-connects* (OXCs). The transport equipment has O-E-Os (optical-electrical-optical) equipment, which implies that any pair of optical channels incident to a common OXC may be cross-connected. (Note that without O-E-Os only channels with the same wavelength may be cross-connected.)

[1] The author was supported in part by Fujitsu Laboratories of America, and the National Science Foundation under grant NCR-9612846.

Full-duplex lightpaths are formed at the WDM layer, by cross-connecting WDM channels at OXCs. The lightpaths are terminated at interface ports of high speed IP routers, that are connected to OXCs as shown in Figure 1. To the IP routers, lightpaths are their physical links.

Since the IP network has high capacity, and thus supports many end-users, it is important that it is survivable under faults. At a minimum, the network should remain connected under any possible fault. As long as the IP network remains connected, it can reorganize its routing to provide some minimum level of packet transport service, and its control plane will remain operational.

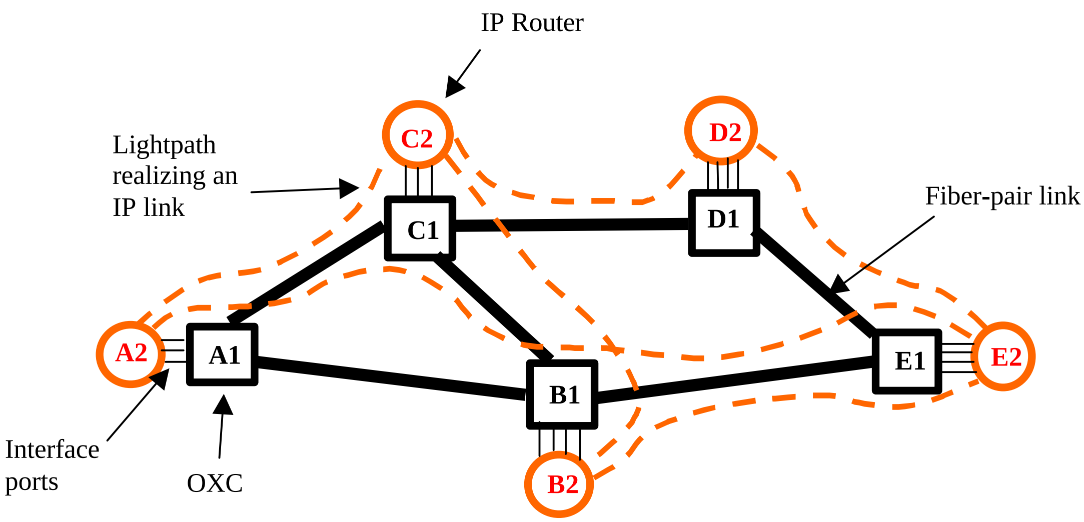

Figure 1. IP network over WDM network.

The type of faults that we consider is the fiber-link failure (e.g., fiber cut). Thus, it is necessary that the IP network topology is *2-connected*, which means that the topology remains connected under any single link failure. Though 2-connectivity is necessary it is not sufficient. If lightpaths (realizing IP links) pass through a common fiber-link then a fiber-link failure will result in multiple IP link failures. This may disconnect the IP network. For example, consider Figure 1. The IP network topology is 2-connected. However, its lightpath layout makes it vulnerable to single fiber-link cuts. In particular, if the fiber-link between OXCs *A1* and *C1* fails then the two IP-links between IP router *A2* and routers *B2* and *C2* fail, leaving IP router *A2* disconnected.

In this paper we consider the problem of laying out the IP network topology over the WDM network so that the former is connected under single fiber-link faults. This was first considered by Crochat and Le Boudec [1], and more recently in [2]. The time complexity of their algorithm for solving the problem can be prohibitive for large networks. We propose simpler algorithms that have lower time complexities.

In Section 2, we describe the network model and survivable layout problem. In Section 3, heuristic layout algorithms with low time complexities are given. Simulation results demonstrating the efficacy are presented in Section 4, and finally concluding remarks are given in Section 5.

2. Network Model and Survivable Layout Problem

There are two network topologies: the IP and WDM network topologies. Both have the same set of nodes V. Each node in V corresponds to an OXC and IP-router pair, where the router is connected to the OXC with interface ports. The fiber-links between the OXCs form the links E_{WDM} of the WDM network. $G_{WDM} = (V, E_{WDM})$ denotes the WDM network topology. The IP network topology is denoted by $G_{IP} = (V, E_{IP})$, where E_{IP} is the collection of IP links. An IP link is realized by a lightpath in the WDM network. A collection of lightpaths that realize the IP links is referred to as a *layout*. We will assume that the WDM network has a large number of wavelengths, so that the number of lightpaths that may traverse a fiber-link is unrestricted.

