High-speed Rerouting Function | QoS Priority Control | IPv6 Support

Technology Reports

Special Articles on All-IP Network Technology —Evolution of Core Network —

Expansion of IP Backbone Network Functions for All-IP Network

As packet traffic increases and FOMA-voice and videophone traffic are superposed on the IP router network as part of a policy toward CS-IP to provide advanced and efficient services, the scalability, quality, and reliability of the IP router network needs to be improved. NTT DOCOMO is adding functions in conjunction with equipment upgrades in the IP router network to not only expand transmission speeds and capacity but also to provide IPv6 support functions, high-speed rerouting functions and OAM functions and to accommodate CS-IP/LTE-related nodes.

Core Network Development Department

Takashi Komuro
Yohei Kaminaga
Takayuki Obayashi[†1]
Hideki Kitahama[†2]
Ryo Saito[†2]

1. Introduction

The IP backbone (IP router network) is a core data-communications network constructed against the background of a growing demand for data communications reflected by the conversion of the packet network to IP, the separation of FOMA-voice and packet services, and the launch of Mzone (public wireless LAN service). It is a VPN type of backbone that combines the features of a wide-area Ethernet[*1] and an IP-Virtual Private Network (IP-VPN). The design requirements established at the time of IP router network construction are summarized below:

- High-speed transmission of IP traffic
- Separation and superposition of logical networks using VPN technology
- Quality guarantees by QoS priority control for critical traffic
- Reliable and high-availability network
- Layer-2 transfer of non-IP traffic
- Diverse connection formats according to the required bandwidths of accommodated nodes

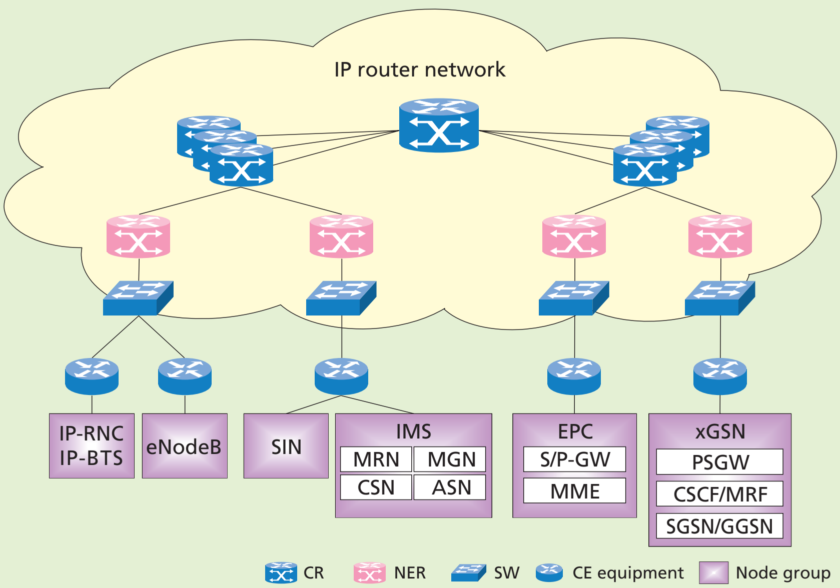
Since its construction in 2004, the IP router network has been expanding to meet the connection needs of diverse systems as a low-cost, high-speed and high-capacity platform for data communications. Recent advances in VoIP technology, moreover, are enabling the IP router network to advance and accommodate voice IP services such as Business mopera IP Centrex as well as FOMA-voice traffic and LTE[*2]-related nodes. The configuration of the IP router network and the IP node groups that it accommodates are shown in **Figure 1**.

This conversion of NTT DOCOMO networks to an All-IP format means that traffic accommodated by the IP router network is going to be increasing

[†1] Currently Network Service Operation Department.
[†2] Currently Core Network Division, DOCOMO Technology, Inc.

[*1] **Wide-area Ethernet**: Technology that connects remote LANs using LAN switches (layer-2 switches) to enable multiple users to communicate simultaneously over a single network.

[*2] **LTE**: Extended standard for the 3G mobile communication system studied by 3GPP. It is equivalent to "3.9G" or Super3G as proposed by NTT DOCOMO.



