小～中規模のネットワーク構築 A to Z

2006年7月20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)

IIJ
Internet Initiative Jap

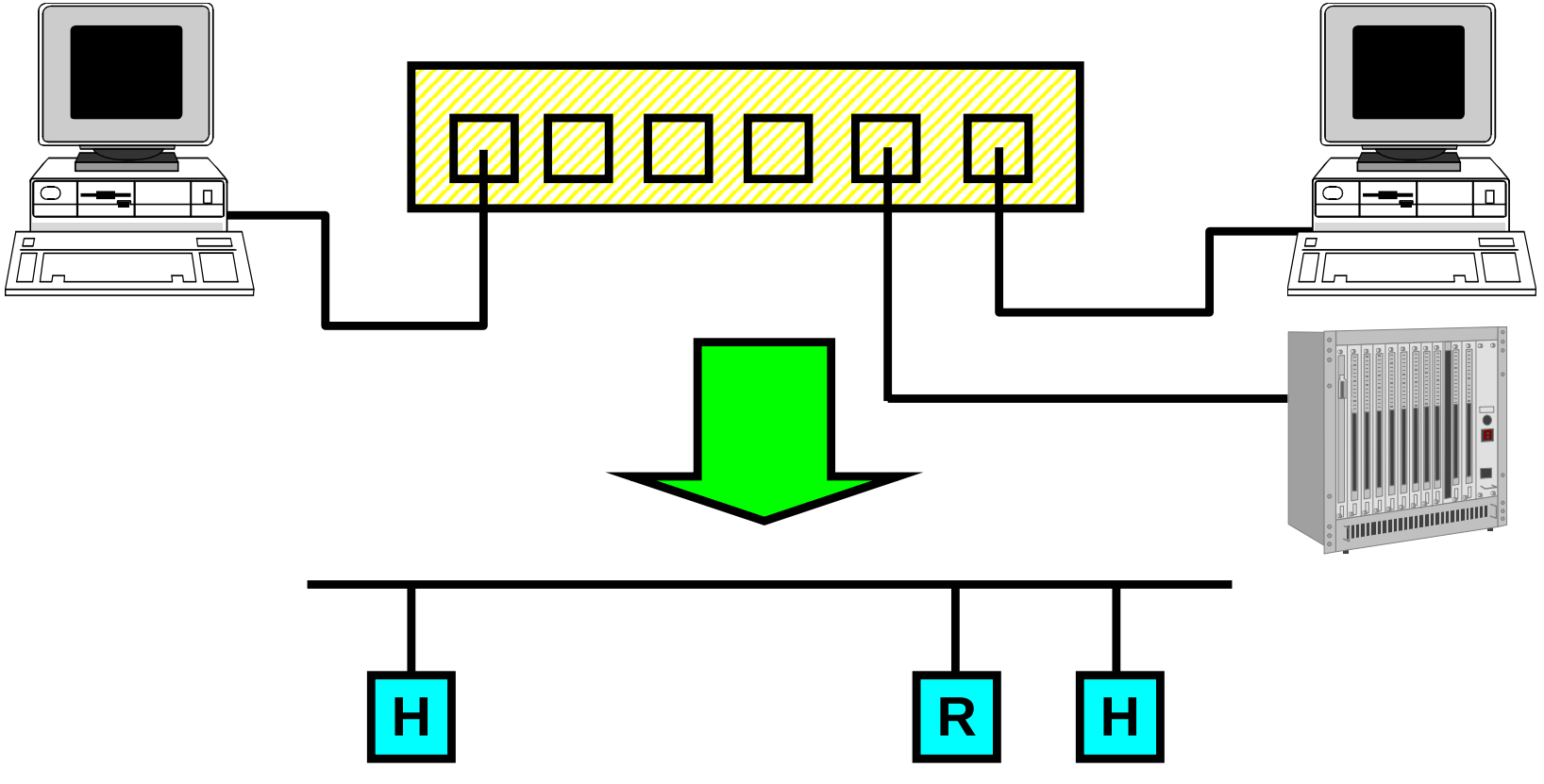
目的

- データリンク層とネットワーク層の役割を理解する
- 障害が起こりにくいネットワークを設計するには
- ネットワークの冗長化を行うには
- ルーティングが必要な理由
- 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

発表内容

- データリンク層とネットワーク層の役割
- ハブ、スイッチ、ルータの違い
- ネットワーク設計
- アドレスの割り当てポリシー
- 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
- ネットワーク構築
- ネットワークトラブルシューティング

ネットワーク表記

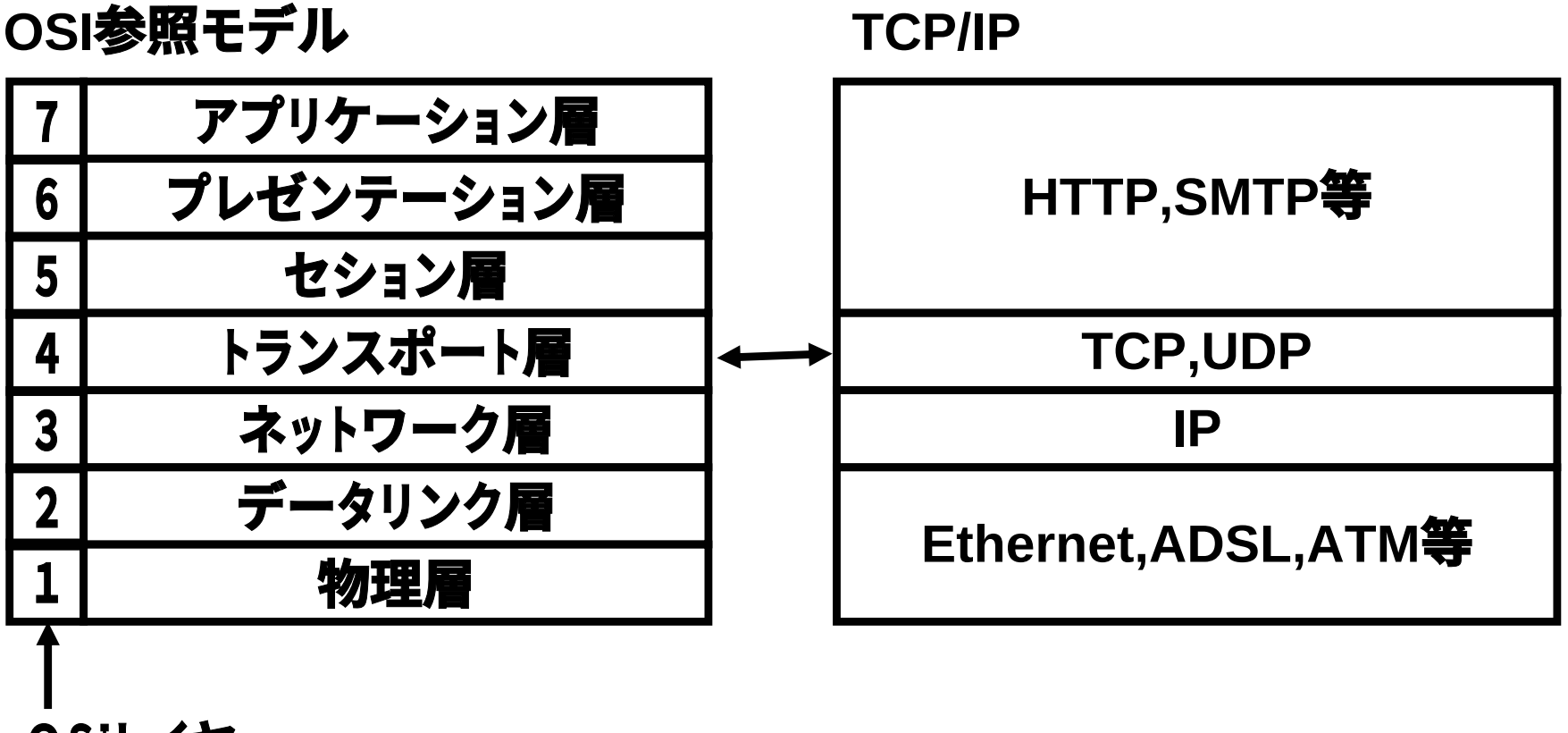


- ハブ、スイッチなどは1本の線またはSWで表わします。
- ホストはH、A、B、C、D等で、ルータはR等で表記します
- レイヤ3スイッチなどは説明中ではルータと区別していません

データリンクフレームとルーティング

- ここではデータリンク層とネットワーク層の役割を解説します
- MACアドレス（イーサネットアドレス）とIPアドレスの両方のアドレスが必要な訳
- データリンク層の種類
- ルーティングがなぜ必要なのか
- ルーティングがなくても通信できるのはなぜか

OSI参照モデルとTCP/IP



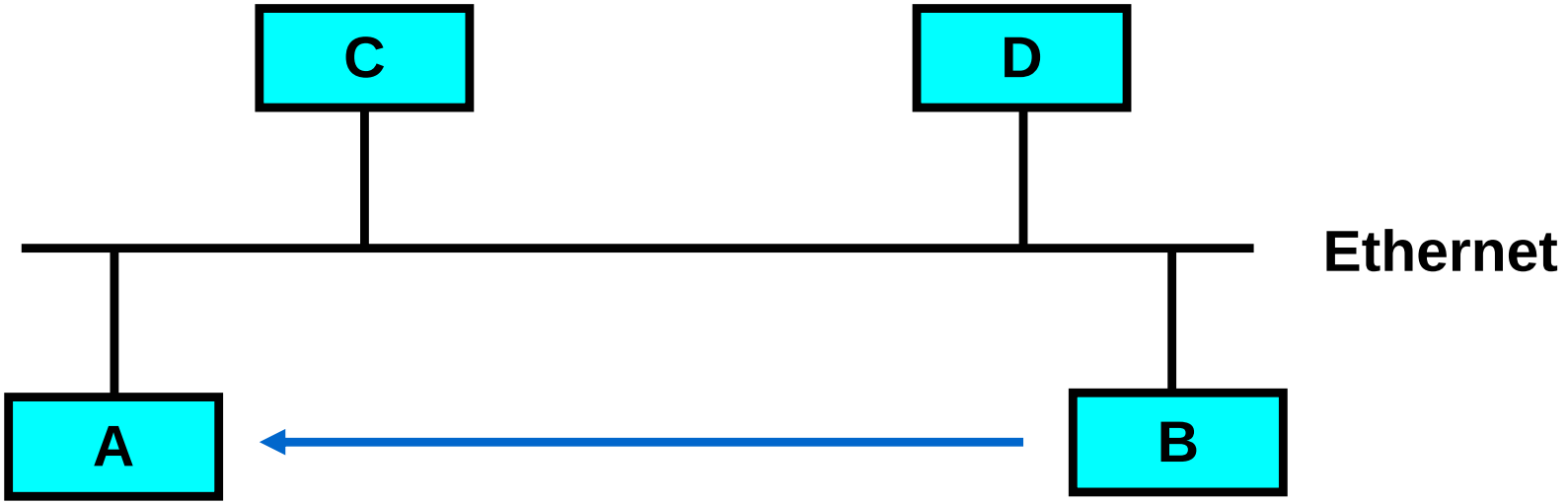
レイヤ2：データリンク層

レイヤ3：ネットワーク層

データリンク層の種類

- Multi Access Media（ARP)
  - MAC(Media Access Control)アドレスを用いて通信を行う
  - MACアドレスとIPアドレスとの対応はARP(Address Resolution Protocol)を用いる
  - Ethernet等
- Multi Access Media (固定)
  - 特定の識別子とIPアドレスに結び付け、固定的に設定を行う
  - フレームリレー、ATM等のMulti Access Mode
  - EthernetでIPアドレスとMACアドレスを固定的に設定
- Point to Point Media
  - 通信相手が物理もしくは仮想I/Fで特定されるもの
  - 64k,128k,1.5M,6M,45M,150M,600M,2.4G,10Gなどの専用線
  - フレームリレー、ATM等のPoint to Point Mode
  - PPPoEを利用したEthernet

ARPの動作ー1



Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	不明
C	192.168.0.3	不明
D	192.168.0.4	不明

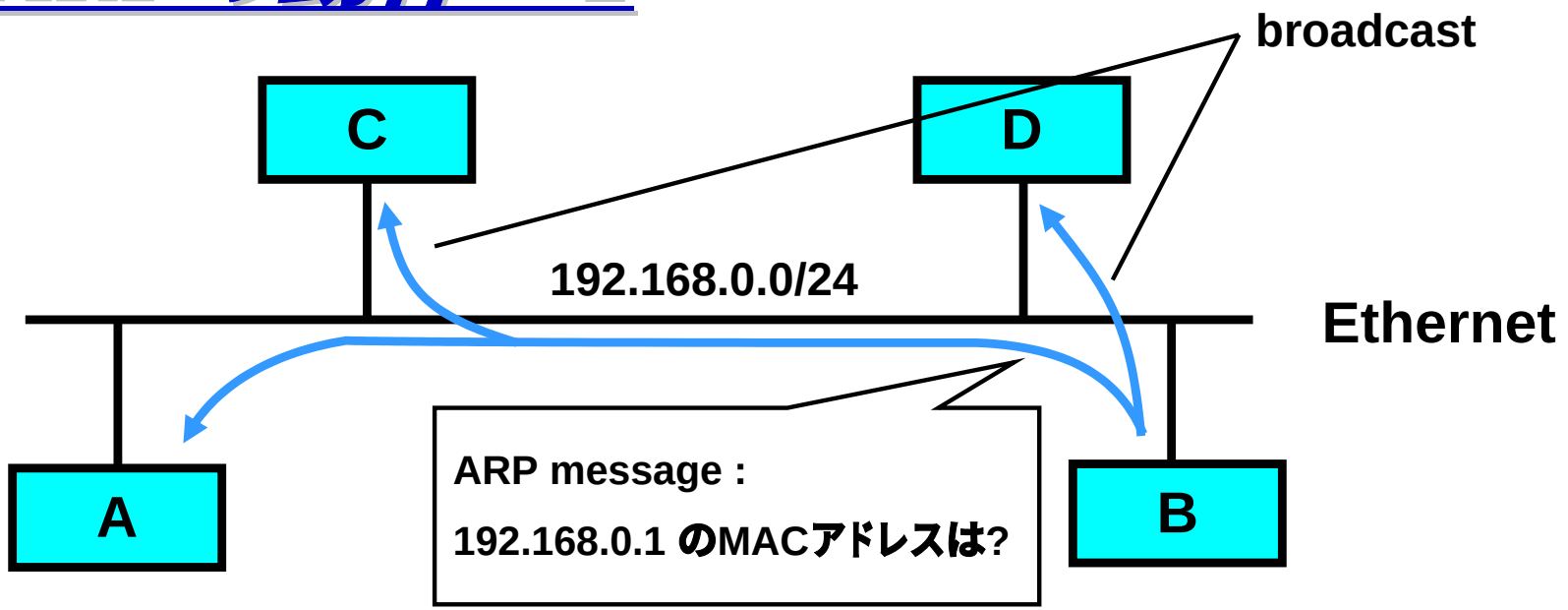
Host AのIP/MACアドレス対応表

Host	IPアドレス	MACアドレス
A	192.168.0.1	不明
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

Host BのIP/MACアドレス対応表

- BはAに通信したいが、BはAのMACアドレスがわからない

ARPの動作ー2



Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	不明
C	192.168.0.3	不明
D	192.168.0.4	不明

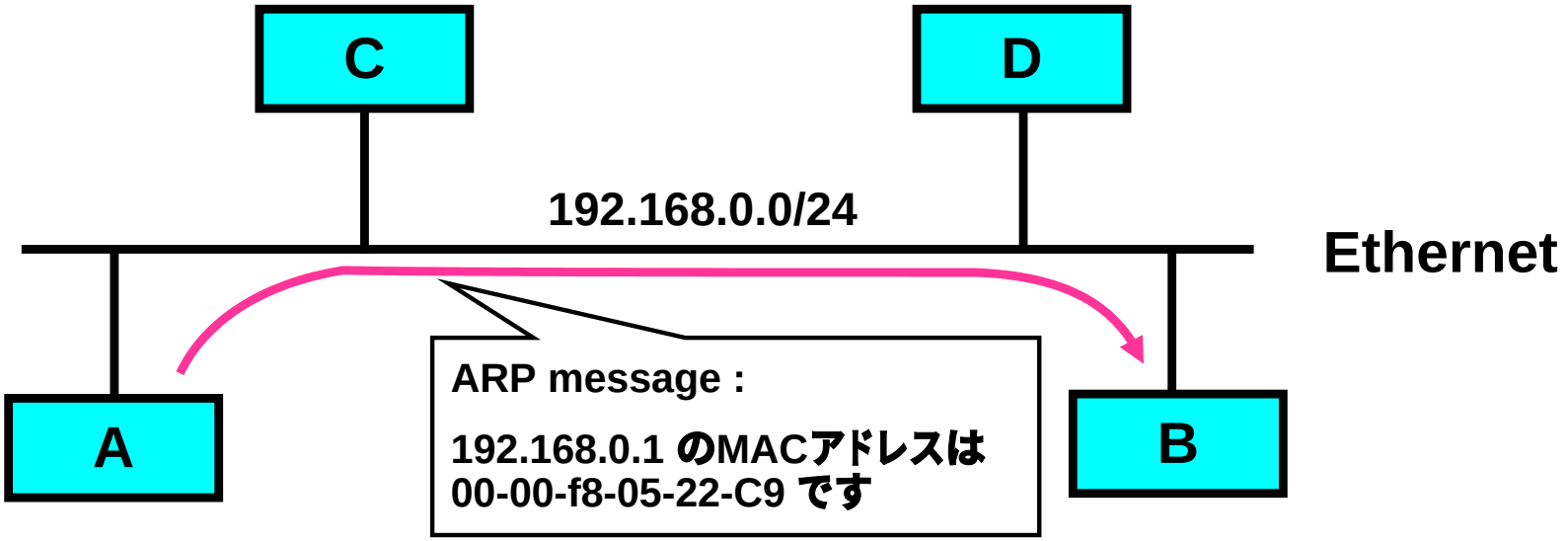
Host AのIP/MACアドレス対応表

Host	IPアドレス	MACアドレス
A	192.168.0.1	不明
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

Host BのIP/MACアドレス対応表

- BはAのMACアドレスを尋ねるメッセージをbroadcastする

ARPの動作ー3



Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	不明
C	192.168.0.3	不明
D	192.168.0.4	不明

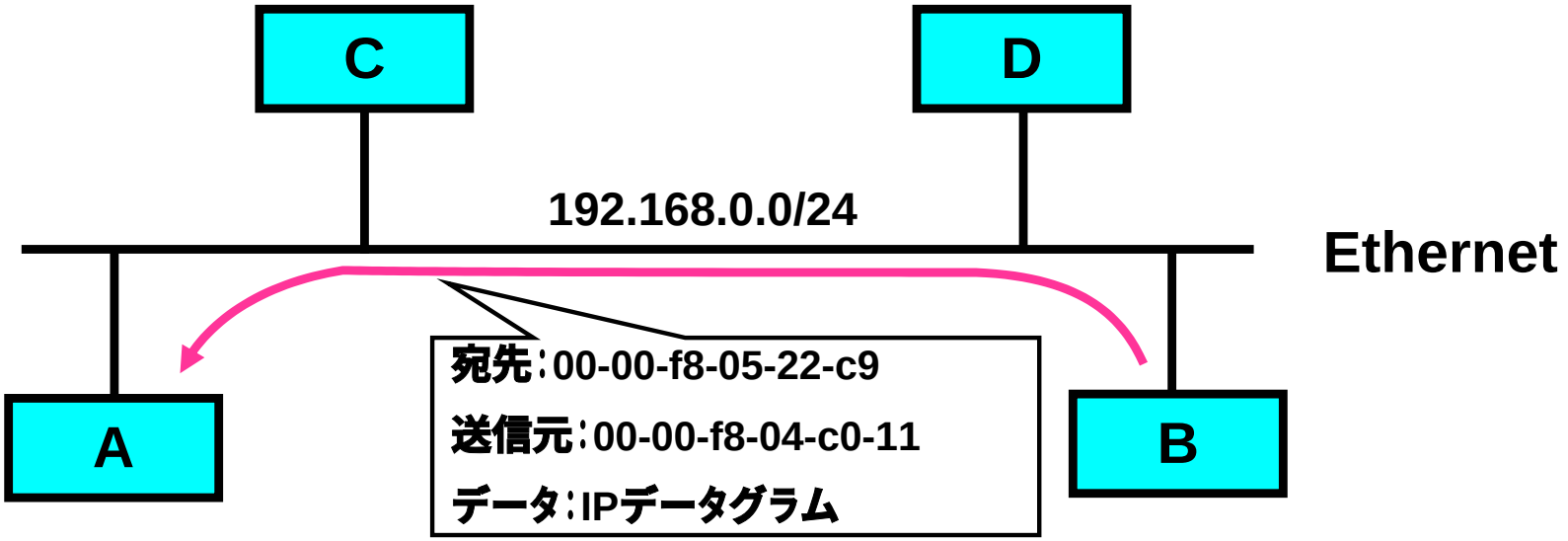
Host AのIP/MACアドレス対応表

Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

Host BのIP/MACアドレス対応表

- Aは自分のMACアドレスをBに返答する

ARPの動作ー4



Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	不明
C	192.168.0.3	不明
D	192.168.0.4	不明

Host AのIP/MACアドレス対応表

Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

Host BのIP/MACアドレス対応表

- BはAに対してデータを送ることができるようになる

Multi Access Media(ARP)－1

- ARP (Address Resolution Protocol)
  - ARPとはIPアドレスとMACアドレスを対応させるためのプロトコル (IP以外のプロトコルでも利用されますが、IPに限って説明します)
- IP/MACアドレス表
  - IP/MACアドレス対応表のことを「ARPテーブル」「ARPキャッシュテーブル」「ARPキャッシュ」などと呼ばれている
- ARPキャッシュ
  - ARPテーブルに登録されたIP/MACアドレスは一定時間保持 (キャッシュ) される
  - ARPテーブルにIP/MACアドレスが存在するときはARPによるbroadcastは行われずに、ARPテーブルにしたがって通信が行われる。
  - 一定時間後、IP/MACアドレスはARPテーブルから削除され、その後通信が行われた場合には再びARPを実施する
  - キャッシュすることで、ARPによるデータリンク層のbroadcastを抑制している

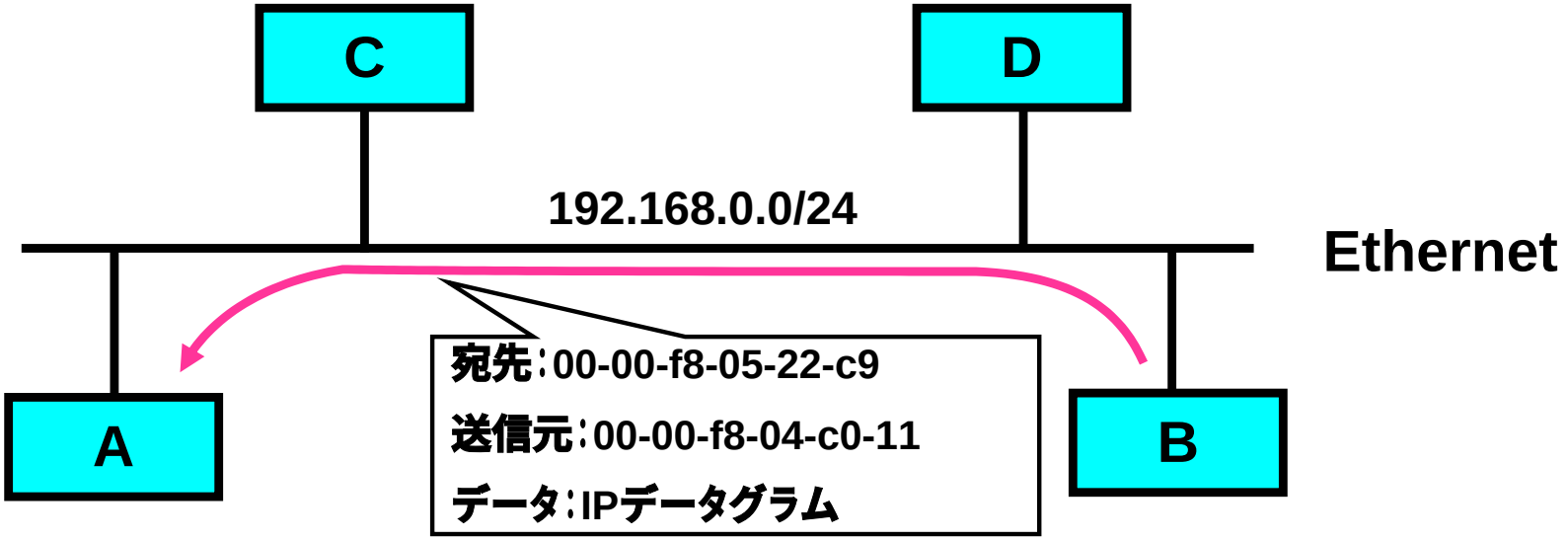
Multi Access Media(ARP)－2

- ARPキャッシュのクリア
  - 機器の交換などでIP/MACアドレス対応に変化がある場合はARPキャッシュをクリアを行う必要がある場合がある
    - arp –d (ホストなど)
    - clear arp (ルータなど)
  - 最近のネットワーク機器やOSは機器交換後に明示的にARPキャッシュをクリアしなくても高速にARPキャッシュ反映されるような実装が増えている

Multi Access Media(ARP)－3

- ARPのメリット
  - 他の機器のIP/MACアドレス対応表を設定する必要が無い
  - 機器交換を行ってもARPキャッシュがクリアされれば自動的に反映される
- ARPの運用上の注意点
  - 機器交換の際にARPキャッシュをクリアしないとすぐに通信できないことがある
  - broadcastが利用されるため大規模なレイヤ2ネットワークでは帯域を圧迫する
  - Globalセグメントで多くの利用されていないアドレスが存在すると、インターネットから未使用アドレスに対するアクセスによりLANが輻輳することがある
    - インターネット上のウイルスに感染したホストなどからのポートスキャンにより発生する(NIMDAなど)
    - 未使用アドレスの個数×リトライ回数のbroadcastが発生する
      →詳しくは後述

固定IP/MACアドレス対応表の動作



Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	00-60-08-07-07-ac
D	192.168.0.4	00-a0-24-4a-7a-12

Host AのIP/MACアドレス対応表（固定）

Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	00-60-08-07-07-ac
D	192.168.0.4	00-a0-24-4a-7a-12

Host BのIP/MACアドレス対応表（固定）

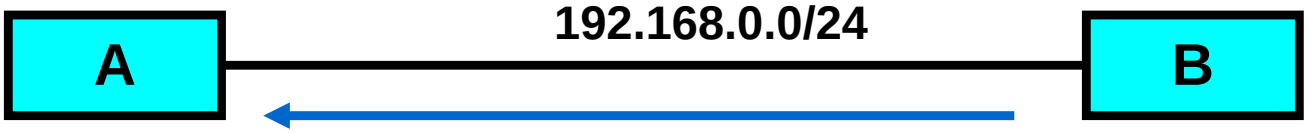
- IP/MACアドレス対応表は事前に固定的に設定されるため、BはAに対してデータを送ることができる

IIJ
Internet Initiative Jap

Multi Access Media(固定)

- ARPを用いず固定的に物理アドレスとIPアドレスを結びつける
- ARPを用いないためbroadcastが発生しない
- broadcastが利用できないため、ARPが利用できない場合に利用
- 機器交換などでIP/MACアドレス対応が変化する場合にはすべての機器の設定を変更する必要がある
- ATMではVPI/VCIを固定的に設定する

Point to Point Mediaの動作



Host	IPアドレス
A	192.168.0.1
B	192.168.0.1以外の 192.168.0.0/24

Host Aの通信先

Host	IPアドレス
A	192.168.0.2以外の 192.168.0.0/24
B	192.168.0.2

Host Bの通信先

- Point to Point Mediaに属しているすべてのネットワークは相手側に送り出す（ARPや固定アドレス表は不要）
- Point to Point Mediaから来たフレームはすべて受け取る
- IP層によってはA、B間をループしてしまうこともある

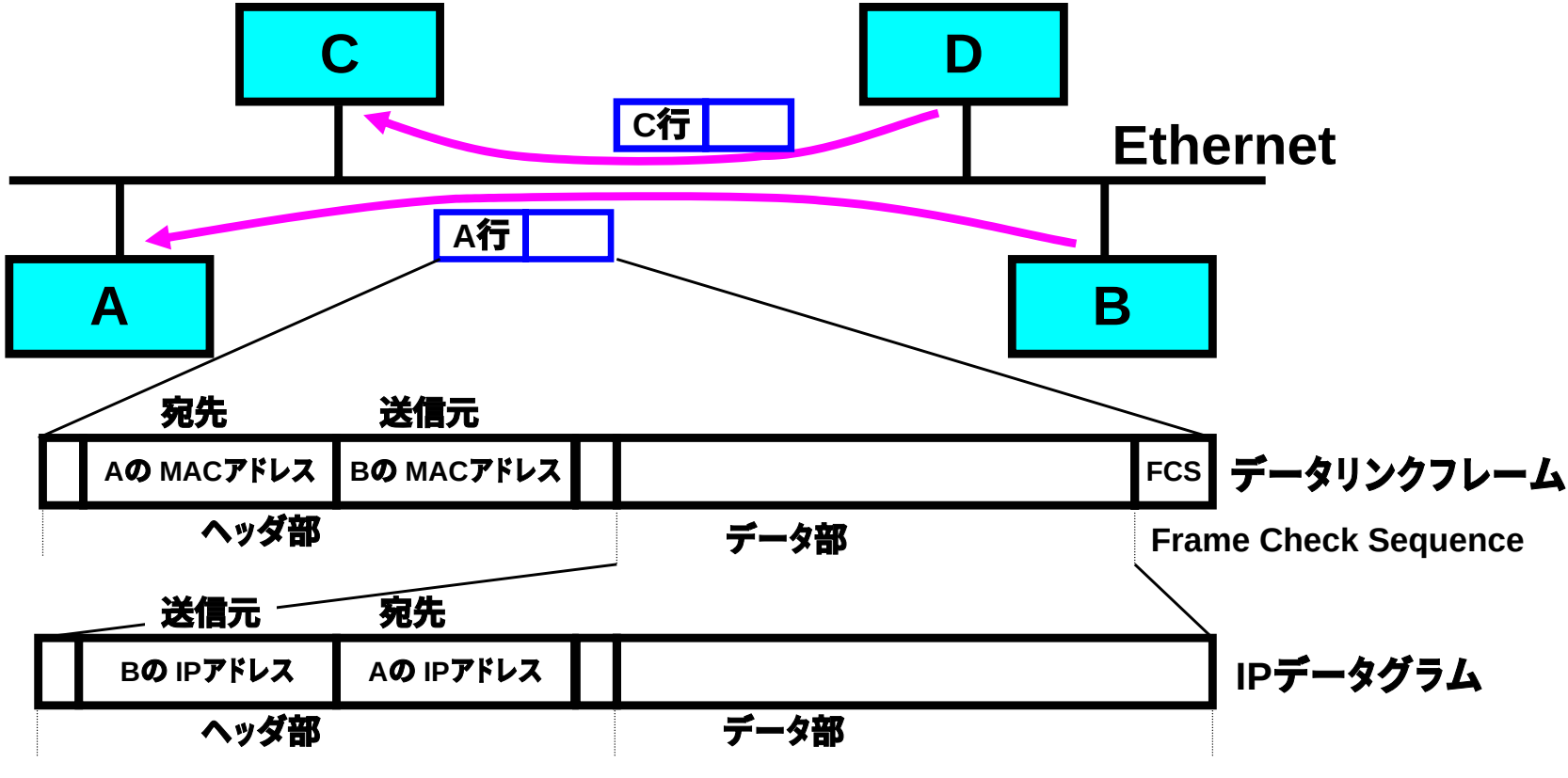
Point to Point Media-1

- 自分以外の属しているネットワークに対するすべての通信をPoint to Point Mediaに送り出す
- Point to Point Mediaから来たフレームはすべて受け取る
- 受け取ったフレームはIP層で評価される
- IP層の評価によってはPoint to Point Mediaでループすることもある
- すべてのフレームを選択せずに送り出し、受け取るためMACアドレス、broadcastは不要

Point to Point Media-2

- ATM専用線も設定によりPoint to Point Mediaとして利用することが可能
- ネットワークは一般的に/30もしくはunnumberedが利用される
  - 192.168.0.0/30 (ネットワーク例)
    - 192.168.0.1 (Router 1)
    - 192.168.0.2 (Router 2)
  - unnumberedインターフェースへのルーティングはインターフェース名などが利用される
    - ip route 172.16.0.0 255.255.0.0 *Serial0/0*

Ethernetを流れるIPデータグラム



Connectedなネットワーク

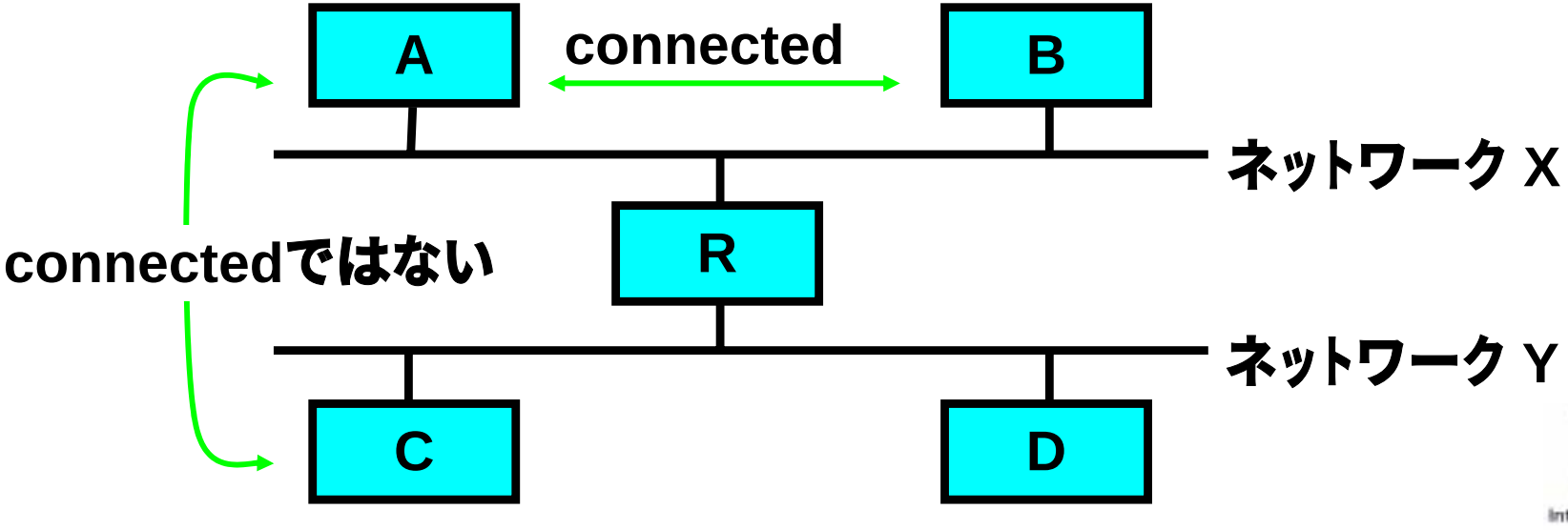
- A、Bは直接同じネットワークに接続している
  - MACアドレス、IPアドレスの対応表を ARP(address resolution protocol) などにより持っている

↓

- これを「connected」な状態という

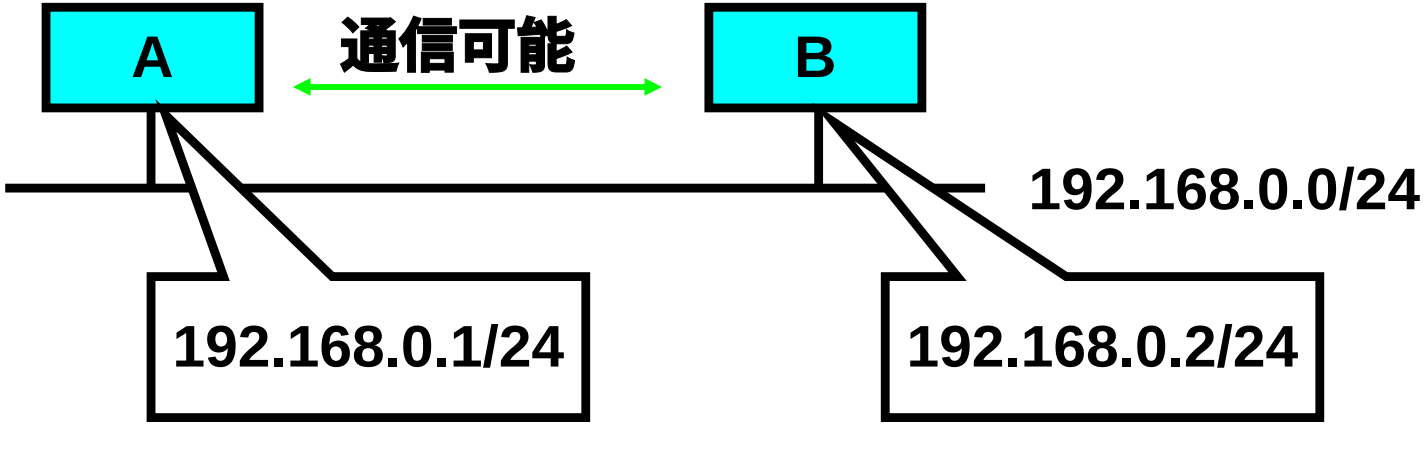
↓

- ルーティング設定が不要で、ハブなどで接続すると通信できる



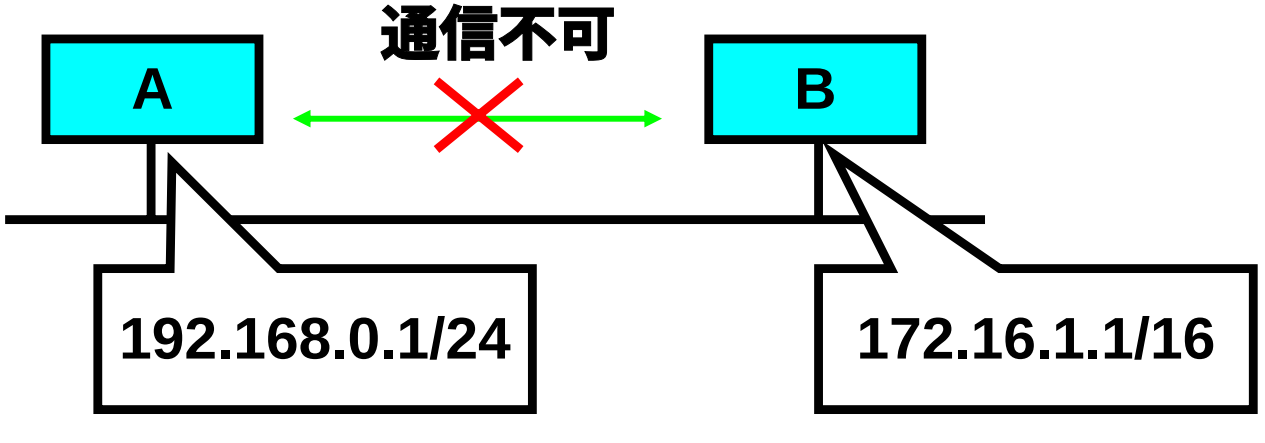
ネットワーク層から見たConnectedなネットワーク－1

- Aのアドレス
  - 192.168.0.1/24
- Aから見たConnectedなアドレス空間
  - 192.168.0.0 ～ 192.168.0.255
- Bに192.168.0.2 ～ 192.168.0.254のアドレスを付ける
  - Bに192.168.0.2を付ける
  - A-B間の通信が可能



ネットワーク層から見たConnectedなネットワーク－2

- Aのアドレス
  - 192.168.0.1/24
- Aから見たConnectedなアドレス空間
  - 192.168.0.0 ～ 192.168.0.255
- Bに192.168.0.2 ～ 192.168.0.254以外のアドレスを付ける
  - A-B間の通信ができない



Connectedではないネットワーク－1

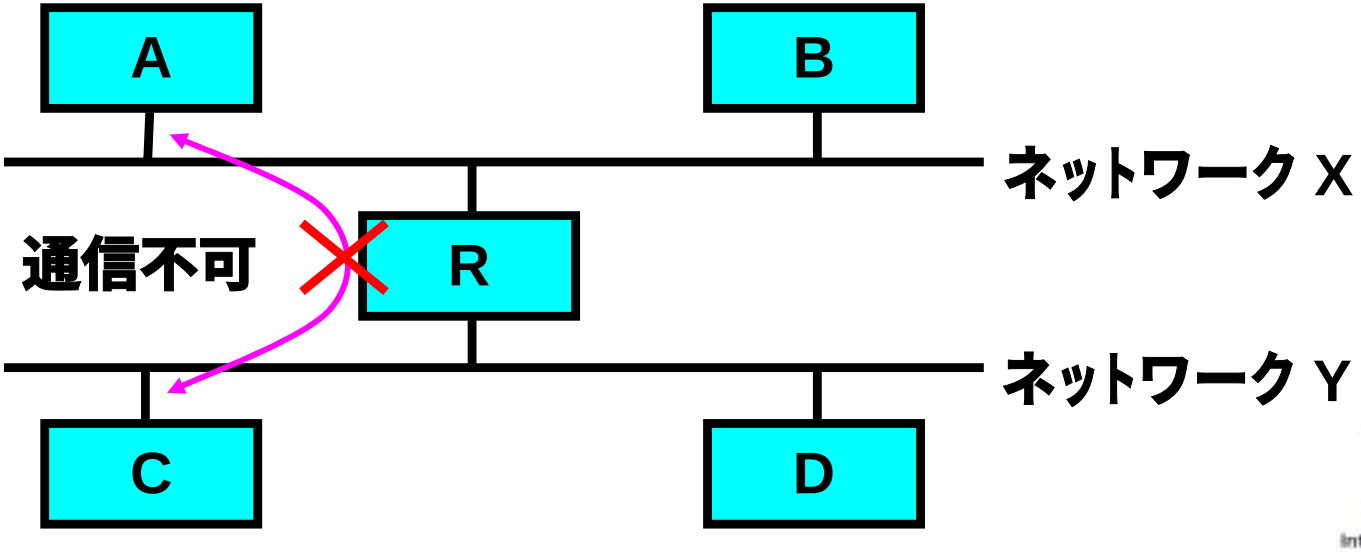
- A、Cはそれぞれ異なるネットワークに接続しているため connectedではない
- ルーティング*設定なし*ではA、C間の通信はできない

Aのルーティングテーブル

destination	Next Hop	到達性
X	Connected	到達可
Y	なし	到達不可

Cのルーティングテーブル

destination	Next Hop	到達性
X	なし	到達不可
Y	Connected	到達可



Connectedではないネットワーク－2

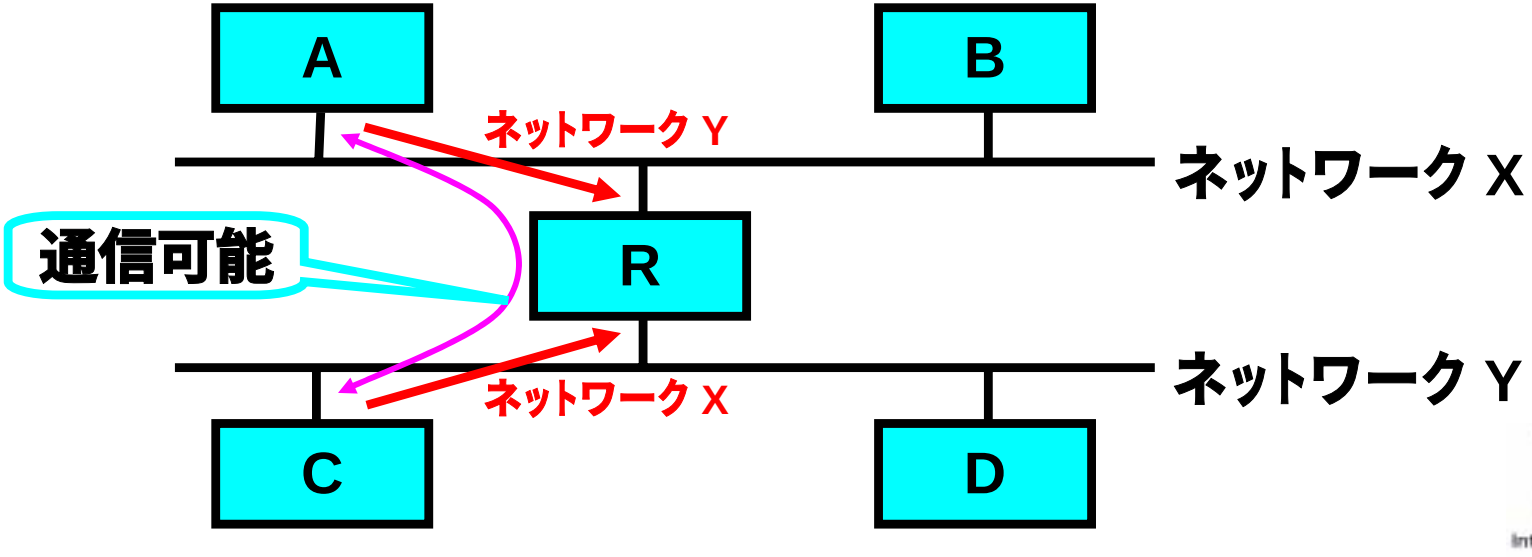
- ルーティング設定を行なう
  - A: ネットワークYを Rにルーティング
  - C: ネットワークXを Rにルーティング
- これにより、A⇔C間の相互通信が可能となる
  - Rは A,C共に connectedなため、アドレスを設定するだけで通信が可能

Aのルーティングテーブル

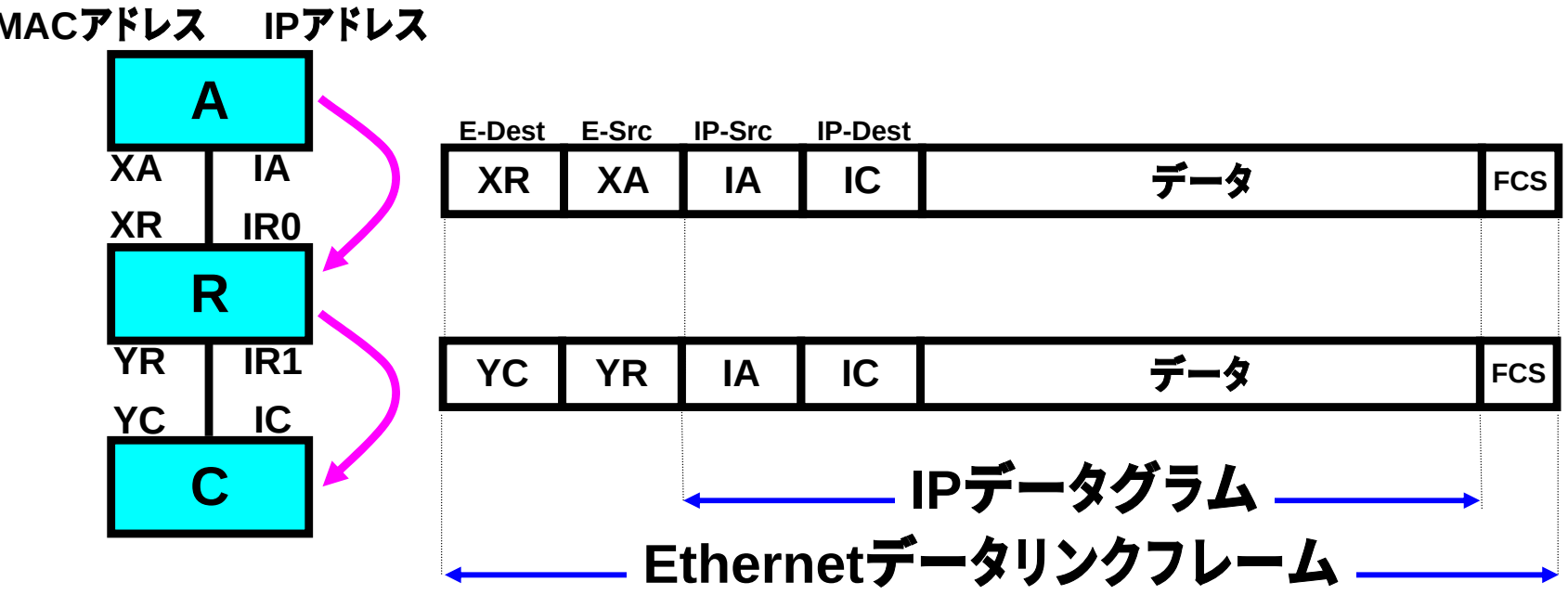
destination	Next Hop	到達性
X	Connected	到達可
Y	R	到達可

Cのルーティングテーブル

destination	Next Hop	到達性
X	R	到達可
Y	Connected	到達可



データリンクフレームの状態



- IPデータグラムの宛先、送信元は途中で変化しない
- データリンクフレームはルータを通過する毎に変化する
- 「データリンクフレームの宛先」＝「IPデータグラムの宛先」とは限らない

ネットワーク用語のまとめ

- Destination、Destination Address
  - 目的地という意味、ネットワークでは文字どおり目適地アドレス、宛先アドレスとして扱われる。Destination（デスティネーション）とそのまま使われることが多い。経路制御ではアドレスだけでなくマスク情報を含んだネットワーク情報もDestinationとして扱われる。
- NEXT HOP、NEXT HOP Address
  - 次に配送すべきアドレス。ルータやホストはDestinationがConnectedでない場合に次に配送すべきアドレス（NEXT HOP）を参照してIPパケットを送信する。IPパケットを受け取ったルータやホストはその次に配送すべきアドレス(NEXT HOP)に送信し、これらを繰り返してDestinationに到達する。
- ルーティング、ルーティング情報
  - 経路。DestinationとNEXT HOPをペアとしたもの。
- ルーティングテーブル
  - ルータやホストが持っているルーティングの一覧
- ルーティングする
  - ルータが正常にルーティングテーブルに基づいてIPパケットを送り出している状態「このルータはきちんとルーティングしている」

データリンクフレームとルーティングのまとめ

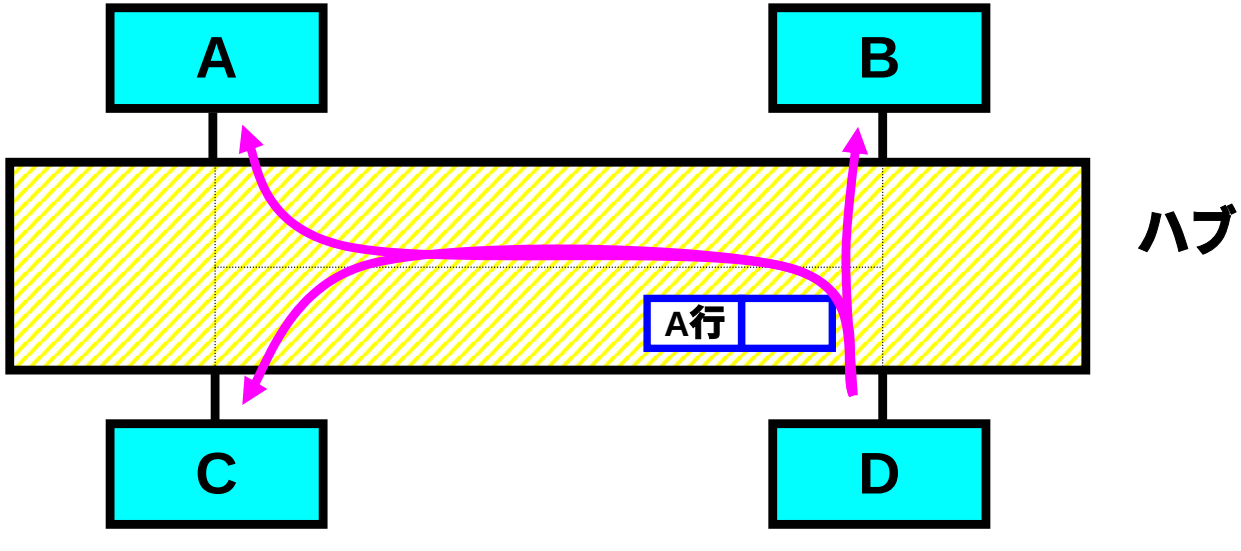
- データリンク層、ネットワーク層共にConnectedな状態であればルーティング設定をせずに通信が可能
- Connectedでないネットワーク、ホストとの通信には必ずルータの設置、ルーティング設定が必要
- IPデータグラムの宛先、送信元は途中で変化しない
- データリンクフレームはルータを通過する毎に変化する
- 「データリンクフレームの宛先」＝「IPデータグラムの宛先」とは限らない

スイッチとルータの機能の違い

- ハブとスイッチの機能の違い
- スイッチを有効に使う方法
- ルータを利用するための設定
- ネットワーク設定の自動化
- スイッチとルータの違い
- スイッチの耐障害性
- ルータの耐障害性
- Broadcast flood問題
- スイッチの耐ウイルス障害性
- ルータの耐ウイルス障害性

