ネットワーク構築 A to Z [I]

～知っているようで知らないネットワークの基礎～

2003年12月3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 Japan

目的

- データリンク層とネットワーク層の役割は
- 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
- ルーティングとは

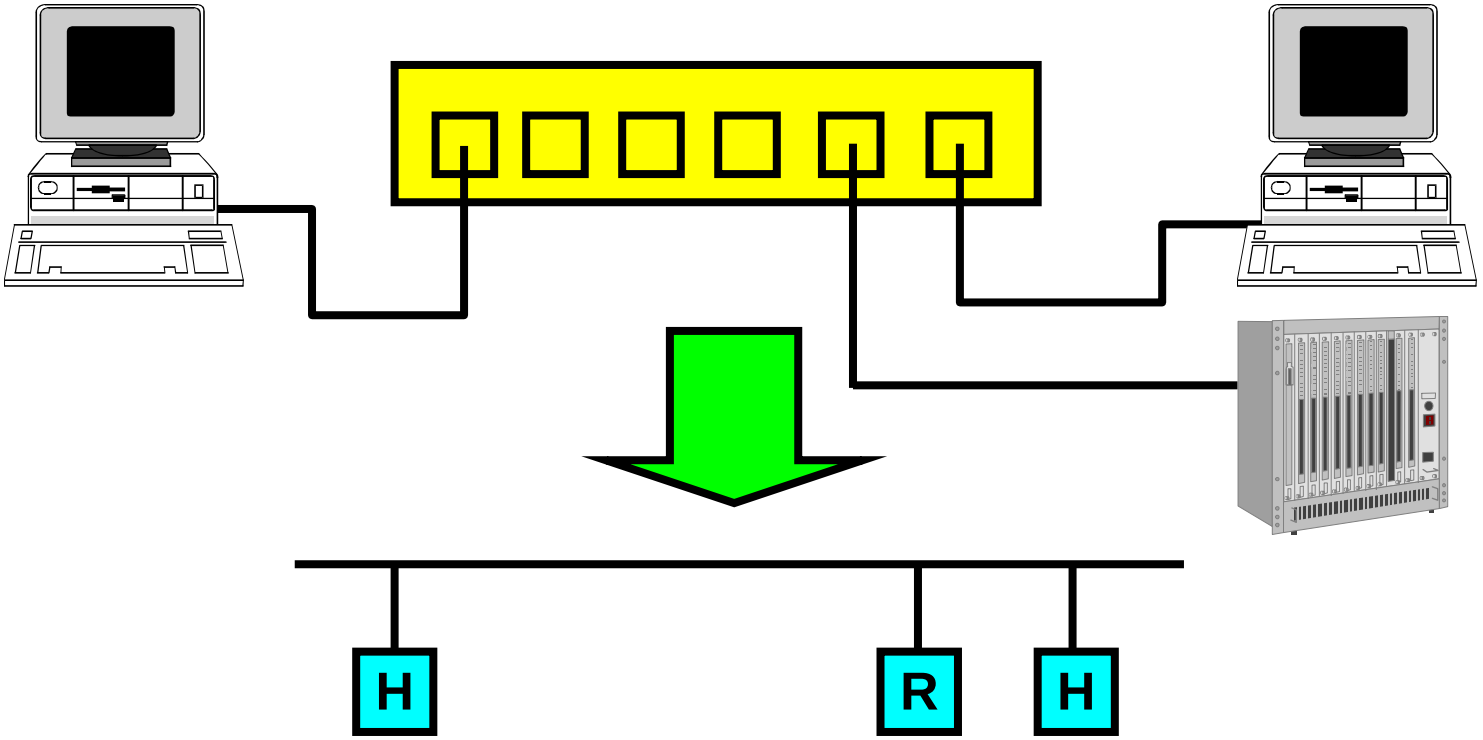
IIJ
Internet Initiative Japan

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
- ネットワークの冗長化
- ネットワークトラブルシューティング

IIJ
Internet Initiative Japan

ネットワーク表記



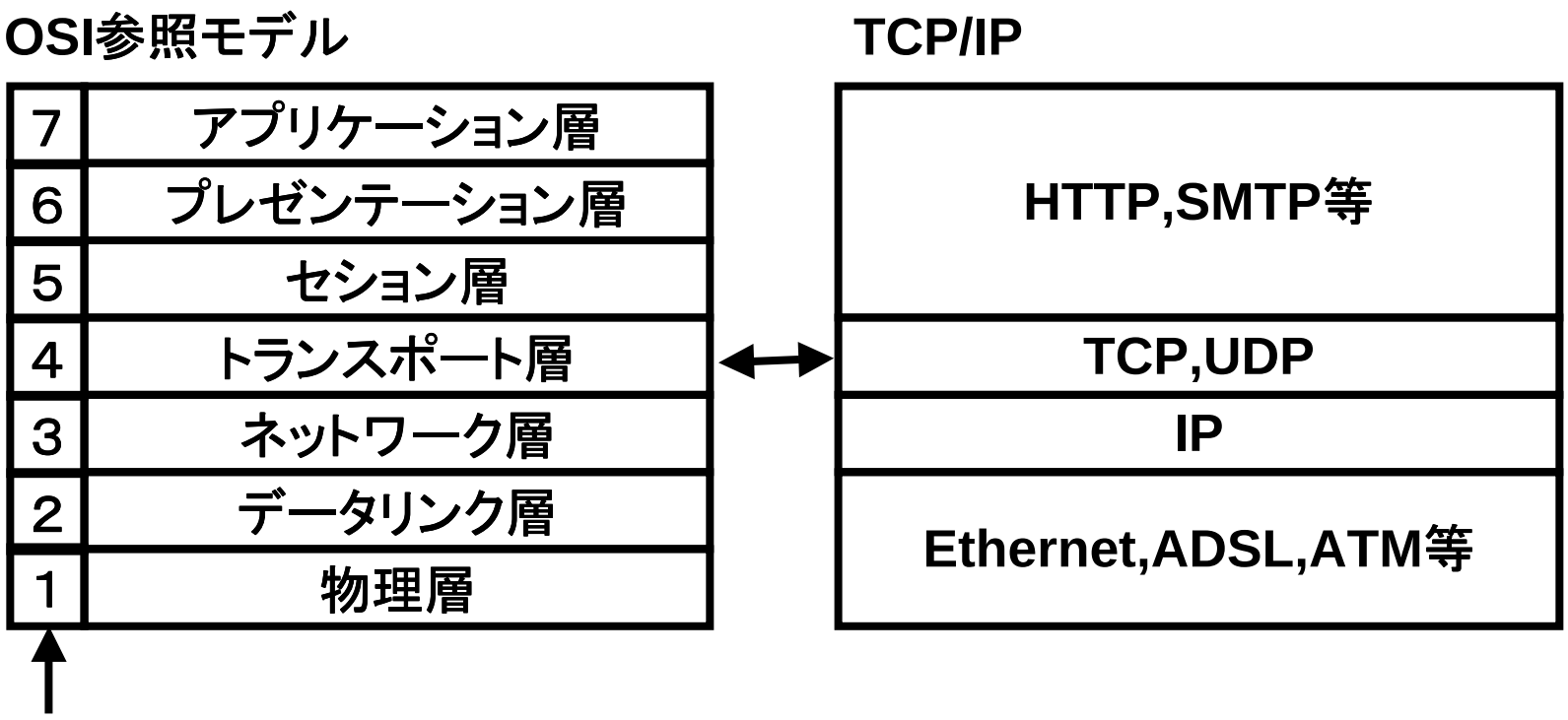
- ハブ、スイッチなどは1本の線またはSWで表わします。
- ホストはH、A、B、C、D等で、ルータはR等で表記します
- レイヤ3スイッチなどは説明中ではルータと区別していません

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

IIJ
Internet Initiative Japan

OSI参照モデルとTCP/IP



レイヤ2：データリンク層

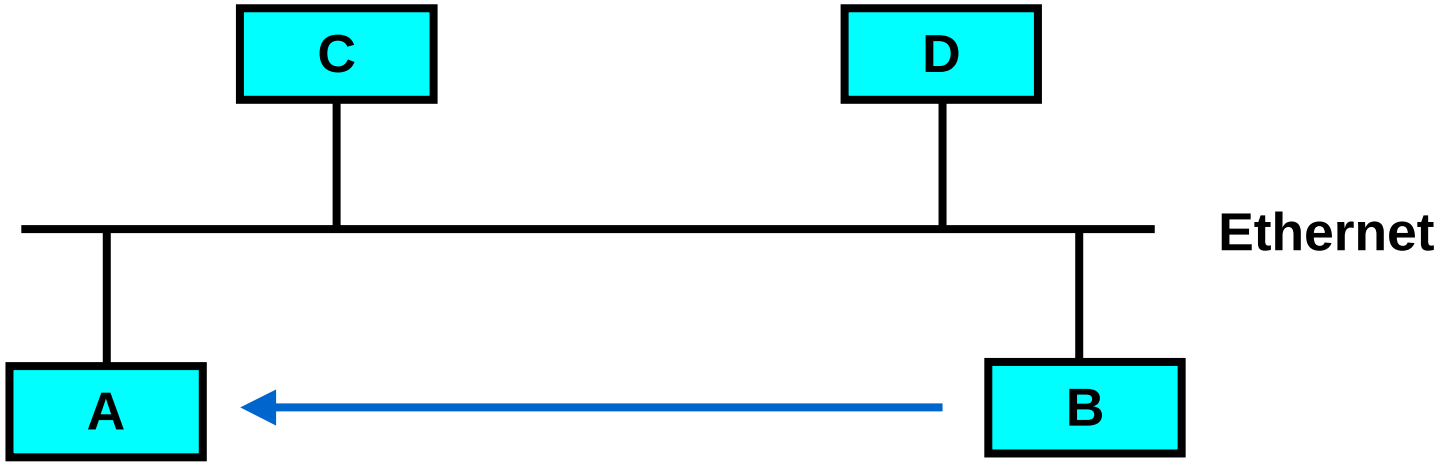
レイヤ3：ネットワーク層

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

IIJ
Internet Initiative Japan

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

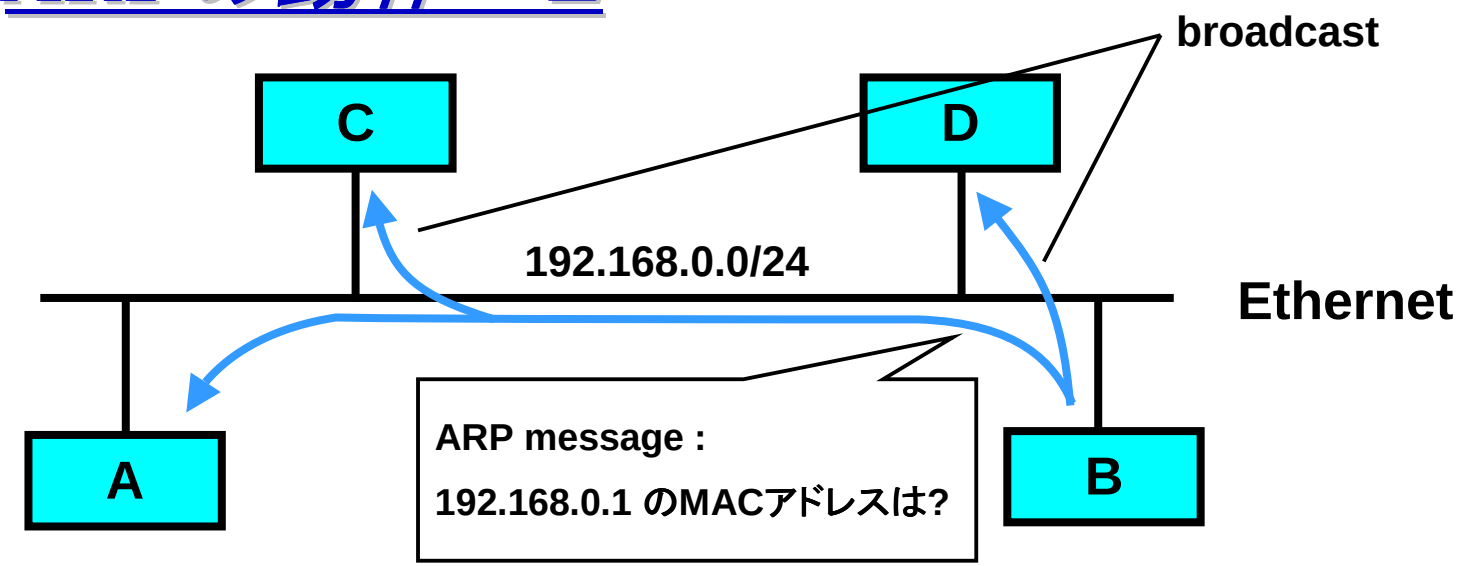
Host	IPアドレス	MACアドレス
A	192.168.0.1	不明
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

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

IIJ
Internet Initiative Japan

ARPの動作—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

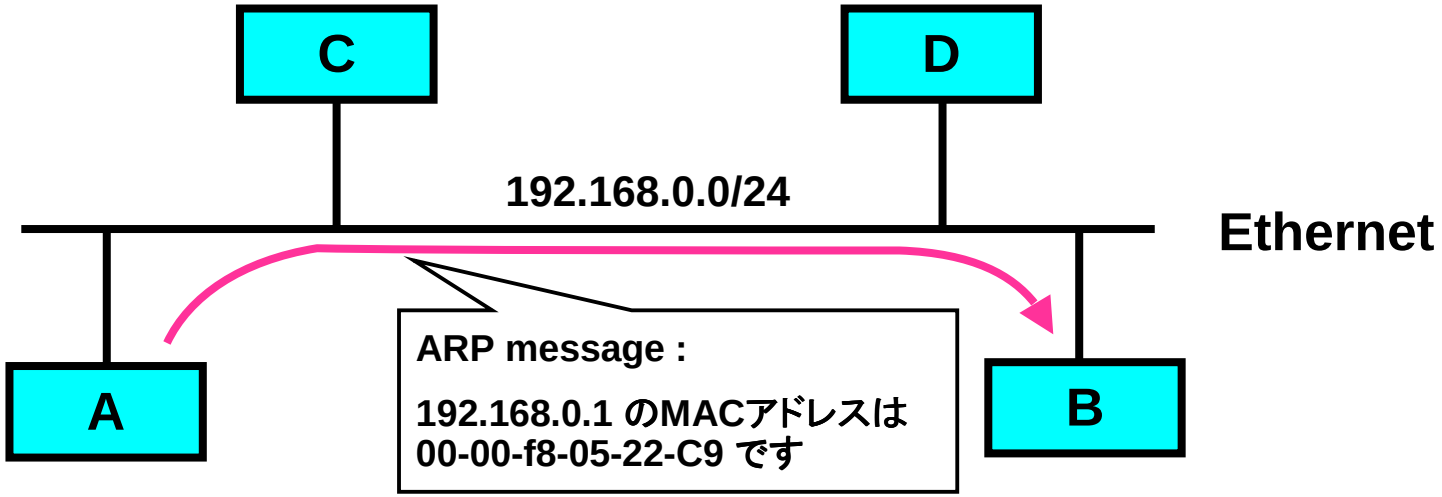
Host	IPアドレス	MACアドレス
A	192.168.0.1	不明
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

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

IIJ
Internet Initiative Japan

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

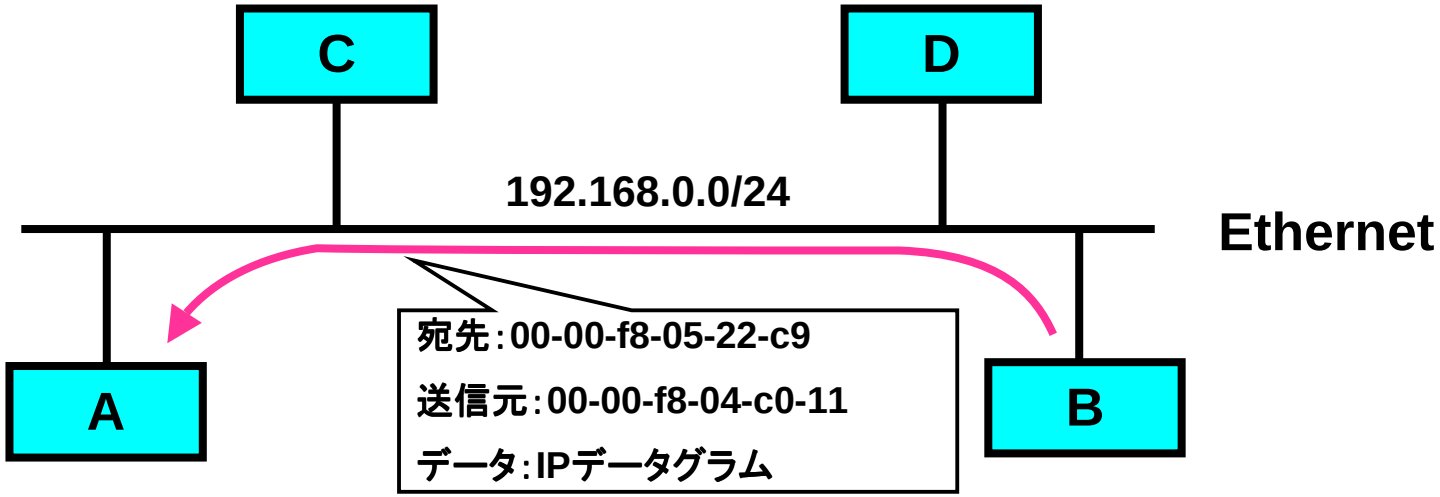
Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

