

A System for Managing Transport-network Recovery

Using Hybrid-backup Operation Planes according to Degree of Network Failure

Toshiaki Suzuki, Hiroyuki Kubo,
Hayato Hoshihara, and Kenichi Sakamoto

Research & Development Group
Hitachi, Ltd.
Kanagawa, Japan

E-mails: {toshiaki.suzuki.cs, hiroyuki.kubo.do,
hayato.hoshihara.dy, and
kenichi.sakamoto.xj}@hitachi.com

Hidenori Inouchi and Taro Ogawa

Information & Telecommunication Systems Company
Hitachi, Ltd.
Kanagawa, Japan

E-mails: {hidenori.inouchi.dw and
taro.ogawa.tg}@hitachi.com

Abstract—A system for managing transport-network recovery using hybrid-backup operation planes according to the degree of a network failure is proposed. Under this management system, an entire network is separated into multiple areas. A network-management server prepares a three-step recovery procedure to cover the degree of network failure. In the first step of the recovery, an inside-area protection scheme is used to recover current data-transmission paths in each area. In the second step, an end-to-end protection scheme is applied to the current data-transmission paths. In the third step, the operation plane is changed. Each assumed operation plane is composed of recovery configurations for restoring failure paths for assumed area-based network failures. If a small network failure occurs, it is recovered by the inside-area protection and end-to-end protection schemes. If a catastrophic network failure (caused by a disaster) that cannot be recovered by those protection schemes occurs, it is recovered by changing the operation plane in accordance with the damaged areas. A prototype system composed of a network-management server and 96 emulated packet-transport nodes was developed and evaluated by configuring 1000 data-transmission paths. In case of a small network failure, 500 data-transmission paths were damaged, and they were reconfigured by the inside-area protection scheme and end-to-end protection scheme in about 5 seconds. If the network failure was not recovered by those protection schemes, 1000 data-transmission paths were reconfigured in about 1.2 seconds after the network-management server decided to change the operation plane. As a result, the proposed system could localize a network failure and recover a transport network according to the degree of network failures.

Keywords - network management; protection; disaster recovery; packet transport

I. INTRODUCTION

Lately, reflecting the rapid growth of the Internet and cloud systems, various services, for example, on-line shopping, net banking, and social-networking services (SNSs), are being provided via networks. Under these circumstances, networks have become an indispensable service supporting daily life. If a network is out of service due to failures of network nodes, people's lives and

businesses would be considerably damaged. Therefore, if a network fails, it should be recovered promptly. Failures of a network can be envisioned as “small” failures (such as a failure of a node or a link) or “extensive” failures (due to natural disasters). It is therefore a crucial issue to develop a scalable network-recovery scheme that can cover recovery from either a small network failure or a catastrophic network failure.

In our previous work presented at INNOV 2015 [1], an entire system architecture was focused on a scalable network-recovery scheme by extending a prior system [2]. In this extended work, a prototype system for multiple tenant users was implemented, and its performance was evaluated in comparison with a conventional system.

As recovery procedures for network failures, two major schemes [3], namely, “protection” and “restoration,” are utilized. As for protection, it is possible to recover from a network failure promptly because a backup path to a current path is prepared in advance. However, to recover from a network disaster, numerous backup paths must be prepared. Protection is therefore useful for small network failures. On the other hand, as for restoration, a recovery path is recalculated after a network failure is detected. It therefore takes much time to recover from a network failure if numerous current paths exist.

In light of the above-described issues, a robust network-management scheme is required. The overall aim of the present study is thus to develop a network-management scheme [1][2] for monitoring and controlling network resources so as to quickly restore network services after a network disaster.

The procedure for recovering from a network failure consists of three steps: the first step is to quickly detect a network failure; the second is to immediately determine how to recover from the failure; the third is to promptly configure recovery paths. The second step is focused on in the present study. In particular, a scalable network-recovery scheme—covering failures ranging from small ones to extensive ones—is proposed. The target network is a transport network, such as a Multi-Protocol Label

Switching - Transport Profile (MPLS-TP) network.

The rest of this paper is organized as follows. Section II describes related work. Section III overviews a previously proposed system and a requirement to apply it to not only catastrophic failures but also small network failures. Section IV proposes a new network-disaster recovery system. Sections V and VI respectively describe an architecture of a prototype system and present some results of evaluations of the system's performance. Section VII concludes the paper.

II. RELATED WORK

Several standardization activities related to reliable networks have been ongoing. The International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) [4] discussed specifications, such as Transport - Multi Protocol Label Switching (T-MPLS), in the first stage of standardization. In the next stage, the ITU-T jointly standardized MPLS-TP specifications with the Internet Engineering Task Force (IETF) [5]. Requests for comments (RFC) on requirements [6] and a framework [7] for MPLS-TP were then issued. In addition, RFCs on a framework for MPLS-TP-related operation, administration, and maintenance (OAM) [8] and survivability [9] were issued. Based on the OAM framework, the previously proposed system can detect network failures promptly.

Several schemes for failure recovery have been proposed. One major scheme, called "fast reroute" [10], prepares a back-up path. Another recovery scheme (for multiple failures) prepares multiple backup paths [11], and another one prepares a recovery procedure for multiple modes [12]. In the case of these protection schemes, to recover from a catastrophic network failure, a huge volume of physical resources for preparing a large number of standby paths is needed. These schemes are useful for limited network failures, namely failures of a few links or nodes.

In the case of restoration schemes, in contrast to protection schemes, recovery paths are calculated after a network failure is detected. Restoration schemes for handling multiple failures [13] and virtual networks [14] have been proposed. A scheme for reducing search ranges by using landmark nodes has also been proposed [15]. It is useful for recovering a seriously damaged network, since all reroutes are calculated after a failure is detected. However, if a large number of current paths exist, it might take much time to calculate all recovery paths.

