Adaptive Reliable Multipath Provisioning in Survivable WDM Mesh Networks

Sheng Huang, Chip Martel, and Biswanath Mukherjee

Abstract—We investigate the problem of adaptive reliable multipath provisioning in next-generation backbone mesh networks employing optical wavelength-division multiplexing (WDM) and channelization techniques such as SONET/SDH, and supporting virtual concatenation (VCAT). VCAT enables multipath provisioning, but also introduces differential delay at destination nodes. How to guarantee service availability, using multipath provisioning, and meet the differential-delay constraint (DDC) is an important research problem. We introduce a new notation $M : N (m)$ for multipath provisioning where a service path for a connection is set up with $M$ primary paths and $N$ backup paths, where each path has a fraction of the bandwidth of the connection, and $(m)$ in this notation denotes “multipath”. With $M : N (m)$ provisioning schemes, we develop an analytical model to analyze the end-to-end connection availability for the full bandwidth request. We propose two types of bandwidth migration methods, which can be implemented by link-capacity adjustment scheme (LCAS) protocol of next-generation SONET/SDH, to optimize resource usage. Based on the $M : N (m)$ analytical model, we develop an adaptive heuristic algorithm to provision a connection subject to DDC while satisfying its service-level agreement (SLA). We show that, for end-to-end connection-availability-guaranteed service, multipath provisioning can achieve better network performance than traditional single-path provisioning. With bandwidth migration, we can further decrease bandwidth blocking ratio by nearly 3 percent in a typical network topology.

Index Terms—virtual concatenation (VCAT), multipath provisioning, availability, SLA, wavelength-division-multiplexing (WDM), mesh network.

I. INTRODUCTION

Optical networks employing wavelength-division multiplexing (WDM) have emerged as the most viable infrastructure for wide-area backbone networks because of their tremendous capacity. A single fiber can support up to 160 wavelength channels, each of which can operate at a data rate of 10 Gbps or 40 Gbps. A single failure (e.g., fiber cut) can lead to a huge loss in data (and revenue) [1]. Therefore, survivability has emerged as a critical concern in network design and its real-time operation. Protection, a proactive procedure, is an important strategy to recover traffic when a failure occurs. The paths carrying traffic in normal operation are referred to as primary paths (or working paths), while the extra backup resources are reserved on backup paths and they will be activated when any primary path fails.

Recent advances and deployment of next-generation SONET/SDH technology, namely, generic framing procedure (GFP), virtual concatenation (VCAT), and link-capacity-adjustment scheme (LCAS) [2], enable flexible and reliable data transport over SONET/SDH networks. GFP, a traffic-adaptation protocol, is a “generic” mechanism to adapt either a physical layer or a logical link layer client traffic into a bit-synchronous or octet-synchronous transmission channel. VCAT is an inverse-multiplexing technique that groups diversely-routed smaller containers at the destination. A number of VCAT members which serve the same connection form a VCAT group (VCG). LCAS is a two-way signaling protocol built on top of VCAT, and it can dynamically adjust the bandwidth of a VCAT connection by adding/deleting VCG members in a hitless manner (i.e., without disrupting the on-going connections). LCAS offers network designers the ability to automatically fine tune the bandwidth based on bandwidth on demand, quality of service, minimum network infrastructure upgrade, dynamic load balancing/dynamic route optimization, and fault-recovery mechanisms. It can dynamically remove failed (or add recovered) VCG members.

Multipath provisioning, which is enabled by virtual concatenation (VCAT), has been shown to be more efficient and flexible; and it significantly improves the performance over conventional single-path provisioning [3]–[5]. However, the diversely-routed VCG members are subject to the differential-delay constraint (DDC) at the destination node, which can impact the QoS of the connection if not considered in the routing [6]. The dominant part of DDC is propagation delay, which is approximately 5 $\mu$s/km. DDC-aware routing in multipath provisioning has been studied in [7]–[9].

Connection availability, which is defined as the probability that the connection will be found in the operating state at a random time in the future, is an important QoS metric describing reliability. Network operators provide connections based on a service-level agreement (SLA), a contract which states the guaranteed connection availability. SLA is a contractual agreement outlining a specific service commitment made between contract parties – a service provider and its customer. It documents, among other details, the availability guarantees and the penalty if such guarantees are violated. The costs of services are differentiated based on services’ bandwidth and SLA. An efficient provisioning scheme to accommodate more service connections simultaneously is desirable for network operators to improve their revenue.

In this work, we investigate adaptive reliable multipath provisioning with guaranteed end-to-end availability of full service bandwidth. The purpose of adaptivity is to achieve availability-aware, differential-delay-aware, and resource-efficiency-aware connection service. We focus on...
DDC-aware link-disjoint paths for primary and backup paths, and exploit the LCAS feature to optimize resources and improve performance by decreasing connection blocking probability. We propose bandwidth migration methods and study their advantages and disadvantages in multipath provisioning.

Our work is different from previous multipath provisioning research in the following aspects: 1) in our work, both the requested bandwidth and availability are strictly satisfied; 2) we calculate link-disjoint paths for multipath provisioning. Obviously, protecting a connection over multiple link-disjoint paths has the advantage of fault tolerance; 3) we consider a differential-delay constraint among multiple provisioning paths; and 4) we choose primary and backup paths adaptively based on the SLA. For mission-critical high availability service requests, it is possible to use multiple backup paths to protect a single primary path. The objective of our work is to design efficient provisioning algorithms to exploit the benefits and constraints of existing technologies, namely, VCAT and LCAS, in multipath provisioning.

