Relaxed Maintenance Protection Architecture by Dynamic Backup Path Configuration

Shohei Kamamura, Takashi Miyamura, and Kohei Shiomoto
NTT Network Service Systems Laboratories, NTT Corporation 9-11, Midori-Cho 3-Chome Musashino-Shi, Tokyo, 180-8585 Japan
E-mail:{kamamura.shohei, miyamura.takashi, shiomoto.kohei}@lab.ntt.co.jp

Abstract: We propose dynamic backup path configuration architecture for maintaining the 1+1 path protection as much as possible after a single failure occurs. Our reliable architecture ensures enough time for repairing the failure component.

OCIS codes: (060.4261) Networks, protection and restoration; (060.4257) Networks, network survivability

1. Introduction

The 1+1 path protection method [1] is effective for reliable optical transport network design. 1+1 path protection provides primary and backup paths, which are link- or node-disjoint paths. Traffic data is transmitted on both paths, and the receiver node switches receiving data; therefore, it can achieve lossless data transmission. However, if the primary or backup path becomes unavailable because of failure or planned shutdown, availability decreases and miss operation probably increases because one-path operation is not common. So, network operator would like to repair the failure component as quickly as possible. However, vital failure such as fiber cut, it accounts for about 11% of failure [2], requires many time for repairing.

Our problem is to maintain 1+1 path protection as much as possible after a single failure occurs for achieving high availability. By maintaining a high reliability even if a failure occurs, we can assign enough time for repairing the failure component. Since bandwidth of current optical backbone is becoming huge (e.g. 40/100Gbps), preparing another backup path is good and practical strategy.

We propose dynamic backup path configuration architecture with a central control server to maintain 1+1 path protection as much as possible. When a failure occurs and primary or backup path failed, we dynamically establish new backup path to maintain 1+1 condition. These processes are controlled by the central control server. The central control server monitors network state and configures a new backup path through a control plane. We also formulate the availability of our dynamic protection architecture considering existing network topologies where three disjoint paths do not always exist.

By introducing our dynamic protection architecture, we can assign enough maintenance time for achieving six's nine (99.9999%) reliability. From our evaluation, our architecture can assign about more than 60 hours for 84% of link failure repairing while conventional 1+1 path protection only assigns about six hours.

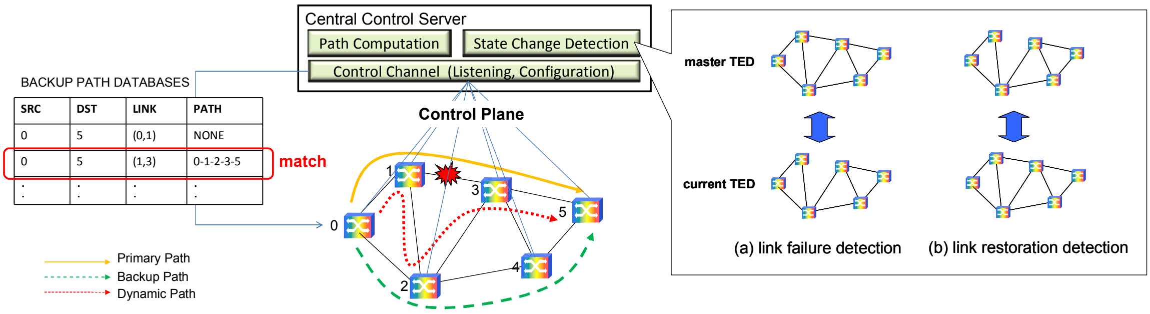
2. Dynamic Protection Architecture

Our architecture is composed of a central control server and network elements (NEs) (Fig. 1). The central control server has i) path computation function, ii) state change detection function, and iii) control channel. Path computation performs when server starts up, and state change detection dynamically detects network state change such as a failure occurrence. The control channel of central control server and each NE are connected by the control plane, which has IP reachability. The central control server monitors network state and configures a new backup path through a control channel. Initial setup processes and dynamic control processes are described as following subsection.

2. 1. Initial Setup

Main processes of initial setup are getting network state and creating the backup path database. On a control plane, adjacency relationship of each NE is advertised by OSPF-TE protocol [3]. The central control server listens to adjacency relationships through control channel, and then concatenates them and creates a traffic engineering database (TED) that represents the overall connection relationships. The TED includes not only connection relationships but also link metrics, and bandwidth information for optimal path computation. The TED created on initial setup phase are saved as a master TED for state change detection function described later.

