

Connection availability analysis of span-restorable mesh networks

Ling Zhou · Marcel Held · Urs Sennhauser

Received: 20 December 2006 / Accepted: 9 May 2007 / Published online: 17 July 2007
© Springer Science+Business Media, LLC 2007

Abstract Dual-span failures are the key factor of the system unavailability in a mesh-restorable network with full restorability of single-span failures. Availability analysis based on reliability block diagrams is not suitable to describe failures of mesh-restorable networks with widely distributed and interdependent spare capacities. Therefore, a new concept of restoration-aware connection availability is proposed to facilitate the analysis. Specific models of span-oriented schemes are built and analyzed. By using the proposed computation method and presuming dual-span failures to be the only failure mode, we can exactly calculate the average connection unavailability with an arbitrary allocation rule for spare capacity and no knowledge of any restoration details, or the unavailability of a specific connection with known restoration details. Network performance with respect to connection unavailability, traffic loss, spare capacity consumption, and dual failure restorability is investigated in a case study for an optical span-restorable long-haul network.

Keywords Connection unavailability · Dual failure restorability · Span restoration · p -Cycles

1 Introduction

An optical network is commonly designed to be 100% restorable for any single-span failure, represented by $R_1 = 1$. However, node failures and multiple failures are not

taken into account. Then a network may not be fully restorable anymore. Considering rates of cable cuts, frequent span maintenance or upgrade operations, and shared risk span groups [1], in which two logically distinct spans are routed through a common duct, span failures can occur much more frequently compared with node failures in practice. According to the concept of “most likely paths to failure” [2], the simplest combinations of elemental failures that lead to an outage state in any system with redundancy will dominate the overall unavailability. Evaluating and summing the probabilities of those dominant classes of failure combinations give a good approximation to the actual failure probability, i.e., the system unavailability. Dual-span failures can then be the next dominant failure mode in a network with $R_1 = 1$. The study of connection unavailabilities considering dual-span failures has actually become a practical issue especially for services in need of high availability.

Earlier studies of availability analysis focused on ring topologies [3–5]. Simulation results of end-to-end unavailability of optical networks implemented with simple protection schemes were presented in [6], but its so-called hybrid simulation and analytical approach was not explained. In [7], a network traffic and availability model was proposed, which is based on the geographic distribution and the number of network components. Integer linear programming-based approaches were provided [8] to find network designs according to the different availability requirements of connection requests. Its connection availability was approximated by neglecting the different connection demands. A method to assess the end-to-end path availability in span-restorable mesh networks was proposed [9] and results about dual failure restorability were given. But its discussion about path availability was only conducted under the assumption that all spans have the same physical unavailability for a path composed of four hops and no case study of path availabilities of

L. Zhou (✉) · M. Held · U. Sennhauser
Swiss Federal Laboratories for Materials Testing and Research
(EMPA), CH-8600 Dübendorf, Switzerland
e-mail: ling.zhou@empa.ch

L. Zhou
e-mail: lingjully@hotmail.com

a real network was presented. Dual failure restorability has been studied [10, 11] and restoration from any second span failure was maximized in networks with 100% restorability to single-span failures. The first analytical consideration of the availability of paths in a network protected by pre-configured protection cycles was presented in [12] without capacity details. Thus we find that viable methods and exact results about connection availability were provided just for unprotected or dedicatedly protected networks and that restoration schemes with shared spare capacity were analyzed by distinct methods from a general point of view of availability consideration only.

We present a new concept for availability analysis, named the restoration-aware connection availability (RACA). When dual-span failures are assumed to be the only failure mode, the RACA concept can be used to exactly analyze the present survivability strategies with dedicated or shared spare capacity. Basic models combined with this concept are built for the schemes of span restoration (SR) and pre-configured protection cycles (p -cycles). The redundancy of a mesh-restorable network with shared spare capacity is allocated to physically distributed spans and paths. Therefore, it is difficult and impractical to explicitly represent each item of a path in such networks with reliability block diagrams (RBDs) [1, 9]. Thus we use Monte Carlo simulations to compute the system unavailability of a basic model and to verify its analytical functions. On the other hand, the analytical computation to find the network performance and check the availability of each connection in a network is desirable and obviously simpler than RBD.

The article is organized as follows. Basic concepts including availability, span-oriented restoration schemes, dual failure restorability, and network redundancy are reviewed first. Then we propose the new availability calculation method for a connection. Based on models of span-oriented schemes, analytical functions are derived and verified by Monte Carlo simulations. Unavailability-related concepts, like average expected down time or traffic loss per year of all connections are introduced. A case study of the US network is then presented. Finally the conclusions are given.

2 Basic concepts

2.1 Availability

Availability is the probability of a repairable system to be in an operating state. Failures and down states occur, but maintenance or repair actions always return the system to an operating state [13]. The familiar equation for availability of a system with a constant failure λ and a repair rate μ is

$$A = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}} = \frac{\mu}{\mu + \lambda}, \quad (1)$$

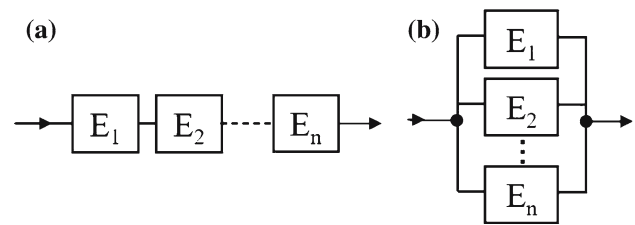


Fig. 1 Basic reliability block diagrams (a) Serial structure; (b) Parallel structure

where A is the availability, $\text{MTTF} = 1/\lambda$ the mean time to failure and $\text{MTTR} = 1/\mu$ the mean time to repair [14].

The availability A_s of a system represented by a serial RBD consisting of n independent elements, shown in Fig. 1a, is expressed as the product of availabilities of the elements:

$$A_s = \prod_{i=1}^n A_i \approx 1 - \sum_{i=1}^n U_i. \quad (2)$$

The probabilistic complement of availability A is the unavailability U , $U \equiv 1 - A$. Unavailability of a serial system U_s is then

$$U_s \approx \sum_{i=1}^n U_i. \quad (3)$$

This is a good approximation for $U_s \ll 1$. Analogously, unavailability of a parallel system U_p is

$$U_p = \prod_{i=1}^n U_i. \quad (4)$$

This is an exact equation for a parallel system [1, 15].

In most availability analyses of communication networks, unavailability quantities or expressions are used based on simplifying assumptions, transforming an exact availability model with intractable complexities to a feasible unavailability model with an acceptable numerical accuracy [1].

2.2 Span-oriented restorations

A span-oriented restoration reroutes failed traffic locally through a set of distinct restoration paths or deploys the pre-planned protection paths between the end nodes of a failed span. In this article, we mainly discuss two span-oriented schemes, i.e., span restoration and p -cycles.

Span restoration is a typical scheme in which restoration paths can be found in real time or be preplanned. The restoration path set may comprise all distinct paths bridging the end nodes of a particular failed span instead of one replacement path only for all failed traffic. The limitation of hop limits or physical length limits may be applied to the distinct restoration paths in the implementation. In this article, we limit the physical lengths of the candidate restoration paths

for each span and sort them in advance according to the rule of shortest paths first.

p-Cycles belong to a special type of span-oriented restoration [1]. *p*-Cycles including all the switching actions are completely preplanned. They are not simply rings. As a kind of span-oriented scheme, it can provide protection to working paths with shortest path routing, unlike rings in which a working path has to be included. *p*-Cycles can restore two types of span failures with different mechanisms. One is called “on-cycle span failure,” which occurs on a span of a *p*-cycle and can be protected by one protection path. The other is a so-called “straddling span failure,” which does not occur on a span of a *p*-cycle, but the failed span straddles on the *p*-cycle, i.e., the end nodes of the failed span are two nonadjacent nodes of the *p*-cycle. The particular *p*-cycle may contribute two protection paths to the restoration of the “straddling span failure.” The concept of “straddling spans” makes an important difference between *p*-cycles and rings. All candidate cycles are also sorted ascendingly by their physical lengths.