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

IIJ
Internet Initiative Japan

ARPの動作—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

IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

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

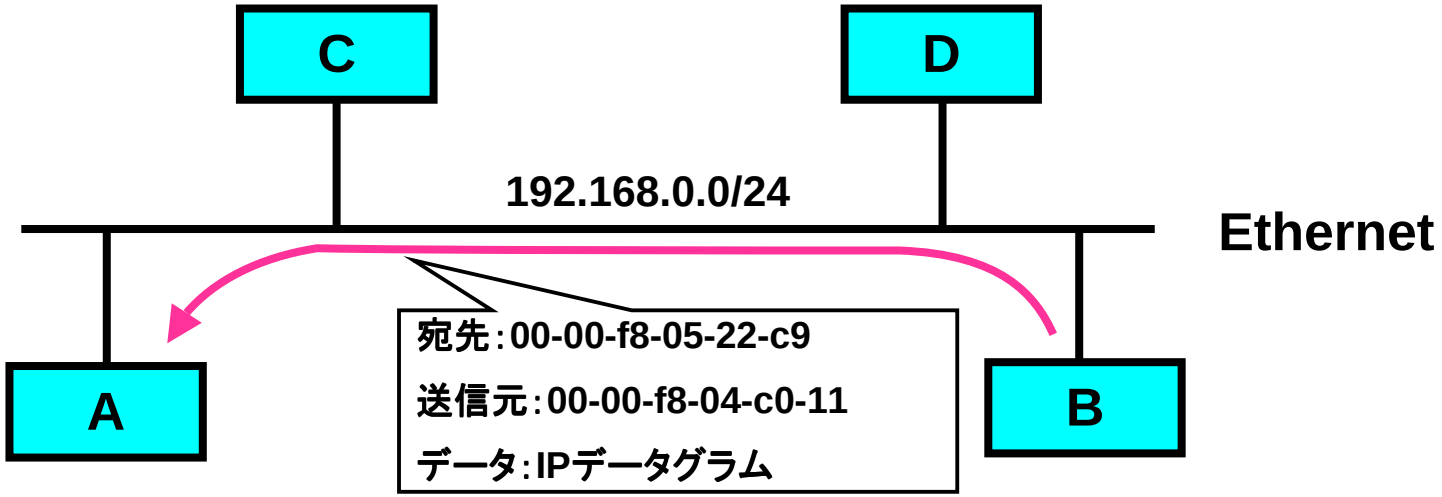
IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

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

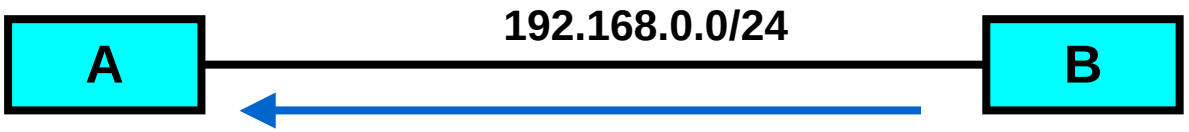
IIJ Internet Initiative Japan

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

IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

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

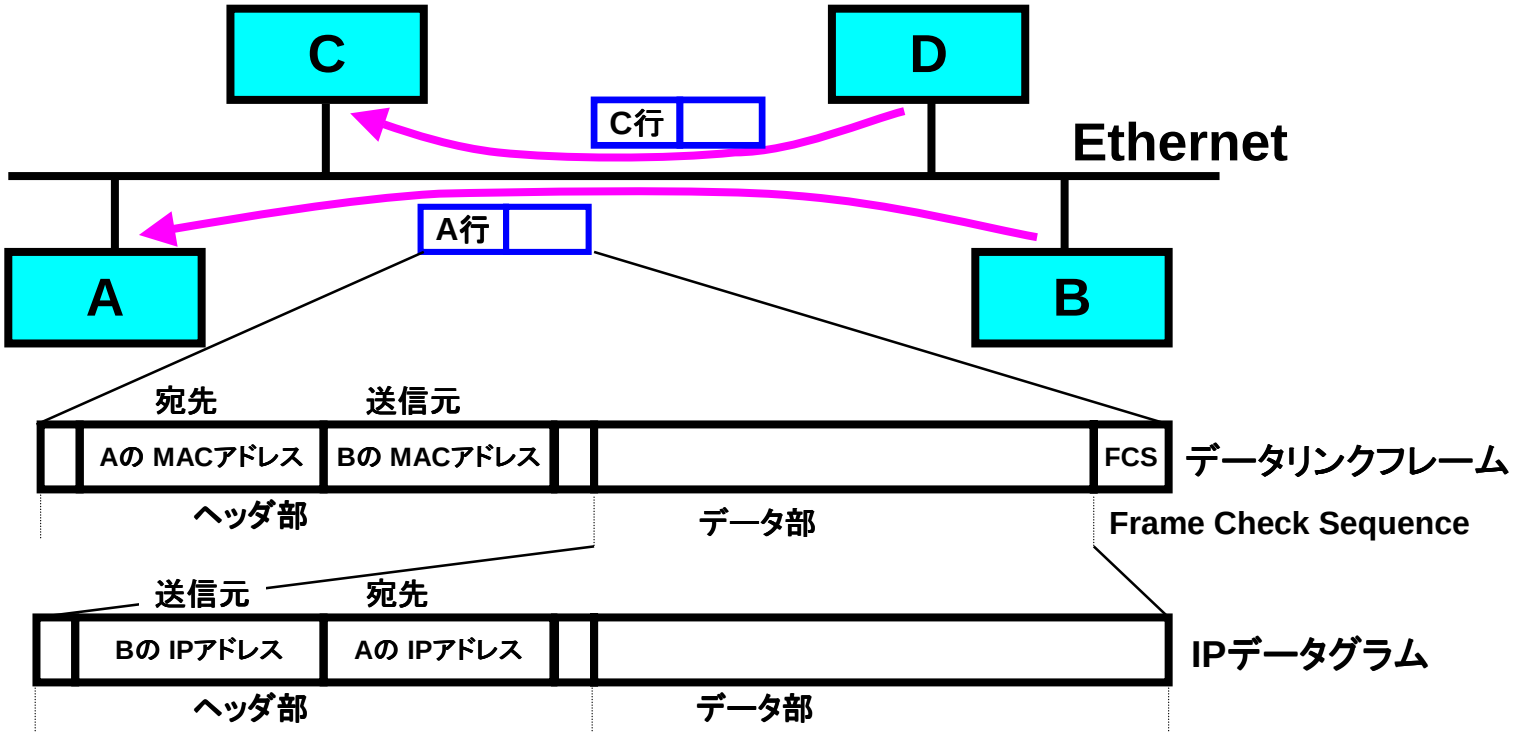
IIJ
Internet Initiative Japan

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*

IIJ
Internet Initiative Japan

Ethernetを流れるIPデータグラム



Connectedなネットワーク

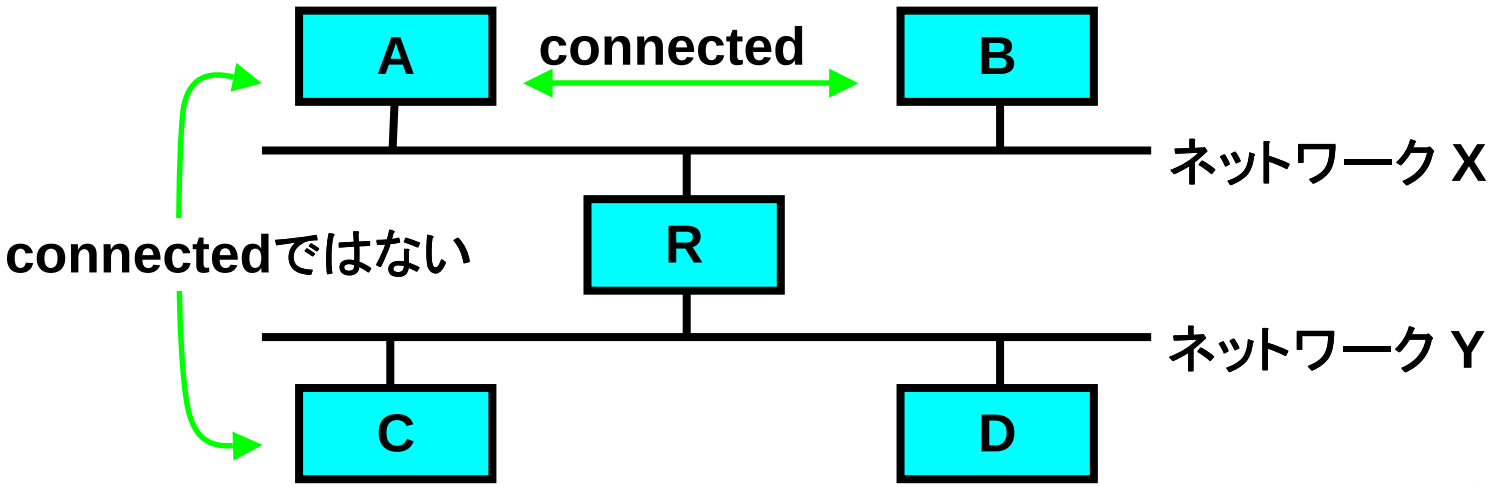
- A、Bは直接同じネットワークに接続している
  - MACアドレス、IPアドレスの対応表を ARP(address resolution protocol)などにより持っている

↓

- これを「connected」な状態という

↓

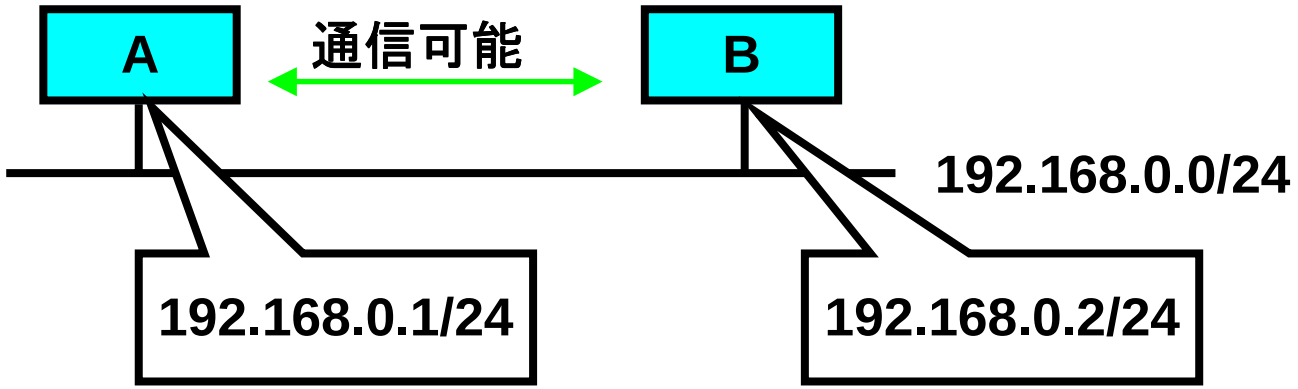
- ルーティング設定が不要で、ハブなどで接続すると通信できる



IIJ
Internet Initiative Japan

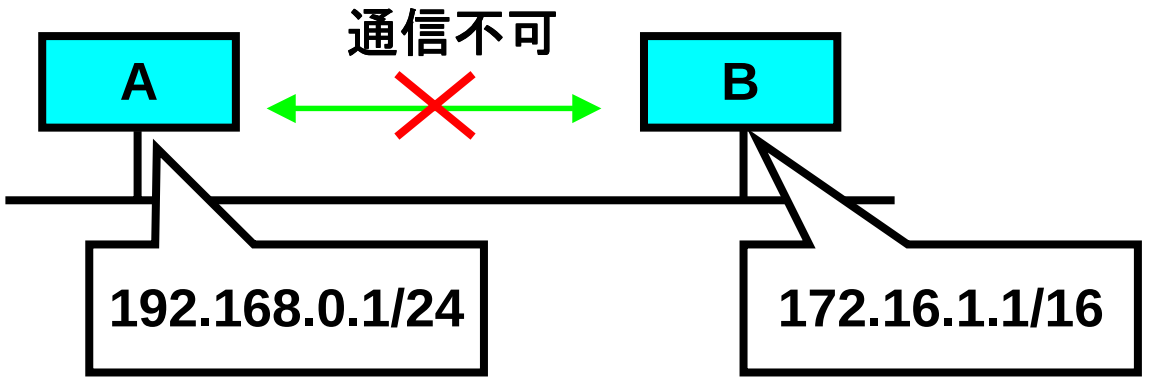
ネットワーク層から見たConnectedなネットワーク－１

- 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

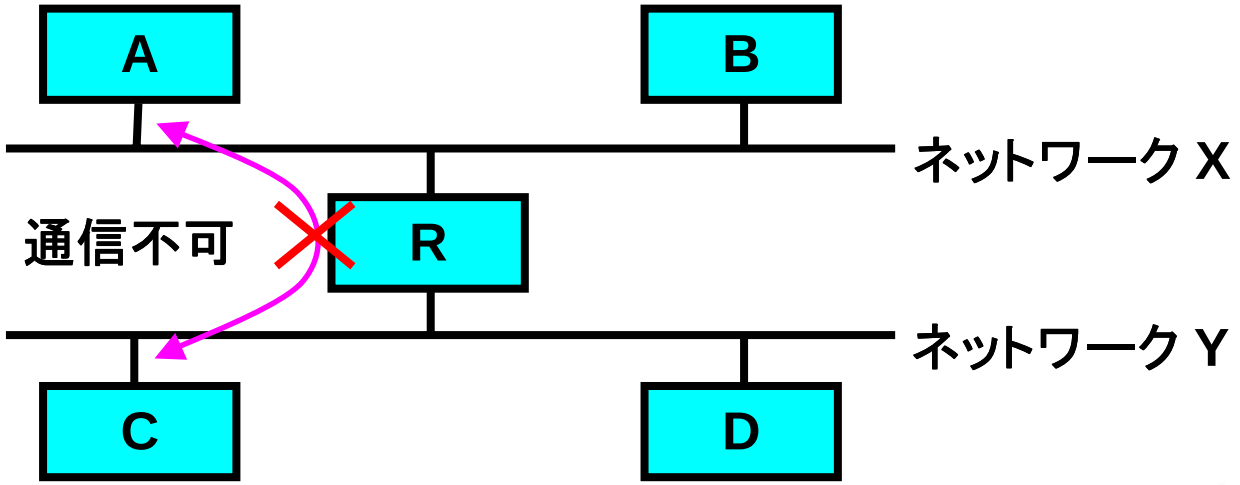
- A、Cはそれぞれ異なるネットワークに接続しているため connectedではない
- ルーティング*設定なし*ではA、C間の通信はできない

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

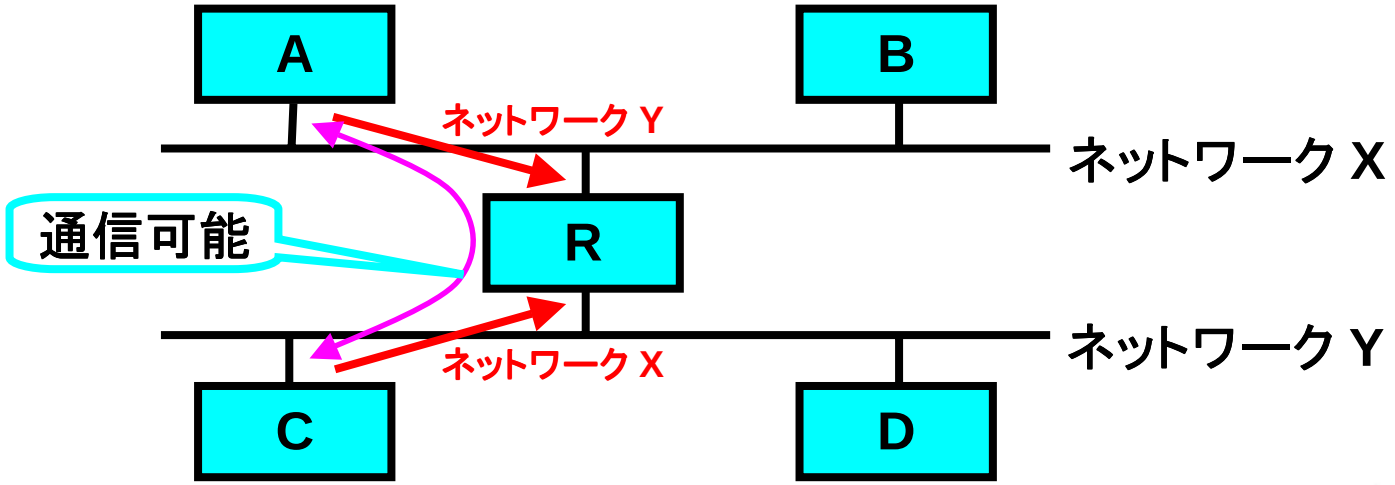
- ルーティング設定を行なう
  - A: ネットワークYを Rにルーティング
  - C: ネットワークXを Rにルーティング
- これにより、A⇔C間の相互通信が可能となる
  - Rは A,C共に connectedなため、アドレスを設定するだけで通信が可能