ハブとスイッチの違い-1

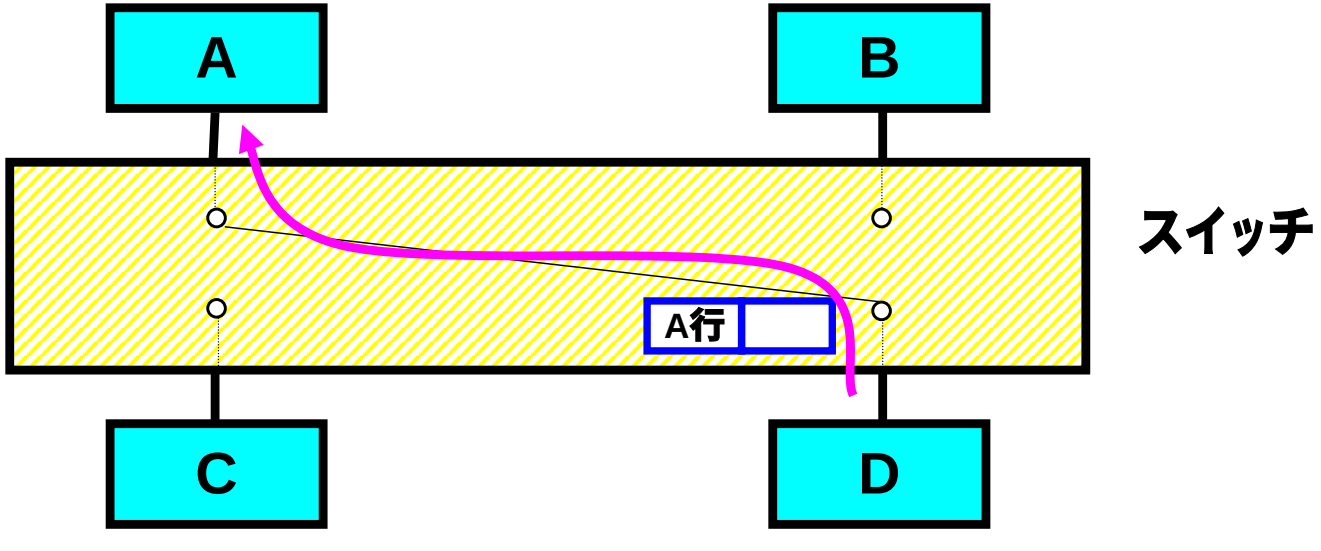
ハブで構成した場合



- ハブは全てのポートが常時接続された状態になっている
- このため異なるポート間の通信を、通信に関係の無い他のポートに伝播して、他の通信を妨げる

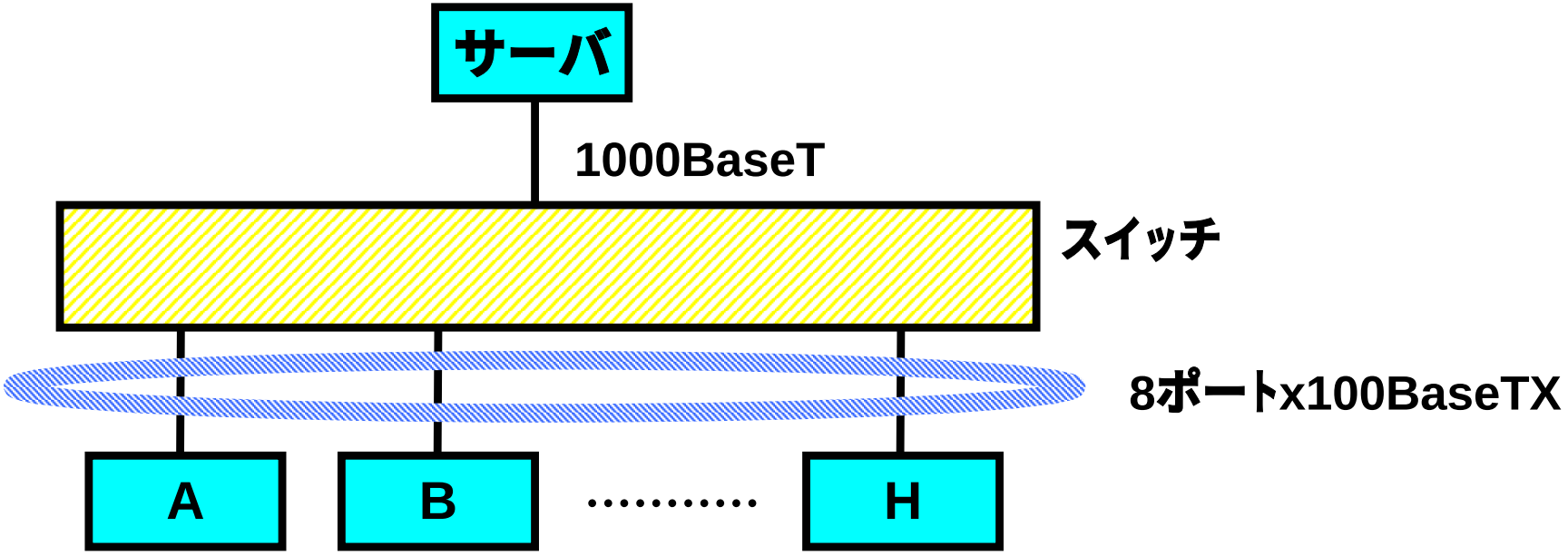
ハブと スイッチの違い-2

スイッチで構成した場合



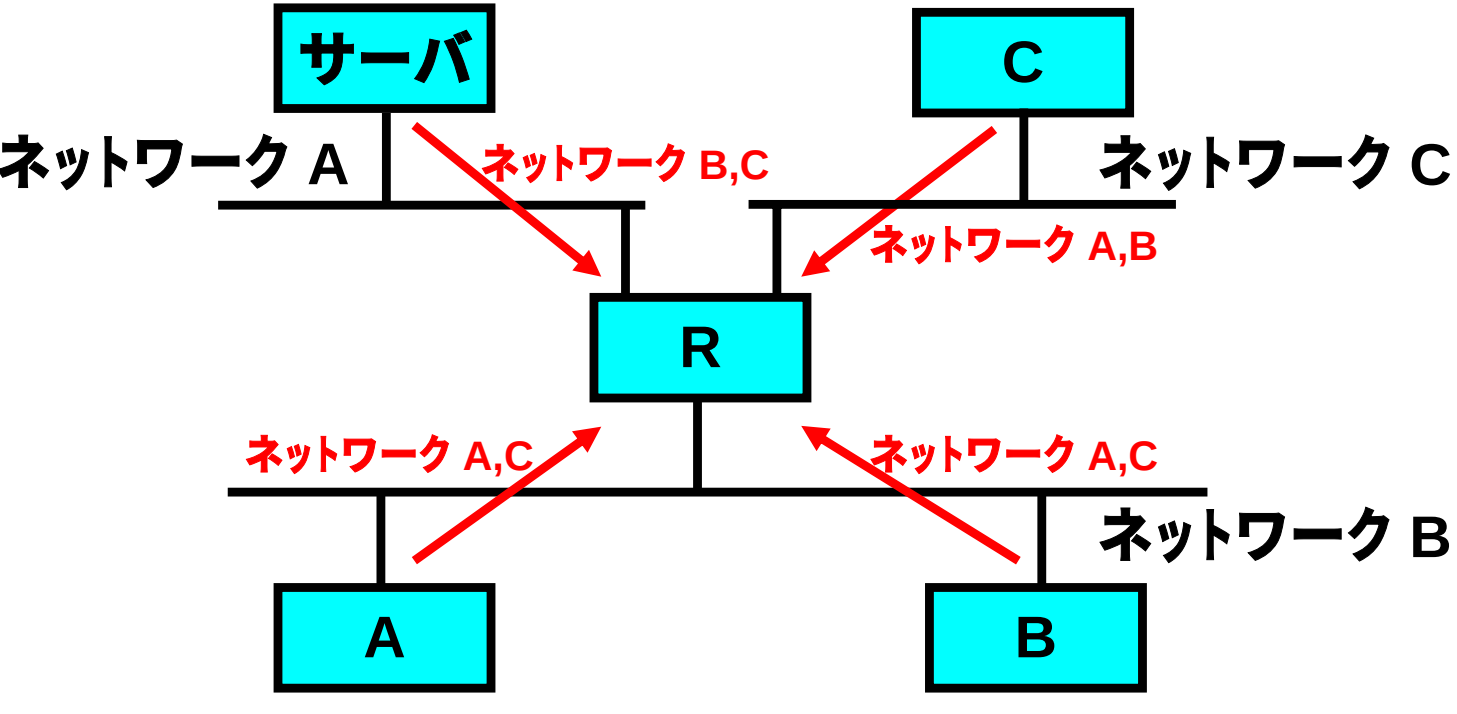
- スイッチは、ポート毎に接続されている機器のMACアドレスを学習し、通信時には必要なポート間のみで通信する

スイッチを有効に使うには



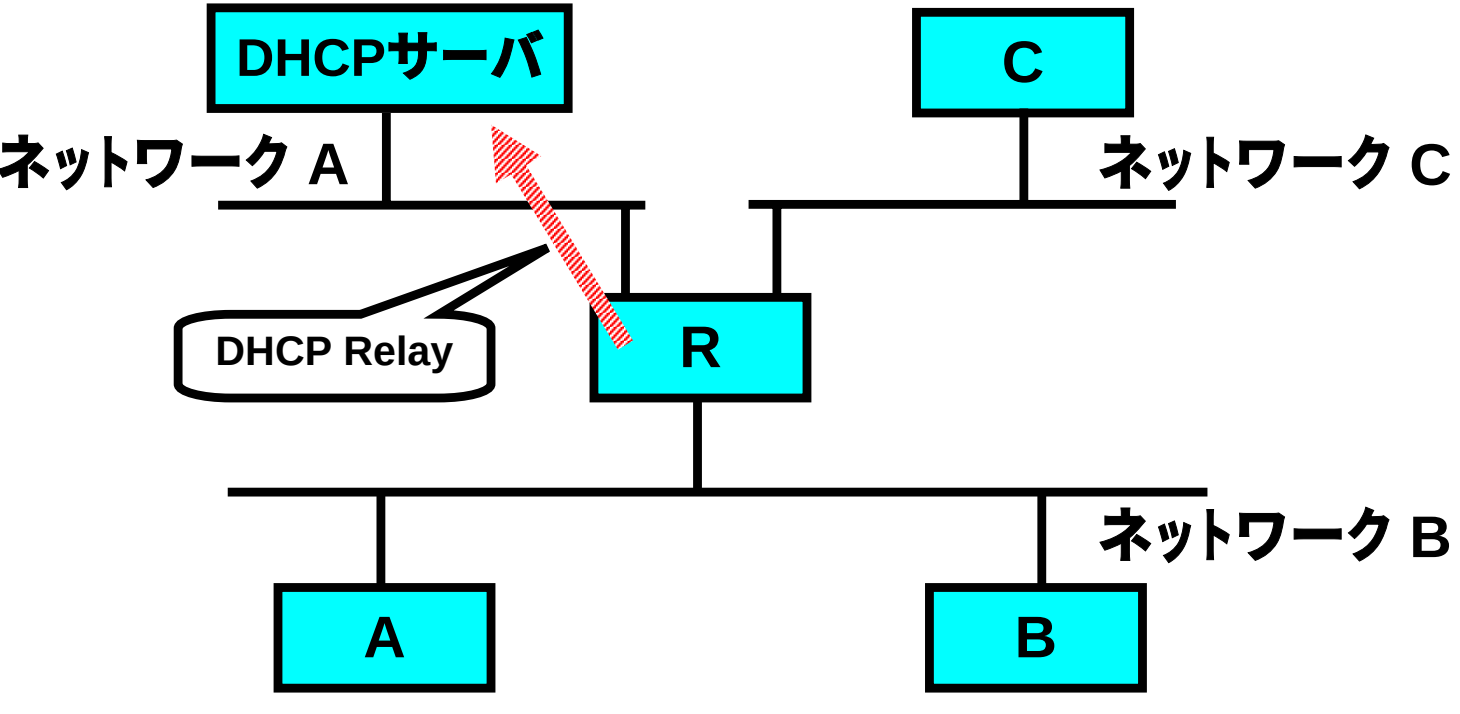
- 主にサーバ、ホスト間のトラフィックの場合に有効
- A ⇔ サーバ
⋮
H ⇔ サーバ
} それぞれ100BaseTXをフルに利用可能

ルータを利用するための設定－1



- ネットワークをサブネットに分割する
- 通信相手のネットワークのルーティングを設定する
  - DHCP,ダイナミックルーティングプロトコルなどで自動化することもできる

ルータを利用するための設定－2

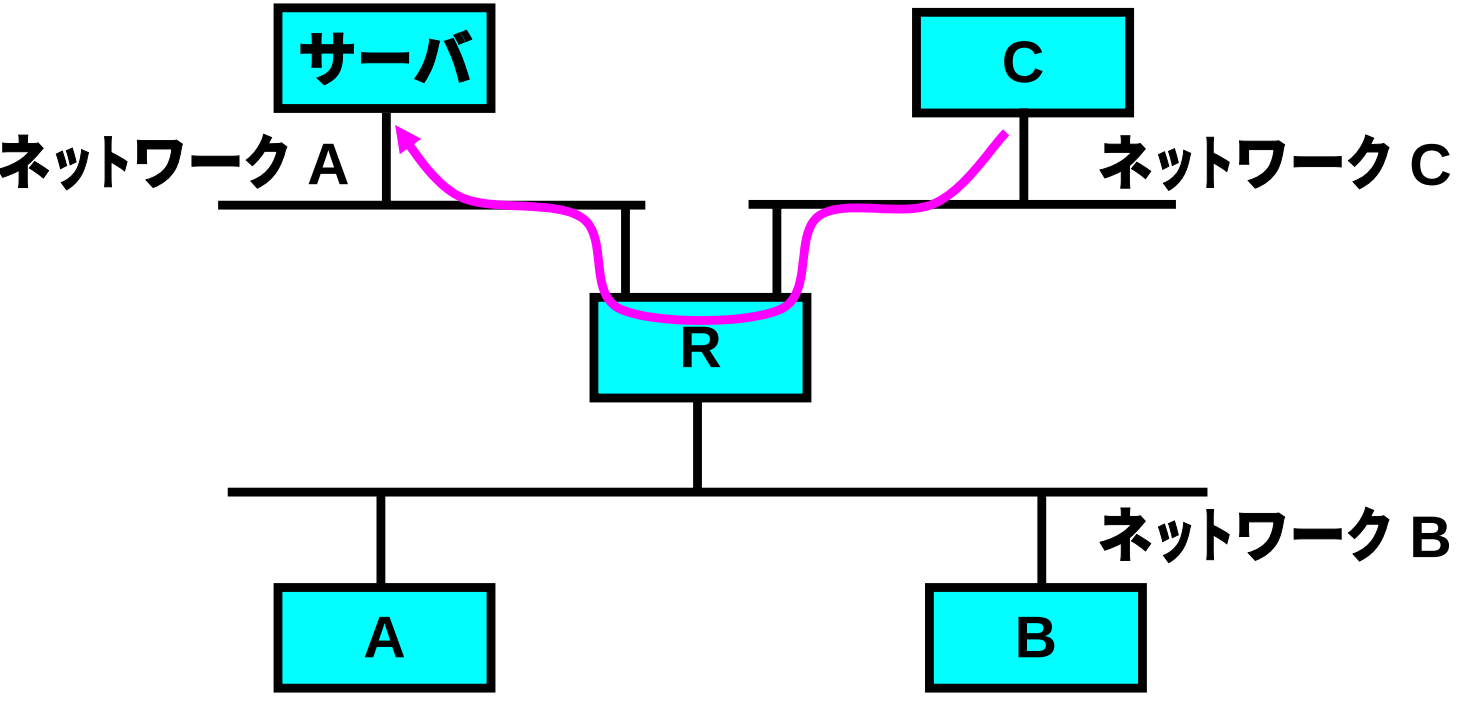


- DHCPサーバ設定
  - DHCPサーバは同一ネットワークに存在する必要がある
  - ルータのDHCP Relay設定により異なるネットワークでもDHCPを利用できる

ネットワーク設定の自動化

- DHCP (Dynamic Host Configuration Protocol)
  - アドレスの自動割り当てを行う
  - RFC2131
  - 主にクライアントで用いられる
  - Renumberを自動的に行うため、ポータビリティがある
- ダイナミックルーティングプロトコル
  - 自動的にルーティングが設定される
  - 主にルータ間で用いられる
  - RIP,RIP2,OSPFなどがある
  - 障害時に迂回路などを自動的に選択する

スイッチとルータの違い



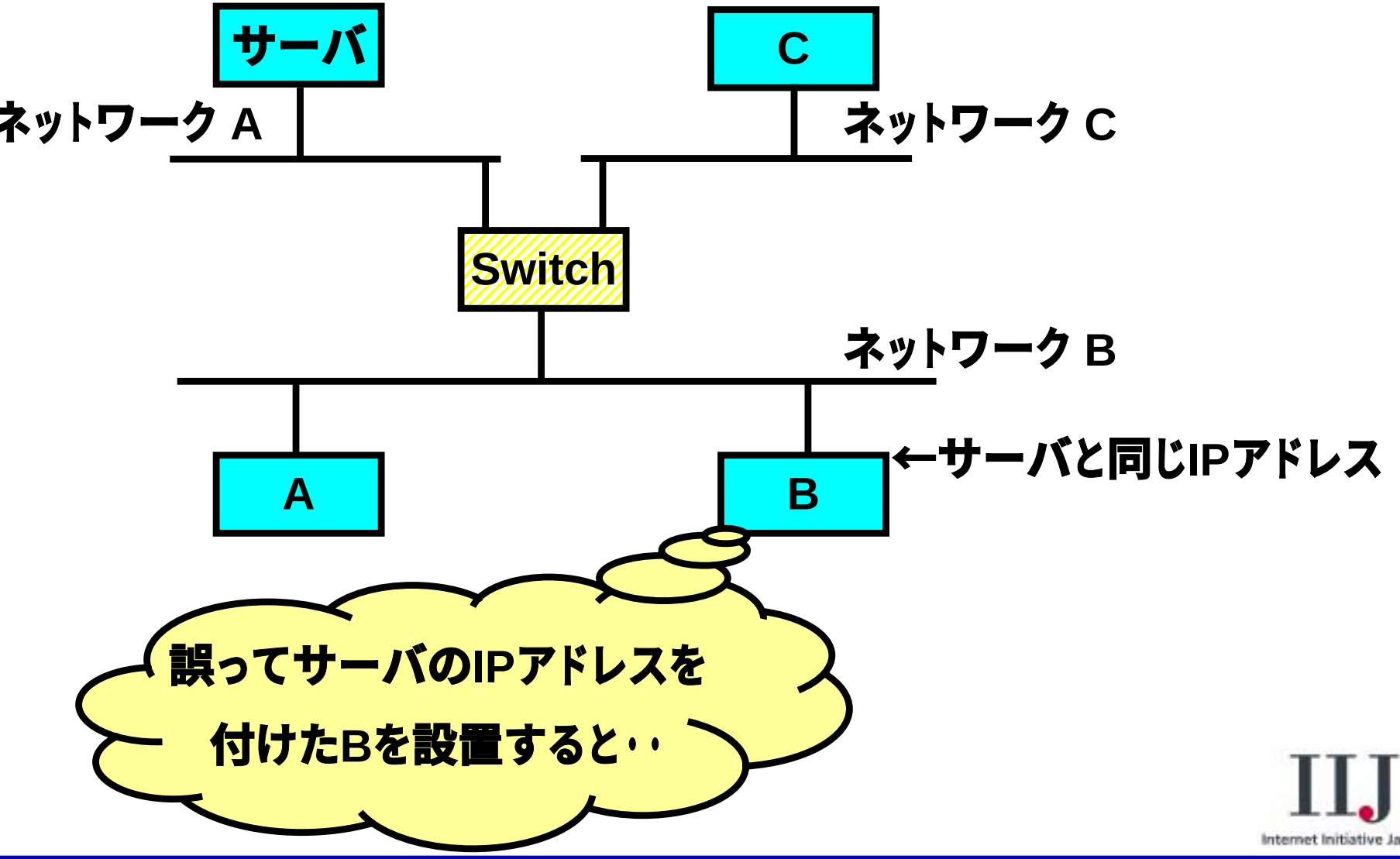
- ルータは、あるネットワーク間の通信を他の関係の無いネットワークに伝播しない

スイッチとルータの機能の違い

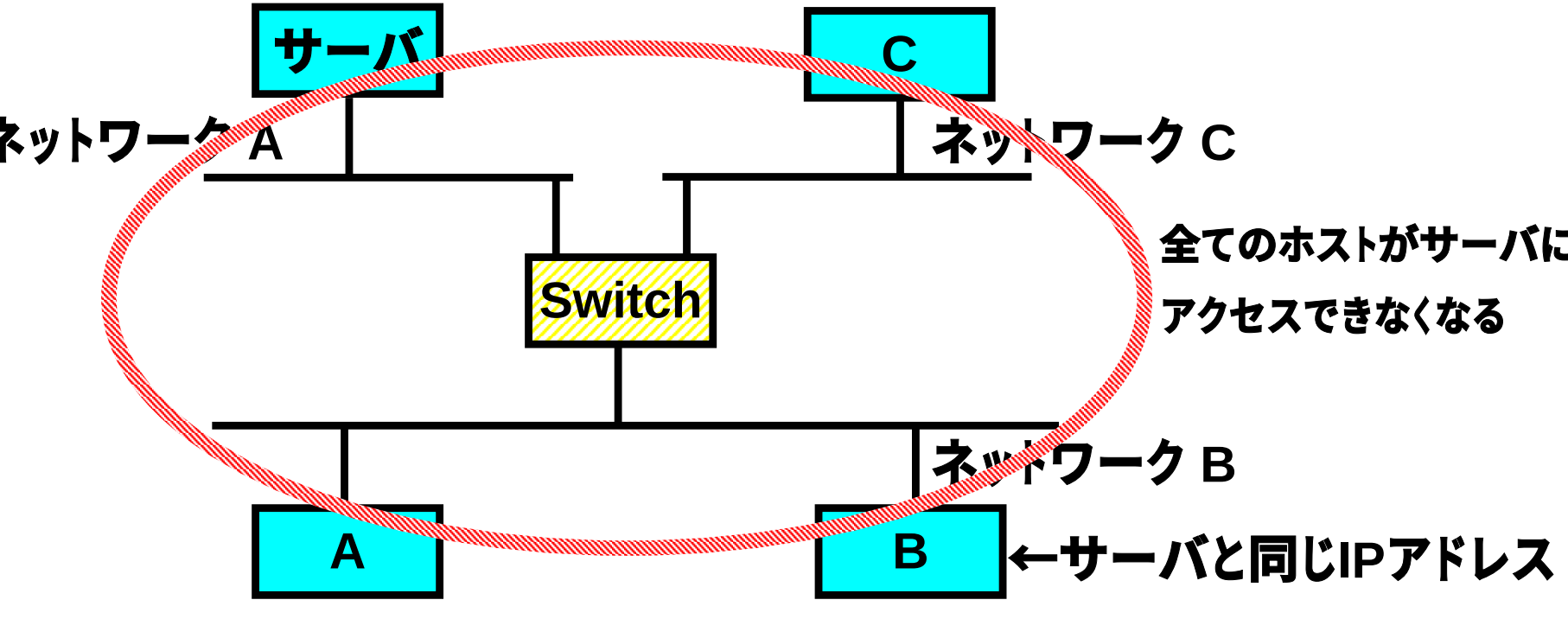
- ハブとスイッチの機能の違い
  - スイッチは異なるポートの通信を他のポートに伝播しない
- スイッチとルータの違い
  - ルータは異なるネットワークの通信を他のネットワークに伝播しない
  - スイッチとは異なり、ルーティングの設定が必要
  - サブネット分割が必要
- スイッチを有効に使うには
  - トラフィックが集中するようなポートにはスイッチを導入する

次に問題点について検討する

スイッチの耐障害性-1

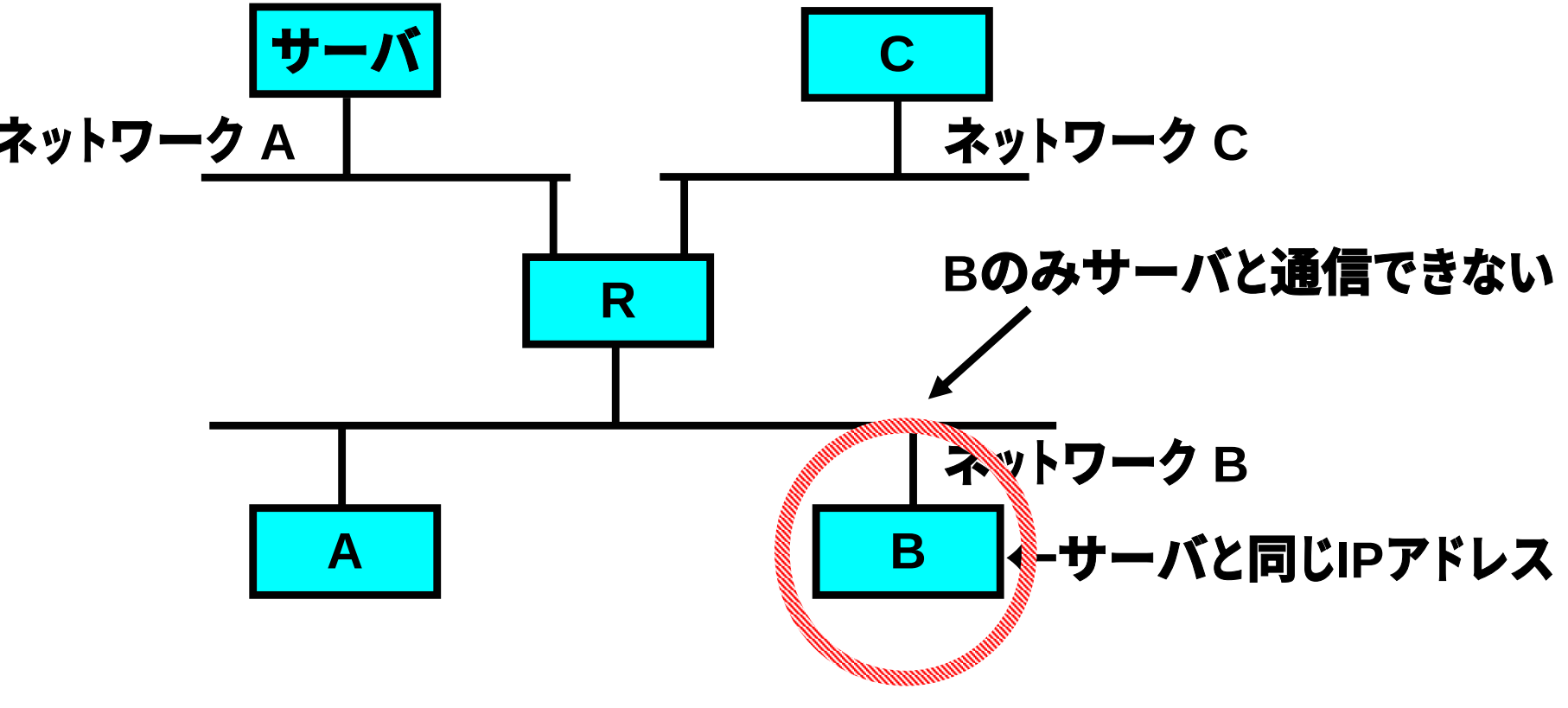

サーバ
C
ネットワーク A
ネットワーク C
Switch
ネットワーク B
A
B
←サーバと同じIPアドレス
誤ってサーバのIPアドレスを
付けたBを設置すると・・

スイッチの耐障害性-2



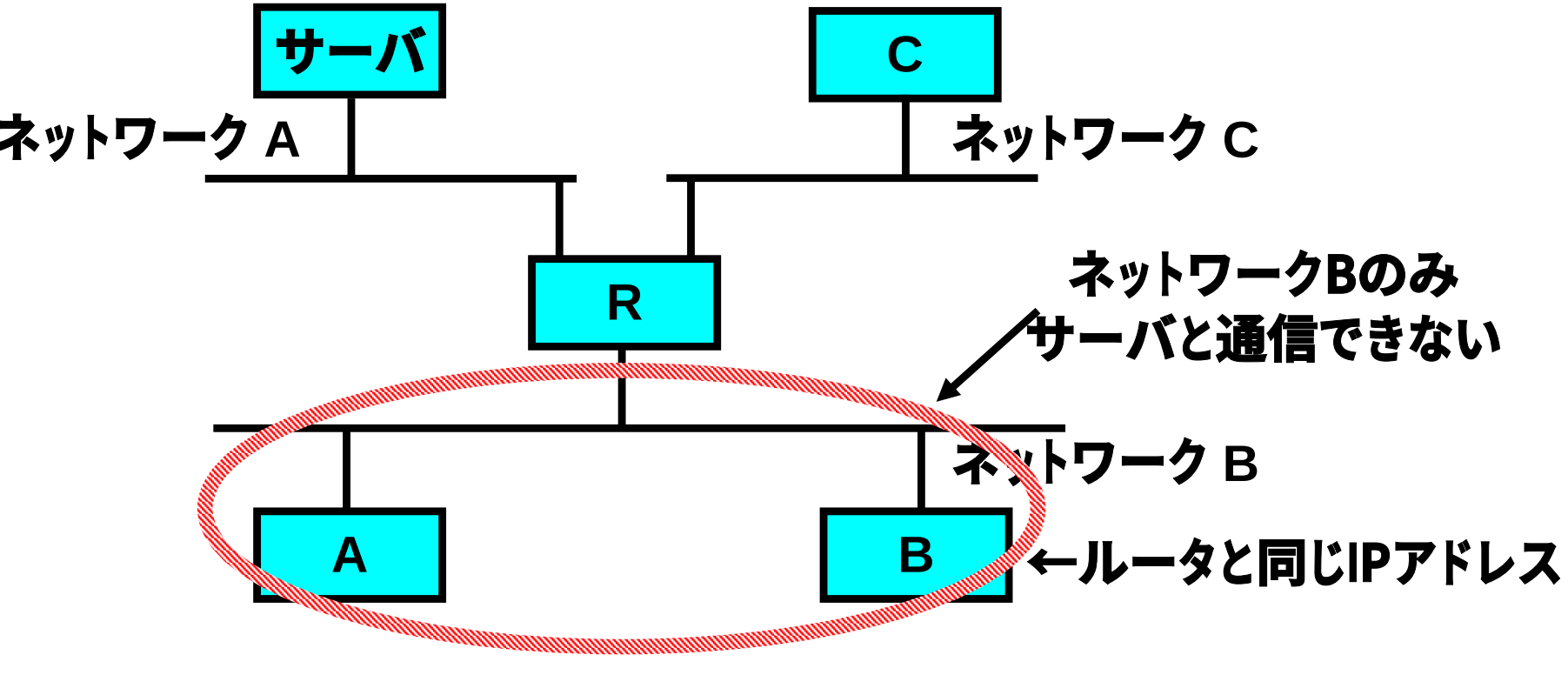
- スイッチでは、1クライアントの間違った設定の影響がネットワーク全体に及ぶ

ルータの耐障害性-1



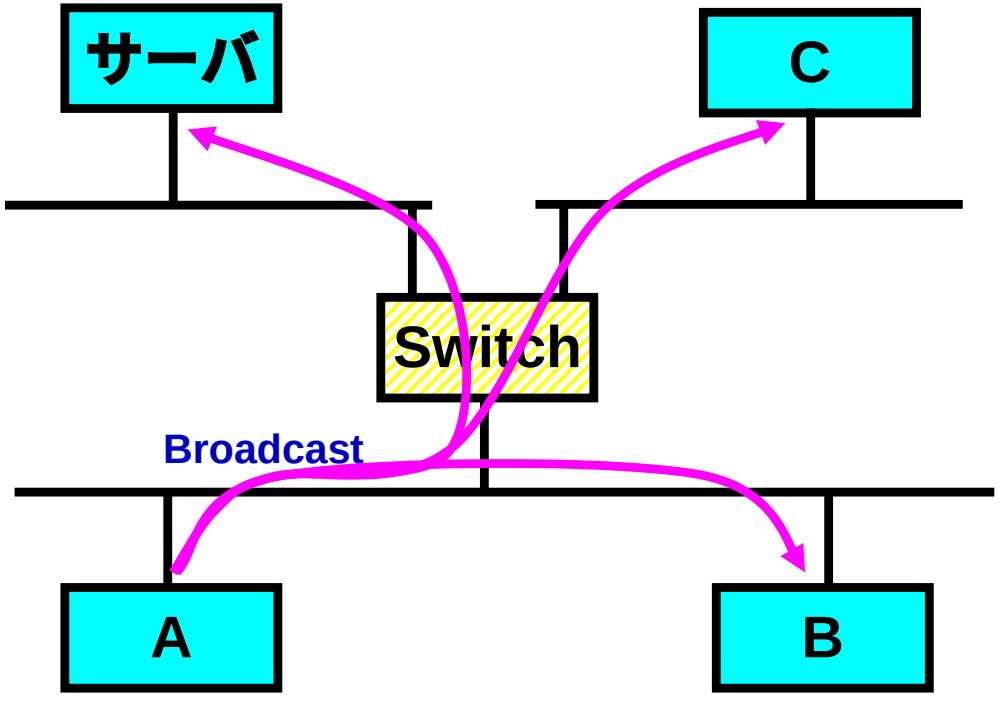
- ルータでは、1クライアントの間違った設定があったとしても、ネットワーク全体に影響を与えることはない

ルータの耐障害性-2



- 最悪の場合でも、ルータでは1クライアントの間違った設定の影響は同一セグメント内にとどまる

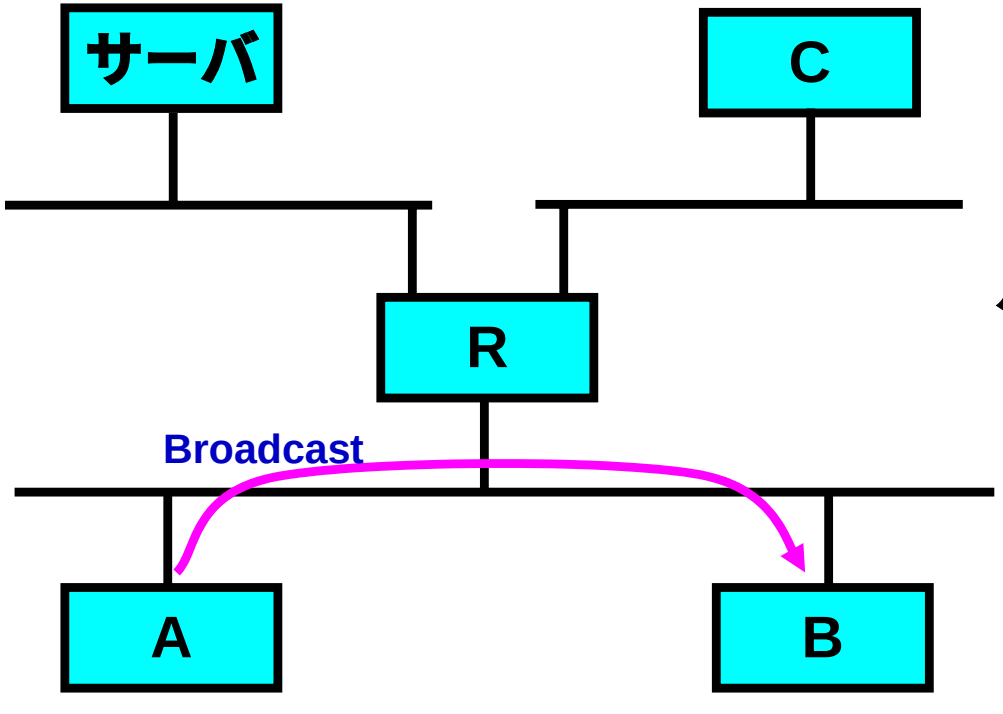
Broadcast Flood-1



Broadcastパケットはスイッチの
全てのポートに影響を与える

- ホスト数が増えると、broadcastパケットも無視できない
  トラフィックとなる
- Windows系の OSはこのようなbroadcastパケットを大量
  に発生させる傾向がある

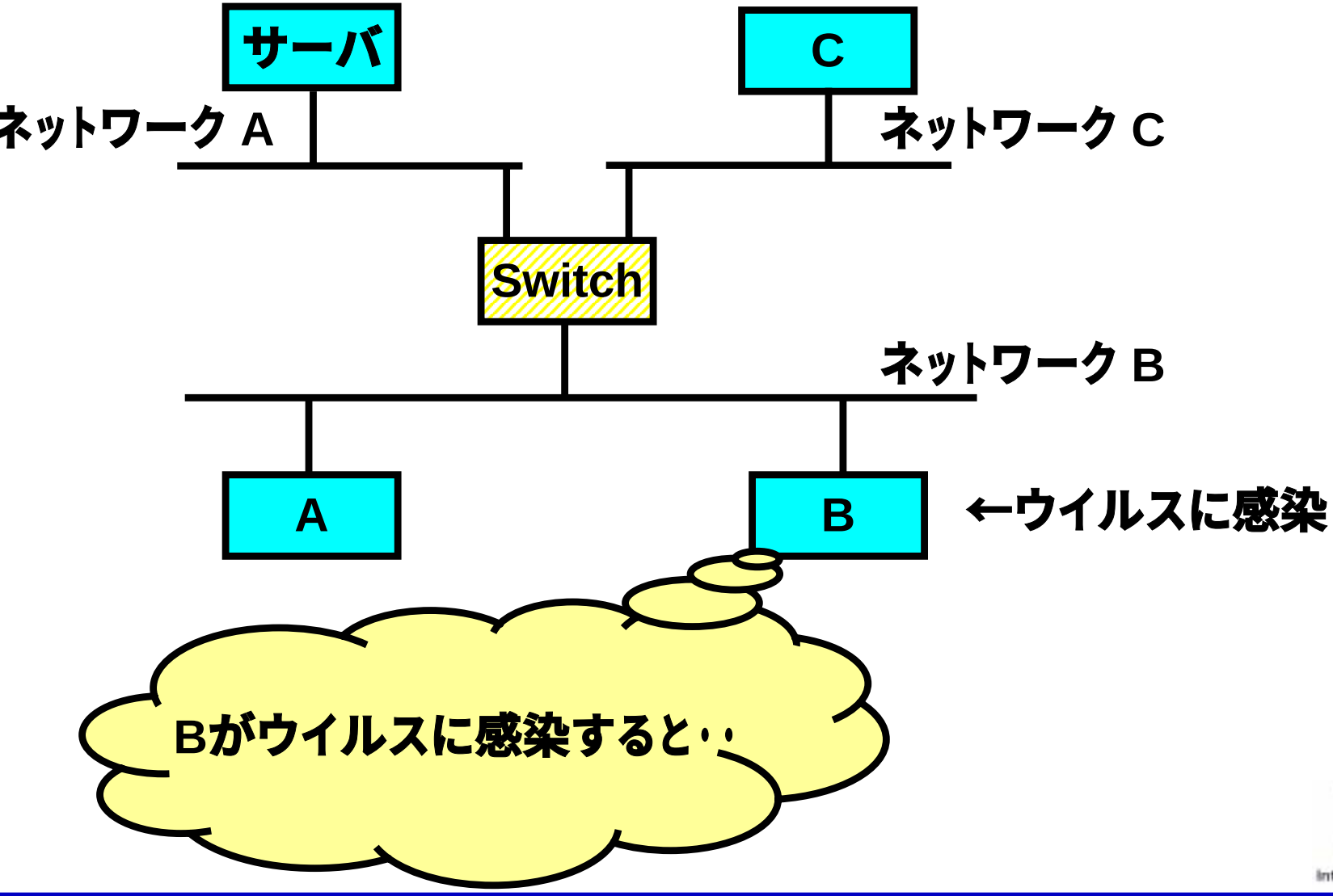
Broadcast Flood-2



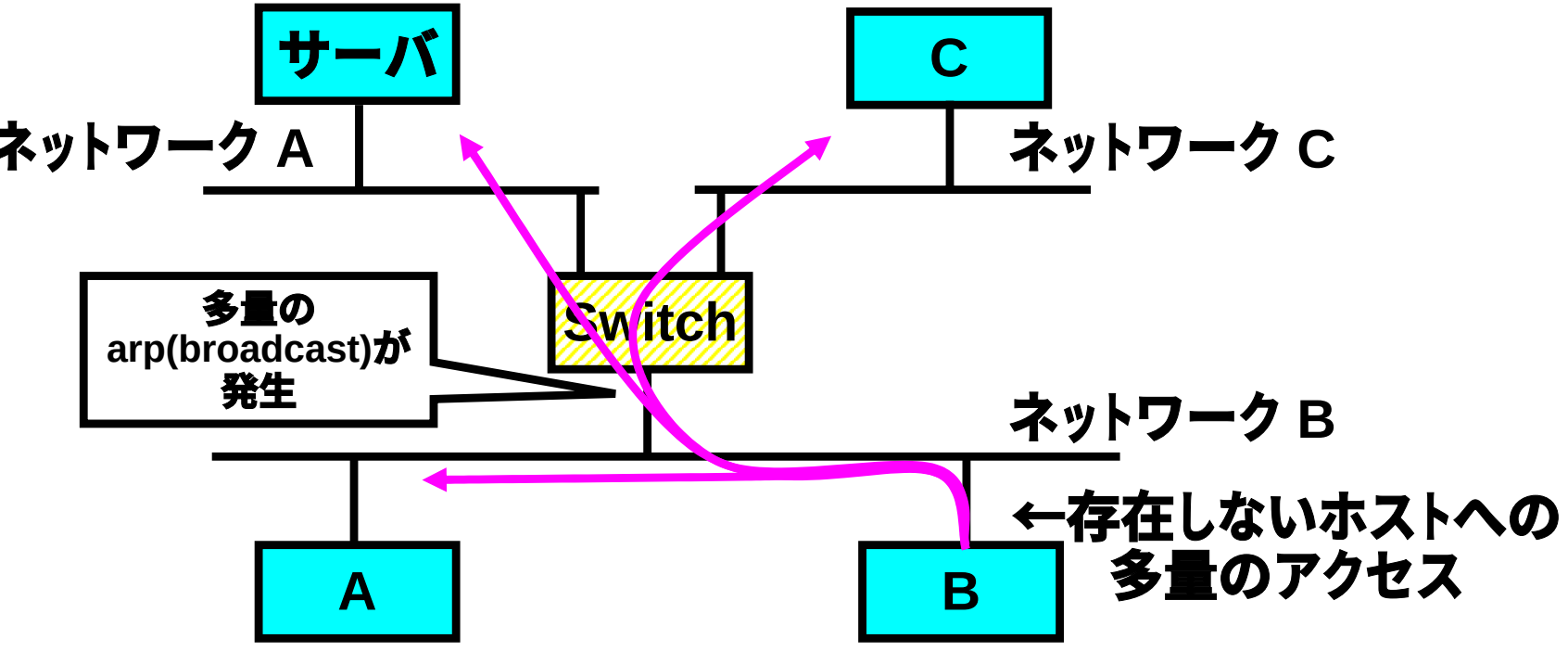
ルータは Broadcastパケットを
他のネットワークに通さない

- Broadcast floodは発生しない
- 大規模ネットワークにも対応

スイッチの耐ウイルス障害性-1

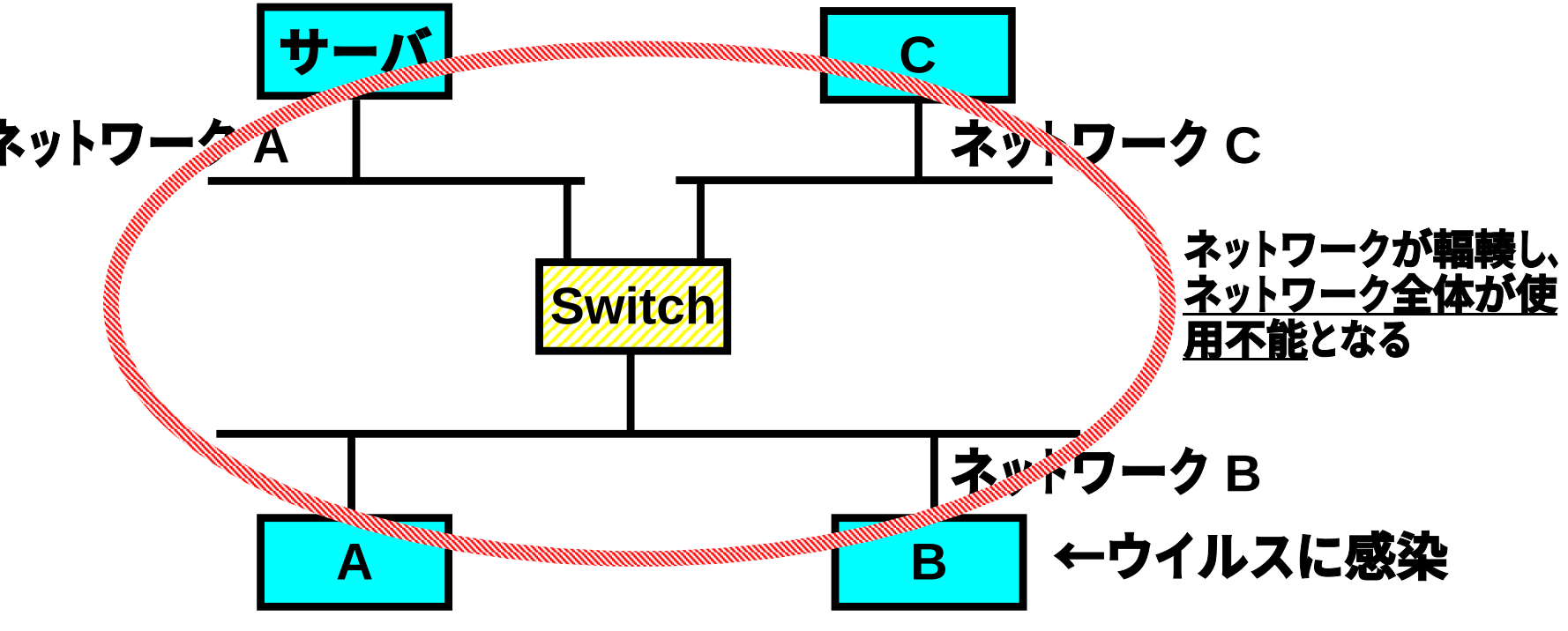


スイッチの耐ウイルス障害性- 2



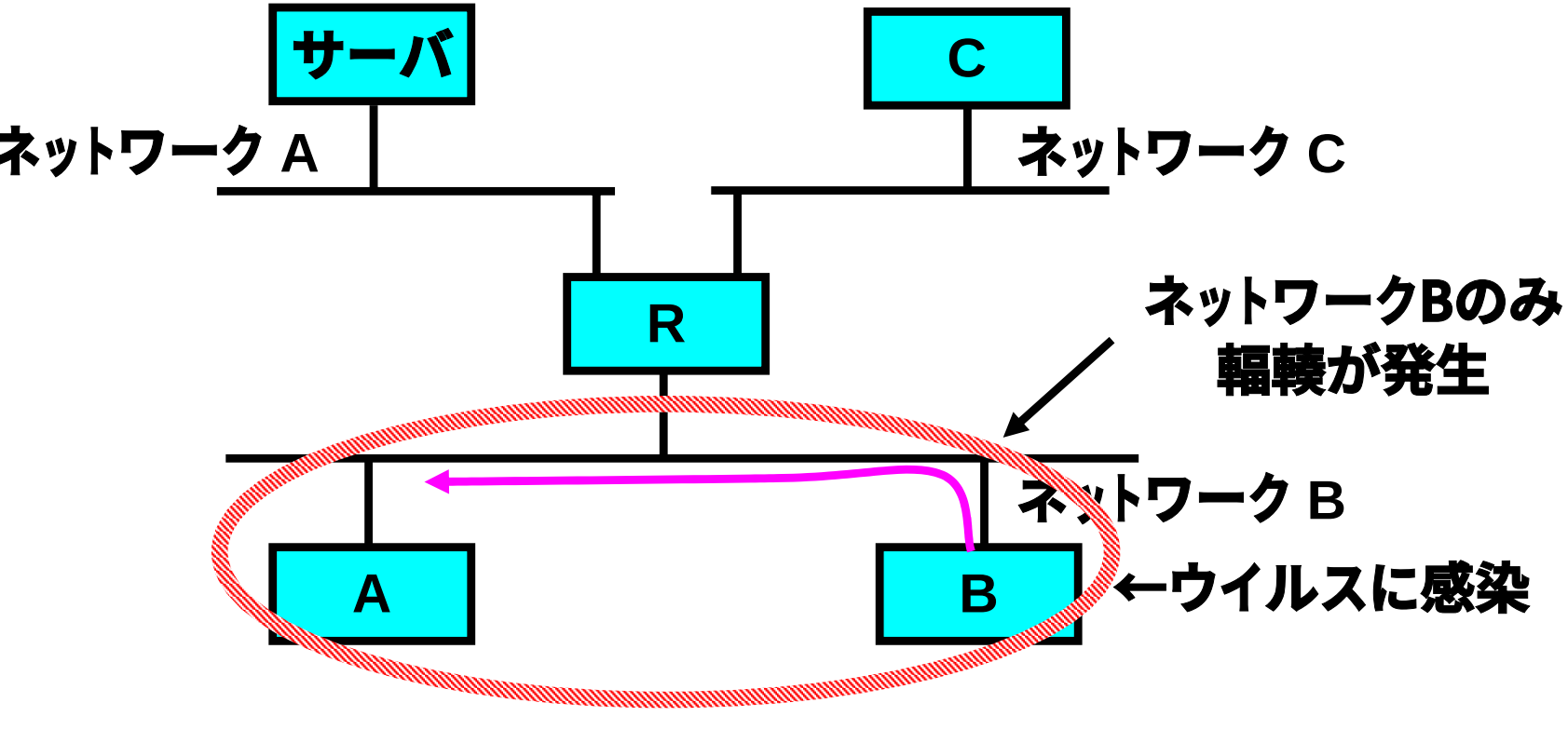
- ウイルスの感染によって存在しないホストへの多量のアクセスが発生する
- 存在しないホストへのアクセスは多量のarp(broadcast)を発生させる
- arp(broadcast)はSwitchのすべてのポートを占有する

スイッチの耐ウイルス障害性-3



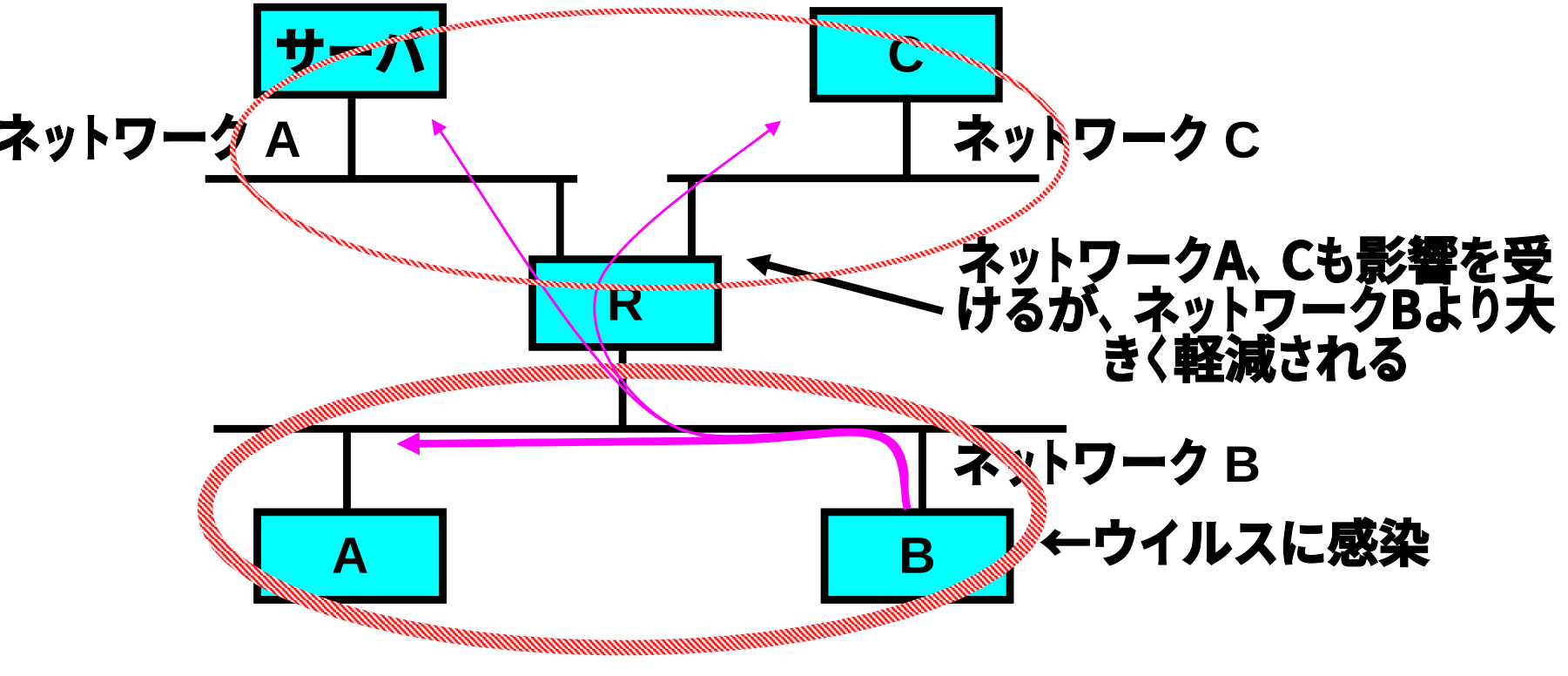
- スイッチのみの構成では、1台のウイルス感染がネットワーク全体を使用不能にする

