ネットワーク設計構築 A to Z [II]

～OSPFを簡単に使う最適WAN設計～

2006年12月5日

株式会社インターネットイニシアティブ

山口 二郎 (jiro-y@iij.ad.jp)

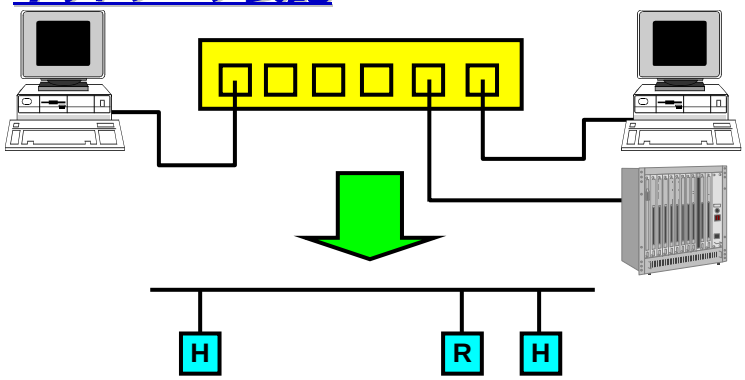
目的

- ダイナミックルーティングが必要な理由
- ダイナミックルーティングの種類と特徴
- 冗長化ネットワークを構築するには
- 広域Ethernetを利用したWANを構築するには
- インターネットVPNを利用したWANを構築するには
- ダークファイバを利用したWANを構築するには
- ポリシーごとにWANを使い分けるには
- フローティングスタティックを利用したバックアップを実現するには
- OSPFをエリア分けしなければならないとき

発表内容

- スタティックルーティングとダイナミックルーティングの違い
- ダイナミックルーティングの動作原理
- ダイナミックルーティングを用いたバックアップ、バランシング
- 広域Ethernetを利用したWAN構築
- インターネットVPNを利用したWAN構築
- ダークファイバを利用したWAN構築
- ポリシーごとのWANの使い分け
- フローティングスタティックを利用したバックアップ
- OSPFエリア構築

ネットワーク表記



- ハブ、スイッチなどは1本の線またはSWで表わします。
- ホストはH、A、B、C、D等で、ルータはR等で表記します
- レイヤ3スイッチなどは説明中ではルータと区別していません

経路制御解説

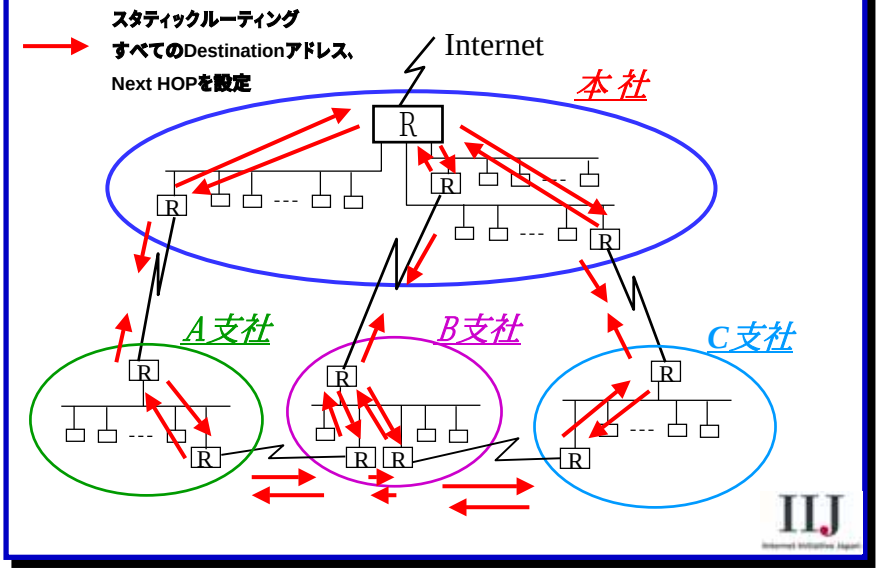
ここではダイナミックルーティングの原理について解説します

- 静的経路制御（スタティック）、動的経路制御（ダイナミック）の特徴
- ダイナミックルーティングの動作原理
- ダイナミックルーティングの種類、特徴
- RIP解説
- VLSM
- OSPF解説
- トラブルシューティング

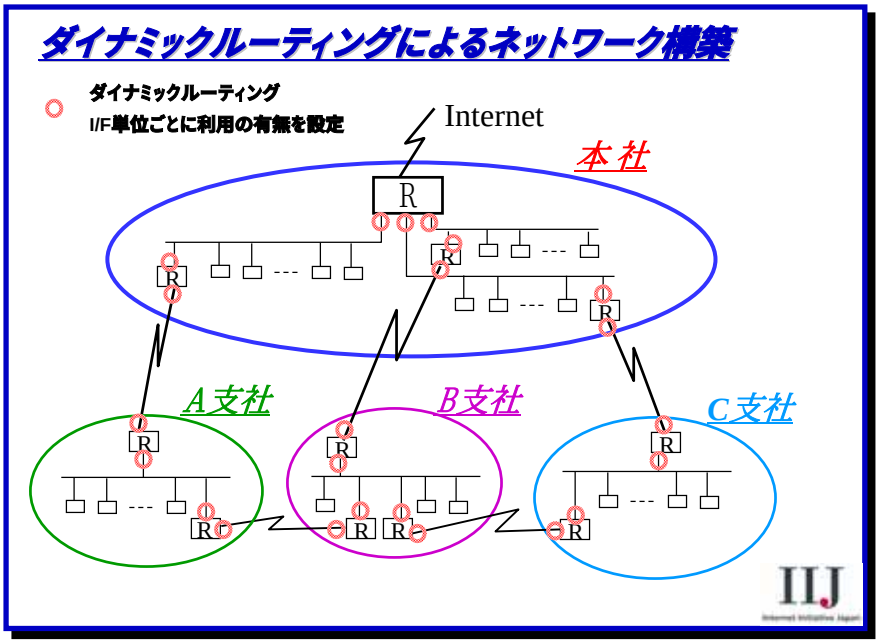
静的な経路制御と動的な経路制御

- 静的（スタティック）な経路制御の特徴
  - 手作業により固定的に経路を設定する
  - 安定している
  - トラフィックや伝送障害の影響を受けない
  - ルーティングプロトコルのためのトラフィックが発生しない
- 動的（ダイナミック）な経路制御の特徴
  - 自動的に経路を設定する
  - ネットワークの変化に対応できる
  - 自動的に最適経路を選択できる
  - 自動的にバックアップ経路を選択できる

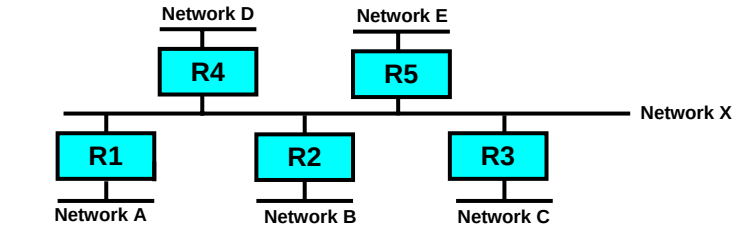
スタティックルーティングによるネットワーク構築



ダイナミックルーティングによるネットワーク構築



スタティックルーティングの設定



R1

Destination	Next Hop
B	R2
C	R3
D	R4
E	R5

R2

Destination	Next Hop
A	R1
C	R3
D	R4
E	R5

R3

Destination	Next Hop
A	R1
B	R2
D	R4
E	R5

R4

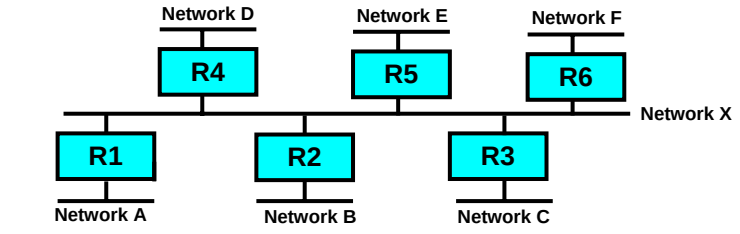
Destination	Next Hop
A	R1
B	R2
C	R3
E	R5

R5

Destination	Next Hop
A	R1
B	R2
C	R3
D	R4

- スタティックルーティングはそれぞれのルータに設定する

スタティックルーティングの追加



R1

Destination	Next Hop
B	R2
C	R3
D	R4
E	R5
F	R6

R2

Destination	Next Hop
A	R1
C	R3
D	R4
E	R5
F	R6

R3

Destination	Next Hop
A	R1
B	R2
D	R4
E	R5
F	R6

R4

Destination	Next Hop
A	R1
B	R2
C	R3
E	R5
F	R6

R5

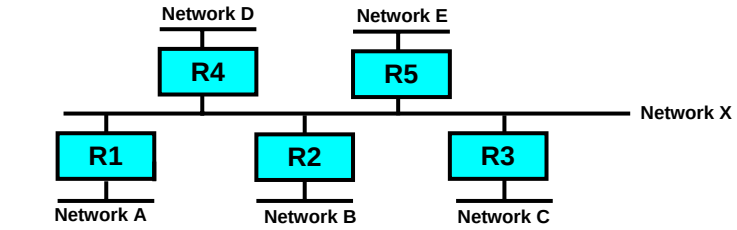
Destination	Next Hop
A	R1
B	R2
C	R3
D	R4
F	R6

R6

Destination	Next Hop
A	R1
B	R2
C	R3
D	R4
E	R5

- ネットワークが追加されると全てのルータに設定を追加する必要がある

ダイナミックルーティングの設定



R1

Protocol	Net
OSPF	X
OSPF(p)	A

R2

Protocol	Net
OSPF	X
OSPF(p)	B

R3

Protocol	Net
OSPF	X
OSPF(p)	C

R4

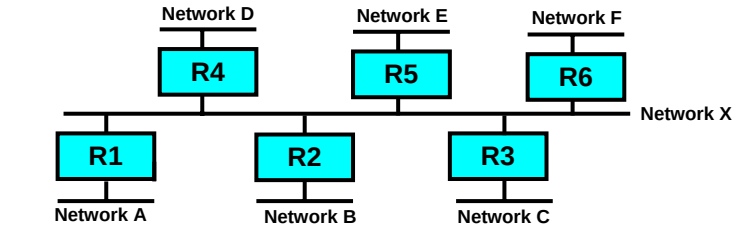
Protocol	Net
OSPF	X
OSPF(p)	D

R5

Protocol	Net
OSPF	X
OSPF(p)	E

- ダイナミックルーティングの設定は利用するプロトコルとネットワークを指定する

ダイナミックルーティングの追加



R1

Protocol	Net
OSPF	X
OSPF(p)	A

R2

Protocol	Net
OSPF	X
OSPF(p)	B

R3

Protocol	Net
OSPF	X
OSPF(p)	C

R4

Protocol	Net
OSPF	X
OSPF(p)	D

R5

Protocol	Net
OSPF	X
OSPF(p)	E

R6

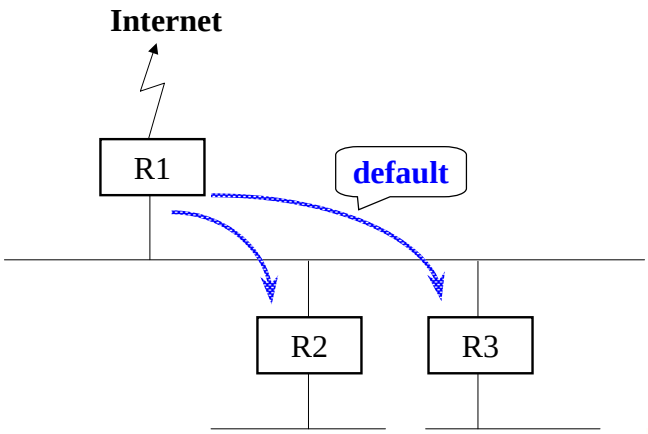
Protocol	Net
OSPF	X
OSPF(p)	F

- ネットワークが追加された場合には追加されたネットワークが接続されているルータのみに設定すればよい

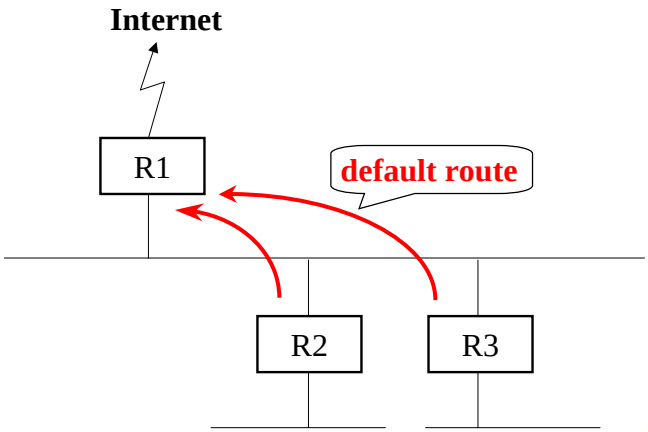
ルーティング設定まとめ

- スタティックルーティングの場合はバックボーンに新しいルータ、ネットワークが接続されると同じバックボーンを利用しているルータ全てに設定を行う必要がある
- ダイナミックルーティングを導入すると新規導入するルータにのみ設定を入れればよい
- ダイナミックルーティングを利用すると自動的にバックアップできる
- 中規模、大規模のネットワークにはダイナミックルーティングを導入したほうが良い

ダイナミックルーティング：経路情報の伝播



ダイナミックルーティング：伝播後の経路情報



ダイナミックルーティングプロトコルの種類

- RIP
  - RFC1058
- RIP 2
  - RFC2453
- OSPF
  - RFC2328
- BGP4
  - RFC1771

RIP

- Routing Information Protocol version 1
- RFC1058
- アドレスのみの伝播
  - VLSM使用不可
- ベクトル距離経路制御
- Broadcastのみ
- UNIXに標準添付されている (routed)

RIP2

- Routing Information Protocol version 2
- RFC2453
- netmaskを伝播できる
  - VLSM使用可能
- ベクトル距離経路制御
- RIPと互換性があり、併用も可能
- Multicastを利用可能
  - ホストの軽減を図る
- 最近では対応したroutedがある

OSPF -1

- Open shortest path first
- RFC2328
- Protocol 89
  - TCP (protocol 6)でもUDP(protocol 17)でもない
- netmaskを伝播できる
  - VLSM利用可能

OSPF -2

- Multicast(224.0.0.5/224.0.0.6)を利用する
- Load-balancingを行う
- UNIX標準で添付されていない
  - gated等をインストールする必要がある

BGP4 -1

- Border Gateway Protocol version 4
- RFC1771
- TCP 179
- EGPとしてのEBGPとIGPとしてのIBGPがある
- AS pathの長さにより経路を選択する

BGP4 -2

- 複数の経路が存在する場合は最適経路のみ伝播する
- Load-balancingは行わない
- Updateプロトコルである
- Aggregateできる。Classless Inter-Domain Routing(CIDR)対応

※ＢＧＰはここでは扱いません

ダイナミックルーティングの解説

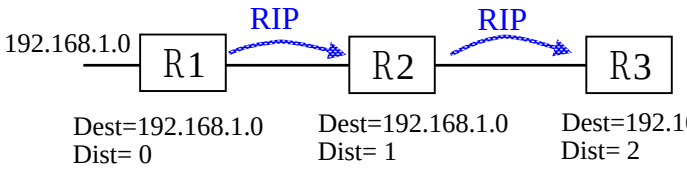
- RIPを理解する
  - RIPを理解すれば、OSPF、BGP4を概念的に理解することは容易
- 現場ではいまだにRIPが使用される場合がある
  - OSPFを利用できないルータが存在するため
  - Defaultだけを流すのでRIPで十分
- OSPF解説
  - RIPの知識をベースに解説します

RIPの動作原理 -1

ベクトル距離経路制御
(vector-distance/Bellman-Ford)

vector=destination(ネットワーク)
distance=HOP count(通過したルータの数)

