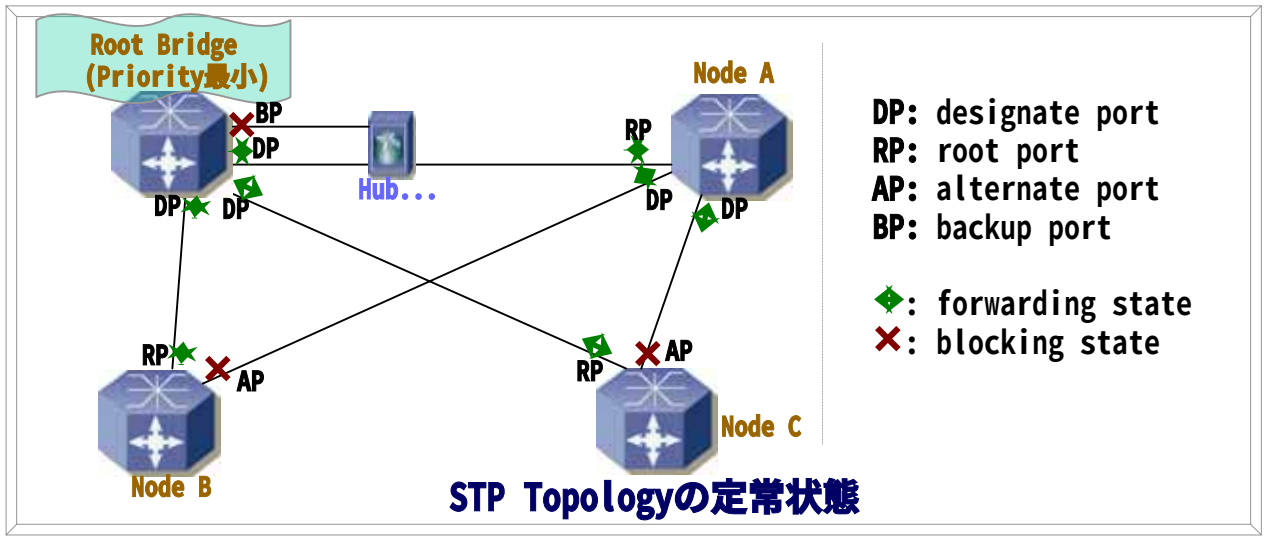
広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

- STPとの戦い(Loop防止)
  - ✓ Loop Guard
  - ✓ Root Guard

- VLAN...
  - ✓ VLAN数の増大を促す潜在的な原因
  - ✓ Tag多重/開放の要望
  - ✓ VLAN数増大による管理面について

広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

- STP（Spanning Tree Protocol）のTopologyについて
  - Root Bridgeについて
    - ✓ Spanning TreeのTopologyにおいては、必ず一つのRoot Bridgeが存在する
    - ✓ Root BridgeはBridge Priorityが最小のものが選出される
  - Nodeにおける各Portの役割について
    - ✓ Designated Port:: Forwarding state
      各Nodeからみて、Root Bridgeへのコストが最小である経路となるPort
    - ✓ Root Port :: Forwarding state
      各Nodeで、Root Bridgeへのコストが最小となるPort
    - ✓ Alternate Port:: Blocking state
      各Nodeで、Root Bridgeへの経路となるPortのうちRoot PortでないPort
    - ✓ Backup Port :: Blocking state
      各Nodeからみて、Root Bridgeへの経路となるPortで、Designated PortでないPort
      (非常に稀なCaseとして存在)