ASN	: Application Serving Node	MRF	: Media Resource Function
BTS	: Base Transceiver Station	MRN	: Media Resource Node
CSCF	: Call/Session Control Function	PSGW	: Packet Switched Gateway
CSN	: Call Session control Node	RNC	: Radio Network Controller
EPC	: Evolved Packet Core	S/P-GW	: Serving/Packet data network-Gateway
GGSN	: Gateway General packet radio service Support Node	SGSN	: Serving General packet radio service Support Node
IMS	: IP Multimedia Subsystem	SIN	: Signaling Interworking Node for 3G access
MGN	: Media Gateway Node	xGSN	: serving/gateway General packet radio service Support Node
MME	: Mobility Management Entity		

Figure 1 Configuration of IP router network

dramatically. For this reason, we have already introduced large-capacity routers, applied a Link Aggregation (LAG) function[*3], and taken measures to improve scalability such as by revamping the topology of the entire network. Plus, for voice traffic, the need is growing not just for greater capacity but also for enhanced communications quality without sacrificing the high-availability and maintainability characteristics provided by the existing FOMA network. It will therefore be necessary to add functions like IPv6 support functions that can support the characteristics of new nodes.

In this article, we describe new technologies introduced into the IP router network to accommodate Circuit Switched over-IP (CS-IP)/LTE-related nodes toward an All-IP network.

2. IPv6 Support Functions

2.1 Background

The IP router network accommodates a variety of systems including NTT DOCOMO's internal system and user-subscribed i-mode. To handle such diverse traffic, NTT DOCOMO adopts Multi-Protocol Label Switching (MPLS)[*4] technology, which encapsulates data using labels so that data of one system can be securely transmitted in a manner separate from data of other systems.

Initially, the IP router network accommodated only IPv4-based systems, but the ability to accommodate systems using IPv6 addresses has become an issue considering that the number of available IP addresses is becoming exhausted as new and varied services come to be developed and that LTE nodes need to be accommodated.

The IP router network must also be able to accommodate IPv6 systems without affecting the settings of the many existing systems currently being accommodated under L2-VPN[*5] and L3-VPN[*6] (IPv4).

We therefore studied ways of accommodating IPv6 systems that would enable operations to continue unimpeded without having to make drastic changes to the facilities, systems, and settings of the existing MPLS-based IP router network. This study led us to adopt the IPv6 VPN Provider Edge (6VPE) method [1].

2.2 6VPE Features

The 6VPE method features technology for accommodating IPv6 systems by converting IPv6 packets to MPLS frames before transfer. This method introduces a new type of Provider Edge (PE)[*7] device called a Next Edge Router (NER)[*8] into the IP router network. With 6VPE, the NER receives packets from the Customer Edge (CE)[*9] and

[*3] **LAG function**: LAN configuration technology that treats multiple physical circuits as one virtual circuit.

[*4] **MPLS**: Technology enabling the high-speed transmission of data by label-based switching.

[*5] **L2-VPN**: A virtual private network operating on layer 2. It interconnects sites of an accommodated system using a point-to-point (1-to-1) format.

[*6] **L3-VPN**: A virtual private network operating on layer 3. It interconnects sites of an accommodated system using a point-to-multipoint (1-to-many) format.

[*7] **PE**: Equipment for terminating VPN services and for providing those services to accommodated systems. Corresponds to NER (see *8) in the IP router network.

judges them to be either IPv6 or IPv4 enabling MPLS labels to be attached. As a result, switching of those packets at the Core Router (CR)*10 can be performed on the basis of those labels without having to worry about differences between IPv6/IPv4 protocols (**Figure 2**).

Another feature of 6VPE is its VPN orientation. The NER can accommodate multiple systems and manage the logical separation of circuits thereby maintaining existing aggregation effects and security. In short, 6VPE makes it possible to mix the IPv6/IPv4 protocols on the same physical circuits.

In addition to allowing existing IP router network facilities and network operation systems to be appropriated, deploying 6VPE with these kinds of features enables IPv6 systems to be accommodated while localizing the effects on existing systems and functions, that is, without affecting the settings of high-speed rerouting functions and Operation, Administration and Maintenance (OAM) functions.

