Monitoring and Alarm Management in Transparent Optical Networks

Sava Stanic, Gokhan Sahin, Hongsik Choi, Suresh Subramaniam, and Hyeong-Ah Choi

Abstract—Rapid fault identification and localization in optical networks are crucial due to high data rates. These problems are more challenging than in traditional electronic networks because of optical transparency. In a transparent optical network which does not regenerate optical signals, a fault may propagate to various parts of the network from the origin, and multiple alarms can be generated for a single failure. In order to reduce the number of redundant alarms, simplify fault localization, and reduce the fault localization time, the number and location of fault monitors that are turned on should be optimized for a given network. In this paper, we formulate a problem on the optimal activation of network monitoring devices and propose a solution approach. First, we provide a brief summary of available physical-layer monitoring devices, and then present a scheme for optimal monitor activation. We show the NP-completeness of the problem and present a mixed-integer-linear-program (MILP) formulation of it. We also present a heuristic, whose performance is evaluated through comparisons with the solutions to the MILP, as well as a naive monitor activation approach.

Index Terms—Transparent optical networks, fault localization, fault detection, monitor activation, alarm processing.

I. INTRODUCTION

Today's optical networks are capable of carrying large amounts of data. Commercial systems have already been deployed with a per-fiber capacity of over 1.6 Terabits per second [16]. This is equivalent to 25 million simultaneous telephone conversations. In lab settings, traffic throughputs of the order of 10 Terabits per second over a single optical fiber have been achieved. As optical network technology advances and higher bandwidth is demanded,

the amount of data transmitted over a single optical fiber is expected to increase even further. Due to such high data rates, even a short service disruption may cause a large amount of data to be affected. Many different types of service disruptions occur frequently in practice. They include bending or cutting of fiber, equipment failure, and human error. Besides faults, optical networks are vulnerable to sophisticated attacks that may not be possible in electronic networks [8]. Therefore, it is of critical importance for such networks to have fast and effective methods for identifying and locating network failures. This is especially important for the physical layer, where any physical failure should be detected, located, and corrected before it is noticed, and acted upon, by upper layer protocols. In this paper, we use the terms fault and failure interchangeably, even though a failure may occur due to an attack on the physical infrastructure rather than faulty equipment.

Fault detection mechanisms in optical networks operate based on alarms that are generated by different types of network monitoring equipment in response to occurrence of unexpected events. Depending on the location and capabilities of these monitoring devices, the network fault manager may receive a large number of redundant alarms for some network failures, while it may not receive any alarms for other network failures. In order for a fault detection and localization mechanism to be fast and effective, it is important to reduce the number of redundant alarms received to the smallest possible number, while providing fault detection capability over the largest possible set of network failures. This will reduce alarm processing time as well as ambiguity in fault localization. Minimizing the number of alarms reported to the network manager without compromising on fault coverage is crucial to the rapid localization of faults as well as the stability of management systems against ever-increasing amounts of alarm data [13]. In this paper, we consider a model where there is a central network manager that can selectively configure the network monitors to (or not to) report alarms to the network manager. This scheme provides a fast and bandwidth-efficient approach, since monitors are dynamically configured as the set of

S. Stanic and S. Subramaniam are with the Dept. of ECE, and H.-A. Choi is with the Dept. of Computer Science, George Washington University, Washington, DC 20052. G. Sahin is currently with the ECE Dept. at Miami University, Ohio. H. Choi is currently with the Dept. of Computer Science, Virginia Commonwealth University, Richmond, VA. {stanic,suresh,hchoi}@gwu.edu,sahin@muohio.edu,hchoi@vcu.edu.

This work was supported in part by DARPA under Grant N66001-00-18949 (co-funded by NSA), and by NSF under grant CNS-0434956.

lightpaths is provisioned. It also avoids the use of test-lightpaths; therefore, eliminating the overhead associated with test-lightpath provisioning, configuration, and associated dedicated bandwidth used by such test-lightpaths. The heuristic algorithm presented in this paper allows for fast real-time re-configuration of available monitor equipment as the set of established lightpaths changes. Furthermore, our approach is based on inexpensive optical power monitors that are commonly available in today's optical network equipment [15]. In this paper we show that use of such inexpensive optical power monitors (when compared to the cost of other monitoring equipment such as per-wavelength monitors, or spectrum analyzers) can still provide sufficient failure detection and localization when used with our scheme. The choice of the monitors that are turned on (or "activated") for the provisioned set of lightpaths determines the alarm processing time, fault-localization capability, and therefore the network response time to failures. Thus, a judicious selection of the set of network monitors to be activated is essential for fast and effective fault detection and localization.

