Fault Recovery Implementation Spec. Draft　(Ver. 2.0)

Edited by PIL Standardization WG
November 30, 2005

Table of contents

1. Introduction

Fault recovery is an indispensable technology for increasing reliability of telecommunication network. In a framework of Multi-Protocol Label Switching (MPLS), a technology called as Fast ReRoute (FRP) has been proposed as RFC (RFC4090) and has being promoted to deploy into the commercial network. Similarly, in GMPLS network that has been constructed by expanding the MPLS network, since it was considered to be necessary to develop a technology for improving reliability, movement to take such technologies as a protection technology that has being utilized in the conventional networks and a restoration technology of which adoption was once taken into consideration but was not actually utilized into GMPLS technology, is being accelerated at present. However, although standardization work in IETF has been once terminated, implementation of relevant technologies and confirmation work of interoperability between different vendors in ISOCORE or UNH (University of New Hampshire) are subjects to be done in the immediate future. The purpose of this IA (Implementation Agreement) is to supplement the insufficient part that is difficult to express in detail in the draft and to establish the detailed specifications that are possible to be implemented. As the result, it is expected that it will accelerate the standardization activities of IETF and make higher the presence of Photonic Internet Laboratory (PIL) in this field.
A part of fault recovery implementation specifications agreed by this IA will be validated in the Technology Validation WG to become an item of interconnectivity tests. Subject of data plane to be discussed in this IA is mainly SDH (Synchronous Digital Hierarchy) and Lambda.

2. Definition
2.1 Restoration and Protection

- Protection
- Usually, also the XCT of non-working LSP is set.
- Signaling for setting up the XCT is not required when switching the non-working LSP to working LSP

- Restoration
- Usually, XCT of non-working LSP is not set.
- Non-working LSP is possible to share the resources (label, bandwidth) with other non-working LSP.
- Signaling for setting up the XCT is required when switching the non-working LSP to working LSP.

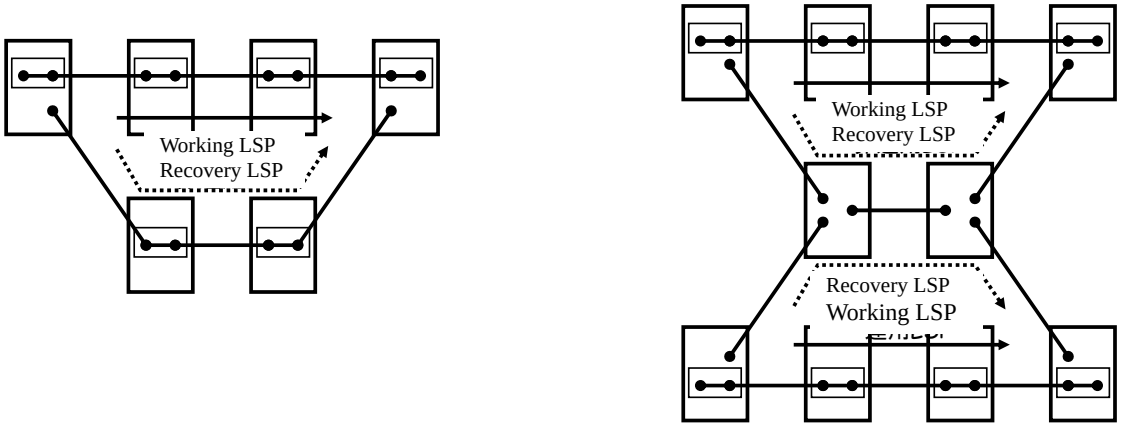


Fig. 1-1 Protection and Restoration

2.1.1 General description about protection function

Protection is a kind of fault recovery that GMPLS is supporting. In general, provisioning for LSP (Label Switched Path) is separated into the following three phases:

(1) Route calculation
(2) Signaling
(3) Cross-connect setting

(1) and (2) are operations for establishing LSP closing within C-Plane, while (3) is an operation for establishing LSP including also D-Plane.

There are two categories in fault recovery function; Restoration and Protection. The difference between them is in the method and timing to provide the recovery LSP.
Protection described in this section has the following features.

① As for the recovery LSP for protecting the currently working LSP, only the state that calculation of route has been completed is allowed.
② When working LSP is operating, only the state that signaling for recovery LSP has been completed is allowed.
③ When working LSP is operating, only the state that cross-connecting for recovery LSP has been completed is allowed.
④ Sharing the part or all of the resources constructing the recovery LSP by multiple working LSPs is not allowed.
⑤ A mode to control switch-over of LSP with D-plane is allowed.

There are following two types in protection to be described in this section when it is seen from the viewpoint of Extra-Traffic.

・1+1 LSP protection ; It is inhibited to flow Extra-Traffic in recovery LSP during the working LSP is operating,
・1:1 LSP protection ; It is allowed to flow Extra-Traffic in recovery LSP during the working LSP is operating,

As for the 1+1 LSP protection, there are two more modes depending on whether it is required or not to cooperate when executing switch over at the transmitting/receiving-end nodes.

・1+1 Uni-directional Protection ; Executes switch over solely at the receiving-end node.
・1+1 Bi-directional Protection ; Executes switch over cooperating each other between transmitting-and receiving-node nodes.

This IA is referencing the “Draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03” that is the latest IETF draft relating to protection functions

at the time of 2005.11.30. Differences in these three modes of protection above described including the correspondence with the form of recovery are shown in Table-2-1.

Table. 2-1 Features of 1+1(uni/bi) and 1:1 protection

No.	Comparison items	1+1 Uni-directional Protection	1+1 Bi-directional Protection	1:N Protection
1	Route calculation of recovery LSP	Completed during the working LSP is operating.		
2	Signaling of recovery LSP	Completed during the working LSP is operating.		
3	Cross-connection of recovery LSP	Completed during the working LSP is operating.		
4	Sharing the recovery LSP resources	Not allowed.		
5	LSP switch over control	Switch over control in D-Plane is allowed		
6	Extra-Traffic	Not allowed even when working LSP is operating.		Allowed to flow Extra-Traffic during working LSP is operating.
7	Cooperation between transmitting/receiving-nodes when switch over is executed.	Operates independently	Cooperates between transmitting/receiving-nodes	Operates independently
	Correspondence with lang-draft(Rev.01:'03.5)	・1+1 Unidirectional Protection	・1+1 Bi-directional Protection	・1:N Protection with Extra-Traffic

Although, in e2e-signaling-03, LSP switching control is taking a form that allows both switch-over/switch-back control in both of the D-plane and C-plane, in this IA, only the switch-over/switch-back control in D-plane is specified and the switch-over/switch-back control in D-plane is not specified being assumed as the function to be expanded in the future.
So, in the succeeding sections of this IA, specification is done focusing on signaling for provisioning the working LSP/recovery LSP and signaling for fault notification.

Relating to recovery operation, we will re-organize the requirements for signaling and routing functions in the followings including the above comparison items.

【RSVP signaling】

RSVP signaling is utilized for provisioning and booting of LSP and the functional requirements in recovery operation are as follows.

(a) Function to determine which should be used either working LSP or recovery LSP.
(b) Function to identify the cross-connect state of recovery LSP that is not currently operating.
(c) Function to determine whether switching control signaling in C-plane is needed or not when a fault occurred.
(d) Function to correlate the recovery LSP to which working LSP is switched with the working LSP from which the recovery LSP is switched each other.
(e) Function to have each node on recovery LSP distinguish the clearly specified route of working LSP.
(f) Function to identify the type of recovery.
(g) Function to suppress the alarm of LSP that is not currently operating.
(h) Function to forcibly inhibit operating/recovery LSP switching.

In order to satisfy the above described functional requirements, the following objects have been defined for signaling relating to recovery operation in GMPLS.

・PROTECTION object (Class-Num=37 / C-Type=2)
・PRIMARY PATH ROUTE object (Class-Num=TBA / C-Type=1) [Abbreviated as PPRO]
・ADMIN_STATUS object (Class-Num=196 / C-Type=1)
・ASSOCIATION object (Class-Num=198 / C-Type=1)

Table.2-2 shows the correspondence list between the fields defined in each object of the above described RSVP signaling and the functional requirement in recovery operation.

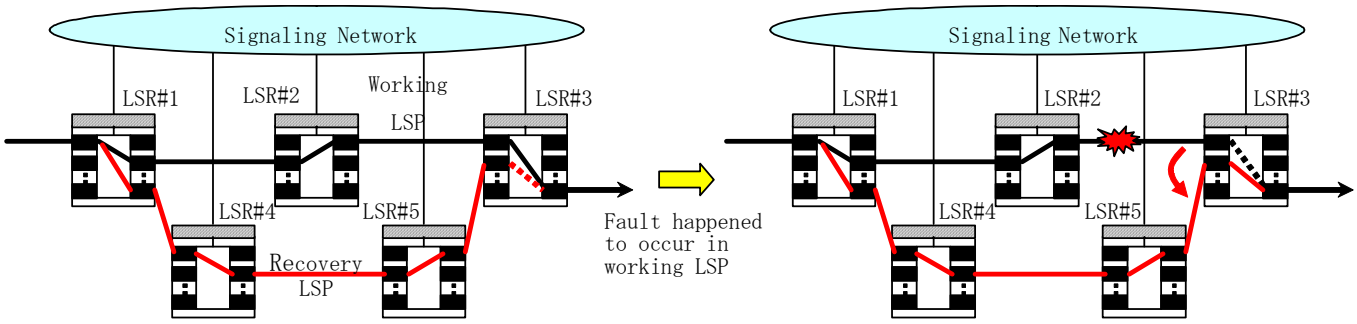
Table.2-2 Correspondence list between the fields defined in each object of the above described RSVP signaling and the functional requirement in recovery operation.

No.	Object name	Message	Field name	Data length	Corresponding functional requirement	Description
1-1	PROTECTION	・PATH ・RESV	Secondary bit (S bit)	1 bit	(b)	0: Cross-connection executed 1: Cross-connection not executed
1-2			Protecting bit (P bit)	1 bit	(a)	0: Working 1: Recovry
1-3			Notification bit (N bit)	1 bit	(c)	0: Switch-over control signaling is executed in C-plane. 1: Switch-over control signaling is not executed in C-plane.
1-4			LSP Flags	6 bit	(f)	0x00; Not protected 0x01; (Full) Re-routing 0x02;1:1 Rerouting (No Extra-Traffic) 0x04;1:1 Protection (Extra-Traffic exists) 0x08;1+1 Uni-directional Protection) 0x10;1+1 Bi-directional Protection)
2-1	PRIMARY PATH ROUTE	・PATH ・RESV	Subobjects	Variable length	(e)	(ERO of the secondary protecting LSP) + (Excerpt from RECORD ROUTE objects of primary protected LSP)
3-1	ADMIN_STATUS	・PATH ・RESV	Administratively Down(A bit)	1 bit	(g)	0; Administratively up 1; Administratively down
3-2			Lock Out(L bit)	1 bit	(h)	0; Normal 1; Lock Out
4-1	ASSOCIATION	・PATH	Association ID	16 bit	(d)	Protected side; Protecting LSP ID Protecting side; Protected LSP ID
4-2			Association Source IPv4	32 bit	(d)	Protected side; Source IPv4 address of Protecting LSP Protecting side; Source IPv4 address of Protected LSP

Name of protection type	S bit	P bit	N bit	LSP Flags	Association ID	PPRO
1+1 Unidirectional Protection	0	1	1	0x08	protected LSP ID	-
1+1 Bi-directional Protection	0	1	1	0x10	protected LSP ID	-
1:N Protection	0	1	1	0x04	protected LSP ID	-

As shown in this Table, the most significant features in each system categorized into protection are that S bit = 0 (cross-connect completed) during operation of working LSP and that N bit = 1 (no switching/switch-back control signaling in C-Plane).

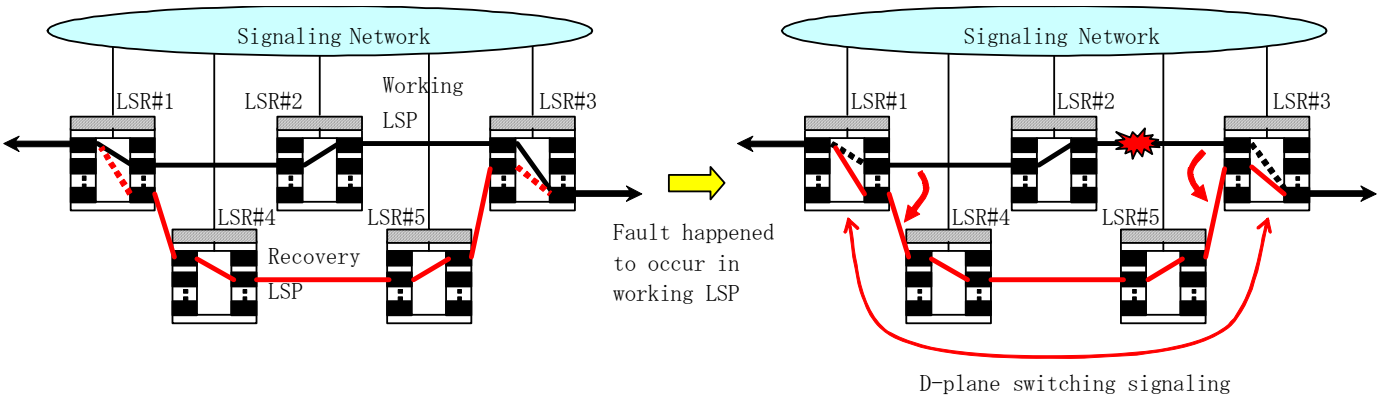
Fig.2-2 shows a concept of 1+1 Uni-directional Protection.
Fig.2-3 shows a concept of 1+1 Bi-directional Protection.
Fig.2-4 shows a concept of 1:N Protection.



<Working LSP is operating>
Recovery LSP;
- doesn't complete cross-connection at only the receive-end node.
- completes cross-connection at the send-end node and the intermediate node.

<Recovery LSP is operating >
Recovery LSP detects the fault of current LSP, and
- Sets up cross-connection (switching) of Recovery LSP at the receive-end node.

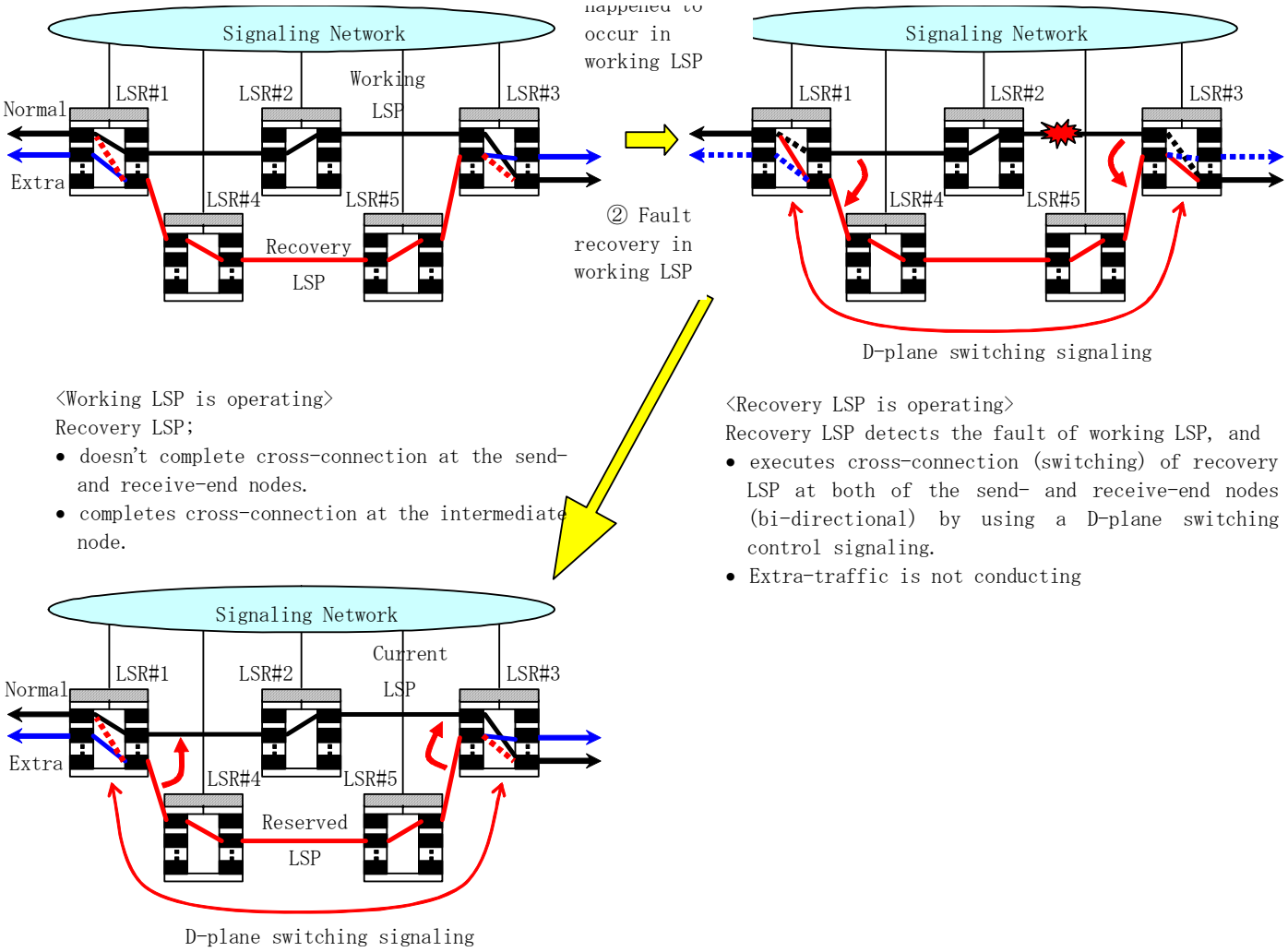
Fig.2-2 Concept of 1+1 Unidirectional Protection



<Working LSP is operating>
Recovery LSP;
- doesn't complete cross-connection at the send- and receive-end nodes.
- completes cross-connection at the intermediate node.

<Recovery LSP is operating>
Recovery LSP detects the fault of working LSP, and
- executes cross-connection (switching) of recovery LSP at both of the send- and receive-end nodes (bi-directional) by using a D-plane switching control signaling.

Fig. 2-3 Concept of 1+1 Bi-directional Protection



<Working LSP is operating>
Recovery LSP;

- doesn't complete cross-connection at the send- and receive-end nodes.
- completes cross-connection at the intermediate node.

<Recovery LSP is operating>
Recovery LSP detects the fault of working LSP, and

- executes cross-connection (switching) of recovery LSP at both of the send- and receive-end nodes (bi-directional) by using a D-plane switching control signaling.
- Extra-traffic is not conducting

<Cut-and-return to recovery LSP>
Recovery LSP detects the fault recovery to working LSP, and

- executes cross-connection (switch-back) of recovery LSP at both of the send- and receive-end nodes (bi-directional) by using a D-plane switch back control signaling.
- Extra-traffic conducts again.

<Legend>
Working (Normal Traffic) Recovery (Normal Traffic) Extra Traffic
(※) As for description of other symbols, refer to <legend> in Fig.4.3

Fig. 2-4 Concept of 1:N Protection with Extra-Traffic

【Routing】
In recovery operation, the role that routing plays is to advertise the Shared Risk Link Group (SRLG) information so that the resource that each route uses doesn't overlap each other in risk management when executing route calculation for operating/recovery LSP. SRLG sub TLV is included in Link TLV (Type=2). In route calculation, each resource constructing the working/recovery LSP should be included in completely independent SRLG.

And, such method that lets current/recovery LSP see as a “special FA” for a client is also under consideration. Also in this case, advertising of FA is executed using a routing function.

2.1.2 General description about Restoration Function

Restoration is one of the of fault recovery modes that GMPLS is supporting.
There are following features in restoration that this section will describe.

① When working LSP as the object of protection is operating, the state that route calculation of recovery LSP that protects the working LSP has not yet been completed is allowed.
② When working LSP is operating, the state that signaling for recovery LSP has not yet been completed is allowed.
③ When working LSP is operating, the state that cross-connection for recovery LSP has not yet been completed is allowed.
④ Sharing the part or all of the resources constructing the recovery LSP by multiple working LSPs is not allowed.
⑤ Switch over control of LSP has to be done in C-Plane.

There are following two types of restoration that is described in this section depending on the allowable range from ① to ⑤.

・Preplanned restoration
・Dynamic restoration

This IA is referencing the “Draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03” that is the latest draft relating to restoration functions at the time of 2005.11.30. Differences in these two modes of restoration above described including the correspondence with the form of recovery are shown in Table.2-4.

Table. 2-4 Features of preplanned restoration and dynamic restorations

No.	Comparison items	Preplanned restoration type	Dynamic restoration type
1	Route calculation of recovery LSP	Completed during the working LSP is operating.	The state that route calculation has not been completed during working LSP is operating is allowed.
2	Signaling of recovery LSP	Completed during the working LSP is operating.	The state that signaling has not been completed during working LSP is operating is allowed.
3	Cross-connection of recovery LSP	The state that cross-connection has not been completed during working LSP is operating is allowed.	The state that cross-connection has not been completed during working LSP is operating is allowed.
4	Sharing the recovery LSP resource	・Allowed in case of Shared Mesh Restoration ・Not allowed in case of 1:1 Restoration	Allowed
5	LSP switch over control	C-Plane	C-Plane
	Correlation with e2e-signaling-03	・Re-routing without Extra-Traffic	・(Full) LSP Re-routing

In implementation, it is possible to assume above two modes. In one LSP, mixing of two modes that the nodes constructing LSP is supporting is allowed.

Relating to recovery operation, we will organize the requirements for signaling functions in the followings.
As for the requirements for routing functions, they are the same as in case of protection. Refer to section 2.1.1.

【RSVP Signaling】
RSVP signaling is utilized for provisioning and booting of LSP and the functional requirements in recovery operation are the same as shown in (a) to (f) in section 2.1.1.

Objects for signaling that are defined in order to satisfy these functional re4quirements are as follows just like the case of restoration.

・ PROTECTION object (Class-Num=37 / C-Type=2)
・ PRIMARY PATH ROUTE object (Class-Num=TBA / C-Type=1) [Abbreviated as PPRO]
・ ADMIN_STATUS object (Class-Num=196 / C-Type=1)
・ASSOCIATION object (Class-Num=198 / C-Type=1)

As for the correspondence list between the fields defined in each object of the above described RSVP signaling and the functional requirement in recovery operation, please refer to Table.2-2.

Among the fields shown in Table.2-2, the values that the recovery LSP can take in each restoration method just after the LSP provisioning (the state that working LSP is operating) are shown in Table.2-5. (as for the dynamic values, please refer to the respective section)

Table.2-5 Recovery related field values that the recovery LSP can take for each restoration type just after the LSP provisioning.

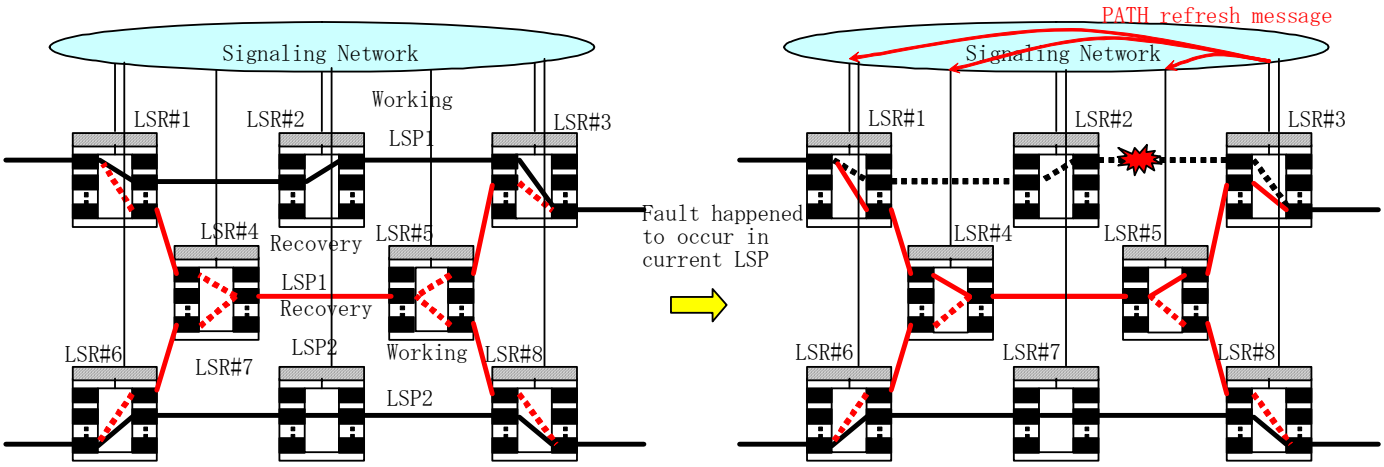
Name of restoration type	S bit	P bit	N bit	LSP Flags	Association ID	PPRO
Shared Mesh Restoration	1	1	0	0x02	protected LSP ID	Yes
1:1 Restoration	1	1	0	0x02	protected LSP ID	No

(*1) Refer to description in Table.2-2 (section 2-1) .
(*2) As for the (Full) LSP Re-routing (Dynamic type), recovery LSP doesn't exist. The value of respective field is the one of working LSP , which was shown only the meaningful value for comparison.

As shown in this table, the greatest features of each method categorized in restoration are that S bit=1 (cross-connect has not yet been completed or recovery LSP itself doesn't exist) during the working LSP is operating and that N bit=0 (switching control signaling in C-Plane is executed). Shared Mesh Restoration and 1:1 Restoration are distinguished from existence of PPRO.