ルータの耐ウイルス障害性-1



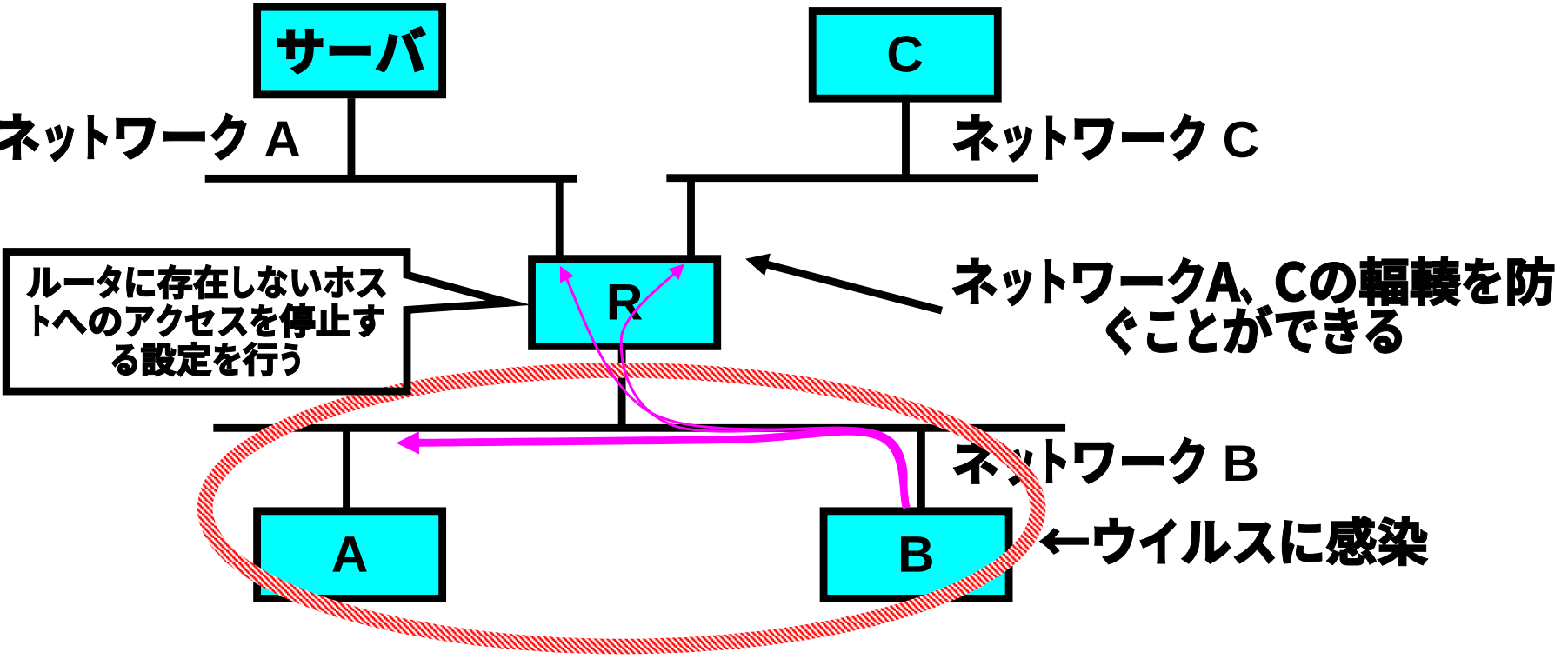
- ルータでは、ウイルス感染ホストによる多量のarp(broadcast)の影響は接続ネットワークに限られる

ルータの耐ウイルス障害性-2



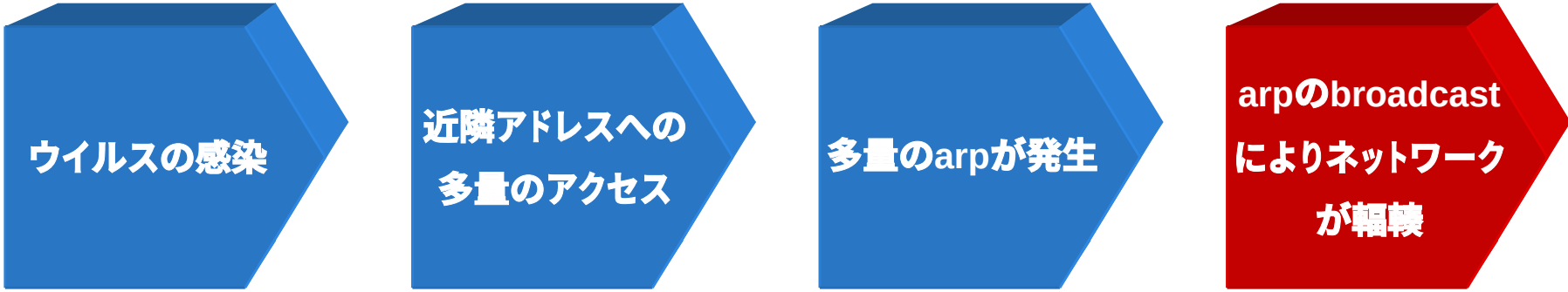
- ルータを越えるパケットも存在するが、ネットワークBに比べるとネットワークA、Cへの影響は小さくなる

ルータの耐ウイルス障害性-3



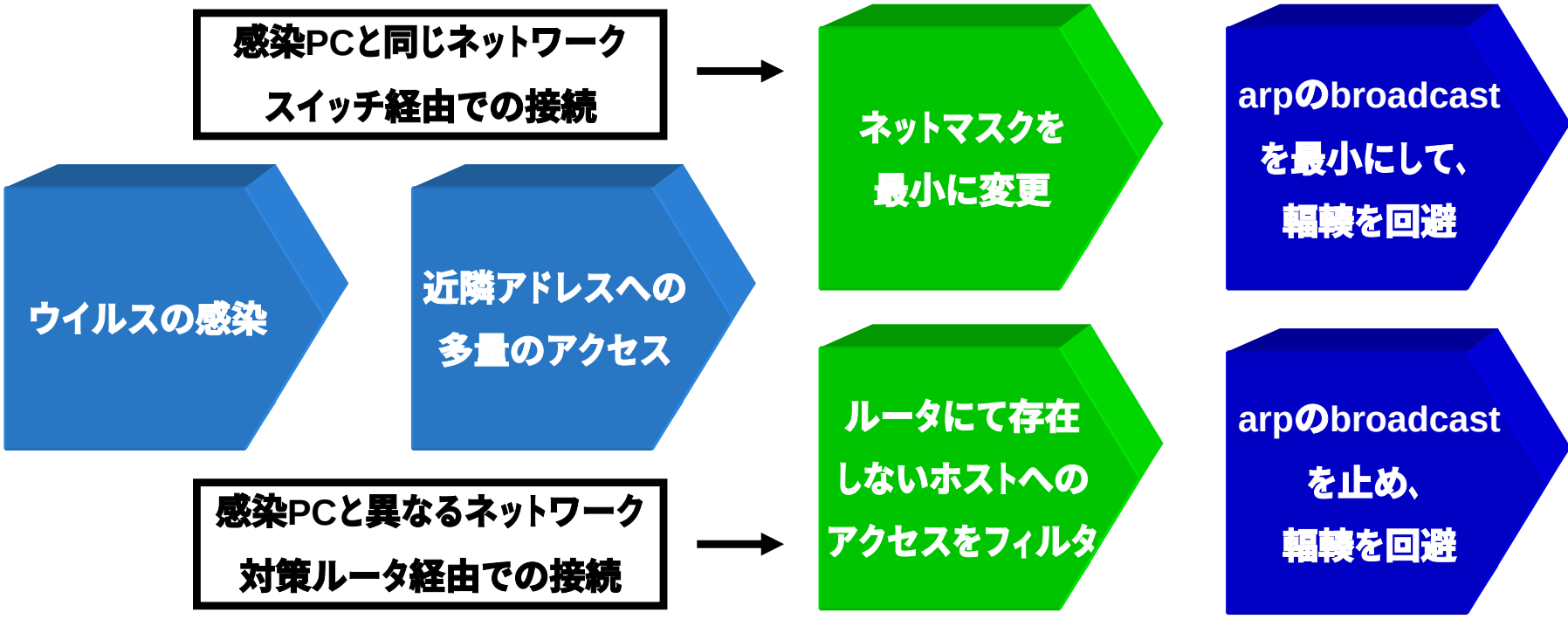
- 存在しないホストに対するフィルタをルータに行うことでネットワークA、Cに対する輻輳を防ぐことができる
- ネットマスクをできる限り小さくすることでフィルタと同様の効果を得ることができる

ウイルス感染によるネットワーク輻輳の仕組み



- ウイルスの感染
  - メール、Web、持ち込みPCなどでLAN内にウィルス感染PCが接続される
- 近隣アドレスへの多量のアクセス
  - ウイルス感染PCよりLANに割り当てられている近隣のアドレスに対して多量のアクセスを試みる
- 多量のarpが発生
  - ネットワークに存在しているホストに対してはarpは1度だけ実行される
  - 接続可能な状態にもかかわらずネットワークに存在していないホストに対しては毎回arpが実行される
- arpのbroadcastによりネットワークが輻輳
  - 未使用アドレスの個数×リトライ回数(通常8回程度)のbroadcastが発生する
  - 50台ほどのホストが接続されている状態でのbroadcastの状況は
    - ネットマスク/24利用で(256-50)×8=1,648回のbroadcast
    - ネットマスク/16利用で(65536-50)×8=523,888回のbroadcast
  - さらに複数感染した場合には感染PC数を乗じたアクセスとなる

ウイルス感染によるネットワーク輻輳の対策

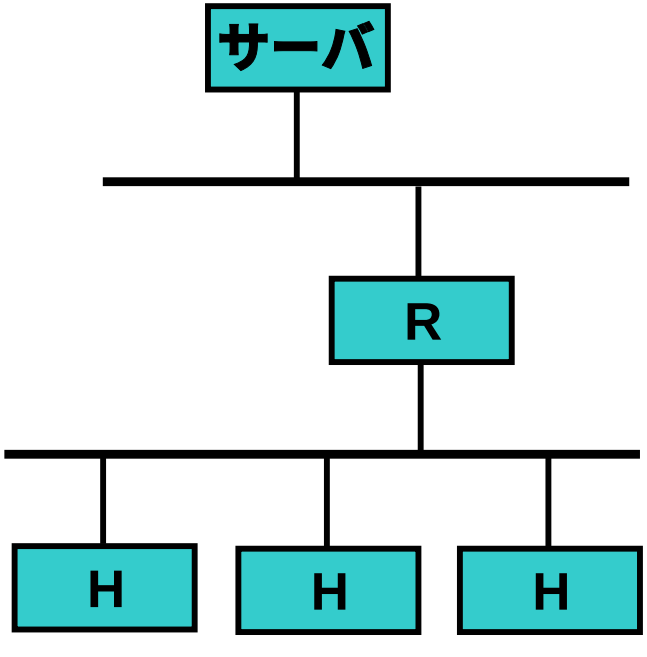


- ルータにて存在しないホストへのアクセスをフィルタ
  - ルータにて存在しないホストへのアクセスをフィルタすることでarpの発生を止め、輻輳を回避することができる
- ネットマスクを最小に変更
  - ネットマスクを最小にすることで、arpの発生を最小限にし、輻輳を回避することができる

スイッチ VS ルータ

- スイッチの利点
  - ルーティングを考慮しなくて良い
  - ハブに比べて効率的なネットワークを構築することができる
- ルータの利点
  - ダイナミックルーティングプロトコルでバックアップ構成が可能
  - Broadcast floodが発生しない
  - ウイルス感染によるネットワーク輻輳を回避できる
  - 規模が大きくなってもスケールする
  - 障害時に被害を最小限に抑えることができる
  - 障害時の切り分け作業が比較的行いやすい
- 結論
  - ルータでサブネット化を行い、トラフィックが集中するようなポートにはスイッチを導入する

ネットワーク設計-1



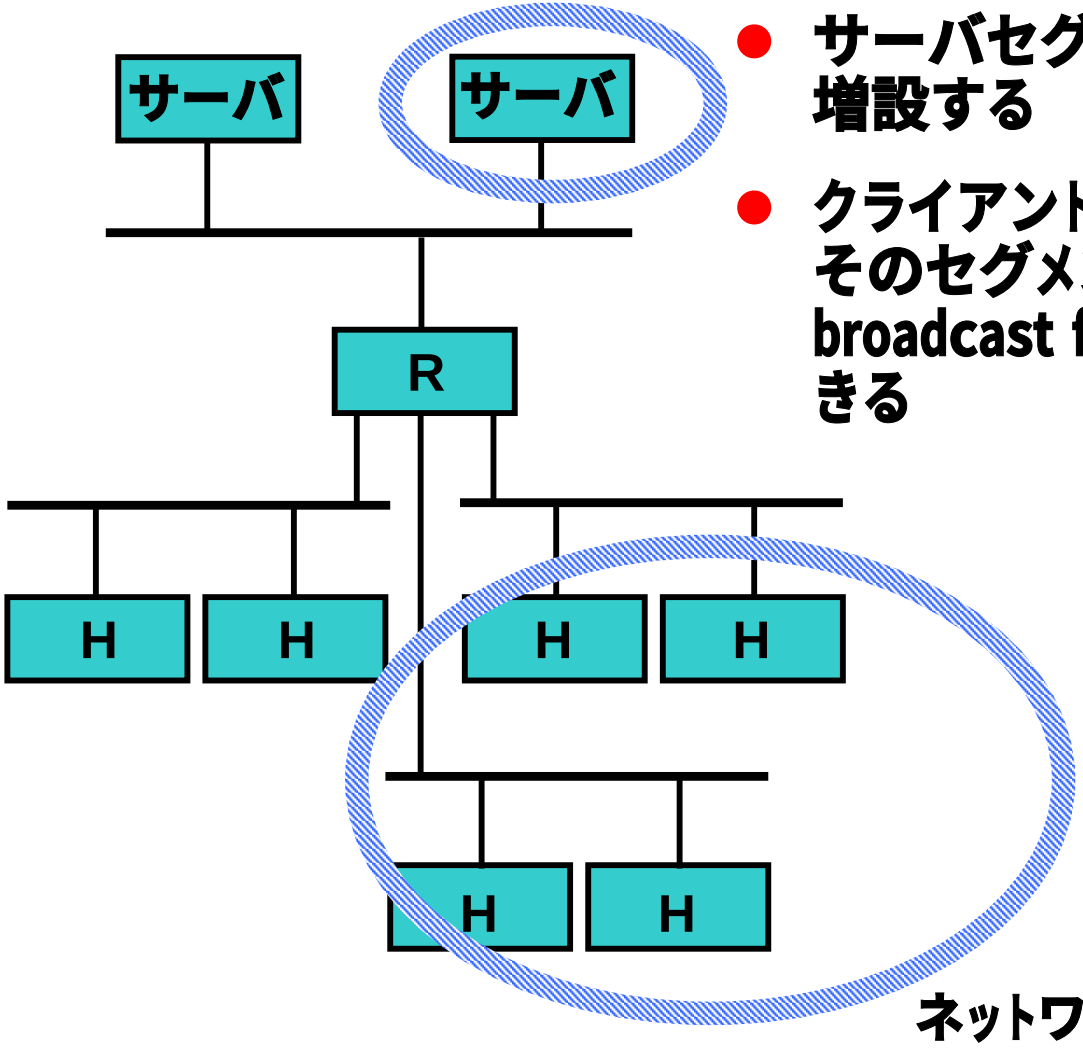
左図ネットワーク構成の特徴

- 小規模であってもサーバのセグメントを分離する

  →サーバの安全性を確保する
- クライアントはDHCPによりアドレスの割り当てとdefault経路を得る
- Broadcast floodのサーバへの影響を防ぐ

ネットワーク設計-2

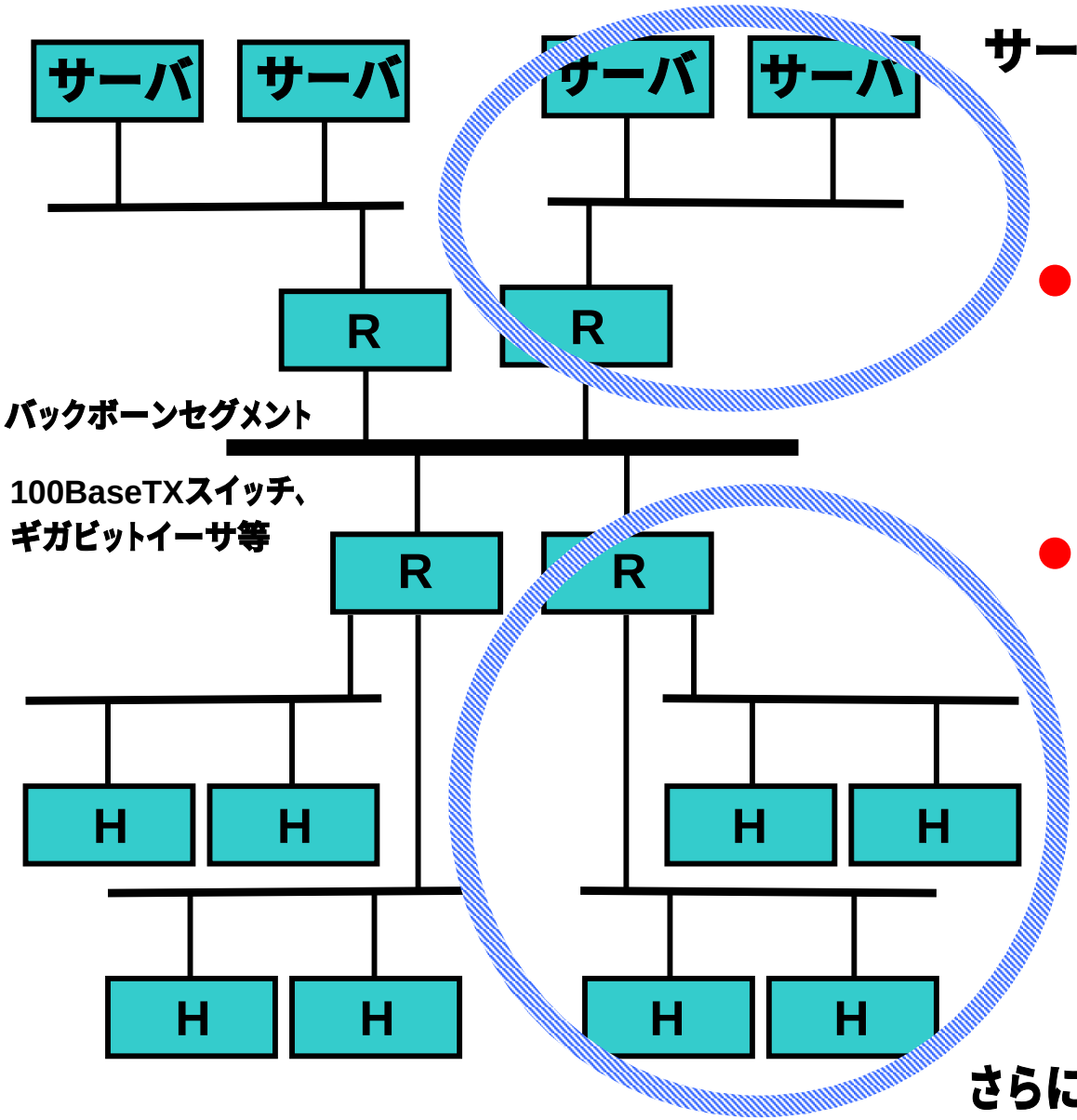
サーバの増設



- サーバセグメントの安全性を保ちつつ増設する
- クライアントセグメントのbroadcast をそのセグメント内に留められるためbroadcast flood現象の発生を抑制できる

ネットワークの追加

ネットワーク設計-3



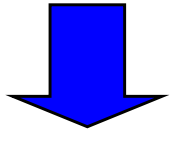
サーバセグメントの増設

- さらにセグメントを増設する際は、バックボーンセグメントを高速化するだけで、対応が可能
- スイッチングベースのネットワークから、左図のような大規模ネットワークに拡張する場合、リナンバリングが避けられない

さらにネットワークを追加

ネットワーク設計のまとめ

- スケーラビリティを考慮するとサブネット化は不可欠
- 安全性を考慮してサーバは別のセグメントに
- トラフィックの集中するサーバ、ルータなどにはスイッチを導入する
- 規模の拡大を見越したネットワークトポロジの設計



ネットワーク規模拡大を考慮したアドレス割り当て

アドレスの割り当てポリシーとは

- 規模の拡大を想定しネットワークアドレスの組織内割り当てを考える

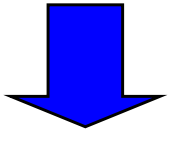
- 最適なネットマスクでの利用
- 最適なIPアドレスの割り当て方法
  - ネットマスク変更法
  - サブネット追加法

最適なネットマスクでの利用

- ホスト数 50台を想定したネットワーク
  - 172.16.0.0/16 利用の場合
    - 未使用IPアドレス 65484
    - ウイルス感染などによるarp broadcast floodの発生状況
      - 65484 × 8 = 52万回程度のbroadcastが発生
  - 192.168.0.0/24 利用の場合
    - 未使用IPアドレス 234
    - ウイルス感染などによるarp broadcast floodの発生状況
      - 204 × 8 = 1632回程度のbroadcastが発生
  - 192.168.0.0/26 利用の場合
    - 未使用IPアドレス 12
    - ウイルス感染などによるarp broadcast floodの発生状況
      - 12 × 8 = 96回程度のbroadcastが発生
- 最適なマスク設定の必要性
  - 拡張性のために大きめのネットワークである/16や/24を設定するとウイルス感染時に脆弱なネットワークとなってしまう
  - broadcastはスイッチであってもすべてのポートを占有するため、少量のパケットでも輻輳が発生しやすい。

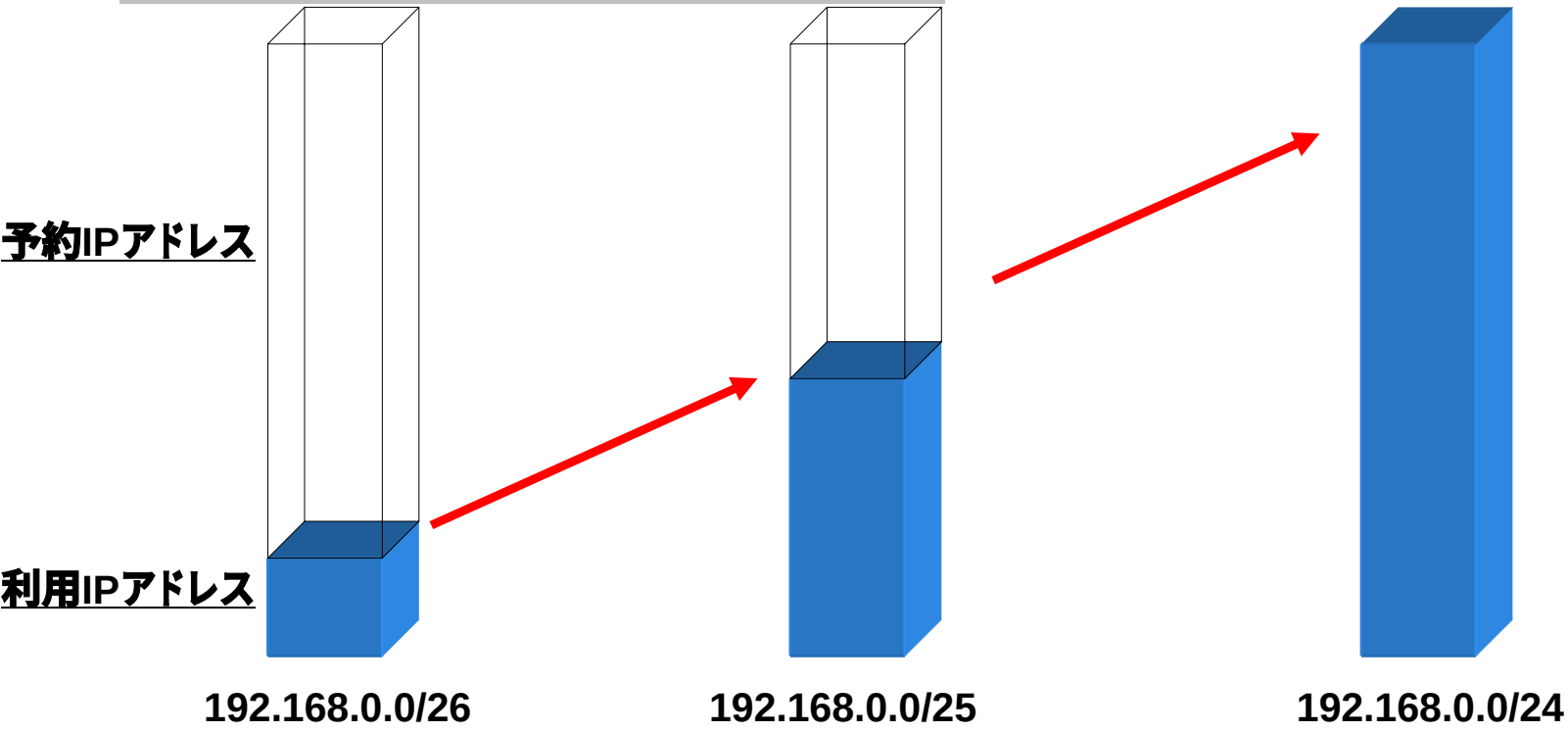
最適なネットマスクでの利用

- 拡張性を持たせつつ、マスクを最長(最少ネットワーク)にするにはどうすれば良いか
  - 拡張性を持たせるには未使用IPアドレス空間を持たせる必要がある
  - 未使用アドレス空間を大きくしすぎるとウイルス感染時に脆弱なネットワークとなってしまう



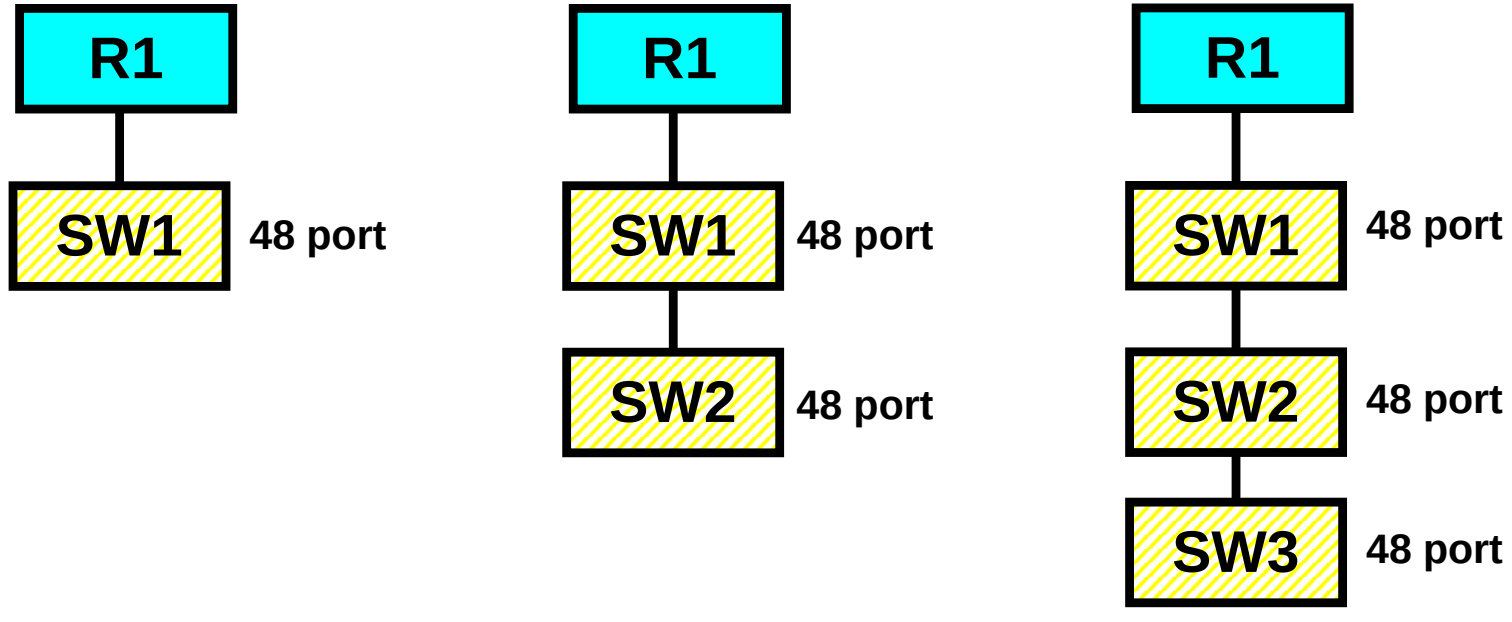
最適なIPアドレス割り当てをどうすべきか？

ネットマスク変更法1



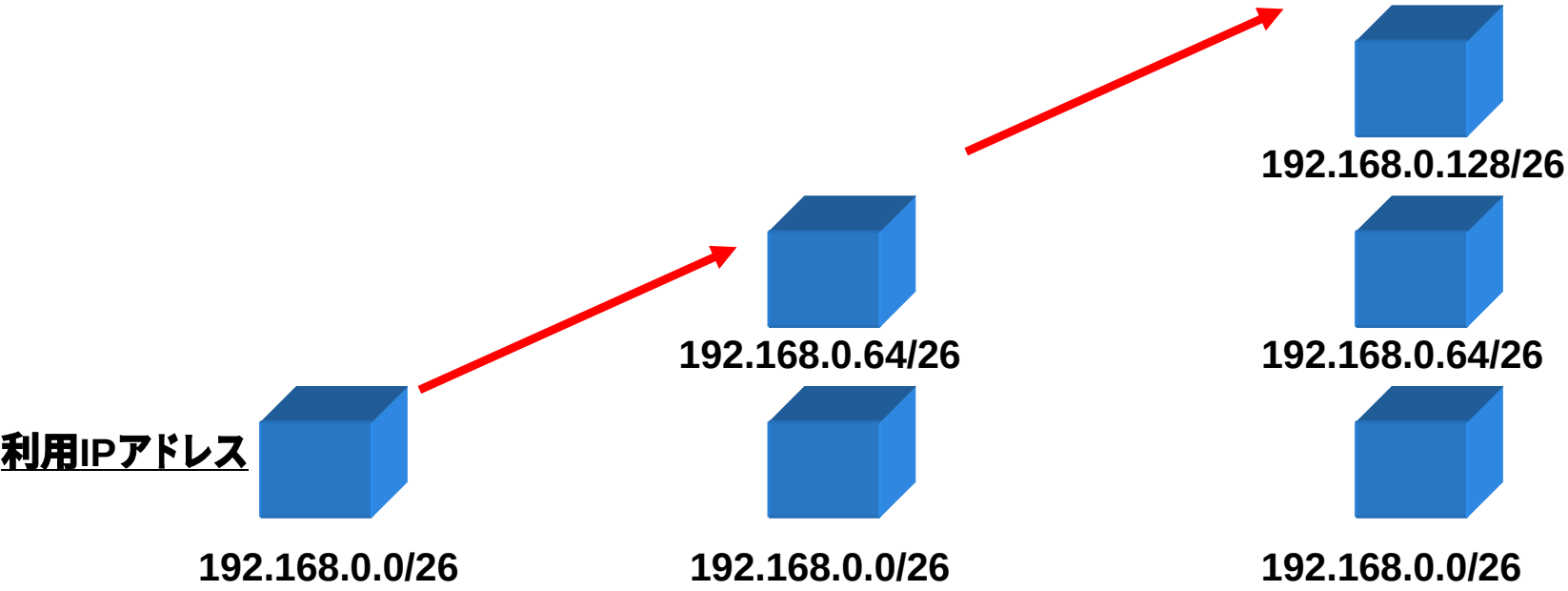
- ネットマスク変更法によるネットワーク拡張
  - 最初から/24程度まで拡張することを前提にIPアドレスを予約する
  - 実際に利用するIPアドレスは/26空間のみとし、/26で不足する場合には/25、/24とマスクを拡張していく
  - /24以上の大きさに対してはサブネット追加を行う

ネットマスク変更法2



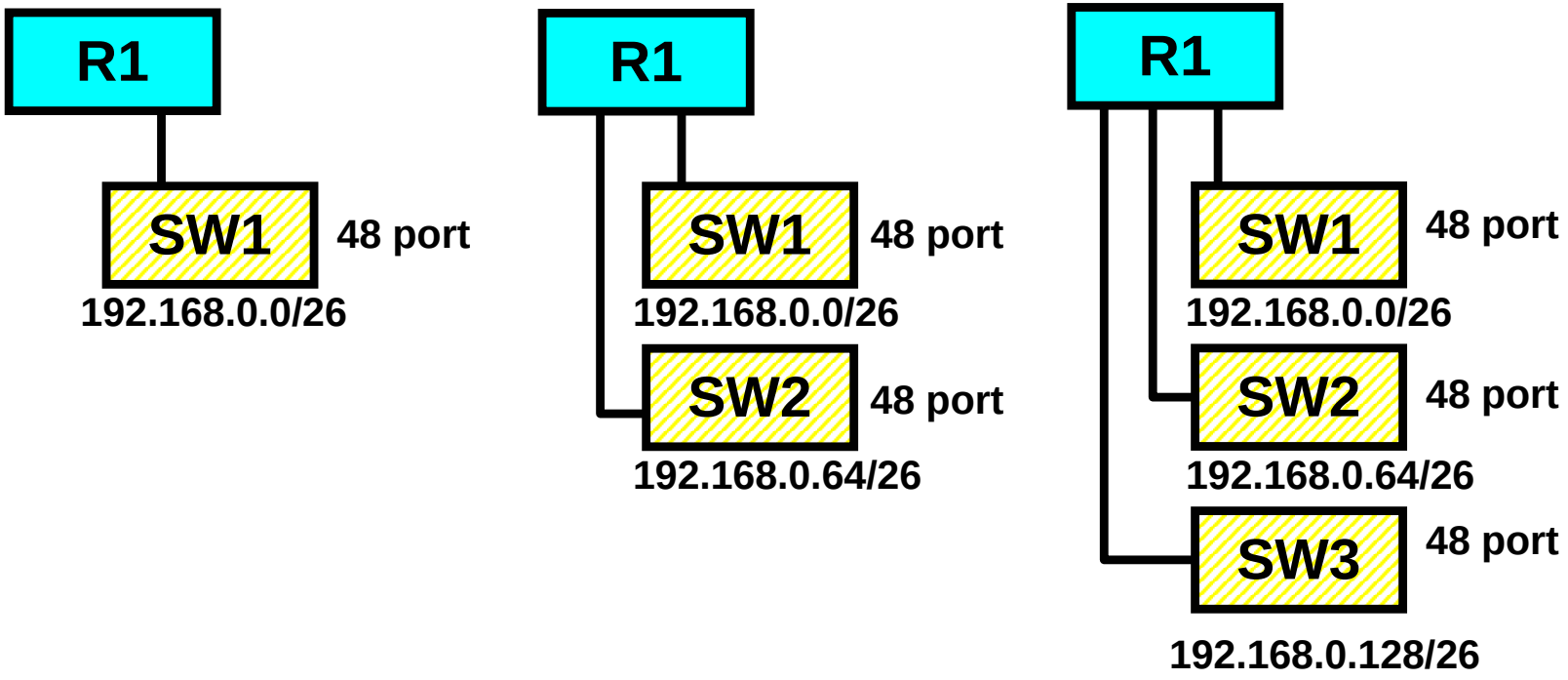
- ネットマスク変更法によるネットワーク拡張
  - マスク変更を前提とした割り当てはL2ネットワークを拡張していく必要があるため、カスケードされたL2ネットワークとなる場合が多い
  - カスケードされたL2ネットワークは障害箇所の特定などが困難となるだけでなく、broadcastドメインを大きくし、パフォーマンスを低下させる

サブネット追加法1



- サブネット追加法によるネットワーク拡張
  - 実際に利用するIPアドレスは/26空間のみとし、/26が不足した場合には/26サブネットを別に構築する
  - 追加されたネットワークはルータなどを経由してトラフィック交換されるため、L2接続に依存したアプリケーションは利用できない

サブネット追加法2



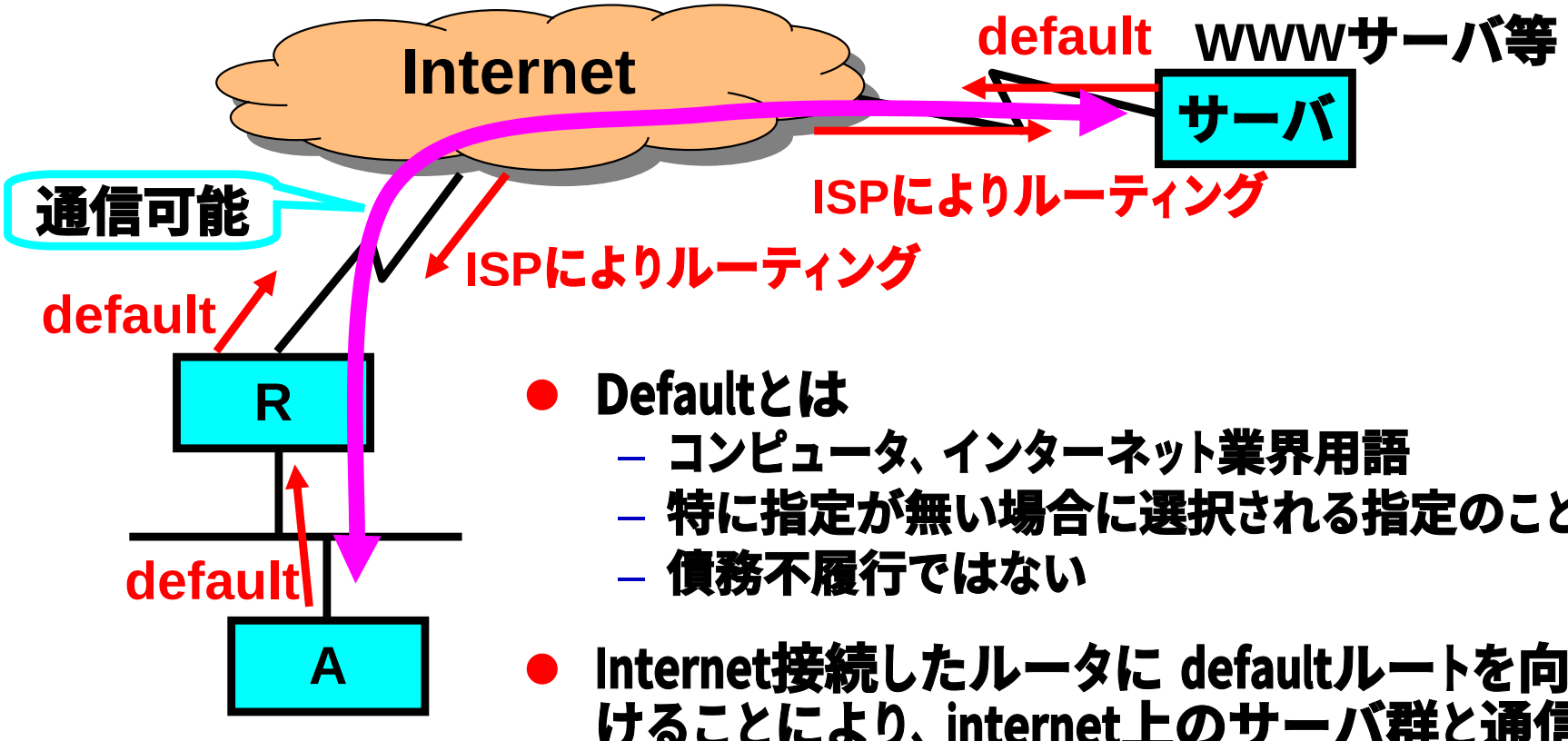
- サブネット追加法によるネットワーク拡張
  - サブネット追加による拡張では、スイッチの物理ポートとサブネットが対応し、スイッチの増設に合わせてサブネットも追加される
  - broadcastドメインの広さ(L2ネットワークの大きさ)を一定の大きさに保つことができるため、安定した拡張を行うことができる
  - 管理しやすくするために/24を予約して、/26だけ使っても良い
    - スペース重視: 192.168.0.0/26, 192.168.0.64/26, 192.168.0.128/26
    - 管理重視: 192.168.0.0/26, 192.168.1.0/26, 192.168.2.0/26

IPアドレスの割り当てポリシー

- ネットマスク変更法によるネットワーク拡張
  - 最初から/24程度まで拡張することを前提にIPアドレスを予約する
  - 実際に利用するIPアドレスは/26空間のみとし、/26で不足する場合には/25、/24とマスクを拡張していく
  - L2ネットワークをそのまま拡張できるため、L2接続に依存したアプリケーションも対応可能
  - マスク変更時には同一サブネットの既存ホストの変更が必要となるため、拡張時に負担がかかる

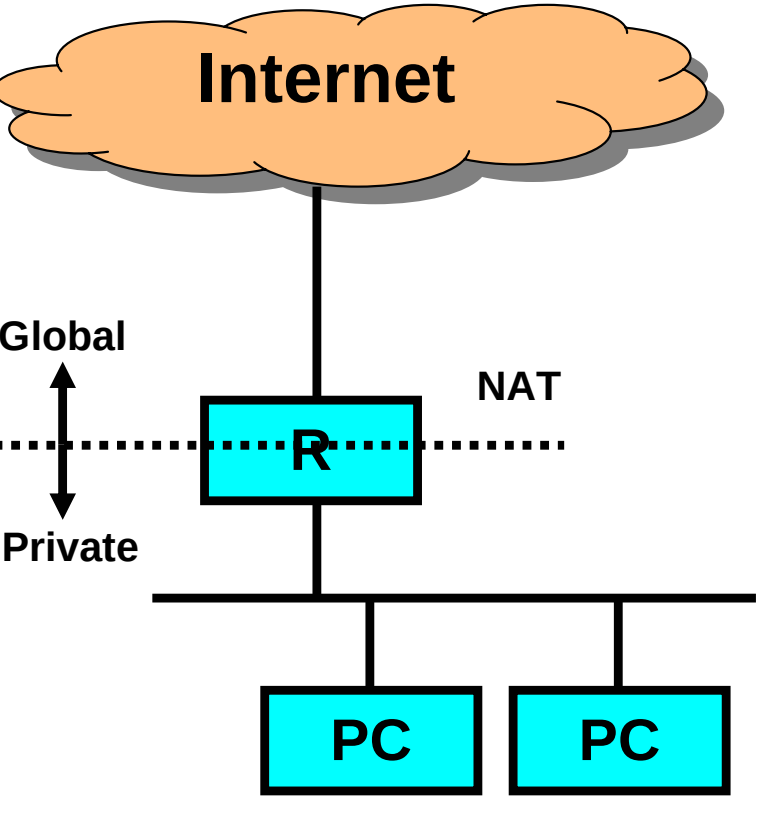
- サブネット追加法によるネットワーク拡張
  - 実際に利用するIPアドレスは/26空間のみとし、/26が不足した場合には/26サブネットを別に構築する
  - 追加されたネットワークはルータなどL3レベルでトラフィック交換されるため、L2接続に依存したアプリケーションは利用できない
    - L2接続に依存したアプリケーションとは、IPアドレスを付与せずにプリンタ出力できたり、ファイル交換するアプリケーションがあげられる。
    - このようなアプリケーションはネットワーク拡張に支障をきたすことが多い
  - ネットワークを設計するうえで、あらかじめ24portや48port HUBごとにサブネット化する仕様としておけば容易に拡張することができる
  - L3レベルでのトラフィック交換を前提としたアプリケーション利用とする必要がある
  - 管理しやすくするために/24を予約して、/26だけ使っても良い

インターネットへの接続形態



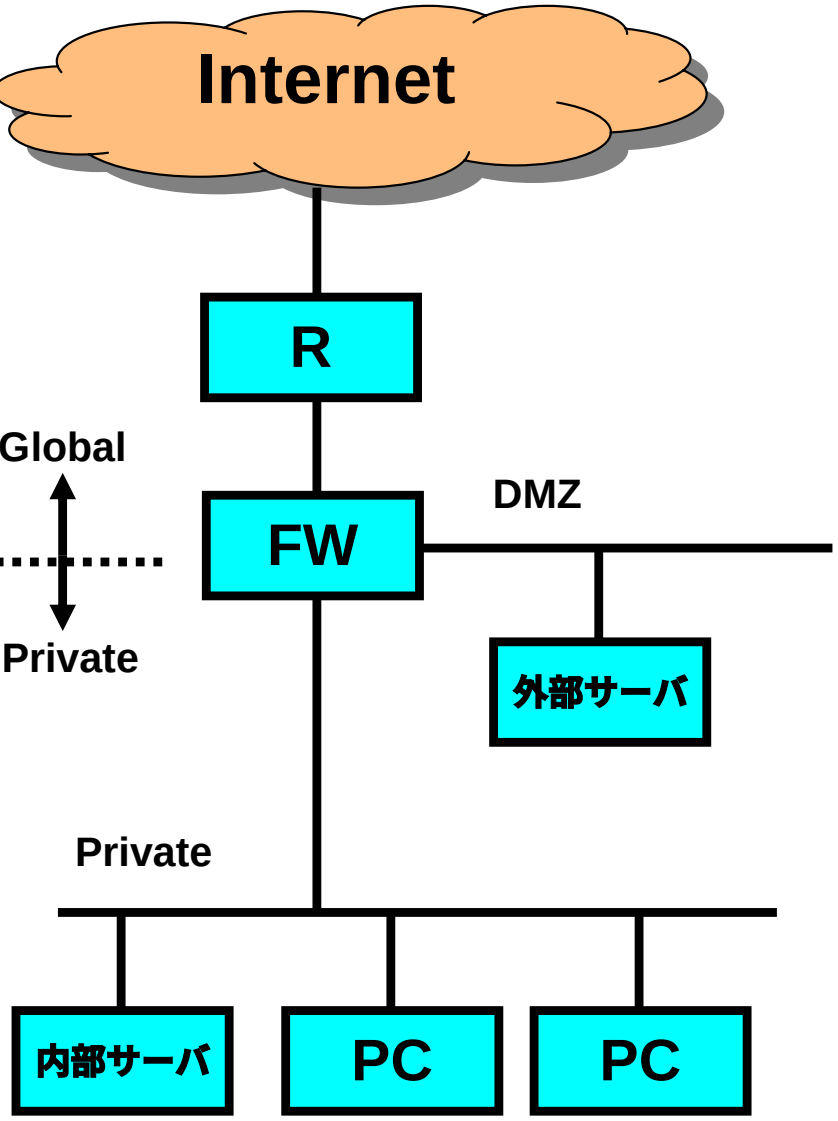
- Defaultとは
  - コンピュータ、インターネット業界用語
  - 特に指定が無い場合に選択される指定のこと
  - 債務不履行ではない
- Internet接続したルータに defaultルートを向けることにより、internet上のサーバ群と通信が行える
- Internet接続にはルーティングは必須

インターネット接続事例ーA



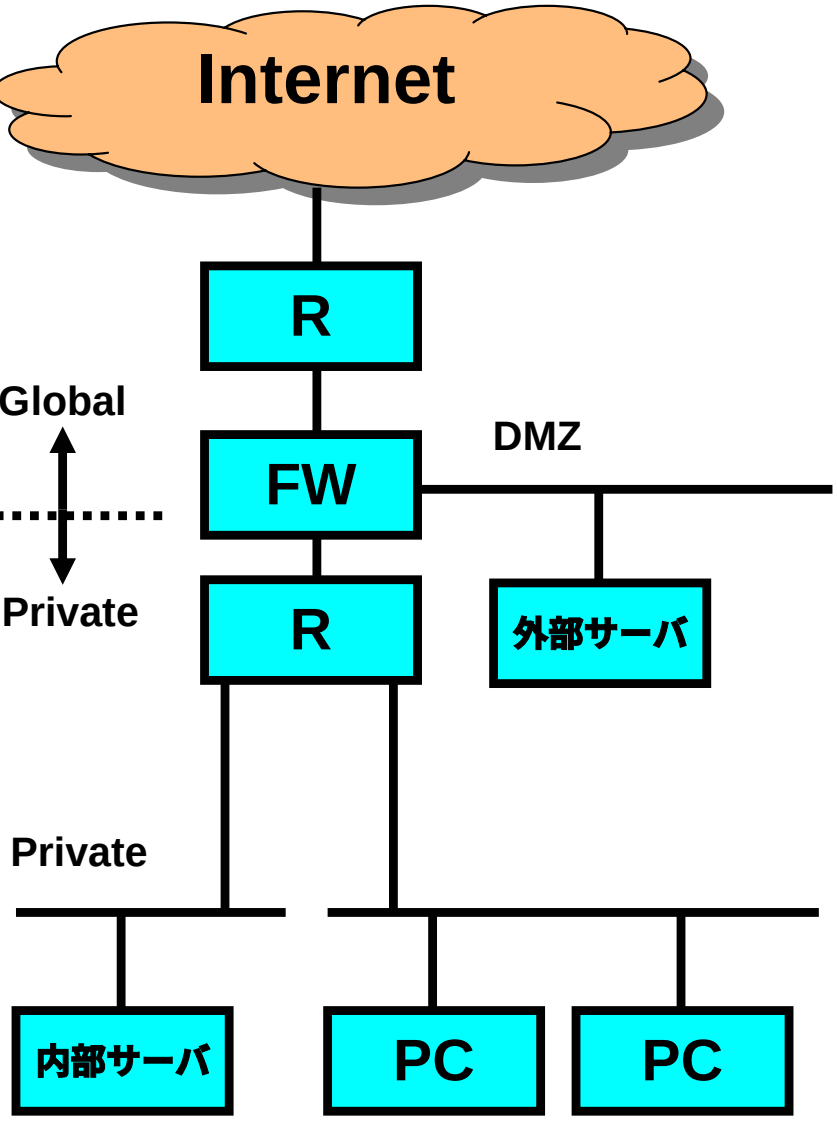
- サーバレス運用
  - ISPのDNS,Mail,Webサービスを利用
- セキュリティ
  - 低い
  - ルータのNAT機能に依存
  - 重要なデータをPCに置けない

