STPファミリ以外の冗長化プロトコル
（リングトポロジー）
RPR
EAPS、MRP、MMRP2

リングトポロジー

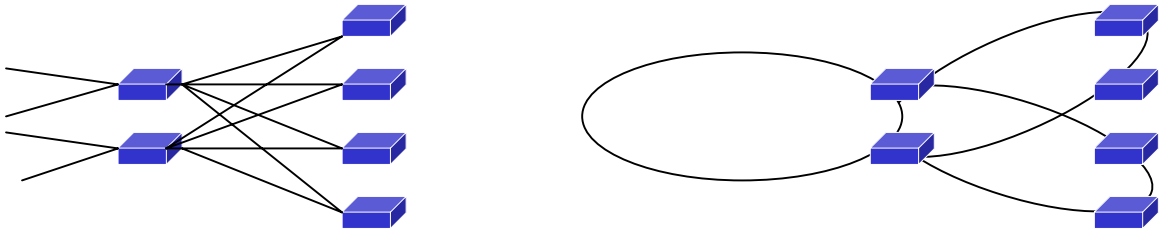
- リングトポロジーはメッシュ（ノード冗長化）トポロジーよりも、伝送路やインターフェースの必要量が少ないと言う特徴がある。

- RPR（Resilient Packet Ring）のように、高度で高価な技術の他にイーサネットスイッチをリング状に配置し、そのリングにHelloパケットを流す事によってリンク断の監視を行い、ブロッキングポートの制御を行うような単純で安価な方式がある。（EAPS、MRP、MMRP2）

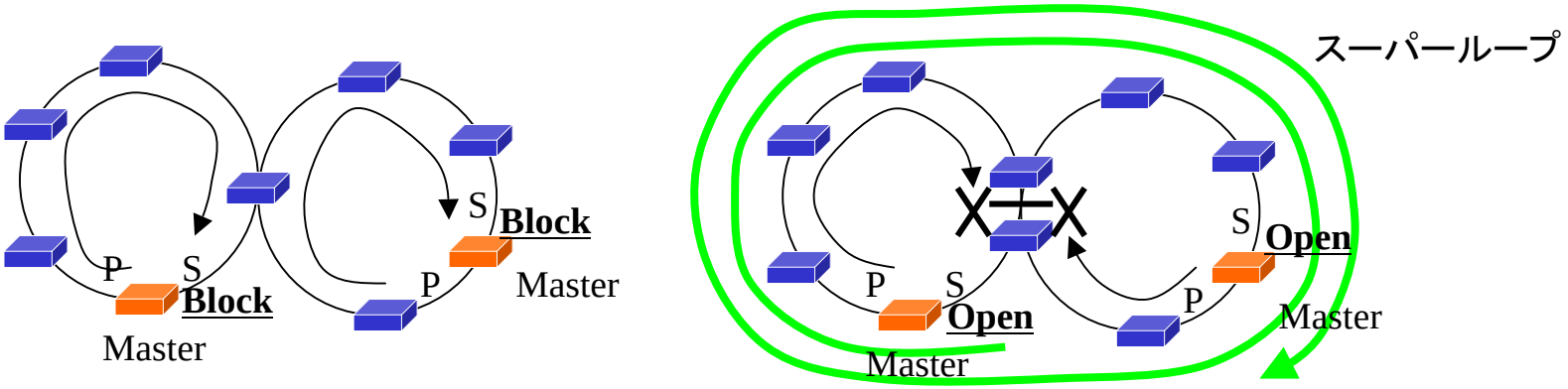
- １つのリングだけでは限界がある事が多く、どうやって、複数のリングを冗長を持たせた形で接続しかつループを起こさないか？が課題の一つとなっている。

ノード冗長化プロトコル　リングマルチ接続

- １つのリングだけではスケーラビリティが限られている為、複数のリングを接続したいと言う要求がある。（大規模な接続を行う場合に、ノード二重化プロトコルを利用するよりも、リンク数を減らせると言う考え方もある、昔、DECのFDDIスイッチが流行した時のようなリングの使い方をやろうとするとこれが必要）

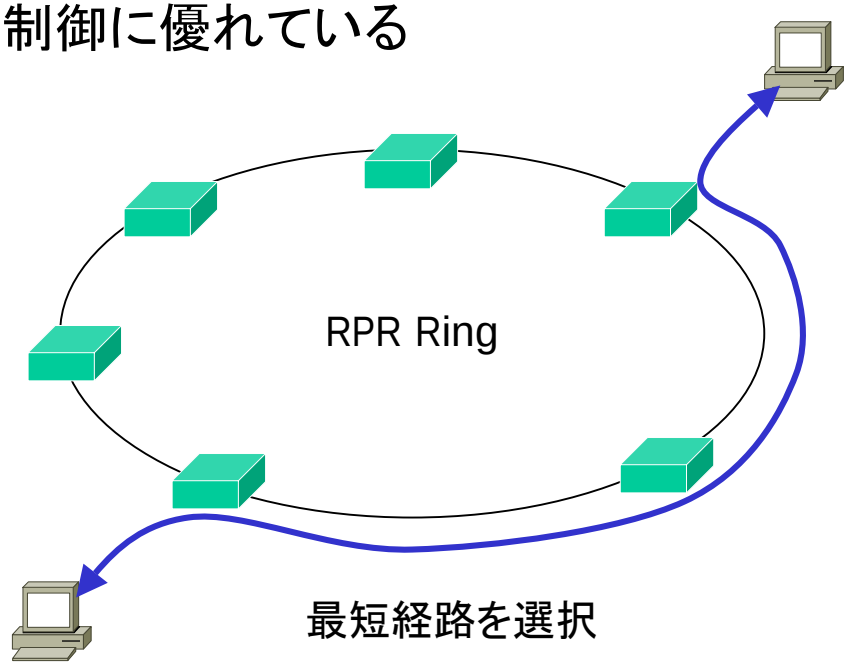


- リングを一箇所で接続するのは問題ないが、冗長の為に2箇所で接続すると、冗長部分の渡りで断が発生した場合にスーパーループが出来てしまう。



RPR(Resilient Packet Ring) 802.17

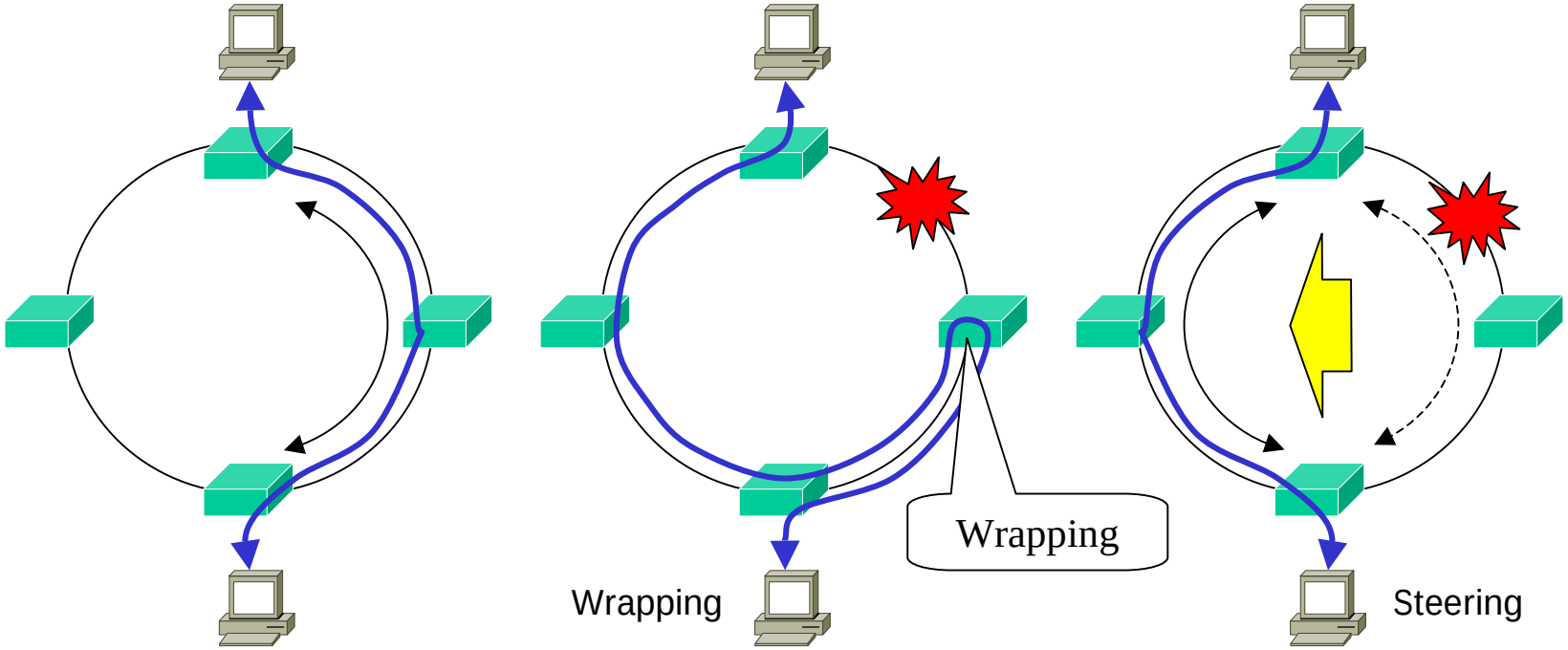
- IEEE 802.17　標準化中
- リング型転送方式（最近のEAPSなどの簡易なリングとは異なる）
- 50msec以内の高速障害回復
- SONET/SDH(C48c、OC192c)で利用可能
- Spatial Reuse 通常時リング内の最短経路で転送を行う（Link State情報）
- 通信の公平性を保つ機構がある
- QoS制御や帯域制御に優れている



RPR(Resilient Packet Ring) 802.17

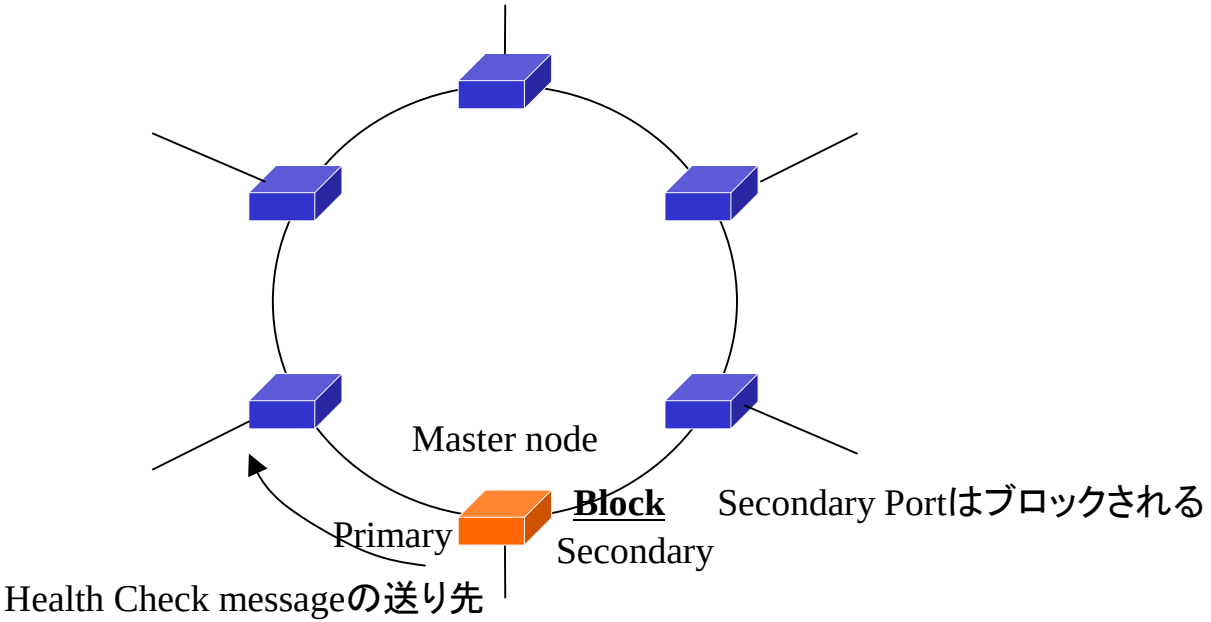
障害発生時の切り替えの方法の種類

- Wrapping
  - 障害が発生部分の直近で折り返したリングを作る事により障害回復を図る。
  - 障害回復がはやい
- Steering
  - 障害部位を通らない方向にリングを切り替える。
  - Wrappingと異なり、切り替え後の遅延の変動も少なく、帯域も有効活用出来る。



ノード冗長化プロトコル　リング構成(EAPS)

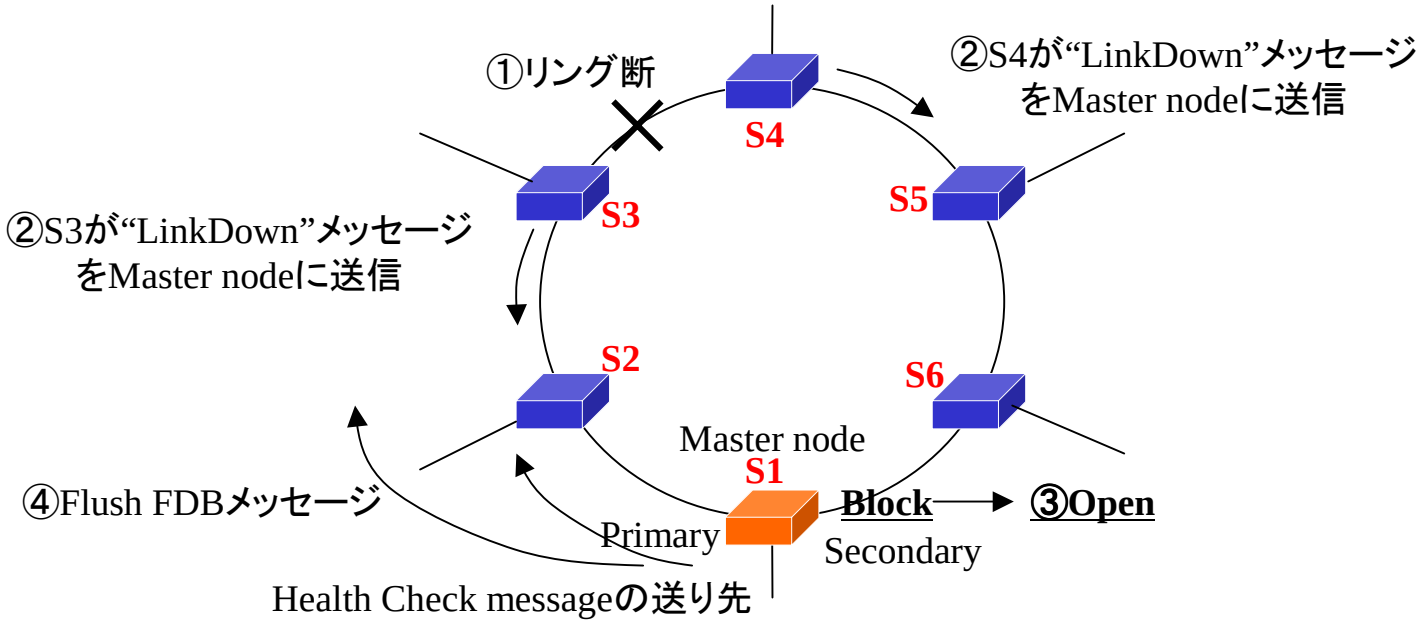
- EAPS（Ethernet Automatic Protection Switching）
  - Extreme社が開発した、リング型冗長化プロトコル
- リング構成で使用する簡易な冗長化プロトコル
  - リングとなるようにスイッチを接続し、その中にMaster nodeを１台指定する（手動）
  - Master nodeのリングに所属するポートの一つをPrimary Portとし、もう一方をSecondary Portとする。
  - Master nodeの Secondary Portをブロッキング状態にする。
  - Primaryポートより、Health Check messageを送信し、Secondaryポートで受信出来るかによってリングの状態を監視する方式



ノード冗長化プロトコル　リング構成(EAPS)

障害発生時の挙動

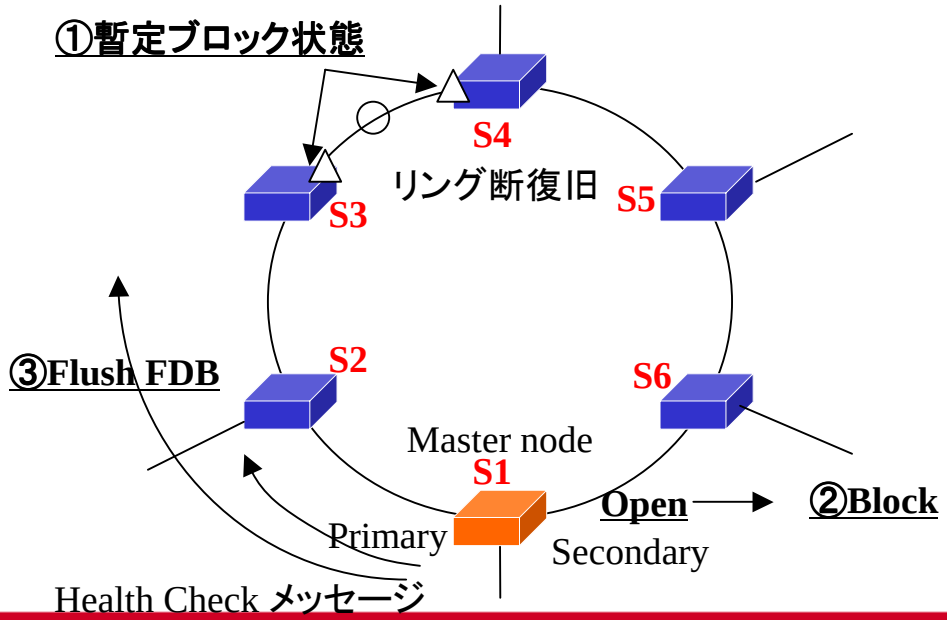
- Master nodeがリングの障害を検出する方法
  - “Link Down”メッセージを転送ノードより受け取る。
  - Health Check messageがSecondaryポートで受信出来なくなる。
- 障害によりリングが切断されたと判断すると、Master nodeはSecondaryポートをBlock状態からOpen(転送)状態に遷移させる。
- トポロジーの変更によるFDBエントリの矛盾を回避する為、Master nodeは転送ノードに対して、“Flush FDB”メッセージを流し、それを受け取った転送ノードはFDBの内容のFlush（消去）を行う。



