

All-IP ネットワークの高度化技術特集—コアネットワークの進化—

All-IP ネットワークを支える IP バックボーンの
ネットワーク機能拡充

近年のパケットトラフィック量の増大や、サービスの高度化・効率化を目指したCS-IP化施策におけるIPルータ網へのFOMA音声/テレビ電話トラフィック重畳などにより、IPルータ網にはさらなるスケーラビリティの拡大と品質・信頼性の向上が求められている。そのため、IPルータ網において機器更改を伴う機能追加を行うことで転送速度・容量の拡張を図るとともに、IPv6収容機能や高速切替機能、OAM機能などの機能拡充を実現し、CS-IP/LTE関連ノードの収容を可能とした。

ネットワーク開発部

こむろ たかし
小室 貴史おおばやし たかゆき
大林 隆之さいとう りょう十
齊藤 良かみなが ようへい
神永 洋平きたはま ひでき
北濱 秀基

1. まえがき

IPバックボーン（IPルータ網）は、FOMAの音声/パケット分離に伴うパケット網のIP化やMzone（公衆無線LANサービス）の開始など、データ通信需要の拡大を背景に構築された基幹データ通信網であり、広域Ethernet^{*1}とIP-VPN（IP-Virtual Private Network）サービスの特徴を併せもつVPNバックボーンである。IPルータ網構築時の設計要件を次に示す。

- ①IPトラフィック高速転送
- ②VPN技術を用いた論理網の分離重畳
- ③重要トラフィックのQoS優先制御による品質保証

- ④ネットワークの信頼性・高可用性の確保
- ⑤非IPトラフィックのレイヤ2転送
- ⑥収容ノードの必要帯域に合わせた多様な接続形態

IPルータ網は、2004年の構築以来、低コストで高速かつ大容量な通信網基盤として多種多様なシステムからの接続要望に応えるべく拡張を続けてきたが、昨今のVoIP技術の発達により、「ビジネスmopera IPセントレックス」などの音声IPサービスの収容をはじめとして、FOMAの音声トラフィックやLTE^{*2}関連ノードも、IPルータ網への収容が進められている。IP系ノード群を収容する

IPルータ網の構成を図1に示す。

これらドコモネットワークのAll-IP化に伴い、IPルータ網の収容トラフィックが飛躍的に増大するため、これまでも、大容量ルータの導入やLAG（Link Aggregation）機能^{*3}の適用に加え、ネットワーク全体のトポロジ改善など、スケーラビリティの向上施策を行ってきた。加えて、音声トラフィックを転送する場合、大容量化対応だけでなく、既存FOMA網で実現してきた高可用性や保守性を損なわないよう、さらなる通信品質の向上が求められるため、IPv6収容機能など、新規収容ノードの特性に応じた機能追加も必要となる。

本稿では、ドコモネットワークの

† 現在、ドコモ・テクノロジー株式会社 コアネットワーク事業部

*1 広域 Ethernet：LANスイッチ（レイヤ2スイッチ）を用い、遠隔地にあるLANどうしを接続し、1つのネットワークを複数の利用者が同時に通信ができるように工夫された技術。

