The Power of Tuning: A Novel Approach for the Efficient Design of Survivable Networks*

Ron Banner and Ariel Orda
Department of Electrical Engineering
Technion – Israel Institute of Technology
Haifa 32000, Israel
{banner@tx, ariel@ee}.technion.ac.il

Abstract— **Current survivability schemes typically offer two degrees of protection, namely *full protection* (from a single failure) or *no protection* at all. Full protection translates into rigid design constraints, i.e. the employment of disjoint paths. We introduce the concept of *tunable survivability* that bridges the gap between full and no protection. First, we establish several fundamental properties of connections with tunable survivability. With that at hand, we devise efficient polynomial (optimal) connection establishment schemes for both 1:1 and 1+1 protection architectures. Then, we show that the concept of tunable survivability gives rise to a novel *hybrid protection* architecture, which offers improved performance over the standard 1:1 and 1+1 architectures. Next, we investigate some related QoS extensions. Finally, we demonstrate the advantage of tunable survivability over full survivability. In particular, we show that, by just slightly alleviating the requirement of full survivability, we obtain major improvements in terms of the "feasibility" as well as the "quality" of the solution.**

I. INTRODUCTION

In recent years, transmission capabilities have increased to rates of 10 Gbit/s and beyond [15]. With this increase, any failure may lead to a vast amount of data loss. Consequently, various survivability strategies have been proposed and investigated (e.g., [2],[7],[8],[11],[14]). These strategies are based on securing an independent resource for each potentially faulty network element [9]. This requirement usually translates into the establishment of pairs of disjoint paths. Two major survivability architectures that employ the use of (link) disjoint paths are the 1+1 and 1:1 protection architectures. In the 1+1 protection architecture, the data is concurrently sent on a pair of disjoint paths. The receiver picks the better path and discards data from the other path. In the 1:1 protection architecture, data is sent only on one (active) path, while the other (backup) path is activated by signaling only upon a failure on the active path.

Under the common single link failure model, the employment of disjoint paths provides full (100%) protection against network failures. However, in practice, this requirement is often too restrictive and requires excessive redundancy. Indeed, in many cases this requirement is infeasible (when pairs of disjoint paths do not exist) and in other cases it is very limiting and results in the selection of inefficient routing paths [15]. Therefore, it has been noted that a milder and more flexible survivability concept is called for, which would relax the rigid requirement of disjoint paths [15]. However, to the best of our knowledge, no previous work has systematically addressed this problem.

In this study, we introduce the concept of *tunable survivability*, which provides a *quantitative measure* to specify the desired level of survivability. This concept allows any degree of survivability in the range 0% to 100% and, in contrast to the rigid requirement of disjoint paths, it offers flexibility in the choice of the routing paths; consequently, it enables to consider valuable *tradeoffs* for designing survivable networks, such as survivability vs. feasibility, survivability vs. available bandwidth, survivability vs. delay performance, etc.

More specifically, tunable survivability enables the establishment of connections that can survive network failures with any desired probability p. We investigate these connections under the widely used single link failure model, which has been the focus of most studies on survivability e.g., [2],[6],[7],[8],[10],[11],[13],[16],[19]. Given a connection that consists of k paths (between some source-destination pair), under this failure model only a failure on a link that is *common* to all k paths can break (fail) the connection[1]. Accordingly, we characterize a set of paths (a connection) as a *p-survivable connection* if there is a probability of at least p to have all common links operational during the connection's lifetime. The following example illustrates the concept of p-survivable connections and demonstrates their power with respect to the traditional scheme of disjoint paths.

Example 1: Consider the network described in Fig. 1. Assume that it is required to establish a connection from S to T; moreover, assume that during the connection's lifetime each link

* Conference version in Proc. IEEE ICNP 2004 (Best Paper Award recipient).

[1] Indeed, a failure on a link that is not common to all k paths maintains at least one operational path; therefore, since only single failures can take place, such failures keep the connection operational.

independently fails with a probability of 0.01; finally, assume that the performance of the link e_3 is poor (e.g., e_3 has high delay or low bandwidth). As no pair of disjoint paths from S to T exists in the network, the traditional survivability requirement is infeasible; hence, there is no full protection against single network failures. Suppose now that we are satisfied with 0.99-survivability against single network failures. In that case, a connection that consists of the paths $\pi_1=(S,a,b,d,T)$ and $\pi_2=(S,c,d,T)$ fits since the only (single) failure that can concurrently damage both paths is a failure in the common link $e_7=(d,T)$; therefore, since the link e_7 fails with a probability of 0.01, the connection (π_1,π_2) is 0.99-survivable. Now suppose that we are satisfied with $(0.99)^2$-survivable connections. In that case it is easy to see that for $\pi_3=(S,a,b,c,d,T)$ and $\pi_4=(S,c,d,T)$ the connection (π_3,π_4) can also be used; thus, increasing the space of feasible solutions and bypassing the inferior link e_3. Finally, assume that we are satisfied with $(0.99)^3$-survivable connections. In that case it is easy to see that the *single* path π_2 turns to be a feasible solution too.

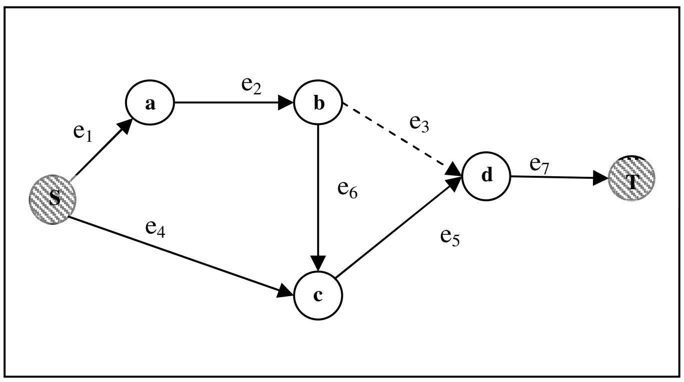

Fig. 1: A reference network for the discussion of *p*- survivable connections.

Through comprehensive simulations we demonstrate the significant advantages of the tunable survivability concept. In essence, we show that, at the price of a *negligible* reduction in the level of survivability, we obtain a *major* increase in the bandwidth as well as in the feasibility of the solutions.

Motivated by the above results, we investigate the tunable survivability concept from several different aspects and for different protection architectures. To that end, we first establish several fundamental properties of *p*-survivable connections under the single link failure model. In particular, we prove that, if it is possible to establish a *p*-survivable connection through *more than two paths*, then it is also possible to establish such a connection (i.e., with the same probability *p*) through *exactly two paths*.[1] Hence, in this study, we focus on survivable connections that consist of exactly two paths. Next, for both the 1+1 and the 1:1 protection architectures, we design efficient schemes for the establishment of *p*-survivable connections. Basically, for each protection architecture, we propose two types of survivability schemes: schemes that aim at *widest p*-survivable connections (i.e., *p*-survivable connections with maximum bandwidth) and schemes that aim at maximum survivability (i.e., connections with the maximum probability to survive single failures). We also show that each of the proposed schemes can be enhanced in order to consider QoS requirements. Finally, we establish that all schemes achieve the *optimal* solution and are *computationally efficient*.

Next, we turn to show that the concept of tunable survivability gives rise to a third protection architecture, which is an hybrid between 1:1 protection and 1+1 protection. This new architecture is shown to have several important advantages over both the 1:1 and the 1+1 protection architectures; moreover, we show that the schemes that we have established for achieving either widest or most survivable connections in the case of 1:1 protection achieve the same goals in the case of hybrid protection.

The rest of this paper is organized as follows. In Section II, we introduce some terminology and formally define the concept of tunable survivability. In Section III, we investigate several properties of connections with tunable survivability. In Section IV, we design efficient schemes that establish *most survivable* and *widest p-survivable* connections for the 1:1 and 1+1 protection architectures. In section V, we introduce the Hybrid Protection architecture, demonstrate its advantages and establish corresponding algorithmic schemes. In Section VI, we show how our schemes can be enhanced in order to consider QoS requirements. Section VII presents simulation results that demonstrate the advantages of tunable survivability. Finally, Section VIII summarizes our results and discusses directions for future research.

II. MODEL AND PROBLEM FORMULATION

A *network* is represented by a directed graph $G(V,E)$, where V is the set of nodes and E is the set of links. Let $N=|V|$ and $M=|E|$. A *path* is a finite sequence of nodes $\pi=(v_0,v_1,\cdots v_h)$, such that, for $0\le n\le h-1$, $(v_n,v_{n+1})\in E$. A path is *simple* if all its nodes are distinct. Given a source node $s\in V$ and a target (destination) node $t\in V$, the set $P^{(s,t)}$ is the collection of all directed paths from the source s to the target t. Each link $e\in E$ is associated with a *bandwidth value* $b_e\in\mathbb{Z}^+$, which corresponds to the available bandwidth in the link. In addition, each link $e\in E$ is assigned a *weight* $w_e\in\mathbb{Z}^+$ that represents some QoS target such as delay, cost, jitter, etc.

Definition 1: Given a (non-empty) path π, its *bandwidth* $B(\pi)$ is defined as the bandwidth of its bottleneck link, namely, $B(\pi)\triangleq\underset{e\in\pi}{Min}\{b_e\}$.

We adopt the *single link failure model*, which assumes that at most one link failure can take place in the network. A link is classified as *faulty* upon its failure and it remains faulty until it is *repaired*. We say that a link $e\in E$ is *operational* if it is not faulty. Likewise, we say that a path π is *operational* if it has no faulty link i.e., for each $e\in\pi$, link e is operational.

[1] While this is a trivial property for *disjoint paths* under the single link failure model, it is far from trivial, and actually quite surprising, for paths that may be non-disjoint.

Definition 2: Given a pair of source and destination nodes *s* and *t*, a *survivable connection* is a pair of paths $(\pi_1,\pi_2)\in P^{(s,t)}\times P^{(s,t)}$.[1]

We say that a connection (π_1,π_2) is *operational* if either π_1 or π_2 are operational. We assume that during the lifetime of (π_1,π_2) each link $e\in E$ independently fails with a probability $p_e\in[0,1]$; we note that these link failure probabilities are often estimated out of the available failure statistics of each network component [6].