インターネット接続事例ーB



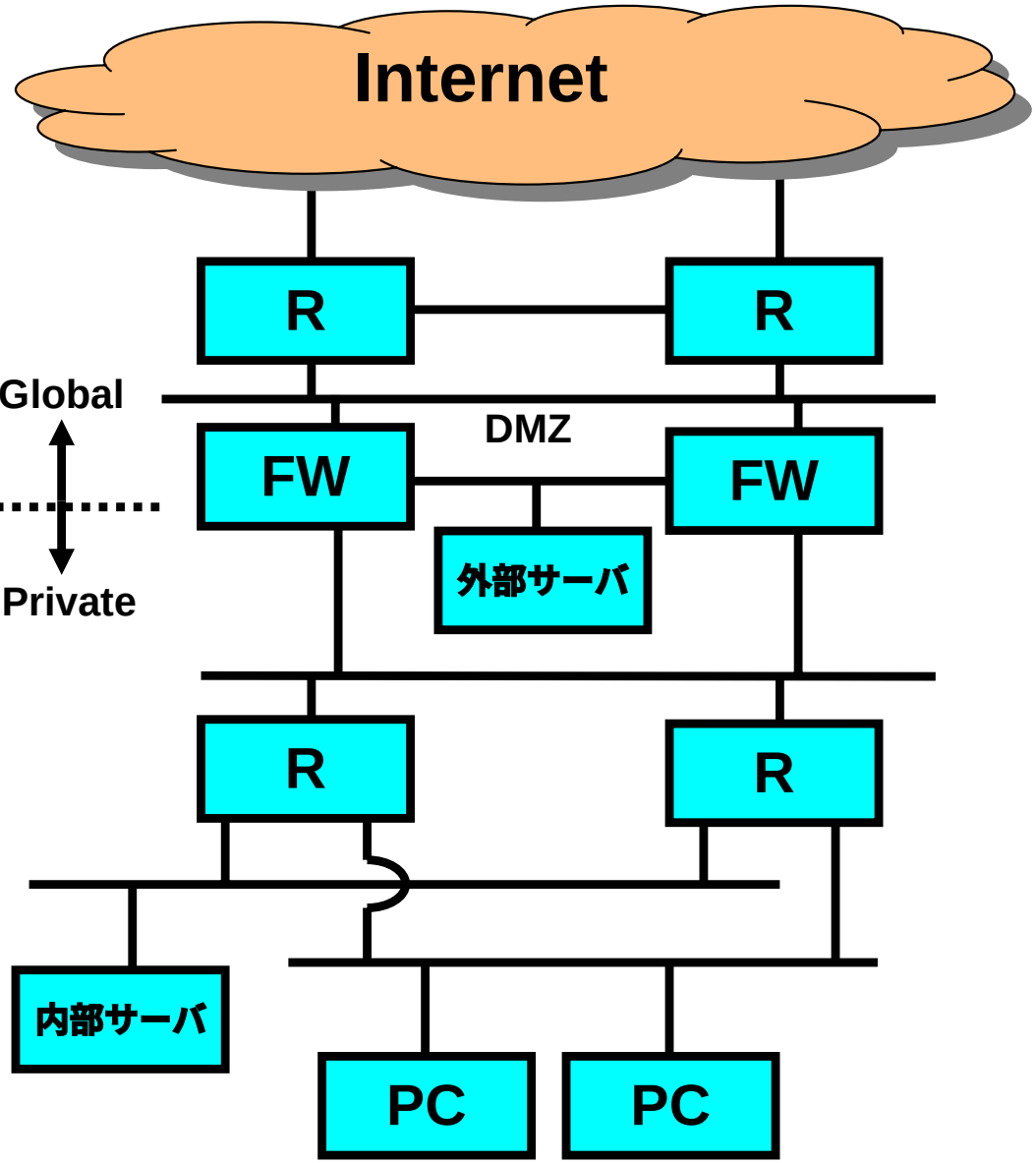
- 自社サーバ運用
- 外部サーバ
  - DMZ (Demilitarized Zone)に設置
  - 公開Web
  - 外部DNS
- 内部サーバ
  - Mail
  - 内部DNS
- セキュリティ
  - 標準的

インターネット接続事例ーC



- 自社サーバ運用
- 外部サーバ
  - DMZ (Demilitarized Zone)に設置
  - 公開Web
  - 外部DNS
- 内部サーバ
  - Mail
  - 内部DNS
- セキュリティ
  - 標準的
- 内部サーバの保護

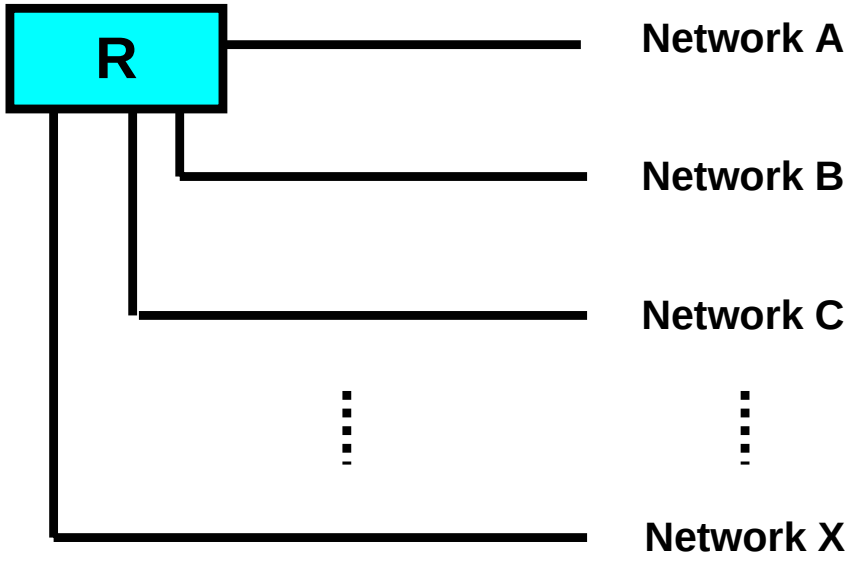
インターネット接続事例ーD



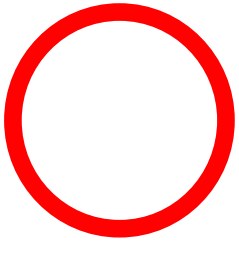
- 自社サーバ運用
- 外部サーバ
  - DMZ (Demilitarized Zone)に設置
  - 公開Web
  - 外部DNS
- 内部サーバ
  - Mail
  - 内部DNS
- セキュリティ
  - 標準的
- 内部サーバの保護
- 回線、機器の二重化
  - HSRP,VRRPの利用

IIJ
Internet Initiative Jap

ネットワーク拡張（スター型）

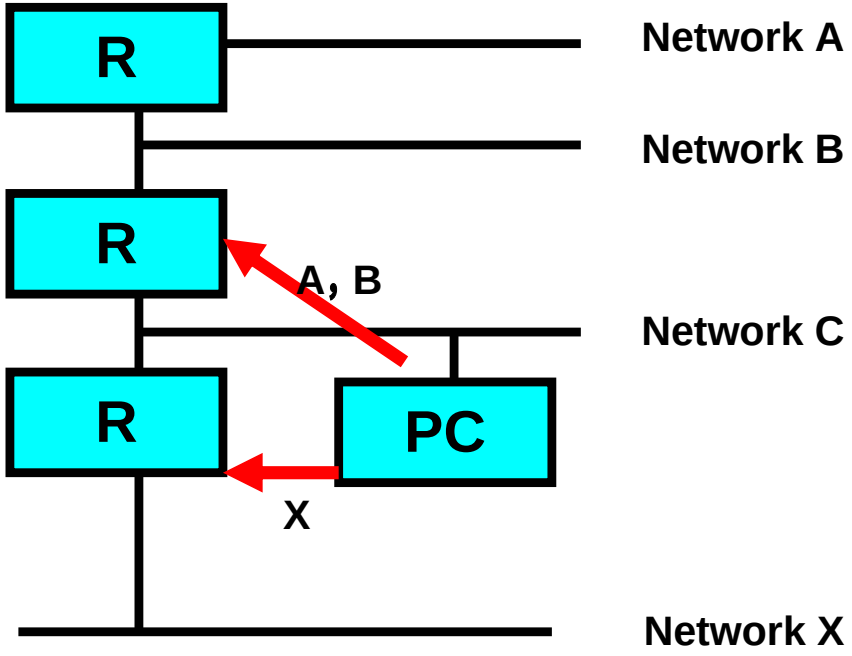


- スター型拡張
  - スター型のネットワーク拡張はルーティングを単純化できるだけでなくポリシー制御も容易なため、小規模から大規模まで幅広く利用されれいる
- 特徴
  - ルーティングが容易
  - ポリシー制御が容易
  - 大規模となると集約されるルータを高性能化する必要がある
  - 多くのネットワークを収容できるルータが必要となるが、VLANなどの利用で安価に構成できるようになった

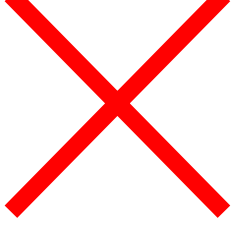


ネットワーク拡張の基本であり、特別な事情がない限り、まずこの形式を検討すべきである

ネットワーク拡張(数珠型)

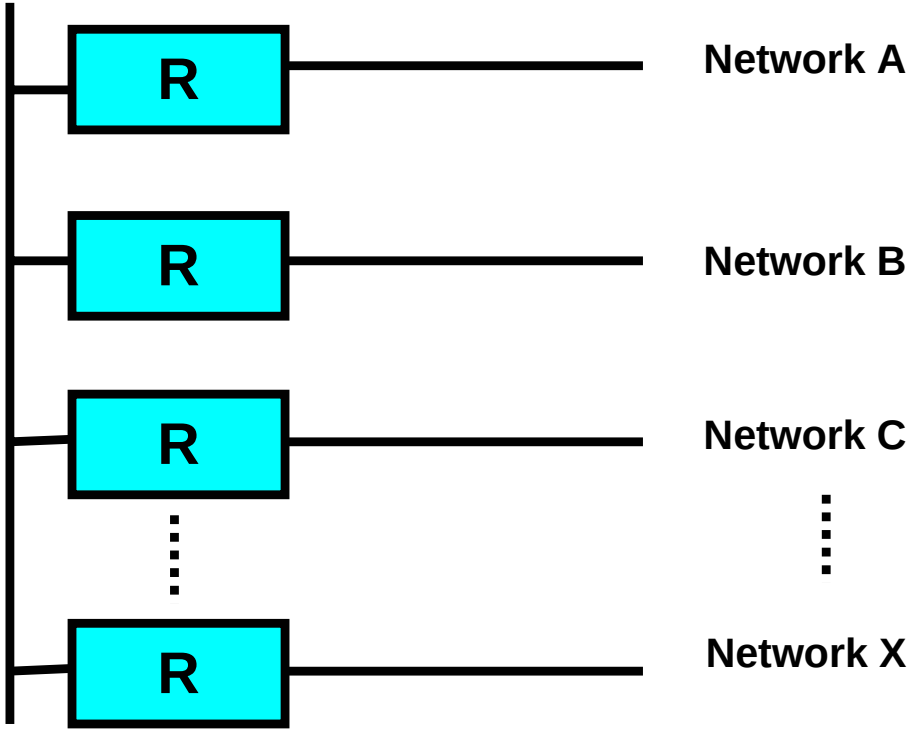


- 数珠型拡張
  - フロアやビル間などを1つのネットワークで構成し、かつ、そのネットワーク上にクライアントが繋がれるモデル
- 特徴
  - 大規模になるにつれてルーティングが複雑になる
  - ダイナミックルーティングとスタティックルーティングが混在し、誤動作する恐れがある



物理的にこの形式しか組めない場合を除いて避けるべき構成である

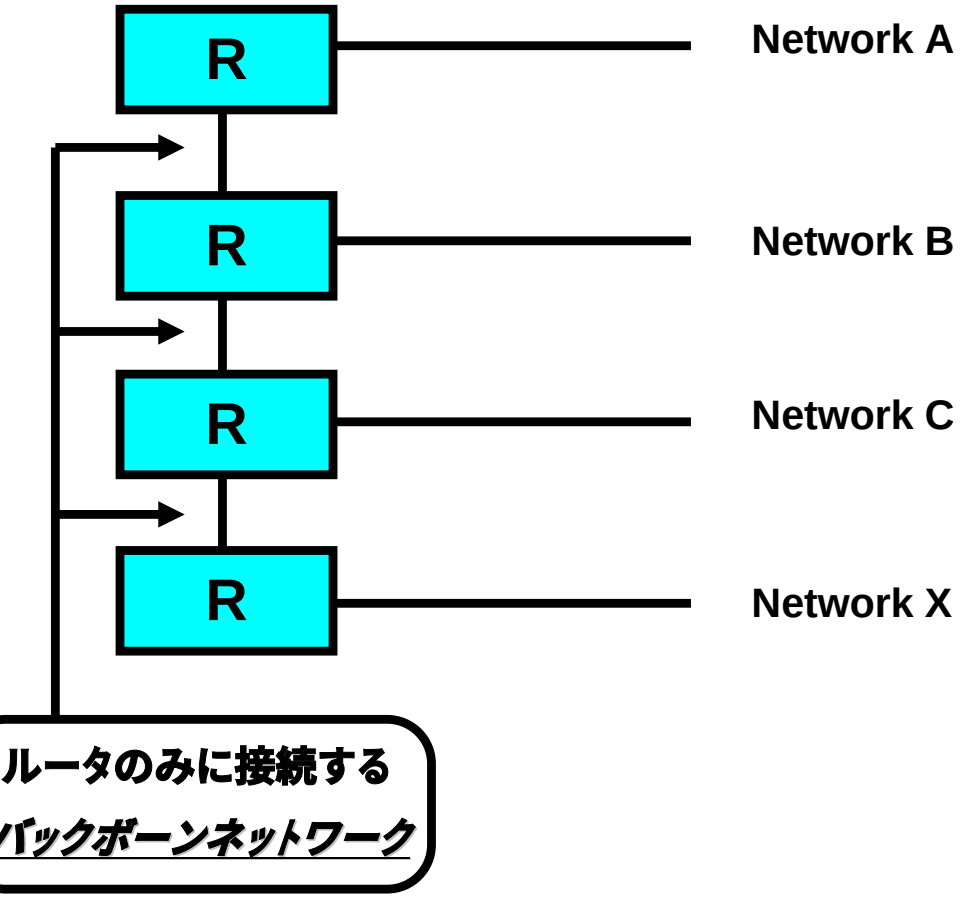
ネットワーク拡張（L2バックボーン型）



同一構内などのLAN接続などに有効に利用できる

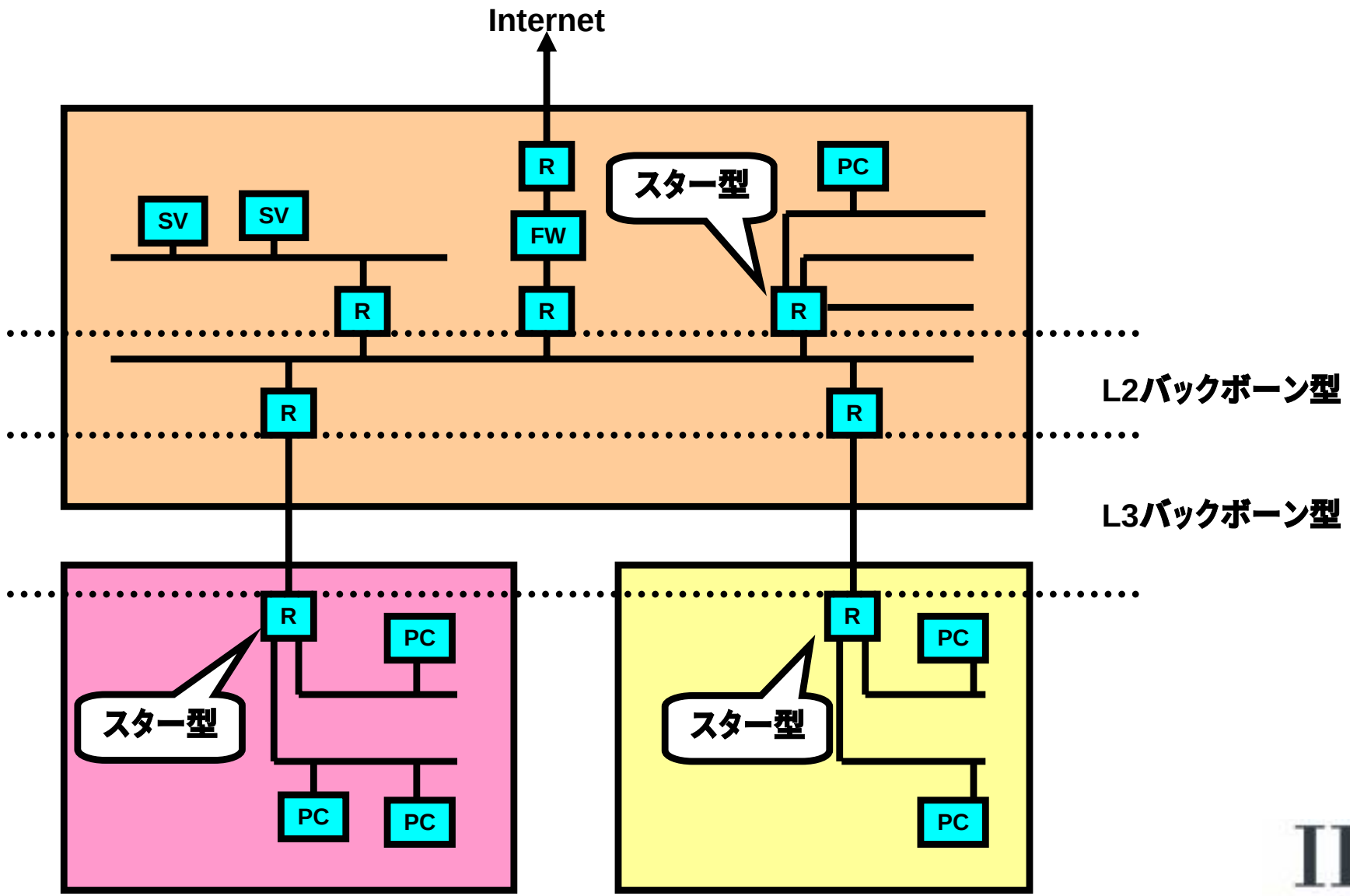
- L2バックボーン型拡張
  - 1つのLayer 2をルータが共有し、PCやサーバとルータを混在させないようにする。
- 特徴
  - ルーティングはルータのみで行えるため、スタティックからダイナミックまで拡張が可能
  - 1つのLayer 2を共有するため、長距離の伝送が難しい
  - 1つのLayer 2が大きくなりすぎる前にバックボーンの階層化を検討する必要がある

ネットワーク拡張（L3バックボーン型）



- L3バックボーン型拡張
  - ルータ間をPoint to Point ネットワークで接続し、るーたのみのバックボーンネットワークを構築する
- 特徴
  - ルーティングはルータのみで行えるため、スタティックからダイナミックまで拡張が可能
  - 専用線などが利用でき、長距離の伝送が容易
  - L2バックボーンに比べて高価なため、拠点間などの長距離に用いる

ネットワーク拡張事例


Internet
R
FW
R
R
スター型
PC
SV
SV
R
R
R
R
L2バックボーン型
L3バックボーン型
R
PC
スター型
PC
PC
R
PC
スター型
PC

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

静的な経路制御と動的な経路制御

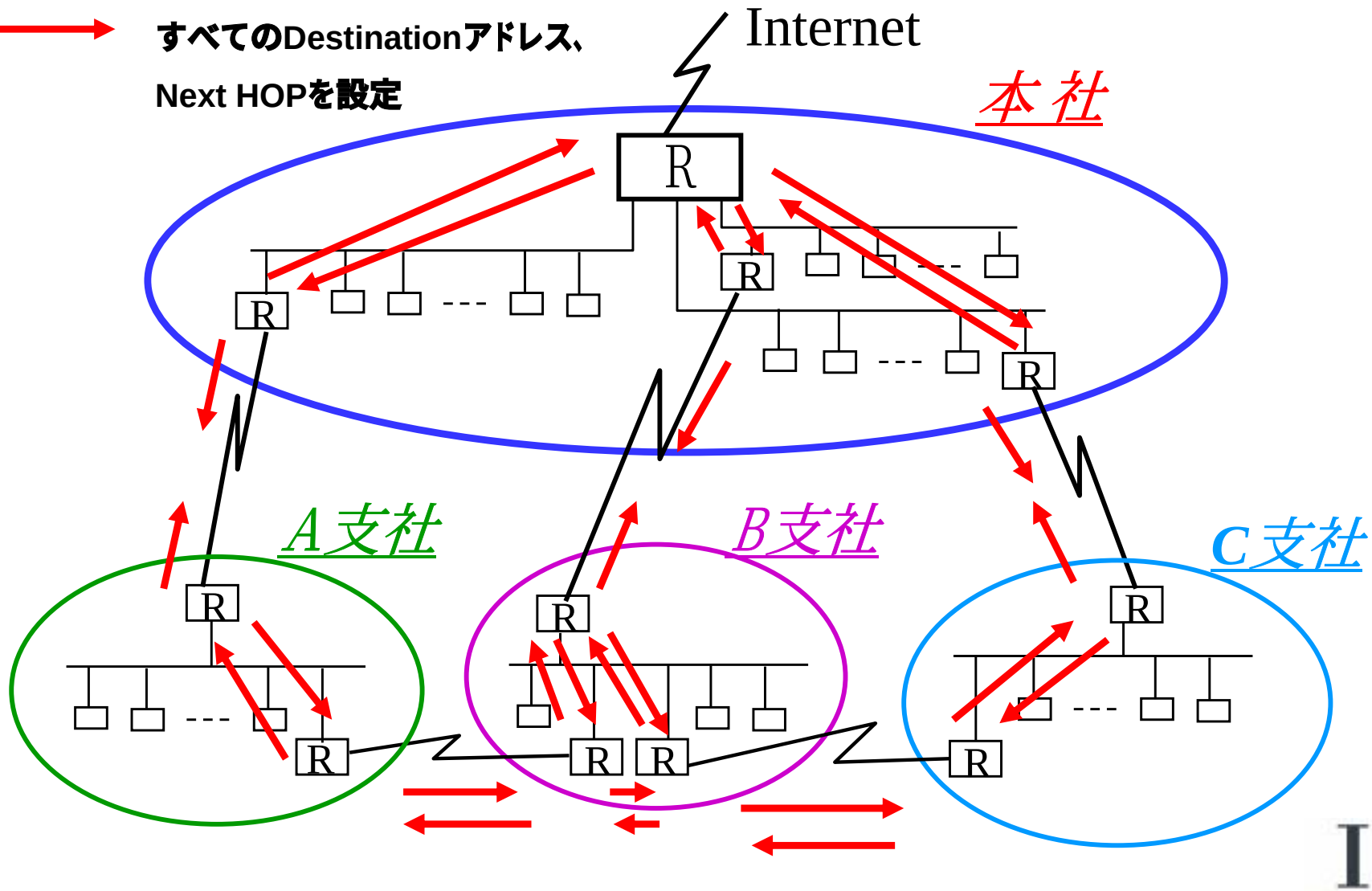
- 静的（スタティック）な経路制御の特徴
  - 手作業により固定的に経路を設定する
  - 安定している
  - トラフィックや伝送障害の影響を受けない
  - ルーティングプロトコルのためのトラフィックが発生しない
- 動的（ダイナミック）な経路制御の特徴
  - 自動的に経路を設定する
  - ネットワークの変化に対応できる
  - 自動的に最適経路を選択できる
  - 自動的にバックアップ経路を選択できる

スタティックルーティングによるネットワーク構築

スタティックルーティング

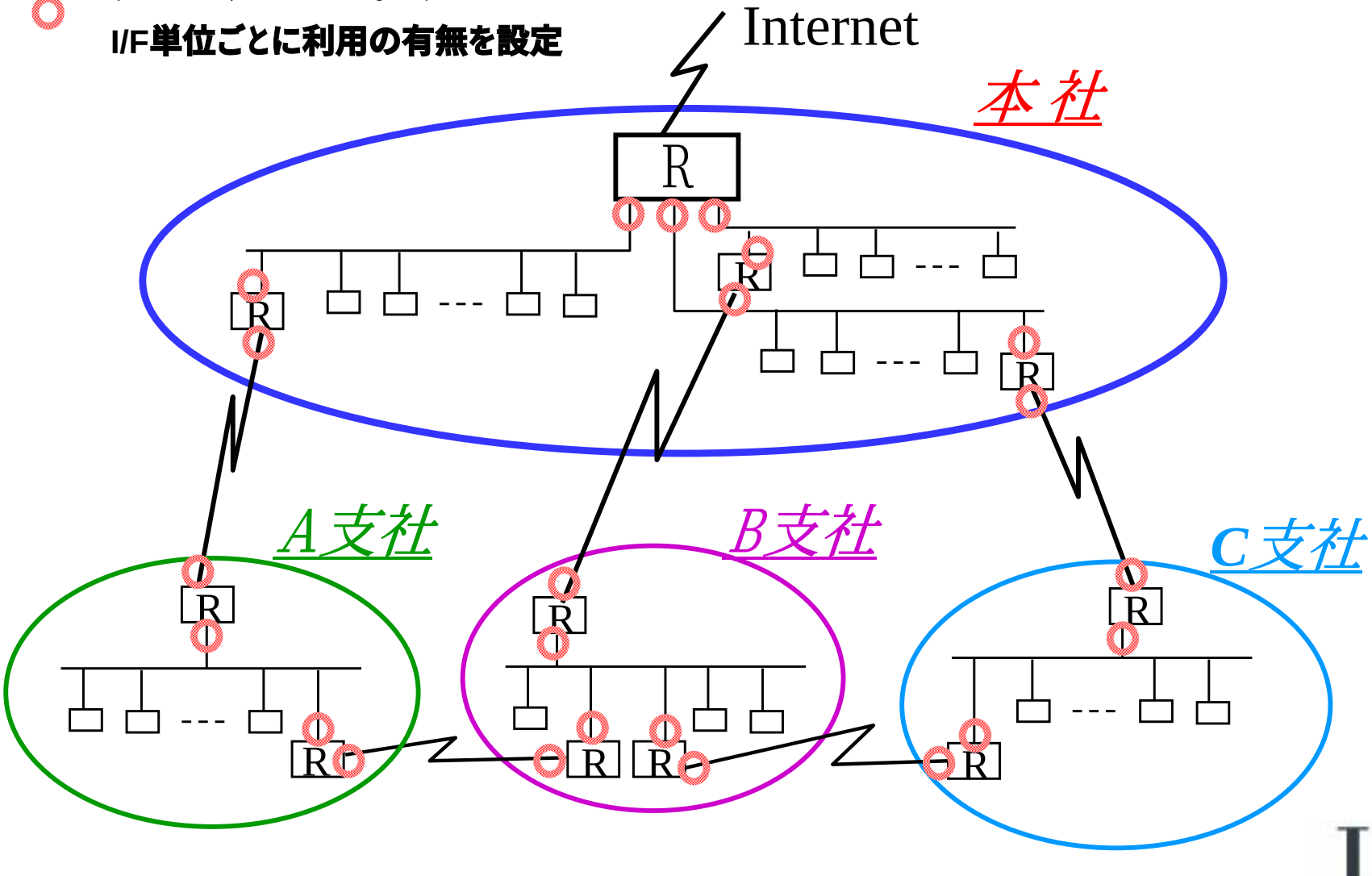
すべてのDestinationアドレス、

Next HOPを設定

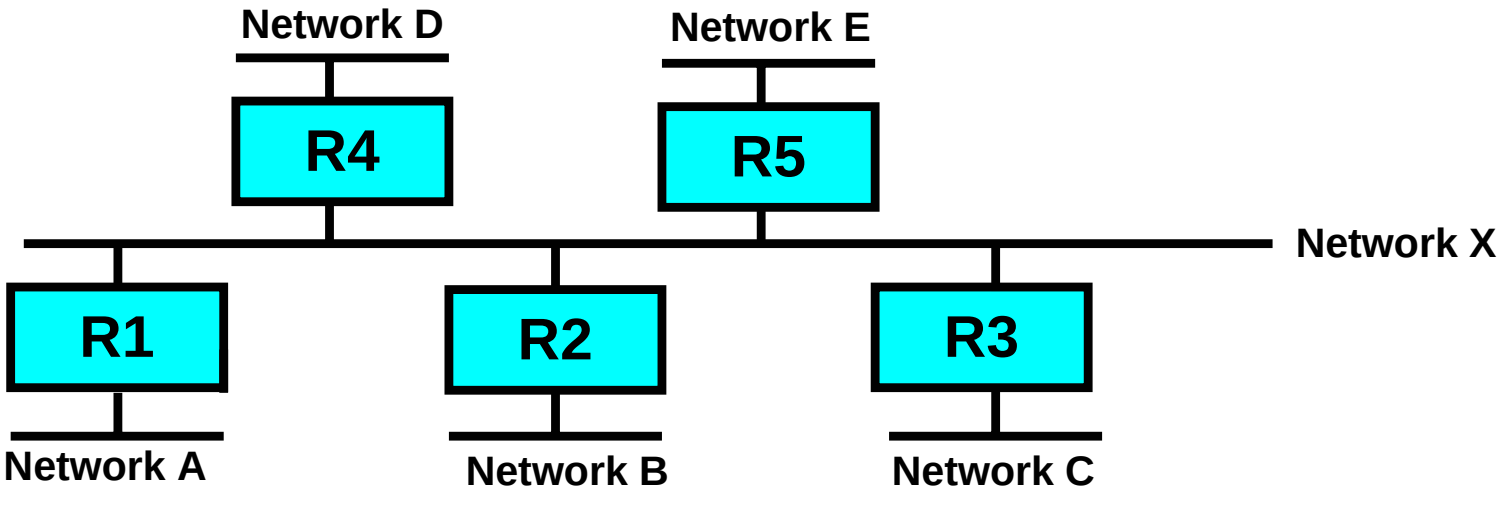


ダイナミックルーティングによるネットワーク構築

- ダイナミックルーティング
  I/F単位ごとに利用の有無を設定



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

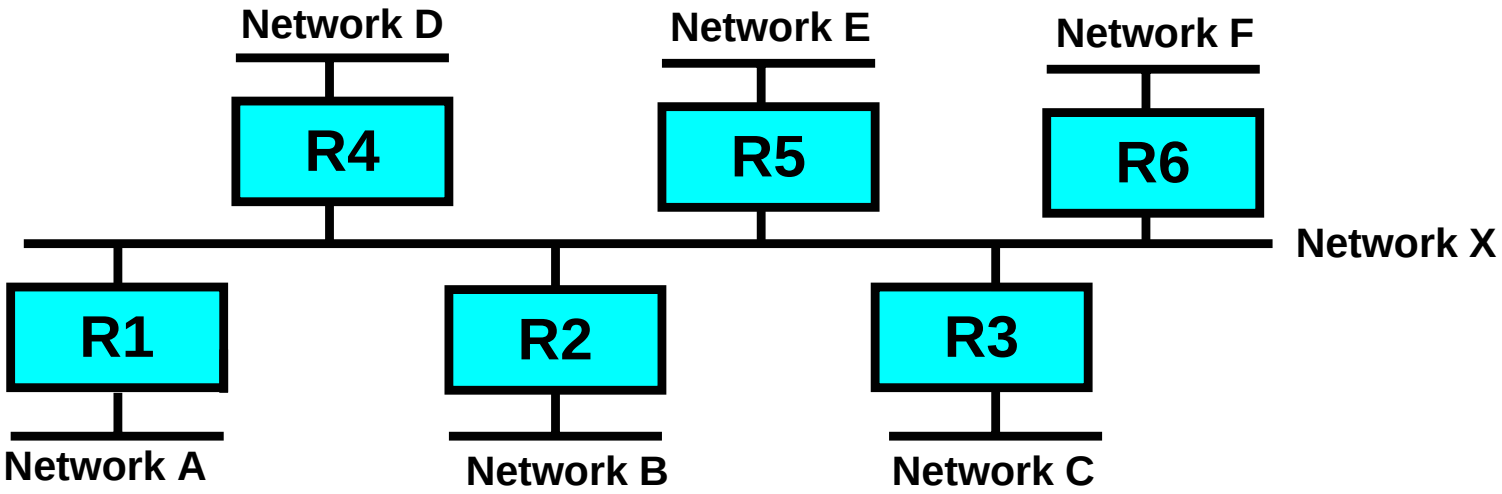
Destination	Next Hop
A	R1
B	R2
C	R3
E	R5

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

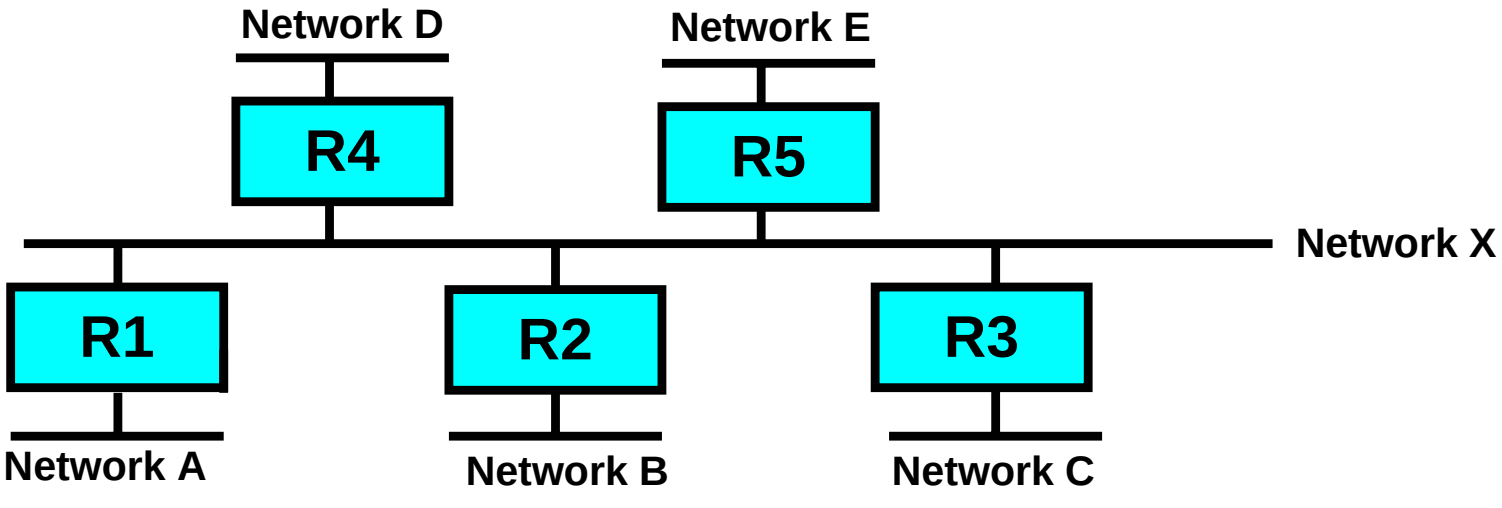
R6

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

IIJ
Internet Initiative Jap

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

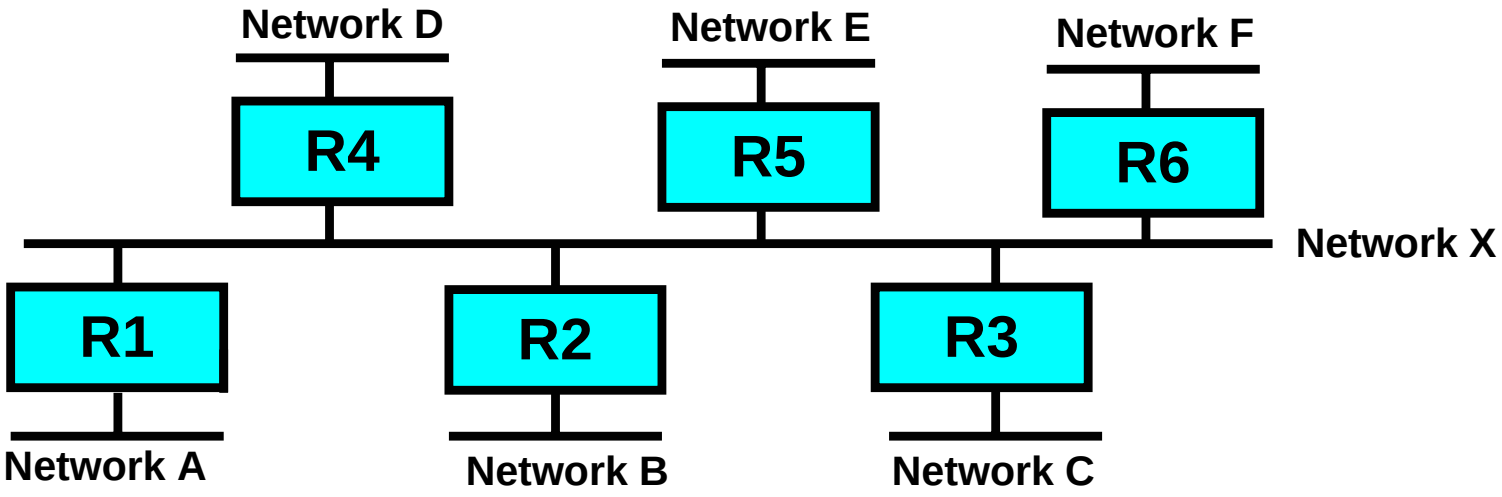
Protocol	Net
OSPF	X
OSPF(p)	D

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

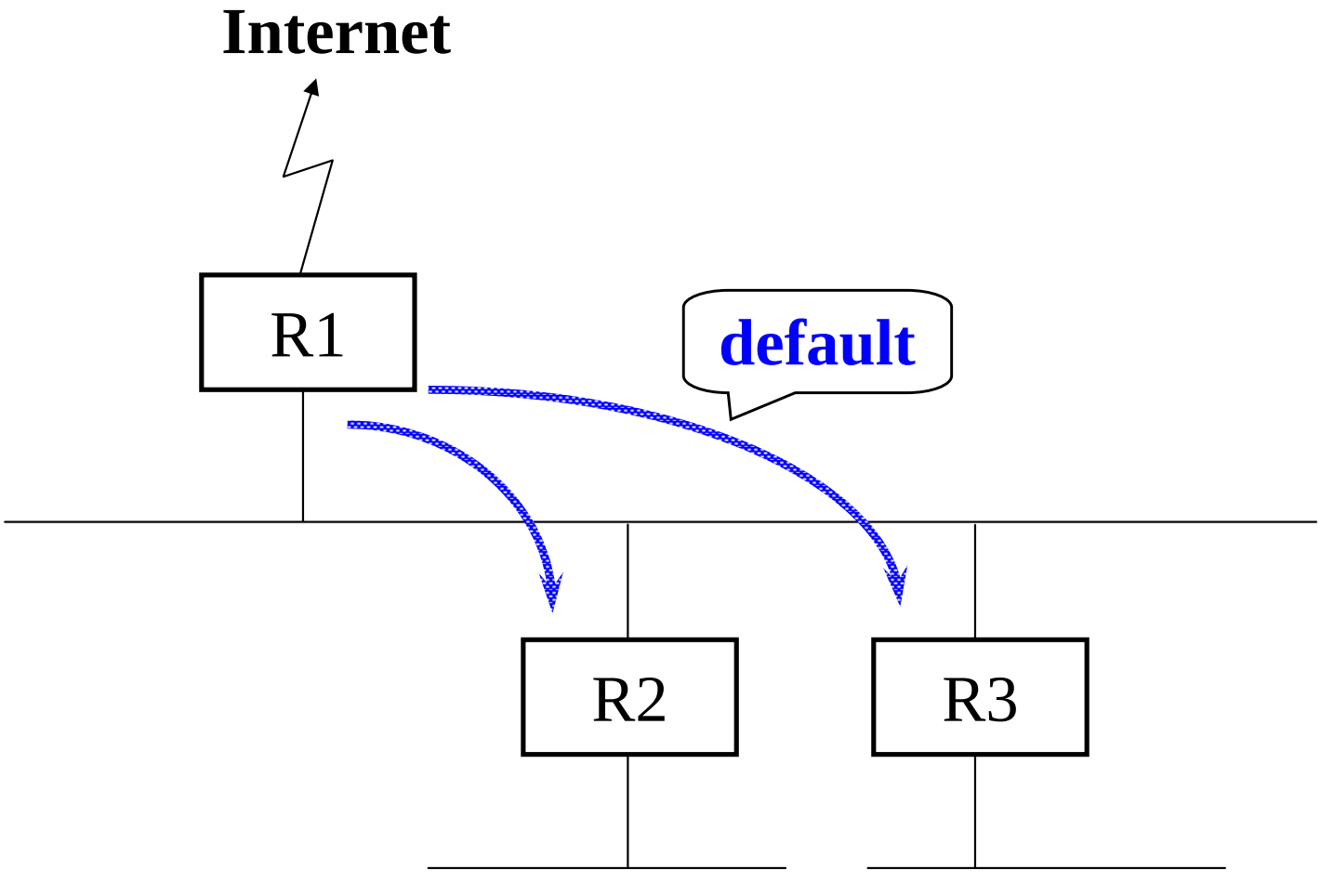
Protocol	Net
OSPF	X
OSPF(p)	F

- 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

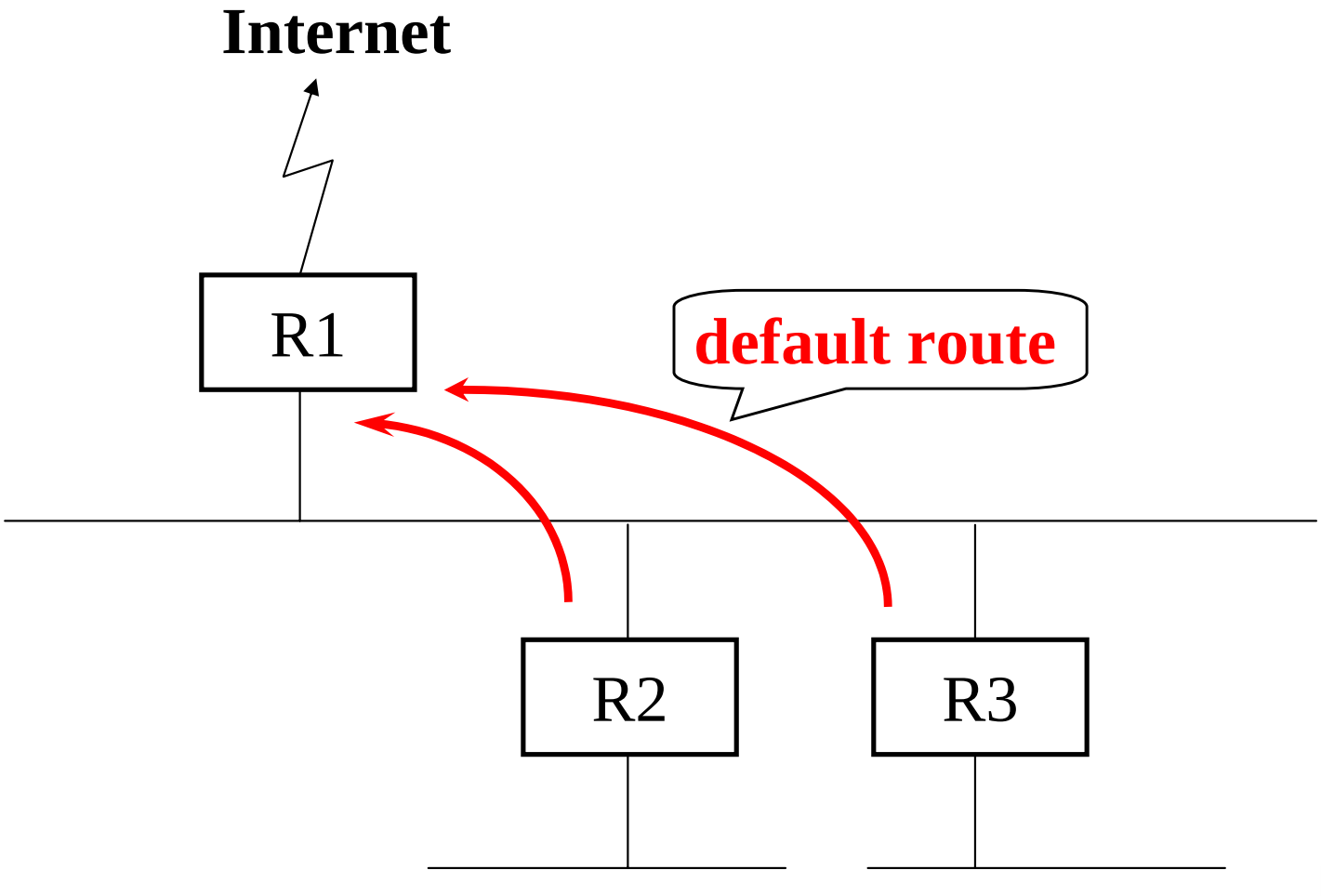
ルーティング設定まとめ

- 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

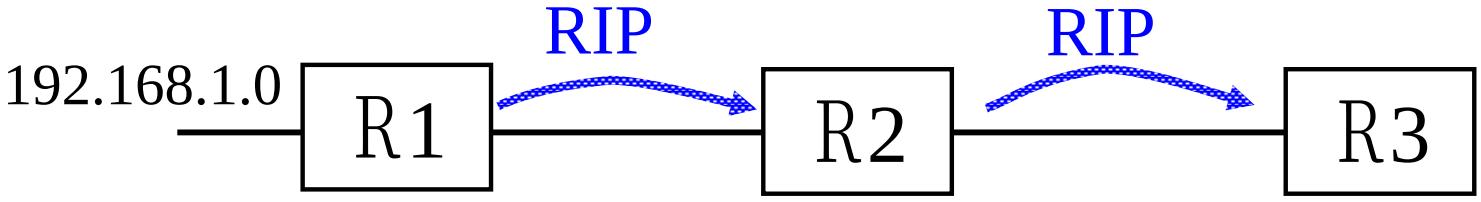
ベクトル距離経路制御

(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



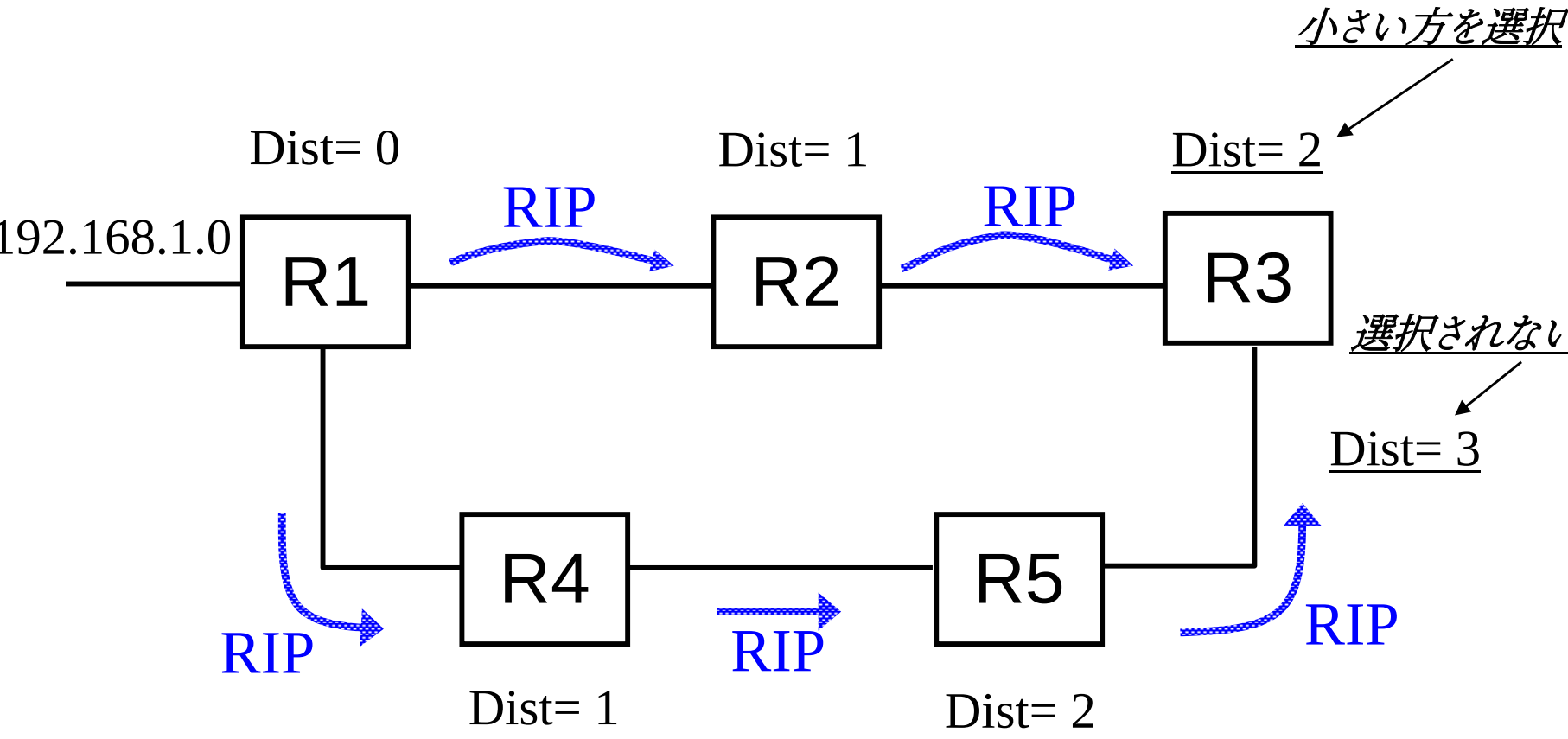
Dest=192.168.1.0
Dist= 0

Dest=192.168.1.0
Dist= 1

Dest=192.168.1.0
Dist= 2

Dest=Destination
Dist= Distance

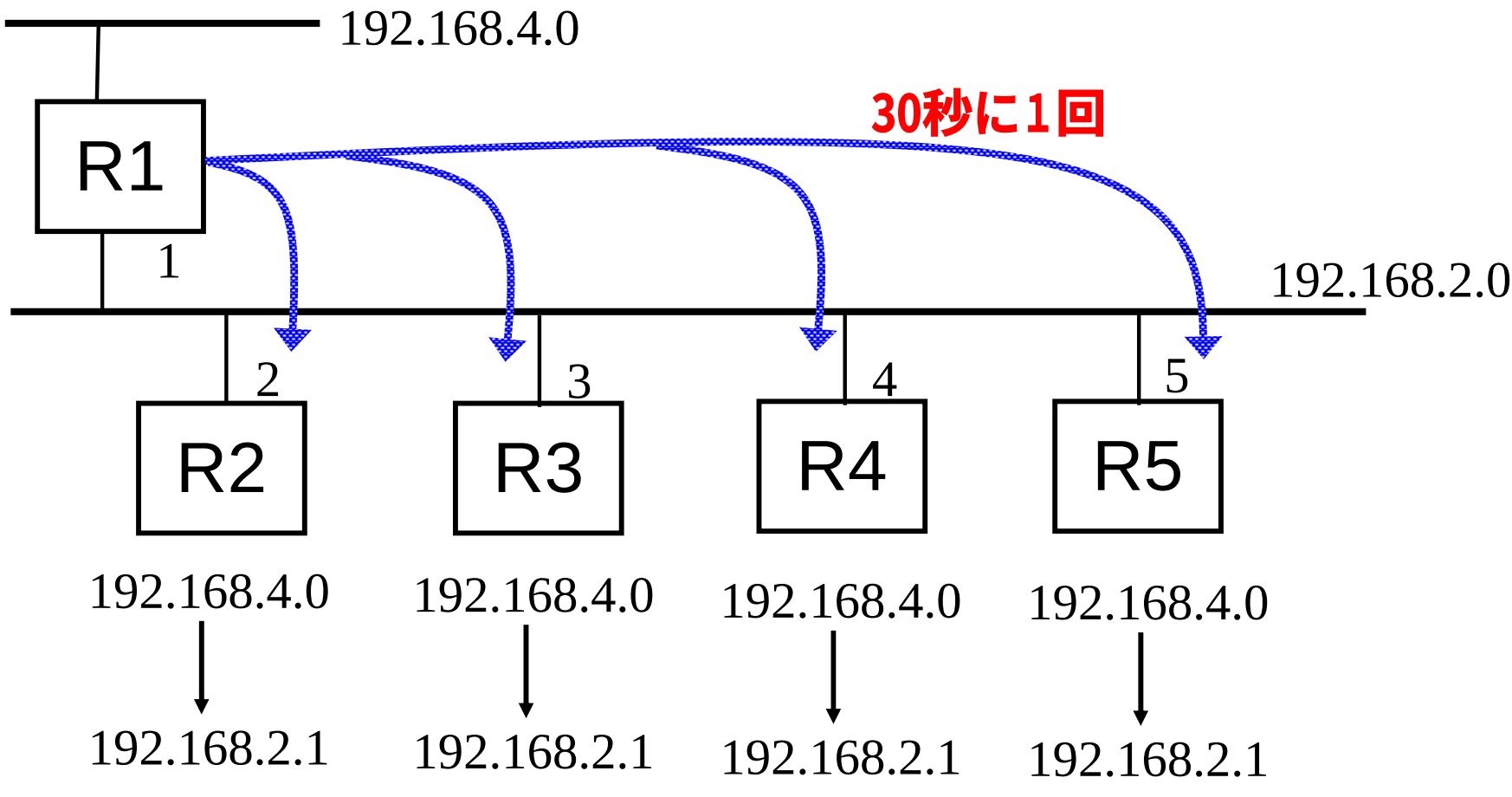
同じdestinationの場合はdistanceが小さい方を選択



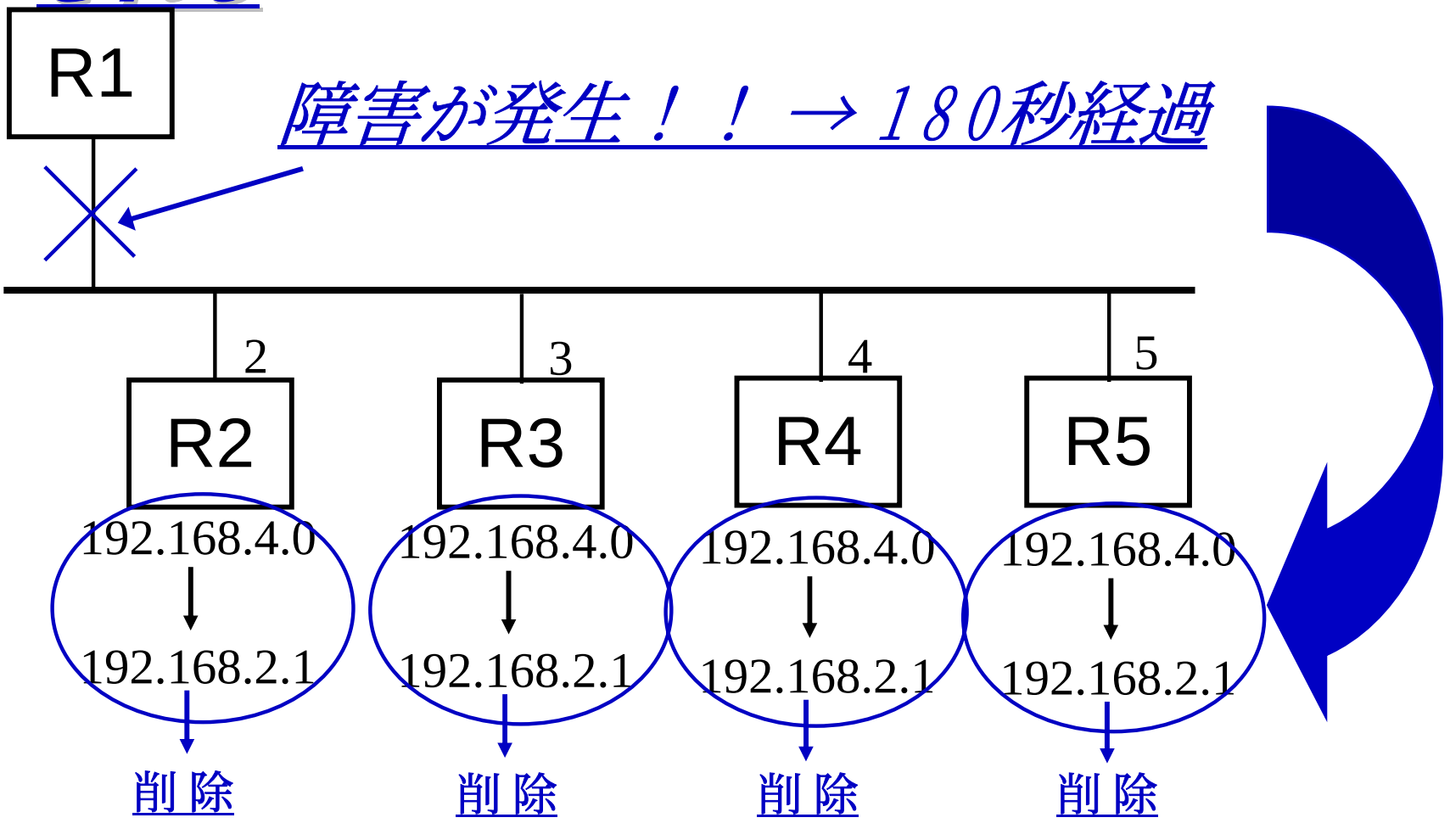
同じDestination同じDistanceの場合は
最初に到着した経路を選択

