

# 経路制御入門

～ネットワーク設計の基本～

2000年12月21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

- 経路制御とは何か
- 経路制御がなぜ必要となるのか
- ネットワークの設計に経路制御をどのように生かせば良いか



## 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
- 静的経路制御と動的経路制御
- 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
- ネットワーク設計
- アドレスの割り当てポリシー

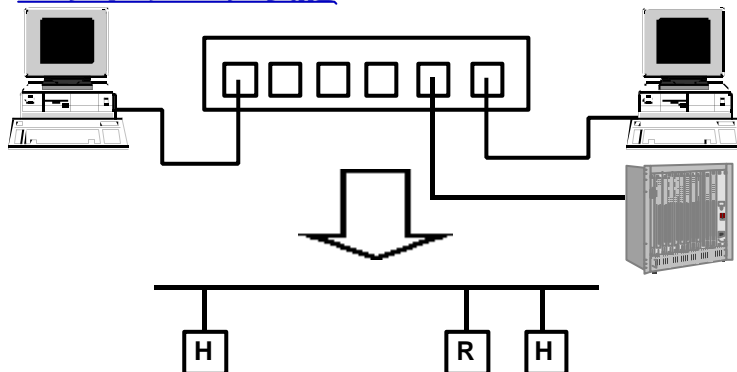


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

3

## ネットワーク表記



- ハブ、スイッチなどは1本の線で表わします。
- ホストはH、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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

4

## 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
- 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

### OSI参照モデル

7	アプリケーション層
6	プレゼンテーション層
5	セッション層
4	トランスポート層
3	ネットワーク層
2	データリンク層
1	物理層

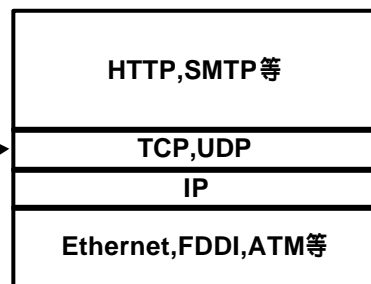


OSIレイヤ

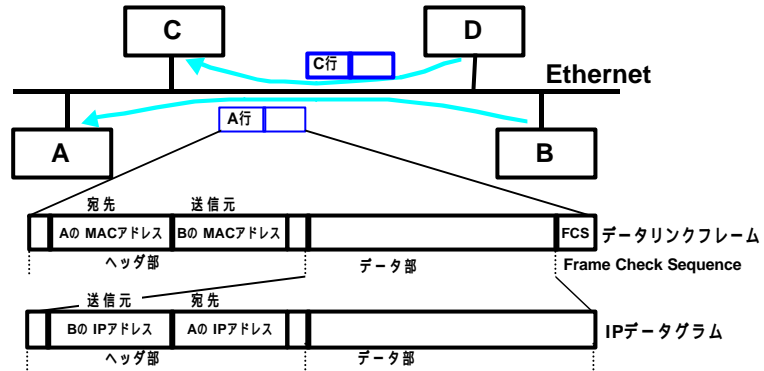
レイヤ2 : データリンク層

レイヤ3 : ネットワーク層

### TCP/IP



## Ethernetを流れるIPデータグラム



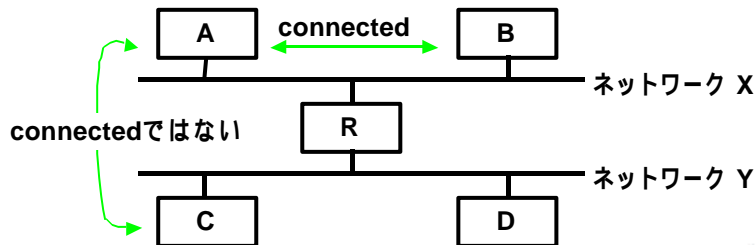
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

7

## 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



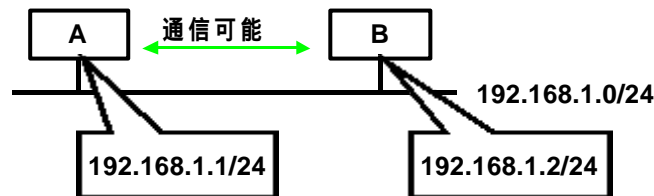
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

8

## 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
  - 192.168.1.1/24
- Aから見たConnectedなアドレス空間
  - 192.168.1.0 ~ 192.168.1.255
- Bに192.168.1.2 ~ 192.168.1.254のアドレスを付ける
  - Bに192.168.1.2を付ける
  - A-B間の通信が可能



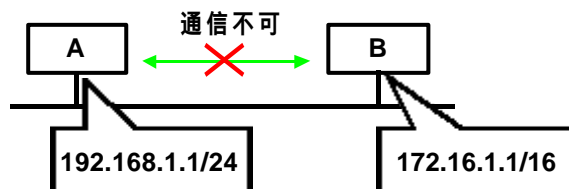
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

9

## 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
  - 192.168.1.1/24
- Aから見たConnectedなアドレス空間
  - 192.168.1.0 ~ 192.168.1.255
- Bに192.168.1.2 ~ 192.168.1.254以外のアドレスを付ける
  - A-B間の通信ができない



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

10

## Connected ではないネットワーク - 1

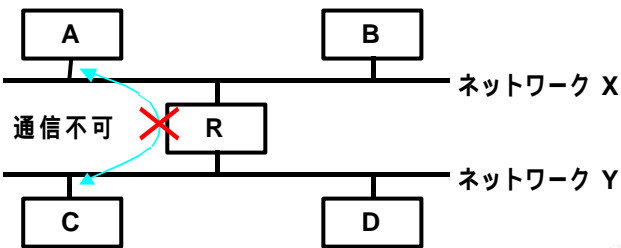
- A、Cはそれぞれ異なるネットワークに接続しているため connected ではない
- ルーティング設定なしではA、C間の通信はできない

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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

11

## Connected ではないネットワーク - 2

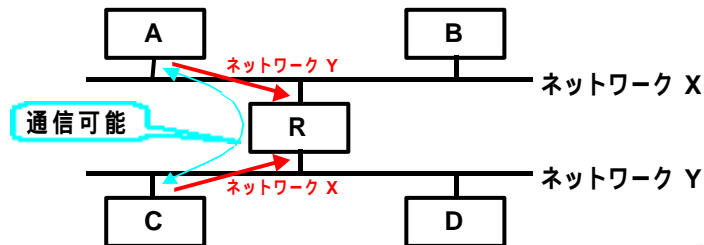
- ルーティング設定を行なう
  - A: ネットワークYをRにルーティング
  - C: ネットワークXをRにルーティング
- これにより、A C間の相互通信が可能となる
  - RはA,C共に connectedなため、アドレスを設定するだけで通信が可能

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

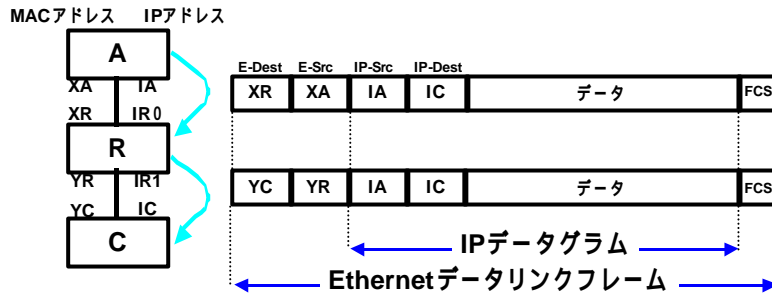


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

12

## 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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

13

## 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的地アドレス、宛先アドレスとして扱われる。Destination(デスティネーション)と宛先アドレスとで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」



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

14

## 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

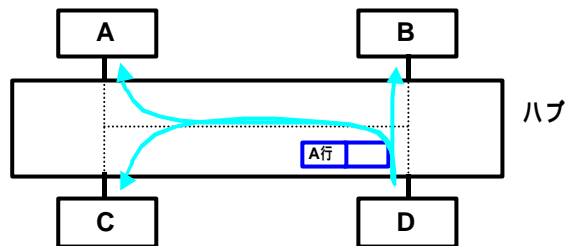
- ハブとスイッチの機能の違い
- スイッチを有効に使う方法
- ルータを利用するための設定
- ネットワーク設定の自動化
- スイッチとルータの違い
- スイッチの耐障害性
- ルータの耐障害性
- Broadcast flood問題





## 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



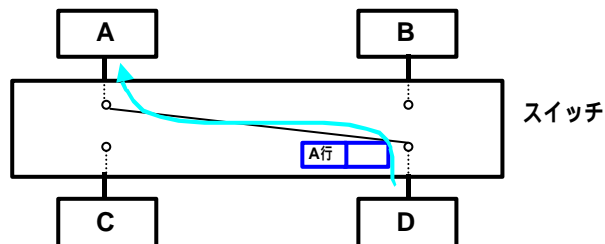
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