3. High-speed Rerouting Functions

The IP router network is used to transfer a wide variety of traffic, and finding ways to perform fault detection and rerouting quickly at the time of equipment failure to shorten service interruption time is a major issue. High availability is a particularly strict requirement for data that demands continuity as in voice traffic, and to shorten interruption time as much as possible during a fault, the IP router network

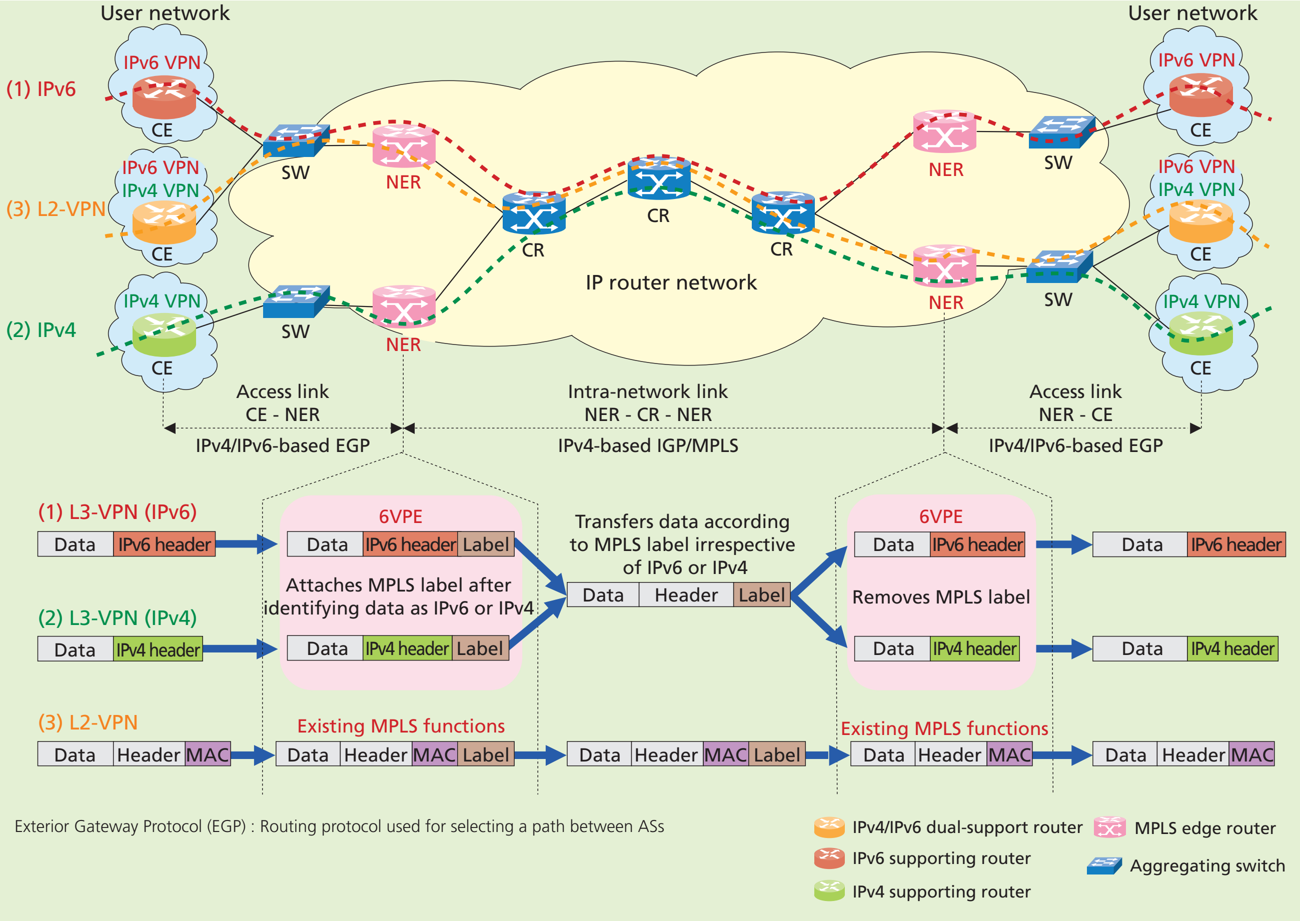


Figure 2 6VPE configuration and operation overview

*8 **NER**: PE equipment in the IP router network. Identifies the destination addresses (IPv4/IPv6) in accommodated systems and attaches labels accordingly.

*9 **CE**: Terminates accommodated systems and connects to VPN services.

*10 **CR**: Aggregates NERs in the IP router network and executes MPLS high-speed label transfers.

adopts rerouting functions that differ according to the characteristics of each link, namely, the access link (CE - NER) and intra-network link (NER - CR - NER).

