

[招待講演]広域イーサネットの運用管理技術

安藤 雅人[†]

[†] KDDI 株式会社 ソリューション商品企画本部 ネットワーク商品開発部

〒102-8460 東京都千代田区飯田橋 3-10-10 ガーデンエアタワー

E-mail: [†] mt-andou@kddi.com

あらまし 広域イーサネットサービスの一つである Powered Ethernet の運用者/設計者としての立場より、広域イーサネットにおける運用管理技術を紹介する。

キーワード 広域イーサネット, VPN, 運用管理

[Invited Talk] Operation and maintenance technology of Wide Area Ethernet

Masato ANDO[†]

[†] Development and Planning Section Data Network Product Development Department, KDDI Corporation GARDEN
AIR TOWER, 3-10-10, Iidabashi, Chiyoda-ku, Tokyo 102-8460 Japan

E-mail: [†] mt-andou@kddi.com

Abstract Introducing operation and maintenance technology of Wide Area Ethernet from operator/architect of “Powered Ethernet” which is one of the most popular Wide Area Ethernet Service in JAPAN.

Keyword Wide Area Ethernet, VPN, Operation and maintenance

広域イーサネットの運用管理技術

KDDI株式会社
安藤 雅人

目次

- はじめに
- 広域イーサネットサービスの概要
- 広域イーサネットシステムの変遷
 - センター方式
 - QinQ方式
 - MinM方式
- 運用管理
 - 資源管理
 - 監視
 - 切分け技術とイーサネットOAM
- おわりに

Copyright © 2004 KDDI Corporation. All Rights Reserved.

Shaping The Future **KDDI**

はじめに

1999年に日本国内で開始された広域イーサネットサービスは現在ではキャリアが提供する主力のVPNサービスの一つとなっている。

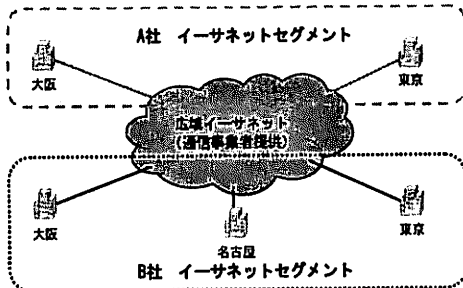
2000年以降キャリアの広域イーサネットシステムの構築及び運用に関わり、さらに、現在国内で最も使われている第三世代の広域イーサネットシステムである、EoE(Ether over Ether)の設計を行った立場から、広域イーサネットシステムの運用管理技術を紹介する。

Copyright © 2004 KDDI Corporation. All Rights Reserved.

Shaping The Future **KDDI**

広域イーサネットサービスの概要

広域イーサネットとは、イーサネットセグメント（ブロードキャストドメイン）を広域で展開するサービスである。利用者は、複数の地理的に離れたイーサネットを一つセグメントとして接続する事が可能となる。



Copyright © 2004 KDDI Corporation. All Rights Reserved.

Shaping The Future **KDDI**

広域イーサネットシステムの変遷

WAN上でイーサネットの利便性を実現

第一世代 センタースイッチ方式

1999年10月～（CIP正式サービス）

WAN上でイーサネット技術の全面採用による経済性を実現

第二世代 QinQ（ダブルタグ）方式

2001年4月～（キャリア系専業サービス） WAN IEEE802.1ad

安定性の追及、大規模化への対応

第三世代 MinM方式 + OAM

MinM: EoE, IEEE802.1ah
OAM: EOP, IEEE802.1ag(Y, 1731)

その他、VPLS方式などを組み合わせる例もある。

Copyright © 2004 KDDI Corporation. All Rights Reserved.