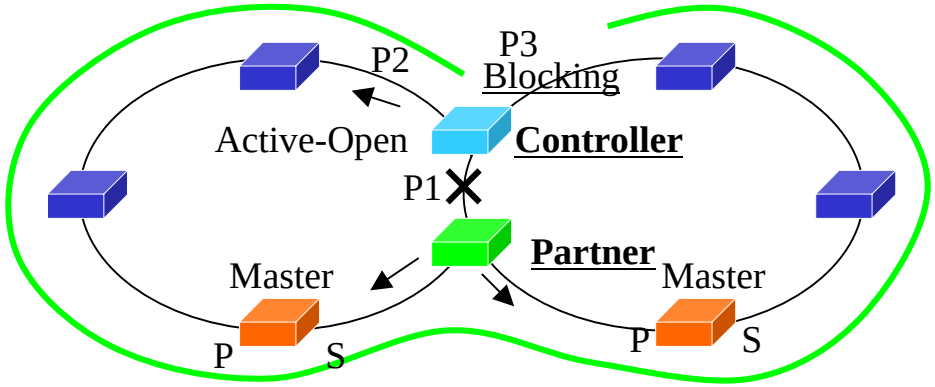
ノード冗長化プロトコル　リング構成(EAPS)

障害復旧時の挙動

- Master nodeはリング障害中もHealth Check メッセージの送信は続け、Secondary Portに届かない限り、障害が継続中であると判断する。
- リングが復旧後、Health Checkメッセージが Secondary Portに到着するまで、ループが発生する可能性がある、これを防ぐ為に、断の復旧を検出した転送ノードはそのポートを暫定ブロック状態とし、実データを通さず、Health Checkのみ通すようにする。
- Master nodeはHealth Check メッセージをSecondary Portで受信すると、Secondary Portをブロッキング状態にし、“Flush FDB”メッセージを送信する。
- 転送ノードは“Flush FDB”メッセージを受信すると、自身のFDBを一旦消去するとともに、暫定ブロック状態のポートを通常の転送状態にし、次のノードに“Flush FDB”メッセージを転送する。



２ノード接続マルチリング構成(EAPS)

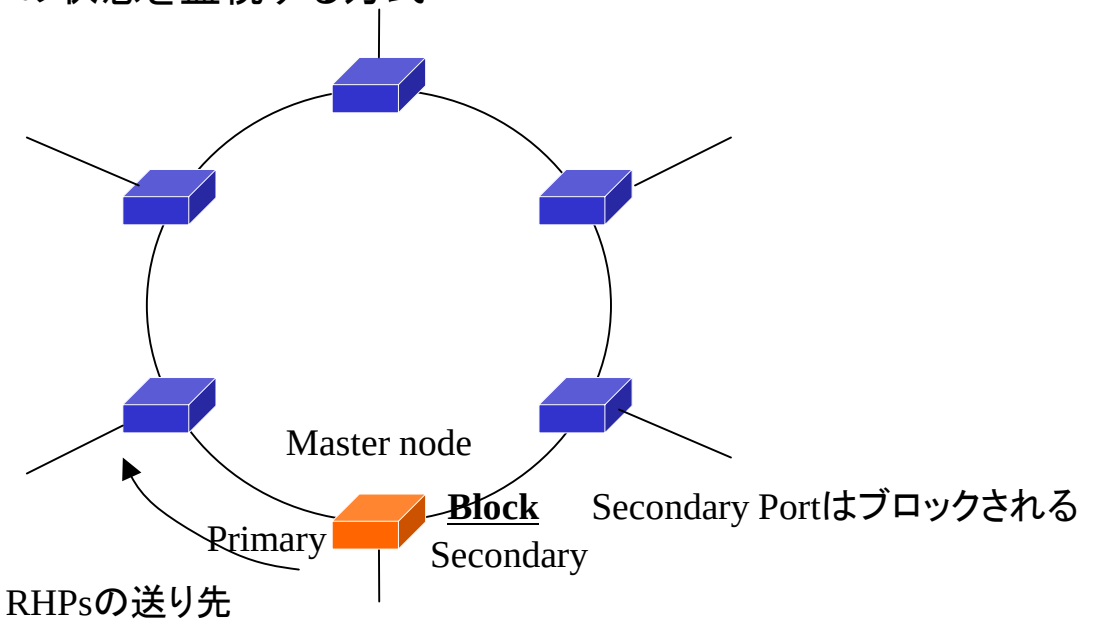


- ２つのリングが共有するリンク部分を挟む形で、Controllerとpartnerを設置しておく（ControllerとPartnerは互いにhelloを交換して共有リンクを監視）
- 共有リンク断
  - 共有リンクの断を検出すると、Controllerは1つのポートをActive-Openと呼ばれる状態にして、他のポートをブロッキング状態にする。
- 共有リンク復旧
  - 共有リンクが復旧すると、ControllerはBlocking Stateにしている部分と共有リンク部分をPreforwarding mode（Masterが流す、health-checkのみ通す）にする。（そのまま転送状態にすると一時的なループを構成してしまう為）
  - 双方のリングのMaster nodeがhealth-checkにより、Secondary Portをブロックにし、“Flush FDB”を送出する。
  - ControllerはMaster nodeがSecondary Portを閉塞した事を示す、“Flush FDB”を受信すると、全てのポートを転送状態にする。

リング構成(MRP)

- MRP（Metro Ring Protocol）
  - Foundry Networks社が開発した、リング型冗長化プロトコル

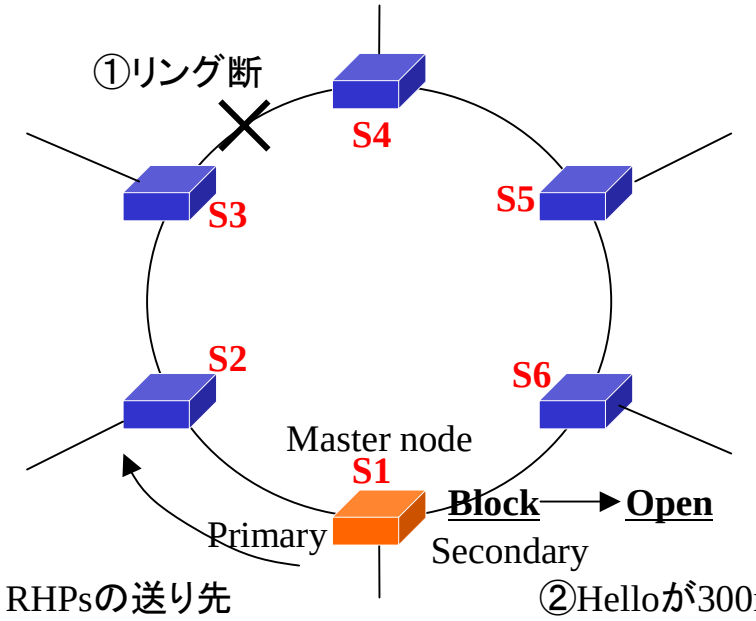
- リング構成で使用する簡易な冗長化プロトコル
  - リングとなるようにスイッチを接続し、その中にMaster nodeを１台指定する（手動）
  - Master nodeのリングに所属するポートの一つをPrimary Portとし、もう一方をSecondary Portとする。
  - Master nodeのSecondary Portをブロッキング状態にする。
  - Primaryポートより、**RHPs(Ring Health Packets)**を送信し、Secondaryポートで受信出来るかによってリングの状態を監視する方式



リング構成(MRP)

障害発生時の挙動

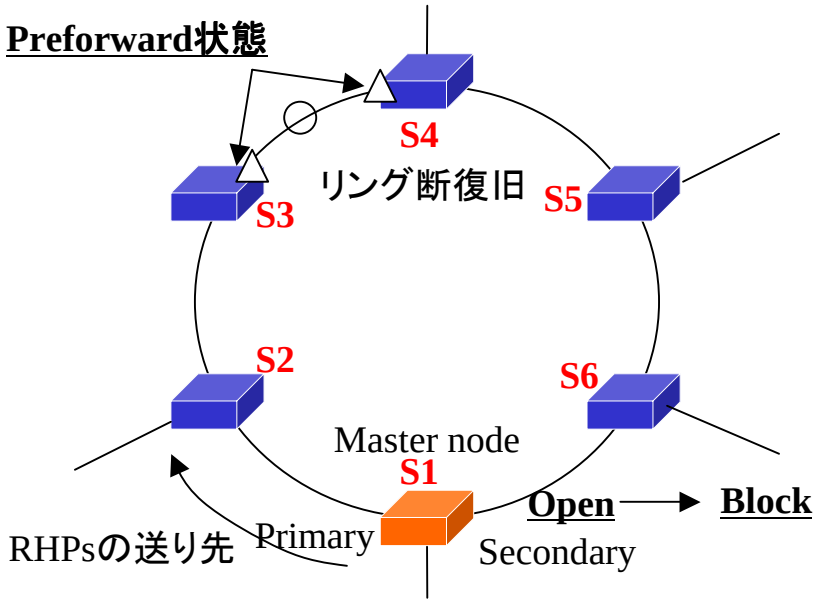
- Master nodeがリングの障害を検出する方法
  - RHPsがSecondaryポートで受信出来なくなる。
  （RHPsは100ms間隔で送信されており、300ms検出されないと異常と見なす。）
- 障害によりリングが切断されたと判断すると、Master nodeはSecondaryポートをBlock状態からOpen(転送)状態に遷移させる。
- トポロジーの変更によるFDBエントリの矛盾を回避する為、Master nodeは転送ノードに対して、“Flush FDB”メッセージを流し、それを受け取った転送ノードはFDBの内容のFlush（消去）を行う。



リング構成(MRP)

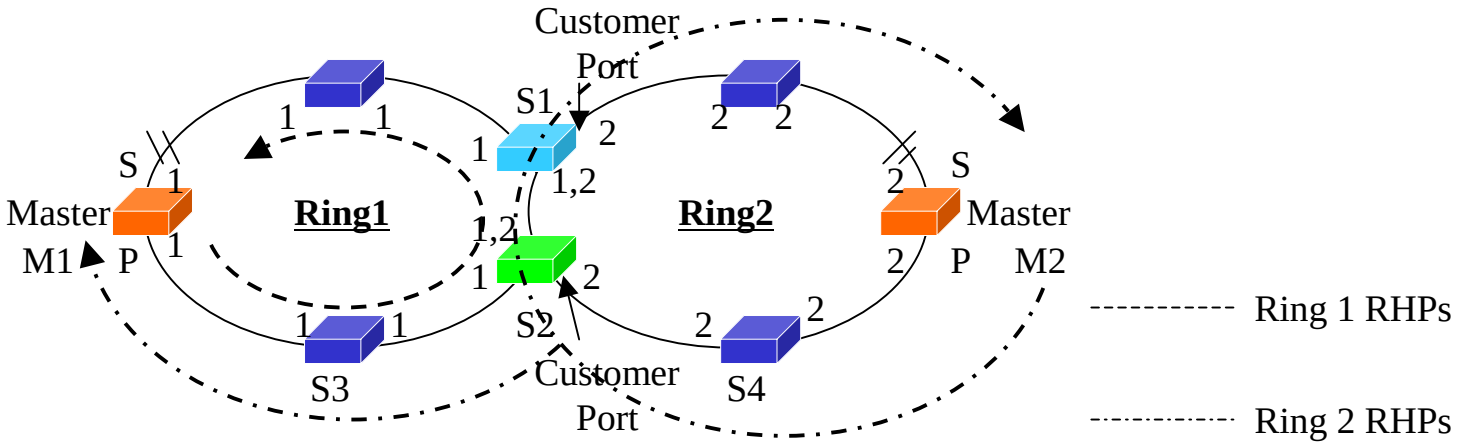
障害復旧時の挙動

- Master nodeはリング障害中もRHPs(Ring Health Packets) の送信は続け、Secondary Portに届かない限り、障害が継続中であると判断する。
- リングが復旧後、RHPsがSecondary Portに到着するまで、ループが発生する可能性がある、これを防ぐ為に、断の復旧を検出した転送ノードはそのポートをPreforward状態とし、実データを通さず、RHPsのみ通すようにする。
- Master nodeは RHPsをSecondary Portで受信すると、Secondary Portをブロッキング状態にし、“Flush FDB”メッセージを送信する。
- 転送ノードは“Flush FDB”メッセージを受信すると、自身のFDBを一旦消去するとともに、Preforward状態のポートを通常の転送状態にし、次のノードに“Flush FDB”メッセージを転送する。



２ノード接続マルチリング構成(MRP)

- 2ノードのマルチリング接続はIronWare Release 07.7.00からサポート
- Ring Priorityと呼ばれる数値が各リングに設定される。
- 二つのリングが接続されている所でRing Priorityの数字が大きい側のリングのポートはCustomer Portと呼ばれる。
- Customer Portから入力された、RHPsだけはRing Priorityの小さい側のリングにも流れ込む。このRHPsは通常Master Nodeで止まる。



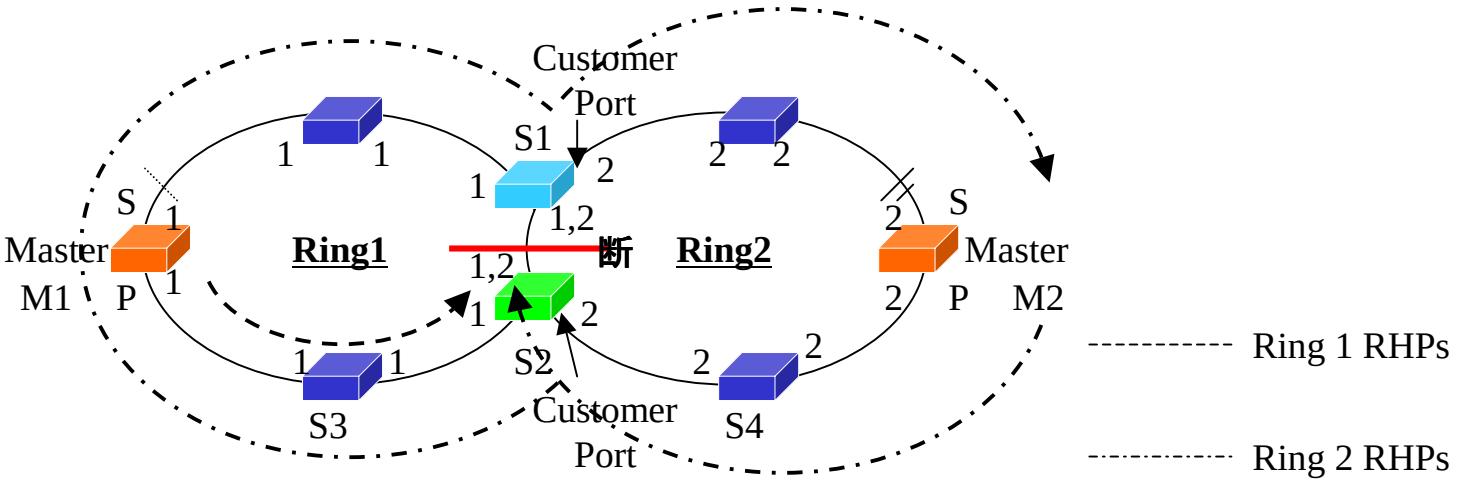
２ノード接続マルチリング構成(MRP)

障害発生時の挙動

- 共有リンク部分S1-S2間での断をM1が検出すると、M1のSecondary portは他のリングのRHPsのみを透過するPreforwarding状態になる。
- この状態ではM2のSecondary Portはブロック状態のままでありこの後、M1がSecondary portをforwarding状態にしても、ループは発生しない。

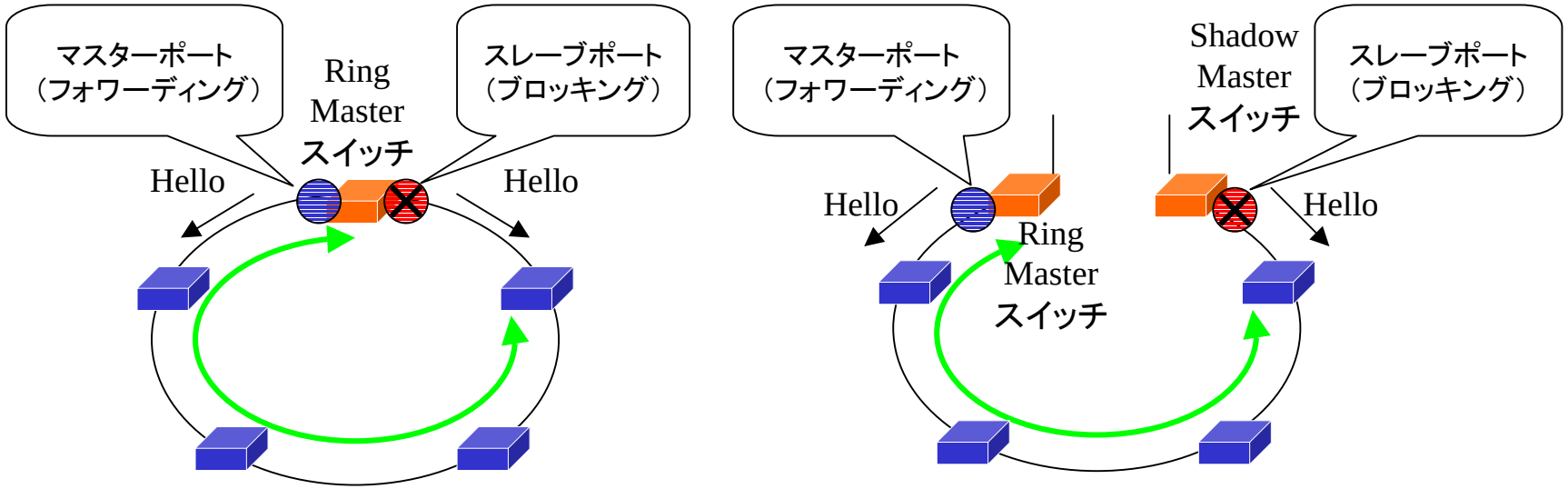
障害復旧時の挙動

- １リング構成の場合と同じ



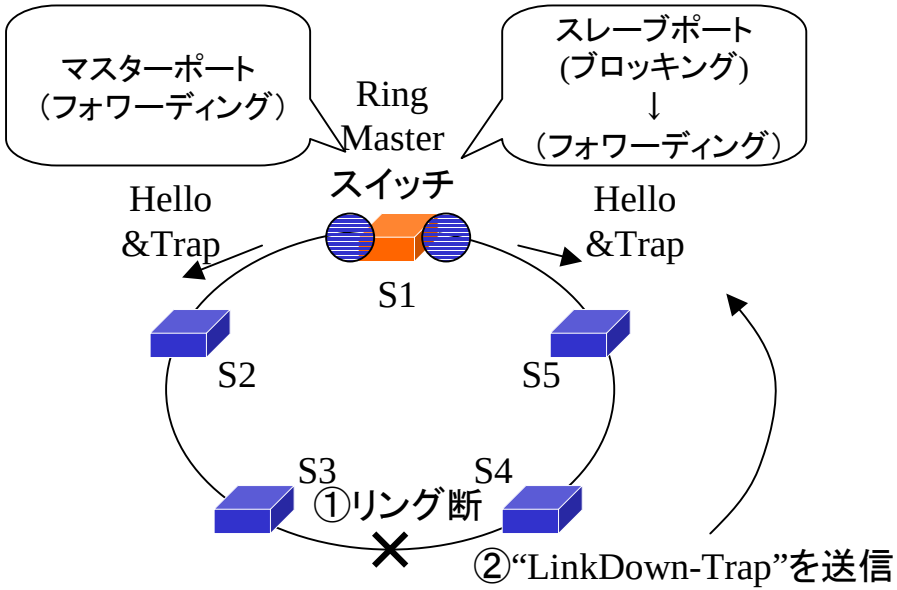
ノード冗長化プロトコル　リング構成(MMRP2)

