Abstract—In optical networks, shared-backup path protection (SBPP) can optimize network resources utilization in scenarios where connection resilience requirements are not achieved without protection, but are exceeded with dedicated-path protection (DPP). However, savings in capacity allowed by backup capacity sharing are achieved at the cost of connection availability degradation. This paper presents a strategy for provisioning connections with guaranteed availability in a dynamic traffic scenario that attempts to minimize the allocation of spare capacity. Connections can be unprotected or protected by SBPP – which in some cases degenerate to DPP – depending on the connection availability requirements. This provisioning strategy employs the previously published matrix-based approach for connection unavailability estimation in SBPP protected networks which offers accurate results for networks of national size. We investigate the performance, in terms of blocking and resource sharing, of three availability classes (availability of 0.999, 0.9999 and 0.99999) that coexist in two representative network topologies without wavelength conversion. The results indicate that SBPP is a viable option for networks of national dimensions, but the backup sharing of high availability connections is strongly limited in networks of continental dimensions.

I. INTRODUCTION

The evolution of photonic technologies in the past two decades enabled an exceptional growth in the aggregate bitrate supported by an optical fiber. A key technology for this development has been Wavelength Division Multiplexing (WDM), through which hundreds wavelength channels carrying data at rates in the order of gigabits per second can be multiplexed into a single fiber. However, the increase in the transported data volume also deepens concerns about the network integrity, since failures can cause severe losses in revenue and image. In this context, protection and restoration schemes are essential to guarantee the network operation under abnormal conditions, such as a fiber break caused by digging work. The survivability requirements for connectivity services are usually stated in the Service Level Agreement (SLA) contract along with revenue and penalty in case of non-fulfillment. Service differentiation in survivable WDM networks can be provided in terms of availability, reliability, protection bandwidth, recovery time and recovery bandwidth. Reference [1] explains some of the concepts used to implement service differentiation, such as DiR (Differentiated Reliability) [2] and QoP (Quality of Protection) [3].

Recently, the most adopted survivability metric has been availability [4], which can be defined as the probability that the system is found in the operating state sometime in the future. From the operator’s point of view, the connection availability should be high enough to fulfill the SLA while allocating minimal spare capacity. In optical networks, shared-backup path protection (SBPP) can optimize network resources utilization in scenarios where connection availability requirements are not achieved without protection, but are exceeded with dedicated-path protection (DPP). However, savings in capacity allowed by backup capacity sharing are achieved at the cost of connection availability degradation. Few publications have approached the issue of provisioning connections with guaranteed availability in shared-backup path protection (SBPP). Reference [5] proposes a strategy for dynamic traffic in a single fiber link failure scenario. Reference [6] limits the sharing of backup capacity introducing a link-based threshold which results in a conservative upper bound on the unavailability of connections in a dynamic traffic scenario. Reference [7] proposes a double-fault shared path protection scheme with constrained connection downtime for static traffic which also uses conservative unavailability estimates. Reference [8] analyzes the availability of connections protected by DPP and SBPP.

This paper investigates the performance of connection classes with guaranteed availability in terms of blocking probability and sharing of backup resources in SBPP protected networks [9] without wavelength conversion. Connections with guaranteed availability are provisioned according to a novel strategy that attempts to minimize the allocation of spare capacity in dynamic traffic scenarios. Connections can be unprotected
or protected by SBPP – which in some cases degenerate to DPP – depending on the connection availability requirements. The provisioning strategy is the first to employ the matrix-based approach for connection unavailability estimation [10] that yields accurate results for networks of national size, which are the networks in which SBPP is most efficiently employed. The paper also discusses the applicability of SBPP in networks of continental size, in which connections can be thousands of kilometers long and a single backup path may be insufficient to guarantee the commonly required availability levels to carrier class connections. In this case other alternatives must be employed, such as link restoration [11], path protection using several backup paths [12], or the use SBPP or DPP in protection domains within the network. The remainder of this paper is organized as follows. Section II presents the general assumptions considered in the paper. Section III introduces the strategy for the dynamic provisioning of connections with guaranteed availability in shared-backup path protected networks. Section IV shows the performance evaluation setup. Section V presents the simulation results. Section VI concludes the paper.