*2 LTE：3GPPで検討されている第3世代

移動通信方式の拡張規格。ドコモがSuper3Gとして提唱したもので、「3.9G」と位置付けられる。

*3 LAG機能：複数の物理回線を仮想的に1本の回線のようにみなすLANの構成技術。

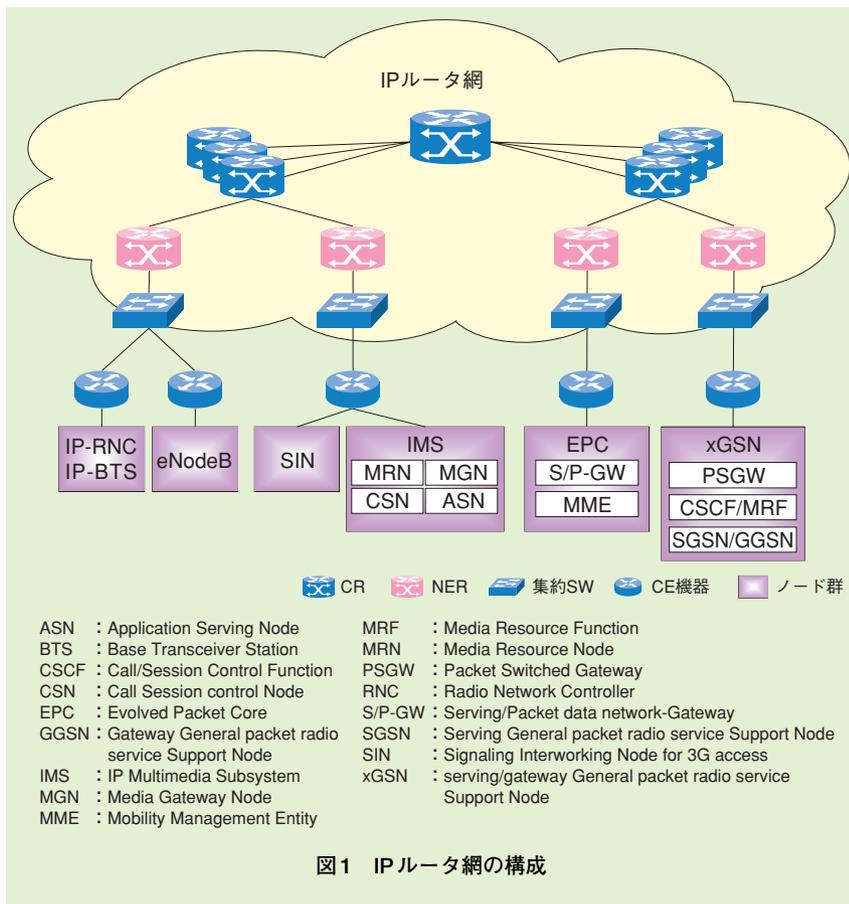


図1 IP ルータ網の構成

All-IP 化に向け、CS-IP (Circuit Switched over-IP) /LTE 関連ノードを収容するために、IP ルータ網に導入した新規技術について解説する。

2. IPv6 システム収容機能

2.1 背景

IP ルータ網では、ドコモの社内システムやユーザの契約する i-mode トラフィックなど、さまざまなシステムを収容しているため、ラベルによりデータをカプセル化することで、収容システム間をセキュアに分離して伝達できる MPLS (Multi-Protocol Label Switching)^{*4} 技術を

採用している。

当初は IPv4 によるシステムのみを収容していたが、今後のさまざまな新サービスの展開による IP アドレス枯渇や LTE ノード収容に対応するために、IPv6 アドレスによるシステム収容が課題となっていた。

さらに IP ルータ網では、L2-VPN^{*5} や L3-VPN^{*6} (IPv4) で収容している既存システムが多数あり、それら既存システムの設定に影響を与えずに IPv6 システムを収容する必要があった。

そこで、IPv6 システム収容に際して、MPLS 技術ベースの IP ルータ網の設備資産や設定を極力変更せずに

運用を継続できるかなどを検討した結果、6VPE (IPv6 VPN Provider Edge) [1]を採用した。

2.2 6VPE の特長

6VPE は、IPv6 システムを収容する際に、該当 IPv6 パケットを MPLS フレームに変換して転送する技術であり、IP ルータ網では新規 PE^{*7} である NER (Next Edge Router)^{*8} に導入している。6VPE の適用により、NER では CE (Customer Edge)^{*9} から受信したパケットが IPv6 か IPv4 かを判別したうえで、MPLS ラベルの付与が可能となり、その結果、CR (Core Router)^{*10} では IPv6/v4 のプロトコル差分を意識する必要がなく、MPLS ラベルによるスイッチングが可能となる (図2)。