Fig.2-5 shows a concept of 1:1 Restoration
Fig.2-6 shows a concept of Shared Mesh Restoration
Fig.2-7 shows a concept of (Full) LSP Re-routing

Fig.2-5 Concept of 1:1 Restoration



<Working LSP1,2 are operating>
Recovery LSP1,2;

- completed route calculation.
- completed signaling.
- cross-connection not completed.

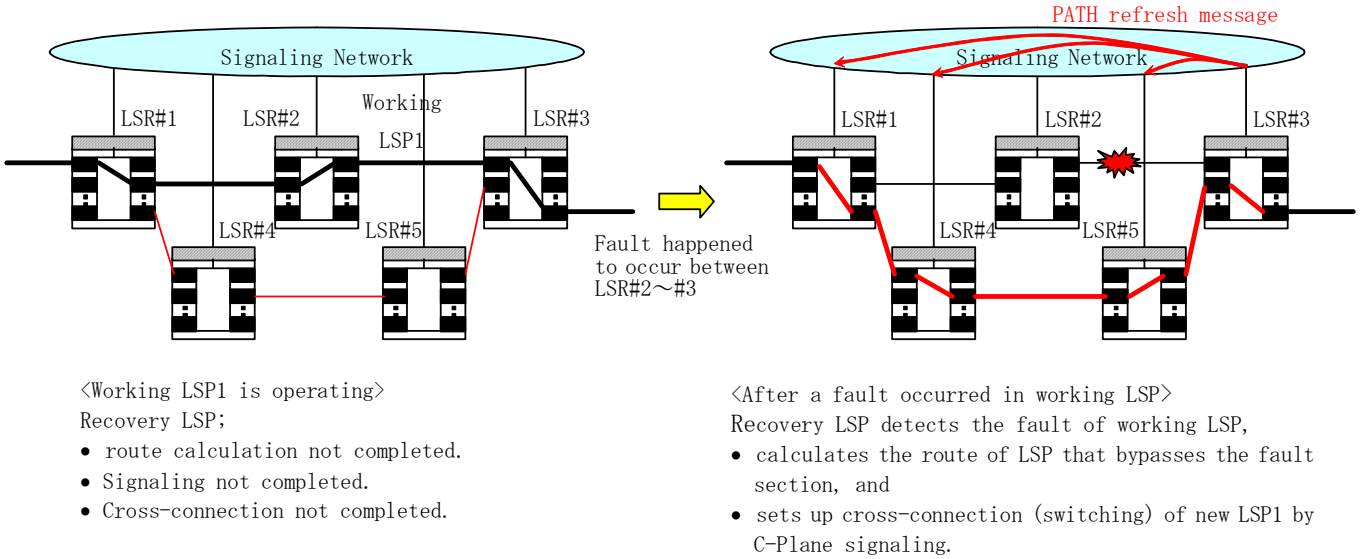
Recovery LSP1,2 share resources between LSR#4～#5.

<Recovery LSP is operating >
Recovery LSP detects the fault of current LSP, and

- sets up cross-connection (switching) of reserved LSP1 by C-Plane signaling.
- There are no change in current LSP2 and reserved LSP2.

When working LSP1,2 failed at the same time, either of them can't be recovered.

Fig.2-6 Concept of Shared Mesh Restoration



<Working LSP1 is operating>
Recovery LSP;
- route calculation not completed.
- Signaling not completed.
- Cross-connection not completed.

<After a fault occurred in working LSP>
Recovery LSP detects the fault of working LSP,
- calculates the route of LSP that bypasses the fault section, and
- sets up cross-connection (switching) of new LSP1 by C-Plane signaling.

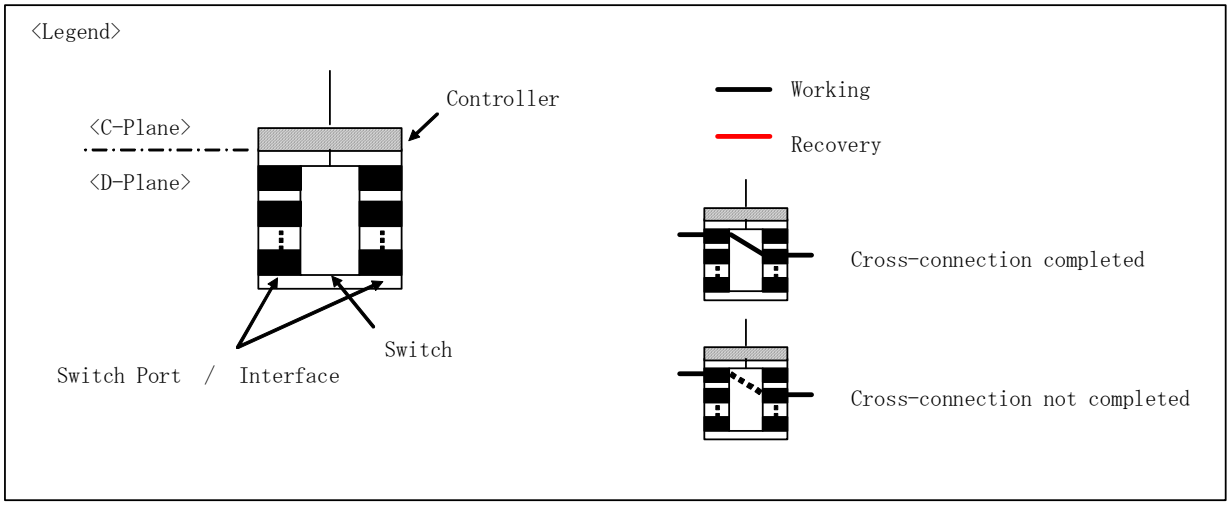


Fig.2-7 Concept of (Full) LSP Re-routing

This IA doesn't specify the dynamic type restoration but specifies only the preplanned restoration.

In the preplanned restoration, setting up of recovery LSP being closed within C-Plane has also been completed at the same time as provisioning of the working LSP.

If we reorganize the features of the preplanned restoration, they will be expressed as the followings.
(1) Calculation of recovery LSP's route → Completed while working LSP is operating
(2) Signaling of recovery LSP → Completed while working LSP is operating
(3) Cross-connection of recovery LSP → Not completed while working LSP is operating
(4) LSP switch over control → Executed in C-Plane

The preplanned restoration can be further categorized into two types from the view point of sharing the recovery LSP resources.

① 1:1 Restoration : This type doesn't share recovery LSP resources among multiple working LSPs.
② Shared Mesh Restoration : This type shares recovery LSP resources among multiple working LSPs.

Depending on degree to which recovery LSP signaling is executed, there is a difference in ratio of network sharing, fault tolerability of LSP, and time to recover from fault. Table.2-6 shows a comparison of each term against the signaling level of recovery LSP.

Table.2-6 Comparison of ratio of network sharing, fault tolerability of LSP and time to recover from fault depending on signaling level of recovery LSP

Item of comparison	Up to label reservation ⟷ Up to bandwidth reservation
Ratio of network sharing	Low → High
Fault tolerability	Low → High
Time to recover from fault	Fast → Slow

In this IA, only the full span restoration (bus restoration) in which SRLG becomes fully independent between working LSP and recovery LSP will be specified, and the span restoration in which only a part of spans is detoured will be excluded from specification.

There are two types in Extra-Traffics as the followings.

- A type that uses all the recovery LSPs excluding the end point (signaling for LSP for Extra-Traffic is not required).
- A type that uses a part of recovery LSP's resources (signaling for LSP for Extra-Traffic is required).

In restoration, both of these are supported.

2.2 Duplication and switching operation of LSP

By duplicating LSP, it is possible to realize a highly reliable GMPLS network.
Tunnel in which high reliability was realized consists of working LSP and recovery LSP. Fig.2-8 shows a relationship between tunnel and LSP. A tunnel is set up between node-A and node-B, and working LSP and recovery LSP are set up as its structural components.
In case of protection, the working LSP and the recovery LSP are set up actually in network. Fig.2-9 shows a state of connection and cross-connection. In protection state, all the relating connections and cross-connections are in operating state in which other LSPs can not utilize them. The traffic divided into two data streams by Ingress or Egress nodes is transferred using both LSPs of current and reserved, and selected by Egress or Ingress nodes at the destination side. When a fault occurred in operating system that is currently selected, it is possible to recover from fault by switching the recovery system to working system.
In case of restoration, working LSP is actually configured, but cross-connection of recovery LSP is not actually configured. Fig.2-10 shows the state of connection and cross-connection in restoration. The working system is assigned to the route passing though Ingress-Core1-Core-2-Egress and the recovery system is assigned to the route passing through Ingress-Core3-Core4-Egress. In the working system, all the cross-connections are in operating state and are possible to be used for transmitting traffic. In the recovery system, all the cross-connections must be in non-operating state. On the other hand, connections are not necessarily required to be in non-operating state. Therefore, even in restoration, such operations as calculating the reserved route (routed state), securing the reserved bandwidth (reserved state), or assigning the label (assigned state) are possible to execute. When a fault happened to occur in the working system, fault is recovered by making all the connections and cross-connections of the recovery system to be in operating state to transmit the traffic through the recovery system. From Fig.2-10, it will be known that the state of current/reserved and operating/non-operating is changed by automatic switching or command switching. Here we explain this in case of describing the state by using S(Secondary)-bit and P(Protecting)-bit in Protection Object described in [recovery-e2e]. As shown in Table.2-7, when S-bit is 1, LSP doesn't execute cross-connection. When S-bit is 0, it indicates that the system is currently operating. When P-bit is 1, it indicates that LSP is in recovery system, and when P-bit is 0, it indicates that LSP is in working system. Table.2-8 shows how S-bit and P-bit change by switching operation. (created by referring [recovery-e2e])
In both cases of protection and restoration, when the working system in which fault occurred was required and became to be available again, it is switched back from the state of recovery system to the state that is possible to operate as a working system. Switch-back operation is optional. Although switch-back operation is not necessarily required, it is desirable to exist from the viewpoint of actual operation.
In case of duplicating LSP, tunnel ID and two LSP-IDs are assigned independently each other from the viewpoint of separating Call and Connection. For example, when duplication of Tunnel (ID:X) was once released and the recovery LSP was setup again, it is not required that the ID of initial recovery LSP coincides with the ID of recovery LSP after it was reconfigured. And for another example, when the initial working LSP was released after the working LSP was switched to recovery LSP by restoration without switch-back, it is possible to use the ID of recovery LSP that has been used so far as the ID of working LSP. And, since the recovery LSP is newly setup, it is also possible to assign an entirely new ID as an ID of new recovery LSP.

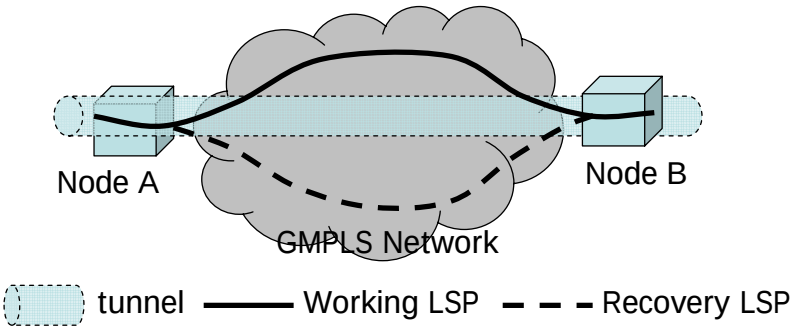


Fig.2-8 Relationship between tunnel and working/recovery LSPs.

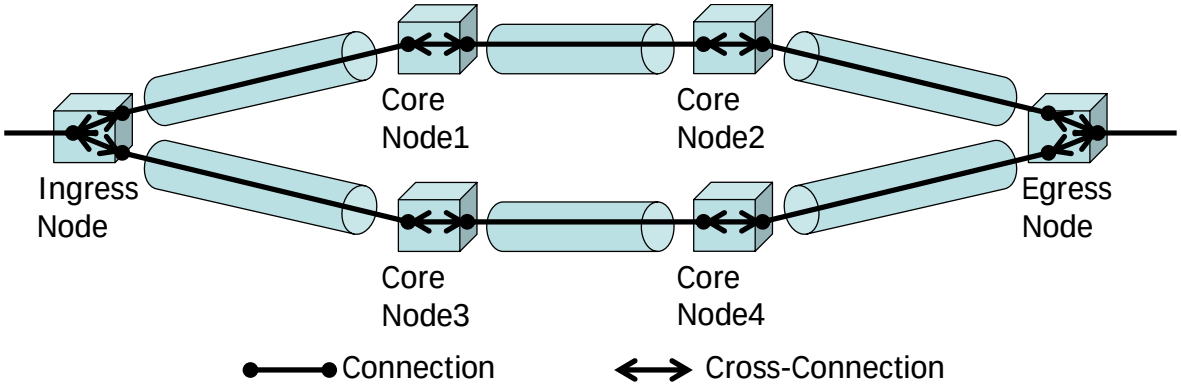


Fig.2-9 Setup of protection state

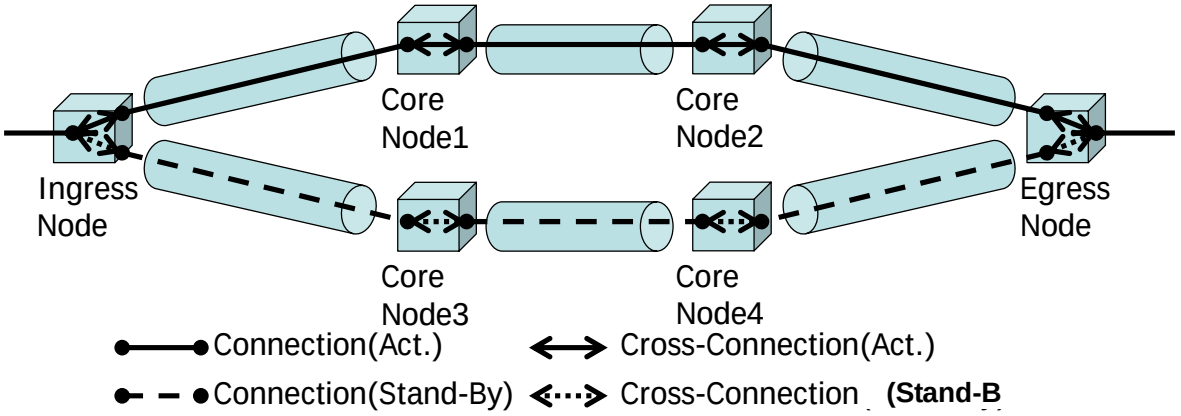


Fig.2-10 Setup of restoration state

Table.2-7 Definition of S-bit, P-bit, and O-bit

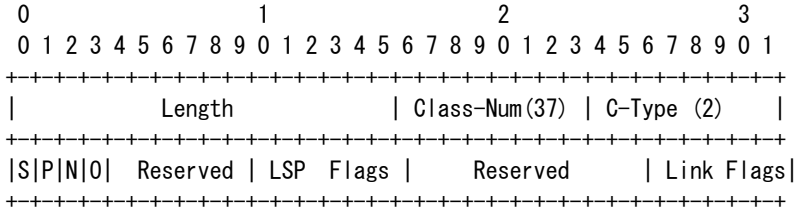
	Value	Meaning
S-bit	0	Cross-connection has been setup
	1	Cross-connection has not been setup
P-bit	0	Working
	1	Recovery
O-bit	0	State other than the recovery LSP is operating
	1	State in which recovery LSP is operating

Table.2-8 Relationship between switching method, S-bit and P-bit

Switch-over method	Status	Working/Recovery	LSP (Protection Type) Flags	S-bit	P-bit	O-bit
1+1 Bi-directional Protection	Before switching	Working	0x10	0	0	0
		Recovery	0x10	0	1	0
	After switching	Working	0x10	0	0	0
		Recovery	0x10	0	1	1
1+1 Unidirectional Protection	Before switching	Working	0x08	0	0	0
		Recovery	0x08	0	1	0
	After switching	Working	0x08	0	0	0
		Recovery	0x08	0	1	1
1:1 Protection with Extra-Traffic (Path and bandwidth protection)	Before switching	Working	0x04	0	0	0
		Recovery	0x04	0	1	0
	After switching	Working	0x04	0	0	0
		Recovery	0x04	0	1	1
1:1 Re-Routing without Extra-Traffic (Path protection only)	Before switching	Working	0x02	0	0	0
		Recovery	0x02	1	1	0
	After switching	Working	0x02	0	0	0
		Recovery	0x02	0	1	1
Shared Mesh	Before switching	Working	0x02	0	0	0
		Recovery	0x02	1	1	0
	After switching	Working	0x02	0	0	0
		Recovery	0x02	0	1	0
Full Re-routing		Working	0x01	0	0	0

【For reference】
(Extracted from “draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03.txt, chap.14, protection object”)

The format of the PROTECTION Object (Class-Num = 37, C-Type = 2 by IANA) is as follows:



Secondary (S): 1 bit

When set to 1, this bit indicates that the requested LSP is a secondary LSP. When set to 0 (default), it indicates that the requested LSP is a primary LSP.

Protecting (P): 1 bit

When set to 1, this bit indicates that the requested LSP is a protecting LSP. When set to 0 (default), it indicates that the requested LSP is a working LSP. The combination, S set to 1 with P set to 0 is not valid.

Notification (N): 1 bit

When set to 1, this bit indicates that the control plane message exchange is only used for notification during protection switching. When set to 0 (default), it indicates that the control plane message exchanges are used for protection switching purposes. The N bit is only applicable when the LSP Protection Type Flag is set to either 0x04 (1:N Protection with Extra-Traffic), or 0x08 (1+1 Unidirectional Protection) or 0x10 (1+1 Bi-directional Protection). The N bit MUST be set to 0 in any other case.

Operational (O): 1 bit

When set to 1, this bit indicates that the protecting LSP is carrying the normal traffic after protection switching. The O bit is only applicable when the P bit is set to 1 and the LSP Protection Type Flag is set to either 0x04 (1:N Protection with Extra-Traffic), or 0x08 (1+1 Unidirectional Protection) or 0x10 (1+1 Bi-directional Protection). The O bit MUST be set to 0 in any other case.

Reference
[recovery-e2e] draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03.txt

2.3 Technical Terms and Abbreviations

○ Active and Stand-by
Active: A state that the resource is working (Client's data are flowing and the client will be affected when a fault happened to occur.)
Stand-by: A state that the resource is waiting to work.

○ Working and Recovery
Working: A resource to be used usually.
Recovery: A resource to be used for recovering from fault.

○ Protection and Restoration
Protection: There are three types of protection from fault; 1+1, 1:n and ring. But, in this IA, ring protection is excluded from discussion.
This is a case in which cross-connection in intermediate node is not required.
Restoration: A fault recovery method in which stand-by is established after a fault happened to occur in working LSP (Mainly in mesh network).
This is a case in which cross-connection in intermediate node is required.
In this IA, an intermediate form is excluded from discussion.

○ Span, Segment and E2E
Span: Traffic path between most adjacent two nodes. For example, M-section in SDH.
Segment: Traffic path that multiple spans are connected in series.
E2E: A LSP segment from Ingress to Egress nodes.

○ Extra-Traffic
A traffic transmitted by using a recovery resource.
This is taken over by traffic passing through a working resource when the resource was switched-back.
In this IA, there is no difference in protection and restoration.

○ Global repair and Local repair
Global repair: End-to-end recovery
Local repair: Local recovery or Segment recovery

○ Full span restoration, Partially span restoration, and Full LSP restoration
Full span restoration: When a fault occurred in a certain span, all the LSPs accommodated in this span are saved.
Partially span restoration: When a fault occurred in a certain span, a part of LSPs accommodated in this span is saved.
Full LSP restoration: Initiator of LSP saves bypassing the fault place. In this case, it is allowed to pass through the original repeater node.

○ State of LSP
State that only the route has been calculated: routed
State that bandwidth was reserved: (label-) reserved
State that label was assigned: (label-) assigned
State that switch was setup: cross-connected

○ TE-Link and FA
TE Link = {Basic TE Link, FA}
TE Link: A link as an object of advertising in routing protocol.
FA: A thing that LSP was advertised as TE Link.
FA-LSP: LSP as an entity of FA
Basic TE Link: TE Link other than FA. Depending on layers, fiber, bundled fibers, wavelength and time-slot are possible to become the basic TE Link.

3. RSVP Signaling

3.1 Preplanned restoration signaling

3.1.1 Signaling sequence

3.1.1.1 Setup of working LSP and recovery LSP

Setup sequence of working LSP (LSP0) and recovery LSP (LSP1) are shown in Fig.3-1.
Between LSP0 and LSP, SESSION Object uses the same value and LSP ID uses different values.

- PROTECTION Object and ASSOCIATION Object are set as the followings.
  LSP Flags = 0x02
  【LSP0】
  P=0, S=0
  Association ID = "LSP1 の LSP　ID"
  【LSP1】
  P=1, S=1
  Association ID = "LSP ID of LSP0"
- When recovery LSP of Shared Mesh Restoration is set up, PRIMARY_PATH_ROUTE Object is included to PATH message. Into the PRIMARY_PATH_ROUTE Object, path of corresponding working LSP is included. In other cases, PRIMARY_PATH_ROUTE Object is not included.
- ADMIN_STATUS Object is optional. If there is no ADMIN_STATUS, all the values are treated as 0. If ADMIN_STATUS Object exists, A-bit is set up as the followings.
  【LSP0, LSP1】
  - 1 st roundtrip: A= 1　(By setting this, false fault detection during setup procedure can be eliminated)
  - 2 nd roundtrip and beyond: A=0
- RESVCONF message is optional.（Egress node can start to make user traffic flow taking the arrival of this message as a trigger.）
- Status of LSP changes as the followings depending on each message.
  【LSP0】
  - PATH: Routed→Reserved
  - RESV: Reserved→Connected
  【LSP1】
  - PATH: Routed→Reserved
  - RESV: Reserved→Reserved or Assigned
- Both of LSP0 and LSP1 execute refreshing.

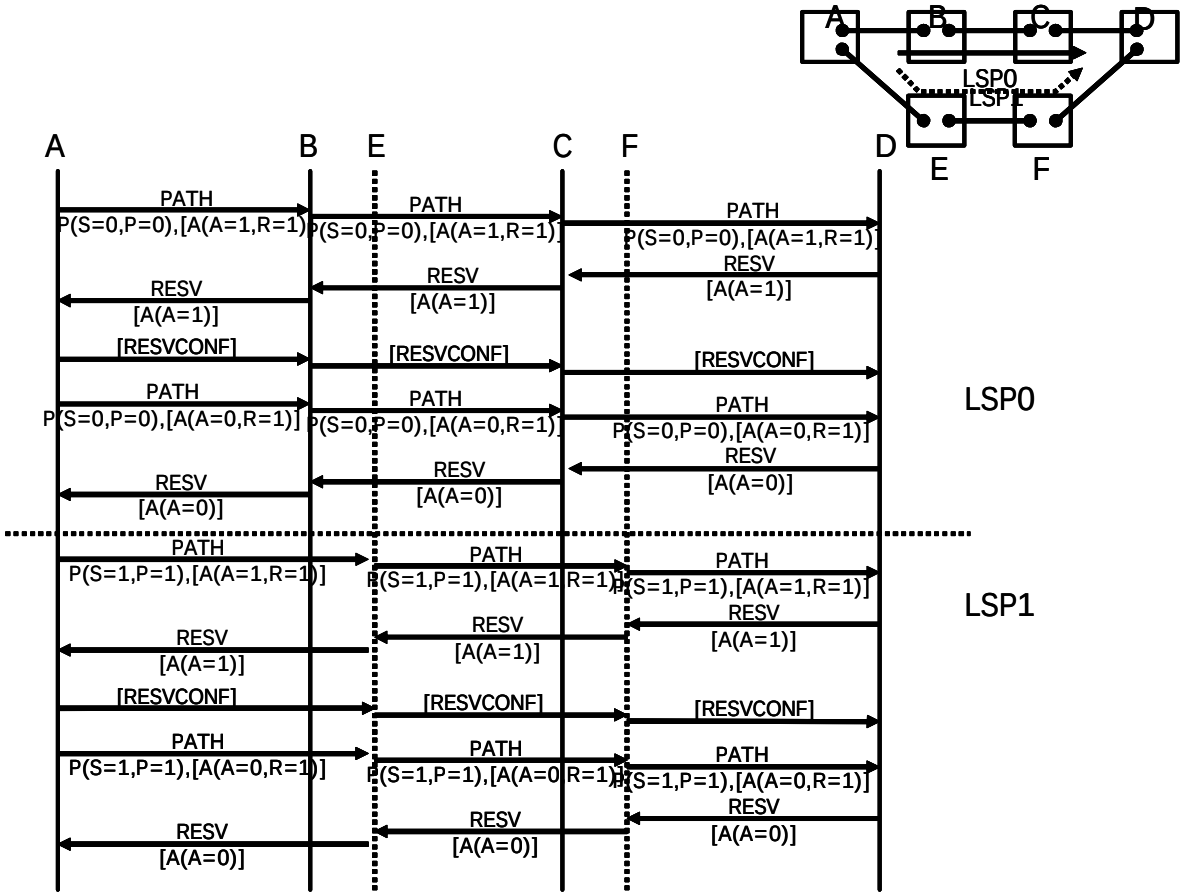


Fig. 3-1 Working/Recovery LSP setup signaling sequence in case of 1:1/Shared Mesh Restoration

3.1.1.2 Fault switching

When fault happened to occur in working LSP, notification of fault is sent to Ingress node (refer to Chapter 5). Ingress node switches the traffic from working LSP to recovery LSP after it changed the traffic status of recovery LSP corresponding to the working LSP of which fault was notified from “Reserved” or “Assigned” to “Connected” (hereafter, this operation is called as “activation”). Fig.3-2 shows the signaling sequence in this activation process.