Shaping The Future **KDDI**

第一世代 センタースイッチ方式

WAN上でイーサネットの利便性を実現

- ・ 専用線を用いてセンターにあるイーサネットスイッチまで回線を引き込みスイッチングする方式



DATA Tag SA DA → DATA Tag SA DA → DATA Tag SA DA

課題

- ・ End-Endの遅延の問題
- ・ センタースイッチの信頼性が重要
- ・ 高価（専用線利用のため）

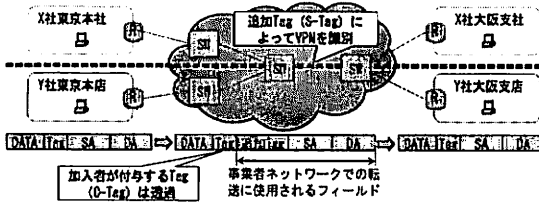
Copyright © 2004 KDDI Corporation. All Rights Reserved.

Shaping The Future **KDDI**

第二世代 QinQ(ダブルタグ)方式

イーサネットの汎用技術をWANのコア技術として採用し経済性を実現

- ・ VLAN-Tag (IEEE 802.1Q) を利用した広域イーサネットシステム
- ・ 加入者のTagを透過出来るように802.1Qの拡張→QinQ(vWAN) 802.1ad



課題

- ・ ループ対策が重要 (イーサネットスイッチ側) ・ 大規模化の限界 (MACアドレス、Tag数等)

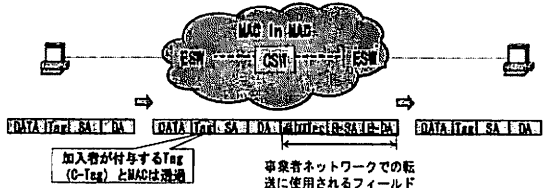
Copyright © 2010 NTT Data Corporation. All Rights Reserved.

Challenging The Future KDDI

第三世代 MinM方式

安定性の追及、大規模化への対応

ユーザのEthernetフレームを事業者のEthernetフレームでカプセル化する方式、EoE (Ether over Ether) : 国内では最も多く実用で使われているMinM方式
IEEE 802.1ah : IEEEにて標準化作業中のMinM方式



MinMによりコアのスイッチにおいて学習しなくなる必要はないMACアドレスの数が劇的に減少し、システムの大規模化と安定性が向上している。

Copyright © 2010 NTT Data Corporation. All Rights Reserved.

Challenging The Future KDDI

運用管理について

運用管理

○主な資源管理項目

- (1) VLAN資源
- (2) 帯域管理
- (3) 学習テーブル

○主な監視項目

- (1) 装置監視 (SNMP, Syslog及び装置に対するPing監視)
- (2) インバンド疎通監視 (主信号疎通の正当性を監視)
- (3) ループ監視/対策

※その他、遅延、ジッタの測定が行われるが顧客に影響が出るようなレベルの現象が出る事はまずない。

○切分け技術とイーサネットOAM

Copyright © 2010 NTT Data Corporation. All Rights Reserved.

Challenging The Future KDDI

資源管理：VLAN資源

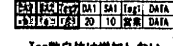
当初、IEEE802.1QをベースとしたVLAN技術(QinQ)を利用。

IEEE802.1Q tag (12bits) = 4096VLAN = 4096VPN

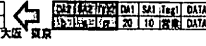
↓ 顧客の増大に伴うVLAN数の不足

OVLAN Cross connect (VXC) によるTag資源の有効活用

エリアごとにTagを付け替え

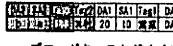


QinQ/MinMにてカプセル化

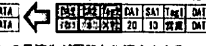


QMinM利用を前提として3段Tag化する事による、コアでのVLAN集約

複数のVLANを異なるTagを追加



QinQ/EoEにてカプセル化



VXCや3段Tagによる対応では、数に限界が来ている。

イーサネットフレームを固定したVLAN空間を拡張 (IEEE802.1ah, EoE)

※TagのBit数を増やしても広大なVLAN空間を握るスイッチ実装の困難が本質的な問題。

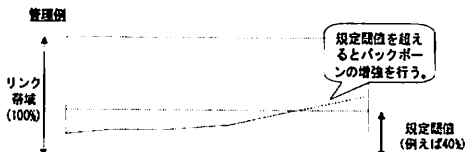
Copyright © 2010 NTT Data Corporation. All Rights Reserved.