3.1 High-speed Rerouting Function on the Access Link

It is imperative that both fault detection and rerouting be performed quickly to shorten the time from fault occurrence to rerouting.

In the IP router network, multiple CEs are aggregated at an aggregating switch (SW) to make efficient use of circuit bandwidth and multiplexed into the physical circuit making up an access link. As a result, the following issues have arisen in terms of fault detection and rerouting (**Figure 3**).

- Issue 1: Fault detection

  Given the occurrence of a fault between a CE and SW, the fault cannot be detected at an NER until the Border Gateway Protocol (BGP)[*11] hold timer[*12] expires, which means that packets will be discarded before rerouting can be performed after a fault occurs. Furthermore, in the event of a fault between an SW and NER, our NER is not equipped with a function for notifying the CE of a fault, which means that packets will be discarded before fault detection on the CE.

- Issue 2: Rerouting

  When a fault occurs between a CE and SW or between that SW and an NER, traffic from opposing NERs will continue to flow to that NER until a path can be changed to an opposing NER by Internal BGP (IBGP) from the NER detecting the fault.

With respect to issue 1, we have achieved fast fault detection by adopting a Bidirectional Forwarding Detection (BFD) function between the CE and NER (**Figure 4**).

To begin with, the BFD function establishes a BFD session and proceeds to transmit BFD control packets periodically between equipment in an adjacent relationship in routing protocol. Then, in the event that this BFD control pack-

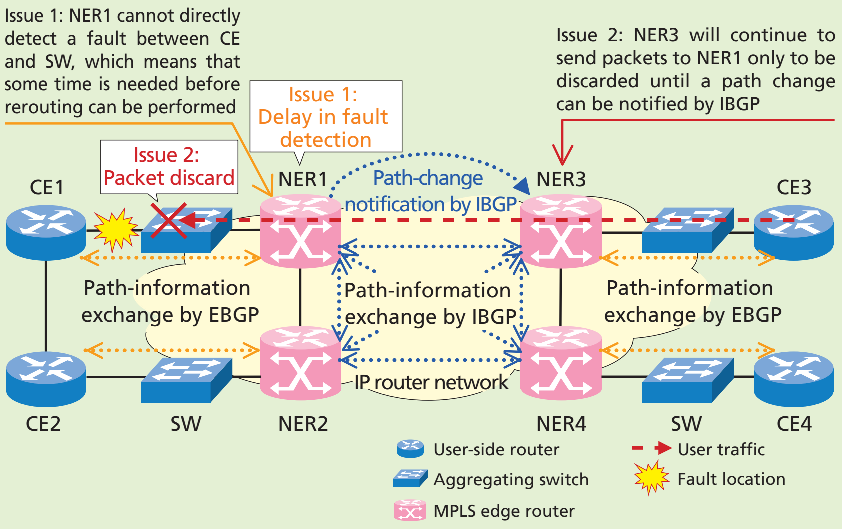

Issue 1: NER1 cannot directly detect a fault between CE and SW, which means that some time is needed before rerouting can be performed
Issue 1: Delay in fault detection
Issue 2: Packet discard
Issue 2: NER3 will continue to send packets to NER1 only to be discarded until a path change can be notified by IBGP
Path-change notification by IBGP
CE1
NER1
NER3
CE3
Path-information exchange by EBGP
Path-information exchange by IBGP
Path-information exchange by EBGP
IP router network
CE2
SW
NER2
NER4
SW
CE4
User-side router
Aggregating switch
MPLS edge router
User traffic
Fault location