- In PATH message of LSP1 (recovery LSP）, S-bit is changed from 1 to 0. By this, nodes on LSP1 know that this signaling is not “refresh” as usual but is “activation”, and activate this LSP.
- In PATH message of LSP0 (working LSP）, any bits are not changed (continue to “refresh” as so far).
- ADMIN_STATUS Object is optional. If there is no ADMIN_STATUS, all the values are treated as 0. If ADMIN_STATUS Object exists, A-bit is set up as the followings.
  【LSP0】
  A=0 or A=1 (If you don’t want to notify to control plane that a fault was detected, it is possible to do so by setting as A=1.)
  【LSP1】
  - 1 st roundtrip after detection of fault: A= 1 (By doing this, false fault detection during activation procedure can be eliminated)
  - 2 nd roundtrip and beyond: A=0
- The states of LSP0 and LSP1 are changed by fault switching as the followings:
  【LSP0】
  Remains as “Connected”
  【LSP1】
  Reserved or Assigned→Connected
- Refreshing of LSP0 and LSP1 is continued also after execution of fault switching. When refreshing was impossible by fault of control channel, etc., soft state is remained according to the procedure described in Chapter-9 of RFC3473.

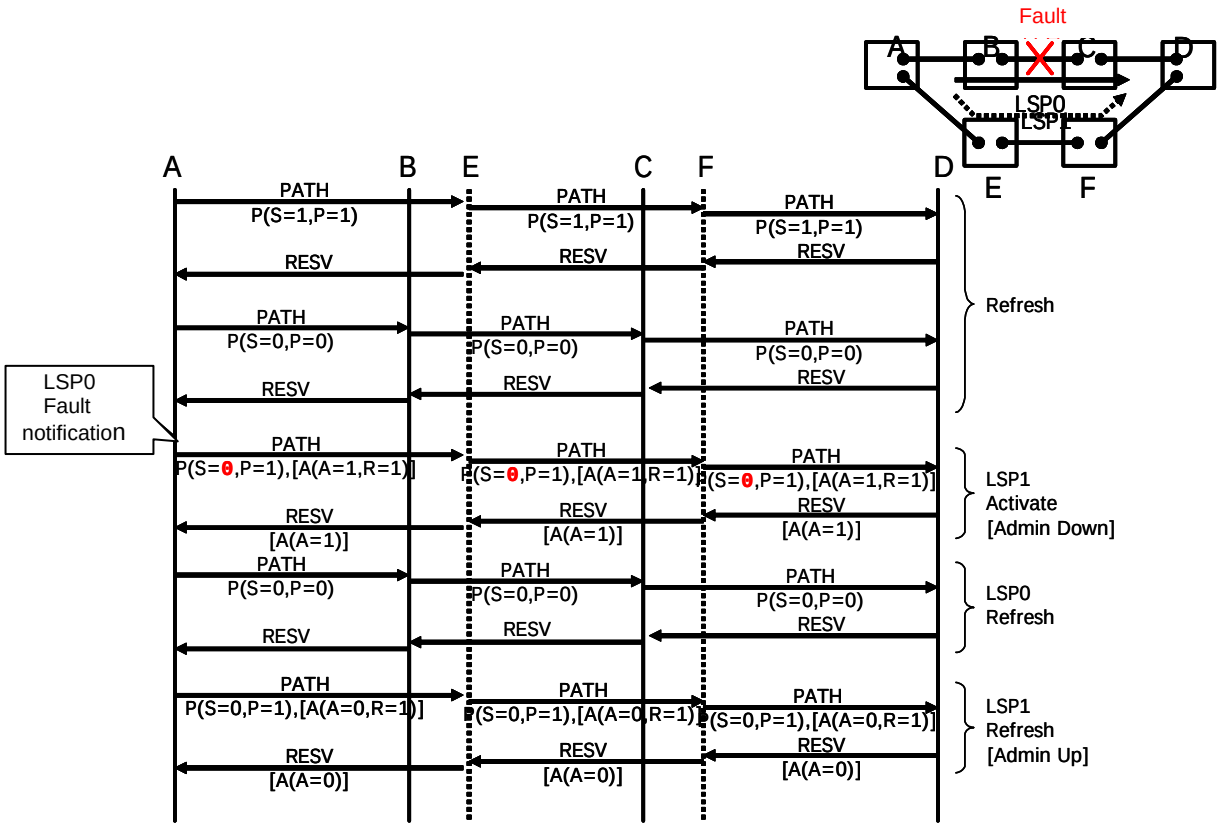


Fig.3-2 Signaling sequence of fault switching in case of 1:1/Shared Mesh Restoration

3.1.1.3 Switch-back

In 1:1/Shared Mesh Restoration, it is possible to switch-back when working LSP was recovered after fault switching was executed. Switch-back can be executed automatically by taking the fault recovery notification of working LSP as a trigger as well as manually. In switch-back operation, first the traffic of recovery LSP is switched to working LSP, and then the state of recovery LSP is changed from "Connected" to "Reserved" or "Assigned" (hereafter this operation is called as "de-activation"). When traffic is switched to working LSP, bridge&select (after flowing the traffic to both of the working LSP and the recovery LSP at the upstream and then switch the traffic to the working LSP at the downstream) is executed to limit the interruption to minimum. Fig.3-3 shows the signaling sequence during this procedure.

- When LSP0 was recovered, node-A sends NOTIFY message having ACK_Desired flag that Error Code/Sub Code built with "Notify Error/LSP Recovered" after switching the traffic of A→D direction of LSP1 so as to flow also to LSP0 (bridge). Node-D that received this message switches the traffic of A→D direction that it has received from LSP1 so as to receive from LSP0 (select), and switches the traffic of D→A direction of LSP1 so as to flow also to LSP0 (bridge). Then, node-D returns NOTIFY ACK signal with ACK_Desired flag to node-A. The Node-A that received this NOTIFY ACK switches the traffic of A→D direction so as to flow only to LSP0 (release bridge), and switches the traffic of D→A direction that has been received from LSP1 so as to be received from LSP0 (select), and return NOTIFY ACK signal to node-D. The node-D that received this NOTIFY ACK signal switch the traffic of D→A direction so as to flow only to the LSP0 (release bridge). After completion of these process, LSP1 is de-activated by PATH/RESV message with S=1.
- By switching-back, states of LSP0 and LSP1 are changed as the followings.
  【LSP0】
  Remained as Connected.
  【LSP1】
  Connected→Reserved or Assigned
- Refreshing of LSP0 and LSP1 are executed.

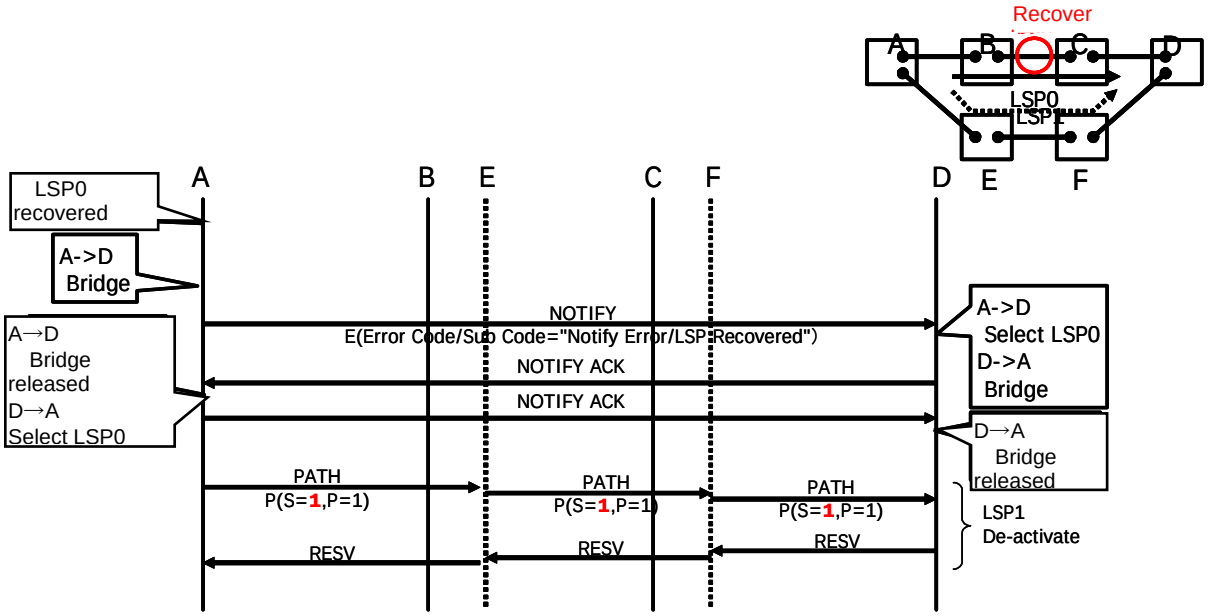


Fig.3-3 Signaling sequence of switch-back in case of 1:1/Shared Mesh Restoration

3.1.1.4 Releasing the working LSP

When switch-back is not executed, working LSP (LSP0) may be released after fault switching. Signaling sequence in this case is the same as usual one for releasing LSP (delete).

3.1.1.5 Forced Switch

When switching the traffic of LSP0 to LSP1: Signaling sequence is the same as the one for fault switching.
When switching the traffic of LSP1 to LSP0: Signaling sequence is the same as the one for switch-back.

3.1.1.6 Manual Switch

When switching the traffic of LSP0 to LSP1: Signaling sequence is the same as the one for fault switching.
When switching the traffic of LSP1 to LSP0: Signaling sequence is the same as the one for switch-back.

3.1.1.7 Lock-out and releasing the Lock-out

In locked-out LSP, both switching of automatic and manual are inhibited. Fig.3-4 shows the signaling sequences for lock-out and releasing the lock-out. By making L-bit=1 of ADMIN_STATUS Object in PATH/RESV message, lock-out is done and by making L-bit =0, rock-out is released. After executing fault switching, working LSP or reserved LSP may be locked-out for inhibiting automatic switch-back. In this case, lock-out should be released to execute switch-back.

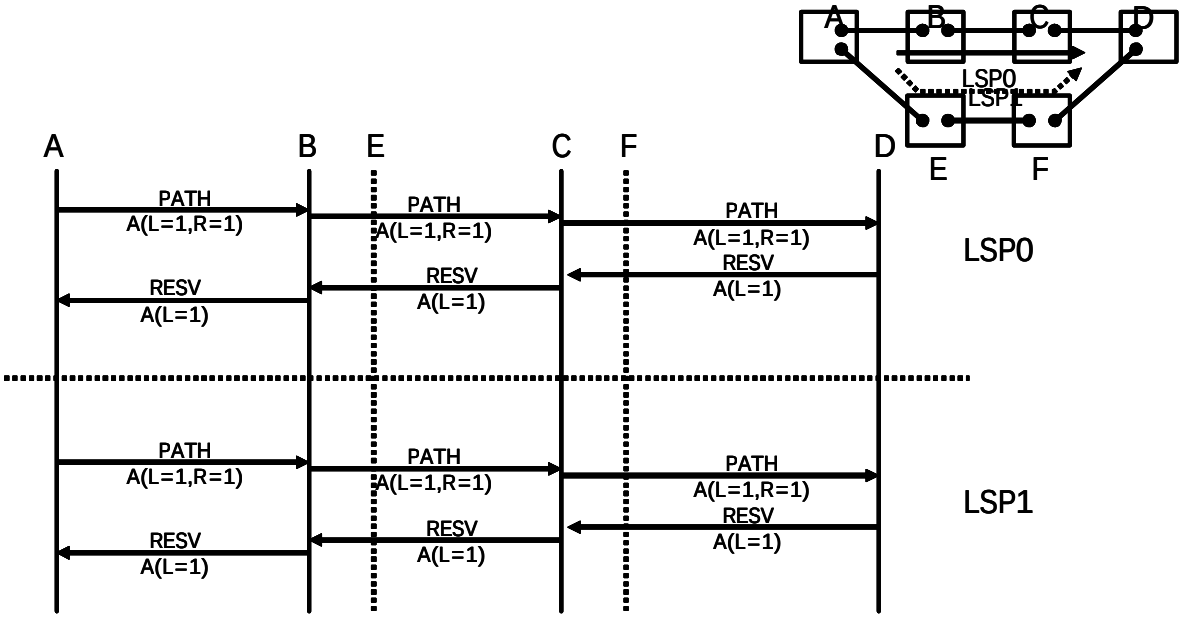


Fig.3-4 Signaling sequence in executing lock-out

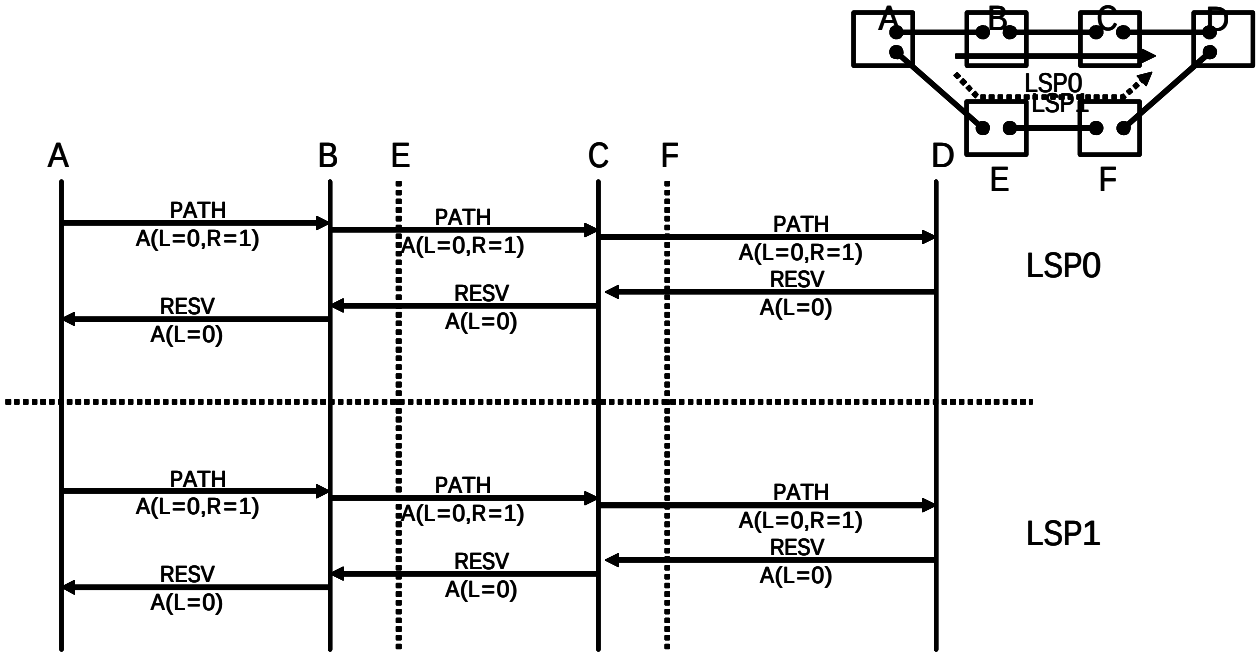


Fig.3-5 Signaling sequence when releasing lock-out

3.1.2 Message processing
3.1.2.1 Message processing at Ingress Node

Message processing at Ingress Node after receiving a fault notification is shown in Fig.3-6.

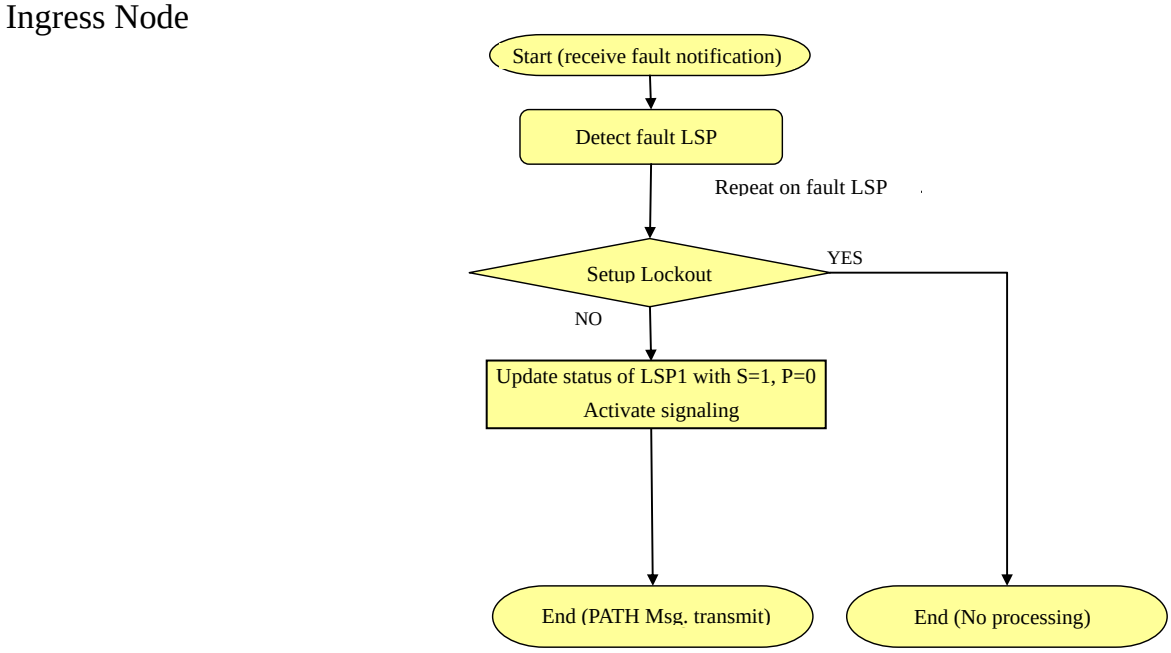


Fig.3-6 Message processing at Ingress Node

3.1.2.2 Message processing at Transit/Egress Node

There are three types of message at Transit/Egress Node as the followings.

- New message: A message for setting up a new LSP.
- Refresh message: A refresh message for sustaining a Soft State Session of RSVP.
- Change message: A message of which parameter has been changed from previously received message. In fault recovery by GMPLS, switching/switch-back of LSP is executed by changing the status of LSP using this Change message.

Fig.3-7 shows a processing flow at each node that received RSVP-TE Path/Resv message.

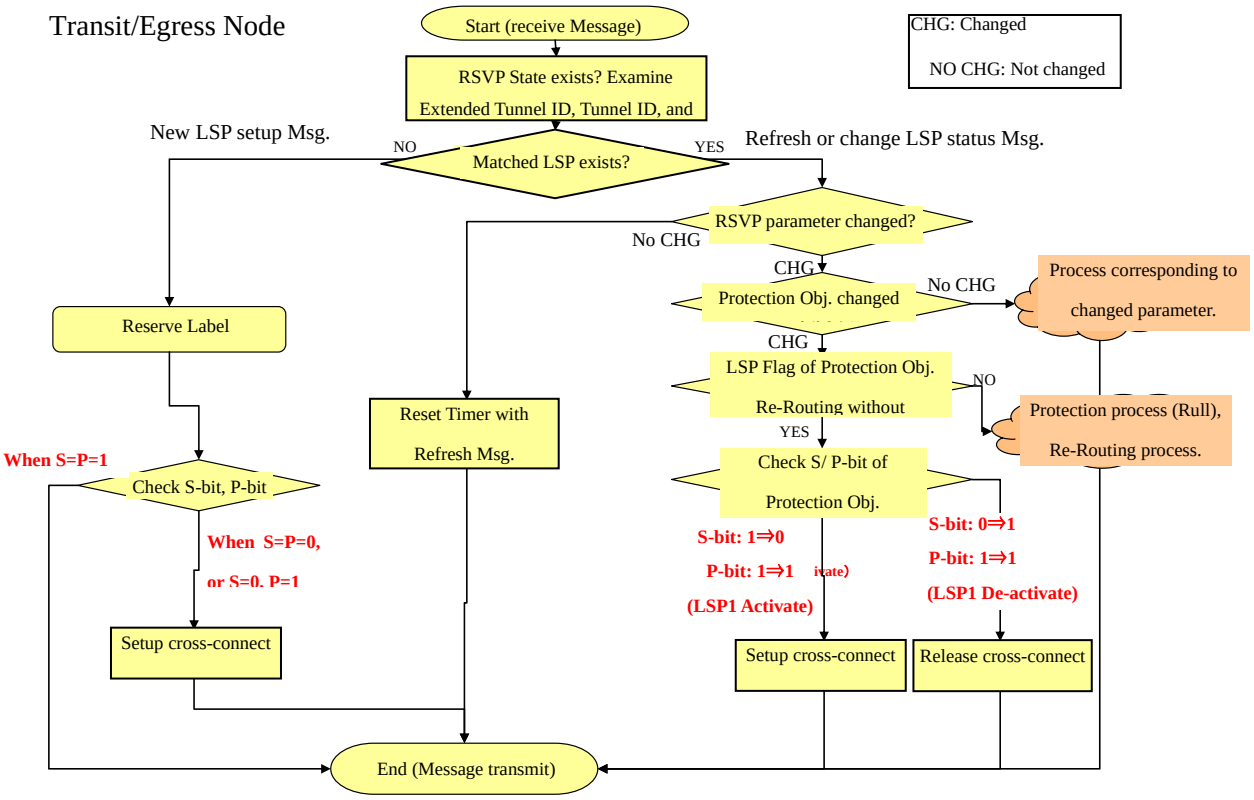


Fig.3-7 Message processing flow

NOTE：Setup of cross-connect may be executed at either timing of PATH/RESV Msg. In case of RESV Msg. timing, cross-connect is set up referring the S/P-bit of Protection Obj. that was received at PATH Msg. timing.

3.1.3 Massage Format

In the Table below, format of PATH, RESV and Notify messages are shown. Parameters used in Technology Verification WG are described except for the parameters related to fault recovery.

【Note】

- In order to improve interoperability, when transmitting the Refresh/Switch/Switch-back messages, make sure to include Object that was used at the time of setup.
- Option flag is an option in PIL Fault Recovery IA, and doesn't synchronize with RFC. Option Object is not required to include in message when transmitting, but must be possible to transfer and receive.

【PATH Message】

Object	Class	CType	Option	Parameter/Sub-Object	Value	Comment
RSVP Header	None	None		RSVP Version	1	IP alert Options not setup.
				Flag	0x00	
				Message Type	1 (Path)	
				Message Checksum	Checksum value	
				Sending TTL	1 or greater	
				Message Length	Path Msg. length	
SESSION	1	7		Destination Address	Node ID of destination	
				Tunnel ID	1 or greater	
				Extended Tunnel ID	Router ID of source	
RSVP HOP	3	3		Neighbor Address	Router ID of C-Plane	
				LIH	IF ID of C-Plane	
				IF_INDEX TLV - IPv4 Address - Interface－ID	 Node ID of D-Plan IF ID of D-Plane	
TIME VALUE	5	1		Refresh Interval	30000 msec（ref. value）	
LABEL REQUEST	19	4		LSP Encoding Type	8 ; Lambda（ref. value）	
				Switching Type	150 ; LSC（ref. value）	
				G-PID	31 ; POS（ref. value）	
PROTECTION	37	??		S, P, N bit	Follow to operation of Protection Restoration（section 3.1）	
				LSP flag	0x02（1:1 Rerouting without Extra Traffic）	
				Link flag	0x02（Unprotected）（ref. value）	
ASSOCIATION	198			Association Type	0x01（Recovery）	
				Association ID	Associated LSP ID	
SESSION ATTRIBUTE	207	7	O	Setup Priority	3（ref. value）	
				Holding Priority	4（ref. value）	
				Flag	0x00（ref. value）	
				Name	Arbitrary	
NOTIFY REQUEST	195	1	O	IPv4 Notify Address	Node ID	
ADMIN_STATUS	196	1	O	R, T, A, D	R: Reflect T: Testing A: Administratively Down D: Deletion in progress	R：expects for response T：Not specified A：Admin Down D: Graceful Deletion
EXPLICIT ROUTE	20	1		ERO Unnumbered Link - Router ID - Interface ID	(Next Hop) D-Plane の Node ID D-Plane の IF ID	
SENDER TEMPLATE	11	7		Sender IPv4 Address	Source Node ID	
				LSP ID	1 or greater	
SENDER TSPEC	12	1		Token Bucket Rate	0	
				Token Bucket Size	0	
				Peak Data Rate	19440000（ref. value）	
				Maximum Policed Unit	0	
				Maximum Packet Size	9180（ref. value）	
UPSTREAM LABEL	35	2		Generalized Label	Label value	In case of Bi-directional LSP
LSP _TUNNEL_IF_ID	227? 193	1	O	Router ID	Node ID of node that sends a message.	
				Interface ID	Interface value assigned to logical IF. 1or greater	Number not used in the existing IF
PRIMARY PATH ROUTE	??	1		ERO Unnumbered Link - Router ID - Interface ID	Working LSP's Node ID of D-Plane IF ID of D-Plane	Included in only 1 : 1 Re-routing without Extra Traffic and setup of recovery LSP.