Related work on fault diagnosis and localization includes [2], [3], [4], [5], [6], [7], [11], [12], [9], [10]. The problem of identifying faults assuming that a fault propagates downstream on all lightpaths from the point of fault origin is considered in [2], [3], [4], [5], [7]. In [2], a central manager is assumed, and an algorithm for single fault identification is presented. The central manager periodically tests all source and destination powers using the routing table information. If some node's power is out of expected bounds, the possible source of the fault is identified. In [3], [4], [5], [7], fault identification through filtering alarms is discussed. Using a fault identification tree of depth equal to the number of alarming components, the set of potential fault sources is narrowed down. A survey of fault detection capability at each layer is presented in [6]. Approaches to the fault location problem are also classified. A problem on monitor placement was considered in [9], [10]. The authors show that a wavelength to be used for probing other nodes from a monitor node is highly likely to be available in dynamic traffic conditions. Based on this, it is claimed that fault detection and identification can be done successfully with high probability with a small number of monitor nodes. A heuristic algorithm for monitor placement that is based on clustering, as well as the idea of setting up additional test connections to collect diagnostic information is proposed. In [11], [12], a fault detection scheme for optical mesh networks based on decomposition of network topology into monitoring-cycles (test-lightpath loops) is presented. The authors present heuristics for construction of m-cycle

cover that minimizes cycle overlap for a given network topology. This approach, however, requires additional network resources to provision, configure, and maintain multiple dedicated-wavelength loops in the network. Also, the presented scheme can only detect failures of network links and nodes, and cannot provide per-wavelength monitoring of failures or per-wavelength signal quality, which can vary across the spectrum of a fiber and therefore result in different signal qualities for different wavelengths on the same fiber. Our scheme provides a general framework for effective optimization of fault-reaction time in transparent optical networks. The importance of this is obvious for large optical networks whose fault-management system must be scalable, dynamic and responsive in order to deal with frequent and large number of possible fault occurrences in a timely manner. In this paper, we present our scheme with optical-power-monitors due to their low cost and common availability in today's networks. However, it can be easily extended to provide more advanced network monitoring capabilities by use of more complex monitoring equipment including spectrum analyzers, wavelength monitors, optical signal-to-noise monitors, etc. Although some previous papers have considered static monitor placement [9], [10] (in order to minimize deployment cost of older optical networks), as far as we are aware, there has not been any work done in considering the optimization of physical-layer fault-reaction time and monitor-scalability in large transparent optical networks by dynamically minimizing activation of available optical monitoring equipment while maximizing the monitoring capability of the network. Our approach considers a fault propagation model for transparent optical network (similar to the one used in [2], [3], [4], [5]) in which faults propagate downstream from the source of a network fault. The rest of the paper is organized as follows. Section II provides a summary of the physical optical layer and commonly used optical monitoring devices. It specifies the set of alarms that are commonly provided by each monitoring device. Section III describes the advantages of activating an optimal set of monitors, and a problem is formulated. In Section IV, we present an optimal monitor activation scheme assuming that the monitoring equipment generates a single type of alarm. Section V shows a simple illustrative example of how the scheme works. Section VI addresses the complexity of the problem and shows its NP-completeness. Section VII presents a mixed-integer-linear-program (MILP) formulation of the problem, which can be used to obtain the optimal set of monitors to be activated. Section VIII presents a heuristic algorithm that provides near-optimal performance. The paper concludes with a summary of advantages achieved

by the use of optimal monitoring equipment activation, and considers possible extensions of this work.

II. PHYSICAL LAYER

In optical networks, the physical layer generally consists of several basic network components such as: optical fibers, amplifiers, multiplexers, switches, wavelength converters, protection switches, 3Rs, etc. These components can be categorized into passive and active components. Passive optical components, such as optical fiber, do not have monitoring equipment capable of detecting and reporting alarms and are therefore unable to report their operation status to the network manager. Active optical components, however, usually have some sort of built-in monitoring equipment and report alarms when some abnormal condition occurs. By considering the fault propagation model in a transparent optical network, it is feasible to locate failures of passive optical components by utilizing the monitoring capabilities of active components. The next two subsections provide a summary of some optical active components that can be used with our scheme. We first review the available optical layer monitoring devices, and then summarize the alarm reporting capabilities of currently available active optical network components.

A. Physical Layer Monitoring Capability

Different network layers provide different monitoring capabilities.

At the physical layer, there are essentially two basic monitor types that provide information about physical network status: optical power meter and optical spectrum analyzer. Both monitors generate alarms when the input power is outside specified threshold levels. An optical power meter generates an alarm when the aggregate power at its input is outside specified threshold levels. An optical spectrum analyzer is capable of monitoring the power level of each active wavelength on the fiber. It generates an alarm for a specific wavelength whose power is outside specified threshold levels. Table I summarizes the alarm characteristics of these two monitor types.

B. Optical Network Components

This section presents a summary of the physical layer monitoring capabilities of commonly used optical network components. Although monitoring capabilities differ slightly from manufacturer to manufacturer, Table II presents the most common alarming capabilities of currently available optical network components. This table is the result of a survey of optical components that are

TABLE I
ALARM CHARACTERISTICS FOR OPTICAL POWER METER AND OPTICAL SPECTRUM ANALYZER

Monitoring Device	Monitoring Capabilities	Possible Fault type	Alarm Type Reported	Alarm Description
Aggregate Optical Power Meter	Total optical power level on the fiber	One or more upstream nodes are transmitting at a power out of specified range	Aggregate optical power out of specified range	Aggregate input power alarm is active when aggregate optical input power is outside of specified range
Optical Spectrum Analyzer	Power level for each wavelength on the fiber	Upstream node is transmitting at a power out of specified range	Specific wavelength power out of specified range	Wavelength input power alarm is active for specific wavelength when the optical power is outside of specified range

offered on the market by several commercial vendors such as JDS Uniphase, Fiberdyne, Nortel, Alcatel, and Evertz.