The *survivable layout problem* is to find a layout for G_{IP} such that any single fiber-link fault in G_{WDM} will leave the G_{IP} topology connected. This problem is a simplified version of the one considered by Crochat and Le Boudec [1, 2]. They consider the case where each layout has a *cost*, and the problem is to find a layout with minimum cost. The cost is zero if and only if the layout is survivable for any single fiber-link fault.

3. Heuristic Algorithms

Before describing our simple layout heuristics, we will briefly discuss the the algorithm of Crochat and Le Boudec [1, 2] for the survivable layout problem. Their algorithm is a *tabu search*, which is an iterative optimization algorithm that searches through possible solutions to a problem. In this case, the possible solutions are lightpath layouts. Each solution has a cost F, which in [1, 2] is a measure of how close it is to solving the survivable layout problem. The cost F is the linear combination of three components: $F1$, $F2$, and $F3$. $F1$ is the sum of fiber-link costs, where the cost of a link e is $\max(0, r(e) - C(e))^2$, $C(e)$ is the capacity of link e, and $r(e)$ is the number of lightpaths on e. (In our case, $C(e)$ is essentially infinite since we assume a very large number of wavelengths.) This cost penalizes fiber-links with load exceeding their link capacities. $F2$ is a measure of the number of IP links broken for a single link fault. $F3$ is the sum of fiber-link costs, where the cost of a link e is $\max(0, r(e) - t(e))^2$ and $t(e)$ is some threshold. The threshold $t(e)$ can be set to 1 to balance lightpath loads across fiber-links.

During one of its iterations, tabu search determines a new layout L_{new} from its current one $L_{current}$. It does this by first creating a collection of layouts that are "perturbations" of $L_{current}$. A perturbation is basically $L_{current}$ with exactly one of its lightpaths randomly rerouted. Then L_{new} is chosen among the perturbations with minimum. Note that each iteration requires $|E_{IP}|$ perturbations, and each perturbation must have its cost computed. Crochat and Le Boudec [1, 2] reported that each iteration takes $O(|E_{WDM}|\,|E_{IP}|\,|V|^2)$ time. The time complexity of an iteration is dominated by the time it takes to evaluate the cost of a candidate layout and in particular cost $F2$.

Thus, the algorithm does not scale well for large networks, especially if it requires many iterations to find a good solution. To address the scalability problem, we investigate simpler layout algorithms that have time complexities on the order of $O(|E_{IP}|\,|V|^2)$. The complexity is dominated by computing $|E_{IP}|$ lightpath routes, each requiring $O(|V|^2)$ shortest path computation.

The algorithms we consider find lightpaths for subsets of IP links at a time. Once a lightpath is computed for an IP link, the IP link is not considered again. The computation of the lightpaths for a subset of size k has time complexity $O(k\ |V|^2)$, which is dominated by k shortest path computations. Thus, the overall time complexity of the algorithms are $O(|E_{IP}|\ |V|^2)$.

Figure 2 shows the basic algorithm. The nodes are considered in some *order*. For each node v, IP links incident to the node, which we denote by $L(v)$, will have their lightpaths computed. Only those IP links that do not yet have lightpaths will have their lightpaths computed, i.e., an IP link will have its lightpath computed only once. Finding lightpaths for the IP links in $L(v)$ can be done by either routing them one-by-one using shortest path routing, or together using *minimum cost link-disjoint path routing* [3]. One-by-one computations are referred to as *single route* computations, while finding the routes together is referred to as a *multiple route* computation. Multiple route computation is implemented by the link-disjoint path routing of [3] and is described in Appendix A.

Survivable Layout Algorithm

1. Order the nodes. Denote the nodes by $v(1), v(2),..., v(n)$, where $n = |V|$ and $v(k)$ is the kth node.

2. For each $k = 1, 2, ..., n$,

- **2a.** Let $L(v(k))$ be the collection of IP links incident to $v(k)$.
- **2b.** Layout lightpaths for the links in $L(v(k))$ that do not have lightpaths. (Note that the layout can be done by single route computations or by a multiple route computation.)

Figure 2. Basic algorithm to compute lightpath layout of an IP network topology.

Step 2b in Figure 2 is the step for single or multiple lightpath route computations. In the case of single route computation, a lightpath is computed by first having *costs* assigned to fiber-links of G_{WDM}. Then a shortest path computation is applied to find the lightpath. In the case of multiple route computation, first costs are assigned to fiber-links of G_{WDM}. Then minimum cost, link disjoint paths are computed (as described in Appendix A).