Survivability is defined to be the capability of the network to maintain service continuity in the presence of failures [12]. Under the single link failure model a survivable connection (π_1,π_2) is operational *iff* the links that are common to both π_1 and π_2 are operational[2]. Accordingly, as the failure probabilities $\{p_e\}$ are independent, we quantify the level of survivability of survivable connections as follows.

Definition 3: Given a survivable connection (π_1,π_2) such that $\pi_1\cap\pi_2\neq\phi$, we say that (π_1,π_2) is a *p-survivable connection* if $\prod_{e\in\pi_1\cap\pi_2}(1-p_e)\geq p$ i.e., the probability that all common links are operational is at least *p*. The value of *p* is then termed as the *survivability level* of the connection.

Definition 3 formalizes the notion of tunable survivability for the single link failure model. Note that when there are no common links between π_1 and π_2 (i.e., paths π_1 and π_2 are disjoint), there is no single failure that can fail (π_1,π_2); for this case (π_1,π_2) is defined to be a 1-survivbale connection.

We now quantify the bandwidth of a survivable connection. We consider first a connection (π_1,π_2) under the standard (full) survivability requirement. This means that π_1 and π_2 are disjoint, namely $\pi_1\cap\pi_2=\phi$. Obviously, for 1+1 protection, the maximum protected traffic rate that can be transferred via (π_1,π_2) is the minimum available bandwidth on any of the two paths. That is, the bandwidth of the connection (π_1,π_2) is $\min\{B(\pi_1),B(\pi_2)\}=\min_{e\in\pi_1\cup\pi_2}\{b_e\}$. However, for connections with tunable survivability, paths are not necessarily disjoint. Therefore, for the 1+1 protection architecture, the total traffic rate that traverses links that belong to both π_1 and π_2 is twice the rate that traverses links that belong to only one out of the two paths. Accordingly, the available bandwidth of a survivable connection with respect to 1+1 protection is defined as follows.

Definition 4: Given a survivable connection (π_1,π_2), its *bandwidth with respect to the 1+1 protection architecture* is the maximum $B\geq 0$ such that $2\cdot B\leq b_e$ for each $e\in\pi_1\cap\pi_2$ and $B\leq b_e$ for each $e\in(\pi_1\cup\pi_2)\setminus(\pi_1\cap\pi_2)$.

In contrast to 1+1 protection, in 1:1 protection only one duplicate of the original traffic is carried at any given time. Therefore, the only restriction here is that traffic rate should not exceed the bandwidth of any of the links in $\pi_1\cup\pi_2$. Accordingly, we formulate the bandwidth of a survivable connection with respect to the 1:1 protection architecture as follows.

Definition 5: Given a survivable connection (π_1,π_2), its *bandwidth with respect to the 1:1 protection architecture* is the maximum $B\geq 0$ such that $B\leq b_e$ for each $e\in\pi_1\cup\pi_2$.

For a source-destination pair, there might be several *p*-survivable connections. Among them, we may be interested in those that have the best "quality". The following definitions correspond to two important quality criteria namely, maximum survivability and maximum bandwidth.

Given a network $G(V,E)$ and a pair of nodes *s* and *t*, we say that a *p*-survivable connection $(\pi_1,\pi_2)\in P^{(s,t)}\times P^{(s,t)}$ is a *most survivable connection* if there is no $\hat{p}$-survivable connection $(\hat{\pi}_1,\hat{\pi}_2)\in P^{(s,t)}\times P^{(s,t)}$ such that $\hat{p}>p$; *p* is then termed the *maximum level of survivability*. Next, we say that a *p*-survivable connection (π_1,π_2) is the *widest p-survivable connection for the 1+1 protection architecture* if it is a *p*-survivable connection that has the largest bandwidth with respect to that architecture. Similarly, we say that (π_1,π_2) is the *widest p-survivable connection for the 1:1 protection architecture* if it is a *p*-survivable connection that has the largest bandwidth with respect to that architecture. In section VI we shall define additional quality criteria.

Finally, note that, whereas the widest *p*-survivable connection depends on the considered protection architecture, a most survivable connection for one architecture is also such for the other architecture.

III. TWO PATHS ARE ENOUGH

Consider a more general protection framework that admits *any* (≥ 2) number of paths. Basically, we show that, in any network and for each survivability constraint $0\leq p\leq 1$, if there exists a - *p*-survivable connection that admits *more than two paths*, then there exists a *p*-survivable connection that admits *exactly two paths*. Moreover, we show that the bandwidth of the widest *p*-survivable connection in protection frameworks where connections are allowed to employ any number of paths *is not larger* than the bandwidth of the widest *p*-survivable connection that is limited to at most two paths.

Remark 1: For completeness, we note that a *p*-survivable connection in protection frameworks that admit more than two paths is a collection of paths $(\pi_1,\pi_2,\cdots,\pi_k)\in P^{(s,t)}\times P^{(s,t)}\times\cdots\times P^{(s,t)}$ such that

[1] As was already mentioned, we will show that there is no advantage in the employment of more than two paths; hence, the definition focuses on two paths.

[2] As mentioned, under the single link failure model, a link that is not common to both paths can never cause a survivable connection to fail; similarly, a failure in a common link, causes a failure of the entire connection.

$\prod_{e\in\pi_1\cap\pi_2\cdots\cap\pi_k}(1-p_e)\geq p$. The bandwidth of such a connection with respect to the 1:1 protection architecture (i.e., in the case where the traffic is sent only over a single path) is the maximum $B\geq 0$ such that $B\leq b_e$ for each $e\in\bigcup_{i=1}^{k}\pi_i$. Similarly, the bandwidth of such a connection with respect to the 1+1 protection architecture (i.e., in the case where the traffic is carried independently over each path) is the maximum $B\geq 0$ such that $n\cdot B\leq b_e$ for each link $e\in E$ that is common to some *n* paths out of $(\pi_1,\pi_2,\cdots,\pi_k)$.

We are now ready to formulate two fundamental properties of survivable connections; the first corresponds to widest *p*-survivable connections and the second to most survivable connections. The proof of both properties can be found in the Appendix.

Property 1: Let $(\pi_1,\pi_2,\cdots,\pi_k)\in P^{(s,t)}\times P^{(s,t)}\times\cdots\times P^{(s,t)}$ be the most survivable connection in $G(V,E)$ and let $(\overline{\pi_1},\overline{\pi_2})\in P^{(s,t)}\times P^{(s,t)}$ be the most survivable connection in $G(V,E)$ that consists of at most two paths. The survivability level of $(\overline{\pi_1},\overline{\pi_2})$ is not smaller than that of $(\pi_1,\pi_2,\cdots,\pi_k)$.

Property 2: Let $(\pi_1,\pi_2,\cdots,\pi_k)\in P^{(s,t)}\times P^{(s,t)}\times\cdots\times P^{(s,t)}$ be the widest *p*-survivable connection in $G(V,E)$ with respect to the 1:1 (alternatively, 1+1) protection architecture. There exists a *p*-survivable connection $(\overline{\pi_1},\overline{\pi_2})\in P^{(s,t)}\times P^{(s,t)}$ that has at least the bandwidth of $(\pi_1,\pi_2,\cdots,\pi_k)$ with respect to the 1:1 (correspondingly, 1+1) protection architecture.

The above observations show that there is no incentive to define survivable connections that consist of more than two paths. Therefore, under the standard single link failure model, this finding indicates an important network design rule in terms of survivability.

IV. ESTABLISHING *P*-SURVIVABLE CONNECTIONS

In this section we show how to construct *p*-survivable connections for the 1+1 and 1:1 protection architectures. In view of the findings of the previous section, we focus on survivable connections that consist of at most two paths. We begin with the establishment of widest *p*-survivable connections and most survivable connections for the 1+1 protection architecture.

A. *Establishing Survivable Connections for the 1+1 Protection Architecture*

The first step towards the establishment of either widest *p*-survivable or most survivable connections is the development of an efficient algorithm that, for any $B\geq 0$, establishes a survivable connection with a bandwidth of at least *B* that has the maximum level of survivability. We term such a connection as the *most survivable connection with a bandwidth of at least B*.

Remark 2 Finding the most survivable connection with a bandwidth of at least *B* is beneficial per se. For example, in cases where the traffic demand γ is known in advance, it may be desired to establish a connection with a bandwidth of at least γ that optimizes the level of survivability of the established connection.

I. Establishing most survivable connections with a bandwidth of at least B

We now establish an efficient algorithm that, for any $B\geq 0$, outputs the most survivable connection that has a bandwidth of at least *B*. Given a network $G(V,E)$, a pair of nodes *s* and *t*, a bandwidth constraint $B\geq 0$, and, for each link $e\in E$, a bandwidth $b_e\geq 0$ and a failure probability $p_e\geq 0$, the algorithm reduces the problem of finding the most survivable connection with a bandwidth of at least *B* into an instance of the Min Cost Flow problem [1]. In essence, the construction is based on a network transformation that considers three different cases, as illustrated in Fig. 2. In the case of a link $e\in E$ with a bandwidth $b_e<B$, it follows by definition (Def. 4) that link *e* cannot be used by any survivable connection that has a bandwidth of at least *B*. Therefore, this link can be discarded from the network without any influence on the optimal solution. On the other hand, each link $e\in E$ that satisfies $b_e\geq 2\cdot B$ can be used concurrently by both of the connection's paths in order to establish a survivable connection with a bandwidth of at least *B*. In that case, the corresponding link is transformed into two parallel links, each with a link bandwidth of *B*; however, whereas the first link is assigned with a zero weight, the other link is assigned with a weight that is a function ($g(p_e)$) of the link's failure probability (p_e). The reason for that follows from the quantification of the survivability level of each connection (Definition 3), which is solely determined by its common links. Specifically, only when both of the connection's paths share the same link *e*, the link's failure probability p_e should be considered. Indeed, a Min Cost Flow (where "cost" is "weight") over the constructed network ensures that, when a single path traverses link *e*, the incurred cost is zero, whereas when both paths traverse through *e*, the cost $g(p_e)$ depends on the failure probability p_e ($g(p_e)$ shall be specified in the following). The third case corresponds to links that satisfy $B\leq b_e<2\cdot B$. In that case, at most one path with a bandwidth *B* can traverse through such a link without violating the link bandwidth b_e. Thus, these links can be transformed into links that have a bandwidth *B* without any effect on the optimal solution. In addition, since these links can be used by at most one path, their failure probabilities should not be considered and therefore the transformed links are assigned zero weight.