TABLE II
ALARMING CAPABILITIES OF CURRENTLY AVAILABLE OPTICAL NETWORK COMPONENTS

Optical Device	Monitoring Capability	Alarm Generated
Transmitter	Per-channel output power level and temperature	Output power alarm Laser temperature alarm
Receiver	Per-channel input power level	Input Power Alarm
ADM	Per wavelength power monitoring	Signal power alarm for dropped, added, and through channels Unable to add channel
OXC	Input and output per-wavelength monitoring	Connection already established Hardware problem Unavailable resources Input wavelength power under/over threshold Output wavelength power under/over threshold
Protection SW	Input power level	Switch from input x to input y due to power level problem
3Rs (regenerator, reshaper, retimer)	Input and output channel power level and temperature	Input/output channel power level under/over threshold Unable to lock to frame Laser temperature out of range
Wavelength Converter	Input and output wavelength power monitor	Input/Output power out of range Wavelength conversion problem Wavelength blocking due to lack of resources on fiber
Optical Amplifier	Aggregate input and output power and pump laser temperature	Aggregate input/output power under/over threshold

Some of the listed components also provide temperature and upper layer monitoring capability, such as BER monitors; however, such information is irrelevant for our approach since it does not provide information

related to hard-failures. It is apparent from the table that the only two monitor types that provide information on hard-failures are aggregate power monitors and optical spectrum analyzers. Therefore our scheme focuses on optimizing the use of this equipment.

III. OPTIMAL MONITOR ACTIVATION AND PROBLEM FORMULATION

All networks require some form of a fault management system which, besides other functions, is capable of detecting and locating failures in a network. The function of such a system is to process alarms received from network components, and based on this information determine all possible locations of failure. Its effectiveness is measured by the speed with which it is able to process the received alarms, as well as by its ability to unambiguously determine location of fault source.

Depending on the type of monitoring equipment available, its location, network type, and network topology, different alarm types may be received by the fault manager. Some of the received alarms may be redundant, while in some cases, no alarms may be generated for a fault. By careful selection of the available network monitors to be activated (i.e., configured to report alarms), it is possible to greatly speed up fault localization by reducing the redundant alarms, while maintaining the same fault-detection coverage. Therefore, optimal selection of the location, type, and number of monitoring components that will be turned on to report alarms will improve the speed and effectiveness with which the fault-manager is able to locate a network failure. Although monitor placement optimization problems have been solved for many different applications they cannot be applied directly to optical networks. This is primarily due to the transparency characteristic of optical networks.

In transparent optical networks, faults propagate in the optical domain, while alarms are generated and processed in the electrical domain. A single fault can propagate throughout an optical network and generate many simultaneous alarms. This complicates the problem of accurately locating the source of the fault. Therefore, we need to take the fault propagation model for a given network into account when determining the optimal monitor activation, in order to reduce the number of redundant alarms that are sent to the fault manager.

In this paper, we develop a scheme for the optimal activation of monitors in an optical network. Our network and traffic model is an arbitrary mesh network with a static set of lightpaths, where the network manager can selectively turn the monitors on and off depending on the provisioned set of lightpaths. Note that we consider

a static set of lightpaths here to model an instance of a more general and practical scenario in which lightpath establishment requests arrive to and depart from the network. In other words, the monitor activation problem must be solved as and when the set of lightpaths in the network changes. A direct implication of this is that the activation algorithm must be fast enough to be used in real-time dynamic lightpath provisioning situations. The fault propagation model we assume is that a fault propagates in the downstream direction of all the lightpaths that pass through the source of the fault. This assumption is consistent with the one used in [2], [3], [4], [5], [7]. Thus, every downstream monitor on every lightpath passing through the fault location is assumed to report an alarm for this fault. We believe this model to be reasonable in a transparent optical network. Nevertheless, other appropriate models can easily be incorporated into the problem formulation. As two other examples, one may consider: (a) a second-order fault propagation model in which the lightpaths which intersect with the lightpaths that pass through the fault origin can also be assumed to propagate the fault, and (b) a fault propagation model in which there are some masking components [5] which do not propagate the fault further downstream. The problem we address here is informally stated below.

The Power Monitor Activation Problem: Given a mesh network and a set of lightpaths, and assuming the fault propagation model above, determine the number and location of the minimum number of active monitors required to achieve the same level of fault localization as the network with monitors activated at every network port.

The goal is to develop a time-efficient algorithm that minimizes the number of activated monitors while maximizing fault coverage, i.e., the number of fault scenarios which can be localized or whose origin can be inferred.