IIJ
Internet Initiative Jap

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

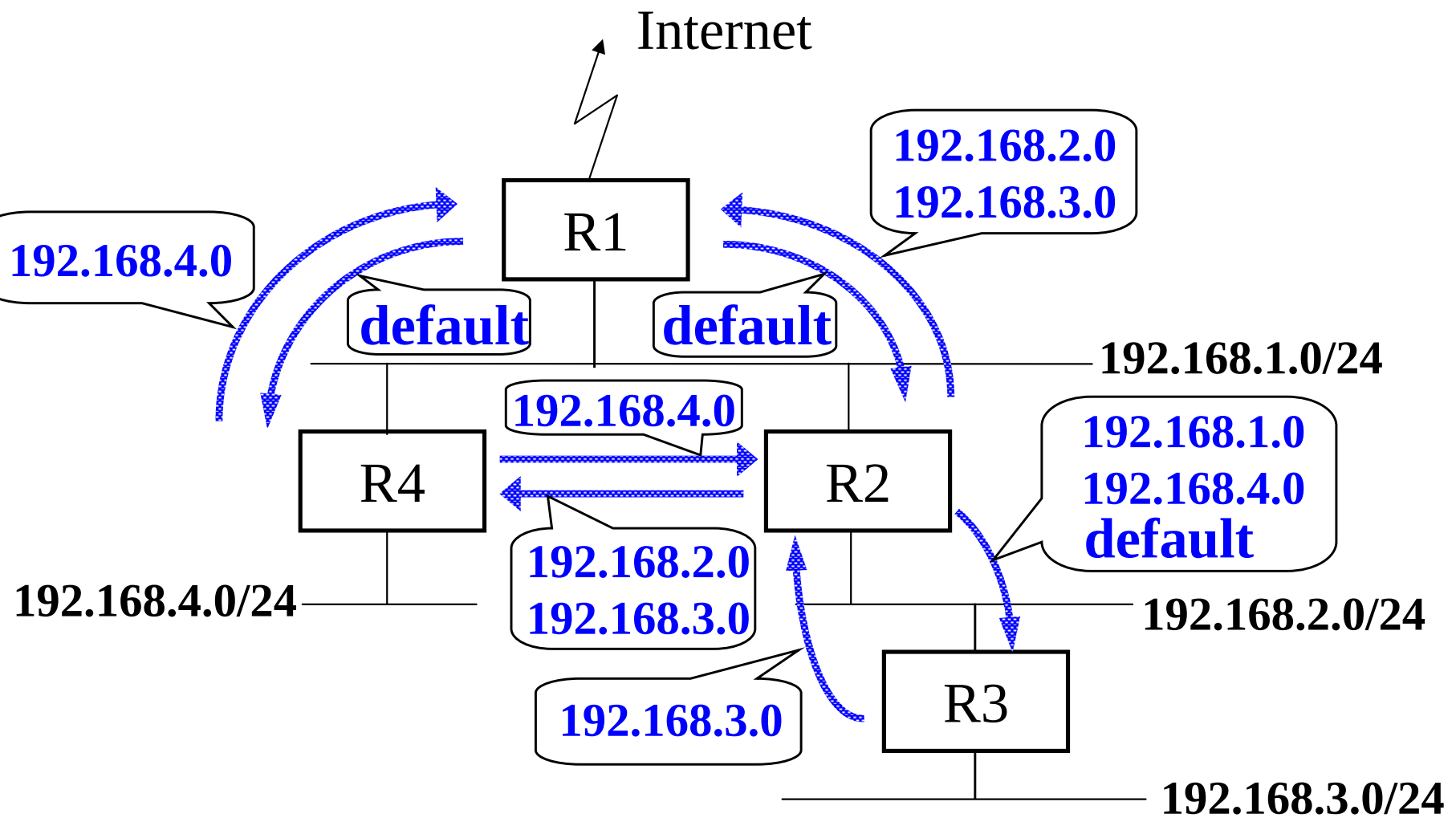


RIPで得られた経路情報は180秒

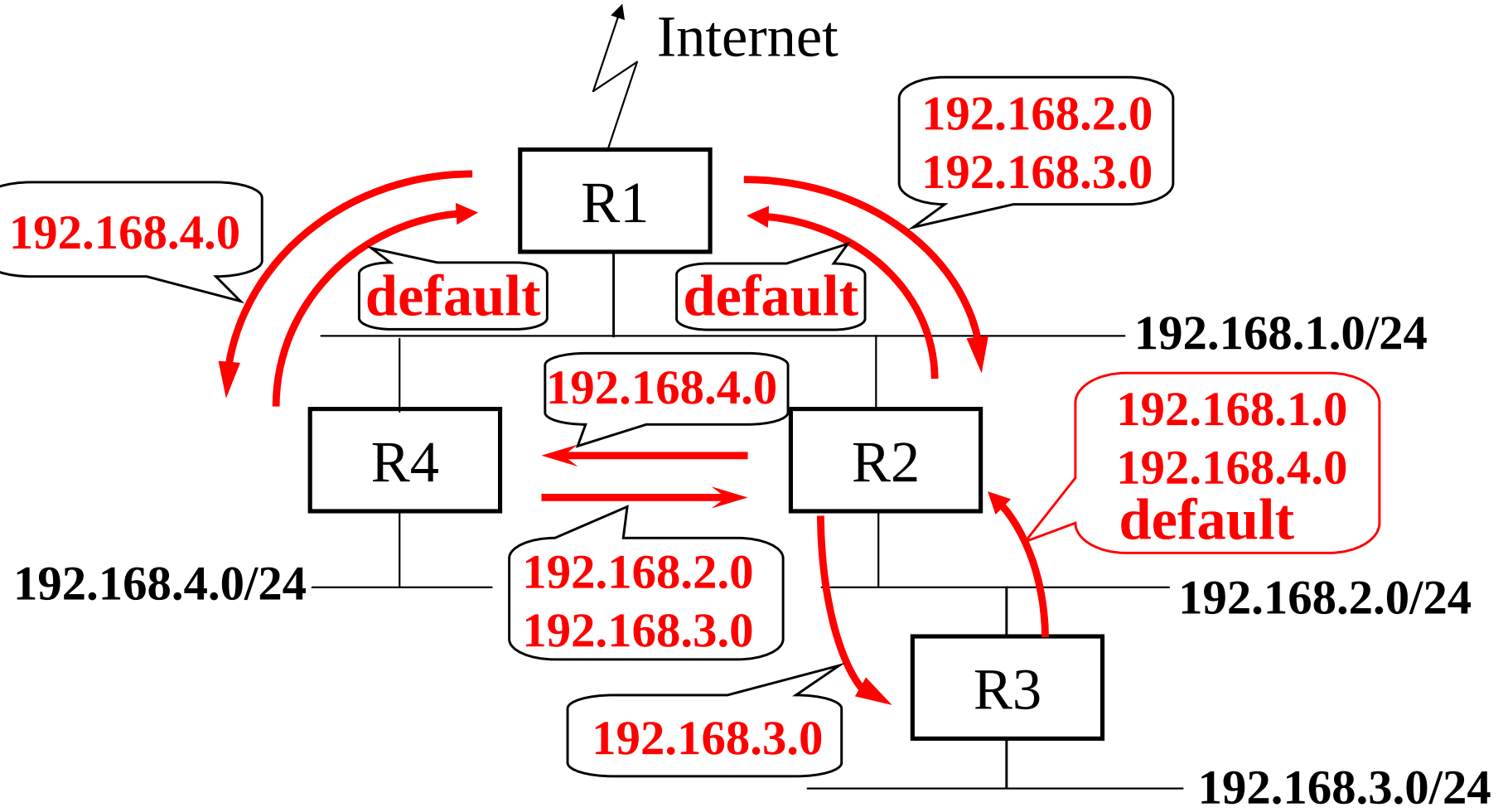
RIPの動作原理-2

- 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


Internet
R1
192.168.2.0
192.168.3.0
192.168.4.0
default
default
192.168.1.0/24
192.168.4.0
R4
R2
192.168.1.0
192.168.4.0
default
192.168.2.0
192.168.3.0
192.168.4.0/24
192.168.2.0/24
192.168.3.0
R3
192.168.3.0/24

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

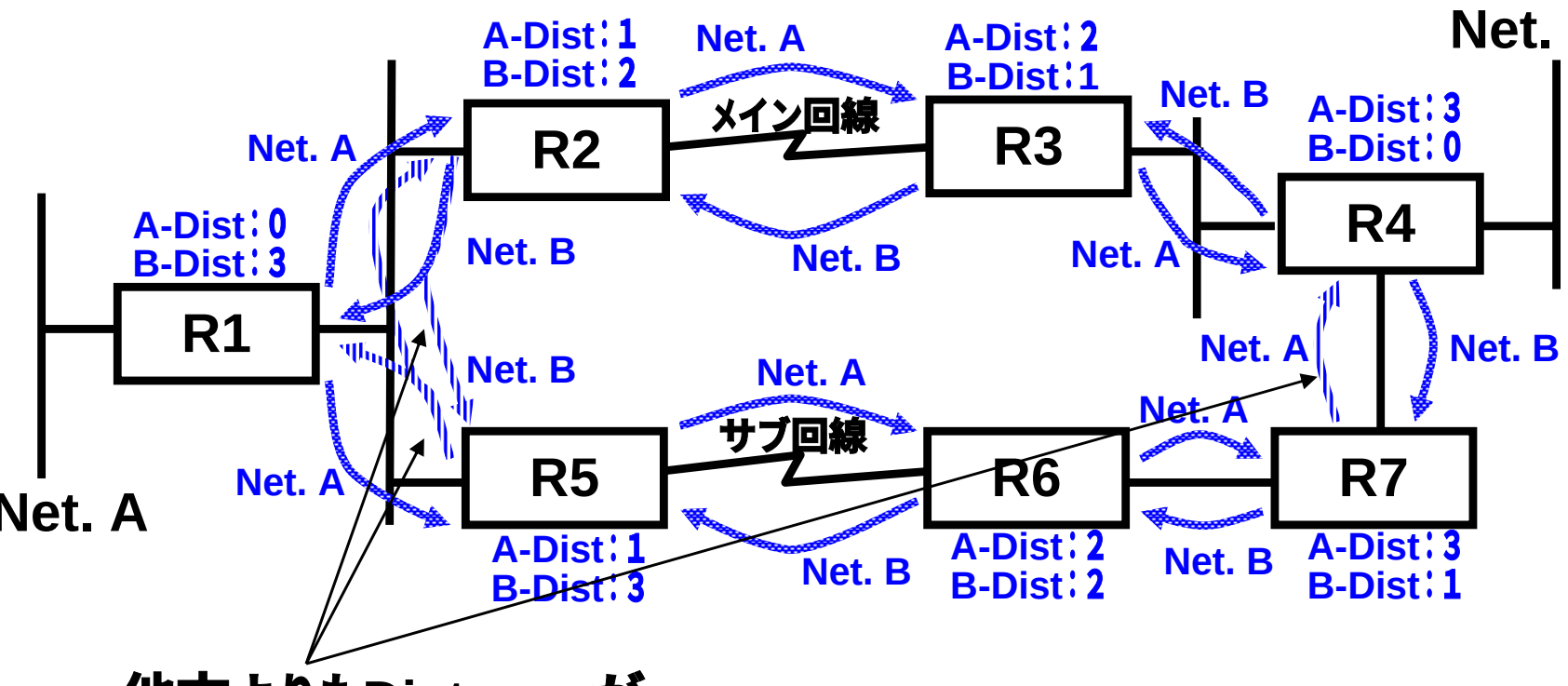
VLSM(Variable Length Subnet Mask)

- 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

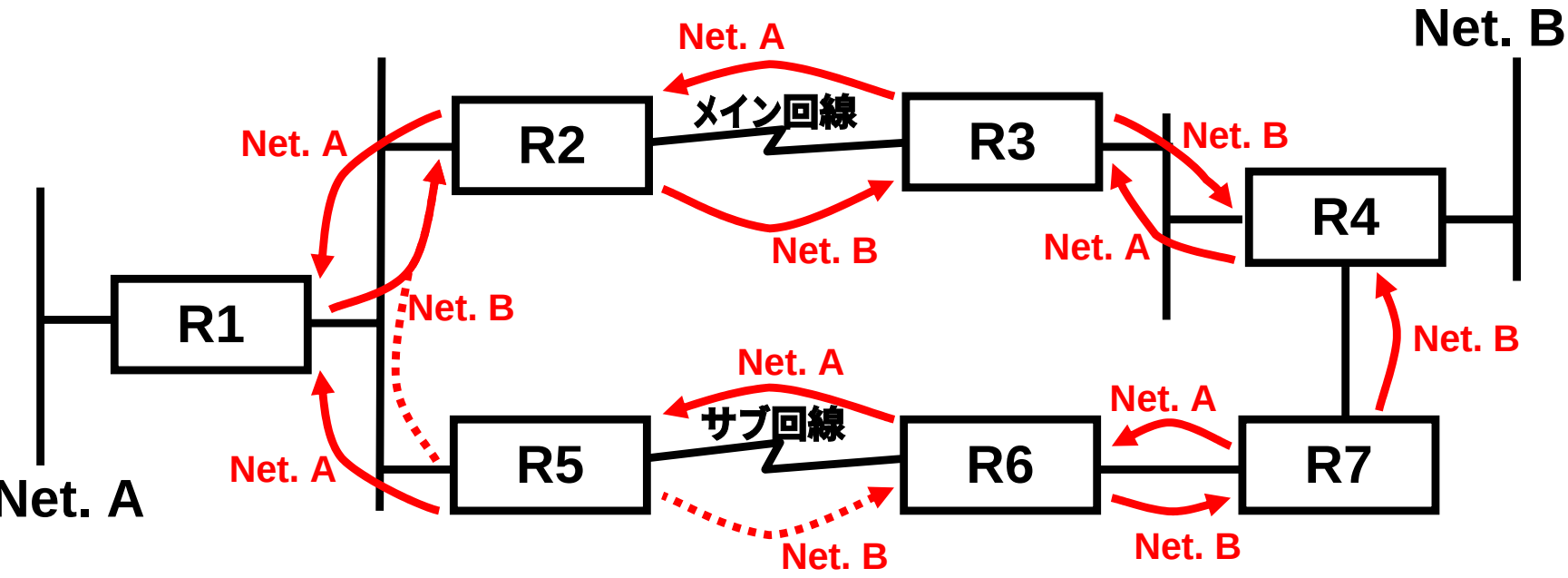
RIPを用いたバックアップ-経路の伝播(定常時)



他方よりもDistanceが
大きいため選択されない

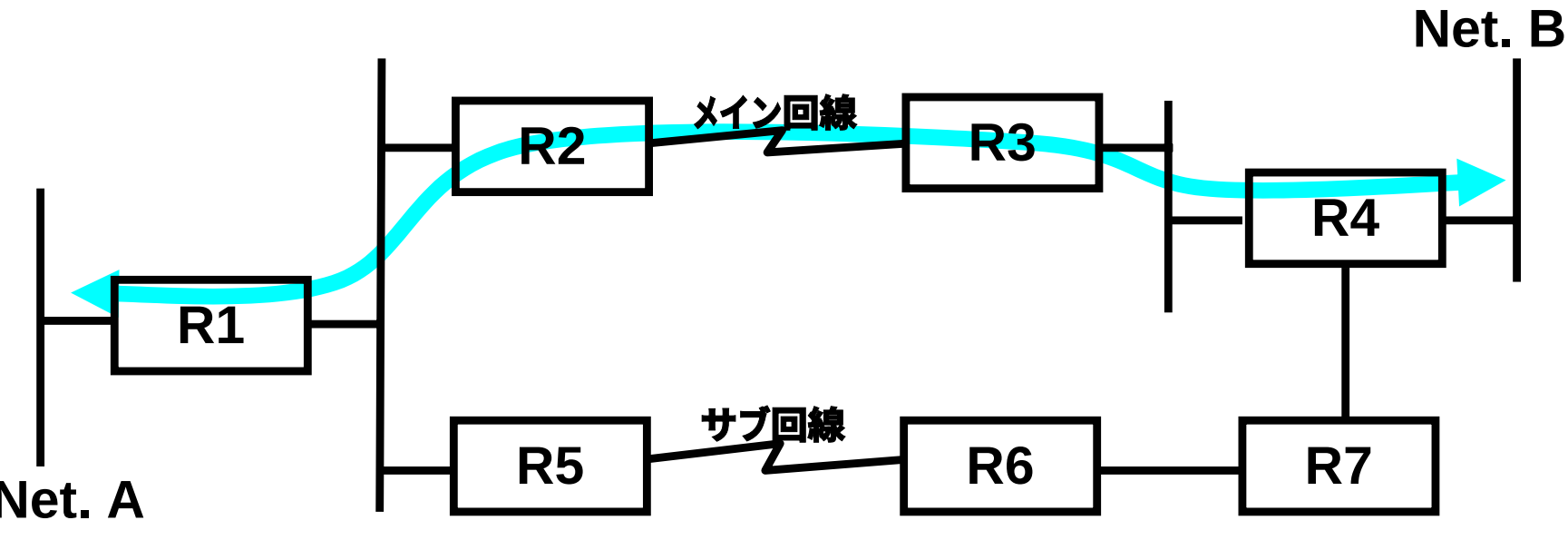
- RIPを利用し、主にバックアップを目的とした構成
- 通常時はメイン回線のみを利用する

RIPを用いたバックアップ -ルーティングテーブル(定常時)



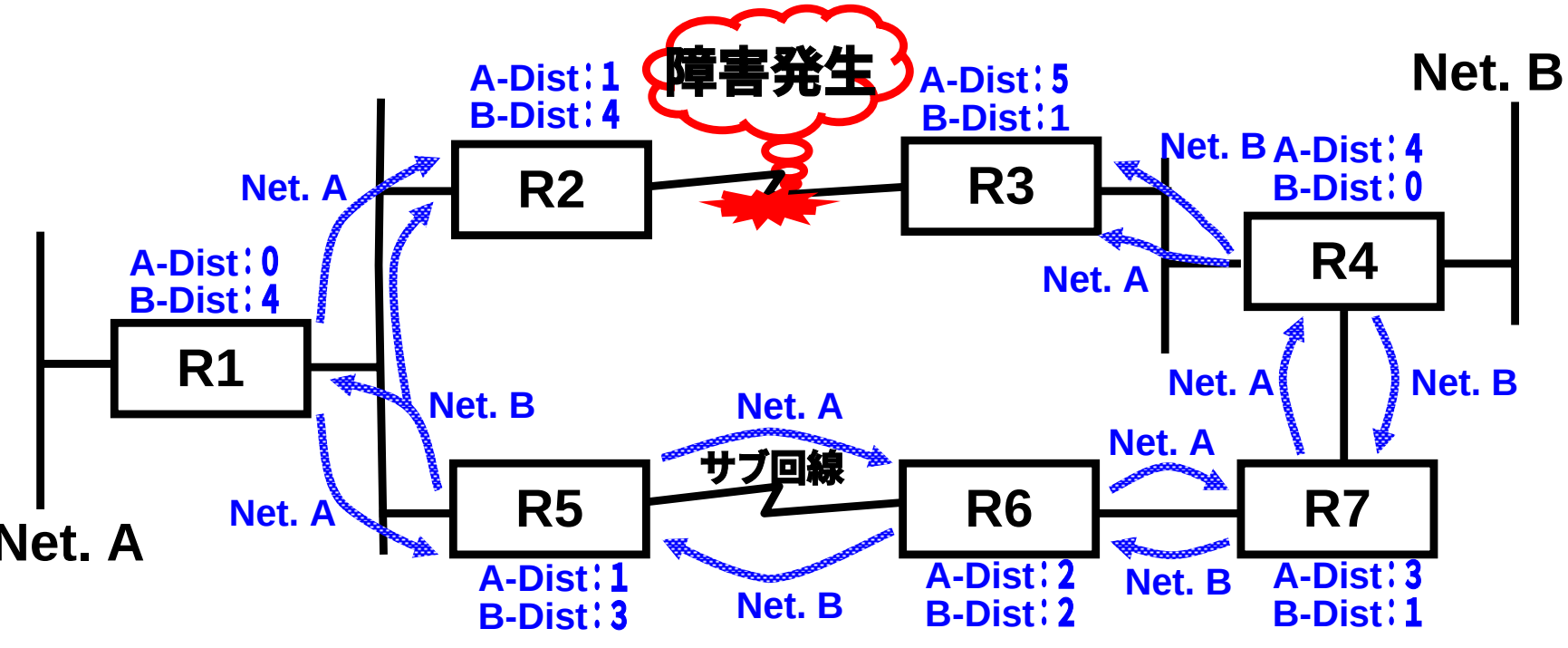
- RIPの経路情報が伝播することにより、各ルータに経路情報が設定される
- Distanceの違いから、メイン回線側の経路が選択される

RIPを用いたバックアップ-トラフィックの流れ（定常時）



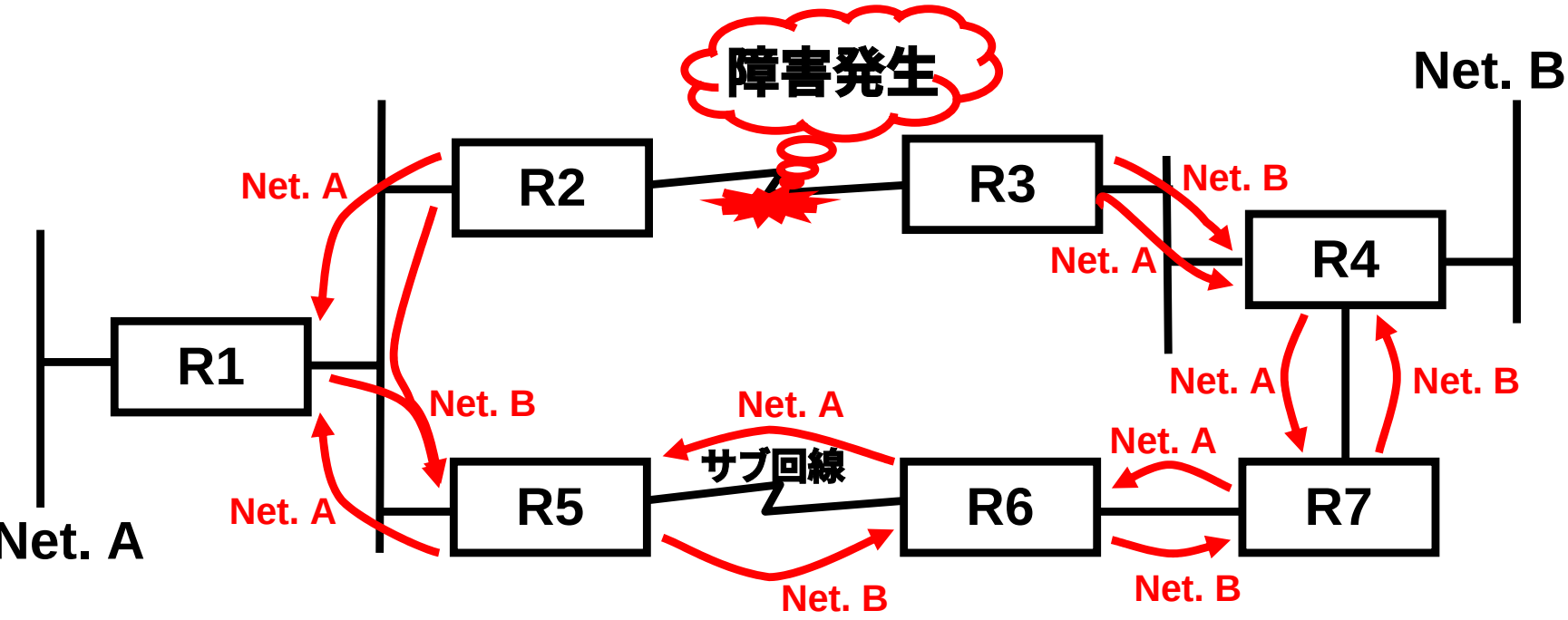
- 通常時はメイン回線のみが利用される

RIPを用いたバックアップ-経路の伝播(障害時)



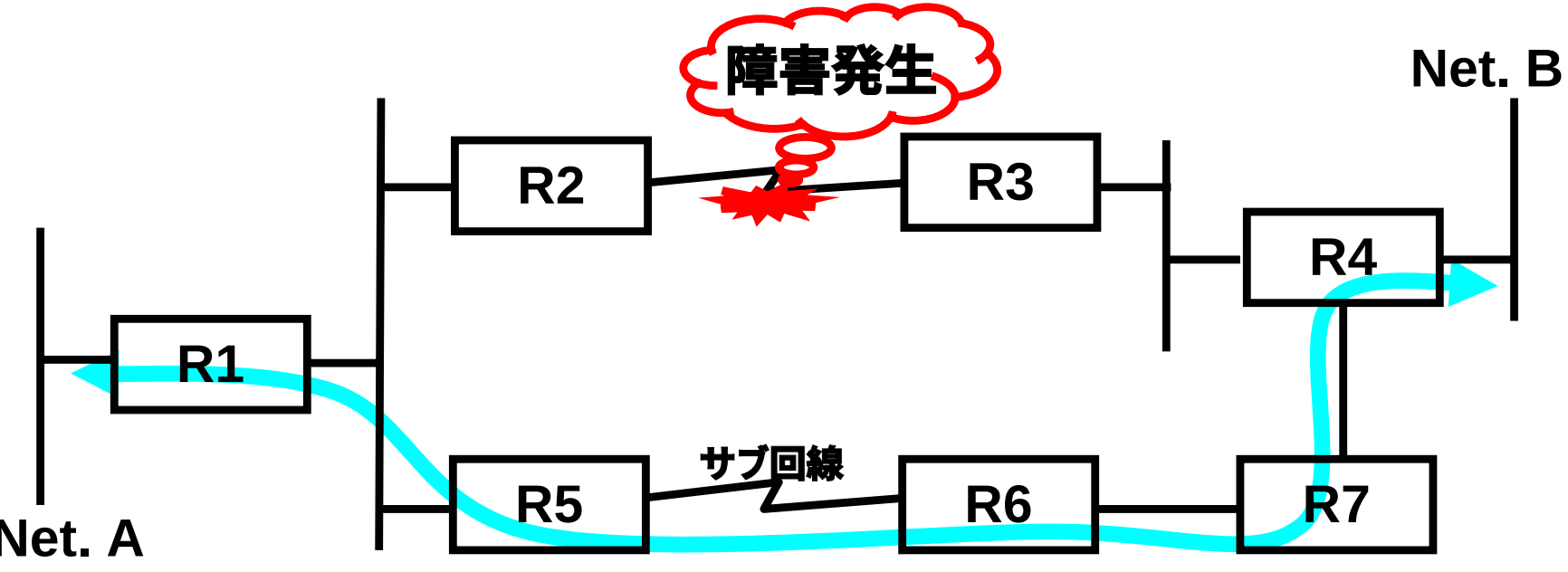
- メイン回線に障害が発生したため、経路情報の伝播が変化する

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解説ー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。

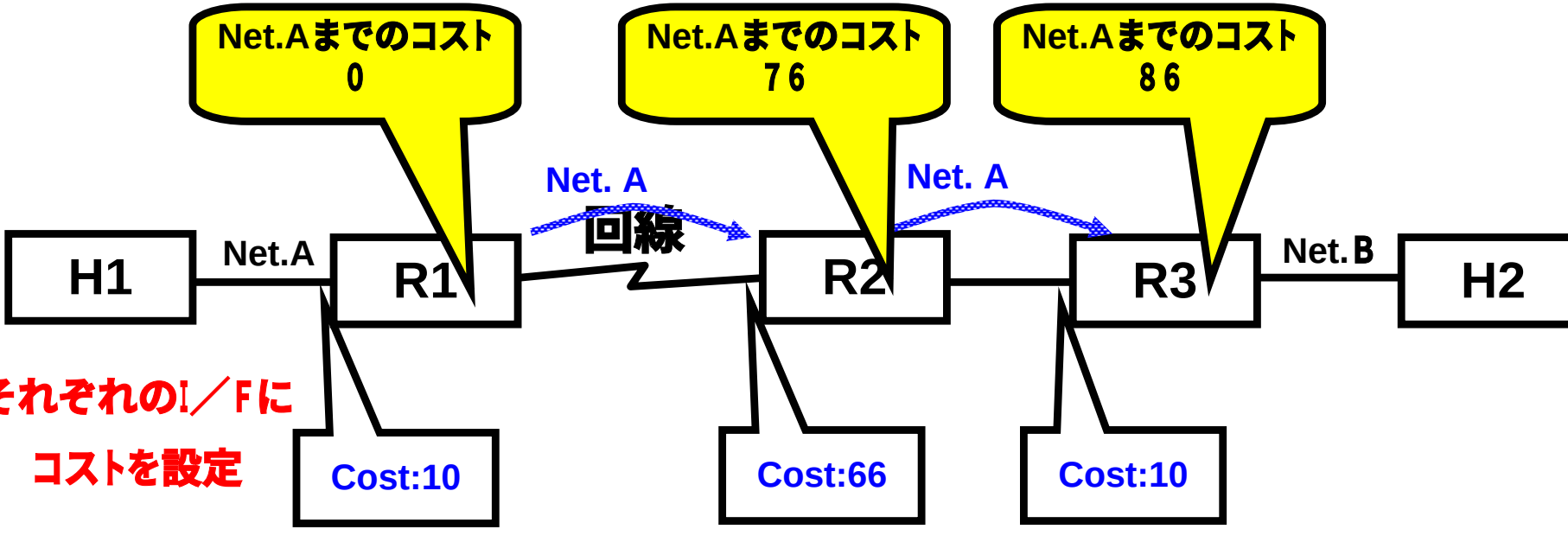
OSPF解説ー2

- 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

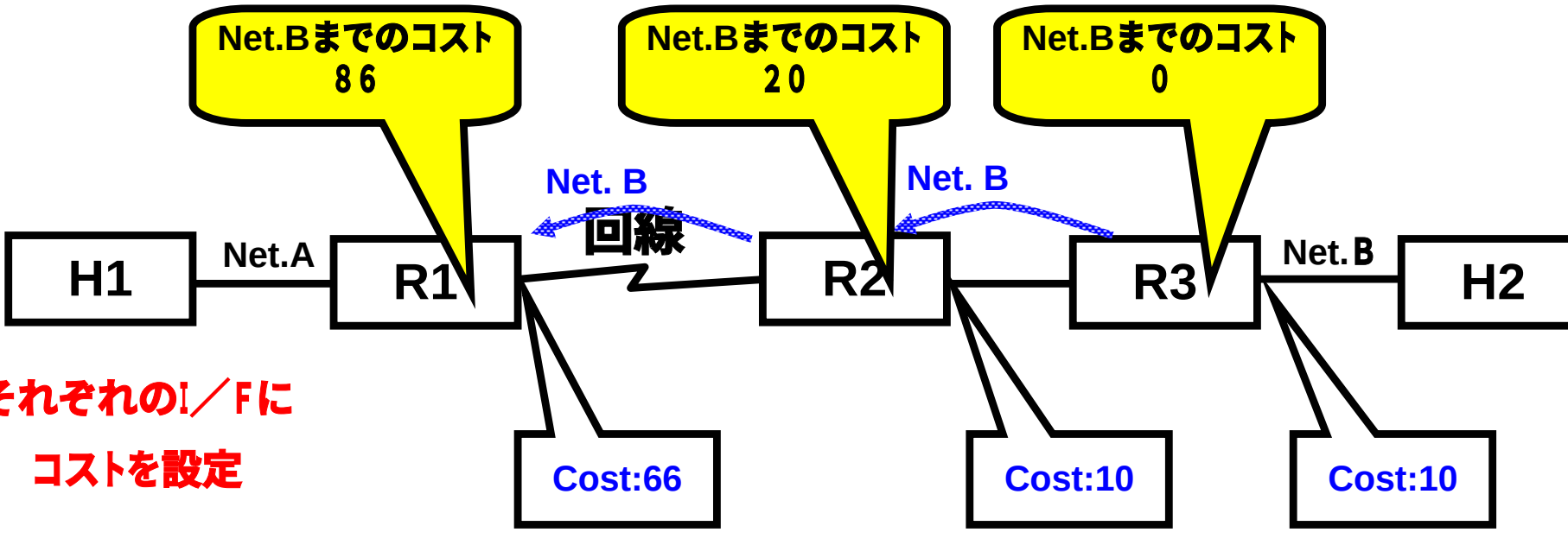
簡単なOSPFコストの計算法－1



- 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

IIJ
Internet Initiative Jap

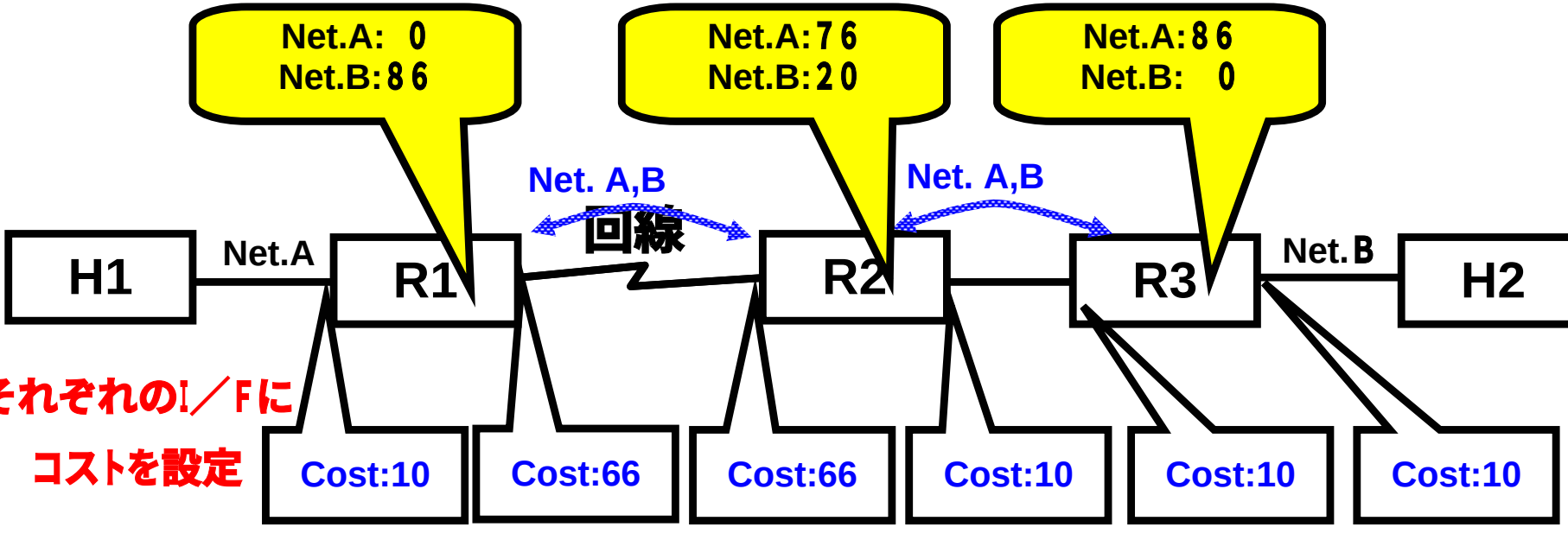
簡単なOSPFコストの計算法－2



- 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

IIJ
Internet Initiative Jap

簡単なOSPFコストの計算法－3

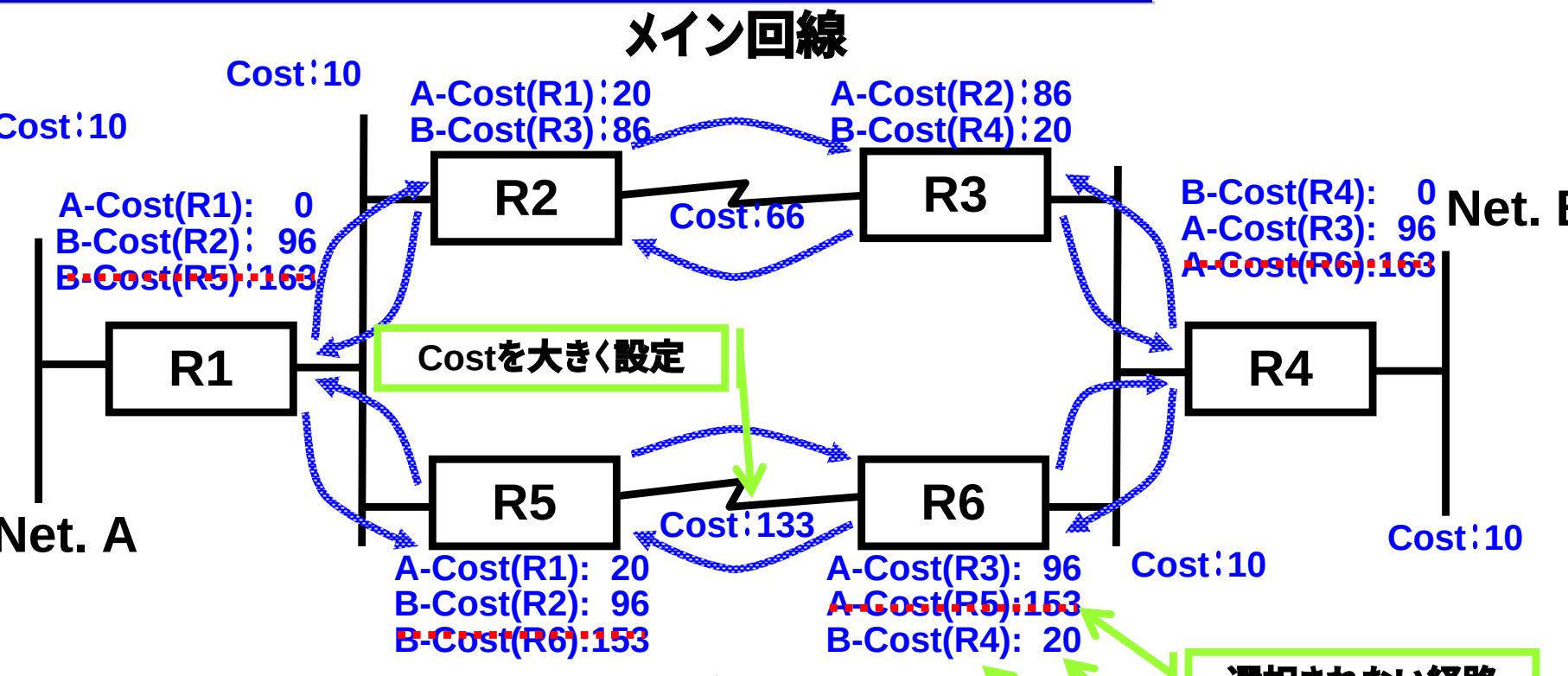


- 同じI/Fに同じコストを付けることにより、行きと帰りのコストを一致させることができる
- 行きと帰りで異なるコストを付与することもできるが、管理が煩雑になるため、理由なく行なうべきではない
- ここで示した図は経路を交換しているように書かれているが、実際はトポロジデータベースの交換により経路を確定している

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

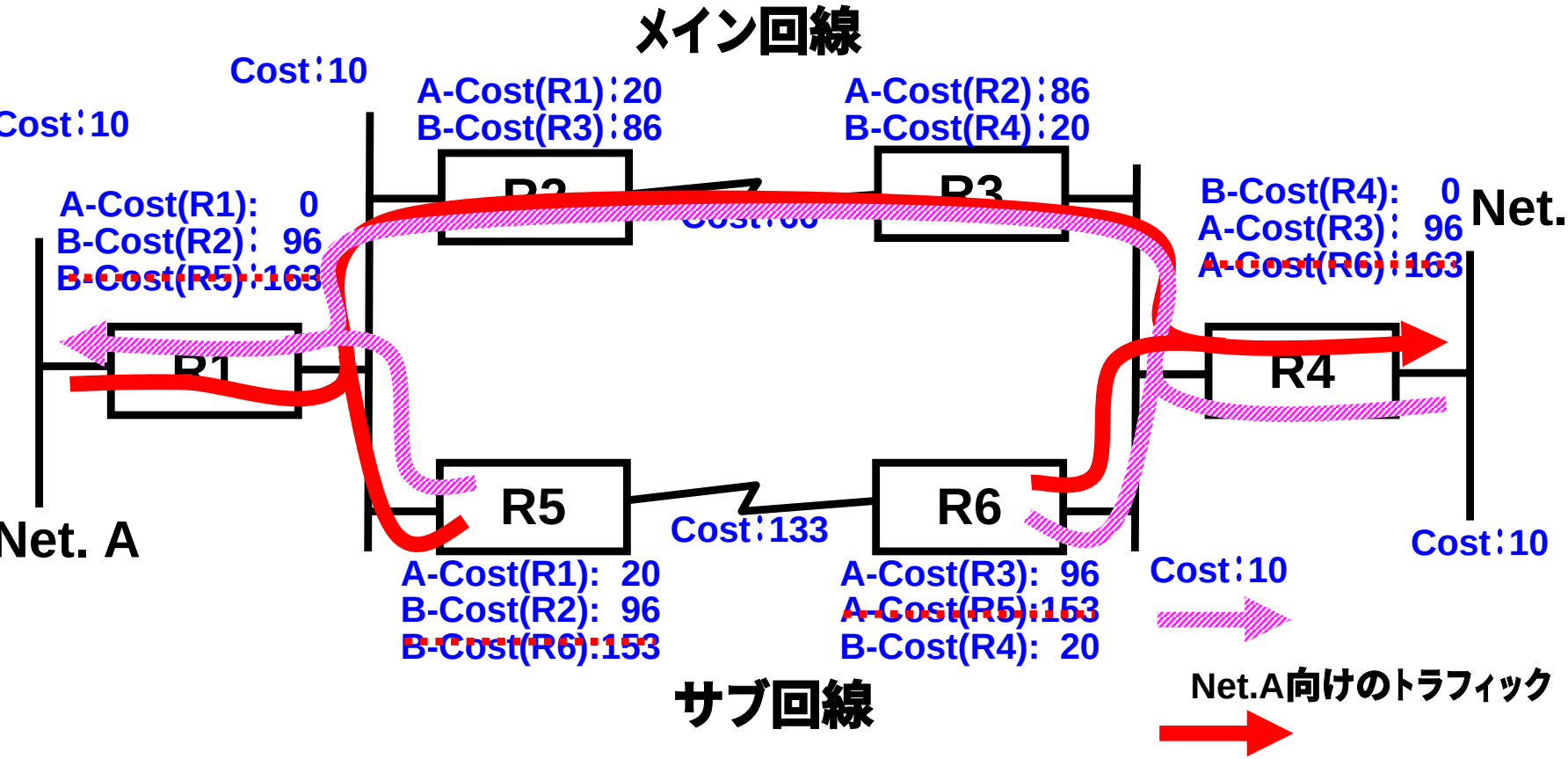
選択されない経路

コスト値

伝播元ルータ名
(NEXT HOP)

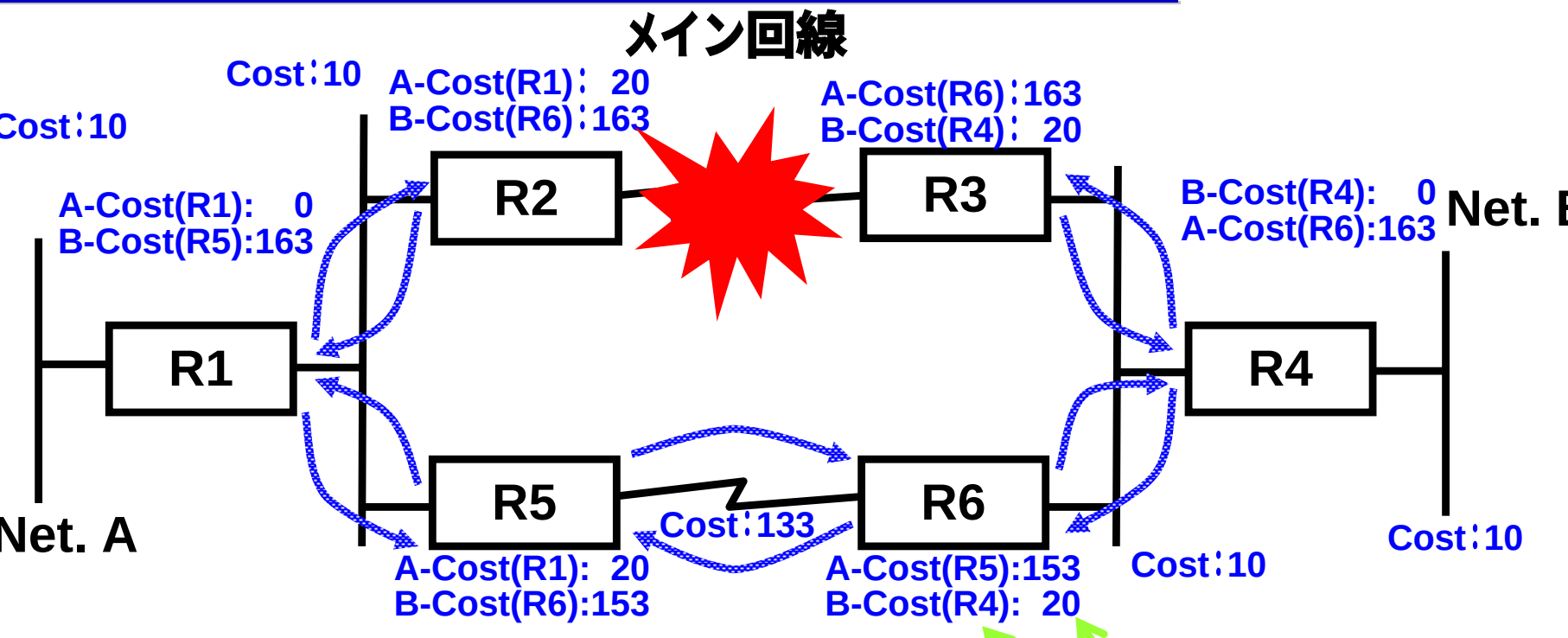
IIJ
Internet Initiative Jap

OSPFを用いたバックアップ-トラフィックの流れ(通常時)



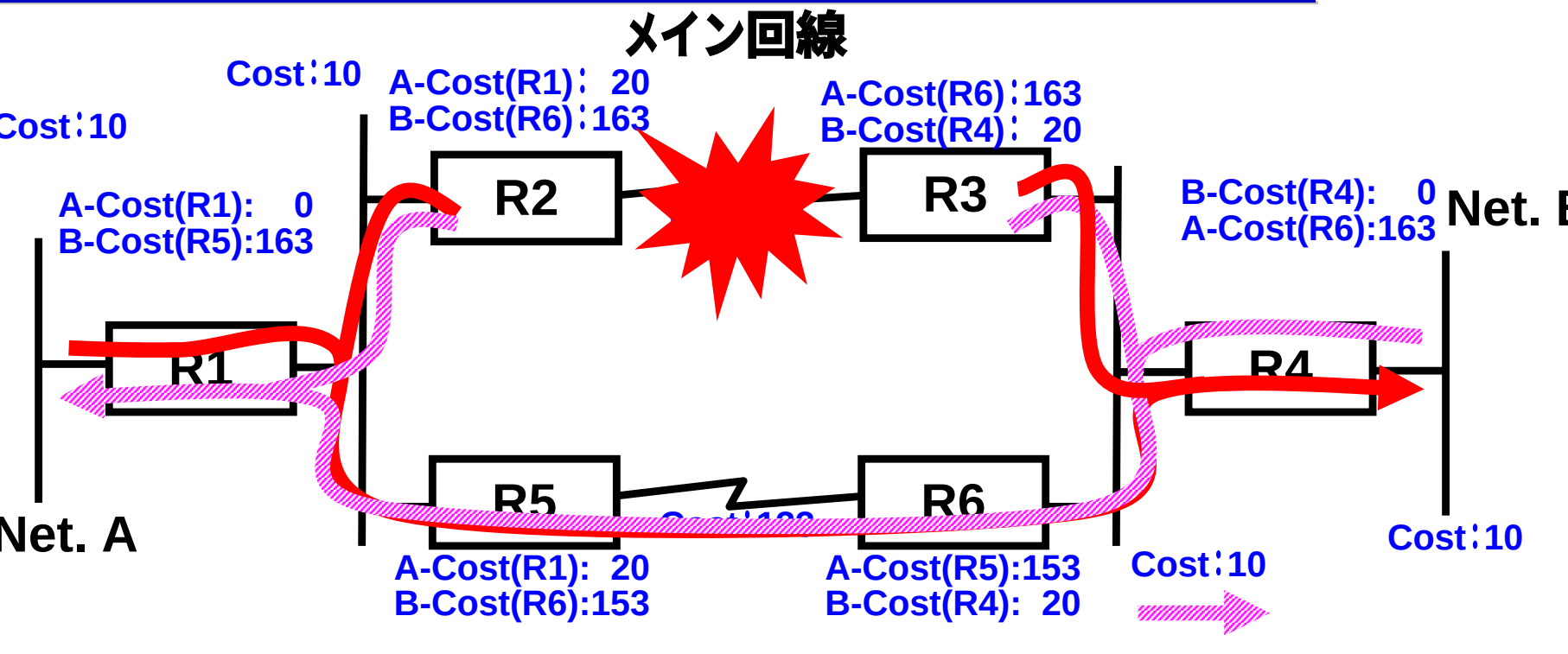
- サブ回線にもOSPF HELLOパケットが流れるため、トラフィックをゼロにはできない

