大規模災害時におけるIP系OSSの信頼性向上

ドコモ・テクノロジ株式会社　ソリューションサービス事業部　北濱（きたはま） 秀基（ひでき）　長澤（ながさわ） 武文（たけふみ）†
西尾（にしお） 進（すすむ）　大谷（おおたに） 知行（ともゆき）

IP系オペレーションサポートシステム（OSS：Operation Support System）*1は，IPバックボーンネットワークを構成する大型ルータ／スイッチ，およびこれらを介して通信する周辺IPネットワーク機器を監視しており，監視対象数は約3万台におよぶ大規模な監視システムである．図1に示すように，今日まで運用してきた旧IP系OSSは，東京拠点ビル内にて2系統の冗長構成を取り，SNMP（Simple

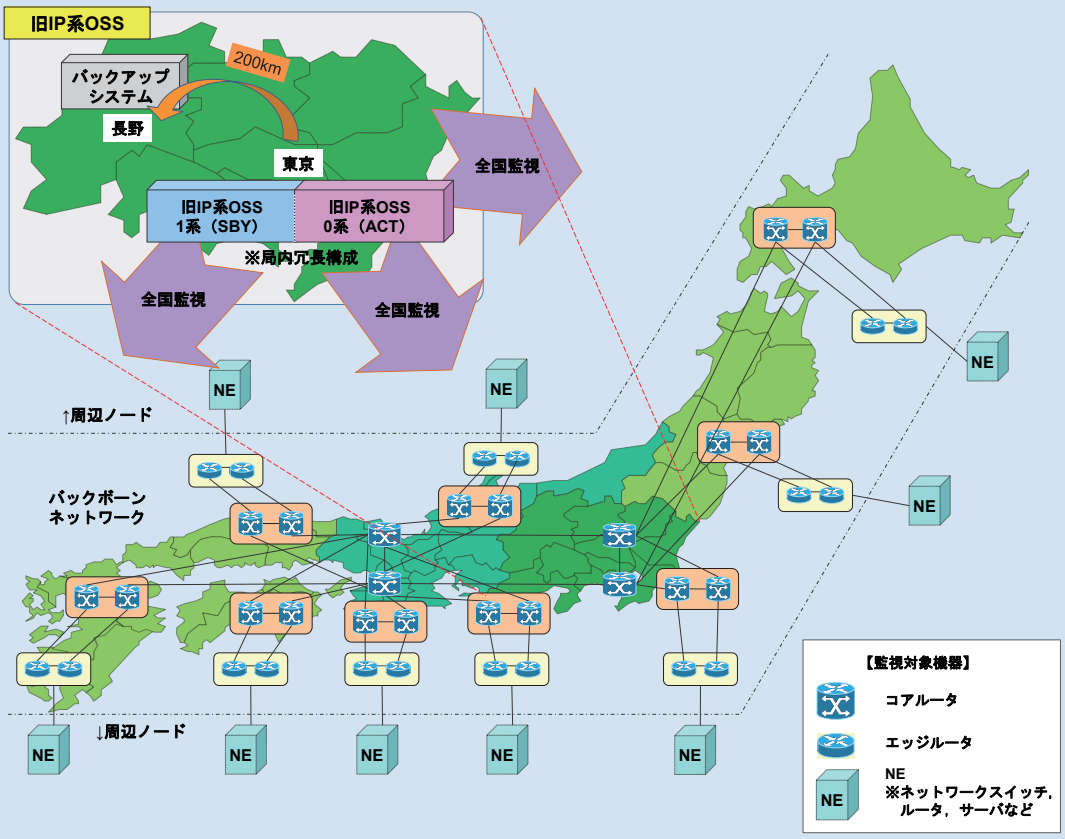

図1　旧IP系OSS監視システムの構成

©2016 NTT DOCOMO, INC.
本誌掲載記事の無断転載を禁じます．

† 現在，株式会社ドコモCS　ソリューションインテグレーション部

*1 オペレーションサポートシステム（OSS）：移動通信網で発生している故障や輻輳の発見とそれに対する制御・措置を行っているシステムのこと．

Network Management Protocol）*2/Syslog*3を用いて全国のNE（Network Element）*4を監視してきた．さらに，両系統が被災などで同時故障した場合を想定し，手動で起動し切り替える長野バックアップシステムを追加構築し，全体で3系統の冗長構成を取っていた．

