

「止まらないシステム」を実現する アラクサラネットワークスのフォールト・トレラント・ネットワーク

1. はじめに

近年、ネットワークを利用したサービスの拡大・成長は目覚ましく、従来型の通信サービス以外にも、政府や自治体の提供する公共系サービス、ネットバンキングやオンライン決済といった金融系サービス、さらには企業の情報システムなど、あらゆるサービスがネットワーク上に構築されるようになってきています。このように、ネットワークの社会インフラとしての重要性は増し続けており、24時間365日の連続運用を可能にする「止まらないシステム」へのニーズが高まってきています。

このようなニーズに応えるネットワークを構築するためには、STPやVRRP、OSPFといったプロトコルを組み合わせるシステムを冗長化するのが一般的です。しかし、この手法ではネットワークが複雑になりすぎて、逆にトラブルの原因となってしまうケースが増えてきています。

アラクサラネットワークスの提案するフォールト・トレラント・ネットワーク (FTN : Fault Tolerant Network) は、この「複雑化によるトラブル」を解消するための新しいアプローチです。FTネットワークによってネットワークをシンプルに構成し、障害の影響範囲を局所化することができます。これは、電気・ガス・電話などの社会インフラと同じ考え方に基づいており、安定かつ確実な「止まらないシステム」の実現を可能にしています。

2. フォールト・トレラント・ネットワーク

IPネットワークの世界では、パケットは経路情報(ルーティング・テーブル)に従って転送されます。つまり、経路情報にエントリが存在しないネットワークは、外部との通信が一切できなくなってしまい、ネットワーク的に陸の孤島状態となるわけです。したがって、経路情報を適切な状態に保ち続けることは、IPネットワークにとって非常に重要なこととなります。

この経路情報を管理するためには、OSPFやBGPなどのダイナミック・ルーティング・プロトコルを使うのが一般的です。ダイナミック・ルーティングでは、個々の装置がお互いの情報をやり取りし、同じアルゴリズムに従って経路情報を計算することで、ネットワーク全体の経路情報を自動構成します。ダイナミック・ルーティングによる経路情報管理には、多くのメリットがあります。

- 全装置の経路情報を、手動設定する必要がなくなる
- 新規ネットワークの追加や、廃止ネットワークの削除が簡単にできる
- 障害時の迂回経路選択を自動化できる

しかし一方で、ネットワーク全体で情報を自動管理するという仕組みのため、デメリットも存在します。それは、ネットワーク内で障害が発生した場合、その周辺の装置でも自律的に経路情報の再計算を行うため、障害の影響が広範囲に広がってしまうということです。通常、この処理はいずれ収束していきませんが、ネットワークの規模や構成、障害発生箇所の組み合わせなどによっては、経路の再構築が収束せずに、ネットワーク全体が不安定になることもあります。

このような「局所的な障害が、ネットワーク全体に波及してしまう」というデメリットが、昨今の複雑化したネットワークで深刻な問題となっています。そのため、ダイナ

ミック・ルーティングを使用した上で、ネットワークを安定化させるような何らかの仕組みが必要とされているのです。

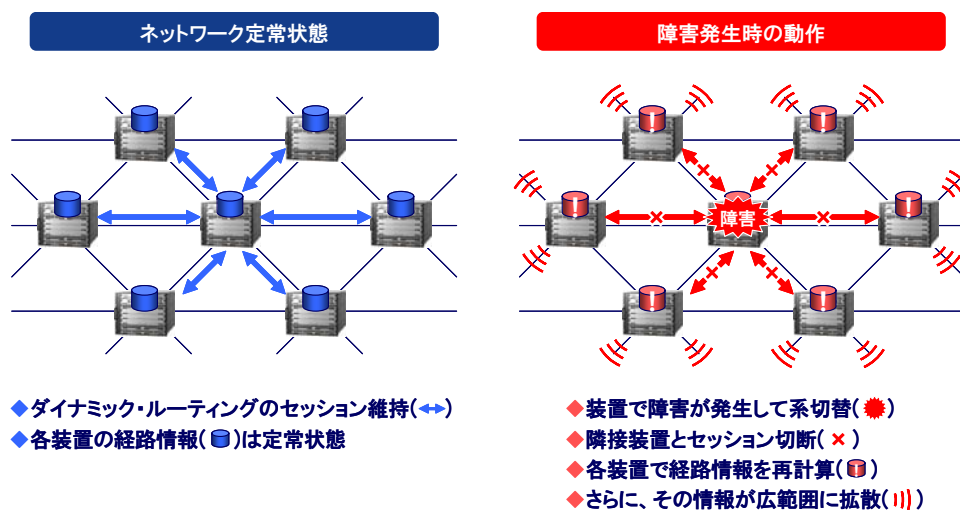


図 2-1 IP ネットワークが抱えている問題

同じような問題は、他の場面でも見ることができます。ブロードキャスト・ドメイン内の冗長を考えた場合、一般的にはL2経路をSTP系のプロトコルによって制御し、ネットワークの出口であるデフォルト・ゲートウェイをVRRPなどで冗長化します。したがって、各プロトコルを正常に動作させた上で、複数のプロトコルを組み合わせるという連携動作まで気をつけなくてはなりません。またSTPを使用することによって、ループ障害に陥ってしまうリスクを抱えこむことになり、最悪の場合はネットワーク全体がダウンしてしまうこともあります。

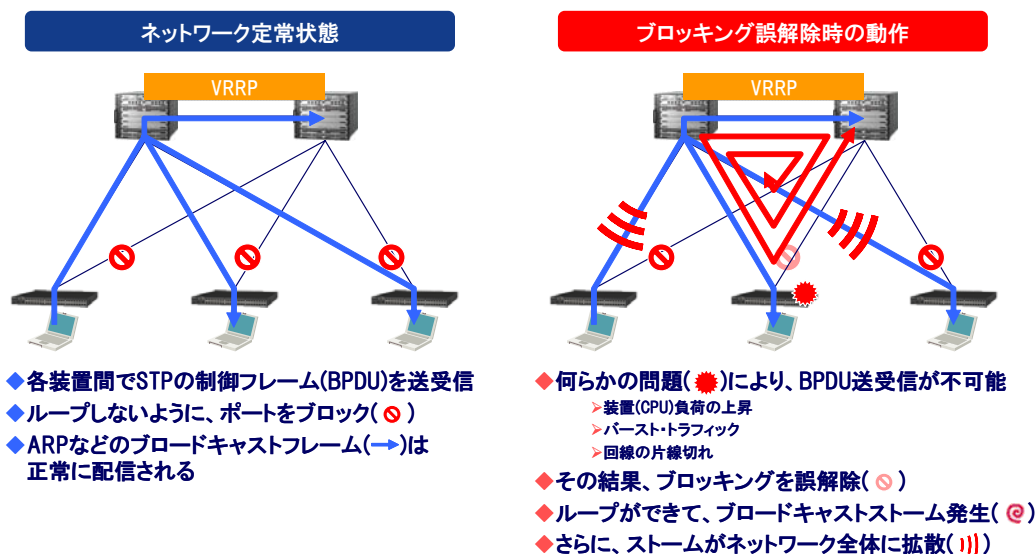


図 2-2 STP の問題点

これらのリスクに対し、「障害の影響範囲を局所化して高速に切り替えることで、システム全体の安定性を高める」というコンセプトのFTネットワークを適用することで、IPネットワークが抱えている問題を解決できます。