OSPFを用いたバックアップ一経路の伝播（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）

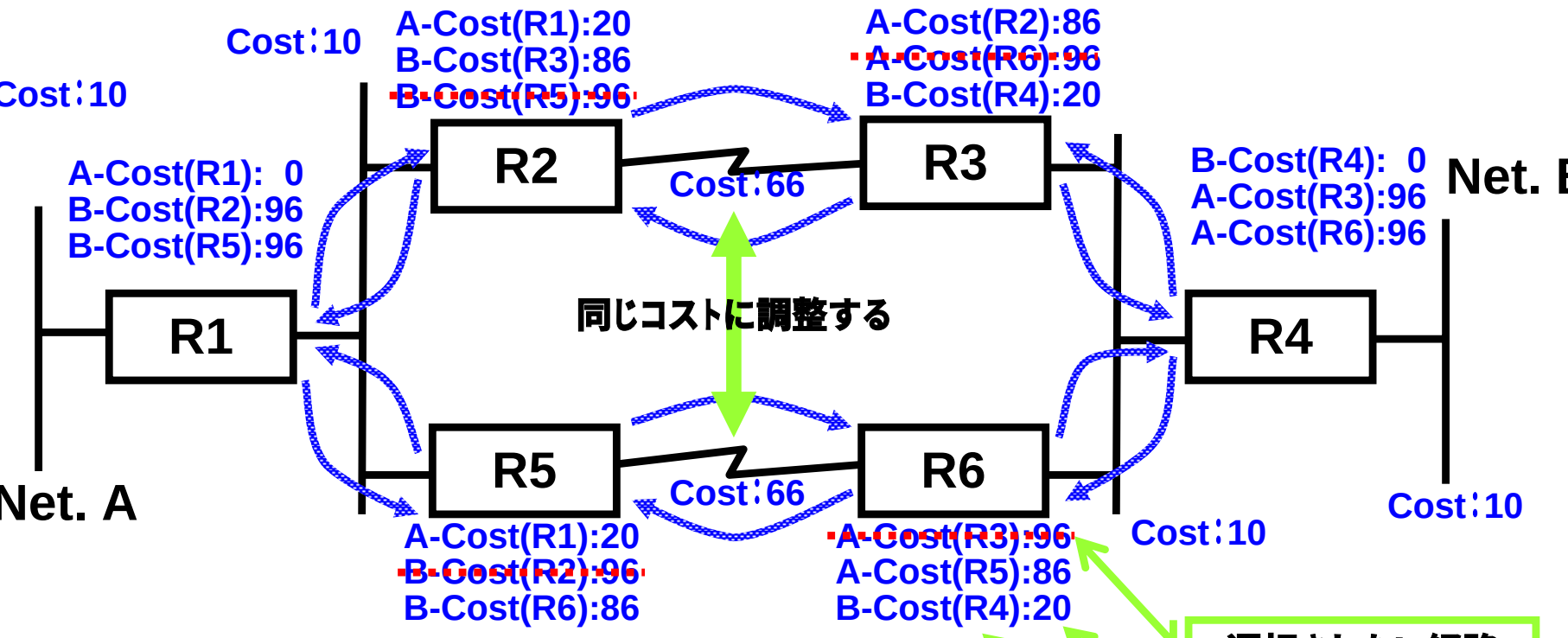


- サブ回線を利用したバックアップが完了

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）



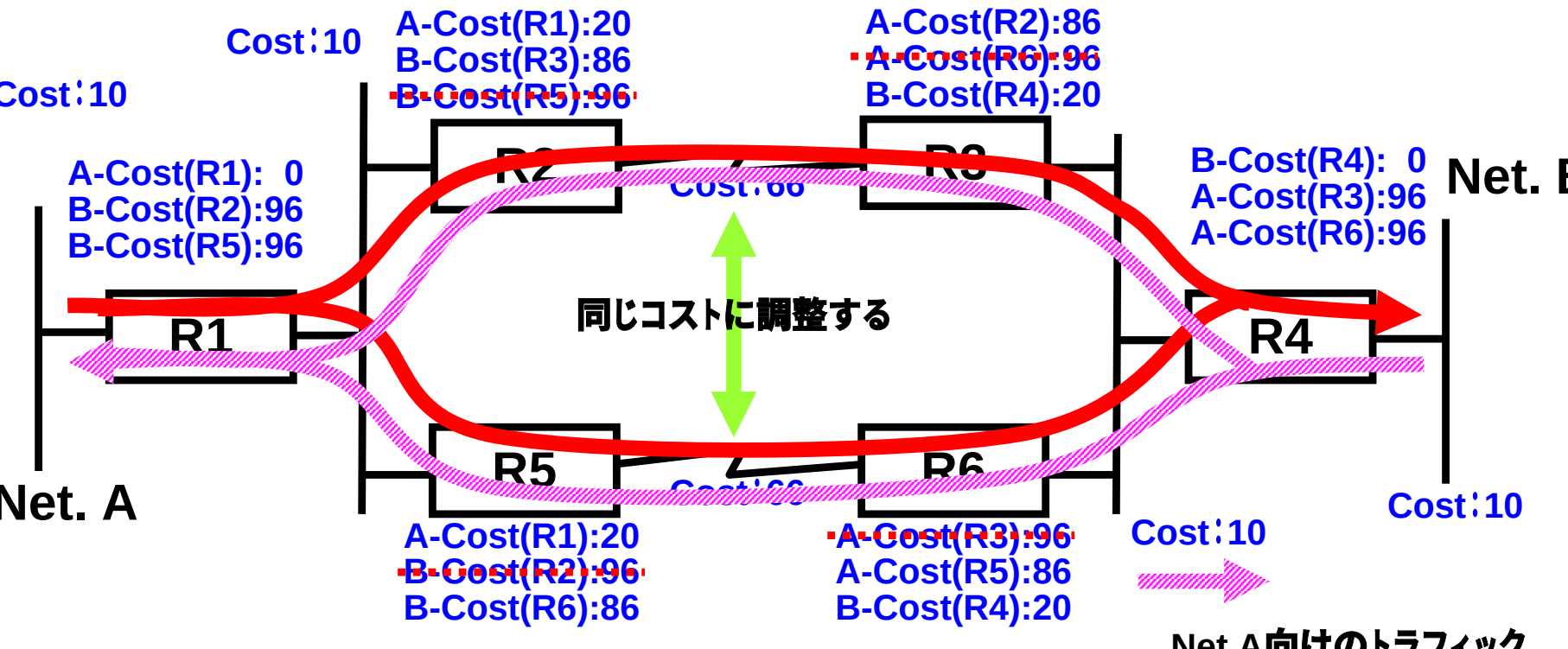
選択されない経路

コスト値

伝播元ルータ名
（NEXT HOP）

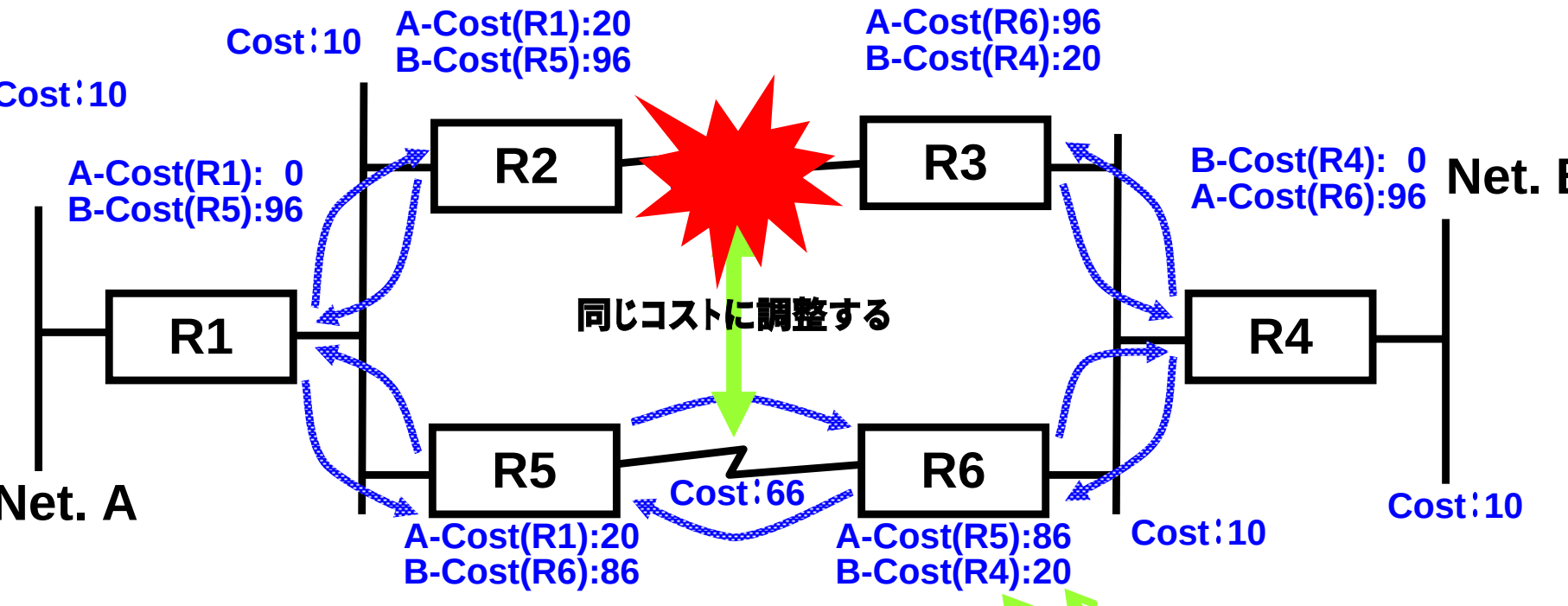
- 2本の回線を同じコストに設定する
- R1からNet.Bに対してR2,R5両方とも同じコストにする
- R4からNet.Aに対してR3,R6両方とも同じコストにする

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）

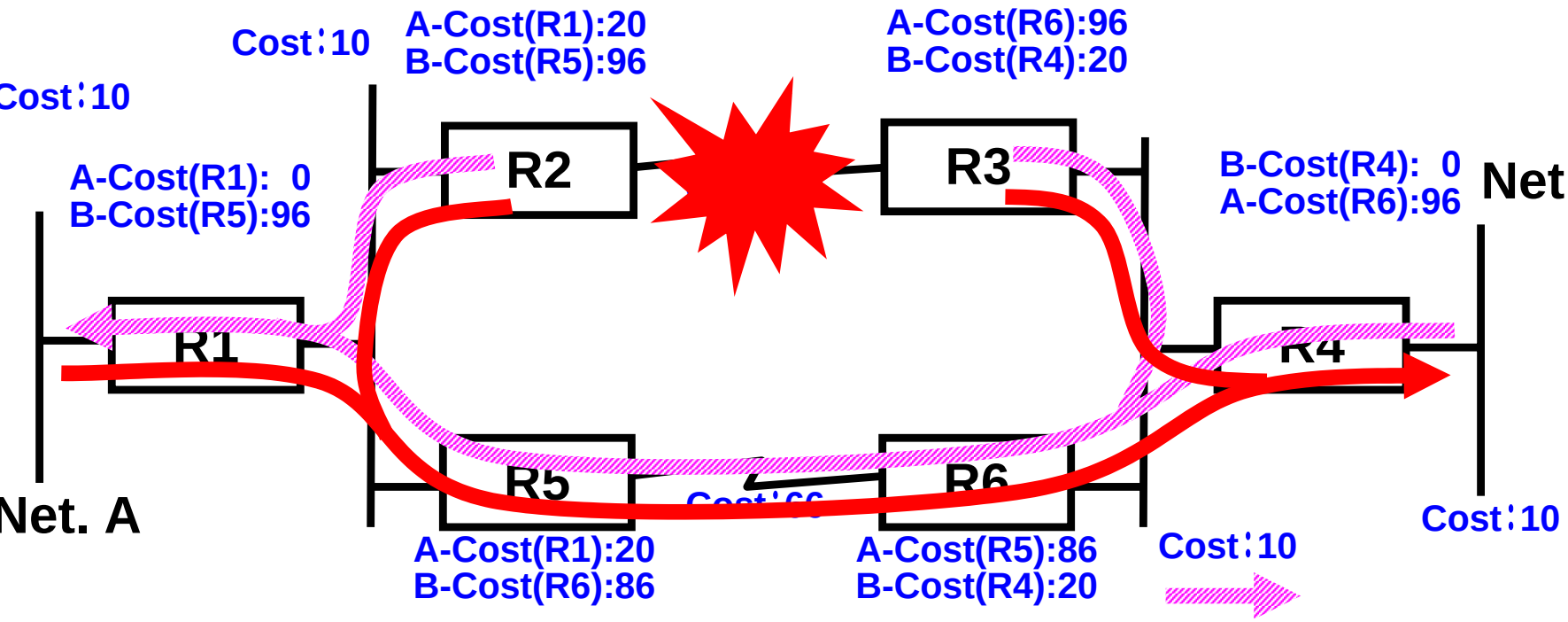


- 障害発生により、R2-R3間のネットワーク情報が削除される

コスト値

伝播元ルータ名
（NEXT HOP）

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

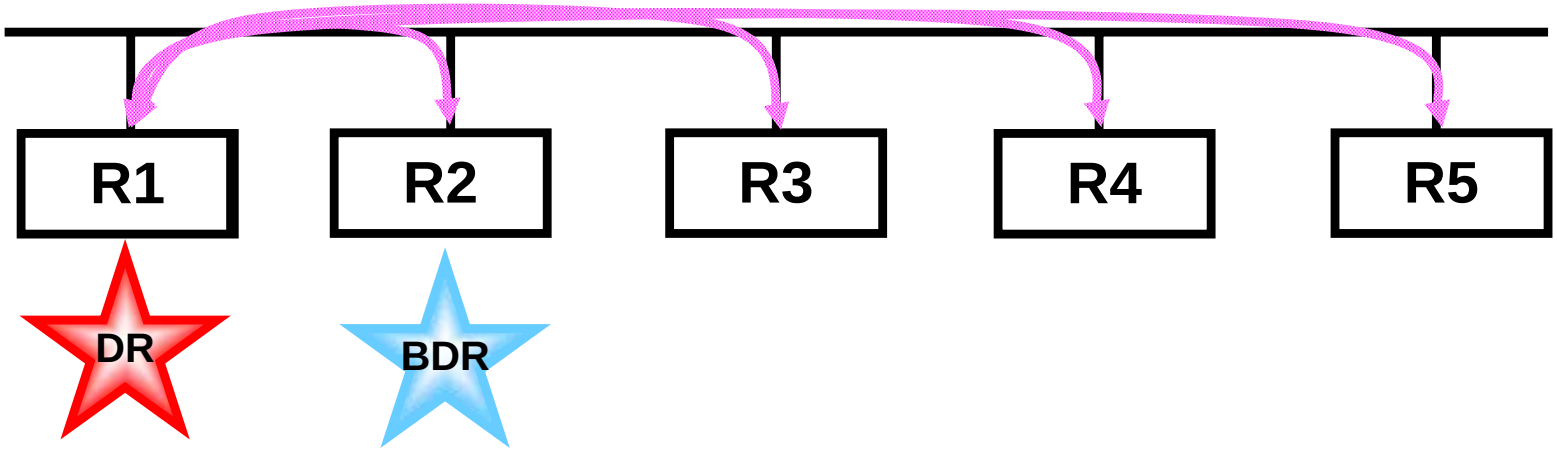
初心者のためのOSPF設定－2

- 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は注入された個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設定が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。

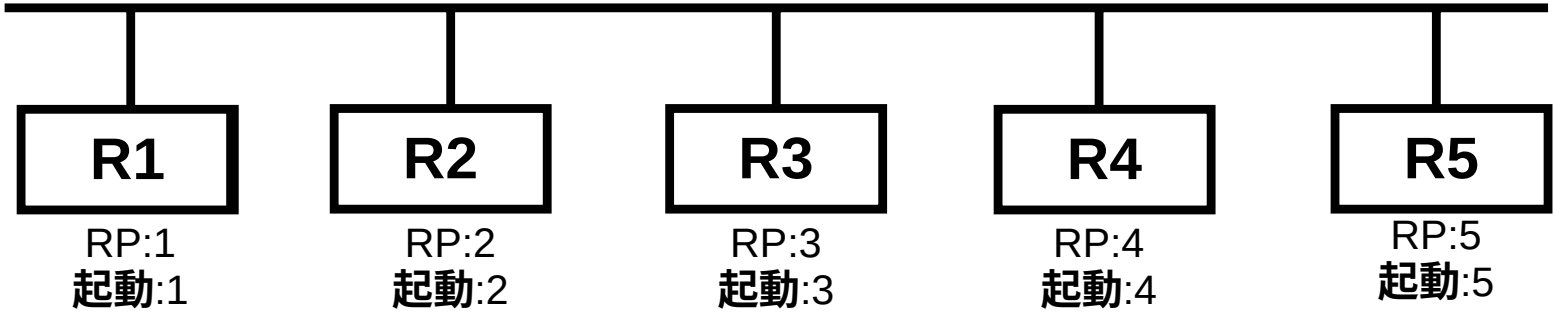
DRとBDRの役割



- 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

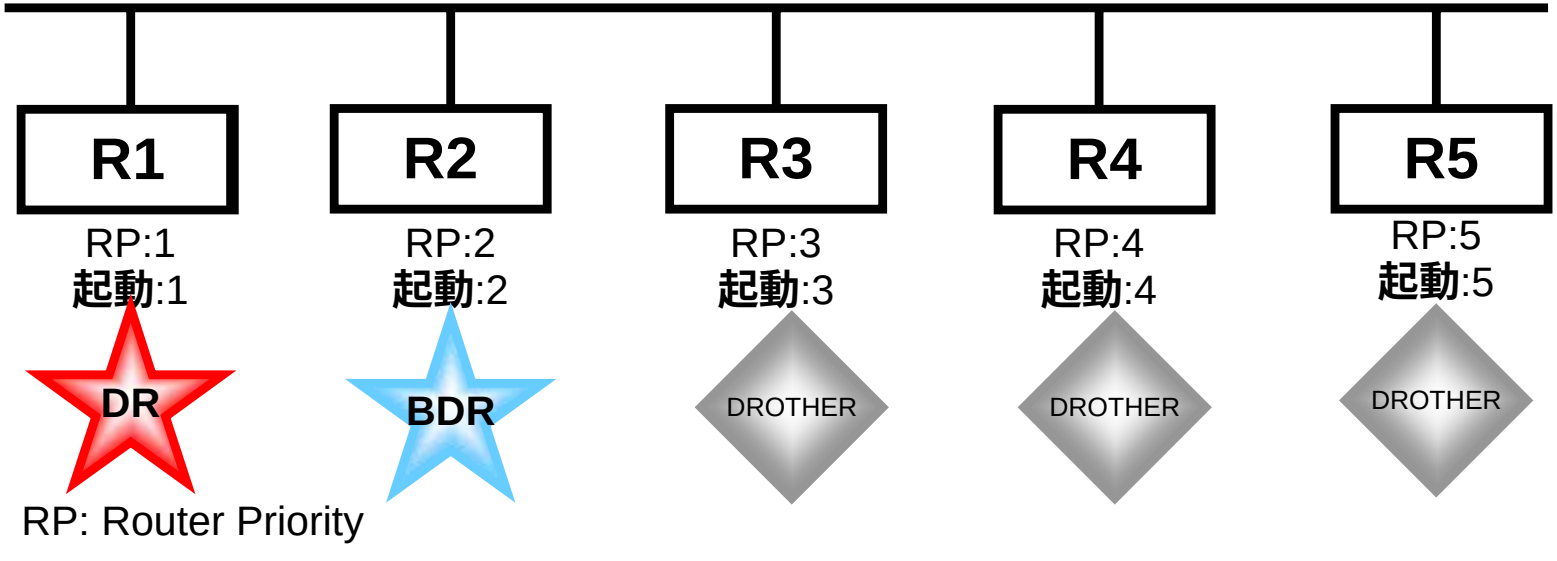
DRとBDRの選出-1



RP: Router Priority

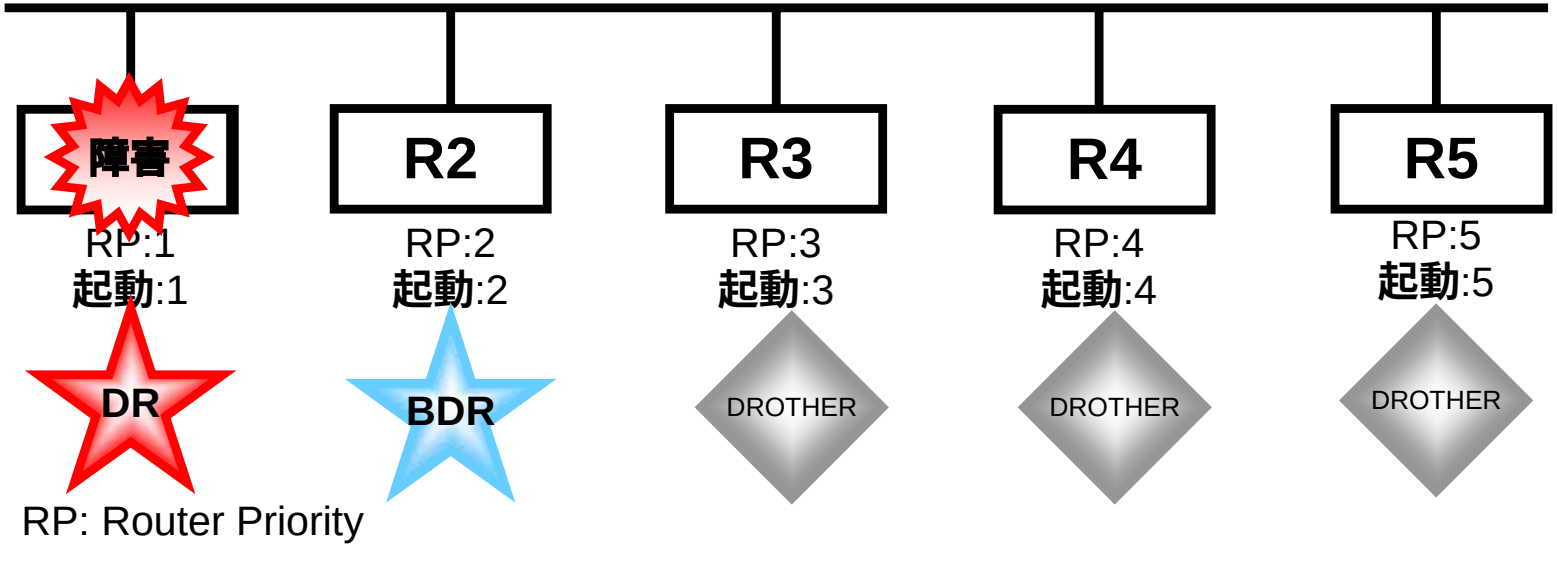
- DRとBDRの選択の方法
  - 最初に起動したルータはDRとなる
  - 2番目に起動したルータはBDRとなる
  - 3番目以降に起動したルータはDROTHERとなる
- 上記ネットワークにおけるDR、BDRはどれか？

DRとBDRの選出-2



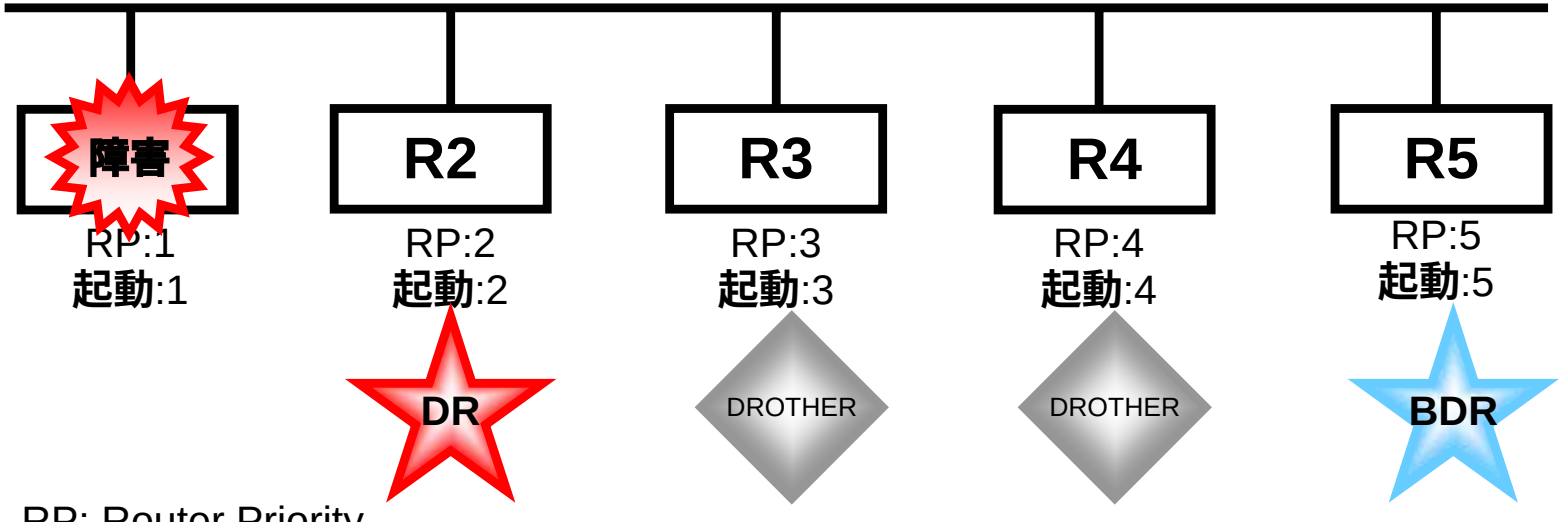
- DRとBDRの選択の結果
  - 最初に起動したR1はDRとなる
  - 2番目に起動したR2はBDRとなる
  - 3番目以降に起動したR3～R5はDROTHERとなる

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？

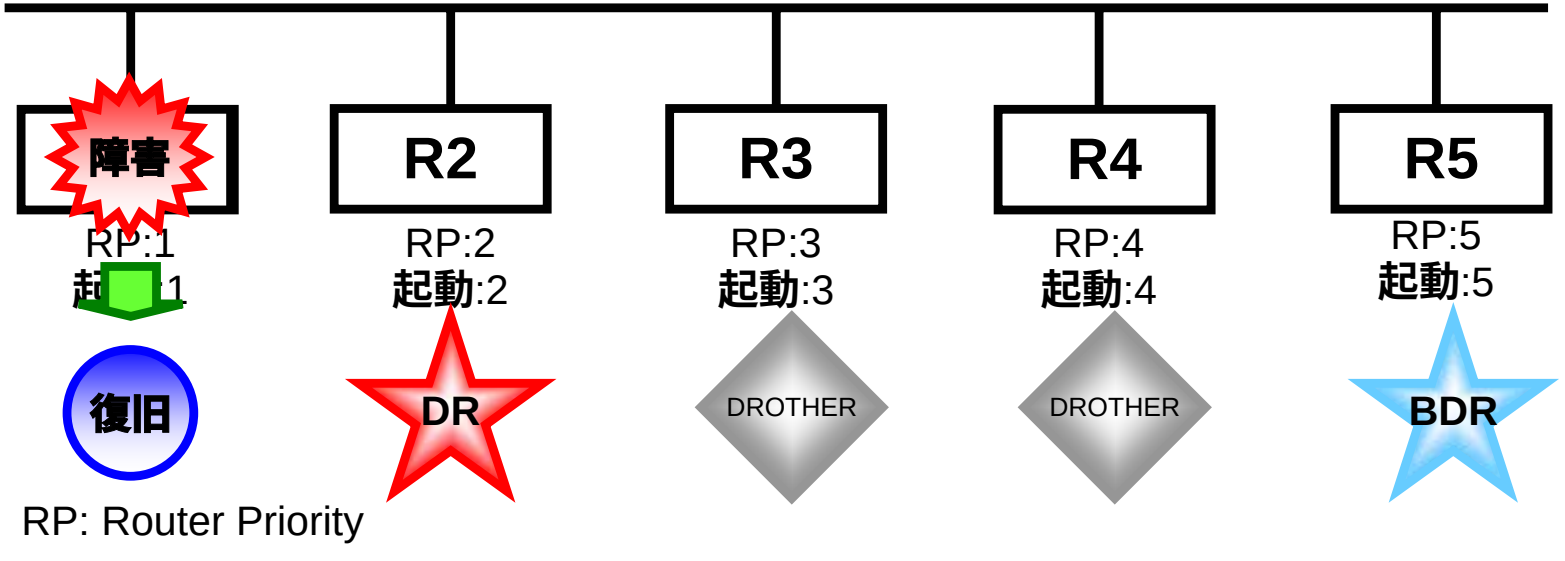
DRとBDRの選出-4



RP: Router Priority

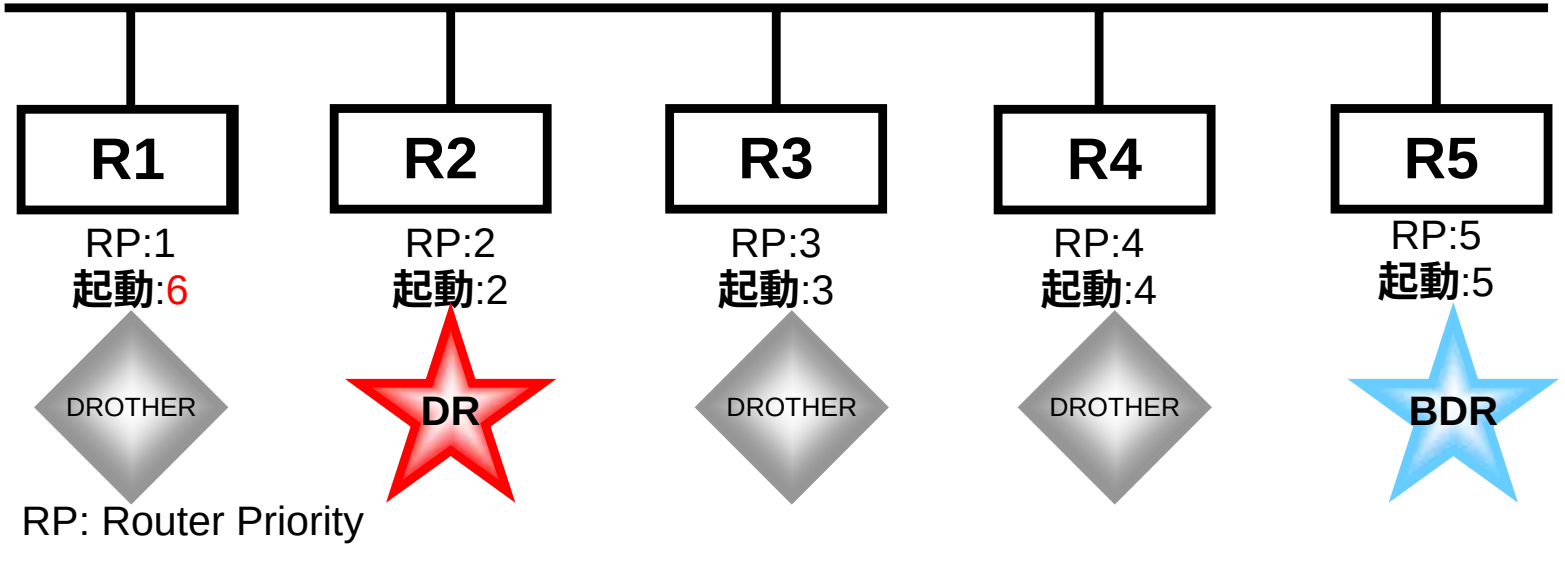
- DRに障害が発生した場合
  - DRに障害が発生するとBDRがDRとなる
  - BDRがDRとなると、新しいBDRがRouter Priorityの大きなルータとなる

DRとBDRの選出-5



- 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



- 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生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

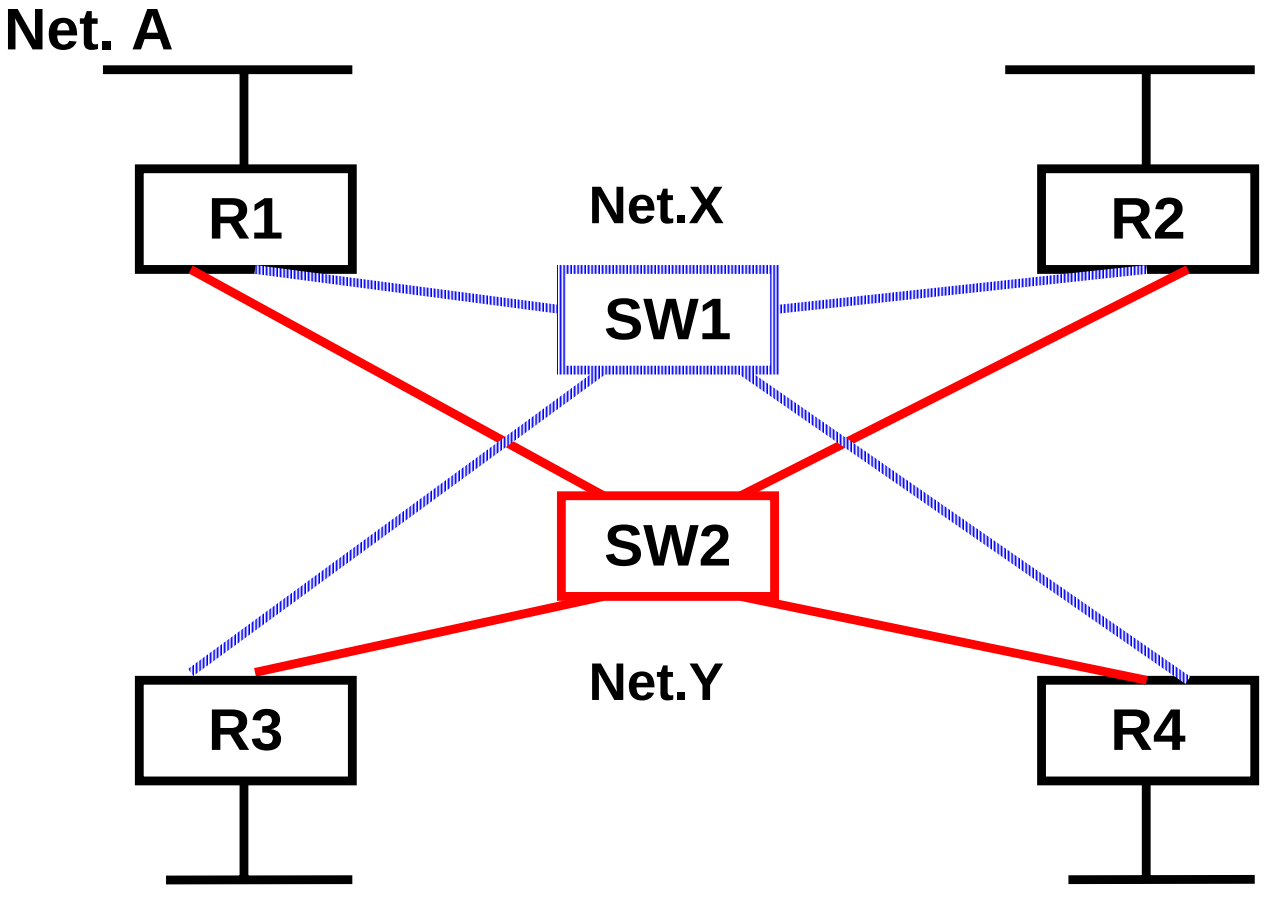
ダイナミックルーティングのまとめ

- 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

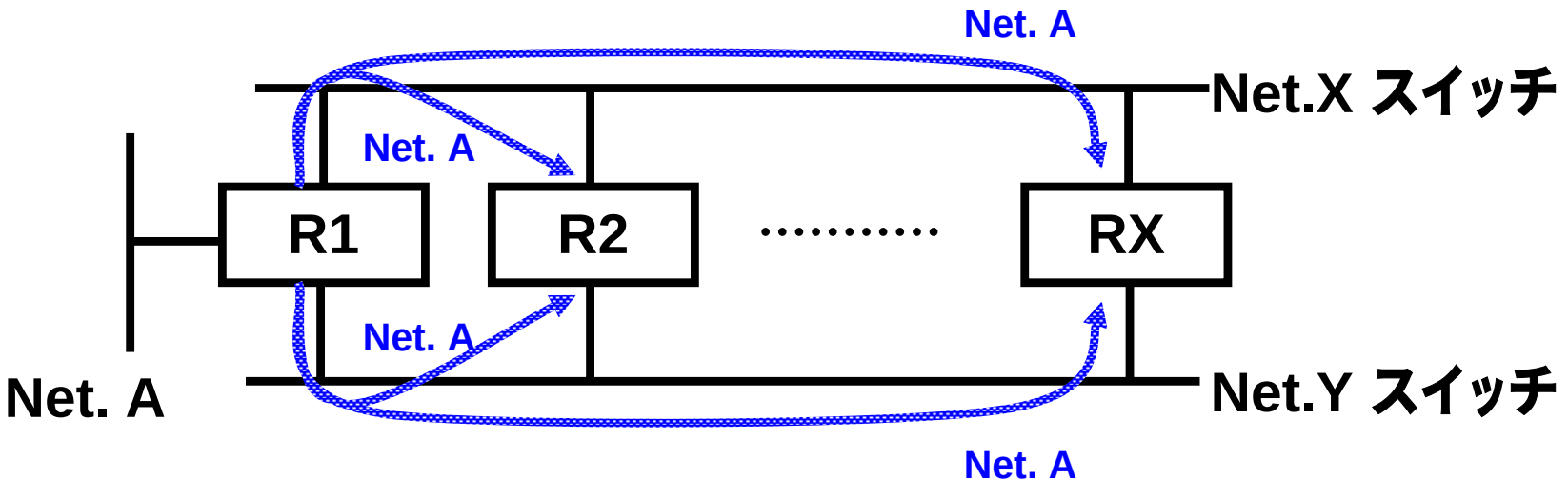
ダイナミックルーティングプロトコルを用いた障害に強いネットワーク構成

- 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

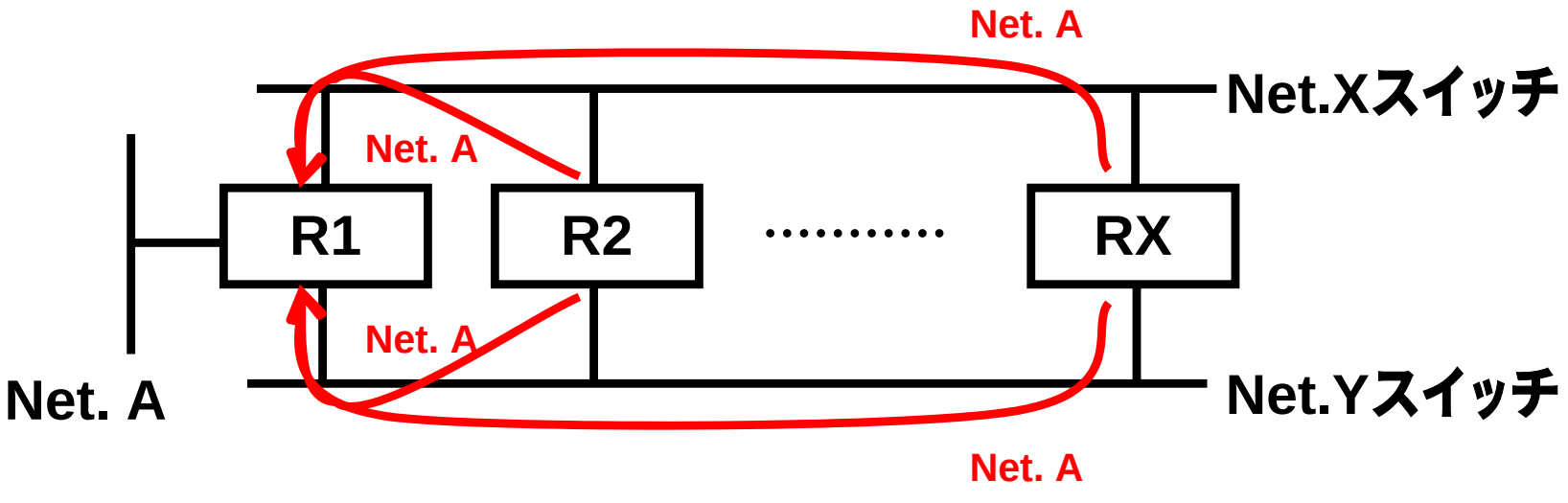


デュアル構成＋OSPFを用いたバックアップ、バランシング-経路の伝播（通常時）



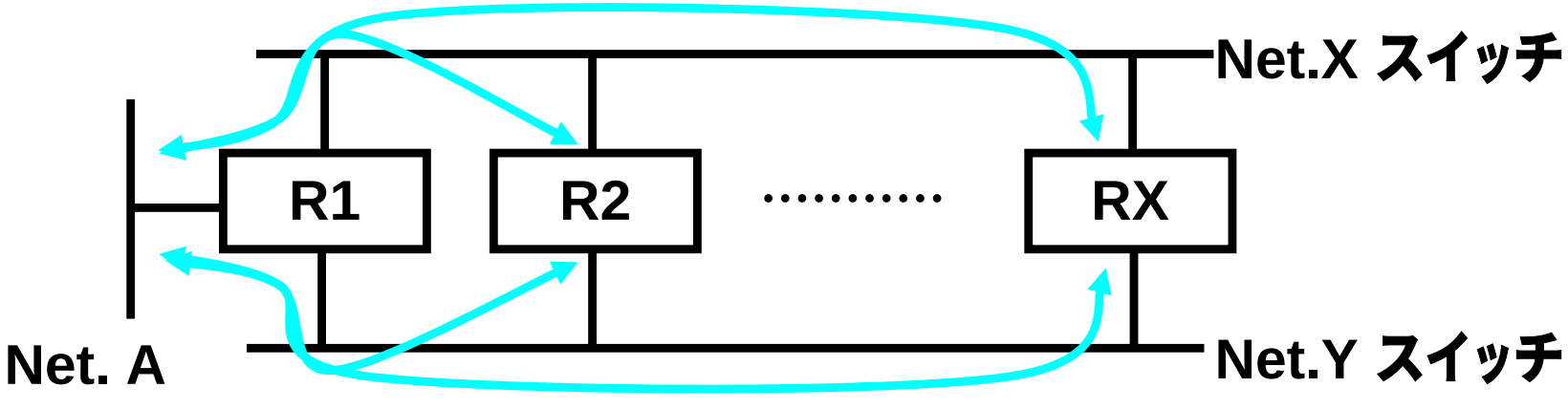
- OSPFで Net.Aの経路情報を広告する
- 経路情報は各ルータに対して、2つのスイッチから等価に伝播する

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）



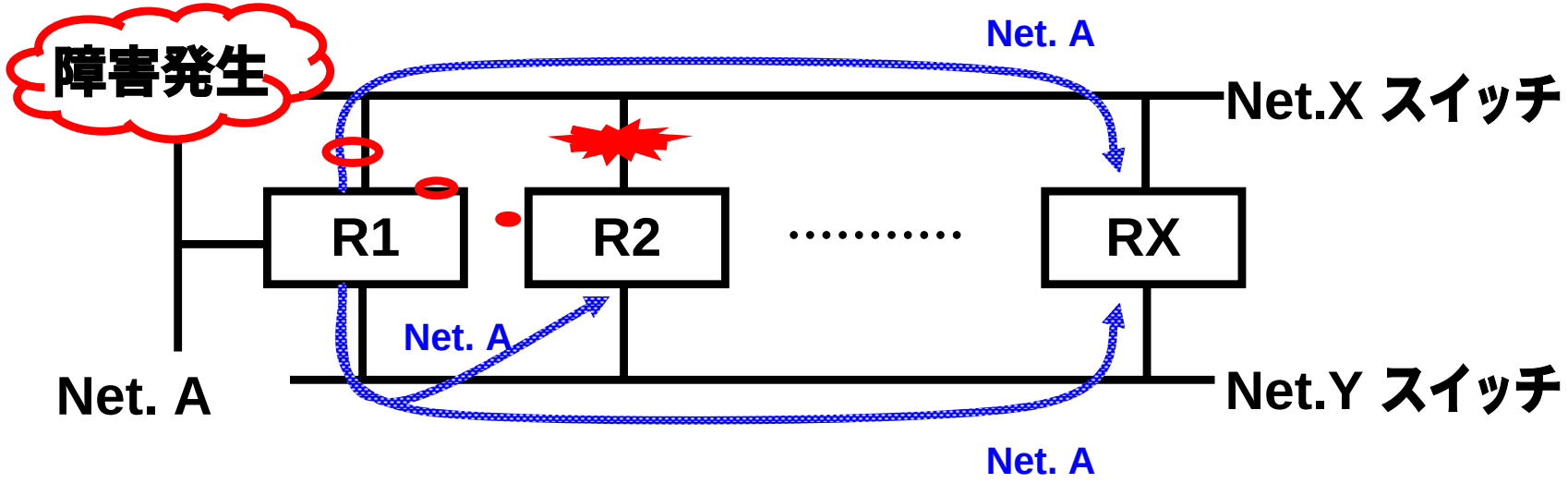
- 伝播した経路情報により、各ルータに経路情報が設定される。
- 2つのスイッチから等価な経路情報が伝播してきたため、2つの経路情報が設定される

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）



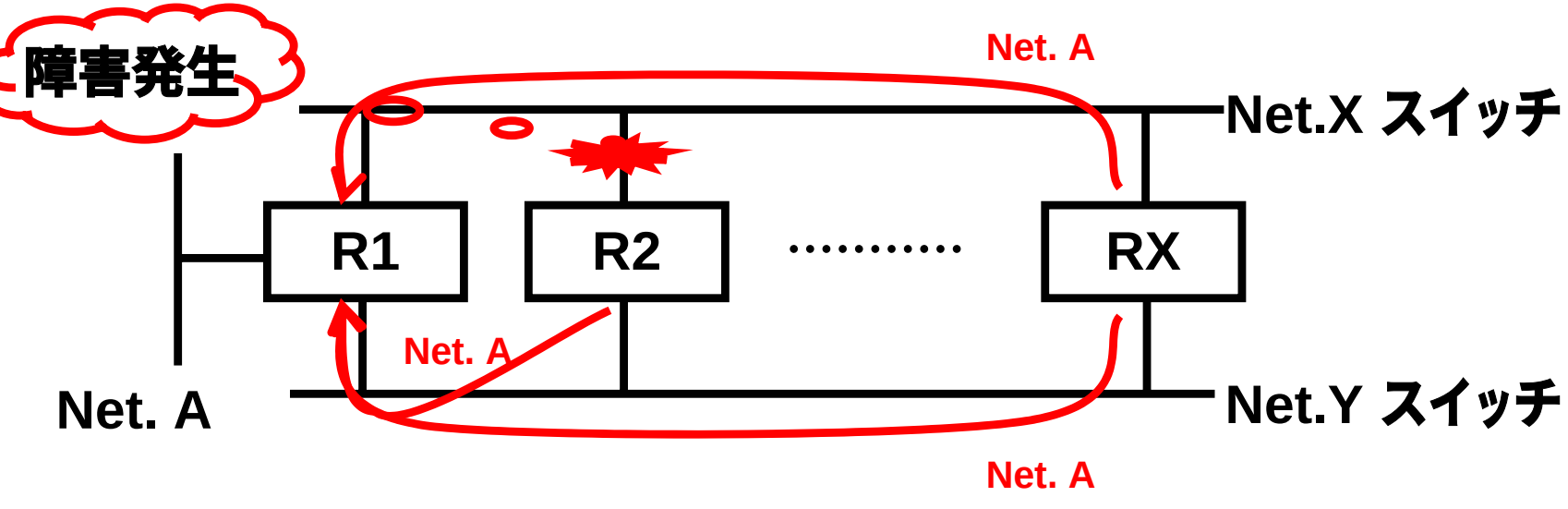
- 通常時には、2つのスイッチを経由するトラフィックがバランスする

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）



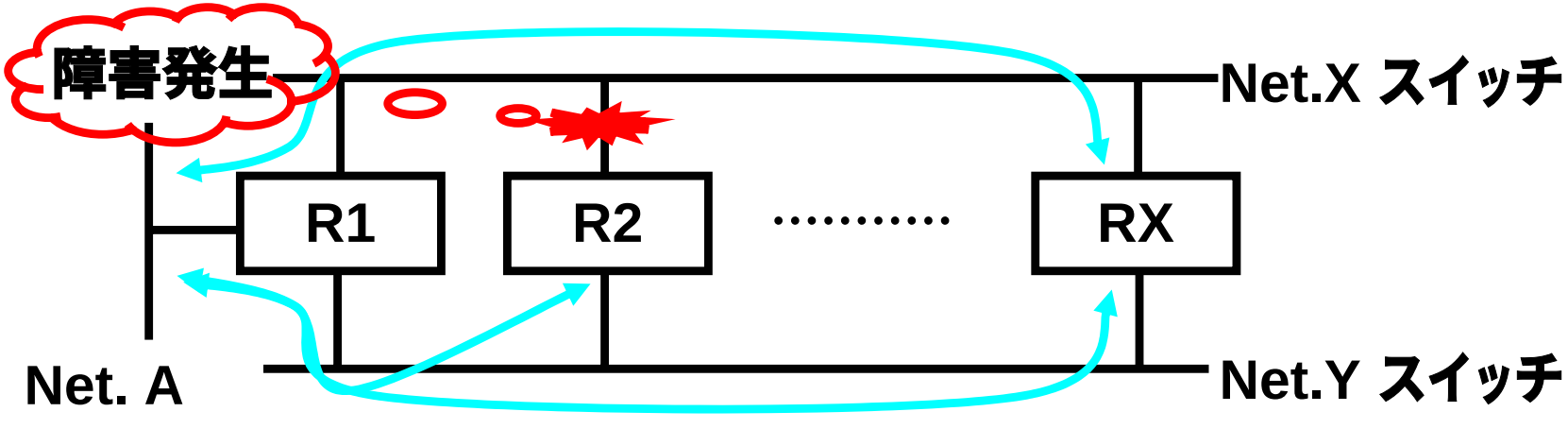
- 障害発生により、経路情報の伝播に一部に変化が生じる

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)



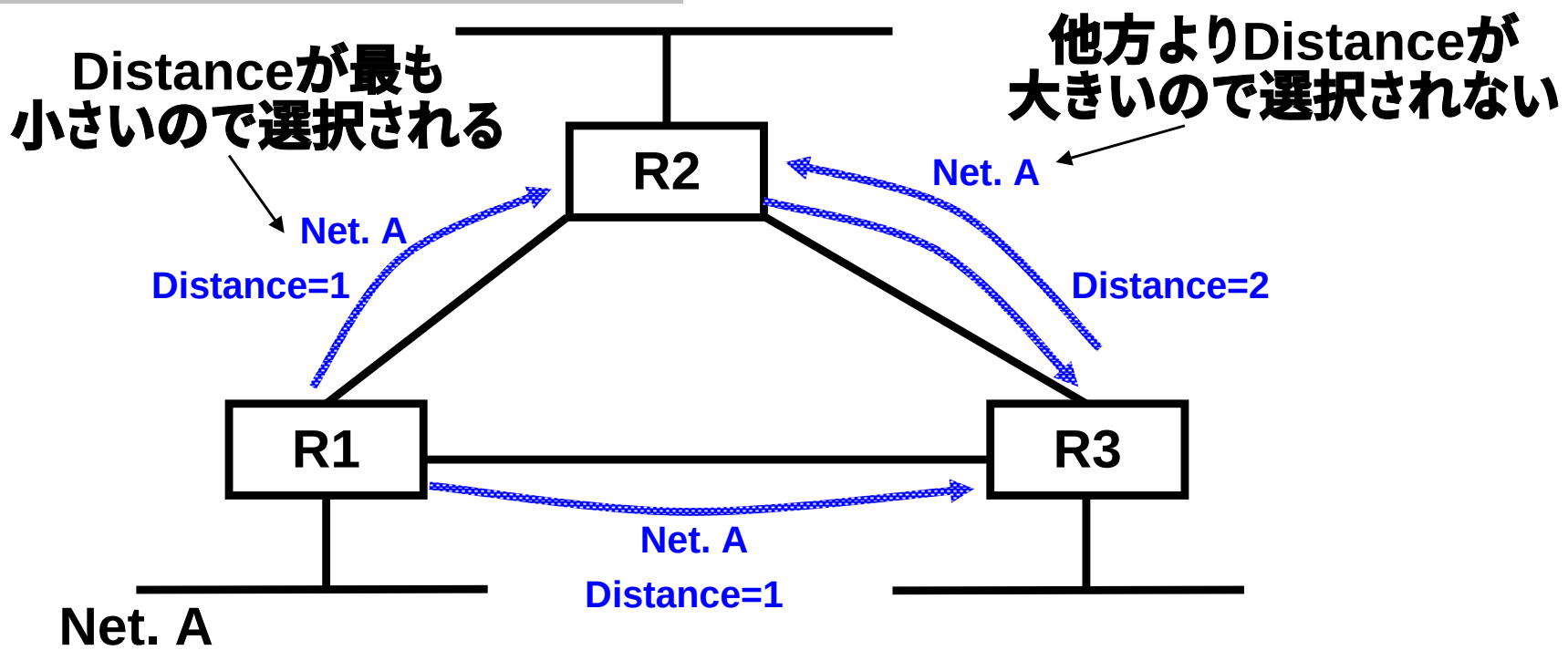
- 伝播する経路情報が変化するため、各ルータに設定されている経路情報も変化する
- 一方のスイッチからの経路が消えても、もう一方のスイッチからの経路でバックアップを行う