また、6VPE のもう1つの特長は VPN の概念を有している点であり、NER で複数のシステムを収容しても論理的に回線を分離した管理が可能で、既存の集約効果とセキュリティを維持したうえで、IPv6/v4 の両プロトコルを同一物理回線上に混在させることを可能とした。

これらの特長をもつ 6VPE の適用により、既設 IP ルータ網設備やネットワーク監視システムが流用できることに加え、高速切替機能/OAM (Operation Administration and Maintenance) 機能などの他機能の設定に影響をおよぼさないなど、既存のシステムや機能に対する影響を局所化しながら、IPv6 システムの収容を可能とした。

*4 MPLS : ラベルに基づいたスイッチングにより高速なデータ転送を行う技術。

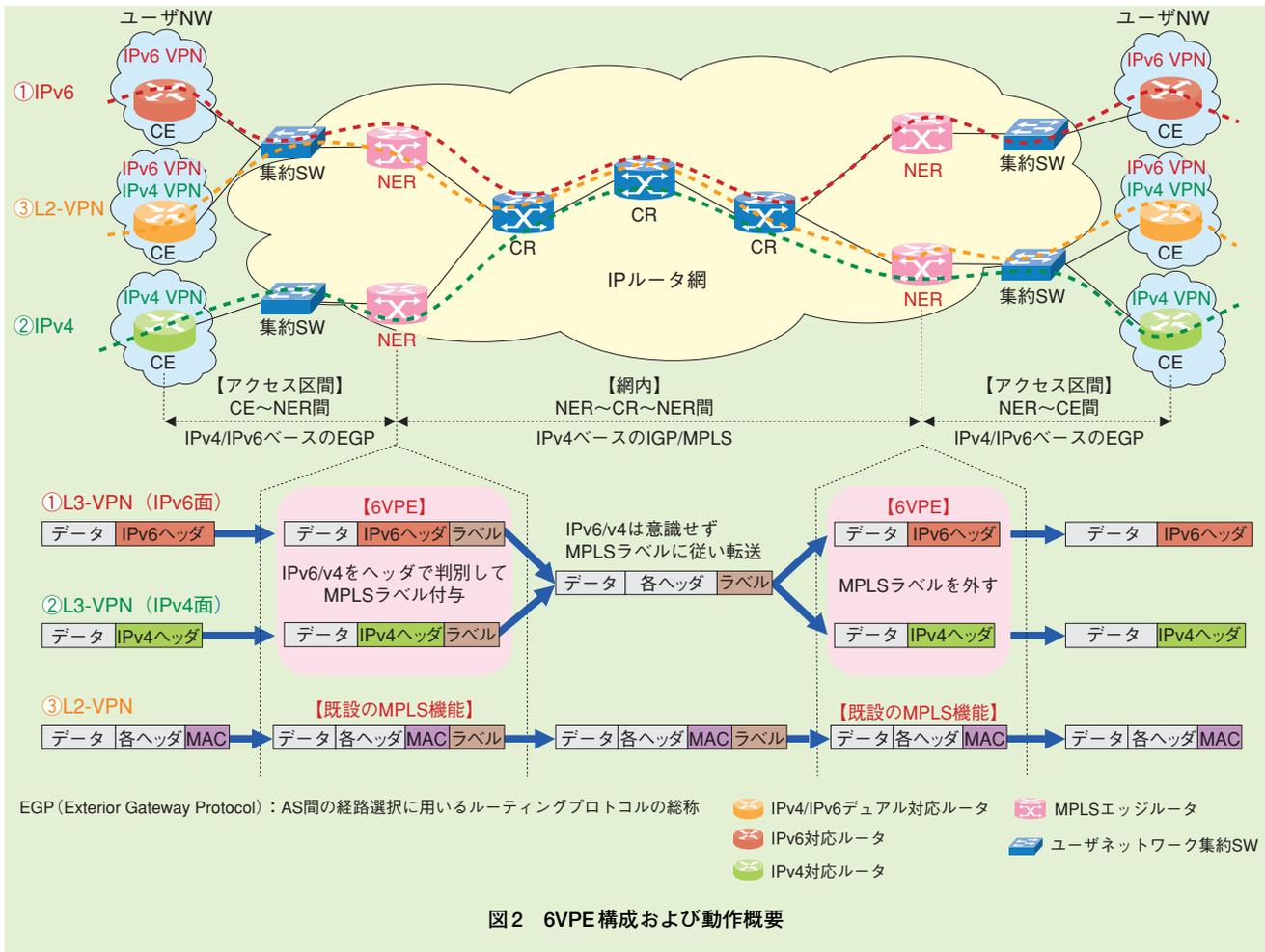
*5 L2-VPN : レイヤ2での仮想プライベートネットワーク。収容システムのサイト間をポイント・ツー・ポイント (1対1) 形態で接続する。