従来のダイナミック・ルーティングで制御されるネットワークに対して、システム全体を安定させる**グレースフル・リスタート(Graceful Restart)**と、装置(スイッチ)自体の障害耐性を高めた**フォールト・トレラント・スイッチ(FTS: Fault Tolerant Switch)**という、2つの要素を追加することでFTネットワークは構成されます。

- **グレースフル・リスタート【障害の局所化】**
障害時に発生する経路情報の再計算処理を、障害発生装置のみに限定化(局所化)する機能。障害の影響を他の装置に波及させないことで、システムの安定化を図ることができる。
- **FTスイッチ【障害箇所の高速切替】**
装置2台の機能を1台に実装することで、装置単体の可用性を高めたスイッチ。装置内での高速・確実な切り替えにより、通信に影響を与えず瞬時復旧する。STPやVRRPなどの、装置や回線を冗長化するためのプロトコルが不要になるため、システムをシンプル化できる。

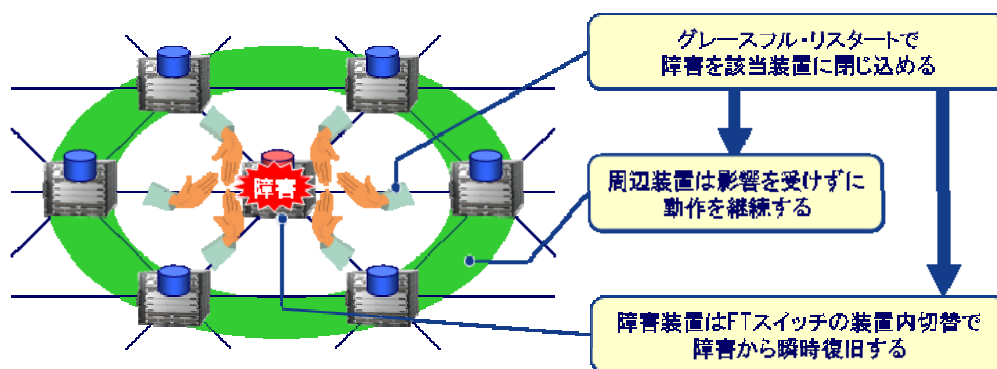


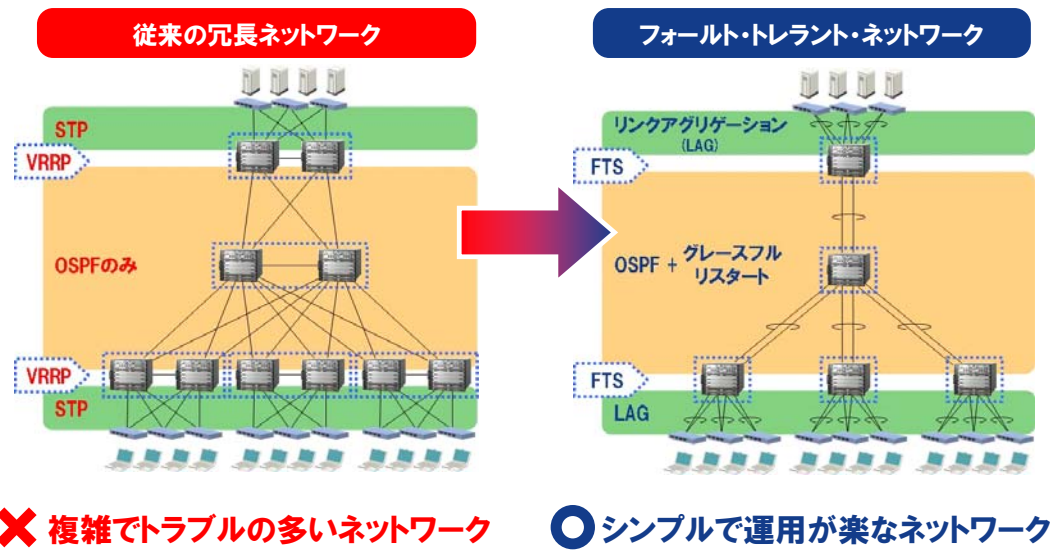
図 2-3 フォールト・トレラント・ネットワークの概要

フォールト・トレラント(耐障害性)という名称が意味する通り、障害が起きても動作を継続し続けることができ、以下のようなメリットがあります。

- グレースフル・リスタートにより、ダイナミック・ルーティングのメリットを活かしながら、ネットワーク全体を安定化できる
- 2台の機能を1台に実装しているので、装置台数を削減できる上に、冗長化の面倒な設定やチューニングをしなくても高速・確実に切り替わる
- 回線の冗長化はSTPではなくリンクアグリゲーションによって行うため、ループ障害から解放される
- ゲートウェイ冗長プロトコル(VRRPなど)が不要になるため、装置のバタツキ障害(フラッピング)から解放される
- 装置台数と使用プロトコルが減少するので、ネットワークがシンプルになって運用管理コストを低減できる
- 無停止バージョンアップが可能
- モジュール交換作業時も通信が止まらない(無停止保守が可能)

以上の通り、シンプルさや高可用性を備えたFTネットワークは、ミッションクリティカル性を求められるシステムに適しています。

このホワイトペーパーでは、FTネットワークを構成するグレースフル・リスタートとFTスイッチという2つの要素を中心に、その考え方や特徴などを解説します。



✗ 複雑でトラブルの多いネットワーク **○ シンプルで運用が楽なネットワーク**

図 2-4 従来の冗長ネットワークとフォールト・トレラント・ネットワークの比較

3. グレースフル・リスタートによる障害の局所化

前述の通り、IP ネットワークにおけるシステム全体の安定化とは、すなわち経路情報（ルーティング・テーブル）の安定化に他なりません。

しかし、ここで大きな壁にぶつかります。それは、ダイナミック・ルーティングによる自動管理を採用すれば、経路情報がダイナミックに変化するのは必然のこととなるため、安定化からは遠のいてしまうという壁です。この壁を乗り越える方法として、スタティック・ルーティングへの回帰が真っ先に思い浮かびますが、大規模化・複雑化した現在のネットワークでは現実的ではありません。