【RESV message】

Object	Class	C-Type	Option	Sub-Object	Parameter	Comment
RSVP Header				RSVP Version	1	
				Flag	0x00	
				Message Type	2（Resv）	
				Message Checksum	Checksum value	
				Sending TTL	1 or greater	
				Message Length	Resv Msg. length	
SESSION	1	7				Same as PATH
RSVP HOP	3	3				Copy the PATH value
TIME VALUE	5	1				Same as PATH
RESV CONF	8	15	O	Receiver Address	Node ID	
NOTIFY REQUEST			O			
ADMIN STATUS			O			Same as PATH
LSP _TUNNEL_IF_ID	227? 193	1	O			Same as PATH
STYLE	8	1		Flag	0x00	
				Style	0x0a(Fixed Style)	
FLOW SPEC	9	2		Token Bucket Rate	0	
				Token Bucket Size	0	
				Peak Data Rate	19440000（ref. value）	
				Maximum Policed Unit	0	
				Maximum Packet Size	9180（ref. value）	
FILTER SPEC	10	7		Sender IPv4 Address	Source Node ID	
				LSP ID	1 or greater	
LABEL	16	2		Generalized Label	Label value	

【Notify message (for use in fault notification, switch-back, and End-End communication in switch-over in 1+1 Protection）】

Object	Class	C-Type	Option	Sub-Object	Parameter	Comment
RSVP Header				RSVP Version	1	
				Flag	0x00	
				Message Type	2（Resv）	
				Message Checksum	Checksum value	
				Sending TTL	1 or greater	
				Message Length	Notify Msg. length	
Message ID	23	1		Message ID	Node-unique value	Necessary Object Ack Desired
ERROR SPEC	6	3		IPv4 Error Node Address	Node ID	
				Flag		
				Error Code	Working Path Failure	
				Error Value	Reversion Request/Response	
				IF_INDEX －IPv4 Address －Interface－ID	 Node ID of D-Plane IF ID of D-Plane	
SESSION	1	7				Same as PATH
SENDER TEMPLATE	11	7				Same as PATH
SENDER TSPEC	12	1				Same as PATH
SESSION	1	7				Same as PATH
FLOW SPEC	9	2				Same as PATH
FILTER SPEC	10	7				Same as PATH

3.2 Protection signaling

Methods of protection to be discussed are as follows. As for definition of these protection methods, refer to [E2E].

1+1 Unidirectional Protection
1+1 Bi-directional Protection
1:1 Protection with Extra-Traffic

At the Ingress node, both of the working LSP and the reserved LSP have to be set up. It doesn't matter which should be set up working LSP or reserved LSP. It is also possible to setup both at the same time.
Route of the working LSP and the reserved LSP must be a node/link/SRLG disjoint route.

3.2.1 Setup of working LSP

(1) Signaling sequence
Signaling sequence in setting up the working LSP is shown Fig.3-8 and Fig.3-9. Fig.3-8 corresponds to a sequence that doesn't use the ResvConf message and ADMIN_SDTATUS. Fig.3-9 corresponds to a sequence that uses the ResvConf message and ADMIN_SDTATUS as an option. These sequences are just examples, and order of transmission of message is not necessarily required to be the same as the order shown in these diagrams. For example, order of setting up the path doesn't matter which LSP is first set up the working LSP (LSP0) or the reserved LSP (LSP1), or at the same time.

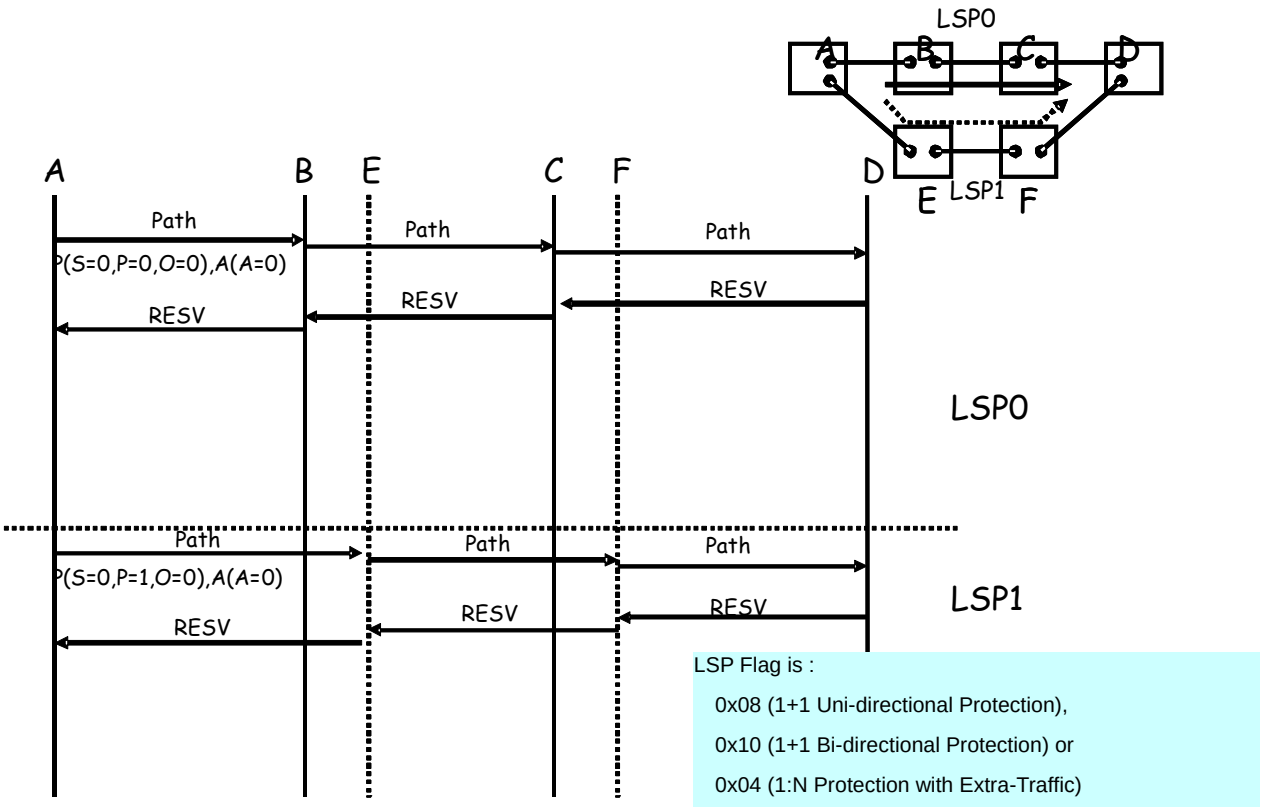


Fig.3-8 Signaling sequence in 1+1 Unidirectional Protection, 1+1 Bi-directional Protection, or 1:1 Protection with Extra-Traffic (in case the ResvConf message and ADMIN_STATUS are not used)

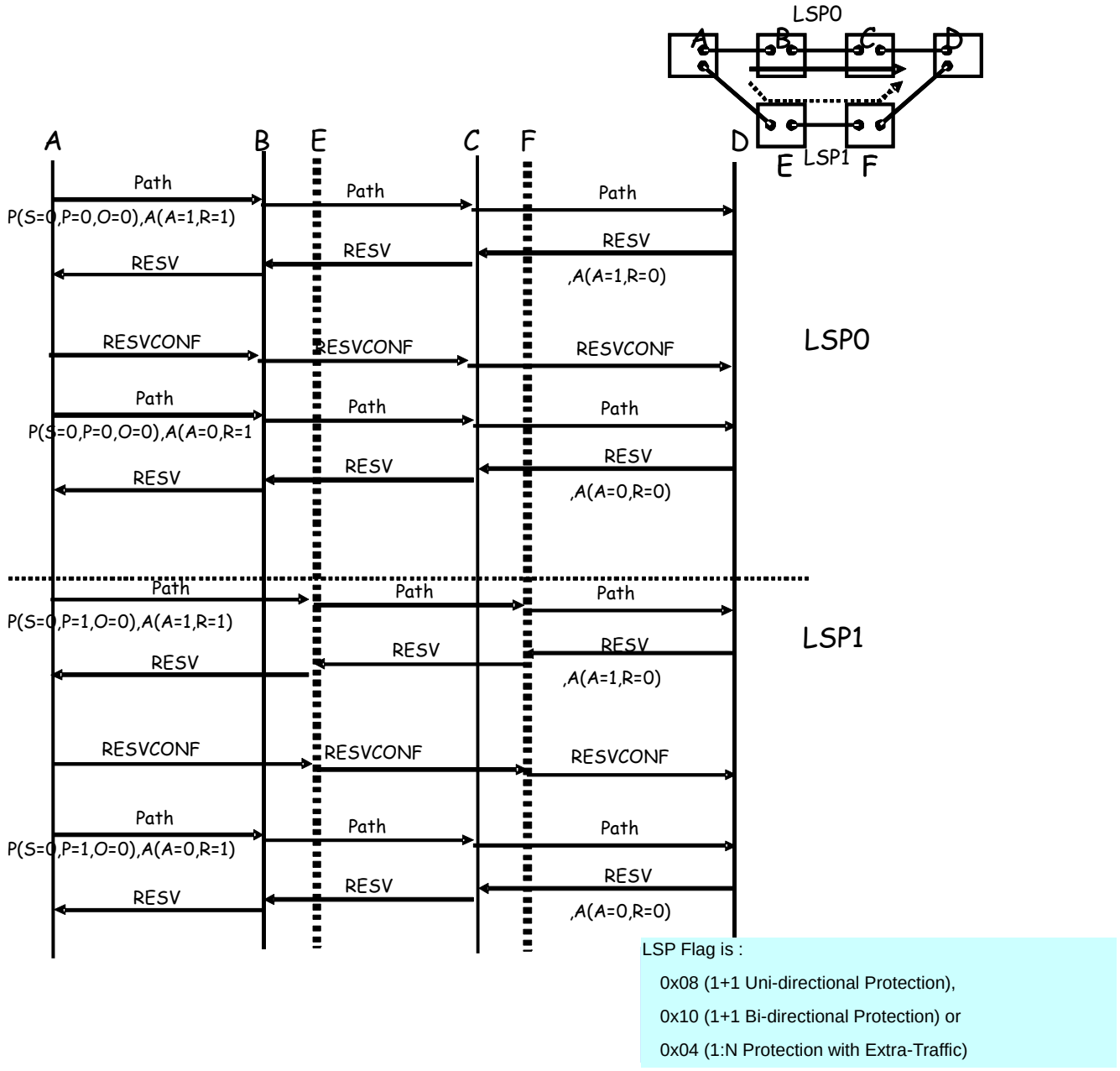


Fig.3-9 Signaling sequence in 1+1 Unidirectional Protection, 1+1 Bi-directional Protection, or 1:1 Protection with Extra-Traffic (in case the ResvConf message and ADMIN_STATUS are not used as an option)

(2) Content of object and parameter definition

Relating to message when Path is set up, contents of main objects relating to protection are shown in the table below. As for the details of definition of each message of RSVP (BNF, TLV bit assign), please refer to [RFC3473].

Into Path message, PROTECTION Object and ASSOCIATION object defined in [E2E] are included.

The same value of SESSION object is set to the working LSP and the recovery LSP. The LSP ID of SENDER_TEMPLATE object or FILTER_SPEC object take a different value.

At the Ingress node and the Egress node, Notify Request is included to Path message and Resv message. But, when executing switching in D-Plane (N bit = 1), fault detection is executed in D-Plane, and if fault notification is not required, Notify Request is not necessary.

Path message

Object	Field/Sub object	Value	Comment
SESSION	Destination address	IP address of Egress node	
	Tunnel ID	Unique value to identify tunnel that is the same as the one of recovery LSP.	

	Extended Tunnel ID	IP address of Ingress node	
SENDER TEMPLATE	Sender IPv4 address	IP address of Ingress node	
	Sender LSP ID	Unique value within tunnel that is different from the one of recovery LSP.	
PROTECTION	Secondary (S)	0	Refer to definition of [E2E]
	Protecting (P)	0	
	Notification (N)	0 or 1	In this IA, only the case of 1 is specified.
	Operational (O)	0	
	LSP (Protection Type) Flags	One of the followings: 0x04(1:N Protection with Extra-Traffic) 0x08(1+1 Unidirectional Protection) 0x10(1+1 Bi-directional Protection)	
	Link Flags	Arbitrary value within specifications (refer to RFC3471)	
NOTIFY_REQUEST	IPv4 Notify Node Address	IP address of Ingress node	Use this object only when notification is required.
ADMIN_STATUS	Reflect (R)		Use this object if necessary (irrelevant to protection sequence)
	Testing (T)		
	Administratively down (A)		
	Deletion in progress (D)		
ASSOCIATION	Association Type	1	Recovery (R)
	Association ID	LSP ID value of recovery LSP	
	Association Source	IP address of Ingress node	

Resv message

Object	Field/Sub object	Value	Comment
NOTIFY_REQUEST	IPv4 Notify Node Address	IP address of Egress node	Use this object only when notification is required.
ADMIN_STATUS	Reflect (R)	0	Use this object if necessary (irrelevant to protection sequence)
	Testing (T)	Copied the same value as the receive Path message.	
	Administratively down (A)		
	Deletion in progress (D)		

(3) Message processing in each of Ingress, Intermediate, and Egress node

(a) Processing at the Ingress node

At the Ingress node, route of working LSP is calculated and Path message is transmitted including the above Object and the determined route to ERO (Explicit Rout Object). When the Path message is transmitted, label of upstream is assigned and setup of cross-connect has been completed after receiving the Resv message.

In case of 1+1 Protection, the Ingress node transfers the traffic to working LSP (LSP0) after receiving Resv message. In case of bi-directional LSP, as for the upstream traffic, the Ingress node selects so as to receive the traffic from the working LSP (LSP0).

In case of 1:1 Protection, the Ingress node transfers the traffic to the working LSP (LSP0). In case of bi-directional LSP, as for the upstream traffic, the Ingress node selects so as to receive the traffic from the working LSP (LSP0).

(b) Processing at the Intermediate node

When transmitting the Path message and Resv message, label of upstream/downstream is assigned to respective message, and setup of cross-connect has been completed.

(c) Processing at the Egress node

When transmitting the Resv message, label of downstream is assigned, and setup of cross-connect has been completed.

In case of 1+1 Protection, the Egress node selects so as to receive the traffic from the working LSP (LSP0). In case of bi-directional LSP, as for the upstream traffic, the Egress node transfers the traffic only to the working LSP (LSP0).

3.2.2 Setup of recovery LSP

(1) Signaling sequence
Signaling sequence is shown in Fig.3-8 and Fig.3-9 in section 3.2.1.

(2) Content of object and parameter definition
Relating to message when LSP is set up, contents of main objects relating to protection are shown in the table below. IPv4 (or IPv6) tunnel sender address of SESSION Object and SENDER_TEMPLATE is set the same value as the one of the working LSP. By doing this, working pair can be distinguished from the recovery pair. LSP ID of SENDER_TEMPLATE/FILTER_SPEC object is used a different value between the working LSP and the recovery LSP.

Path message

Object	Field/Sub object	Value	Comment
SESSION	Destination address	IP address of Egress node	
	Tunnel ID	Unique value to identify tunnel that is the same as the one of recovery LSP.	
	Extended Tunnel ID	IP address of Ingress node	
SENDER TEMPLATE	Sender IPv4 address	IP address of Ingress node	
	Sender LSP ID	Unique value within tunnel that is different from the one of working LSP.	
PROTECTION	Secondary (S)	0	Refer to definition of [E2E]
	Protecting (P)	1	
	Notification (N)	0 or 1	In this IA, only the case of 1 is specified.
	Operational (O)	0	
	LSP (Protection Type) Flags	One of the followings: 0x04(1:N1 Protection with Extra-Traffic) 0x08(1+1 Unidirectional Protection) 0x10(1+1 Bi-directional Protection)	
	Link Flags	Arbitrary value within specifications (refer to RFC3471)	
NOTIFY_REQUEST	IPv4 Notify Node Address	IP address of Ingress node	Use this object only when notification is required.
ADMIN_STATUS	Reflect (R)		Use this object if necessary (irrelevant to protection sequence)
	Testing (T)		
	Administratively down (A)		
	Deletion in progress (D)		
ASSOCIATION	Association Type	1	Recovery (R)
	Association ID	LSP ID value of working LSP	
	Association Source	IP address of Ingress node	

Resv message

Object	Field/Sub object	Value	Comment
NOTIFY_REQUEST	IPv4 Notify Node Address	IP address of Egress node	Use this object only when notification is required.
ADMIN_STATUS	Reflect (R)	0	Use this object if necessary (irrelevant to protection sequence)
	Testing (T)	Copy the same value as the one of received Path message	
	Administratively down (A)		
	Deletion in progress (D)		

(3) Message processing in each of Ingress, Intermediate, and Egress node
(a) Processing at the Ingress node
At the Ingress node, recovery LSP is determined. Route of recovery LSP is selected as link/node/SRLG-disjoint route with the route of working LSP. Path message is transmitted including the above Object and the determined route to ERO. When Path message is

transmitted, label of upstream is assigned. In case of 1+1 Protection, the Ingress node copies the traffic and transfers it also to the recovery LSP (LSP1) after receiving Resv message.

(b) Processing at the Intermediate node

When transmitting the Path message and Resv message, label of upstream/downstream is assigned to respective message. When Resv message is transmitted, set-up of cross-connect has been completed.

(b) Processing at the Egress node

When transmitting the Resv message, label of downstream is assigned. In case of 1+1 Protection and bi-directional LSP, as for the upstream traffic, the Egress node copies the traffic and transfers it also to the recovery LSP (LSP1).

3.2.3 Switching

In this IA, it is presupposed that switching is executed by a sequence on D-Plane. Foe example, in case of OTN-based Lambda Switch LSP, it is possible to use the method of linear protection (APS channel of ODU) that is defined in G.873.1.
In SDH, there is a specification of MSP (K1/K2 byte) and VC trail (K3/K4, protocol has not yet been defined) in G.841.
As for the sequence of switching in C-Plane (N-bit is 0), it is excluded from this IA. For brief description, refer to [E2E]
In the following, change of signaling state is shown presupposing that switching is executed by a sequence on D-Plane.

(1) Signaling sequence

Signaling sequence of switching is shown in Fig.3-10. This sequence is just an example. Order of updating the signaling statues doesn't matter which LSP is first updated the working LSP (LSP0) or the recovery LSP (LSP1), or at the same time.

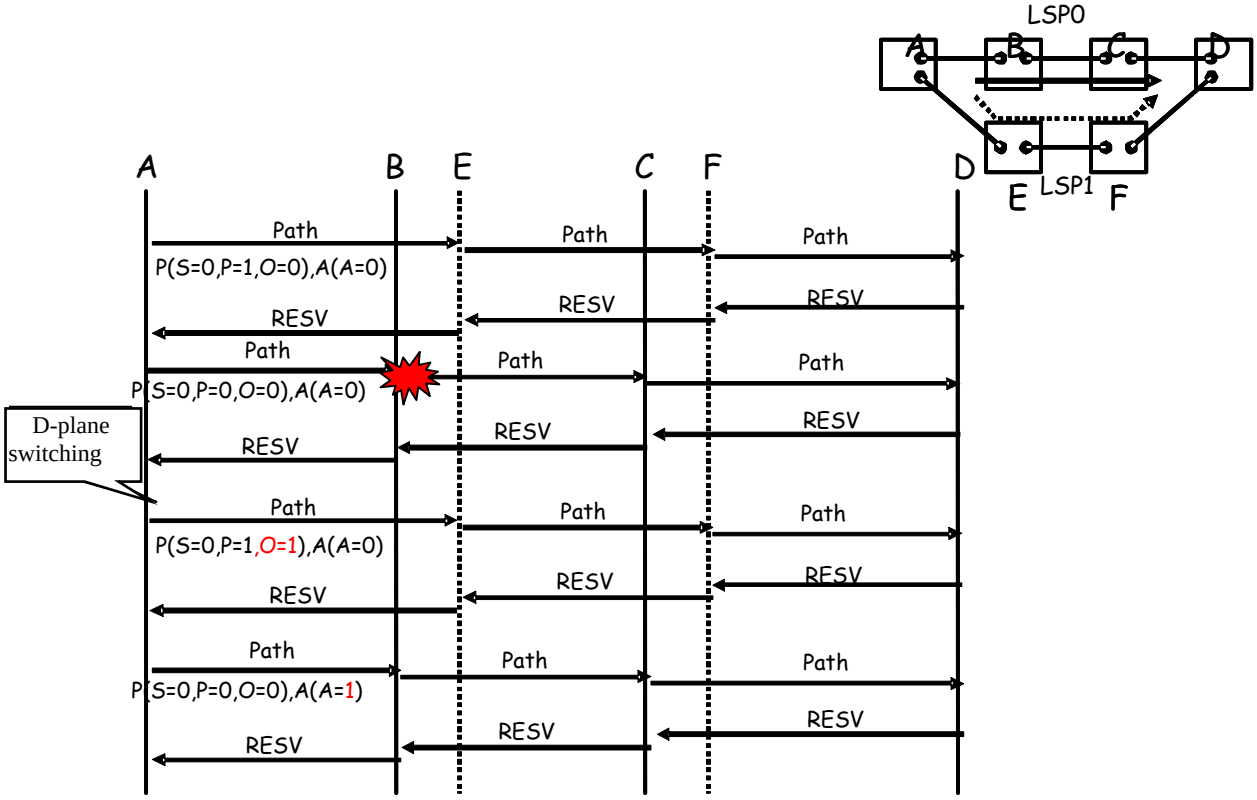


Fig.3-10 Signaling sequence in case of 1+1 Uni-directional Protection, 1+1 Bi-directional Protection, or 1:1 Protection with Extra-Traffic (A procedure for setting up the A-bit of working LSP to 1 in order to restrain fault monitoring is optional.)

(2) Content of object and parameter definition

Since it is possible to know which LSP is currently working if the Ingress/Egress nodes are understanding the signaling status of working LSP and recovery LSP after completion of switching, the status of working LSP (LSP0) is not changed and only O-bit of recovery LSP (LSP1) is changed from 0 to 1.

At the intermediate node, it is necessary to know which LSP is currently working considering about the fault monitoring. But, since it is possible to know which LSP is currently working if the intermediate node understands the switching sequence of D-Plane, there is no problem if value of 0-bit was not changed.

It is also possible to set up A-bit of ADMIN_STATUS of LSP that has been a working LSP before switching to 1 in order to retrain fault

monitoring, which is optional. If the fault was recovered, A-bit can be returned to 0 at an arbitrary time point.

(3) Message processing in each of Ingress, Intermediate, and Egress node
Changing of bridge/select is executed at the both end nodes, and there is no change in assigned status of label. Status of cross-connect is not changed at the Intermediate node.

(a) Processing at the Ingress node
In case of 1+1 Protection and bi-directional, the Ingress node that has been receiving the upstream traffic from the working LSP (LSP0) becomes to receive it from the recovery LSP (LSP1) after switching.
In case of 1:1 Protection, the Ingress node becomes to transmit traffic to the recovery LSP (LSP1). In case of bi-directional, the Ingress node becomes to receive upstream traffic from the recovery LSP (LSP1).

(b) Processing at the Intermediate node
Status of cross-connect is not changed.

(c) Processing at the Egress node
In case of 1+1 Protection, the Egress node that has been receiving the traffic from the working LSP (LSP0) becomes to receive traffic from the recovery LSP (LSP1).
In case of 1:1 Protection, the Egress node becomes to receive traffic from the recovery LSP (LSP1). In case of bi-directional, the Egress node becomes to transmit upstream traffic to the recovery LSP (LSP1).

3.2.4 Switch-back operation

In this IA, it is presupposed that switch-back is executed on D-Plane similarly to the switching operation. Also in switch-back operation, like the switching operation, it is possible to apply OTN linear protection (APS channel of ODU) defined in G.873.1 or SDH MSP (K1/K2 byte).

As for the sequence of switching in C-Plane (N-bit is 0), it is excluded from this IA. For brief description, refer to [E2E]
In the following, change of signaling state is shown presupposing that switching is executed by a sequence on D-Plane.

(1) Signaling sequence
Signaling sequence of switch-back is shown in Fig.3-11.
Order of updating the signaling status doesn't matter which LSP is first updated either the working LSP (LSP0) or the recovery LSP (LSP1), or at the same time.

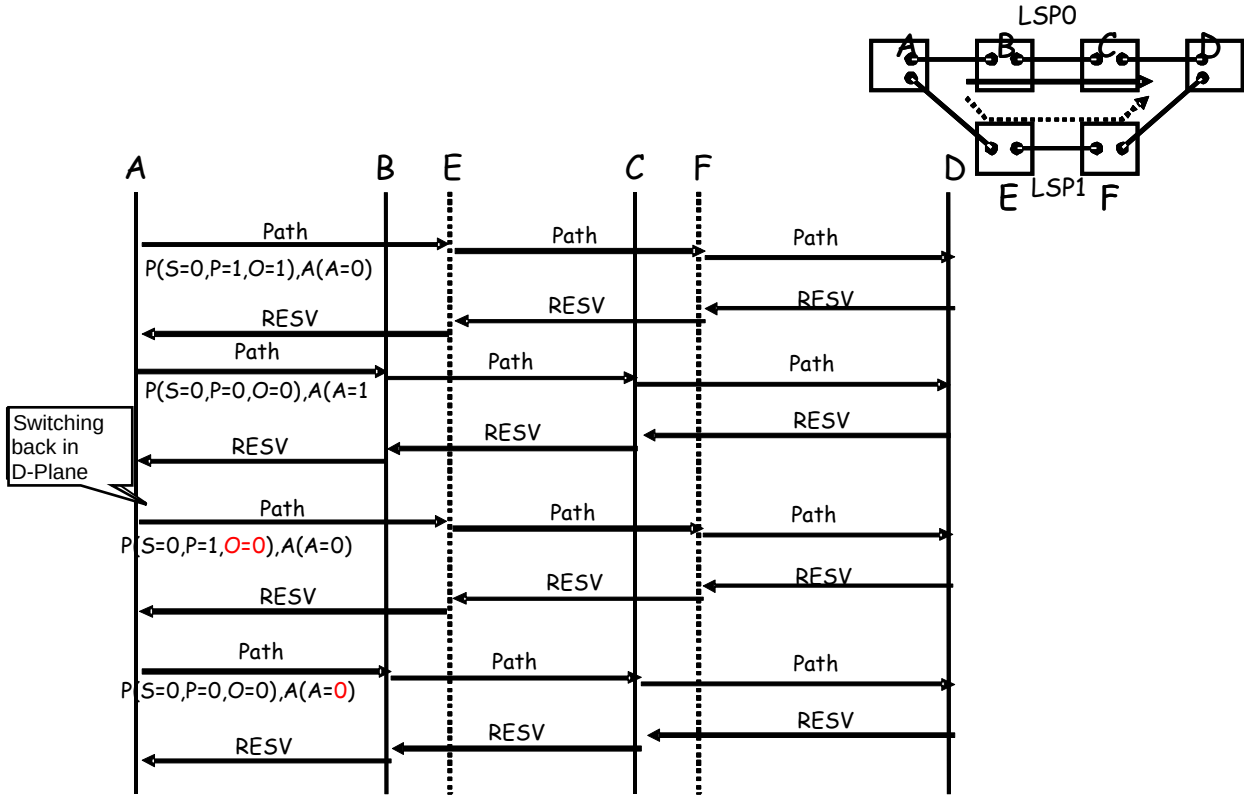


Fig.3-11 Signaling sequence in case of 1+1 Uni-directional Protection, 1+1 Bi-directional Protection, or 1:1 Protection with Extra-Traffic (Since restraining the fault monitoring has been executed, A-bit of working LSP was set to 1. The procedure to change A-bit to 0 is optional.)