Challenging The Future KDDI

資源管理：帯域管理

帯域管理は「じゃぶじゃぶ論」が主流

バックボーンの各リンクの利用帯域を定期的にモニタし、規定の帯域を越えると伝送路やスイッチの増強する事により顧客のトラフィックが溢れる事が無いようにコントロール。



※LAGと10GbEによる帯域拡張には限界があるためより高速度メディアが求められている。

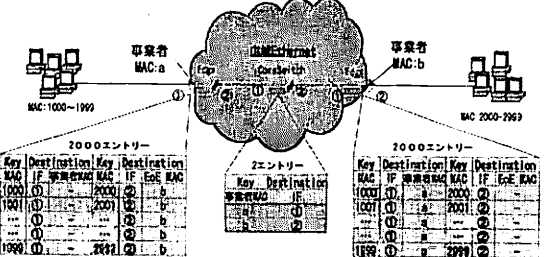
Copyright © 2010 NTT Data Corporation. All Rights Reserved.

Challenging The Future KDDI

資源管理：学習テーブル

学習テーブルの肥大化の問題はMinMの導入により概ね解決されている

MinMの導入により、広域イーサネットコアでのMACアドレス学習量は飛躍的に減少。ただし、運用上の理由により、スイッチ単位ではなく、ポート単位で事業者内部のMACを持つ場合は同様数分のMACがコアスイッチで学習されるため、モニタが必要。



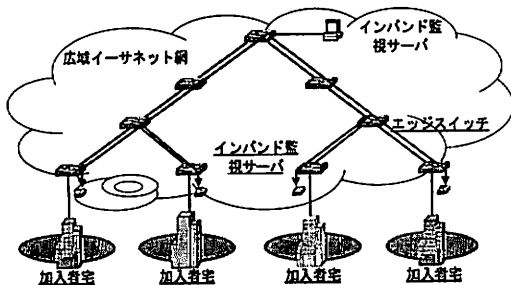
Copyright © 2010 NTT Data Corporation. All Rights Reserved.

Challenging The Future KDDI

インバンド監視：インバンド疎通監視イメージ

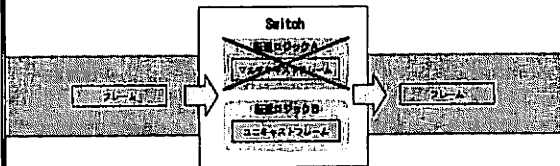
インバンド監視の目的

- 機器が自立的に検出出来ない機器の不具合を検出する
- 設定等の不具合を検出する



インバンド監視：インバンド疎通監視の注意点 1

- フレームの種類により、機器内部で処理が異なる場合がある
特定の転送ロジックのバグや、故障、過負荷が発生すると、特定のフレームが通過しない現象が発生する可能性がある。例えば、ユニキャストフレームは疎通するが、マルチキャストフレームは疎通しないような現象が発生する。

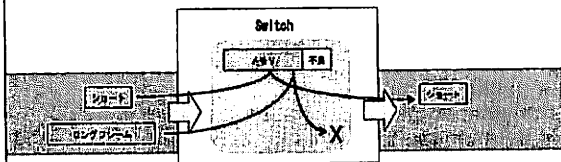


インバンド監視は、ユニキャストだけでなくマルチキャストでも行う。

インバンド監視：インバンド疎通監視の注意点 2

フレーム長による疎通の問題

イーサネットスイッチや伝送装置のメモリ不良、MTUの設定間違い等で、ショートフレームは疎通するが、ロングフレームは疎通しないような場合がある。

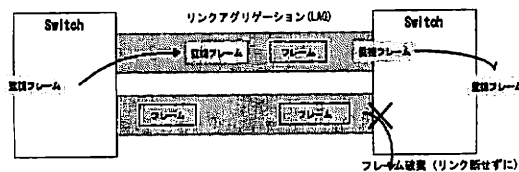


インバンド監視は、最大長フレームで行う。

インバンド監視：インバンド疎通監視の注意点 3

リンクアグリゲーションでの疎通確認の問題

スイッチ間でリンクアグリゲーションを利用している場合、インバンド監視のフレーム1物理リンクしか監視出来ない事がある。可能であれば、物理リンクごとに監視出来るようなインバンド監視を行う。