- MMRP（Multi Master Ring Protocol 2）
  - 日立電線が開発した、リング型冗長化プロトコル
  - Ring Master スイッチには、Masterポート及び、Slaveポートがあり両方のポートで、Health Checkフレームを投げる。
  - Health チェックが相手のポートに届いているかどうかでリングの状況を確認する。



ノード冗長化プロトコル　リング構成(MMRP2)

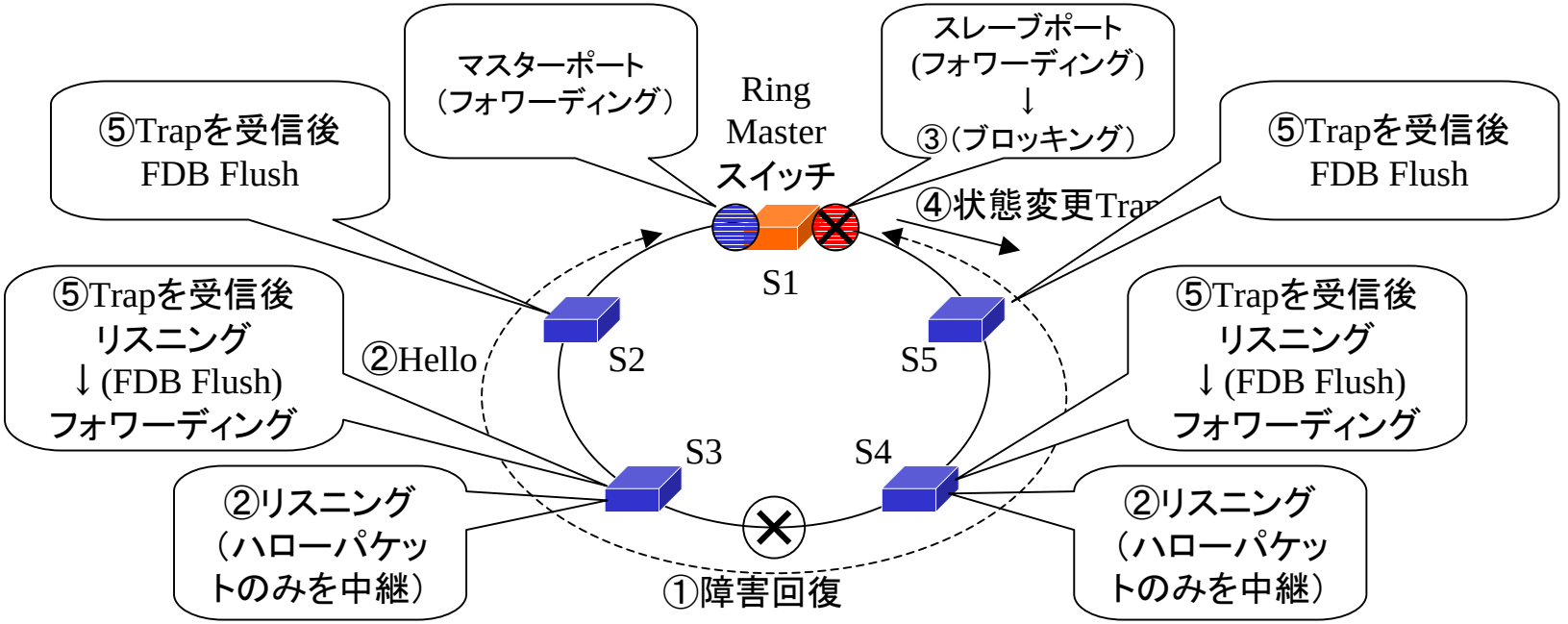
- 障害発生時の挙動
  - 障害の検出方法(Ring Master node)
    - スレーブポートがHello packetを n秒以上連続して受信しない場合
    - LinkDown-Trapを受信した場合
  - Ring nodeでの障害の検出とFDBフラッシュ
    - 直接LinkDownを検出した場合
    - Ring Masterスイッチがスレーブポートをフォワーディング状態に変更したTrapを受信した場合
    - どちらかのHello Packetをn秒以上連続して受信しない場合



ノード冗長化プロトコル　リング構成(MMRP2)

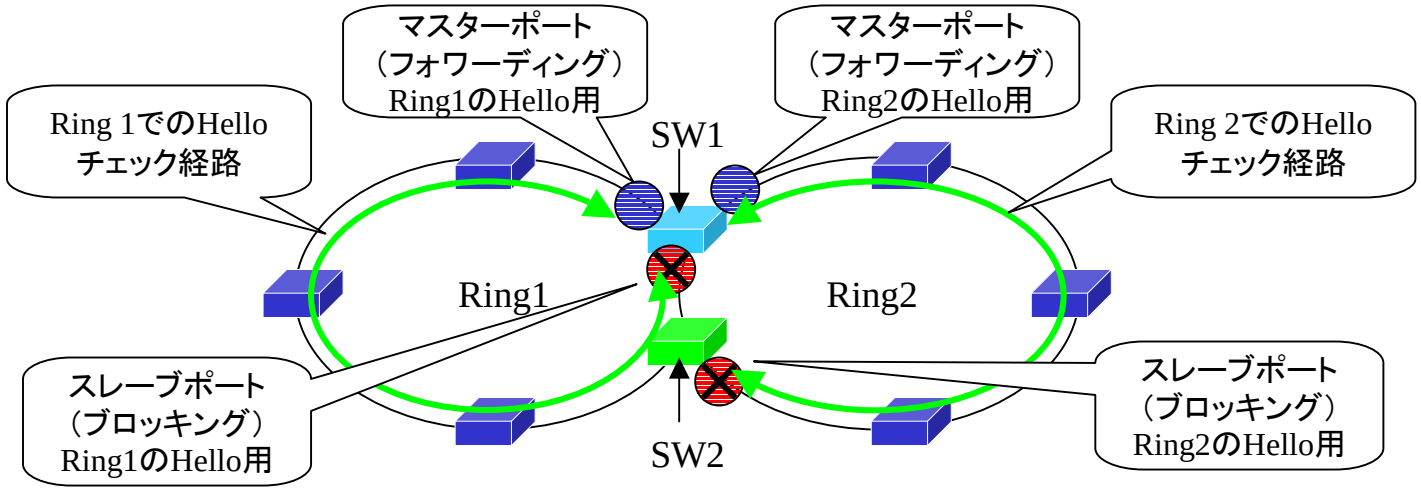
- 障害回復時の挙動
  - Master nodeでの障害回復の検出
    - Helloパケットを受信した場合

    →スレーブポートをブロッキングに切り替え、ブロッキングにした事を示すTrapを送信
  - Ring nodeでの障害回復の検出とFDBフラッシュ
    - 直接LinkUpを検出したノードはそのポートをすぐに転送状態にはしないで、ハローパケットのみ通す、リスニング状態にする。
    - Ring Masterスイッチのスレーブポートがブロッキングになった事を示す、trapを受信後、リスニング状態のポートがあれば、フォワーディングに変更し、FDBをFlushする。



２ノード接続マルチリング構成(MMRP2)

- MMRP2によるマルチリング
  - Masterスイッチを分散して設置出来る事を利用して、マルチリング接続を行う。
  - 共有部分のリンク断によって、ループ構成とならないように、MasterとShadow Masterの配置を行えばそれだけで、特別な機能は使わずにマルチリングの構成を構築する事が出来る。

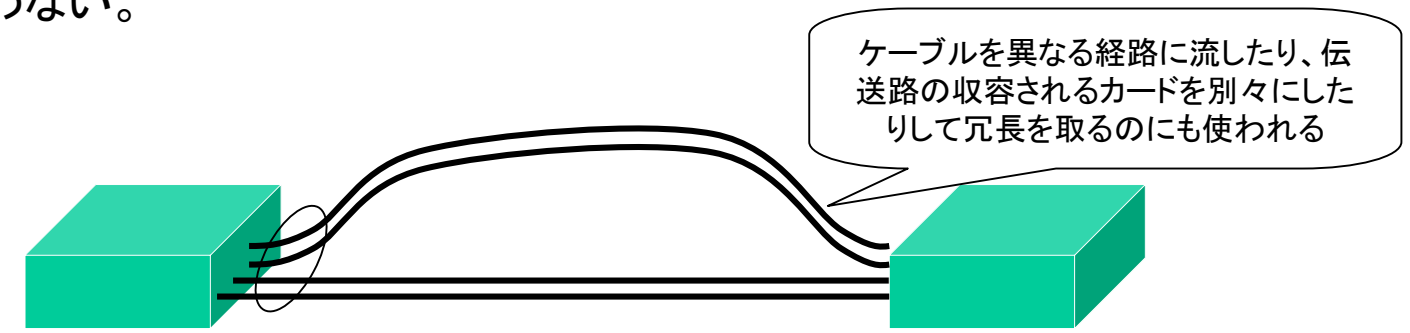


SW1：Ring1のシングルMasterスイッチ、リング2のShadow Masterスイッチ
SW2：Ring2のShadow Masterスイッチ

リンク冗長としてのLink Aggregation

Link Aggregation 802.3ad

- IEEE 802.3adとして標準化
- スイッチ間の複数の物理リンクを論理的に1本にまとめて使う機能
- 負荷分散の他、伝送路の冗長を確保する為にも用いられる
- 制御プロトコルとして、(LACP:Link Aggregation Control Protocol)を規定
  - 論理チャネルを動的に組み上げるためのプロトコルで、リンクの状態のチェックや接続の間違いをチェック出来る。
  - 実装されていない場合は、手動で設定する。
- トラフィックの振り分けは、MAC、IP、ポート番号、入力ポートなどのハッシュやラウンドロビンなどがある。
  - 平均的に分散するわけではないので、リンク数に比例したパフォーマンスを期待出来るわけではない。
  - ラウンドロビンはパケットの順番入れ替えが発生する可能性があるのであまり使わない。

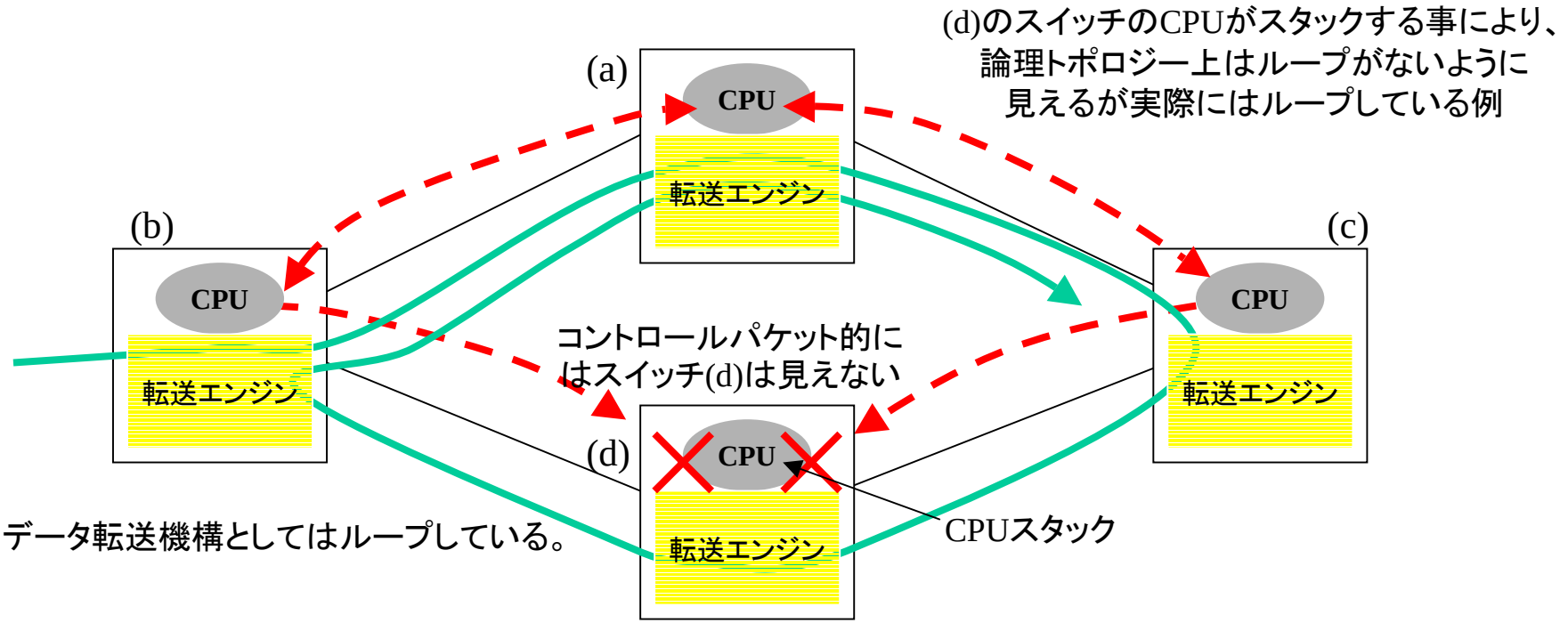


複数の伝送路をまとめて使用

ループを検出し、論理トポロジーに働きかける機構

ループを検出し、論理トポロジーに働きかける機構

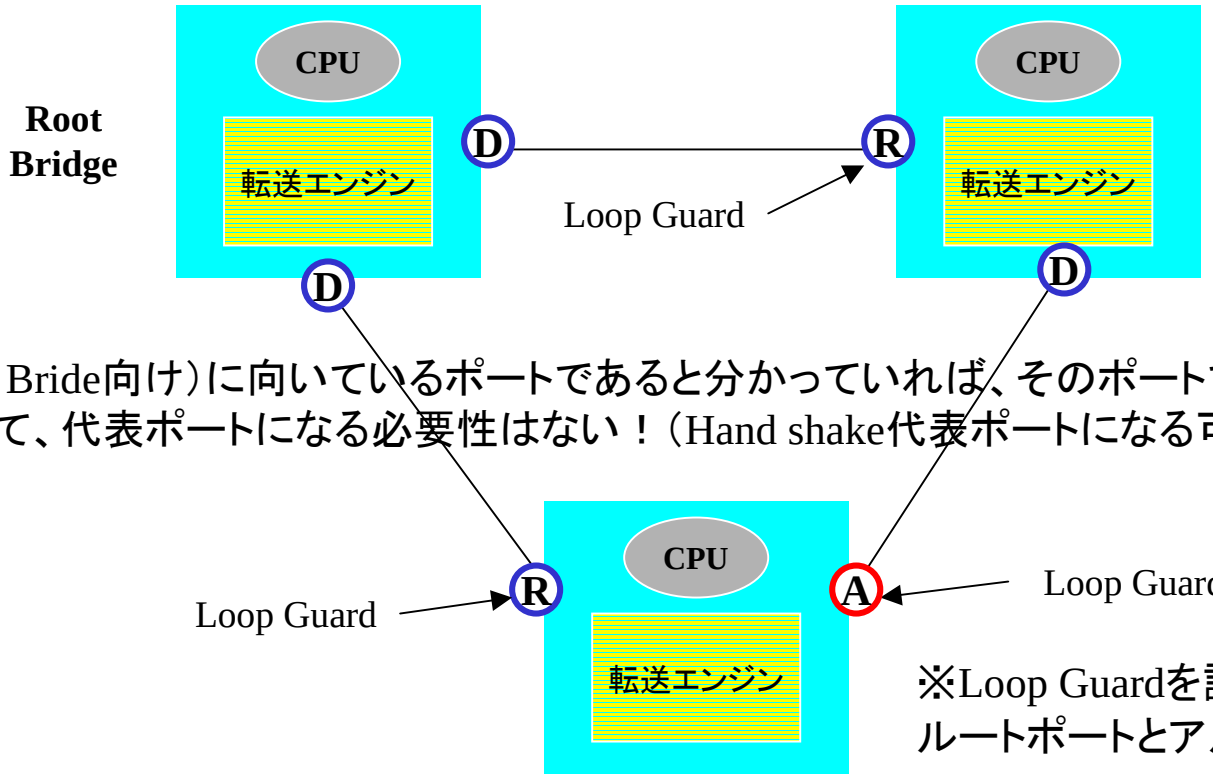
- 論理トポロジーがアルゴリズム通り維持出来ない可能性もある。
  - 多くのスイッチが、データ転送を行う部分（転送エンジン）と、制御パケット（BPDU）などをコントロールする部分（CPU）が分かれている為、CPUの過負荷やソフトウェアの障害により、制御パケットがCPUに転送されるが、処理させず、データの転送のみが実行される状況が発生する可能性が存在する。
  - その他、リブートのタイミングで、設定情報の読み込みに失敗して立ち上がってくるようなスイッチが存在する可能性も存在する。