(2) Content of object and parameter definition

As for the status of signaling of working LSP and recovery LSP, a value of 0-bit of reserved LSP changes after completion of switching operation. That is, 0-bit of the recovery LSP (LSP1) that has been working before switch-back changes from 1 to 0. A-bit of ADMIN_STATUS of LSP0 that has been in fault before switch-back was has been set to 1, this is set to 0 after switch-back.

(3) Message processing in each of Ingress, Intermediate, and Egress node

Changing of bridge/select is executed at the both end nodes, and there is no change in assigned status of label. Status of cross-connect is not changed at the Intermediate node.

(a) Processing at the Ingress node

In case of 1+1 Protection and bi-directional, the Ingress node that has been receiving the upstream traffic from the recovery LSP (LSP1) becomes to receive it from the working LSP (LSP0) after switching.

In case of 1:1 Protection, the Ingress node becomes to transmit traffic to the working LSP (LSP0). In case of bi-directional, the Ingress node becomes to receive upstream traffic from the working LSP (LSP0)

(b) Processing at the Intermediate node

Status of cross-connect is not changed.

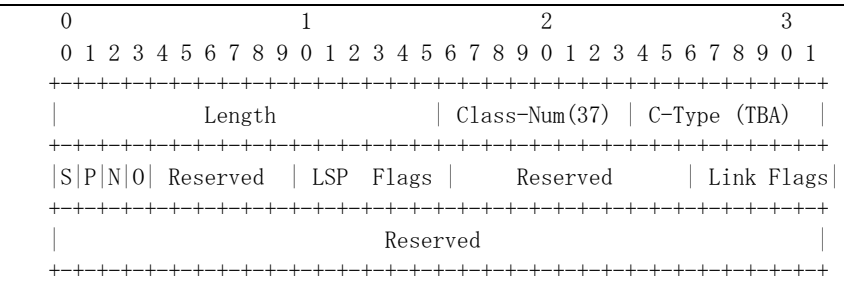
(c) Processing at the Egress node

In case of 1+1 Protection, the Egress node that has been receiving the traffic from the recovery LSP (LSP1) becomes to receive traffic from the working LSP (LSP0) after completion of switching.

In case of 1:1 Protection, the Egress node becomes to receive traffic from the working LSP (LSP0). In case of bi-directional, the Egress node becomes to transmit upstream traffic to the working LSP (LSP0).

3.3 Object

3.3.1（Extended）protection object



Secondary (S): 1 bit
When set to 1, this bit indicates that the requested LSP is a secondary LSP . When set to 0 (default), it indicates that the requested LSP is a primary LSP .

Protecting (P): 1 bit
When set to 1, this bit indicates that the requested LSP is a protecting LSP . When set to 0 (default), it indicates that the requested LSP is a working LSP . The combination, S set to 1 with P set to 0 is not valid.

Notification (N): 1 bit
When set to 1, this bit indicates that the control plane message exchange is only used for notification during protection switching. When set to 0 (default), it indicates that the control plane message exchanges are used for protection switching purposes. The N bit is only applicable when the LSP Protection Type Flag is set to either 0x04 (1:N Protection with Extra-Traffic), or 0x08 (1+1 Unidirectional Protection) or 0x10 (1+1 Bi-directional Protection). The N bit MUST be set to 0 in any other case.

Operational (O): 1 bit
When set to 1, this bit indicates that the protecting LSP is carrying the normal traffic after protection switching. The O bit is only applicable when the P bit is set to 1 and the LSP Protection Type Flag is set to either 0x04 (1:N Protection with Extra-Traffic), or 0x08 (1+1 Unidirectional Protection)or 0x10 (1+1 Bi-directional Protection). The O bit MUST be set to 0 in any other case.

Reserved: 5 bits
This field is reserved. It MUST be set to zero on transmission and MUST be ignored on receipt. These bits SHOULD be passed through unmodified by transit nodes.

LSP (Protection Type) Flags: 6 bits
Indicates the desired end-to-end LSP recovery type. A value of 0 implies that the LSP is "Unprotected". Only one value SHOULD be set at a time. The following values are defined. All other values are reserved.

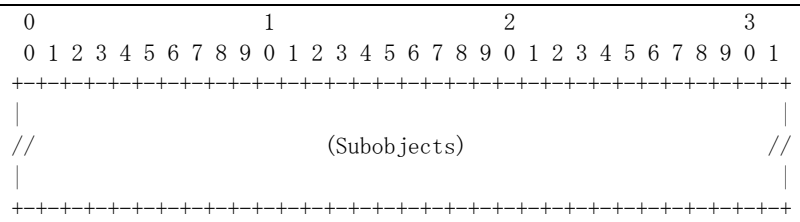
0x00	Unprotected
0x01	(Full) Re-routing
0x02	Re-routing without Extra-Traffic
0x04	1:N Protection with Extra-Traffic
0x08	1+1 Unidirectional Protection
0x10	1+1 Bi-directional Protection

Reserved: 10 bits
This field is reserved. It MUST be set to zero on transmission and MUST be ignored on receipt. These bits SHOULD be passed through unmodified by transit nodes.

Link Flags: 6 bits
Indicates the desired link protection type (see [RFC3471]).

Reserved field: 32 bits
Encoding of this field is detailed in [SEGREC].

3.3.2 Primary Path Route Object



The contents of a PRIMARY_PATH_ROUTE object are a series of variable-length data items called subobjects (see Section 15.3).

To signal a secondary protecting LSP , the Path message MAY include one or multiple PRIMARY_PATH_ROUTE objects, where each object is meaningful. The latter is useful when a given secondary protecting LSP must be link/node/SRLG disjoint from more than one primary LSP (i.e. is protecting more than one primary LSP).

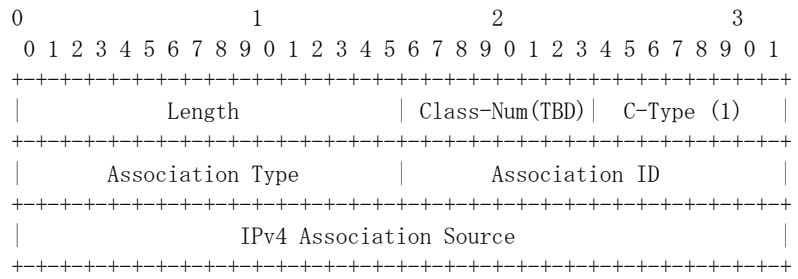
Sub-Objects:

The following subobjects are currently defined for the PRIMARY PATH ROUTE object:

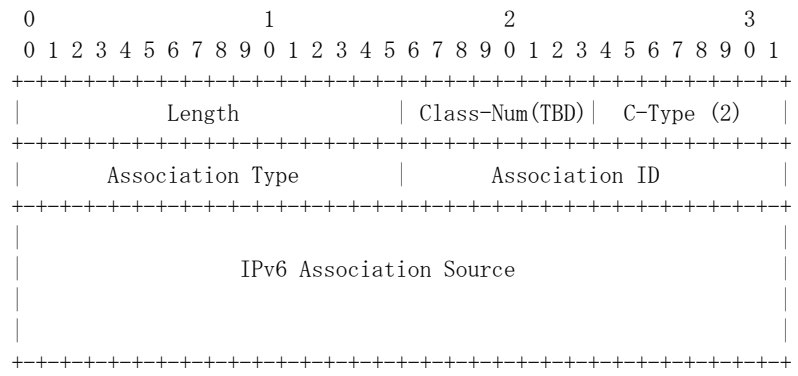
- Sub-Type 1: IPv4 Address (see [RFC 3209])
- Sub-Type 2: IPv6 Address (see [RFC 3209])
- Sub-Type 3: Label (see [RFC-3473])
- Sub-Type 4: Unnumbered Interface (see [RFC-3477])

3.3.3 Association Object

IPv4 ASSOCIATION object:



IPv6 ASSOCIATION object:



Association Type: 16 bits

Indicates the type of association being identified. Note that this value is considered when determining association. The following are values defined in this document.

Value	Type
0	Reserved
1	Recovery (R)

Association ID: 16 bits

A value assigned by the LSP Ingress. When combined with the Association Type and Association Source, this value uniquely identifies an association.

Association Source: 4 or 16 bytes

An IPv4 or IPv6 address, respectively, that is associated to the node that originated the association.

3.3.4 Definition of S, P, O-bit in Protection Obj

- S-bit
  S-bit indicates whether cross-connect in D-Plane is set up or not in setting up of LSP.
  0： cross-connect is set up
  1： cross-connect is not set up.

- P-bit
  P-bit indicates working or recovery of LSP. Working/Recovery is fixed regardless of existence of fault (unless the name was changed).
  0： Working LSP
  1： Recovery LSP

- O-bit
  O-bit indicates whether the recovery LSP (P=1) of the following Protections are used (operated) for fault recovery.
  0x04 1:N Protection with Extra-Traffic
  0x08 1+1 Unidirectional Protection
  0x10 1+1 Bi-directional Protection

  0：Not-working
  1：Working

4. Routing
4.1 General description about routing
4.1.1 Expansion for fault recovery

Among the Sub-TLVs added to expand the OSPF-TE and GMPLS, Sub-TLVs that are deeply relating to fault recovery are shown in Table.2-1. [OSPF-TE, GMPLS-OSPF, SHARABLE-OSPF]

Table.2-1

Sub-TLV Type	length	Name
1	1	Link type [OSPF-TE]
5	4	Traffic engineering metric [OSPF-TE]
6	4	Maximum bandwidth [OSPF-TE]
7	4	Maximum reservable bandwidth [OSPF-TE]
8	32	Unreserved bandwidth [OSPF-TE]
11	8	Link Local/Remote Identifiers [GMPLS-OSPF]
14	4	Link Protection Type [GMPLS-OSPF]
15	variable	Interface Switching Capability Descriptor [GMPLS-OSPF]
17	variable	Shared Risk Link Group [GMPLS-OSPF]
TBD	variable	Sharable Bandwidth [SHARABLE-OSPF]

General utilization methods of Sub-TLV are described in the followings.

Link type
Link type indicates a type of link (Point-to-point or Multi-access). [OSPF-TE]

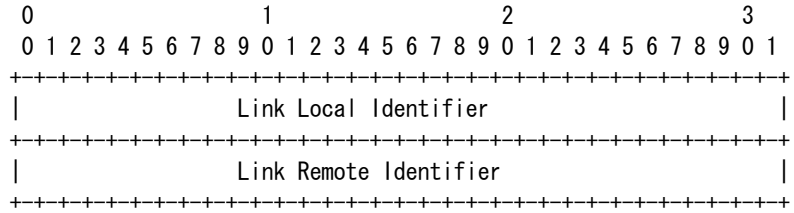
Traffic engineering metric
This is a metric of link used for traffic engineering, and is different from ordinary metric information of OSPF. [OSPF-TE]

Maximum bandwidth
This indicates the maximum bandwidth available in relevant link. [OSPF-TE]

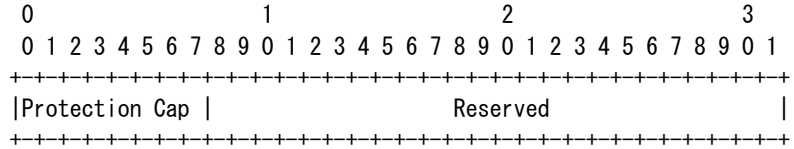
Maximum reservable bandwidth
This indicates the maximum bandwidth reservable in relevant link. [OSPF-TE]

Unreserved bandwidth
This indicates the bandwidth that is unreserved in relevant link with 0～7 of priority level. [OSPF-TE]

Link Local/Remote Identifiers
This is the information to identify the unnumbered type interface.
Pair of local-side link and remote-side link is advertised. The format is shown below. [GMPLS-OSPF]



Link Protection Type
This is used for identifying the type of Protection of relevant link. The format is shown below. [GMPLS-OSPF]



” Protection Cap” in the first octet shows the type of Protection as the followings.

0x01 Extra Traffic
0x02 Unprotected
0x04 Shared

0x08 Dedicated 1:1
0x10 Dedicated 1+1
0x20 Enhanced
0x40 Reserved
0x80 Reserved

Shared Risk Link Group (SRLG)

SRLG is used for identifying a group of multiple links to which the fault influences. The format of SRLG is shown in Fig.4-1. [GMPLS-OSPF]

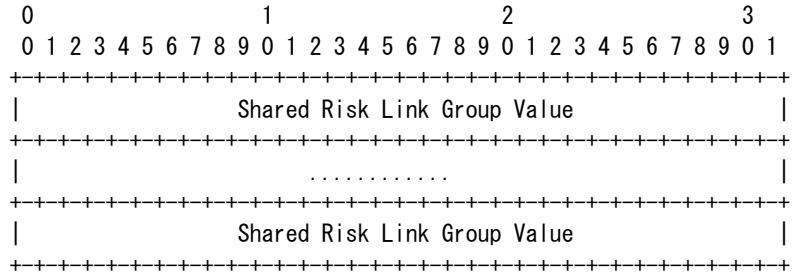


Fig.4-1

Interface Switching Capability Descriptor

This is used for advertising information such as the type of interface and available bandwidth, etc. The format of this descriptor is shown in Fig.4-2. [GMPLS-OSPF]

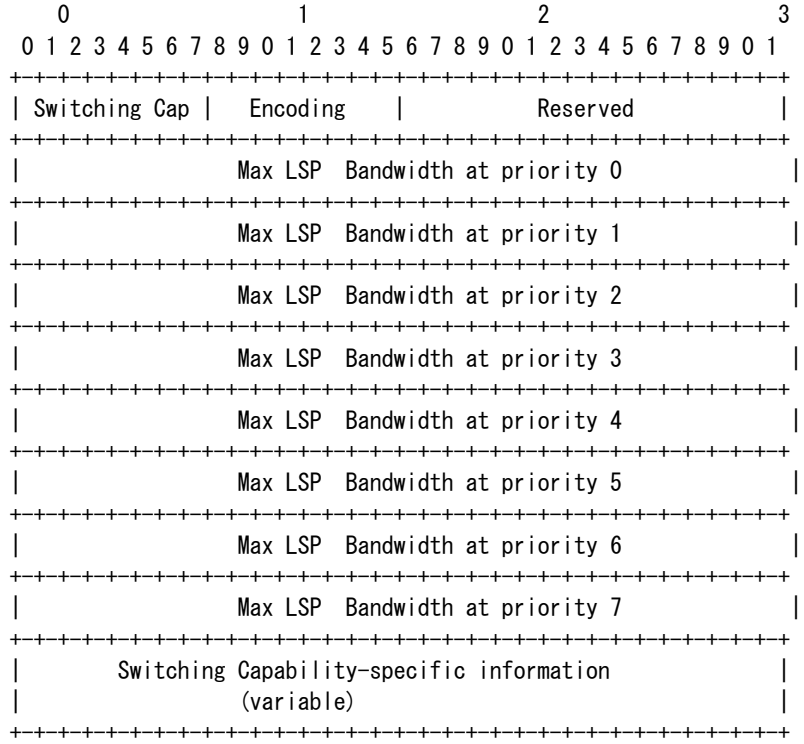


Fig.4-2

In case the switching capability is SONET/SDH, a field shown below is added to this descriptor to be used for utilizing the minimum bandwidth. [GMPLS-OSPF]

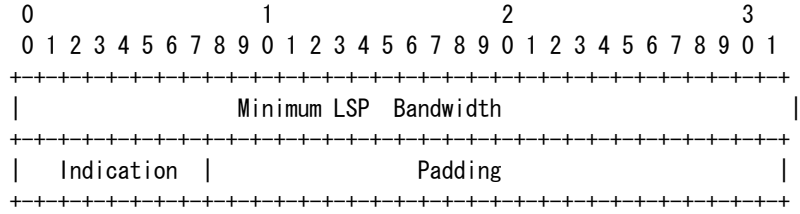


Fig.4-3

Sharable Bandwidth

Sharable Bandwidth is used for advertising the sharable bandwidth. The format of this is shown in Fig.4-4. [SHARABLE-OSPF]

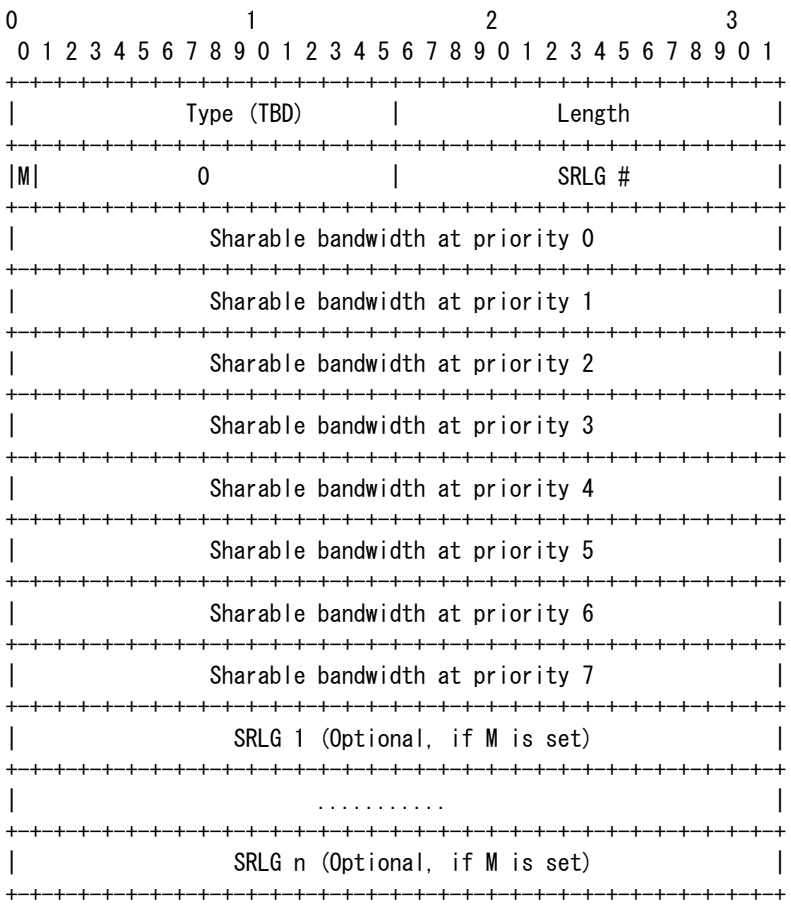


Fig.4-4

Forwarding adjacency

This is an attribute of FA.

General utilization method of Sub-TLV that is deeply relating to fault recovery among the Sub-TLVs advertised by OSPF is shown below.

Link type

Link type of FA becomes point-to-point. [LSP-HIER]

Traffic engineering metric

In case the Ingress of FA is not set up, a value obtained by subtracting 1 from metric of FA-LSP path is set as the metric value. [LSP-HIER]

Maximum bandwidth

In case the Ingress of FA is not set up, a value of FA-LSP path is set as the value of FA's bandwidth. [LSP-HIER]

Maximum reservable bandwidth

This value is set as the initial value of the maximum reservable bandwidth of F-LSP. [LSP-HIER]

Unreserved bandwidth

This value is set as the initial value of the unreserved bandwidth of F-LSP. [LSP-HIER]

Link Local/Remote Identifiers

When setting up the LSP, Ingress of LSP assigns the local Link Identifier and notifies it to each node on the path with LSP _TUNNEL_INTERFACE_ID in PATH message of RSVP. When Egress of LSP received the PATH message that includes the LSP

_TUNNEL_INTERFACE_ID, it assigns the Link Identifier of remote side FA and notifies it with a responding RESV message. Link Local/Remote Identifiers in FA is advertised the value assigned by the above described procedure. [RFC3477]

Link Protection Type

This indicates the type of protection. For more details, refer to Chapter-6.

SRLG

Aggregate of SRLG in TE-Link constructing the FA-LSP is advertised as SRLG of FA. [LSP-HIER]
As for the SRLG in case the duplicated LSP is advertised with FA, refer to Chapter-6.

Interface Switching Capability Descriptor

To the Interface Switching Capability Descriptor of FA, a value of the first link of FA-LSP path. When the Interface Switching Capability Descriptor is TDM, the maximum bandwidth among the Minimum LSP Bandwidths in FA-LSP path is set to FA.

Sharable Bandwidth

As the sharable bandwidth of FA, the sharable bandwidth of FA-LSP is set.

【Note】 About the attribute of FA

Protection type

- Local/remote link identifier
- Interface switching capability
- Maximum bandwidth
- Maximum reservable bandwidth：Overall assignment
- Maximum LSP bandwidth （Residual bandwidth)：Overall residue
- Maximum sharable bandwidth：Assignment of backup
- Residual sharable bandwidth：Residue of backup
- Maximum extra bandwidth：Assignment of Extra
- Residual extra bandwidth：Residue of Extra

4.2 Utilizing method of FA for fault recovery

About the Sub-TLV of each OSPF in FA used for fault recovery and duplicated LSP, refer to Chapter-6. About the Extra-Traffic, refer to Chapter-7.

4.3 Management of FA and LSP

When duplicating LSP, these LSP are advertised as one FA. FA is taken as non-number. Identifiers relating to FA and LSP are shown in Table.4--2.
To the LSP comprising the same FA, the same Tunnel ID is assigned. LSP is identified by LSP-ID.

Table.4-2 Identifiers relating to FA and LSP

	Attribute	Description
FA	Link local identifier	Link Identifier that the node provides by itself
	Link remote identifier	Link Identifier that the opposing node provides
LSP	IPv4 Tunnel end point address	Specifies the Router ID of Egress side
	Tunnel ID（Session Obj）	This is used to identify the session in RSVP. Here, session is the same as FA.
	Extended Tunnel ID（Session Obj）	This is used to identify the session in RSVP. Here, session is the same as FA. In order to restrict session identification between Ingress and Egress, Router ID of Ingress may be specified. Usually, this is all “0”.
	LSP ID（Sender Template Obj or Filter Spec Obj）	This is used for identify the LSP that comprises the session (FA). A unique value is assigned at the originating node.
	Interface ID (LSP TUNNEL INTERFACE ID Obj)	This is used for exchanging the Link Identifier that the node itself and the opposing node provide. The one that the Ingress node provided is included in Path message, and the one that the Egress node provided is included in Resv message. If this is exchanged by RSVP, it is advertised as a Link Local/Remote Identifier when FA is advertised in OSPF.

Link Identifier

FA is taken as non-number. When LSP is bi-directional, each originating-side node advertises respective FA. Each node provides respective Link Identifier of FA. As for the FA that the node itself becomes originating side, a Local Link Identifier is provided. As for the FA that the node itself becomes receiving-side, a Remote Link Identifier is provided.

Fig.4-1 shows an example that a duplicated LSP was set up between node-A and node-D. The node-D advertises FA directed to the node-A. Link Identifier of node-A side is provided by node-A (In Fig.4-5, this is expressed by X). Link Identifier of node-D side is provided by node-D (in Fig.4-5, this is expressed by Y). Node-A and node-D execute exchanging of Link Identifiers using RSVP-TE [RFC3473,RFC3477]. Node-A loads LSP _TUNNEL_INTERFACE_ID object to Path message, set the Router-ID of the node itself to the RouterID field and the provided Link Identifier (in this case, X) to the InterfaceID field, respectively, and transmits them to node-D. When Node-D received the Path message, it recognizes that the Link Identifier that was provided by node-A is X. Then, node-D loads LSP _TUNNEL_INTERFACE_ID object to Resv message, set the Router ID of the node itself to the Router ID field and the provided Link Identifier (in this case, Y) to the Interface ID, respectively, and transmits them to node-A. When node-A received the Resv message, it recognizes that the Link Identifier provided by node-D is Y. Like this way, node-A and node-D exchange their respective Link Identifier provided. After this, node-A and node-D advertise respective FA. Node-A advertises FA directed from node-A to node-D, and node-D advertises FA directed from node-D to node-A.