2.3 Dual failure restorability

The “dual-span failure restorability” $R_2(i, j)$ of a pair of failed spans *i* and *j* is defined [1] as the ratio of the restored capacities to the total working capacities on spans *i* and *j*, if both spans failed.

$$R_2(i, j) \equiv 1 - \frac{N(i, j)}{w_i + w_j} = 1 - \frac{N_i + N_j}{w_i + w_j}, \tag{5}$$

where

$N(i, j)$ is the total non-restorable capacities if a dual failure of spans *i* and *j* failed (span *i* failed first and span *j* second);

w_i and w_j are the working capacities of spans *i* and *j*, respectively, and $w_i + w_j$ is not zero;

N_i and N_j are the individual non-restorable capacities of spans *i* and *j*, respectively.

The average R_2 of a network, reflecting the overall demand-impact, can be stated as [1]:

$$R_2 \equiv \frac{\sum_{(i,j) \in S^2 | i \neq j} (w_i + w_j) \cdot R_2(i, j)}{\sum_{(i,j) \in S^2 | i \neq j} (w_i + w_j)}, \tag{6}$$

where S is the set of spans of the network.

Generally a different order of occurrence of the failure pair may lead to a different restorability for each span and possibly for the dual failure scenario. $R_2(i, j)$ is assumed to denote the restorability of a dual-span failure (*i, j*), where span *i* is the first failed span and span *j* is the second failed span.

For span restoration, we can compute $R_2(i, j)$ for the failed span pair (*i, j*) in two steps. In any variant form of span restoration, we have a determined set of backup paths for each failed span when computing R_2 . In step 1 we restore the traffic of the first failed span *i* by its intact backup paths, which are not hit by the second failed span *j*, and seize the necessary spare-capacity on the backup paths. If span *j* hits some of the backup paths of span *i*, the traffic of span *i* is no longer fully restored. N_i records the traffic loss. In step 2, the same procedure of restoring span *j* with the remaining spare-capacity of the network is applied. The corresponding traffic loss is recorded in N_j . If two span failures isolate a degree-2 node, no restoration can be made and $R_2(i, j) = 0$. For this case we can avoid going through the above two steps.

p-Cycles, as a span-oriented scheme, have a similar R_2 computational approach as that of span restoration. However, since it distinguishes between on-cycle and straddling spans, more complex judgments are required. The scheme of *p*-cycles applies a set of protection cycles for the whole network according to an optimization calculation. We need some criteria to find the cycles which are responsible to protect one specific span, i.e., to determine the set of backup paths for each span. We assume that traffic is restored in the order of the cycle with the biggest spare capacity first. If two cycles happen to have the same spare capacity, we use the shorter cycle first. If the failed span is an on-cycle span, the judgment is similar to the above span restoration. If the failed span is a straddling span, we have to judge whether both backup paths are required for its restoration. The shorter restoration path is always selected first.

2.4 Network redundancy

Network redundancy indicates how efficient the design is in its use of spare capacity. Standard network redundancy R is defined as [1]:

$$R = \frac{\left\{ \sum_{i \in S} (w_i + s_i) - \sum_{i \in S} w_i^* \right\}}{\sum_{i \in S} w_i^*}, \tag{7}$$

where $\sum w_i^*$ is the total working capacity needed to support shortest-path routing of all demands; w_i is the actual amount of working capacity on span *i*; s_i is the amount of spare capacity on span *i*. The capacity can be counted by the units of channels or by the product of channels and the lengths of their spans. A span is defined as the physical entity to collect all capacity channels in parallel between adjacent nodes [1]. In this article, we use the latter and compute the network redundancy. Therefore,

$$\sum_{i \in S} w_i = \sum_{i \in S} n w_i \cdot d_i, \tag{8}$$

$$\sum_{i \in S} s_i = \sum_{i \in S} n s_i \cdot d_i, \quad (9)$$

where $n w_i$ is the number of working channels on span i , $n s_i$ is the number of spare channels on span i , and d_i is the length of span i . In this article, $\sum w_i^*$ and $\sum w_i$ are identical for both span-oriented schemes.

3 Restoration-aware connection availability

3.1 General concept of RACA

We developed a new concept, called restoration-aware connection availability, or unavailability, to facilitate the assessment of connection unavailability in various survivable networks with dedicated or shared spare capacities.

A path in an unprotected network is a concatenation of spans. According to Eq. 3, the path unavailability U_{path} is [1,9]

$$U_{\text{path}} \approx \sum_{i=1}^n U_i^{\text{phy}}, \quad (10)$$

where U_i^{phy} is the physical unavailability of span i of the path.

A working path in a survivable network has a similar serial structure. However, each failed span of the path can be restored by a designated survivability scheme. Considering the protection or restoration effect, we use U_i^* as the equivalent unavailability of span i instead of U_i^{phy} in Eq. 10. The unavailability of a connection in such a network is then [1,9]

$$U_{\text{conn}} \approx \sum_{i=1}^n U_i^*. \quad (11)$$

In a mesh-restorable network designed with 100% single-span failure restorability, dual-span failure combinations dominate the system unavailability and spans fail independently. If the working demand of a connection is lost completely due to two failed spans i and j with either failing sequence (i, j) or (j, i) , the occurrence probability is the product of physical unavailabilities of the two failed spans, i.e., $U_i^{\text{phy}} \cdot U_j^{\text{phy}}$ [1]. Let $d_{s,t}$ represent the traffic demand of connection $s-t$ and NT_i the unrestored working traffic of span i of the working path due to a dual failure ($NT_i \in [0, d_{s,t}]$). Note that here NT_i is different from N_i in Eq. 5. The latter N_i includes the unrestored traffic of all affected connections due to a specific dual-span failure. If only a part of the traffic on span i is lost, the dual failure of spans i and j contributes $U_i^{\text{phy}} \cdot U_j^{\text{phy}} \cdot NT_i/d_{s,t}$ to the equivalent unavailability U_i^* of span i . The fraction of $NT_i/d_{s,t}$ represents the

occurrence probability of the non-restorable traffic of span i . Therefore we can express U_i^* as follows:

$$U_i^* = U_i^{\text{phy}} \cdot \sum_{j \in S | j \neq i} (c_j \cdot U_j^{\text{phy}} \cdot NT_i/d_{s,t}), \quad (12)$$

where $c_j = 0.5$ or 1. The particular value of c_j can be determined in each basic model of survivability schemes.

The computation of connection unavailability is based on physical unavailabilities of spans, which can be calculated according to $U = 1 - A$ and Eq. 1. Thus the physical unavailability of span i is:

$$U_i^{\text{phy}} = \frac{\text{MTTR}}{\text{MTTF} + \text{MTTR}} = \frac{\lambda_S}{\lambda_S + \mu_S} \approx \lambda_S \cdot \text{MTTR}, \quad (13)$$

where λ_S and μ_S are the failure rate and the repair rate of the span, respectively. The failure rate λ_S of a span is its length L multiplied by the sum of the cable failure rate per kilometer λ_{cable} and the optical amplifier failure rate per kilometer λ_{OA} [16,17]:

$$\lambda_S = L \cdot (\lambda_{\text{cable}} + C_{OA} \cdot \lambda_{OA}), \quad (14)$$

where C_{OA} is an optical amplifier spacing constant per kilometer.