2002年のシステム導入以来，監視不能に陥るような大きな障害に遭遇することなく安定運用を継続し，OSSの災害対策は万全と考えられていた．しかし，2011年の東日本大震災では，被災地域がこれまでの常識を超えて広範囲に至ったことから，2つのケースを想定する必要が出てきた．

①まず，200kmしか離れていない東京と長野では，最悪の事態として両拠点同時壊滅が考えられる．また，長野バックアップシステムが直接被災を免れたとしても，ライフラインや交通の麻痺による保守要員の確保が困難となり，起動を大幅に遅延させる事態も想定される（図2(a)）．

②2つめのケースとして全国のNEを収容するNE収容ネットワーク伝送路の分断が挙げられる．この場合，分断ポイントより先に収容されているNEが監視不能となり，影響範囲は極めて大きい（図2(b)）．

---

*2 **SNMP**：IPネットワーク上のネットワーク機器をモニタリング・制御するための情報の通信方法を定めるプロトコル．TRAP受信／MIB情報が取得できる．

*3 **Syslog**：システム動作状況やエラーメッセージなどを記録し，ネットワークを通じて他のコンピュータとその記録を送受信するためのプロトコル．

*4 **NE**：電気通信サービスを提供するにあたり，必要となる機能を実現する機能ブロックを指す．具体的には交換機，伝送装置，無線装置などの電気通信装置のこと．