ループを検出し、論理トポロジーに働きかける機構

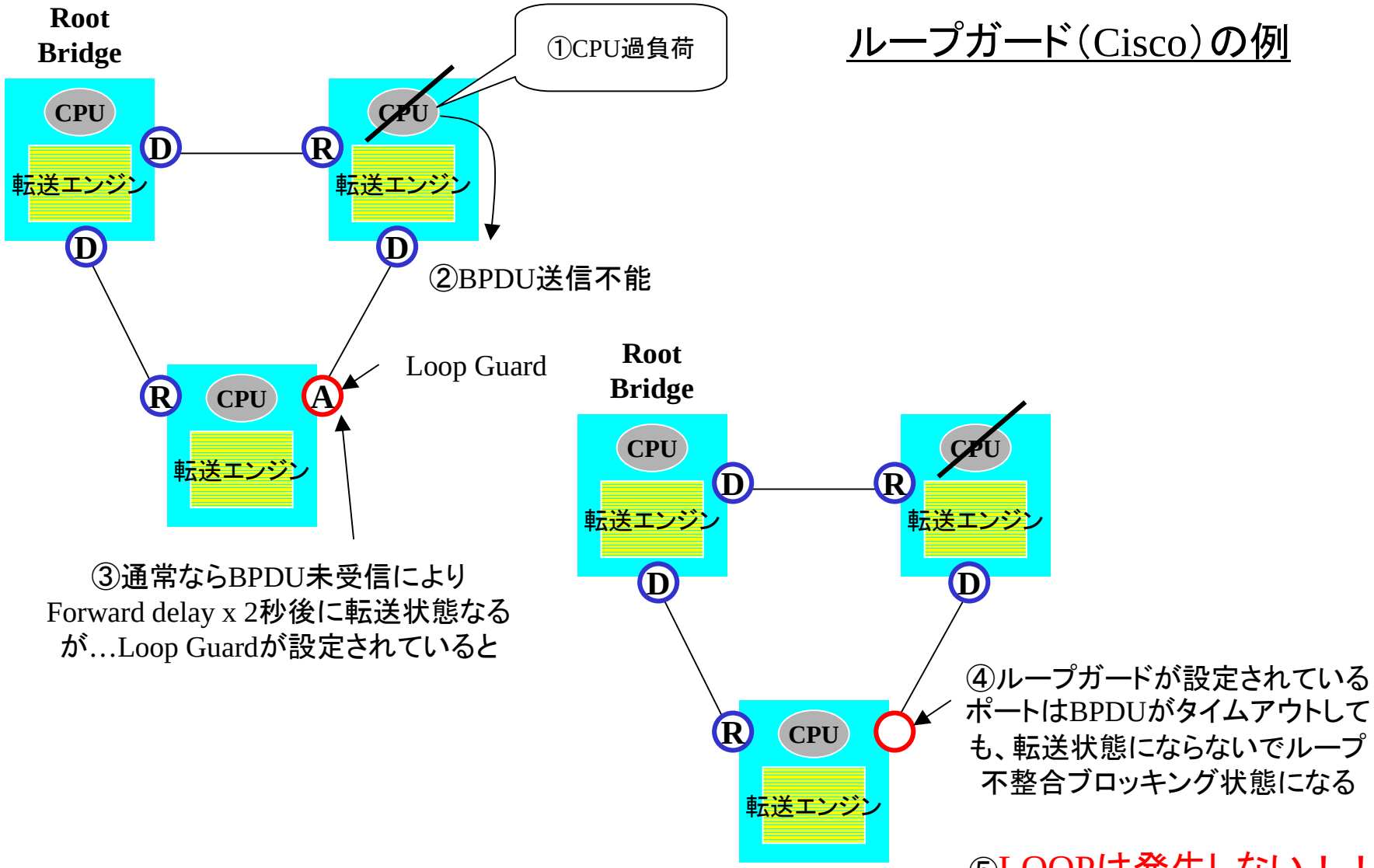
ループガード（Cisco）の例

- ループガード（R-STP用）
  - 隣接したスイッチのCPU異常や単方向リンクの発生によりBPDUが受信出来ない場合に、ループの発生を防ぐ機能
  - ループガードを設定したポートでBPDUを受信しなくなっても、代表ポートにはせず、ループ不整合ブロッキングステートにする。



※上位（Root Bride向け）に向いているポートであると分かっていれば、そのポートでBPDUを受信しなくなる事によって、代表ポートになる必要性はない！（Hand shake代表ポートになる可能性はある）

ループを検出し、論理トポロジーに働きかける機構



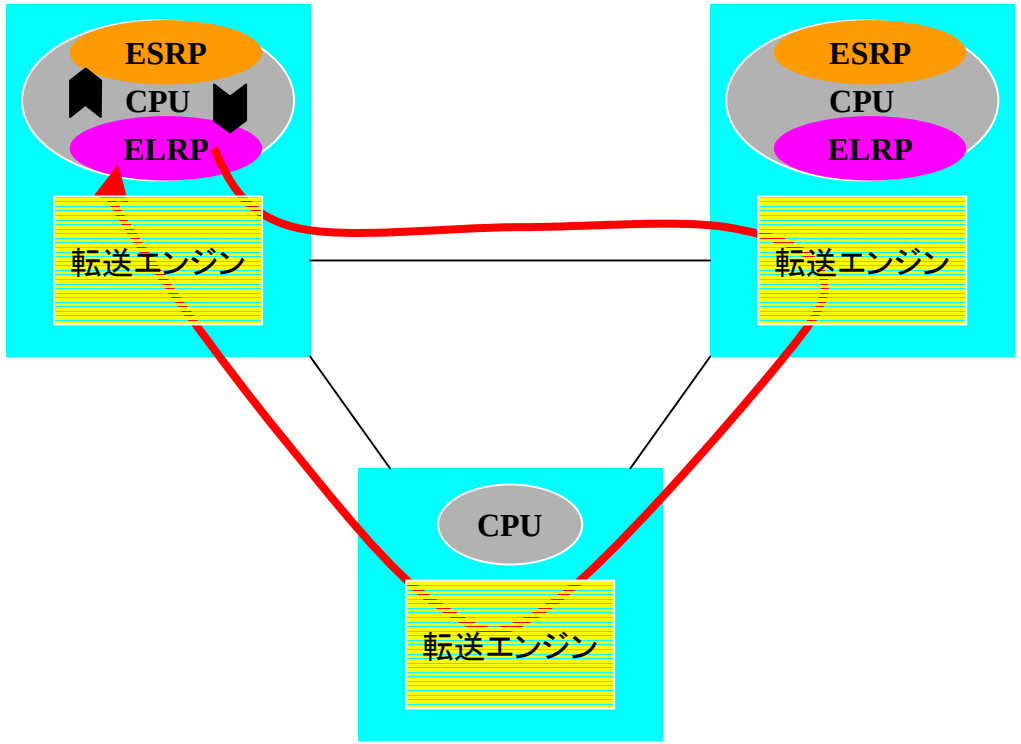
ループガード（Cisco）の例

Loop Guard

ループを検出し、論理トポロジーに働きかける機構

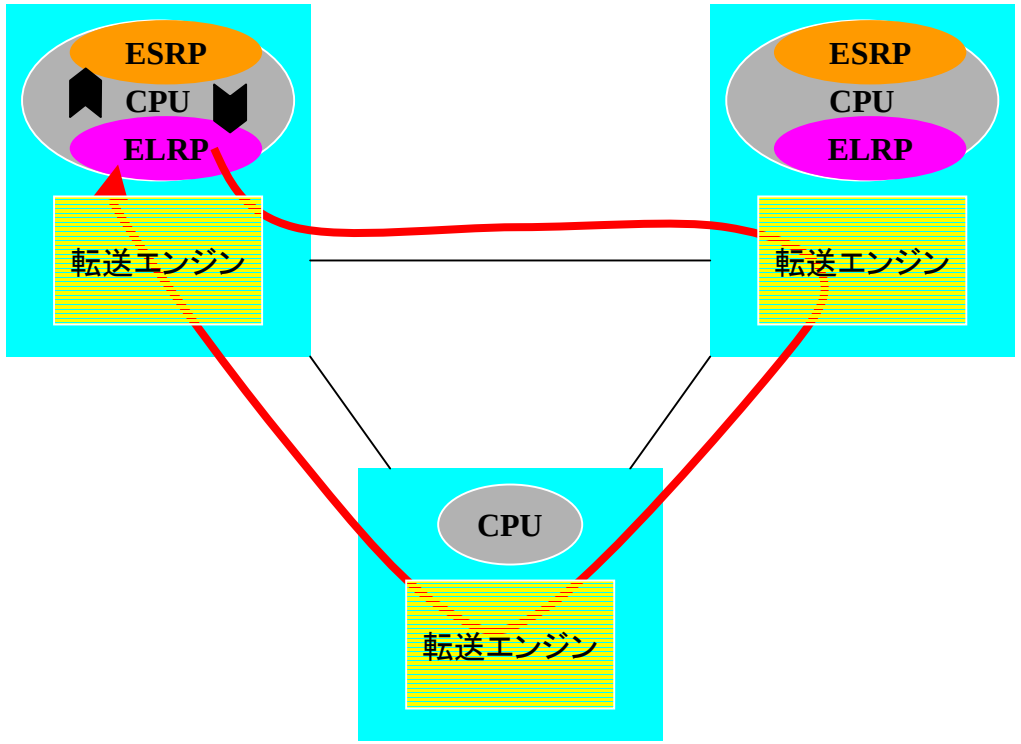
Extremeの例

- ELRP（Extreme Loop Recovery Protocol）
  - ループ検出を行い、論理トポロジーに働きかける機構として、ELRPと呼ばれる機構を提供している。この機構はESRPと組み合わせて利用される。
  - この機能はESRPのMaster/Slaveの関係の中で、両方のスイッチがMasterにならないよう、ループの防止/検出をする。



ループを検出し、論理トポロジーに働きかける機構

- ELRPの動作概要
  - ELRPパケットを呼ばれる監視パケットを送出し、ループの有無を検出
    - 送出されたパケットを受信した場合は、ループありと判定
    - 送出されたパケットを受信しない場合は、ループなしと判定（正常）
  - ELRPのパケットは発信スイッチ以外のスイッチではCPUに転送されず、転送エンジン内で転送される為、他のスイッチのCPUの状況に左右されない。



ループを検出し、論理トポロジーに働きかける機構

- ESRPとELRPの連携
  - ELRPはループ検出を行うが、トポロジー維持機構のESRPと連係して動作をする事により、ループを検知し、ループ回避を行う事が出来る。

- ELRP Master-poll機能
  - Master スイッチが定期的に両系Master検出用パケット（宛先マルチキャスト）を送出し、戻りを検出した場合にループが発生していると判断し、Slaveに落ちる機能。

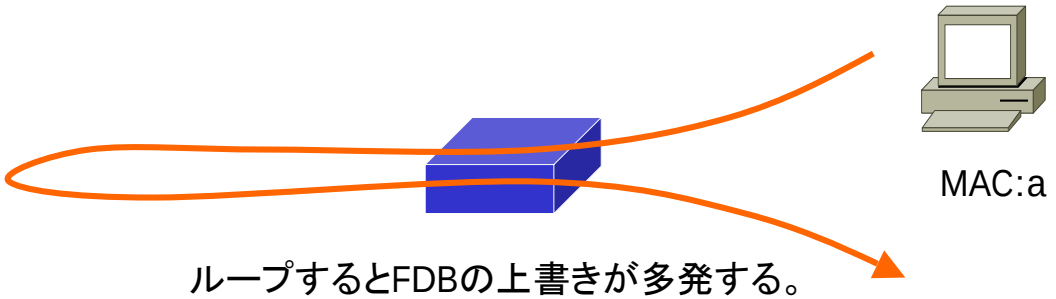
- ELRP premaster-poll機能
  - スイッチがESRPのMasterに遷移する直前に、両系Master検出用パケット（宛先マルチキャストを送出し、戻りを検出した場合に自身がMasterに遷移した場合にループになる事を事前に察知し、Masterにならない機能。

ループ検出機構

ループ検出機構

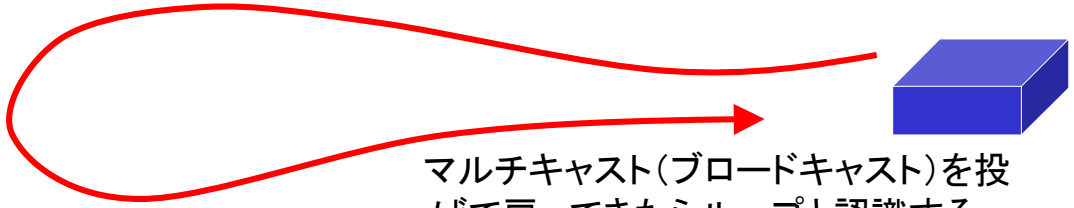
- ループ検出の必要性について
  - 様々な機構を駆使して、ループを防止したとしても、設定ミスなどの可能性やケーブリングのミスなど、ループする可能性を0にする事は難しい。
  - 被害を最小限に食い止める為にループが発生している事にすぐに認識し対策をうつ必要がある。

- ループ検出機構
  - FDBの書き換わりを見張る方法
  - ループするとFDBの書き換わりが多発する、ある程度の時間内にある程度の回数書き換えが発生すると、ループの疑いがあるとして、警報を上げる。（SEIKOのキャリア向けスイッチなどに実装）



ループ検出機構

- マルチキャストHelloポーリング
  - マルチキャスト宛てにHelloフレームを投げ、それが戻って来るかどうかを確認する事により、ループがないか確認する方法。（制限はあるがCiscoなどが実装）

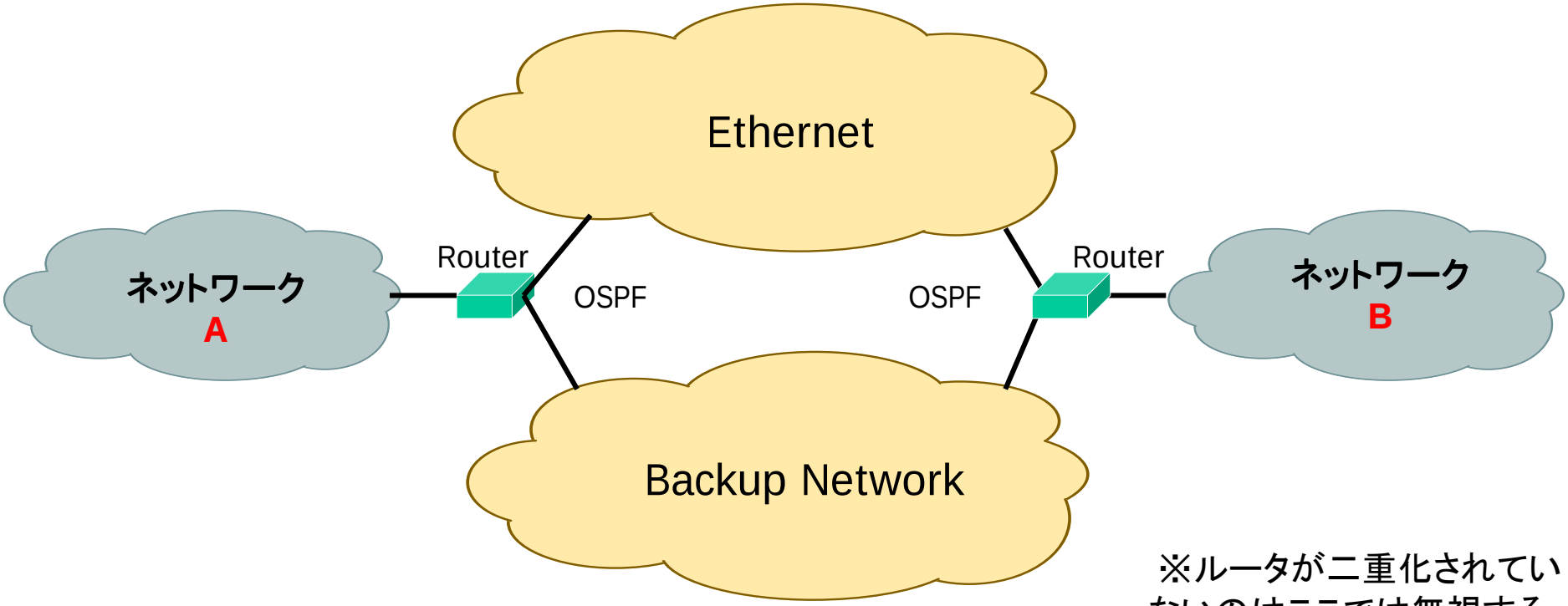


- マルチキャスト、ブロードキャスト、ユニキャストの流量観察
  - ネットワーク上に流れている、マルチキャスト、ブロードキャスト、ユニキャストの定常的な流れのバランスをモニタしておき、マルチキャストやブロードキャストの流量の急激な増加、ユニキャストの減少などよりループを検知する。
  - 運用レベルの検出方法。

大規模イーサネットを介したネットワーク冗長について

大規模イーサネットを介した冗長を組む場合

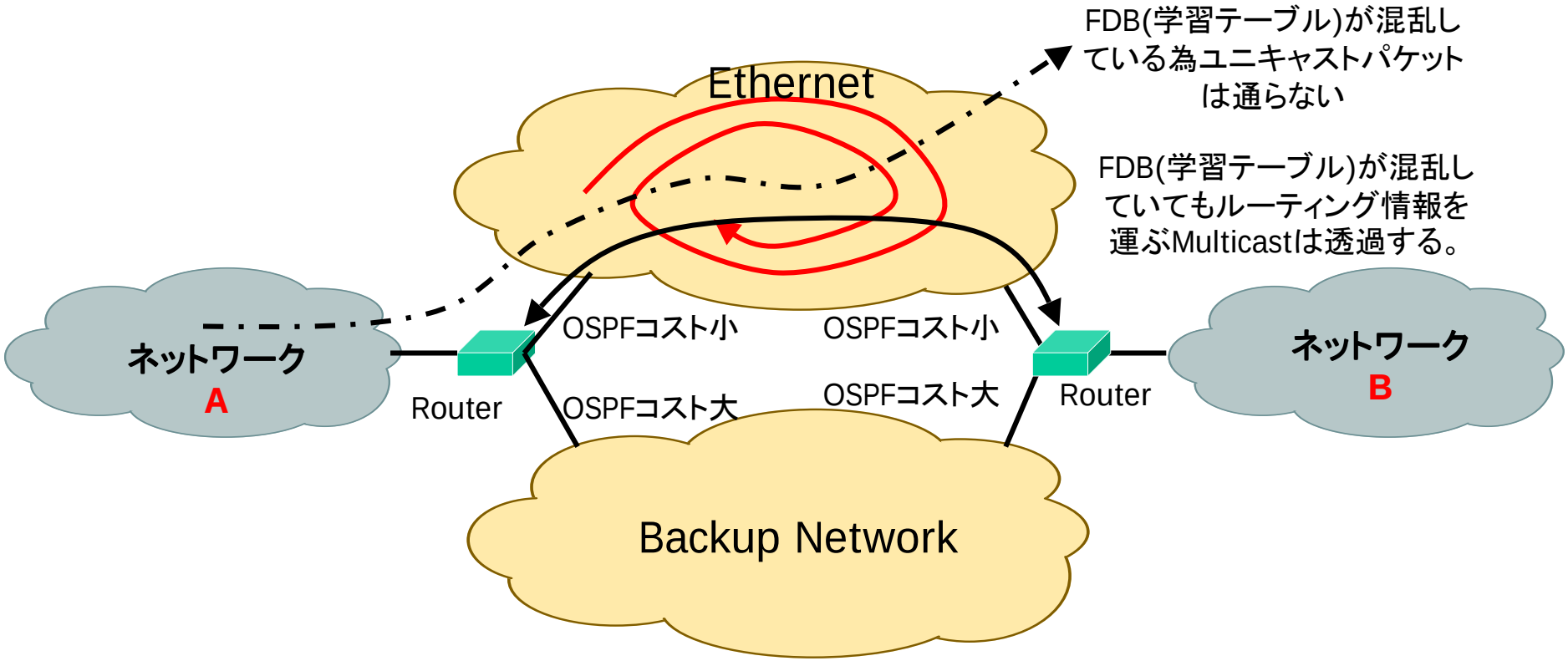
- 広域イーサネットや大規模なイーサネットを利用し重要なサービスを動作させるような場合、バックアップの通信経路を持たせるような事がしばしば行われる。
- 一般には、OSPFやEIGRPなどをそのまま動作させる方法が取られる。