A. Related Work

For a single primary path, various provisioning schemes to satisfy the availability requirement have been proposed in [10]–[16]. When no primary path, that meets the availability requirement, can be found, a backup path and resources will be reserved to increase the connection availability. The availability analysis for differentiated single-primary-path provisioning has been presented in [12]. The backup resources can be shared among different connections if their primary paths are disjoint. However, backup-resource sharing can impact the existing connections’ availabilities, which can lead to SLA violation [10]. Backup re-provisioning is studied to improve the availability satisfaction ratio [15].

Survivability with multipath provisioning is also an important research problem [3]–[5], [17]. However, compared to single-path survivability, it is a relatively un-explored area. The effective multipath bandwidth concept was proposed in [4] to exploit the degraded service feature. By finding better sets of paths that preserve network capacity, a better provisioning algorithm, in the sense of capacity efficiency, was presented in [5]. The papers in [4], [5] investigate multipath provisioning using a very different SLA objective: effective multipath bandwidth, and thus, the results are not comparable to ours. Two heuristic algorithms for survivable virtual concatenation were proposed in [3]. However, the DDC of multipath provisioning was not considered. The DDC in multipath provisioning has been studied in [7], [9]. The authors in [7] studied the problem of minimizing the differential delay in a virtually-concatenated Ethernet-over-SONET (EoS) system by suitable path selection. A DDC-aware link-disjoint path routing algorithm was presented in [8]. Reference [18] presents the availability analysis for an optical backbone network. However, none of these works considered the connection availability of multipath provisioning with backup paths.

B. Our Contribution

We extend the study of multipath provisioning and addresses some of the limitations of previous research. We provide a framework, which includes an availability analysis model, route calculation, and a novel adaptive algorithm, for availability-guaranteed multipath provisioning. Our algorithm shows better usage of the network resources. We also propose and study two bandwidth migration methods for multipath provisioning.

We define successful $M : N(m)$ provisioning between a source and a destination as the following:

- Connection service bandwidth in normal operation is carried by $M (M \geq 1)$ primary paths.
- The pre-reserved backup capacity is carried in $N (N \geq 0)$ backup paths.
- Capacity of each backup path is allocated the maximum of any primary path, so a single backup path can protect any single primary path failure.
- All $M$ primary paths and $N$ backup paths are link-disjoint.
- If $M + N \geq 2$, the differential delay among all paths meets the service request, and network-wide DDC requirement.
- End-to-end connection availability meets the SLA of the service request.
- Total resource (capacity multiplied by the fiber distance) of primary and backup paths is minimized for the current network state (when the connection request arrived).

The backup paths are shared among the same VCG members. Backup sharing among different VCGs is not considered in this study. So, we do not need to consider vulnerable connections caused by excessive backup sharing [10].

For the $M : N(m)$ provisioning scheme, we develop a model to analyze the end-to-end connection availability. We exploit the benefit of LCAS and propose two types of bandwidth migration methods to optimize resource usage. Based on our $M : N(m)$ analytical model, we develop an adaptive heuristic algorithm to provision a connection subject to DDC while satisfying its SLA.

C. Organization

This paper is organized as follows. Section II provides availability-guaranteed multipath provisioning schemes and an analytical model. Section III introduces the bandwidth migration method with LCAS. Section IV presents the general problem statement, and proposes a novel multipath provisioning heuristic algorithm. In Section V, we evaluate the algorithm on a representative backbone topology, compare and discuss the simulation results. Section VI concludes this work.

II. CONNECTION AVAILABILITY-ANALYSIS MODEL FOR MULTIPATH PROVISIONING

We assume that a network component can only have two states (available or unavailable), network components fail independently, and for any network component, the normal operating and repair times are independent processes with known mean values (namely, Mean Time To Failure (MTTF) and Mean Time to Repair (MTTR)) [10]. The availability of an individual component is given by $\frac{MTTF}{MTTF + MTTR}$ [12], [19]. The availability of a multipath end-to-end connection from source to destination is
defined as the product of the availabilities of all the constituent components of all paths [18].

In this study, we refer to the individual components of a path to be links only, and we only consider link-disjoint paths for intra-VCG shared protection. In addition, we define full service availability and degraded service availability for multipath provisioning. Our notations follow.

$R$ a connection request between source and destination;

$a_j$ availability of link $j$;

$L_i$ all links along path $i$;

$M$ number of primary paths;

$N$ number of backup paths;

$S_{i}^{(M,k)}$ $i$th subset of $k$-subset of $\{1, 2, \cdots , M\}, \{k\}$;

$A_{R}$ connection availability (full service bandwidth availability);

$A_{R-p}$ degraded service (partial bandwidth) connection availability with $k$ paths;

$A_{p,i}$ availability of primary path $i$;

$A_{b,i}$ availability of backup path $i$;

$A_{p,i,k}$ availability of any $k$ out of $M$ primary paths (any $k$ paths available and $M-k$ paths unavailable);

$A_{b,i,k}$ availability of any $k$ out of $N$ backup paths (any $k$ backup paths available and $N-k$ backup paths unavailable).

A. Multipath Provisioning without Backup

The availability of a single primary path $A_{p,i}$ is the product of the availabilities of all the links along path $i$:

$$A_{p,i} = \prod_{j \in L_i} a_j$$  

(1)

Assuming $M$ primary paths, the end-to-end connection availability $A_{R}$, can be computed as:

$$A_{R} = \prod_{i=1}^{M} A_{p,i}$$  

(2)

Equation (2) calculates the full-service availability of connection $R$ (namely the expected fraction of the time all paths are working). In multipath provisioning, we can also provide degraded service with higher connection availability. The availability of degraded service, which can be provisioned with at least $k^{'\prime}$ ($k^{'\prime} \leq M$) primary paths, can be calculated as:

$$A_{R-p}^{k^{'\prime}} = \prod_{i=1}^{M} A_{p,i} + \sum_{q=k^{'\prime}}^{M} \sum_{i=1}^{M} \prod_{j \in S_{(M,q)}} A_{p,j} \prod_{l \neq j} (1 - A_{p,l})$$  

(3)

where the first term is the availability of using all $M$ primary paths, and the second term is the summation of the availability of all combinations of $q$ ($q \geq k^{'\prime}$) out of $M$ paths that are available and $M - q$ paths that are unavailable. $\binom{M}{q}$ is the binomial coefficient.