17

## 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

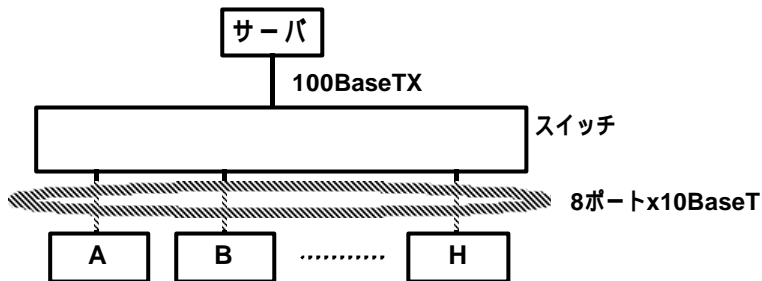


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

18

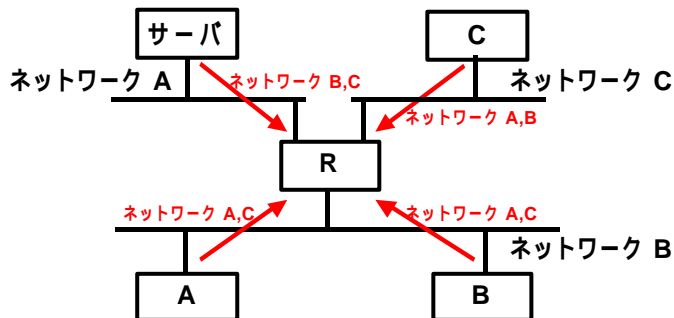
## スイッチを有効に使うには



- 主にサーバ、ホスト間のトラフィックの場合に有効
- A } それぞれ10BaseTをフルに利用可能  
 : }  
 H } サーバ



## ルータを利用するための設定



- 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



## 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

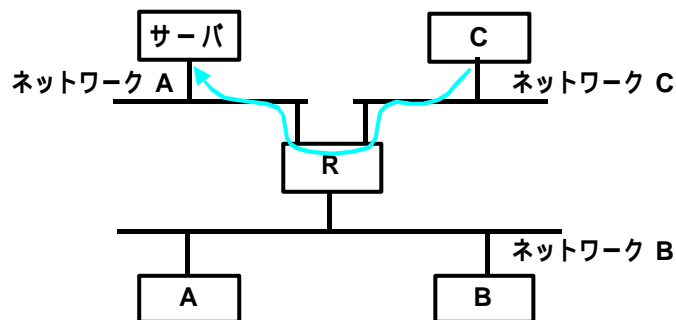


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

21

## 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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

22

## スイッチとルータの機能の違い

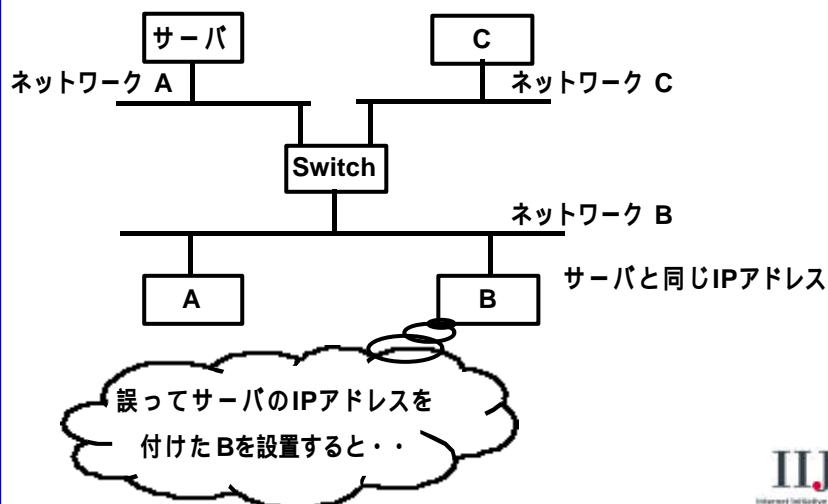
- ハブとスイッチの機能の違い
  - スイッチは異なるポートの通信を他のポートに伝播しない
- スイッチとルータの違い
  - ルータは異なるネットワークの通信を他のネットワークに伝播しない
  - スイッチとは異なり、ルーティングの設定が必要
  - サブネット分割が必要
- スイッチを有効に使うには
  - トラフィックが集中するようなポートにはスイッチを導入する



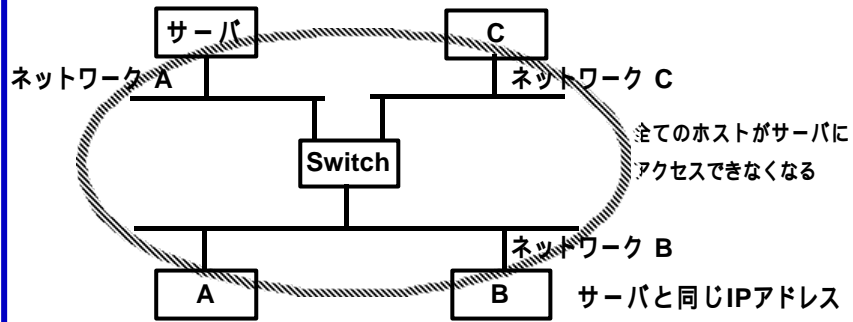
次に問題点について検討する



## スイッチの耐障害性-1



## 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

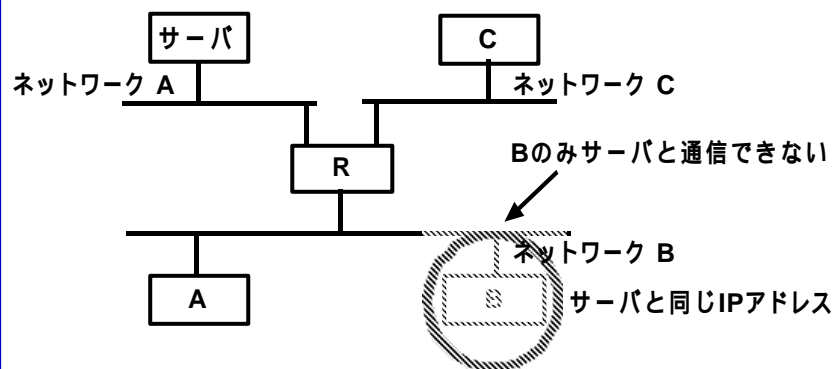


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

25

## 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

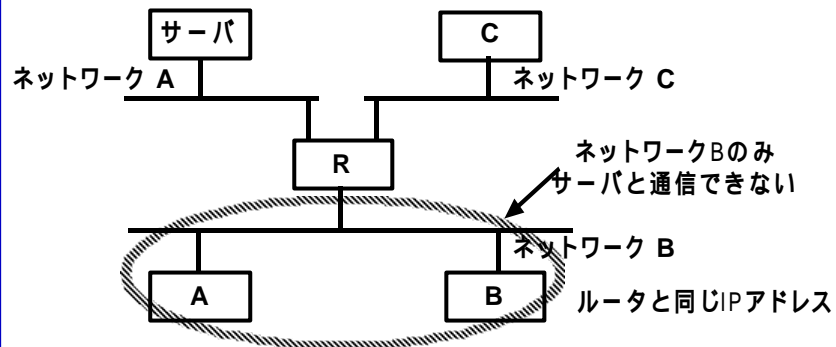


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

26

## 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

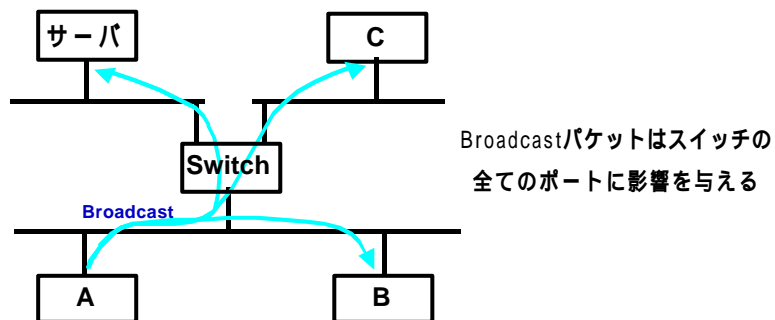


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

27

## Broadcast Flood - 1



- 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

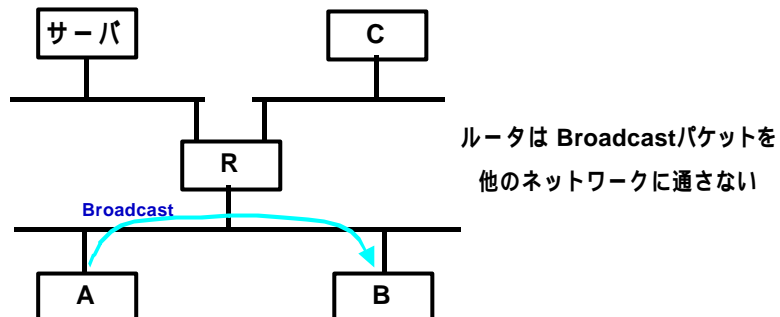


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

28

## Broadcast Flood-2



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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

29

## 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
  - 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しやすい
- 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

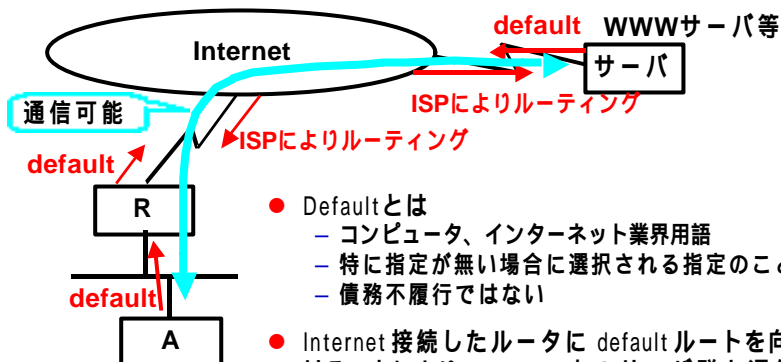


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

30

## 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



## 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

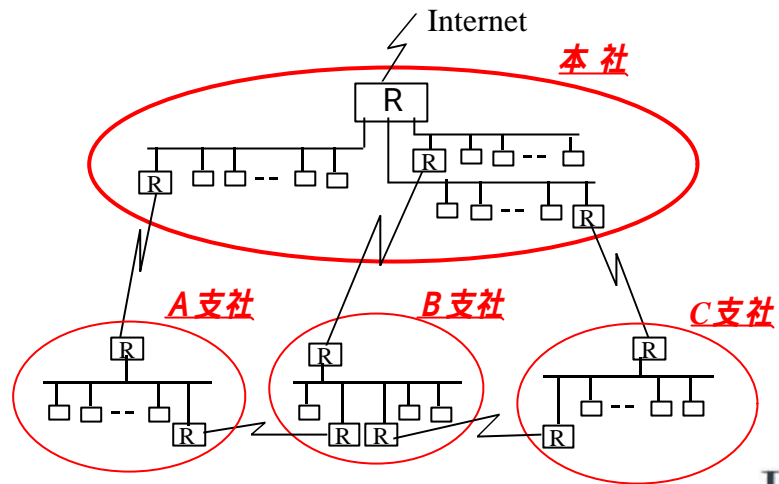


## 動的な経路制御を選択しなければならない理由-1

- ネットワークの変化に対応しなければならない
  - 一部の追加が全体の変更になることを防ぐ
- 責任者の異なるネットワークを接続する必要がある
  - 複数の管理ネットワークとの接合
- ルータの設定を容易にする
  - 大規模ネットワークを手作業で管理することは難しい