II. THE MATRIX-BASED APPROACH TO CONNECTION UNAVAILABILITY ESTIMATION IN SBPP

Shared-backup path protection (SBPP) is a strong candidate to be the dominant protection scheme in future transport mesh networks [9]. In its original conception, a working path and a backup path are assigned to each connection, and capacity can be shared among backup paths if their corresponding working paths do not traverse common fiber links. However, savings in capacity allowed by backup capacity sharing are achieved at the cost of connection resilience degradation. When a customer contracts a connectivity service from a network service provider, the connection resilience performance is specified in the Service Level Agreement (SLA). The strategy proposed in this paper attempts to provision connections with guaranteed availability while minimizing the allocated spare capacity. The connection availability estimation uses the matrix-based approach proposed in [10]. The approach is derived from a continuous time Markov model which assumes that not more than two simultaneous fiber link failures occur in the network. This approximation, well-accepted in the literature, is valid for most wide-area networks, since higher order failures are extremely unlikely. Its derivations are based on the following assumptions:

1) A two-state “working-failed” model describes the status of all fiber links.
2) The network nodes have availability equal to one.
3) All fiber links fail independently.
4) The repair time and the time to failure of a fiber link are memoryless, exponentially distributed random processes with constant means MTTR (mean time to repair) and MTTF (mean time to failure).
5) In SBPP, the first path to fail holds the protection resources until it is completely repaired.
6) At most two fiber links can be simultaneously in the failed state within the network.

Assumptions 1-4 are typical for transport networks. Assumption 5 is a plausible implementation of SBPP. Notice that by assumption 2 network nodes are treated as atomic entities that do not fail. Indeed, nodes are very reliable compared to links, and their availability can be always controlled by internal redundancy. Assumption 6 makes the approach applicable to most (even large) networks. If connections are protected against single link failures, double link failures are the dominant source of connection outages [9]. For an Italian Network [13], upon the occurrence of multiple link failures, there is approximately 99% probability that only two links are in the “failed” state for a MTTR = 20 h and 200 FIT/km as value for 1/MTTF (1 FIT = 1 failure in 10^9 h). In continental networks this probability is lower, e.g., 95% for the NSFNet Network [14].

The matrix-based approach offers for a specific network two sorts of quantities which can be used to connection unavailability estimation. The first one are the $\pi_i$, values, which represent the probability that the network has a single failure in fiber link $i$. The second one are the $\pi_{ij}$ values, which represent the probability that the network has a double failure in fiber links $i$ and $j$, and fiber link $i$ failed before fiber link $j$. In the original conception of SBPP, a working path $w$ and a backup path $b$ are assigned to each connection $c$, and capacity can be shared among backup paths if their corresponding working paths do not traverse common fiber links (Shared Risk Link Group constraint, or SRLG constraint). This group of working paths of which backup paths share some capacity with $b$ is called here $SG$. In Figure 1 the thick lines correspond to working paths and the thin lines to backup paths. The $SG$ group of connection $C$ consists of connections $C2$ and $C3$, but connection $C1$ may not be part of it due to the SRLG constraint (notice that by definition connection $C$ is not part of its $SG$ group).

Let $L$ be the set of all fiber links and $f_w$, $f_b$ and $f_{SG}$ the set of fiber links traversed by $w$, $b$ and $SG$. The availability of connection $c$ can be computed as follows:

$$A_{cp} = 1 - \sum_{m: f_w \neq f_b} \pi_{ml} \sum_{m: f_w \neq (f_{SG} \cup f_b)} \pi_{lm}.$$  \hspace{1cm} (1)

The first summation in Equation 1 accounts for all double fiber link failures which affect first a fiber link traversed by the working path and subsequently a fiber link traversed by the backup path. The second summation accounts for all double

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure1.png}
\caption{Shared-backup path protection.}
\end{figure}
fiber link failures which affect first a fiber link traversed by
the backup path or any path in SG and subsequently the
working path. Note that if SG is empty the method falls back
to dedicated-path protection.
The availability of an unprotected connection \( c \) (in this case
there is no backup path) can be calculated by the following
equation:

\[
A_{cu} = 1 - \sum_{m \in f_w} \pi_m - \sum_{m \in f_w, l \in f_w} (\pi_{lm} + \pi_{ml}) - \sum_{m \in f_w, l \in f_w, m \neq l} \pi_{ml}.
\]

The first summation in Equation 2 accounts for all single
failures that affect fiber links traversed by the working path.
The second summation accounts for all double failures which
affect a fiber link traversed by the working path and other fiber
link which is not traversed by the working path. The third
summation accounts for all double failures which affect two
fiber links traversed by the working path.