*6 L3-VPN : レイヤ3での仮想プライベートネットワーク。収容システムのサイト間をポイント・ツー・マルチポイント (1対多) 形態で接続する。

*7 PE : VPN サービスを終端し、収容システムに対して VPN サービスを提供する装置。

置。IP ルータ網では NER (*8 参照) が該当する。

*8 NER : IP ルータ網における PE 機器。収容システムのあて先アドレス (IPv4/IPv6) の判別およびラベルの付加を行う装置。



3. 高速切替機能

IPルータ網で多種多様なトラフィックを転送するうえで、装置故障時の障害検知/経路切替えを迅速に行い、サービス中断時間をいかに短縮していくかは、重要な課題である。特に、音声トラフィックのような連続性を求められるデータの場合、高可用性に対する要求が最も厳格であり、障害時の中断時間を極力短縮するために、IPルータ網では、「アクセス区間 (CE～NER間)」および「網内 (NER～CR～NER間)」の各

区間の特性に応じて異なる切替機能を採用している。

3.1 アクセス区間 高速切替機能

障害発生から経路切替えを行うまでの時間を短縮するには、「障害検知」と「経路切替え」の両面で高速化を図る必要がある。

IPルータ網では、回線帯域と物理ポートの有効活用のために複数のCEを集約スイッチ (SW) で束ね、アクセス区間の物理回線上に多重する構成をとるため、「障害検知」と

「経路切替え」の両面でそれぞれ次のような課題があった (図3)。

課題①：障害検知

CE～集約SW間で障害が発生した場合、NERではBGP (Border Gateway Protocol)^{*11} holdタイマ^{*12}の満了まで障害検知が行えないため、障害検知から経路切替え動作まで時間を要していた。また、集約SW～NER間の障害については、新規PEとして導入したNERではCE側へ障害通知を行う機能を具備していないため、CE側での障