IV. A SOLUTION APPROACH

In our approach, we assume that only common aggregate power monitors are available in a network. We assume that any input or output port of a network node may be assigned an aggregate power monitor. Our algorithm determines the optimal activation of these monitors in network components on a per-port basis.

Our proposed power-monitor activation algorithm consists of two phases. The first phase is a preprocessing phase. During this phase, a fault-reporting *alarm matrix* is formed by considering fault propagation in a given network and determining the set of alarms reported for the failure of each network component. Each matrix row represents a binary vector of alarms that are reported

by the monitors for a specific failure scenario, a 1 denoting a reported alarm and a 0 denoting no alarm. The alarm vector is determined for each failure scenario by using information about the network topology and the established lightpaths. Initially, we assume that the power monitor on each used port on a network node is activated. Therefore, the resulting alarm matrix represents an “upper bound” on fault reporting for each network node. After such a matrix is formed, row-wise redundant information is removed by deleting all zero-vector rows and by grouping all identical rows into a single row. This is because, if a row consists of all zeros then no monitor detects the fault corresponding to that row. If two rows are identical, then the set of alarms reported for the two corresponding fault scenarios are identical, and therefore these faults cannot be unambiguously determined to have occurred. The time requirement for this preprocessing phase is negligible, assuming that the central network manager already has information on network topology and the established lightpaths.

The second phase of the algorithm is the monitor activation optimization phase. During this phase, the optimal set of power monitors to be activated is determined. The optimal activation of power monitors is determined by eliminating redundant monitors from the upper-bound fault-reporting alarm matrix that was constructed during the preprocessing phase. This corresponds to the removal of the maximum number of matrix columns (monitors) such that two row-wise constraints are met; namely, (a) a column may be removed only if after such removal none of the matrix rows become zero-vectors, and (b) all matrix rows remain distinct. By removing all possible columns under these two constraints, we can deactivate redundant monitors and therefore achieve the optimal set of monitors to be turned on in a given network. The next section provides a simple example that illustrates the application of this algorithm in a small network.

V. A SIMPLE OPTIMAL MONITOR ACTIVATION EXAMPLE

In this example we consider a simple network with eight optical nodes as shown in Figure 1, and show how our algorithm can be used to determine the optimal placement of aggregate-power monitors. For this example, we assume only single node-fault scenarios for simplicity, whereas in the rest of the paper we consider all single port-failure, node-failure, and bidirectional link-failure scenarios.

It is assumed that the following three lightpaths $LP1$, $LP2$, and $LP3$ have been established in the network shown above:

Fig. 1. An example network.

$LP1$: Node1 $\rightarrow$ Node2 $\rightarrow$ Node3,
 $LP2$: Node1 $\rightarrow$ Node4 $\rightarrow$ Node2 $\rightarrow$ Node5 $\rightarrow$ Node3, and
 $LP3$: Node6 $\rightarrow$ Node4 $\rightarrow$ Node2 $\rightarrow$ Node5 $\rightarrow$ Node8.

Using this information about established lightpaths, we can simplify the network graph as shown in Figure 2.

Fig. 2. Three lightpaths in a network.

We start solving the optimal monitor activation problem by determining the upper-bound activation scenario. This results in the activation of aggregate power monitors on every currently used input port of the network nodes. We label these initial set of activated monitors as $M_1, \dots, M_8$ as shown in Figure 2.

Using the information about established lightpaths, network topology, and the active monitor locations, we create the alarm matrix as shown in Table III.

To remove redundant information from the alarm matrix, we remove all zero-vector rows and group all identical rows into a single row. This results in the simplified alarm matrix shown in Table IV. In this case, faults in nodes 3, 7, and 8 cannot be identified because no alarms are reported. All other node faults can be unambiguously located because no two sets of alarms are identical.

Finally, redundant monitors (matrix columns) can be removed (deactivated) under the two constraints:

TABLE III
ALARM MATRIX

Faults	M1	M2	M3	M4	M5	M6	M7	M8
FN1	1	1	1	1	1	1	0	0
FN2	0	1	0	0	1	1	0	1
FN3	0	0	0	0	0	0	0	0
FN4	0	0	0	1	1	1	0	1
FN5	0	0	0	0	0	1	0	1
FN6	0	0	0	1	1	0	1	1
FN7	0	0	0	0	0	0	0	0
FN8	0	0	0	0	0	0	0	0

TABLE IV
SIMPLIFIED ALARM MATRIX

Faults	M1	M2	M3	M4	M5	M6	M7	M8
FN1	1	1	1	1	1	1	0	0
FN2	0	1	0	0	1	1	0	1
FN4	0	0	0	1	1	1	0	1
FN5	0	0	0	0	0	1	0	1
FN6	0	0	0	1	1	0	1	1

1. None of the matrix rows can be zero vectors.
2. All matrix rows must remain distinct.

Using these constraints, we can reduce the above matrix to the optimal alarm matrix as shown in Table V.