Denote the transformed network as $\widetilde{G}(\widetilde{V},\widetilde{E})$. The algorithm computes a min-cost flow $\{f_e\}$ with a flow demand of $2\cdot B$ units over the network $\widetilde{G}(\widetilde{V},\widetilde{E})$ by employing any standard Min Cost Flow algorithm that returns an integral link flow when all

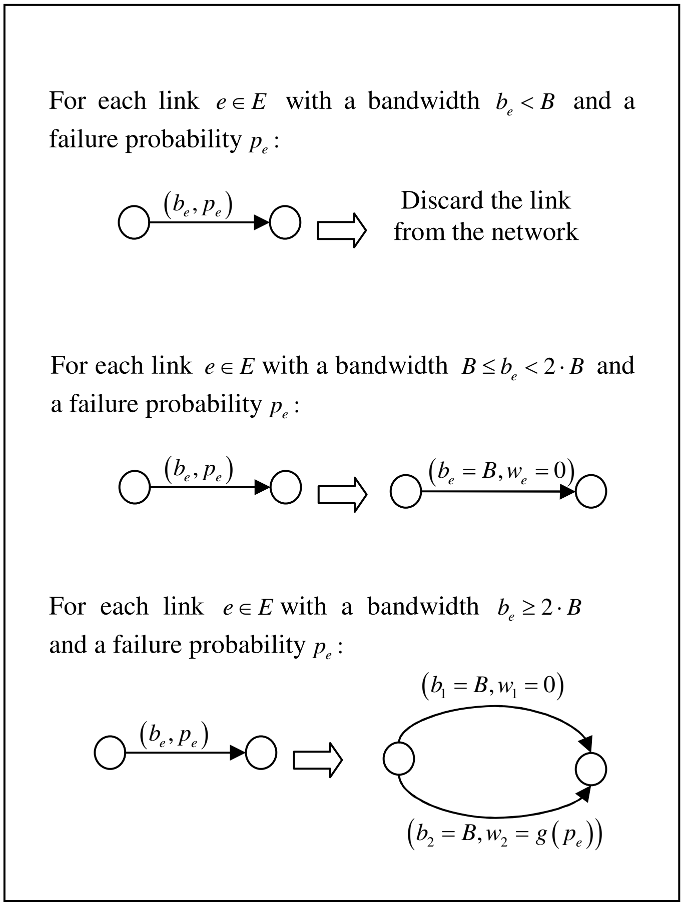

Fig. 2: Finding the most survivable connection with a bandwidth of at least *B* (for the 1+1 protection architecture) by a reduction to the Min Cost Flow problem.

link bandwidths $\{b_e\}$ are integral (see [1]). Since all link bandwidths in $\widetilde{G}(\widetilde{V}, \widetilde{E})$ are integral in *B*, the link flow $\{f_e\}$ is *B*-integral i.e., f_e is a multiple of *B* for each $e \in E$. Therefore, since the total traffic equals to $2 \cdot B$ flow units, the *flow decomposition algorithm* [1] can be applied in order to decompose the link flow $\{f_e\}$ into a flow over two paths π_1, π_2 such that each carry *B* flow units from *s* to *t*. Moreover, since the flow has minimum cost, it follows that $\sum_{\tilde{e} \in \tilde{E}} f_{\tilde{e}} \cdot w_{\tilde{e}} = \sum_{e \in \pi_1 \cap \pi_2} B \cdot g(p_e) = B \cdot \sum_{e \in \pi_1 \cap \pi_2} g(p_e)$ has minimum value. Thus, $\sum_{e \in \pi_1 \cap \pi_2} g(p_e)$ has minimum value. Finally, if we define $g(p_e) \triangleq -\ln(1 - p_e)$ for each $e \in E$, the algorithm defines a pair of paths π_1, π_2 that minimizes $-\sum_{e \in \pi_1 \cap \pi_2} \ln(1 - p_e) = -\ln \prod_{e \in \pi_1 \cap \pi_2} (1 - p_e)$ and therefore maximizes $\ln \prod_{e \in \pi_1 \cap \pi_2} (1 - p_e)$. Thus, it maximizes $\prod_{e \in \pi_1 \cap \pi_2} (1 - p_e)$ i.e., the connection level of survivability. The formal description of the algorithm appears in Fig. 3.

The following theorem shows that, for every $B \ge 0$, our algorithm establishes the most survivable connection with a bandwidth of at least *B*; the proof appears in the Appendix.

Theorem 1: Given are a network $G(V, E)$, a pair of nodes *s* and *t*, a bandwidth constraint $B \ge 0$, and, for each link $e \in E$, a bandwidth $b_e \ge 0$ and a failure probability $p_e \ge 0$. If there exists a survivable connection with a bandwidth of at least *B*, then *Algorithm B-Width Most Survivable Connection* returns the most survivable connection with a bandwidth of at least B; otherwise, the algorithm fails.

II. Establishing most survivable and widest p-survivable connections

Finally, we are ready to construct most survivable connections and widest *p*-survivable connections for the 1+1 protection architecture. As is easy to see, the most survivable connection with a bandwidth of at least *B*=0 is in essence also a most survivable connection. As the corresponding problem is a special case of the problem that was addressed in the previous subsection, in this section we only focus on the establishment of widest *p*-survivable connections.

In order to establish the widest *p*-survivable connection, we employ Algorithm B-Width Most Survivable Connection. Specifically, given a network and a survivability constraint *p*, we search for the largest value of *B* such that the most survivable connection with a bandwidth of at least *B* is a *p*-survivable connection Obviously, this strategy is attractive only if we consider a small number of bandwidth constraints before we get to the bandwidth of the widest *p*-survivable connection. Fortunately, in the following we show that it is sufficient to consider just $O(\log N)$ bandwidth constraints in order to find the bandwidth of the widest *p*-survivable connection.

First, we observe that, for every given network, the bandwidth of the widest *p*-survivable connection belongs to a set of at most $2 \cdot M$ values. To see this, recall that the bandwidth of each survivable connection (π_1, π_2) with respect to the 1+1 protection architecture, is defined as the maximum $B \ge 0$ such that $2 \cdot B \le b_e$ for each $e \in \pi_1 \cap \pi_2$ and $B \le b_e$ for each $e \in (\pi_1 \cup \pi_2) \setminus (\pi_1 \cap \pi_2)$. Hence, if the survivable connection (π_1, π_2) admits a link $e \in E$, then by definition, its bandwidth with respect to the 1+1 protection, is not larger than either $\frac{b_e}{2}$ (for $e \in \pi_1 \cap \pi_2$) or b_e (for $e \in (\pi_1 \cup \pi_2) \setminus (\pi_1 \cap \pi_2)$). Moreover, it follows by definition that there exists at least one link $e \in \pi_1 \cup \pi_2$ such that the bandwidth of (π_1, π_2) is either $\frac{b_e}{2}$ or b_e. Therefore, each survivable connection in $G(V, E)$ has a link $e \in E$ whose bandwidth is either $\frac{b_e}{2}$ or b_e. In particular, the bandwidth of the widest *p*-survivable connection in the network, denoted as B^*, must belong to the set $\mathbb{B} \triangleq \left\{ \frac{b_e}{k} \middle| e \in E,\ k = 1, 2 \right\}$,

Algorithm B - Width Most Survivable Connection $(G,\{s,t\},\{b_e\},\{p_e\},B)$

Parameters:

$G(V,E)$ – network

s – source

t – target (destination)

$\{b_e\}$ – link bandwidth values

$\{p_e\}$ – failure probabilities

B – bandwidth constraint

Variables:

$\tilde{G}(\tilde{V},\tilde{E})$ – network

$\{b_{\tilde{e}}\}$ – link bandwidth values

$\{w_{\tilde{e}}\}$ – weights

$\{f_{\tilde{e}}\}$ – link flow

$\tilde{\gamma}$ – flow demand

1. Construct an instance $\langle \tilde{G}(\tilde{V},\tilde{E}),(\tilde{s},\tilde{t}),\{b_{\tilde{e}}\},\{w_{\tilde{e}}\},\tilde{\gamma}\rangle$ of the Min Cost Flow Problem as follows:
   a. $\tilde{V} \leftarrow V$.
   b. For each link $e: u \rightarrow v \in E,\ b_e \in [B, 2\cdot B)$ construct a link $\tilde{e}$ between $\tilde{u}$ and $\tilde{v}$. Assign it with a bandwidth B and a zero weight.
   c. For each link $e: u \rightarrow v \in E,\ b_e \geq 2\cdot B$:
      - Construct a *cheap* link $\tilde{e}_1$ between $\tilde{u}$ and $\tilde{v}$. Assign it with a bandwidth B and a zero weight.
      - Construct an *expensive* link $\tilde{e}_2$ between $\tilde{u}$ and $\tilde{v}$. Assign it with a bandwidth B and a weight of $-\ln(1-p_e)$.
   d. $\tilde{\gamma} \leftarrow 2\cdot B$.
2. Solve the instance $\langle \tilde{G}(\tilde{V},\tilde{E}),(\tilde{s},\tilde{t}),\{b_{\tilde{e}}\},\{w_{\tilde{e}}\},\tilde{\gamma}\rangle$ of the Min Cost Flow Problem using the *Cycle Canceling Algorithm* [1].
3. If there is no feasible solution for the instance, then return Fail. Otherwise let $\{f_{\tilde{e}}\}$ represent the solution.
4. Construct a link flow $f: E \rightarrow \{0, B, 2\cdot B\}$ as follows:
   - For each link $e: u \rightarrow v \in E,\ b_e \in [B, 2\cdot B)$ set $f_e \leftarrow f_{\tilde{e}}$.
   - For each link $e: u \rightarrow v \in E,\ b_e \geq 2\cdot B$ set $f_e \leftarrow f_{\tilde{e}_1} + f_{\tilde{e}_2}$.
5. Employ the *Flow Decomposition Algorithm* [1] over link flow $\{f_e\}$, in order to obtain a pair of paths π_1, π_2 such that each path carries B flow units from s to t.
6. Return connection (π_1, π_2).