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

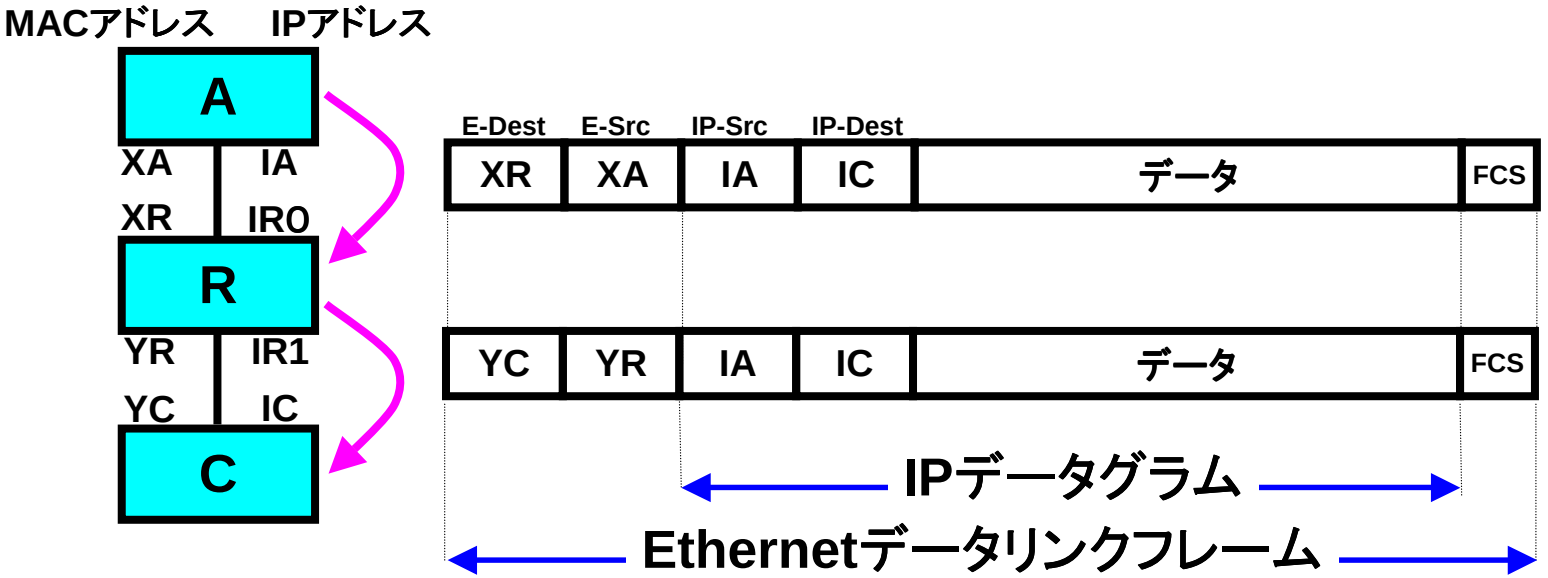
Cのルーティングテーブル

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



IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

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」

IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

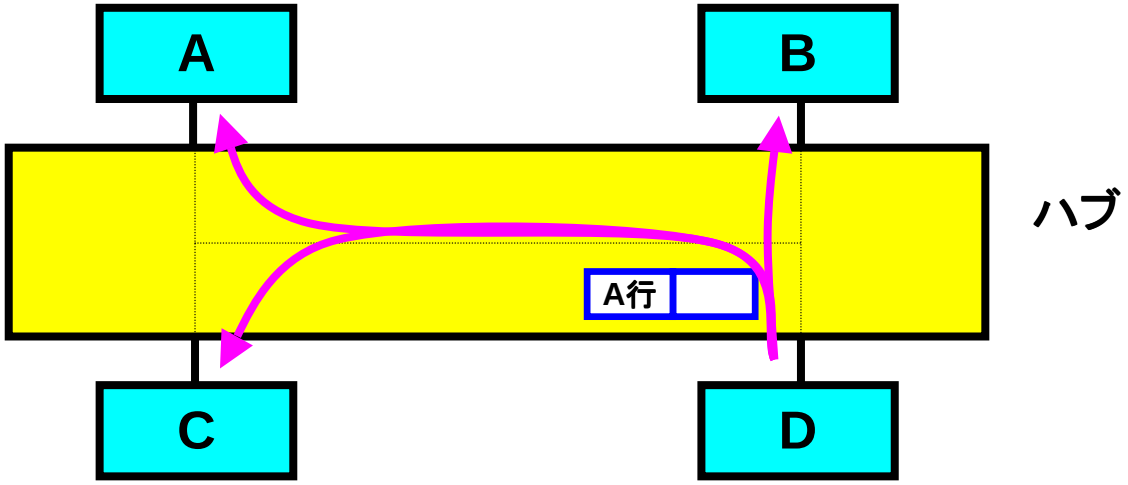
スイッチとルータの機能の違い

- 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

IIJ
Internet Initiative Japan

ハブとスイッチの違い-1

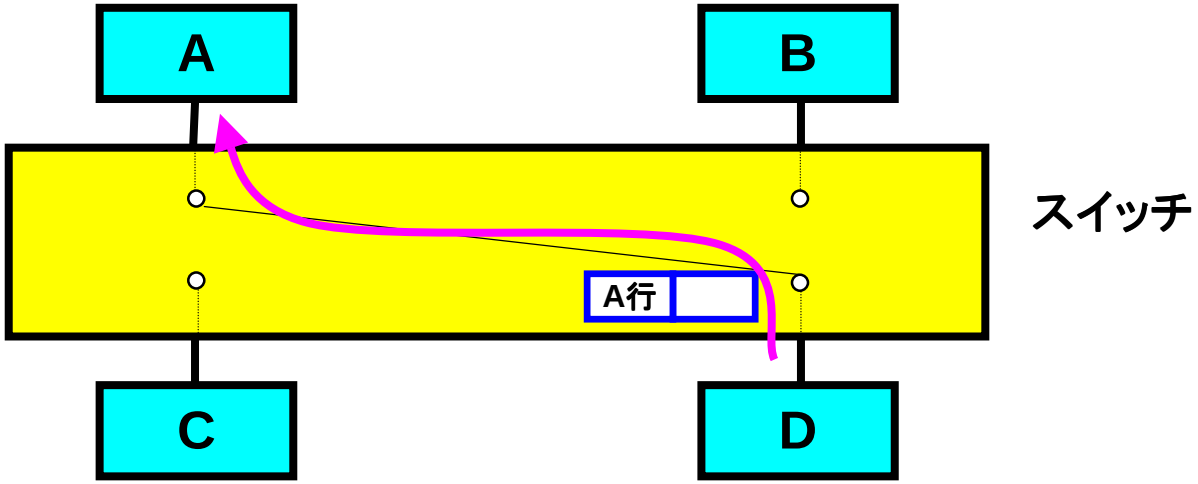
ハブで構成した場合



- 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

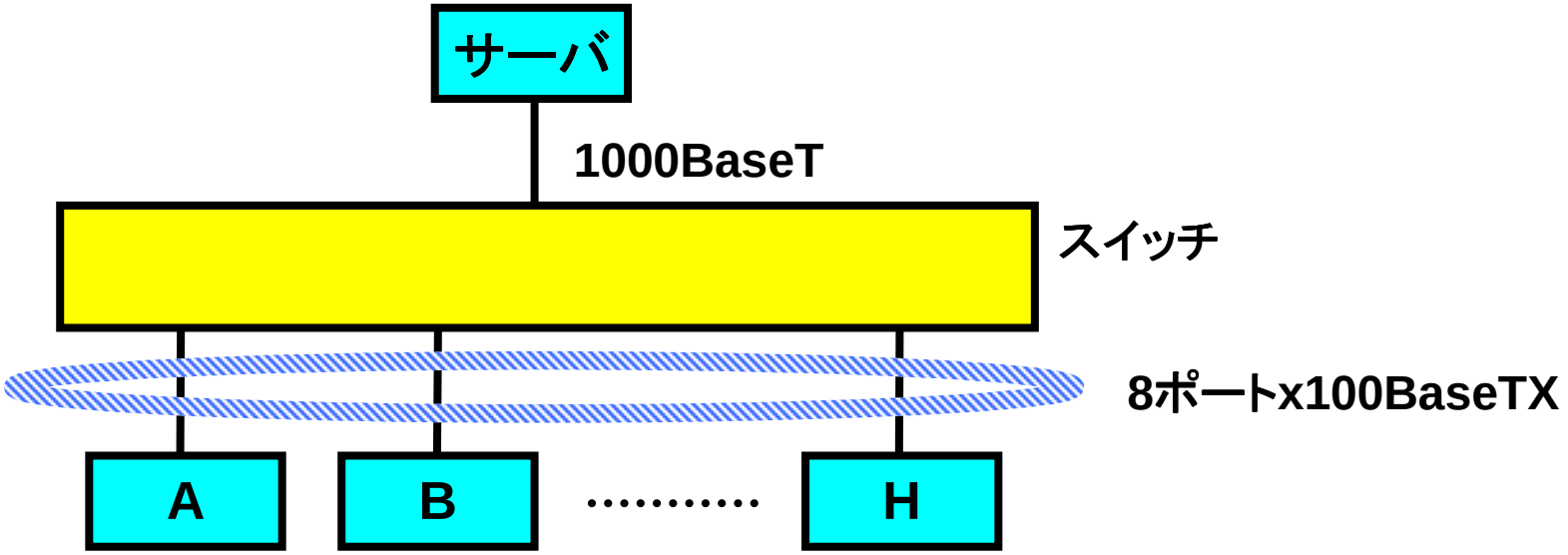
ハブとスイッチの違い-2

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

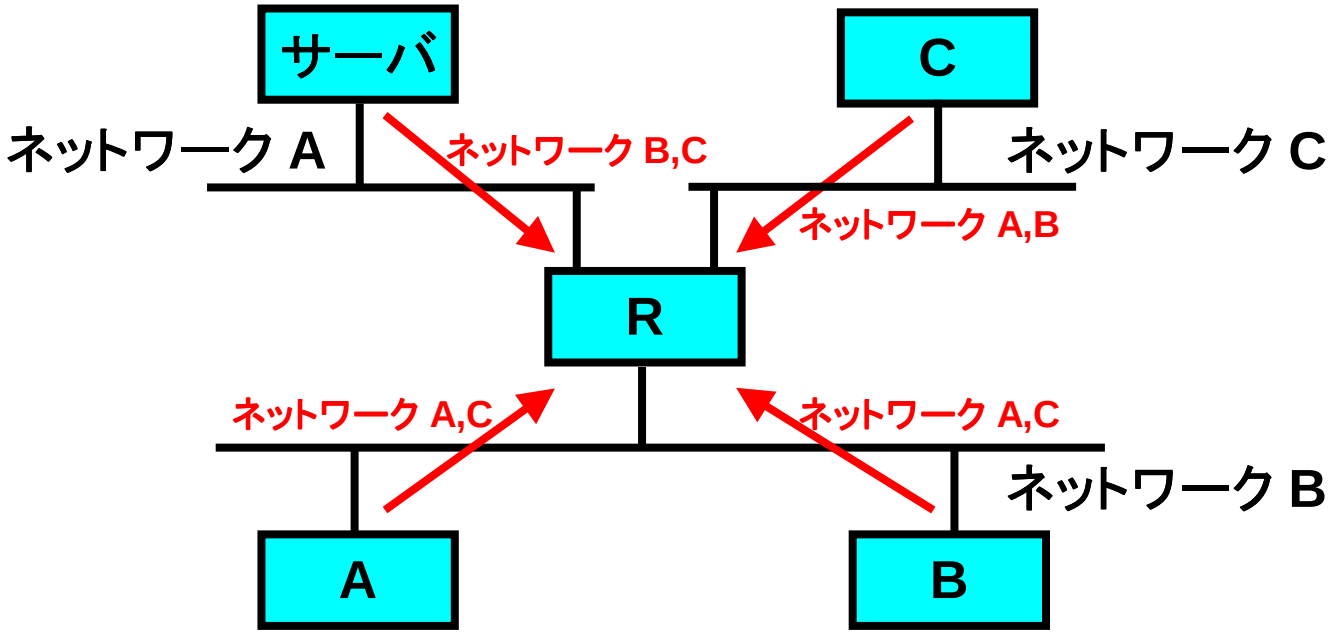
スイッチを有効に使うには



- 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

IIJ
Internet Initiative Japan

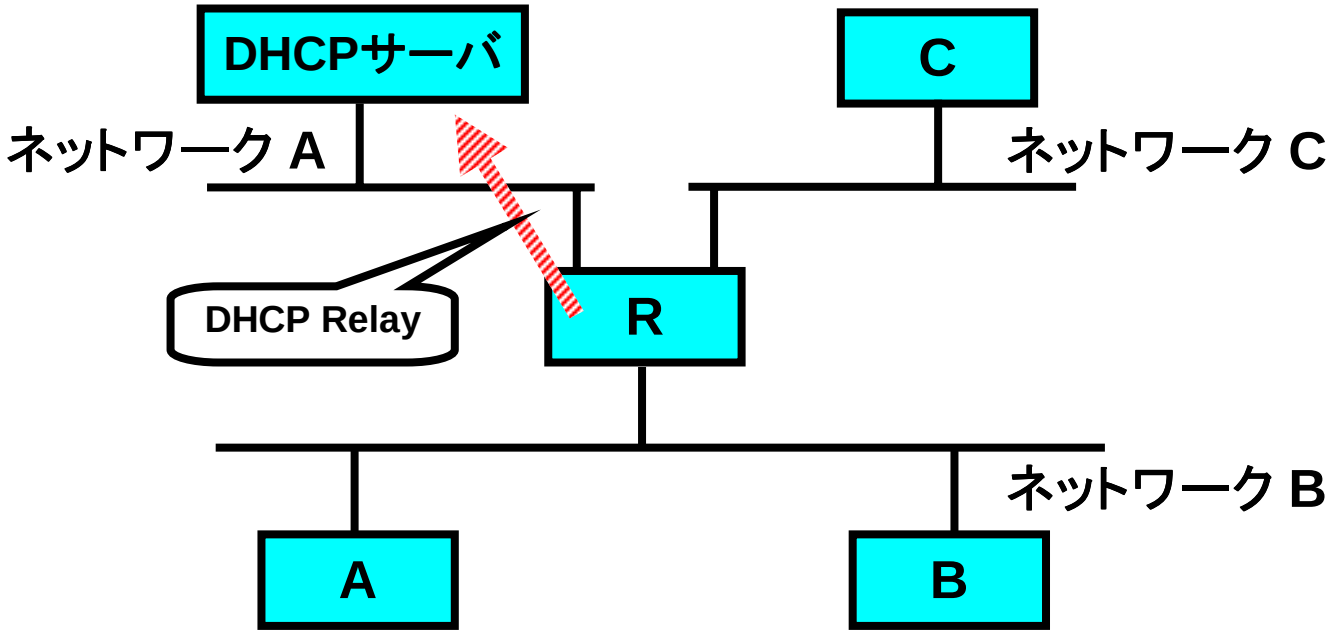
ルータを利用するための設定－1



- 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

IIJ
Internet Initiative Japan

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

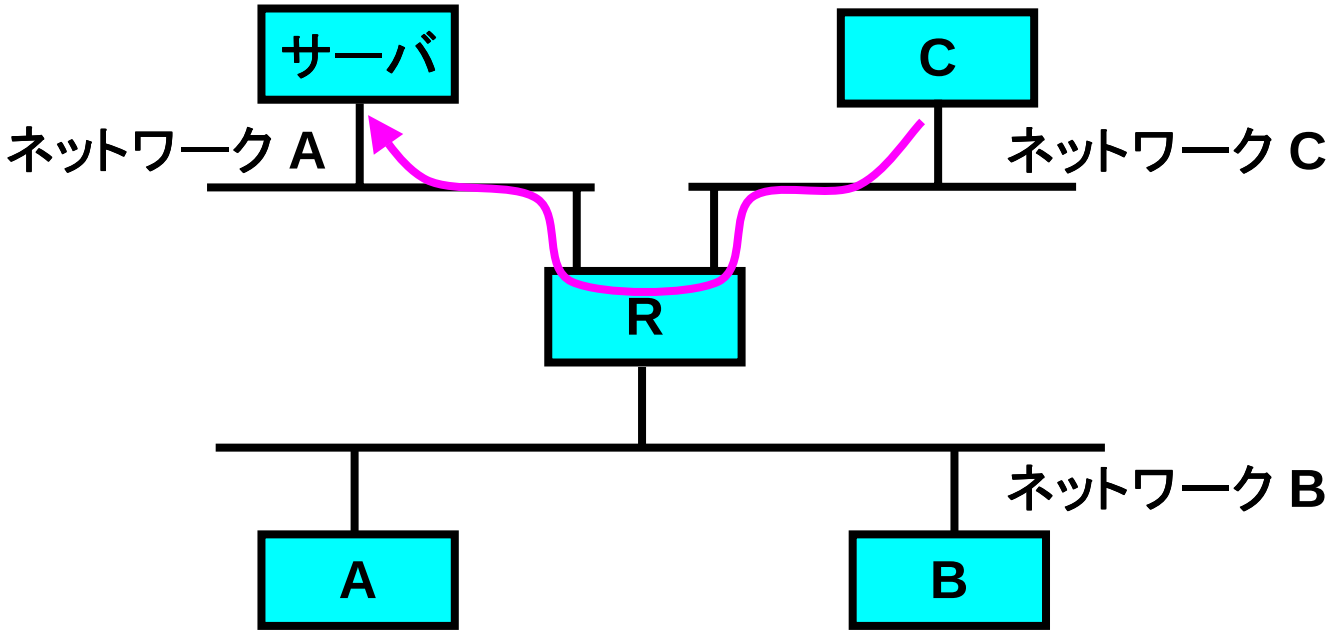
IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