## 複雑に延びるネットワーク



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

35

## 動的な経路制御を選択しなければならない理由-2

- 自動的に最適経路を選択できる
  - 管理できないほど複雑なネットワークポロジ-
- 自動的にバックアップ経路を選択できる
  - 死守するネットワークが存在する
  - 障害時に強い構成を考える

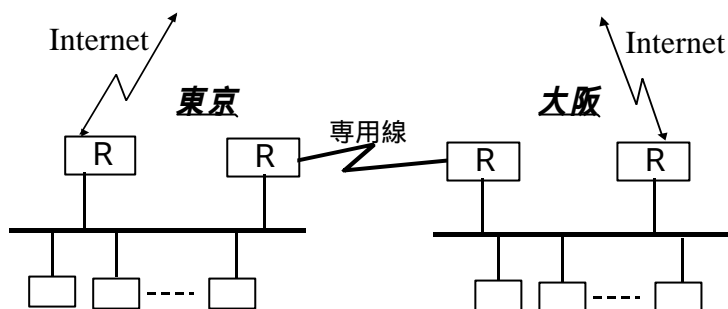


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

36

## 東京、大阪バックアップ

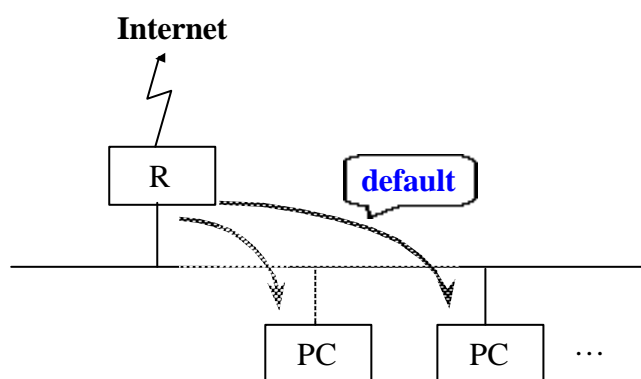


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

37

## ダイナミックルーティング：経路情報の伝播

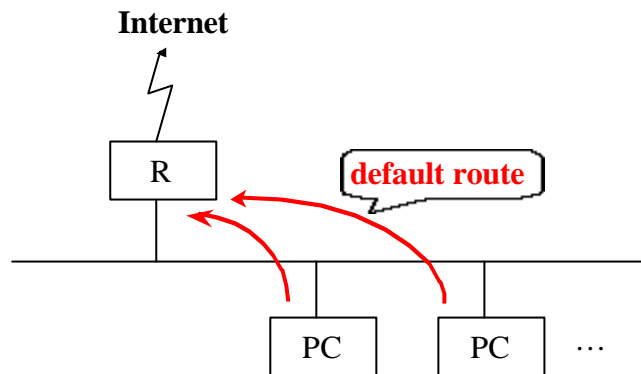


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

38

## ダイナミックルーティング：伝播後の経路情報



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

39



## ダイナミックルーティングプロトコルの種類

- RIP
  - RFC1058
- RIP2
  - RFC2453
- OSPF
  - RFC2328
- BGP4
  - RFC1771

2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

40



## 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



## 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)

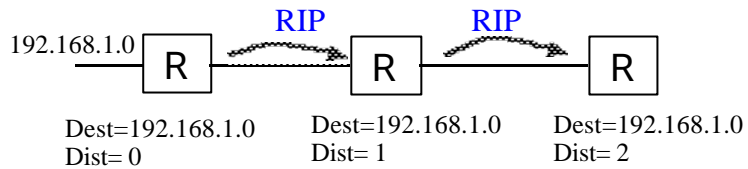
vector=destination(ネットワーク)

distance=HOP count(通過したルータの数)