B. Multipath Provisioning with Backup

If the SLA cannot be satisfied with only primary paths, we use backup paths and allocate to the backup paths the maximum carried capacity of any primary path, so the backup path can protect any primary path in case of a failure. There could be multiple backup paths to satisfy high-availability requirements. Let us consider $M$ primary paths and $N$ backup paths. In case of any primary path failure, we can re-route the bandwidth of the failed primary path to any backup path. So, the full-service connection availability $A_{R}$ can be calculated as:

$$A_{R} = A_{p}^{M} + \sum_{i=1}^{\min(M,N)} \left( A_{p}^{M-i} \sum_{j=i}^{N} A_{b}^{j} \right)$$  

(4)

where $A_{p}$ is the availability of using all primary paths and the second term is the summation of the availabilities of all possible combinations of backup paths replacement. Below, we give a general availability-analysis model for $M: N(m)$ provisioning. In $M: N(m)$ provisioning, $N = 0$ means there is no backup path.

The connection availability ($A_{R}$) can be represented as:

$$A_{R} = \begin{cases} A_{p}^{M} + \sum_{i=1}^{N} \left( A_{p}^{M-i} \sum_{j=i}^{N} A_{b}^{j} \right) & \text{if } N \leq M; \\ A_{p}^{M} + \sum_{i=1}^{N} \left( A_{p}^{M-i} \sum_{j=i}^{N} A_{b}^{j} \right) & \text{if } N > M. \end{cases}$$  

(5)

where:

$$A_{p}^{k} = \sum_{i=1}^{\binom{M}{i}} \left( \prod_{j \in S_{i}} A_{p,j} \prod_{l \neq j} (1 - A_{p,l}) \right)$$  

(6)

$$A_{b}^{k} = \sum_{i=1}^{\binom{N}{k}} \left( \prod_{j \in S_{i}} A_{b,j} \prod_{l \neq j} (1 - A_{b,l}) \right)$$

In Eqn. (6), all possible ways of selecting $k$ primary paths and backup paths are considered in $A_{p}^{k}$ and $A_{b}^{k}$, respectively.

C. An Example

As shown in Fig. 1, a service request from $s$ to $d$ is 150 Mbps. Based on different SLA, we can have several different provisioning schemes using link-disjoint paths, three of which have no backup (Fig. 2) and three have backup (Fig. 3). However, with multipath provisioning, in the following situations (i.e., $M \geq 2$), we need to consider DDC, namely $3:0(m), 2:0(m), 2:1(m)$.

For illustration purposes, let us consider all link availabilities to be 0.999 (the path availability will be 0.999 $\times$ 0.999 $\times$ 0.999 $\simeq$ 0.997). Please note that, in our model, the links availabilities don’t have to be same. In Fig. 2(a), the availability of 150 Mbps is 0.997$^{3}$ $\simeq$ 0.99, as an extra benefit of multipath provisioning, the availability of 100 Mbps is...
0.997 + 0.997(1 − 0.997) ≃ 0.9999, and the availability of 50 Mbps is 1 − (1 − 0.997) 3 ≃ 0.9999999. Similarly, in Fig. 2(b), the availability of 150 Mbps is 0.997 2 ≃ 0.99, and availability of 75 Mbps is 1 − (1 − 0.997) 2 ≃ 0.999999. In Fig. 3(a), the connection availability is 0.997 2 + 2 × 0.997 2 − (1 − 0.997) ≃ 0.9999, in Fig. 3(b), the connection availability is 0.997 + 0.997 − (1 − 0.997) ≃ 0.999999, the availability of Fig. 3(c) is 0.997 + (1 − 0.997) × (0.997 × 0.997 + 0.997 × (1 − 0.997) + (1 − 0.997) × 0.997) ≃ 0.9999999. Table I summarizes the connection availabilities of different provisioning schemes.

With M : N (m) provisioning, we can meet much wider range of SLA (from 0.99 to 0.9999999) and much higher availability for degraded bandwidth (e.g., 100 Mbps, 50 Mbps in Fig. 2(a) and 75 Mbps in Fig. 2(b) and Fig. 3(a)). In the case of single primary-path provisioning, the availability will be 0.99 and 0.999999 only (Fig. 2(c) and Fig. 3(b)).

III. BANDWIDTH MIGRATION WITH LCAS

In traditional SONET/SDH networks, to change the capacity of an in-service channel, we need to tear down the connection, adjust capacity, and re-start traffic, which can have a considerable downtime. However, in next-generation SONET/SDH networks, in addition to VCAT, LCAS is defined as an adjustment scheme that hitherto increases or decreases the capacity of a container that is transported in a SONET/SDH network using VCAT. LCAS offers the ability to automatically adjust the bandwidth on-the-fly based on bandwidth on demand. In this section, we introduce the idea of bandwidth migration (BM), which can be implemented by the LCAS protocol that often co-exists with VCAT. The bandwidth, at the granularity of a single VCAT capacity, can be migrated within the same VCG without interrupting the existing connection. Since BM won’t impact the end-to-end full-service connection availability, it is helpful to reduce “bandwidth blocking” and new reserved backup capacity. We specify two types of bandwidth migration based on its purpose. We use \((s, d, B, SLA, DDC, h_i)\) to represent a connection request \(R\), where \(s\) and \(d\) are source and destination nodes, \(B\) is the requested bandwidth, \(DDC\) is differential-delay constraint (in milliseconds), \(SLA\) is Service-Level Agreement specified availability requirement, and \(h_i\) is request’s holding time. The capacity of a path can be classified as working, free, and backup [20]. We refer to a group of VCGs that have a joint service path as a Joint Multipath Provisioning Group (JMPG).