ルータを通る度にdistanceが1追加される



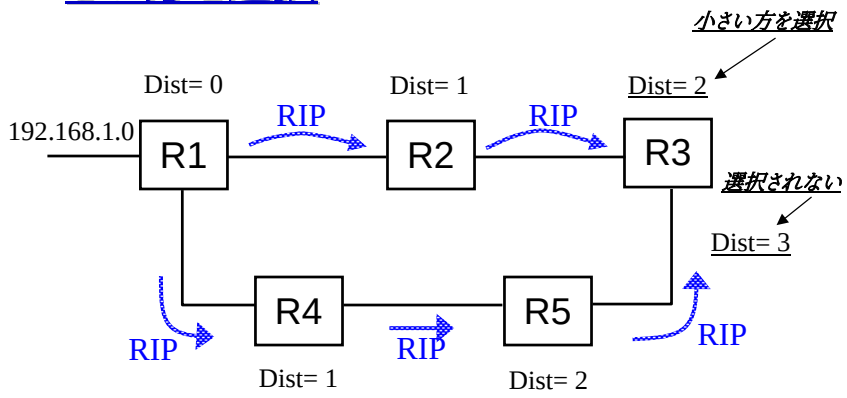
Dest=192.168.1.0
Dist= 0

Dest=192.168.1.0
Dist= 1

Dest=192.168.1.0
Dist= 2

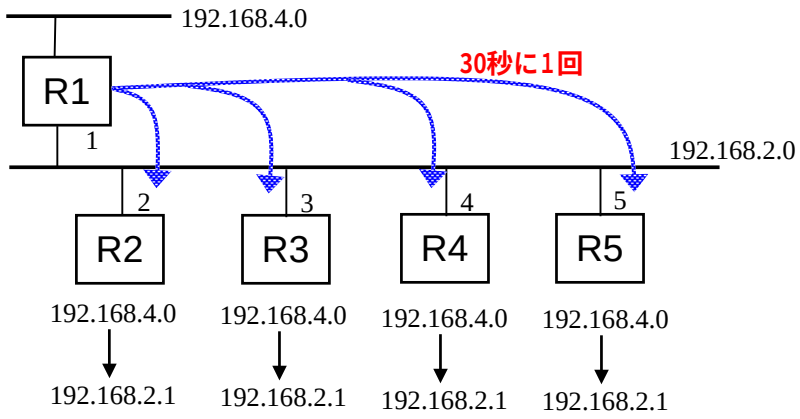
Dest=Destination
Dist= Distance

同じdestinationの場合はdistanceが小さい方を選択

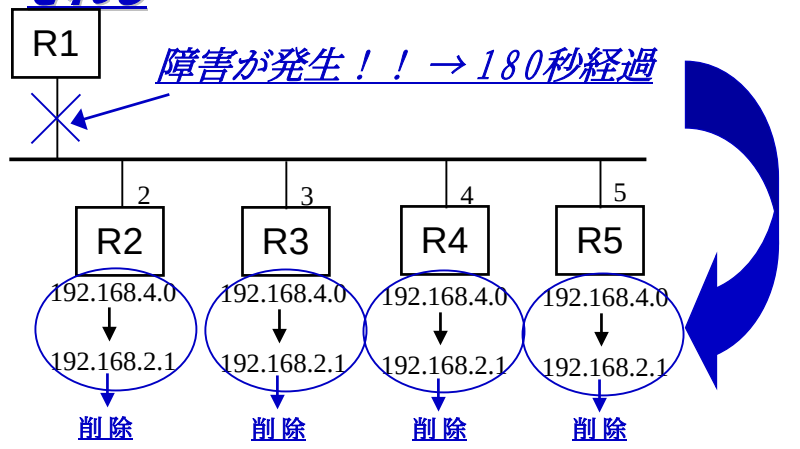


同じDestination同じDistanceの場合は
最初に到着した経路を選択

30秒ごとにbroadcastされる



3分間経路が到着しないと経路は削除される

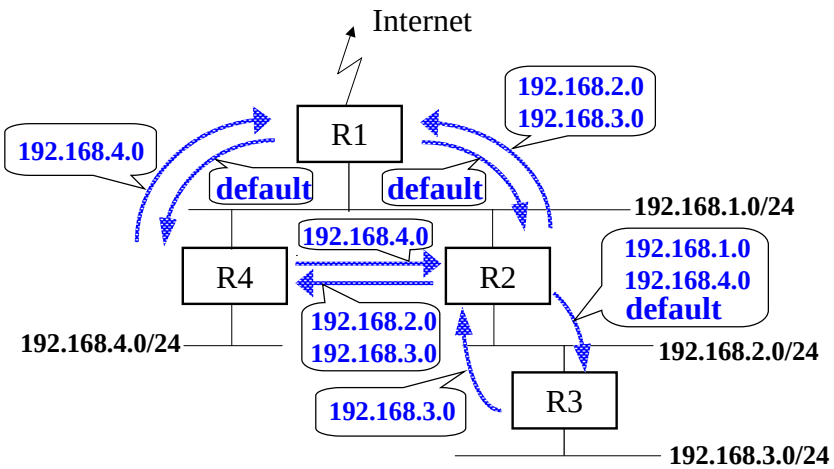


RIPで得られた経路情報は180秒

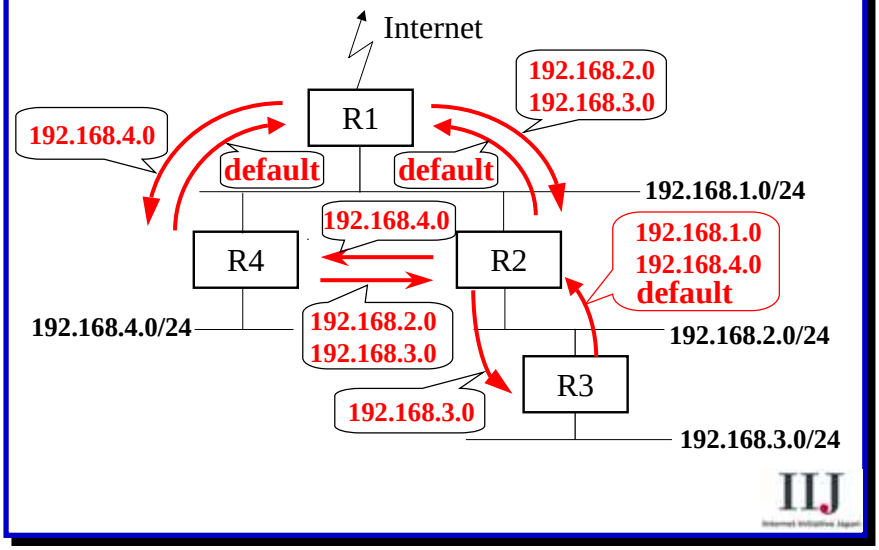
RIPの動作原理-2

- ネットワーク障害時には3分間で経路が切り替わる。 複数ルータがある場合には3分×ルータ数
- RIPはネットマスクを伝播しない
- クラスフルなマスクと見なされる
  - 利用可能な例
    - 192.168.1.0/24
    - 172.16.0.0/16
    - 10.0.0.0/8

RIP伝播



RIP伝播後の経路情報



RIPの動作原理-3

- 利用不可能な例
  - 192.168.1.0/26
  - 172.16.0.0/24
- 0.0.0.0というアドレスはdefaultとして機能する

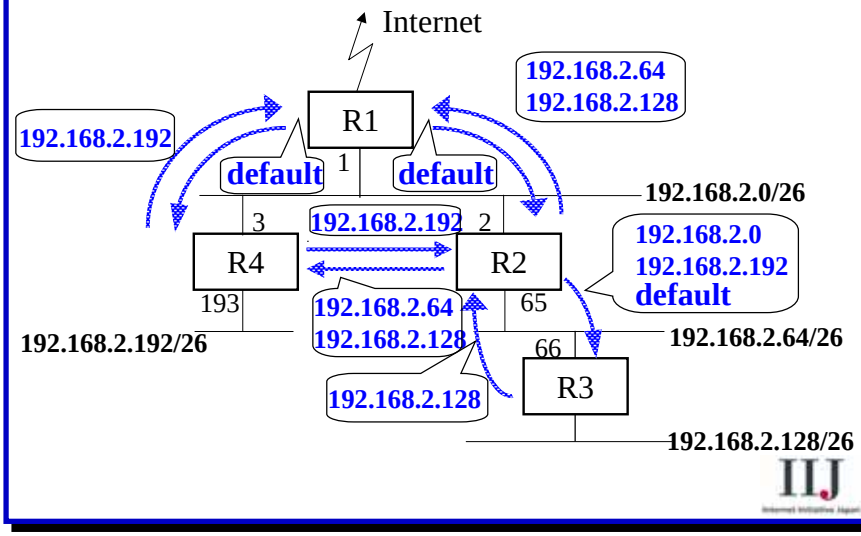
RIPのまとめ-1

- ベクトル距離経路制御(vector-distance/bellman-ford)
  - Vector=destination(ネットワーク)
  - Distance=hop count(通過したルータの数)
- ルータを通る度にdistanceが1追加される
- 同じdestinationの場合はdistanceが小さい方を選択
- 同じdestination同じdistanceの場合は最初に到着した経路を選択

RIPのまとめ-2

- 30秒ごとにbroadcastする
- 3分間経路が到着しないと経路は削除される
- ネットワーク障害時には3分間で経路が切り替わる。
  - 複数ルータがある場合には3分×ルータ数

Subnetmaskありのネットワーク構成



RIPでSubnetmaskを利用する場合-1

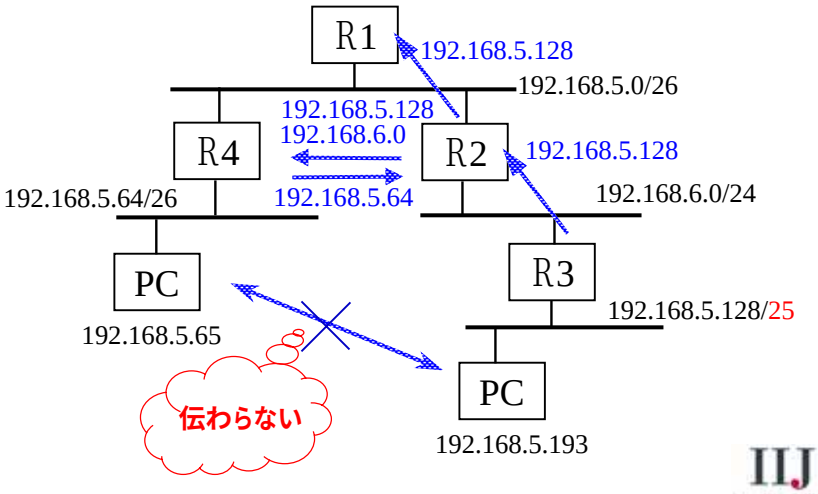
- インターフェースに設定されているnetmaskを適用
- 192.168.2.1/26 ルータのアドレス、マスクの場合

RIPで得られたdestination	ルーティングテーブル
192.168.2.64	192.168.2.64/26
192.168.2.65	192.168.2.65/32
192.168.2.128	192.168.2.128/26
192.168.2.192	192.168.2.192/26
192.168.3.0	192.168.3.0/24
192.168.3.64	192.168.3.64/32

RIPでSubnetmaskを利用する場合-2

- インターフェースに設定されているnetmaskが適用できない場合、RIPでは経路制御できない

VLSMありのネットワーク構成



VLSM(Variable Length Subnet Mask)

- ネットワーク例
  - 192.168.5.0/26
  - 192.168.5.64/26
  - 192.168.5.128/25
- 192.168.5.1が192.168.5.128を受け取った場合
  - 192.168.5.128/26と誤認する
  - 192.168.5.192～192.168.5.255がルーティングされない
- RIPだけではVLSMに対応できない
  - VLSM対応には RIP2、OSPFを利用

ルータでのRIP制御

- 聞く　広告
  - ○　○　RIPのみで運用可能
  - ×　○　defaultのみ広告を行うなどで利用
  - ○　×　defaultを告知しない場合に利用

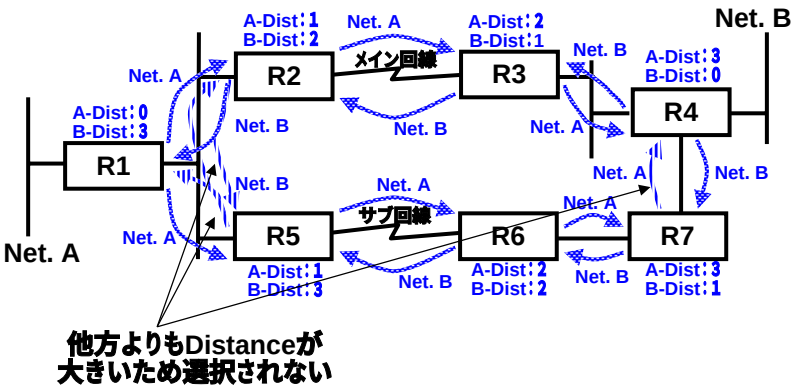
トラブルシューティング- RIPが伝播しない-1

- 同じbroadcastアドレスを利用していない
  - Broadcastアドレスが異なっている場合
  - 192.168.1.0/24を利用の場合
    - 192.168.1.255　network+all-1
    - 192.168.1.0　network+all-0
    - 255.255.255.255　all-1
    - 0.0.0.0　all-0
- 古いルータやワークステーション等はall-0,all-1固定の場合がある

トラブルシューティング- RIPが伝播しない-2

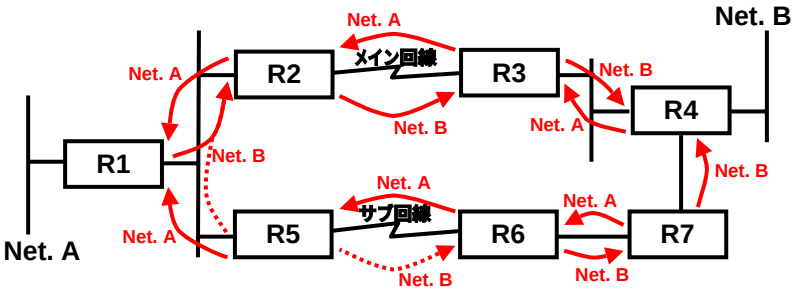
- Broadcastアドレスがfilterされている
  - 255.255.255.255,0.0.0.0などがインターフェースのoutputでfilterされていないか？
- プロトコル、ポートがfilterされている
  - UDP 520がfilterされていないか？
- Unnumberedのi/fでbroadcastを伝播できない
  - unicastで広告するように設定する
  - unicastで広告して良いのか？

RIPを用いたバックアップ-経路の伝播（定常時）



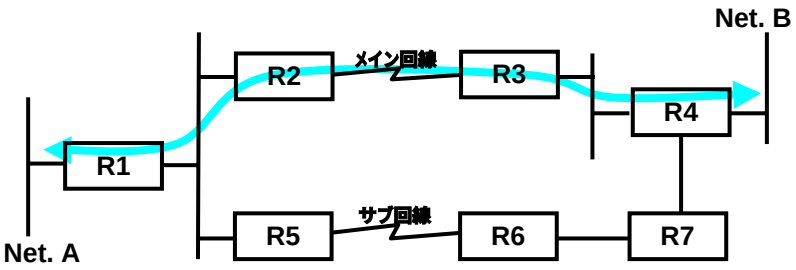
- RIPを利用し、主にバックアップを目的とした構成
- 通常時はメイン回線のみを利用する

RIPを用いたバックアップ -ルーティングテーブル（定常時）



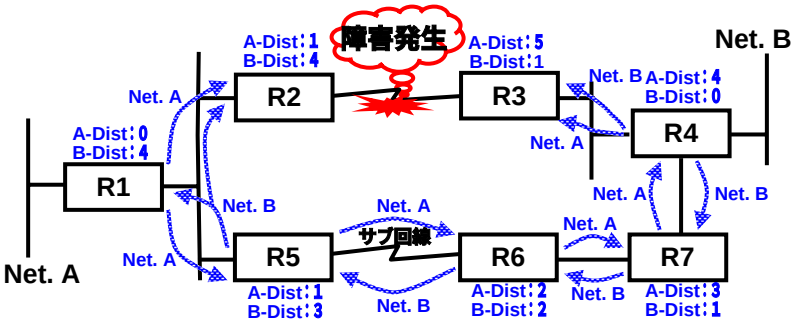
- RIPの経路情報が伝播することにより、各ルータに経路情報が設定される
- Distanceの違いから、メイン回線側の経路が選択される

RIPを用いたバックアップ-トラフィックの流れ（定常時）



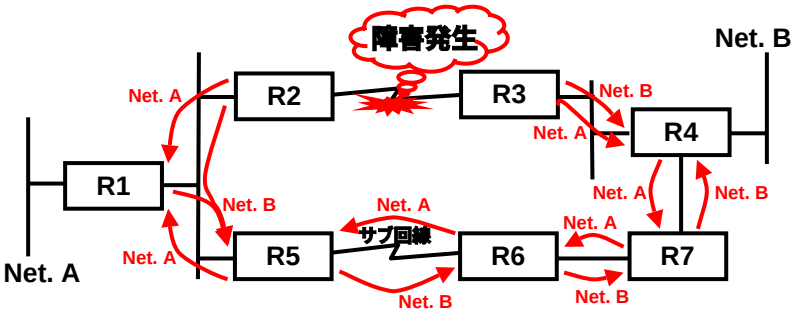
- 通常時はメイン回線のみが利用される

RIPを用いたバックアップ-経路の伝播（障害時）



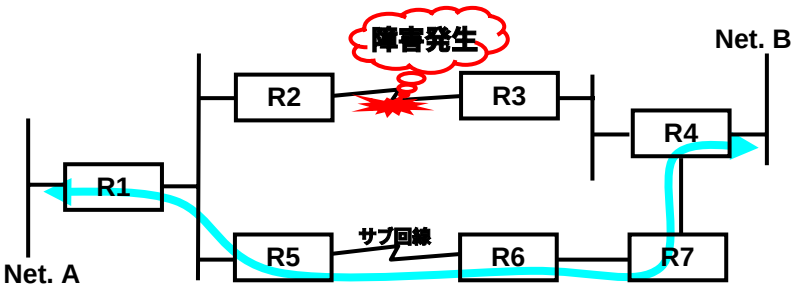
- メイン回線に障害が発生したため、経路情報の伝播が変化する

RIPを用いたバックアップ-ルーティングテーブル(障害時)



- 経路情報の伝播が変化するため、各ルータに設定されている経路情報が変更される

RIPを用いたバックアップ-トラフィックの流れ(障害時)



- メイン回線に障害が発生しているため、トラフィックの流れも変化する
- サブ回線を利用して、通信のバックアップを行う

OSPF解説－1

- 解説方針
  - ここではOSPFを知らない方のために一般的な利用法について解説します。
  - わかりやすさを重視して説明するため、RFCで定義されている厳密なOSPFの定義とは異なる部分もありますが、ご了承願います。
  - 大規模ネットワークではBGPとの連携は欠かせませんが、ここでは説明しません。

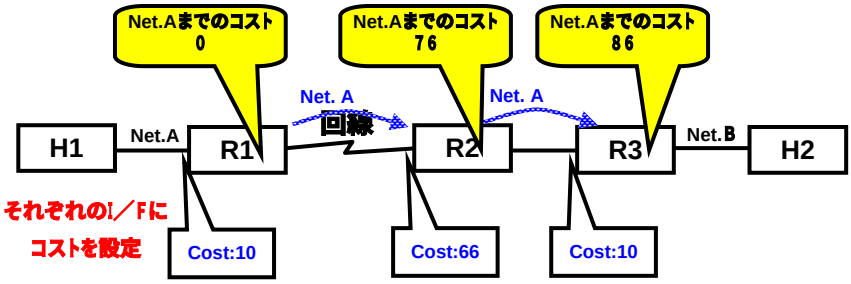
OSPF解説－2

- Link State型ルーティングプロトコル
  - ネットワークトポロジをLSA(Link State Advertisement)と呼ばれる形式でデータベース化し、最適な経路を選択する。
    - RIPやBGPと異なり、単純な経路交換を行なわないため、経路フィルタをかけることは難しい
  - トポロジに変更が合った場合にすぐ変更がかかる
  - ルータ故障検出も可能
    - HELLOパケットによりルータの故障を検出し、バックアップ経路を選択できる。
    - 切り替え時間がRIPよりずっと早い(数秒～1分程度)

OSPFコストとは

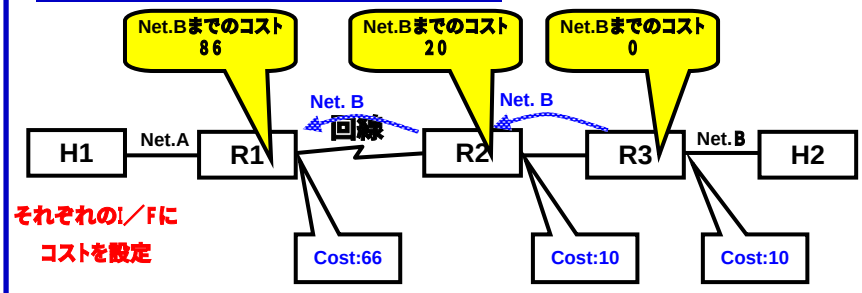
- OSPFではRIPでいうDistanceの変わりにコストを利用する
  - OSPFコストは0～65535の値を取る
  - インターフェース毎に自由にコストを設定することができる
  - コストは小さければ小さいほどネットワーク的に近距離に見せられる
  - ルータによっては回線速度に応じて自動的にコストを付与するものもあるが、ネットワークの高速化などに対応できなくなるだけでなく、運用が困難になるため、明示的に設定したほうが良い