TABLE V
OPTIMAL ALARM MATRIX

Faults	M2	M4	M6
FN1	1	1	1
FN2	1	0	1
FN4	0	1	1
FN5	0	0	1
FN6	0	1	0

We have deactivated redundant aggregate-power monitors M_1 , M_3 , M_5 , M_7 , and M_8 in the network, and therefore reduced the number of alarms that need to be processed by the network fault manager. We have achieved over 60% reduction in the number of monitors that need to be turned on in this example. The resulting optimal set of activated monitors are shown in bold in Figure 3 (the deactivated monitors are not shown).

VI. COMPLEXITY ANALYSIS

Once an alarm matrix is introduced (e.g., Table IV), finding an optimal alarm matrix (e.g., Table V) is equivalent to finding a set of redundant monitors of maximum cardinality whose removal (i.e., deactivation) satisfies the following properties:

Fig. 3. Optimal monitor activation for the example.

1. the maximum fault detection coverage is still achieved, and
2. each fault is reported by a unique set of alarms.

We call this problem the Redundant Monitor Deactivation Problem and formally state it below.

Redundant Monitor Deactivation Problem (RMDP): Given an integer m , a set of k monitors $M = \{M_1, \dots, M_k\}$ and a set of n faults $F = \{f_1, \dots, f_n\}$ such that $A_M(f_i) \neq A_M(f_j)$ for any $1 \leq i \neq j \leq n$, where $A_M(f_i)$ denotes the set of monitors in M that raise alarms when fault f_i occurs, the problem is to find a subset $D \subseteq M$ such that

- (A1) $A_{M-D}(f_i) \neq A_{M-D}(f_j)$ for any $1 \leq i \neq j \leq n$
- (A2) $A_{M-D}(f_i) \neq \emptyset$ for any $i, 1 \leq i \leq n$ and
- (A3) $|D|$ is maximum.

In what follows, we show the NP-completeness of this problem by sketching a polynomial time transformation to the decision version of the RMDP (i.e., decide whether there exists a subset $D \subseteq M$ such that $|D| = m$ for a given integer m) from the Exact Three Cover (X3C) problem which is known to be NP-complete [1]. The X3C problem is given below.

X3C Problem:

Given a set of n elements $X = \{x_1, \dots, x_n\}$ (where n is a multiple of three). and a set of m elements $C = \{c_1, \dots, c_m\}$ where $c_j \subset X$ and $|c_j| = 3$ for all $j, 1 \leq j \leq m$, the objective is to find a subset $C_0 \subset C$ such that $\cup_{c_j \in C_0} c_j = X$ and $|C_0| = n/3$.

A. Polynomial Time Transformation

Let X and C be an arbitrary input to the X3C problem. Define $C(x_i) = \{c_j \mid x_i \in c_j\}$. From X and C , we construct a monitor set M , a fault set F , and a set of alarms $A_M(f_i)$ corresponding to each $f_i \in F$ as follows.

The monitor set is constructed as $M = \{c_1, \dots, c_m\} \cup \{d_1, \dots, d_n\}$. The fault set is constructed as $F =$

$\{x_1, \dots, x_n\} \cup \{g_1, \dots, g_n\}$. The alarm set is constructed as $A_M(x_i) = C(x_i) \cup \{d_i\}$ for $1 \leq i \leq n$ and $A_M(g_i) = \{d_i\}$ for $1 \leq i \leq n$.

Consider the following example as an input to the X3C problem: $X = \{x_1, \dots, x_6\}$ and $C = \{c_1, c_2, c_3\}$, where $c_1 = \{x_1, x_2, x_3\}$, $c_2 = \{x_2, x_3, x_5\}$, $c_3 = \{x_4, x_5, x_6\}$. The input to the RMDP is then constructed as: $M = \{c_1, \dots, c_3, d_1, \dots, d_6\}$, $F = \{x_1, \dots, x_6, g_1, \dots, g_6\}$ and $A_M(x_1) = \{c_1, d_1\}$, $A_M(x_2) = \{c_1, c_2, d_2\}$, $A_M(x_3) = \{c_1, c_2, d_3\}$, $A_M(x_4) = \{c_3, d_4\}$, $A_M(x_5) = \{c_2, c_3, d_5\}$, $A_M(x_6) = \{c_3, d_6\}$, $A_M(g_1) = \{d_1\}$, $A_M(g_2) = \{d_2\}$, $A_M(g_3) = \{d_3\}$, $A_M(g_4) = \{d_4\}$, $A_M(g_5) = \{d_5\}$, and $A_M(g_6) = \{d_6\}$.

B. Correctness of Transformation

Lemma 1: Let X and C be an input to the X3C problem, and let $I_{X,C}$ instance of the input (i.e., M , F , and $A_M(F)$) to the RMDP constructed by following the above procedure. Then, there exists a solution to the X3C problem if and only if there exists a solution $D \subseteq M$ to $I_{X,C}$ such that $|D| = m - n/3$ satisfying the constraints (A1) and (A2) defined in the RMDP.