*9 CE : 収容システムを終端し、VPNサービスに接続する装置。

*10 CR : IPルータ網内でNERを集約し、MPLS高速ラベル転送を実行する装置。

*11 BGP : ルーティングプロトコルの1つ、自律システム (AS : Autonomous System) 間で交換する情報に基づいて経路選択する。

*12 holdタイマ : BGPセッションを確立しているBGPスピーカ間で最後にメッセージをやり取りしてからセッションが切断されたかと判断するまでの間隔。

害検知に時間を要していた。

課題②：経路切替え

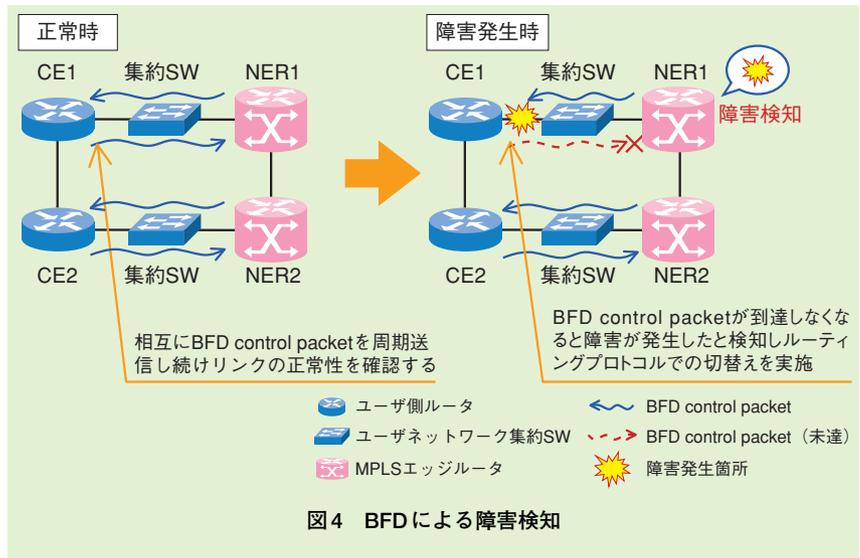
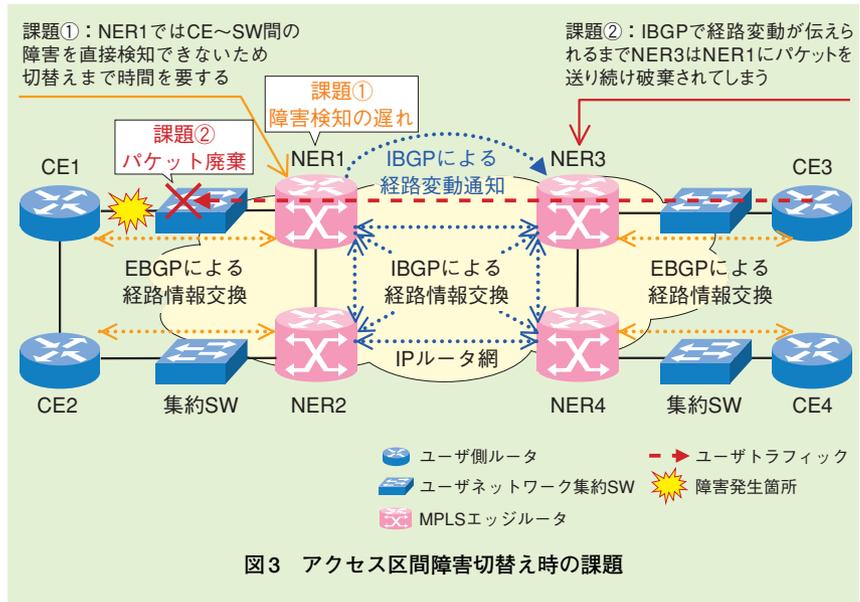
CE～集約SW～NER間で障害が発生した場合、障害を検知したNERからIBGP（Internal BGP）で対向NERへ経路変動を伝えられるまでの間、対向NERからのトラフィックは、障害が発生したNERへ向けて流れ続けてしまっていた。

課題①に対し、CE側とNERとで協調してBFD（Bidirectional Forwarding Detection）機能を採用することで、高速な障害検知を実現している（図4）。

BFDでは、まずルーティングプロトコルの隣接関係に対してBFD sessionの確立を行い、BFD control packetを周期的に送信する。このBFD control packetを対向機器から複数回受信できない場合、対向機器間とのリンクに障害が発生したと判断し、ルーティングプロトコルに対して障害発生を通知する。

次に課題②に対しては、NERの機能であるLLC（Local Link Convergence）を用いて、経路の切替動作を行っている最中のトラフィック救済を実現している（図5）。