III. PREVIOUS SYSTEM AND REQUIREMENTS

The previously proposed network-recovery system is shown in Figure 1 [2]. The target network is composed of packet transport nodes (PTNs), such as those in an MPLS-TP network. The system only focuses on recovery from multiple area-based network failures on PTN networks. A critical issue in the case of a network disaster is the time consumed in recovering the numerous established paths

(shown as solid blue arrows) in packet networks. (Note that "path" means a label-switched path (LSP) [16] and a pseudo wire (PW) [17].) A user is connected to one of the PTNs through a network such as an IP network. A server located in a data center (DC) is also connected to one of the PTNs through an IP network.

The previously proposed system could promptly recover from a catastrophic failure of a network by using prepared back-up paths (shown as dotted red arrows). However, it significantly changes network configurations, even if a network failure is small, since network conditions are managed on the basis of divided network areas. It must therefore be enhanced so that it can recover from a catastrophic network failure, as well as a small network failure, by using fewer configurational changes based on the degree of damage due to the network failure.

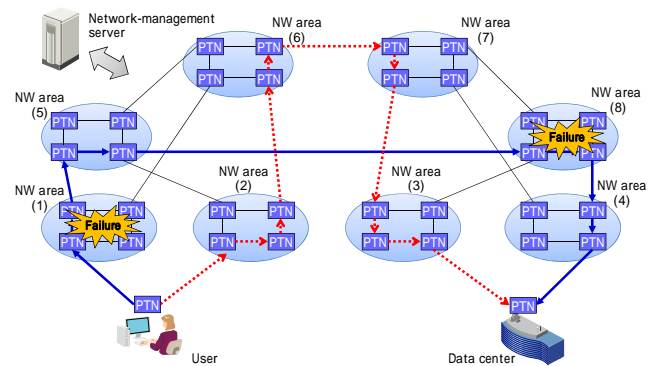


Figure 1. Previously proposed network-recovery system

IV. PROPOSED TRANSPORT NETWORK-RECOVERY SCHEME

To meet the above-described requirements, a three-step recovery procedure for covering the degree of network failures is proposed. The first step of the procedure is to execute "an inside-area protection scheme" to recover small failures, such as a node failure in each area formed by separating an entire network into small areas. The second step is to execute "an end-to-end protection scheme" to recover small failures, such as a failure of a link between areas not recovered by the inside-area protection scheme. The third step is to execute "an operation-plane change scheme" to recover extensive failures, such as network failures of multiple areas.

A. Path protection for small network failures in each area

The proposed system should promptly recover a network from a small failure, such as a link failure between PTNs or a PTN failure. A scheme called "inside-area protection"—for localizing and swiftly recovering from a small network failure—is overviewed in Figure 2. Using a conventional scheme (such as cluster analysis), the network-management server divides an entire PTN network into multiple (e.g.,

eight) areas, which it then manages. It configures a current path (shown as solid black arrows in the figure), composed of a LSP and a PW, for transmitting data from a sender to a receiver according to requests by end users. The network-management server configures a backup path for each current path, namely, an inside-area protection path (shown as dotted red arrows), between one edge PTN and another edge PTN in every area. Specifically, the network-management server finds an edge PTN pair that is related to the current path in every area. For example, PTN 14 and PTN 11 are the edge PTN pair in area (1), since packet data from PE1 are received by PTN 14 and then transmitted to area (5) by PTN 11, as shown in Figure 4. In addition, PTN 54 and PTN 53, PTN 84 and PTN 83, and PTN 42 and PTN 43 are the edge PTN pairs that are related to the current path. The network-management server calculates a detour path for each edge PTN pair by excluding network links that are parts of the current path. For example, a detour path between PTN 14 and PTN 11 through PTN 13 and PTN 12 is calculated as the backup path. All calculated detour paths in each area become the inside-area protection paths.

In each area, both edge PTNs exchange OAM packets to check if a disconnection exists between the PTNs. If a disconnection is detected, they send an alert to the network-management server, which keeps the received alert and monitors the degree of failures, namely, numbers of link and PTN failures, and damaged areas.

In the case shown in Figure 2, a link failure between PTN 14 and PTN 11 is assumed to occur in area (1). PTN 14 and PTN 11 detect the link failure, which is recovered by the inside-area protection. Specifically, a direct data-transmission path from PTN 14 to PTN 11 is changed to a backup transmission path through PTN 13 and PTN 12. On the other hand, the path between PTN 14 and PTN 11 is a part of an end-to-end path between provider-edge 1 (PE1) and PE2. The link failure between PTN 14 and PTN 11 is therefore temporarily detected by PE1 and PE2, since both PEs also exchange OAM packets. However, both PEs wait for 100 milliseconds to see whether the link failure is recovered by the inside-area protection. Therefore, when the link failure is recovered by the inside-area protection, neither PE executes further recovery action.

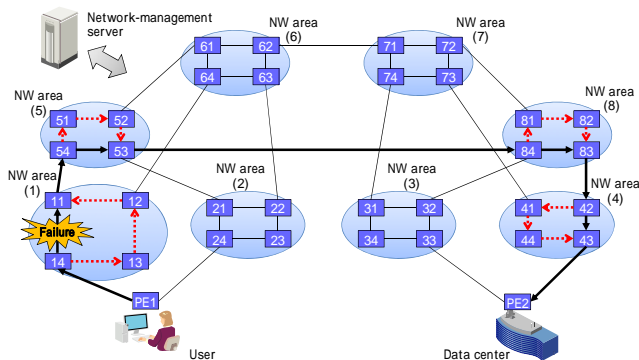


Figure 2. Configuration of path protection in each NW area

B. End-to-end path protection for small network failures

The proposed system should be able to immediately recover from a small failure that is not recovered by the above-described protection (such as a link failure between areas). A scheme called “end-to-end protection” to promptly recover from a failure that is not restored by the inside-area protection is overviewed as follows. The network-management server configures a backup path (called an “end-to-end protection path”) for each current path between PE1 and PE2. PEs exchange OAM packets to check whether a disconnection exists between them.

Specifically, as shown in Figure 3, the network-management server configures a current path (shown as solid black arrows) between PE1 and PE2 [through areas (1), (5), (8), and (4)] for transmitting data packets between a user and a DC. In addition, the network-management server configures a backup path called an “end-to-end protection path (shown as dotted red arrows)” between PE1 and PE2. The end-to-end protection path is established so as not to travel through the same areas used by the current path as much as possible. In Figure 3, the backup path is configured to transmit data through areas (2), (6), (7), and (3).