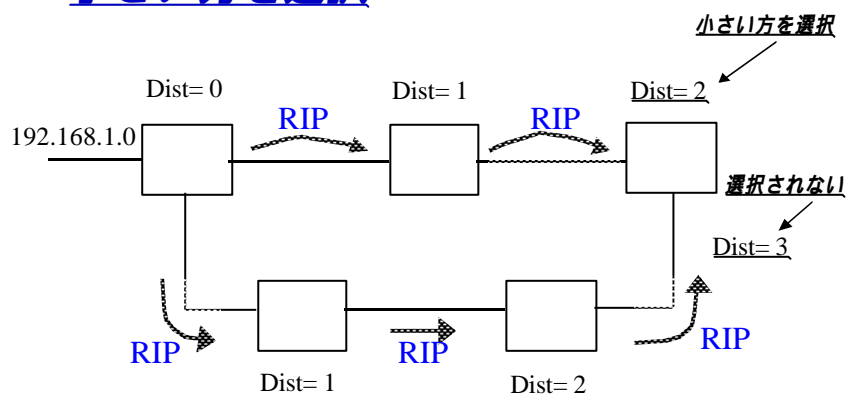
## ルータを通る度にdistanceが1追加される



Dest=Destination  
Dist= Distance



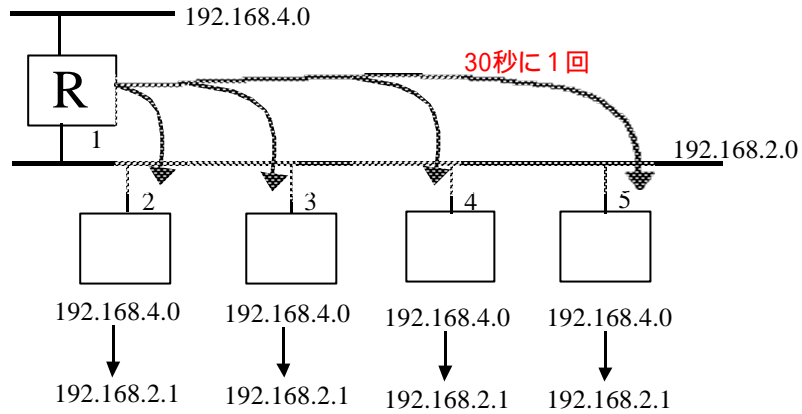
## 同じdestinationの場合はdistanceが小さい方を選択



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

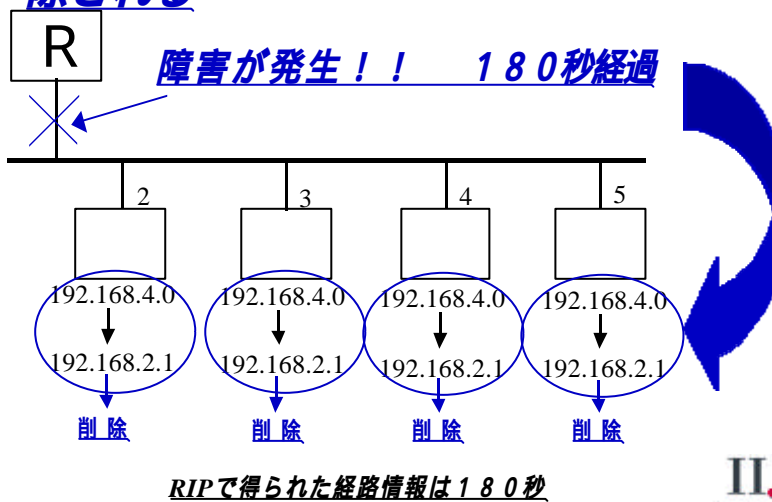


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

51

## 3分間経路が到着しないと経路は削除される



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

52

## 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

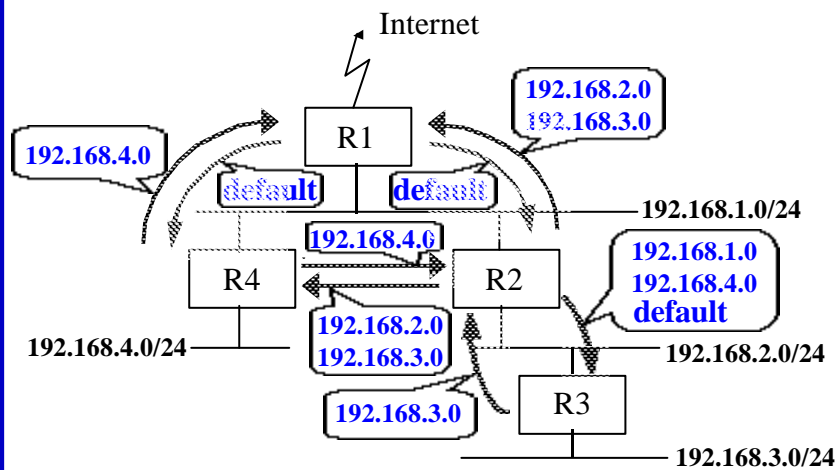


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

53

## RIP伝播

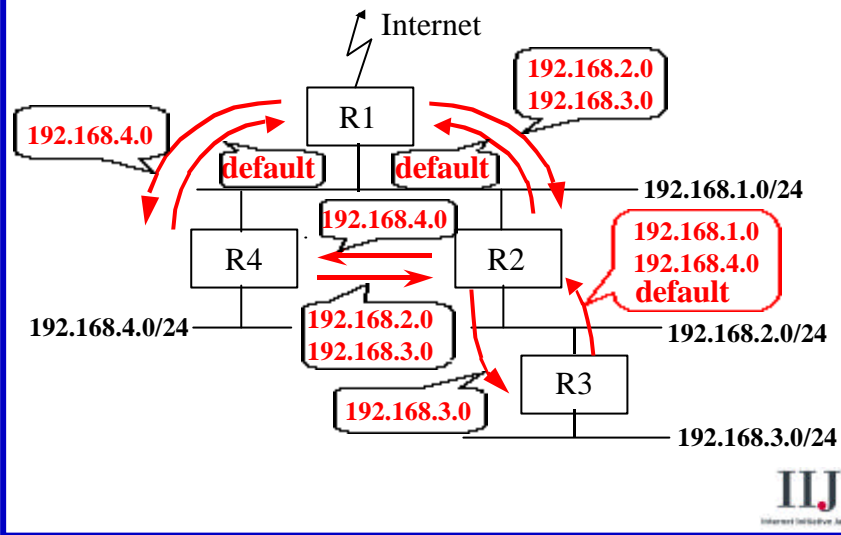


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

54

## RIP 伝播後の経路情報



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

55

## 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

IIJ  
Internet Initiative Japan

2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

56

## 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

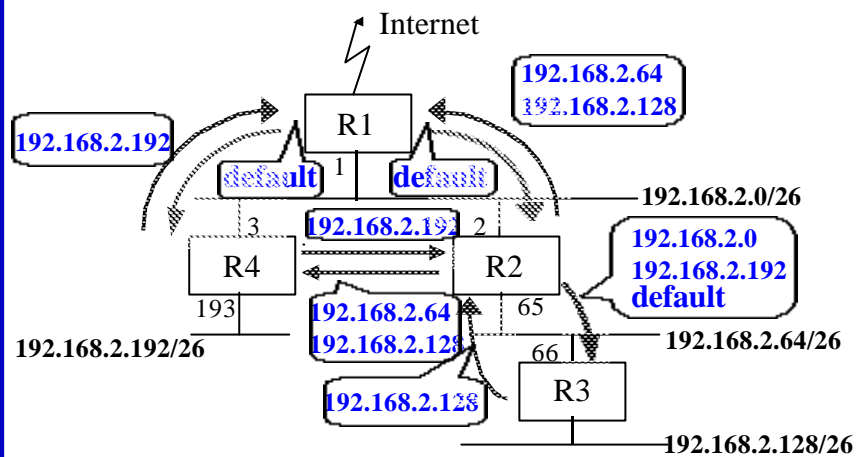


## RIP のまとめ-2

- 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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

59

## 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

2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

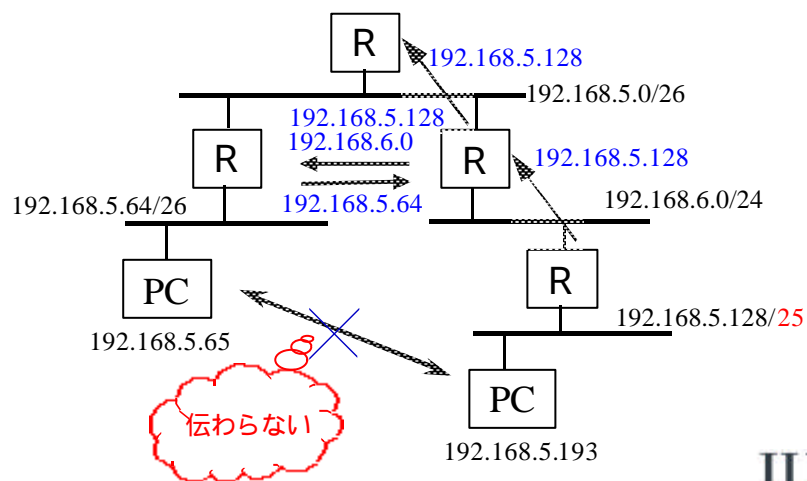
60

## 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

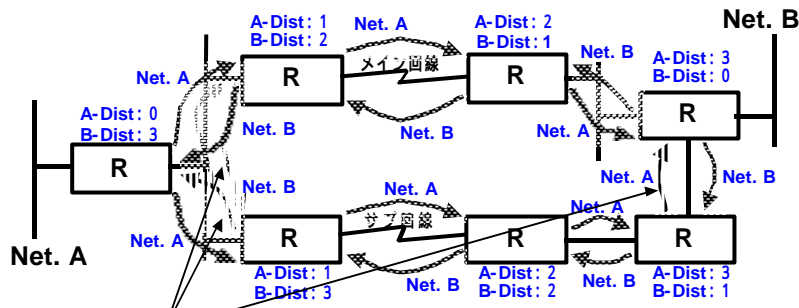


## トラブルシューティング- RIP が伝播しない-2

- 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)

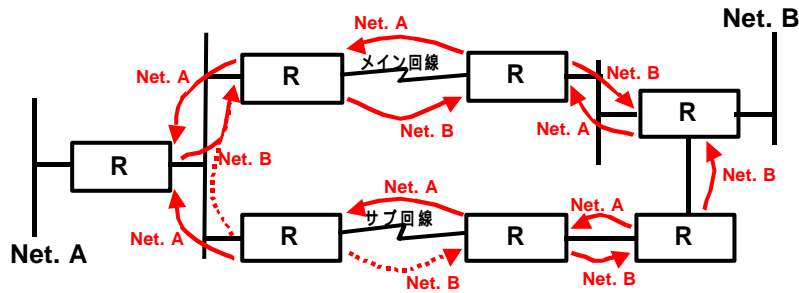


他方よりも Distance が  
大きいため選択されない

- 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



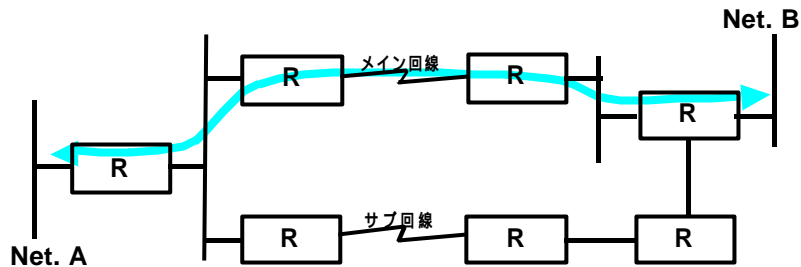
### RIPを用いたバックアップ-ルーティングテーブル(定常時)



- 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

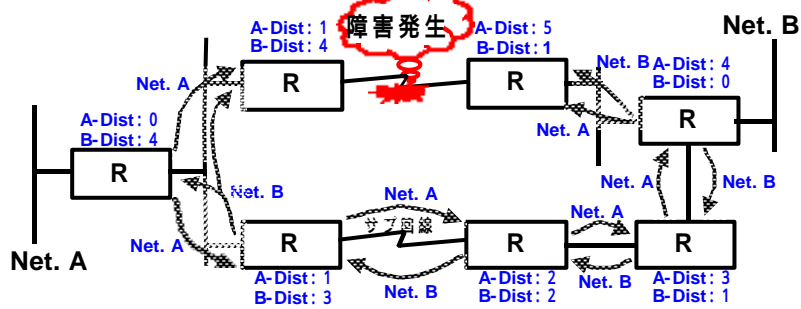


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

69

### 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

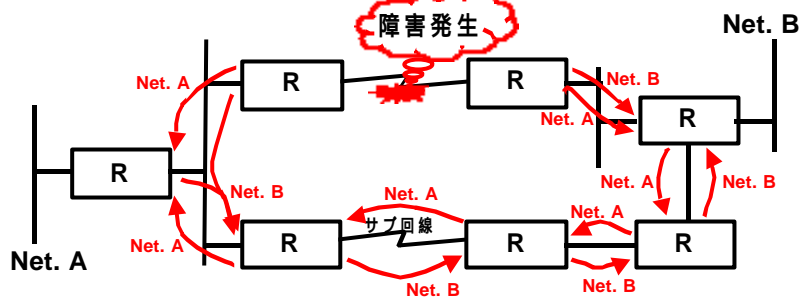


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

70

### 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

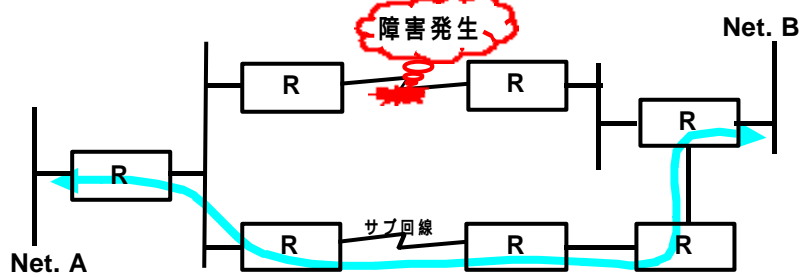


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

71

### 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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

72

## 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の代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恐れがあるため、バックボーンなど重要なインターフェースは明示的に設定したほうが安全