STP Topologyの定常状態

広域イーサネットでの実運用においての改善点と問題点 ～信頼性向上へ向けて～

e-VLAN

❋ STPにおけるLoop防止策

➢ Loop Guard (Cisco)の採用

✓ Loop Guardとは? = CPU異常や単方向Link障害時のLoopを文字通り防ぐ機構 =

- 正しいRoot Bridge情報を含んだBPDUを一定時間(max ageの設定値)に受信できなかった場合のAlternate Portの動作は、

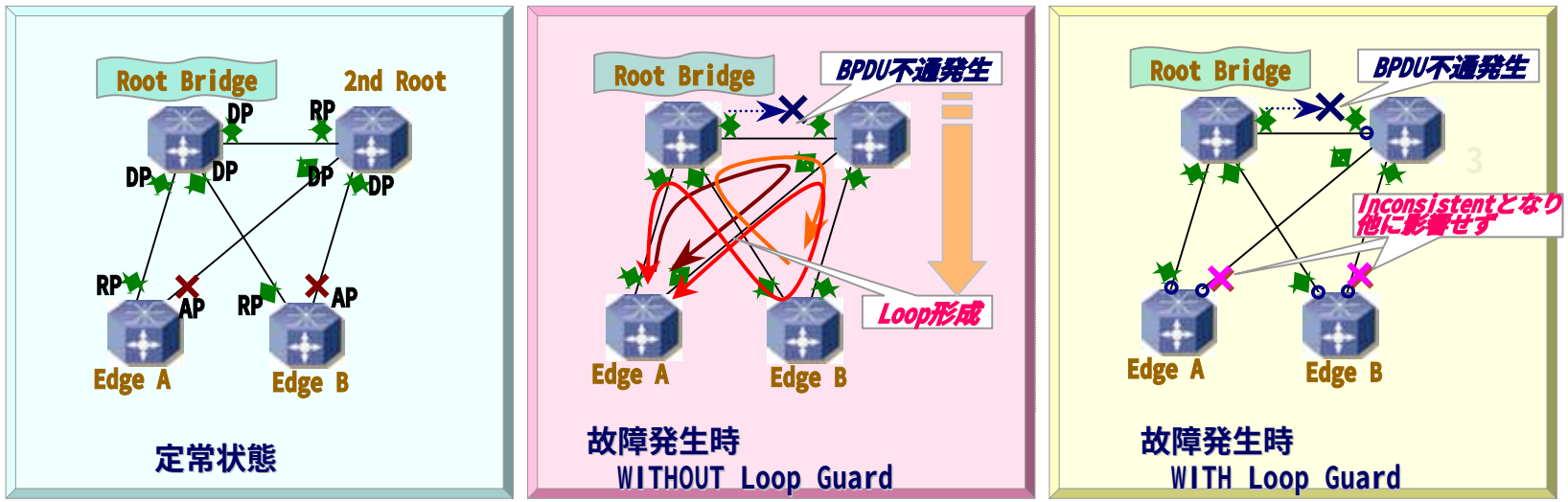
= WITHOUT Loop Guard

Listening → Learning → Forwardingへと状態遷移

= WITH Loop Guard

inconsistent (Blocking = NOT Forwarding) へと状態遷移

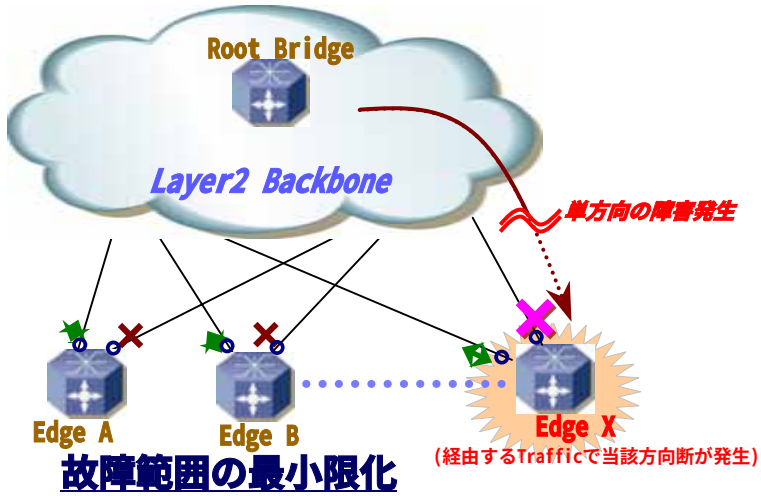
となり、Loop形成の抑制が可能



広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

e-VLAN

- STPにおけるLoop防止策
  - Loop Guard (Cisco)の採用
    - ✓ 適用効果
      - 故障範囲を最小限にとどめる事は実現できる…と確信