この壁を乗り越えるためには、FT ネットワークの構成要素であるグレースフル・リスタートの併用が有効です。グレースフル・リスタートは障害の影響が回りの装置に波及するのを防ぐ技術で、障害を局所化してシステム全体を安定化できます。

「図 2-1 IPネットワークが抱えている問題」のようなダイナミック・ルーティングのみのネットワークでは、装置間でお互いの状態情報を交換し、各装置で最適な経路を計算することで全体の経路情報を自律的に構築しています。このネットワーク内の 1 台の装置で障害が発生すると、周辺装置はセッションの切断を検出します。その結果「隣の装置に関係する経路が使えなくなった」と判断し、経路の再計算処理を開始します。さらに、この情報はネットワークを伝播して広がっていきます。

通常は、経路情報の再計算はいずれ収束して、ネットワークは安定な状態に復旧しますが、復旧するまでの間は通信が途切れてしまう場所もあります。また、ネットワークの規模や構成、障害発生箇所の組み合わせなどによっては、特定の装置に再計算処理が集中して、復旧時間が長くなることもあります。さらに、その負荷が装置の処理能力を超えてしまうと、最悪の場合には、障害とは関係ないはずの装置がダウンするという 2 次障害に発展してしまうこともあります。

これに対し、グレースフル・リスタートを併用したネットワークで障害が発生した場合、障害装置から周辺装置に対してリスタート通知が送信されます。先ほどと同様に、周辺装置はセッションの切断を検出しますが、リスタート通知を受信している場合は「何もなかったかのように」経路情報を維持し続けます。結果として、障害装置以外では経路情報の再計算処理は発生しないので、通信が途切れることもありません。リスタート通知とはすなわち「すぐに復活するから、今の状態のままで転送を続けてください」という依頼であり、その依頼通りに周辺装置が協力することで、ネットワーク全体の安定性を保つことができるのです。

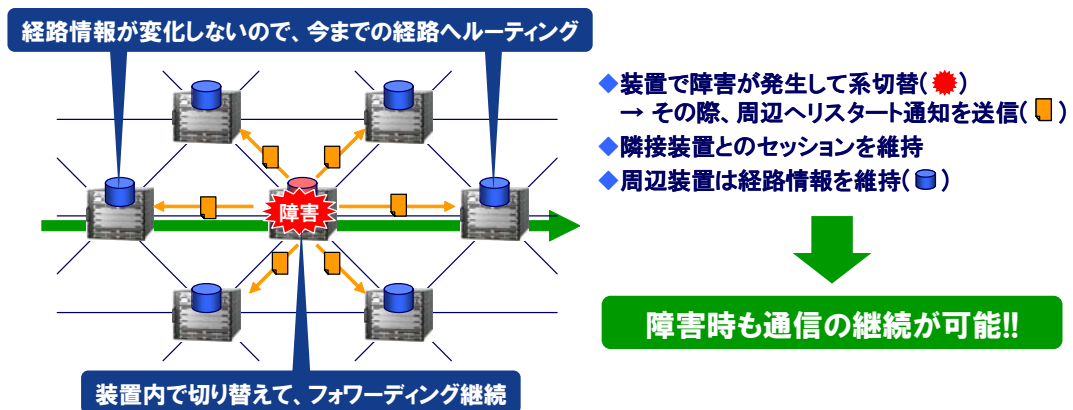


図 3-1 グレースフル・リスタートを併用したネットワーク

このときに注意しなければならないのは、装置の耐障害性です。グレースフル・リスタートによって周辺装置の経路情報は維持されるので、それまで障害装置を経由していた通信は、迂回することなく障害装置へ転送され続けます。したがって、障害装置が転送可能な状態に復帰するまでが通信断時間となるため、装置の切替時間がサービス停止時間として見えてくることとなります。IP ネットワークのインフラとしての重要性を考えた場合、この時間は短ければ短いほど良いということになります。

このようなニーズに応えることができるのが、FT スイッチである AX6700S です。最短 50msec での高速切替を確実に実施できる AX6700S であれば、障害発生時でも瞬時に復旧して通信を継続することができます。そこで次章では、この FT スイッチについて解説したいと思います。

- ※ グレースフル・リスタートは以下の RFC および IETF ドラフトに準拠しています
 - OSPF : RFC2370, RFC3623
 - draft-kompella-ospf-opaquev2-00.txt
 - draft-ietf-ospf-ospfv3-graceful-restart-04.txt
 - BGP : draft-ietf-idr-restart-13.txt

4. フォールト・トレラント・スイッチで障害箇所を高速切替

FT ネットワークのもう一つの構成要素は、装置(スイッチ)自体の障害耐性を高めた FT スイッチの AX6700S です。グレースフル・リスタートによって局所化した障害から、FT スイッチの高速切替によって復帰します。

FT ネットワークがシステムとしての耐障害性コンセプトなのに対し、FT スイッチは装置としての耐障害性コンセプトです。そのため、基本的な考え方はサーバやストレージと通じています。すなわち、装置内部で主要部品を冗長化・同期運転することで、装置障害時でも動作を継続するという考え方です。サーバであれば FT サーバが、ストレージであれば RAID が、そしてネットワーク装置であれば FT スイッチが、これに該当します。さらに、最高レベルの可用性を持った交換機の技術エッセンスを取り入れることで、従来の LAN スイッチを大きく上回る高可用性を実現しています。

FT スイッチを実現するための技術をフォールト・トレラント・アーキテクチャ (FTA : Fault Tolerant Architecture) と呼び、「①完全冗長化」「②マイクロ・モジュール」「③クロス・コネクション」「④ハードウェア・ミラーリング」という 4 つの技術コンポーネントで構成されています。

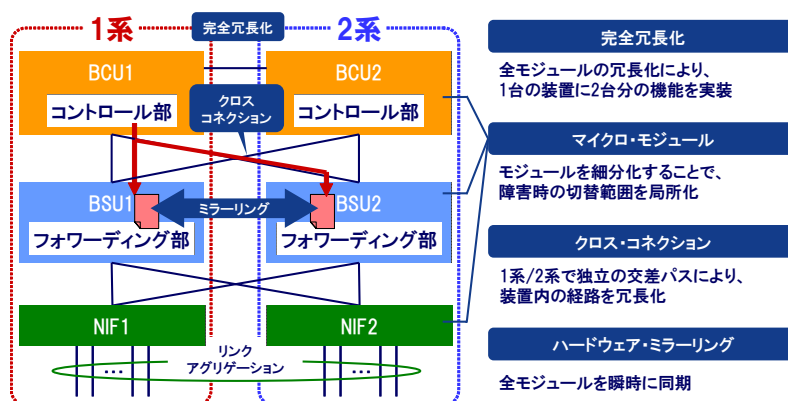


図 4-1 フォールト・トレラント・アーキテクチャ

また、隣接装置との接続はリンクアグリゲーションによって冗長化するため、L2 経路制御プロトコルである STP が不要になります。したがって、STP 使用時に大きな懸念事項となる、ループ障害のリスクを考慮する必要がなくなるため、ネットワークの安定化という面でも非常に大きな効果があります。