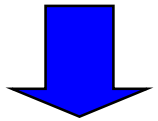
スイッチとルータの違い



- 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

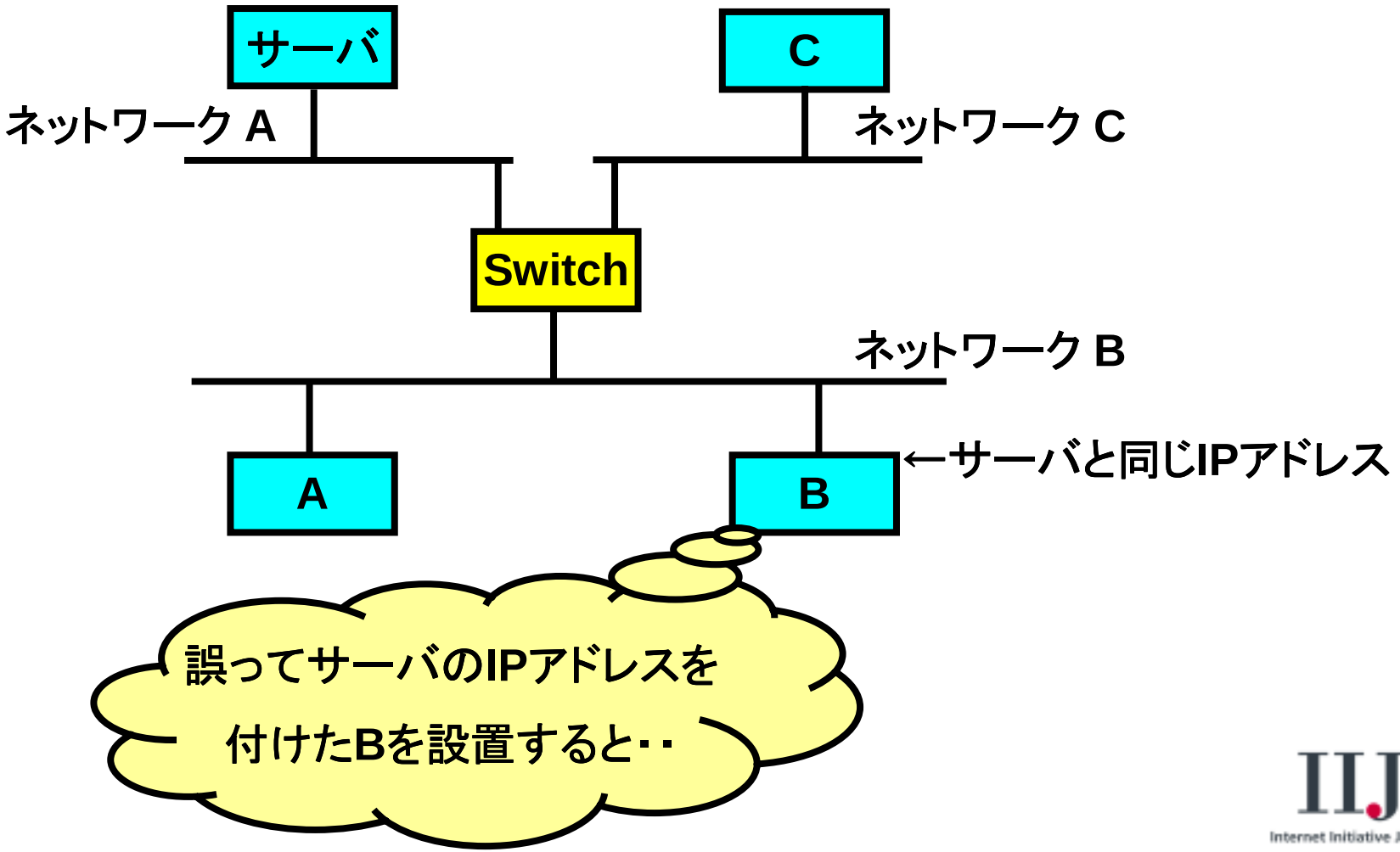
- ハブとスイッチの機能の違い
  - スイッチは異なるポートの通信を他のポートに伝播しない
- スイッチとルータの違い
  - ルータは異なるネットワークの通信を他のネットワークに伝播しない
  - スイッチとは異なり、ルーティングの設定が必要
  - サブネット分割が必要
- スイッチを有効に使うには
  - トラフィックが集中するようなポートにはスイッチを導入する