簡単なOSPFコストの計算法－1



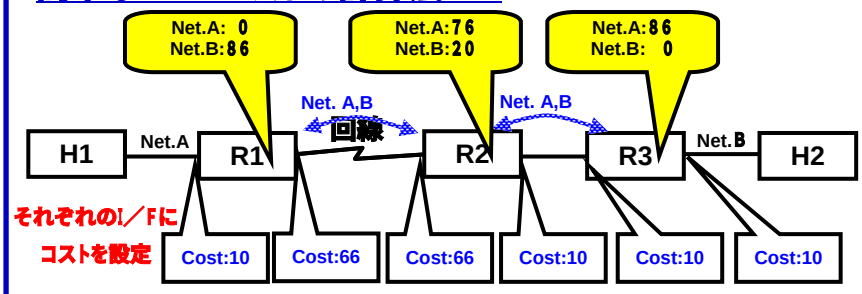
- R1から見たH1への経路
  - R1は直接Net.Aに接続されているため、同じNet.Aに接続されているH1はコスト0として見える
- R2から見たH1への経路
  - R2からは(R1のI/Fに設定されたNet.Aのコスト+R1と接続するI/Fに設定されたコスト)となる
- R3から見たH1への経路
  - R3からは(R2から見たNet.Aのコスト+R2と接続するI/Fに設定されたコスト)となる

簡単なOSPFコストの計算法－2



- R3から見たH2への経路
  - R3は直接Net.Bに接続されているため、同じNet.Bに接続されているH2はコスト0として見える
- R2から見たH2への経路
  - R2からは(R3のI/Fに設定されたNet.Bのコスト+R3と接続するI/Fに設定されたコスト)となる
- R1から見たH2への経路
  - R1からは(R2から見たNet.Bのコスト+R2と接続するI/Fに設定されたコスト)となる

簡単なOSPFコストの計算法－3

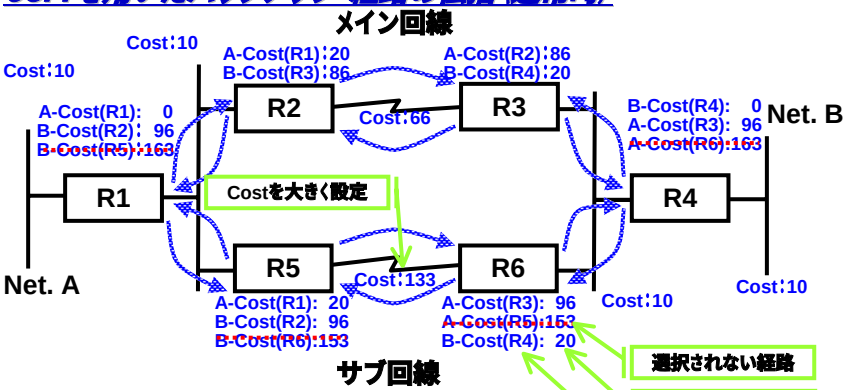


- 同じI/Fに同じコストを付けることにより、行きと帰りのコストを一致させることができる
- 行きと帰りで異なるコストを付与することもできるが、管理が煩雑になるため、理由なく行なうべきではない
- ここで示した図は経路を交換しているように書かれているが、実際はトポロジデータベースの交換により経路を確定している

バックアップ、バランシングを行なうには

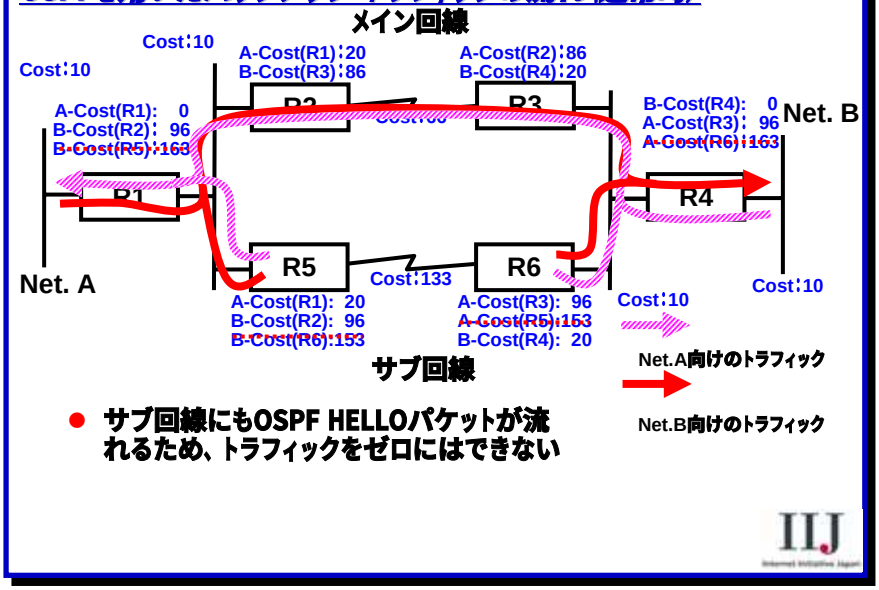
- OSPFでは複数の経路を持った場合にバックアップやバランシングを行なうことができる
- 異なるコストの経路がある場合
  - コストが小さい経路をメインとして利用しコストが大きい経路をバックアップとして利用できる
- 同じコストの経路がある場合
  - バランシングを行ない、トラフィック分散することが可能
  - バランシングを行なっている経路の1つが切断されても残った経路でバックアップすることも可能

OSPFを用いたバックアップ-経路の伝播 (通常時)



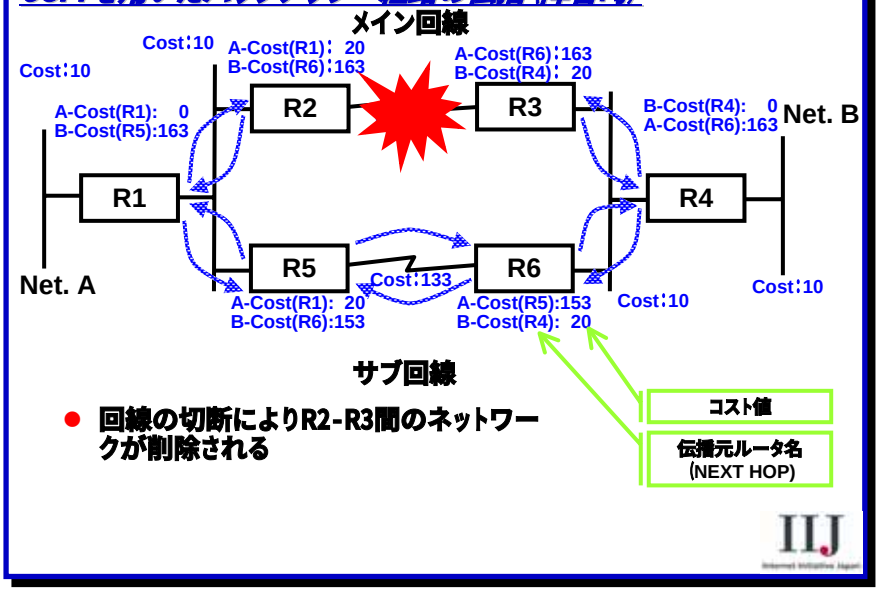
- OSPFを利用して、通常時はメイン回線のみを利用する
- 障害時にはサブ回線を利用してバックアップを行う

OSPFを用いたバックアップ-トラフィックの流れ(通常時)



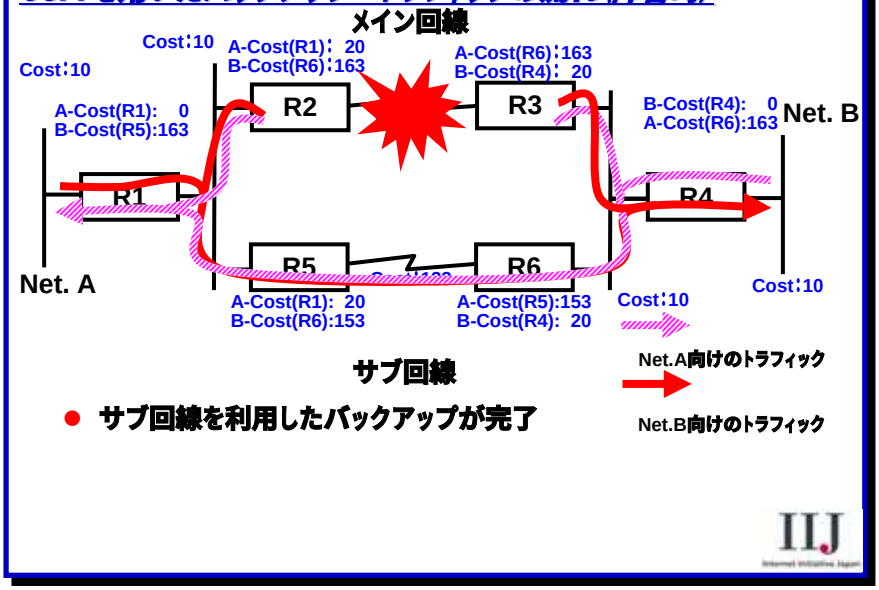
- サブ回線にもOSPF HELLOパケットが流れるため、トラフィックをゼロにはできない

OSPFを用いたバックアップ-経路の伝播(障害時)



- 回線の切断によりR2-R3間のネットワークが削除される

OSPFを用いたバックアップトラフィックの流れ(障害時)

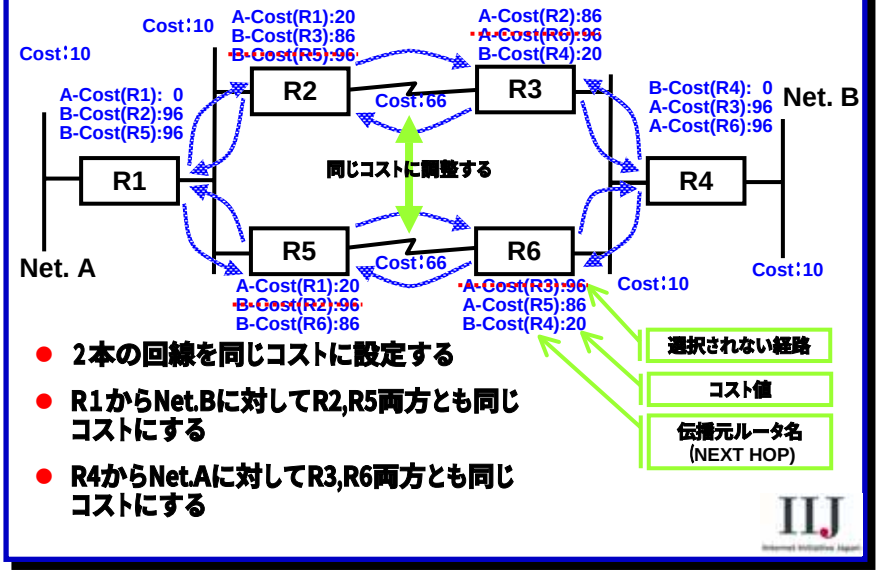


- サブ回線を利用したバックアップが完了

OSPFバックアップルーティングの特徴

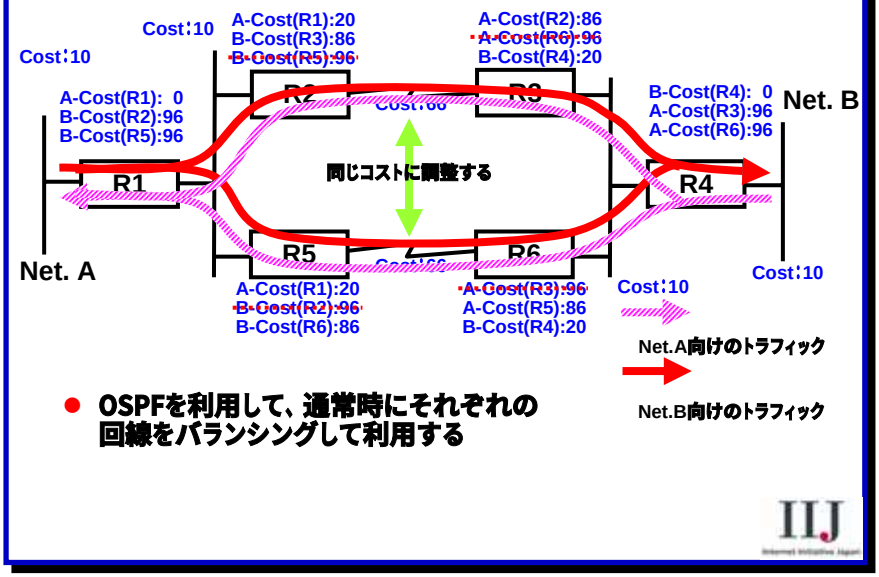
- RIPとは異なり、すばやいバックアップが可能
- バックアップ用の回線上もOSPF HELLOが流れるため、サブ回線を切断することはできない
  - ISDNなどでバックアップさせるにはOSPFだけのチューニングでは難しい
- 2本の回線を別々の用途に利用して障害時にそれぞれバックアップとして利用することが可能

OSPFを用いたバックアップ、バランシング-経路の伝播(通常時)



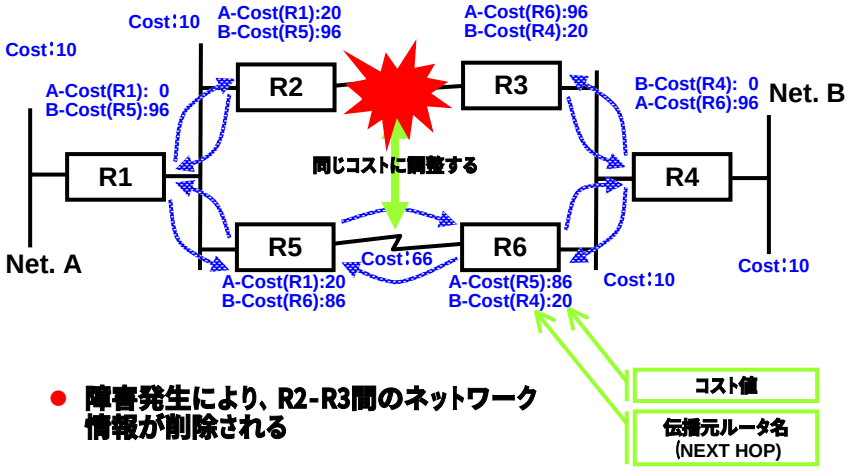
- 2本の回線を同じコストに設定する
- R1からNet.Bに対してR2,R5両方とも同じコストにする
- R4からNet.Aに対してR3,R6両方とも同じコストにする

OSPFを用いたバックアップ、バランシング-トラフィックの流れ(通常時)



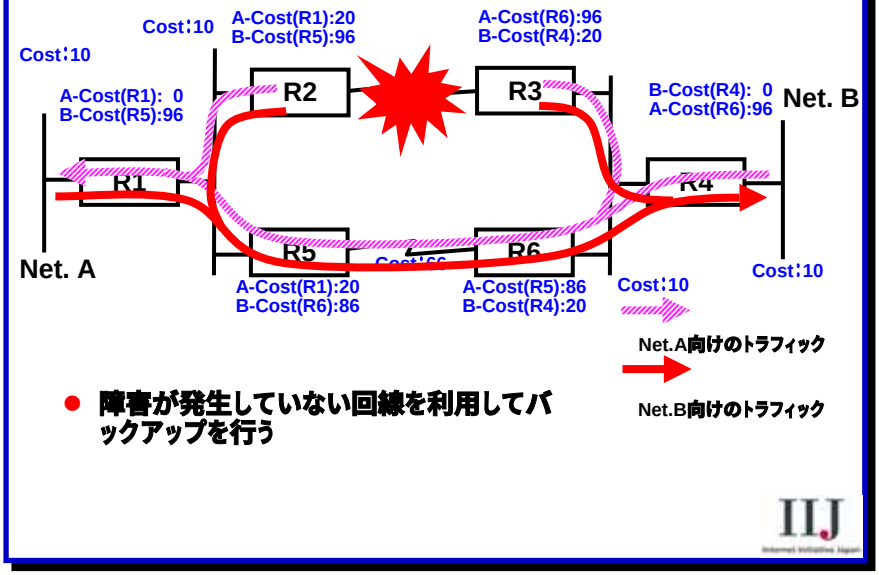
- OSPFを利用して、通常時にそれぞれの回線をバランシングして利用する

OSPFを用いたバックアップ、バランシング-経路の伝播(障害時)



- 障害発生により、R2-R3間のネットワーク情報が削除される

OSPFを用いたバックアップ、バランシング-トラフィックの流れ(障害時)



- 障害が発生していない回線を利用してバックアップを行う

バックアップ、バランシングの特徴

- 障害発生時には50%の帯域でバックアップ
- バランシング(ECMP:Equal Cost Multipath)は基本的に1:1でバランスするため、速度の異なる回線をバランスさせることは難しい
- 2本の回線を有効に利用し、回線コストを抑えることができる
- LAN等に利用すると100Mbpsメディアを200Mbpsメディアとして利用することもできる
- バランシングの精度はルータの機能に依存するため、1:1のバランシングとならない場合がある
- バランシング（ECMP）は一部のルータやL3SW等では機能しないため、注意が必要

初心者のためのOSPF設定－1

- エリア
  - 必ず0を設定する
    - OSPFでは経路の集約のためにエリアという概念があるが、小規模なネットワークではバックボーンエリア=エリア0だけで構築すればよく、エリアを分けて構築する必要はない
    - エリア0以外のエリアは必ずエリア0と接している必要があるため、むやみにエリア分けをするとバックボーンの拡張が難しくなる
    - ISPなど大規模ネットワークとなるとBGP+OSPFが主流であり、経路の集約という観点ではBGPのほうが優れているため、バックボーン以外のエリアを積極的に使っていくことはあまりない
    - 使用機器などの制限によりBGPが利用できず、OSPFで多くの経路を扱う場合にはエリアを利用して経路集約を図る必要がある
- デフォルトルート
  - 必ずstaticなどでデフォルトルートを確保してからOSPFでデフォルトルートを流す
    - 余力があればExternal Type 1で流す