During network operation, the end-to-end protection is executed when the data transmission between PEs is disconnected for a while (for example, 100 milliseconds). In the case of Figure 3, a link failure between areas (5) and (8) is assumed. This failure is not recovered by the inside-area protection; instead, it is recovered by the end-to-end protection because it occurs between areas. Specifically, a data-transmission path is changed from the current path (shown as solid black arrows) to a backup path (shown as dotted red arrows).

The end-to-end protection scheme is similar to a conventional protection scheme. In the case of a conventional scheme, the protection is immediately executed after one of the PEs detects a disconnection. However, in the case of the proposed end-to-end protection scheme, it is not executed for 100 milliseconds so that it can be checked whether a failure has been recovered by the inside-area protection or not.

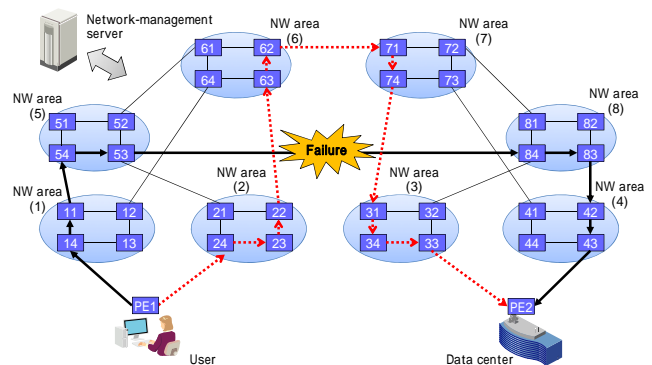


Figure 3. Configuration of path protection for end-to-end transmission

C. Changing operation plane for network-disaster recovery

The proposed system should be able to promptly recover not only failures inside a network area and between network areas but also catastrophic failures. A recovery scheme that changes the operation plane to recover from area-based network failures is overviewed in Figure 4. Before starting network operations, the network-management server prepares multiple backup operation planes for handling possible area-based network failures. Each backup operation plane is composed of recovery configurations for restoring failure paths due to assumed network failures. During network operation, if network failures are not recovered by both the inside-area protection and the end-to-end protection, the failures are recovered by changing an operation plane.

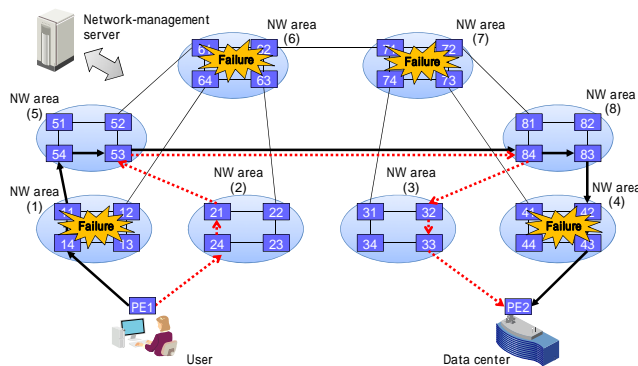


Figure 4. Configuration for changing an operation plane for network-disaster recovery

In Figure 4, as an example, the network-management server configures multiple current paths [through areas (1), (5), (8), and (4)] for transmitting data packets between a user and a DC. It calculates all recovery paths preliminarily by assuming all possible area-based network failures. The number of possible combinations of areas is 256 (i.e., 2^8), and it includes a pattern by which no area-based network failure occurs. The network-management server therefore prepares 255 backup operation planes. It then assigns a unique recovery identifier (ID) for each backup operation plane, and sends all recovery IDs and recovery configurations to each PTN, which stores all received recovery IDs and configurations.

An example area-based network-failure recovery procedure is shown in Figure 4. In the figure, area-based network failures are assumed to occur in areas (1), (4), (6), and (7). In this case, PE1 (namely, an edge node of the current path) detects a disconnection between PE1 and PE2. PE1 waits 100 milliseconds to check whether the failures are recovered by the inside-area protection. It also checks the availability of the end-to-end protection path (which is not shown in Figure 4) by using OAM packets. If the failures are not recovered in 100 milliseconds and the end-to-end protection path is not available, PE1 sends an alert to the network-management server to inform it that the end-to-

end protection is not available. The network-management server then checks which areas are not available. In this example, by receiving many alerts sent by multiple PTNs, the network-management server determines that area-based network failures occur in areas (1), (4), (6), and (7). By using the determined network-failure information, it then determines the most suitable backup operation plane to recover. To change an operation plane, the network-management server sends a recovery ID specifying the most-suitable backup operation plane to related PTNs, which change data transmissions according to the received recovery ID. By means of the above-described procedures, the operation plane is changed, and catastrophic network failures are swiftly recovered.

V. ARCHITECTURE OF PROTOTYPE SYSTEM

In this section, the architecture of a prototype system is described. Specifically, the structure of the prototype system is shown first. Then, recovery procedures are overviewed. After that, calculation procedures for the inside-area protection paths and the end-to-end protection paths are described. (Note that calculation procedures for the backup operation planes are not described since they are explained in a previous work [2].) At the end of this section, an implemented viewer is depicted.

A. Structure of prototype system

A prototype system was implemented by using three servers. The structure of the prototype system—composed of an application server, a control server, and a node simulator server—is shown in Figure 5. Specifically, implemented software components are shown in the figure.

The application server is in charge of the entire network management. Specifically, it manages calculation and configuration of current paths and protection paths by sending commands. In addition, it calculates and configures back-up operation planes with multiple detour paths by assuming possible node failures or area-based network failures. Besides, it receives alerts and determines the degree of network failures. It then selects a recovery back-up operation plane and sends it an identifier specifying it to network nodes.