3.2 Methodology of simulation

Since the RBD method is not suitable for survivable networks with shared spare capacity, we use a Monte Carlo simulation to compute the unavailabilities and to find their analytical functions. In a simulation, we could collect actual outage data from random trials and compute the unavailability. In this article, we apply simulations simply to a model of one span-restorable connection consisting of a few spans. The general methodology is as follows:

1. Generate a very long “time-line” of randomly occurring failures for each span of a network (in this article, only for spans of a connection). Simulate the interval of failures by using a random number generator of a negative exponential distribution with a parameter λ , i.e., the failure rate of a span. If $|S|$ is the number of spans in a network, we have to generate $|S|$ time-lines from the same time point $t = 0$.
2. The repair time is regarded as the outage time due to a failure. If a failure occurs, the repair time is generated through a random number generator of a negative exponential distribution with a parameter $\mu = 1/\text{MTTR}$, where MTTR is the mean time to repair of a span.
3. Find dual failures in a simulated system. If there is an overlapping time of two span failures, a dual-span failure has occurred. As shown in Fig. 2, the failure f_2 of span j occurs during the repair time of the failure f_1 of span

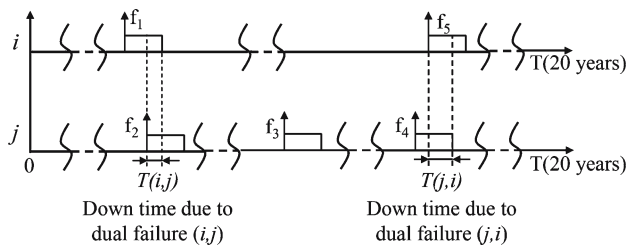


Fig. 2 Dual-span failures in a simulation. A certain repair time follows the occurrence of each failure. A dual failure appears if the repair times of two independent failures overlap each other. The expression of (i, j) indicates the sequence of failure occurrence of the two spans, i.e., span i fails first and span j second

- i . (i, j) is then a dual failure. $T(i, j)$ represents the possible down time caused by the dual-span failure (i, j) . The expression of (i, j) indicates the failing sequence of the two spans, i.e., span i fails first and span j second.
- 4. Record the traffic loss $LT(i, j)$. We do not record the outages caused by single-span failures to which full restorability is designed in a survivable network. If the network with built-in redundancy can restore the dual-span failure completely, there will be no traffic loss and no down time for the connection due to the dual failure; otherwise we have to record its outage time.
- 5. Check whether the dual failures affect the traffic for a connection. The unavailability of connection $s - t$ with demand $d_{s,t}$ can be computed with the following equation:

$$U_{\text{conn}}^{s,t} = \frac{\sum_k T_k(i, j) \cdot LT_k(i, j) / d_{s,t}}{T}, \tag{15}$$

where T is the total operating time, e.g., 20 years. k denotes the k th dual failure during the period of T . $LT_k(i, j)$ is the lost traffic due to dual-span failure (i, j) and $T_k(i, j)$ the down time due to (i, j) . Note that Eq. 15 can also be used to calculate the equivalent unavailability of part of connection $s - t$, such as for one or more spans considering their restoration effect.

- 6. Compute an average unavailability by running the simulation 1,000 times.

Triple-span or even higher-order span failures may appear in addition to dual-span failures. But their contributions to a connection unavailability is negligible compared with that of dual failures [1].

3.3 RACA model for span-oriented schemes

In span-oriented schemes including span restoration and p -cycles, a failed span i may affect the traffic of several connections traversing it and usually has multiple restoration paths at its disposal. In the implementation of span restoration or

p -cycles, a selection among the possible restoration paths has to take place to determine the backup paths of a specific failed span of a connection, which has been illustrated in Sect. 2.3

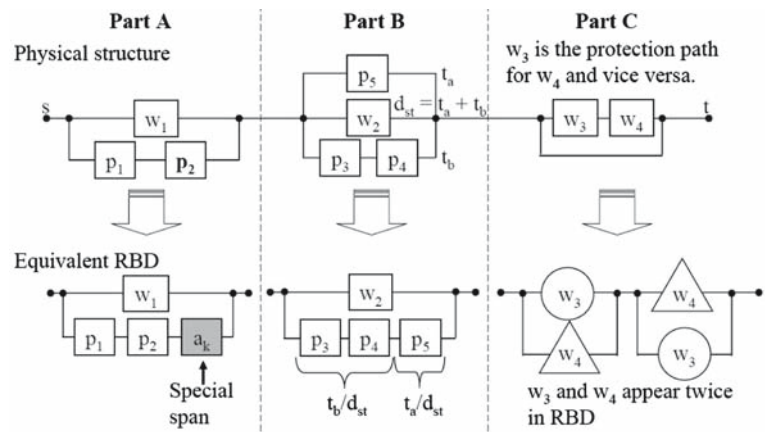
To determine the unavailability of a connection, we use the approach of transferring a physical structure to an equivalent RBD of a connection for better illustration. The upper portion of Fig. 3 shows the typical physical structure of a connection in a span-restorable network, which can be divided into three basic parts, actually representing three elementary types. The lower portion of Fig. 3 shows the equivalent RBD of each part. Span w_i represents a span of the working path and span p_j a span of the protection path.

“Part A” represents that the traffic of span w_1 of connection $s - t$, i.e., $d_{s,t}$, can be restored by a backup path over spans p_1 and p_2 with sufficient spare capacity if w_1 fails first. We assume that span a_k is of the working path of another connection $o - d$ and carries traffic $d_{o,d}$. The backup path of a_k also includes span p_1 and/or p_2 . If span a_k fails first and w_1 fails second in a dual failure scenario, capacity contention may occur and the traffic of w_1 may not be fully restored.

If the traffic $d_{s,t}$ of span w_1 can be fully restored with its direct protection path built by spans p_1 and p_2 , the physical structure of “part A” is then a valid RBD and the equivalent unavailability $U_{w_1}^*$ of span w_1 can be calculated directly. If span a_k is the first failure and its traffic $d_{o,d}$ is restored first, the traffic of span w_1 may not be fully restored with the remaining spare capacity of the span(s) p_1 and/or p_2 shared with the restoration path of a_k . Therefore, span a_k does affect the restoration of span w_1 . The unrestored traffic of span w_1 is thus called $NT_{w_1,ct}^{(a_k,w_1)}$. “ct” represents contention. (a_k, w_1) indicates that span a_k fails first. Thus we can draw an equivalent RBD by adding an equivalent block for span a_k to the backup path of span w_1 . The “gray” block of span a_k , with the unavailability of $U_{a_k}^{\text{eqv}} = 0.5 \cdot U_{a_k}^{\text{phy}} \cdot NT_{w_1,ct}^{(a_k,w_1)} / d_{s,t}$, is different from the original block of span a_k with the unavailability $U_{a_k}^{\text{phy}}$. Since only one failing sequence (a_k, w_1) may cause unrestored traffic of span w_1 , the factor of 0.5 has to be added to reflect the probability. If there are more spans playing the same role as a_k to the connection $s - t$, we can add all of them into the equivalent block diagram in the same way. Generally, the equivalent unavailability of a working span i with the feature of “part A” is:

$$U_{i,partA}^* = U_i^{\text{phy}} \cdot \left(\sum_{j \in R_i} U_j^{\text{phy}} + \sum_k U_{a_k}^{\text{eqv}} \right) = U_i^{\text{phy}} \cdot \left[\sum_{j \in R_i} U_j^{\text{phy}} + \sum_k \left(0.5 \cdot U_{a_k}^{\text{phy}} \cdot \frac{NT_{i,ct}^{(a_k,i)}}{d_{s,t}} \right) \right], \tag{16}$$

Fig. 3 Basic models of the equivalent RBD of span-oriented schemes. The upper portion shows the typical physical structure of a connection in a span-restorable network, which has three basic parts A, B, and C. The lower portion shows the equivalent RBD of each part



where R_i is the set of spans of the restoration paths of span i .

Figure 4 shows an example of how to transfer the physical structure with the feature of “part A” to its equivalent RBD. Span 1 is the working span of connection $s - t$ while the backup path consisting of spans 2 and 3 is responsible for restoring the traffic $d_{s,t}$ of span 1. Span 4 is the working span of connection $o - d$ and its backup path also includes span 3. The physical characteristics of spans in Fig. 4 are listed in Table 1. There are two situations of dual-span failures which can cause traffic loss of span 1 in this model. The first situation is that one failed span is span 1 and the other one is of its backup paths, i.e., span 2 or 3. The second situation, i.e., span 4 fails first and span 1 fails second, can cause spare-capacity contention and lead to traffic loss. Span 4 is the so-called “special span” in Fig. 4. In the simulation, we collect the outage due to these two types of dual-span failures. If the dual failure belongs to the first type, we need to include the outage caused by both failing sequences of two spans; if it is of the second type, we only collect the outage caused by the sequence of (4, 1) with the ratio of the lost traffic NT_{ct} to the demand $d_{s,t}$. (4, 1) means that the special span, e.g., span 4 in the figure, fails first and the working span second. Finally the equivalent unavailability of span 1 is computed with Eq. 15.