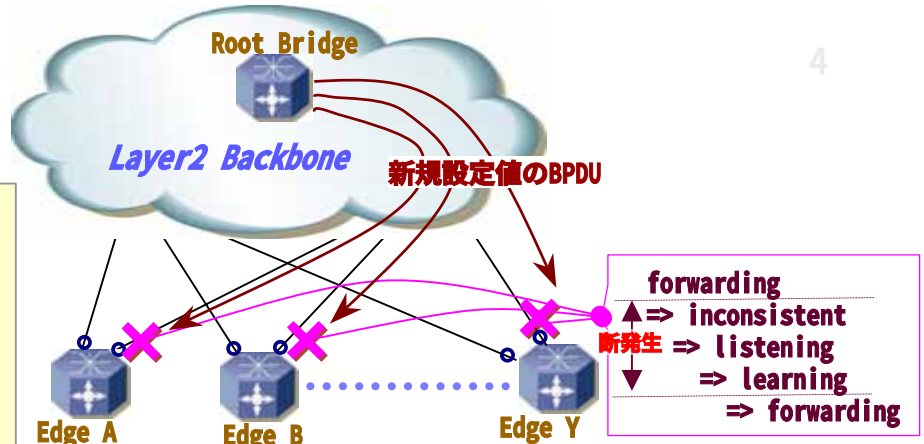


故障範囲の最小限化

1: Layer1での単方向Link障害が発生
2: 障害の事実を検知/対処するまでの間、当該Linkの当該方向のTrafficはLost
3: が、Loop Guardの効能でLayer2 Loopの形成は免れる

- 弊害とまでは言いませんが、、

  構成変更等の作業では追加項目が要



追加作業が必要...

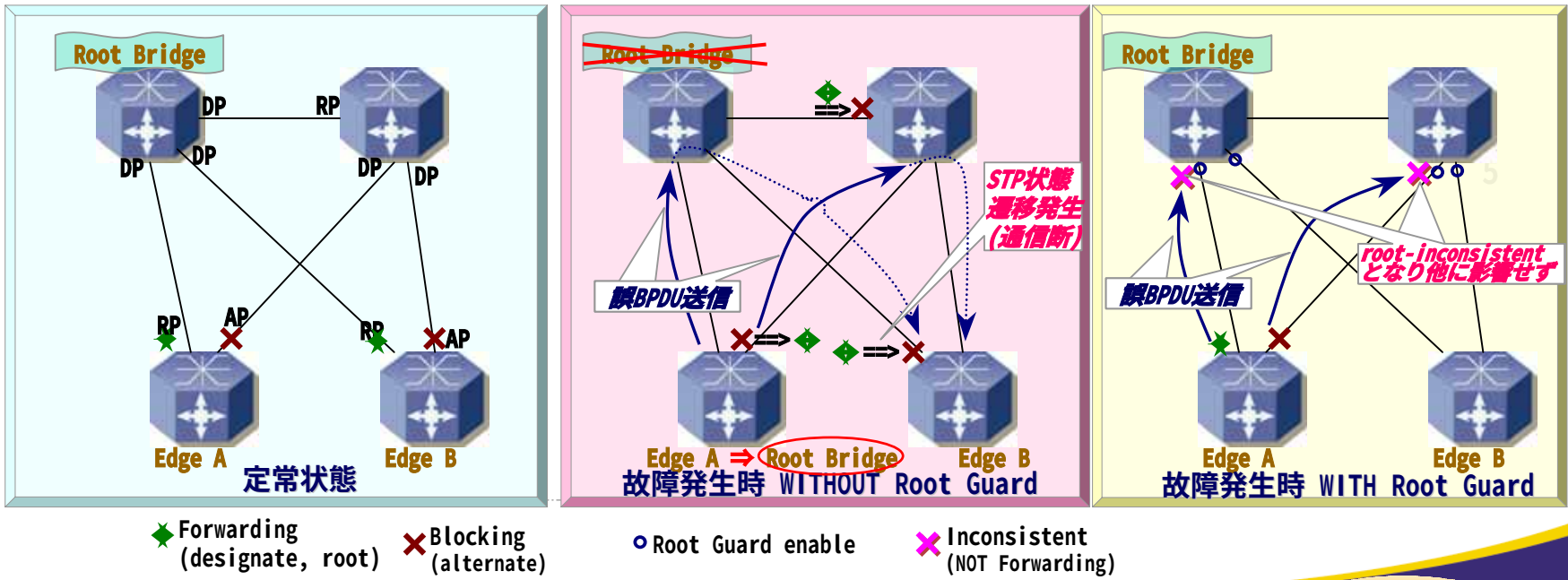
例としてRoot Bridge STPパラメータを変更する場合
= パラメータ値変更後の値に合わせたBPDUを送信
= 右図の各Edgeでは、Loop Guardを有効にしたままだとinconsistentとなり、forwardingまでの間、通信断となってしまう