Figure 3 Issues in access-link fault rerouting

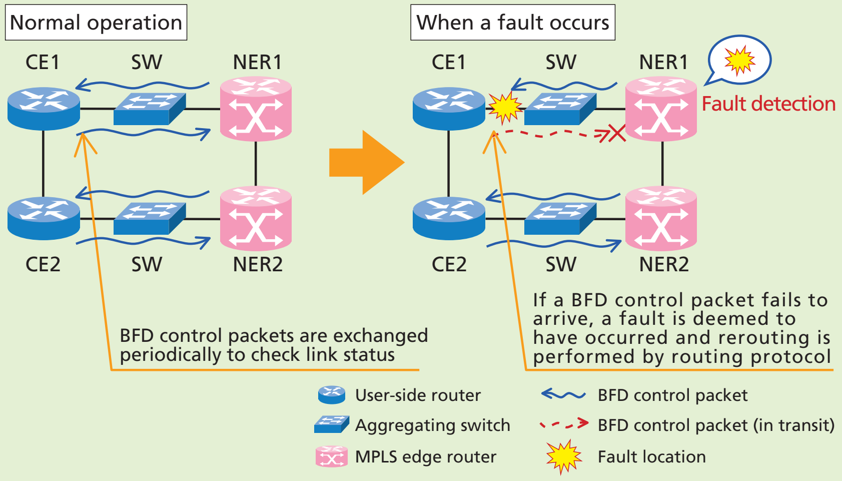

Normal operation
When a fault occurs
CE1
SW
NER1
CE2
SW
NER2
CE1
SW
NER1
Fault detection
CE2
SW
NER2
BFD control packets are exchanged periodically to check link status
If a BFD control packet fails to arrive, a fault is deemed to have occurred and rerouting is performed by routing protocol
User-side router
Aggregating switch
MPLS edge router
BFD control packet
BFD control packet (in transit)
Fault location

Figure 4 Fault detection by BFD

[*11] **BGP**: A type of routing protocol that selects a path based on information exchanged between Autonomous Systems (ASs).

[*12] **Hold timer**: The interval from the last message exchanged between BGP speakers in a BGP session to the point at which the session is deemed to be terminated.

et is not received a certain number of times, the function infers that a fault has occurred in the link between those opposing pieces of equipment and notifies routing protocol of that fault.

Next, with respect to issue 2, the NER Local Link Convergence (LLC) function is used to provide traffic relief while performing rerouting operations (**Figure 5**).

With LLC, two BGP paths are prepared beforehand with respect to some address, and among these, the path from the CE based on External BGP (EBGP) is taken to be the best path for the moment. If a fault should then be detected in the link between CE and NER, the current best path is considered to be down and the other saved BGP path is instantaneously adopted as the best path. This enables packets that have arrived at the NER in question to be diverted along this detour thereby preventing packets from arriving at the fault location during a rerouting operation and being lost.

In the above way, fault detection by BFD and high-speed rerouting by LLC are combined to deal with a fault in the link between CE and NER.

3.2 High-speed Rerouting Function within the IP Router Network

For high-speed fault detection within the IP router network (NER - CR - NER), we apply extended Link Aggregation Control Protocol (LACP), which is a standard protocol using the LAG function. This extension is called enhanced-LACP (e-LACP).

Typical high-speed fault detection functions in this link include BFD described above and the Open Shortest Path First (OSPF)[*13] fast-hello[*14] system. These techniques, however, perform detection on the IP layer. As mentioned earlier, the IP router network makes use of the LAG function for expanding transfer capacity. Here, though, only one IP address is assigned to a logical link that bundles multiple physical links, which prevents a fault detection system on the IP layer from performing high-speed fault detection on all physical links.

The e-LACP function, in contrast, sends and receives LACP packets and detects faults on all physical links enabling individual physical links to be identified. What's more, LACP packets are sent and received at high-speed intervals of several tens of milliseconds making for high-speed fault detection. The e-LACP function, moreover, can also be applied to links that do not use the LAG function (that do not require the bundling of multiple physical links), which enables e-LACP to run and high-speed fault detection to be performed on all links within the IP router network. This is why we adopted e-LACP for use in the IP router network.

Next, considering that the IP router network is based on MPLS technology, we are achieving high-speed rerouting by using Traffic Engineering (TE), a technology that has found widespread use on MPLS.

With TE, relay routes are explicitly defined independent of path selection based on an Interior Gateway Protocol (IGP)[*15] metric as in OSPF. TE also provides an additional function called Fast Reroute (FRR), which is used in the IP router network to achieve

