Adaptive Reliable Multi-Path Provisioning in WDM Mesh Networks

Sheng Huang
Digital Enterprise Group
Intel Corporation
Email: sheng.huang@intel.com

Biswanath Mukherjee
Department of Computer Science
University of California, Davis
Email: mukherje@cs.ucdavis.edu

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). With VCAT, a connection request can be split, diversely routed, and inversely multiplexed on to multiple paths, a feature that has many advantages over conventional single-path provisioning, such as improved reliability, load balancing, etc. However, these diversely routed traffic components are subject to the differential delay constraints (DDC), which could be limited by the destination node’s delay-compensation capability, network operations, administration, maintenance, and provisioning (OAM&P), or connection quality-of-service (QoS).

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 may have a fraction of the bandwidth of the connection, and \( (m) \) in this notation denotes “multipath”. We present the flexibility and benefits of multipath provisioning and develop an analytical model to analyze the connection availability under \( M \times N(m) \) provisioning schemes. 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. We develop two adaptive heuristic algorithms 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 improve the performance.

I. INTRODUCTION

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 quality-of-service (QoS) metric describing reliability. Network operators can provide differentiated service connections based on a service-level agreement (SLA), a contract which states the guaranteed availability. Recent advances and deployment of next-generation SONET/SDH technology, especially virtual concatenation (VCAT), enables multipath provisioning to achieve more resource efficiency and flexibility for new data services. 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). Used with VCAT, link-capacity adjustment scheme (LCAS) provides management ability of VCG members for bandwidth allocation [1]. However, these diversely-routed group members are subject to the differential delay constraints (DDC) at the destination node, which can impact the QoS of the connection if not considered in the routing [2].

Most prior works [3]–[5] employ single-path provisioning to satisfy the availability requirement of a connection or they did not consider DDC of multipath provisioning [6], [7]. The author in [8] studied the problem of minimizing the differential delay in a virtually-concatenated Ethernet-over-SONET (EoS) system by suitable path selection. The dominant part of DDC is propagation delay, which is approximately 5 \( \mu s/km \) [2]. The effective multipath bandwidth concept was proposed in [6] with the assumption that degraded services are acceptable. In this study, 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, DDC-aware, and resource-efficient-aware connection service. We introduce \( M : N(m) \) notation to represent generic multipath provisioning and then propose the availability analysis model. The backup paths are shared among same VCG members. We exploit the LCAS function and propose bandwidth migration method to optimize resource usage without impacting the connection availability.

The paper is organized as follows. Section II provides availability-guaranteed multipath provisioning schemes and analysis model. Section III introduces the bandwidth migration method. Section IV proposes two multipath provisioning heuristic algorithms. In Section V, we compare and discuss the simulation results. Section VI concludes this study.

II. MULTIPATH PROVISIONING AND CONNECTION AVAILABILITY ANALYSIS

The availability of 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 [6], where the availability of an individual component is given by \( \frac{MTTF}{(MTTF + MTTR)} \) [3], [5], [6]. In this study, without loss of generality, we refer to the individual components of a path to be links only. If the SLA cannot be satisfied with only primary paths, we use backup paths and allocate to the

This work is supported in part by NSF grant CNS-0520190. Sheng Huang is a component design engineer at Intel Folsom CA. He is also working towards his PhD at UC Davis.
backup paths the maximum carried capacity of any primary path, so the backup path can protect any primary path in case of failure. There could be multiple backup paths to satisfy high availability requirements.

In this study, we only consider link-disjoint paths for intra-VCG shared protection. 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). In general, we use the \( M : N (m) \) protection scheme for a connection, where \( M \) is the number of primary paths, \( N \) is the number of backup paths [1], and \( (m) \) refers to multipath provisioning, to represent these provisioning schemes (\( N = 0 \) means there is no backup path). However, with multipath provisioning, in the following situations (i.e., \( M \geq 2 \)), we need to consider DDC, namely \( 3 : 0 (m) \), \( 2 : 0 (m) \), \( 2 : 1 (m) \).

![Fig. 1. Multipath provisioning example.](image)

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

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

For illustration purposes, let us consider all link availabilities to be 0.999 (the path availability will be 0.999 * 0.999 * 0.999 ≈ 0.997). Please note that, in our study, the links availabilities don’t have to be same. In Fig. 2(a), the availability of 150 Mbps is 0.997^3 ≈ 0.99, as an extra benefit of multipath provisioning, the availability of 100 Mbps is 0.997^3 + 3 * 0.997^2 * (1 − 0.997) ≈ 0.9999, and the availability of 50 Mbps is \( 1 − (1 − 0.997)^3 \approx 0.999999 \). Similarly, in Fig. 2(b), the availability of 150 Mbps is 0.997^2 ≈ 0.99, and availability of 75 Mbps is \( 1 − (1 − 0.997)^2 \approx 0.9999 \). 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 \( 0.997 + 0.997 * (1 − 0.997) \approx 0.999999 \). The availability of Fig. 3(c) is \( 0.997 + (1 − 0.997) * (0.997 * 0.997 + 0.997 * (1 − 0.997) + (1 − 0.997) * 0.997) \approx 0.99999999 \). 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 0.99999 only (Fig. 2(c) and Fig. 3(b)).