→上記を防ぐためには、Loop Guardを一旦無効化させ、パラメータ値の変更を実施した後に再度Loop Guardを有効化させる必要がある

広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

e-VLAN

- STPにおけるLoop防止策
  - Root Guard (Cisco)の採用
    - ✓ Root Guardとは? = Root Bridgeの配置を固定させるための機構 =
      - Root Bridge以外のNodeから誤ったBridge PriorityのBPDUが送信された場合
        = WITHOUT Root Guard
        Root Bridgeが変更され、Root Bridgeへの経路が変更
        →他NodeにてSTPの状態遷移(⇒通信断)が発生
        = WITH Root Guard
        root-inconsistent (Listening = NOT Forwarding) へと状態遷移
        →他Nodeへは影響を及ぼさない



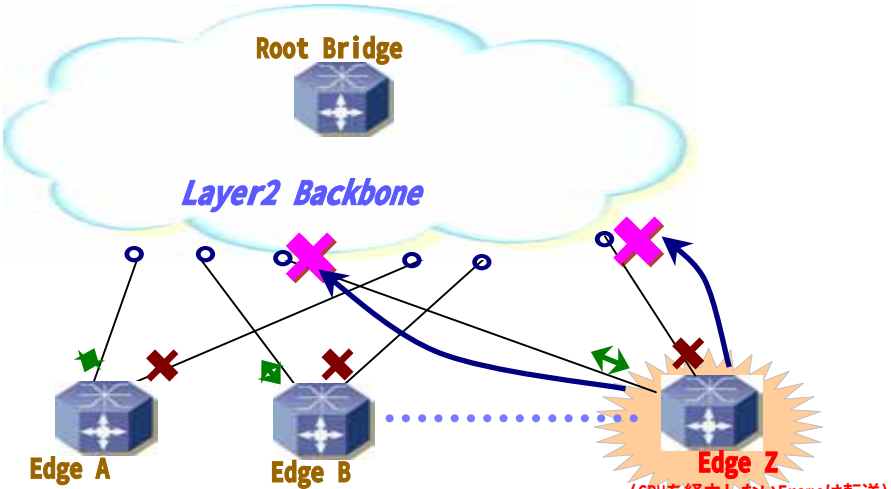
広域イーサネットでの実運用においての改善点と問題点 ～信頼性向上へ向けて～

❋ STPにおけるLoop防止策

➢ Root Guard (Cisco)の採用

✓ 適用効果

- **故障範囲を最小限にとどめる事は実現できる･･･**と確信



故障範囲の最小限化

1: Edge ZでCPU異常が発生
2: Edge Zでは、STP処理不能に陥る(CPU経由しないFrameについては転送可能)
3: STP処理不能に陥ったため、誤ったBPDUをUp Link側へ送出
4: Root Guardにより、root-inconsistentとなり網全体のTopologyは崩壊せず
(Edge Zを経由するTrafficについては、CPU異常が復旧するまでLost)