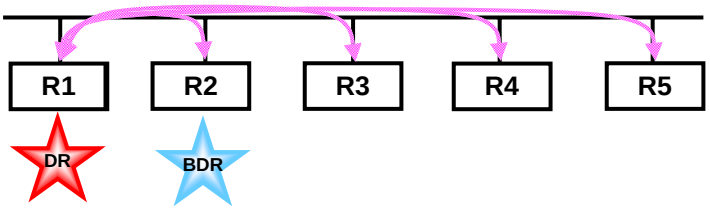
初心者のためのOSPF設定－2

- Staticからの経路注入
  - デフォルトルートなどと同じくExternal Type 1で流す
    - OSPFではOSPF以外のstaticやRIPなどから経路を注入するときにExternal Type 1とExternal Type 2が選べるようになっている
    - External Type 1とは
      - 注入時に付与したコストに、注入された場所から実際にOSPFの経路を受けるルータまでのOSPFコストを加えて評価する。同じ経路が複数注入されたときに最も近い出口から出るように制御するために使われる。Staticは注入された箇所が最も近いと判断できるため、Type 1が向いている。
    - External Type 2とは
      - 注入時に付与したコストをそのまま維持する。同じ経路が複数注入されたときに注入の際に付けられた優先順位に基づいて評価される。これはBGPなど他のプロトコルの情報をOSPFで実現するために有効な手法だが、現状BGPをそのままOSPFには流せないため、あまり意味がない
      - CiscoのルータはデフォルトがExternal Type 2であるため、注意が必要
    - External Type 1とExternal Type 2を混ぜない
      - OSPFコストとは別にExternal Type 1 > External Type 2という優先順位があるため、障害の切り分けが難しくなる

初心者のためのOSPF設定－3

- ルータID
  - 小規模では特に気にしなくても良いが、loopbackインターフェースを設定したほうが良い。
    - OSPFではルーター間通信にルータID(ルータについているIPアドレス)を用いる。
    - 通常はloopbackインターフェースを設定するとそのアドレスが使われる
    - 同じアドレスを複数のルータのloopbackインターフェースに付けると誤動作するため、注意が必要
- ルータを立ちあげる順番
  - 能力が高く、負荷が低いルータを先に立ち上げたほうがよい。
    - OSPFではDR(Designated Router)「指定ルータ」、BDR(Backup DR)、DROTHERが立ち上がった順に決まり、Ethernetなどマルチアクセスメディアの通信はDRが情報を管理するため、処理能力の余裕があるルータに行なわせたほうが良い。
    - 小規模では意識しなくても問題が発生しないことがほとんど。

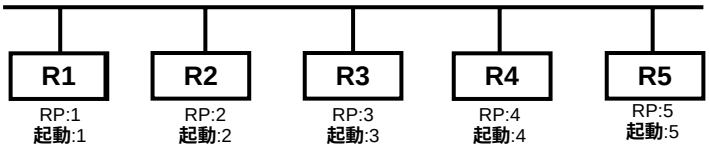
DRとBDRの役割



- DRの役割
  - DRはEthernetなどのマルチアクセスメディア利用時に、同じセグメントの代表して経路交換を行う
  - DRが存在することで、経路交換数は接続ルータ数に比例した量に抑えることができる
    - DRといった概念が無い場合には接続ルータ数の二乗に比例する
- BDRの役割
  - BDRはDRに障害が発生したときにすみやかにDRとなる役割を持つ

DR: Designated Router
BDR: Backup Designated Router

DRとBDRの選出-1

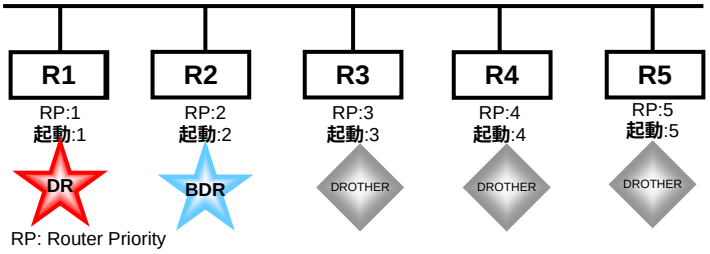


R1	R2	R3	R4	R5
RP:1 起動:1	RP:2 起動:2	RP:3 起動:3	RP:4 起動:4	RP:5 起動:5

RP: Router Priority

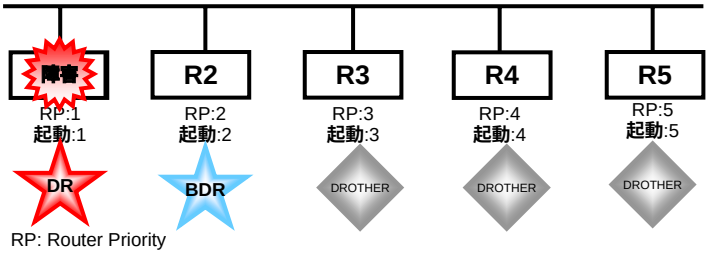
- DRとBDRの選択の方法
  - 最初に起動したルータはDRとなる
  - 2番目に起動したルータはBDRとなる
  - 3番目以降に起動したルータはDROTHERとなる
- 上記ネットワークにおけるDR、BDRはどれか？

DRとBDRの選出-2



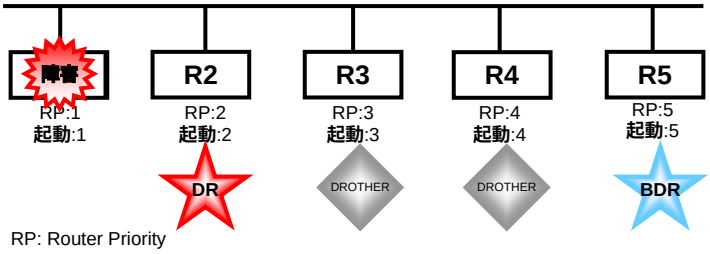
- DRとBDRの選択の結果
  - 最初に起動したR1はDRとなる
  - 2番目に起動したR2はBDRとなる
  - 3番目以降に起動したR3～R5はDROTHERとなる

DRとBDRの選出-3



- DRに障害が発生した場合
  - DRとBDRはどのように選択されるのか？

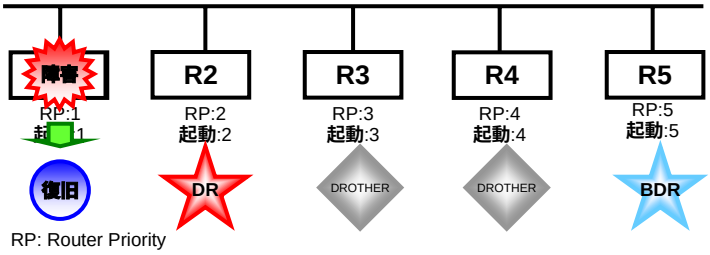
DRとBDRの選出-4



RP: Router Priority

- DRに障害が発生した場合
  - DRに障害が発生するとBDRがDRとなる
  - BDRがDRとなると、新しいBDRがRouter Priorityの大きなルータとなる

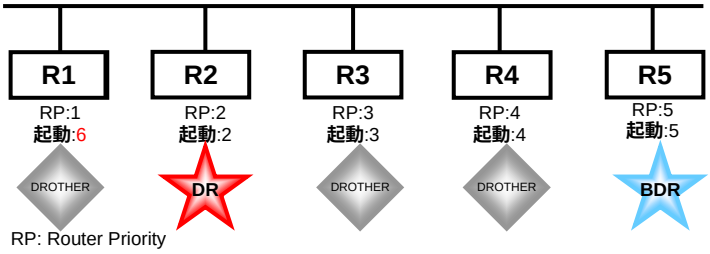
DRとBDRの選出-5



RP: Router Priority

- 障害が復旧した場合
  - 障害が復旧した場合DRは変化するか？

DRとBDRの選出-6



RP: Router Priority

- 障害が復旧した場合
  - DRに変動はない
  - R1は障害発生のため、起動順番が最後となる
    - 起動順番が3以降の場合は、起動順番がDR選択に影響を与えることはない

DRとBDRのまとめ

- DRとBDRの役割
  - DRはEthernetなどのマルチアクセスメディア利用時に、同じセグメントの代表して経路交換を行う
  - DRが存在することで、経路交換数は接続ルータ数に比例した量に抑えることができる
    - DRといった概念が無い場合には接続ルータ数の二乗に比例する
  - BDRはDRに障害が発生したときにすみやかにDRとなる役割を持つ
- DRとBDRの選出
  - DRとBDRは起動順に決定される
  - DR、BDRに障害が発生した場合にはRouter Priorityが高いルータが選出される
  - 常にDR、BDRを希望のルータにしておくことは困難
  - Router Priorityを0にすることで、DR、BDRにならないルータを作ることができる
    - 広域Ethernetの小規模拠点に有効
- 初心者のためのDR、BDR
  - LANでOSPFを利用している場合にはそれほど意識する必要は無い
  - 広域EthernetなどのWAN利用の際には小規模拠点のRouter Priorityを0としたほうが良い

OSPF利用上の注意点

- アドレスの重複には細心の注意を払う
  - loopbackアドレスはOSPF Router IDとして利用されるため、重複した場合にはOSPFデータベースが正常に維持できず、経路障害となる
  - LAN IPアドレスの重複が発生した場合にもデータベースが混乱し、該当ネットワークへの到達性が失われるだけで無く、多量のOSPF更新情報が流れ続けるなどの障害が発生する
  - shutdown状態のインターフェースであってもOSPFデータベースに登録されてしまう場合があるため、移行作業時などのIPアドレスの消しこみは速やかに実施した方がよい

トラブルシューティング-RIPv2とOSPFが伝播しない

- ルータのfilter等でmulticastアドレスや、protocol、portなどが制限されていないか注意する
  - RIP2
    - 224.0.0.9
    - UDP 520
  - OSPF
    - 224.0.0.5/224.0.0.6
    - Protocol 89
- Multicastをサポートしない場合
  - OSによってはmulticastを受けられない場合がある このときはbroadcastにて代用する

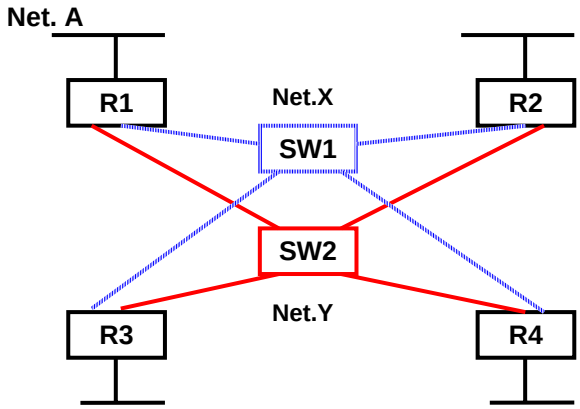
ダイナミックルーティングのまとめ

- VLSMを考慮するとRIP2,OSPFを利用すべき
- 単純なネットワーク構成はstaticを選択
- Defaultのみを利用する場合はRIPでも十分
- バランシングなどを行なう場合はOSPFを用いる

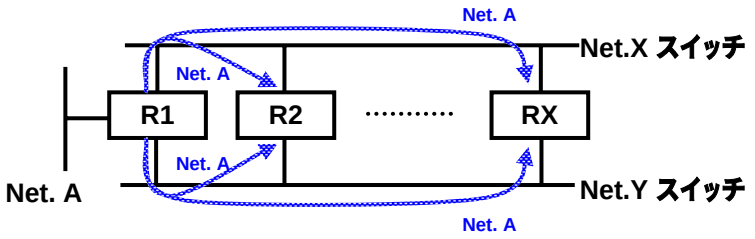
ダイナミックルーティングプロトコルを用いた障害に強いネットワーク構成

- デュアル構成＋OSPFによるバックアップ、バランシング
- リングトポロジによるバックアップ

デュアル構成＋OSPFを用いたバックアップ、バランシング接続図

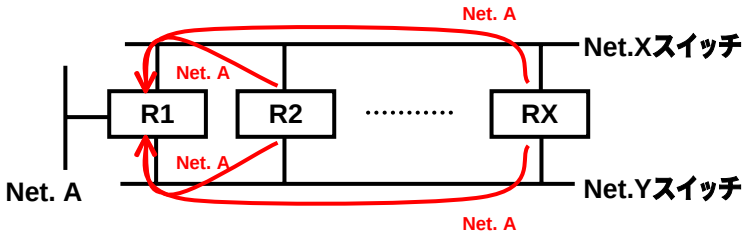


デュアル構成＋OSPFを用いたバックアップ、バランシング-経路の伝播（通常時）



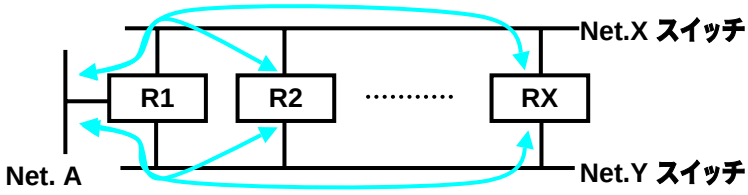
- OSPFで Net.Aの経路情報を広告する
- 経路情報は各ルータに対して、2つのスイッチから等価に伝播する

デュアル構成＋OSPFを用いたバックアップ、バランシング -ルーティングテーブル (通常時)



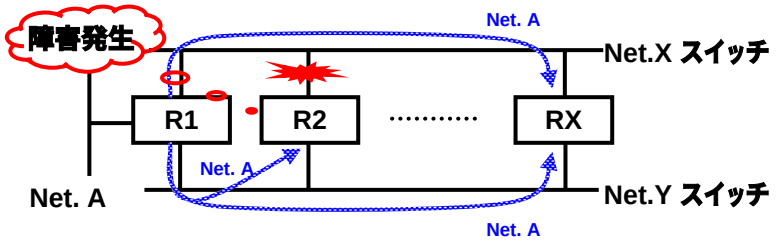
- 伝播した経路情報により、各ルータに経路情報が設定される。
- 2つのスイッチから等価な経路情報が伝播してきたため、2つの経路情報が設定される

デュアル構成＋OSPFを用いたバックアップ、バランシング -トラフィックの流れ (通常時)



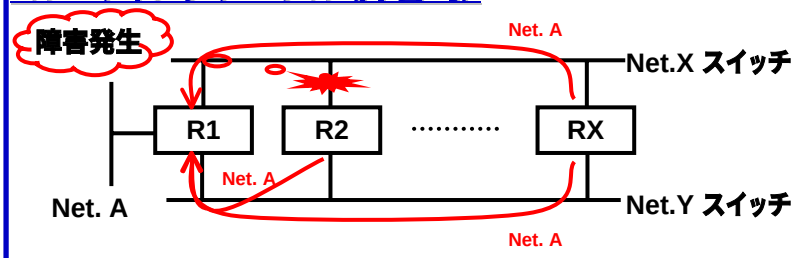
- 通常時には、2つのスイッチを経由するトラフィックがバランスする

デュアル構成＋OSPFを用いたバックアップ、バランシング -経路の伝播（障害時）



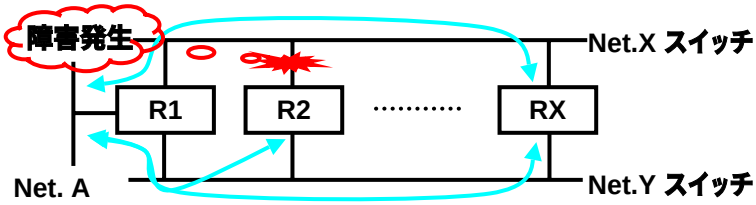
- 障害発生により、経路情報の伝播に一部に変化が生じる

デュアル構成＋OSPFを用いたバックアップ、バランシング -ルーティングテーブル（障害時）



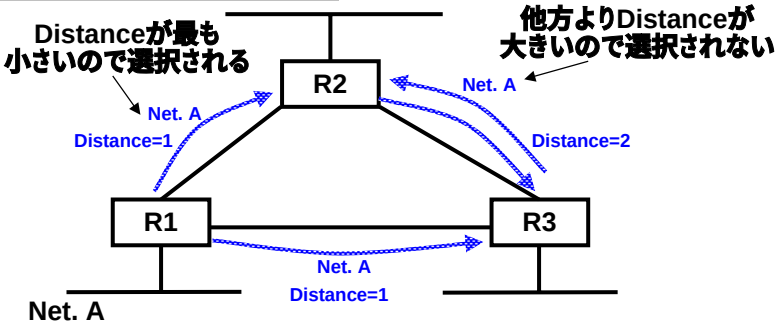
- 伝播する経路情報が変化するため、各ルータに設定されている経路情報も変化する
- 一方のスイッチからの経路が消えても、もう一方のスイッチからの経路でバックアップを行う

デュアル構成＋OSPFを用いたバックアップ、バランシング -トラフィックの流れ（障害時）



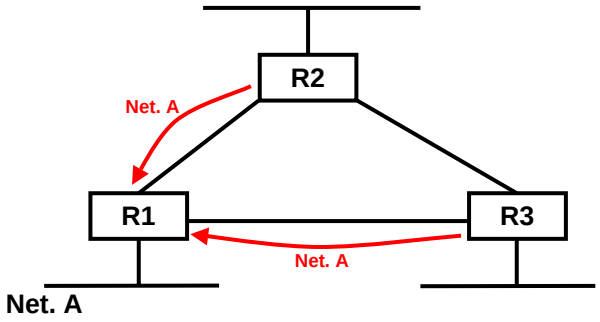
- 障害時には、2つのスイッチどちらかを利用して障害を迂回することができる

リングトポロジによるバックアップ -経路の伝播（通常時）



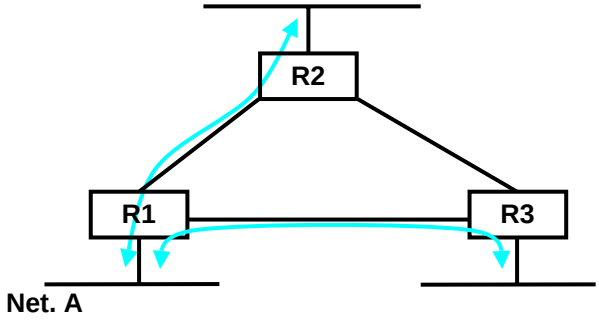
- RIPで Net.Aの経路情報を広告する
- 通常時は最短な経路が優先される

リングトポロジによるバックアップ -ルーティングテーブル(通常時)



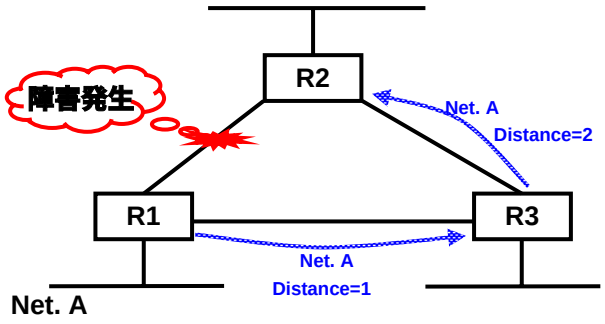
- 伝播した経路情報から、各ルータに経路情報が設定される

リングトポロジによるバックアップ - トラフィックの流れ(通常時)



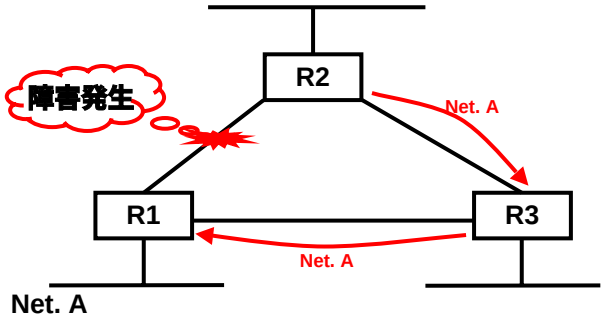
- 通常時は最短な経路が優先されて、通信が行われる

リングトポロジによるバックアップ -経路の伝播(障害時)