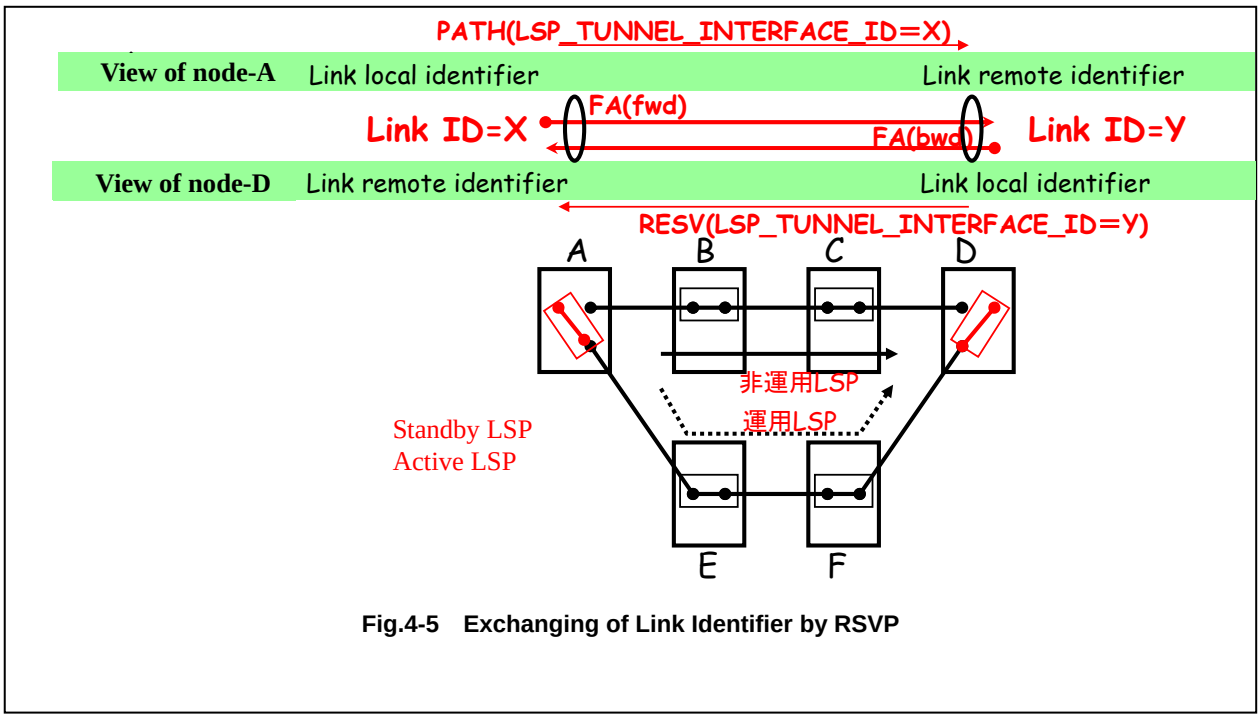


Fig.4-5 Exchanging of Link Identifier by RSVP

Protection type

Protection type applied when LSP is advertised as FA is determined by a Protection Type (LSP Flags of PROTECTION Obj) of the LSP [E2E]. Table.4-3 shows the relationship between LSP Flags of PROTECTION Obj and Protection type of FA.

Table.4-3 Relationship between LSP of PROTECTION Obj and Protection type of FA

LSP （Flags of PROTECTION Obj）	FA（Protection type）
0x00 Unprotected	0x02 Unprotected
0x01 (Full) Re-routing	NA
0x02 1:1 Re-routing without Extra-Traffic	0x04 Shared
TBD 1:1 Re-routing with Extra-Traffic	0x04 Shared
0x04 1:1 Protection with Extra-Traffic	0x08 Dedicated 1:1
0x08 1+1 Unidirectional Protection	0x10 Dedicated 1+1
0x10 1+1 Bi-directional Protection	0x10 Dedicated 1+1

Dedicated 1+1

Advertization of FA in case of “Dedicated 1+1” is shown in Fig.4-6.

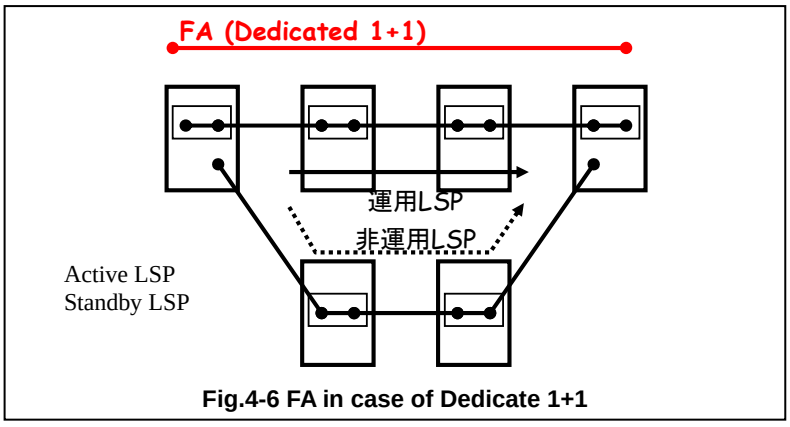


Fig.4-6 FA in case of Dedicate 1+1

Dedicated 1:1

Advertization of FA in case of “Dedicated 1:1” is shown in Fig.4-7.

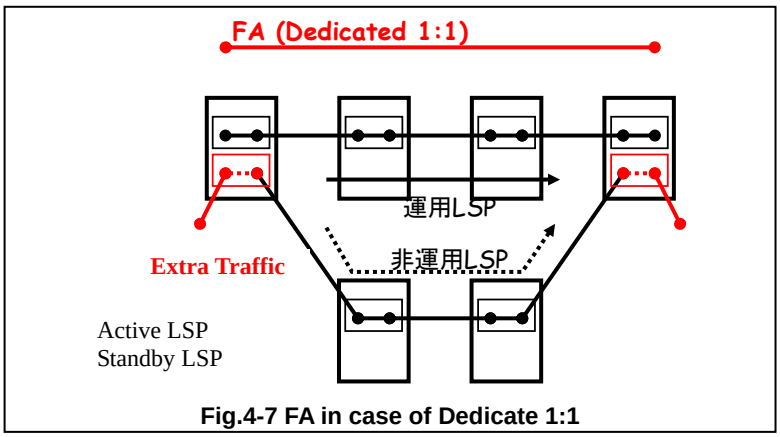


Fig.4-7 FA in case of Dedicate 1:1

Shared

SAdvertization of FA in case of “Shared” is shown in Fig.4-8.

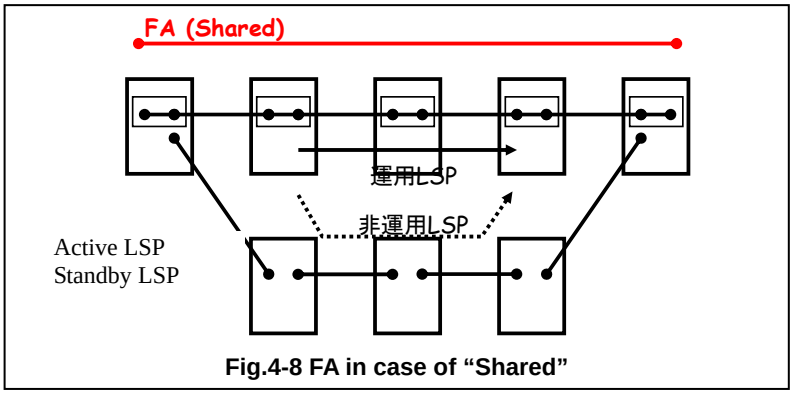


Fig.4-8 FA in case of “Shared”

Unprotected

Advertization of FA in case of “Unprotected” is shown in Fig.4-9.

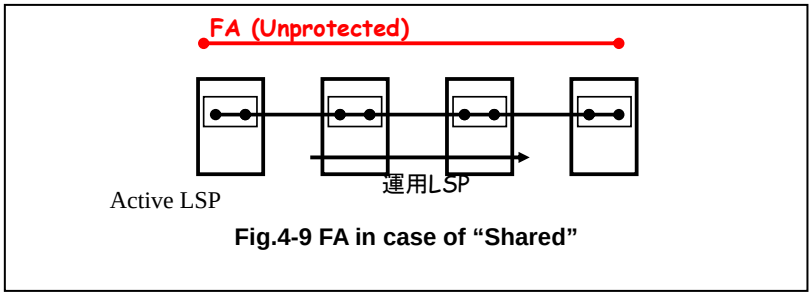


Fig.4-9 FA in case of “Shared”

SRLG

SRLG of LSP consists of TE links that comprise the LSP.
When SRLG advertises a duplicated LSP as FA, it is assigned independently of LSP that is constructing FA.
When SRLG advertises a single LSP as FA, SRLG of the LSP is used.
Fig.4-10 is an example when a duplicated LSP is set up between node-A and node-D and it is advertised as FA. SRLG of 0-system LSP is {1,2,3} that is an aggregate of SRLG of TE links that the traffic pass through. On the other hand, SRLG of 1-system LSP is {4,5,6} that is an aggregate of SRLG of TE links that the traffic pass through. When these are bounded together and advertised as FA, SRLG of FA is newly provided and {100} is advertised.

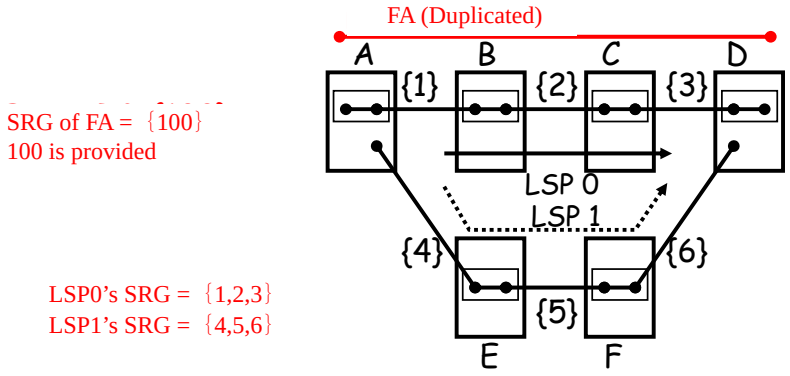


Fig.4-10 SRLG

[RFC3477] Signaling Unnumbered Links in Resource ReSerVation Protocol -Traffic Engineering (RSVP-TE), 1/03
[RFC3473] Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions, RFC3473, 1/03
[E2E] RSVP-TE Extensions in support of End-to-End GMPLS-based Recovery
<draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03.txt>, 5/03

5. Fault Notification

5.1 Fault notification during restoration

(1) Steps to recover from fault

This section describes about how to recover from fault when fault happened to occur. Fig.5-1 shows the steps to recover from fault.

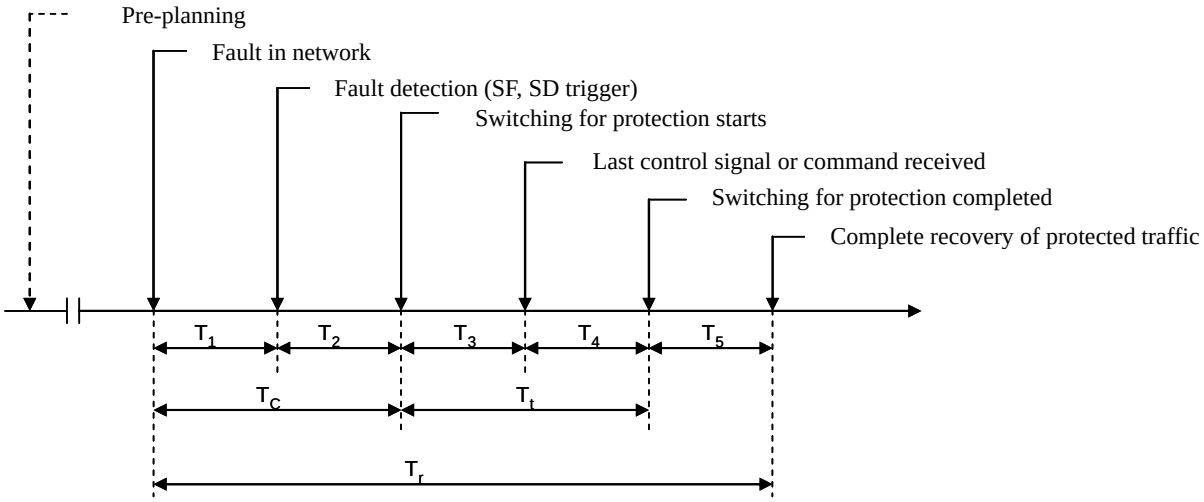


Fig.5-1 Steps to recover from fault

Details of each step are described below.

【Pre-planning】

Management system pre-supposes a certain fault and designs a spare LSP route.

【T1: Fault detection time】

Time required for detecting a fault. An adjacent node to the location of fault detects a fault (SD, SF) from loss of input signal in light channel or from bit-error-rate exceeding a threshold value.

【T2: Hold-off time)】

Latency from detection of fault to start of fault recovery process. In case that the lower layer in data plane executes fault recovery, T2 has a role as a timer for judging whether fault recovery should be done in the upper layer or not.

【T3: Protection switching operation time)】

Time for transferring and processing the control signal that is required to execute protection switching.

【T4: Protection switching transfer time】

Time required for switching all the protection switches.

【T5: Recovery time)】

Time required until the recovered main stream signal becomes the same communicating state as the one before the fault occurred. In order to validate the connectivity of protection path that was set up at each node, a pre-defined byte for path tracing of light channel is utilized.

【Tc: Fault confirmation time】

Time from occurring of network fault to completion of detecting SF or SD: Tc = T1 + T2

【Tt: Transfer time】

Time from start to completion of switching of the protection switch. Tt = T3 + T4

【Tr: Protected traffic restoration time】

Time from occurring of fault to complete recovery of traffic. Tr = T1 + T2 + T3 + T4 + T5 = Tc + Tt + T5

In case of dynamic restoration, time to calculate the route for bypassing the fault is required. This time to do so differs depending on how to execute the dynamic restoration. In case of End-to-end recovery, the time required is in between T3 and T4, because routing is executed at the path-end node. On the other hand, in case that the fault is recovered at the fault-end, the calculation time required is in between T2 and T3 because the route is calculated at the fault end. In ITU-T recommendation G.841, a protection switching time in SDH network has been specified under the certain conditions such as extra-traffic is not flowing, etc. and is specified as 50ms. Also in Photonic IP network, when supposing that it replaces the existing SONET/SDH network, it may be required to assure the same switching time. That is, if we assume that the time required for fault recovery is Tr (for example, 50ms), the following condition must be satisfied to assure the Tr.

$$T1 + T2 + T3 + T4 + T5 \leqq Tr$$

(2) Fault detection method

As the method of fault detection using GMPLS control plane, there are two methods as shown below. Although it is fundamental to utilize a fault detection method by hardware, when considering the case that fault detection doesn't function by damage of hardware or control node, fault detection method that detects a fault on GMPLS control plane is inevitable. As for the fault detection method by hardware, refer to section 2.2.

Method-1: A method utilizing a RSVP Hello message

It is possible to verify the reach ability to adjacent node by using the RSVP Hello message. RSVP Hello is especially effective to detect a fault between adjacent nodes. Node transmits Hello Request to adjacent node with a certain time interval. If Hello is running in the adjacent node, the node returns Hello Ack. If the node didn't receive Ack message four times consecutively (Cisco spec.), or if the node received a wrong message, the node judges that the adjacent node failed.

There are two setting parameters as the followings.

- Hello interval (ip rsvp signaling hello refresh interval command) (default: 200ms (Cisco spec.))
- Number of transmission of Ack signal by which the node judges that the adjacent node failed (ip rsvp signaling hello refresh misses command)(default: 4)

Method-2: A method utilizing a LMP Hello message [LMP]

LMP (Light Management Protocol) has LMP Hello and a function that is possible to rapidly detect a fault. If a control channel has been once established, Hello protocol is used to maintain connectivity with adjacent node. Parameters are exchanged through Config Message. When a node transmitted or received the ConfigAck message, it starts Hello message communication. Hello message is transmitted with a time interval of HelloInterval. If the node did not receive Hello message for a period of HelloDeadInterval, it judges that the control channel failed.

Relating to fault detection, there are two setting parameters as the followings:

- HelloInterval (default: 150ms)
- HelloDeadInterval (default: 500ms, necessary to take a value greater than 3 times of HelloInterval)

(3) Identification of fault location

Identification of fault location is based on localization by hardware. But, in case of all-optical network, there is a possibility that multiple alarms by loss of light may be generated on the optical path. In LMP, in order to identify such a fault location, it is designed so as to be possible to localize the fault location by using a ChannelStatus message. ChannelStatus message can be used to identify the location of single data channel fault, multiple data channel fault and the all TE-link fault.

A downstream node that detected a data link fault sends a ChannelStatus message to the adjacent upstream node bundling all the notifications of fault data links together. A node that received the ChannelStatus message must send a ChannelStatusAck message that indicates the receipt of the ChannelStatus message to downstream node. (MUST) The upstream node should make a correlation of faults in order to confirm if that fault is detected also in the corresponding LSP(s) (including the ingress side). For example, when a fault was not detected on input channel of the upstream node or internally, the upstream node will be able to identify the fault location. If a fault was once localized, the upstream node should send a ChannelStatus message that indicates the channel is normal or failed to the downstream node. (SHOULD) If the ChannelStatus message was not received at the downstream node, it should send a ChannelStatusRequest message because there may exist a fault in this channel. If a fault was once detected, a signaling protocol can be used to commence the span or path-protection/restoration procedure.

When all the datalinks on TE-link failed, the upstream node may notify the fault of TE-link without notifying the fault of respective data link in failed TE-link. (MAY) When doing this, it is possible to notify the fault by sending a ChannelStatus message specifying the TE-link without including any interface IDs within the CHANNEL_STATUS object.

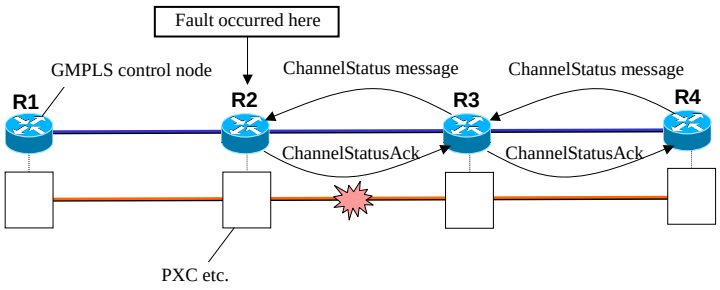


Fig.5-2 Identification of fault location (localization)

(4) Method of fault notification

As a method of fault notification using a GMPLS control plane, there are four methods as described below. Although a method by hardware notification is available, when considering the multi-domain and non-OTN network, it is indispensable to adopt a method that notifies a fault on GMPLS control plane. As for a method of fault hardware by hardware, refer to section 5.2.3.

In case that data plane provides an automated Protection Switching ability (for example, please refer to ITU-T G.841 Recommendation), Notification (N) bit is defined to Protection Object. This is done for distinguishing it from Protection signaling through control plane or data plane.

Method-1: A method using a RSVP-TE Path Error message

This is a method that uses a Path Error message of RSVP-TE for fault notification. In MPLS Fast Reroute (FRR), a Path Error is used for fault notification. Also in GMPLS, similar method can be available (see Fig.5-3) In this case, when a fault occurred, usually Resv Tear message is flown to both the upstream and downstream directions, while in MPLS FRR, Path Error message is flown to upstream direction and the path-end in downstream (R4) waits for refresh message. At the path-end in upstream (R1), it starts recovery operation triggered by Path Error message. (In MPLS FRR, it is specified so that path-end R1 may do optimization of working LSP triggered by

Path Error message at the same time when RLR starts operation.)

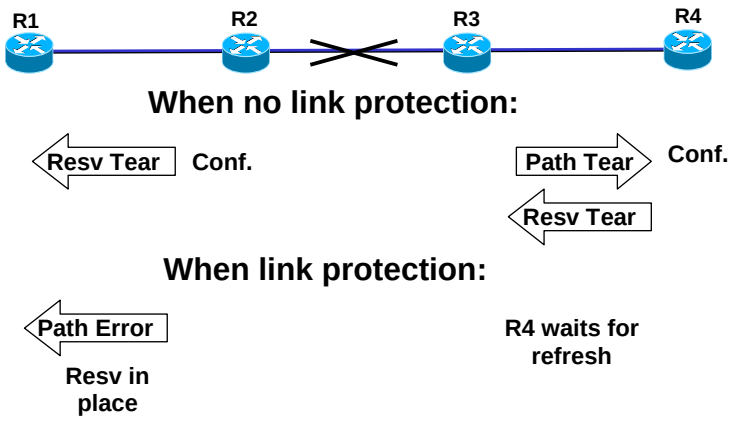


Fig.5-3 Fault notification using a Path Error message

Method-2: A method using a RSVP-TE Notify message

In GMPLS signaling protocol, a function called as “Notify Message” is supported. Notify Message is a function that is possible to directly notify a status of LSP to the node on working LSP that is not an adjacent node. By expanding this Notify Message, it is possible to direct the node that relays a protection path to switch. If there is no need to notify to the relay node, it is possible to use a standard Notify message. As an advantage of this method, there is a fact that it is easy to implement this method because a signaling protocol that is managing label information executes fault notification. While, as for a weak point, there is a fact that it is not suitable for rapid fault recovery because multiple notifications are required (for example, for each wavelength) even for a single fault resulted from break-off of fiber cable, etc.

Method-3: A method using an OSPF flooding (unique method)

In case of OSPF/IS-IS, control information is notified to overall network using a flooding. It becomes possible to notify fault information by setting this flooding framework of layer-3 as a base. If we say it more clearly, this method uses Opaque LSA of OSPF and writes clearly fault information in Opaque relevant fault.

As an advantage of this method, there is a fact that it is possible to notify many fault information resulted from single fault with single flooding operation. While, as for a weak point, there is a fact that, since flooding process in current OSPF is very heavy, it is highly possible that a requirement of 50ms for protection time specified in SONET/SDH network can’t be satisfied.

Method-4: A method using a flooding dedicated for fault notification (unique method)

Generally speaking, flooding is a promising method because it reduces number of fault notification messages. Therefore, by executing flooding at the layer close to data plane, it becomes possible to notify rapidly the fault information.

Brief description of this method is shown below.

① It gathers information of relationship between adjacent nodes the time of startup.
② It securely establishes a connectivity using a Hello message, etc.
③ It specifies two types of message; FaultNotify and FaultNotifyAck.
④ It stores information on fault location and a sequence number in FaultNotify message.

Although it is best to realize these functions by specifying a new protocol for fault notification, it is also possible to implement them into such protocols as LMP (Link Management Protocol) that is under standardized in sub IP area of IETF. This method is a method in which the conventional flooding functions were specialized to fault notification, and even if there is a subject of standardization, it is thought to be a realistic method as the one that rapidly notifies a fault using a control plane.

Fig.5-4 shows a concept of fault notification method using a flooding. A fault occurred in data plane is detected at the node in the data plane and notified to control plane. A controller in the data plane notifies the fault information to the network using a flooding. A controller in the control plane that received the fault notification message makes the settings of node of the data plane if it judged this was a fault information. Especially in edge-node, it switches a path from working system to recovery system.

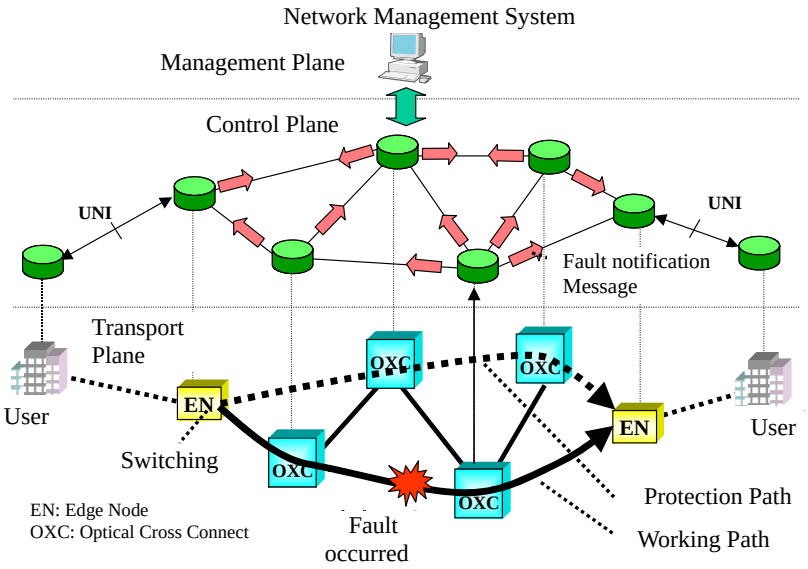


Fig.5-4 Fault notification method using a flooding

Table.5-1 shows the applicability of each fault notification method to the place to be notified (Ingress node, Egress node, and repeater node)

Table.5-1 Summary of fault notification methods

	Notify to Ingress node	Notify to Egress node	Notify to repeater node	Comment
Method-1: Path Error	○	Possible if message was specified	×	RSVP is used for activating repeater node.
Method-2: Notify	○	○	Message expansion is required	RSVP is used for activating repeater node.
Method-3: OSPF expansion	○	○	○	Unique method
Method-4: Dedicated flooding	○	○	○	Unique method

As above described, as for fault notification using a control plane, although there are various possible methods are possible, in this IA, it is based on the fault notification method using a GMPLS Notify message that is currently being standardized in IETF. In the followings, details of Notify message will be described.

(5) Fault notification procedure using a Notify message

(5-1) Operation procedure

In e2e recovery, there are two types of fault notification, the one is a case that merely to notify a fault and another is a case that executes fault notification and executes switching of protection switch. When N-bit of Protection object has been set to “0”, protection operation is triggered, and when N-bit has been set to “1”, only notification of CP is executed.

○ In case that only fault notification is executed (in case of 1+1 Unidirectional Protection)

In case of 1:N Protection with Extra-Traffic, 1+1 Unidirectional Protection, or 1+1 Bi-directional Protection, N-bit of Protection object can be set to “1”, and in other cases than these, N-bit must be set to “0”.