広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

e-VLAN

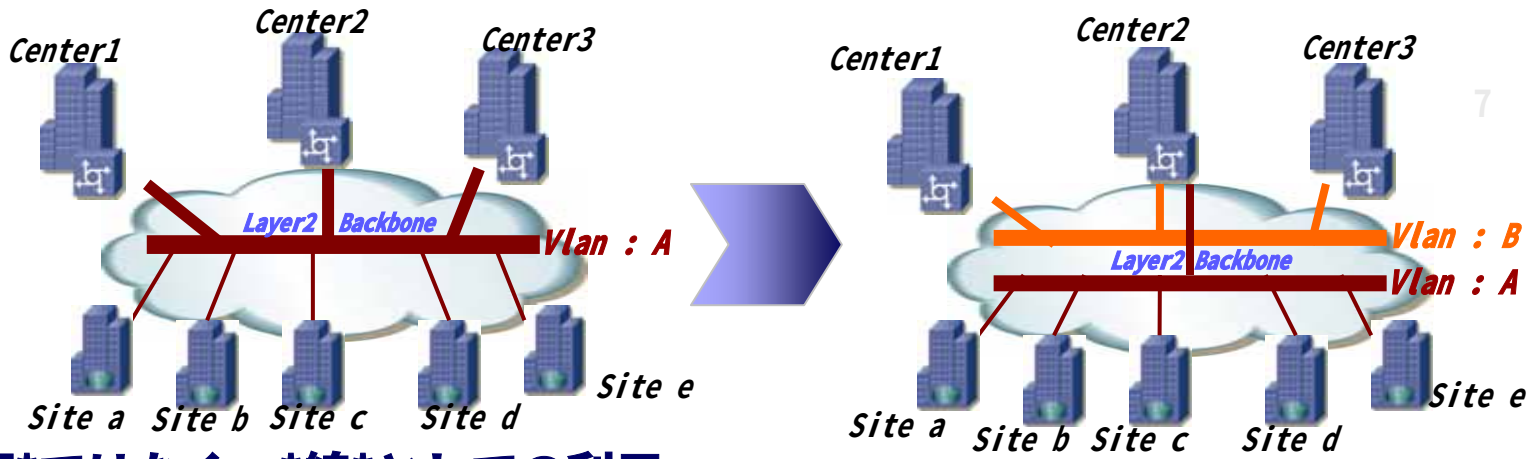
❋802.1q Tagged-VLANでの問題点

➢VLAN数の増大を促す潜在的な原因

✓ 帯域設計におけるMulticast/Broadcastの抑制

- 広域Ethernetは、Layer2のフラットな構成
- 高速～低速のアクセス品目に関係なく同一のBroadcast Domain
- Broadcast/Multicast(*後述)が全てのSiteへ送信されてしまう
- 低速アクセス品目Siteへの無駄なTrafficを抑制するために
  高速アクセス品目Siteと低速アクセス品目Siteを分離

☆ **VLAN数を余分に使用してしまう**



✓ *面*ではなく、*線*としての利用
極端な例として、完全にLayer3のためのPoint-to-Pointの形態も・・・

広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

e-VLAN

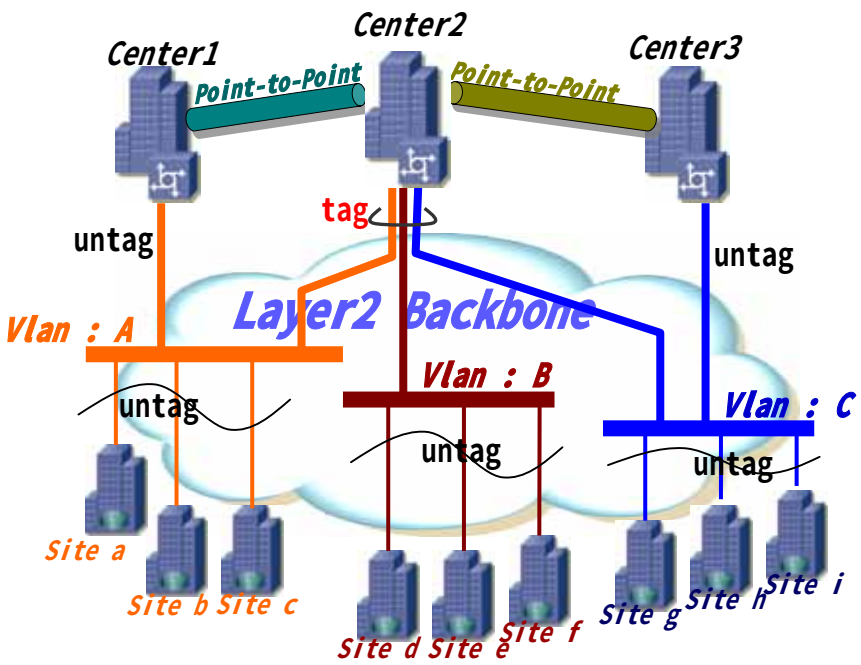
✵802.1q Tagged-VLANでの問題点

➢Tag多重/開放の要望

- ✓vlanの分割においては、Centerはtag多重させたい
- ✓一般のSiteでは、運用/管理負荷の軽減のためuntagで繋ぎたい
- ✓可能であればCenterのvlan IDは、Userで任意に設定したい
  ⇒キャリア側でつけるvlan IDそのままの提供希望もあり