By appropriately choosing the costs, one can control the type of paths that will be found. In the case of single route computations, the fiber-links have the following costs.

- *Unit link (UNIT) cost:* Each fiber-link has unit cost. This leads to shortest hop lightpaths, which minimize the amount of bandwidth use. However, this may also lead to an imbalance of lightpath loads on fiber-links. Such an imbalance should be avoided because a fault along a heavily loaded fiber-link will result in faults on many IP links.

- *Quadratic load (QUAD) cost:* Each fiber-link e has cost $(l_e + 1)^2$, where l_e is the number of lightpaths using e. This tends to distribute lightpaths evenly over fiber-links, reducing multiple IP link failures per fiber-link failure. This is similar to cost *F3* of Crochat and Le Boudec [1, 2].

- *Existing Lightpaths-Single Route Computation (EL-S) cost:* This cost depends on the lightpath being computed. The fiber-link costs are as follows. Suppose that *s* and *t* are the terminating nodes of the lightpath being computed. A fiber-link incident to *s* (resp., *t*) will have high cost K (= 2 $|E_{IP}|$ $|V|$) if it carries a lightpath that terminates at *s* (resp., *t*). The other links have unit cost. This tends to keep lightpaths that terminate at a common node from using the same links.

  As an example, consider Figure 1 with its existing collection of lightpaths. Suppose a new lightpath is to be computed between routers at *A2* and *E2*. Then the fiber-links (*A1*,*C1*), (*B1*,*E1*), and (*D1*,*E1*) will have high cost *K* because they are incident to OXC *A1* or *E1*, and they carry lightpaths terminating at those OXCs.

In the case of multiple route computations, we use the following fiber-link cost.

- *Existing Lightpaths-Multiple Route Computation (EL-M) cost:* This cost depends on node *v(k)* of Step 2 of the algorithm (see Figure 2). A fiber-link incident to *v(k)* that has a lightpath terminating at *v(k)* has high cost $K = 2\ |E_{IP}|\ |V|$ (same *K*, as in the *EL-S* cost). The other fiber-links have unit cost. Like the *EL-S* cost, this tends to keep lightpaths that terminate at a common node from using the same links.

  As an example, consider Figure 1 with its existing collection of lightpaths. Suppose new lightpaths were to be computed from the router *A2*. Since the existing lightpaths that terminate at *A2* use fiber-link (*A1,C1*), the link has cost *K* and the other links have unit cost.

We should note that with multiple route computation, the survivable layout algorithm may end up with some IP links without lightpaths. This can occur if there are many IP links, but not many fiber-link disjoint paths out of nodes. To take care of the IP links without lightpaths, the survivable lightpath layout algorithm is run again for these links but now for single route computations and *EL-S* cost.

We consider two types of node orderings.

- *Fixed ordering (FIXED):* The ordering is arbitrary and fixed. In this paper, we order according to node IDs.

- *Smallest degree first (DEGR):* This is an ordering according to node degrees, where the degree of a node is the number of IP links incident to it. Presumably nodes with smaller degrees are more likely to be disconnected from a fault, so they are considered first.

4. Performance Results

To compare the options of the survivable lightpath layout algorithm in the previous section, we ran simulations. For each option we ran simulations on three topologies of G_{WDM} shown in Figure 3. Topology *G1* was created by hand, while topologies *G2* and *G3* were randomly created. The random topologies were generated iteratively starting from initial topologies, that have no fiber-links. At each iteration, a fiber-link is added between a random pair of nodes, each pair being equally likely until the topology is 2-connected. However, we disallow multiple links between nodes.

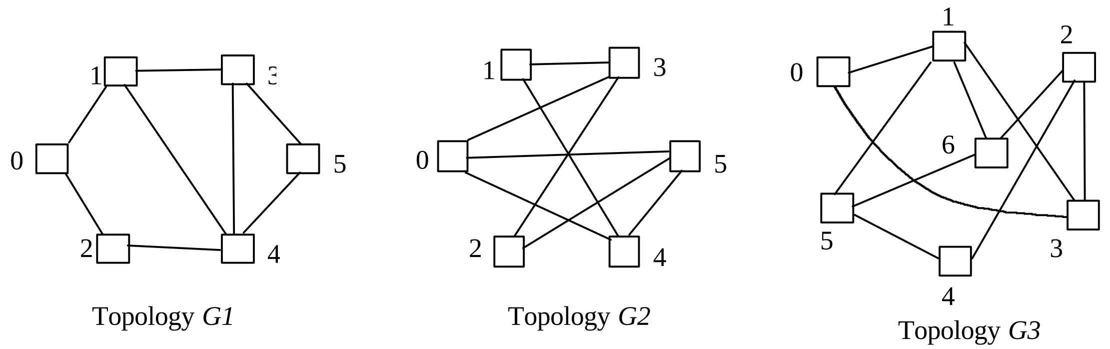

Figure 3. Three WDM network topologies.