※ルータが二重化されていないのはここでは無視する。

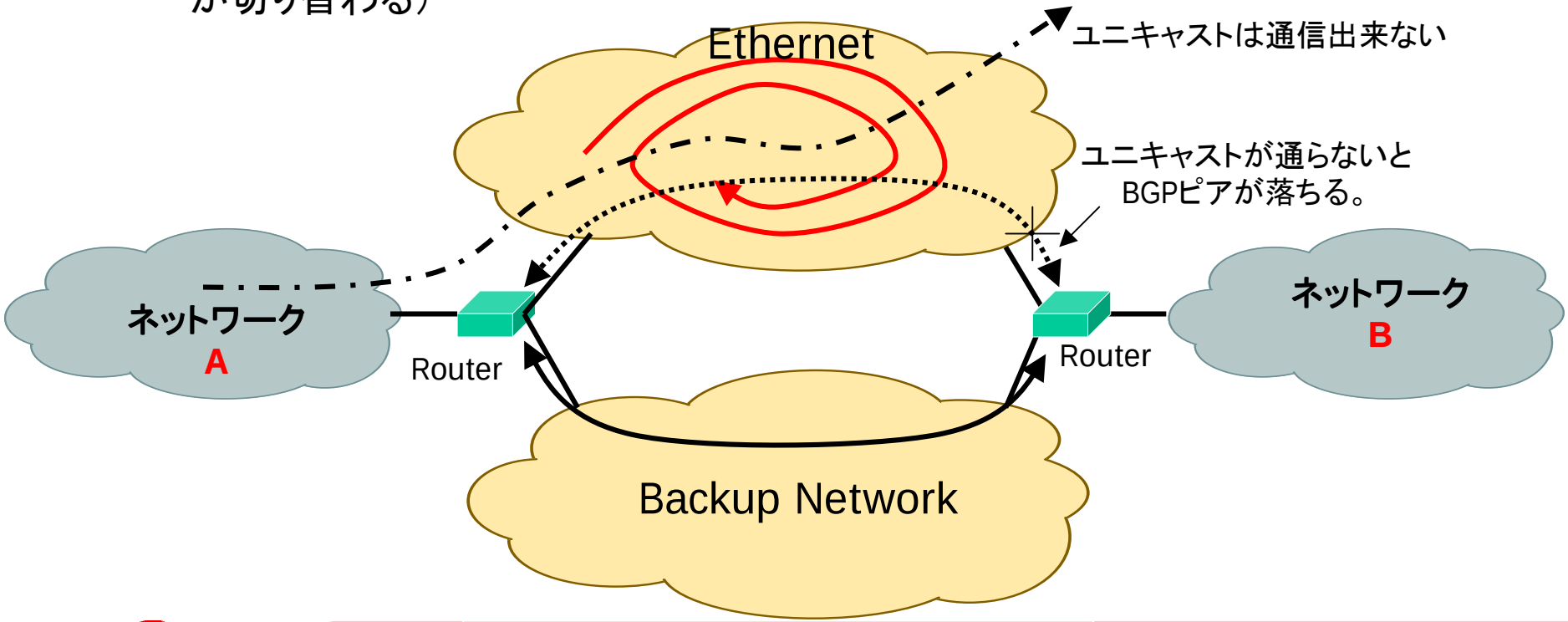
大規模イーサネットを介した冗長とループ

図のEthernet部分でループが発生した場合、
FDBの内容が狂う為、条件によっては、マルチキャストは通るが、ユニキャストが通らないと言う状況が発生する。OSPFはルータ間のプロトコルのやり取りをMulticastで行っている為、Ethernet1側でループによって接続性が不安定になっている事を検出出来ず、結果的に、ネットワークAからネットワークBへの通信に影響を与える可能性がある。



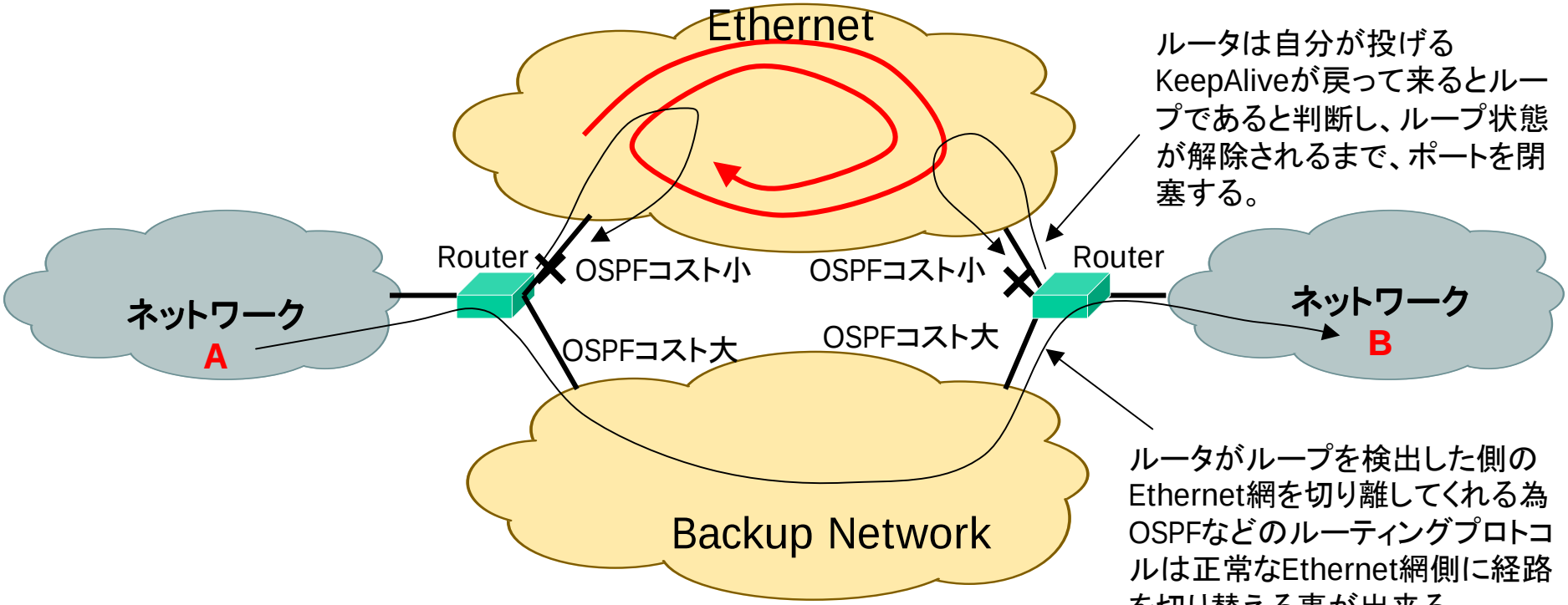
大規模イーサネットを介した冗長とループ

- 拠点数が少ない場合は拠点間でBGPピアを張る場合もある。
- BGPで張る場合のメリット
  - Ethernet網の中間で切れた場合に、検出しやすい。
  - ループの発生によって、使えなくなっているEthernetを検知出来る場合がある。（BGPはユニキャストでルータ間にTCPピアを張るので、FDBが狂っている場合にピア自体が落ちてくれる場合がある、そうするとBACK Up側にネットワークが切り替わる）



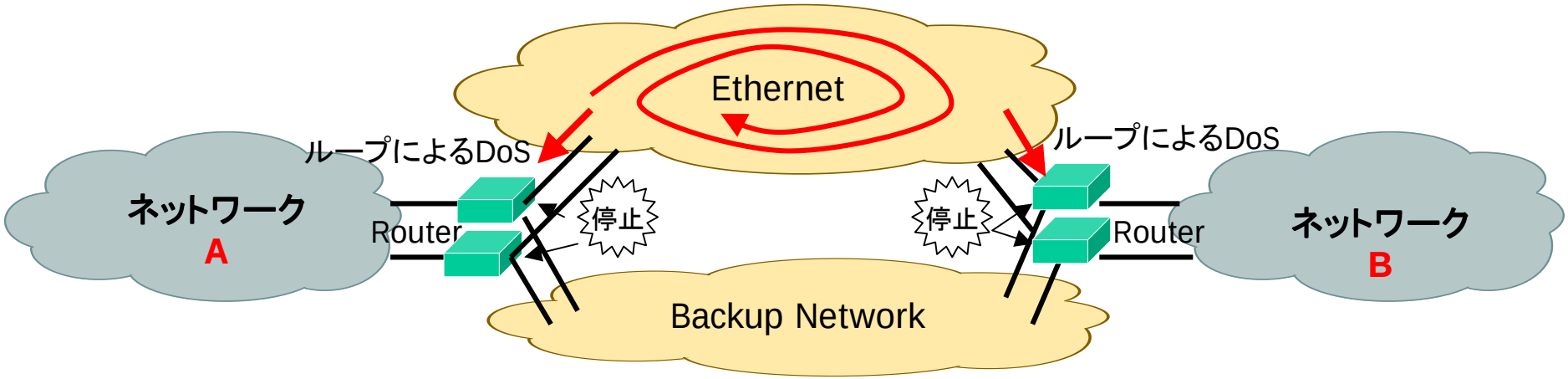
大規模イーサネットのループを検出し経路を切りかえる

例えばCiscoの一部のL3スイッチの機能には、Keep Aliveを定期的にEthernet網側に投げ、そのKeep Aliveが折り返して来た場合にそのEthernet網がループを発生させていると判定してポートに閉塞をかける機能があるものがある。(ループが解除されてしばらくたつとポートの閉塞を解除するような機能もあるものもある)このような機能を利用する事により、複数のEthernet網を使って、冗長を組んだ場合に、ループの発生を検出して、経路の切り替えを、効果的に行う事が出来る可能性がある。このような機能は、ループによる、フレーム増殖より、アプリケーションを守ると言う側面もある。
※ただし、Ethernet網側では、Keep Aliveを透過するような設定をしておかなくてはならない。

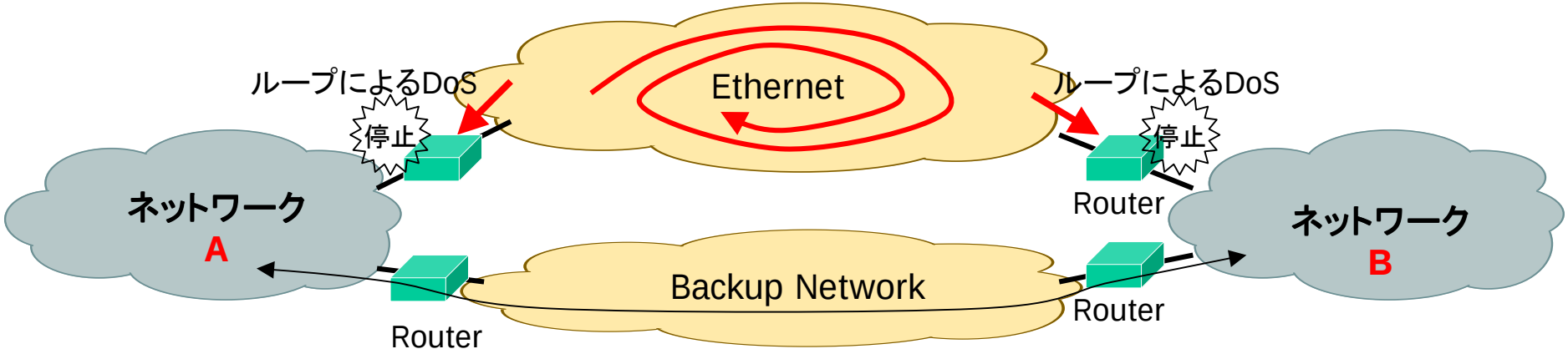


大規模イーサネットでのループによるルータ停止対策

- ループによって増殖したフレームでルータが死ぬ場合もある。

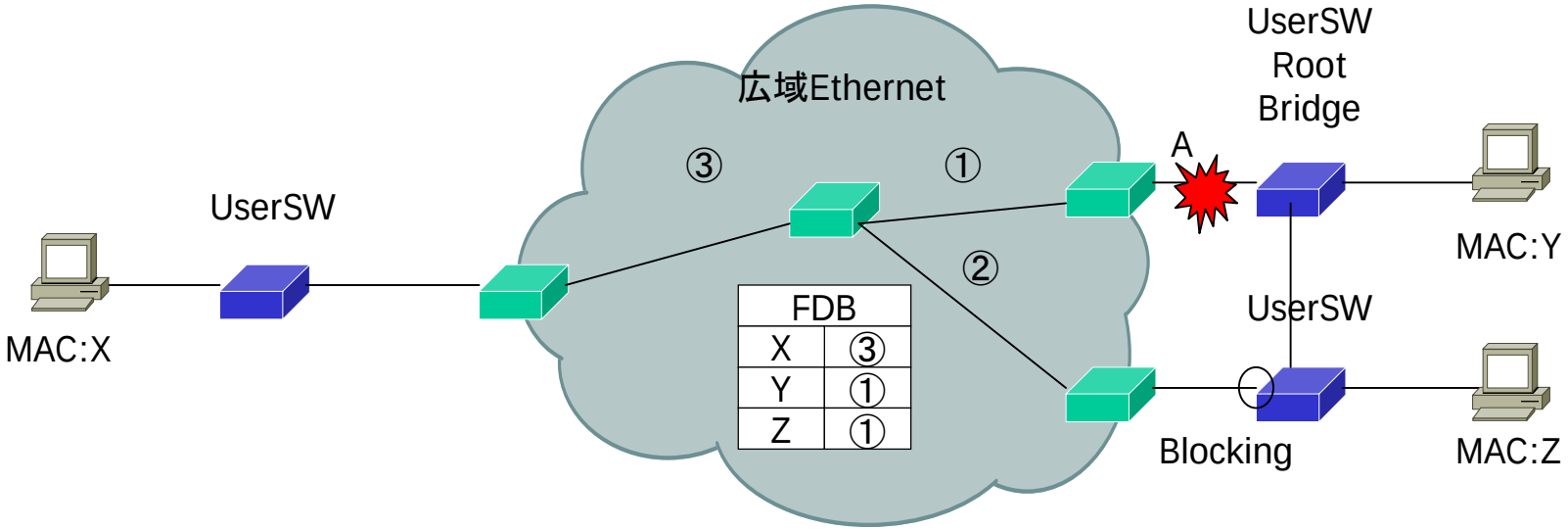


- 一箇所のループで、影響を受けないルータの組を考えたほうがいい。



広域イーサネットをまたいだSTP

- 広域イーサネットをまたいだSTP
  - 全てのキャリアが対応しているわけではない。(あんまりおすすめじゃない)
  - 広域イーサネットのスイッチが、ユーザSTPの切り替えを認識しない為切り替え動作にFDBが Age outするまで(一般には5分)待たなくてはならない。(広域イーサネットに接続された機器が定期的にマルチキャストやブロードキャストを送信していればこの問題はある程度解決する)



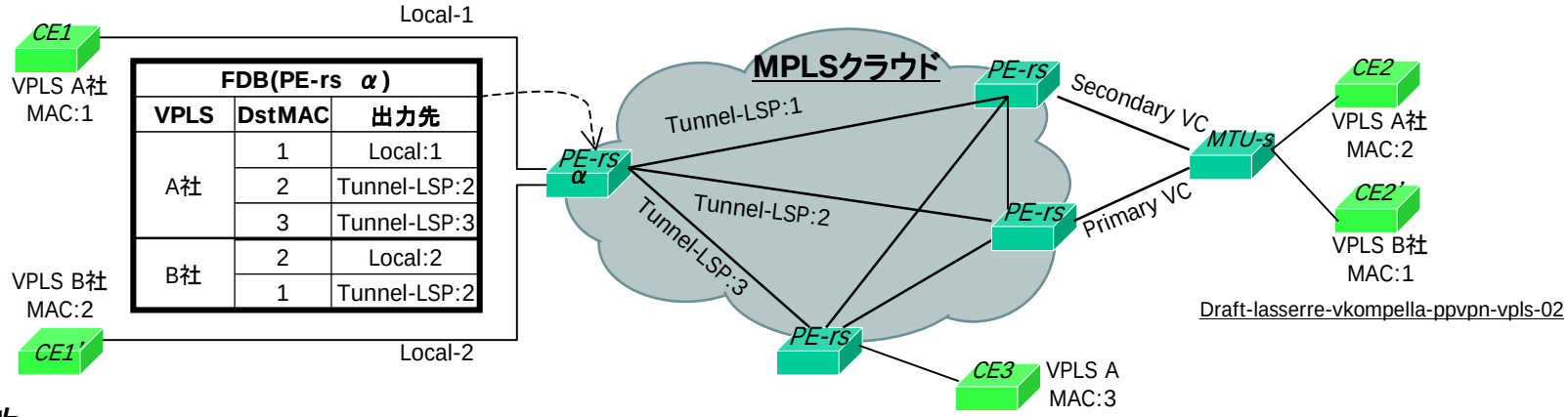
このような構成でAの部分が切断されると、UserSWのトポロジー変更が上手くいったとしても、広域Ethernet網内のBridgeTableが古い状態のまま保持されてしまう為、AgeOutするまで通信が出来なくなってしまう事がある。

付録1 : VPLS（Virtual Private LAN Service）

ルーティングを使ってイーサネットの冗長を実現する例

VPLS(Virtual Private LAN Service)

- **EthernetフレームをMPLSを使ってMultipoint to Multipointで転送する技術**
  - PE間でフルメッシュLSPを張り、ブリッジングはPEで行い、MPLSのコアのではラベルスイッチングのみを行う。