Fig. 3: Algorithm B-Width Most Survivable Connection. The algorithm establishes the most survivable connection that has a bandwidth of at least *B*.

which consists of at most $2\cdot M$ members. In Fig. 4 we provide the formal specification of the algorithm.

Remark 3: Note that we can employ a *binary search* over the set $\mathbb{B}$ in order to find the value of B^*.[1] Indeed, for each $B \in \mathbb{B}$, if the most survivable connection with a bandwidth of at least B is a *p*-survivable connection then so are all the other most survivable connections with bandwidths of at least $B', B' \leq B$; on the other hand, when the most survivable connection with a bandwidth of at least B is *not* a *p*-survivable connection, then none of the most survivable connections with bandwidth of at least $B'', B'' > B$, is a *p*-survivable connection.

We consider now the complexity incurred by the establishment of most survivable connections and widest *p*-survivable connections. To that end, we denote by $T(N,M)$ the running time of any standard min-cost flow algorithm for an *N*-nodes *M*-links network. Since Algorithm B-Width Most Survivable Connection solves a single instance of the min-cost flow problem, the complexities of establishing most survivable connections and widest *p*-survivable connections are $O(T(N,M))$ and $O(T(N,M)\cdot \log N)$, respectively.

Remark 4: We note that it is possible to solve the min-cost flow problem in $O((M\cdot \log N)\cdot(M+N\cdot\log N))$ operations [1]; hence, we can establish widest *p*-survivable connections and most survivable connections within a total complexity of $O(M^2\cdot\log^2 N + M\cdot N\cdot \log^3 N)$ and $O(M^2\cdot \log N + M\cdot N\cdot\log^2 N)$, respectively.

Finally, we note that, while the algorithms presented above consider a single survivable connection, multiple connections can be handled by executing the respective algorithm sequentially, each time considering a different connection. For the establishment of Widest *p*-Survivable Connections, it is required to update the values of available bandwidth after the establishment of each connection.

B. *Establishing Survivable Connections for the 1:1 Protection Architecture*

We turn to establish survivable connections for the 1:1 protection architecture. Obviously, the most survivable connection in the 1+1 protection architecture is the same as that of the 1:1 protection architecture; therefore, we will only consider the establishment of widest *p*-survivable connections for the 1:1 protection architecture. Moreover, as the establishment of the widest *p*-survivable connection with respect to the 1:1 protection architecture is quite similar as for the 1+1 protection architecture, we only sketch the main ideas.

As before, we begin by finding a solution to the dual problem of establishing the most survivable connection with a bandwidth of at least *B* (however, this time the bandwidth is computed according to the 1:1 protection architecture). This is based on a reduction similar to that of Fig 2. However, as the bandwidth of any survivable connection (π_1, π_2) for the 1:1 protec-

[1] Thus, considering only O(logN) bandwidth constraints before we find B^*.

Algorithm Widest p-Survivable Connection $(G,\{s,t\},\{b_e\},\{p_e\},p)$

Parameters:

G − network

s − source

t − target (destination)

$\{b_e\}$ − link bandwidth values

$\{p_e\}$ − failure probabilities

p − survivability constraint

Variables:

B − bandwidth constraint

(π_1,π_2) − survivable connection

1. Perform a binary search over the set $\mathbb{B} \triangleq \left\{\frac{b_e}{k} \middle| e\in E,\ k=1,2\right\}$ in order to find the largest $B\in\mathbb{B}$ that satisfies $\prod_{e\in\pi_1\cap\pi_2}(1-p_e)\geq p$, where (π_1,π_2) is obtained by $(\pi_1,\pi_2)\leftarrow$ Algorithm **B-Width Most Survivable Connection** $(G,s,t,\{b_e\},\{p_e\},B)$.

2. **If** the search failed

Return Fail.

Else

Return (π_1,π_2).

Fig. 4: Algorithm Widest *p*-Survivable Connection.

tion architecture is defined as the largest $B\geq 0$ such that $B\leq b_e$ for each $e\in\pi_1\cup\pi_2$, it follows that only two cases should be considered in the reduction, namely $b_e<B$ and $b_e\geq B$. More specifically, as before, all the links that satisfy $b_e<B$ should be discarded from the network since they cannot be used in order to construct a survivable connection with a bandwidth of at least *B*. However, in contrast to the solution of the 1+1 protection architecture, all other links can be concurrently employed by the pair of paths that constitute the survivable connection. More precisely, the only difference between the reduction that corresponds to the 1+1 protection architecture and the reduction that corresponds to the 1:1 protection architecture, is the type of links that can be used by both paths; namely, whereas in the 1+1 protection architecture the most survivable connection with a bandwidth of at least *B* cannot employ a link $e\in E$ that satisfies $B\leq b_e<2\cdot B$ for both paths, in the 1:1 protection architecture such a link can be common to both paths. The reduction for the 1:1 protection architecture is illustrated in Fig. 5.

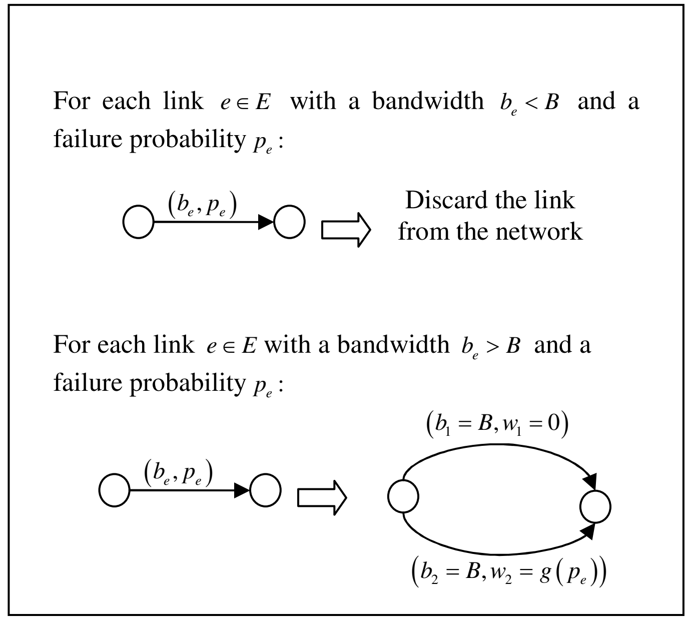

Fig. 5: Finding the most survivable connection with a bandwidth of at least *B* for the 1:1 protection architecture by a reduction to the Min Cost Flow problem.

As before, given a scheme for constructing most survivable connections with a bandwidth of at least *B*, we employ a binary search in order to find the largest *B* such that the most survivable connection with a bandwidth of at least *B* is a *p*-survivable connection. However, this time the bandwidth of the widest *p*-survivable connection belongs to the set $\{b_e|e\in E\}$, which consists of at most M elements (as opposed to the previous case where it belongs to a set of at most $2\cdot M$ elements). To see this, note that, by definition, the bandwidth of the survivable connection (π_1,π_2) with respect to the 1:1 protection architecture is the bandwidth of its bottleneck link i.e., $\min_{e\in\pi_1\cup\pi_2}\{b_e\}$. Therefore, the bandwidth of each survivable connection with respect to the 1:1 protection architecture is determined by some link in $e\in E$ i.e., it belongs to $\{b_e|e\in E\}$.

V. A Hybrid Protection Architecture

Thus far, we have focused on the 1+1 and 1:1 protection architectures. However, the tunable survivability concept gives rise to an efficient third protection architecture, which is a *hybrid* approach that combines the 1:1 and 1+1 protection architectures. More specifically, given a survivable connection (π_1,π_2) with a traffic demand γ, we present a new architecture that, for a connection (π_1,π_2), transfers γ flow units over the links in $\pi_1\cap\pi_2$, as in 1:1 protection, while over the links in $(\pi_1\cup\pi_2)\setminus(\pi_1\cap\pi_2)$, it transfers γ flow units, as in 1+1 protection. This new architecture is illustrated through the following example.

Example 2: Consider the network depicted in Fig. 6. Suppose that we are given a survivable connection (π_1,π_2) such that

$\pi_1 = (e_1, e_3, e_4)$ and $\pi_2 = (e_2, e_3, e_5)$. Hybrid Protection transfers one duplicate of the original traffic through link $e_1 \in \pi_1$ and another duplicate through link $e_2 \in \pi_2$. While both duplicates arrive to node *u*, only the *first* to arrive is assigned to link $u \to v$ and the other one is discarded. When the duplicate that was assigned to $u \to v$ arrives to *v*, Hybrid Protection transfers one duplicate through link $e_4 \in \pi_1$ and another through link $e_5 \in \pi_2$. Finally, as with 1+1 protection, node *t* considers only the duplicate that is the first to arrive. Note that such an assignment of traffic to links is *not* a flow.

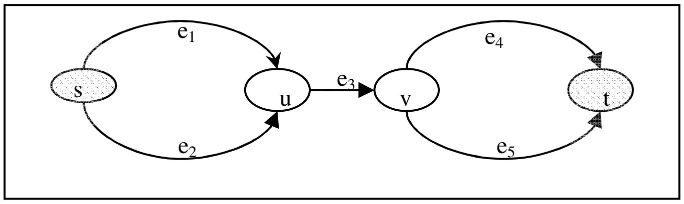

Fig. 6: The Hybrid Protection Architecture

Hybrid Protection has several important advantages. First, it reduces the congestion of all links that are shared by both paths with respect to 1+1 protection. At the same time, upon a link failure, it has a faster restoration time than 1:1 protection. Finally, it provides the *fastest* propagation of data with respect to the propagation time of *all* paths that can be constructed out of the links in $\pi_1 \cup \pi_2$. We demonstrate this property on the above example. Assume that the link propagation delays satisfy $d_{e_1} < d_{e_2}$ and $d_{e_5} < d_{e_4}$. Then, by construction, node *u* assigns the incoming flow of link e_1 over link e_3, and node *t* considers only the duplicate of link e_5. Thus, data is propagated through the path $\pi = (e_1, e_3, e_5)$, which has the *minimum propagation delay* among all the paths that can be constructed out of the links in $\pi_1 \cup \pi_2$.