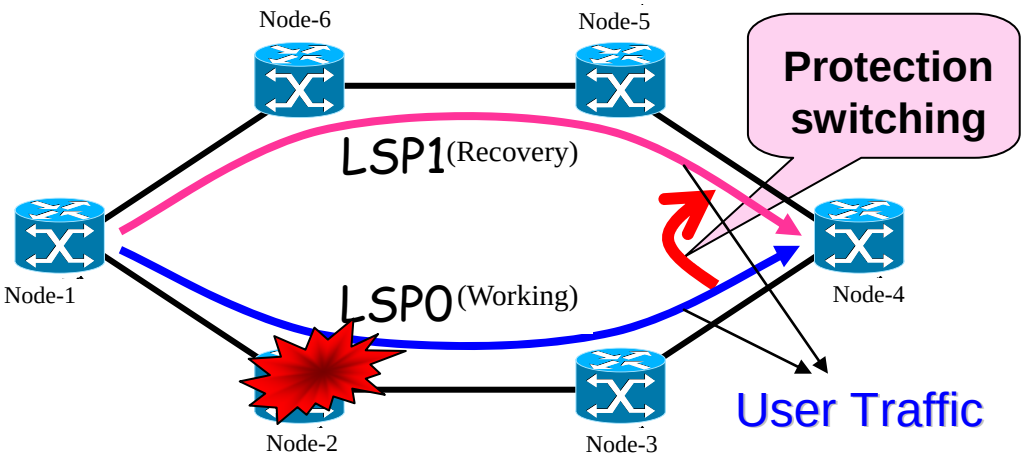
When N-bit is “1”, protection switch is not triggered by fault notification.

In Fig.5-5, in case that a working LSP (LSP0) and a recovery LSP (LSP1) of 1+1 Unidirectional Protection were setup between node-1 and node-4 and a fault happened to occur at node-2, a fault in LSP0 is detected at the node-4 of Egress side and the traffic is switched to LSP1 side that is currently transferring normal traffic.

In case of Unidirectional, switching is executed independently in each direction like this way.

When PathErr message is generated to failed LSP0 after the protection switch was switched, Path_State_Removed flag of RROR_SPEC object is not set up. (RECOMMEND)

When switching of protection switch was completed and PathErr message was received, LSP1 sets up O-bit to “1” to indicate that Protecting LSP is transferring normal traffic. (SHOULD) In LSP0, it is also possible to set up A-bit of ADMIN_STATUS object.



1+1 Unidirectional Protection

Fig.5-5 Concept of 1+1 Unidirectional Protection

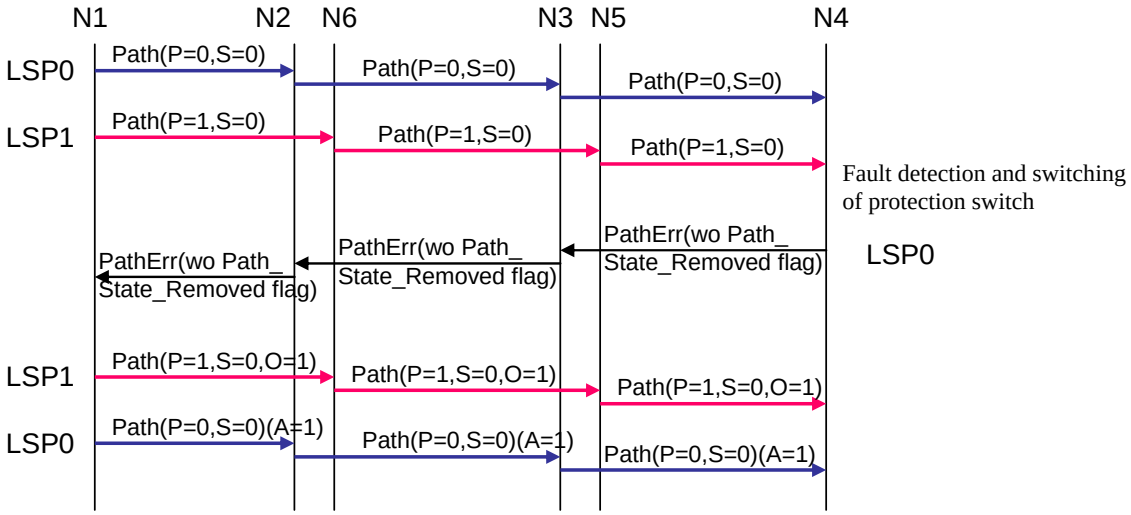


Fig.5-6 Sequence chart of 1+1 Unidirectional Protection

○ In case that protection switch controlling is involved (in case of 1+1 Bidirectional Protection)

1. When Egress node detected a fault of working LSP (LSP0) (or degradation of signal performance of working LSP) or received a Notify message that included a SESSION in <upstream/downstream session list> and a new error code/sub-code "Notify Error/LSP Locally Failed" in (IF_ID)_ERROR_SPEC object, it must commence receiving in recovery LSP (LSP1) (switching of the protection switch). (MUST) That is, in case that intermediate node that detected a fault notifies a fault, it must generate a Notify message that includes a SESSION in <upstream/downstream session list> and a new error code/sub-code "Notify Error/LSP Locally Failed" in (IF_ID)_ERROR_SPEC object.
   A node that receives Notify message has to use <sender descriptor> or <flow descriptor> in Notify message in order to solve a conflict of Notify messages after recognizing the failed LSP, because there is a case that the node may receive multiple Notify messages for the same fault.
   A node that executed switching of protection switch must have a new error code/sub-code "Notify Error/LSP Failed"(Switchover Request) showing that the working LSP is failed and send a Notify message including a MESSAGE_ID object to other end-point node with a reliable method. This Notify message must set an ACK_Desired flag in MESSAGE_ID object and send it to receiver in order to request the receiver for sending the acknowledgement of receiving the message.
   This Notify message (Switchover request) is possible to send an identifier of fault link and relating information by using a IF_ID ERROR_SPEC object. (MAY) In this case, ERROT_SPEC object of Notify message can be replaced by IF_ID ERROR_SPEC object, or can be sent with PathErr/ResvErr message.

2. When a node received a Notify message (Switchover request), the node must commence receiving from recovery LSP (LSP1). (MUST)
   This end-node must send a MESSAGE_ID object and a Notify message (Switchover response) including a MESSAGE_ID_ACK object

for confirming receipt of Notify message (Switchover Request) to other end-node with a reliable method.
The Notify message (Switchover response) may send an identifier of fault link and relating information by using a IF_ID ERROR_SPEC object. (MAY)
The end-node must send an Ack message to other end-node to confirm the receipt of Notify message when it received the Notify message. (MUST)

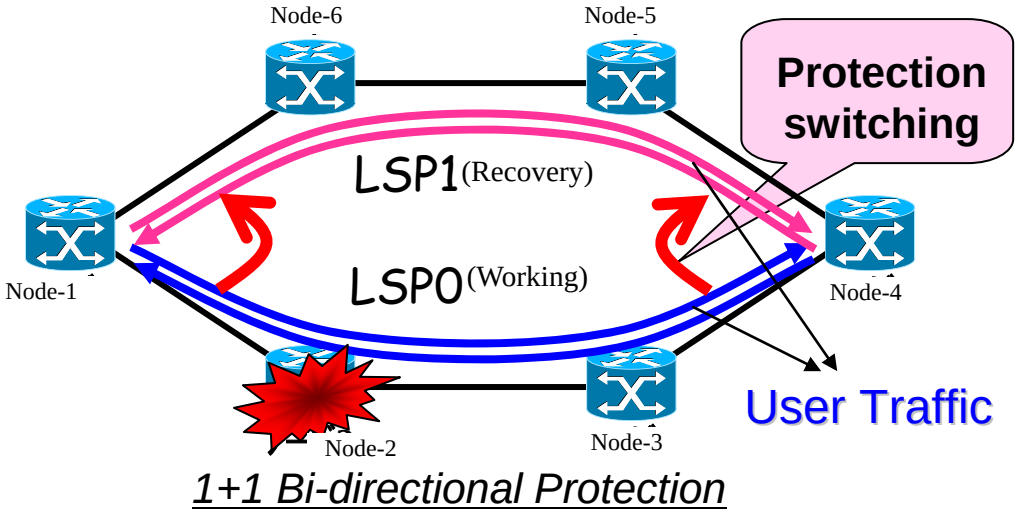


Fig.5-7 Concept of 1+1 Bi-directional Protection

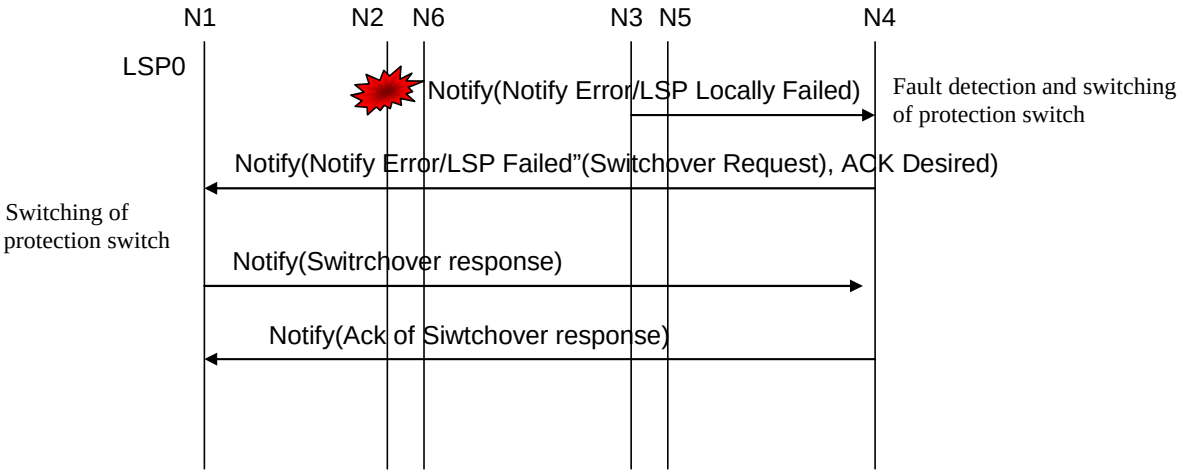


Fig.5-8 Concept of 1+1 Bi-directional Protection

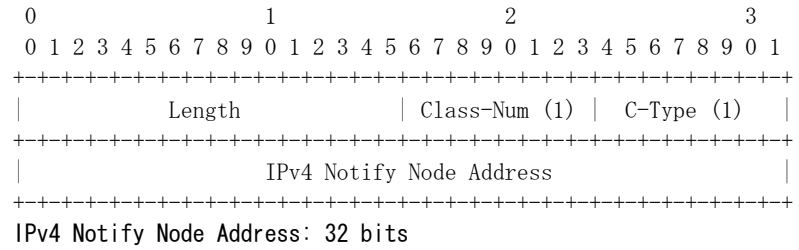
In case that an intermediate node has GMPLS RSVP-TE signaling ability, each node adjacent to the fault may directly generate a Notify message to either the start-point or the end-point of LSP or to both of them. (MAY) Therefore, a node that is terminating these LSPs (may detect a fault from data plane) is expected to supply a mechanism that correctly associate the message to fault or a disposal mechanism of multiple Notify messages corresponding to the same session in order to avoid the repetition of above procedures. In addition, each node does not set a Path_State_Removed flag of ERROR_SPEC object for fault LSP when generating a PathErr message. (RECOMMENDED)

After completion of switching of the protection switch (Step-2) and receiving the PathErr message, the recovery LSP (LSP1) must set the O-bit and execute signaling so as not to miss the direction from which LSP it is receiving the signal. (SHOULD) Previous working LSP0 may set A-bit of ADMIN_STATUS object and execute signaling. (MAY)

In case that N-bit was set, Switchover request/response in end-to-end is used only in control plane and doesn't trigger an action to data plane.

(5-2) Object utilized [RFC3473]
(5-2-1) Notify Request Objects
Notify Request object is transferred in Path or Resv messages. Notify Request class number is 195 (a form of 11bbbbbb).
Format of Notify Request's as follows:

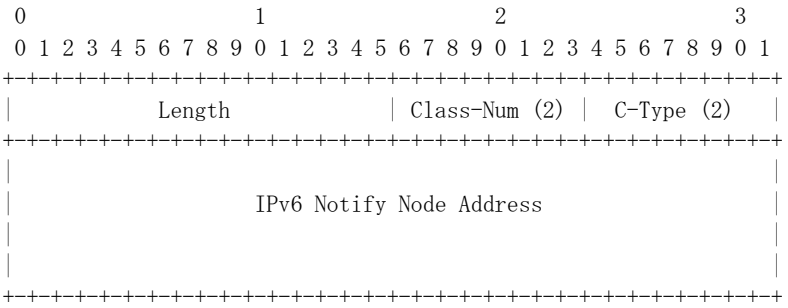
IPv4 Notify Request Object



IPv4 Notify Node Address: 32 bits

Above address indicates the IP address of node that should be notified when error message is generated.

IPv6 Notify Request Object



IPv6 Notify Node Address: 16 bytes

Above address indicates the IP address of node that should be notified when error message is generated.

If a message contains multiple Notify_Request objects, only the first object has a meaning. The rest of Notify_Request objects may be neglected (MAY) or shouldn't be transferred (SHOULD NOT).
Notify_Request objects may be inserted in Path or Resv messages in order to indicate the address of node to which LSP fault should be notified. As mentioned previously, notification may be requested from both directions of upstream and downstream. Notification to upstream direction is indicated by involving the Notify Request object in corresponding Path message. Notification to downstream direction is indicated by involving the Notify Request object in corresponding Resv message.
A node that received a message including Notify Request object should store the Notify_request address in a corresponding state block. (SHOULD) If the node is a passing node, it should involve the Request object in output Path message or Resv message. (SHOULD) Output Notify Node address may be updated by local policy. (MAY)
Please notice that, if Notify_Request object is included, it isn't guaranteed that Notify message is generated

(5-2-2) Notify message

Notify message provides a mechanism for notifying the event relating to LSP to the nodes that are not neighboring to it. Notify message is generally generated only after the Notify_Request message was received. Notify message is different from error messages (that is, PathErr and ResvErr messages) that have already been defined in such points that it is possible to notify to the nodes other than the neighboring nodes in upstream side or downstream side and that it is a conventional notification mechanism. Notify message doesn't replace the existing error messages. Notify message is sent by either methods of the followings: (a) Like the ResvConf process in [RFC2205], Notify message is sent to non-target node by just forwarding the Notify message addressed to the target node, (b) Notify message is sent by being encapsulated with a new IP header having the same destination address as the target IP address. Not depending on transmission mechanism, a node that received a Notify message that is not addressed to itself forwards the message to the target node without modifying it.
In order to support reliable delivery of Notify message, Ack message [RFC2961] for confirming the receipt of Notify message is utilized. For details relating to reliable delivery of RSVP message, refer to [RFC2961].

Necessary information
Notify message is a conventional notification message. IP destination address is set to IP address of intended receiving node. Notify message is sent without a router alert option. One Notify message may include a notification sent together with the respective listed sessions of both of the upstream and downstream directions.
Notify message is a message type 21. Format of Notify message is as follows:

<Notify message>::= <Common Header> [<INTEGRITY>]
[[<MESSAGE_ID_ACK> | <MESSAGE_ID_NACK>] ...]
[<MESSAGE_ID>]

<ERROR_SPEC> <notify session list>

<notify session list>::= [<notify session list>]
<upstream notify session> |
<downstream notify session>

<upstream notify session>::= <SESSION> [<ADMIN_STATUS>]
[<POLICY_DATA>...]
<sender descriptor>

<downstream notify session> ::= <SESSION> [<POLICY_DATA>...]
<flow descriptor list>

ERROR_SPEC object indicates that there is an error and includes an IP address of either the node that was detected an error or the failed link. About definition of ERROR_SPEC, refer to [RFC2205]. MESSAGE_ID and the relating object are defined in [RFC2961] and utilized when refresh reduction is supported.

- Procedure

Notify message is most generally generated at the node that detected an error that triggers generation of PathErr or ResvErr messages. If PathErr message has been generated and Notification_Request object has been received in Path message, Notify message addressed to the node that has been recorded in it should be generated. (SHOULD) If ResvErr message has been generated and a Notification_Request object has been received in Resv message, Notify message addressed to the node that has been recorded in it should be generated. (SHOULD) As previously described, one error may generate Notify messages addressed to both of the upstream and the downstream nodes. Notify message must not be generated when an appropriate Notify_Request object has not been received. (MUST NOT)
When a node generates Notify message, it tries to bind the notifications addressed to the same Notify node and to share the same ERROR_SPEC in one Notify message. A measure to determine which information can be bound depends on implementation. Implementation may use events, timer-based, or other approaches. If it takes timer-based approach, implementation should make enable users to set up what range of notifications are bound. (SHOULD) When it takes timer-based approach, a default "notification interval" value of 1ms should be used. (SHOULD) Notify message should be delivered using a reliable message delivery mechanism defined in [RFC2961]. (SHOULD) When a node received a Notify message, the Notify node should send a corresponding Ack message. (SHOULD)

5.2 Fault notification during protection

Relating to fault notification, in case that data plane provides an automated Protection Switching capability (for example, refer to ITU-T G.841 Recommendation), Notification (N) bit is defined in Protection Object. This is executed to distinguish the fault notification from protection signaling via control plane or data plane.
In this section, we describe about fault notification in data plane. Fault notification in control plane is available also to protection. About fault notification in data plane, refer to section 5.1.

Fault notification method in optical layer (OTN): [G.709]
Fig.5-5 shows a layer structure of OTN (Optical Transport Network)

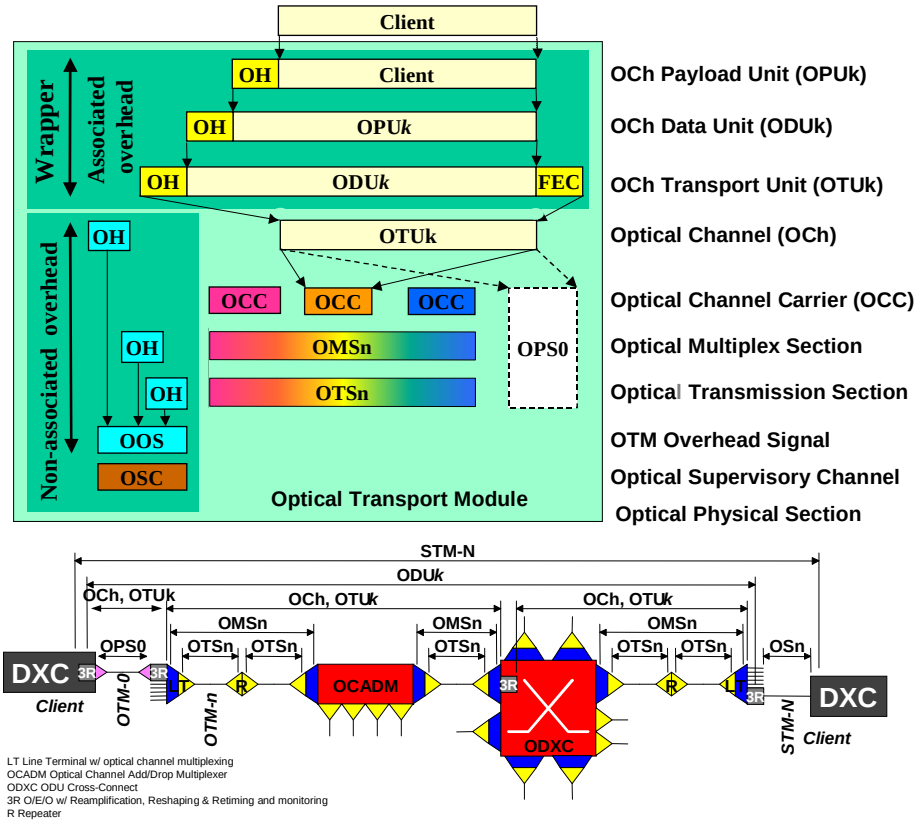


Fig.5-5 Layer structure of OTN

Fig.5-6 shows a structure of OTN maintenance signal. In principle, maintenance signal is transferred from lower layer to upper layer. In OTN, it is called as FDI (Forward Defect Indication)/AIS(Alarm Indication Signal) that fault is notified in the same direction as the main signal. As we know from this figure, alarm information directed to OMS-OCh is called as FDI, and alarm information directed to the upper layer than OMS layer is called as AIS. FDI is divided more into FDI-P and FDI-O. OMSn-FDI-P is a method that notifies a fault using a status of Payload, and OMSn-FDI-O is a method that notifies a fault using Overhead of OMS. PMI(Payload Missing Indication) is a signal sent to downstream direction in order to indicate that no signal is included in OCCp (OCC with full functionality Payload) at the source locating in upstream of OMS. That is, downstream nodes can recognize the fault in upstream nodes by looking at this PMI signal. AIS is a method of fault notification in layer that processes signals electrically, that is, in upper layers than OTUk level. In the upper layer than ODUk level, maintenance signal is interworked to each layer (SONET/SDH, ATM, MPLS, etc.).

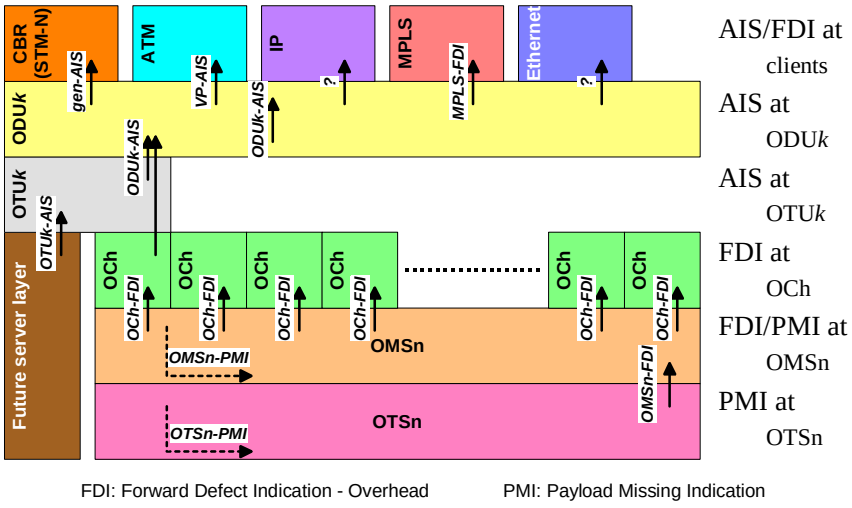


Fig.5-6 Structure of maintenance signal in OTN (Forward Defect Indication)

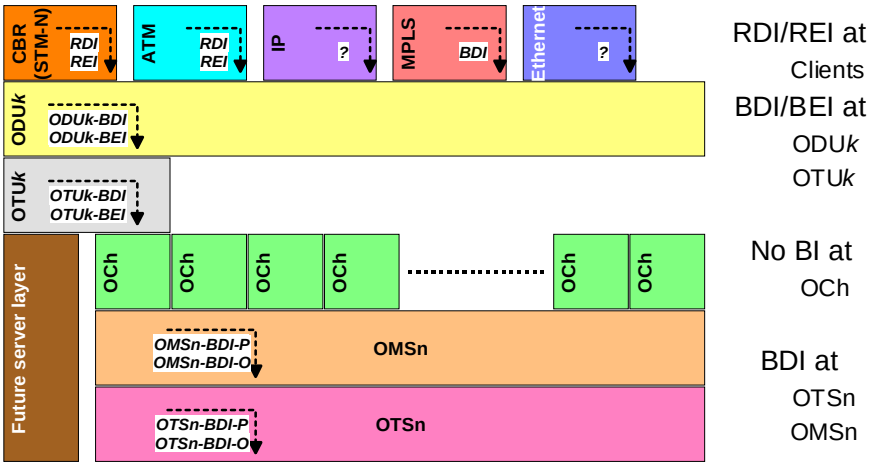


Fig.5-7 Structure of maintenance signal in OTN (Backward Defect Indication)

When recovery operation in optical layers is executed by using this maintenance signal, a signal of Och layer is used as a trigger. If recovery operation is limited in optical layers (lower than ODUk level), when FID is executed at the time of fault occurrence, forward defect indication in OCh layer or fault notification as ODUk-AIS at the end-point of ODUk is done, and backward notification to Ingress node at the ODKk end-point of far end is done. This is because that ODUk is possible to flow optical signal end-to-end that is terminating the optical signal.

(2) Summary of fault detection and notification by hardware

About SONET/SDH refer to ITU-T recommendation G.783. About MPLS, refer to ITU-T recommendation Y.1711/Y.1720. Table.5-2 is a summary of above described methods.

Table.5-2 Fault detection and notification methods by hardware

Overhead layer	Fault detection	Fault notification	Comments
OTS Layer OMS Layer OCh Layer		Refer to previous section	G.709
SONET RS Section	Possible to detect by A1, A2 (frame synchronization) and B1 (error monitoring) in section overhead		
SONET MS Section	Possible to detect by B2 (error monitoring) and LOS (Loss of Signal) in section overhead.	Notify with K1, K2 byte in section overhead. Priority has been assigned to the events of signal failure and signal deterioration.	G.783
SONET Path layer	Possible to detect by J1 (conduction monitoring) in path overhead.	Switching is controlled by K3,K4, and alarm to station is execute by G1 in Path overhead.	
MPLS	MPLS OAM packet is available. Possible to check the normality by Connectivity Verification.	MPLS OAM packet is available. Possible to specify FDI and BDI.	Y.1711 / Y.1720
Ethernet	Discussion on Mac level OAM has been started in MEF or ITU-T	Discussion on Mac level OAM has been started in MEF or ITU-T	

(3) Fault notification to OLS (Optical Line System) using LMP-WDM [RFC4209]

In LMP-WDM (Link Management Protocol-Wavelength Division Multiplexing), it has been made possible to use functions of LMP by expanding conventional LMP's function to make LMP run between node and OLS. (refer to Fig.5-8) OLS means optical line systems, for example, optical transmission system like WDM, SONET/SDH systems, devices equipped with Ethernet port, etc. Like LMP, OLS has four major functions: Control Channel Management, Link Property Correlation, Link Verification, and Fault Management.