In the following, we give a general availability analysis model for \( M : N (m) \) provisioning. We first select “good” link-disjoint paths which meet DDC and SLA, and then allocate bandwidth for every primary path and backup path. The bandwidth allocation is not included for the connection availability analysis because the goal is to meet the availability of full requested bandwidth. The backup paths are for connection protection and they can be shared by multiple primary paths within same VCG. The notations we use are:

- \( A_t \): connection availability;
- \( M \): number of primary paths;
- \( N \): number of backup paths;
- \( A_{pi} \): availability of primary path \( i \);
- \( A_{bi} \): availability of backup path \( i \);
- \( A_p^k \): availability of \( k \) primary paths; \( k \) paths available and \( M − k \) paths unavailable);
- \( A_b^k \): availability of \( k \) backup paths; \( k \) backup paths available and \( N − k \) backup paths unavailable).

The connection availability \( A_t \) can be represented as:

\[
A_t = \begin{cases}
A_p^M + \sum_{i=1}^{N} \left( A_p^{M-i} * \sum_{j=1}^{i} A_b^j \right) & \text{if } N \leq M; \\
A_p^M + \sum_{i=1}^{M} \left( A_p^{M-i} * \sum_{j=1}^{i} A_b^j \right) & \text{if } N > M.
\end{cases}
\]

where:

\[
A_p^k = \sum_{i=1}^{M-k} \left( \prod_{j=1}^{i+k} A_{pj} * \prod_{l \neq j}^{1} (1 - A_{pj}) \right)
\]

\[
A_b^k = \sum_{i=1}^{N-k} \left( \prod_{j=1}^{i+k} A_{bi} * \prod_{l \neq j}^{1} (1 - A_{bi}) \right)
\]

In the above formula, all the possible ways of selecting \( k \) primary paths and backup paths are considered in \( A_p^k \) and \( A_b^k \), respectively. When a connection arrives, we first find \( M : N (m) \) scheme following the priority from high-to-low as (1) multipath without backup \((M : 0 (m), M \geq 2)\); (2) single-path without backup \((1 : 0 (m))\); and (3) multipath with backup \((M : N (m), N \geq 1)\). Then, we assign the path type and allocate the bandwidth. If a connection request \((s, d, B, SLA)\), is accepted, it means the following: (1) we found DDC-aware \( M : N (m) \) provisioning with \( A_t \geq SLA \); and (2) the sum of all primary paths’ carried capacity is \( B \), the backup paths’ reserved capacity is set as \( \text{Max}(C_i) \), 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 or there is not enough free capacity on the paths. We refer to the blocked cases as “availability blocking” and “bandwidth blocking”.

III. BANDWIDTH MIGRATION WITH LCAS

In this section, we introduce the idea of bandwidth migration (BM), which can be implemented by LCAS protocol which is often co-existing with VCAT. The bandwidth, as the granularity of single VCAT capacity, can be migrated within the same VCG. Since BM won’t impact the end-to-end connection availability, it is helpful to reduce “bandwidth blocking” and reserved backup capacity. We use \( (s, d, B, SLA) \) to represent a connection \( R \). We specify two types of BM.

Type 1: Local bandwidth migration among primary paths. When a new connection arrives, in case of “bandwidth blocking”, we can migrate part of the “bottleneck” link’s capacity within an existing multipath provisioning group. This may enable us to accept the connection.

![Fig. 4. Type 1 bandwidth migration.](image)

Figure 4 shows an example of type 1 bandwidth migration. The paths’ working capacity, free capacity, and availability are listed beside the paths. In Fig. 4(a), \( (s_1, d_1, 96, 0.99) \) is an existing connection, and due to bandwidth limitation on link \( a-b \), the new connection \( (s_2, d_2, 48, 0.99) \) cannot be accepted although the connection availability can meet the SLA (0.99). After bandwidth migration as shown in Fig. 4(b), the new connection request can be accepted.

Type 2: Global bandwidth migration for lower backup capacity. Some request B’s 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 network state changes (e.g., connection departure), we can re-allocate bandwidth among primary paths so we can save the reserved bandwidth on backup path.