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



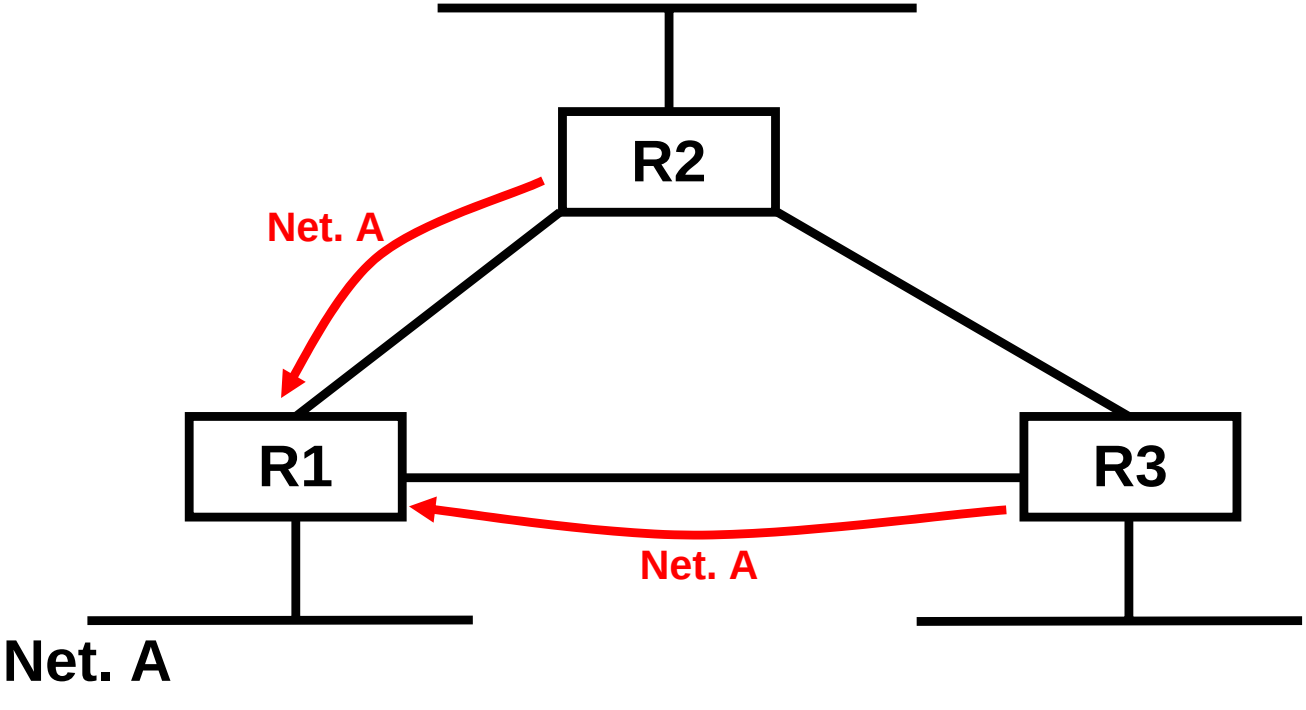
- 障害時には、2つのスイッチどちらかを利用して障害を迂回することができる

リングトポロジによるバックアップ経路の伝播（通常時）



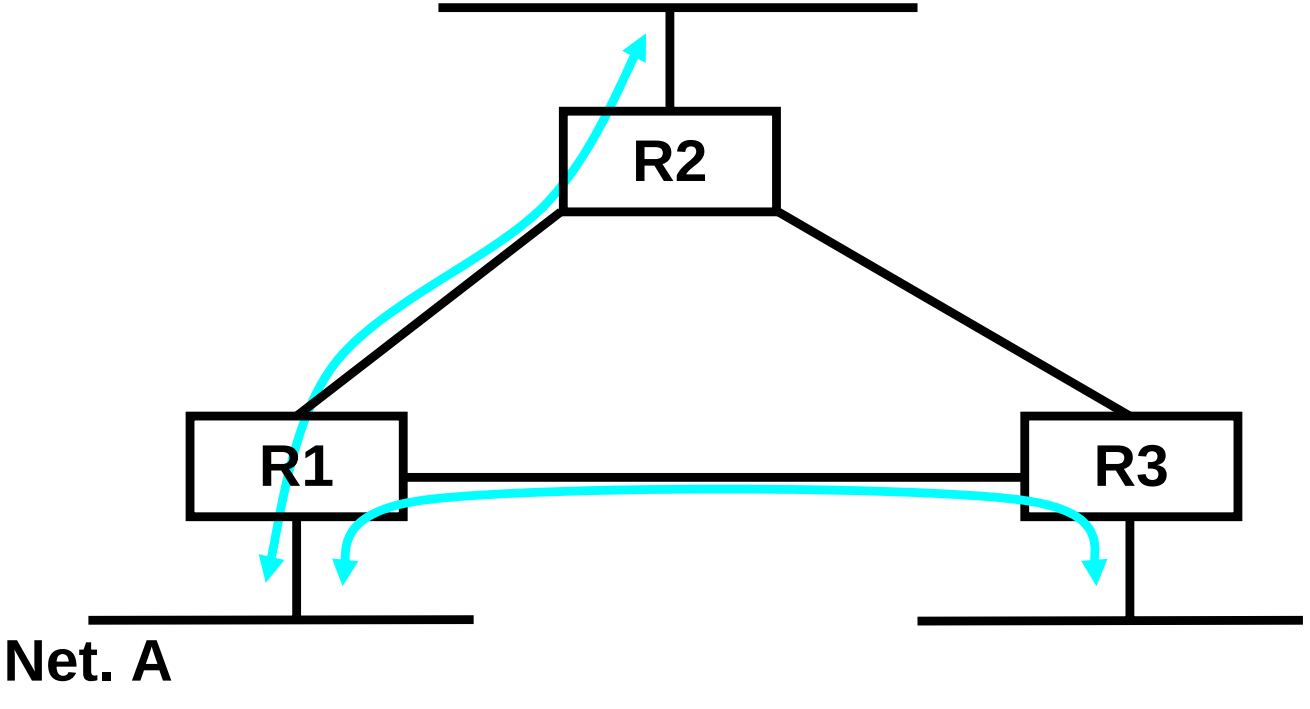
- RIPで Net.Aの経路情報を広告する
- 通常時は最短な経路が優先される

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)



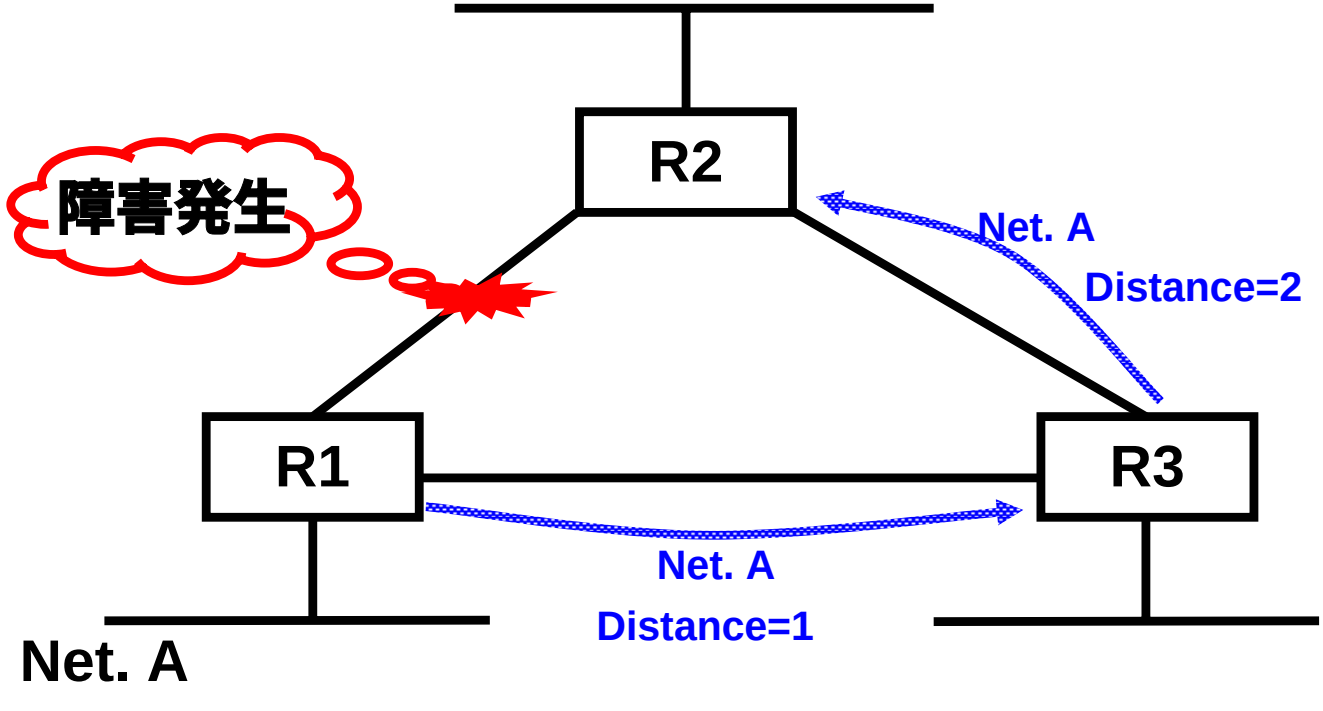
- 伝播した経路情報から、各ルータに経路情報が設定される

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)



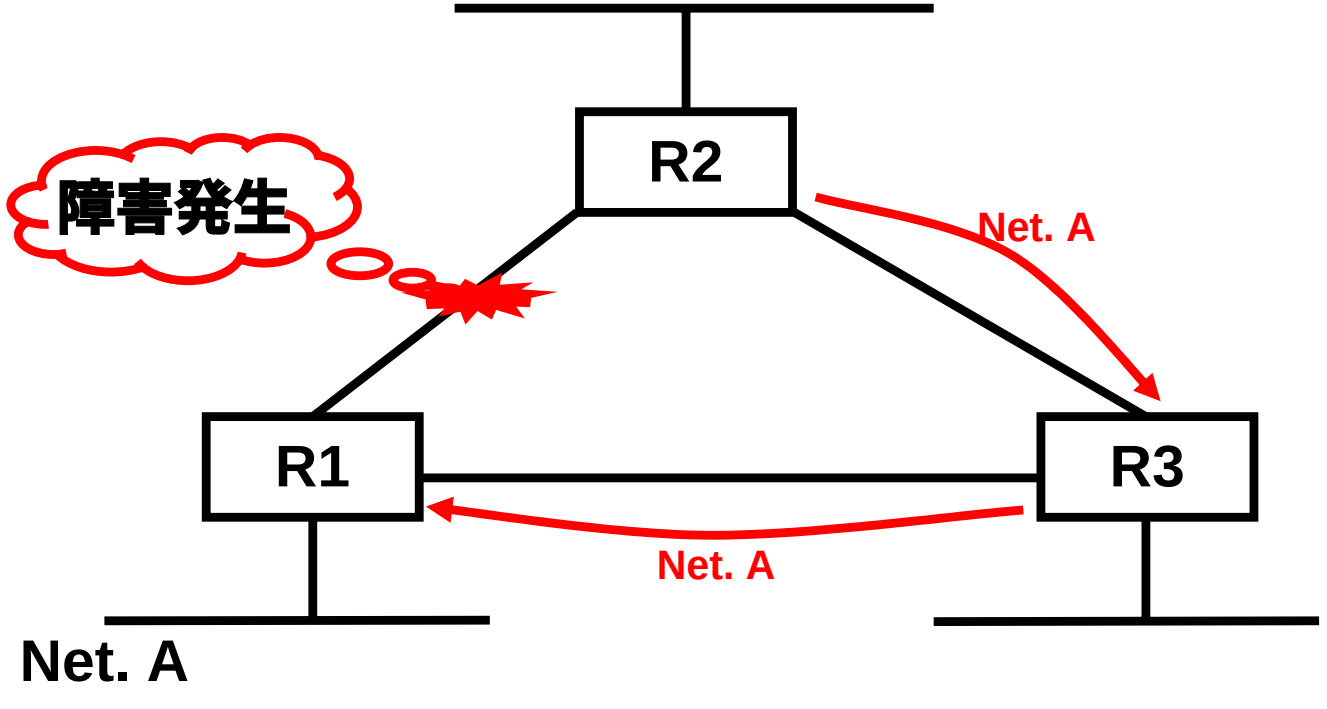
- 通常時は最短な経路が優先されて、通信が行われる

リングトポロジによるバックアップ -経路の伝播（障害時）



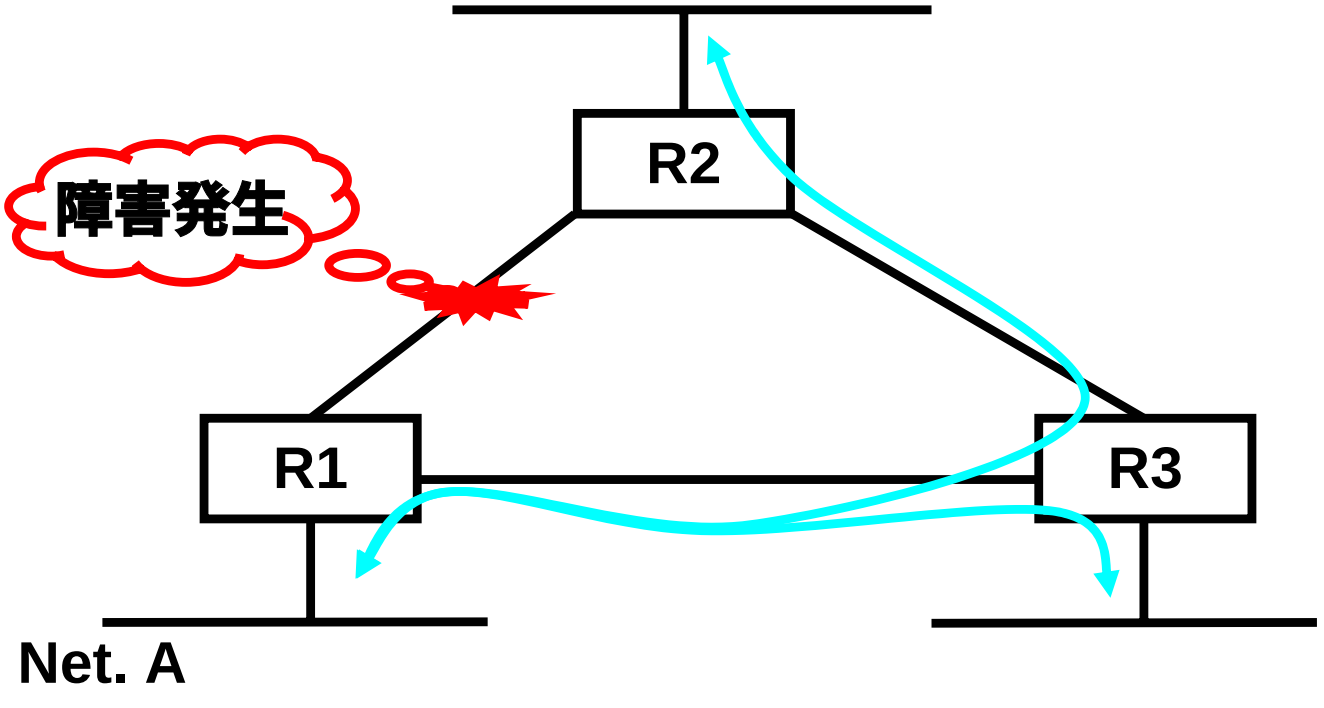
- 障害により、経路情報の伝播に変化が生じる

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

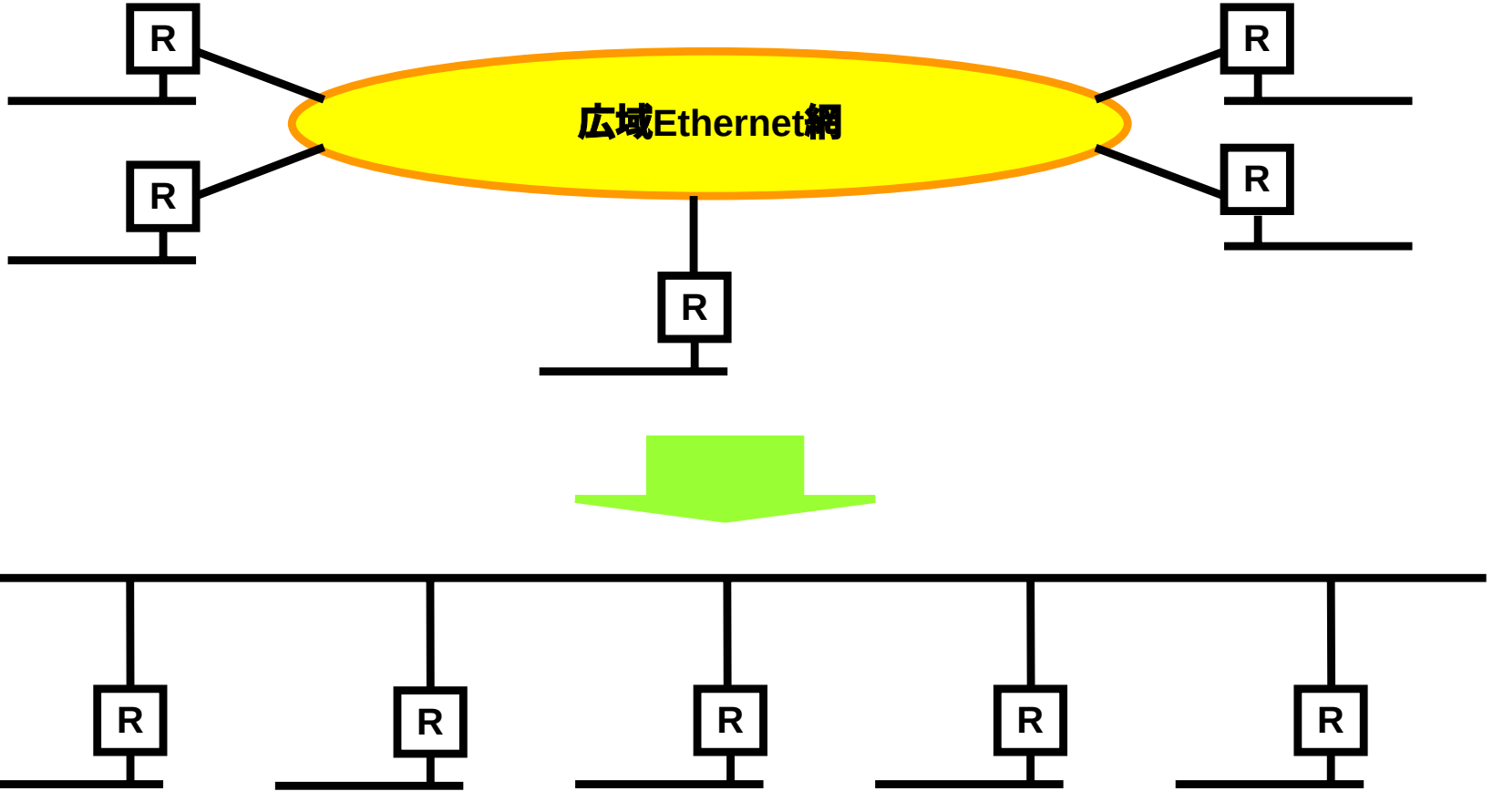
ここでは利用回線別のWAN構築の方法について解説します

- 広域Ethernetを利用したWAN
- インターネットVPNを利用したWAN

広域Ethernetを利用したWAN

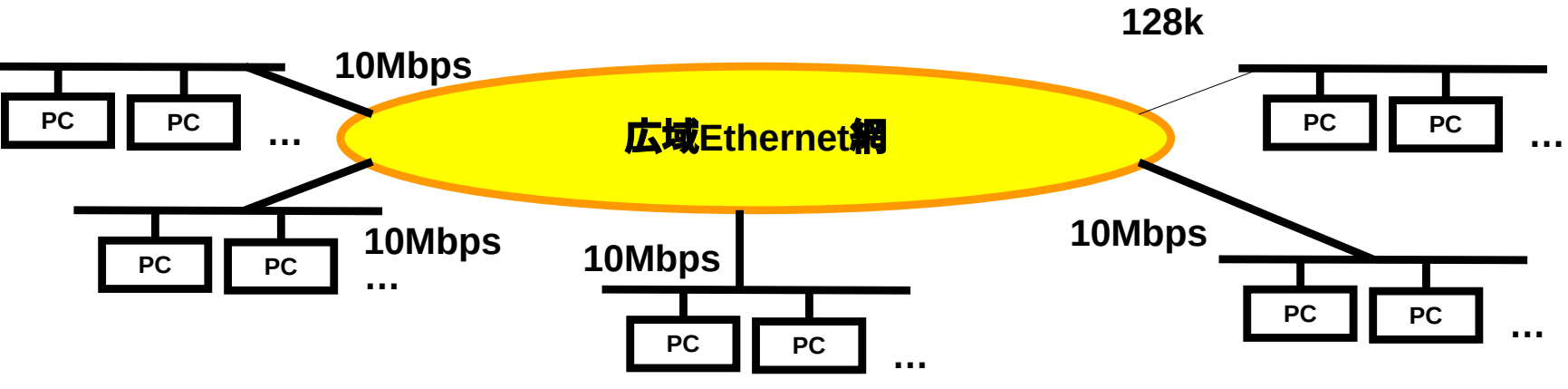
- 広域Ethernetを利用する理由
  - 安価
  - IP以外のパケットが通る(SNAなど)
  - ATMやPOSなどの高価なWAN I/Fが不要
  - ルータを利用せずにHUBだけでネットワークが構築できる
  - Tag VLANを利用して複数のVLANを複数拠点に容易に持っていくことができる
- 今回とりあげるポイント
  - IPのみを利用、ルータを利用、ダイナミックルーティング

ネットワーク構成



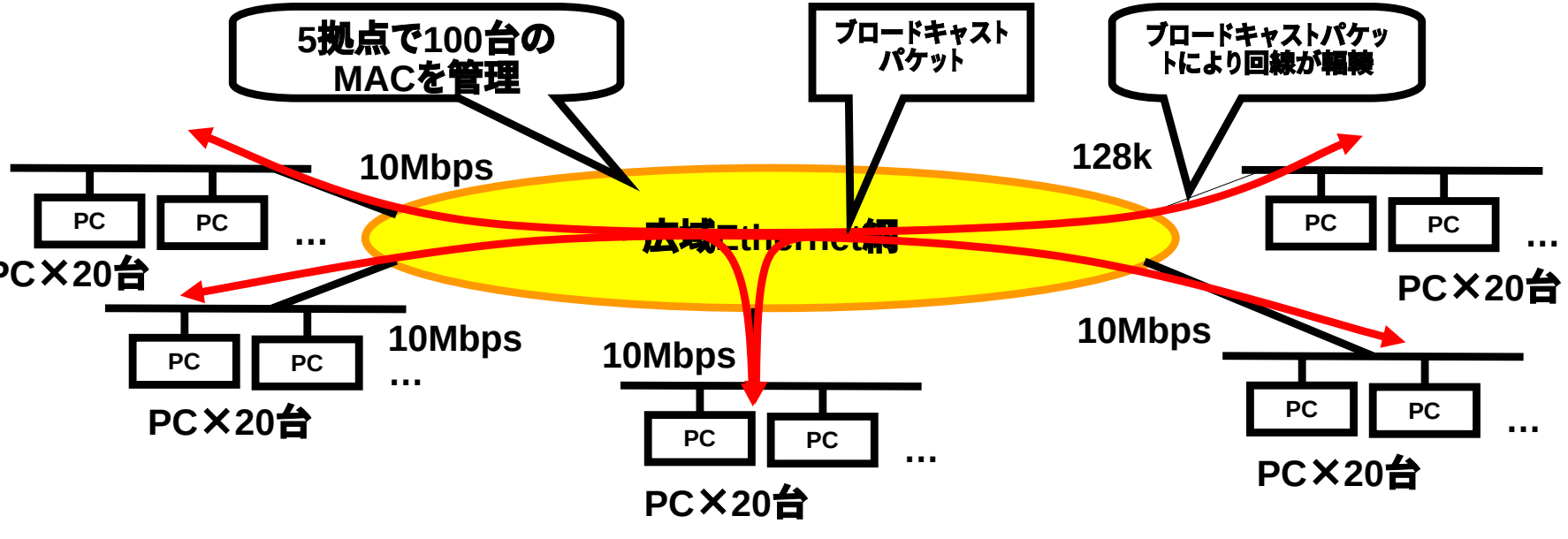
- 広域Ethernet網はLANのEthernetと同様に見える
- 基本的にはLANと同じ設計手法が使える

HUBのみで構成した場合



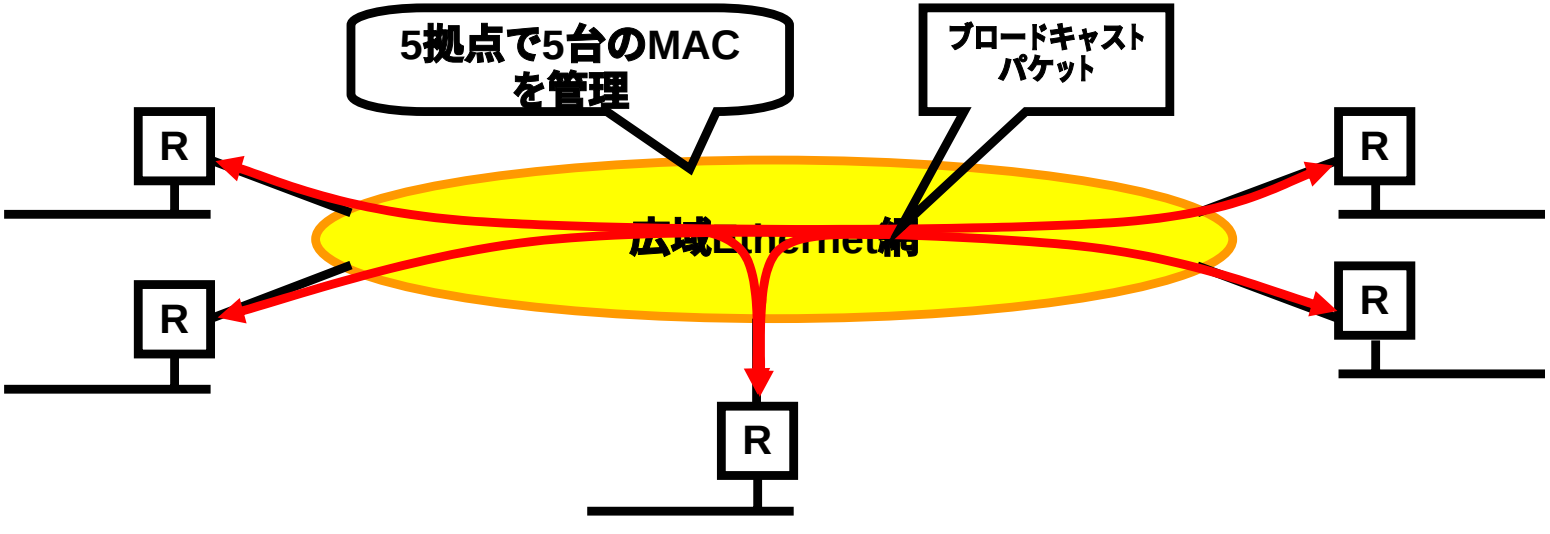
- 広域EthernetはLANと同様にHUBのみでもネットワークを構成することができる。

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

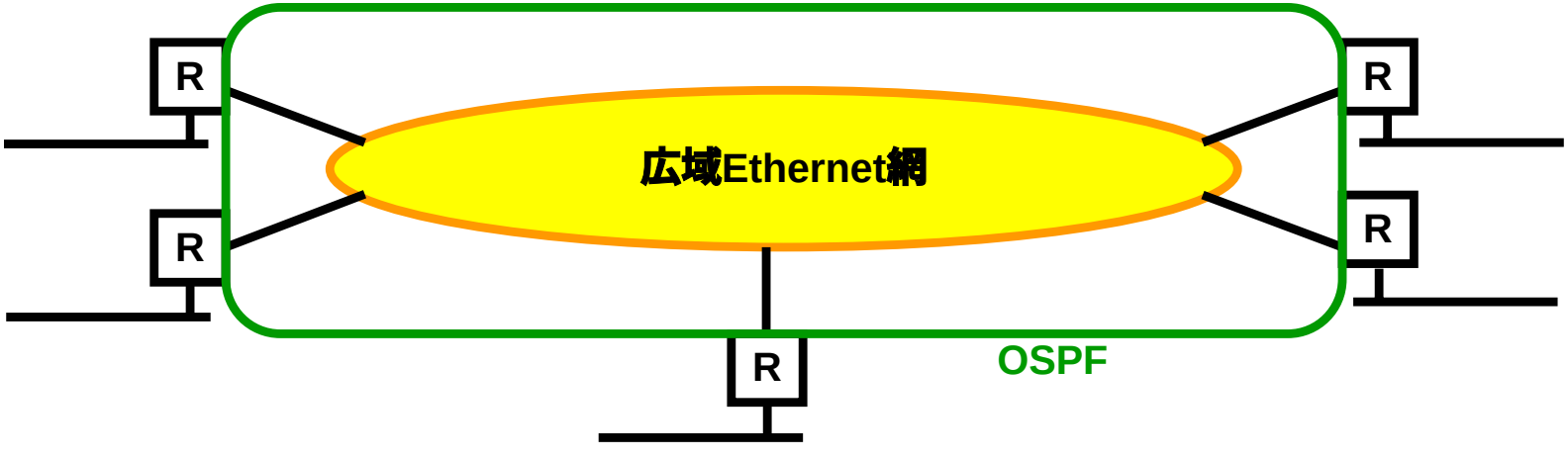


- ルータを設置した場合には広域Ethernet内でのMAC はルータの台数に限られるため、ブロードキャストの増加を防ぐことができる。

ルータを設置すべきか

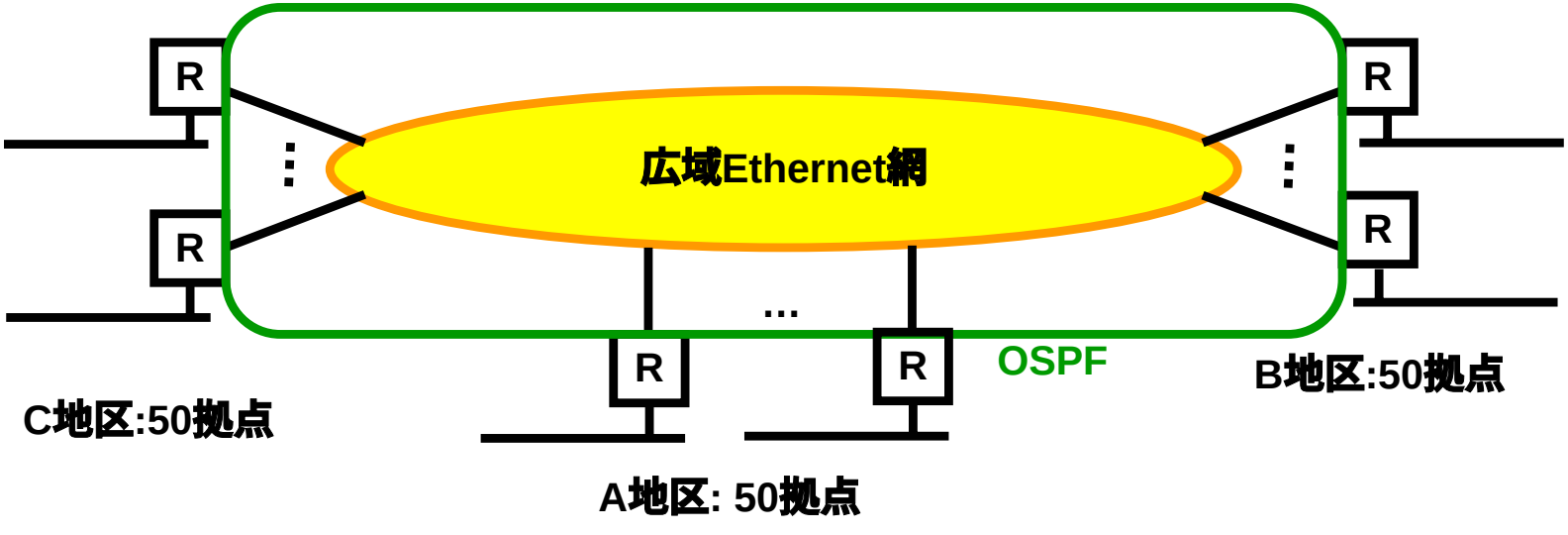
- ルータを設置すべきか、HUBのみで構築すべきか
  - 広域EthernetはHUBだけで容易にネットワークを構築できるが、スケールするネットワークとするためにはルータを設置すべきである。
  - 小規模拠点などHUBのみで構築が必要な場合にはルータ接続拠点とは異なるVLANで構築することが望ましい。

広域Ethernetでの OSPF の利用



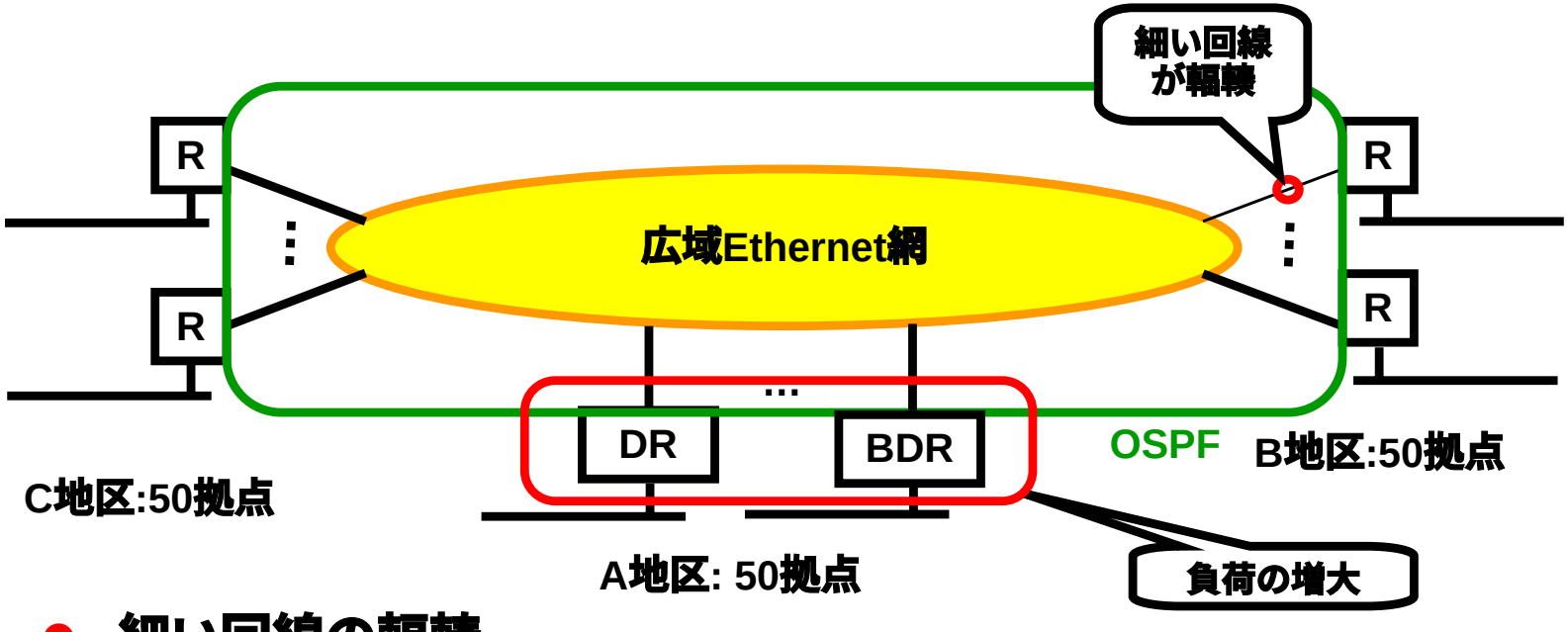
- 広域Ethernet網でのOSPFの利用
  - 広域EthernetではLANと同様にダイナミックルーティングを利用できるが、一般的にはOSPFを用いられることが多い。
  - 広域Ethernet網でのOSPF利用のポイントについて解説する

多拠点でのOSPFの利用



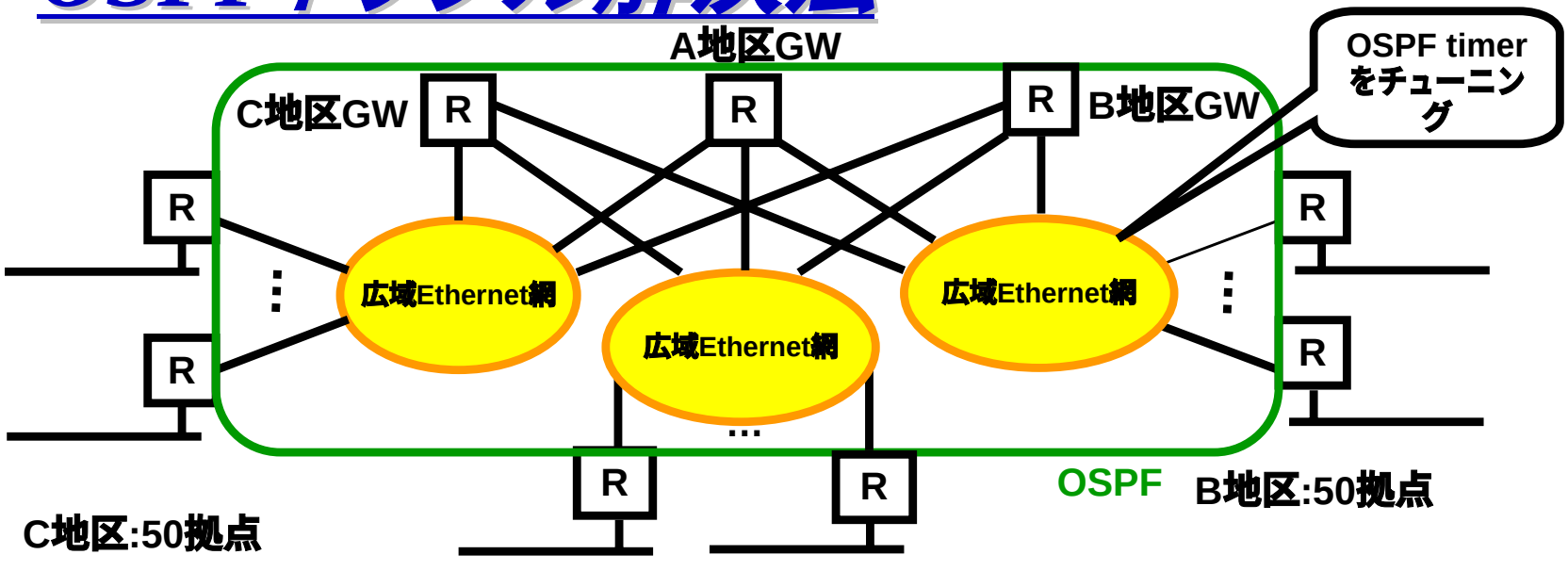
- 多くの拠点を1つの広域Ethernet網で結び、OSPFを動かす。
- 図ではA,B,C地区それぞれに50拠点接続しており、合計150台のルータが同一の広域Ethernet網を利用している。

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

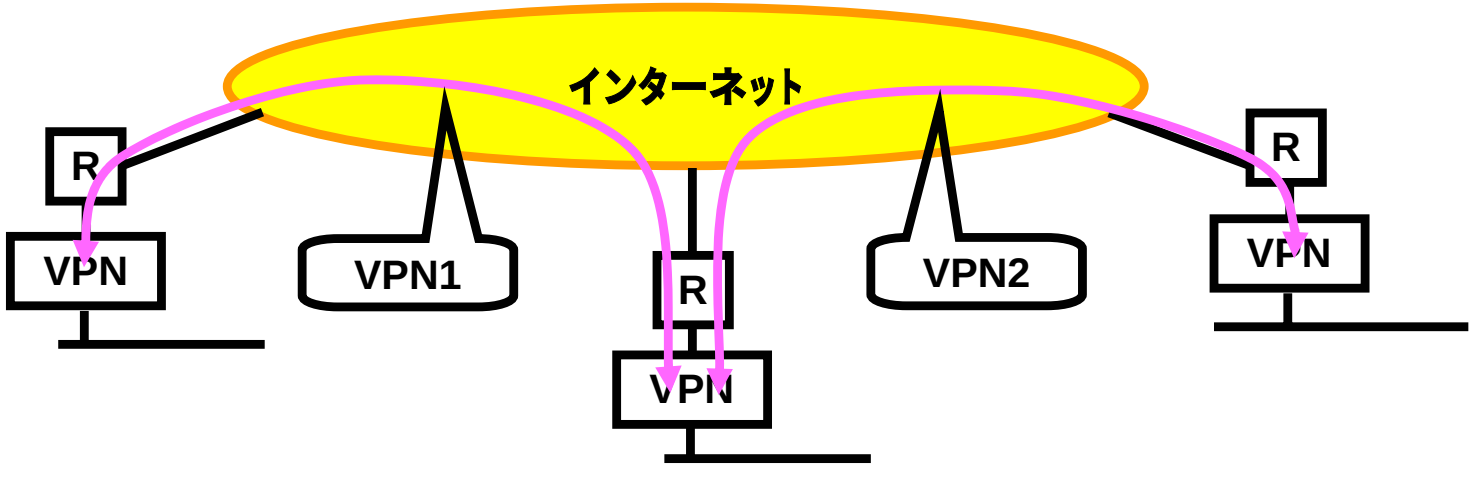
広域Ethernetを利用したWAN: まとめ

- 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。

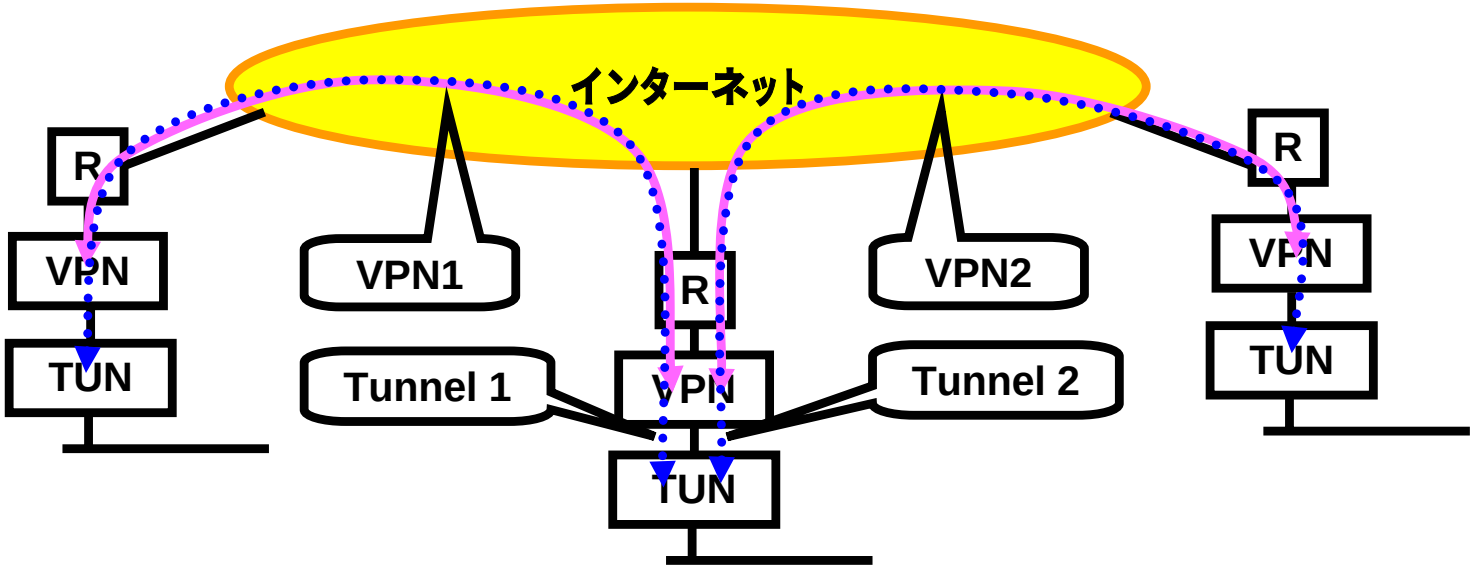
一般的なインターネットVPN構成



- 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

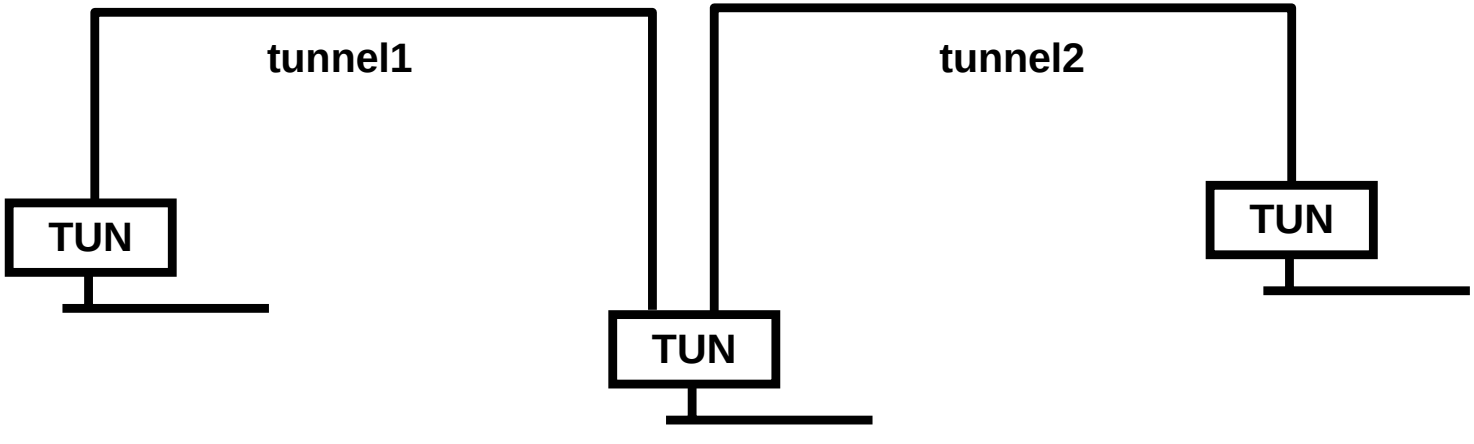


- ネットワーク化されたインターネットVPN構成
  - インターネット接続ルータの下にVPN装置を設置し、VPN装置間でVPNを張る(IPsecなど)
  - VPN装置の内側にtunnelルータを設置し、VPN上にtunnelを張る(GRE、ipipなど)
  - Tunnelは暗号化する必要はない
  - Tunnelは専用線と同等に見えるため、ダイナミックルーティングが利用できる
  - VPN装置はダイナミックルーティングを利用できなくてよい

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



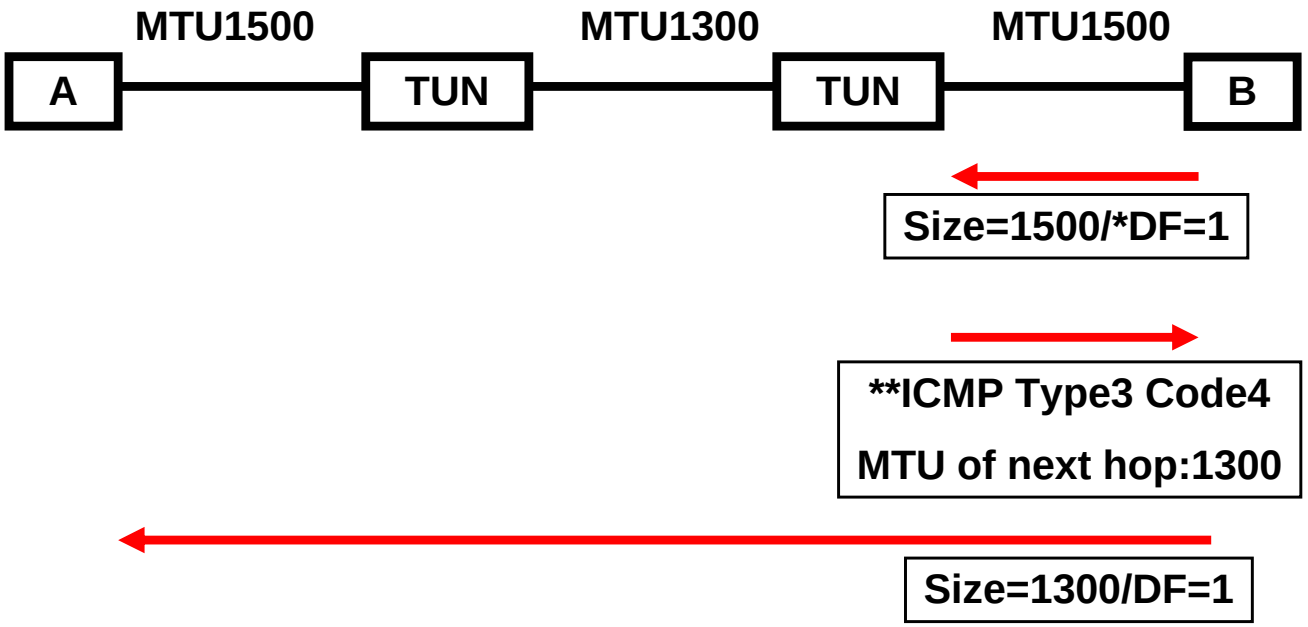
- プライベート側からみたネットワーク構成
  - Tunnelルータ間はtunnel1およびtunnel2の2つの専用線で接続されていることと同様に扱える
  - OSPFなどのダイナミックルーティングを容易に扱うことができる