Table 2 compares the equivalent unavailability results of span 1 in Fig. 4 computed by the simulation and analytical approaches, if the special span is span 4, 3, or 2, respecti-

vely. The column “Lost traffic ratio” lists the corresponding $NT_{ct}/d_{s,t}$ for the case of different special span. The analytical results were computed according to Eq. 16. The simulation results may vary in some range due to the finite simulation time, but its average values of just a few trials are always close to the analytical results in each case, which demonstrates the validity of the analytical functions.

In “part B,” there are two restoration paths which are responsible for the restoration of span w_2 . Each restoration path has a different traffic flow ($t_a \neq t_b$ and $t_a + t_b = d_{s,t}$). We put the spans of the two restoration paths into one path in the equivalent RBD of “part B” with modified probabilities, $t_a/d_{s,t}$ and $t_b/d_{s,t}$, respectively. If the dual failure of spans w_2 and p_5 occurs, the traffic of t_a is lost and only t_b on span w_2 can be transmitted successfully. If the dual failure of spans w_2 and p_3 or p_4 occurs, the traffic of t_b is lost and t_a can survive. Therefore the equivalent unavailability of span w_2

Table 1 Assumed span physical unavailability with MTTR = 24h

Span ID	Span failure rate λ_s (FITs)	Span unavailability
1	1,72,900	4.13E-03
2	86,240	2.07E-03
3	1,75,140	4.19E-03
4	1,50,220	3.59E-03
5	4,40,440	1.05E-02

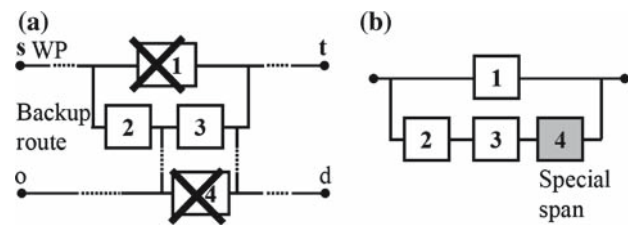


Fig. 4 Transition from the physical structure of a “part A” example to its equivalent RBD. (a) The physical structure of the “part A” example where span 3 is shared by the protection paths of connections $s - t$ and $o - d$. (b) The equivalent reliability block diagram

Table 2 Simulation and analytical unavailability results (Part A)

Special span ID	Lost traffic ratio $NT_{ct}/d_{s,t}$	Simulation result U_1^*	Analytical result U_1^*
4	0.5	2.93E-05	2.95E-05
3	0.25	2.54E-05	2.55E-05
2	0.7	3.49E-05	3.51E-05

is

$$U_{w_2}^* = U_{w_2}^{\text{phy}} \cdot \left(U_{p_3}^{\text{phy}} \cdot \frac{t_b}{d_{s,t}} + U_{p_4}^{\text{phy}} \cdot \frac{t_b}{d_{s,t}} + U_{p_5}^{\text{phy}} \cdot \frac{t_a}{d_{s,t}} \right). \tag{17}$$

Here $t_b/d_{s,t}$ or $t_a/d_{s,t}$ is in fact the ratio of the traffic loss to the working demand. More generally, for a working span i with multiple restoration paths as span w_2 in “part B,” we have

$$U_{i,\text{partB}}^* = U_i^{\text{phy}} \cdot \sum_{j \in R_i} \left(0.5 \cdot U_j^{\text{phy}} \frac{NT_i^{(i,j)}}{d_{s,t}} + 0.5 \cdot U_j^{\text{phy}} \frac{NT_i^{(j,i)}}{d_{s,t}} \right), \tag{18}$$

where R_i is the set of spans of the restoration paths of span i . $NT_i^{(i,j)}$ and $NT_i^{(j,i)}$ are the unrestored traffics due to dual failures (i, j) and (j, i) , respectively. When $NT_i^{(i,j)} = NT_i^{(j,i)} = NT_i$, Eq. 18 becomes:

$$U_{i,\text{partB}}^* = U_i^{\text{phy}} \cdot \sum_{j \in R_i} \left(U_j^{\text{phy}} \cdot \frac{NT_i}{d_{s,t}} \right). \tag{19}$$

Figure 5 shows an example of the “part B” model of span-oriented schemes. Figure 5a shows its physical structure in which span 1 is the working span carrying the demand $d_{s,t}$ and has two restoration paths. Spans 2 and 3 are of a restoration path of span 1 with flow t_a and spans 4 and 5 of the other restoration path with flow t_b ($t_a + t_b = d_{s,t}$). The physical characteristics of the spans in Fig. 5 are listed in Table 1. Figure 5b shows the equivalent reliability block diagram. In the simulation, we consider all the dual-span failures including span 1. If the other failed span is span 2 or 3, the dual failure outage has to be multiplied with $t_a/d_{s,t}$ before being added into the total downtime, because the dual failure only causes $t_a/d_{s,t}$ traffic loss. Similarly, the outage has to be multiplied with $t_b/d_{s,t}$, if the other failed span is span 4 or 5.

Table 3 compares the system unavailability results computed by the simulation and analytical approaches with three groups of $t_a/d_{s,t}$ and $t_b/d_{s,t}$ ratios. We apply Eq. 19 in the analytical computation of equivalent unavailability of span 1 in Fig. 5. Thus

$$U_{1,\text{partB}}^* = U_1^{\text{phy}} \cdot \left[(U_2^{\text{phy}} + U_3^{\text{phy}}) \frac{t_a}{d_{s,t}} + (U_4^{\text{phy}} + U_5^{\text{phy}}) \frac{t_b}{d_{s,t}} \right]. \tag{20}$$

The values in the table show that simulation and analytical solutions for “part B” are in good agreement.

“Part C” represents a basic model which often appears in a span-restorable network, i.e., two spans are both part

of a working path and also of the restoration path of each other. This is a kind of RBD in which at least one element appears more than once. This situation can be investigated with the key item method introduced in [14]. Thus the system availability of the “Part C” example in Fig. 3 is

$$A_{\text{partC}} = A_{w_3}^{\text{phy}} + A_{w_4}^{\text{phy}} - A_{w_3}^{\text{phy}} \cdot A_{w_4}^{\text{phy}}, \tag{21}$$

and its system unavailability

$$U_{\text{partC}} = U_{w_3}^{\text{phy}} \cdot U_{w_4}^{\text{phy}}. \tag{22}$$

If we apply our RACA concept, we can get the same equation as follows:

$$U_{\text{partC}}^* = U_{w_3}^* + U_{w_4}^* = 0.5U_{w_3}^{\text{phy}} U_{w_4}^{\text{phy}} + 0.5U_{w_4}^{\text{phy}} U_{w_3}^{\text{phy}} = U_{w_3}^{\text{phy}} U_{w_4}^{\text{phy}}. \tag{23}$$

In fact the equivalent unavailabilities of the working spans w_3 and w_4 are $U_{w_3}^* = 0.5U_{w_3}^{\text{phy}} U_{w_4}^{\text{phy}}$ and $U_{w_4}^* = 0.5U_{w_3}^{\text{phy}} U_{w_4}^{\text{phy}}$, respectively. The factor 0.5 in the equation has a different meaning from what we have met in the analysis of “part A” model. In the previous equations, 0.5 is used to adjust the probability $U_i^{\text{phy}} U_j^{\text{phy}}$ of two failed spans i and j if only one failure sequence can cause traffic loss. In Eq. 23 of “part C,” we apply the factor 0.5 to avoid double counting of the probability of $U_{w_3}^{\text{phy}} U_{w_4}^{\text{phy}}$ in the system unavailability.

If we assume that the characteristics of spans w_3 and w_4 of “part C” are the same as those of spans 3 and 4 listed in Table 1, i.e., $U_{w_3}^{\text{phy}} = 4.19E - 3$ and $U_{w_4}^{\text{phy}} = 3.59E - 3$, we can use the software Relex [18], a tool providing reliability and availability analysis, to compute the unavailability of “part C” through its RBD directly, where $U = 1 - A$ and Eq. 21 are used. The unavailability calculated according to its RBD is $U_{\text{partC}}^{\text{RBD}} = 1.50366E - 5$. The unavailability from the analytical equation with the RACA concept, Eq. 23, is $U_{\text{partC}}^{\text{analy}} = 1.50366E - 5$. Thus $U_{\text{partC}}^{\text{RBD}} \approx U_{\text{partC}}^{\text{analy}}$. Even if we change the values of span physical unavailabilities, the results directly from RBD and from the analytical methods are in good agreement with each other. Therefore, the RACA concept and the factor of 0.5 are used correctly in the analytical function for “part C.”