Then, the central control server computes primary paths and backup paths by referring to the TED. The primary path and backup path are computed as link- or node-disjoint paths that do not share the same links or nodes [4]. In this paper, we only consider link-disjoint path and a link failure because of space limitation. But our architecture can also handle a single node failure in the same way. The primary and backup paths are saved to 1+1 path database whose record is composed of *(src, dst, primaryPath(up/down), backupPath(up/down))*. Then, dynamic paths are



SRC	DST	LINK	PATH
0	5	(0,1)	NONE
0	5	(1,3)	0-1-2-3-5
:	:	:	:

Fig.1 Overview of our dynamic protection architecture

computed assuming possible link failure. When a link that is included in the primary path (backup path) fails, a shortest path excluding a failed link and links on backup path (primary path) are computed as dynamic path. Figure 1 shows an example. A primary path is 0-1-3-5, and a backup path is 0-2-4-5 from node 0 to node 5. Then, dynamic path becomes 0-1-2-3-5 after link 1-3 fails. Dynamic paths are stored in the backup path database whose record is composed of *(src, dst, failed/repaired link, dynamicPath)*. In this paper, dynamic path is computed on an initial setup phase for reducing dynamic control time, but if we would like to handle more complicated failure (e.g. concurrent failures), dynamic path may be computed on a dynamic control phase.

2. 2. Dynamic Control

Main processes of dynamic control are network state change detection and path configuration. Path configuration is triggered by recovery process when failure occurs and reversion process when a failed link is repaired. A purpose of reversion is to keep path quality because quality of former primary and backup path is higher than dynamic path.

Network state change is detected by comparing a master TED and currently advertised TED. For example, when a link 1-3 fails (Fig.1), adjacency relationship is advertised as that node 1 and node 3 has no connectivity. On the other hand, on a master TED, there is connectivity between node1 and node 3. From this difference, the control server can detect a failure of link 1-3. After a failure detection, the master TED is overwritten in the current TED.

After a single link failure occurs, new backup paths are selected by referring to the backup path database with a failed link ID as a key. Figure 1 also illustrates a backup path database, and there one record matches with a key. Usually, multiple flows are included in one link; therefore, multiple records match with a key. Primary or backup paths of these flows are down. For these down flows, down flags of their primary or backup paths are set to true on 1+1 path database. After selecting new backup paths, the central control server orders the signaling for path establishment to source nodes of these flows. RSVP-TE [5] is used as a signaling protocol.

After a failure is repaired, our architecture performs reversion processes. A repaired link is detected in the same way for failed link detection. And then, withdrawal dynamic paths are chosen by referring to the backup path database with a repaired link as a key. Firstly, these paths are released because they share the resources with former primary or backup paths. Then, primary or backup paths, whose down flag are true, are reconfigured. After path establishments, down flags of reconstructed paths are set to false, and then the master TED is overwritten in the current TED.

3. Analysis of Maintenance Time

We formulate the availability of our dynamic protection architecture, and then evaluate effectiveness in terms of maintenance time.

3.1. Mathematical Model

A network is denoted as directed graph $G=(V,E)$ with nodes V and links E. A mean time between failure (MTBF) of link $e \in E$ is denoted as $MTBF(e)$, and a mean time to repair (MTTR) is denoted as $MTTR(e)$.Then, availability of link e is denoted as $MTBF(e)/(MTBF(e)+MTTR(e))$ [6]. Network has multiple source and destination pairs (flows), therefore, we modeling the availability for flows included in a certain link. The set of flows that are included in link $x \in E$ is denoted as F_x. On flow $f \in F_x$, the set of links that compose primary path, backup path, and dynamic path are denoted as $E_p(f)$, $E_B(f)$, and $E_D(f)$ respectively. Then, availabilities of primary path A_p, backup path A_B, and dynamic path A_D become direct product of availability of each link set, $E_p(f)$, $E_B(f)$, and $E_D(f)$ respectively. Availability of 1+1 path protection is denoted as $A_{1+1}(x)$ that is parallel

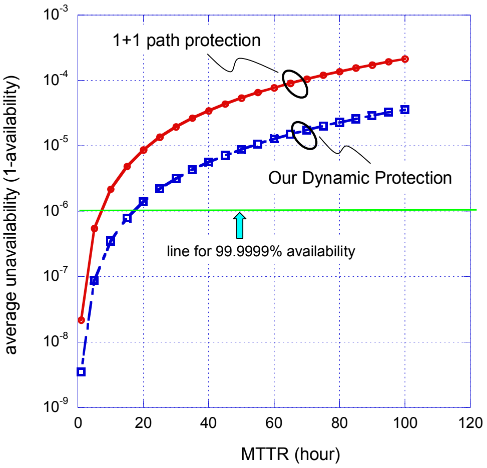


Fig.2 Numerical result of unavailability. X-axis means repairing time (MTTR). Simulation topology is COST266 [7] model (26 node, 98 link, average node degree=3.76). MTBF is set to 26280 hour (3 year).