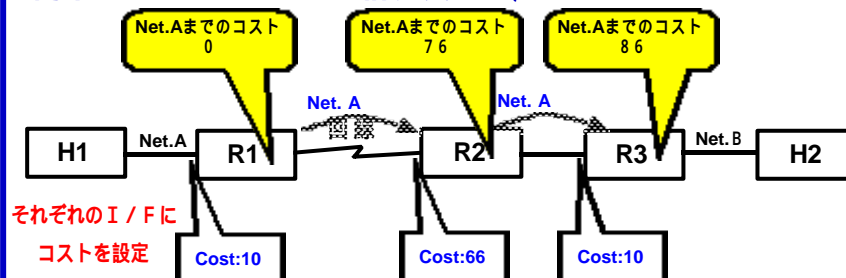


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

75

### 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

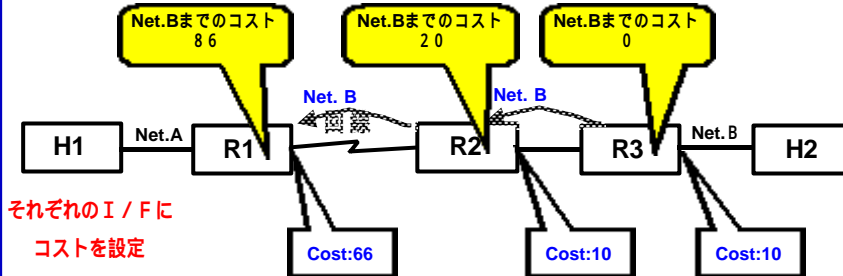


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

76

### 簡単なOSPFコストの計算法 - 2



それぞれのI/Fに  
コストを設定

- 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

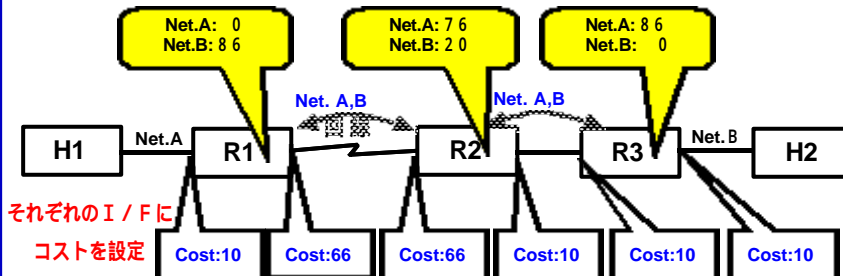


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

77

### 簡単なOSPFコストの計算法 - 3



それぞれのI/Fに  
コストを設定

- 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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

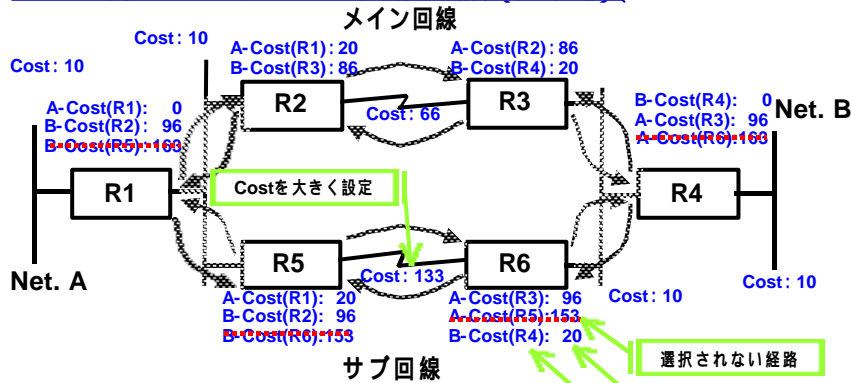
78

## 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)

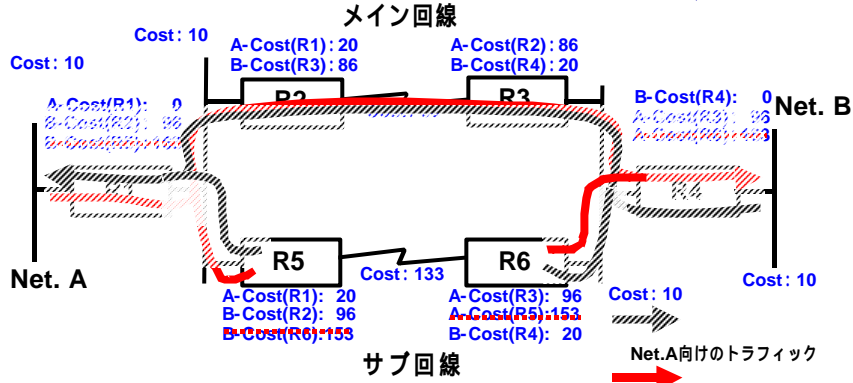


- OSPFを利用して、通常時はメイン回線のみを利用する
- 障害時にはサブ回線を利用してバックアップを行う





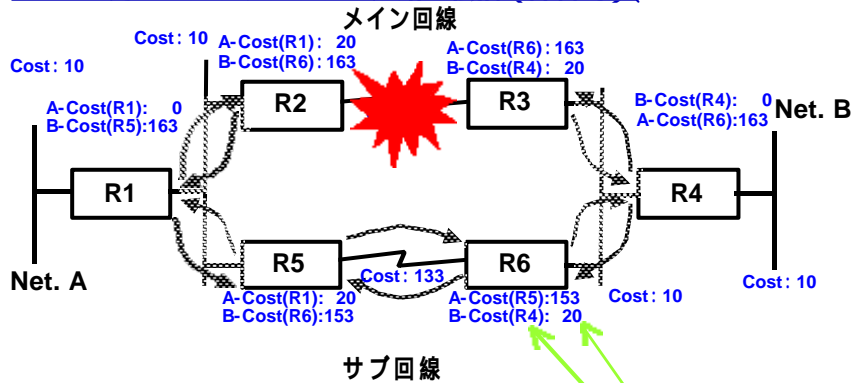
### OSPF を用いたバックアップ-トラフィックの流れ (通常時)



- 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)

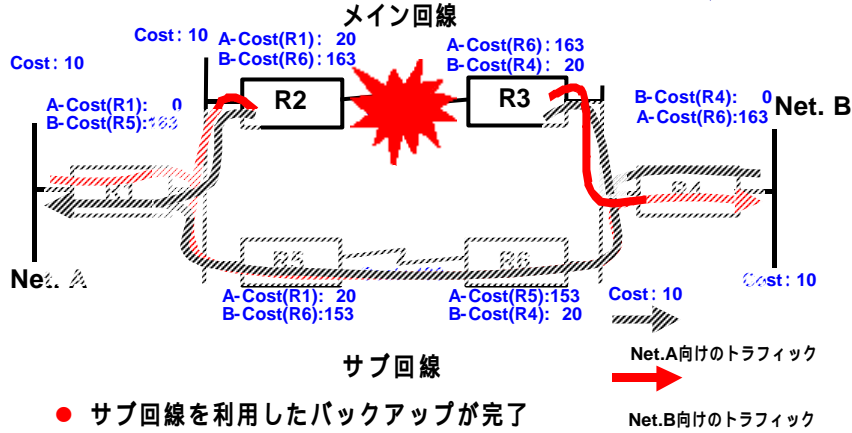


- 回線の切断によりR2-R3間のネットワークが削除される

コスト値
伝播元ルータ名 (NEXT HOP)



### OSPF を用いたバックアップトラフィックの流れ (障害時)



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

83

### OSPF バックアップルーティングの特徴

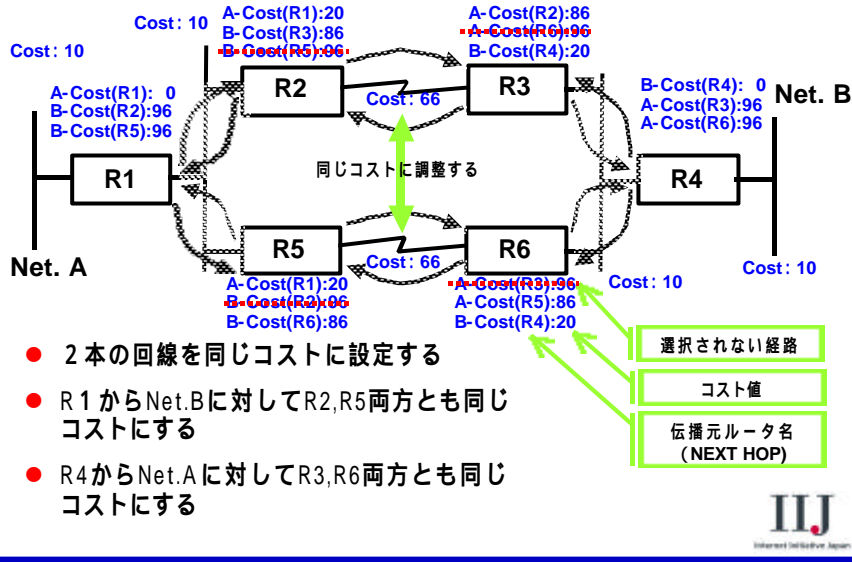
- RIPとは異なり、すばやいバックアップが可能
- バックアップ用の回線上もOSPF HELLOが流れるため、サブ回線を切断することはできない
  - ISDNなどでバックアップさせるにはOSPFだけのチューニングでは難しい
- 2本の回線を別々の用途に利用して障害時にそれぞれバックアップとして利用することが可能

2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

84

### 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)



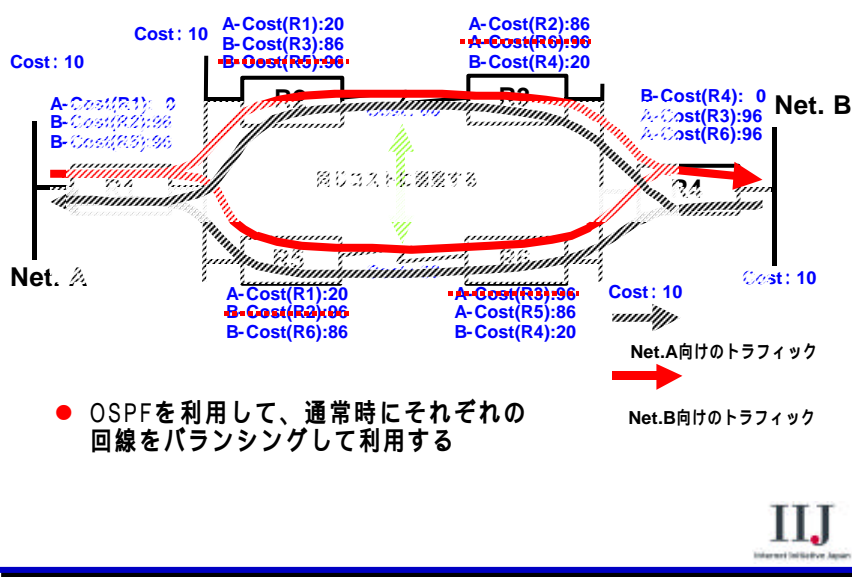
- 2本の回線を同じコストに設定する
- R1からNet.Bに対してR2,R5両方とも同じコストにする
- R4からNet.Aに対してR3,R6両方とも同じコストにする