図 4-2 STP とリンクアグリゲーションの特徴比較

4.1. 主要部品の完全冗長化

シャーシ型(モジュラー型)スイッチの AX6700S では、同一モジュールを複数搭載することで、ハードウェアの冗長化が可能になります。モジュールの種類は下表の通りで、全てのモジュールが冗長化可能です。

表 4-1 AX6700S のモジュール一覧

#	モジュール種別	概要	最大搭載数
1	BCU (Basic management Control module)	コントロールエンジン CPU 等が搭載され、ソフトウェアが動作します	2
2	BSU (Basic packet Switching module)	フォワーディングエンジン パケット転送処理を行います	3
3	NIF (Network Interface module)	回線(インタフェース)を 収容します	8
4	PS (Power Supply)	電源の供給を行います	8(AC) 4(DC)
5	FAN (FAN module)	1つのモジュールに、 3個のファンが実装されています	4

また以下の通り、BCU/BSU/NIF は装置前面に、PS と FAN は装置背面に実装する構造となっています。

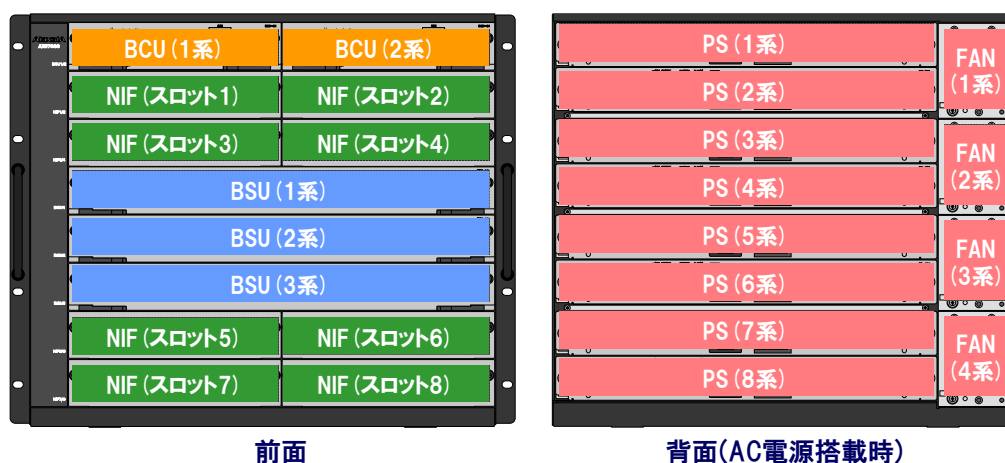


図 4-3 AX6700S モジュール搭載位置

一般のシャーシ型スイッチでも、複数モジュールを搭載することで装置のハードウェア冗長性を確保できるように考えられますが、実際にはそれだけでは十分とはいえません。何故ならば、モジュール化されていない部品については単一障害点(SPOF: Single Point of Failure)としてのリスクを抱えているためです。この問題に対し、AX6700S は主要部品の冗長設計を徹底することで、SPOF リスクの低減を図っています。具体的には、筐体(シャーシ)自体を除くほとんどの主要部品を完全に冗長化することで、SPOF リスクを限りなくゼロに近づけています。

また、NIF には BCU/BSU のような運用系(ACT)/待機系(SBY)という概念がありませんが、同一のモジュールを2枚搭載した上で、NIF 間でのリンクアグリゲーションを設定することで冗長化が可能です。

このように、AX6700S は主要部品について完全冗長化設計を行っているため、通常の LAN スイッチを大幅に上回る信頼性を、装置単体で達成できているのです。

4.2. マイクロ・モジュール

一般的に、モジュール化の単位を細分化すると、障害発生時の切替単位が小さくなるため、切替時間の短縮や障害影響範囲の局所化、メンテナンス性の向上といった様々なメリットを得ることができます。その一方で、コストアップというデメリットもあるため、細分化の徹底が最良の選択であるとは限りません。信頼性重視か、コストパフォーマンス重視かで、モジュール分割の方針は全く異なってきます。

これをまとめたのが表 4-2 ネットワーク装置のモジュール分割方針です。通常のLANスイッチは2モジュール分割であることが多いのに対し、信頼性を重視した交換機の場合は4モジュール分割となっています。そこでAX6700Sは3モジュール分割というマイクロ・モジュール(MM：Micro Module)方式を採用し、通常のLANスイッチ以上に信頼性を重視した設計を行っています。

表 4-2 ネットワーク装置のモジュール分割方針

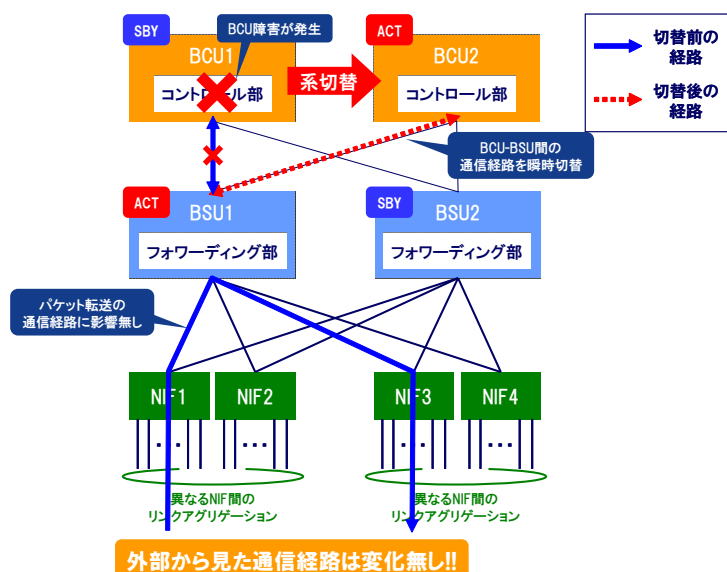
装置種別	転送系のモジュール分割方針
交換機	CPU/スイッチ/メモリ/インタフェースの4モジュール
AX6700S	CPU/(スイッチ+メモリ)/インタフェースの3モジュール BCU → CPU / BSU → スイッチ+メモリ / NIF → インタフェース
通常のLANスイッチ	(CPU+スイッチ+メモリ)/インタフェースの2モジュール

4.3. クロス・コネクション

交換機では交絡(交差)接続という方式が採用されており、モジュール同士が専用経路で相互接続されています。これはクロス・コネクション(CC：Cross Connection)としてAX6700Sにも採用されており、大きな特徴となっています。