LLCでは、あらかじめ任意のあて先に対して2つのBGPパスをもっておき、CEからのEBGP（External BGP）による経路がベストパスとなるようにしておく。CE～NER間のリンクで障害を検知すると、現在のベストパスを利用不可とし、保持しているもう1つのBGPパスをベスト

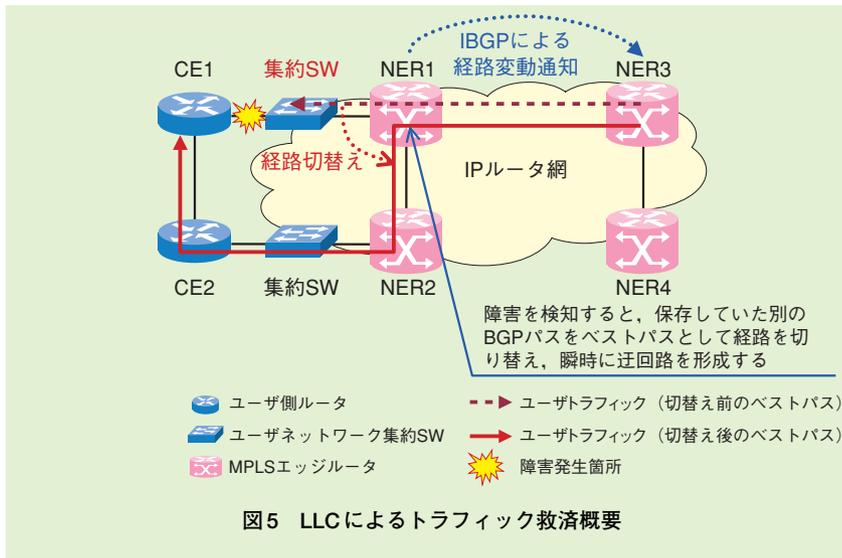


パスとして採用する。そうすることで、NERまで到達したパケットは迂回路を経由することが可能となり、網内の切替動作中に障害箇所へ到達したパケットロス回避することが可能となる。

このように、CE～NER間の障害においてはBFDでの障害検知とLLCでの高速切替機能を組み合わせている。

3.2 IP ルータ網内高速切替機能

IPルータ網内（NER～CR～NER間）の「高速障害検知」に関しては、LAG機能で使用している標準プロトコルであるLACP（Link Aggregation Control Protocol）を独自に拡張した技術（以下、e-LACP：enhanced-LACP）を適用している。



一般に障害の高速検知機能としては、アクセス区間で適用している前述のBFDやOSPF (Open Shortest Path First)^{*13}のfast-hello^{*14}などが挙げられるが、これらはIPレイヤでの検知技術となっている。一方IPルーター網では、転送容量拡大のためにLAG機能を用いるが、複数の物理リンクを束ねた論理リンクに対し1つのIPアドレスの付与を行っているため、IPレイヤの障害検知方式では、すべての物理リンクで高速に障害を検知することは不可能である。

e-LACPではLACPパケットをすべての物理リンクで送受信し障害を検知しているため、複数の物理リンクを個々に判別できる。またLACPパケットを数十ミリ秒という高速な単位で送受信することにより、障害の高速検知を可能としている。さらに、e-LACPはLAG機能を適用していない(複数本の物理リンクを束ねる必要がない)区間においても適用可能であるため、IPルーター網内の

全区間でe-LACPを動作させての高速障害検知を可能としている。そのため、IPルーター網ではe-LACPを採用している。

次に「高速経路切替え」においては、IPルーター網がMPLS技術をベースとしていることから、MPLS上で広く確立された技術であるTE (Traffic Engineering)を用いて実現している。