The control server is in charge of transmitting command messages from the application server to the simulator server. Specifically, it receives calculated route information of the current paths, protection paths, and back-up operation planes and distributes it to the simulated multiple network nodes. In addition, it monitors state of connections between not only current paths but also protection paths. When it detects a disconnection, it prompts the node simulator server to activate an alert. On the other side, it transmits alert information from the simulator server to the application server.

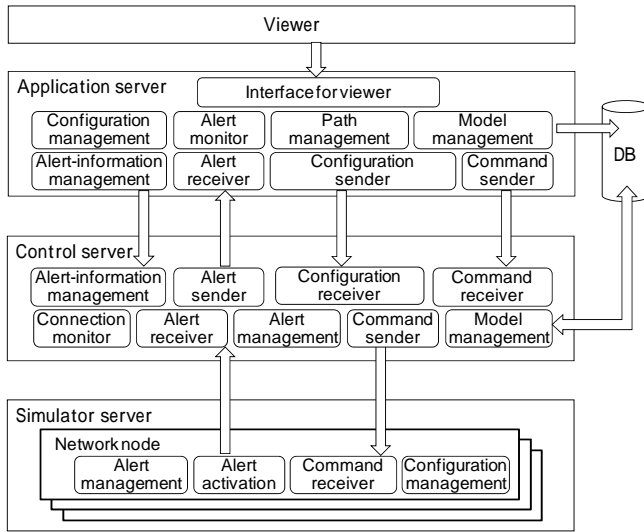


Figure 5. Structure of prototype system

The simulator server emulates certain parts of the functions of MPLS-TP network nodes. It receives configurations of LSP and PW paths and sets data-transmission paths on the basis of the received path information. When it is requested to activate alerts, it sends an SNMP trap to the control server.

B. Overview of recovery procedures

The structure of the proposed transport network-recovery scheme is similar to the previously proposed scheme (shown in Figure 1). Namely, it is composed of a network-management sever and multiple PTNs. The network-management server centrally manages the whole network. However, the recovery procedures differ from those of the previous system.

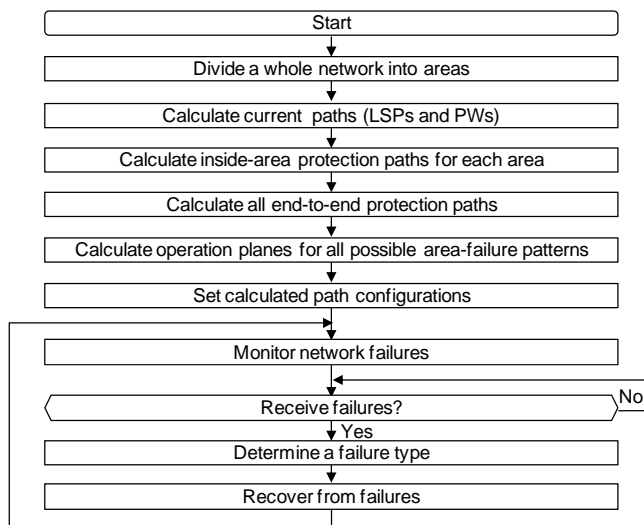


Figure 6. Overview of proposed recovery procedure

A flow chart of the new recovery procedure is shown in Figure 6. First, after starting a network-management function, the network-management server divides the whole network into multiple areas. It calculates current paths (composed of LSPs and PWs) for transmitting data from a sender node to a receiver node according to inputs by a network manager. The network-management server calculates “inside-area protection paths” for each area and “end-to-end protection paths” to recover current paths in case of network failures. In addition, it calculates virtual operation planes for all possible area-failure patterns. The protection paths and virtual operation planes are described in detail in later sections. The network-management server sets the entire configuration of the calculated paths to all network nodes and starts to monitor the network for failures. When it detects a network failure, it determines the type of failure, namely, an area-based or node-based failure. The network-management server then executes the appropriate failure-recovery procedures according to the determined failure degree.

C. Calculation of inside-area protection paths

An implemented flow chart of the calculation procedure of inside-area protection paths is shown in Figure 7.

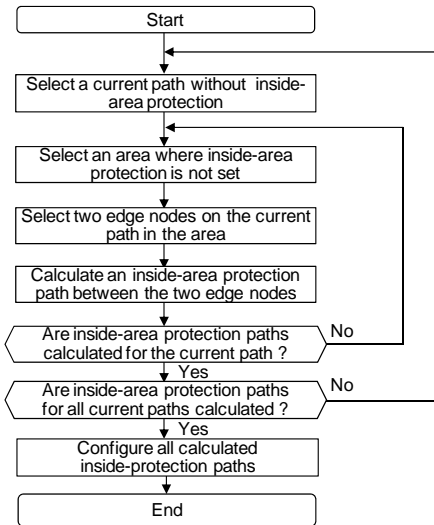


Figure 7. Calculation procedure for inside-area protection paths

The application server selects a current path that does not have an inside-area protection path. In addition, it selects an area where inside-area-protection for the selected current path is not set. Then, it selects two edge nodes on the selected current path in the selected area. One is the start point and the other is the end point for the selected path in the selected area. The application server calculates the inside-area protection path between the two selected edge nodes and stores it. After that, it checks whether the inside-area protection paths related to the selected current path are calculated or not. If they are calculated, it checks whether

all inside-area protection paths for all current paths are calculated or not. If they are not calculated, it calculates other inside-area protection paths for the remaining current paths. If all inside-area protection paths are calculated and stored, it terminates the calculation procedures.

D. Calculation of end-to-end protection paths

An implemented flow chart of the calculation procedure for end-to-end protection paths is shown in Figure 8. The application server selects a current path that does not have an end-to-end protection path. In addition, it sets a high cost value for links in areas that the current path passes through. It then calculates an end-to-end protection path for the selected current path to minimize the cost of the summation of the links composing the protection path. After that, it checks whether all end-to-end protection paths for all current paths are calculated or not. If they are not calculated, it calculates other end-to-end protection paths for the remaining current paths. If they are calculated and stored, it terminates the calculation procedure.

