

東京エレクトロン デバイス

メディアエクステンジが顧客向けハウジングサービスの コアスイッチとして「BlackDiamond 12804R」を採用

停止することが許されない データセンターネットワーク

多くの企業の中核となるデータセンターにおいて、サーバ群を収容するユーザースイッチの上位に位置するコアスイッチは、全ての通信が行き来するところであり、ミッションクリティカルなシステムをハウジングする企業にとって、ネットワークのダウンはビジネスの停滞を招くことになりかねない。そこで、ダウンを防ぐためコアスイッチを二重化し、高可用性のネットワークを構築することが多いが、それだけの対策では不十分である。確かにハードウェア障害に対しては二重化しているおかげで通信が止まることはないが、ソフトウェアに対しては不安である。例えば、早急にパッチを適用しなければならぬ場面や新しいサービス開始のためにソフトウェアをアップグレードする場合、ユーザーのことを考えると「通信を止めずにソフトウェアを変更したい」とシステム管理者は思うはずである。今回メディアエクステンジ (MEX) が採用したエクストリーム ネットワークス社のL3スイッチ「BlackDiamond 12804R (BD12R)」は、このような課題を解決した。

ネットワークに求めた大容量、高速化、安定性

1997年の創業以来、インターネット関連のサービス事業を提供しているMEXでは、システム要件を満たしながら安定性を提供するバックボーン構築を最重要課題として取り組んでいる。そのため、リプレース時などの製品選択における機器検証に対しては厳しい対応を図ってきた。MEXの取締役最高技術責任者 兼最高情報責任者 兼技術部長の高田寛氏は、当初、エクストリーム

ネットワークス社の製品について、必ずしも評価は高くなかったと語っている。

MEXでは、一時は特定の製品やベンダーに偏ってきたが、故障した時のリスクが大きいと判断し、再度マルチベンダーを採用する方針に変更した。またその頃、顧客が提供するサービス自体の多様化かつ大容量化により、広帯域の接続サービスを求める声も高まっていたため、10Gビットイーサネット接続サービスを開始した。しかし、障害からの復旧が早いことで採用していた顧客収容ルータが販売終了という新たな問題が持ち上がり、新しい製品の検討をした結果、東京エレクトロン デバイス (TED) が提供しているエクストリーム ネットワークス社のBD12Rの採用に至った。

ゼロパケットロスを実現する機能

MEXが採用したBD12Rは「マネジメントモジュール (MSM)」、「電源」、「冷却ファン」の全てのコンポーネントが冗長化されており、SPOF (Single Point of Failure) が発生しても通信を継続することができるようになっている。

さらにBD12Rの最大の特長は、ゼロパケットロスを実現する機能、ヒットレスフェールオーバーが搭載されていることである。この機能をうまく利用することで、ソフトウェアをヒットレスでアップグレードすることが可能になる。ヒットレスフェールオーバーは、コントロールプレーン機能とデータプレーン機能の両方をゼロパケットロスの状態でスタンバイシステムにフェールオーバーさせることができる。データプレーンでは、実際のパケットをハードウェア処理するためのキャッシュ化されたフォワーディングテーブルを持ち、コントロールプ

レーンではルーティングプロセスの実行、プロトコル単位での独自テーブル（例：OSPFネイバーテーブル等）などを管理している。

一般的に、コアスイッチでは、標準化された冗長化プロトコル（VRRP）やベンダー独自の冗長化プロトコル（ESRP等）、OSPFなどのダイナミックルーティングプロトコルを使用することが多いが、ヒットレスフェールオーバーは、このようなプロトコルで持つステータステーブルやネイバーテーブルなどのコントロールプレーンで管理するテーブルを保持し続けながらスタンバイシステムに引き継ぐことができる。

真のヒットレスソフトウェアアップグレード

ヒットレスアップグレードは、ソフトウェアの不具合を改修した、もしくは新しい機能を搭載したソフトウェア（ExtremeXOS）を、通信を止めずにスイッチに適用したい場合に威力を発揮する機能である。前述したように、ヒットレスフェールオーバー機能を利用して（MSMを切り替えて）実現する。

例えば、図1の高信頼システム構成例のような環境下でコアスイッチのソフトウェアを通信を止めずにアップグレードしたい場合は、OSPFエリア内でグレースフルOSPFリスタート機能を使用すれば、MSMの切り替わりが発生してもBD12RのOSPFネイバーテーブルやOSPF経路情報がクリアされることはない。また、VRRPのステータステーブル（Master/Backup）もクリ

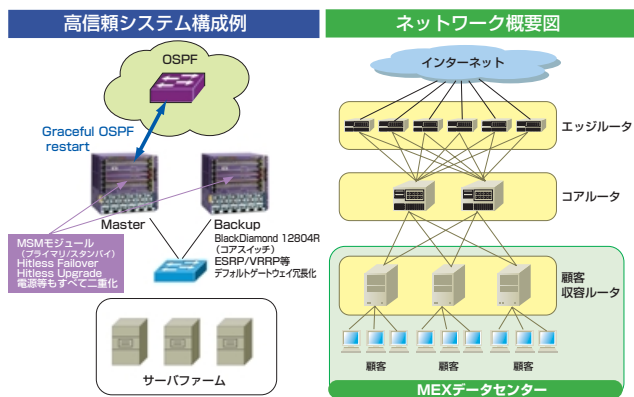


図1 高信頼システム構成例（左）とネットワーク概要図（右）

アされることはないので、冗長化されたコアスイッチ間でのバタつきによるパケットドロップは発生しない。

このようなコントロールプレーンで管理するテーブル情報がそのままスタンバイMSMに引き継がれるので、ヒットレスでソフトウェアをアップグレードすることができ、24時間365日のノンストップサービスを提供することができるようになる。

製品の長期にわたる念入りの検証を実施

MEXでは、TEDにBD12Rの貸し出しを依頼し、次のような3つの評価ポイントを設けて検証を実施した。

◆ポイント1 ネットワークが障害時に瞬時に切り替わるか？：カタログスペックにあることが実際できるかどうか。例えば、記載されている帯域の確保、経路制御の正常動作、冗長構成の正しい作動など。BD12Rで特に注目したのは、冗長構成で障害時の切り替えの際、パケットロスなくマネジメント機能を瞬時に切り替える「ヒットレスフェールオーバー」と、ネットワークへの影響を最小限に抑えながらExtremeXOSをアップグレードする「ヒットレスアップグレード」だった。

◆ポイント2 想像通りの動きをするか？：エンジニアが作業したときに、想像通りに動作することは重要である。特にネットワーク機器を触る時は、障害復旧などの緊急時など、迅速な対応に迫られている場合が多くある。

◆ポイント3 ハードウェアは安定しているか？：安定性という面では、ポートが壊れた場合に迅速な交換をするためにボードの抜き差しが容易であることも重要である。エクストリームネットワークス社の製品は、エアフローの設計を改良し、ファン制御による省エネ設計を実現している。

2007年12月、以上のような厳しい検証要件を満たしたBD12Rは、導入してから技術部の基幹スイッチとして安定稼働を続けている。MEXでは手順書などが作成でき次第、ハウジングサービスの顧客収容ルータとして正式採用を予定している。

東京エレクトロン デバイス(株) CN事業本部
ネットワーク事業部 TEL：03-5908-1962