In Fig. 6 a complicated model of “part C,” which cannot be solved by Relex with RBD directly, is drawn. The span characteristics are still the same as those in Fig. 5 and listed in Table 1. The working traffic of span 1, $d_{s,t}$, can be restored by one path consisting of spans 2 and 3. The working span 2 has two restoration paths. One path consists of span 5 with t_a flow and the other path includes spans 1 and 4 with t_b flow. Thus the dual-span failure including spans 1 and 2 may cause the traffic of $d_{s,t}$ or t_b to be lost. Since $d_{s,t} > t_b$ and spans 1, 2 are part of the same working path, we find that the traffic loss due to dual failure (1,2) or (2,1)

Fig. 5 Transition from the physical structure to its equivalent RBD for the “part B” model of span-oriented schemes. (a) Physical structure where span 1 is the working span. (b) Equivalent reliability block diagram

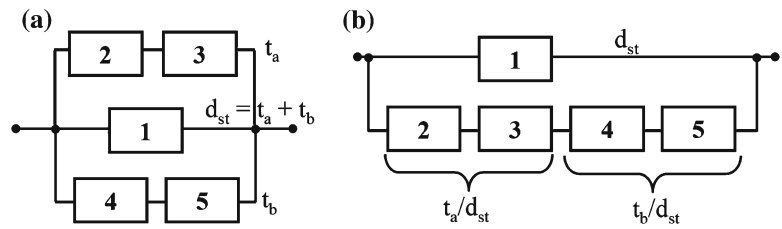


Table 3 Simulation and analytical unavailability results (part B)

Trial	Ratio $t_a/d_{s,t}$	Ratio $t_b/d_{s,t}$	Simulation result U_1^*	Analytical result U_1^*
1	0.25	0.75	5.04E-05	5.00E-05
2	0.5	0.5	4.20E-05	4.20E-05
3	0.6	0.4	3.87E-05	3.87E-05

Table 4 Simulation and analytical unavailability results of span-oriented schemes (part C)

Trial	Ratio $t_a/d_{s,t}$	Ratio $t_b/d_{s,t}$	Simulation result $U_{conn}^{s,t}$	Analytical result $U_{conn}^{s,t}$
1	0.2	0.8	3.56E-05	3.61E-05
2	0.5	0.5	4.07E-05	4.03E-05
3	0.7	0.3	4.32E-05	4.32E-05

is always $d_{s,t}$. Then we conclude that *for the working spans appearing twice in an equivalent RBD, the biggest traffic loss determines their failure probabilities, though they may cause different outages in the diverse parts of the RBD*. The rule can help us make simulations and derive the corresponding analytical functions correctly. In the simulation, we collect all the outage caused by dual-span failures (2,1), (1,2), (1,3), and (3,1). If a dual-span failure including spans 2 and 5 occurs, the outage has to be multiplied with $t_a/d_{s,t}$. Similarly, the outage caused by dual failures (2,4) or (4,2) has to be multiplied with $t_b/d_{s,t}$. According to the above description, the equivalent unavailabilities of the working spans 1 and 2 can be as follows:

$$U_1^* = U_1^{phy} \left(0.5U_2^{phy} + U_3^{phy} \right), \tag{24}$$

$$U_2^* = U_2^{phy} \left(0.5U_1^{phy} + U_4^{phy} \frac{t_b}{d_{s,t}} + U_5^{phy} \frac{t_a}{d_{s,t}} \right), \tag{25}$$

where factor 0.5 is used to avoid double counting of the probability $U_1^{phy}U_2^{phy}$, because spans 1 and 2 appear twice in the equivalent RBD.

The equivalent system unavailability shown in Fig. 6 is

$$U_{partC}^* = U_1^* + U_2^*. \tag{26}$$

Table 4 compares the system unavailabilities shown in Fig. 6, which were computed by the simulation and analytical

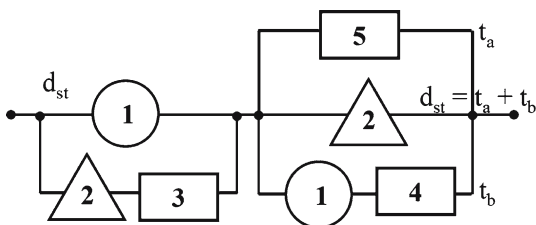


Fig. 6 An equivalent RBD of a complicated example for span-oriented schemes (part C)

approaches with three ratio groups of $t_a/d_{s,t}$ and $t_b/d_{s,t}$, respectively. The simulation results are in good agreement with the analytical results. Therefore, the analytical functions for “part C” are valid.

We also checked the frequency of triple-span failures in each simulation. Even if triple-span failures can bring traffic loss, they only occur with a very low probability compared with dual-span failures and can be neglected in the computation. For example, when we ran the simulation 1,000 times for 20 years in a network composed of five spans, 99% of them did not have any triple-span failures while 96% of the simulations did have dual-span failures. The number of triple-span failures in one simulation can never exceed three.

We sum up the described equations for the three basic parts and obtain the total equivalent unavailability U_i^* . The restoration-aware connection unavailability for span-oriented schemes, U_{conn}^{SO} , is as follows:

$$U_{conn}^{SO} = \sum_{i \in S | i \in W} U_i^*, \tag{27}$$

where S is the set of all spans in a network and W is the span set of the working path of a connection.

The case of an element appearing more than once in a RBD is only discussed in “part C.” We have to note that the same spans may appear in the restoration paths of two working spans of a connection, especially for adjacent working spans. We now discuss this situation to show that it does not affect the computation of connection unavailability according to the RACA concept.

In Figure 7, spans 2 and 3 are two adjacent working spans, whose restoration paths both include span 1. The physical characteristics of Fig. 7 are listed in Table 1. According to our RACA concept, the equivalent unavailabilities of spans 2 and 3 are:

$$U_2^* = U_2^{phy} \cdot \left(U_1^{phy} + U_4^{phy} \right), \tag{28}$$

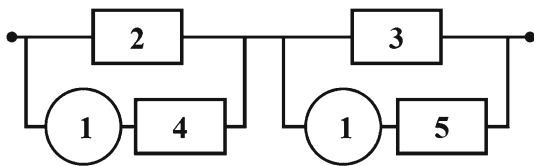


Fig. 7 Span 1 appears in both the restoration paths of the working spans 2 and 3

and

$$U_3^* = U_3^{\text{phy}} \cdot (U_1^{\text{phy}} + U_5^{\text{phy}}), \tag{29}$$

respectively. The system unavailability U_{sys} of Fig. 7 is thus:

$$U_{\text{sys}}^* = U_2^* + U_3^* = 7.70E - 5. \tag{30}$$

By the key item method introduced in [14], the system availability is:

$$\begin{aligned} A_{\text{sys}} &= A_1^{\text{phy}} \cdot (A_2^{\text{phy}} + A_4^{\text{phy}} - A_2^{\text{phy}} A_4^{\text{phy}}) (A_3^{\text{phy}} + A_5^{\text{phy}} \\ &\quad - A_3^{\text{phy}} A_5^{\text{phy}}) + (1 - A_1^{\text{phy}}) A_2^{\text{phy}} A_3^{\text{phy}} \\ &= 0.9999232. \end{aligned} \tag{31}$$

Thus its unavailability is $U_{\text{sys}} = 1 - A_{\text{sys}} = 7.68E - 05$, which can also be verified through the software Relex. The system unavailability U_{sys} is almost the same as the equivalent unavailability U_{sys}^* (see Eq. 30) calculated by the RACA concept, considering the unavailability approximation of serial structures. If we make an even simpler RBD, the two unavailability values calculated by different methods can be completely the same. However, the RBD method is too complex to draw reliability block diagrams and make their formulations to be applied for all the connections of a network.