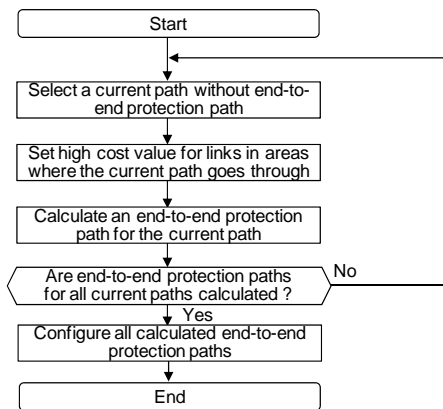


Figure 8. Calculation procedure for end-to-end protection paths

E. Calculation of recovery paths for operation-plane change

An implemented flow chart of the calculation procedure for recovery paths when changing the operation plane is shown in Figure 9. The detailed calculation procedure is described in our previous study [2]. The application server selects one of all possible patterns of area failures that does not have a backup operation plane. In addition, it excludes all nodes in the selected failure pattern of area failures. It then selects a current path that does not have a recovery path and calculates a recovery path for the selected current path. After that, it checks whether all recovery paths for the selected pattern of area failures are calculated or not. If they are not calculated, it calculates other recovery paths for the selected pattern of area failures. If they are calculated, it calculates recovery paths for other patterns of area failures. If all recovery paths for all possible patterns of area failures are calculated, it terminates the calculation procedure.

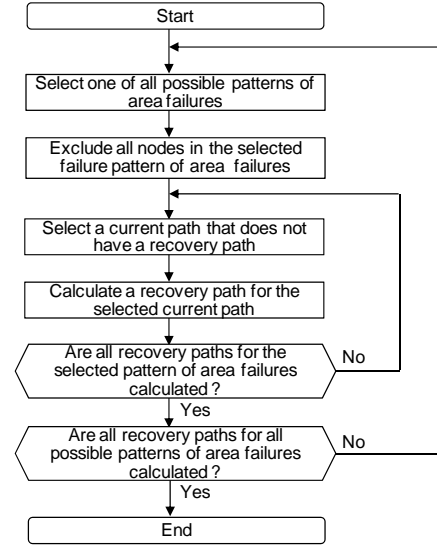


Figure 9. Calculation procedure for recovery paths for changing operation plane

F. Implementation of viewer

The primary-screen layout of the prototype system’s viewer is shown in Figure 10. The function menu allows selection of a topology view or a system-configuration view. The operation ID means the number of a selected operation planes. If no failure occurs, the number zero is used. The recovery indicator shows conditions after execution of one of the recovery procedures, namely, inside-area protection, end-to-end protection, and a selection of a backup operation plane. The condition panel shows current operational status of the system. The topology tree shows a list and structure of connected nodes. The alert panel shows a list of failures, such as node failures. The “area object” tag indicates an existence of each area. The “user terminal” tag indicates each user terminal. The map location indicates the position of the displayed network.

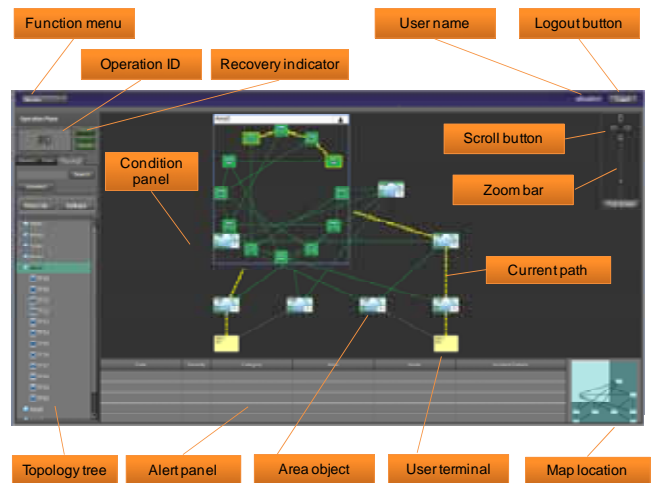


Figure 10. View of primary-screen layout

The “current path” tag highlights the currently used path. The “zoom bar” button provides a function to change the size of the displayed network. The “scroll” button provides a function to change the position of the displayed network. The “user name” tag shows the name of a current user. The “logout” button is used to terminate network management.

VI. PERFORMANCE EVALUATION AND RESULTS

The above-described recovery procedures were evaluated in the case of a small network failure and a catastrophic network failure by using the prototype system. In the evaluation, the times needed to calculate and to configure a table for current data-transmission paths (composed of PWs and LSPs) were evaluated. In addition, the times taken to configure recovery paths in the case of a failure of a PTN or an area-based failure were evaluated.

A. Evaluation system

The system used for evaluating the proposed recovery procedures is shown in Figure 11. It is composed of a network-management server and 96 PTNs. As shown in the figure, an entire PTN network is divided into eight areas. Each network area is composed of 12 PTNs, as shown in NW area (7). In each area, PTNs are connected in a reticular pattern. In addition, each user terminal is connected to PTN-network areas (1) and (2) through PE1 or PE3, and each application server in DC1 or DC2 is connected to PTN-network areas (3) and (4) through PE2 or PE4.

Note that the PTN networks (composed of 96 PTNs) are emulated by a physical server. The user terminal and application server are also emulated by the physical server, whose specification is listed in Table I. Another physical server, which executes the network-management function, has the same specifications as the former server.