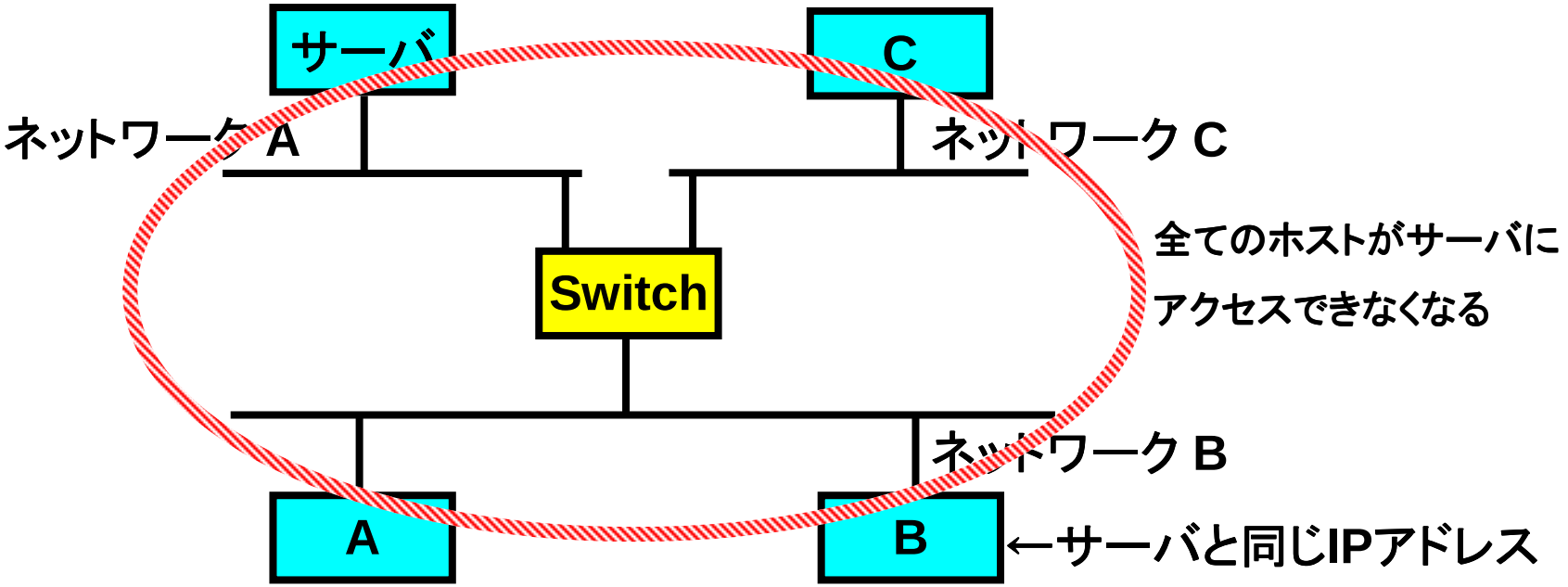
次に問題点について検討する

IIJ
Internet Initiative Japan

スイッチの耐障害性-1

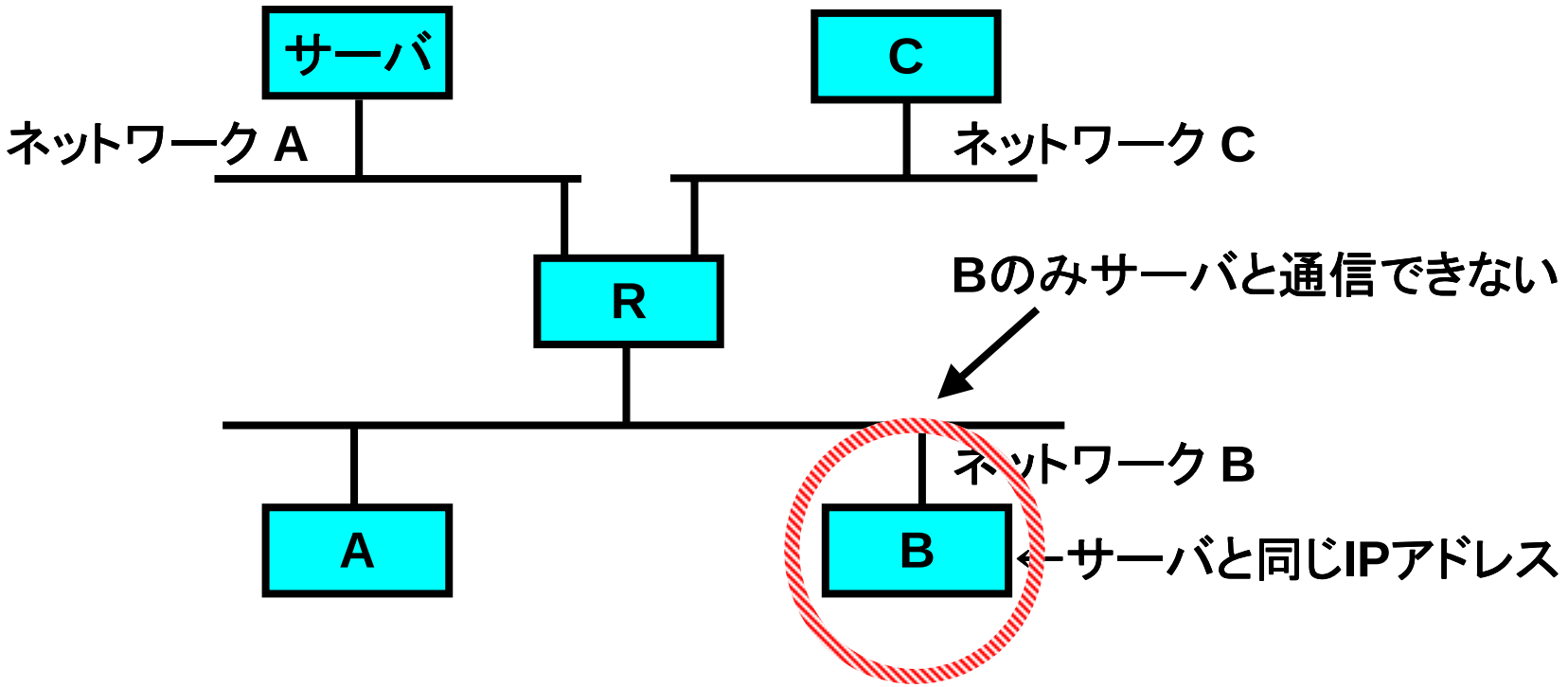


スイッチの耐障害性-2



- 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

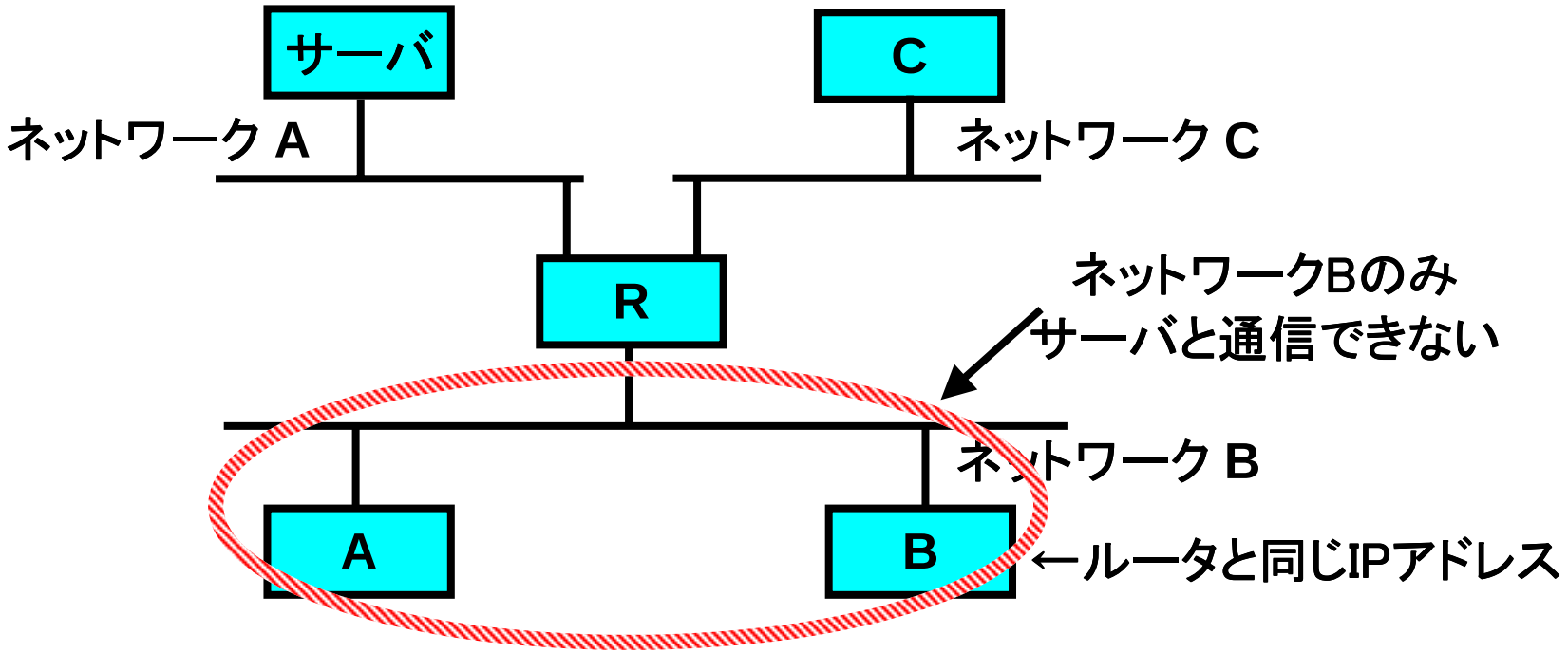
ルータの耐障害性-1



- 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

IIJ
Internet Initiative Japan

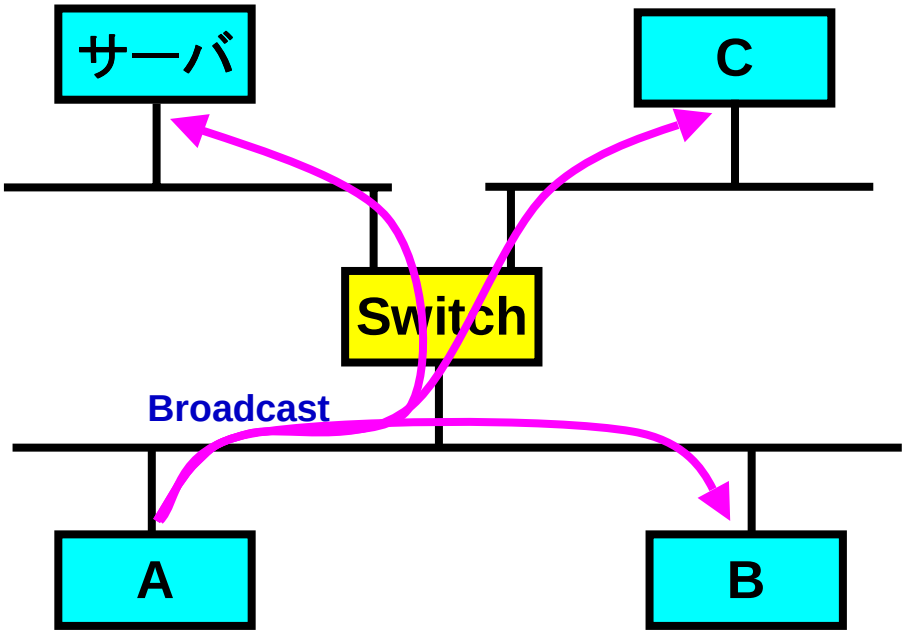
ルータの耐障害性-2



- 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

IIJ
Internet Initiative Japan

Broadcast Flood-1

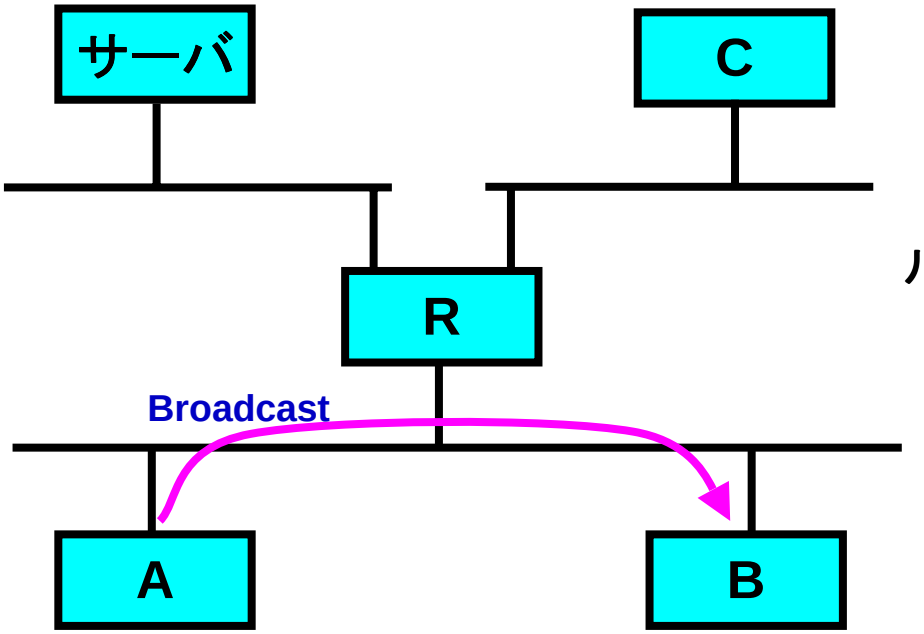


Broadcastパケットはスイッチの
全てのポートに影響を与える

- 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

IIJ
Internet Initiative Japan

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

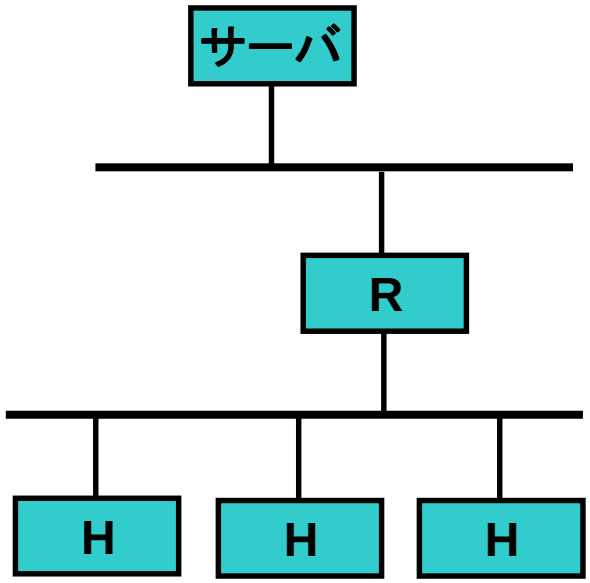
IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

ネットワーク設計−1



左図ネットワーク構成の特徴

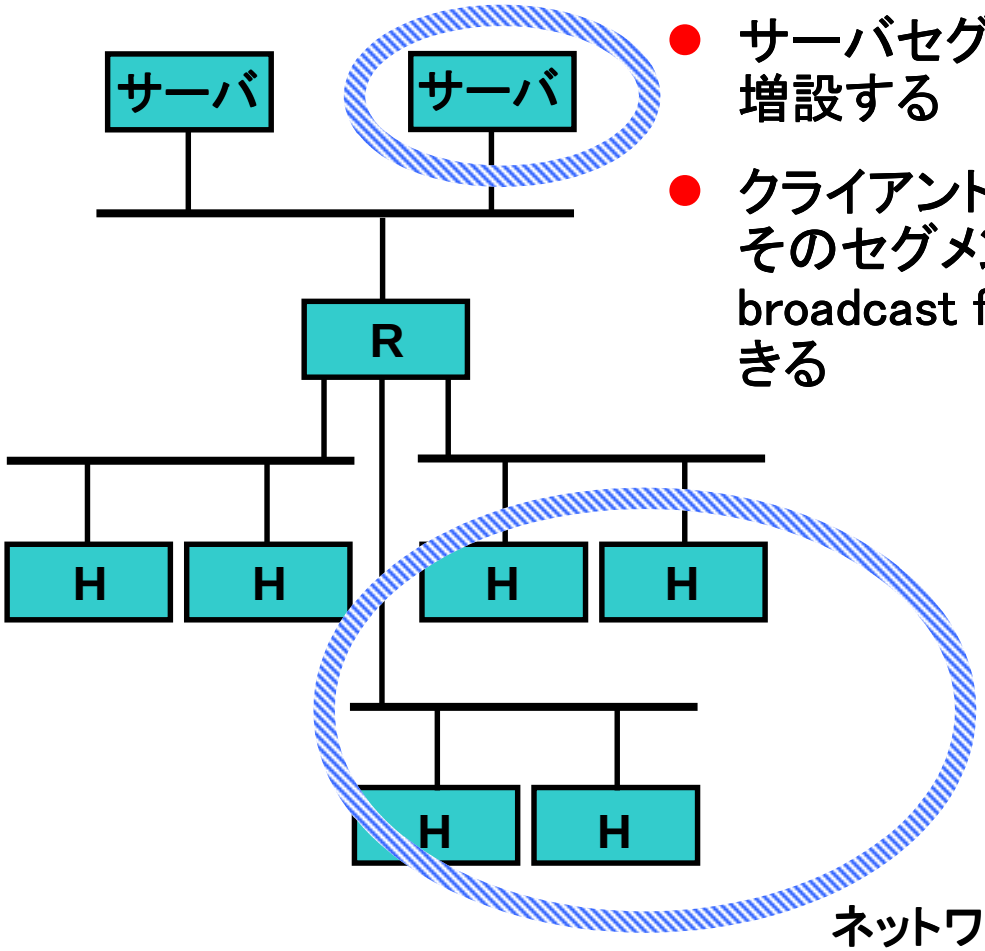
- 小規模であってもサーバのセグメントを分離する

  →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

IIJ
Internet Initiative Japan

ネットワーク設計-2

サーバの増設

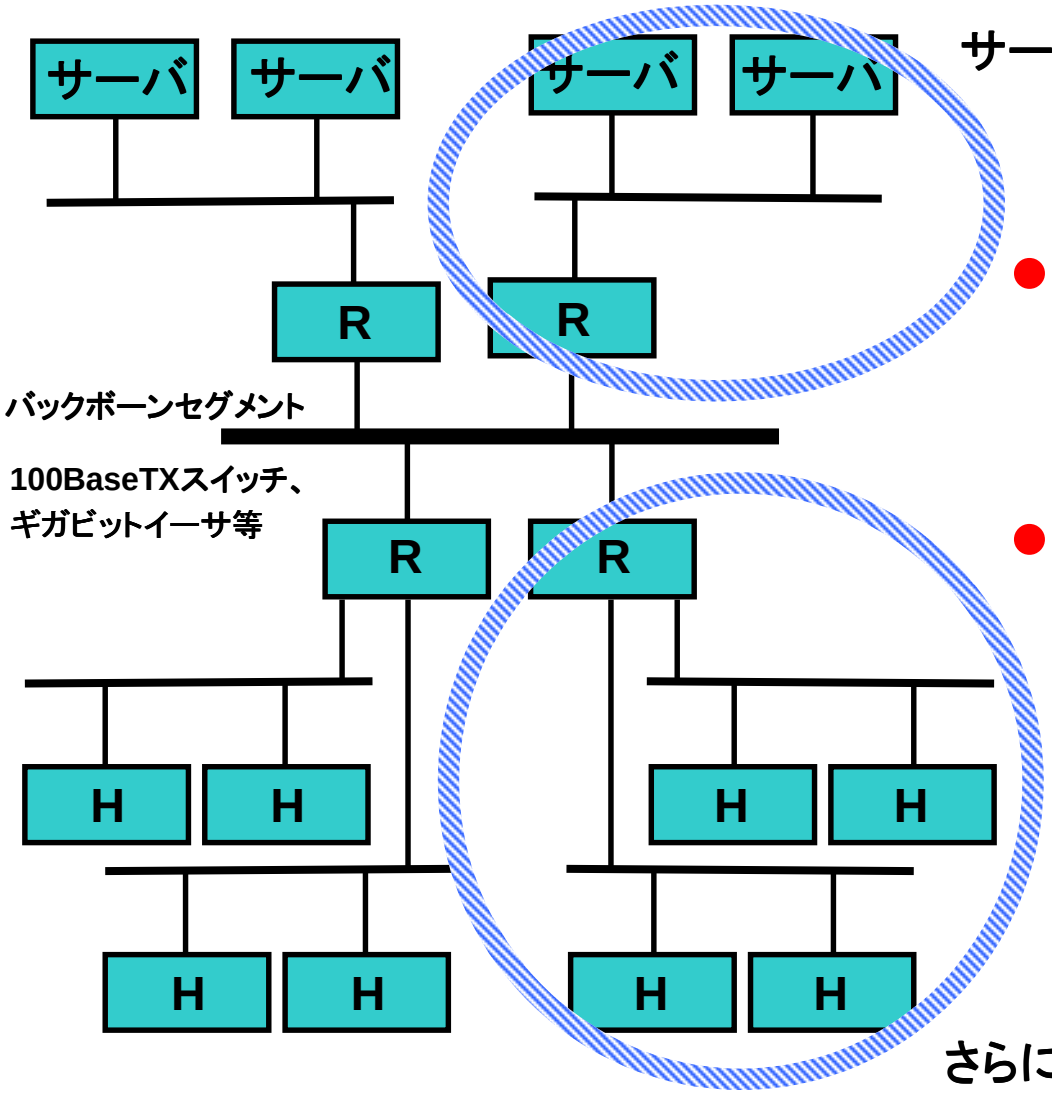


- サーバセグメントの安全性を保ちつつ増設する
- クライアントセグメントのbroadcast をそのセグメント内に留められるため broadcast flood現象の発生を抑制できる