We have analyzed the basic models of span-oriented schemes and derived their proper analytical functions, which all have the form presented in Eq. 12. In a practical unavailability computation of a connection in a span-restorable network, more complicated situations may appear, e.g., a combination of part A and B (i.e., there is also a spare-capacity contention on the basis of the “part B” model). We can always apply the rules concluded from the basic models and make a correct analysis of connection unavailability.

If $EDC_{s,t}$ represents the expected down time per year of a connection between nodes s and t , then

$$EDC_{s,t} = U_{\text{conn}}^{s,t} * 365 * 24, \quad s, t \in N. \tag{32}$$

After we have $EDC_{s,t}$ for each connection, we can define the average connection downtime of a network, i.e., average expected downtime per year of all connections (AEDC) (hour/year) [19].

$$AEDC = \frac{1}{|N|(|N| - 1)} \cdot \sum_{(s,t) \in N^2 | s \neq t} EDC_{s,t}, \tag{33}$$

where $|N|$ is the number of nodes of the network.

To calculate the connection traffic loss in a span-restorable network, we introduce a concept of restoration-aware connection traffic loss (RCTL), to sum up the product of the traffic loss due to a dual-span failure and its occurrence probability for all dual failure cases. If $RCTL_i^*$ is the equivalent traffic loss of each working span of connection $s - t$. The loss of traffic of connection $s - t$, $RCTL_{s,t}$ is then:

$$RCTL_{s,t} = \sum_{i \in S | i \in W} RCTL_i^*, \quad s, t \in N, \tag{34}$$

and the average value, i.e., average restoration-aware connection traffic loss (ARCTL), is

$$ARCTL = \frac{1}{|N|(|N| - 1)} \cdot \sum_{(s,t) \in N^2 | s \neq t} RCTL_{s,t}. \tag{35}$$

3.4 The assumption of allocation priority

After the development of the basic connection availability concept for networks with span-oriented survivability schemes, we apply the concept to networks optimized for single-span failures. One must be aware that the result of optimization merely consists of the provisioning of enough spare capacity for the case of single-span failures and the span capacity allocation to restore all the traffic of a span. However, this contains no information about connection backup routing, i.e., about the allocation of restoration paths of a failed working span of a connection to restore its connection demand, which will be necessary in a real implementation. The situation gets even more complicated with the occurrence of the second failure if contention about spare capacity can occur. Decisions on connection-related spare capacity allocation are required as well. The exact information about which channels are allocated to which restoration path and which restoration paths of a span are responsible to restore its connection-related working traffic is what we call the restoration details.

In practice, allocation rules can be based on criteria such as traffic priority levels, the lengths of working paths of connections, etc. In order to demonstrate the RACA concept for span-oriented schemes, we have implemented allocation rules to determine restoration details as follows: a fixed number is assigned to each connection, or connections are given numbers according to the numbering of their end nodes. Then a priority criterion is applied, e.g., the lowest number first. This is of course a quite arbitrary rule but it is sufficient for the evaluation of the RACA concept and can be easily expanded to rules with a more practical background, for instance, when priorities of connections or certain traffic demands can be defined for the network. Therefore, we have to make assumptions on the unknown information by the following two steps.

Step 1 is to define backup paths of spans of each connection by assuming a certain allocation rule. In our case study, we define four rules with the following codes written in C:

• **Ascending sequence:**

```
for( $n_1 = 1$ ;  $n_1 < \#nodes$ ;  $n_1 ++$ )
    for( $n_2 = n_1 + 1$ ;  $n_2 \leq \#nodes$ ;  $n_2 ++$ )
```

• **Descending sequence:**

```
for( $n_1 = \#nodes - 1$ ;  $n_1 \geq 1$ ;  $n_1 --$ )
    for( $n_2 = \#nodes$ ;  $n_2 > n_1$ ;  $n_2 --$ )
```

• **Mixed sequence 1:**

```
for( $n_1 = \#nodes$ ;  $n_1 \geq 2$ ;  $n_1 --$ )
    for( $n_2 = n_1 - 1$ ;  $n_2 \geq 1$ ;  $n_2 --$ )
```

• **Mixed sequence 2:**

```
for( $n_2 = 2$ ;  $n_2 \leq \#nodes$ ;  $n_2 ++$ )
    for( $n_1 = 1$ ;  $n_1 < n_2$ ;  $n_1 ++$ )
```

The demand matrix of a test network is symmetric, i.e., $d_{n_1, n_2} = d_{n_2, n_1}$, where n_1, n_2 are the end nodes of a connection. Therefore it is sufficient to check half of the demand matrix only. If the node-ID $n_1 < n_2$, we check (n_1, n_2) ; otherwise, we check (n_2, n_1) only. Table 5 shows an example of each rule with various allocation sequences, if the total number of nodes is four.

Now we can use an allocation rule to restore each failed connection. Figure 8 illustrates the operation of step 1.

Table 5 Four assumed allocation rules

Sequence	Ascending	Descending	Mixed 1	Mixed 2
Example	(1,2)	(3,4)	(3,4)	(1,2)
total	(1,3)	(2,4)	(2,4)	(1,3)
nodes=4	(1,4)	(2,3)	(1,4)	(2,3)
(n_1, n_2)	(2,3)	(1,4)	(2,3)	(1,4)
	(2,4)	(1,3)	(1,3)	(2,4)
	(3,4)	(1,2)	(1,2)	(3,4)

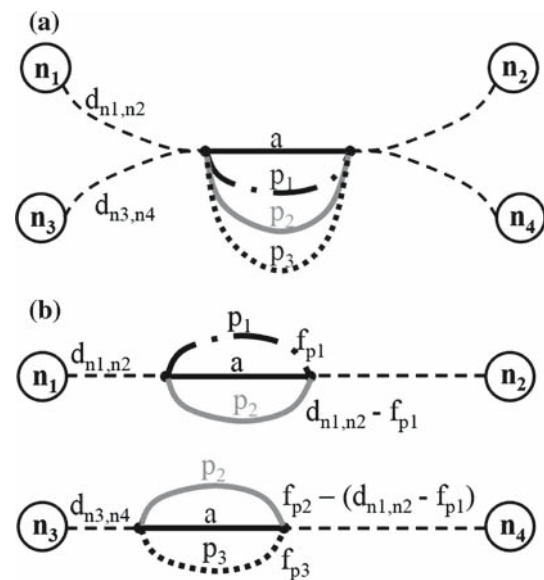


Fig. 8 Determination of restoration paths for each connection under an ascending sequence. (a) A physical structure, where span a carries the working demands (d_{n_1, n_2} and d_{n_3, n_4}) of connections $n_1 - n_2$ and $n_3 - n_4$ only. (b) Restoration paths in the equivalent RBD of each connection ($0 < d_{n_1, n_2} - f_{p_1} < f_{p_2}$ where f_{p_i} is the flow on path p_i)

Span a , which is protected by three backup paths p_1 , p_2 , and p_3 , carries the working demands of connections $n_1 - n_2$ and $n_3 - n_4$ only. The optimized results of an integer linear programming (ILP) formulation give this routing information directly. However, the backup paths, responsible for the restoration of each connection, can be determined only after an allocation rule has been assumed. If we use the “ascending sequence,” connection $n_1 - n_2$ will have higher priority than connection $n_3 - n_4$ and $n_1 - n_2$ can be provided as much spare capacity as possible first. If we use the “descending sequence,” connection $n_3 - n_4$ will have higher priority. In Fig. 8, the “ascending sequence” is used. Fig. 8b shows that backup paths p_1 and p_2 are responsible to restore the working demand d_{n_1, n_2} between nodes n_1 and n_2 under the condition that $0 < d_{n_1, n_2} - f_{p_1} < f_{p_2}$. The spans of path p_1 and its restoration flow, f_{p_1} , is recorded as part of backup paths in the equivalent RBD of connection $n_1 - n_2$. The spans of path p_2 and its capacity contribution, $(d_{n_1, n_2} - f_{p_1})$, need to be recorded for connection $n_1 - n_2$ as well. Thus backup paths p_2 and p_3 with their corresponding remaining flows are responsible for the restoration of connection $n_3 - n_4$. In this way, we can record all the restoration paths with their contribution of spare capacity for each working span of a connection in a span-restorable network.