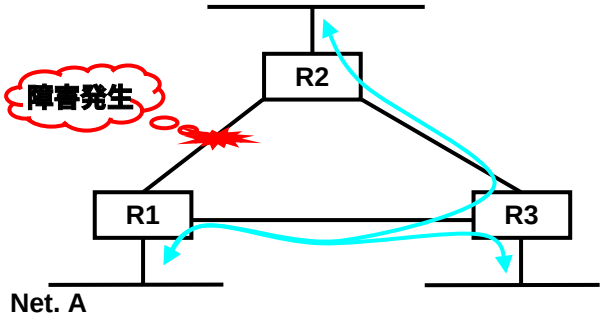
- 障害により、経路情報の伝播に変化が生じる

リングトポロジによるバックアップ -ルーティングテーブル(障害時)



- 伝播する経路情報の変化により、ルータに設定されている経路情報も変化する

リングトポロジによるバックアップ -トラフィックの流れ(障害時)



- 障害時には、遠回りな経路を利用して通信をバックアップする

WAN構築

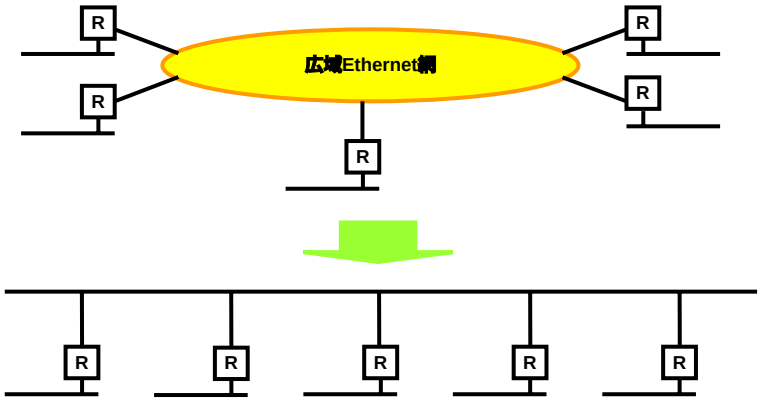
ここでは利用回線別のWAN構築の方法について解説します

- 広域Ethernetを利用したWAN
- インターネットVPNを利用したWAN
- ダークファイバを利用したWAN

広域Ethernetを利用したWAN

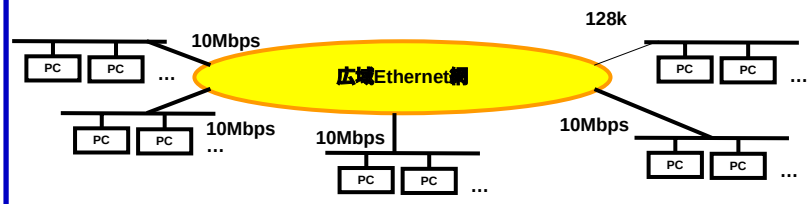
- 広域Ethernetを利用する理由
  - 安価
  - IP以外のパケットが通る(SNAなど)
  - ATMやPOSなどの高価なWAN I/Fが不要
  - ルータを利用せずにHUBだけでネットワークが構築できる
  - Tag VLANを利用して複数のVLANを複数拠点に容易に持っていくことができる
- 今回とりあげるポイント
  - IPのみを利用、ルータを利用、ダイナミックルーティング

ネットワーク構成



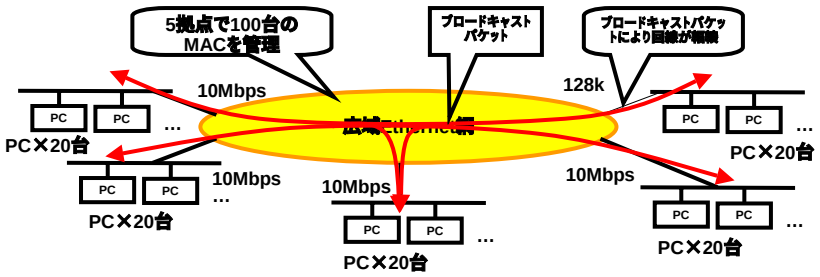
- 広域Ethernet網はLANのEthernetと同様に見える
- 基本的にはLANと同じ設計手法が使える

HUBのみで構成した場合



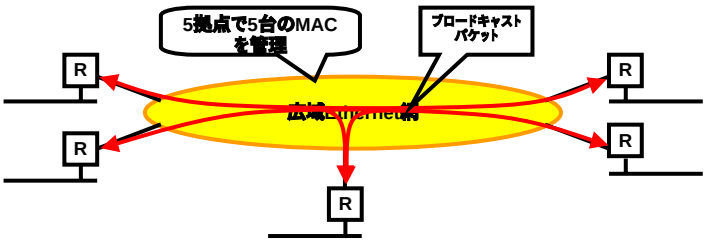
- 広域EthernetはLANと同様にHUBのみでもネットワークを構成することができる。

HUBのみで構成した場合の問題点



- HUBのみで構築し、PC端末を直接広域Ethernetに接続すると広域Ethernet内で管理すべきMACが増加する。
- これによりARPやWindows系のブロードキャストが増加し、回線の細い拠点が輻輳する。

ルータを設置した場合

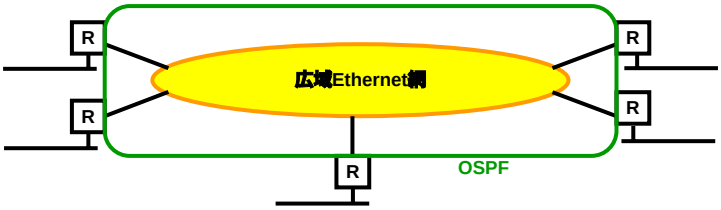


- ルータを設置した場合には広域Ethernet内でのMAC はルータの台数に限られるため、ブロードキャストの増加を防ぐことができる。

ルータを設置すべきか

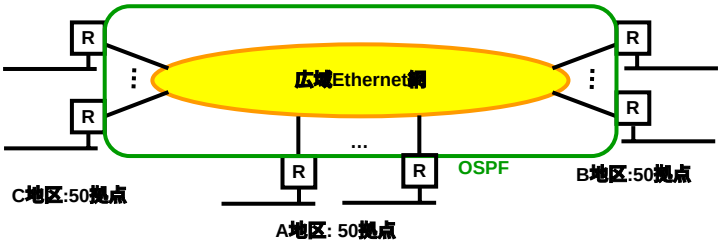
- ルータを設置すべきか、HUBのみで構築すべきか
  - 広域EthernetはHUBだけで容易にネットワークを構築できるが、スケールするネットワークとするためにはルータを設置すべきである。
  - 小規模拠点などHUBのみで構築が必要な場合にはルータ接続拠点とは異なるVLANで構築することが望ましい。

広域Ethernet でのOSPF の利用



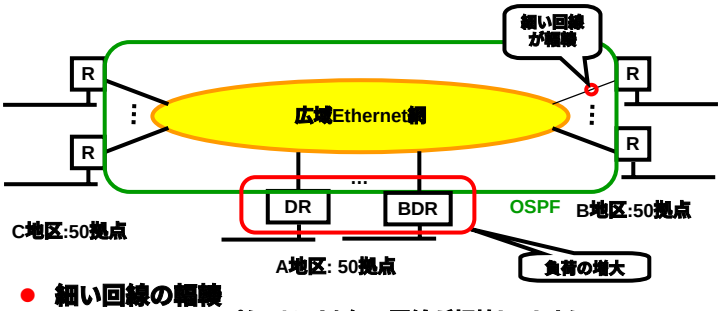
- 広域Ethernet網でのOSPFの利用
  - 広域EthernetではLANと同様にダイナミックルーティングを利用できるが、一般的にはOSPFを用いられることが多い。
  - 広域Ethernet網でのOSPF利用のポイントについて解説する

多拠点でのOSPF の利用



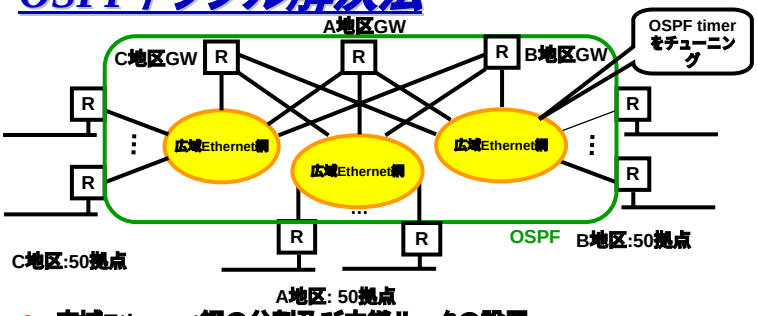
- 多くの拠点を1つの広域Ethernet網で結び、OSPFを動かす。
- 図ではA,B,C地区それぞれに50拠点接続しており、合計150台のルータが同一の広域Ethernet網を利用している。

OSPF利用時のトラブルシューティング



- 細い回線の輻輳
  - OSPFのHelloパケットにより細い回線が輻輳してしまう
- DR/BDRの負荷増大
  - DRおよびBDRに負荷が集中し、不安定となる
  - DR,BDRに高いスペックのルータが必要となる

OSPFトラブル解決法



- 広域Ethernet網の分割及び中継ルータの設置
  - 巨大すぎる広域Ethernet網を1セグメント50台程度として分割
  - それぞれの広域Ethernet網を中継するルータを設置
  - 広域Ethernetを分割することで、OSPFマルチキャストを減らすことができる
  - セグメントの分割により、DR、BDRを分散配置でき、負荷を下げることができる
- OSPF timerのチューニング
  - 細い回線を収容している広域Ethernet網のOSPFのHello間隔を伸ばし、輻輳を回避する

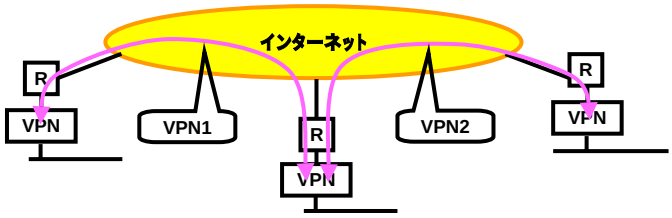
広域Ethernetを利用したWAN: まとめ

- HUBのみで構築すべきか、ルータを設置すべきか
  - ルータを設置したほうがスケールする
  - HUBのみの構成とルータ設置の構成が混在する場合にはTag VLANなどを利用して異なるネットワークに収容する
- ルータの設置台数が50台を超えるようであれば広域Ethernetを分割して、それぞれのネットワークを接続する中継ルータを用意する
- 細い回線を利用する場合にはOSPF timerをチューニングして輻輳しないようにする

インターネットVPNを利用したWAN

- インターネットのブロードバンド化と低価格化、VPN装置の低価格化と高性能化により、急激にインターネットVPNが普及してきている。
- ここではネットワークという切り口からWANとしてインターネットVPNを利用することを前提とする。
- VPNにはさまざまなプロトコル、暗号化技術、認証システムなどの要素があるが、プライベートネットワーク間で影響を受ける部分にのみ着目して解説する。暗号化されたパケットの状態など、プライベートネットワーク間では隠される要素に関してしはここではブラックボックスとして扱うものとする。

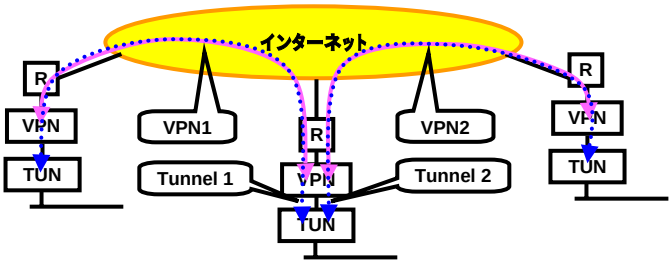
一般的なインターネットVPN構成



- 一般的なインターネットVPNの構成
  - インターネット接続ルータの下にVPN装置を設置し、VPN装置間でVPNを張る
  - VPNに流せるパケットはVPN依存
  - 暗号化パケットのスループットの高いVPN装置ではVPN間でOSPFなどのダイナミックルーティングが利用できないことが多い

スループットとダイナミックルーティングなどを両立するためには

ネットワーク化されたインターネットVPN構成

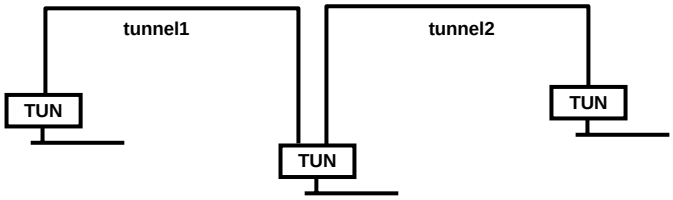


- ネットワーク化されたインターネットVPN構成
  - インターネット接続ルータの下にVPN装置を設置し、VPN装置間でVPNを張る(IPsecなど)
  - VPN装置の内側にtunnelルータを設置し、VPN上にtunnelを張る(GRE、ipipなど)
  - Tunnelは暗号化する必要はない
  - Tunnelは専用線と同等に見えるため、ダイナミックルーティングが利用できる
  - VPN装置はダイナミックルーティングを利用できなくてよい

VPNで利用されるプロトコル

- IPsec
  - IP Security Protocol
  - 認証、暗号化を行う
  - RFC2401〜2412,2451,2857,3526,3554,3566,3602
  - Protocol 50(ESP:encapsulating security payload)
  - Protocol 51(AH:authentication header)
- GRE
  - Generic Routing Encapsulation
  - レイヤ3 tunnelを行う
  - RFC1701,1702
  - Protocol 47(GRE)
- IPIP
  - IP Encapsulation within IP
  - レイヤ3 tunnelを行う
  - RFC2003
  - Protocol 4(IP-ENCAP:IP encapsulated in IP)
- GIF
  - Generic Tunnel Interface
  - IPIPをUNIXなどで扱うときに利用される
  - IPIP tunnelのことをGIF tunnelと呼ぶこともある

ネットワーク化されたインターネットVPN構成



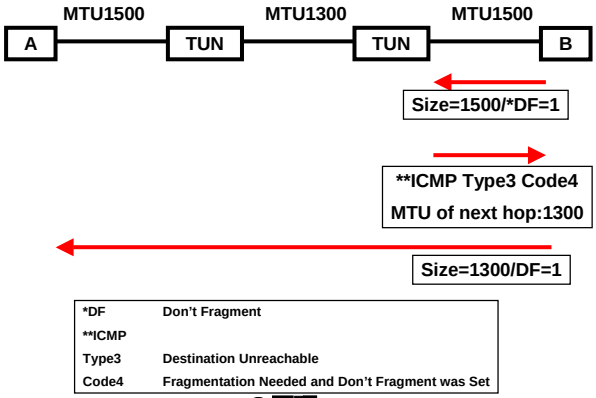
- プライベート側からみたネットワーク構成
  - Tunnelルータ間はtunnel1およびtunnel2の2つの専用線で接続されていることと同様に扱える
  - OSPFなどのダイナミックルーティングを容易に扱うことができる

インターネットVPNの解決すべき問題点

- MTUが1500より小さくなることによる問題点
  - インターネットVPNなど既存のネットワークの上にtunnelを張って利用する場合にはMTU(Maximum Transmission Unit)が1500より小さくなることによる問題が発生する
  - Path MTU Discovery Blackhole問題
    - RFC1191 Path MTU Discovery
    - RFC2923 TCP Problems with Path MTU Discovery

    →詳しくは後述
  - Path MTU Discovery Blackhole以外のMTU1500を必要とするアプリケーションの問題
    - DF=1で送信を行うLANアプリケーション
  - フラグメントによるパフォーマンスの低下
    - 小さいMTUの個所を通過する際にフラグメントが許可されていれば、フラグメントすることによりすべてのサイズのIPパケットを通過させることができる
    - ただし、フラグメントによりスループットが低下する恐れがある

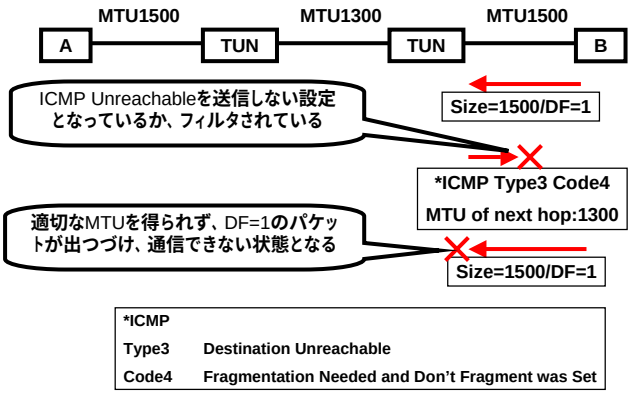
Path MTU Discovery の動作原理



*DF	Don't Fragment
**ICMP	
Type3	Destination Unreachable
Code4	Fragmentation Needed and Don't Fragment was Set

- Path MTU Discoveryの原理
  - DF=1としたIPパケットを送出し、Destination Unreachableが戻ってきたときに適切なIPパケットサイズに調整して送信することで、エンド-エンド間の最大で最適なMTUを利用する仕組み

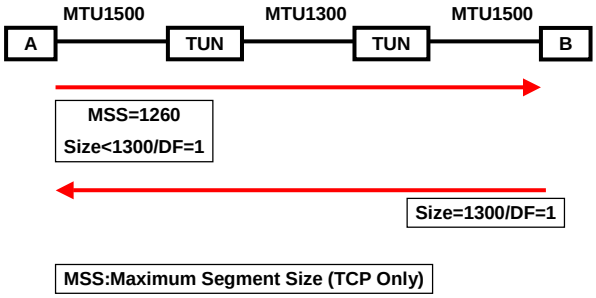
Path MTU Discovery Black hole問題



*ICMP	
Type3	Destination Unreachable
Code4	Fragmentation Needed and Don't Fragment was Set

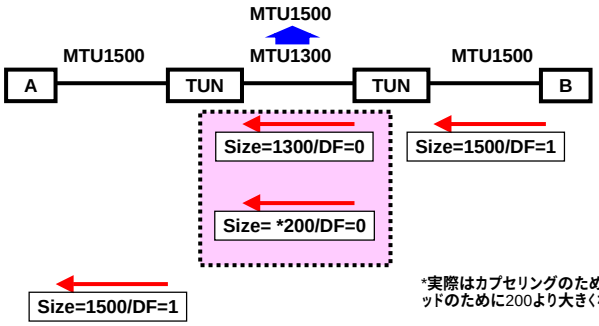
- Path MTU Discovery Black holeとは
  - Path MTU Discoveryの原理で重要な役割を持つICMP Unreachableがフィルタされることで、適切なIPパケットの送信が行えず、通信できなくなる状態のこと