ネットワークの追加

IIJ
Internet Initiative Japan

ネットワーク設計−3



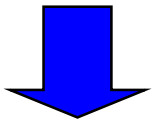
サーバセグメントの増設

- 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

IIJ
Internet Initiative Japan

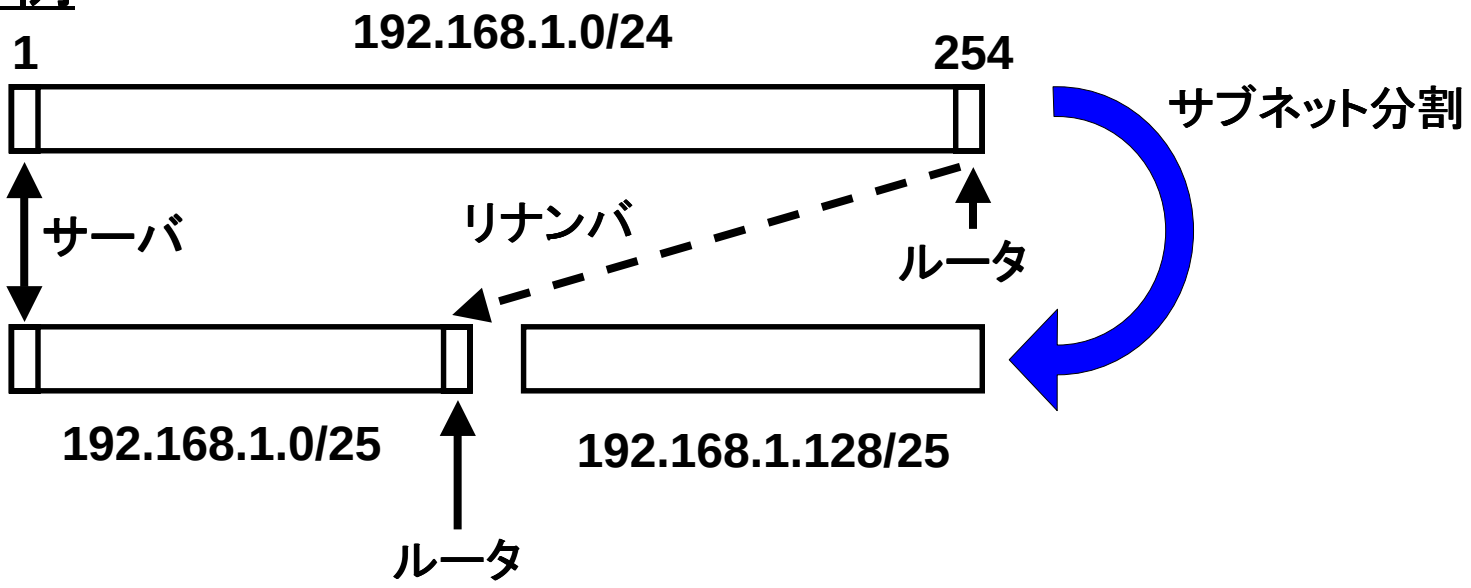
アドレスの割り当てポリシーとは

- 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良いのか

IIJ
Internet Initiative Japan

組織全体でのアドレスの割り当て-1

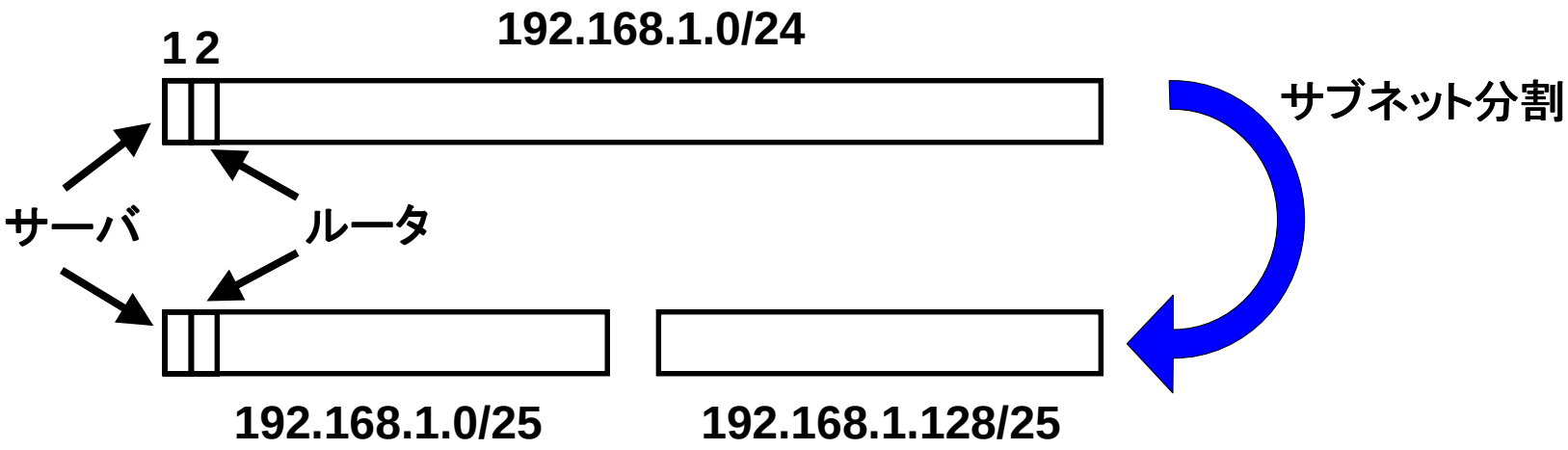
悪い例



- 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。

組織全体でのアドレスの割り当て-2

良い例

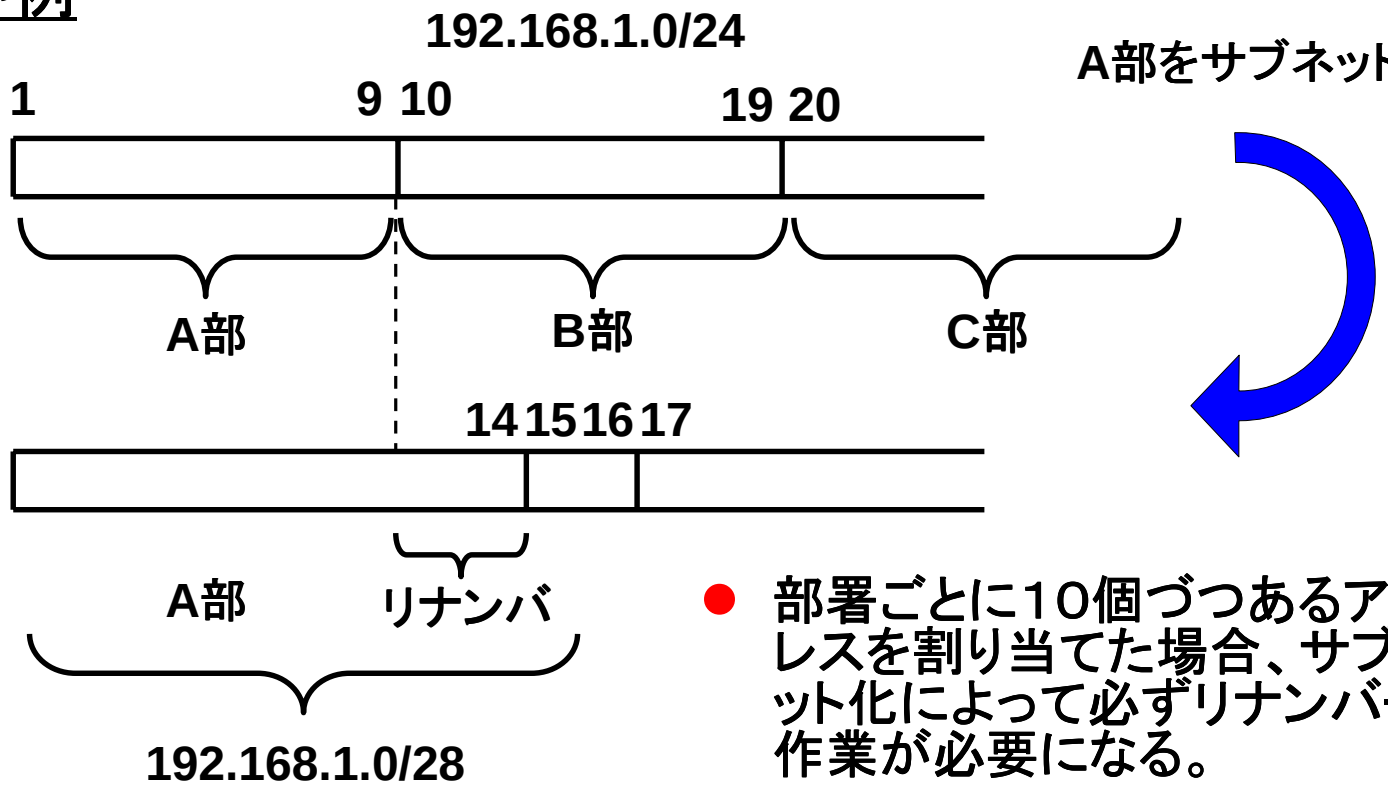


- アドレスを前詰めで使用した場合、サブネット分割を行っても、リナンバー等の無駄な作業を行う必要がない。

IIJ
Internet Initiative Japan

部署ごとのアドレスの割り当て-1

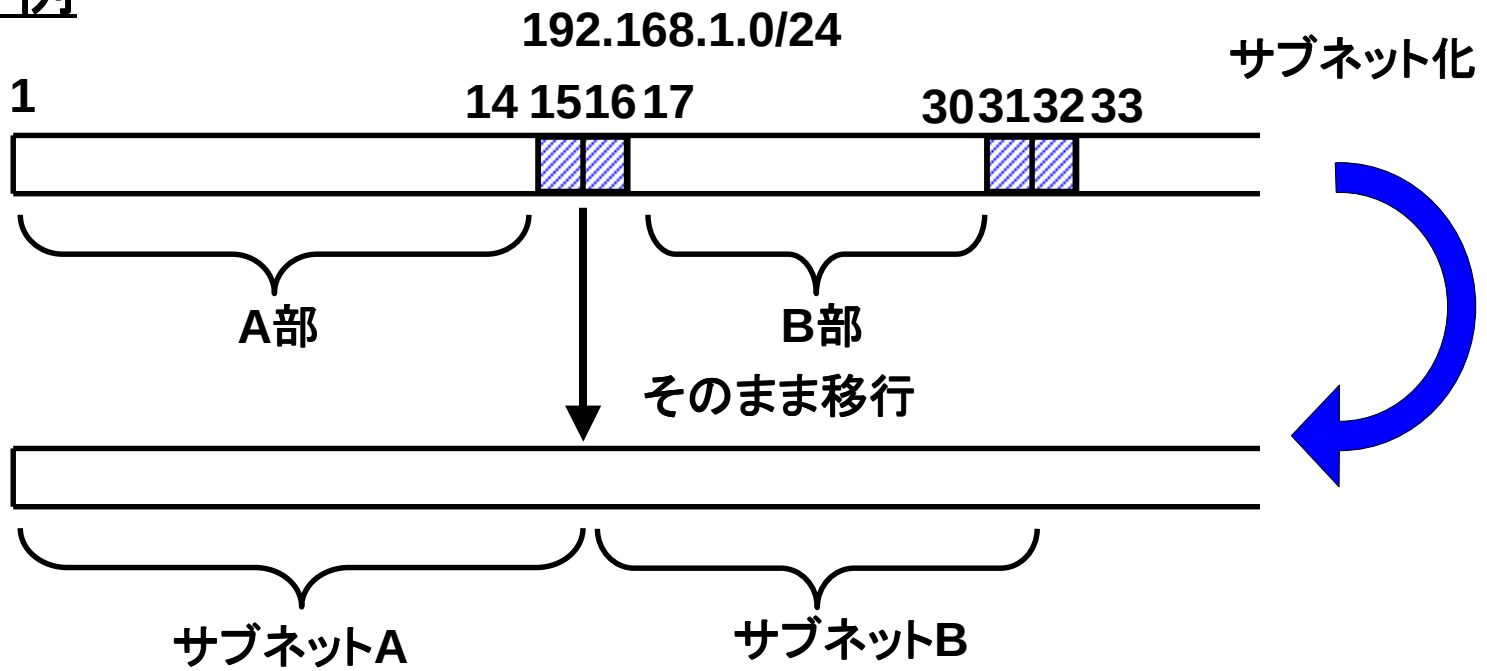
悪い例



- 部署ごとに10個づつあるアドレスを割り当てた場合、サブネット化によって必ずリナンバー作業が必要になる。

部署ごとのアドレスの割り当て-2

良い例

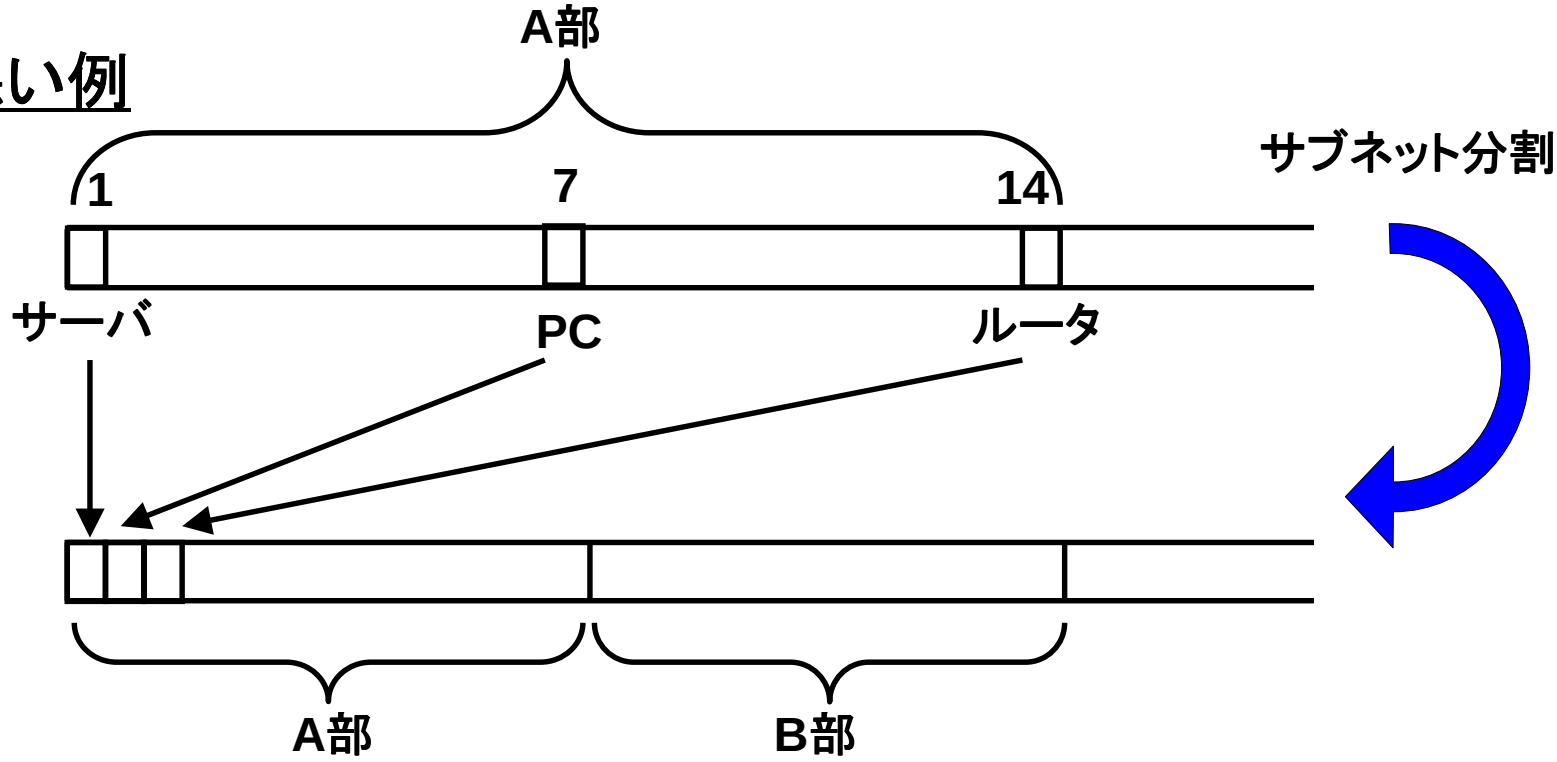


- 予めサブネットを意識して部署毎にアドレスを割り当てることにより、リナンバー作業を防ぐことができる

IIJ
Internet Initiative Japan

部署内でのアドレスの割り当て-1

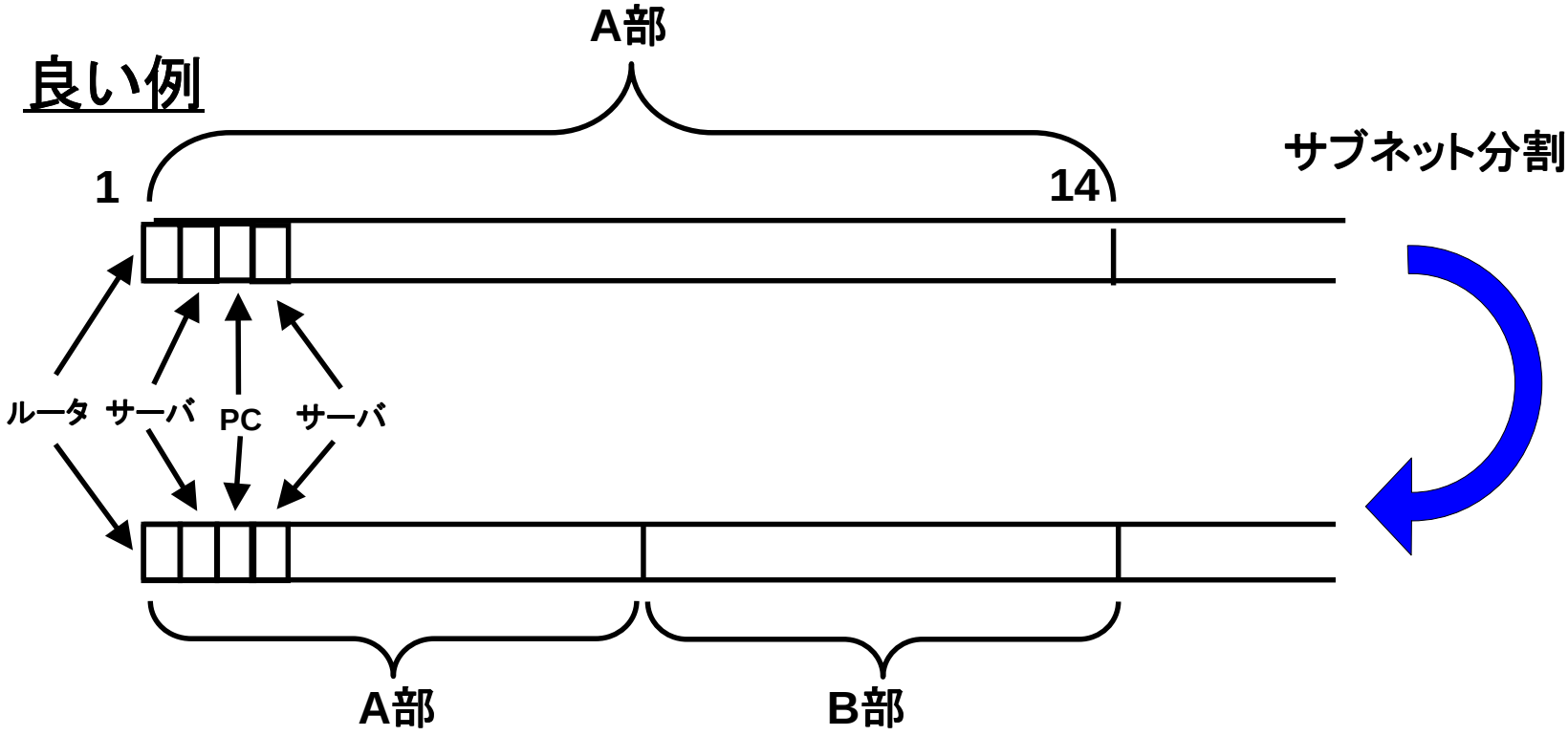
悪い例



- 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。

IIJ
Internet Initiative Japan

部署内でのアドレスの割り当て-2



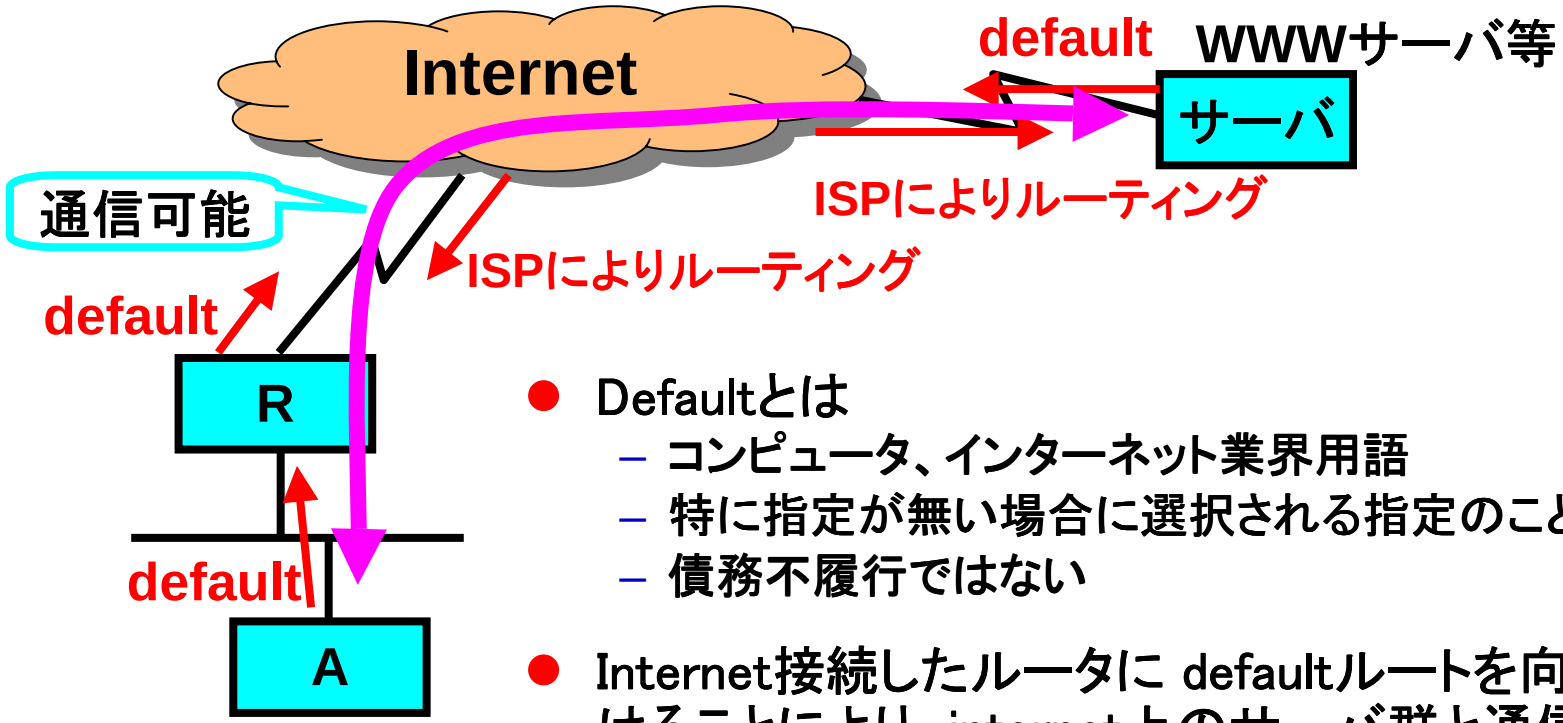
- アドレスを前詰めで使用すればさらなるサブネット化にもスムーズに対応できる