### III. A STRATEGY FOR THE DYNAMIC PROVISIONING OF
CONNECTIONS WITH GUARANTEED AVAILABILITY IN
SHARED-BACKUP PATH PROTECTED NETWORKS

Reference [15] demonstrated that the problem of finding an
eligible path pair \( w-b \) under shared-path-protection constraints
for a lightpath request with respect to existing lightpaths is
NP-complete. We therefore developed a heuristic method to
compute a feasible solution \( w-b \) for a connection request \( c \)
under availability constraints. Let a fiber link be a bidirectional
fiber connecting two nodes and transporting wavelength links.
Consider the notation in Table 1.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>number of active connections.</td>
</tr>
<tr>
<td>( L )</td>
<td>set of all fiber links.</td>
</tr>
<tr>
<td>( f_w )</td>
<td>set of fiber links traversed by ( w ).</td>
</tr>
<tr>
<td>( f_b )</td>
<td>set of fiber links traversed by ( b ).</td>
</tr>
</tbody>
</table>
| \( f_{SG} \) | set of fiber links traversed by working paths of connections that
share some spare capacity with \( c \). |
| \( G \) | set of all wavelength links. |
| \( W \) | set of all wavelength links used for routing working paths. |
| \( B \) | set of all wavelength links used for routing backup paths. |
| \( w_c \) | set of wavelength links used for routing \( w \). |
| \( b_c \) | set of wavelength links used for routing \( b \). |
| \( s_c \) | source node of \( c \). |
| \( d_c \) | destination node of \( c \). |
| \( a_c \) | availability lower bound of \( c \). |
| \( \epsilon_{ac} \) | estimated availability of connection \( c \). |
| \( \pi_m \) | probability that the network contains a single failure of link \( m \). |
| \( \pi_{lm} \) | probability that the network contains a double failure of links \( m \)
and \( l \), and link \( l \) failed before link \( m \). |

### A. Problem Statement

Each connection request \( c \) is associated to 3 inputs: \( s_c, d_c \)
and \( a_c \). The aim of the heuristic method is to route request \( c \)
from \( s_c \) to \( d_c \) trying to allocate minimal spare capacity while
respecting the availability constraints:

1) \( \epsilon_{ac} > a_c \).
2) \( \epsilon_{an} > a_n \), for all \( n=1,2,3...N \) active connections.

If this is not possible, the request is blocked.

### B. Algorithm

1) Choose \( w_c \) in the shortest path with random wavelength
using free wavelength links, i.e., \( w_c \subseteq B \cup W \). If there
is no available path return: request blocked.

2) Estimate \( ea_c \) using Equation 2.

3) If \( ea_c > a_c \)
   \( \text{then \{ no backup path is needed; \}} \)
   \( \text{return } w_c: \text{ request accepted} \}
   \( \text{else \{ a backup path is needed; \}} \)
   \( \text{go to step 4} \} \)

4) Remove from \( G \) all wavelength links that cannot be
used for routing \( b_j \):
   i) erase all wavelength links that belong to \( W \cup w_c \);
   ii) erase all wavelength links that belong to \( B \) and
are used by backup paths of connections whose
 corresponding working paths traverse fiber links which
transport \( w_c \) (SRLG constraint);
   iii) for each connection index \( j \) already active in the
network:

\[
\text{for } < j = 1 \text{ to } < j = N > \text{ do } \}
\]

\[
ea_j = 1 - \sum_{m \in f_{w_j}, l \in f_{b_j}} \pi_{ml} - \sum_{m \in f_{w_j}, l \in f_{SG_j}} \pi_{lm};
\]

\[
\text{if } ea_j < a_j \}
\text{then \{ erase all wavelength links from } b_j \} \}
\]

5) Set the cost of the remaining wavelength links that belong
to \( B \) to zero, and set the cost of the remaining links that
do not belong to \( B \) to one.

6) Find the \( k \) minimum-cost alternatives for \( b_c \). If there is
no available path return: request blocked.

7) Let \( s_j \) be the set of wavelength links traversed
by the \( j^{th} \) minimum-cost alternative for \( b_c \);
for \( < j = 1 \text{ to } < j = k > \text{ do } \}
\text{Estimate } ea_c \text{ using Equation 1;}
\text{ if } ea_c > a_c \}
\text{then \{ \}
\text{return } w_c \text{ and } b_c: \text{ request accepted} \}
\}

8) Return: request blocked.