Path MTU Discovery Black hole問題の解決法1



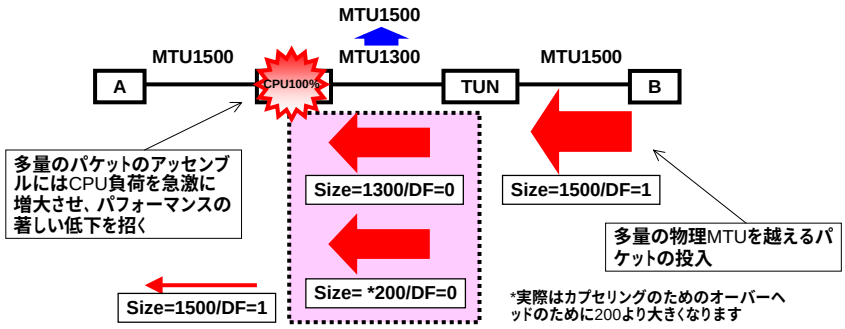
- TCPのMSSを利用する
  - TCPでは1度に送出する最大のセグメントサイズMSSを指定することができる。このパラメータをMTUが小さくなるポイントで書き換えることで、TCPに限ってPath MTU Discovery Black holeを解決することができる

Path MTU Discovery Black hole問題の解決法2-1



- TunnelのMTUを1500に拡張する
  - TunnelのMTUを1500に拡張することで、tunnel区間をパケットを分割して通過させることができる
  - TCP(Protocol:6)以外のUDP(Protocol:17)やESP(Protocol:50)などの1500バイトパケットを通すことができる
  - Path MTU Discovery以外の要因によるDF=1のIPにも対応が可能

Path MTU Discovery Black hole問題の解決法2-2

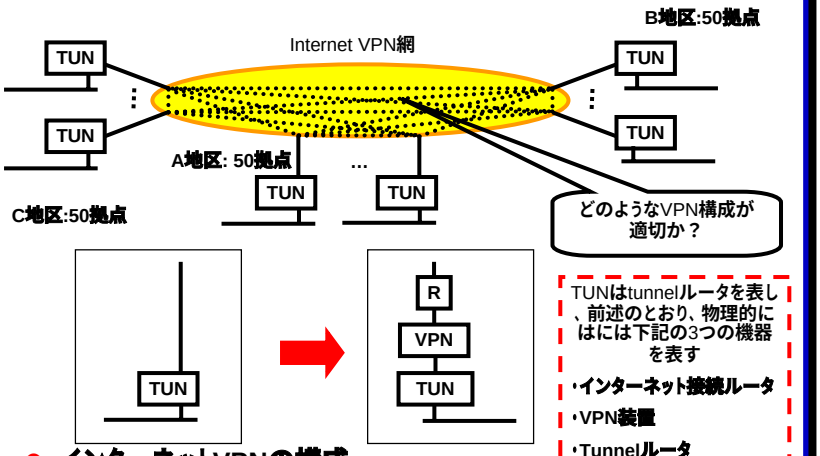


- TunnelのMTUを1500に拡張した場合の問題点
  - 多量の物理MTUを越えるパケットをMTU拡張したVPNに流すとTunnelの出口でルータがパケットをアッセンブルする負荷が増大し、パフォーマンスの著しい低下を招くことがある。
- 対策
  - TCPのMSS機能を併用して物理MTUを越えないパケットを流すようにする
  - MTUを物理MTUに合わせ、ICMP Unreachableを通すように設定し、Path MTU Discoveryを正常に機能させる

インターネットVPNの問題点の解決

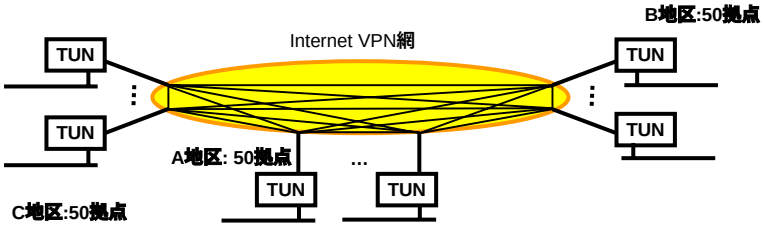
- MTUが1500より小さくなることによる問題点の解決
  - TCPについてはMSS調整により解決を行う
    - Path MTU Discovery Black holeの解決とスループット低下の防止が同時に行われる
    - 多くのアプリケーションがTCPを利用しているため、MSS調整により問題が解消することが多い
  - TCP以外のプロトコルはMTU拡張により解決を行う
    - 暗号化パケットのESPやUDPなどTCPでないプロトコルの解決にはMTU拡張を行い、パケットを分割して通すようにする。
    - パケットを分割することでパフォーマンスは低下するが、すべてのIPパケットを通すことができる
    - MTU拡張による解決法では多量のトラフィックをさばくことができないため、MSS調整の併用やPath MTU Discoveryが正常に動作するネットワークを構築する必要がある
  - 2つの手法の併用
    - 「MSS調整」と「MTU拡張」を同時に設定することですべてのIPパケットが通るだけでなく、TCPは効率よく通すことができる。
    - MTUを設定すると自動的にMSS値が決定するようなVPN機器、tunnelルータは2つの手法の併用はできない。
      - Tunnel MTU=1500→MSS=1460ではMTU1300の物理I/Fに対しMSS調整されたパケットがフラグメントしてしまい効率よく転送することができない
      - Tunnel MTU=1300→MSS=1260ではTCP以外のDF=1のIPパケットが通らない。1300バイトより大きいDF=1のUDP、ESPが通らない

多拠点でのインターネットVPN接続



- インターネットVPNの構成
  - インターネットVPNは各拠点間を自由に結ぶことができる
  - どのような接続法が望ましいのか検討する

VPNフルメッシュ構成

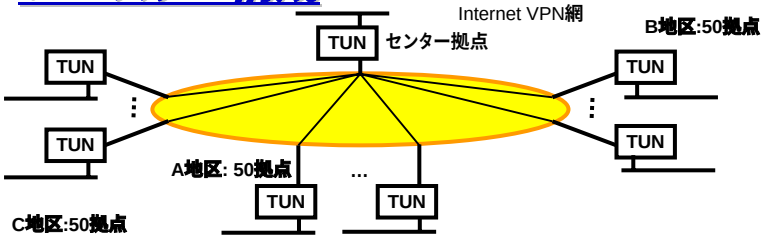


- フルメッシュ構成
  - すべての拠点間をVPNで結ぶ
  - 拠点の数:mとしたときのVPNの数

$$\text{VPNの数} = \frac{m(m-1)}{2} = \frac{m^2 - m}{2}$$

  - VPNの数が拠点数の二乗に比例するため多拠点の管理が煩雑
  - 1拠点追加ですべての拠点のVPN機器の変更が必要
  - 拠点数が増えるとすべての拠点のVPN機器の性能を上げる必要がある

VPNスター構成

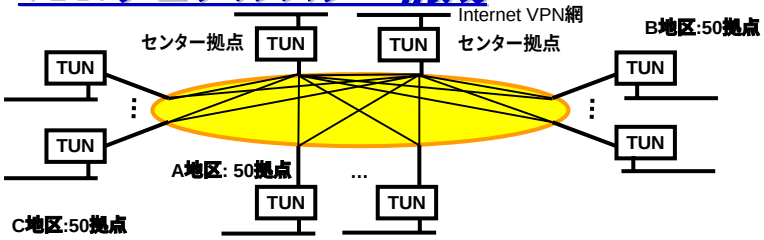


- スター構成
  - 各拠点とセンター拠点との間をVPNで結ぶ
  - 拠点の数:mとしたときのVPNの数は

$$\text{VPNの数} = m - 1$$

  - VPNの数は拠点数に比例するため多拠点の管理が容易
  - 1拠点追加でセンター拠点のVPN機器の変更だけで済む
  - スターの中心となるセンター拠点のtunnelルータの障害ですべての拠点の通信ができなくなる
  - 拠点数が増えるとセンター拠点に多数のVPNを収容する必要があり、高性能なVPN機器が必要となる

VPNデュアルスター構成

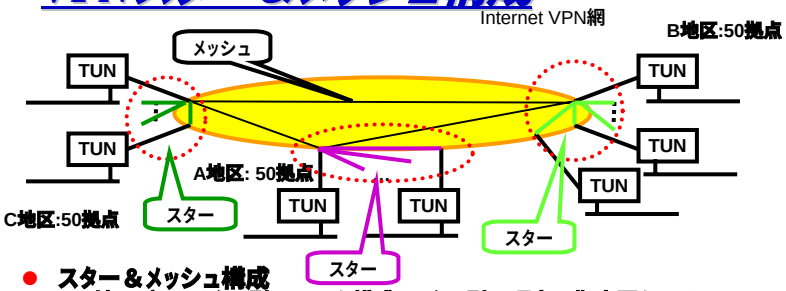


- デュアルスター構成
  - 各拠点から2箇所のセンター拠点にVPNで結ぶ
  - 拠点の数:mとしたときのVPNの数は

$$\text{VPNの数} = 2(m-1)$$

  - VPNの数は拠点数に比例するため多拠点の管理が容易
  - 1拠点追加でセンター拠点のVPN機器の変更だけで済む
  - スターの中心となるセンター拠点のtunnelルータに障害が発生してもバックアップされる
  - 拠点数が増えるとセンター拠点に多数のVPNを収容する必要があり、高性能なVPN機器が必要となる
  - 地区ごとにデュアルスター構成を行い、さらにセンター拠点ともデュアル構成とすることでVPN機器の負荷の集中を防ぐことができる

VPNスター＆メッシュ構成



- スター＆メッシュ構成
  - 地区ごとにスター型にVPNを構成。スター型の頂点の拠点間をメッシュ状にVPNで結ぶ
  - 拠点の数:m、スターの頂点をnとしたときのVPNの数は

$$\text{VPNの数} = m - n + \frac{(n^2-n)}{2} = \frac{n^2-3n+2m}{2}$$

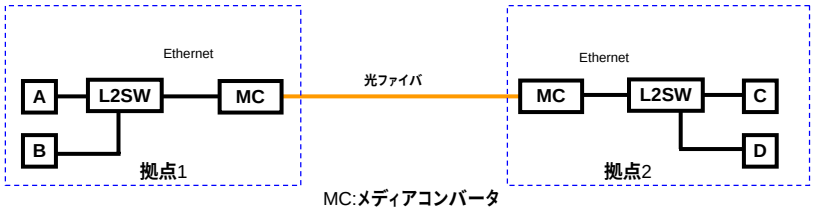
  - VPNの数は拠点数に比例するため多拠点の管理が容易
  - 頂点の数はVPNの数の二乗に比例するが、多く設置する必要は無いためVPNの数への影響は少ない
  - 1拠点追加でセンター拠点のVPN機器の変更だけで済む
  - スターの頂点のtunnelルータの障害が発生しても全体ではなく局所化した障害となる
  - スター構成をデュアルスター構成に変更すればバックアップも可能

ダークファイバを利用したWAN

- ダークファイバを利用する理由
  - 広帯域、安価
  - IP以外のパケットが通る(SNAなど)
  - Ethernetだけでなく ATMやPOSとしても利用可能
  - Tag VLANを利用して複数のVLANを複数拠点に容易に持っていくことができる
  - STPなどを用いてL2冗長化ネットワークを構築することができる
- 今回とりあげるポイント
  - IPのみを利用、L2SW、L3SWを利用、STP
  - ダイナミックルーティングによる冗長構成は広域Ethernetと同様に構成することができるため、ここではSTPによるL2冗長化ネットワークについて解説する

STP: Spanning Tree Protocol (IEEE802.1D)

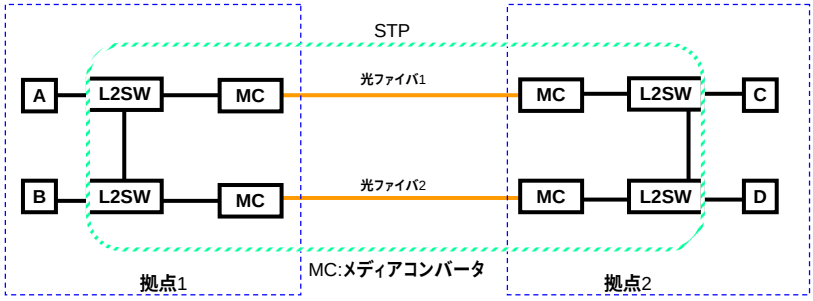
ダークファイバWANネットワーク構成



MC:メディアコンバータ

- ダークファイバによるネットワーク構成では光ファイバにMC(メディアコンバータ)を接続し、利用しやすいEthernetに変換する
- 拠点内LANではL2SWでネットワークを構成することで、同じL2ネットワークを複数の拠点で共有する
- 広域Ethernetとは異なり、Point to Pointネットワークとなる

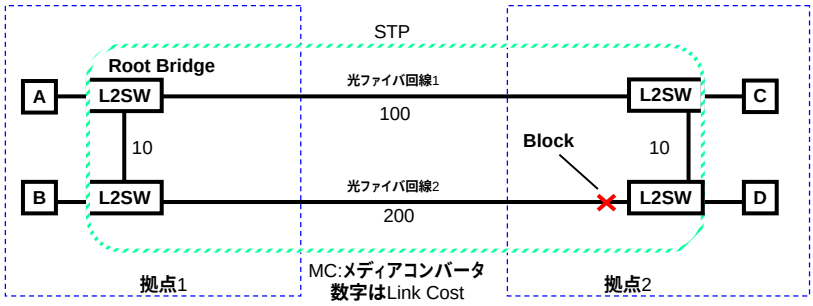
ダークファイバWAN冗長化ネットワーク物理構成



- ダークファイバを利用し、L2レベルでの冗長化を行うにはSTPを利用する必要がある
- STPはほとんどのキャリアサービスでは利用できないが、ダークファイバでは利用することができる。このため、L2レベルでの冗長化を行うことが可能となる
- ダークファイバを利用していてもL3レベルでの冗長が可能な場合にはダイナミックルーティングを利用した冗長構成とした方がよい

STP: Spanning Tree Protocol (IEEE802.1D)

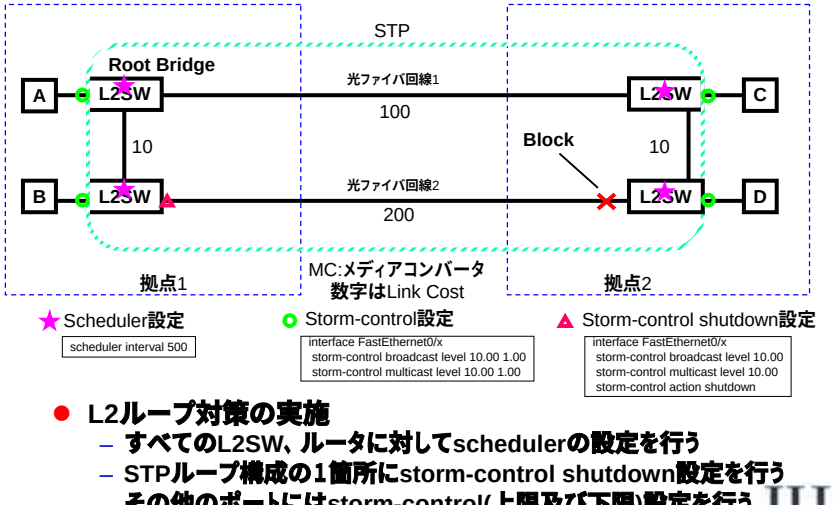
ダークファイバWAN冗長化ネットワーク論理構成



- メディアコンバータはリピータタイプを想定し、BPDUを透過するものとして論理構成には記載していない
- STPを利用するとRoot Brigdeからツリーが形成され、BPDU Root Path Costの大きいリンクがBlockされ、ループを回避する
- L2ループ対策としてschedulerやStorm-Controlの設定も同時に実施する

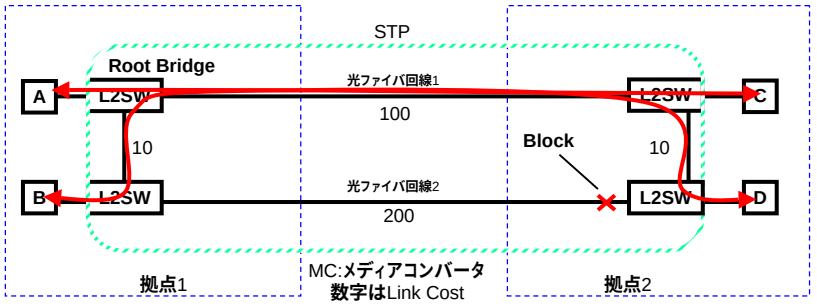
BPDU: Bridge Protocol Data Unit

ダークファイバWAN冗長化ネットワークL2ループ対策



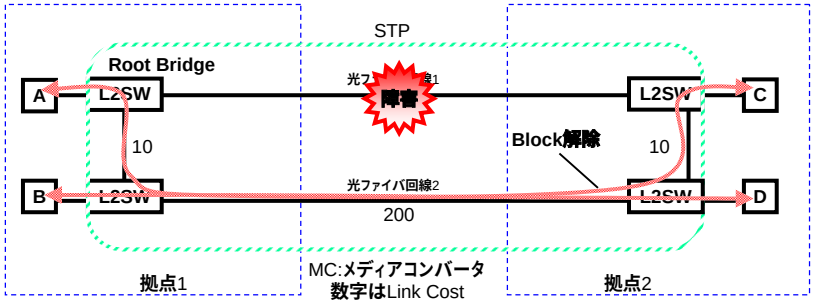
- L2ループ対策の実施
  - すべてのL2SW、ルータに対してschedulerの設定を行う
  - STPループ構成の1箇所にstorm-control shutdown設定を行う
  - その他のポートにはstorm-control(上限及び下限)設定を行う

ダークファイバWAN冗長化構成　通常時



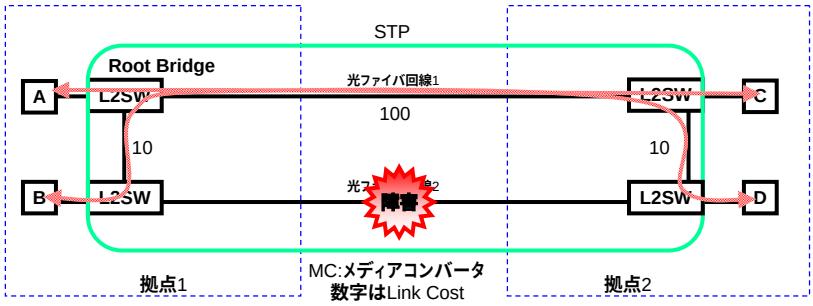
- 通常時は光ファイバ回線1のみ利用される
- STPにより光ファイバ2回線はBlockされ、通常時トラフィックは流れない