クロス・コネクションによって、各モジュール間の通信経路が冗長化されるため、どのモジュールで障害が発生しても、通信の継続が可能になります。

- BCU(コントロール部)障害時の通信経路
クロス・コネクションで、BCU-BSU間の通信経路が切り替わります。パケット転送の通信経路(BSU-NIF間)には影響が無く、外部との通信経路は変化しません。



- BSU(フォワーディング部)障害時の通信経路
クロス・コネクションで、BCU-BSU/BSU-NIF 間の通信経路が切り替わります。内部的にはパケット転送の通信経路が変化しますが、外部との通信経路は変化しません。

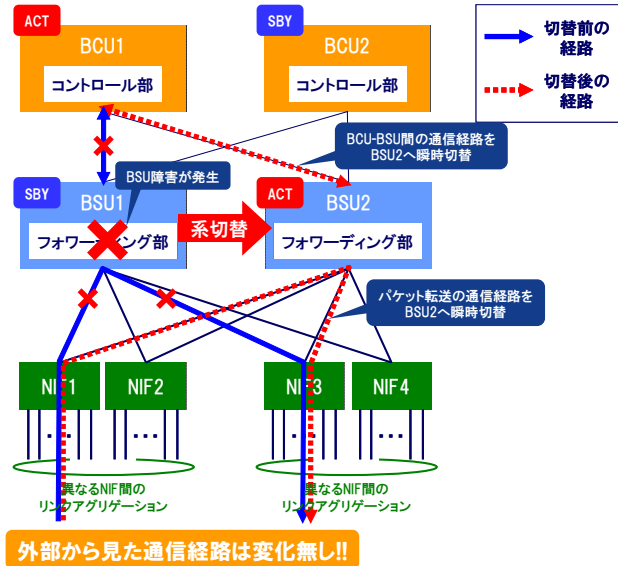


図 4-5 BSU 障害時の経路切替

- NIF/回線/トランシーバ(回線収容部)障害時の通信経路
クロス・コネクションで、BSU-NIF 間の通信経路が切り替わります。BCU-BSU 間の通信経路には影響ありません。内部的にはパケット転送の通信経路が変化しますが、異なる NIF 間のリンクアグリゲーションが縮退するだけで、外部との論理的な通信経路は変化しません。

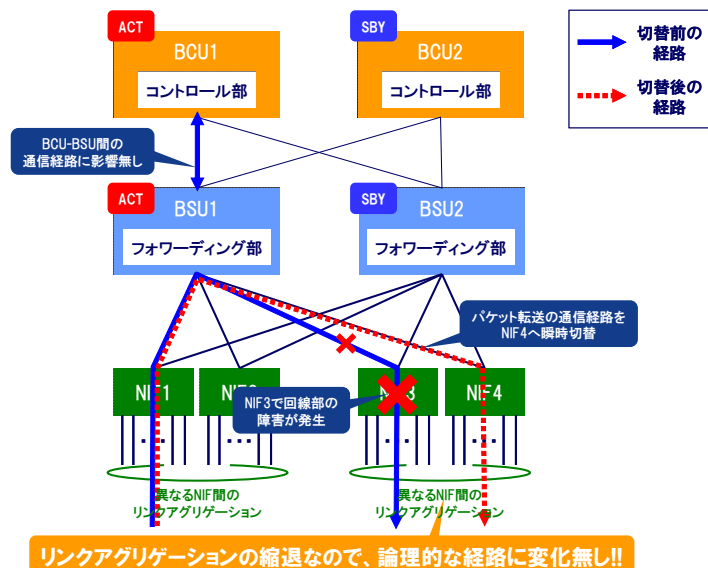


図 4-6 NIF/回線/トランシーバ障害時の経路切替

4.4. ハードウェア・ミラーリング

一般的に、LAN スイッチでハードウェアによる高速転送を行うためには、パケット転送を行うモジュール(フォワーディング部)にフォワーディング・テーブル(転送先情報)を設定する必要があります。フォワーディング・テーブルはコントロール部で動作しているソフトウェアによって自動計算された後、全てのフォワーディング部に対して同じ情報が設定されます。

通常の LAN スイッチで全てのフォワーディング部へ設定を行う場合、運用系(ACT)のソフトウェアから待機系(SBY)のコントロール部に情報を送信し、実際の設定は待機系(SBY)のソフトウェアが実施することになります。これをソフトウェア方式と呼びます。

これに対して AX6700S では、クロス・コネクションを使って、全てのフォワーディング部へのダイレクト設定が可能なハードウェア方式を採用しています。これをハードウェア・ミラーリング(HM: Hardware Mirroring)と呼び、ソフトウェア方式と比べると待機系ソフトウェアを中継する必要が全く無くなるため、全テーブルを瞬時に完全同期させることが可能です。

なおAX6700Sでは、フォワーディング・テーブルの同期についてはハードウェア方式を用いていますが、場合によってはソフトウェア方式も使用しています。その特徴や使い分け方の詳細については、「5.3 切替後の情報引き継ぎと再開処理」を参照してください。

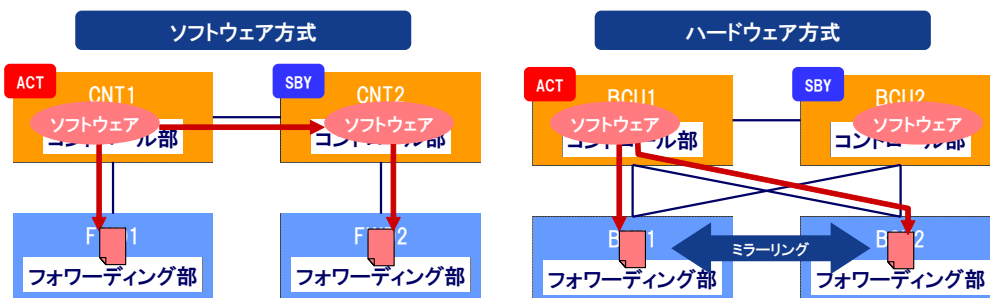


図 4-7 フォワーディング・テーブルの同期方式

5. フォールト・トレラント・アーキテクチャ以外の高可用性技術

AX6700S では、FT アーキテクチャ以外にも様々な可用性を高める技術を採用しています。ここでは、それらの技術を説明します

5.1. 確実な障害検出と切替処理の重要性

FT アーキテクチャの採用により、AX6700S はハードウェアの冗長レベルを大きく高めることに成功しました。しかし、そもそも障害の発生を検出できなかったり、切り替え動作に問題があっては意味がありません。つまり、確実に障害を検出し、確実に切り替えを実施するための仕組みが必要になります。

AX6700S では、EMA(Emergency Action circuit)と呼ばれるエマージェンシー回路をハードウェアで実装することで、どのような場合でも確実な障害検出と切替処理を実現しています。エマージェンシー(非常時、緊急事態)という名称が示す通り、装置が最悪な状態になるのを防ぐ、「最後の砦」の役割を果たしています。