Steps 1, 2 and 3 of the algorithm verify if the availability re-
quirements of the incoming connection request can be achieved
with just an unprotected working path. If this is not the case,
step 4 removes from the network topology all the wavelength
paths that cannot be used for routing the backup path. Case
4i accounts for all wavelength links that are used by any
working path of active connections, or the working path of the
incoming connection request. Case 4ii removes the wavelength
links that cannot be used due to the SRLG constraint. Case
4iii removes the wavelength links used by backup paths of
connections that cannot share protection capacity with the
incoming request because they would have their availability
constraint violated. Equation 3 recalculates the unavailability
of each active connection to verify if backup capacity may be
shared with the incoming request. Note that knowledge about
the working path (and not about backup path) of the incoming request is enough to calculate how would be the unavailability of each active connection if they shared capacity with the incoming request. After step 4 only the free wavelength links and the backup wavelength links which can be used for routing \(b_e\) remained. Since the aim of the algorithm is to allocate minimal spare capacity, step 5 sets the cost of the remaining wavelength links that belong to \(B\) to zero, and the cost of the remaining wavelength links that do not belong to \(B\) to one. Steps 6 and 7 find the \(k\) minimum-cost alternatives for \(b_e\) and sequentially checks, for all minimum-cost alternatives, if the chosen \(w_{c-b_e}\) pair satisfies the unavailability constraint. Once a satisfactory \(w_{c-b_e}\) pair is found the connection request is accepted, otherwise blocked.

A possible implementation bottleneck for the algorithm is the off-line utilization of Yen’s algorithm, which assumes full knowledge of connections routed within the whole network. However, the aim of this paper is not to assess the efficiency of the algorithm itself, but to analyze the coexistence of connections with different availability classes in networks of national and continental dimensions.

C. Computational Complexity

The routing algorithms used in this paper, the Dijkstra’s algorithm and the Yen’s algorithm, rely on the mapping of the network topology onto a graph representation with nodes and edges. Although this paper does not consider wavelength conversion, the algorithm implementation was prepared to take it into account in future studies (“Wavelength Graph with Extra Nodes” implementation [16]). Each of the \(n\) network nodes is regarded as an atomic entity directly mapped onto a graph node. Each graph node representing a network node is connected via graph edges to other \(\lambda\) graph nodes (which we call “virtual” graph nodes) that represent the wavelengths that can be switched by the network node. These graph edges are weighed according to the conversion cost: since no wavelength conversion is provided, the edges connected to source or destination nodes have cost 0, and the others have cost infinity. Each fiber link is mapped onto edges that connect virtual graph nodes that correspond to the same wavelength and are attached to neighboring network nodes. Therefore, the final graph has \((\lambda n + n \approx \lambda n)\) nodes and \((\lambda S + \lambda n)\) edges, where \(S\) is the number of fiber links. As stated before, the weight of these graph edges is dynamically updated whenever a connection is established or finished.

The worst-case computational complexity of finding a shortest path in step 1 without wavelength conversion and using Dijkstra’s algorithm is approximately \(O((\lambda n)^2)\). This complexity can be reduced to \(O(\lambda n \log(\lambda n))\) using more efficient implementations. Since the connection unavailability estimation using Equations 1, 2 and 3 involves just additions of terms which correspond to specific double fiber link failures, its computational complexity is \(O(S^2)\). The most computationally expensive part is step 6, which finds the \(k\)-minimum cost alternatives for \(b_e\) using Yen’s algorithm with complexity \(O(k(\lambda n)^3)\) [17].

IV. Performance Evaluation

The performance of connection classes with guaranteed availability is evaluated using an event-based simulator written in the Java programming language. Uniform and homogeneous traffic is assumed. The arrival of connection requests follows a Poisson process, and the holding time is exponentially distributed. Connection requests are equally divided into three availability classes: class 1 with 0.999, class 2 with 0.9999, and class 3 with 0.99999 availability. We simulated two representative network topologies: an Italian network [13] with a MTTR = 20 h and 200 FIT/km as value for 1/MTTF (conservative hypothetical values, reference [18] assumes MTTR = 21 h and 100 FIT/km), and the NSFNet network [19] with a MTTR = 12 h and 310 FIT/km as value for 1/MTTF (Bellcore Figures [20]). In general, link failure rates are obtained from operator statistics and the repair time is a parameter that can be controlled by the network operator. All links are bidirectional with 10 wavelengths. No wavelength conversion is provided. The curves are calculated by averaging the results of at least 10 rounds of 500 connection requests each. The results provided by the first round are discharged to simulate a steady-state network occupancy. The results analysis was based on the following metrics, which apply individually to each availability class:

\[
\begin{align*}
B_W &\quad \text{Fraction of connection requests blocked in step 1 of the algorithm, i.e., no working path was found.} \\
B_B &\quad \text{Fraction of connection requests blocked in step 6 of the algorithm, i.e., a working path was found, but no backup path was available.} \\
B_{AC} &\quad \text{Fraction of connection requests blocked in step 8 of the algorithm, i.e., working and backup paths were found, but the availability constraints for the current connection could not be fulfilled.} \\
B_T &\quad \text{Total blocking probability: } B_T = B_W + B_B + B_{AC}. \\
M_{SG} &\quad \text{SG is the set of working paths of which backup paths share some capacity with the backup path of a specific connection. } M_{SG} \text{ is the average, over all connections that use protection, of the maximal size of set } SG \text{ during the existence of a connection. In DPP } M_{SG} \text{ would be zero, since connections do not share protection resources.}
\end{align*}
\]