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良いのか
  - 先頭から前詰めで使用する

IIJ
Internet Initiative Japan

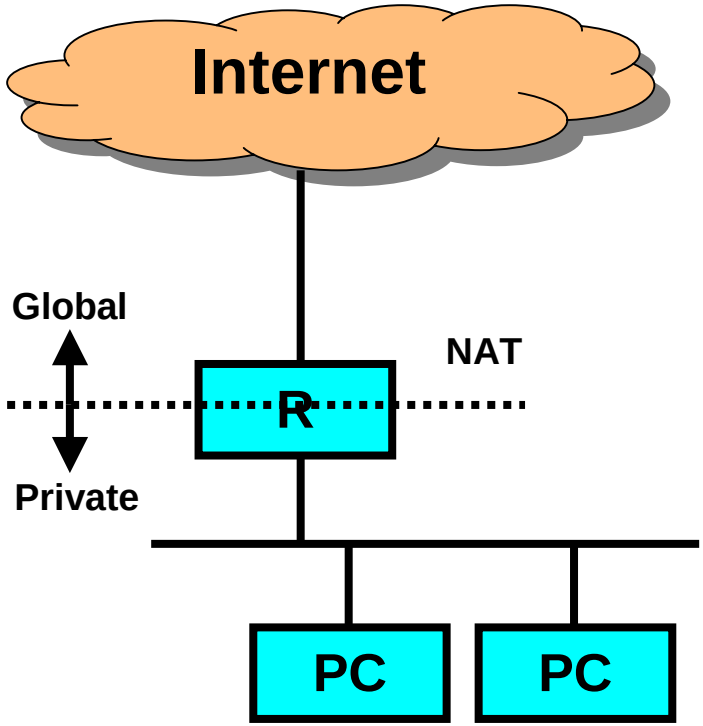
インターネットへの接続形態



- 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

IIJ
Internet Initiative Japan

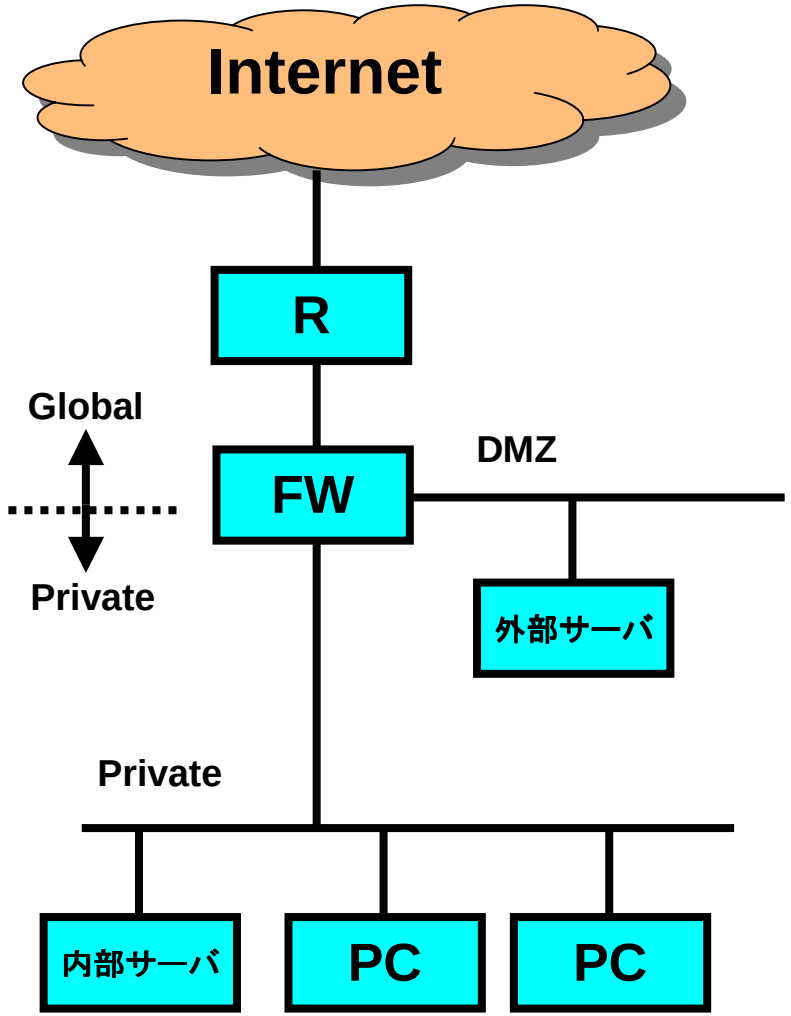
インターネット接続事例ーA



- 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

IIJ
Internet Initiative Japan

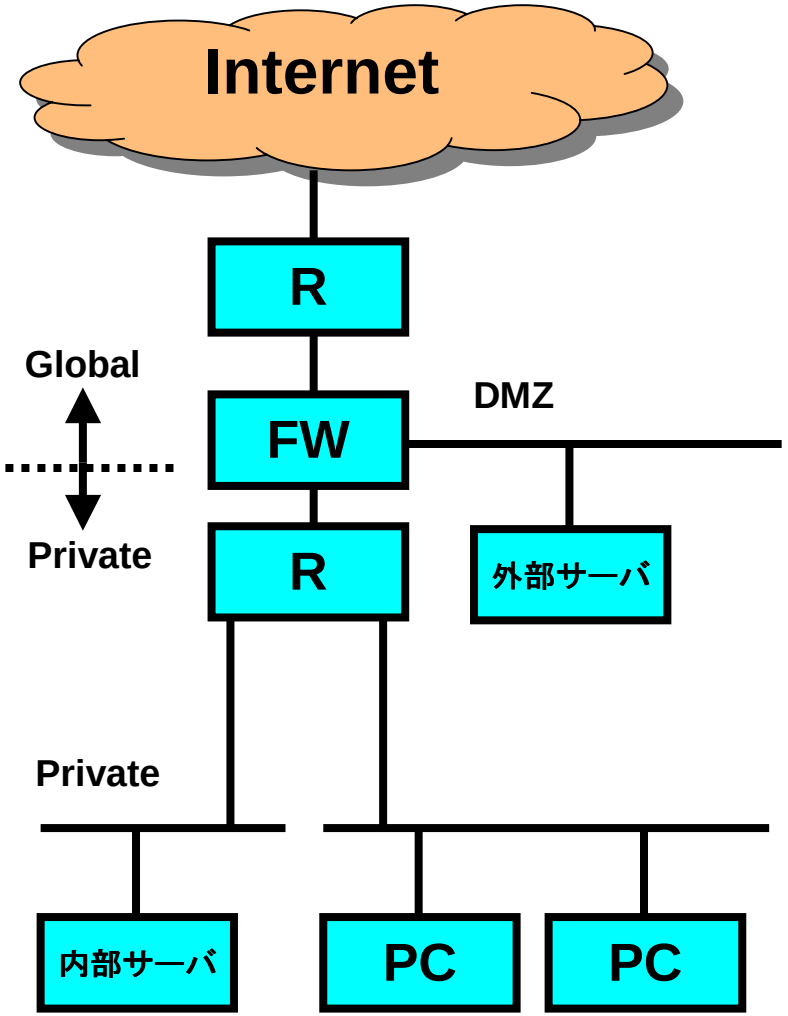
インターネット接続事例―B



- 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

IIJ
Internet Initiative Japan

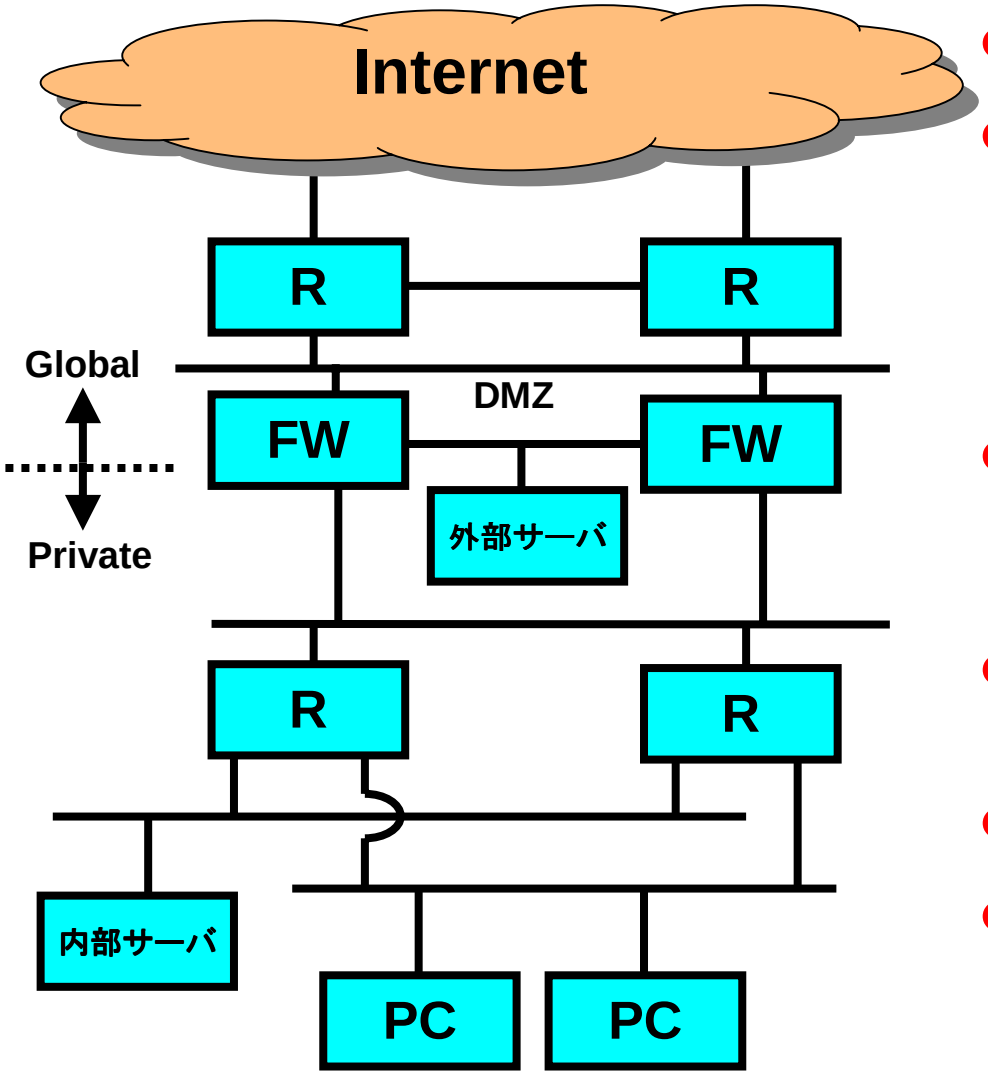
インターネット接続事例—C



- 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

IIJ
Internet Initiative Japan

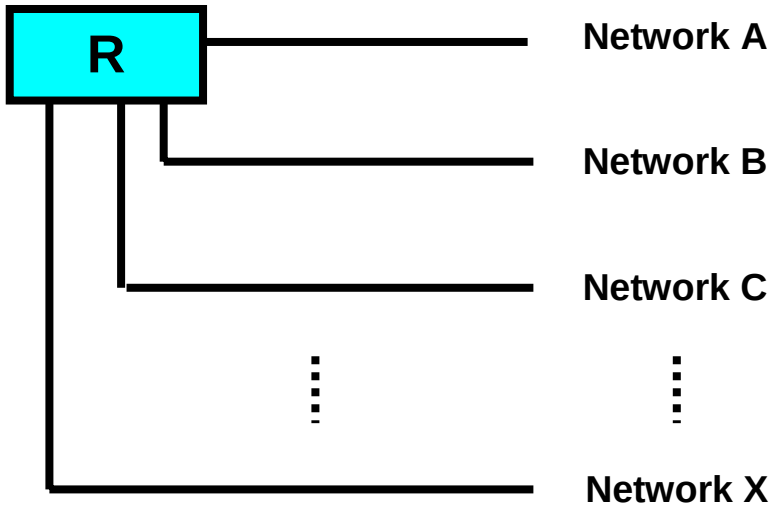
インターネット接続事例—D

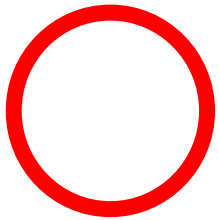


- 自社サーバ運用
- 外部サーバ
  - DMZ(Demilitarized Zone)に設置
  - 公開Web
  - 外部DNS
- 内部サーバ
  - Mail
  - 内部DNS
- セキュリティ
  - 標準的
- 内部サーバの保護
- 回線、機器の二重化
  - HSRP,VRRPの利用

IIJ
Internet Initiative Japan

ネットワーク拡張（スター型）



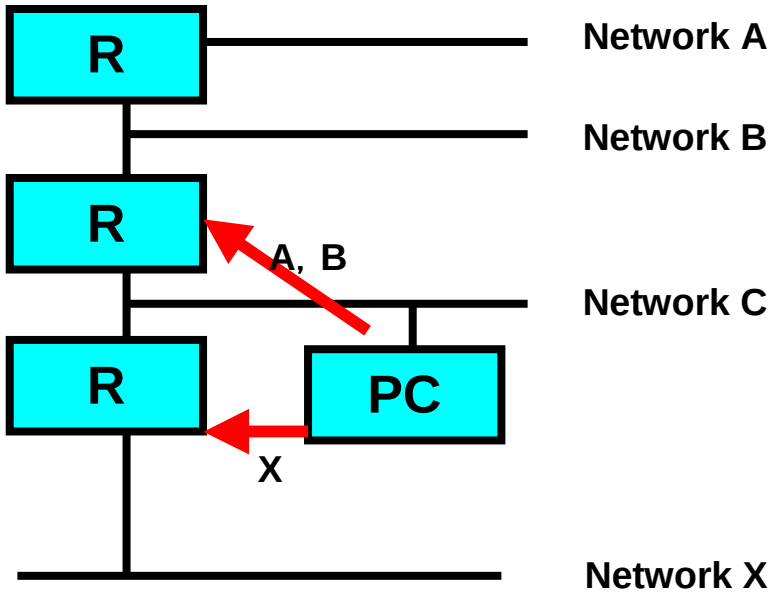


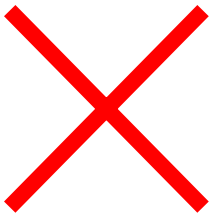
ネットワーク拡張の基本であり、特別な事情がない限り、まずこの形式を検討すべきである