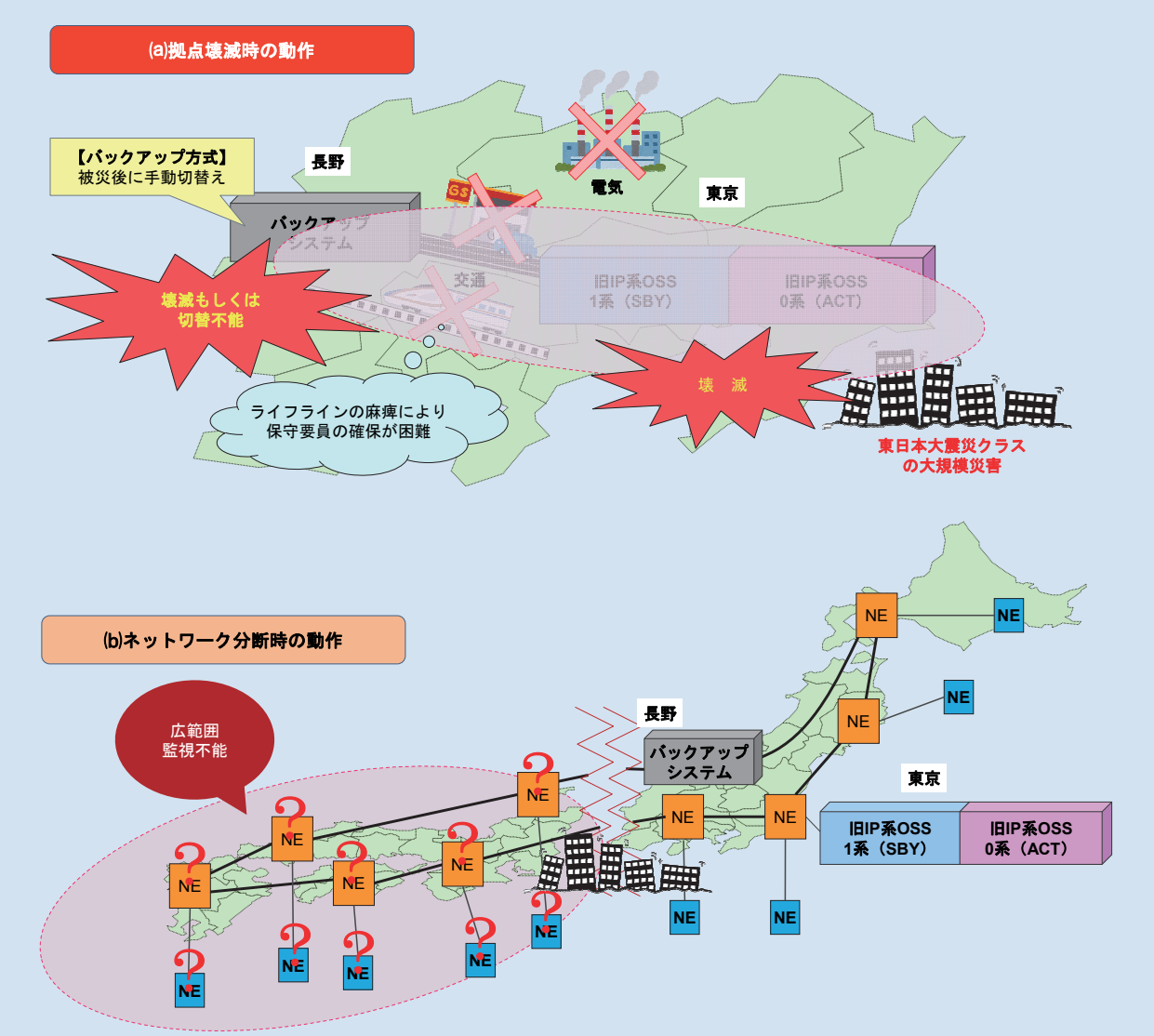

図2　旧IP系OSSの概要と大規模災害時の問題

このように従来のバックアップシステムは，平常時は非運用であっても1系統システム分のハードウェアリソースを必要としコスト高の要因となっているにもかかわらず，大規模災害時の動作に不安が残るため，高度な耐被災性を実現できる新IP系OSSへ移行することとなった．新IP系OSSでは，高コスト要因のバックアップシステムを廃止しつつ，低コストに，つまり系統数を増やすことなく，同時被災しない距離において冗長化することが目標となる．そこで図3のように，24時間保守体制を確立できる大阪拠点へSBY系を移動して東京～大阪間で遠隔冗長構成を取り，両系同時運用を可能とする2拠点1システム化方式[1]を採用することとした．

この新IP系OSSの構成と動作について，図4を用いて旧システムと対比しながら解説する．新IP系OSSでは，専用で用意された広帯域（数十Gbps）の東西拠点間伝送ネットワークを介して両拠点システム間で遠隔クラスタを組み，キープアライブ*5を含む拠点間システムメッセージの送受信や，両拠点システム冗長化のためのデータベース同期を行っている（図4上段）．

平常時には東西各拠点をACT/SBY*6動作させる（厳密には機能サーバ単位でACT/SBYが切り替わるが説明簡略化のため，ここでは全機能サーバを片方の拠点へ寄せると仮定して説明を行う）．NEは常に両方の拠点システムへ監視メッセージを送信するようにし，被災による障害発生直後のルーティング混乱期を含め，いかなる状況においても監視メッセージを取りこぼさないよう考慮されている（図5）．これらNEからの監視メッセージはACT側で処理され，処理結果はSBY側へ同期される．

①1つめの大規模災害ケースの拠点壊滅が発生した場合，旧IP系OSSでは監視不能に陥るのに対し，新IP系OSSではACT～SBY間のキープアライブ受信不可を契機に即座にSBY側がACT状態へ移行し監視を継続できる（図4中段）．

②2つめのケースの東西拠点システム間伝送用ネットワーク分断発生時は，旧IP系OSSでは分断ポイント以遠のNEが監視不能に陥る一方，新IP系OSSではキープアライブによる死活確認に失敗すると，SBY状態の西日本拠点システムがACT動作を開始し，両拠点システムが同時ACT動作を行うことで，全国のNEの監視を継続することが可能である．東西間ネットワークは完全に分断されているため，東西各拠点システムはそれぞれの管内のNEのみを監視することとなる（図4下段）．

なお，この2拠点同時ACT動作を導入したことにより新たな課題が発生した．新IP系OSSでのネットワーク分断時動作パターンにおいて，運良くNE収容ネットワークが分断を免れた場合，これを経由

*5 **キープアライブ**：ネットワーク上で装置間の接続が有効であることを確認するために定期的に行われる通信のこと．

*6 **ACT/SBY**：同一機能を搭載したサーバを2台設置し，1台を運用し（ACT系），残りの1台を待機状態（SBY系）にすること．ACT系サーバ故障時に即座に待機系に処理を引き継ぐことでサービス停止を防止するシステム構成．なお，SBY系サーバは切替えに備えて，常時ACT系サーバと同じ状態を保持している．