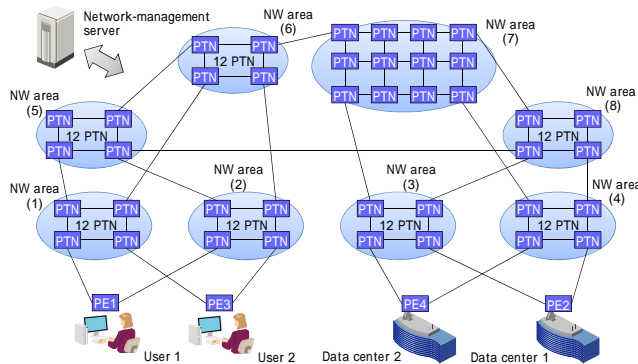


Figure 11. Evaluation system

TABLE I. SPECIFICATION OF SERVER

#	Item	Specification
1	CPU	1.8 GHz, 4 cores
2	Memory	16 Gbytes
3	Storage	600 Gbytes

TABLE II. EVALUATED ITEMS

#	Item	Evaluation specification
1	Current-path calculation time	Time to calculate 100 (=50+50), 500 (=250+250), and 1000 (=500+500) PWs
2	Current-path distribution time	Time to distribute all calculated current paths in case of 100 (=50+50), 500 (=250+250), and 1000 (=500+500) PWs
3	Protection-path calculation time for each area	Time to calculate all protection paths in each area for 100 (=50+50), 500 (=250+250), and 1000 (=500+500) PWs
4	Protection-path calculation time for end-to-end protection paths	Time to calculate all protection paths for all end-to-end current paths for 100 (=50+50), 500 (=250+250), and 1000 (=500+500) PWs
5	Recovery-path calculation time for changing operation plane	Time to calculate recovery 100 (=50+50), 500 (=250+250), and 1000 (=500+500) PWs for all possible area-failure patterns
6	Recovery-configuration time	Time to configure all protection paths after detecting path failures for 100 (=50+50), 500 (=250+250), and 1000 (=500+500) PWs
7	Recovery-ID distribution time	Time to distribute a recovery ID after detecting an area failure for 100 (=50+50), 500 (=250+250), and 1000 (=500+500) PWs

B. Evaluation conditions

The times taken to calculate multiple PWs between PE1 and PE2 and between PE3 and PE4 were evaluated. Each PW was included in a LSP. If a transmission path of a PW differed from the path of an already setup LSP, a new LSP was setup, and the PW was included in the new LSP. The evaluations were executed according to the patterns listed in Table II. Specifically, the times taken to calculate current paths, to distribute their configuration to all PTNs, and to calculate the inside-area protection paths and end-to-end protection paths were evaluated by changing the number of PWs (namely, 50+50, 250+250, and 500+500 for two users). In addition, the times taken to calculate recovery paths for changing the operation plane, to configure protection paths, and to distribute the recovery ID were evaluated.

C. Evaluation results

1) Current-path calculation time

The times taken to calculate current PWs requested by the two users are plotted in Figure 12. User 1 accesses a server in DC1 through PE1 and PE2. User 2 accesses a server in DC2 through PE3 and PE4. A scalability evaluation was executed by changing setup PWs for each user. As shown in the figure, the times taken to calculate 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs were respectively about 152, 570, and 1079 milliseconds.

2) Distribution time for configuring current paths

The times taken to distribute all configurations of the calculated current paths to all PTNs are plotted in Figure 13. As shown in the figure, the times taken to distribute all configurations of the 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs are respectively about 25, 330, and 923 milliseconds.

3) Protection-path calculation time for all current paths in each area

The times taken to calculate protection paths corresponding to all current PWs in each area are plotted in Figure 14. As shown in the figure, the times required for calculating all the inside-area protection paths for 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs are respectively about 600, 1392, and 2095 milliseconds.

4) Protection-path calculation time for all end-to-end current paths

The times taken to calculate end-to-end protection paths to all current PWs are plotted in Figure 15. As shown in the figure, the times taken to calculate all the end-to-end protection paths for 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs are respectively about 230, 952, and 1983 milliseconds.

5) Recovery-path calculation time for operation-plane change

The times taken to calculate all recovery PWs for 255 possible area-based network-failure patterns are plotted in Figure 16. As shown in the figure, the times taken to calculate all recovery PWs for 255 area-based network-failure patterns and 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs are respectively about 12.2, 44.8, and 85.5 seconds.

6) Recovery-configuration time required by both protection schemes for each area and end-to-end path

The times taken to set recovery configuration and to store a configured network topology by the inside-area protection and end-to-end protection schemes after detecting a path disconnection are plotted in Figure 17. Specifically, recovery configuration time was evaluated by intentionally invoking a node failure in area (5). In this case, half of the PWs were damaged and recovered. In the evaluation, if a disconnected path is not recovered for 100 milliseconds by the inside-area protection, it is automatically recovered by the end-to-end protection. Actually, disconnected paths were recovered by the end-to-end protection. As shown in the figure, the times to set recovery configurations for 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs by both protections are respectively about 0.8, 2.2, and 4.7 seconds.

7) Recovery-ID distribution time for changing operation plane

The times taken to distribute the recovery ID to related PTNs and recover after the last area-based network failure is detected in the case of 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs

are plotted in Figure 18. Three area-based network-failure patterns, namely, failures of network areas (1) and (6), failures of network areas (1), (6), and (4), and failures of network areas (1), (6), (4), and (7), were evaluated. As shown in the figure, in the case of 100 (=50+50) current PWs, the times taken to recover from the last failure for the three area-based network-failure patterns are respectively about 167, 177, and 167 milliseconds. In the case of 500 (=250+250) current PWs, the times taken to recover from the last failure for the three area-based network-failure patterns are respectively about 533, 569, and 564 milliseconds. In the case of 1000 (=500+500) current PWs, the times taken to recover from the last failure for the three area-based network-failure patterns are respectively about 1227, 1134, and 1205 milliseconds. As a result, tables that are used for data transmission on 1000 (=500+500) PWs are reconfigured by changing an operation plane in about 1.2 seconds.

In Figure 18, the proposed method is compared with a conventional restoration method in terms of the time taken to calculate and configure PWs. With the conventional method, the times to set recovery configurations for 100 (=50+50) current PWs, 500 (=250+250) current PWs, and 1000 (=500+500) current PWs are respectively about 177, 900, and 2002 milliseconds.