2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

85

### 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)



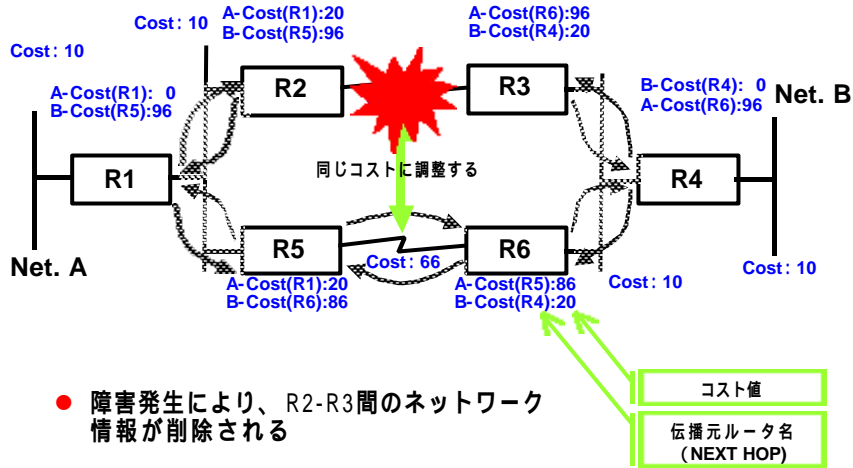
- OSPFを利用して、通常時にそれぞれの回線をバランシングして利用する

2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

86

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

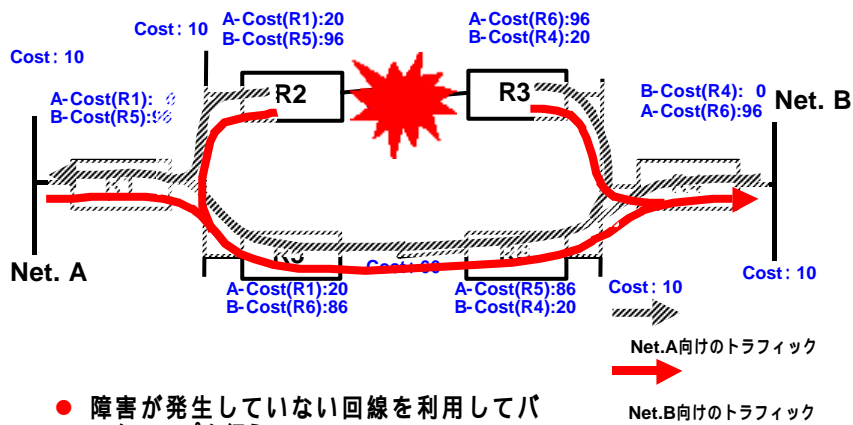


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

87

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



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

88

## 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



## 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
- デフォルトルート
  - 必ずstaticなどでデフォルトルートを確認してから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の経路を受け取るルータまでの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で表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

### ● ルータを立ち上げる順番

#### – 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るほうが良い。
- 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。



## トラブルシューティング-RIP v2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への移行が望まれる
- 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
- ATM障害検出

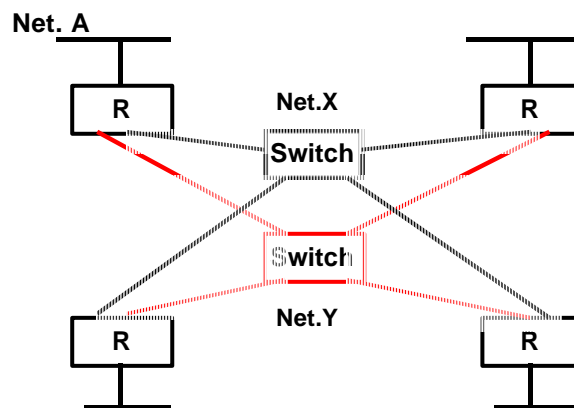


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

95

## 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



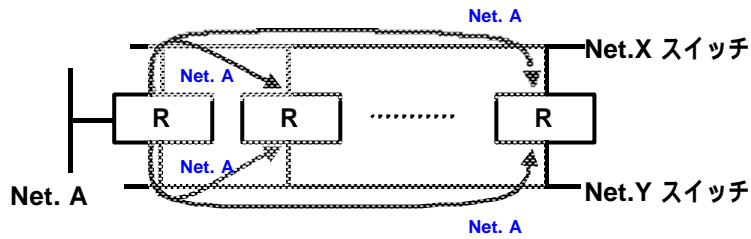
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

96



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

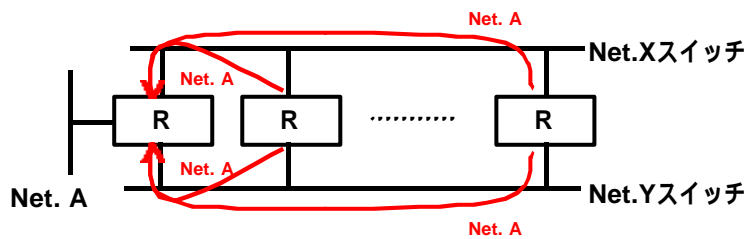


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

97

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

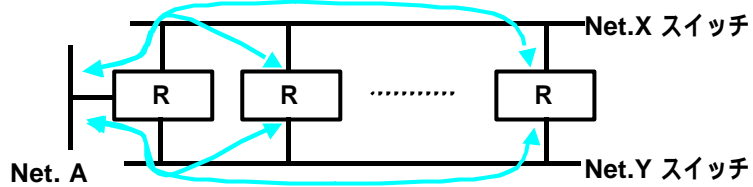


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

98

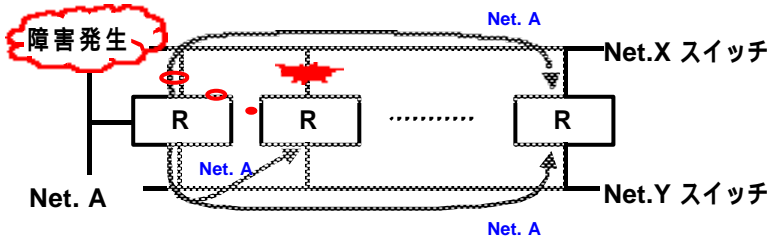
デュアル構成 + OSPF を用いたバックアップ、バランシング  
- トラフィックの流れ (通常時)



- 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



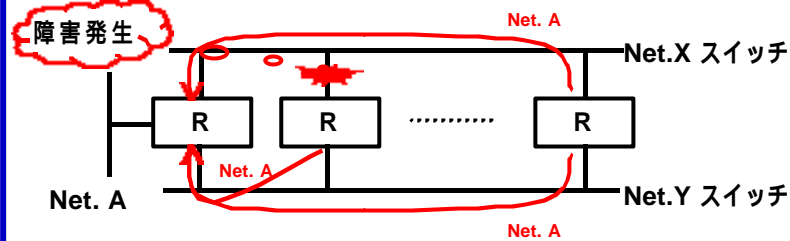
デュアル構成 + OSPF を用いたバックアップ、バランシング  
- 経路の伝播 (障害時)



- 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、บาลancing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