Type 1: Local bandwidth migration.

When a new connection request will be rejected because of “bandwidth blocking”, we first identify the links whose free capacity is less than their assigned capacity. These links are referred to as “bottleneck” links. If the number of “bottleneck” links is within a pre-defined threshold, for each bottleneck link, we find its migration connection (MC) by searching its JMPG, and then shift part of its working capacity to other paths among it’s MC. If all bottleneck links are migrated, it may enable us to accept the connection. It could be the case that there are multiple JMPG for a single bottleneck link of a connection, the first JMPG that is considered as the MC is used for bandwidth migration. We refer to this kind of local bandwidth migration as Type 1 BM. The algorithm of Type 1 BM is stated in Algorithm 1.

Figure 4 shows an example of Type-1 bandwidth migration with one bottleneck link. The paths’ working capacity, free capacity, and availability are listed beside the paths, as: (working, free, availability). Request \(R_1\) and \(R_2\) form a joint multipath provisioning group (JMPG). In Fig. 4(a), \(⟨s_1, d_1, 96, 0.99, 8, 1⟩\) is an existing connection, and due to bandwidth limitation on link \(a − b\), the new connection \(⟨s_2, d_2, 48, 0.99, 8, 1⟩\) can not be accepted although the connection availability can meet the SLA (0.99). This kind of blocking is referred to as bandwidth blocking. After bandwidth migration as shown in Fig. 4(b), 6 units (in the granularity of VCAT) of bandwidth from the path \(s_1\)-\(a\)-\(b\)-\(d_1\) are migrated to another path \(s_1\)-\(e\)-\(f\)-\(d_1\). So, the free capacity of path \(s_1\)-\(a\)-\(b\)-\(d_1\) will increase 6 units from 15 to 21. For request \(R_2\), which requests 48 units of bandwidth, 27 of them can be carried by another link-disjoint path \(s_2\)-\(c\)-\(d_2\), and 21 of them can be carried by paths \(s_2\)-\(a\)-\(b\)-\(d_2\). So the new connection request \(R_2\) can be accepted after migration.

Type 2: Global bandwidth migration.

![Fig. 2: Multipath provisioning without backup.](image)

![Fig. 3: Multipath provisioning with backup.](image)

**TABLE I: Connection Availability of Different Multipath Provisioning.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3 : 0 (m) Fig. 2(a)</td>
<td>0.99</td>
<td>100 Mbps: 0.9999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50 Mbps: 0.99999</td>
</tr>
<tr>
<td>2 : 0 (m) Fig. 2(b)</td>
<td>0.99</td>
<td>75 Mbps: 0.99999</td>
</tr>
<tr>
<td>1 : 0 (m) Fig. 2(c)</td>
<td>0.997</td>
<td>N/A</td>
</tr>
<tr>
<td>2 : 1 (m) Fig. 3(a)</td>
<td>0.99999</td>
<td>75 Mbps: 0.9999999</td>
</tr>
<tr>
<td>1 : 1 (m) Fig. 3(b)</td>
<td>0.99999</td>
<td>N/A</td>
</tr>
<tr>
<td>1 : 2 (m) Fig. 3(c)</td>
<td>0.99999</td>
<td>N/A</td>
</tr>
</tbody>
</table>

![Fig. 4: Type-1 bandwidth migration.](image)
Algorithm 1 Type-1 BM Algorithm

Input: Blocked request \( R_b = \{s, d, B, SLA, DDC, h_1\} \), \( P \): a set of paths meet DDC and SLA, \( A \): a set of assigned capacity for each path. \( F \): a set of current free capacity for all links. \( Q \): connection queue with sorted bandwidth from high to low. \( T_{BL} \): The threshold of maximum number of bottleneck links to do Type-1 BM, \( D_{bm1} \): search depth in \( Q \) for Type-1 BM.

Output: The \( P \) with enough free capacity to accept \( R_b \).

1: Find all bottleneck links \( l_1, l_2, \ldots, l_n \). The link \( l_i \) is considered as a bottleneck link, if its free capacity is less than its assigned capacity.
2: If the total number of bottleneck links is more than \( T_{BL} \), stop. No BM-1 done.
3: For a given bottleneck link \( l_i \), find its migration capacity (i.e., the difference of assignment and free capacity) \( \phi \). Find its migration connection \( MC^l_i \) from its JMPG using following criteria:
   1) \( MC^l_i \) has multiple paths and all of them are working path;
   2) \( MC^l_i \) is a JMPG of \( R_b \), and has only one joined path which includes link \( l_i \);
   3) The carried working capacity of \( MC^l_i \) in the path passing \( l_i \) is greater than \( \phi \);
   4) There is enough free capacity on other working paths of \( MC^l_i \) to jointly carry \( \phi \) (migrate to one or more other paths within same VCG).
4: Search all connections in \( Q \) until \( D_{bm1} \) is reached.
5: Repeat above Step 3-4 for all bottleneck links.
6: Reserve capacity and accept the request \( R_b \).

Some requests’ capacity cannot be evenly distributed among all primary paths at the moment when it arrives due to different available free capacity on each path. Since the backup capacity is reserved as the maximum carried capacity of primary paths, when the network state changes (e.g., connection departure), we can re-allocate bandwidth among primary paths to save the reserved bandwidth on backup path. This kind of migration is referred to as Type 2 BM. It’s stated in Algorithm 2.

Algorithm 2 Type-2 BM Algorithm

Input: \( Q \): connection queue containing requests with sorted bandwidth from high to low, \( D_{bm2} \): search depth in \( Q \) for Type-2 BM, and all connections’ information and network state.