The above advantages notwithstanding, the implementation of the Hybrid Protection architecture requires additional nodal capabilities in comparison with the 1+1 and 1:1 architectures. To see this, note that node *u* in the example must be able to discard all the duplicates that it encounters for the second time i.e., the duplicates that contain data that was already sent to node *v*. This is in contrast to the 1+1 protection architecture, where such functionality is required only from the destination, and the 1:1 protection architecture, where such functionality is not required at all.

Finally, note that the Hybrid Protection architecture transfers through each link (that belongs to the survivable connection) exactly one duplicate of the original traffic. Hence, the maximum traffic rate that can be transferred through a survivable connection (π_1, π_2) with respect to Hybrid Protection is bounded by $\min_{e \in \pi_1 \cup \pi_1} \{b_e\}$. In other words, the bandwidth of the survivable connection (π_1, π_2) with respect to Hybrid Protection is the maximum $B \geq 0$ such that $B \leq b_e$ for each $e \in \pi_1 \cup \pi_2$. Since this is precisely the definition of bandwidth with respect to 1:1 protection, a widest *p*-survivable connection with respect to 1:1 Protection is also a widest *p*-survivable connection with respect to Hybrid protection. Hence, we can employ the solution for 1:1 protection in order to establish widest *p*-survivable connections for Hybrid Protection. Nevertheless, it is important to note that, while 1:1 protection assigns traffic only to the links that belong to either π_1 or π_2, Hybrid Protection assigns traffic to *all* the links in $\pi_1 \cup \pi_2$.

VI. QUALITY OF SERVICE EXTENTIONS

Between a pair of nodes, there might be several widest *p*-survivable connections as well as several most survivable connections. Among them, we may be interested in those that optimize some QoS target, such as end-to-end delay, jitter, cost, etc. Such (additive) metrics can be represented by *weights* $\{w_e\}$. Given a network and a survivability constraint *p*, we show in the Appendix how to establish, for the 1+1, 1:1 and Hybrid Protection architectures, widest *p*-survivable connections as well as most survivable connections that *minimize the total weight* $\sum_{e \in \pi_1 \cup \pi_1} w_e$.

VII. SIMULATION RESULTS

The goal of this section is to demonstrate *how much* we gain by employing tunable survivability instead of traditional "full" survivability. To that end, we first compare between the maximum bandwidth of survivable connections that consist of a pair of disjoint paths (i.e.,1-survivable connections) and the maximum bandwidth of *p*-survivable connections, where $p \in [0,1)$. Then, we compare between the feasibility of both approaches i.e., the incidences where the establishment of pairs of disjoint paths is impossible and the incidences where the establishment of *p*-survivable connections is impossible. Through comprehensive simulations, we show that, at the price of a marginal reduction in the common requirement of 100% protection, a major increase in bandwidth as well as in feasibility is accomplished.

Remark 5: In Section V we have shown that the bandwidth of survivable connections with respect to the hybrid protection architecture is equal to that of the 1:1 protection architecture. Therefore, it is sufficient to conduct the simulations only for the 1:1 and 1+1 protection architectures. All the results of the 1:1 protection architecture also apply to the Hybrid Protection architecture.

A. *Setup*

We generated two types of random networks: network topologies that follow the four power laws defined by [5] (henceforth: *power-law* topologies), and networks with a *flat topology* i.e., Waxman networks [18] (henceforth: *flat* topologies). Then, we constructed 10,000 random networks for each combination of the following three items: (a) the level of survivability $p \in [0,1]$; (b) the type of protection architecture (i.e., either 1+1 or 1:1); and (c) the class of random networks (i.e., either power-law or flat). For each network, we identified a single source-destination pair. We then conducted the following measurements: (1) We measured the number of networks $N(p)$ that admits a *p*-survivable connection (between the selected source-destination pair) among the 10,000 networks; we then derived the *feasibility ratio* $\rho_N(p) \triangleq \frac{N(p)}{N(1)}$; (2) for each of the $N(1)$ networks that admit 1-survivable connections, we measured the ratio $\frac{B(p)}{B(1)}$, where $B(p)$ denotes the bandwidth of the widest *p*-survivable connection, and derived the corresponding *bandwidth ratio* $\overline{\rho_B(p)}$, which is the average value of $\frac{B(p)}{B(1)}$ over the corresponding $N(1)$ networks.

In all runs, we assumed that the link bandwidths are distributed uniformly in [5,150] MB/sec and the failure probability of each link is distributed normally with a mean of 1% and a standard deviation of 0.3%.

Topology Generation: We first specify the way we generated flat topologies. Our construction follows the lines of [18]. Initially, we located the source and the destination at the diagonally opposite corners of a square area of unit dimension. Then, we randomly spread 198 nodes over the square. Finally, we introduced a link between each two nodes *u* and *v*, with the following probability, which depended on the distance between them, $\delta(u,v)$:

$$p(u,v) = \alpha \cdot \exp\left[\frac{-\delta(u,v)}{\beta \cdot \sqrt{2}}\right],$$

using $\alpha=1.8$ and $\beta=0.05$. This resulted with 200 nodes and approximately 1800 links per network topology.

We turn to specify how we generated power-law topologies. Our construction followed the lines of [17]. First, we randomly assigned a certain number of *out-degree credits* to each node, using the power-law distribution $\beta \cdot x^{-\alpha}$, where $\alpha = 0.756$ and $\beta = 110$. Then, we connected the nodes so that every node obtained the assigned out-degree. More specifically, we randomly picked a pair of nodes *u* and *v*, and assigned a directed link from *u* to *v* if *u* had some remaining out-degree credits and link $u \rightarrow v$ had not been defined already. Whenever a link $u \rightarrow v$ was placed between the corresponding nodes, we also decremented the out-degree credit of node *u*. On the other hand, when the selected pair of nodes was not suitable for a link, we continued to pick pairs of nodes until finding one that was suitable. This resulted with 200 nodes and approximately 1200 links per network topology.

B. *Results*

First, we note that the value $N(1)$ i.e., number of networks that admitted 1-survivable connections, was in the range 4,000-7,000 (out of 10,000), hence the samples were always significant. In Figs. 7 and 8 we depict the bandwidth ratio $\overline{\rho_B(p)}$ versus the level of survivability $p \in [0.95,1]$ for 1:1 protection and 1+1 protection, respectively. In particular, for 1:1 protection (Fig. 7), we show that with a reduction of 2% in the requirement of full survivability, the bandwidth is increased *by 51%* for Waxman networks and *by 100%* for power law networks. When we consider the same reduction in survivability for 1+1 protection (Fig. 8), we see that the bandwidth is increased by 18% for Waxman networks and by 41% for power-law networks.

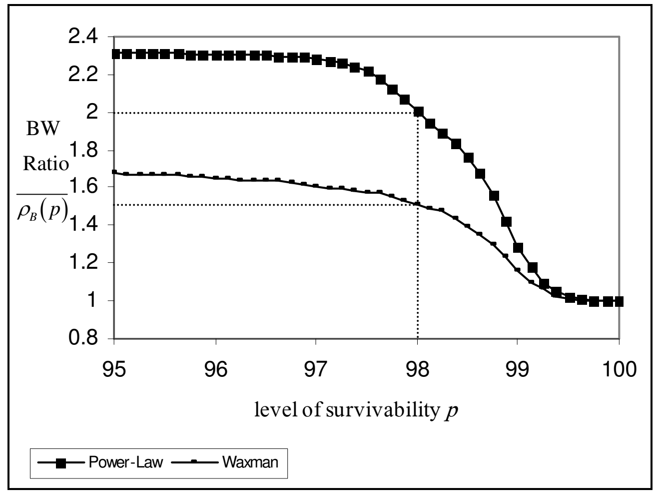

Fig. 7: The average ratio between the bandwidths of widest *p*-survivable connections and widest 1-survivable connections in the 1:1 protection architecture.

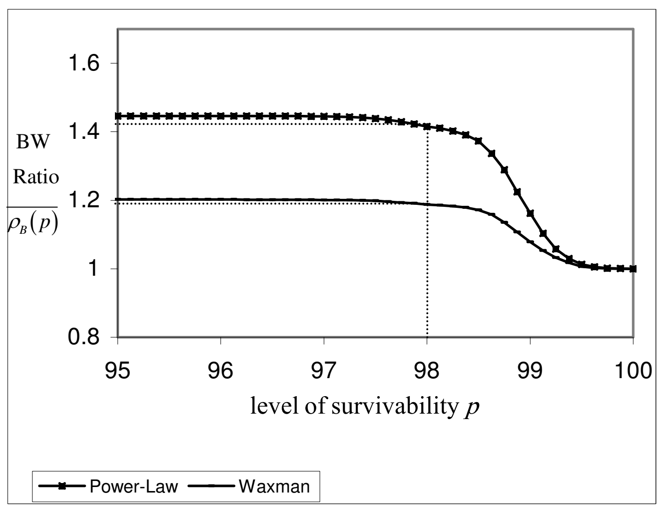

Fig. 8: The average ratio between the bandwidths of widest *p*-survivable connections and widest 1-survivable connections in the 1+1 protection architecture.

In Fig. 9, we depict the ratio between the number of networks that have at least one feasible *p*-survivable connection and the number of networks that have at least one feasible 1-survivable connection; to that end, we present the feasibility ratio $\rho_N(p)$ versus the level of survivability $p \in [0.95,1]$. Note that the feasibility ratio is independent of the employed protection architecture; therefore, the corresponding results hold for both protection architectures. We observe that, with a reduction of 2% in the requirement of full survivability, the feasibility ratio is increased *by 54%* for Waxman networks and *by 127%* for power law networks.