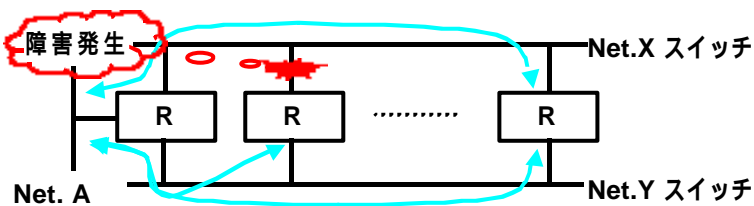


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

101

## デュアル構成+OSPFを用いたバックアップ、บาลancing-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

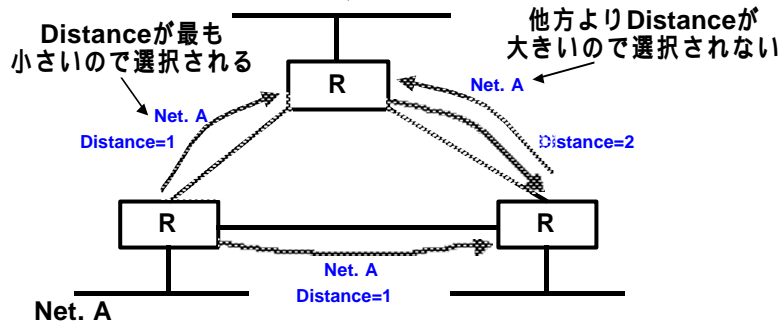


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

102

## 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

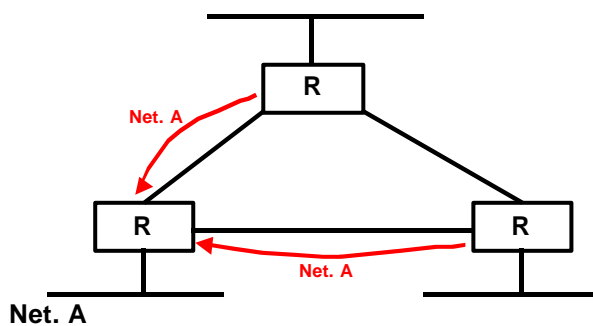


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

103

## 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

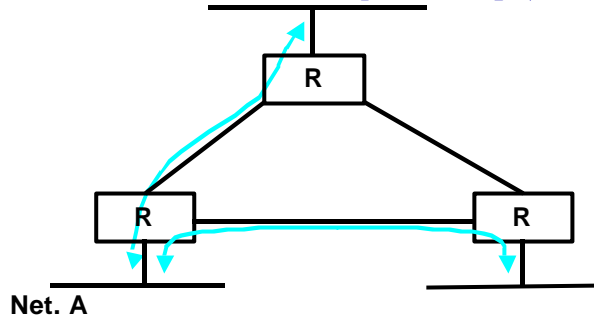


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

104

## 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

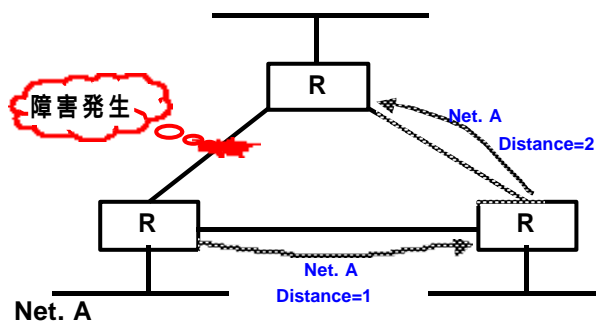


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

105

## 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

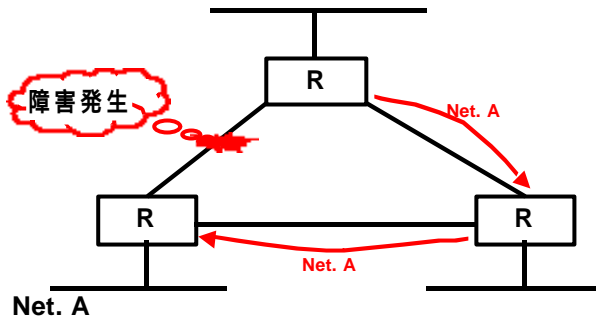


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

106

## 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

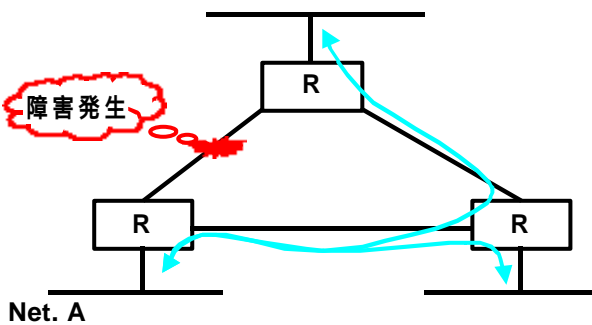


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

107

## 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

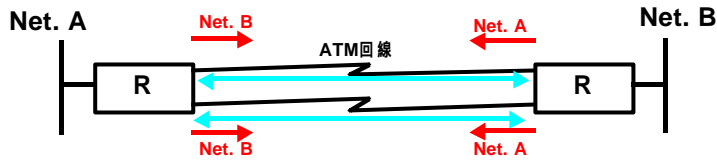


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

108

## ATM障害検出-1



- VPのdownを検出して自動的にインターフェースをshutdownすることができない(cisco IOS11.X)
- このため、上記のようにstaticにルーティングを設定し、2本のATM回線を束ねて利用する場合に、望むバックアップ動作が行われない



- この例では概ね50%のパケットをこぼしてしまう。

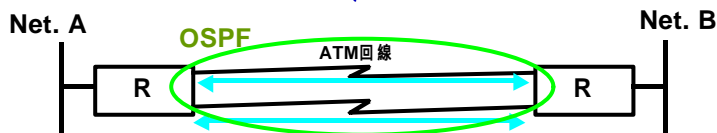


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

109

## ATM障害検出-2



- OSPFを用いて動的に経路制御してバランシングを行えば、ATM回線でも障害検出が可能となる



- OSPFが障害を検出し、回線を利用しなくなるのでパケットをこぼさない



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

110

## ダイナミックルーティング以外のバックアップ、バランシング技術

- STP(spanning tree protocol)
  - レイヤ2での冗長構成
  - 障害の発生から spanning tree変更までには10秒程度必要
- FDDI DAS(dual attachment station)
  - レイヤ2での冗長構成
  - ほぼ瞬時に切り換わる
- I/F downと static
  - I/Fの downを検出するとその i/fに向いている routingが消えることを利用した backup
  - ATM専用線等では回線断が I/F downとならないため、利用できない
    - OAMセルを keepalive代わりに利用することで、回線断を検出可能にしようとしている (IOS12.X)
- HSRP
  - 一つの仮想的な MACアドレスを複数のルータで共有することで、サーバ等でダイナミックルーティングを利用せずに障害時の切り換えを行う

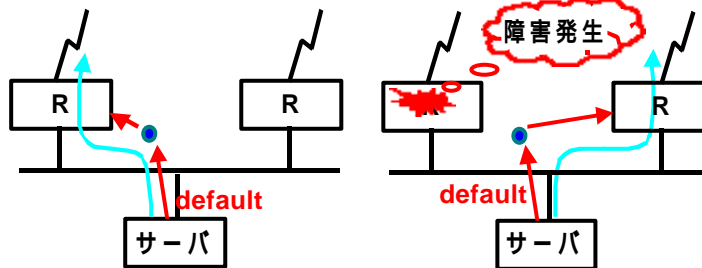


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

111

## HSRP-1



- 障害時には MACアドレスとルータの対応を変更するため
  - デフォルト設定では以下の停止が伴う
    - 切り換えに 10秒 (最近では3秒)
    - 切り戻しに 30秒 (最近では9秒)
  - スイッチ等にルータを接続している場合には、ポート、MACアドレスの対応に食い違いが生じるため、さらに切り換えに時間を要する場合がある



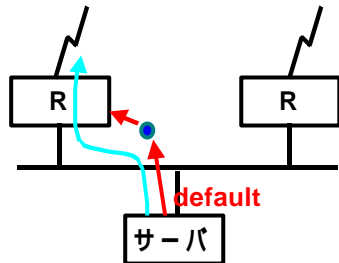
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

112



## HSRP-2



- HSRP+Interface Tracking (通常運用時)

- Interfaceの downを検出して、Trackingすることで回線障害時にactiveルータの切り換えを行う

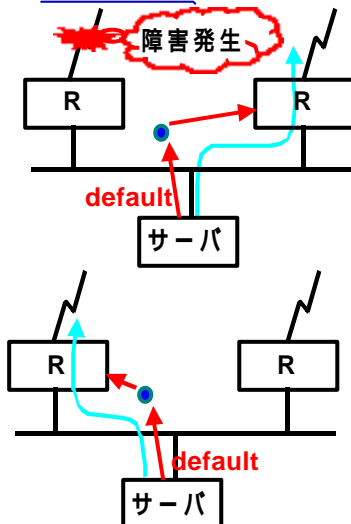


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

113

## HSRP-3



- HSRP+Interface Tracking (障害発生時)

- Interface Trackingにより切り替え
- 10秒間停止 (最近では3秒間)

- HSRP+Interface Tracking (障害復旧時)

- 復旧により切り戻しが発生
- 30秒間停止 (最近では10秒間)
- 切り戻し停止時間をなくすHSRP Delay機能が最新のファームウェアで実現されてきている

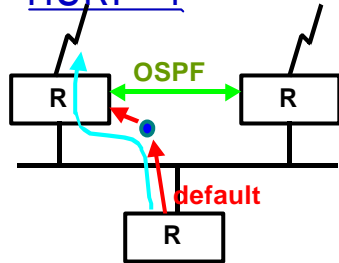


2000/12/21

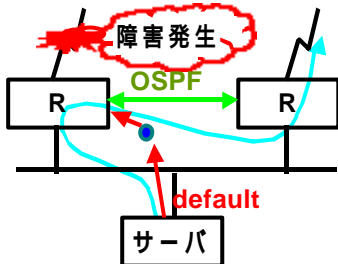
Copyright © 2000 Internet Initiative Japan Inc.

114

## HSRP-4



- HSRP+OSPF(通常運用時)
  - ルータ障害時のみHSRPを利用して、回線障害時にはOSPFを利用してバックアップを行う



- HSRP+OSPF(障害発生時)
  - 回線断により、OSPFにて切り換え
  - 1 ~ 3 秒間停止

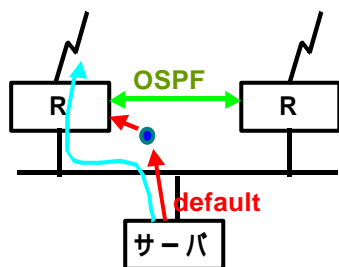


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

115

## HSRP-5



- HSRP+OSPF(障害復旧時)
  - 復旧したために切り戻し発生
  - 経路の切り換えのみなので停止なし

- ダイナミックルーティングは切り戻しに停止を伴わないため、ルータ間はOSPF等のダイナミックルーティングを利用したほうが良い



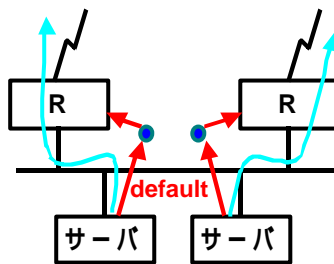
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

116

## MHSRP - 1

- マルチグループを用いて MHSRP を利用すれば、サーバ毎にトラフィックを分ける事ができる



- MHSRP (通常運用時)
  - それぞれのサーバは対応する HSRP の仮想アドレスに default を向ける

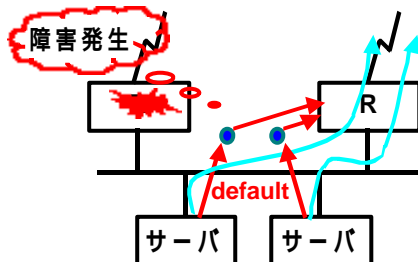


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

117

## MHSRP - 2



- MHSRP (障害発生時)

- なお、MHSRP にはグループ ID 衝突問題があるため、オープンなネットワークでの利用には注意が必要



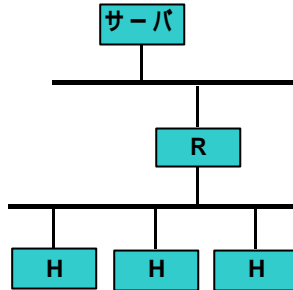
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