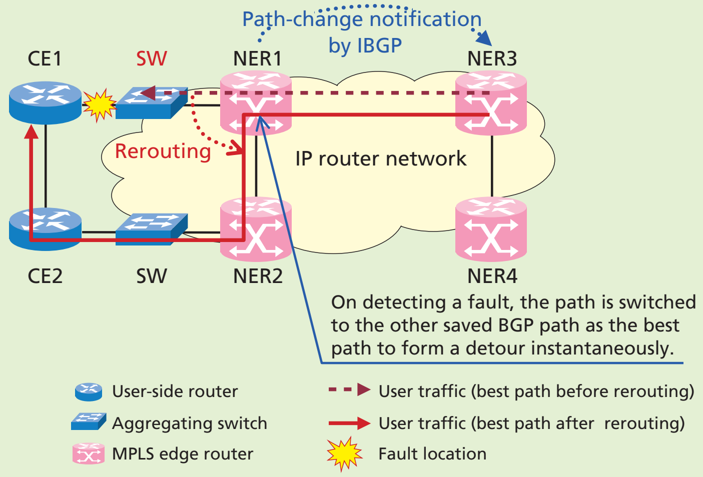

Path-change notification by IBGP
CE1
SW
NER1
NER3
Rerouting
IP router network
CE2
SW
NER2
NER4
On detecting a fault, the path is switched to the other saved BGP path as the best path to form a detour instantaneously.
User-side router
Aggregating switch
MPLS edge router
User traffic (best path before rerouting)
User traffic (best path after rerouting)
Fault location

Figure 5 Traffic relief by LLC

[*13] **OSPF**: A routing protocol that selects the minimum cost path based on numerical data that indicates interface weighting values called cost.

[*14] **Fast-hello**: A high-speed fault detection system that sends and receives control packets (hello packets) at high speed.

[*15] **IGP**: A routing protocol used for path selection within an AS.

high-speed rerouting by setting backup tunnels beforehand and rerouting traffic to such a backup tunnel immediately on detection of a fault. This backup operation is shown in **Figure 6**. As shown in the figure, a fault detected on the normally used path (Tunnel 1) will result in rerouting to another previously set path (Tunnel 2).

Combining the above techniques in the IP router network achieves a high-speed rerouting function in response to a fault in the network.

4. OAM Functions

Ethernet technology has recently come to be incorporated even in operator networks, but as this technology was originally designed for LAN use, it has been deficient in a link fault troubleshooting function that operators need when operating a wide-area network. The ping[*16] function on the IP layer has traditionally been used for checking connectivity in the network. However, even if ping should detect the presence of a fault, it cannot immediately determine whether the cause of that fault lies on the IP layer or on the Ethernet layer itself. For this reason, we introduced OAM functions in the IP router network for troubleshooting faults on the Ethernet layer.

Specifically, we introduced MPLS-OAM [2] within the IP router network (NER - CR - NER), which is the MPLS-applicable link, and Ethernet-OAM in the MPLS-non-applicable access link (**Figure 7**).

4.1 MPLS-OAM Functions

The MPLS-ping/trace[*17] function (IP connectivity check within the core network) that specifies an IP address for a destination and the VPN-ping/trace function (inter-VPN connectivity check) have already been incorporated in existing PE equipment as MPLS-OAM functions for use in the IP router network (NER - CR - NER). Now, with the introduction of NER equipment, it has also become possible to apply Label Switched Path (LSP)[*18]-ping/trace and Virtual Circuit Connectivity Verification (VCCV)[*19] [3].

The LSP-ping/trace function speci-

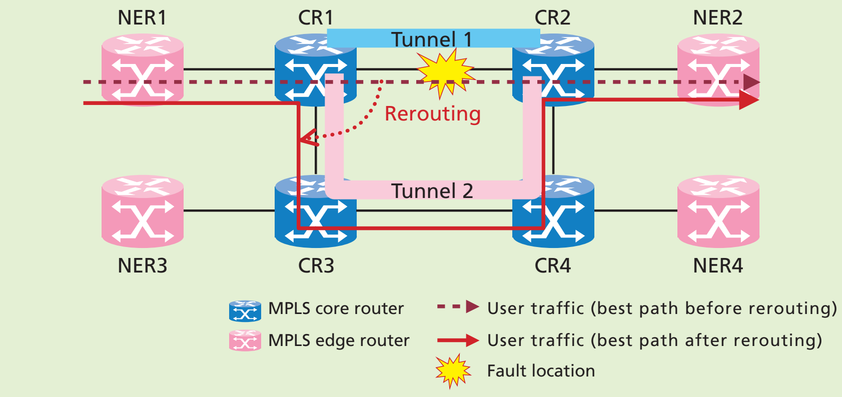
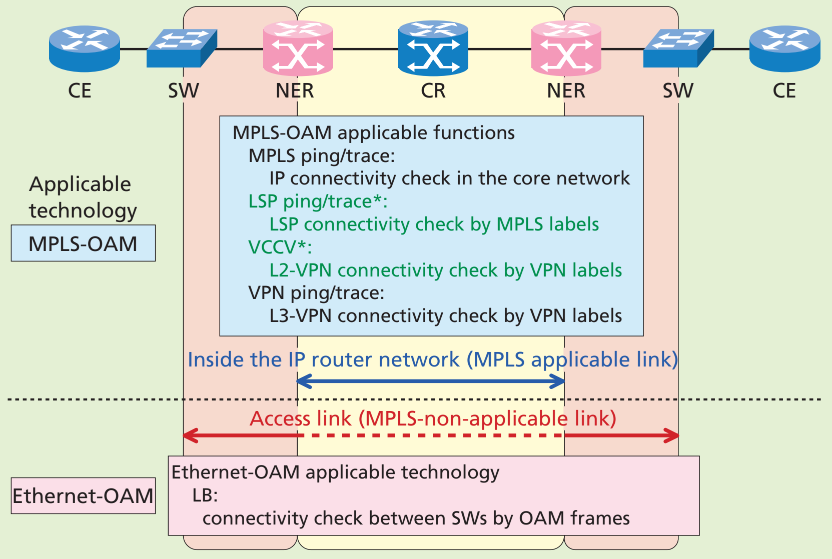