Output: Balanced assignment of primary capacity for qualified connections.

1: Find Type-2 BM connection: search request \( R \) in \( Q \) using following criteria:
   1) There are at least three paths in \( R \).
   2) At least two are working paths, and at least one is a backup path.
   3) The difference among carried-capacity on working paths is more than 1.
   4) The available free capacity on each working path allows migration.
2: A connection that meets above conditions can do Type-2 BM.
3: Adjust carried capacity on working paths and re-allocate backup capacity on backup path.
4: Search \( Q \) until \( D_{bm2} \) is reached.

in Fig. 5(b) by re-allocating primary path capacities, so the backup capacity can be set as 24. So, we can save 4 units of capacity along backup path.

Type-1 BM can be performed when a new connection arrives and in case of “bandwidth blocking”. We find the “bottleneck” links first and then search its JMPG to migrate bandwidth among its primary paths. Type-2 BM can be performed networkwide when the network state changes (e.g., when a large connection departs). The network operator can also use some other criteria to trigger Type-2 BM.

IV. GENERAL PROBLEM STATEMENT AND PROPOSED ALGORITHM

In this section, we define the notations and then formally state the adaptive reliable multipath provisioning problem with differentiated service availability-guaranteed \( M : N(m) \) protection schemes. The following inputs are given:

1) \( G(V, E, A, D, C, W) \) is a weighted, directed graph which represents a network topology, where \( V \) is the set of nodes, \( E \) is the set of bidirectional fibers (referred to as links), \( A : E \rightarrow (0, 1] \) is the availability of each link (where \( (0, 1] \) denotes the set of positive real numbers between 0 and 1), \( D : E \rightarrow R^+ \) specifies link distance in km (where \( R^+ \) denotes positive real numbers), \( C : E \rightarrow R^+ \) specifies link cost (where \( R^+ \) denotes positive real numbers), and \( W : E \rightarrow Z^+ \) specifies the number of wavelengths on each link (where \( Z^+ \) denotes the set of positive integers). Also, we assume that each node in the graph has full wavelength-conversion capability.

2) Network state. \( NS = \{ f_e, w_e, b_e \} , e \in E \), where \( f_e \), \( w_e \), and \( b_e \) are free, working, and backup capacity of link \( e \), respectively, in units of \( g \), where \( g \) is the granularity of VCAT. For each link, we maintain its state as \( NS \).

Figure 5 shows an example of Type-2 bandwidth migration. In Fig. 5(a), the primary paths’ capacities are not evenly allocated due to free capacity restriction on path \( P_{w2} \), and the backup capacity is reserved as 28. After the network state changes, we can perform bandwidth migration as shown.
3) \( R = (s, d, B, SLA, DDC, h_s) \) is an incoming connection request, where \( s \) is the source, \( d \) is the destination, \( B \) is its requested bandwidth which is represented as multiples of VCAT granularity \( g \) (where \( g \) can be low-order \( VT1.5s/VT2s \) or high-order \( STS - 1s/STS - 3cs \)), \( DDC \) is the constraint of differential delay among all diversely-routed sub-connections, \( SLA \) is the availability requirement, and \( h \) is the holding time of connection \( R \). Under a dynamic traffic pattern, multiple paths which have been set up between a node pair to satisfy a connection request are taken down after a period of time \( h \) (the connection’s holding time). In multipath provisioning, the requested bandwidth of connection \( R \) could be various from sub-wavelength capacity (represented as multiples of \( g \)) to a full wavelength channel capacity.

Our goal is to develop algorithms to provision differentiated services, i.e., \( M : N(m) \) provisioning. The multipath provisioning is adaptive to connection \( R \), such that its SLA requirement is met, \( DDC \) requirement is met (if there are multiple primary paths), and the backup-resource investment is minimized. With the support of LCAS protocol, the capacity of service paths can be adjusted dynamically, in hitless manner with granularity of \( g \), among primary paths so that the blocking probability of \( R \) and backup resources is minimized.

We propose heuristic algorithms for availability-guaranteed adaptive multipath provisioning with and without LCAS support. They are referred to as Multi-Path Provisioning with Availability Guarantee (MPAG) (without BM) and Multi-Path Provisioning with Availability Guarantee using Bandwidth Migration (MPAG-BM). Both are identical. MPAG-BM is MPAG with bandwidth migration.

A. Calculate Meet-SLA-Path and Link-disjoint Paths

In a special case of multipath provisioning, single path provisioning is possible. In this case, we do not need to consider the DDC, and look for the path which meets the SLA requirement and also takes the cost (i.e., shortest distance in this study) into consideration. We refer to this as Meet-SLA-Path (MSP). We define MSP for request \( R \) the cheapest path between \( s - d \) pair among a small number of paths (i.e. \( MaxSP \)) that meets the SLA requirements of \( R \). A request’s MSP is calculated adaptively based on its SLA.

To benefit from the resource efficiency of multipath provisioning, we first try multiple link-disjoint paths without backup. The \( M \) link-disjoint paths are found with DDC-awareness. We use the differential-delay-constrained \( K \) (\( K \geq 2 \)) link-disjoint path algorithm (DDCKDP) [8] to find the \( K \) shortest link-disjoint paths which satisfy DDC.

B. Multi-Path Provisioning with Availability Guarantee using Bandwidth Migration (MPAG-BM)

In this study, we restrict the maximum number of link-disjoint paths between a node pair to \( MaxDP \). The maximum number of shortest paths to calculate MSP is \( MaxSP \). When a connection arrives, we find the appropriate \( M : N(m) \) scheme using the following order:

1) multipath without backup \( (M : 0(m), M \geq 2) \), start \( M \) with the sequence \( MaxDP, (MaxDP - 1), \cdots , 2; \)

2) single-path without backup \( (1 : 0(m)) \), calculate MSP (using \( MaxSP \) as introduced in Section IV-A);

3) multipath with backup \( (M : N(m), N \geq 1) \), start \( N \) with the sequence \( 1, 2, \cdots , (MaxDP - 1) \), while \( M = MaxDP - N \).