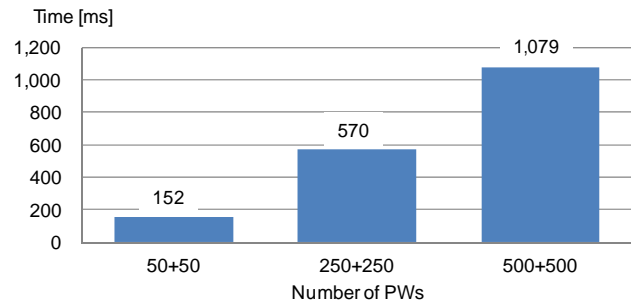


Figure 12. Calculation time for current paths

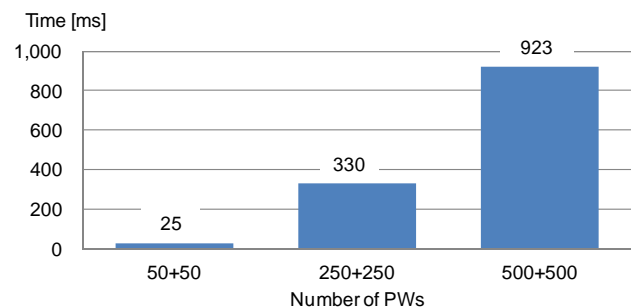


Figure 13. Distribution time for current-path configuration

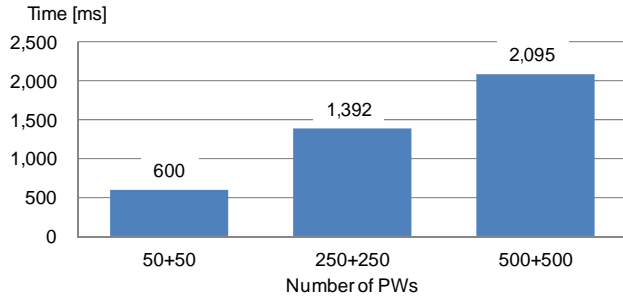


Figure 14. Calculation time for protection paths in each area

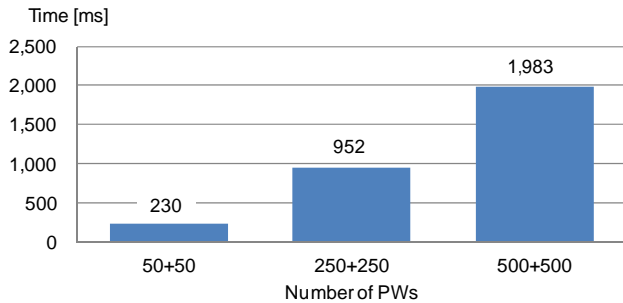


Figure 15. Calculation time for end-to-end protection paths

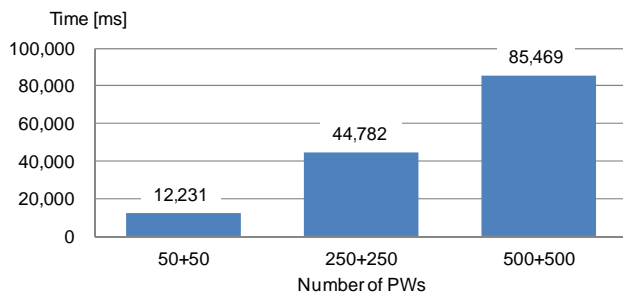


Figure 16. Calculation time for changing operation plane

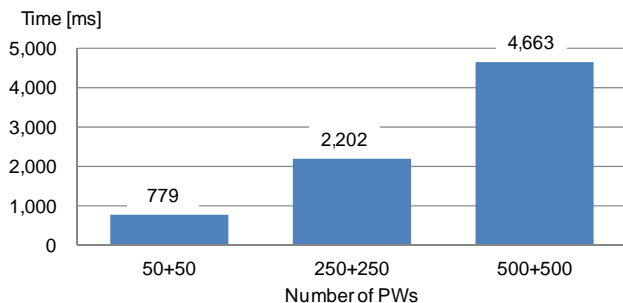


Figure 17. Recovery-configuration time in the cases of using protection paths in NW areas and end-to-end protection paths

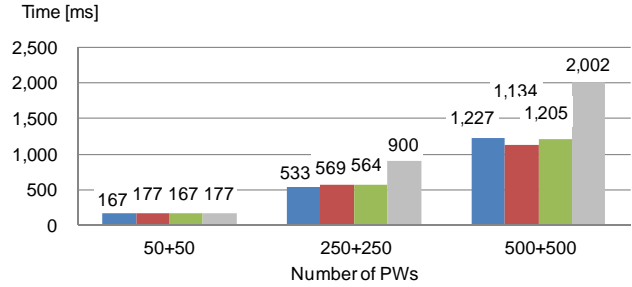


Figure 18. Recovery-configuration time in the case of changing operation plane

D. Discussion

The times taken to recover from failures, such as disconnection of paths, are plotted in Figure 17. In this evaluation, a PTN failure was intentionally invoked in area (5). As a recovery procedure, inside-area protection is expected to be appropriate, since the failure was invoked in area (5). However, end-to-end protection was also used. As for the proposed system, updated PWs and LSPs are always stored after changing data-transmission paths by one of the recovery procedures, such as inside-area protection. In addition, if a failure that is not recovered by the inside-area protection for 100 milliseconds occurs, it is recovered by the end-to-end protection. Over 100 milliseconds were taken to store the PWs and LSPs updated by the inside-area protection; therefore, the PTN failure in area (5) was recovered by both the inside-area protection and the end-to-end protection. The PTN failure was recovered in about 5 seconds for 1000 (=500+500) current PWs, which is a little longer than expected recovery times as a protection, since recovery paths are sequentially configured one by one by using emulated nodes on a server. In addition, a configured network topology was detected and stored. Therefore, the times taken by both protection schemes to recovery are a little longer compared to the recovery time by changing an operation plane. In future work, the times taken to manage multiple updated PWs and LSPs should thus be shortened.

The times taken to distribute the recovery ID and store updated PWs and LSPs are shown in Figure 18. As shown in the figure, the times taken to recover are almost independent of the number of area-based network failures, although they are dependent on the number of setup PWs. In the case of 96 PTNs, tables for data transmission on 1000 (=500+500) current PWs were reconfigured in about 1.2 seconds. The times for recovery are short because the times for setting up real PWs are not included; instead, the times for configuring tables to transmit data are included. In addition, all tables for data transmission are changed at once by switching the operation plane. According to the results of this evaluation, the proposed system can provide a faster recovery procedure

than recalculating and transmitting recovery paths to PTNs (since it omits the recalculation process).