==Tag多重/開放の例==

```
Center2(tag)  :vlan A,B,C
-------------------------
Center1(untag):vlan A
Site a (untag):vlan A
Site b (untag):vlan A
Site c (untag):vlan A
- - - - - - - - - - - - -
Site d (untag):vlan B
Site e (untag):vlan B
Site f (untag):vlan B
- - - - - - - - - - - - -
Center3(untag):vlan C
Site g (untag):vlan C
Site h (untag):vlan C
Site I (untag):vlan C
```



NTT Communications

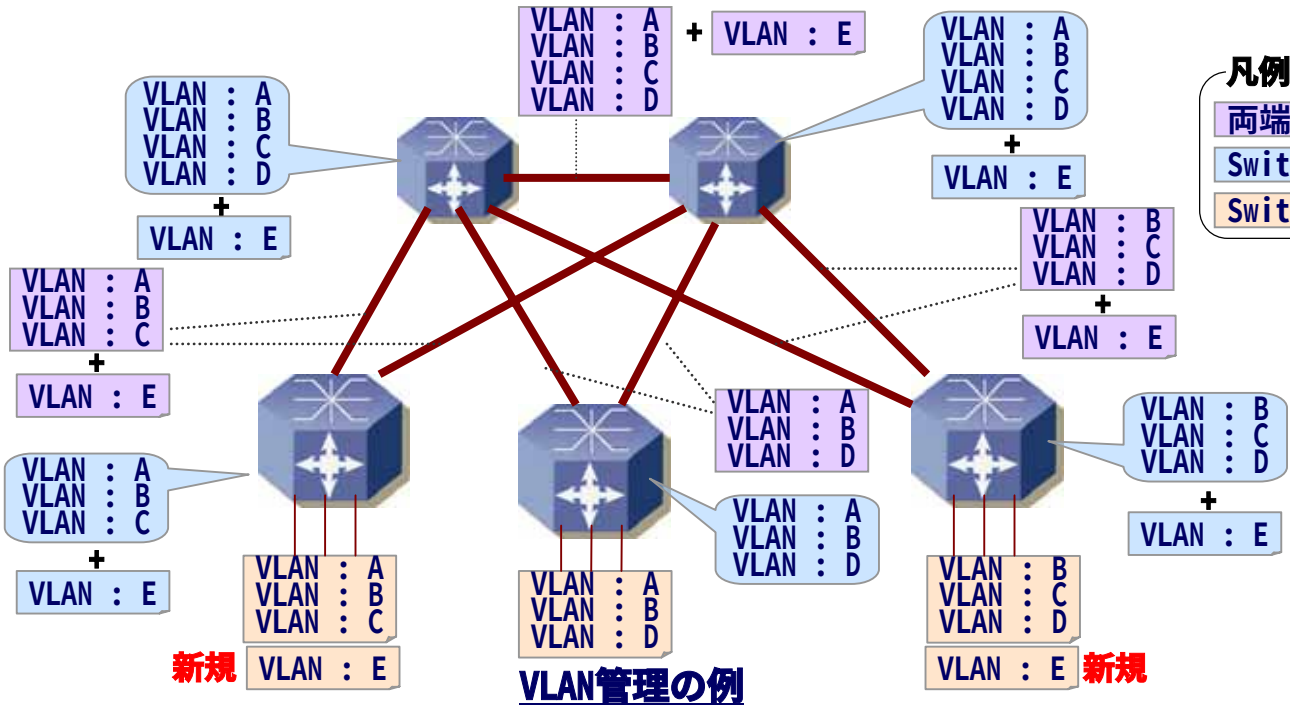
広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

e-VLAN

✵802.1q Tagged-VLANでの問題点

➢VLAN数増大による管理面について

- ✓VLAN毎に経由する/しないSwitchの管理(設定)が必須
- ✓同様にL2 Backbone内で経由するPortの管理(設定)も必須
  ⇒常にVLAN毎のTopology管理が必要不可欠となっている
  市販の運用Toolもあるにはあるが、カスタマイズは必要
- ✓VTP(VLAN Trunking Protocol:Cisco社)はSwitch間でVLAN管理を実現
  ⇒機種依存/VLAN数に制限があり、全ての問題解消には至らず…