TEでは、OSPFのようなIGP (Interior Gateway Protocol)^{*15}のMetricに基づいた経路選択に依存せず、中継ルートを明示的に定義することが可能である。さらに、IPルーター網ではTEの付加機能であるFRR (Fast Reroute)を使用し、事前にバックアップトンネルを設定しておくことにより、障害が発生した際に直ちにバックアップトンネルへとトラフィック転送経路を切り替えることで、高速切替えを実現している。バックアップの動作概要を図6に示す。図6のように通常使用するパス (Tunnel 1)にて障害を検知すると、事前

に設定していた別のパス (Tunnel 2)へ経路切替えを行う。

IPルーター網では、これら複数の技術を組み合わせることにより、網内の障害時の高速切替機能を実現している。

4. OAM 機能

近年Ethernet技術はキャリア網でも取り入れられるようになったが、当該技術は元来LAN用技術として設計されたため、キャリアが広域ネットワークを運用するにあたり必要となる「回線障害切分け機能」が不足していた。ネットワークの疎通性確認には、これまでIPレイヤのping^{*16}などが使用されてきたが、pingでは障害が検出できても、その原因がIPレイヤとEthernetのいずれにあるのかは、即時に判別できなかった。そこで、Ethernetのレイヤ内で障害切分けに利用できるOAM機能^{*17}をIPルーター網に導入した。

MPLS適用区間であるIPルーター網内 (NER ~ CR ~ NER間)にはMPLS-OAM[2]を、MPLS非適用のアクセス区間についてはEthernet-OAMを導入した (図7)。

4.1 MPLS-OAM 機能

IPルーター網内 (NER ~ CR ~ NER間)のMPLS-OAM機能として、あて先にIPアドレスを指定するMPLS-ping/trace^{*18} (コア網内IP疎通確認)とVPN-ping/trace (VPN間疎通確認)については、既存PE機器にてすでに実現していたが、今回NERの導入に伴い、新たにLSP (Label Switched