Proof: Suppose there exists a solution C_0 to the X3C problem. (Note that if C_0 is a exact cover for X , then $|C_0| = n/3$.) We then define $D = C - C_0$, i.e., monitors corresponding to elements in $C - C_0$ will be deactivated. It is then easy to see that each fault x_i ($1 \leq i \leq n$) will have a unique alarm set $A_{M-D}(x_i) = \{d_i, c_j\}$, where $x_i \in c_j$ for some $c_j \in C_0$. Each fault g_i ($1 \leq i \leq n$) will also have a unique alarm set $\{d_i\}$ of size one. (In the example above, $D = \{c_2\}$ and $C_0 = \{c_1, c_3\}$.)

Now, assume that there exists a solution D for $I_{X,C}$ where $|D| = m - n/3$. We note that $d_i \notin D$ as otherwise fault g_i cannot be detected for $1 \leq i \leq n$. We also note that for each x_i , at least one monitor in $C - D$ should remain activated to monitor x_i since otherwise x_i and g_i both will only be monitored by the unique alarm d_i . This implies that $C - D$ must be a solution to the X3C problem. $\square$

We thus have the following result.

Theorem 1: The RMDP is NP-complete.

VII. AN MILP FORMULATION FOR MINIMIZING THE NUMBER OF ACTIVATED MONITORS

It was shown in Section VI that the problem of optimally activating monitors is NP-complete. In this section, we present a mixed-integer-linear-program (MILP) formulation which minimizes the number of monitors that are activated, while maintaining the fault-identification capability. Given the requirements on the speed of the

monitor activation algorithm in dynamic traffic settings and the computational intensity of solving MILPs of even moderate sizes, we do not suggest solving the MILP as a monitor activation algorithm. Our goal in introducing the MILP here is to make performance comparisons with a simple and fast heuristic that *can* be used. The input to the problem is an alarm matrix which indicates the set of monitors that generate alarms in the event of each failure. Let $\{a_{f,m}\}$ be the elements of the alarm matrix A , where $a_{f,m}$ takes the value 1 if monitor m generates an alarm when failure f occurs, and 0 otherwise. For simplicity, we assume that A is the reduced alarm matrix obtained at the end of the preprocessing phase of our solution approach described in Section IV, and thus all the rows are non-zero, and identical rows are eliminated. The decision variables are $\{u_m\}$, where u_m takes the value 1 if monitor m is activated, and the value 0 otherwise. The MILP attempts to minimize the number of monitors that are activated such that the following two constraints are satisfied by the matrix consisting of only the columns corresponding to the activated monitors: **i)** Rows corresponding to different failures should be non-identical, and **ii)** Each row should be non-zero. The optimization problem can be described by the following MILP.

The objective of the MILP is to minimize the number of activated monitors:

$$\text{Minimize } Z = \sum_m u_m$$

subject to:

Constraint Set 1. Uniqueness constraints: Each row, consisting of the elements corresponding to the set of activated monitors, should be different from all other rows. Thus, for every pair of rows corresponding to failures f_i and f_j , where $i \neq j$ we have,

$$\sum_{m=1}^M (a_{f_i,m} - a_{f_j,m})^2 u_m > 0. \quad (1)$$

Note that $a_{f,m}$ are parameters and not decision variables, and the problem is not non-linear as (1) might appear to suggest.

Constraint Set 2. Non-zero constraints: For each failure scenario, there should be at least one activated monitor that generates an alarm in the event of that failure to ensure that the failure can be detected. Thus, each row should contain at least one non-zero element in the columns corresponding to the activated monitors. For each f_i ,

$$\sum_{m=1}^M a_{f_i,m} u_m > 0. \quad (2)$$

Constraint Set 3. Binary Decision Variable Constraints. All decision variables $\{u_m\}$ should be binary.

$$u_m \in \{0, 1\} \quad \forall m.$$

VIII. AN EFFICIENT HEURISTIC ALGORITHM

As mentioned earlier, the optimal monitor activation problem is NP-complete. In practice, optimal solutions to the problem are useful only if the established lightpaths remain constant during its computation. Depending on lightpath traffic dynamics as well as available processing power, it might not be always possible to compute an optimal solution before the lightpath configuration changes. Accordingly, our heuristic algorithm described below will be more useful as it provides near-optimal solutions with much less (polynomial) time complexity. We first provide a short summary of the heuristic, after which we compare the simulation results obtained using the heuristic with the optimal solutions obtained by solving the MILP formulation of Section VII using CPLEX.

A. A Greedy Heuristic Algorithm

The heuristic algorithm we use for this problem is based on a greedy method. Redundant monitors are removed (turned-off) based on a simple rule of first deleting a monitor whose column in the alarm matrix contains the smallest number of 1's. The idea behind this approach is that such a monitor is least effective in terms of partitioning the fault set and may be a good candidate for deactivation. Every time such a column deletion is to be performed, two constraints on the alarm matrix are checked. If the deletion of a monitor column in the alarm matrix causes any of the matrix rows to become zero-vectors or if any two rows of the matrix become identical, the column must not be deleted and it is selected as a required monitor. Otherwise, if constraints hold after monitor column deletion, such a monitor is assumed not to be required and is deactivated. To demonstrate an application of this algorithm (called Greedy Min), we use it to solve the monitor activation example in Section V. Processing the simplified alarm matrix in Table IV using this algorithm results in the near-optimal solution of Table VI. Comparing this with the optimal solution in Table V, we see that the greedy heuristic requires the activation of one more monitor for this example. Pseudocode for the heuristic is presented below.