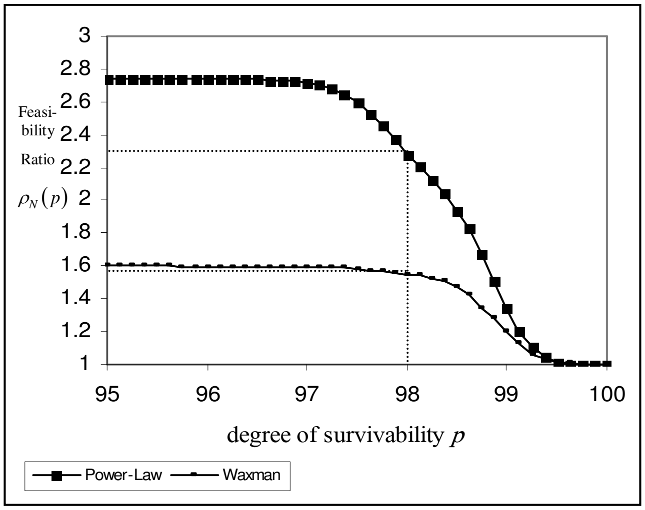

Fig. 9: The ratio between the number of networks with at least one feasible *p*-survivable connection and the number of networks with at least one feasible 1-survivable connection.

VIII. CONCLUSIONS

Standard survivability schemes enhance the ability to recover from network failures by establishing pairs of disjoint paths. However, in practice, this approach is too restrictive and often leads to the selection of poor routing paths (if any). In this work, we have proposed a novel quantitative approach for network survivability. The new approach allows to alleviate the rigid path disjointedness requirement, which considers only full (100%) protection, into a weaker requirement, which can be tuned to accommodate any desired degree (0%-100%) of survivability. Just as in the standard approach, we have shown that the new approach can also be accommodated by efficient polynomial (optimal) schemes. However, as opposed to the traditional approach, the new approach allows a flexible choice of the desired degree of survivability, hence enabling to consider important tradeoffs. Moreover, since a 1-survivable connection is also *p*-survivable (for any value of *p*), our approach always offers a solution of at least (and usually higher) quality than the traditional approach.

We have characterized several properties of the new approach. In particular, we established that, under the single link failure model, there is no benefit in establishing survivability schemes that employ more than two paths per connection. Since the single link failure assumption is practically valid in many cases of interest, this finding suggests an important network design rule in terms of survivability.

We evaluated the power of the new approach through comprehensive simulations. Our results clearly demonstrate the advantages of tunable survivability over full survivability. In particular, all measurements have shown that, by alleviating the traditional requirement of full survivability by just 2%, we obtained major improvements in the quality of the solutions. Effectively, this indicates that (traditional) full protection levies an excessive price.

Finally, we have shown that the tunable survivability approach gives rise to a new protection architecture that poses several advantages over current architectures; moreover, the new architecture was shown to admit computationally efficient optimal schemes.

The above notwithstanding, the practical deployment of the tunable survivability approach still posses several challenges. As mentioned, the hybrid protection architecture requires additional capabilities from transit nodes. The efficient implementation of these capabilities is an interesting issue for future work. More generally, although our algorithmic schemes are of polynomial complexity, in some cases simpler solutions might be called for. Therefore, it is of interest to investigate simpler heuristic schemes, which would be based on the insight provided by this study. Similarly, while our work focused on centralized algorithms, the employment of distributed schemes is often preferable, in particular in large-scale networks. Therefore, the distributed implementation of our algorithmic schemes is yet another interesting subject for future work. While much is still to be done towards the actual deployment of the tunable survivability approach, we believe that this study provides ample and firm evidence of its major benefits and potential practical feasibility.

APPENDIX

The Appendix contains the proofs of Properties 1 and 2 of Section III, the Proof of Theorem 1 of Section IV and the solution scheme for establishing survivable connections with QoS constraints (Section VI). We begin with the proofs of Properties 1 and 2.

A. Two paths are enough

The proof of Properties 1 and 2 immediately follows from the following theorem, which focuses on 1:1 protection (hence, also on Hybrid Protection). The corresponding proof for 1+1 protection goes along similar lines and is therefore omitted.

Theorem 2 Given are a network $G(V,E)$, a pair of nodes $\{s,t\}$ and, for each $e \in E$, a failure probability $p_e \geq 0$. Let $(\pi_1,\pi_2,\cdots,\pi_k) \in P^{(s,t)} \times P^{(s,t)} \times \cdots \times P^{(s,t)}$ be a *p*-survivable connection with a bandwidth of *B* with respect to the 1:1 protection architecture. There exists a *p*-survivable connection $(\overline{\pi_1},\overline{\pi_2}) \in P^{(s,t)} \times P^{(s,t)}$ that has a bandwidth of at least *B* with respect to the 1:1 protection architecture.

Proof: Let $\widehat{E} \triangleq \left\{ e \,\middle|\, e \in \bigcup_{i=1}^{k} \pi_i \right\}$ i.e., the collection of all links that are employed by the paths of the given survivable connection $(\pi_1,\pi_2,\cdots,\pi_k)$. We shall construct a survivable connection $(\overline{\pi_1},\overline{\pi_2}) \in P^{(s,t)} \times P^{(s,t)}$ such that $\overline{\pi_1} \cup \overline{\pi_2} \subseteq \widehat{E}$. Since by definition, the bandwidth of $(\pi_1,\pi_2,\cdots,\pi_k)$ with respect to the 1:1 protection architecture is determined by the bandwidth of its bottleneck link namely, $\min_{e \in \widehat{E}}\{b_e\}$ (see remark 1), it follows that

$\min_{e\in\hat{E}}\{b_e\}=B$; hence, since we shall construct the survivable connection $(\overline{\pi_1},\overline{\pi_2})$ *only* from links in $\hat{E}$, it follows that the bandwidth of $(\overline{\pi_1},\overline{\pi_2})$ with respect to 1:1 protection is at least *B* i.e., $\min_{e\in\overline{\pi_1}\cup\overline{\pi_2}}\{b_e\}\geq\min_{e\in\hat{E}}\{b_e\}=B$.

We now construct a pair of paths $\overline{\pi_1},\overline{\pi_2}\in P^{(s,t)}$ from links in $\hat{E}$ such that the level of survivability of $(\overline{\pi_1},\overline{\pi_2})$ is not less than the level of survivability of $(\pi_1,\pi_2,\cdots,\pi_k)$. Denote by $\overline{E}$ the set of all links that are common to the paths $\pi_1,\pi_2,\cdots,\pi_k$ i.e., $\overline{E}\triangleq\left\{e\,\middle|\,e\in\bigcap_{i=1}^{k}\pi_i\right\}$. Recall that the level of survivability of $(\pi_1,\pi_2,\cdots,\pi_k)$ is $\prod_{e\in\overline{E}}(1-p_e)$ and the level of survivability of $(\overline{\pi_1},\overline{\pi_2})$ is $\prod_{e\in\overline{\pi_1}\cap\overline{\pi_2}}(1-p_e)$. Thus, in order to establish the theorem we only need to show the existence of a pair of paths $\overline{\pi_1},\overline{\pi_2}\in P^{(s,t)}$ that satisfies $\overline{\pi_1}\cup\overline{\pi_2}\subseteq\hat{E}$ and $\overline{\pi_1}\cap\overline{\pi_2}\subseteq\overline{E}$.

Erase all links that are not used by the paths of $(\pi_1,\pi_2,\cdots,\pi_k)$ i.e., all links that are not in $\hat{E}$. We have to show that there exists a pair of paths $\overline{\pi_1},\overline{\pi_2}\in P^{(s,t)}$ over $G(V,\hat{E})$ such that $\overline{\pi_1}\cap\overline{\pi_2}\subseteq\overline{E}$. To that end, we employ the following construction that transforms $G(V,\hat{E})$ into a *flow network* [1]. Assign to each $e\in\overline{E}$ two units of bandwidth, and assign to each $e\in\hat{E}\setminus\overline{E}$ one unit of bandwidth. We now show that, if it is possible to define an integral link flow[1] that transfers two flow units from *s* to *t* over $G(V,\hat{E})$ then there exists a pair of paths $\overline{\pi_1},\overline{\pi_2}\in P^{(s,t)}$ over $G(V,\hat{E})$ such that $\overline{\pi_1}\cap\overline{\pi_2}\subseteq\overline{E}$.

Assume it is possible to define an integral link flow that transfers two flow units from *s* to *t* over $G(V,\hat{E})$. Hence, by the flow decomposition theorem [1], it is possible to define a pair of paths such that each path transfers one flow unit from *s* to *t* over $G(V,\hat{E})$. Moreover, the corresponding paths can intersect only at the links that have two units of bandwidth; hence, by construction, these paths intersect only at links that belong to $\overline{E}$. Thus, there exists a pair of paths $\overline{\pi_1},\overline{\pi_2}\in P^{(s,t)}$ over $G(V,\hat{E})$ such that $\overline{\pi_1}\cap\overline{\pi_2}\subseteq\overline{E}$.

Hence, in order to prove the theorem, it remains to be shown that it is possible to define an integral link flow that transfers two flow units from *s* to *t* over $G(V,\hat{E})$. However, since all the links have an *integral* bandwidth, the maximum flow that can be transferred from *s* to *t* under the integrality restriction is equal to the maximum flow that can be transferred from *s* to *t* when the integrality restriction is omitted [4][2]; hence, it is sufficient to show that it is possible to transfer two flow units from *s* to *t* over $G(V,\hat{E})$.