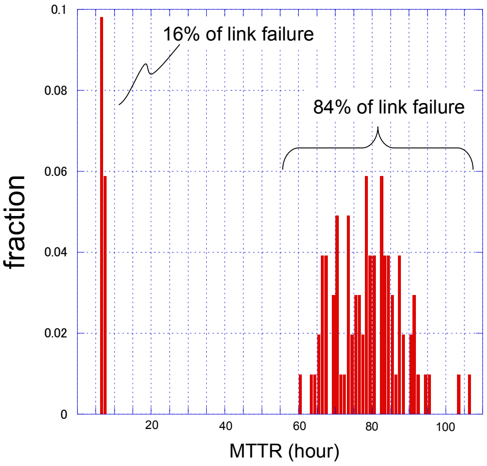


Fig.3 MTTR of each link for achieving six's nine availability.

structure of A_p and A_B, and it of our dynamic architecture on an ideal environment is denoted as $A_d^{ideal}(x)$ that is parallel structure of A_p, A_B, and A_D. Ideal environment means that topology is 3-edge connected graph. However, the dynamic path does not always exist for flow f; if the minimum cut between source node and destination node is less than three, we cannot recover the link failure that is composed of the minimum cut (e.g. link 0-1, 0-2 on Fig. 1). Therefore, we define the fraction β_x that are fraction of flows that have own dynamic path, and are included in link x. Then, availability of our architecture $A_d(x)$ is denoted as follows;

$$A_d(x) = \beta_x \times A_d^{ideal}(x) + (1-\beta_x) \times A_{1+1}(x) \qquad (1)$$

3.2. Numerical Result

Figure 2 shows an average unavailability ($1-A_{1+1}, 1-A_d$) with variable $MTTR(e)$, and an average unavailability is computed by assuming all possible link $x \in E$ failure. Our dynamic architecture is more redundant than 1+1 path protection, therefore, unavailability of ours always becomes low on the same parameters ($MTBF(e)$, $MTTR(e)$).

Figure 3 shows MTTR of each link for achieving six's nine (99.9999%) availability. Using our dynamic architecture which has high availability, MTTR can be set to high value for achieving a six's nine availability; we can assign enough time for repairing. For example, while MTTR for 16% of link failure is 6 or 7 hour (that is same with it of 1+1 path protection), it for 84% of link failure is more than 60 hour. The ratio of links with high MTTR depends on network topology. If topology is a ring graph, MTTR of every link becomes low. However, existing topology tend to be partial 3-edge connected, and dynamic architecture provides not path-disjoint but failed-link disjoint path, therefore, the ratio of links with high MTTR tends to increase. From our evaluation on other existing topologies, the ratio of links with high MTTR is about 99% (Italian National Network, 33 nodes, 136 links) as the best case, and is still about 63% in the worst case (German Test Case, 17 nodes, 52 links).

Relaxing maintenance time provides the OPEX reduction; with conventional 1+1 path protection, network operator should rapidly go failed place even if failure occurs at night. On the other hand, our architecture tolerates time-rich repairing for most failure.

4. References

[1] S. Ramamurthy and B. Mukherjee, "Survivable WDM Mesh Networks, Part I – Protection," in proceedings of IEEE INFOCOM, vol. 2, pp. 744-751, Mar. 1999.

[2] A. Markopoulou, et al., "Characterization of Failures in an IP Backbone," in Proc. IEEE INFOCOM, vol. 4, pp. 2307-2317.Mar. 2004.

[3] D. Katz, K. Kompella, D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2," IETF RFC 3630 Sep 2003.

[4] J. Suurballe, "Disjoint paths in a network," Networks, vol. 4, pp. 125 –145, 1974.

[5] D. Awduche, et al., "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, IETF, December 2001.

[6] ITU Recommendation G.911, "Parameters and calculation methodologies for reliability and availability of fiber optic systems," ITU-T Standardization Organization, April 1997.

[7] R. Inkret, A. Kuchar, and B. Mikac, "Advanced Infrastructure for Photonic Networks European Research Project," in Extended Final Report of COST266 Action, ISBN 953-184-064-4,p.20, 2003.