For each WDM topology, 100 random IP topologies (G_{IP}) were created. These topologies were created iteratively in the same way topologies *G2* and *G3* were created. Each IP topology has a layout computed for the WDM topology using the survivable layout algorithm. Table 1 shows the results for single route computations when the node ordering is *FIXED* or *DEGR*, and the fiber-link costs are *UNIT*, *QUAD*, or *EL-S*. There are results for each WDM topology. Each table entry has the number of IP topologies that result in survivable layouts. For example, for WDM topology *G1*, the number of IP topologies that result in survivable layouts is 43 if the node order is *FIXED* and the link cost is *QUAD*.

	WDM Network Topology								
	G1			***G2***			***G3***		
	Link Cost			**Link Cost**			**Link Cost**		
Order	*UNIT*	*QUAD*	*EL-S*	*UNIT*	*QUAD*	*EL-S*	*UNIT*	*QUAD*	*EL-S*
FIXED	33	43	70	33	50	48	34	75	84
DEGR	33	60	84	33	69	69	34	90	92

Table 1. Performance results for single route computation.

The table shows that node ordering and link costs can significantly improve performance. In the case of node ordering, *DEGR* performs at least as well as *FIXED*, and in most cases the performance difference is significant. In the case of link costs, *EL-S* performs best. Thus, the combination of options *DEGR* and *EL-S* works best for single route computation.

Table 2 compares single and multiple route computation. The data for the single route computation for the options *DEGR and EL-S*, which we have shown work best. The table shows that multiple route computations performs about the same or better.

Routing	WDM Network Topology		
	G1	*G2*	*G3*
Single	84	69	92
Multiple	84	85	91

Table 2. Performance results for single versus multiple route computation.

5. Concluding Remarks

We propose simple algorithms to compute a survivable lightpath layout for an IP network topology. It is an iterative approach that computes lightpaths in sequence. Link cost and node ordering options were compared for single route computation. It was shown that the choice of node ordering and link costs can result in significant improvement. In the case of node ordering, it is better to first layout lightpaths of nodes with a small number of IP links since they are more vulnerable to faults. In the case of link costs, *UNIT* is the worse since it will not spread lightpaths over links. *QUAD* is better since the link costs tend to balance lightpath load over links. *EL-S* is a little better than *QUAD*, perhaps because its link costs tend to spread lightpath load and it is also specific to the lightpath being routed. We finally compared single and multiple route computation, and found that multiple route computation is about the same or better.

A. Link disjoint path algorithm

In this appendix we discuss how multiple route computations are done for IP links $L(v)$, which are the IP links incident to node v. The algorithm is basically a minimum cost, link disjoint path algorithm of [3] on an appropriately modified network topology. The algorithm will compute lightpaths for those IP links in $L(v)$ that do not have one. Let n denote the number of such IP links, and denote the links by $L'(v) = \{ (v, u_1), (v, u_2), ...,(v, u_n) \}$.

The algorithm of [3] computes minimum cost, link disjoint paths in a graph between a pair of nodes. If the number of disjoint paths is k then the algorithm's time complexity is dominated by the time complexity of k shortest path computations. Thus, the algorithm's time complexity is $O(k |V|^2)$. We apply the link disjoint path algorithm to a modified version of network topology G_{WDM}. In particular, a new (fictitious) node t is added. Also added are new links (u_1, t), (u_2, t),..., (u_n, t), i.e., these are new links connecting the nodes of $\{u_1, u_2,..., u_n\}$ to node t. The new links have zero cost. Then the minimum cost, disjoint path algorithm of [3] is applied to the modified topology G_{WDM}, where the paths terminate at nodes v and t. Note that the computed paths can be truncated at the nodes $\{u_1, u_2,..., u_n\}$ to form link disjoint lightpaths for a subset of the IP links in $L'(v)$. Those IP links in $L'(v)$ that do not get a lightpath are left for a future lightpath computation.

References

1. O. Crochat, J.-Y. Le Boudec, "Design protection for WDM networks," *IEEE J. Selected Areas on Communications*, vol. 16, no. 7, pp. 1158-1165, Sept. 1998.
2. O. Crochat, J.-Y. Le Boudec, O. Gerstel, "Protection interoperability for WDM optical networks," unpublished manuscript. http://lrcwww.epfl.ch/crochat.
3. J. Surballe, "Disjoint paths in a network," *Networks*, vol. 4, 1977, pp. 125-145.