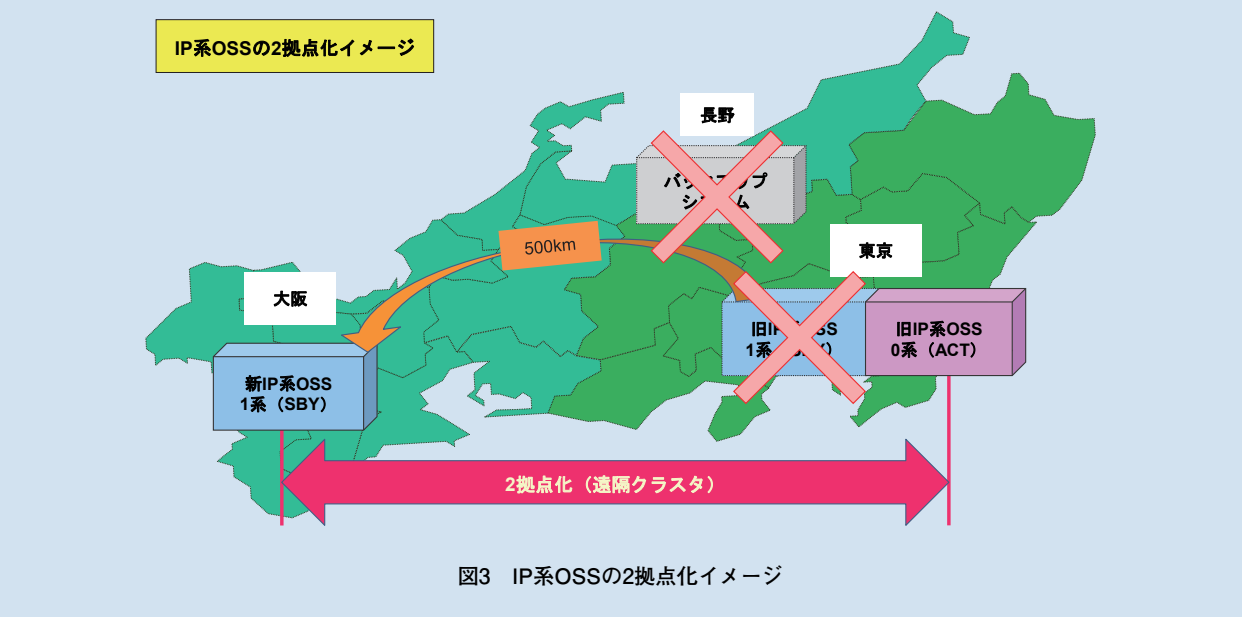

図3　IP系OSSの2拠点化イメージ

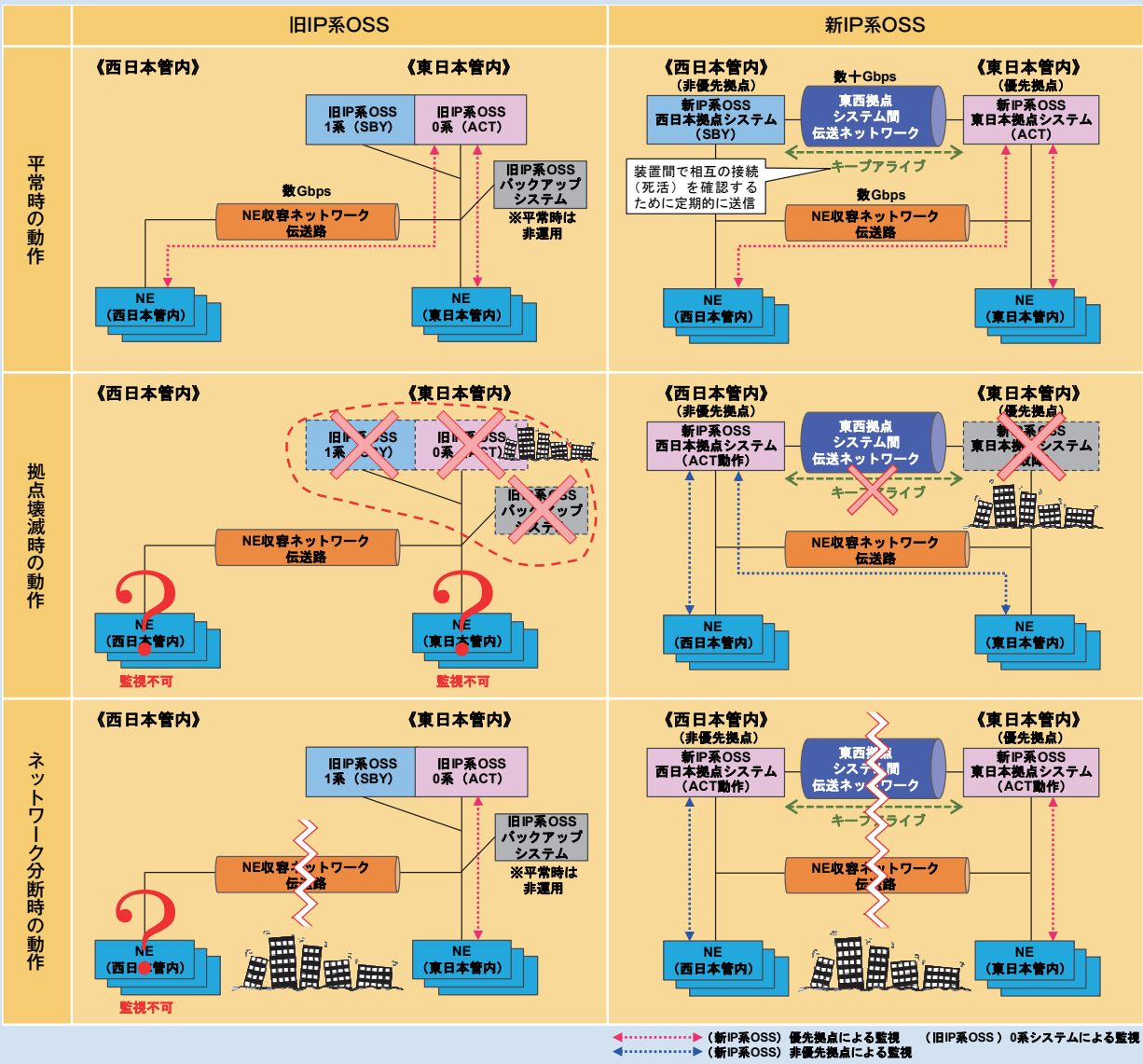

図4　新旧IP系OSSのシステム構成と災害時動作