In LMP-WDM, as functions of LinkSummary, it is possible to evaluate the bit error rate and to exchange information such as total span length between node and OLS to check if the protection is supported in OLS. And, as functions of Fault Management, it is possible to execute fault detection, localization of fault location, and fault notification in the same way as in LMP. Also, it is possible to notify a fault to devices on span by using a ChannelStatus message.

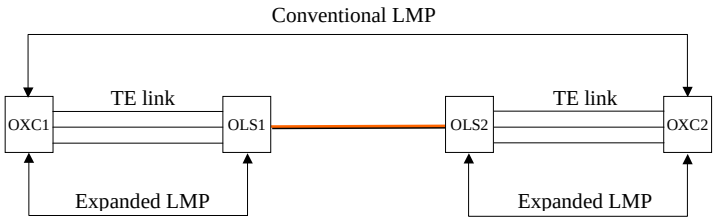


Fig.5-8 Expanded LMP model

References

[G.808.1] ITU-T Draft Recommendation G.808.1, “GENERIC PROTECTION SWITCHING - LINEAR TRAIL AND SUBNETWORK PROTECTION”
[RFC4204] RFC4204, “Link Management Protocol (LMP)”, October 2005.
[RFC3473] RFC3473, “Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions”, January 2003.
[G.709] ITU-T Recommendation G.709
[RFC4209] RFC4209, “Link Management Protocol (LMP) for Dense Wavelength Division Multiplexing (DWDM) Optical Line Systems”, , October 2005.

6. Cooperating operation between Routing and Signaling

6.1 General description about functions

Duplication of LSP is realized by setting up two LSPs between Ingress node and Egress node. Here, we call these two LSPs as 0-system and 1-system.

Path of respective LSP of 0-system and 1-system is calculated at the Ingress node. Path is calculated by using a traffic engineering database (TED). TED consists of FA advertised by routing protocol OSPF.

Paths of 0-system and 1-system calculated at the Ingress node are stored in ERO (Explicit Route Object) of Path message when Ingress node sets up LSP of 0-system and 1-system using a signaling protocol RSVP. The paths in which LSPs of 0-system and 1-system were set up are collected using RRO (Record Route Object).

In an intermediate node that LSP passes through when setting up LSP, reservation of necessary resources is executed. If a resource status of TE link between the intermediate links, the TE link is re-advertised as FA using OSPF. Attributes of TE link at this time is included to FA being advertised this time.

It is possible to advertise the set up LSP itself as FA by OSPF. By advertising the set up LSP as FA, it becomes possible to execute routing of LSPs of multiple tiers.

Calculation of path

Path of respective LSP of 0-system and 1-system is calculated at the Ingress node. Path is calculated by using a traffic engineering database (TED). TED consists of FA advertised by expanded OSPF-TE [GMPLS-OSPF] from GMPLS. Respective path of LSP of 0-system and 1-system can be calculated so as to be SRLG-Disjoint each other, and in case of Shared mesh restoration, it is also possible to calculate taking sharing of reserved bandwidth and recovery performance into consideration. In order to notify which backup LSP the reserved bandwidth is assigned to, PRIMARY PATH ROUTE object is utilized. [E2E] This object is used to transfer path information through which working LSP passes.

6.2 Signaling

Path of respective LSP of 0-system and 1-system calculated at the Ingress node is included to ERO of Path message when Ingress node sets up LSP of 0-system and 1-system using expanded RSVP-TE[RFC3473] signaling procedure from GMPLS [RFC3209]. LSPs of 0-system and 1-system are set up according to the collected path information into ERO. The paths in which LSPs of 0-system and 1-system were set up are collected using RRO. Ingress node involves its IP address into RRO of Path message and sends it. When intermediate node received Path message including RRO, it involves its own IP address into RRO stack and sends the Path message to next hop. In this way, the Path message including RRO is transferred to the Egress node. When the Egress node received the Path message including RRO, it involves RRO in RESV message and sends it to opposite direction of the Path message. At this time, the Egress node involves its own IP address in RRO of RESV message and sends it.

6.3 Resource management

In an intermediate node that LSP passes through when setting up LSP, reservation of necessary resources is executed. Resources of TE link between the intermediate node and the neighboring node are reserved. By this, resource status of TE link is changed. When the resource status of TE link is changed, TE link is re-advertised as FA. Attributes of TE link at this time is included to FA being advertised this time. If FA is advertised at every time when resource status of TE link was changed, amount of packets of OSPF increases. Therefore, frequency of advertising FA should be possible to control. Besides, advertising of FA should be executed based on change ratio of resource status.

6.4 Advertization as FA

It is possible to advertise the setup LSP itself as FA. FA is advertised by expanded OSPF-TE [GMPLS-OSPF] from GMPLS.

[GMPLS-OSPF] OSPF Extensions in Support of Generalized MPLS <draft-ietf-ccamp-ospf-gmpls-extensions-09.txt>, 12/02

[RFC3473] Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions, RFC3473, 1/03

[RFC3209] RSVP-TE: Extensions to RSVP for LSP Tunnels, RFC3209, 12/01

[E2E] RSVP-TE Extensions in support of End-to-End GMPLS-based Recovery <draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03.txt>, 5/03

7. Extra Traffic LSP

In this IA, we will expand the method of shared mesh restoration that has been defined in [E2E] and propose the LSP settings for extra traffic that is applicable to the expanded shared mesh restoration. We call this LSP as “extra LSP”. Aldo, we will reorganize the definition and usage method for extra traffic that is applicable to 1:1 protection.

7.1 Definition of Extra Traffic

In this section, we will define the relationship between extra traffic, LSP and resource management in each protection typ.

7.1.1 Shared mesh restoration

In shared mesh restoration defined in [E2E], extra traffic has not been supported, and the assigned bandwidth to recovery LSP is in unused state that has not transferred any traffic in faultless state. In this IA, in order to use the network resources more effectively, we propose the LSP setting for extra traffic that is applicable to shared mesh restoration. The specification of signaling for this LSP setting has been partly reflected to [E2E].

In extra traffic of this method, it is not required that the route is the same as the one of specific recovery LSP like 1:1 protection. It is required to set up the LSP (in this IA, it is defined as “extra LSP”) that carries extra traffic by using a bandwidth assigned to multiple and optional recovery LSPs. Fig.7-1 shows an example of extra LSP. ① and ② are a pair of working LSP and recovery LSP to protect a certain traffic. ③ and ④ are extra LSPs and are transferring different extra traffic respectively. Extra LSP ③ is using a bandwidth assigned to recovery LSP ② on the link between node-E and node-F. Extra LSP ④ is using a bandwidth assigned to recovery LSP ② on the link between node-F and node-G.

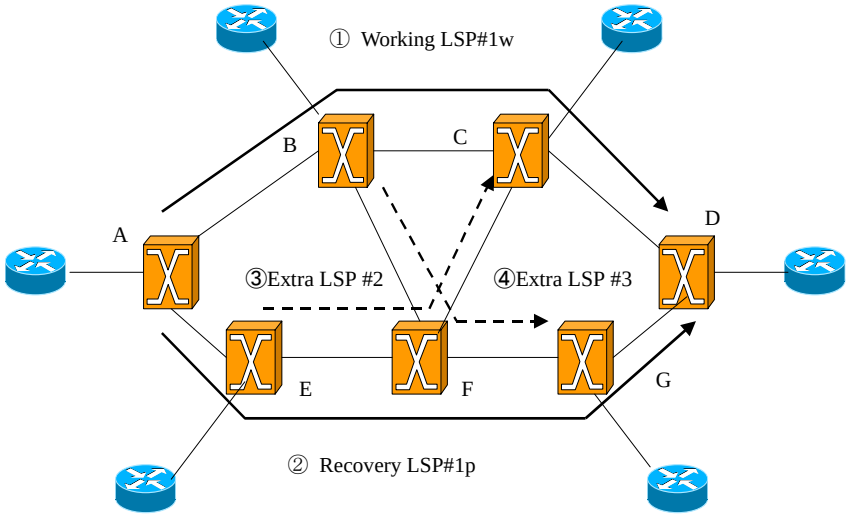


Fig.7-1 Extra LSP in shared mesh restoration

Fig.7-2 shows the content of cross-connect settings in ①(Working LSP #1w), ②(Recovery LSP #1p) and ③(Extra LSP #2) of Fig.7-1.

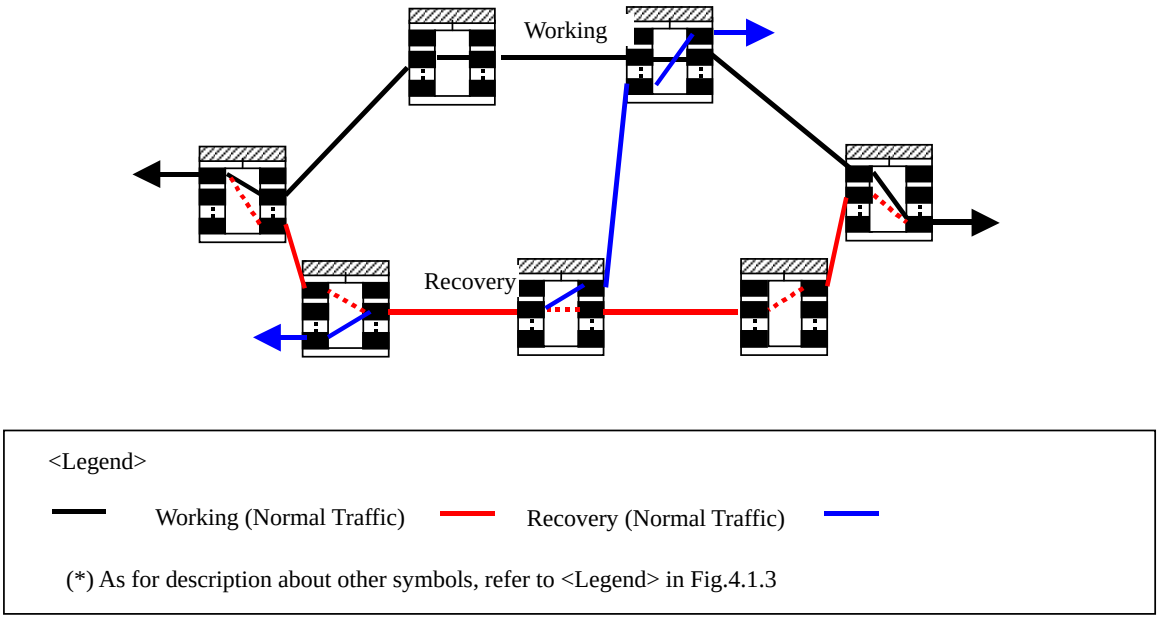


Fig.7-2 Cross-connect settings of extra LSP in shared mesh restoration

Extra LSP uses the bandwidth assigned to arbitrary recovery LSP or unreserved bandwidth. That is, unreserved bandwidth can be assigned to extra LSP, and protecting bandwidth assigned to recovery LSP can be assigned also to extra LSP.
According to [E2E], extra LSP can be realized as a LSP with lower priority than recovery LSP. Handling of protection bandwidth follows also to [E2E]. That is, it is possible to assign protecting bandwidth to the LSP with a holding priority lower than the priority of recovery (holding) LSP.
Bandwidth assigned to recovery LSP of 1:1 protection is not available to use.

7.1.2 1:1 protection with extra traffic

Extra traffic is transferred using the same route and resource (bandwidth, etc.) as those of specific recovery LSP. At this time, signaling and routing for extra traffic are not executed.
Fig.7-3 shows the status of cross-connect of extra traffic, working traffic and recovery traffic. Input port and output port of extra traffic are different from those of normal traffic that is secured by protection. Since signaling for extra traffic is not executed, Ingress node becomes impossible to specify the output port of extra traffic to the Egress node and usage of extra traffic becomes limited. Thus, a method that advertises the recovery LSP as FA-LSP by setting as protection type = extra traffic has been devised and proposed.

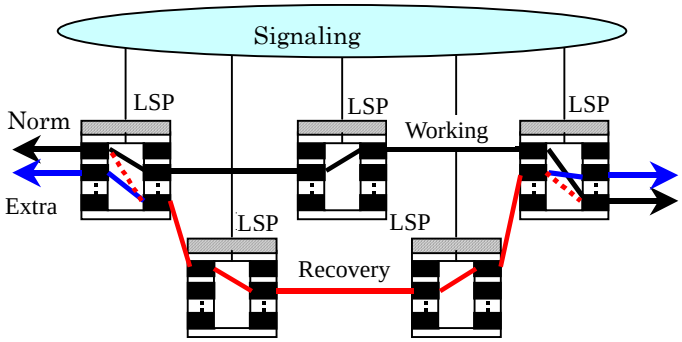


Fig.7-3 Cross-connect settings of extra LSP in 1:1 protection

7.2 Signaling

7.2.1 shared mesh restoration

In shared mesh restoration, LSP is setup to carry the extra traffic. Values of SESSION and SENDER_TEMPLATE are independent on any working LSP and recovery LSP, and the LSP set up is an independent LSP.
According to [E2E], extra LSP is set up as a LSP having a lower priority than recovery LSP. That is, signaling of holding priority of extra LSP is executed with a lower priority (with a greater value) than the holding priority of recovery LSP. Here, a setup priority must be lower than the holding priority of recovery LSP (otherwise, recovery LSP will be disconnected).
At this time, in each processing node of signaling, it is possible to assign a protecting bandwidth that has been assigned to recovery LSP to extra LSP.

7.2.2 1:1 protection with extra traffic

In 1:1 protection with extra traffic, signaling is not executed. At the time point when recovery LSP was setup, both end-nodes of LSP becomes possible to receive/transfer extra traffic from/to recovery LSP.

7.3 Routing

7.3.1 Shared mesh restoration

According to [E2E], protecting bandwidth that has been assigned to recovery LSP can be assigned to the extra LSP as a lower priority bandwidth and be advertised. That is, bandwidth with a lower priority (with a greater value) than holding priority of recovery LSP doesn't decrease even if the recovery LSP was established. When recovery LSP was established, only the bandwidth with the same priority as that holding priority is decreased (other priority values are not changed), and Unreserved bandwidth and Max LSP bandwidth are advertised.

7.3.2 1:1 protection with extra traffic

Advertization of available bandwidth for extra traffic is not executed.
When advertising the recovery LSP as TE link (as FA-LSP), protection type becomes "extra traffic".

7.4 Switching

When switching occurred, it is required to change the cross-connect status with an appropriate timing in order to avoid miss-connection (to inhibit the working traffic from flowing into LSP of extra traffic, and vice versa.).

7.4.1 Shared mesh restoration

When activating the recovery LSP, it is required to appropriately release and setup cross-connect status so as not to flow traffic of extra LSP into recovery LSP by mistake and vice versa. For example, if we setup cross-connect (selection of received traffic) when transferring the Path of activation at the Ingress node of recovery LSP, traffic of extra LSP will flow into the Ingress node. Specification for inhibiting such a mistake is shown in "10. LSP Preemption" of "8.3 Signaling Secondary LSP s" of [E2E].

For example, it is possible to avoid miss-connection by releasing the cross-connect status of extra LSP when transferring Path message, and by setting up the cross-connect status of recovery LSP when transferring Resv message. For doing this, PROTECTION object is included to distinguish that Resv is activated from the case of refreshing.

Since D-plane of LSP that is carrying extra traffic is disconnected by switching, each node providing extra LSP on recovery LSP sends PathErr with a Path State Removed flag and PathTear about extra LSP and releases also the C-plane status. Error Code of PathErr is set to "Policy Control failure/Hard Pre-empted". In addition, in order to suppress alarm notification when extra LSP was disconnected, it is also possible to use ADMIN_STATUS before releasing cross-connect.
Since the LSP carrying extra traffic is disconnected, the service carrying extra traffic is not restored when switch-back occurred. Therefore, it is required to set up again the LSP for extra traffic.

7.4.2 1:1 protection with extra traffic

Switching should be executed following to the procedure written in [E2E].

7.5 Switch-back

7.5.1 Shared mesh restoration

In shared mesh restoration, restoration of extra LSP is not executed when switch-back is executed.

7.5.2 1:1 protection with extra traffic

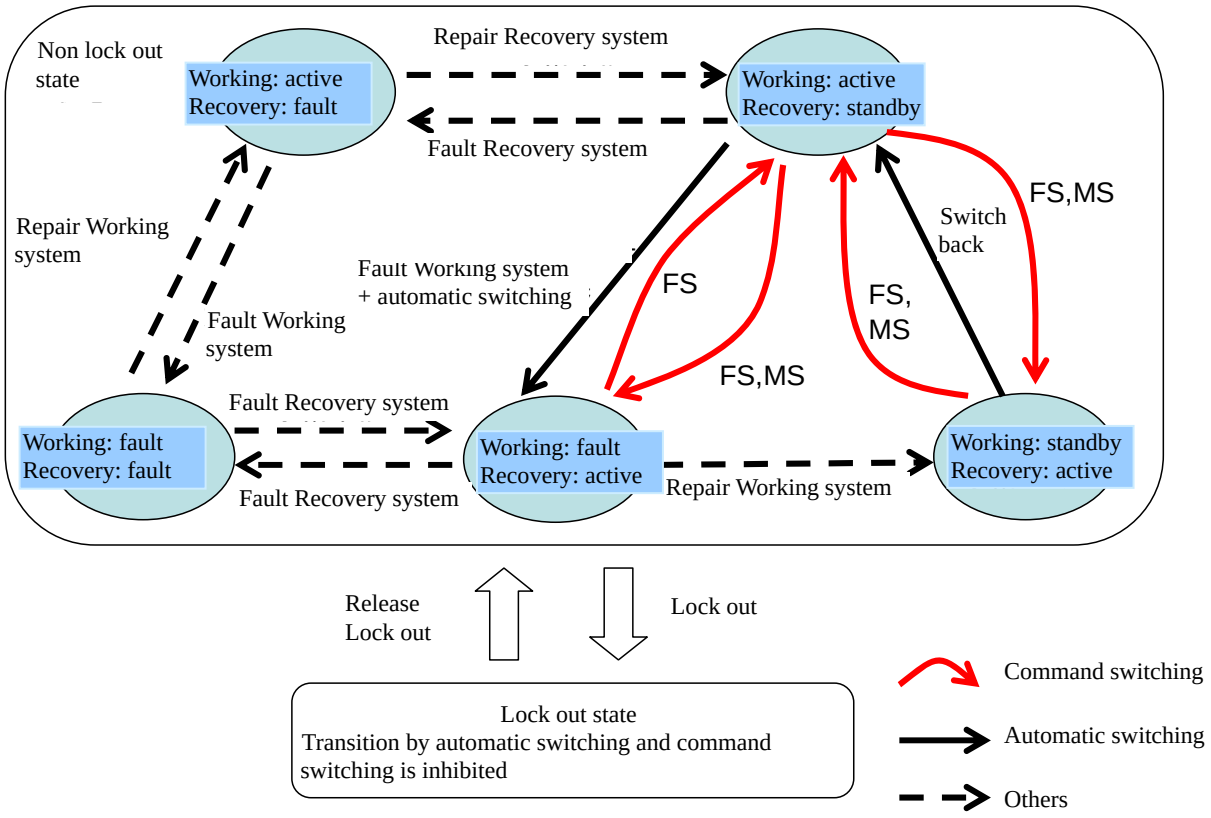
Switching should be executed following to the procedure written in [E2E].

8. External Commands

As the external commands relating to fault recovery, following three commands should be implemented.

- Lock Out
  This command inhibits the switching operation of either automatic or manual.
- Forced Switch
  This command switches the working system to non-working system regardless of the status of destination of switching, instead of locked out state.
- Manual Switch
  This command switches the working system to non-working system when destination of switching is in normal state or is possible to be switched, instead of locked out state.

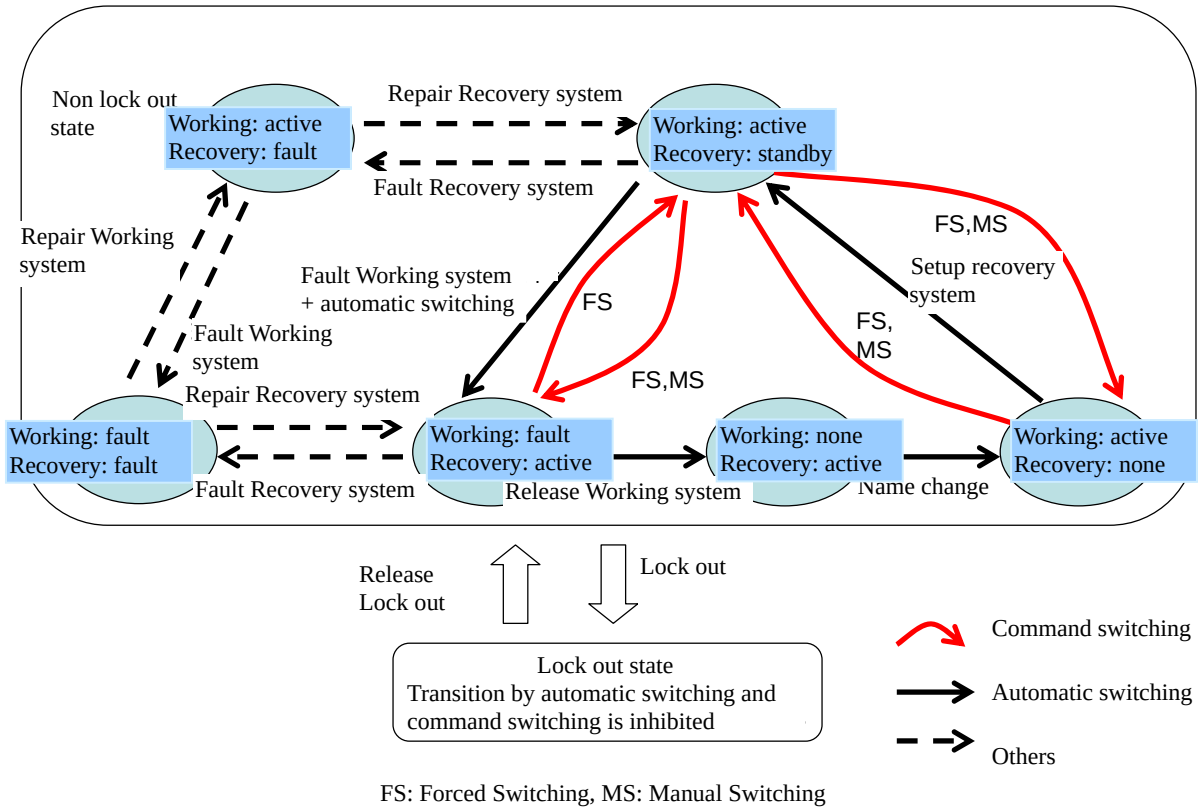
Fig.8-1 shows a state transition diagram in restoration in case that protection and switch-back are executed. In lock-out state, state transition by automatic switching and command switching is inhibited. Lock Out state can be released only by lock out release command.



FS: Forced Switching, MS: Manual Switching

Fig.8-1 State transition in restoration with protection and switch-back

Fig.8-2 shows a state transition diagram in restoration in case that switch-back is not executed. In restoration without switch-back, it is required to do name change operation because recovery system is operated as a working system after the fault recovery operation. In addition, resources of working system must be released. New recovery system is re-setup for new working system.



FS: Forced Switching, MS: Manual Switching

Fig.8-2 An example of state transition in restoration without switch-back

9. References

- RSVP-TE Extensions in support of End-to-End GMPLS-based Recovery <draft-ietf-ccamp-gmpls-recovery-e2e-signaling-03.txt >
- Generalized MPLS Recovery FunctionaLSP specification < draft-ietf-ccamp-gmpls-recovery-functional-04.txt >
- Recovery (Protection and Restoration) Terminology for GMPLS <draft-ietf-ccamp-gmpls-recovery-terminology-06.txt>
- Analysis of Generalized MPLS-based Recovery Mechanisms (including Protection and Restoration) <draft-ietf-ccamp-gmpls-recovery-analysis-05.txt >
- Optical Network Failure Recovery Requirements <draft-czezowski-optical-recovery-reqs-01.txt>
- Fault Notification Protocol for GMPLS-Based Recovery <draft-rabbat-fault-notification-protocol-02.txt>
- RSVP extensions for GMPLS restoration signaling <draft-shimano-imajuku-gmpls-restoration-00.txt>
- Extensions to LMP for Flooding-based Fault Notification <draft-soumiya-lmp-fault-notification-ext-00.txt>
- Extensions to RSVP-TE for Supporting Multiple Protection and Restoration Types <draft-suemura-gmpls-restoration-signaling-00.txt>s
- Protection of Hierarchical LSP s <draft-suemura-protection-hierarchy-00.txt>
- Extensions to OSPF-TE for supporting shared mesh restoration <draft-yagyu-gmpls-shared-restoration-routing-00.txt>

10. About This Document

10.1 Authors

NEC: Takehiko Suemura
NEC: Itaru Nishioka
Fujitsu Labs: Tishio Munemiya
Fujitsu Labs: Shinya Kanoh
Fujitsu Labs: Keiji Miyazaki
Furukawa Denko: Dai Mutoh
Mitsubishi Electric: Shoichiro Senoo
Mitsubishi Electric: Eiichi Horiuchi
Hitachi Communication Technology: Akio Nogi
Hitachi: Kenji Kataoka
NTT: Kouhei Shiomoto
NTT: Katsuhiro Shimano
NTT: Wataru Imajuku

10.2 Revision history

2003.6.4 Ver. 0.7 (Organized first the each company's draft)
2003.6.13 Ver. 0.9 (Results of discussion in PIL standardization WG held on 2003.6.10 was reflected)
2003.6.18 Ver. 0.92 (Revised preparing for the meeting of 2003.6.19)
2003.7.1 Ver. 0.95 (Contents of Saku-meeting was reflected)
2003.9.16 Ver. 1.0
2005.11.30 Ver. 2.0 (Revised reflecting the latest Internet Draft)