Given \( M \) link-disjoint paths, we first look for multipath without backup \( (M : 0(m), (M - 1) : 0(m), \cdots , 2 : 0(m)) \). If the SLA cannot be satisfied with \( M : 0(m) \), we calculate MSP, as introduced above. If the MSP cannot be found, we look for multiple paths with backup provisioning (e.g., \( (M - 1) : 1(m), (M - 2) : 2(m), \cdots , 1 : (M - 1)(m) \)). Then, we assign the path type and allocate the bandwidth.

If a connection request \( R = (s, d, B, SLA, DDC, h_s) \) is accepted, it means the following: (1) we found DDC-aware \( M : N(m) \) provisioning with \( AR \geq SLA \); and (2) the sum of all primary paths’ carried capacity is \( B \), each backup paths’ reserved capacity is set as \( Max\{C_i\} \), where \( C_i \) is the carried capacity of the primary paths. Similarly, when a connection request is blocked, it means that we cannot find “good” paths (availability blocking) or there is not enough free capacity on the paths (bandwidth blocking).

We present the flow chart of Algorithm 3 in Figure 6.

![Fig. 6: Flow chart of MPAG-BM.](image)

The detailed steps of MPAG-BM are stated in Algorithm 3.

C. Computational Complexity Analysis for MPAG-BM

The following refers to the algorithm description below. The computational complexity for Step 1 is \( O(|MaxDP|(|E| + |V| \log |V|)) \), Step 2 is \( O(|MaxDP||E|) \), Step 3 is \( O(|V|^2 + |MaxSP||V|) \) using Yen’s algorithm [21], Step 4 is \( O(|MaxDP|^2|E|) \) considering all possible connection availabilities, Step 5 is \( O(|E|) \), Step 6 is \( O(|E|^2) \), Step 7 is \( O(|E|) \), and Step 8 is \( O(|E|^2) \). Since \( MaxDP \) and \( MaxSP \) are
Algorithm 3 MPAG-BM Algorithm

**Input:** Graph \(G(V, E, A, D, C, W)\) (refer to Section IV for the meaning of each set), Network State: \(NS = \{f_e, w_e, b_e\}, e \in E\); a connection request: \(R = \{s, d, B, SLA, DDC, h_t\}\).

**Output:** Link-disjoint \(M (1 \leq M \leq MaxDP)\) primary paths and \(N (0 \leq N \leq MaxDP - 1)\) backup paths to provide bandwidth \(B\) satisfying connection request \(R\)'s SLA and DDC.

1: Compute DDC-aware \(M (M \leq MaxDP)\) link-disjoint paths between s-d pair [8], using the following link-cost function:

\[
\text{Cost}(e) = \begin{cases} 
+\infty & f_e \leq 2B \gamma \\
 d_e & f_e > 2B \gamma 
\end{cases}
\]

where \(d_e\) is the distance of link \(e\), and \(\gamma\) is the fraction of free capacity of link \(e\) divided by \(B\) (we use \(\gamma = 0.3\) in our numerical examples based on average nodal degree of sample topology).

2: Try no-backup multipath provisioning in the following sequence \(-\ 1: M : 0(m), (M - 1) : 0(m), \ldots, 2 : 0(m)\), using Eqn. (5) to calculate \(A_R\). If \(A_R \geq SLA\), exit and go to Step 5; if no success, continue to the next step.

3: Compute \(MaxSP\) shortest paths using Yen's K-shortest-path algorithm [21] but set availability as the cost metric (take logarithm to use Multiplication-to-Summation) [12]. Found \(MaxSP\) paths which meet SLA and then select the one with smallest distance as MSP. If MSP is found, go to Step 5; otherwise, continue to the next step.

4: Try multipath with backup provisioning following the sequence with resource overbuild from low to high (e.g., for \(M=4\), try \(3 : 1(m)\), \(2 : 1(m)\), \(2 : 2(m)\), \(1 : 1(m)\), \(1 : 2(m)\), \(1 : 3(m)\), exit if \(A_R \geq SLA\). For each case, use Eqn. (5) to calculate \(A_R\). If found multipath with backup which meet SLA, continue to the next step. Otherwise, block the request ("availability blocking"), and return NULL.

5: Assign requested capacity on multiple paths or single path. If there are multiple primary paths, allocate bandwidth on multiple paths with minimum difference. If there are backup paths, reserve backup path capacity as the maximum carried capacity of primary paths. If provisioning with single path, just allocate \(B\) on it. If successfully assigned bandwidth, go to Step 7; otherwise, continue to the next step.

6: Perform Type-1 bandwidth migration. First find “bottleneck” links, and then search JMPG for each bottleneck link. If success, go back to Step 5 to re-assign capacity. If failed, undo Type-1 BM and block the request ("bandwidth blocking"), and return NULL.

7: Accept the connection request, allocate capacity, and update network state.

8: Check Type-2 BM flag (the flag can be set following an event, e.g., large request departure), and perform Type-2 BM if the flag is set. If success, update the network state, and clear the flag.

small values, the computation complexity of Algorithm 3 is \(O(|E|^2 + |V|^2) = O(|E|^2)\).

### D. Multi-Path Provisioning with Availability Guarantee (MPAG)

The MPAG algorithm is similar to MPAG-BM except that there are no bandwidth migration (BM) steps (Steps 6 and 8 in Algorithm 3). MPAG is used in case of no LCAS support. Removing Steps 6 and 8 of Algorithm 3, the computation complexity of MPAG is \(O(|E| + |V|^2)\) for small MaxDP and MaxSP values.