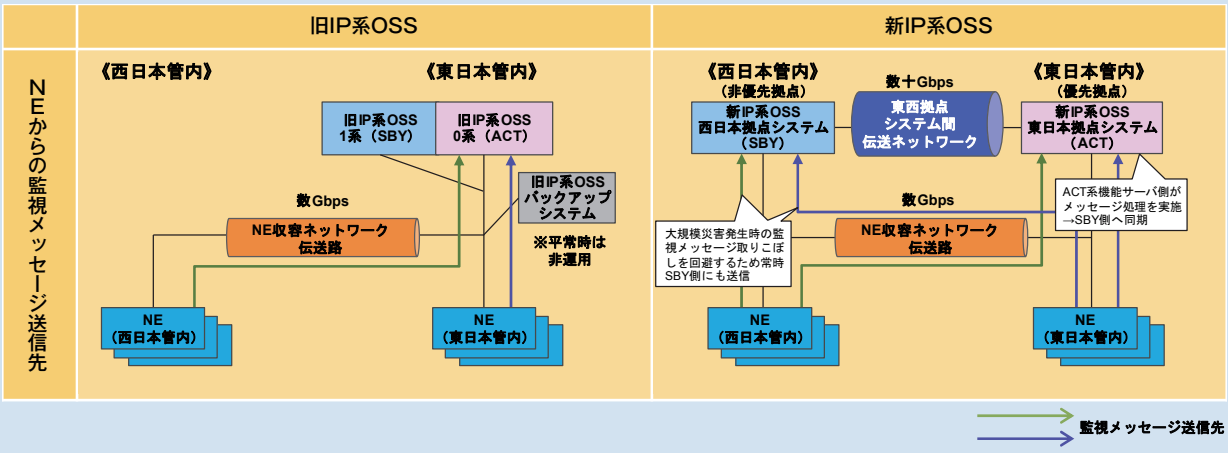

図5　NEから新旧IP系OSSへの監視メッセージ送信動作

してNEは両拠点システムからの二重監視を受け，監視処理負荷が倍増する（図6(a)）．ほとんどの機器においては大きな影響はなかったが，膨大な数の監視ポイントをもつ超大型ルータにおいて，正常動作