After allocating connection-related restoration paths for each span, we go to step 2 to find the special spans which affect the restoration of a connection due to spare-capacity contention. Step 2 is conducted in the process that individual dual failure restorabilities are computed. Dual-span

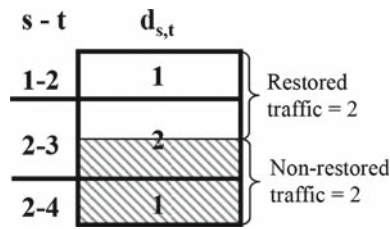


Fig. 9 Impact of dual failure (a, b) on connections and its restoration with an “ascending sequence.” $LT_{1,2}(a, b) = 0 \Rightarrow 0 \rightarrow U_{\text{conn}}^{1,2}$, $LT_{2,3}(a, b) = 1 \Rightarrow 0.5U_a^{\text{phy}}U_b^{\text{phy}} \cdot 1/2 \rightarrow U_{\text{conn}}^{2,3}$, $LT_{2,4}(a, b) = 1 \Rightarrow 0.5U_a^{\text{phy}}U_b^{\text{phy}} \cdot 1/1 \rightarrow U_{\text{conn}}^{2,4}$

failure (a, b) may cause part of the working traffic on a span $(a$ or $b)$ to be lost. The survivable spare capacity will be allocated for the connection with higher priority first. The dual failure (a, b) affects the unavailability of a connection with the probability of $0.5U_a^{\text{phy}}U_b^{\text{phy}} \cdot LT_{s,t}(a, b)/d_{s,t}$, where $0.5U_a^{\text{phy}}U_b^{\text{phy}}$ denotes the occurrence probability of (a, b) ; $LT_{s,t}(a, b)$ is the lost traffic of connection $s - t$ due to dual failure (a, b) ; $d_{s,t}$ is the working demand of connection $s - t$. Figure 9 shows the impact of dual failure (a, b) on three connections and the corresponding restoration when the “ascending sequence” is applied. Connections 1–2, 2–3, 2–4 are affected due to the dual failure (a, b) . The working demands of the three connections are $d_{1,2} = 1$, $d_{2,3} = 2$ and $d_{2,4} = 1$, respectively. Two units of the traffic can be restored and two units are lost. If the “ascending sequence” is applied, spare capacity is allocated first to connection 1–2, second to 2–3, and last to 2–4. Because $d_{1,2} = 1 < 2$ units of restored traffic, connection 1–2 can be restored completely ($LT_{1,2}(a, b) = 0$) and the dual failure (a, b) does not contribute to its connection unavailability $U_{\text{conn}}^{1,2}$. Connection 2–3 can only be restored half ($LT_{2,3}(a, b) = 1 < d_{2,3} = 2$) and $0.5U_a^{\text{phy}}U_b^{\text{phy}} \cdot 1/2$ is added to its connection unavailability $U_{\text{conn}}^{2,3}$. The traffic of connection 2–4 is lost completely ($LT_{2,4}(a, b) = 1$) due to dual failure (a, b) with the probability of $0.5U_a^{\text{phy}}U_b^{\text{phy}}$.

Combining these two steps, we can draw a complete equivalent reliability block diagram for a connection of a span-restorable network. Step 1 is aiming to find the direct backup paths of a connection. Step 2 is to find the special spans which affect the restoration of a connection due to spare-capacity contention. We call the method including the two steps “fixed backups” (FB), which reflects the realistically expected restoration.

4 Case study of connection unavailability

We made an experiment on the US network [20], consisting of 19 nodes, 28 spans, and 171 bidirectional connections. The

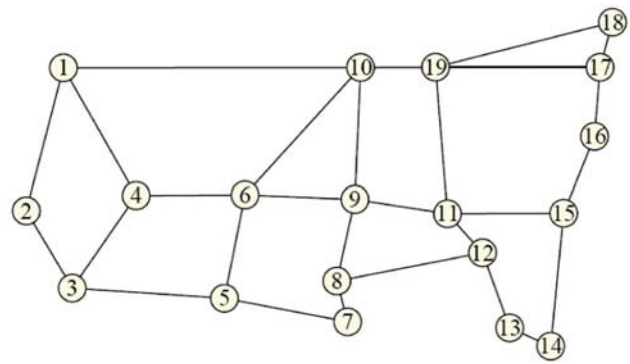


Fig. 10 The topology of the US network

Table 6 Reference reliability data

Parameter	Reference value	
λ_{cable}	100	FIT/km
$MTTR_S = 1/\mu_S$	24h	
C_{OA}	0.02/km	
λ_{OA}	2,000	FIT

network is investigated from the availability point of view, as shown in Fig. 10, when protected with span restoration and p -Cycles, respectively. Node failures are neglected because the node equipment usually comprises a redundancy and its repair time is very short. Analytical calculations are based on reference data [21, 22], as given in Table 6, which lists all the parameters of Eq. 14. $MTTR_S$ applies to the complete span, i.e., repair times of optical amplifiers and fiber optical cables are considered to be equal. The length of spans and the demand matrix are from the data in [20]. A connection demand $d_{s,t}$ between nodes s and t ranges from 1 to 3 (10 Gb/s).

Table 7 shows the connection unavailability-related results of span restoration and p -cycles. The working path of a connection is the shortest path for the US network with span restoration or p -cycles. In the computation of span restoration, at most ten shortest backup paths of each working span were exploited, resulting in $R_2 = 0.74$. In the computation of p -cycles, 50 candidate protection cycles were used in the optimization, leading to $R_2 = 0.75$. The column “R” represents the “standard network redundancy” as described in Sect. 2.4. The column “Allocation rules” lists four allocation rules “Ascending sequence,” “Descending sequence,” “Mixed sequence 1” and “Mixed sequence 2,” respectively. For both computations of span restoration and p -cycles, the average results under four allocation rules are almost unchanged. Therefore, we can conjecture that the average connection unavailability-related measures of a network will not change when computed with any allocation rule and can reflect the performance of a span-restorable network. Thus a computation under one presumed allocation rule is enough.

Table 7 Average connection performance of the US network with span restoration or p -cycles under four presumed rules

Schemes	Allocation rules	R	R_2	Average U_{conn}	AEDC (hour/year)	ARCTL (10 Gb/year)
Span Res.	ascending	1.08	0.74	1.34E-04	1.18	8.09E+03
	descending			1.31E-04	1.15	8.01E+03
	mix1			1.31E-04	1.15	8.00E+03
	mix2			1.36E-04	1.19	8.13E+03
p cycles	ascending	1.26	0.75	1.60E-04	1.40	9.67E+03
	descending			1.60E-04	1.40	9.61E+03
	mix1			1.61E-04	1.41	9.71E+03
	mix2			1.61E-04	1.41	9.82E+03

In Table 7, the corresponding values of U_{conn} , AEDC or ARCTL of span restoration and p -cycles are of the same order of magnitude when their dual failure restorabilities R_2 are similar. But p -cycles are less capacity efficient with about 17% more network redundancy compared with that of span restoration.

5 Conclusions

We developed an availability analysis concept, restoration-aware connection availability (RACA), to analyze the unavailabilities of connections in span-restorable mesh networks. According to this new concept, basic computation models were built for span-oriented survivability schemes such as span restoration and p -cycles. The core of our concept is the transformation of the physical structure of a span and its backup paths equipped with spare capacity to an equivalent reliability block diagram by taking into account the positions and sequence of dual failures. Therefore we obtain generic computational functions for connection unavailability affected by all kinds of dual failure scenarios. The functions have been verified by simulations.

With the developed concept combined with restoration details, i.e., the definite and predefined allocation of related spare capacity units to the particular backup paths of a span, we are able to exactly compute connection unavailabilities as the concatenation of equivalent span unavailabilities for networks with shared spare capacity. Previously this could only be conducted by a rough estimate of its availability performance. Even if we do not know any restoration details of an optical transport network, we can still use this concept to find average connection availability by presuming an allocation rule. We can use the average value to compare the overall availability performance of networks with different survivability schemes.

The case study shows that the investigated survivability schemes span restoration and p -cycles with similar levels of dual failure restorability R_2 do not yield a similar availa-

bility performance. Average connection unavailability, the dependent down time and traffic loss measures are about 18% lower for span restoration compared to p -cycles even though span restoration consumes about 14% less spare capacity. This underlines the importance of exact and restoration-aware availability assessment for optical transport networks beyond just comparisons on dual failure restorability.