### V. Illustrative Numerical Examples

We simulated a dynamic network environment, with uniform distributed connection requests on two typical U.S. nationwide network topologies. We only show the results for one topology in Fig. 7 since the results on the other topology are very similar. There are 16 wavelengths per unidirectional fiber in each topology. The connection arrival process is Poisson and the connection holding time follows a negative exponential distribution. The capacity of each wavelength is OC-192 (\(\approx 10\) Gbps). The connection requests follow the bandwidth distribution \(100M : 150M : 600M : 1G : 2.5G : 5G : 10G = 50 : 20 : 10 : 10 : 4 : 2 : 1\) (which is a typical bandwidth distribution in a practical network). The VCAT granularity is STS-1. Load (in Erlang) is defined as the product of the connection-arrival rate, the average connection-holding time, and a connection's average bandwidth normalized in the unit of OC-192. For our illustrative examples, the connection availability requirements are chosen uniformly from the set \([0.99, 0.999, 0.9999]\). The availabilities of links are assumed to be uniformly distributed over \([0.999, 0.9999, 0.99999]\).

We simulated MPAG-BM and MPAG with three different DDC values \([8ms, 18ms, 50ms]\), and compared their performance with single-path provisioning (using a link-disjoint backup path if needed). The single path is selected using the MSP method in Step 3 of Algorithm 3. If MSP can not be found, we just use two shortest link-disjoint paths and set one as the backup path. We set \(T_{BL}\) as 3, \(D_{bm1}\) as 2000, \(D_{bm2}\) as 2000, \(MaxDP\) as 4 and \(MaxSP\) as 5 to reduce the complexity. From experiments, we found that the results do not vary between any node pair if DDC is 8ms or more, the 50ms DDC is selected because it’s comparable to the case without DDC consideration (DDC is \(\infty\)).
not improve much for higher values of these parameters.

A. Bandwidth Blocking Ratio

The bandwidth blocking ratio (BBR), which is defined as the total amount of bandwidth blocked over the total amount of bandwidth requested, is used to evaluate the effectiveness of the algorithms. Since connections have different bandwidth requirements, the connection blocking probability, defined as the percentage of connections blocked, does not accurately reflect the effectiveness of the algorithm. We further breakdown the BBR as multipath bandwidth blocking ratio (MBBR), availability blocking ratio (AB-BBR), and bandwidth blocking ratio (BB-BBR).

Figure 9 compares the BBR with BM and with different DDC values. With bandwidth migration, we can get nearly a 3 percent lower BBR when network load is above 200 Erlang. From simulation, we show that, in different network loads, DDC and topologies, we can carry more requested bandwidth simultaneously with BM. Another important advantage of using BM is that we can accept a higher priority mission-critical connection request without dropping existing connections. Figure 9 also compares the MPAG BBR with single path provisioning. The BBR of MPAG is obviously lower than that of single path provisioning scheme with different DDC values.

1) Multipath Bandwidth Blocking: Since connections are provisioned with single or multiple primary paths, to evaluate the performance of multipath provisioning ($M : N(m), M \geq 2$), we use the multipath bandwidth blocking ratio (MBBR). MBBR is defined as the amount of bandwidth blocked by $M : N(m) (M \geq 2)$ provisioning over the total amount of bandwidth requested. Figure 10 compares the MBBR with different DDC with and without BM. From this figure, we can conclude that 1) The MBBR is much lower (MBBR is less than 8 percent) than overall BBR. This further proves...
that, with multipath provisioning, the possibility of rejection a connection request is lower; 2) The improvements of MBBR with BM are more obvious since BM is only applicable in \( M : N(m) (M \geq 2) \) connections; 3) With different DDC, the MBBR values are different. Contrary to what we expected, the MBBR is higher for higher DDC in this case. Generally, with higher DDC, we can find more link-disjoint paths between a source-destination pair [8], and thus more options for \( M : N(m) \) provisioning. However, more \( 1 : N(m) \) \((N \geq 2) \) connections (they are not counted for MBBR in our study) are also established to meet high SLA requests. So, given the same capacity setup, the bandwidth blocking for \( M : N(m) \) \((M \geq 2) \) is higher.

2) Availability Blocking vs. Bandwidth Blocking: The connection blocking is classified as two situations: availability blocking (AB-BBR) and bandwidth blocking (BB-BBR). The availability blocking means there is no path, that meets SLA, can be found between source and destination. Bandwidth blocking means that a path that meets SLA can be found, but it fails on bandwidth allocation (i.e., there is no enough available bandwidth on the path). Classifying these two metrics is an effective way to evaluate how good the algorithm is in selecting a suitable path. If an algorithm is very adaptive, the chance of rejecting a request due to availability blocking should be very low. Figure 11 shows these two blocking situations with a typical 8 ms DDC value. We can see that the dominant part of connection blocking is due to the bandwidth limitation. The AB-BBR is very small relative to the overall blocking ratio. It means that our adaptive algorithm is very flexible in finding qualified path(s) to meet request’s SLA. The opportunity of blocking due to lacking of “good” path(s) is very low since the algorithm is very robust in selecting paths to meet wide range of SLA requirements. The blocking due to bandwidth limitation is unavoidable when network load is high. From Fig. 11, we can see that the improvement of BBR with bandwidth migration comes from BB-BBR.

B. Resource Overbuild

To quantify the extra resources needed for providing protection, we employ a performance metric called resource overbuild (RO) [3]. Resource overbuild is defined as the amount of backup capacity (weighted by the average hop distance of backup VCG members) over the amount of working capacity (weighted by the average hop distance of working VCG members). Figure 12 compares the RO with BM. We show that, at all loads and DDC values, we can also get slightly lower RO with BM. The purpose of Type-2 BM is to reduce the backup bandwidth by balancing working bandwidth on multiple paths. Since we also choose single paths in provisioning, the overall improvement of RO with BM is small. However, the results show that the RO is significantly lower in all cases with multipath provisioning than that of the single path provisioning.