Suppose, by way of contradiction, that it is impossible to transfer two flow units from *s* to *t* over $G(V,\hat{E})$. Thus, according to the max-flow min-cut theorem [4], there exists a cut (S,T) with $s\in S$ and $t\in T$ such that $B(S,T)\triangleq\sum_{x\in S,y\in T}b_{x\to y}<2$ (where $b_{x\to y}$ denotes the bandwidth of link $x\to y\in\hat{E}$). Therefore, since the bandwidth of all links are integral, it follows that $B(S,T)\leq 1$. Thus, since each link has *at least* one unit of bandwidth, it follows that *at most* one link $x\to y\in\hat{E}$, such that $x\in S$ and $y\in T$, crosses (S,T). Denote this link by e.[3] Obviously, each path from *s* to *t* must traverse through the link *e*; in particular, all the paths of $(\pi_1,\pi_2,\cdots,\pi_k)$ must traverse the link *e*. Hence, by definition, it follows that $e\in\overline{E}$. However, since $b_e\leq 1$, the latter (i.e., the conclusion that $e\in\overline{E}$) contradicts our construction that assigns two units of bandwidth for each link $e\in\overline{E}$. Thus, it is possible to transfer two flow units from *s* to *t* over $G(V,\hat{E})$. ■

B. *Proof of Theorem 1*

The proof of Theorem 1 follows from the following two lemmas.

Remark 6: *Algorithm B-Width Most Survivable Connection* (Fig. 3) is said to *succeed* whenever it does not return Fail.

Lemma 1*:* Given are a network $G(V,E)$, a pair of nodes *s* and *t*, a bandwidth constraint $B\geq 0$, and, for each link $e\in E$, a bandwidth $b_e\geq 0$ and a failure probability $p_e\geq 0$. If *Algorithm B-Width Most Survivable Connection* succeeds for the input $\langle G,\{s,t\},\{b_e\},\{p_e\},B\rangle$, then the returned connection is a most survivable connection with a bandwidth of at least B.

Proof: Since Algorithm B-Width Most Survivable Connection succeeds for the input $\langle G,\{s,t\},\{b_e\},\{p_e\},B\rangle$ it follows by construction that there is a feasible solution for the Min Cost Flow instance $\langle\tilde{G}(\tilde{V},\tilde{E}),(\tilde{s},\tilde{t}),\{b_{\tilde{e}}\},\{w_{\tilde{e}}\},\tilde{\gamma}\rangle$ of Step 1. Therefore, since both the demand $\tilde{\gamma}$ and the link bandwidths $\{b_{\tilde{e}}\}$ of the min cost flow instance are integral in *B*, it follows that the output of the Cycle Canceling Algorithm (link flow $\{f_{\tilde{e}}\}$), executed in step 2, is integral in *B* (see [1]). Moreover, since all

[1] An integral link flow assigns over each link $e\in E$ a flow f_e that has an integer value.

[2] Indeed, according to the *Integrality Theorem* [4] if all link bandwidths are integral there exists at least one maximum flow that transfers over each link $e\in E$ an integral flow f_e.

[3] Obviously, there must exists at least one link that connects some node in *S* to some node in *T* since all paths in $(\pi_1,\pi_2,\cdots,\pi_k)$ are from s to t.

links in the network $\widetilde{G}\left(\widetilde{V},\widetilde{E}\right)$ have a capacity of B, it follows that, with the link flow $\left\{f_{\tilde{e}}\right\}$, each link transfers either zero or B flow units. Thus, as $\widetilde{\gamma}=2\cdot B$, it follows by the flow decomposition theorem [1], that link flow $\left\{f_{\tilde{e}}\right\}$ can be decomposed into a pair of paths in $\widetilde{G}\left(\widetilde{V},\widetilde{E}\right)$ that corresponds to a pair of paths $\pi_1,\pi_2\in P^{(s,t)}$ in $G\left(V,E\right)$, such that each transfers B flow units without violating the capacity constraints. We only need to show that $\left(\pi_1,\pi_2\right)$ has the maximum level of survivability with respect to all connections with a bandwidth of B.

To that end, we employ the fact that the total cost of the link flow $\left\{f_{\tilde{e}}\right\}$ is minimal with respect to all link flows that transfer $2\cdot B$ flow units from $\tilde{s}$ to $\tilde{t}$ in network $\widetilde{G}\left(\widetilde{V},\widetilde{E}\right)$. Therefore, since we have shown that $f_{\tilde{e}}\in\left\{0,B\right\}$ for each $\widetilde{e}\in\widetilde{E}$, it follows that $\sum_{e\in\widetilde{E}} f_{\tilde{e}}\cdot w_{\tilde{e}}=B\cdot\sum_{f_{\tilde{e}}\neq 0} w_{\tilde{e}}$ is minimal with respect to all link flows that transfer $2\cdot B$ flow units from $\tilde{s}$ to $\tilde{t}$; hence, $\sum_{f_{\tilde{e}}\neq 0} w_{\tilde{e}}$ is also minimal with respect to all link flows that transfer $2\cdot B$ flow units from $\tilde{s}$ to $\tilde{t}$.

Consider now an expensive link $\widetilde{e_x}:\tilde{u}\rightarrow\tilde{v}\in\widetilde{E}$ (as defined in Step 1.c), where $f_{\widetilde{e_x}}\neq 0$. It follows by construction that there exists a parallel cheap link $\widetilde{e_y}:\tilde{u}\rightarrow\tilde{v}\in\widetilde{E}$ such that $w_{\widetilde{e_y}}=0$. Therefore, since $\left\{f_{\tilde{e}}\right\}$ is a min-cost flow, it follows that $f_{\widetilde{e_y}}=B$. Thus, by construction, there exists a link $e_{xy}\in E$ that corresponds to the links $\widetilde{e_x}$ and $\widetilde{e_y}$ in $\widetilde{G}\left(\widetilde{V},\widetilde{E}\right)$ such that $e_{xy}\in\pi_1\cap\pi_2$. Conversely, since π_1 and π_2 are decomposed out of the link flow that corresponds to $\left\{f_{\tilde{e}}\right\}$, it follows that, if $e_{xy}\in\pi_1\cap\pi_2$, then there exists a corresponding pair of parallel links $\widetilde{e_x},\widetilde{e_y}\in\widetilde{E}$ such that $f_{\widetilde{e_x}}\neq 0, f_{\widetilde{e_y}}\neq 0$. Thus, it holds that $e_{xy}\in\pi_1\cap\pi_2$ in network $G\left(V,E\right)$ *iff* the corresponding expensive link $\widetilde{e_x}$ in network $\widetilde{G}\left(\widetilde{V},\widetilde{E}\right)$ satisfies $f_{\widetilde{e_x}}\neq 0$. Therefore, since the expensive links are the only non-zero cost links in $\widetilde{E}$, and since each expensive link $\tilde{e}\in\widetilde{E}$ that corresponds to the link $e\in E$ has a cost $w_{\tilde{e}}=-\ln\left(1-p_e\right)$, it follows that $\sum_{f_{\tilde{e}}\neq 0} w_{\tilde{e}}=\sum_{e\in\pi_1\cap\pi_2}-\ln\left(1-p_e\right)$. Thus, since we have shown that $\sum_{f_{\tilde{e}}\neq 0} w_{\tilde{e}}$ has the minimal value with respect to all B-integral link flows that transfer $2\cdot B$ flow units from $\tilde{s}$ to $\tilde{t}$ in $\widetilde{G}\left(\widetilde{V},\widetilde{E}\right)$, it follows by construction that the decomposed path flow (that assigns B flow units over each of the paths π_1 and π_2) has the minimum value for $\sum_{e\in\pi_1\cap\pi_2}-\ln\left(1-p_e\right)$ among all B-integral path flows that transfer $2\cdot B$ flow units from s to t in $G\left(V,E\right)$. Thus, we conclude that $\sum_{e\in\pi_1\cap\pi_2}-\ln\left(1-p_e\right)=-\ln\prod_{e\in\pi_1\cap\pi_2}\left(1-p_e\right)$ is minimal and therefore $\ln\prod_{e\in\pi_1\cap\pi_2}\left(1-p_e\right)$ is maximal with respect to all the pairs of paths that transfer B flow units (each) from s to t in $G\left(V,E\right)$. Thus, $\prod_{e\in\pi_1\cap\pi_2}\left(1-p_e\right)$ is maximal and, by definition, (π_1,π_2) is a most survivable connection with a bandwidth of at least B. Thus, the Theorem is established. ■

Lemma 2 Given are a network $G\left(V,E\right)$, a pair of nodes s and t, a bandwidth constraint $B\geq 0$, and, for each link $e\in E$, a bandwidth $b_e\geq 0$ and a failure probability $p_e\geq 0$. If *Algorithm B-Width Most Survivable Connection* fails for the input $\left\langle G,\left\{s,t\right\},\left\{b_e\right\},\left\{p_e\right\},B\right\rangle$, then there is no survivable connection with a bandwidth of at least B.

Proof: Assume that Algorithm B-Width Most Survivable Connection fails for the input $\left\langle G,\left\{s,t\right\},\left\{b_e\right\},\left\{p_e\right\},B\right\rangle$. Hence, it follows by construction that there is no feasible solution for the Min Cost Flow instance $\left\langle\widetilde{G}\left(\widetilde{V},\widetilde{E}\right),\left(\tilde{s},\tilde{t}\right),\left\{b_{\tilde{e}}\right\},\left\{w_{\tilde{e}}\right\},\widetilde{\gamma}\right\rangle$ of Step 1. Thus, it is impossible to transfer $2\cdot B$ flow units from $\tilde{s}$ to $\tilde{t}$ in the network $\widetilde{G}\left(\widetilde{V},\widetilde{E}\right)$. Hence, by construction, it is impossible to identify a B-integral link flow that transfers $2\cdot B$ flow units from s to t over the network $G\left(V,E\right)$ without violating the capacity constraints; thus, it is impossible to identify a pair of paths from s to t that carry B flow units each without violating the capacity constraints of $G\left(V,E\right)$. ■

C. *Quality Of Service Extensions*

In this appendix we specify our solution scheme for establishing widest p-survivable connections as well as most survivable connections that minimize the total weight $\sum_{e\in\pi_1\cup\pi_1} w_e$. We focus on 1:1 protection; the solution for 1+1 protection can be addressed in a similar way. The solution is based on the following observation.

Lemma 3: Given are a network $G(V,E)$, a pair of nodes $\{s,t\}$ and, for each link $e\in E$, a failure probability $p_e\geq 0$. If (π_1,π_2) and (π_3,π_4) are most survivable connections with a bandwidth of at least B, then $\pi_1\cap\pi_2=\pi_3\cap\pi_4$.