Acknowledgements This work was partly financed by the Swiss Federal Office for Education and Science BBW (project C01.0087).

Appendix

$N(i, j)$	—	Total non-restorable capacity if a dual failure of span i and j failed (span i failed first and span j second);
N_i	—	Individual non-restorable capacity of span i ;
U_i^{phy}	—	Physical unavailability of span i ;
U_i^*	—	Equivalent unavailability of span i considering restoration effects;
U_{conn}	—	Unavailability of a connection;
$U_{\text{conn}}^{s,t}$	—	Unavailability of the connection $s - t$;
$T(i, j)$	—	Down time caused by the dual-span failure (i, j) in a simulation;
$T_k(i, j)$	—	Down time caused by the k^{th} dual-span failure (i, j) in a simulation;
$\text{LT}(i, j)$	—	Traffic loss caused by the dual-span failure (i, j) in a simulation;
$\text{LT}_k(i, j)$	—	Traffic loss caused by the k^{th} dual-span failure (i, j) in a simulation;
$\text{LT}_{s,t}(a, b)$	—	Lost traffic of connection $s - t$ due to dual failure (a, b) in a simulation;
$\text{EDC}_{s,t}$	—	Expected down time per year for a connection between nodes s and t (hours/year);
AEDC	—	Average expected down time per year of all connections (hours/year);

- RCTL_{s,t} — Restoration-aware connection traffic loss per year between nodes *s* and *t* (Gb/year);
- ARCTL — Average restoration-aware connection traffic loss (Gb/year).

References

- [1] Grover, W. D. (2004). Mesh-based survivable networks, options and strategies for optical, MPLS, SONET, and ATM Networking. New Jersey: Prentice Hall PTR.
- [2] Willebeek-LeMair, M., & Shahabuddin P. (1997). Approximating dependability measures of computer networks: An FDDI case study. *IEEE/ACM Transactions on Networking*, 5(2), 311–327.
- [3] To, M., & Neusy, P. (1994). Unavailability analysis of long-haul networks. *IEEE Journal on Selected Area in Communications*, 12, 100–109.
- [4] Grover, W. D. (1999). High availability path design in ring-based optical networks. *IEEE/ACM Transactions on Networking*, 7, 558–574.
- [5] Zhou, D., & Subramaniam, S. (2000). Survivability in optical networks. *IEEE Network*, 14(6), 16–23.
- [6] Chu, K. C.-H., Meyhoudi, M., & Hu, Y. (2002). Comprehensive end-to-end reliability assessment of optical network transports. *Proceedings of IEEE/OSA Optical Fiber Communication Conference*, (OFC'02) (Anaheim, CA, USA, March 2002) 228–230.
- [7] Guo, W., Guo, J., Jin, Y., Sun, W., & Hu, W. (2005). Optical network design with geographic distribution information. *Proceedings of ISPRS Workshop on Service and Application of Spatial Data Infrastructure*, 36(4/W6), 243–248 (Hangzhou, China, Oct. 2005).
- [8] Zhang, J., Zhu, K., Zang, H., & Mukherjee, B. (2003). Service provisioning to provide per-connection-based availability guarantee in WDM mesh networks. *Proceedings of IEEE/OSA Optical Fiber Communication*, 2, 622–624 (OFC' 03) (Atlanta, GA, USA, March 2003).
- [9] Clouqueur, M., & Grover, W. D. (2002). Availability analysis of span-restorable mesh networks. *IEEE journal on selected areas in communications*, 20(4), 810–821.
- [10] Fredericks, M. T., Datta, P., & Somani, A. K. (2004). Evaluating dual-failure restorability in mesh-restorable WDM optical networks. *Proceedings of 13th international conference on computer communications and networks*, (ICCCN) (Chicago, USA, Oct. 2004) 309–314.
- [11] Sivakumar, M., & Sivalingam, K. M. (2004). On surviving dual link failures in path protected optical wdm mesh networks. *Tech. Rep.*, DAWN Networking Research Laboratory, (June).
- [12] Clouqueur, M., & Grover, W. D. (2005). Availability analysis and enhanced availability design in p-Cycle-based networks. *Photonic Network Communications*, 10(1), 55–71.
- [13] Billinton, R., & Allan, R. N. (1992). Reliability Evaluation of Engineering Systems: Concepts and Techniques. New York and London: Plenum press.
- [14] Birolini, A. (1999). Reliability Engineering — Theory and Practice. Berlin: Springer-Verlag.
- [15] Freeman, R. L. (2002). Reference manual for telecommunications engineering. John Wiley & Sons.
- [16] Zhou, L., Held, M., & Wosinska, L. (2004). Analysis and optimization of connection availabilities in optical networks with different protection strategies. *Proceedings of SPIE Photonics Europe, Reliability of Optical Fiber Components, Devices, Systems, and Networks II*, (Strasbourg, France, September 2004) 5465, 157–167.
- [17] Zhou, L., & Held, M. (2006). Optimization of path availability of span-restorable optical networks. *Proceedings of SPIE Photonics Europe, Reliability of Optical Fiber Components, Devices, Systems, and Networks III*, (Strasbourg, France, May 2006) 6193, 619312.
- [18] <http://www.relexsoftware.com/>.
- [19] A Technical Report on Network Survivability Performance, T1A1.2 Working Group on Network Survivability Performance, Report No. 24., November (1993)
- [20] Rawaswami, R., & Sivarajan, K. (2002). Optical networks, a practical perspective. *Morgan Kaufmann Publishers*, 6, 667–696.
- [21] Mikac, B., & Inkret, R. (2000). Availability model of WDM optical networks. *Proceedings of the Int. Workshop on the Design of Reliable Communication Networks*, DRCN (Munich, Germany, April 2000) 80–85.
- [22] Wosinska, L., & Pedersen, L. (2001) Scalability limitations of optical networks due to reliability constraints. *Proceedings of NFOEC'01* (Baltimore, MD, USA, July 2001) 1529–1540.



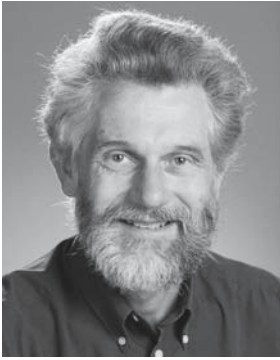
Ling Zhou received the B.S.E.E. and M.S.E.E. degrees in electronic engineering from Southeast University, Nanjing, China, in 1998 and 2001, respectively. She is currently working toward the Ph.D degree with the department of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology Zurich (ETHZ) and the Electronics/Metrology Laboratory and Reliability Center, Swiss Federal Laboratories for Materials Testing and Research (EMPA), CH-8600

Duebendorf Switzerland. She will finish her Ph.D study in the beginning of 2007. She is working on a topic of Availability Analysis and Optimization in Optical Transport Networks. Her research is in the area of optical transport networks including network architectures and protocols (multilayer networks, IP/MPLS networks, Ethernet networks), routing and wavelength assignment, recovery methods (protection and restoration), availability and reliability analysis, network optimization, and network planning.



Marcel Held received the Engineer degree for Communication Technology and Informatics at the Zurich University of Applied Science in Winterthur in 1988. He specialized in reliability engineering at the Reliability Laboratory of the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, mainly in the field of reliability and availability of microelectronic and power electronic components and systems. Since 1998, he has been a Senior Engineer for the Reliability Center at the Swiss

Federal Laboratories for Materials Testing and Research (EMPA), Duebendorf, Switzerland and Lecturer on reliability at ETH, Zurich. He is a member of the Swiss IEEE Reliability Chapter.



Urs Sennhauser is heading the Electronics/Metrology Laboratory and the Reliability Center at the Swiss Federal Laboratories for Materials Testing and Research (EMPA). He has a Ph. D. degree in Physics from the Swiss Federal Institute of Technology (ETH). He is lecturing “Reliability of Devices and Systems” and “Physics of Failure and Failure Analysis of Electronic Circuits and Devices” at ETH-Zurich. In addition to national affiliations he is an active member

in international professional organizations like IEEE, APS and SPIE.