リンクアグリゲーション利用時のインバンド監視はどの物理回線を監視しているのか注意する必要がある。

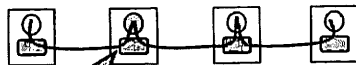
インバンド監視：インバンド疎通監視の注意点 4

インバンド監視は、何を監視しているかを尋ねて使う必要あり。

Simple OAM (IP-Ping, ECP)



Relayed OAM (802.1ag Link Trace)



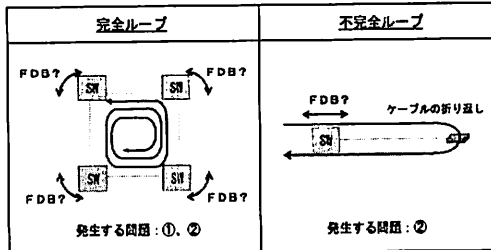
通常のフレームが疎通する回路をモニタ出来ない可能性があるので注意

○ CPU ■ 転送エンジン

ループ監視：ループの弊害

- フレームが増殖する
→ 帯域の圧迫、機器に対するDoS的な過負荷
- FDB (Forwarding Data Base=MAC学習テーブル) の擾乱
→ ユニキャスト通信が不可となる

ループ対策が広域イーサでは適用上最も重要なポイント



ループ監視：検出及び対策	切分け機能とイーサネットOAM技術
ループ検出方式 - スイッチの学習テーブルの書き換え回数に着目する方法 - マルチキャスト/ブロードキャストを網に送信しそのフレームが戻ってくる事をトリガとする方法（発生箇所検出のアルゴリズムは応用例多数） - マルチキャスト/ブロードキャスト/ユニキャストのバランスの変化を見る方法 ループ検出防止方式 - 拡張イーサネットにおけるTTLによるループフレーム破壊とループ検出 - WinBoxにおける階層化Macアドレスリングと階層化フィルタリングによるループフレーム破壊とループ検出	お客様との頃との間ではIPレベルの切分けを行う事も多いが、網内は専用のL2切分けプロトコルを利用するケースが増えている。 □IEEE802.1agのOAM機能 【LB: Loopback】 ループバック試験（Ping相当） 【CC: Continuity Check】 接続性チェック（定期的なインバンド監視） 【LT: Link Trace】 トレースルート □EoE Control Protocolの機能（現状使っているもの） 【EoE Ping】 PE 間あるいはEoE終端機能を付った装置間のPing（疎通確認） 【EoE Traceroute】 TTLを用いて EoE経路のトレースを行う 【EoE フラッシュ】 MAC学習テーブルの全部又は一部のエントリのクリア要求をスイッチ間で交換 今後具体的な利用方法が現場レベルで検討されたいと思われる。

おわりに	EoE階層化MACアドレッシングによるループ防止
広域イーサネット運用上の課題と注意点 - 資源管理面では、VLAN数の拡張とイーサネットメディアの高速化が課題。 - インバンド疎通監視は、何を監視すべきが明らかにした上で漏れなく行う必要あり。 - ループ対策には万全を期する必要がある。（人間がミス犯してもシステムが自動的にループ防止するなどの仕組みが求められる）	エリアごとにEoEのMACアドレスを階層的に割り当てておき、中継ノードはフレームの送り元EoE MACアドレスを調べる事によりFDBの振乱を発生させるようなループフレームを検出し、破壊する。 (1) フィルタによってFDBの誤入がエリアAに波及しない (2) TTLによりループフレームの無意味増殖を防止する

文 献

- [1] 阿留多伎 明良 瀬戸康一郎 園洋志 兵頭弘一 森茂人 岩田淳 安藤雅人, 広域イーサネット技術概論, 阿留多伎 明良 (編), pp.111-153 (社) 電子情報通信学会, 東京, 2005.