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

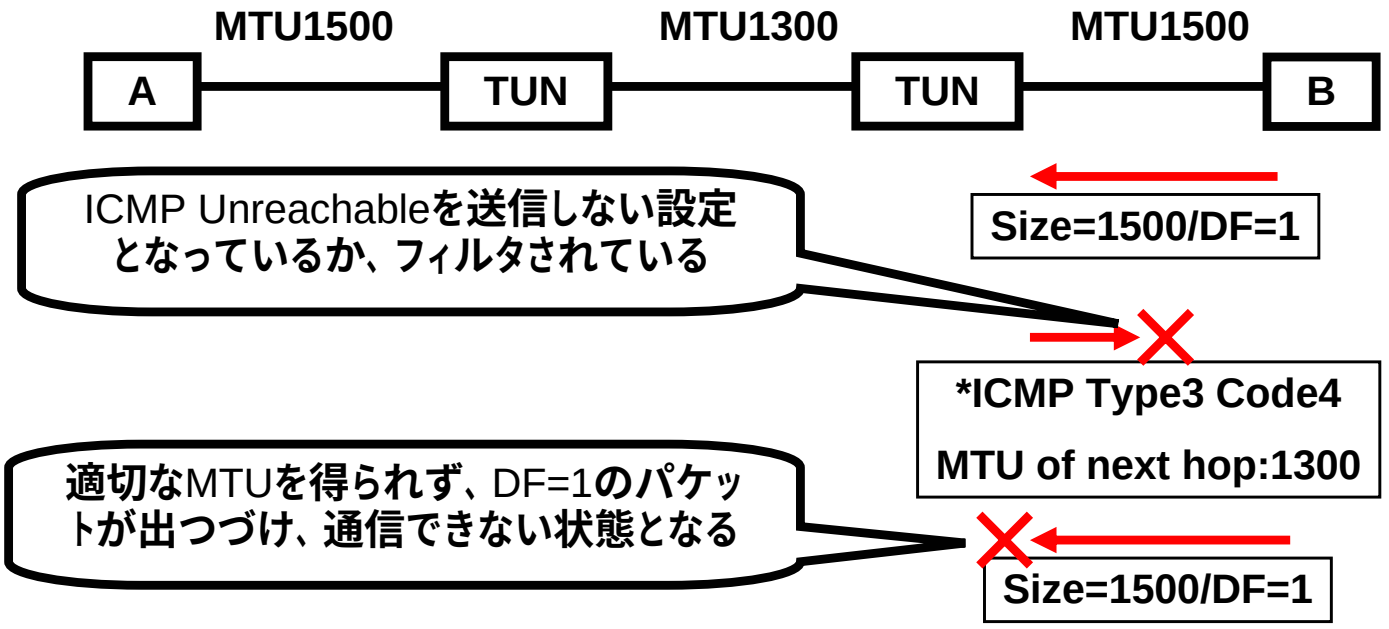
Path MTU Discovery の動作原理



*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

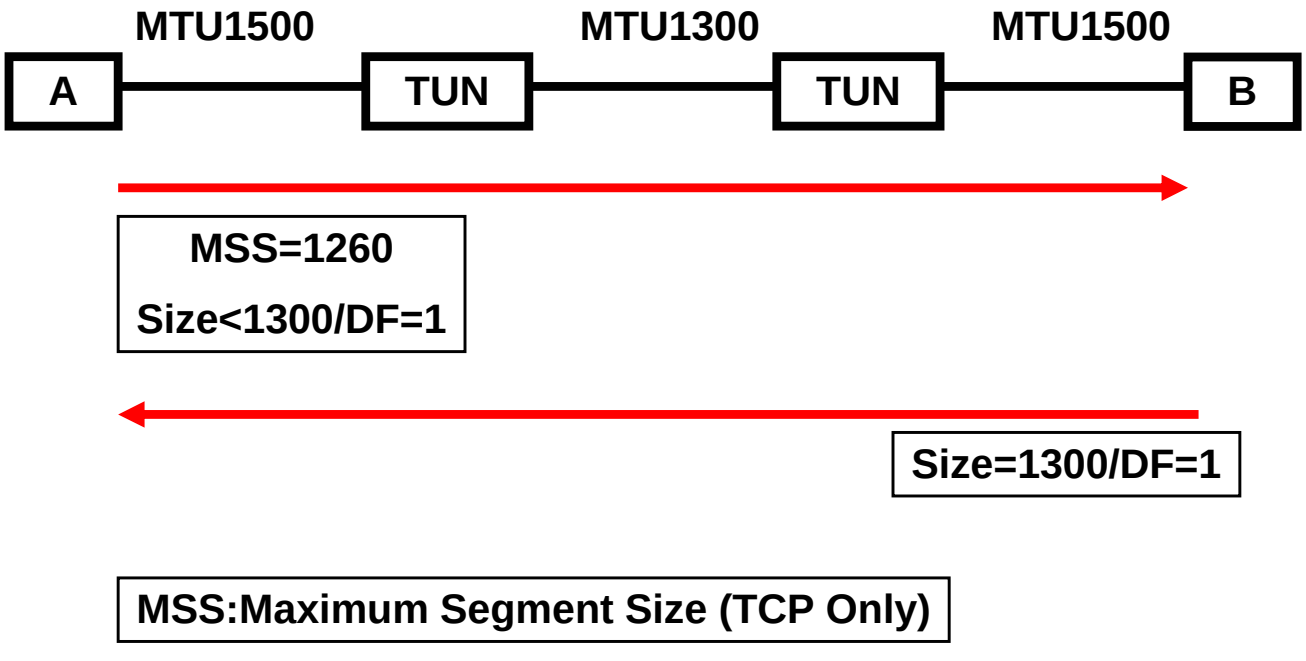
Path MTU Discovery Black hole問題



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

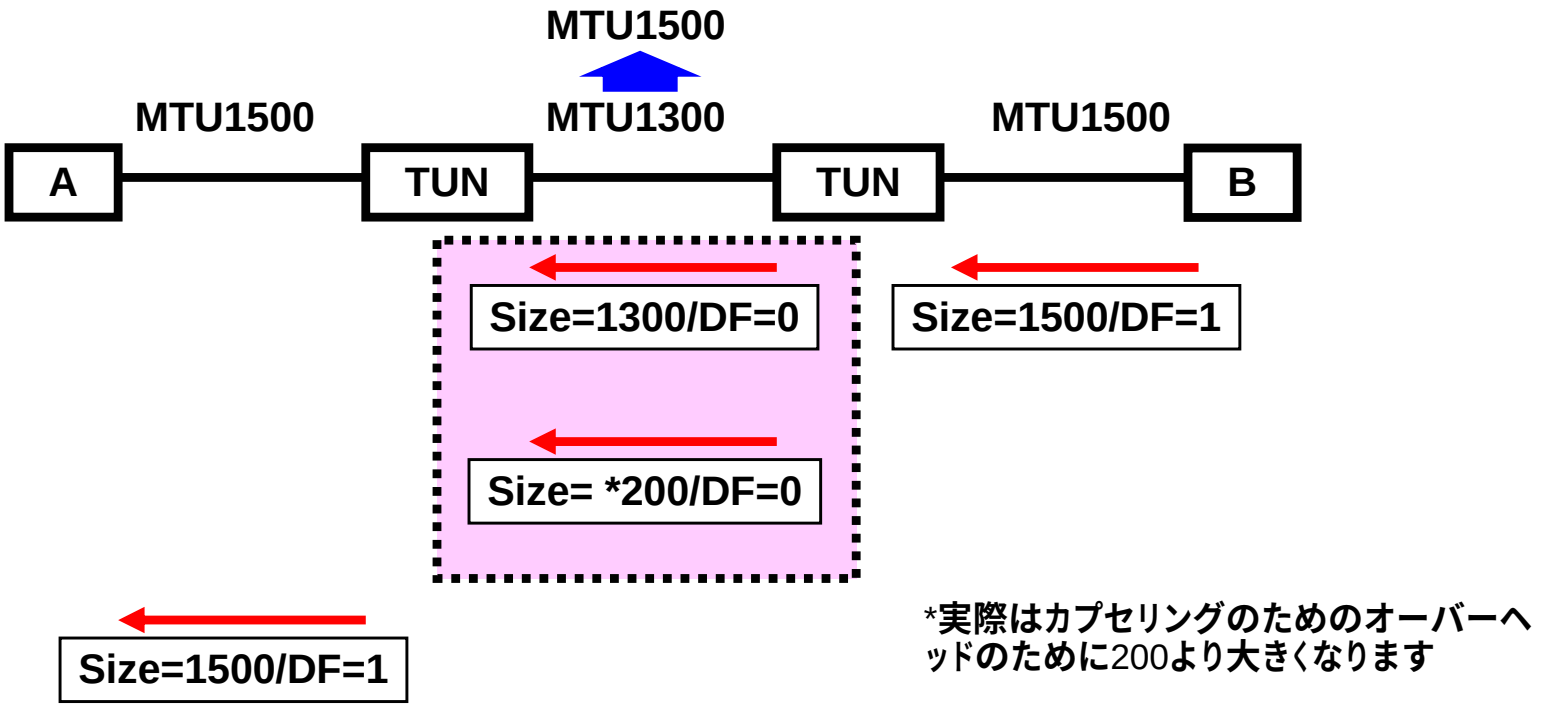
- Path MTU Discovery Black holeとは
  - Path MTU Discoveryの原理で重要な役割を持つICMP Unreachableがフィルタされることで、適切なIPパケットの送信が行えず、通信できなくなる状態のこと

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

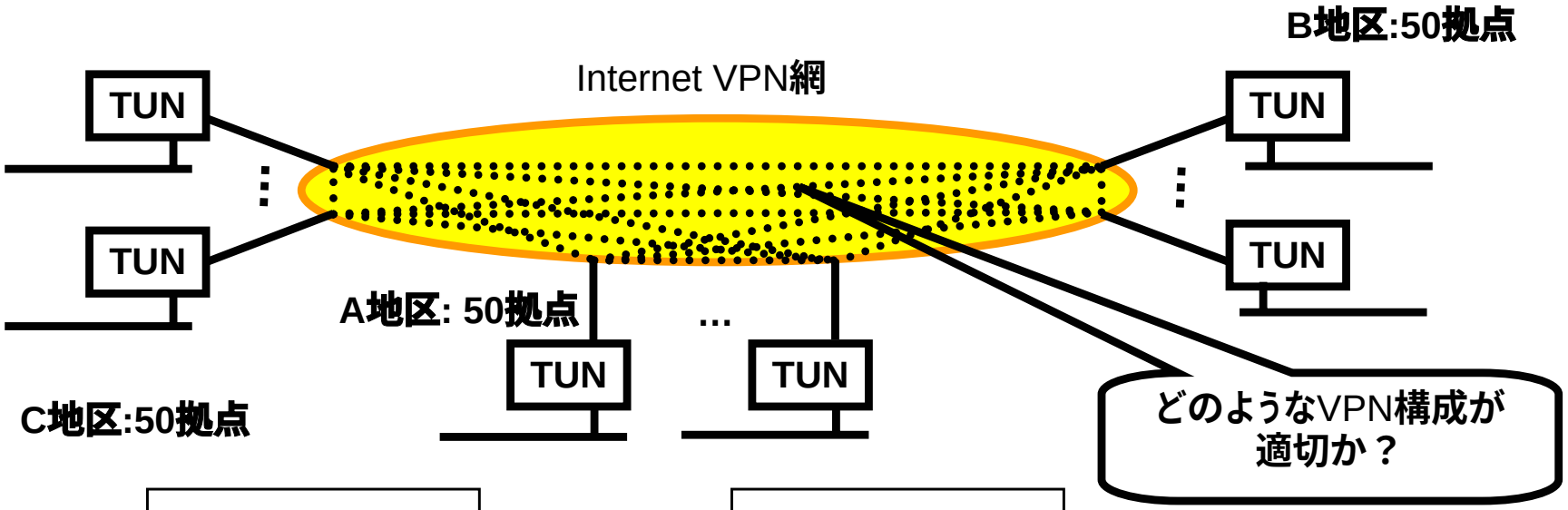


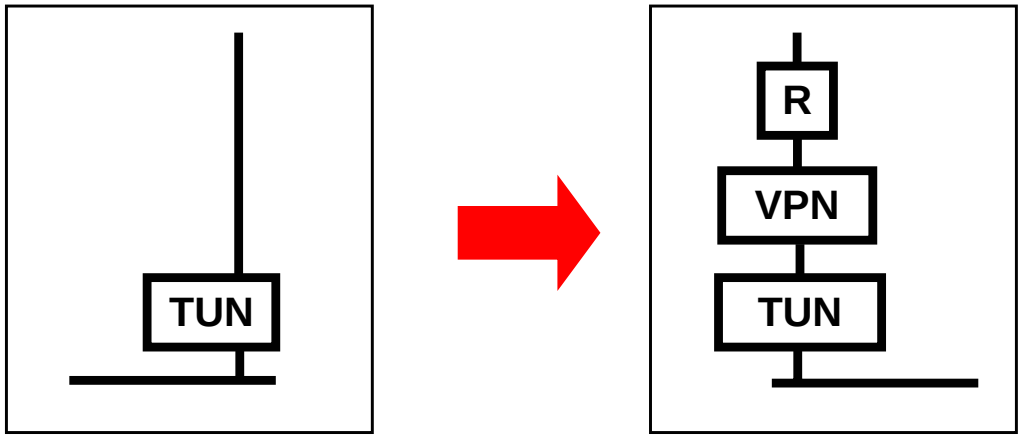
- TunnelのMTUを1500に拡張する
  - TunnelのMTUを1500に拡張することで、tunnel区間をパケットを分割して通過させることができる
  - TCP(Protocol:6)以外のUDP(Protocol:17)やESP(Protocol:50)などの1500バイトパケットを通すことができる
  - Path MTU Discovery以外の要因によるDF=1のIPにも対応が可能

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
  - 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



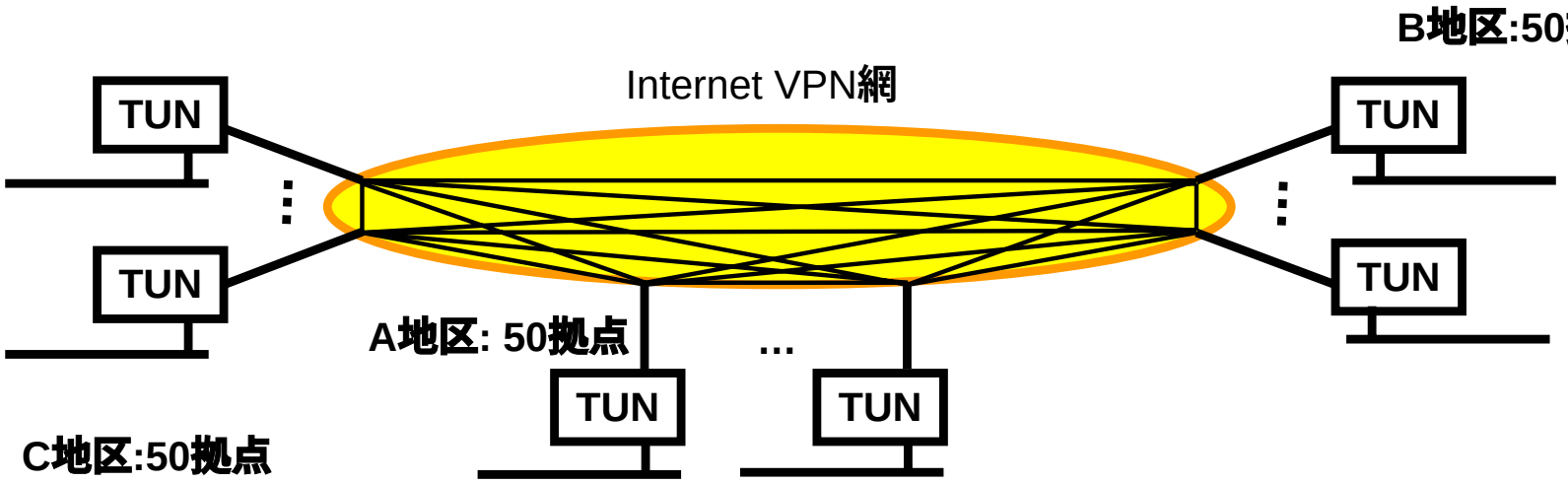


TUNはtunnelルータを表し、前述のとおり、物理的にはには下記の3つの機器を表す

- インターネット接続ルータ
- VPN装置
- Tunnelルータ

- 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

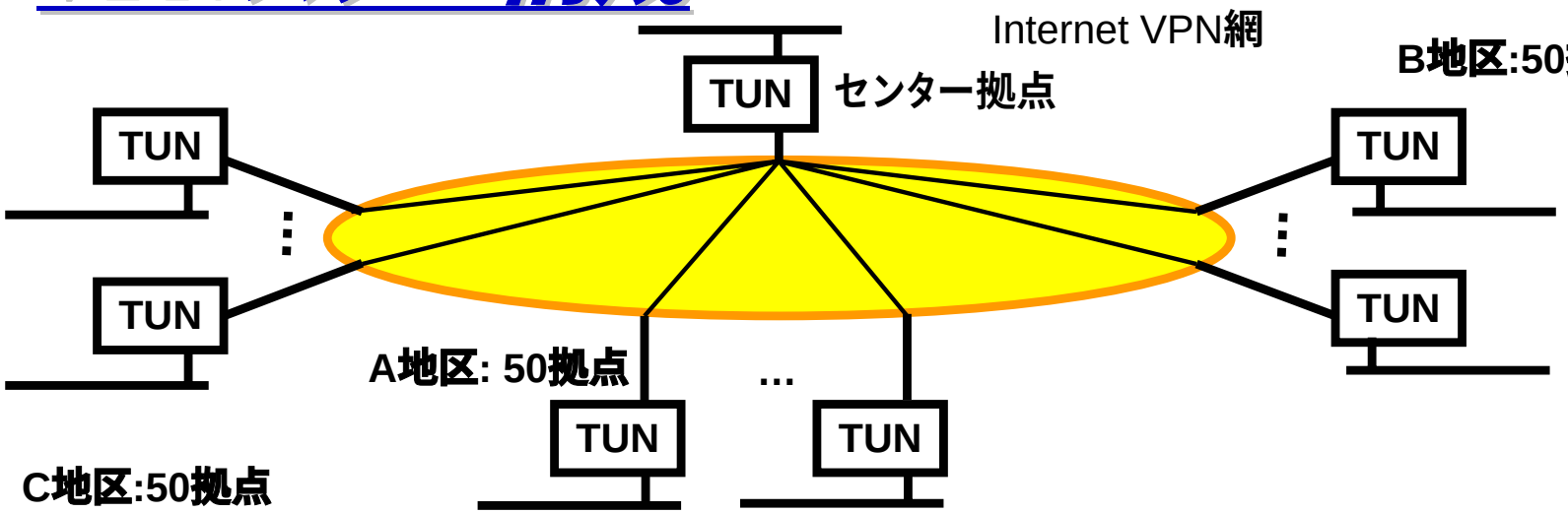


- フルメッシュ構成
  - すべての拠点間をVPNで結ぶ
  - 拠点の数:mとしたときのVPNの数

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

  - 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

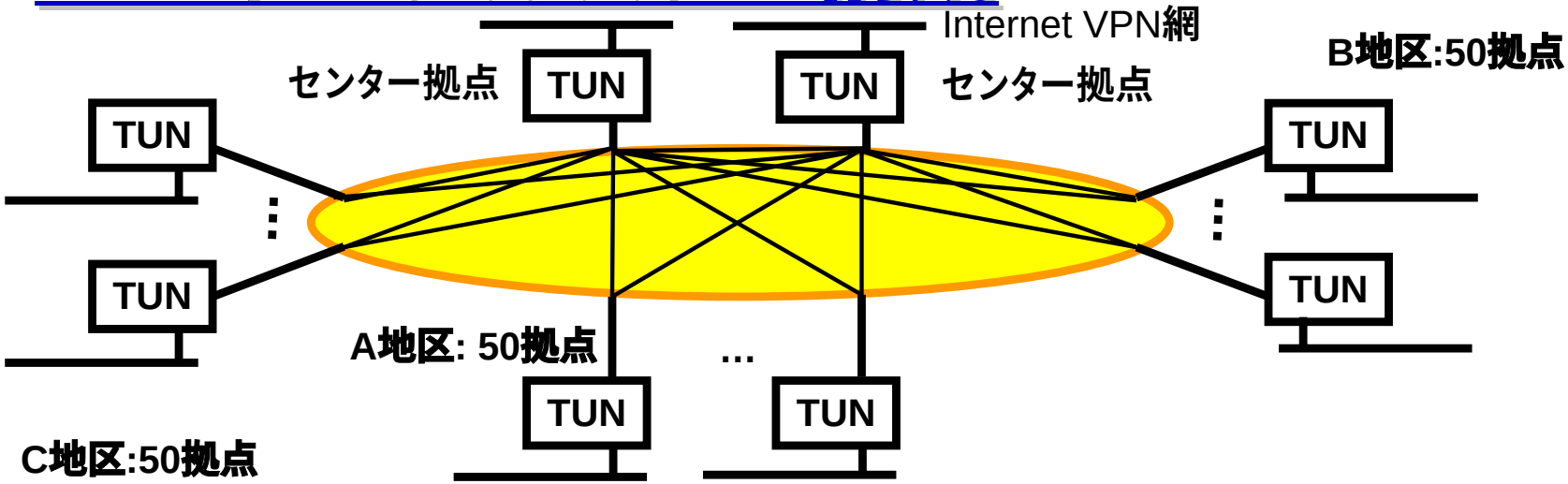


- スター構成
  - 各拠点とセンター拠点との間をVPNで結ぶ
  - 拠点の数:mとしたときのVPNの数は

$$\text{VPNの数} = m - 1$$

  - 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

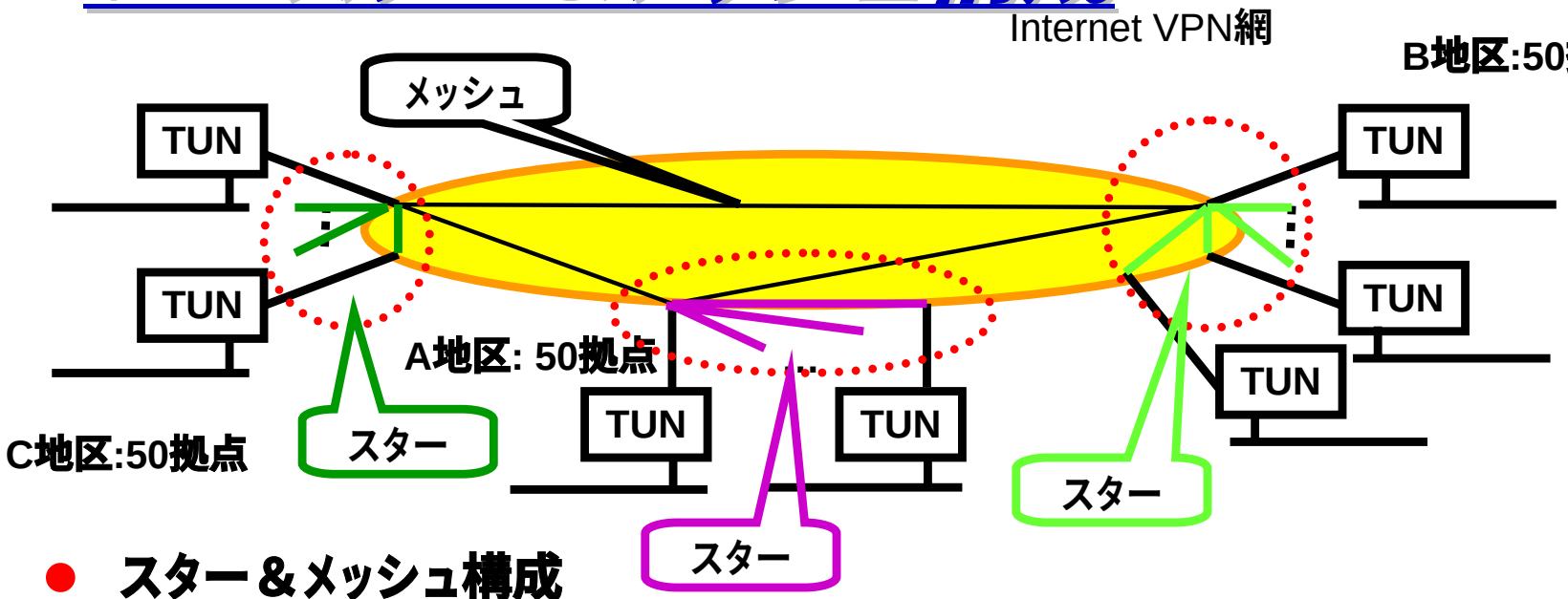


- デュアルスター構成
  - 各拠点から2箇所のセンター拠点にVPNで結ぶ
  - 拠点の数:mとしたときのVPNの数は

$$\text{VPNの数} = 2(m-1)$$

  - 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

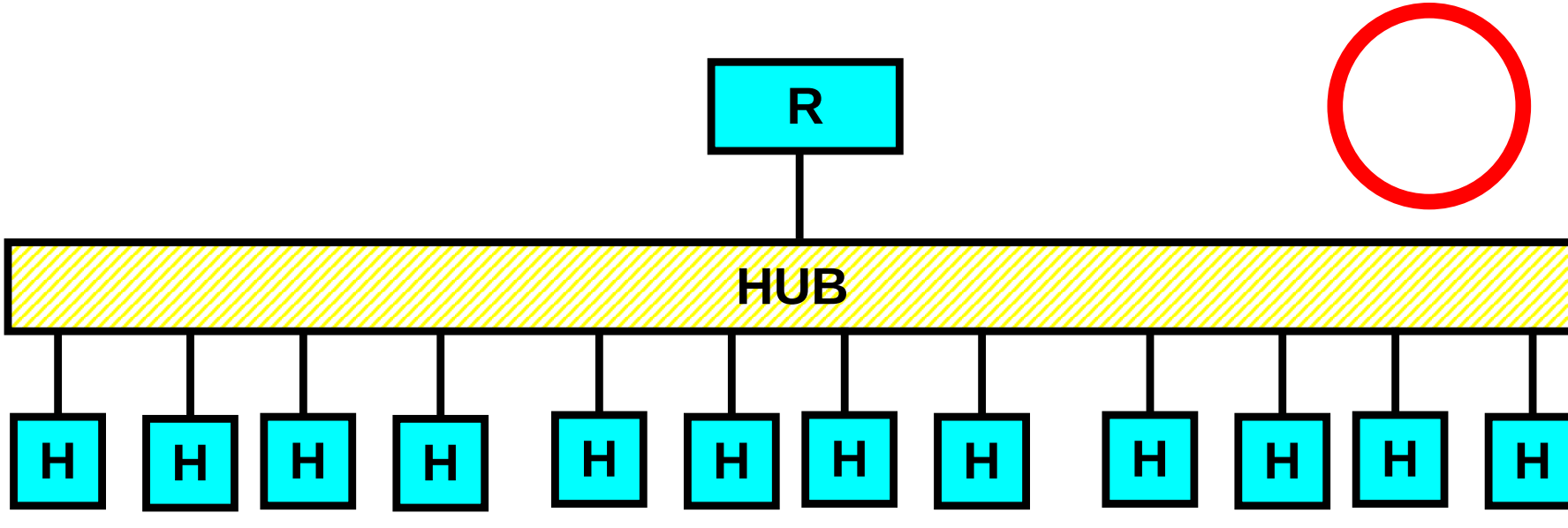
$$\text{VPNの数} = m - n + \frac{(n^2 - n)}{2} = \frac{n^2 - 3n + 2m}{2}$$

  - 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

ネットワーク構築

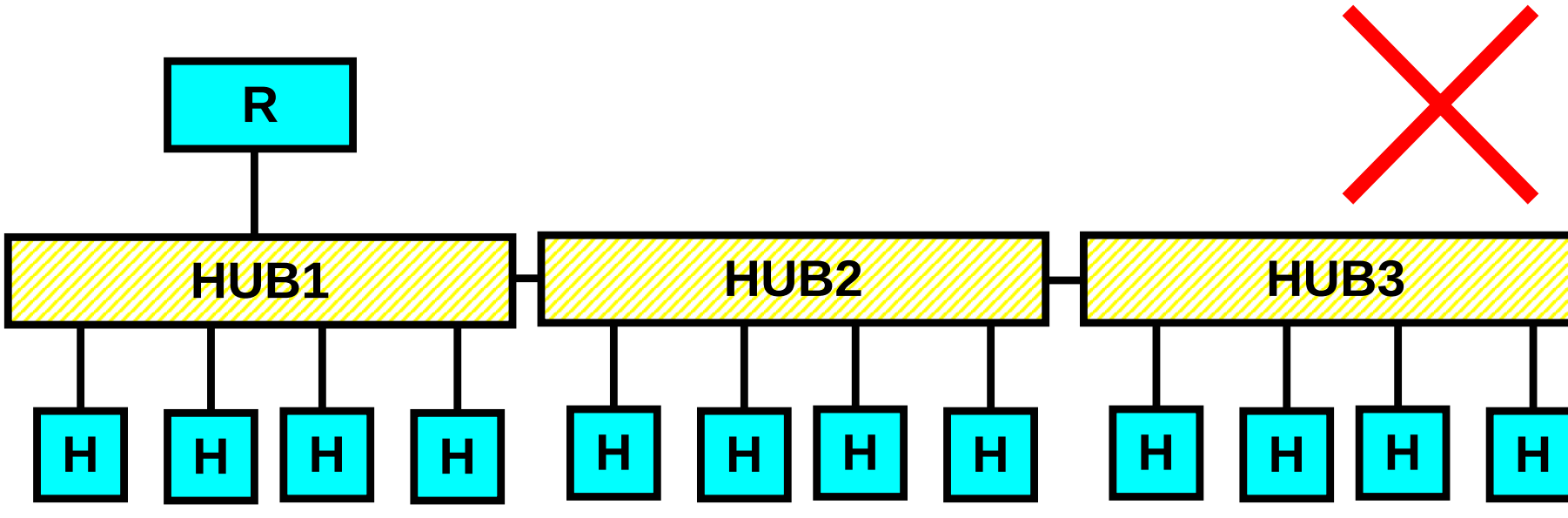
- ネットワーク構築に必要なL2ネットワークの構成法について解説します
- 配線に必要な部材について紹介します

L2ネットワーク構成 カスケードなし



- 標準的なカスケードの無いL2ネットワーク
- ホストの数だけ多ポートのHUBを用意する必要がある
- 遅延は少なく、L2障害発生の可能性も低く安定している
- HUBのカスケード接続は許可しないため、L2ループなどの障害発生を防止できる

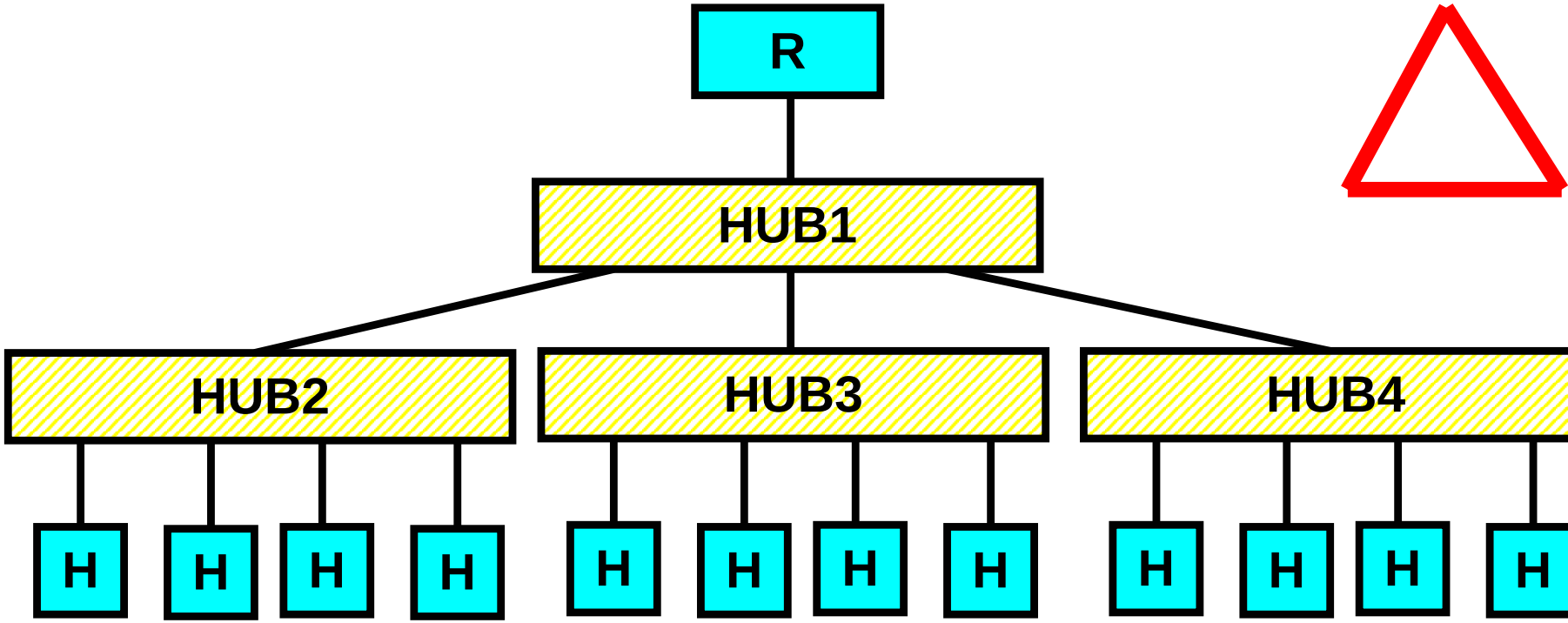
L2ネットワーク構成 横繋ぎカスケード



- 横繋ぎカスケードによるL2ネットワーク
- 安価なHUBをカスケード接続利用できる
- 障害発生時に障害箇所の特定が困難
- カスケードポイントに障害が発生するとL2ネットワークが分断し、正常に冗長化できない
- 誤ったカスケード構成によりループが発生する可能性がある

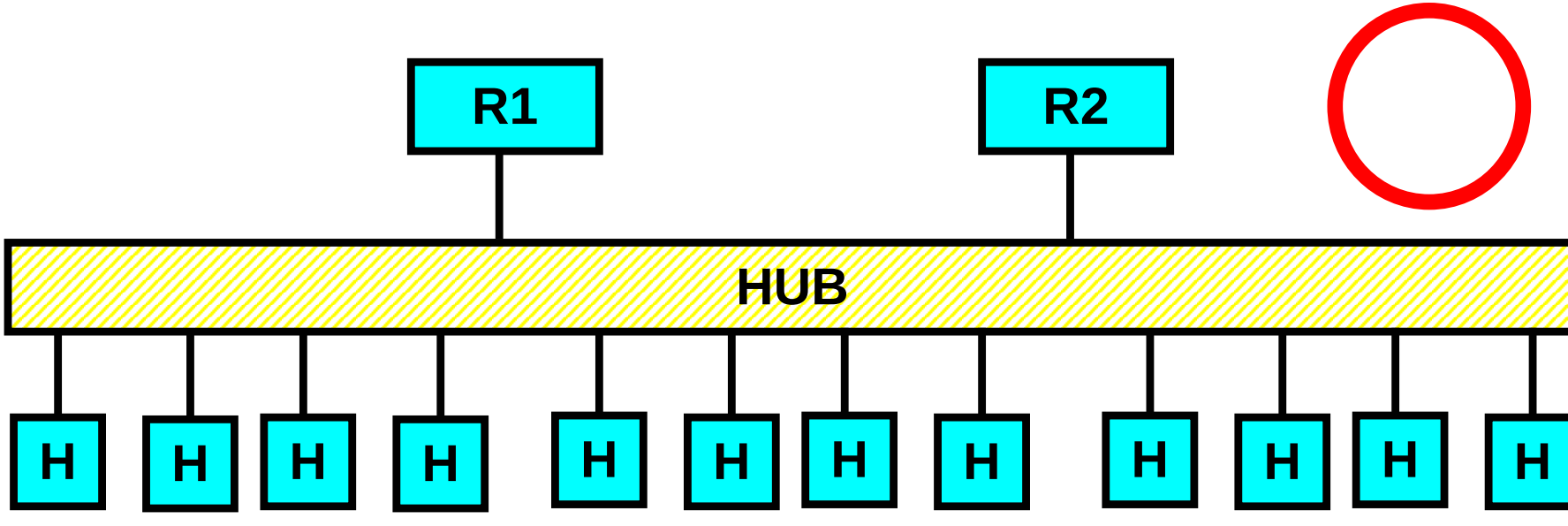
IIJ
Internet Initiative Jap

L2ネットワーク構成 ツリー型カスケード



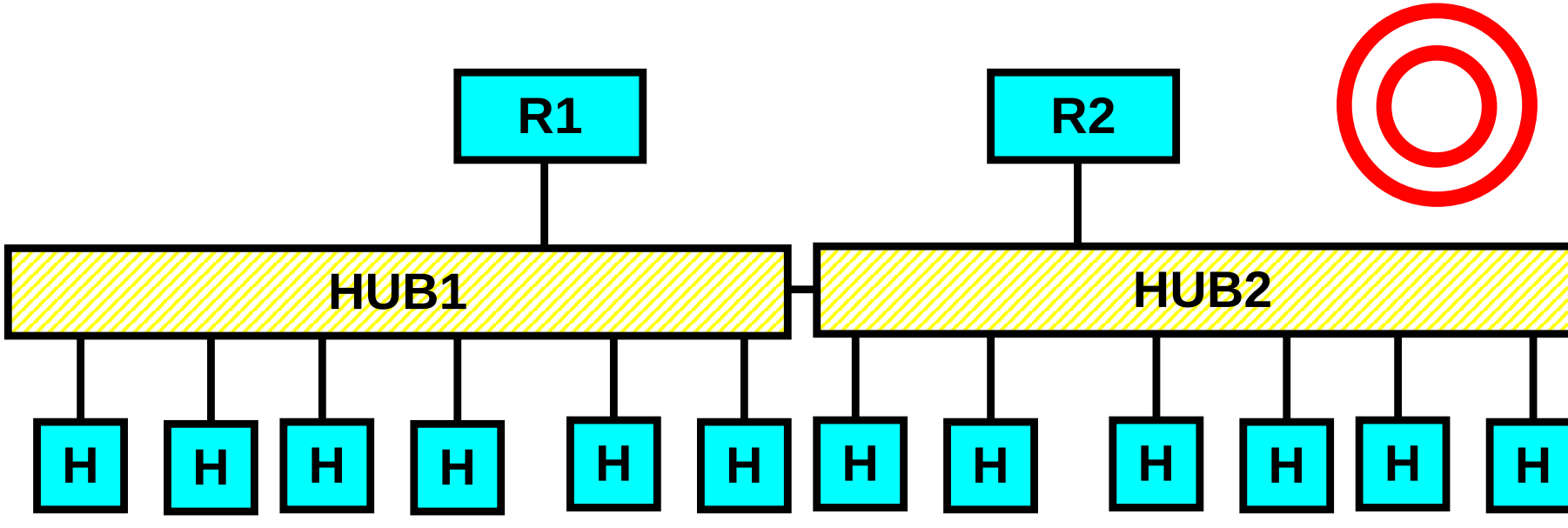
- ツリー型カスケードによるL2ネットワーク
- 安価なHUBをカスケード接続利用できる
- 障害箇所をある程度局所化できる
- カスケードポイントに障害が発生するとL2ネットワークが分断し、正常に冗長化できない
- 誤ったカスケード構成によりループが発生する可能性がある

冗長化L2ネットワーク構成 カスケードなし



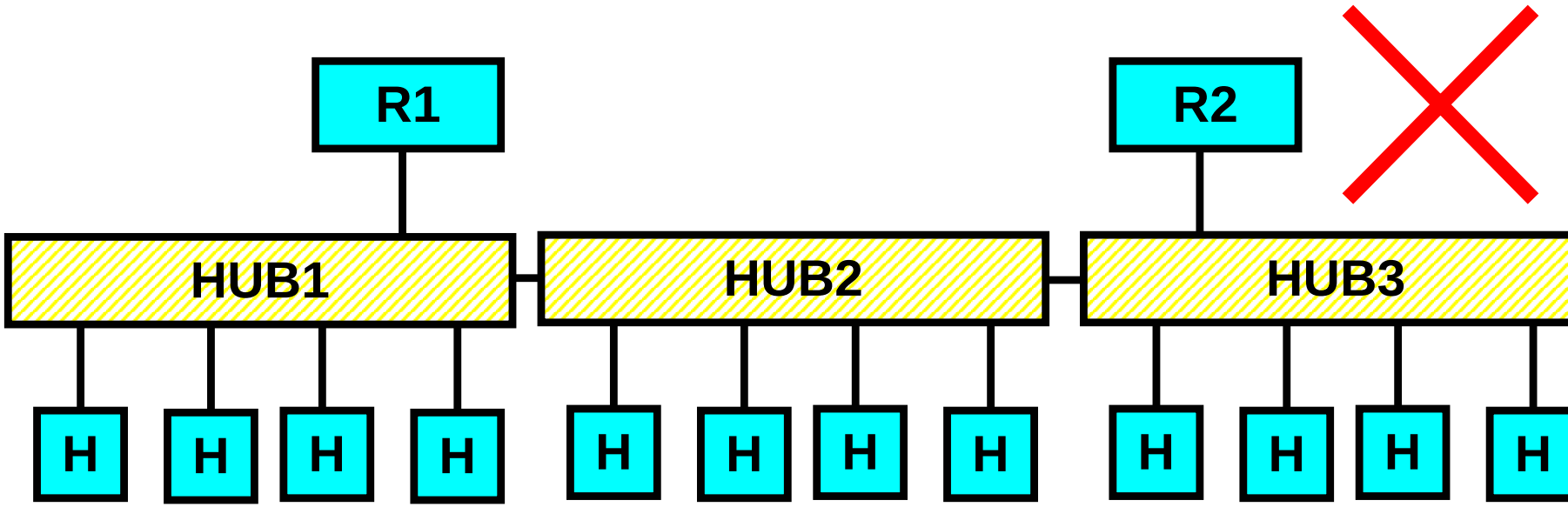
- 標準的なカスケードの無いL2ネットワーク
- ホストの数だけ多ポートのHUBを用意する必要がある
- 遅延は少なく、L2障害発生の可能性も低く安定している
- HUBのカスケード接続は許可しないため、L2ループなどの障害発生を防止できる
- 1台のHUBの障害で1つのL2ネットワークが全滅してしまうため、完全な冗長化とはならない

冗長化L2ネットワーク構成 2台カスケード



- 2台カスケードによるL2ネットワーク
- カスケードポイントに障害が発生するとL2ネットワークが分断し、正常に冗長化できない
  - この区間のみSTPやチャネル接続などにより冗長化をはかる方法もある
  - STPを利用することで、STP要因による障害も発生する可能性があるため、カスケードケーブルが短い場合には冗長化せずに接続する方がよい
- 1台のHUBが故障してももう1台のHUBによりL2ネットワークの一部が動作し続ける
- ホストも冗長化し、2台のHUBに接続すれば完全な冗長化を図ることが可能

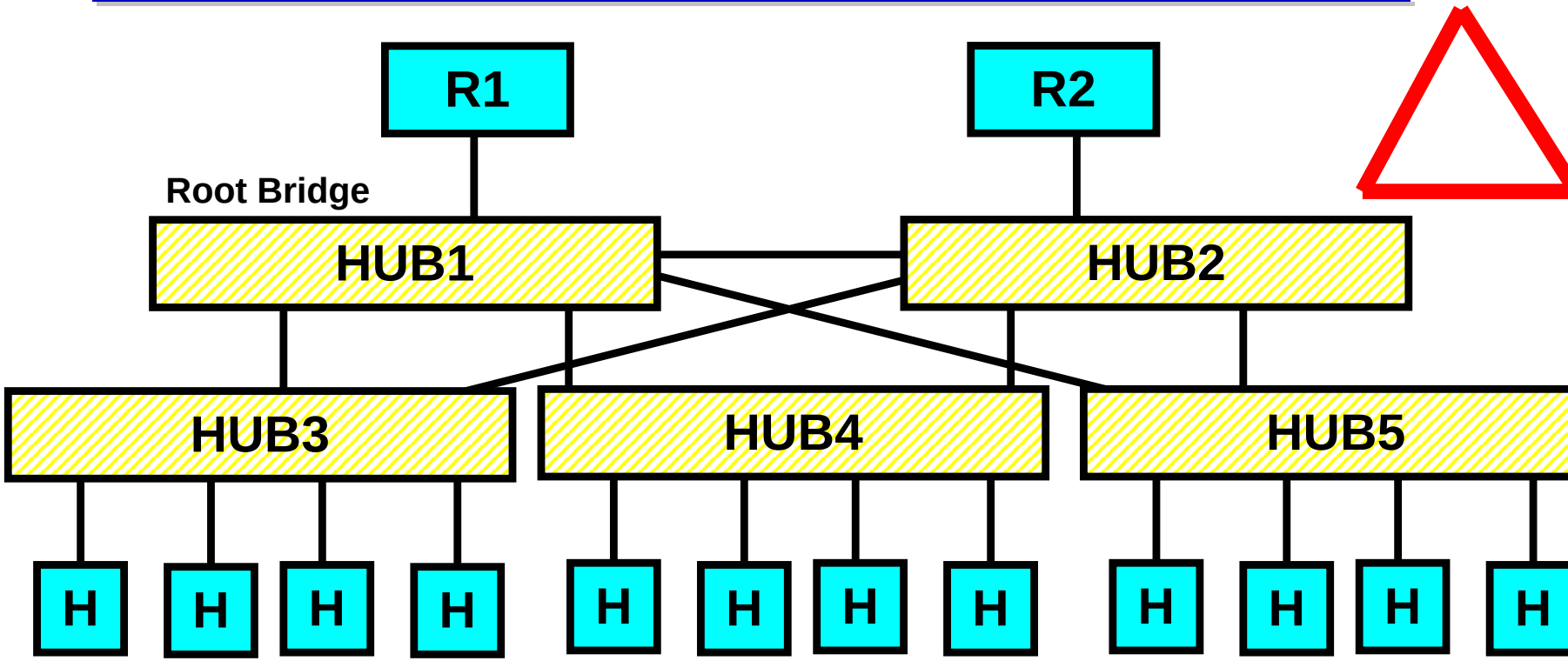
冗長化L2ネットワーク構成 3台以上カスケード



- 3台以上カスケードによるL2ネットワーク
- 安価なHUBをカスケード接続利用できる
- 障害発生時に障害箇所の特定が困難
- HUB2に障害が発生するとL2ネットワークが分断し、正常に冗長化できない
- 3台以上をカスケードし、冗長化を図るにはSTPが必要になる

IIJ
Internet Initiative Jap

冗長化L2ネットワーク構成 STP利用



- STPを利用したカスケードによるL2ネットワーク
- 安価なHUBをカスケード接続利用できる
- 障害発生時にはSTPの状態を確認するスキルが必要
- カスケードポイントに障害が発生してもSTPにより冗長化される
- STPの特性により、障害時の切り替え時間、切り戻り時の通信不通時間が長くなる

IIJ
Internet Initiative Jap

L2ネットワーク構成のまとめ

- L2ネットワークは極力カスケードしないほうがよい
- 多くのホストを収容する場合には多ポートのHUBを利用する
- 多ポートのHUBでも収容できない場合にはサブネットを分割し、異なるL2/L3ネットワークに収容する
- やむを得ずカスケードする場合にはツリー型カスケードとする
- L2ループを防止するため、カスケードするHUBを管理する
- 冗長化ネットワーク構成を組む場合にはカスケードは2台までとする
- 3台以上のカスケードを行う必要がある場合にはSTPを利用する必要がある

配線部材（メタル）

Ethernet規格と対応ケーブル

	カテゴリ3	カテゴリ5	カテゴリ5e	カテゴリ6
10BASE-T	○	○	○	○
100BASE-TX		○	○	○
1000BASE-T			○	○
1000BASE-TX				○

※カテゴリ5e:エンハンスト・カテゴリ5

- エッジスイッチ-端末間
  - カテゴリ5eを推奨
  - カテゴリ6であればすべての規格に対応するが、コスト高となる
- ルータ-エッジスイッチ間
  - カテゴリ6を推奨
  - 1000BASE-TXを利用しない場合でも今後のLANの高速化への対応と下位規格での安定した品質を提供できる

配線部材（光ファイバ）

Ethernet規格と対応光ファイバ

	MMF	SMF	DSF
100BASE-FX	○		
1000BASE-SX	○		
1000BASE-LX	△	○	
1000BASE-ZX		△	○

MMF: Multi-mode Fiber

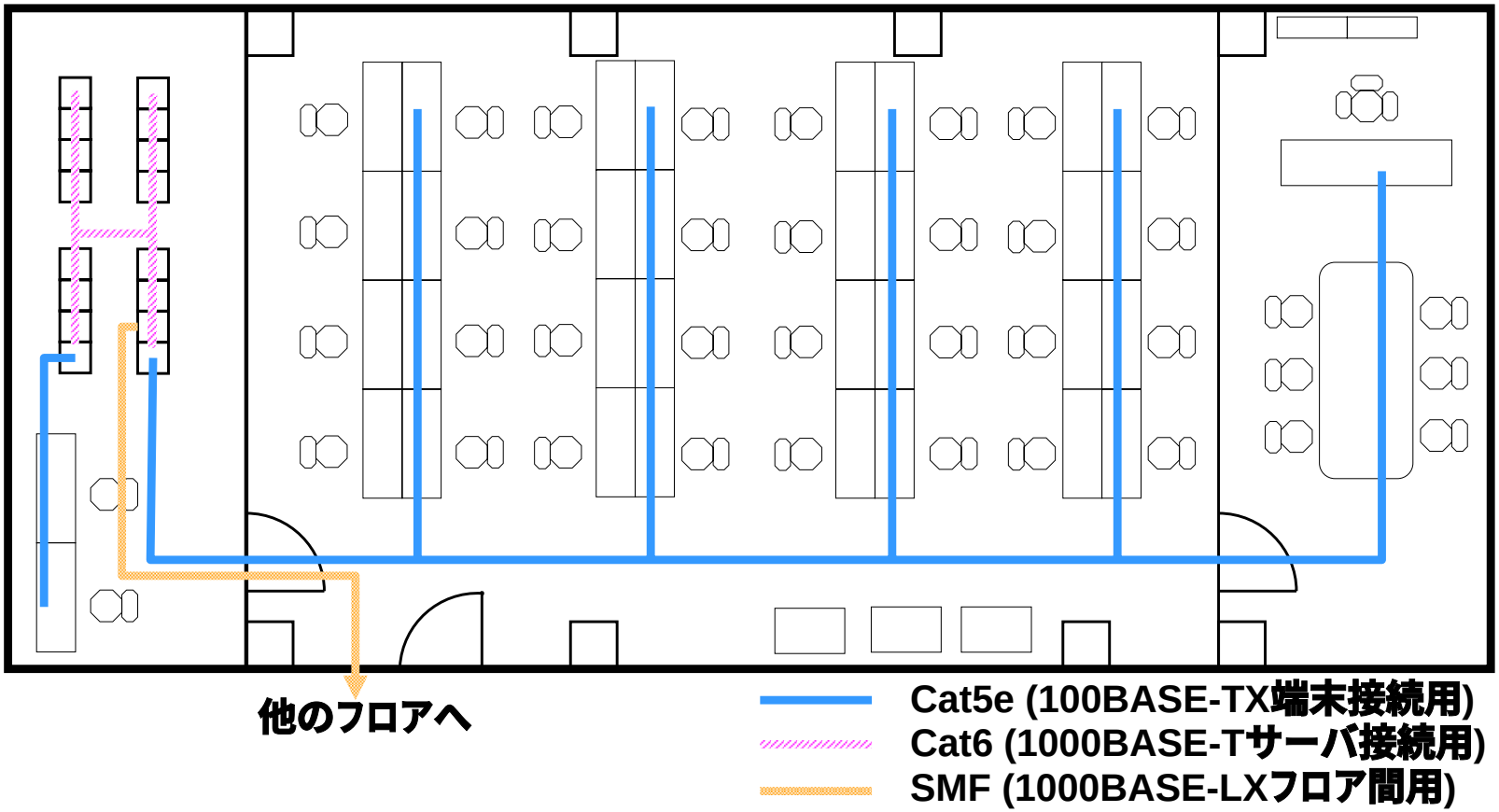
SMF: Single-mode Fiber

DSF: Dispersion Sifted Single-mode Fiber

- 同一ラック内、同一フロア内
  - MMFを推奨
  - SMFでも支障はないが、安価な1000BASE-SX対応スイッチ利用できなくなるため、コスト高となる
  - MMF新規敷設であれば50/125μm GI(Graded Index)ファイバを選択
- 異なるフロア間
  - SMFを推奨
  - MMFでは長距離を引き伸ばすことができない
  - SMFはWAN回線の延長や事前敷設、10GbEにも利用可能
  - SMFは多くのメディアコンバータにも対応しており、100BASE-TXや1000BASE-Tを安価に延長できる
- 異なるビル間
  - SMFを推奨。長距離となる場合にはDSFを検討しても良い。

IIJ
Internet Initiative Jap

配線例



- 適切なケーブル敷設により低コストと高い拡張性を実現する
  - 端末接続用にはCat5eで100BASE-TXを利用　(1000BASE-Tまで対応可能)
  - サーバ接続用にはCat6で1000BASE-Tを利用
  - フロア間接続にはSMF(光ファイバ)で1000BASE-LXを利用

ネットワークトラブルシューティング1

- Ethernetを利用したLANや回線で遅かったり、エラーが出る
  - Duplexミスマッチ
    - 対向となる通信機器のDuplexが異なることによる通信エラー
    - Late collisionなどが検出される
    - Full Duplex-Full Duplex(全二重同士)/Half Duplex-Half Duplex(半二重同士)など、おなじDuplexに設定することで問題は解消する
    - Autoに設定するとHalf Duplexとなる機器
    - Autoに設定しないとFull Duplex動作しない機器
    - Full Duplexに設定するとエラーが出る機器
  - ケーブル不良
    - ケーブル自体、コネクタ、継ぎ手、パッチパネルなどの不良による品質劣化による通信エラー
    - CRCエラーなどが検出される
    - ケーブルの交換、継ぎ手やパッチパネル区間を無くすか品質の高いものに交換することで問題は解消する
    - 継ぎ手やパッチパネルなどは極力なくして配線したほうがよい
    - 市販ケーブルであってもクロストークなどが測定できるケーブルテスターでテストを行う

ネットワークトラブルシューティング2

- Ethernetを利用したLANや回線で遅かったり、エラーが出る(続き)
  - STPに起因する問題
    - 利用していないSTPによる通信エラー
    - ネットワーク高負荷時にCRCエラーが0.01%程度観測される
    - STPをoffにすることで問題は解消する
- HSRP/VRRPが一時的に両方がアクティブなり、誤動作する
  - STPに起因する問題
    - STPネゴシエーション時にリンクはあがっている状態にも関わらず通信できない状態が発生し、このとき、HSRP/VRRPが誤動作する
    - STPをoffもしくはportfastに設定する
  - VLAN trunkに起因する問題
    - VLAN trunkのネゴシエーション時に、リンクはあがっている状態にも関わらず通信できない状態が発生し、このとき、HSRP/VRRPが誤動作する
    - VLAN trunk上でHSRP/VRRPを利用しないようにネットワークを変更する
    - HSRP/VRRP timerを調整し、hold timeをネゴシエーション時間より長くする

まとめ－1

- データリンク層とネットワーク層の違い
  - データリンクフレームは中継が起こる毎に変化する
  - IPデータグラムは変化しない
  - データリンクフレームの宛先＝IPデータグラムの宛先とは限らない
- ハブとスイッチ、スイッチとルータの違い
  - スイッチはハブに比べLANを有効に利用できる
  - スイッチのみのネットワークはbroadcastに脆弱で、ウイルスを原因としたLAN輻輳を起こしてしまう
  - ルータを利用することでbroadcast floodを回避し、ウイルスを原因としたLAN輻輳を回避でき、障害に強ネットワークを構築することができる
- インターネット接続にはルーティングは必須
- サーバなどの安全性を要求されるものは別のセグメントに配置する
- ネットワークの拡張を考慮したアドレス割り当てポリシーで運用する

まとめー2

- 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－3

- 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
- L2ネットワークでは極力カスケード接続を避ける
- 冗長化L2ネットワークでは2台カスケード構成とする
- 配線はできるがきりパッチパネルを利用せず、I/Fのエラーの状況からトラブルを解決する