118

## ネットワークの拡張を考慮した設計-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



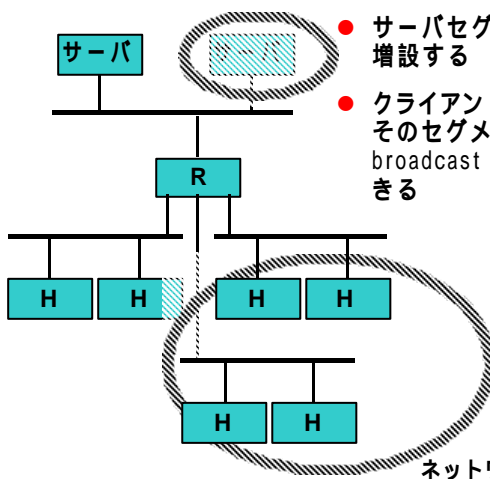
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

119

## ネットワークの拡張を考慮した設計-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

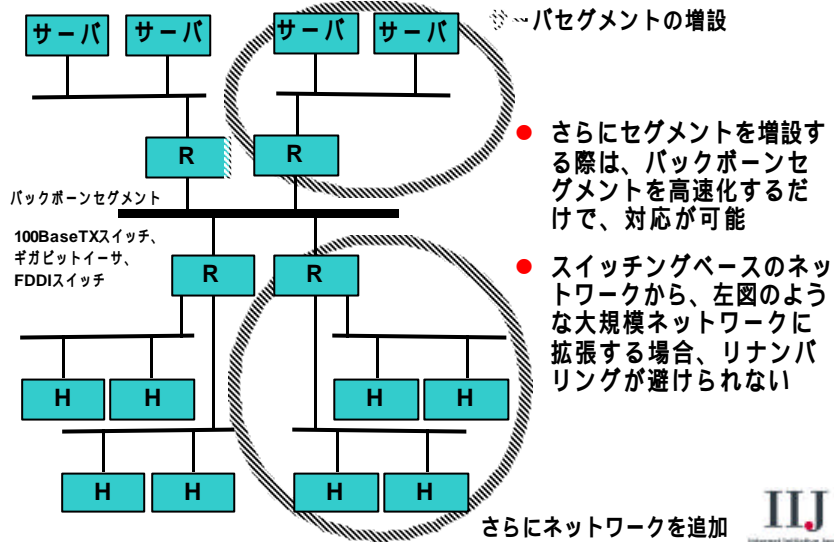


2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

120

## ネットワークの拡張を考慮した設計-3



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

121

## ネットワーク設計

- 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

IIIJ  
Internet Initiative Japan

2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

122

## 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
- アドレスを先頭から詰めて使用するべきか、それとも先頭と後ろから使用していくべきか
- 各部署に割り当てする時は、どのように割り当てていけばいいのか
- 各部署内で各ホストに割り当てる場合にどのように割り当てていけばいいのか



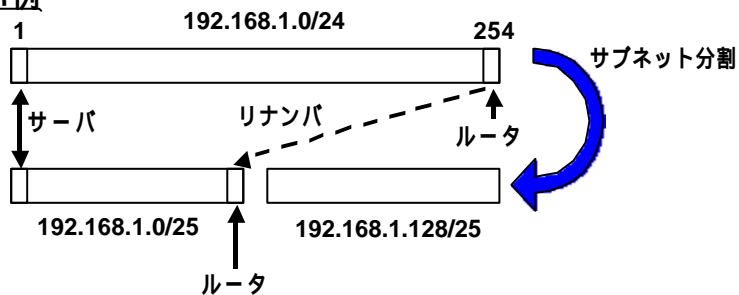
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

123

## 組織全体でのアドレスの割り当て-1

悪い例



- アドレスを先頭と後ろから使用した場合、サブネット分割を行う必要が生じた場合に、リナンバー作業を行う必要が出てしまう。



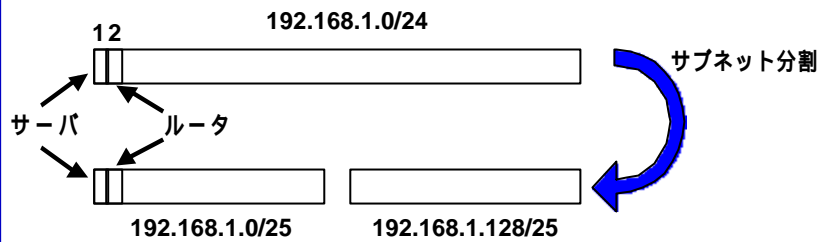
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

124

## 組織全体でのアドレスの割り当て-2

### 良い例



- アドレスを前詰めで使用した場合、サブネット分割を行っても、リナンバー等の無駄な作業を行う必要がない。



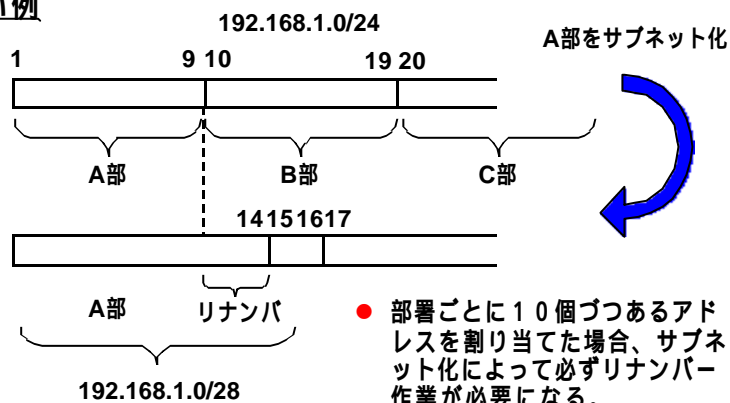
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

125

## 部署ごとのアドレスの割り当て-1

### 悪い例



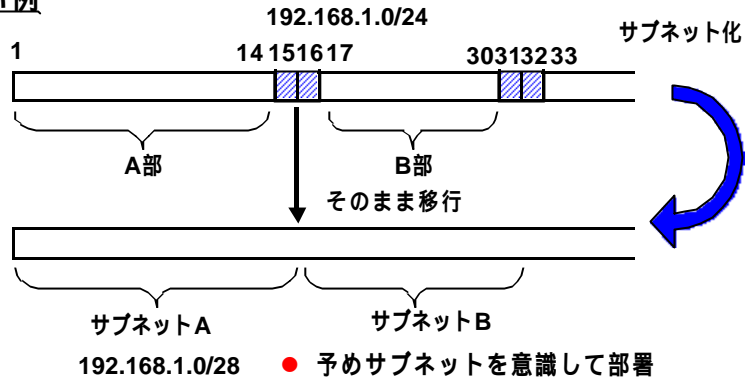
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

126

## 部署ごとのアドレスの割り当て-2

良い例



- 予めサブネットを意識して部署毎にアドレスを割り当てることにより、リナンバー作業を防ぐことができる



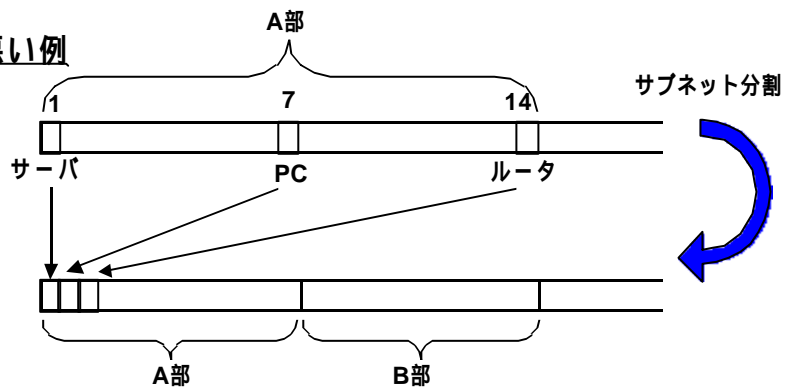
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

127

## 部署内でのアドレスの割り当て-1

悪い例



- 部署内でルータやサーバ等利用目的別に割り当てるアドレス空間を決めてしまうと、さらなるサブネット化に対応できず、リナンバー作業が発生してしまう。



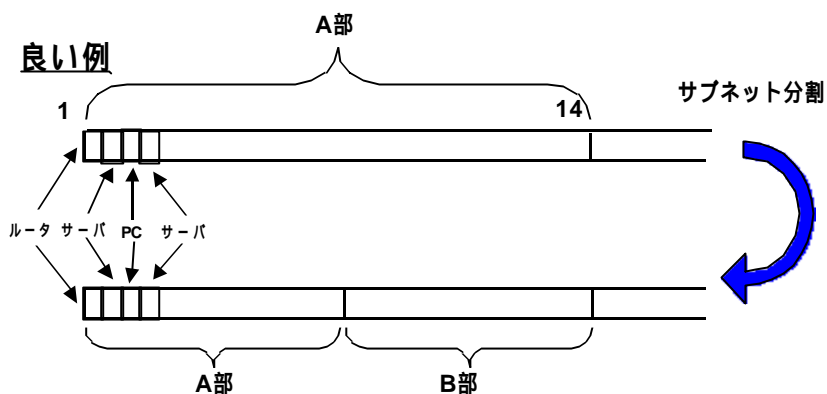
2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

128



## 部署内でのアドレスの割り当て-2



- アドレスを前詰めで使用すればさらなるサブネット化にもスムーズに対応できる



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

129

## アドレスの割り当てポリシーとは

- アドレスを先頭から詰めて使用するべきか、それとも先頭と後ろから使用していくべきか
  - 先頭から詰めて使用する
- 各部署に割り当てる時は、どのように割り当てていけばいいのか
  - サブネット化を考慮して、例えばA部に1~14、B部に17~30のように割り当てる
- 各部署内で各ホストに割り当てる場合にどのように割り当てていけばいいのか
  - 先頭から前詰めで使用する



2000/12/21

Copyright © 2000 Internet Initiative Japan Inc.

130

## 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
  - それぞれを有効に配置する
- インターネット接続にはルーティングは必須
- ダイナミックルーティングは基本を理解すれば応用できる



## まとめ - 2

- VLSM導入にはRIP2、OSPFを利用
- 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
- 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