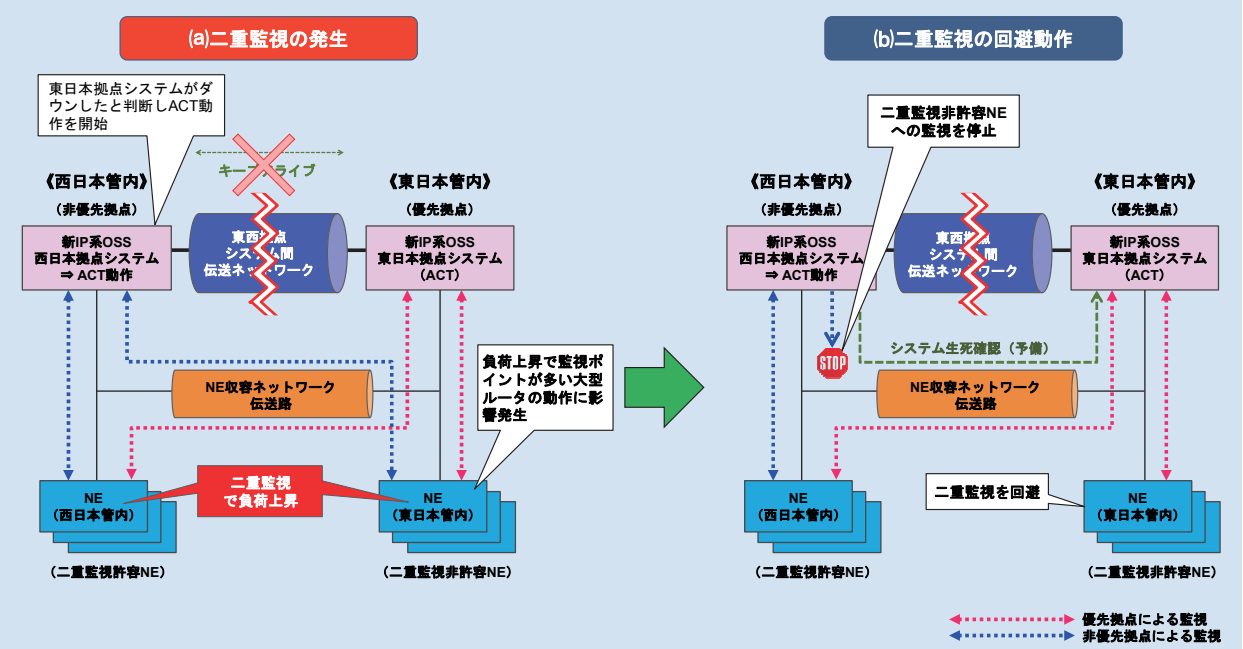

図6　新IP系OSSにおける二重監視の回避動作

を妨げるほどの深刻な影響が見られた．超大型ルータは処理性能に余裕があると思われがちであるだけに注意が必要である．新IP系OSSでは本課題に対して，拠点システムがSBYからACT状態へ移行した時点で，念のためNE収容ネットワーク経由で追加の生死確認を行い，2拠点同時ACT状態を検出した場合は，非優先拠点側において，二重監視非許容NEへの監視をストップする機能を追加することで対策を行った（図6(b)）．

本稿では，新IP系OSSの構成と動作について解説した．本システムは2016年2月に商用導入され運用を開始しており，大規模災害時でも安定した監視を期待できる．今後の課題としては，現行では一部手動となっている2拠点同時ACT動作解消後のデータベースマージ処理を自動化していきたい．

文　献

[1]　高橋　和秀：“ディザスタリカバリを経済的に実現するOSS構成技術，”電子情報通信学会技術研究報告，NS，ネットワークシステム112（392），pp.61-66，Jan. 2013.