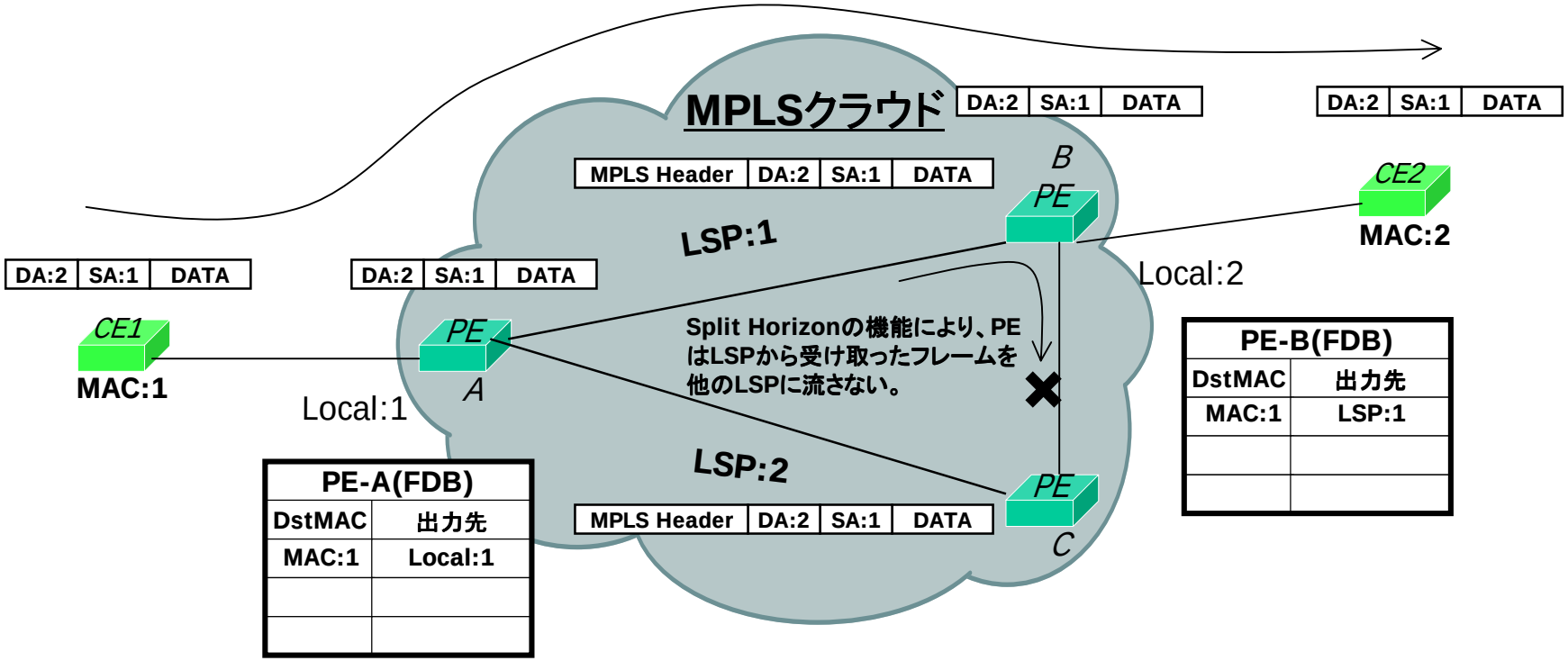


- 特徴
  - PE間でフルメッシュにトンネルを張る。
  - PEは関連したトンネルに関してMACの学習機能を持つ。
  - MPLS網側から受け取ったフレームをMPLS網に戻さない、**Split Horizon** の機能により、Loopを防止する。
  - MPLSベースの強力な冗長化機能が使える。

VPLS(Virtual Private LAN Service)基本動作(1)

FDBにMACアドレスが学習されていない場合のVPLS上のフレーム転送

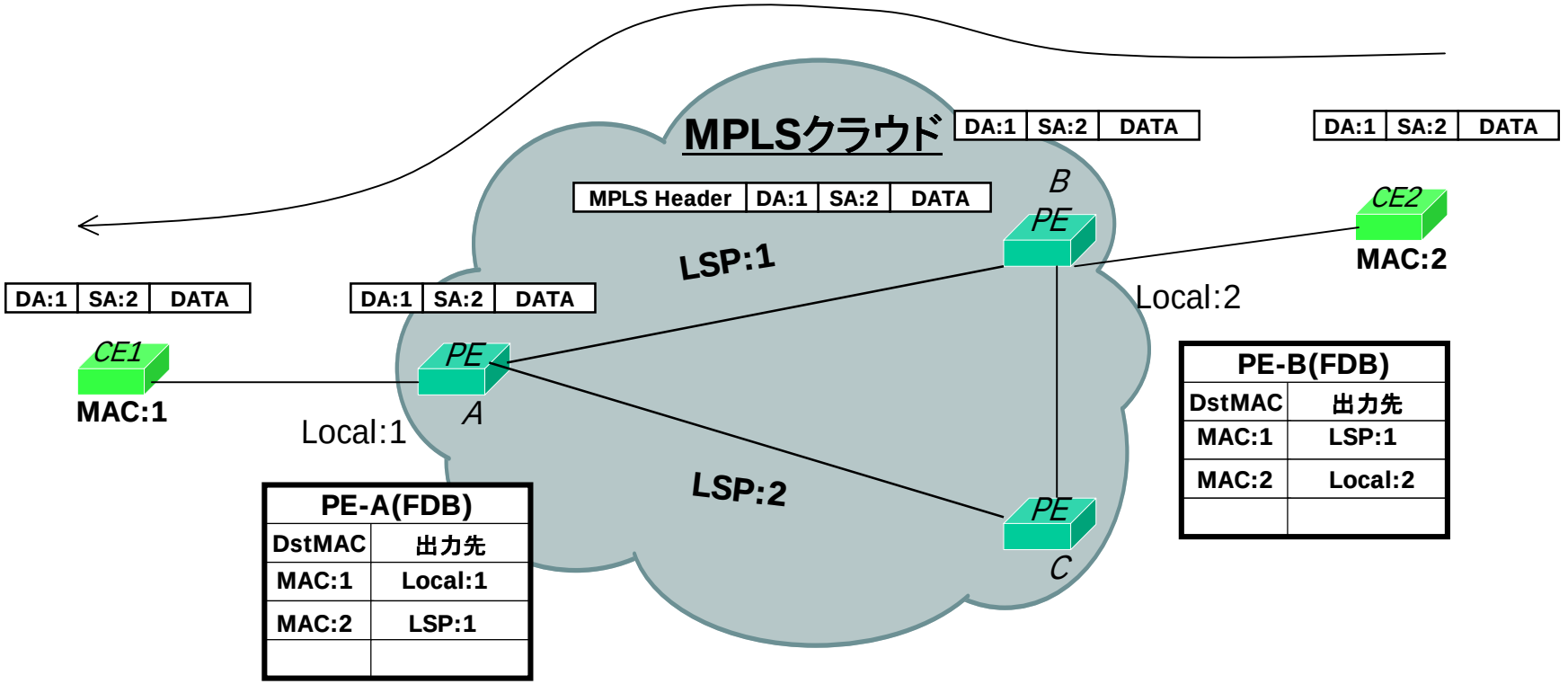
- イングレス（入力側）PEにてMACアドレスが学習されていない場合、PEはフレームをフラッドする。
- イグレス（出力側）PEにてMACアドレスが学習されていない場合、PEはローカルポートにのみ送信する。（Split Horizon)



VPLS(Virtual Private LAN Service)基本動作(2)

FDBにMACアドレスが学習されている場合のVPLS上のフレーム転送

- PEのFDBにMACアドレスが学習されている場合は、学習の内容にそって、フレームが転送される。



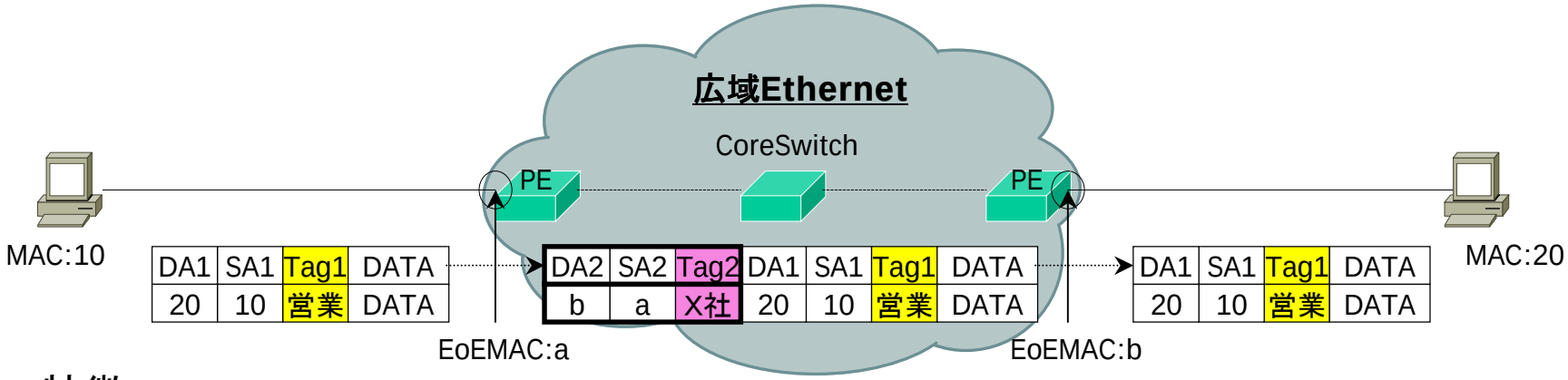
付録２：Ether over Ether（Interop発表資料より）

ループしたフレームを検出し、フレームを破棄する例

階層化ブリッジング(Ether over Ether)

802.1Q Tag VLANを使ったVLAN VPNの改良方式

- PEの加入者向けポートそれぞれにユニークなEoEMACアドレスを定義し、加入者から受け取ったEthernetフレームをその入力ポートに定義されたEoEMACアドレスをソースとし送り先のPEのポートのEoEMACアドレスをディスティネーションとするEthernetフレームでカプセル化して転送する方式。

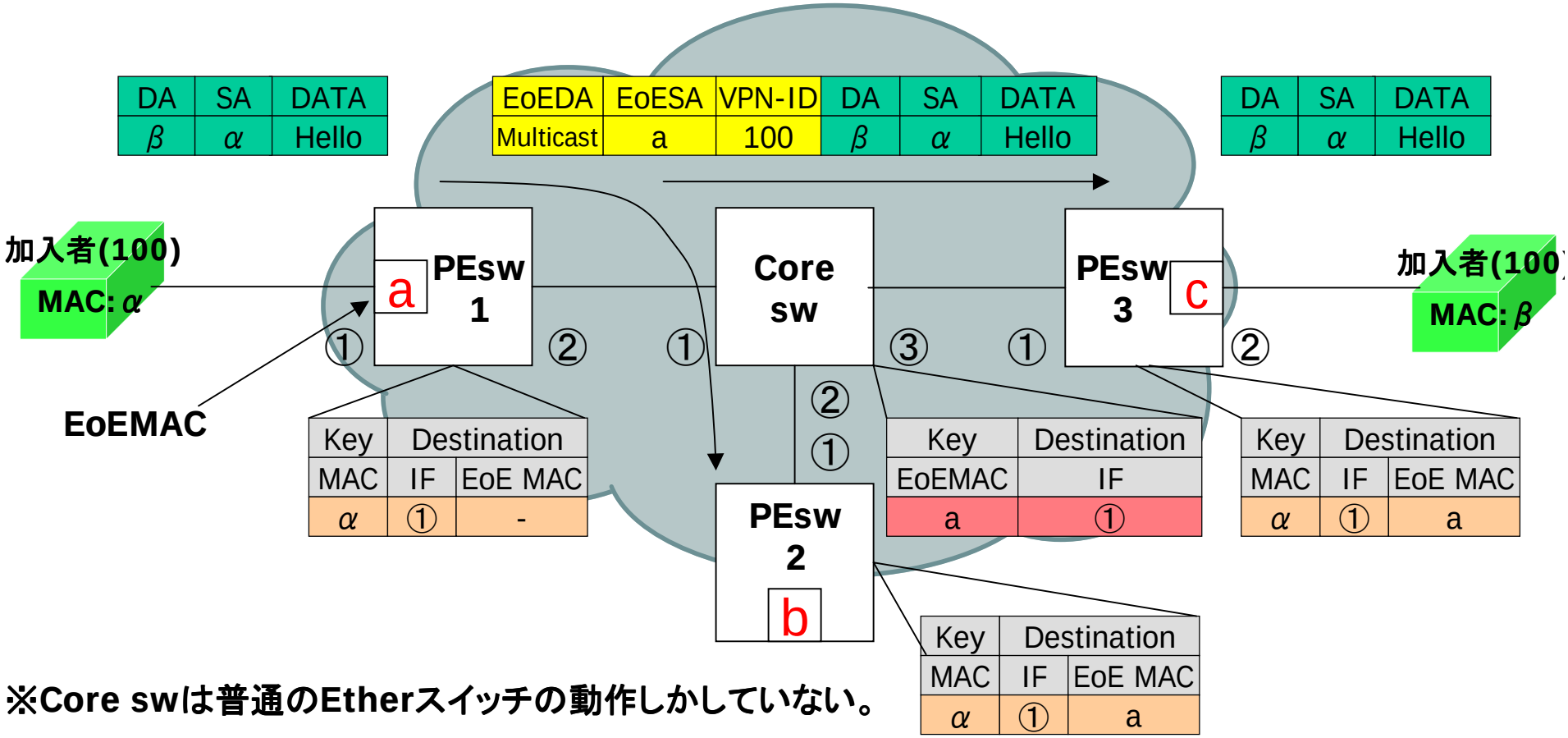


- 特徴
  - コアスイッチで学習しなくてはならないMACアドレスを劇的に減らす事が出来る。
  - EoEMACアドレスを階層的に割り振る事により、ループトラフィックが防止出来る。
  - 特殊な処理を意味するあて先MACアドレスを持つパケットを安全に転送する。
  - コアスイッチは単にジャンボフレームを転送出来る普通のスイッチでかまわない。（過去の資産の継承）

EoE(Ether over Ether)基本動作(1)

FDBにMACアドレスが学習されていない場合のEoE上のフレーム転送

- イングレス（入力側）PEにてMACアドレスが学習されていない場合、PEはEoEDAにマルチキャストアドレスをセットしフレームをフラッドする。

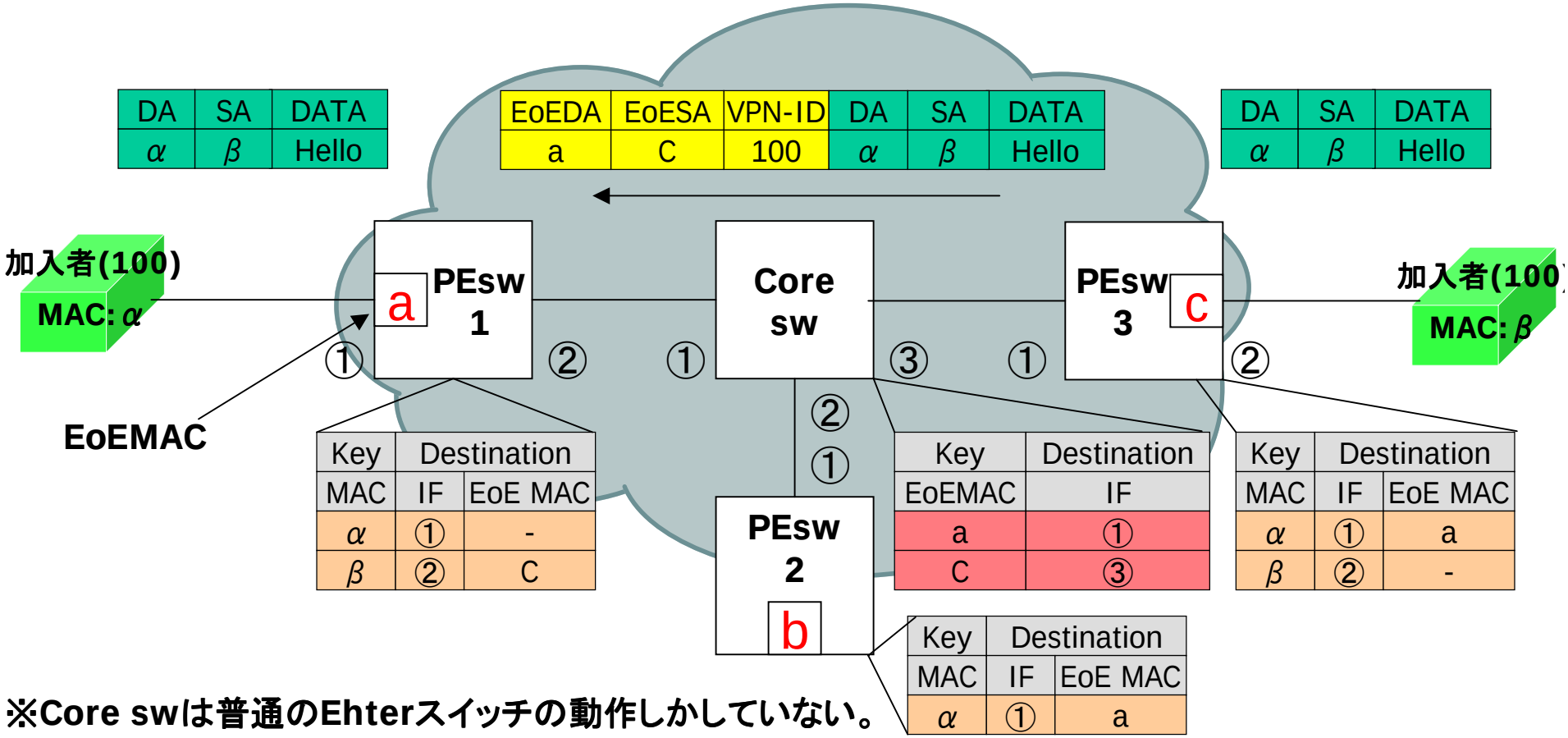


※Core swは普通のEtherスイッチの動作しかしていない。

EoE(Ether over Ether)基本動作(2)

FDBにMACアドレスが学習されている場合のEoE上のフレーム転送

- FDBにMACアドレスが学習されている場合は、学習の内容にそって、フレームが転送される。



※Core swは普通のEhterスイッチの動作しかしていない。

EoEフレームフォーマット

(802.1Qにカプセル化する場合)

7bytes	1bytes	6bytes	6bytes	2bytes	2bytes	2bytes	48～1520bytes or more	4bytes
プリアンブル	SFD	宛先EoE MACアドレス (DstEoE)	送信元EoE MACアドレス (SrcEoE)	TPID	TCI	EoE TPID	ペイロード (EoEフレーム)	FCS
プリアンブル								

※1: TPIDについては、802.1Qを利用する場合0x8100となりますが、VMANスイッチなどで転送する場合は0x9100を使用する事も出来ます。

※1

16bits	3bits	1bits	12bits	16bits
0x 8100	Pri	CFI	VPN-ID	0x E0E0

※2

※2:EoEデータフレームのTPID/Typeは0xE0E0になります。これにより、EoEカプセルフレームを判定します

(UntagEthernetフレームにカプセル化する場合)

7bytes	1bytes	6bytes	6bytes	2bytes	48～1520bytes or more	4bytes
プリアンブル	SFD	宛先EoE MACアドレス (DstEoE)	送信元EoE MACアドレス (SrcEoE)	TYPE	ペイロード (EoEフレーム)	FCS
プリアンブル						