In summary, a transport-network-recovery management system, which can recover from both a small network failure and a major network disaster, was proposed and evaluated. Specifically, for small failures, inside-area protection and end-to-end protection were proposed. In addition, for major failures, an area-based recovery procedure was proposed. As described above, updated data-transmission paths of PWs and LSPs are always stored in a database. Therefore, transmission paths composed of PWs and LSPs updated by changing the operation plane are also stored in the database. As a result, the times taken to recover from the network disaster by changing the operation plane depend on the number of PWs. However, as shown in Figures 16 and 17, the proposed system could promptly recover from both a small network failure and a catastrophic network failure (which is not covered by conventional network-recovery schemes).

VII. CONCLUSION

A system for managing transport-network recovery based on the degree of network failures is proposed. Under this management scheme, an entire network is separated into multiple areas. A network-management server executes a three-step recovery procedure. In the first step, an inside-area protection scheme is applied to the current data-transmission path in each area. In the second step, an end-to-end protection scheme is applied to the current data-transmission path. In the third step, the operation plane is changed. Each assumed operation plane is composed of recovery configurations for restoring failure paths under the assumption of area-based network failures. If a small network failure occurs, it is recovered and localized by the inside-area protection and end-to-end protection schemes. If a catastrophic network failure (due to a disaster) that is not recovered by the protection schemes occurs, it is recovered by changing the operation plane according to damaged areas.

A prototype system composed of a network-management server and 96 emulated packet-transport nodes was developed and evaluated by configuring 1000 (=500+500) data-transmission paths. In the case of a small network failure, 500 data-transmission paths composed by LSPs and PWs were damaged and reconfigured by the inside-area protection and end-to-end protection schemes in about 5 seconds. If a network failure was not recovered by the protection schemes, all tables for 1000 (=500+500) data transmission paths were reconfigured to recover from the failure by changing the operation plane in about 1.2 seconds. As a result, the proposed system could provide a faster recovery procedure than recalculating and transmitting recovery paths to PTNs. In addition, it could localize and recover a network failure according to the degree of network failures.

Although the protection scheme could recover 500 data

transmission paths from a small network failure, it took the network-management server about 5 seconds to configure and store changed-data transmission paths. If numerous current paths exist, it will take too much time to assess changed paths. Accordingly, the protection scheme will be further developed so that it can promptly manage a large number of recovered paths.

ACKNOWLEDGMENTS

Part of this research was done within research project O3 (Open, Organic, Optima) and programs, "Research and Development on Virtualized Network Technology," "Research and Development on Management Platform Technologies for Highly Reliable Cloud Services," and "Research and Development on Signaling Technologies of Network Configuration for Sustainable Environment" supported by MIC (The Japanese Ministry of Internal Affairs and Communications).

REFERENCES

- [1] T. Suzuki et al., "A system for managing transport-network recovery according to degree of network failure," *The Fourth International Conference on Communications, Computation, Networks and Technologies (INNOV 2015)*, Nov. 2015, pp. 56-63.
- [2] T. Suzuki et al., "A network-disaster recovery system using multiple-backup operation planes," *International Journal on Advances in Networks and Services*, vol. 8 nos. 1&2, July 2015, pp. 118-129.
- [3] E. Mannie and D. Papadimitriou, "Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS)," IETF RFC 4427, Mar. 2006.
- [4] International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) <http://www.itu.int/en/ITU-T/Pages/default.aspx> [retrieved: Nov. 2016].
- [5] The Internet Engineering Task Force (IETF), <http://www.ietf.org/> [retrieved: Nov. 2016].
- [6] B. Niven-Jenkins, D. Brungard, M. Betts, N. Sprecher, and S. Ueno, "Requirements of an MPLS transport profile," IETF RFC 5654, Sept. 2009.
- [7] M. Bocci, S. Bryant, D. Frost, L. Levrau, and L. Berger, "A Framework for MPLS in transport networks," IETF RFC 5921, July 2010.
- [8] T. Busi and D. Allan, "Operations, administration, and maintenance framework for MPLS-based transport networks," IETF RFC 6371, Sept. 2011.
- [9] N. Sprecher and A. Farrel, "MPLS transport profile (MPLS-TP) survivability framework," IETF RFC 6372, Sept. 2011.
- [10] P. Pan, G. Swallow, and A. Atlas, "Fast reroute extensions to RSVP-TE for LSP tunnels," IETF RFC 4090, May 2005.
- [11] J. Zhang, J. Zhou, J. Ren, and B. Wang, "A LDP fast protection switching scheme for concurrent multiple failures in MPLS network," *2009 MINES '09. International*

- Conference on Multimedia Information Networking and Security*, Nov. 2009, pp. 259-262.
- [12] Z. Jia and G. Yunfei, "Multiple mode protection switching failure recovery mechanism under MPLS network," *2010 Second International Conference on Modeling, Simulation and Visualization Methods (WMSVM)*, May 2010, pp. 289-292.
- [13] M. Lucci, A. Valenti, F. Matera, and D. Del Buono, "Investigation on fast MPLS restoration technique for a GbE wide area transport network: A disaster recovery case," *12th International Conference on Transparent Optical Networks (ICTON)*, Tu.C3.4, June 2010, pp. 1-4.
- [14] T. S. Pham, J. Lattmann, J. Lutton, L. Valeyre, J. Carlier, and D. Nace, "A restoration scheme for virtual networks using switches," *2012 4th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Oct. 2012, pp. 800-805.
- [15] X. Wang, X. Jiang, C. Nguyen, X. Zhang, and S. Lu, "Fast connection recovery against region failures with landmark-based source routing," *2013 9th International Conference on the Design of Reliable Communication Networks (DRCN)*, Mar. 2013, pp. 11-19.
- [16] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol label switching architecture," IETF RFC 3031, Jan. 2001.
- [17] S. Bryant and P. Pate, "Pseudo wire emulation edge-to-edge (PWE3) architecture," IETF RFC 3985, Mar. 2005.