The network is represented as \( G(V, E, A, C, \lambda, g) \), \( A : E \rightarrow (0, 1) \) specifying link availability, \( C : E \rightarrow Z^+ \) specifying link cost, and \( g \) is the granularity of VCAT. For each link, we maintain state as \( NS = \{ f_e, w_e, b_e \} \), where \( f_e \), \( w_e \) and \( b_e \) are free, working and backup capacity of link \( e \) in unit of \( g \). Connection request is represented as \( (s, d, B, DDC, SLA, ht) \), where \( ht \) is connection holding time. We state MPAG-BM in Algorithm 1.
Algorithm 1 MPAG-BM Algorithm

Input: Graph $G (V, E, A, C, \lambda, g)$, Network State: $NS = \{f_e, w_e, b_e\}$, $e \in E$; a connection request: $R (s, d, B, DDC, SLA, ht)$.

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 $M$ (upbounded as $MaxDP = 4$) link-disjoint paths between s-d pair, using the following link-cost function:

$$Cost (e) = \begin{cases} +\infty & f_e \leq \gamma \ast B \\ d_e & f_e > \gamma \ast B \end{cases}$$

where $d_e$ is distance of link $e$, and $\gamma$ is fraction of free capacity of link $e$ over $B$ (we use $\gamma = 0.3$ in our numerical examples).

2: Try no-backup multipath provisioning in the following sequence – $M : 0 (m), (M - 1) : 0 (m), ..., 2 : 0 (m)$, using Equation (1) to calculate $A_t$. If $A_t \geq SLA$, 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 [9] but set availability as the cost metric (take logarithm to use Multiplication-to-Summation) [5]. 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 $A_t \geq SLA$). For each case, use Eqn. (1) to calculate $A_t$. 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. If success, go back to Step 5 to re-assign capacity. If failed, un-do 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.

B. Multi-Path Provisioning with Availability Guarantee (MPAG)

The MPAG algorithm is similar to MPAG-BM except that there are no BM steps (Steps 6 and 8 in Algorithm 1). MPAG is used in case of no LCAS support.

V. ILLUSTRATIVE NUMERICAL EXAMPLES

We simulated a dynamic network environment on the topology shown in Fig. 6, with 16 wavelengths per fiber. 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 100$M : 150M : 600M : 1G : 2.5G : 5G : 10G = 50 : 20 : 10 : 10 : 4 : 2 : 1$ (which is a typical bandwidth distribution in a practical network). The VCAT granularity is STS-1. The request arrivals are uniformly distributed among all node pairs. 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 $\{0.999, 0.9999, 0.9999\}$.

We simulated MPAG-BM and MPAG with two different DDC values $\{8ms, 18ms\}$ and compared their performance with single-path provisioning (using a link-disjoint dedicated backup if needed). The single path is selected using the MSP method in Step 3 of Algorithm 1. If MSP cannot be found, we use two shortest link-disjoint paths and set one as backup path. We set $MaxDP$ as 4 and $MaxSP$ as 5 to reduce the complexity. The bandwidth blocking ratio (BBR) and resource overbuild (RO) [11] are used as performance metrics.

Figure 7(c) compares the BBR of the three algorithms with network load. Figure 8(c) compares the RO. Figures 7(a) and 8(a) compare the MPAG performance for the two DDC values, while Figs. 7(b) and 8(b) compare the MPAG-BM

1 The $8ms$ DDC is the typical industry value, and the $18ms$ DDC is selected because, in the sample topology, we can get two or more link-disjoint paths between any node pair if DDC is up to $18ms$. 
performance with the two DDC values. Our results show that both MPAG and MPAG-BM perform significantly better than single-path provisioning (in terms of lower BBR and lower RO). Also, MPAG-BM performs better than MPAG, as expected. In addition, a larger DDC allows more flexibility in connection provisioning, leading to improved network performance (lower BBR and lower RO).

VI. CONCLUSION

We proposed the availability analysis model for generic \( M : N \) multipath provisioning for next-generation high-capacity backbone networks which support VCAT. We developed two availability-aware and DDC-aware adaptive heuristic algorithms for \( M : N \) provisioning and introduced bandwidth migration methods to optimize resource. We viewed the connection bandwidth as a whole to satisfy its availability. Our results show that both MPAG and MPAG-BM perform significantly better than single-path provisioning. Also, as an extra benefit, the operator can provide much higher availability for degraded service bandwidth. The DDC value impacts the performance in both algorithms. With bandwidth migration, MPAG-BM performs better than MPAG with different network load and DDC.

REFERENCES