VLAN管理の例

NTT Communications

広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

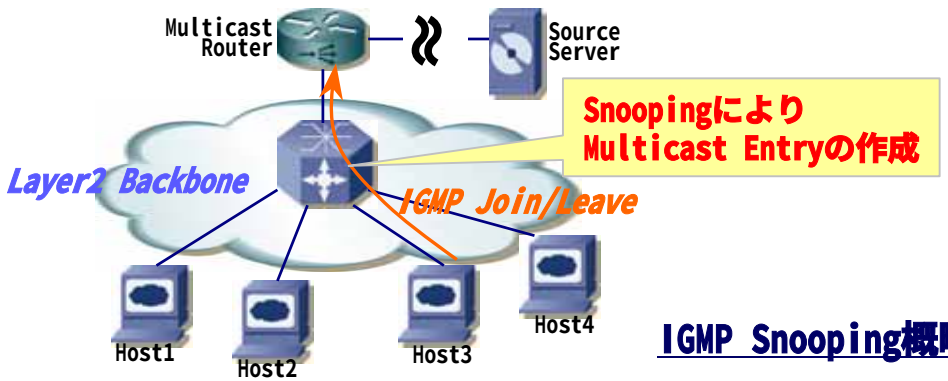
e-VLAN

✵L2での”美しい”Multicastの限界

➢“美しい”(IP)Multicastの転送を実現するには、、、

L2 SwitchでのMulticastの転送は、
= MAC Address Entry有り:　該当Portへ転送
= MAC Address Entry無し:　同一VLAN内へFlooding
L2 SwitchでEntryを作る(美しい転送をする)には、
* IGMP Snooping
* RGMP (Cisco社)
の活用が必要となる

✓ IGMP Snooping
Multicast RouterとHostとのIGMP ReportをSnoopしEntryを作る



IGMP Snooping概略

* CPU Resourceを一部使ってEntry作成
⇒そのため、動作上の不具合もいくつか報告あるが確実に改善中↑
* 実際のUser環境下ではほぼ全SiteでRouterが繋がるため有効度は低…

広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

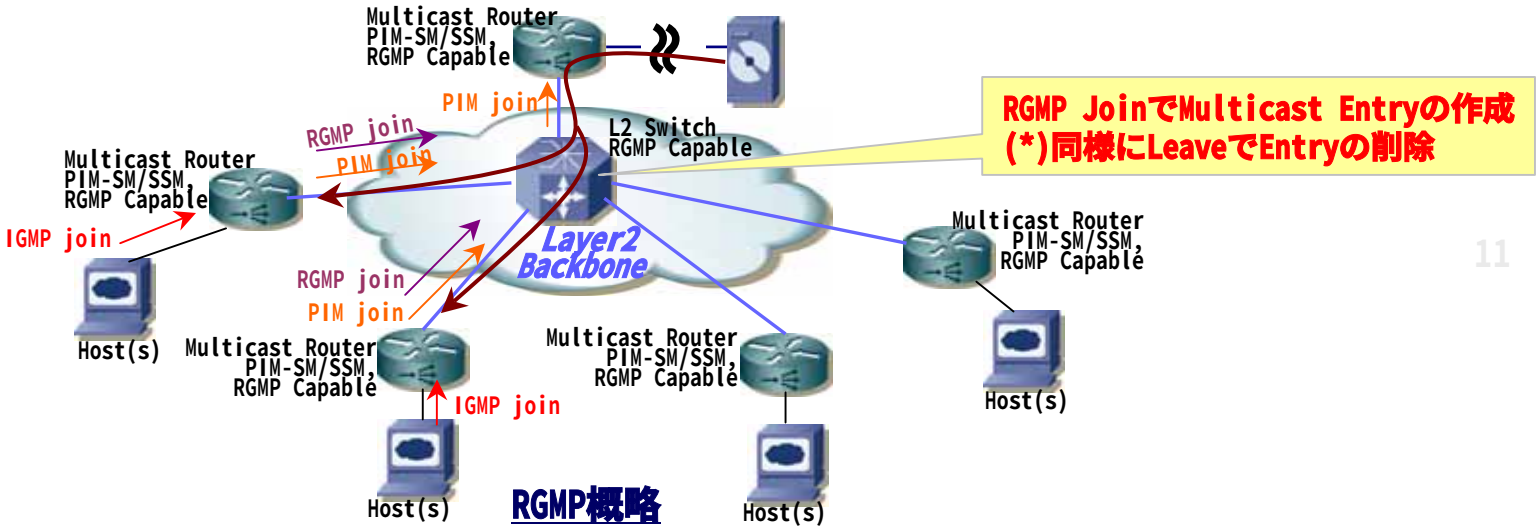
e-VLAN

✵L2での”美しい”Multicastの限界

- ➢“美しい”(IP)Multicastの転送を実現するには、、、

  Multicast Routerに囲まれたL2上ではRGMPが有効 ではある....