*13 OSPF：ルーティングプロトコルの1つ。コストと呼ばれるインタフェースの重み値を示す数値情報に基づいて、最小コストとなる経路を選択する。

*14 fast-hello：障害の高速検知方式の1つ。送受信される制御パケット (helloパケッ

ト)を高速に送受信することで、障害の高速検知を行う。

*15 IGP：AS内部での経路選択に用いるルーティングプロトコルの総称。

*16 ping：通信相手のホストにパケットを発行し、その応答を受信することにより疎

通確認を行う機能。

*17 OAM機能：ネットワークにおける保守運用管理機能。

*18 trace：あるノードから指定したノードまで、どのような経路を通過しているかを確認する機能。

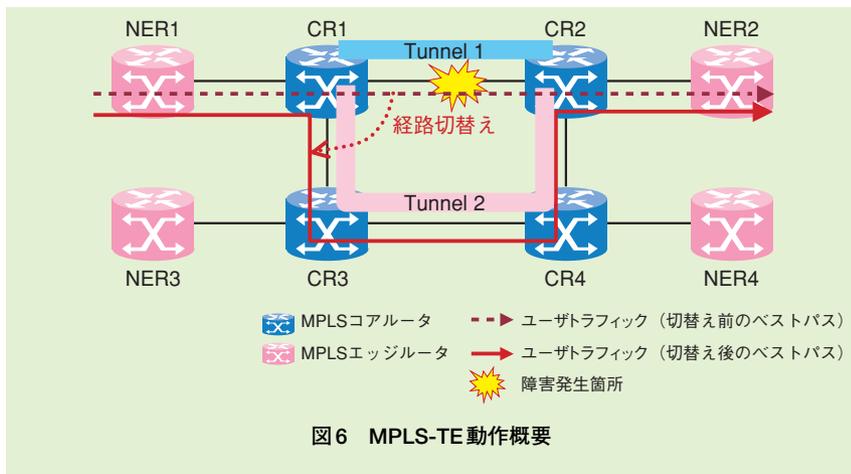


図6 MPLS-TE動作概要

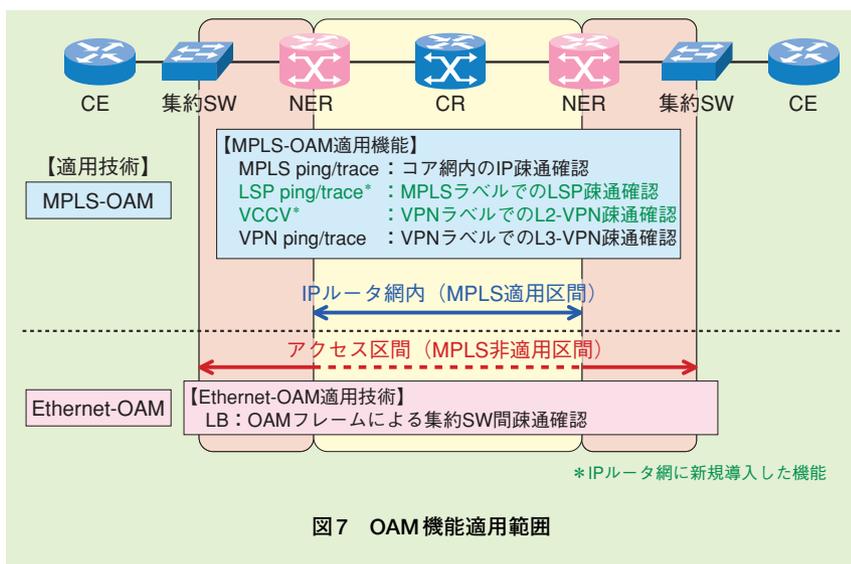


図7 OAM機能適用範囲

Path)^{*19}-ping/traceとVCCV (Virtual Circuit Connectivity Verification)^{*20}[3]の適用が可能となった。

LSP-ping/traceでは、MPLSのFEC (Forwarding Equivalence Class)^{*21}を指定してMPLSラベルでのLSP疎通確認が可能であり、MPLSパケットが実際に通るパスを確認できる。VCCVでは、PWID (Pseudowire ID)^{*22}まで指定してあて先を細分化でき、VPNラベルまでを含んだVPN疎通確認が可能である。これら

の機能により、例えば、ノード間にて通常のIP転送は可能であるが、MPLS転送ではパケットロスが発生するなどの障害が検知可能となる。

4.2 Ethernet-OAM機能

MPLS非適用区間であるアクセス区間(集約SW間)は、Ethernet-OAM[4]の疎通確認用のLB (Loop Back)機能を適用した。

Ethernet-OAMのLB機能は、IP pingと同様の機能で、IPアドレスで

はなくMAC (Media Access Control) アドレスを指定し、Ethernetレベルで疎通確認を行う機能である。

Ethernet-OAMでは、保守ドメインを規定することにより、特定地域のドメイン内に閉じた確認が可能となり、ドメイン外の収容ノードへの影響を排除できる。

5. あとがき

本稿では、本格的なAll-IPネットワーク構築に向け、CS-IP/LTE関連ノードをIPルータ網に収容するために導入した機能群について述べた。ほかにも、QoS優先制御機能を拡充し、輻輳時でも緊急呼の優先疎通を可能にするなど、収容サービスの特性に応じた的確な機能提供を実現している。

今後も、キャリアの基幹データ通信網として、さらなるトラフィックの増大に対応すべくスケラビリティの向上を図るとともに、新規サービスの収容に応じたタイムリーな機能提供に向け、検討を進めていく。

文献

- [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.

* 19 LSP: MPLSラベルに基づきルータ間に設定されるパス。
 * 20 VCCV: MPLS網におけるVPNラベルまでを含んだVPN疎通確認。
 * 21 FEC: MPLS網内で同じあて先、あるいは同じ扱いをさせたいパケットのグループ。

* 22 PWID: MPLSなどのトンネリング技術により構築された仮想回線 (Pseudowire) を識別するためのID。