ダークファイバWAN冗長化構成　障害時



- 光ファイバ回線1障害時は光ファイバ回線2によりバックアップされる

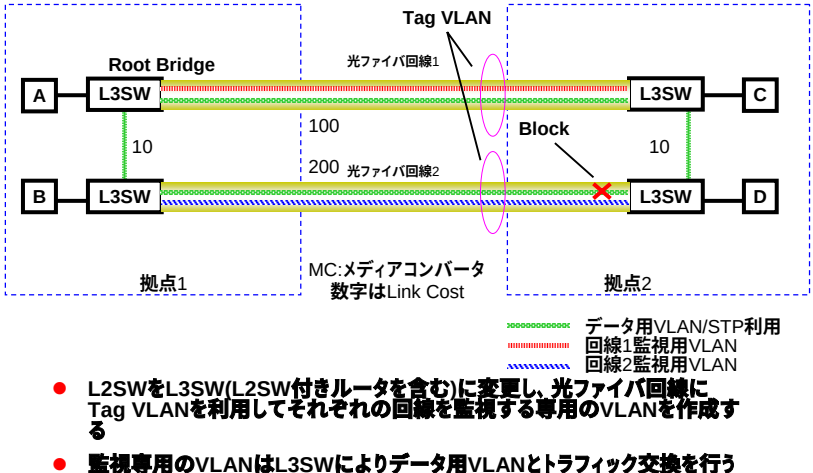
ダークファイバWAN冗長化構成　バックアップ回線障害時



- バックアップ回線(Blockされている回線)に障害が発生しても通信に影響を与えることはない
- L2冗長化構成では回線のみをpingで監視することは難しい

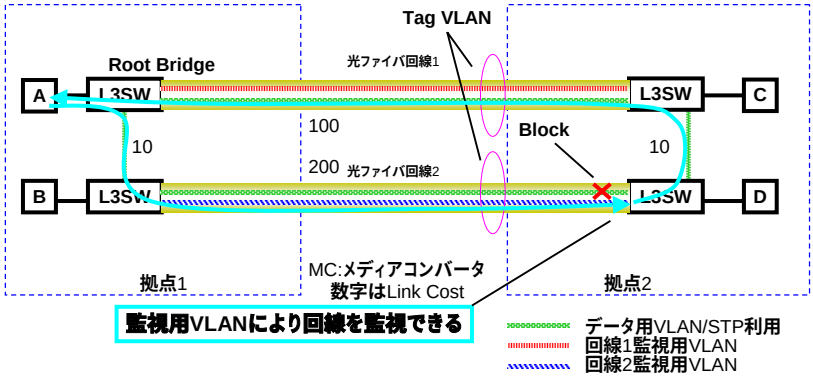
回線部分をpingで監視するためには

ダークファイバWAN冗長化ネットワーク監視構成



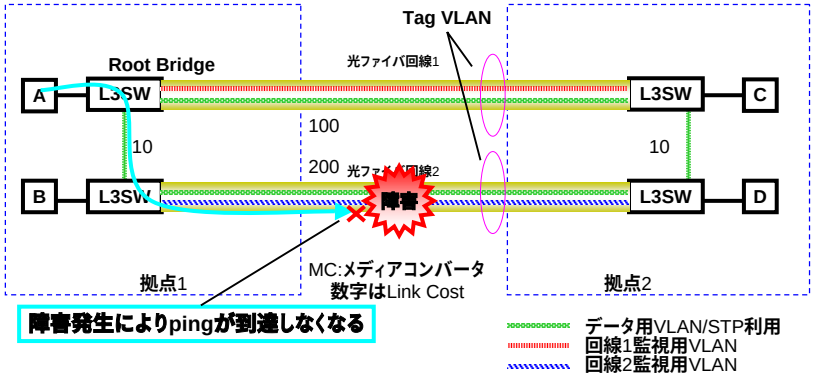
- L2SWをL3SW(L2SW付きルータを含む)に変更し、光ファイバ回線にTag VLANを利用してそれぞれの回線を監視する専用のVLANを作成する
- 監視専用のVLANはL3SWによりデータ用VLANとトラフィック交換を行う
  - データ用VLANを利用してルーティングを行う必要がある
  - L2SWではVLAN毎にIPアドレスを付けたり、ルーティングを行うことはできないため、L3SWを利用する必要がある

ダークファイバWAN冗長化監視状態



- 監視用VLANに付けられたL3SWのアドレスを利用すると、pingにより監視を行うことができる
- Pingの返答はデータ用VLANから戻ってくる

ダークファイバWAN冗長化障害検知



- 障害が発生すると回線監視用VLANへの到達生が失われ、障害を検知することができる
- 光ファイバ回線1も同様にpingによる監視と障害検知を行うことが可能

ダークファイバを利用したWAN: まとめ

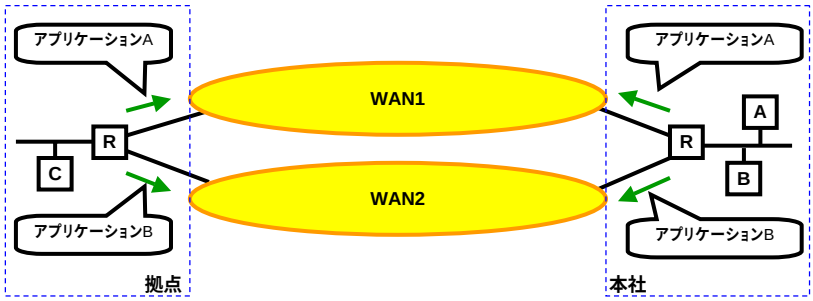
- L3ネットワーク構成の場合
  - ダークファイバであってもL3ネットワーク構成を組むことができる場合には専用線と同じくダイナミックルーティングによりWANを構成した方がよい
  - 広域Ethernetと同様に利用することができるが、ネットワーク的には専用線と同じくPoint to Pointの構成となる
- L2ネットワーク構成の場合
  - L2ネットワーク構成とする場合にはSTPを利用して冗長化ネットワークを構成することができる
- L2ネットワーク冗長構成でping監視を行う場合
  - L2ネットワーク冗長構成でping監視を行うにはTag VLANを利用して回線ごとに監視用VLANを作成する
  - 監視用VLANに個別にIPアドレスを付与するためにL3SW(L2SW付きルータを含む)が必要となる
  - 監視用VLANに付与したIPアドレスにping監視を行うことで、回線障害を検知することができる

ポリシーごとのWANの使い分け

VoIPや基幹系データなど高信頼性を要求される通信とインターネットアクセスなど信頼性よりも広帯域を要求される通信を複数のWANを使って使い分ける方法を解説する

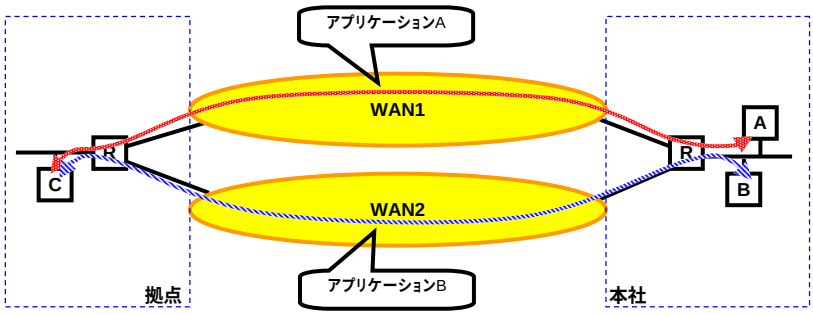
- PBRを利用したWANの使い分け
- ネットワークトポロジによりWANの使い分け

PBRを利用したWANの使い分け設定



- 本社側のアプリケーションサーバAはWAN1を利用し、アプリケーションサーバBはWAN2を利用するようにルータにてPBR(Policy-Based Routing)を設定する
- PBRはIPアドレスごと、利用アプリケーション（利用ポート）などによりNext Hopを指定することができる
- 通常のルーティングではdestination（宛先アドレス）ごとに経路を設定することしかできないが、PBRによりsource(送信元アドレス)や利用ポートなどにより経路を設定することができる

PBRを利用したWANの使い分け

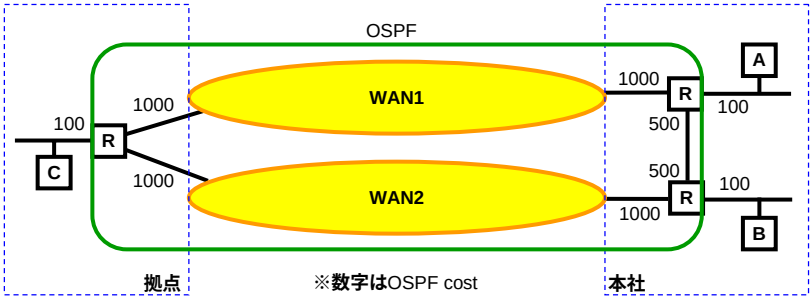


- 拠点側クライアントCよりアプリケーションサーバAおよびBにアクセスする回線を使い分けることができる
- PBRはIPアドレスごと、利用アプリケーション（利用ポート）などによりNext Hopを指定することができる

PBRを利用したWANの使い分けのまとめ

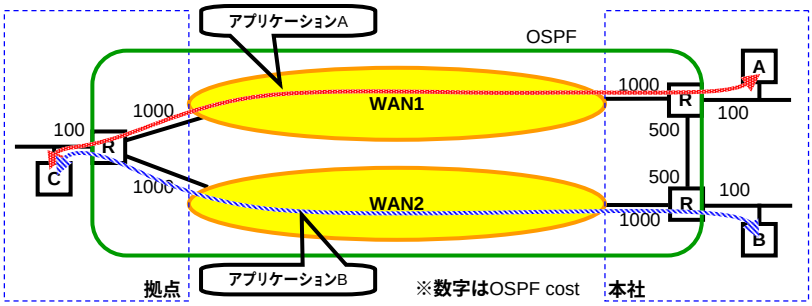
- PBR(Policy-Based Routing)はIPアドレスごと、利用アプリケーション（利用ポート）などによりNext Hopを指定することができる
- Next Hopを異なったWANに設定することでアプリケーションごとに利用回線を使い分けることができる
- 通常のルーティングではdestination（宛先アドレス）ごとに経路を設定することしかできないが、PBRによりsource(送信元アドレス)や利用ポートなどにより経路を設定することができる
- PBRはstaticと同様に固定的に経路を設定するため、2つのWANの相互の冗長化は難しい

ネットワークトポロジによるWANの使い分け設定



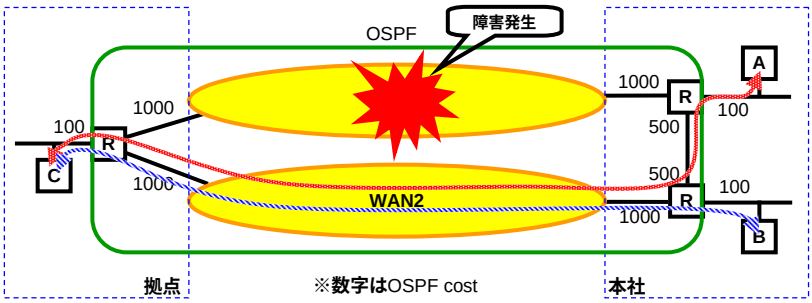
- 本社側のアプリケーションサーバAはWAN1を収容しているルータに接続し、アプリケーションサーバBはWAN2を収容しているルータに収容する。
- 本社ルータ間は高めのOSPF cost(500)を設定する
- WANにはLANより高いOSPF cost(1000)を設定する
- 拠点側ルータのWAN側は本社側と同じOSPF cost(1000)を設定する

ネットワークトポロジによるWANの使い分け



- 拠点側ルータから見るとAに対してはWAN1を経由しBに対してはWAN2を経由するように経路制御される
- 本社側WAN1収容ルータはCに対してWAN1を経由し、WAN2収容ルータはCに対してWAN2を経由するように経路制御される
- 実質的にアプリケーションごとにWANを使い分けることができる

ネットワークトポロジによるWANの使い分け　障害時



- WAN1に障害が発生した場合にはWAN2を利用してバックアップすることができる
- 障害復旧時は自動的に復旧する
- WAN2障害時にも同様にバックアップを行うことが可能

ネットワークトポロジによるWANの使い分けのまとめ

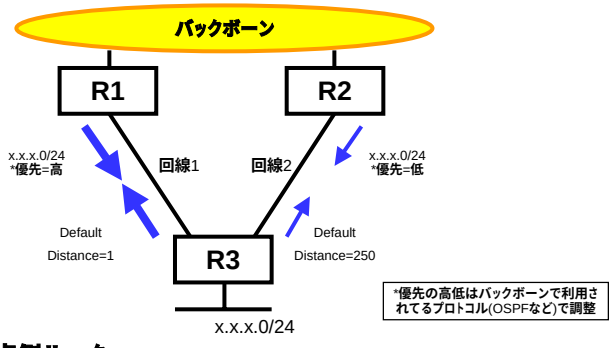
- 本社側のアプリケーションサーバAはWAN1を収容しているルータに接続し、アプリケーションサーバBはWAN2を収容しているルータに収容する。
- WAN収容ルータ間は高めのOSPF costを設定し、WANにはLANより高いOSPF costを設定する
- OSPF cost調整により、実質的にアプリケーションごとにWANを使い分けることができる
- OSPFなどのダイナミックルーティングを利用できるため、冗長化が可能
- RIPなどのOSPF以外のダイナミックルーティングでも実現可能
- 既存ネットワークによってはIPアドレスのリナンバなどの構成変更が必要になる

フローティングスタティックを利用したバックアップ

フローティングスタティックを利用したバックアップ手法を解説する

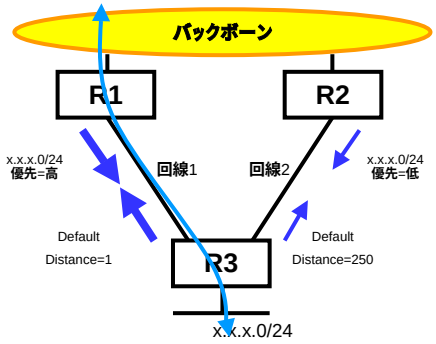
- スタティック・スタティック　バックアップ
  - スタティックルーティングだけでバックアップを実現
- ダイナミック・スタティック　バックアップ
  - ダイナミックルーティングとスタティックルーティングを併用してバックアップを実現

スタティック・スタティックバックアップ　設定



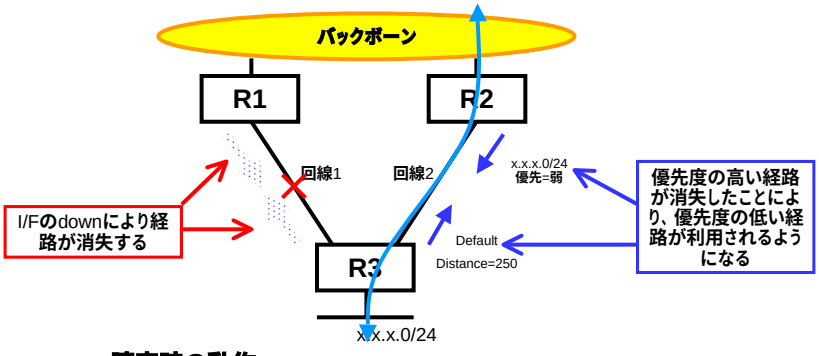
- 拠点側ルータ
  - 優先度の高いdefault(distance=1)を回線1に設定し、優先度の低いdefault(distance=250)を回線2に設定する。
- センター側ルータ
  - 優先度の高くした拠点向け経路(x.x.x.0/24)を回線1に設定し、優先度を低くした拠点向け経路を回線2に設定する
  - バックボーン側の優先度はOSPF costなどで調整

スタティック・スタティックバックアップ　通常時



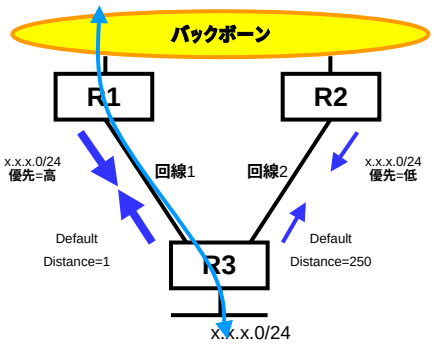
- 通常時の動作
  - 回線1に設定された優先度の高い経路が有効となる
- 通常時のトラフィック
  - 回線1のみ利用される

スタティック・スタティックバックアップ　障害時



- 障害時の動作
  - 回線障害などの要因によりI/Fがdownし、優先度の高い経路が消失する
  - 優先度の高い経路が消失することにより優先度の低かった経路が選択され、利用されるようになる
- トラフィック
  - 回線2によりバックアップされる

スタティック・スタティックバックアップ　復旧時

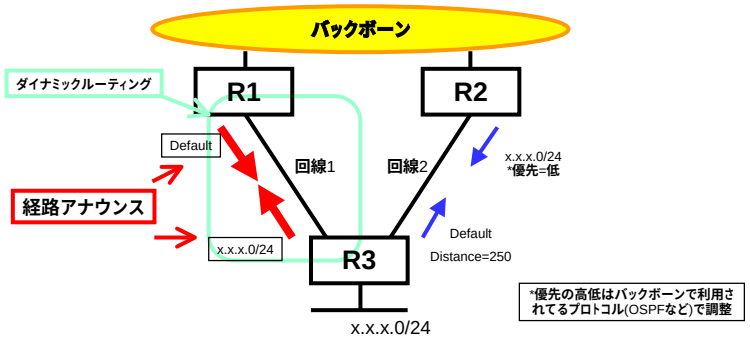


- 復旧時の動作
  - 回線が復旧し、I/Fがupすることで回線1に設定された優先度の高い経路が有効となる
- 通常時のトラフィック
  - ふたたび回線1のみ利用されるようになる

スタティック・スタティックバックアップ　特徴

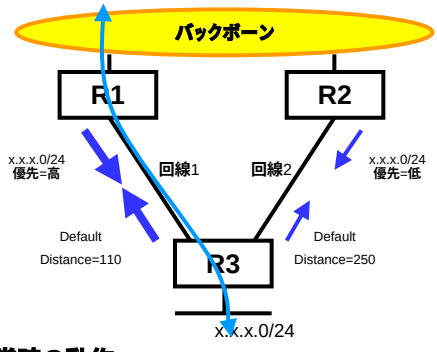
- スタティック・スタティックバックアップ
  - フローティングスタティックを利用したスタティック・スタティックバックアップはI/F downにより回線障害を検知し、優先度の低いスタティックを有効にすることでバックアップを実現する
  - ダイナミックルーティングを利用せずに容易にバックアップが実現できる
  - 回線障害時にI/Fがdownしない回線は利用できない
    - Ethernet専用線
    - PPPoEなどを利用した回線
    - VPN/tunnel
    - HUBを経由したLAN