  - ✓RGMP--Router Group Management Protocol--(Cisco社)の例
    Mcast Routing ProtocolがPIM-SM/SSMの場合のみ活用可能
    当然ですが、、、RGMP Capable RouterとSwitchで動作する
    ⇒Cisco社RouterとSwitchの組合せが必須 ___ なので汎用度が低



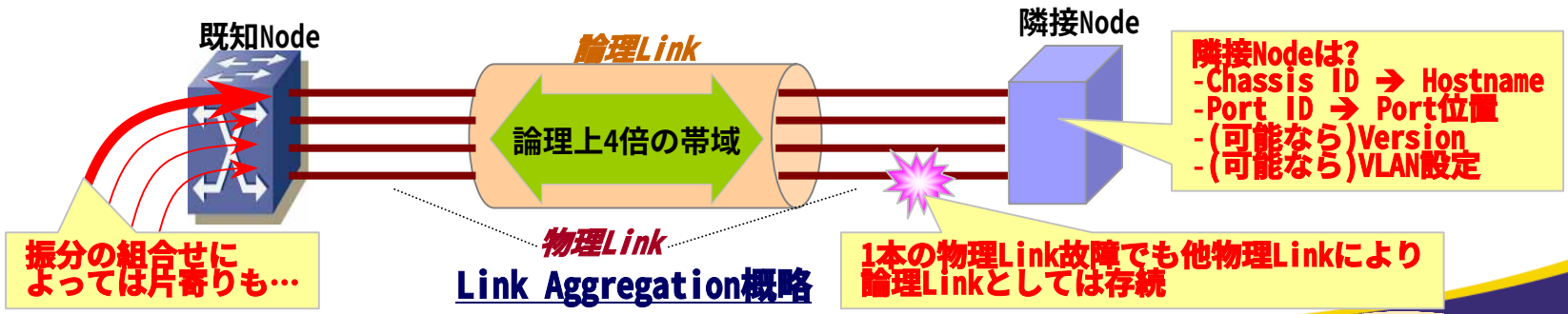
RGMP概略

  - ✓結局IP LayerのAssistが必要
    現状では美しく流すことには制限があり、VLAN分割 + IPのAssistが要

広域イーサネットでの実運用においての改善点と問題点
～信頼性向上へ向けて～

e-VLAN

- Link Aggregationにおいて...
  - Link Aggregationとは、
    Link Aggregation:複数の物理Linkを一つの論理Linkとみなす
    - ✓ 耐障害性の向上
      1つの物理Linkの故障でも論理Linkとしては確立しているため、
      完全断までは発生しない
    - ✓ 有効帯域の確保
      複数の物理Linkを束ねることにより帯域は × n倍
  - Link Aggregationで内在する問題
    - ✓ 設定でのコツ…
      - AggregationさせるPort(Link)間の論理設定の順番/同一設定が肝
      - 隣接Nodeとも当然ですが論理設定を同一にさせる事は
        ⇒お隣さんは誰? を動的に管理できる仕組みを熱望(e.g. CDP,EDP⇒LLDP)
    - ✓ 論理Link内での振り分け方法の潜在的な問題
      - IPアドレス振分/MACアドレス振分/固定なので組合せによっては
        片寄る懸念は内在



Link Aggregation概略

✵まとめ

- STPとの戦いは、使う限りまだまだ続く(Bug潰しの王道?)
- 時間と共にVLAN数の限界が見える前の対処が必須
  - ✓網構成の対処
  - ✓管理面の対処
- Multicastで、早くLayer2とLayer3の協業の実現、、、、(キラーアプリも)
- Link Aggregationに関連して汎用性のあるLLDPの実装