Proof: Assume by the way of contradiction that $\pi_1\cap\pi_2\neq\pi_3\cap\pi_4$. Hence, $\pi_1\cap\pi_2\cap\pi_3\cap\pi_4\subset\pi_1\cap\pi_2$; in particular,

$$\prod_{e\in\pi_1\cap\pi_2\cap\pi_3\cap\pi_4}(1-p_e) > \prod_{e\in\pi_1\cap\pi_2}(1-p_e) \qquad \text{(A1)}$$

Next, since the bandwidth of both (π_1,π_2) and (π_3,π_4) is at least B (with respect to 1:1 protection), it follows by definition that $\min_{e\in\pi_1\cup\pi_2}\{b_e\}\geq B$ and $\min_{e\in\pi_3\cup\pi_4}\{b_e\}\geq B$; in particular,

$$\min_{e\in\pi_1\cup\pi_2\cup\pi_3\cup\pi_4}\{b_e\}\geq B. \qquad \text{(A2)}$$

Consider now the survivable connection (π_1,π_2,π_3,π_4). Recall (remark 1) that the bandwidth of (π_1,π_2,π_3,π_4) is $\min_{e\in\pi_1\cup\pi_2\cup\pi_3\cup\pi_4}\{b_e\}$ and its level of survivability under the single link failure model is $\prod_{e\in\pi_1\cap\pi_2\cap\pi_3\cap\pi_4}(1-p_e)$. Therefore, from (A1) and (A2) it follows that (π_1,π_2,π_3,π_4) is a survivable connection with a bandwidth of at least B that has a level of survivability larger than (π_1,π_2). However, this contradicts the selection of (π_1,π_2). Indeed, (π_1,π_2) is a most survivable connection with a bandwidth of at least B; moreover, from theorem 2 it follows that (π_1,π_2) is a most survivable connection with a bandwidth of at least B also among survivable connections that can admit any number of paths; hence, the level of survivability of (π_1,π_2) cannot be smaller than the level of survivability of (π_1,π_2,π_3,π_4). ■

Define a *B-connection* as a survivable connection with a bandwidth of at least B. In the following we use lemma 3 in order to establish an algorithm that constructs in a polynomial time a B-connection with the maximum level of survivability as a primary objective and the minimum total weight $\sum_{e\in\pi_1\cup\pi_1} w_e$ as a secondary objective. Given are a network $G(V,E)$, a source-destination pairs s, t, link weights $\{w_e\}$, link bandwidths $\{b_e\}$ and failure probabilities $\{p_e\}$. In the first phase, the algorithm finds a most survivable connection with a bandwidth of at least B (using the algorithm described in Fig. 3); let (π_1,π_2) be the returned survivable connection. Note that since the level of survivability is solely determined by the common links (Def. 3), every B-connection that has the links $\pi_1\cap\pi_2$ as common links has the maximum level of survivability among all B-connections. On the other hand, according to Lemma 3, all B-connections with the maximum level of survivability must have the links $\pi_1\cap\pi_2$ as the common links; in particular, the most survivable connection with a bandwidth of at least B that minimizes the total weight $\sum_{e\in\pi_1\cup\pi_1} w_e$ must have $\pi_1\cap\pi_2$ as the links that are common to both of its paths. Therefore, we conclude that it is sufficient to find a B-connection that consists as common links only the links in $\pi_1\cap\pi_2$ such that $\sum_{e\in\pi_1\cup\pi_2} w_e$ is minimized. This is done as follows:

Transform the given network $G(V,E)$ into a network $G'(V,E')$ as follows. Each common link $e\in\pi_1\cap\pi_2$ is transformed into a pair of parallel links e_{p1}, e_{p2} each with a negative weight K, $K<-\sum_{e\in E} w_e$. The rest of the links in $G(V,E)$ that have a bandwidth of at least B (i.e., the links in $E\setminus(\pi_1\cup\pi_2)$ such that $b_e\geq B$) are assigned with a weight w_e. Finally, all links that have a bandwidth smaller than B are discarded from the network.

In the second phase, the algorithm identifies a pair of disjoint paths π'_1,π'_2 (from s to t) in $G'(V,E')$ such that $\sum_{e\in\pi'_1\cup\pi'_2} w_e$ is minimized[1]. Note that such a pair of disjoint paths must exist; indeed, in the returned survivable connection (π_1,π_2) all the links that are common to both paths are split in $G'(V,E')$ into a pair of parallel links. Next, note that the negative value of K is small enough so that all parallel links would be selected by π'_1,π'_2; indeed, for $K<-\sum_{e\in E} w_e$, the total weight of every pair of disjoint paths that does not consist all links from $\{e_{p1},e_{p2}\mid e\in\pi_1\cap\pi_2\}$ is suboptimal[2]. Finally, note that the paths π'_1,π'_2 in $G(V,E)$[3] has the minimum total weight $\sum_{e\in\pi_1\cup\pi_2} w_e$ among all paths that consist as common links the links in $\pi_1\cap\pi_2$; hence, the paths that correspond in $G(V,E)$ to the paths π'_1,π'_2 constitute a most survivable connection with a bandwidth of at least B that has the minimum total weight (alternatively, these paths constitute a B-connection with the maximum level of survivability as a primary objective and the minimum total weight $\sum_{e\in\pi_1\cup\pi_1} w_e$ as a secondary objective).

The most survivable connection and the widest p-survivable connection that minimize the total weight are established as in Section IV.A.2. Specifically, the most survivable connection with the minimum weight is established by requiring that B=0; the widest p-survivable connection with the minimum weight is established by searching for the largest $B\in\{b_e\mid e\in E\}$, such that the most survivable connection with a bandwidth of at least B is a p-survivable connection; since each iteration produces the most survivable connection with a bandwidth of at least B such that the total weight is minimized, this strategy produces the widest p-survivable connection with the minimum total weight.

REFERENCES

[1] R. K. Ahuja, T. L. Magnanti and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, 1993.

[2] M. Alicherry and R. Bhatia, *Pre-provisioning networks to support fast restoration with minimum over-build*, in *Proc.* IEEE INFOCOM, March 2004.

[1] Finding such a pair of paths is straightforward and can be done for example, by assigning each link in $G'(V,E')$ a unit capacity; then, any min cost flow (which can be computed in polynomial time) with a flow demand of 2 units can be decomposed into a pair of disjoint paths that has the minimum total weight [1].

[2] E.g., it is easy to see that the weight of the pair of paths in $G'(V,E')$ that corresponds to the returned paths π_1,π_2 in $G(V,E)$, is smaller than the weight of each pair of paths that does not consist one or more parallel links.

[3] I.e., the paths that correspond to the paths π'_1,π'_2 in $G(V,E)$.

[3] A.L. Barabasi and R. Albert. *Emergence of Scaling in Random Networks*, SCIENCE, 286:509-512, October 1999.
[4] T. Cormen, C. Leiserson, R. Rivest, Introduction to Algorithms, MIT press, Cambridge, 1990.
[5] M. Faloutsos, P. Faloutsos, and C. Faloutsos, *On Power-law Relationships of the Internet Topology*, in Proceedings of ACM SIGCOMM, cambridge, MA, September 1999.
[6] A. Fumagalli and M. Tacca, *Optimal Design of Optical Ring Networks with Differentiated Reliability (DiR)*, in Proceedings of the International Workshop on QoS in multiservice IP networks, Rome, Italy, Janurary 2001.
[7] M. T. Frederick and A. K. Somani, *A Single-Fault Recovery Strategy for Optical Networks using Sub-Graph Routing*, 7th IFIP Working Conference on Optical Networks Design and Modeling (ONDM 2003), Budapest, Hungary, Feh 2003.
[8] W.D. Grover and D. Stamatelakis, *Bridging the ring-mesh dichotomy with p-cycles*, In Proceedings of *DRCN Workshop*, 2000.
[9] P. Ho, J. Tapolcai, and H. T. Mouftah, *On Achieving Optimal Survivable Routing for Shared Protection in Survivable Next-Generation Internet*, IEEE Transactions on Reliability, March 2004.
[10] G. Italiano, R. Rastogi, and B. Yener. *Restoration Algorithms for Virtual Private Networks in the Hose Model*. In Proceedings of IEEE INFOCOM, New York, NY, June 2002.
[11] M. Kodialam and T. V. Lakshman, *Restorable Dynamic Quality of Service* Routing, IEEE Communication Magazine, vol. 40, no. 6, June 2002.
[12] W. Lai, Ed. and D. McDysan, Ed., *Network Hierarchy and Multilayer Survivability*, IETF RFC 3386, November 2002.
[13] G. Li, D. Wang, C. Kalmanek and R. Doverspike, *Efficient distributed restoration path selection for shared mesh restoration*, IEEE/ACM Transactions on Networking, Vol. 11, No. 5, October 2003.
[14] Y. Liu, D. Tipper, and P. Siripongwutikorn, *Approximating Optimal Space Capacity Allocation by Successive Survivable Routing ,* IEEE/ACM Transactions on Networking, Vol. 13, No. 1, February 2005.
[15] G. Maier, A. Pattavina, S. De Patre, and M. Martinelli, *Optical Network Survivability: Protection Techniques in the WDM Layer*, Photonic Networks Communications, vol. 4, no. 3-4, July-Dec. 2002.
[16] G. Mohan and A. K. Somani, *Routing Dependable Connections with Specified Failure Restoration Guarantees in WDM Networks,* in Proceedings of IEEE Infocom, Tel Aviv, Israel, April 2000.
[17] C. R. Palmer and J. G. Steffan, *Generating Network Topologies that Obey Power Laws*, in Proceedings of the Global Internet Symposium, (Globecom) 2000, San Fransisco, CA, November 2000.
[18] B. M. Waxman, *Routing of Multipoint Connections*, IEEE Journal on Selected Areas in Communications, 6:1617-1622,1988.
[19] D. Xu and C. Qiao, *Distributed Partial Information Management (DPIM) Schemes for Survivable Networks - Part 2*, in Proceedings of IEEE Infocom, NY, June 2002.