例えば、カーネルパニックやデッドロックなどによってソフトウェア処理自体が不可能になった場合や、コントロール部(BCU)が突然引き抜かれるといった不正オペレーションの場合でも、EMA が障害を検出して切替を実施するので、装置が制御不能な状態になることはありません。

つまり、ソフトウェアがダメな場合に備えて、ハードウェアで救う仕組みを織り込んでいるのです。EMA というハードウェアは、AX6700S の高可用性を支える大きな柱の一つとなっています。

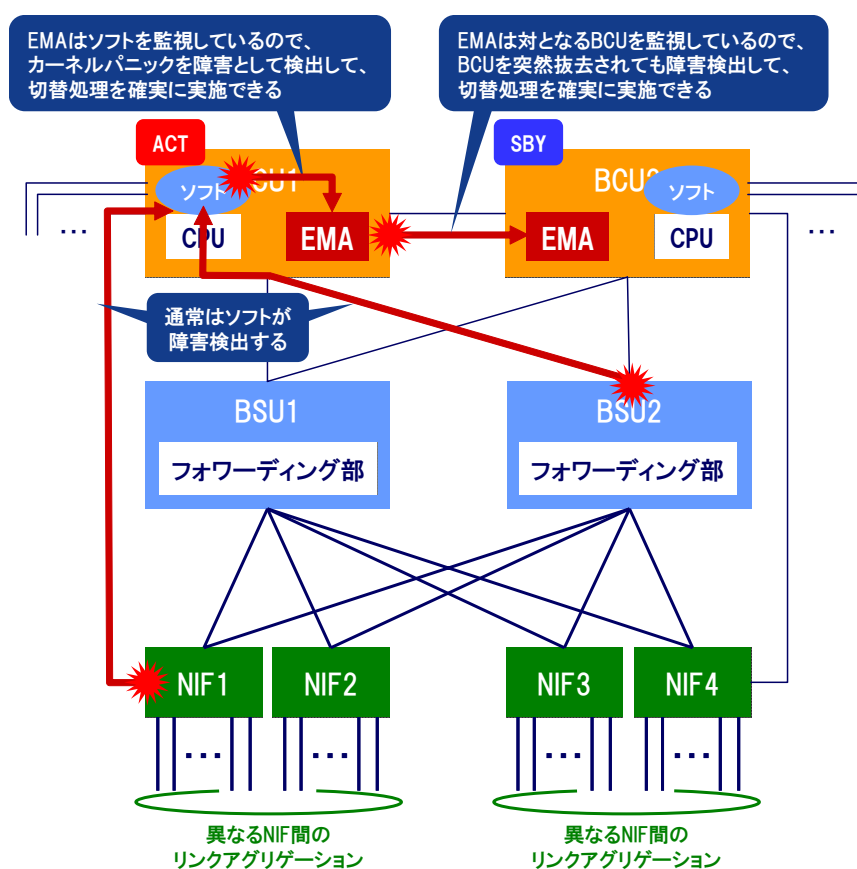


図 5-1 EMA の役割

5.2. サイレント故障対策

障害発生時に確実な切り替わりが実施されたとしても、切り替わり先のモジュールが故障しては意味がありません。このように、切り替わった先の待機系が故障していた状態をサイレント故障と呼び、特に連続運用が求められるケースでは対策する必要があります。

サイレント故障を防止するには、運用系(アクティブ)/待機系(スタンバイ)という状態に関わらず、実装している全てのモジュールについて正常性をチェックする必要があります。交換機の場合には、定期的(例えば24時間ごと)に運用系と待機系を切り替えるという機能が実装されています。アクティブな状態になれば通常の障害検出ができるので、問題を回避できるという理屈です。

対する AX6700S では、複数の手法を併用することでサイレント故障対策を行っています。

5.2.1. BSU のダブルアクト/トリプルアクト運転

複数のモジュールが同時に運用(アクティブ)状態で動作できるようになっていれば、サイレント故障から解放されます。AX6700Sはこの考えを取り入れており、複数のBSUモジュールを同時にアクティブな状態で動作させることができます。

表 5-1 BSU の運転モード

機能名称	概要	BSU 実装数	備考
シングルアクト	BSU×1 アクティブ	1 枚以上	2 枚目以後の BSU はスタンバイ
ダブルアクト	BSU×2 アクティブ	2 枚以上	3 枚目の BSU はスタンバイ
トリプルアクト	BSU×3 アクティブ	3 枚	全ての BSU がアクティブ

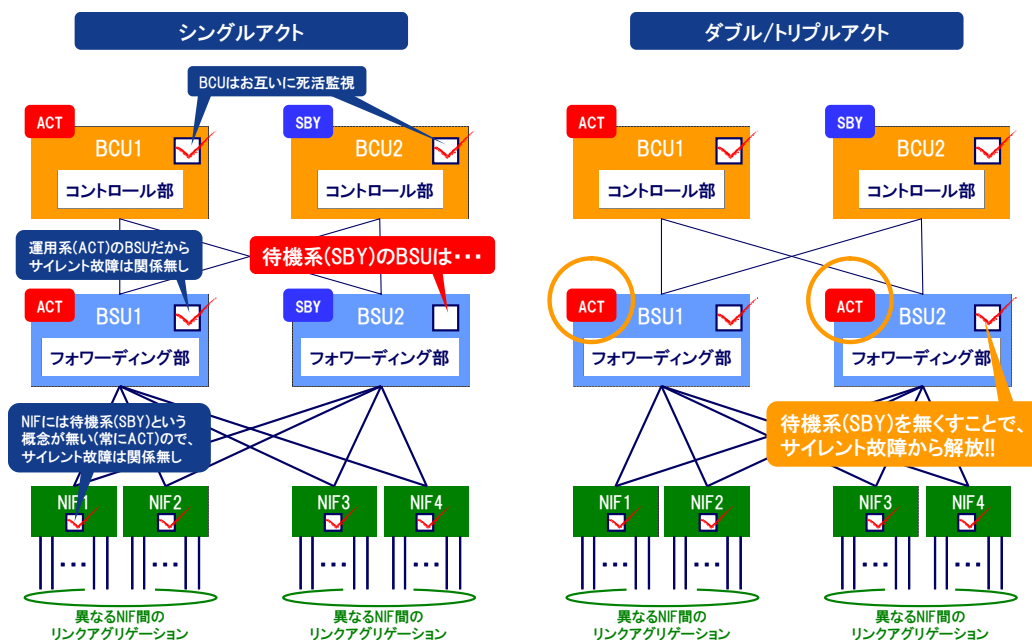


図 5-2 ダブルアクトトリプルアクトのメリット

つまり、装置に実装している BSU 全てをアクティブにして運用することで(ダブルアクトやトリプルアクト)、BSU では待機系という状態を無くしてしまうのです。その場合、BSU 障害時には縮退動作を行うこととなりますが、「障害時でも止まら