C. Bandwidth Migration Overhead and Delay

The bandwidth migration needs LCAS [2] support. The efficiency of bandwidth utilization comes for a price, namely, management overhead and time delay due to LCAS protocol. The migration needs LCAS protocol to remove one or more VCAT members from one path and add them back to another path under control of a management system. The dominant part of delay is to add the members back to the VCAT group since the dropping process is straightforward and the adding process can be started simultaneously.

In LCAS protocol design, synchronization of changes in the capacity of the transmitter (So) and the receiver (Sk) is achieved by a control packet. Each control packet describes the state of the link during the next control packet. Changes are sent in advance, so that the receiver can switch to the new configuration as soon as it arrives. The control packet consists of fields dedicated to a specific function. The Multi Frame Indicator field (MFI): used to determine the differential delay between members of the same VCG. Sequence Indicator field (SQ): contains the sequence number assigned to a specific number. Control field (CTRL): used to transfer information from So to Sk and synchronize the Sk with the So. Group Identification bit (GID): used for identification of the VCG. Member status field (MST): information from Sk to So about the status of all members of the same VCG. It reports the member status of OK or FAIL from Sk to So. Re-Sequence Acknowledge bit (RS-Ack): the toggling of the RS-Ack bit will validate the MST in the preceding multiframe. When a member is added it will always be assigned a sequence
number greater than the currently highest sequence number that has EOS in the CTRL code. Following an ADD command the first member to respond with MST=OK is allocated the next highest sequence number and changes its CTRL code to EOS coinciding with the currently highest member changing its CTRL code to NORM. When the CTRL=ADD is sent to initiate the addition of a new member, it will be sent continuously until the MST=OK is received. In case more than one member is being added, then the allocation of sequence indicators is arbitrary provided they are the next x sequence numbers after the current highest sequence number.

The high order mode operation is required in our algorithm, the delay introduced by migration could be calculated as follows: 1) LCAS management software drops the members directly and changes the member state from IDLE to ACTIVE ($t_1$); 2) Source generates and sends an ADD command and transmit it ($t_2$); 3) Receiver sends acknowledge via MST bits ($t_3$). For HO mode, this can take up to 32 H4 bytes; 4) Source changes CTRL word from ADD to EOS and send it ($t_4$); 5) Waiting for RS-ACK toggle reception from destination to source ($t_5$). The $t_1$ is 2 ms, $t_2$ is 4 ms, which contains 2 ms to generate and 2 ms to transmit it, $t_3$ is $32 \times 16 \times 125\mu s = 64ms$, $t_4$ also contains 2 ms to generate and 2 ms to send, and $t_5$ is 2 ms. So the total delay time is $t_1 + t_2 + t_3 + t_4 + t_5 = 2 + 4 + 64 + 4 + 2 = 76ms$. Any usage of LCAS for “bandwidth on demand” causes some delay (up to 76ms in high order mode) following above protocol analysis. In our work, the delay of BM only means fluctuation, the service always keeps on (it drops part of its bandwidth and then add back 76ms later). Since such fluctuation does not happen frequently (only one or two times for a service connection), the impact is small.

---

**Fig. 11:** Availability Blocking vs. Bandwidth Blocking.

**Fig. 12:** Comparison of Resource Overbuild with BM.
We proposed the availability-analysis model for generic \( M : N (m) \) multipath provisioning for next-generation high-capacity backbone networks which support VCAT. We developed availability-aware and DDC-aware adaptive heuristic algorithms for \( M : N (m) \) provisioning. By categorizing the blocking situation, with extensive simulation, we show the robustness of our algorithm in selecting suitable paths between source and destination. We introduced two kinds of bandwidth migration methods to optimize resource usage, and also evaluate the signaling overhead of BM. We showed a way to exploit LCAS protocol for better resource usage. The reason for the small improvement of BM is that we also choose single MSP path, and a considerable number of connections are delivered using single MSP path. We viewed the connection bandwidth as a whole to satisfy its availability. Also, as an extra benefit, the operator can provide much higher availability for degraded service bandwidth. The DDC value impacts the performance of our algorithms. MPAG-BM performs better than MPAG for different network loads and DDC values. The BM can be used to avoid rejecting a high priority request by migration of the bandwidth of existing connections. However, as latency becomes critical to next-generation applications, the BM should be used carefully considering its signaling overhead and delay. For most service connections, the BM can be used for better efficiency since it uses existing LCAS protocol which most often co-exists with VCAT.

VI. CONCLUSION

We proposed the availability-analysis model for generic \( M : N (m) \) multipath provisioning for next-generation high-capacity backbone networks which support VCAT. We developed availability-aware and DDC-aware adaptive heuristic algorithms for \( M : N (m) \) provisioning. By categorizing the blocking situation, with extensive simulation, we show the robustness of our algorithm in selecting suitable paths between source and destination. We introduced two kinds of bandwidth migration methods to optimize resource usage, and also evaluate the signaling overhead of BM. We showed a way to exploit LCAS protocol for better resource usage. The reason for the small improvement of BM is that we also choose single MSP path, and a considerable number of connections are delivered using single MSP path. We viewed the connection bandwidth as a whole to satisfy its availability. Also, as an extra benefit, the operator can provide much higher availability for degraded service bandwidth. The DDC value impacts the performance of our algorithms. MPAG-BM performs better than MPAG for different network loads and DDC values. The BM can be used to avoid rejecting a high priority request by migration of the bandwidth of existing connections. However, as latency becomes critical to next-generation applications, the BM should be used carefully considering its signaling overhead and delay. For most service connections, the BM can be used for better efficiency since it uses existing LCAS protocol which most often co-exists with VCAT.

REFERENCES