EoE フレーム

※3: TTL(Time to Live)はEoE Awareなスイッチを通過するたびに減算され、0になったフレームは転送されず破棄されます。(最大255)

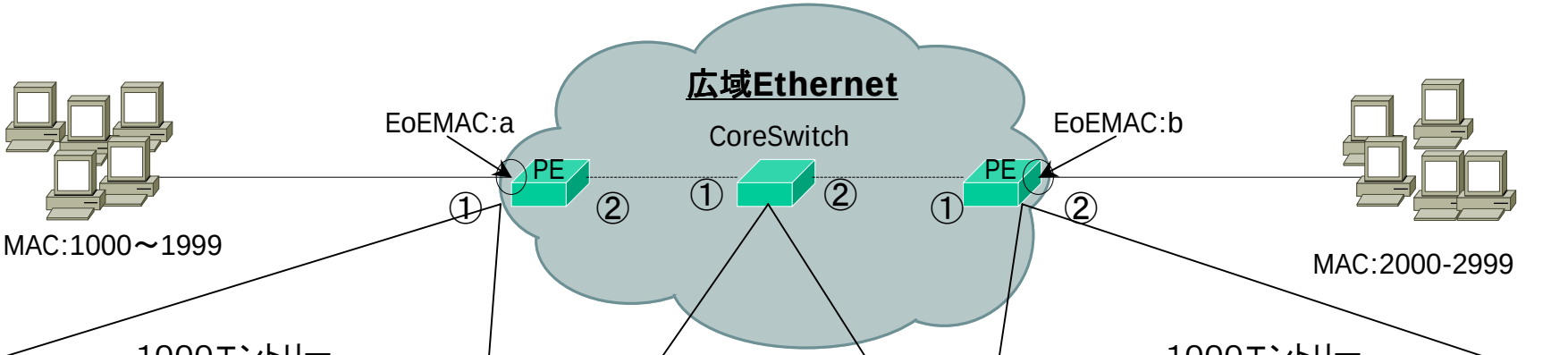
1bytes	1bytes	46～1518bytes or more
EoE TTL	予約	ペイロード (加入者フレーム)

※3

EoE ver1.0より

EoEによる、CoreSwitchでのMAC学習の低減

- 加入者が使用するMACアドレスの量が増えたとしても、コアスイッチが学習するMACアドレスの量は変わらない。（コアスイッチは数千VPNを扱う為、VPNごとのMAC学習数を減らしたい。）
- エッジスイッチでは、多くてもポート数程度の数のVPNしか存在しない為、スケールし易い。



1000エントリー

Key	Destination		Key	Destination	
MAC	IF	EoE MAC	MAC	IF	EoE MAC
1000	①	-	2000	②	b
1001	①	-	2001	②	b
...	①	-	...	②	b
...	①	-	...	②	b
1999	①	-	2999	②	b

2エントリー

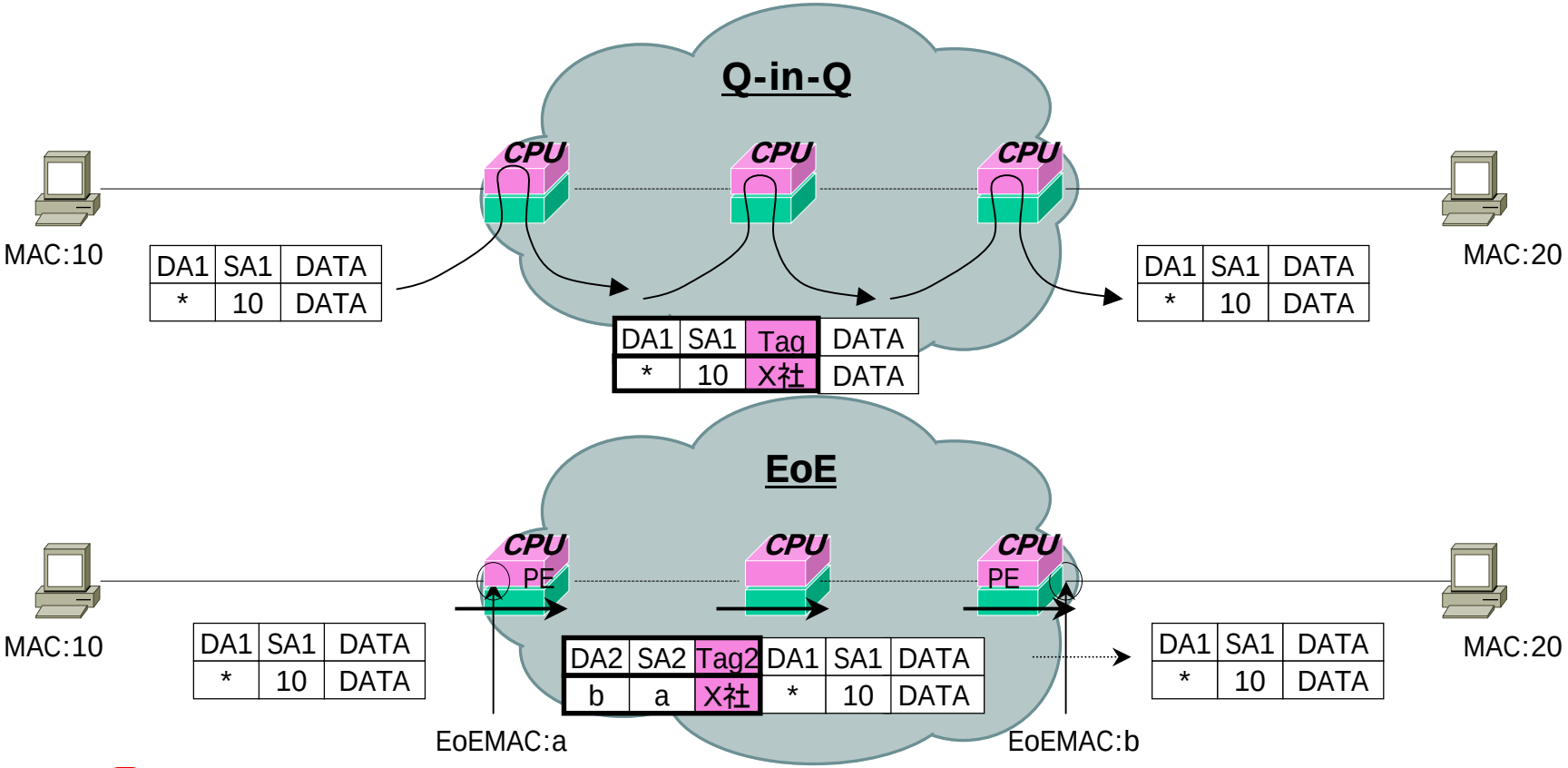
Key	Destination
EoEMAC	IF
a	①
b	②

1000エントリー

Key	Destination		Key	Destination	
MAC	IF	EoE MAC	MAC	IF	EoE MAC
1000	①	a	2000	②	-
1001	①	a	2001	②	-
...	①	a	...	②	-
...	①	a	...	②	-
1999	①	a	2999	②	-

EoEによる制御パケットの安全な透過

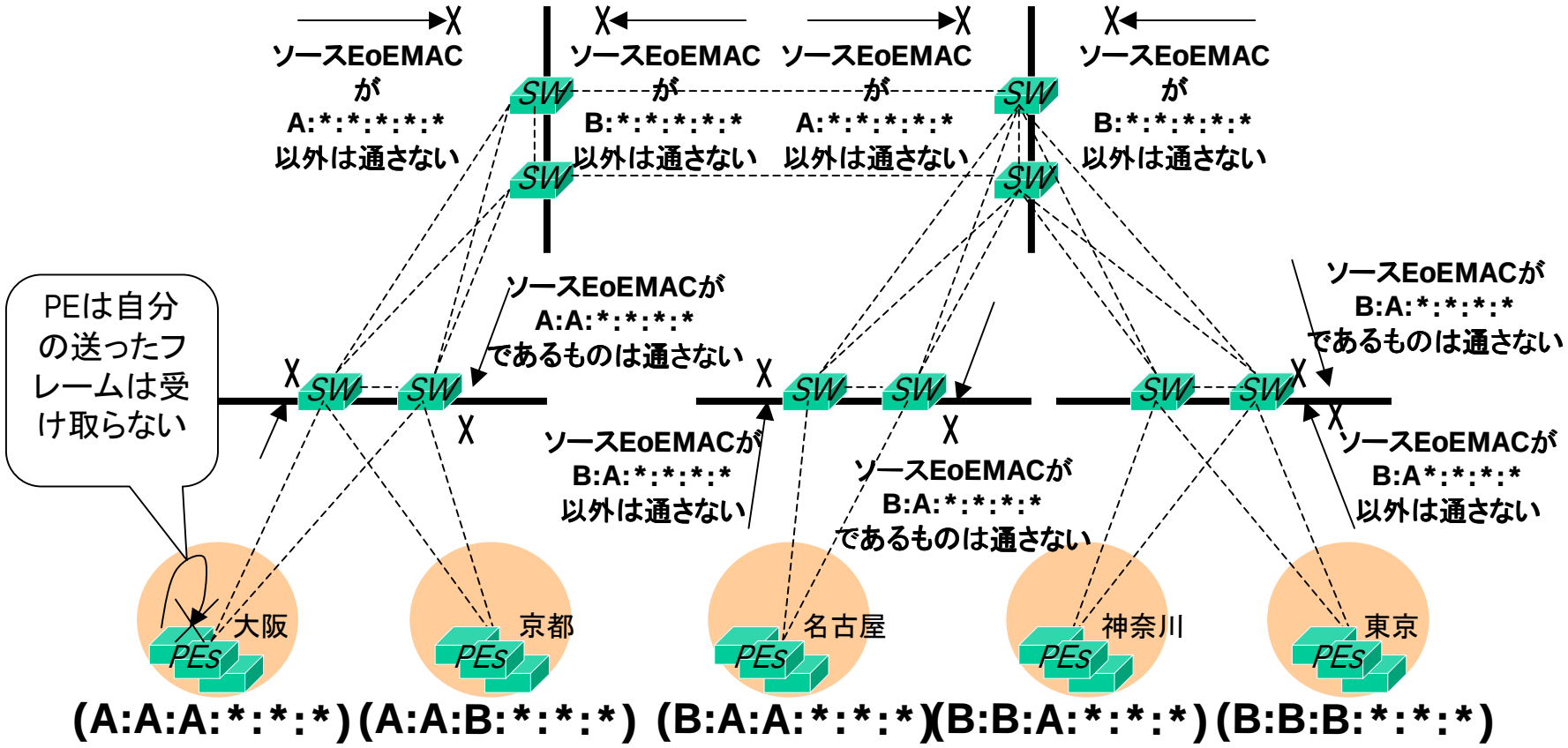
- Q-in-Qでは、特定の制御プロトコルと同じマルチキャストアドレスを宛先に持ったフレームをユーザが送信すると、スイッチのCPUに転送されたり、ブロックされる場合がある。
- EoEではユーザが送信したマルチキャストのアドレスは隠蔽されるので、CPUに転送されず、透過する。



EoE階層化MACアドレッシングによるループ防止(1)

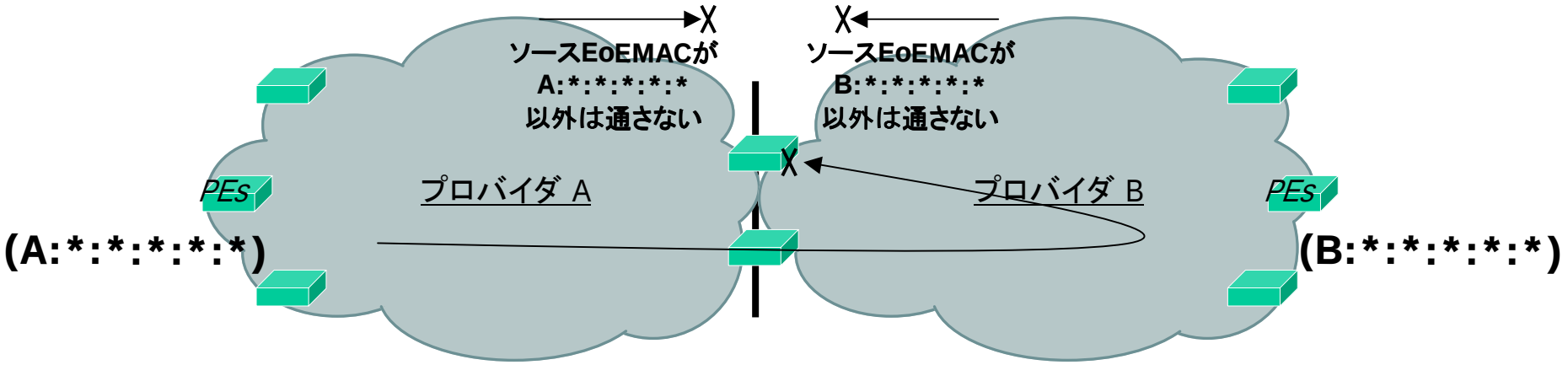
EoEMACアドレスを階層的に割り振る事により、ループの発生を防止する

- EoEMACアドレスを階層的にふり、マスク付きMACアドレスフィルタを使って、ループを防止出来る。(ストリクトなフィルタの例)

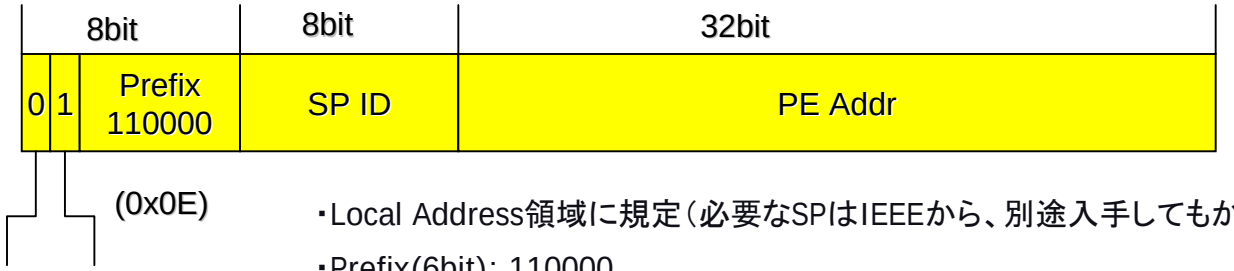


EoE階層化MACアドレッシングによるループ防止(2)

- サービスプロバイダ相互接続点



サービスプロバイダEoE MACアドレスの構造



・Local Address領域に規定（必要なSPはIEEEから、別途入手してもかまわない）

・Prefix(6bit): 110000

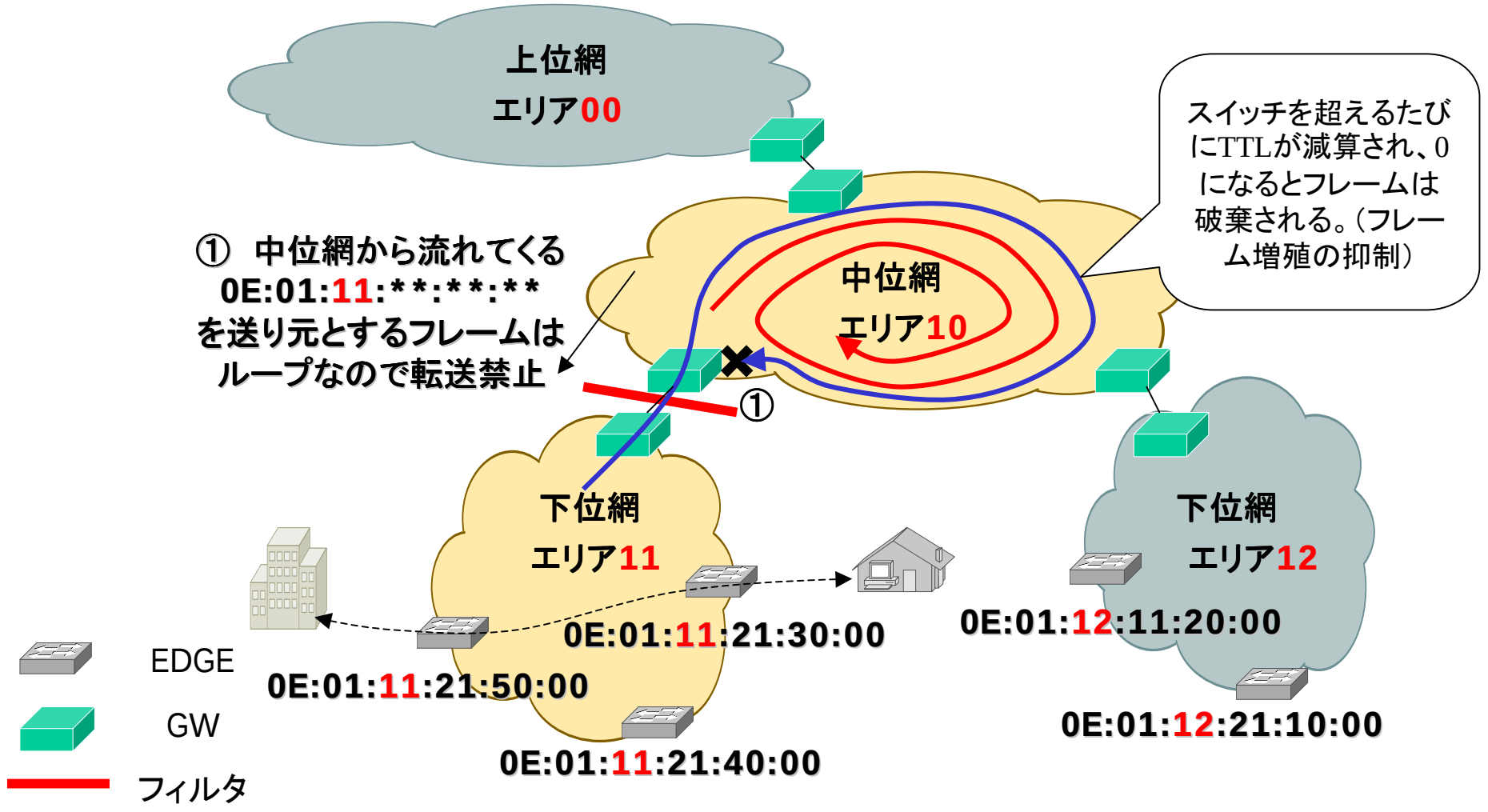
・SP ID(8bit): Service Providerの識別番号

・PE Addr: Provider Edgeスイッチのアドレス

PEスイッチそのもの、又はPEスイッチの内部の1エンティティ(ラインカードやポート)を表す場合が有る (使用方法はService Provider毎に任意)