V. Simulation Results

Two network topologies were simulated: an Italian network representing topologies of national dimensions, and the NSFNet network representing topologies of continental dimensions. As the network dimensions grow, constraints on the connection unavailability become more restrictive. In the following figures for blocking values the error intervals appear at a confidence level of 95%. Figure 2 shows the blocking results for three classes that coexist in the Italian network. Figure 2a presents the blocking performance of class 1 connections. In this case it is interesting to notice that for all investigated traffic loads – even for high load values and blocking probabilities – \(B_{AC}\) remained equal to zero. This means that for class 1 connections an unprotected path or a single backup path is enough to
Figure 2: Blocking Probabilities for an Italian network [13]. Solid line with circles: class 1 connections. Dashed line with triangles: class 2 connections. Dash-dotted line with stars: class 3 connections. Dot-dotted line with circles: class 3 connections. Dotted line with squares: class 3 connections.

Figure 3: Average largest-SG size $M_{SG}$ for the Italian network [13]. Solid line with circles: class 1 connections. Dash-dotted line with empty squares: class 2 connections. Dotted line with solid squares: class 3 connections.

guarantee the availability requirements. Figure 2b presents the blocking performance of class 2 connections, which is only slightly different to the performance of class 1 connections. As expected, the curves for $B_W$ are equivalent, but for high network loads $B_B$ is higher for class 2 than for class 1. This can be explained by the fact that less class 2 unprotected connections could be accepted in the network. Figure 2c shows the blocking performance of class 3 connections. At low network loads about 30% of class 3 connection requests are blocked because the availability constraint could not be fulfilled with a single backup path. This is therefore a lower bound on the blocking probability of class 3 connections. As the network load increases $B_W$ and $B_B$ start to contribute to $B_T$. It is also noticeable that $B_B$ is equivalent for class 2 and class 3 connections, which means for both a reduced acceptance of unprotected connections. Although at least 30% of class 3 connections are blocked, if priced accordingly these accepted connections with “carrier-class availability” can be an important source of revenue.

In some cases, it may be worth to control the admission of class 1 or class 2 connections to release network capacity for routing class 3 connections. For the investigated case study this sort of control would only have an influence for more than 1.5 Erl traffic load per node, when class 3 connections start to be blocked not only because of $B_{AC}$, but also $B_W$ and $B_B$. Figure 3 shows the performance of the $M_{SG}$ metric for the three connection classes which coexist in the Italian network. Connections class 1 and 2 performed similarly, and for high traffic load values (full network) the $M_{SG}$ value stabilizes in about 4.5 connections. For class 3 connections this value is about 4 connections, which reflects the sharing constraints imposed by the provisioning algorithm.

Figure 4 shows the blocking results for three classes that
VI. CONCLUSION

This paper compares the performance of connection classes with guaranteed availability in SBPP protected networks. Connections are provisioned according to a novel strategy that attempts to reduce the allocation of spare capacity while guaranteeing the availability requirements. The connection availability estimation relies on the matrix-based approach for connection unavailability estimation in SBPP protected networks which offers accurate estimates for networks of national size.
Two network topologies have been simulated: an Italian network representing networks of national dimensions, and the NSFNet network representing networks of continental dimensions. The results indicate that SBPP can be efficiently employed in the Italian network, in which the network dimensions and the availability requirements allow backup capacity sharing. On the other hand, in the NSFNet network high availability classes exhibit a degraded performance in terms of blocking probability and bandwidth sharing. This can be explained by the fact that long connections are highly susceptible to double link failures which can cause outages in protected connections. In these cases other alternatives can be employed, such as link restoration, path protection using several backup paths, or the use of SBPP or DPP in protection domains within the network. The work reinforces the applicability of the matrix-based approach for connection unavailability estimation in SBPP, which offers accurate results for networks of national dimensions.

ACKNOWLEDGEMENT

This work has been partially supported by FAPESP and CNPq. The authors would like to thank Gustavo Souza Pavani for providing the implementation of Yen’s algorithm.

REFERENCES