ずに動き続ける」という FT スイッチの機能は損なわれません。

また、サイレント故障とは違う観点になりますが、ダブルアクト/トリプルアクトでは複数 BSU が同時運転することになるため、負荷分散という点でも効果があります。

5.2.2. 他モジュールやバスのヘルスチェック

AX6700S では、BCU から各モジュールに対するヘルスチェックも実施しています。したがって、サイレント故障を防止することができます。具体的には、運用系ソフトと待機系ソフトの間で定期的な死活監視を行うことで、CPU のサイレント故障を防止しています。また、各モジュールへの定期的なヘルスチェックによって、待機系制御バスのサイレント故障を防止しています。

さらに、NIF には運用系 (ACT) / 待機系 (SBY) という概念が無く、常に運用状態 (ACT) となるため、BSU-NIF 間のバスはサイレント故障がありません。

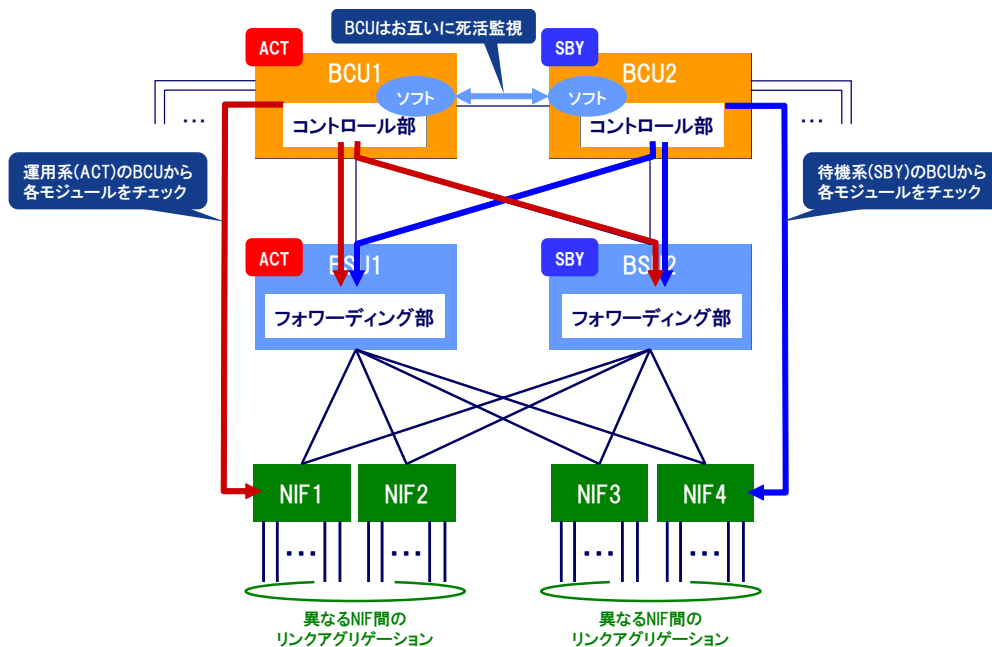


図 5-3 BCU からのヘルスチェック

5.3. 切替後の情報引き継ぎと再開処理

障害時にハードウェアを切り替えた場合でも、ソフトウェアが処理していた情報が失われてしまった場合は、サービスに影響が出てしまいます。そのため、切り替えの前後で情報を引き継いだ上で、適切な再開処理を実施しなくてはなりません。

交換機の場合、運用系と待機系をハードウェア的に完全同期させるという方式を採用しています。交換機で扱う情報(経路情報)は基本的にスタティックなため、運用系と待機系で状態を同期させるのは比較的容易なことでした。しかし、IP ネットワークの経路情報は基本的にダイナミック・ルーティングによって管理されるため、装置同士が連携することで構築されます。したがって、この情報をハードウェアで完全に同期させるのは、非常に困難なこととなります。

そこでAX6700Sでは、装置自身が持つ情報はハードウェアによって引き継ぎ、IPネットワークが持つ経路情報はグレースフル・リスタートという周辺装置との連携機能で引き継ぐ手法を採用しています。グレースフル・リスタートの詳細については、「3 グレースフル・リスタートによる障害の局所化」を参照してください。

5.3.1. BCU の情報引き継ぎ

BCU を冗長化した場合、待機系の BCU は装置管理情報やコンフィギュレーションの同期化を行い、ホットスタンバイ状態で待機します。その際、同期する情報の種類に応じて、2つの引き継ぎ方式を使い分けています。各方式の良いところを活かすことで、確実かつ高速な同期化を実現しています。使い分けの観点、データ量(大量のデータを同期するスピード)と即時性(同期完了までのオーバーヘッド)が大きなポイントとなります。

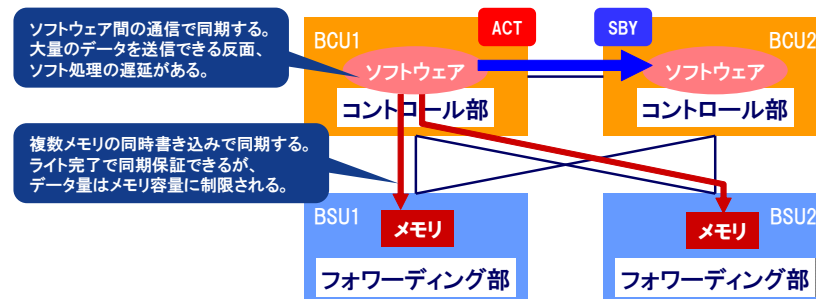


図 5-4 BCU の情報同期化

- ソフトウェア方式

BCU 同士を接続した通信経路を経由して、運用系と待機系のソフトウェア間で連続データ送信を行います。大量のデータを同期する用途に向いていますが、ソフトウェア処理のため即時性は犠牲になります。例えば、ログ情報などの同期に使用されます。

- ハードウェア方式

BSU上のメモリに情報を書き込み、系切替時にはそのメモリから情報を読み出して引き継ぎを行います。メモリに書き込むだけで同期が完了するため、即時性は高くなります。また、別モジュール上にメモリが実装されているため、BCUダウン時も情報が失われない上に、全てのBSUに同じ情報を書き込むことで冗長化しています。「4.4 ハードウェア・ミラーリング」にある通り、フォワーディング・テーブルはこの方式で同期されており、BSUとBCUの両方から参照されます。

5.3.2. BSUの情報引き継ぎ

ハードウェア方式で説明した通り、BCUから設定される情報は、全てのBSUに対して同時に書き込みされるため、常に同期状態が保たれます。詳細は「4.4 ハードウェア・ミラーリング」で説明した通りで、ハードウェア転送に不可欠なフォワーディング・テーブルを瞬時に完全同期させることが可能です。