TABLE VI
THE NEAR-OPTIMAL ALARM MATRIX

Faults	M4	M5	M6	M8
FN1	1	1	1	0
FN2	0	1	1	1
FN4	1	1	1	1
FN5	0	0	1	1
FN6	1	1	0	1

Algorithm 1: GREEDY MIN

Input : $M_{m,n}$
 // $M_{i,j} = 1$ if fault i causes alarm j //
Output : $M_{m,n'}$ **where** $\lceil \log n \rceil \leq n' \leq n$
 1. **while** There is unmarked column **Do**
 2. find column with smallest sum ;
 3. if column deletion is allowed
 4. delete the column ;
 5. else mark this column;
 7. **endwhile**

We also present results for a similar algorithm called Greedy Max in which a column with highest sum is picked first as a candidate for deletion. It can be easily verified that the time complexity of the heuristics (given the alarm matrix) is $O(nk)$ where n is total number of faults and k is the total number of monitors in the network.

B. Experimental Results

We have implemented the heuristic algorithms in C, and solved the MILP formulation of the problem using CPLEX, a commercially available optimization software. A comparison between the number of required active monitors obtained by the heuristic algorithm and the optimal solution given by the MILP was performed. The network topology we assumed for the simulation was the NSFnet topology shown in Figure 4. This network has 42 directed links and 182 ordered pairs of nodes. The simulation results are given in Figures 5, 6, and 7. In our experiments, we considered all single node-fault, port-fault, and (bidirectional) link-fault scenarios. Note, however, that the problem formulation and the solution approach themselves do not depend on any specific fault model. The fault model affects only the alarm matrix that is generated.

In Figures 5 and 6, we show the number of monitors to be activated obtained by the heuristic algorithms (Greedy

Max and Greedy Min) as well as the optimal solution (MILP) for various lightpath set sizes. For each set size, we generated 100 random sets picked using a uniform distribution for the sources and destinations, and both average and worst-case results over these 100 sets are presented. Also shown are the maximum number of possible active monitors (Naive), i.e., the number of ports that are in use, for the various lightpath set sizes. In addition, we also report the total number of possible faults arising from nodes/links/ports that are actually used, and the number of the faults that can be detected by the set of monitors activated using the algorithms (Figure 7). It can be seen from results that the heuristics provide significant reductions in the number of active monitors with respect to the naive approach. For example, when there are 110 lightpaths, the average number of active monitors can be reduced from 83.5 to 16.4 through our Greedy Min heuristic (which is 92% of optimal reduction to 10.7 monitors) (Figure 5). Note that this reduction comes without any compromise in the fault-detection capability, i.e., the solutions obtained by the naive approach, the greedy algorithm, and the MILP can all identify the same percentage of failure cases. We also note that the percentages of failure cases that can be uniquely identified are quite high for moderate to large number of lightpaths. In particular, when the number of lightpaths is larger than 30, more than 90% of the failure cases can be identified (Figure 7). It is interesting to note that the fault-detection capability increases rapidly as the number of lightpaths increases. The increase is in part due to the fact that additional lightpaths may provide additional information since they may lead to alarms in more monitors. Also, as the number of lightpaths increases over 20, the number of undetectable faults remains below 12 on average. These undetectable faults represent rows of the original alarm matrix that are either zero-rows or identical rows. In other words, these undetectable faults are specific to the network topology, fault propagation model, and equipment placement; and are undetectable without addition of extra monitors to supplement those already provided by the network devices. Note also that while it is possible to obtain further reduction in the number of active monitors by using the MILP instead of the greedy heuristic, the performance difference between the two approaches is relatively small compared with the improvements through the greedy algorithm over the naive approach. In other words, most of the potential reduction in the number of active monitors can be achieved through the greedy algorithm. This shows the efficiency of the greedy algorithm, especially given its simplicity compared with the MILP.

During our simulations we also measured the time re-

quired for each of the three monitor activation approaches. Here we briefly summarize the obtained results, and explain how they affect the overall network's responsiveness to failures. Assuming that a central network manager already has information of provisioned lightpaths, the naive approach does not take any time to execute; however, as mentioned earlier, this approach is inefficient in large transparent optical networks and will create a storm of redundant alarms, slowing down the network's overall response to failures and complicating fault localization. Solving the optimal monitor activation problem through MILP took up to several hours for a single set of lightpaths on a standard PC. Obviously, this is unacceptable in networks with dynamically provisioned lightpaths, and is useful only in scenarios where provisioning information for new lightpaths is available at least several hours before lightpaths are needed. Finally, the greedy algorithm provided close to optimal results within a fraction of a second (on the average less than half a second). This allows for monitor activation optimization to be performed dynamically as lightpaths are provisioned, making it useful for actual real-time deployment in today's and future optical networks.