ダイナミック・スタティックバックアップ　設定



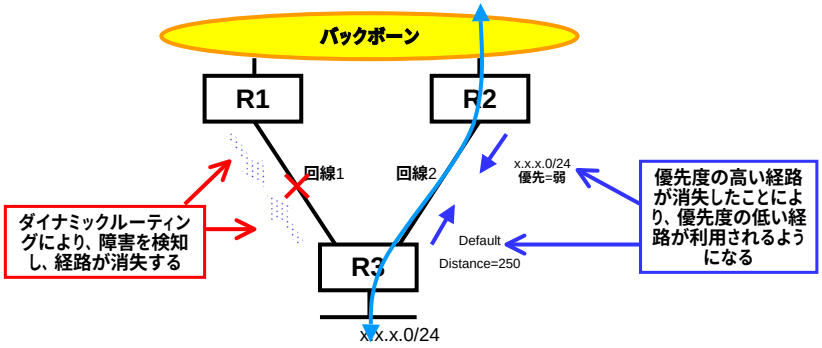
- 拠点側ルータ
  - ダイナミックルーティングを利用して回線1に拠点側経路(x.x.x.0/24)をアナウンスする。優先度の低いdefault(distance=250)を回線2に設定する。
- センター側ルータ
  - ダイナミックルーティングを利用してdefaultを回線1にアナウンスする。優先度を低くした拠点向け経路を回線2に設定する
  - バックボーン側の優先度はOSPF　costなどで調整

ダイナミック・スタティックバックアップ　通常時



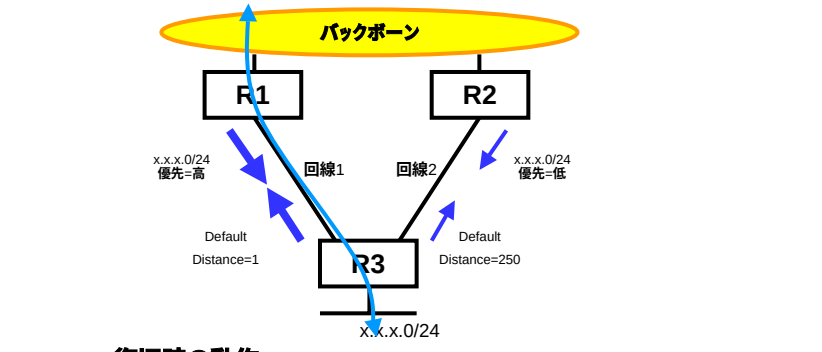
- 通常時の動作
  - 回線1を利用したダイナミックルーティングにより優先度の高い経路が設定される。
- 通常時のトラフィック
  - 回線1のみ利用される

ダイナミック・スタティックバックアップ　障害時



- 障害時の動作
  - 回線障害などの原因をダイナミックルーティングが検知し、優先度の高い経路が消失する
  - 優先度の高い経路が消失することにより優先度の低かった経路が選択され、利用されるようになる
- トラフィック
  - 回線2によりバックアップされる

ダイナミック・スタティックバックアップ　復旧時



- 復旧時の動作
  - 回線が復旧し、ダイナミックルーティングが再び有効となり、回線1に優先度の高い経路が設定される
- 通常時のトラフィック
  - ふたたび回線1のみ利用されるようになる

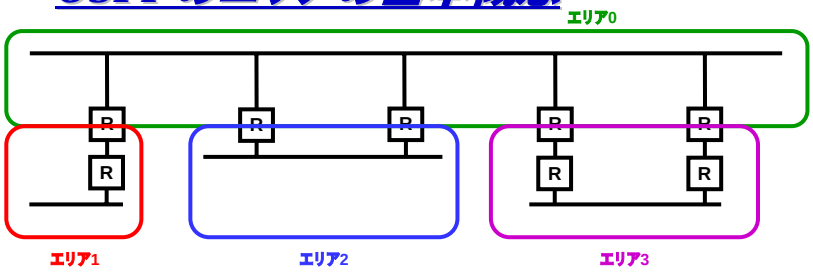
ダイナミック・スタティックバックアップ　特徴

- ダイナミック・スタティックバックアップ
  - フローティングスタティックを利用したダイナミック・スタティックバックアップはダイナミックルーティングにより回線障害を検知し、優先度の低いスタティックを有効にすることでバックアップを実現する
  - 回線障害時にI/Fがdownしない回線であっても利用が可能
    - Ethernet専用線
    - PPPoEなどを利用した回線
    - VPN/tunnel
    - HUBを経由したLAN
  - バックアップ側回線のトラフィックを通常時ゼロにすることができるため、ISDNなどの回線をバックアップに利用することが可能

OSPFのエリアを利用した構築

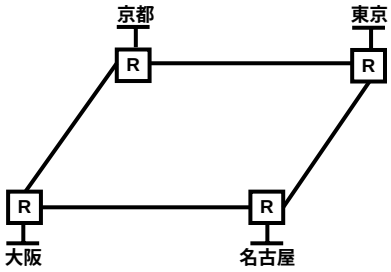
- OSPFのエリア
  - OSPFは経路集約のためにネットワークをいくつかのエリアに分けることができる。
  - 先の説明にあったように小規模から中規模のネットワークではエリアを分ける必要はほとんどない
  - 利用機器の制約からBGPを利用することができず、OSPFのみで大規模なネットワークを設計する場合にはエリア分けを検討する価値はある
- 間違いやすいOSPFのエリア
  - 間違いやすいOSPFのエリアの概念を具体例を示して解説する

OSPFのエリアの基本概念



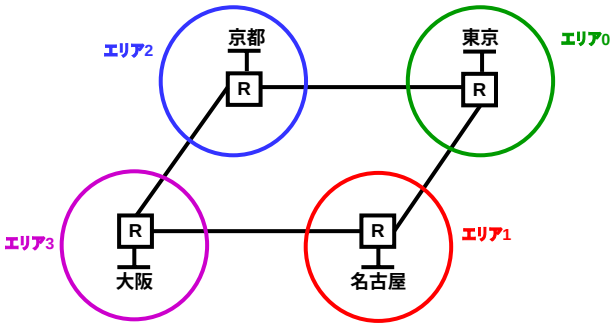
- バックボーンエリア
  - バックボーンエリアと呼ばれるエリア0は必ず存在しなければならない
- バックボーン以外のエリアはバックボーンエリアに接していなければならない
  - エリア1、エリア2、エリア3は必ずエリア0に接している
- エリア境界はルータになる
  - ネットワークがエリア境界となることはない
  - 同じネットワークは同じエリアになる
- 同じI/Fに複数のエリアを設定することはできない
  - 1つのルータのI/Fには1つのエリアにだけに属することができる
  - Passive interface設定であっても1つのエリアにだけ属する

OSPFエリア構築事例1



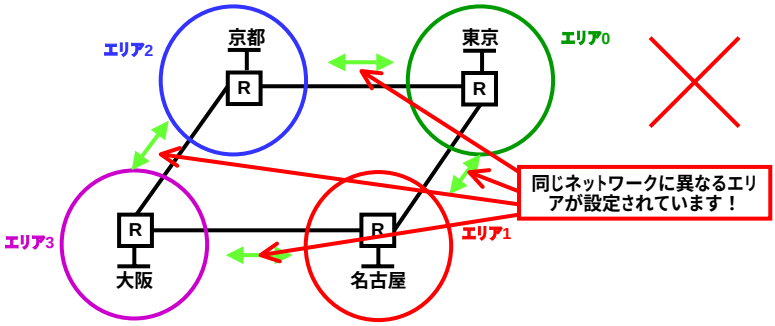
- 4拠点を結ぶバックボーンを適切にエリア分けする
- 各拠点には実際には多くのネットワークがあることを想定
- イメージとしては各拠点にエリアを分ければよいように見える

OSPF エリア構築事例1 回答例1



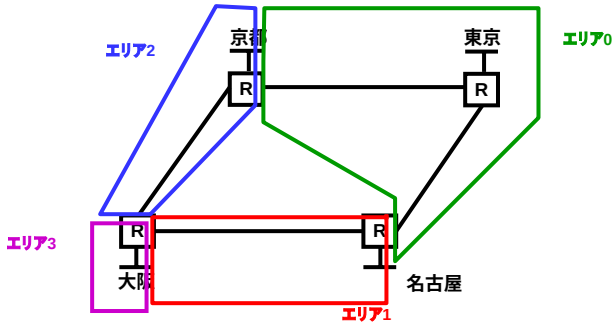
- 拠点ごとにそれぞれをエリアとした場合
- 一見よさそうに見えますが？

OSPF エリア構築事例1 回答例1



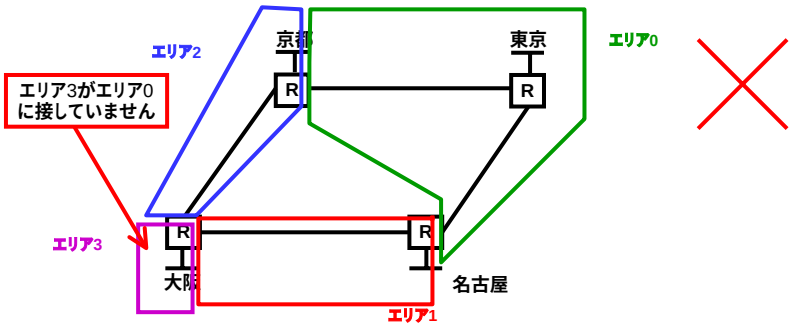
- 同じネットワークに異なるエリアが設定されているため動作しません！

OSPFエリア構築事例1 回答例2



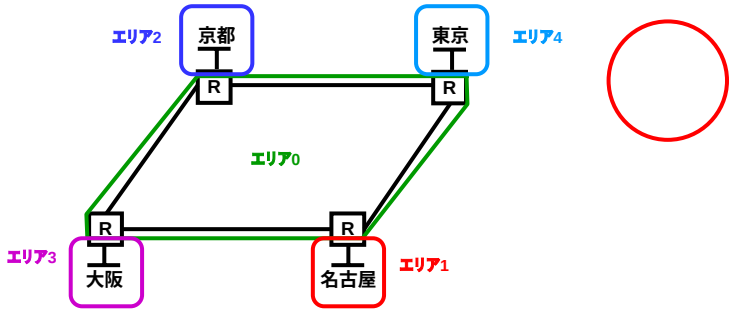
- 東京側から大阪側にかけて同じインターフェースを同じエリアにするように設定
- 一見よさそうに見えますが？

OSPFエリア構築事例1 回答例2



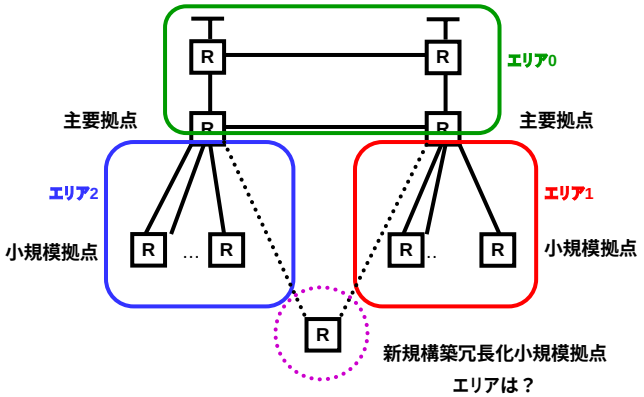
- エリア3がエリア0に接していないため、動作しません
- VL(Virtual Link)という技術により仮想的にエリア0に接するように見せ、動作させることもできますが、適切なエリア分けをした場合に比べて煩雑でバックアップに問題があります。東京に障害が発生すると正常にバックアップされません

OSPFエリア構築事例1 正解



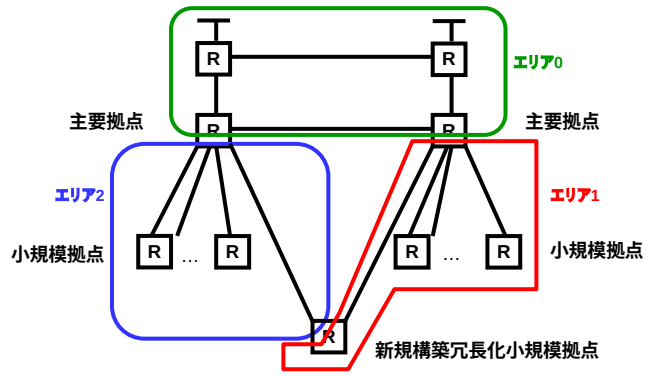
- 各拠点を結ぶネットワークにはエリア0を設定する
  - すべての拠点にエリア0が存在するため、各拠点に自由にエリアを追加することができる
  - 東京が障害となってもバックアップが可能
- 各拠点には各拠点に閉じるエリアを設定する
  - 各拠点に閉じる経路を集約することができる

OSPFエリア構築事例2



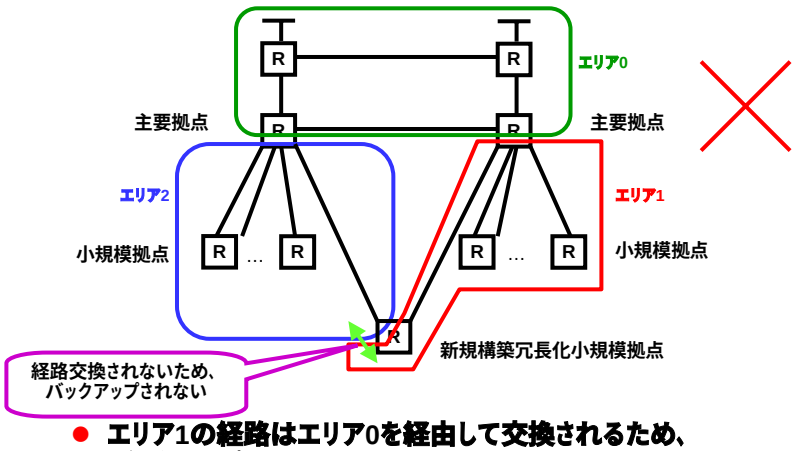
- 事例1に従い、バックボーンエリアと各拠点ごとに適切にエリアを分け、小規模拠点を主要拠点にまとめて収容している
- 冗長化のために複数の主要拠点から小規模拠点を接続する場合に、どのようにエリアを分ければよいか？

OSPFエリア構築事例2 回答例1



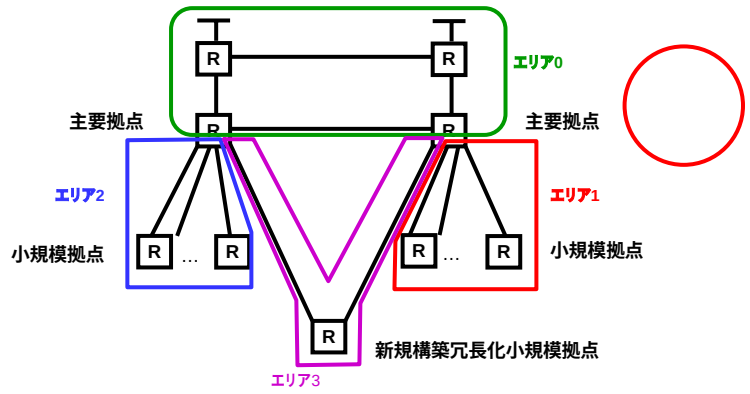
- 冗長化拠点をエリア1に属するように設定。
- エリア2からも1本冗長化拠点に伸ばす
- すべてのエリアがエリア0に接していて、一見よさそうに見えますが？

OSPFエリア構築事例2 回答例1



- エリア1の経路はエリア0を経由して交換されるため、バックアップされません

OSPFエリア構築事例2 正解



- 冗長化拠点用に新たなエリア3を設定
- エリア3は2箇所でエリア0に接し、それぞれが異なる回線を通るようにする
- 同様の構成の冗長化小規模拠点は同じエリア3に設定が可能

OSPFエリア利用時の注意点

- OSPFエリア番号の一意性確保
  - OSPFエリアはネットワーク全体で一意に確保されなければならない
  - 同じOSPFエリア番号が管理されずに別の場所で使われると正常にOSPFが動作しない
    - 異なるベンダーとの同時構築などでは注意が必要
    - エリア1,2,3,10,20,30,100,200,300などのきりの良い数字は要注意
  - 特定拠点に閉じるOSPFエリアは代表的なルータのIPアドレスを利用すると良い
    - エリア192.168.0.1といった具合となる
    - 代表的なルータ「エリア0との境界ルータ」のloopbackアドレスを利用すると良い
    - エリア番号をIPアドレスにするとtracerouteなどで、エリアの場所を簡単に特定できる

OSPFのエリアを利用した構築：まとめ

- OSPFエリア分けの方法
  - 主用拠点間など迂回経路となりうるすべてのルータにエリア0を設定する必要がある
  - 特定の拠点にのみ接続する小規模拠点などはエリアを分けることで経路集約が可能
- OSPFエリア分けの必要性
  - OSPFエリア分けの方法のひとつとして「すべてエリア0で構成する」という選択肢を忘れてはいけない
  - 大規模なネットワークはBGPなどOSPF以外のプロトコルを利用して多くの経路を交換することも検討すべき
- OSPFエリア番号の一意性確保
  - OSPFエリア番号の重複が発生しないようにエリア0境界ルータのIPアドレスを利用する

まとめ-1

- 一定の規模を超えるとスタティックルーティングよりダイナミックルーティングの方が容易に管理することができるようになる
- ダイナミックルーティングは「経路の流れる方法と、経路が向く方向が逆になる」という基本法則を理解すればどのようなIGPを利用してもネットワーク設計をすることができる
- ダイナミックルーティングを利用すれば障害に強いネットワークを構築できる
- OSPFを利用すればバランシングとバックアップを同時に実現可能

まとめ-2

- 広域Ethernetを利用した大規模なWANでは適切な規模ごとにネットワークを分離して細い回線の輻輳を防止する必要がある
- インターネットVPNではＶＰＮ装置だけでなくtunnelルータを設置することで専用線と同様にダイナミックルーティングを利用したネットワークを構築することができる
- インターネットVPNではPath MTU Discovery Block holeの解決をはかる必要がある
- ダークファイバを利用したWANであってもL3ネットワーク構成を組むことができる場合には専用線と同じくダイナミックルーティングによりWANを構成した方がよい
- ダークファイバを利用したWANでL2ネットワーク冗長構成を組むには、Tag VLANを利用し、データ用VLANはSTPで冗長化を図り、障害検知のため監視用VLANを回線ごとに作成する

まとめ-3

- PBRやネットワークトポロジにより複数のWAN回線をアプリケーションごとに使い分けることができる
- フローティングスタティックとダイナミックルーティングを併用してISDNなどを利用したバックアップを実現できる
- OSPFのエリアは通常はエリア0だけを利用して構築すればよいが、OSPFのみで大規模なネットワークを構築する際は適切なエリア分けを行わなければならない
- OSPFエリア番号の重複が発生しないようにIPアドレスをOSPFエリア番号として利用する