5.3.3. NIF および回線部の情報引き継ぎ

FT スイッチでは、同一 NIF を 2 枚搭載した上で、NIF 間でのリンクアグリゲーションを設定することで、回線部を冗長化します。したがって、回線や NIF で障害が発生した場合、リンクアグリゲーションが縮退するだけで、論理的な通信経路に変更はありません。そのため、縮退時間以上の通信断が発生することはありません。

5.3.4. AX6700S の再開処理

これまで、各モジュール切替時の情報引き継ぎについて技術的な説明してきましたが、最終的にユーザに見えるのは「切替後の再開時に周囲が受ける影響」となります。そこで、以下のように再開処理レベルを定義して整理します。

- レベル 1 再開
周囲に影響を与えずに、切り替えが可能です。

表 5-2 レベル 1 再開の機能 (AX6700S)

#	分類	機能
1	ネットワークインタフェース	ポート状態(アップ/ダウン)
2		フィルタ & QoS
3		リンクアグリゲーション
4	レイヤ 2 中継	MAC アドレステーブル
5		ポート VLAN
6		プロトコル VLAN
7		MAC VLAN
8		VLAN Tag 変換
9		VLAN トンネリング
10		BPDU フォワーディング
11		EAPOL フォワーディング
12		IGMP/MLD Snooping
13	MLD snooping	
14	ネットワーク認証	IEEE802.1X 認証
15		Web 認証(ダイナミック VLAN)
16	高信頼化	IEEE802.3ah/UDLD (※1)
17	IPv4 パケット中継	ARP テーブル
18		DHCP サーバ (※1)
19	IPv4 ユニキャスト・ルーティング	スタティック・ルーティング
20		OSPF (※2)
21		BGP4 (※2)
22	IPv6 パケット中継	NDP テーブル
23		DHCP サーバ (※1)
24	IPv6 ユニキャスト・ルーティング	スタティック・ルーティング
25		OSPFv3 (※2)
26		BGP4+ (※2)
27	IPv6 マルチキャスト・ルーティング	PIM-SSM

(※1) プロトコル動作の関係から、切り替えによって通信が切れることはありません

(※2) グレースフル・リスタートを併用した場合です

詳細については「3 グレースフル・リスタートによる障害の局所化」を参照してください

- レベル 2 再開
周囲に影響があるため、通信無停止での切り替えはできませんが、機能自体の使用については特に制限はありません。

表 5-3 レベル 2 再開の機能 (AX6700S)

#	分類	機能
1	レイヤ 2 冗長プロトコル	STP (Single/PVST+/MSTP) (※1)
2		アラクサラリングプロトコル
3	ネットワーク認証	Web 認証 (固定 VLAN)
4		MAC 認証
5	高信頼化	GSRP (※1)
6		VRRP (※1)
7	IPv4 ユニキャスト・ルーティング	RIP/RIPv2 (※2)
8	IPv4 マルチキャスト・ルーティング	PIM-SM (※3)
9		PIM-SSM (※3)
10	IPv6 ユニキャスト・ルーティング	RIPng (※2)
11	IPv6 マルチキャスト・ルーティング	PIM-SM (※3)

(※1) FT ネットワークでは STP や VRRP の代わりに、
完全冗長化とリンクアグリゲーションを使います

(※2) FT ネットワークでは OSPF/OSPFv3 が推奨です

(※3) マルチキャスト・ルーティングは PIM-SSM (IPv6) が推奨です

5.4. バックアップ機能

ここまで、どのようにして装置の可用性を高めているのかを詳しく解説してきました。しかし、それらの手段によってもリカバリーできない重大な障害が発生する可能性も、考慮しなければなりません。すなわち、装置の状態を容易に復旧できるように、設定などをバックアップする機能が必要となります。

- FTP や TFTP によってサーバへ、設定ファイルを転送する
- 内部フラッシュメモリや外部メモリ (SD カード) へ、設定ファイルを保存する
- show コマンドによって、設定をコンソール画面に表示する

通常の LAN スイッチと同様に、AX6700S もこれらの標準的な手段によって設定などをバックアップすることができます。しかし、バックアップやリストア作業の際にはコマンドラインへアクセスする必要があるため、熟練したオペレータ以外にとっては、難解な作業になりがちです。また、作業手順書を準備することで作業を容易化しても、オペレーションミスの可能性はなくなりません。

この保守に関する問題は、SD カードスクリプト機能で解決することができます。これは、SD カードに作業手順を記述したスクリプトを格納しておけば、その SD カードを装置に挿すだけで作業を自動実行できるという機能です。コマンド入力の必要が無いため、装置に精通していないユーザでも、短時間で簡単・確実にバックアップやリストアなどの作業を実施することができます。

また、設定以外のソフトウェアなども一括でバックアップ/リストアすることができるので、通常の装置交換作業時に必要となるバージョン合わせ作業が不要になります。通常の運用中に SD カードスクリプトでバックアップしておけば、障害時には交換した装置へ SD カードを挿すだけで、完全に元の状態へ復帰することができます。結果として、オペレーションミスの可能性を排除し、装置が故障した場合の交換作業時間を最短にすることができます。

6. まとめ ～フォールト・トレラント・ネットワークの効果～

ネットワークが複雑になりすぎて、トラブルの原因となってしまうケースが増えているという問題を解決するために、アラクサラネットワークスは**フォールト・トレラント・ネットワーク (FTN : Fault Tolerant Network)**というコンセプトを提案してまいりました。FT ネットワークでは、複雑になりすぎたネットワークをシンプルに再構成し、障害の影響範囲を局所化することで、安定かつ確実な「止まらないシステム」を実現しています。

「複雑化によるトラブル」とはすなわち、IP ネットワークが抱えている「局所的な障害が、ネットワーク全体に波及してしまう」という問題に他なりません。FT ネットワークは、システム全体を安定させる**グレースフル・リスタート (Graceful Restart)**と、装置(スイッチ)自体の障害耐性を高めた**フォールト・トレラント・スイッチ (FTS : Fault Tolerant Switch)**という2つの要素によって、この問題に対処しています。

ミッションクリティカル用途にも適用可能な装置自体の高可用性に加え、グレースフル・リスタートとの連携によって「止まらないシステム」を構築可能な FT スイッチ AX6700S は、信頼性において他の追随を許さない製品となっています。

Alaxala

2008年6月20日 第1版発行

アラクサラネットワークス株式会社

〒212-0058

川崎市幸区鹿島田 890 番地 新川崎三井ビル西棟

URL : <http://www.alaxala.com/>

問い合わせ URL : <http://www.alaxala.com/jp/contact/index.html>