Fig. 4. The NSFnet topology.

Fig. 5. Average number of activated monitors.

Fig. 6. Worst-case number of activated monitors.

Fig. 7. Number of faults, and fault-localization coverage.

IX. CONCLUSIONS AND FUTURE WORK

In this paper, we presented the problem of optimal monitor activation in optical networks. We described the types of optical monitoring equipment currently available, and the types of alarms generated by currently deployed optical networking components. A simple example illustrated potential reduction in the number of alarms by optimally activating monitors, and an approach to choosing the set of monitors to be turned on was presented. The problem of minimizing the number of activated monitors was proven to be NP-hard, and a heuristic algorithm that gave good performance in our experiments was also presented.

This work may be extended in the following ways. In this paper, we assumed a static lightpath scenario as a snapshot of a network with dynamic lightpath traffic, and the monitor activation algorithm was based on the alarm matrix derived for this set of lightpaths. An interesting extension of this work is the efficient updation of the alarm matrix itself when the lightpath configuration changes. One may also consider the problem of minimal reconfiguration of monitors (i.e., changing the status of a

minimum number of monitors from the previous lightpath configuration). Furthermore, one may consider various other kinds of monitoring equipment and fault propagation models. Designing more efficient algorithms for monitor activation with provable worst-case guarantees is another possible topic for future study.

REFERENCES

- [1] M. R. Garey and D. J. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, New York, 1979).
- [2] I. Katzela, G. Ellinas, W. S. Yoon, and T. E. Stern, "Fault diagnosis in optical networks", *Journal of High Speed Networks*, vol. 10, no. 4, 2001, pp. 269-91.
- [3] C. Mas, O. Crochat, and J.-Y. Le Boudec, "Fault Localization in an Optical Network," *All-Optical Networking: Architecture, Control, and Management Issues (VV09), SPIE'98 Voice, Video, and Data Communications*, November, 1998.
- [4] C. Mas and J.-Y. Le Boudec, "An alarm filtering algorithm for optical communication networks", *Proc. Management of Multimedia Networks and Services, IFIP/IEEE TC6/WG6.4/WG6.6 International Conference*, 1998, pp. 205-18.
- [5] C. Mas and P. Thiran, "An Efficient Algorithm for Locating Soft and Hard Failures in WDM Networks", *IEEE J. Select Areas Commun.*, vol. 18, no. 10, pp. 1900-1911, Oct. 2000.
- [6] C. Mas and P. Thiran, "A review on fault location methods and their application to optical networks", *Optical Networks Magazine*, pp. 73-87, Jul./Aug. 2001.
- [7] C. Mas, I. Tomkos, O. Tonguz, "Failure Location Algorithm for Transparent Optical Networks", *IEEE J. Select Areas Commun.*, vol. 23, no.8, pp. 1508-1519, Aug. 2005.
- [8] M. Medard, D. Marquis, R. A. Barry, and S. G. Finn, "Security issues in all-optical networks," *IEEE Network*, vol. 11, no. 3, pp. 42-48, May-June 1997.
- [9] W. Tao and A. K. Somani, "Attack monitoring and monitor placement in all-optical networks," *Proc. Gigabit Networking Workshop*, Anchorage, Alaska, in conjunction with INFOCOM '01.
- [10] W. Tao and A. K. Somani, "Attack monitoring and Localization in all-optical networks," *Proc. Opticomm 2002*, Boston, 2002, pp. 235-248.
- [11] H. Zeng, C. Huang, A. Vukovic, and M. Savoie, "Fault Detection and Path Performance Monitoring in Meshed All-Optical Networks", *IEEE Globecom'04*, Dallas, Nov. 2004.
- [12] H. Zeng, C. Huang, and A. Vukovic, "A Novel Fault Detection and Localization Scheme for Meshed All-Optical Networks Based on Monitoring-Cycles", *Photonic Network Communications*, Volume 11, No.3, pp. 277-286, May 2006.
- [13] R. D. Gardner and D. A. Harle, "Pattern discovery and specification techniques for alarm correlation," *Network Operations and Management Symposium*, 1998, pp. 713-722.
- [14] E. Horowitz, S. Sahni, and S. Rajasekaran, "Computer algorithms," Computer Science Press, New York, 1997.
- [15] H. Agawa, M. Nagata, M. Araragi, "A high performance optical power meter", *Instrumentation and Measurement Technology Conference*, pp. 622-624., May 1991.
- [16] L. Labrunie, F. Boubal, E. Brandon, L. Buet, N. Darbois, D. Dufournet, V. Havard, P. Leroux, M. Mesic, L. Piriou, A. Tran, and J. P. Blondel, 1.6 Terabits/s (160 x 10.66 Gbit/s) unrepeated transmission over 321 km using second-order pumping distributed Raman amplification, in *Proc. Optical Amplifiers and Their Applications Conf.*, Stresa, Italy, 2001, Postdeadline Paper PD3-1.