Figure 6 Overview of MPLS-TE operation


CE
SW
NER
CR
NER
SW
CE
Applicable technology
MPLS-OAM
MPLS-OAM applicable functions
MPLS ping/trace:
IP connectivity check in the core network
LSP ping/trace*:
LSP connectivity check by MPLS labels
VCCV*:
L2-VPN connectivity check by VPN labels
VPN ping/trace:
L3-VPN connectivity check by VPN labels
Inside the IP router network (MPLS applicable link)
Access link (MPLS-non-applicable link)
Ethernet-OAM
Ethernet-OAM applicable technology
LB:
connectivity check between SWs by OAM frames

* Newly introduced functions in the IP router network

Figure 7 Application scope of OAM functions

[*16] **Ping**: A function for checking connectivity by sending a packet to the host computer of the other party and receiving a reply to that packet.

[*17] **Trace**: A function for verifying the path that will be crossed from a certain node to a specified node.

[*18] **LSP**: Path set between routers based on MPLS labels.

[*19] **VCCV**: VPN connectivity check up to the VPN label in an MPLS network.

fies an MPLS Forwarding Equivalence Class (FEC)[*20] enabling an LSP connectivity check to be performed using MPLS labels and the actual path taken by MPLS packets to be verified. The VCCV function enables addresses to be segmented and specified up to the Pseudowire ID (PWID)[*21] and for a VPN connectivity check up to the VPN label to be performed. The above functions make it possible, for example, to detect that a fault such as a packet loss has occurred in an MPLS transfer even though a normal IP transfer can be made between the nodes involved.

4.2 Ethernet-OAM Functions

The MPLS-non-applicable access link (between SWs) applies the Ethernet-OAM [4] Loop Back (LB) function for performing a connectivity check. This function is similar to the IP ping function, and it specifies a Media Access Control (MAC) address instead of an IP address enabling connectivity checks to be performed on the Ethernet level.

Ethernet-OAM enables maintenance domains to be specified so that checks that are closed to a domain in a specific region can be made without affecting nodes outside of that domain.

5. Conclusion

We described a function group that NTT DOCOMO has introduced to accommodate CS-IP/LTE-related nodes in the IP router network toward the construction of a full-scale All-IP network. NTT DOCOMO has also expanded QoS priority control functions, enabled emergency calls to be given priority at times of congestion, and provided other functions in unison with the characteristics of accommodated services.

Looking forward, we plan to improve scalability of the IP router network as an operator's core data communications network to deal with further increases in traffic. We will also continue to study the timely provision of functions in conjunction with the accommodation of new services.

References

[1] IETF RFC 4659: "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN," 2006.

[2] IETF RFC 4379: "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures," 2006.

[3] IETF RFC 5085: "Pseudowire Virtual Circuit Connectivity Verification (VCCV) - A Control Channel for Pseudowires," 2007.

[4] IEEE 802.1ag: "Virtual Bridged Local Area Networks - Amendment 5: Connectivity Fault Management," 2007.

[*20] **FEC**: A group of packets for which the same address or handling is desired within the MPLS network.

[*21] **PWID**: The ID identifying a virtual circuit (pseudowire) constructed by MPLS or another tunneling technique.