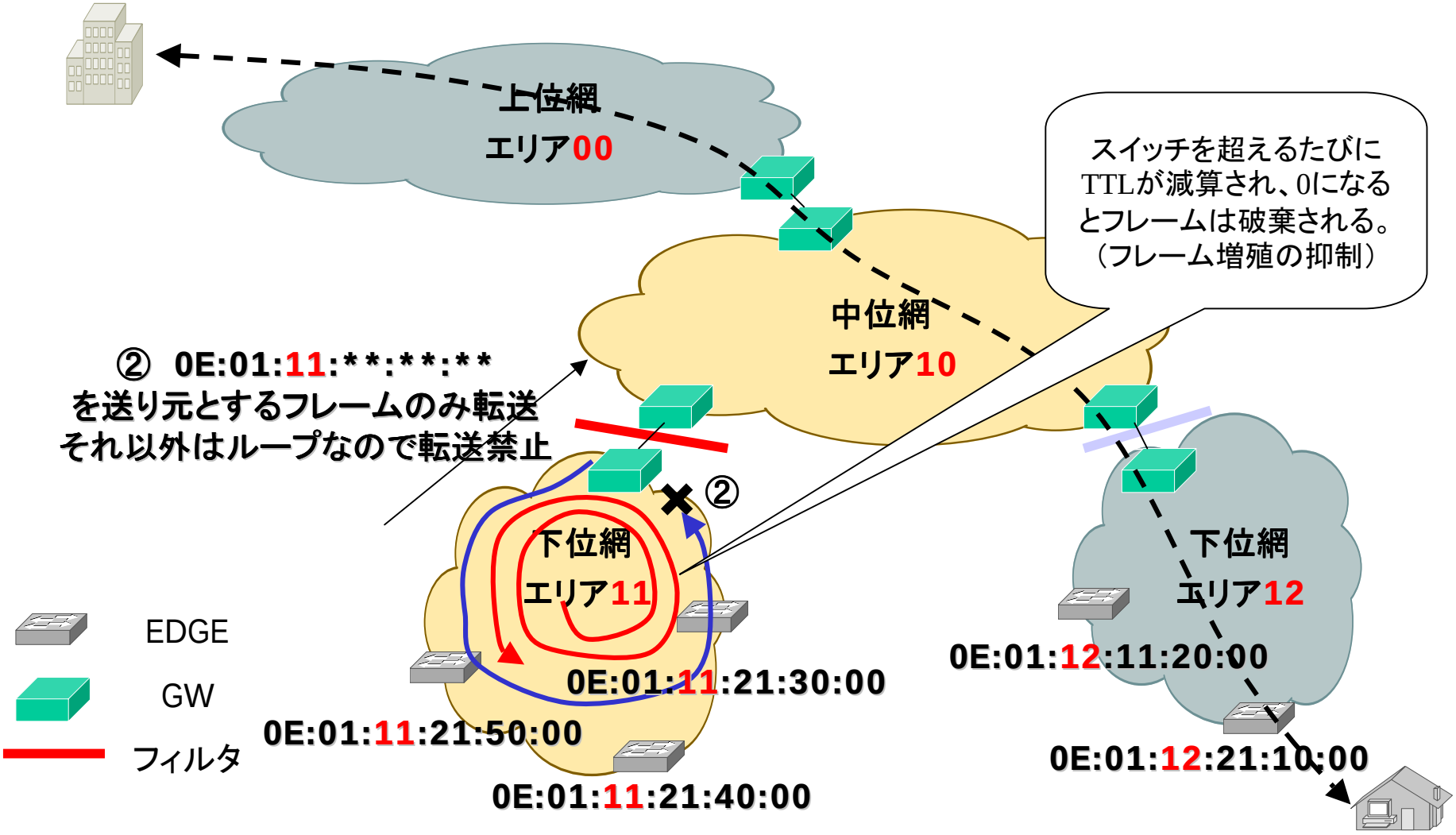
ループ発生時の流入防止（ルーズなフィルタの例）

中位網でループが発生しても下位網内部の通信には影響を与えないようにフィルターを設定する。



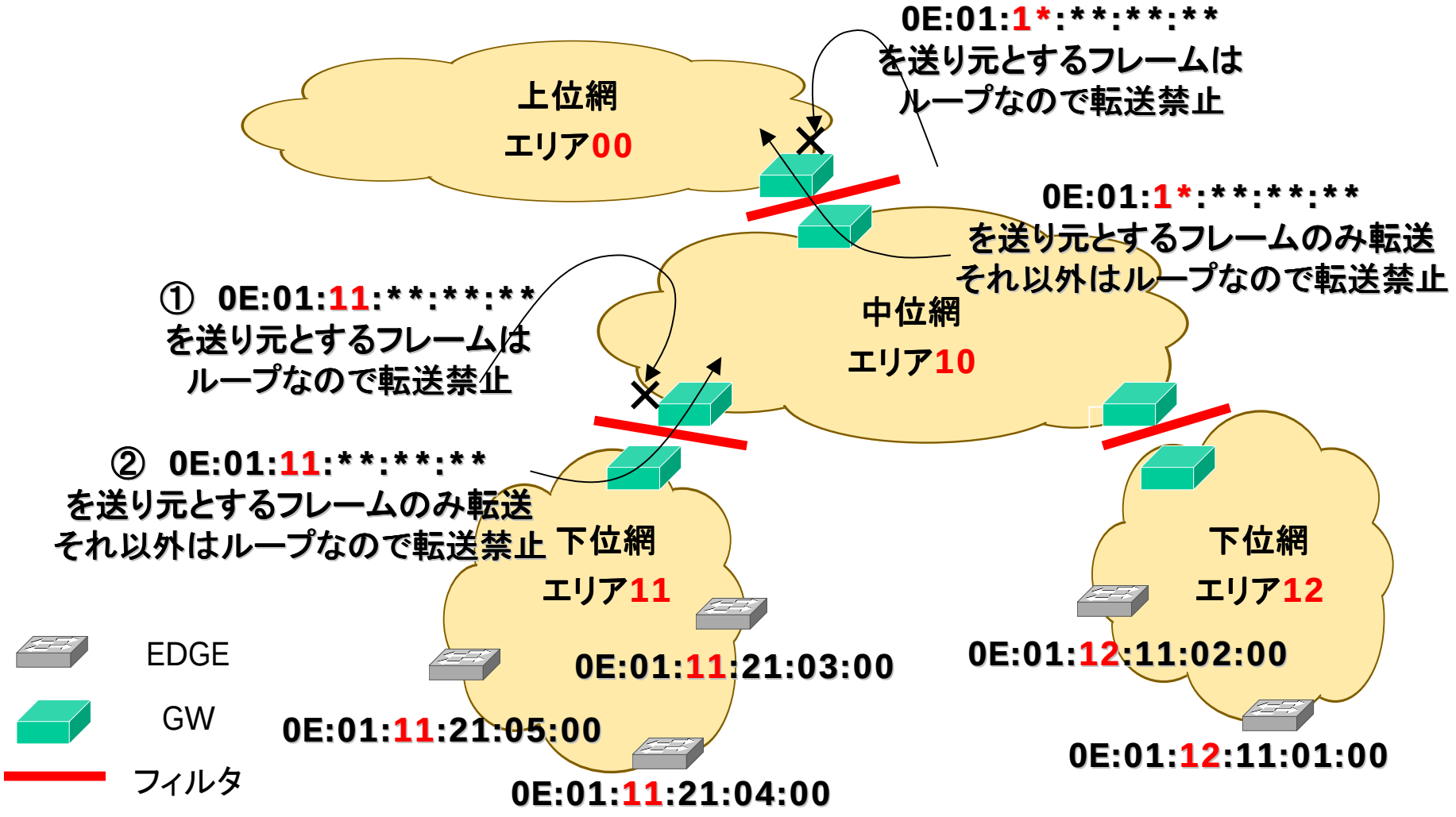
ループ発生時の流出防止（ルーズなフィルタの例）

下位網でループが発生しても他網の通信には影響を与えないようにフィルターを設定する。



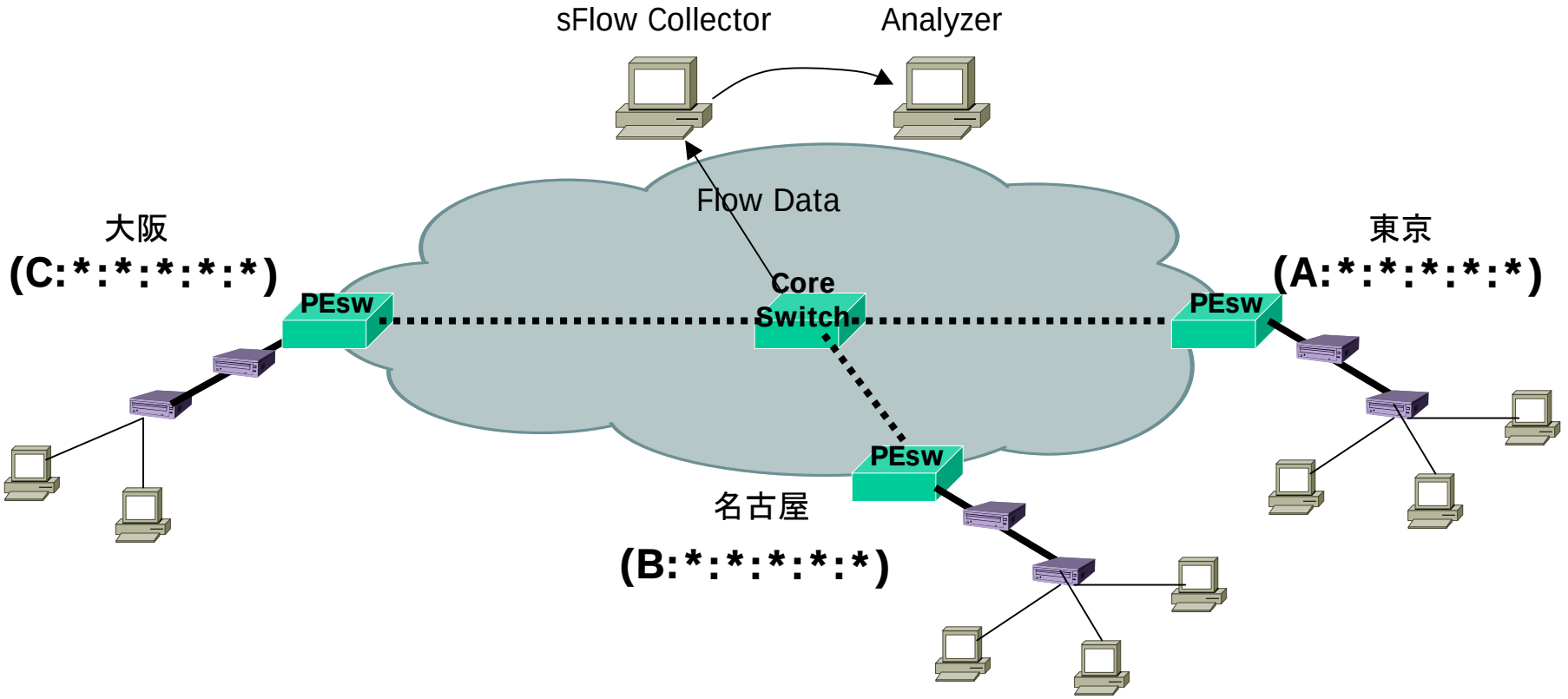
階層エリアのフィルタ（ルーズなフィルタの例）

2種類のフィルタの組み合わせにより、ループの影響範囲の限定を行ってみる例。



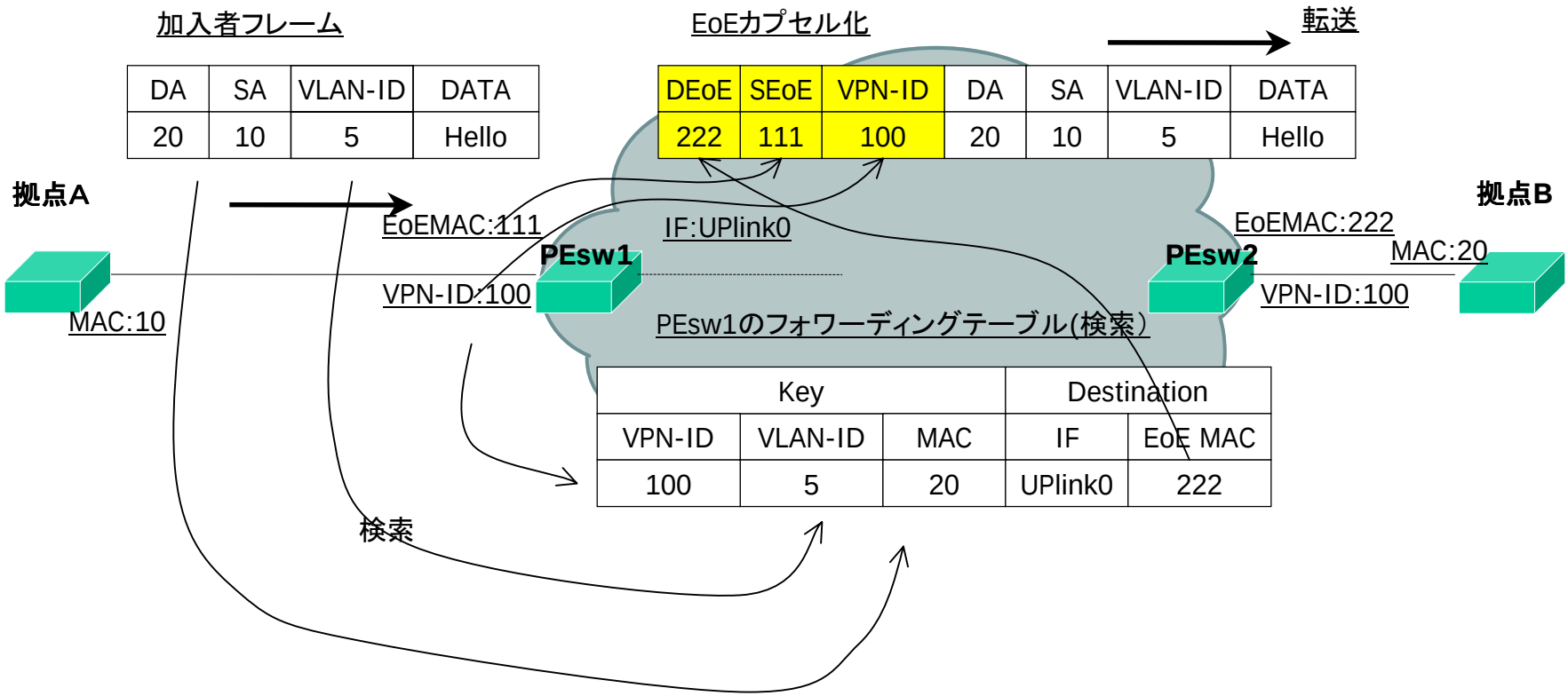
EoE階層化MACアドレッシングとsFlow（NetFlow）

- 階層化MACアドレッシングとsFlowの様なトラフィック分析システムを組み合わせると、地域間のトラフィックの流れを分析する事が出来る。
- 大阪と東京の間のトラフィックがどれくらいあるのか？など。



EoEとTag学習について

- 一般にQ-in-Qでは、加入者が設定したVLAN内で、MACアドレスの重複は許されない。
- EoE では、エッジでMACと共にTag-IDを学習する事によって、加入者の設定したVLAN内でのMACの重複を許容する方式を導入。



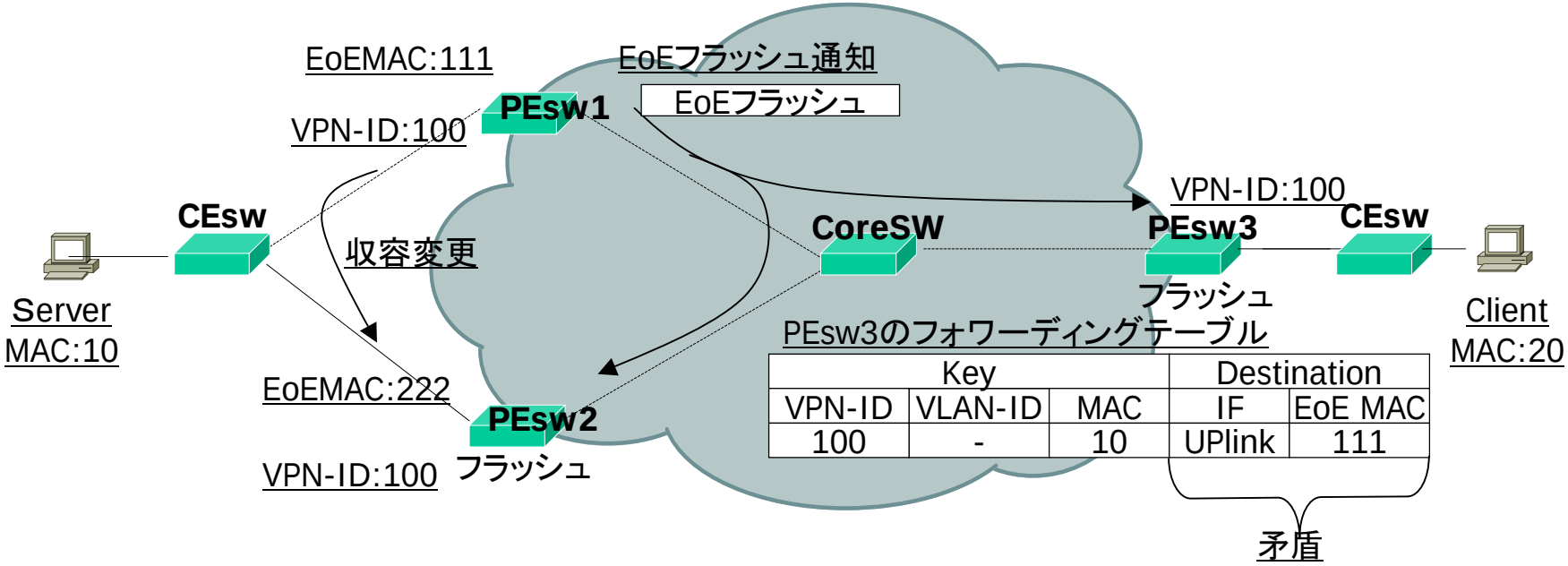
EoE OAM

- EoE　Ping

  PE SW間あるいはEoE終端機能を持った装置間のPing（疎通確認）
- EoE Traceroute

  TTLを用いて EoE経路のトレースを行う（専用の中継Ethernet SWが必要）
- EoE フラッシュ

  EoEのFDBの全部又は一部のエントリのクリア要求をPEスイッチでやりとり



Key			Destination	
VPN-ID	VLAN-ID	MAC	IF	EoE MAC
100	-	10	UPlink	111

Questions？

masaty@pwd. ad. jp