- 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

IIJ
Internet Initiative Japan

ネットワーク拡張（数珠型）



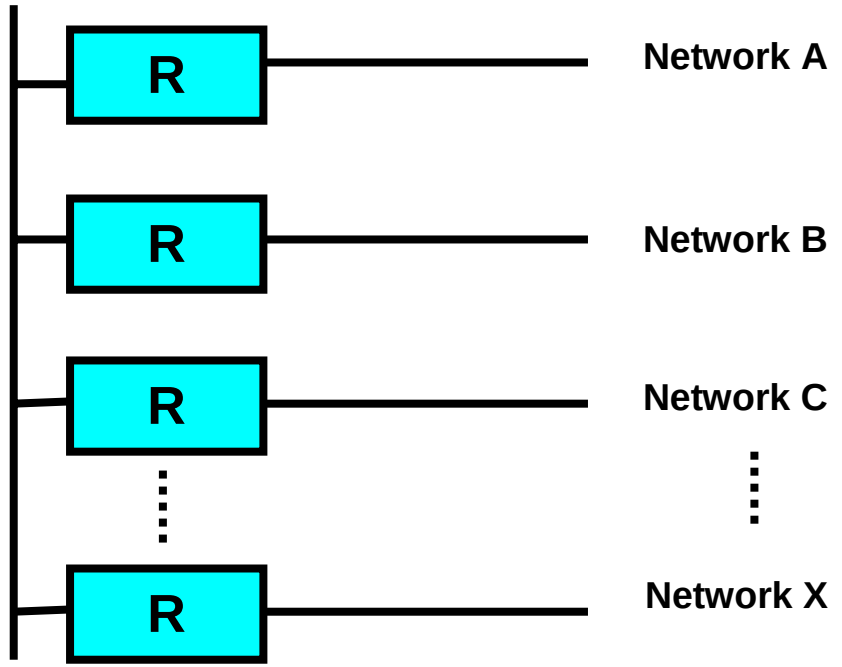


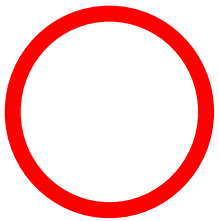
物理的にこの形式しか組めない場合を除いて避けるべき構成である

- 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

IIJ
Internet Initiative Japan

ネットワーク拡張（L2バックボーン型）



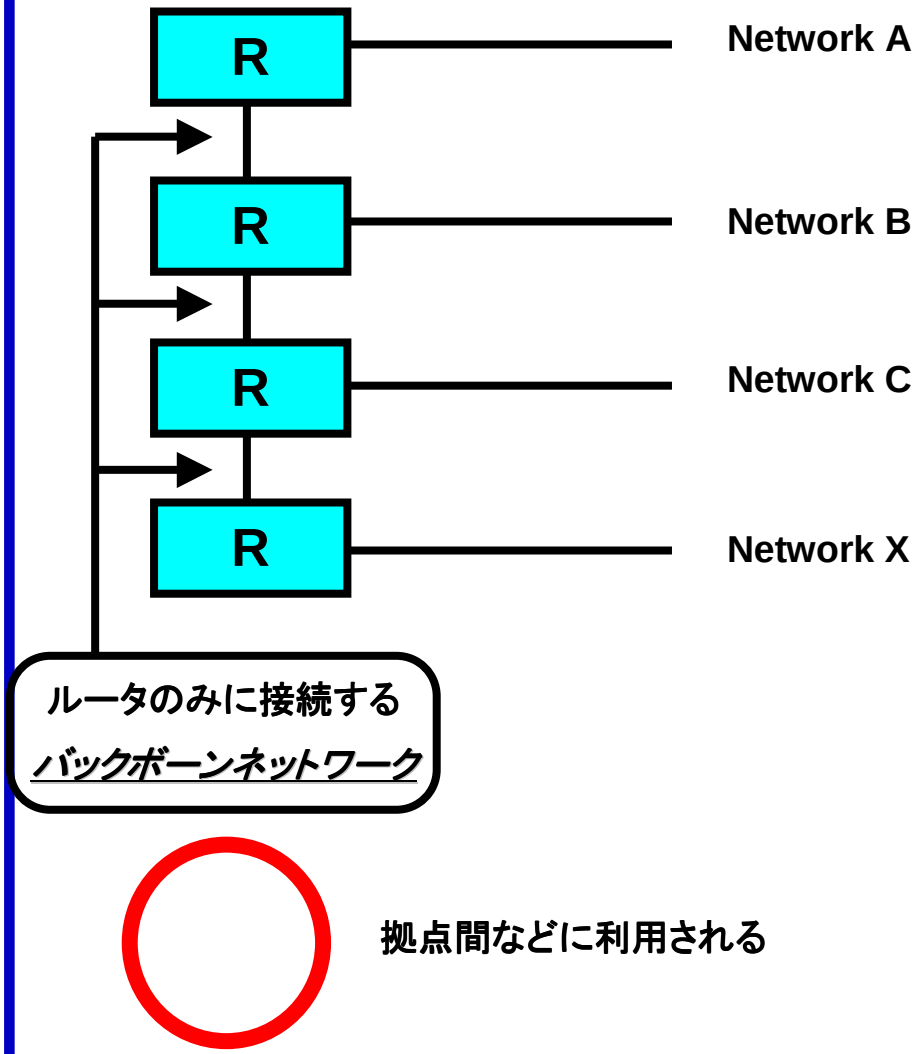


同一構内などのLAN接続などに有効に利用できる

- 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

IIJ
Internet Initiative Japan

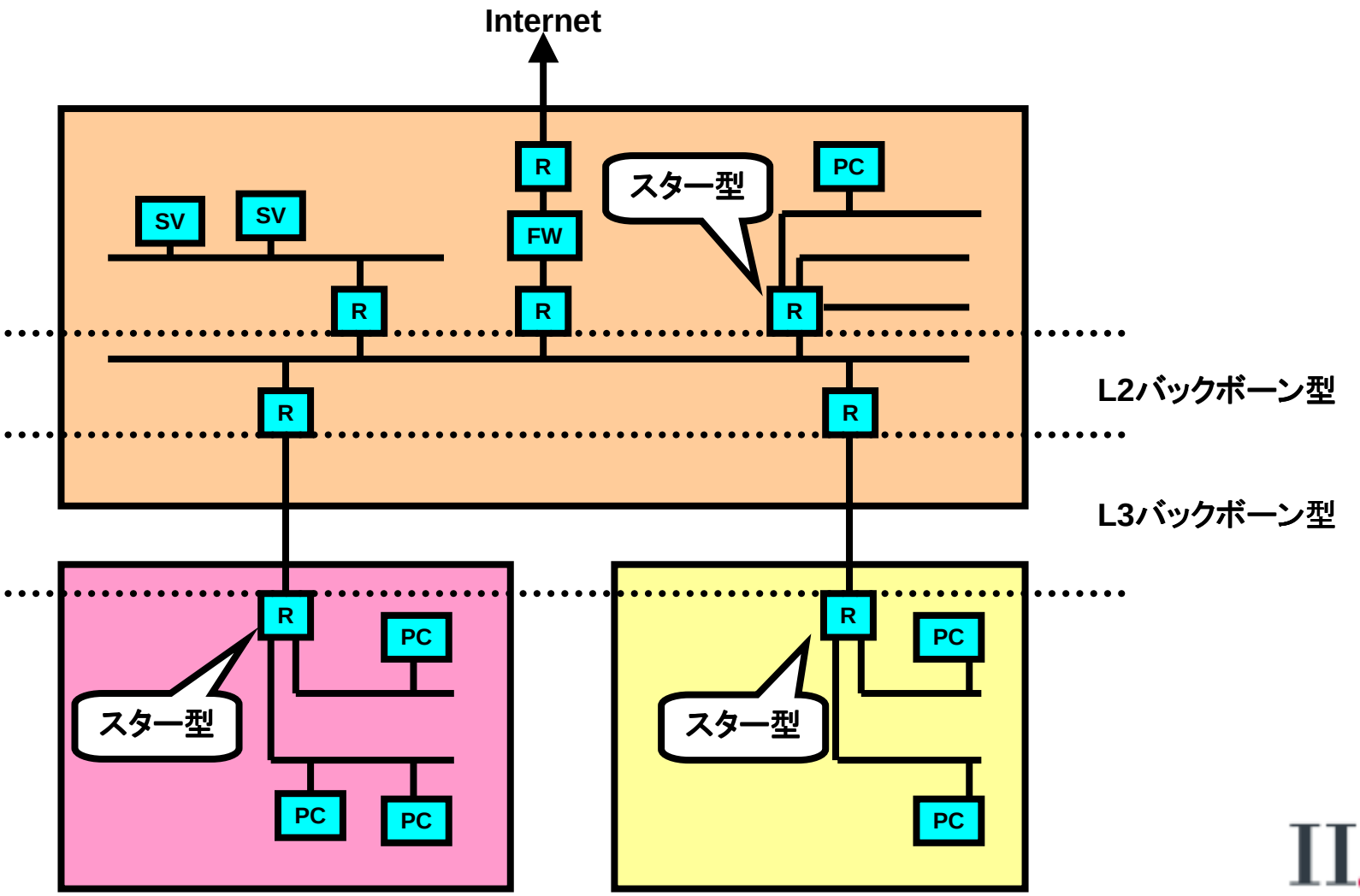
ネットワーク拡張（L3バックボーン型）



- 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

IIJ
Internet Initiative Japan

ネットワーク拡張事例



ネットワーク拡張のまとめ

- 小規模な同一構内のネットワークにはスター型を用いる
- 中規模の同一構内のネットワークにはL2バックボーン型を用いる
- 拠点間を結ぶネットワークにはL3バックボーン型を用いる
- 数珠型接続はできる限り避けるようにする

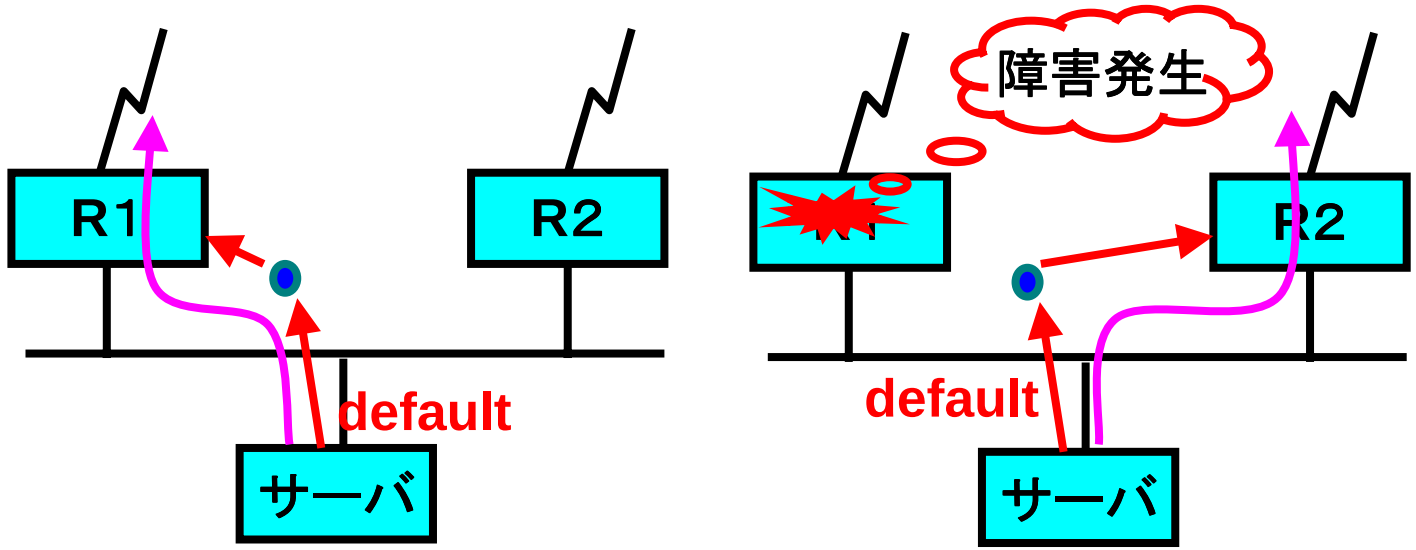
IIJ
Internet Initiative Japan

ネットワーク構築に利用される冗長化の仕組み

- 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～60秒程度必要
- 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
  - Ethernet専用線等のdownしないI/Fでは利用できない。
- HSRP/VRRP
  - 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

IIJ
Internet Initiative Japan

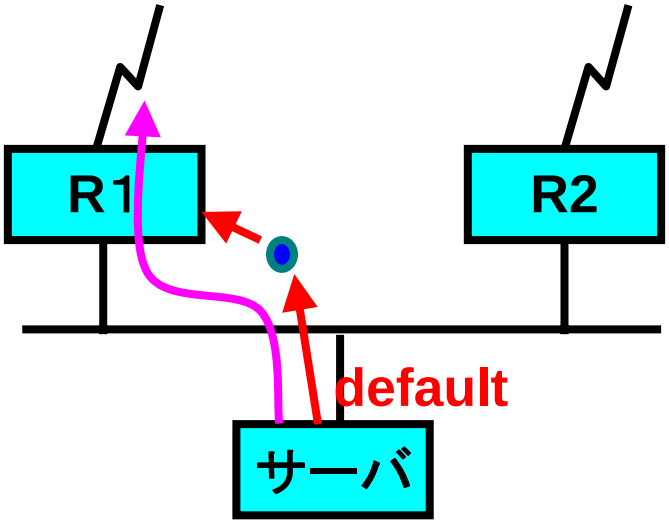
HSRP/VRRP-1



- 障害時には仮想MACアドレスがR1からR2に切り替わる
  - 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

IIJ
Internet Initiative Japan

HSRP/VRRP-2

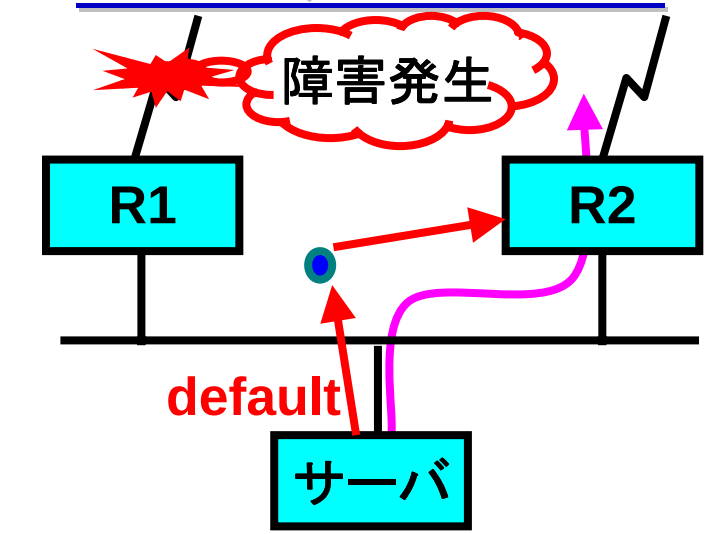


- HSRP+Interface Tracking（通常運用時）

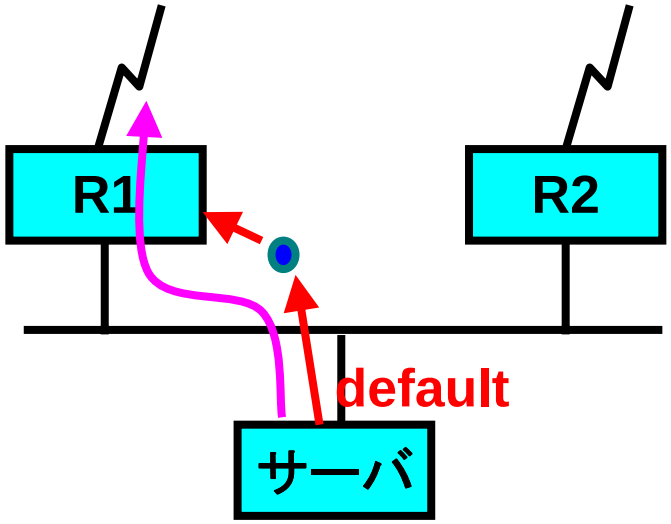
- Interfaceの downを検出して、Trackingすることで回線障害時にactiveルータの切り換えを行う

IIJ
Internet Initiative Japan

HSRP/VRRP-3



- HSRP+InterfaceTracking（障害発生時）
  - Interface Trackingにより切り替え

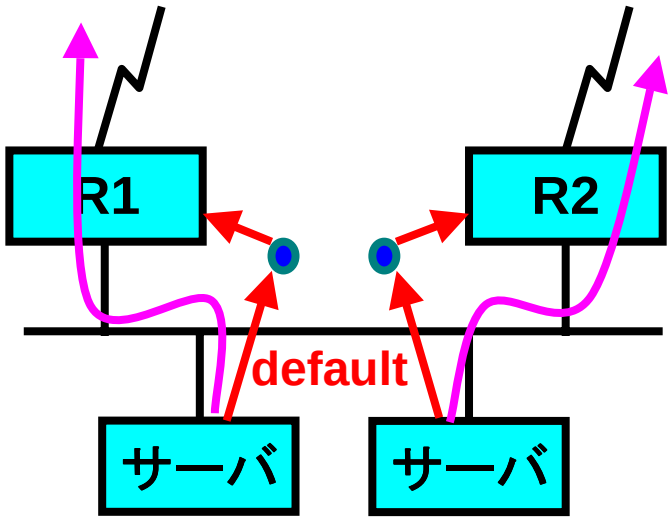


- HSRP+Interface Tracking（障害復旧時）
  - 復旧により切り戻しが発生

IIJ
Internet Initiative Japan

MHSRP-1

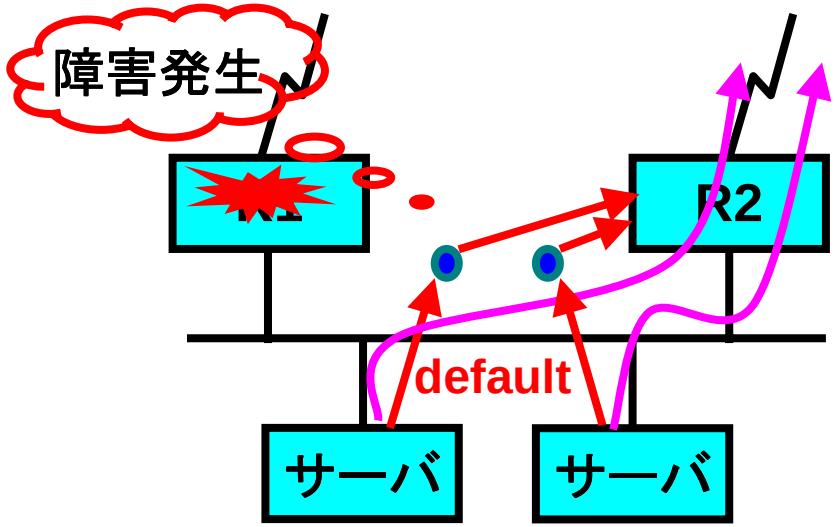
- マルチグループを用いて MHSRPを利用すれば、サーバ毎にトラフィックを分ける事ができる



- 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

IIJ
Internet Initiative Japan

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

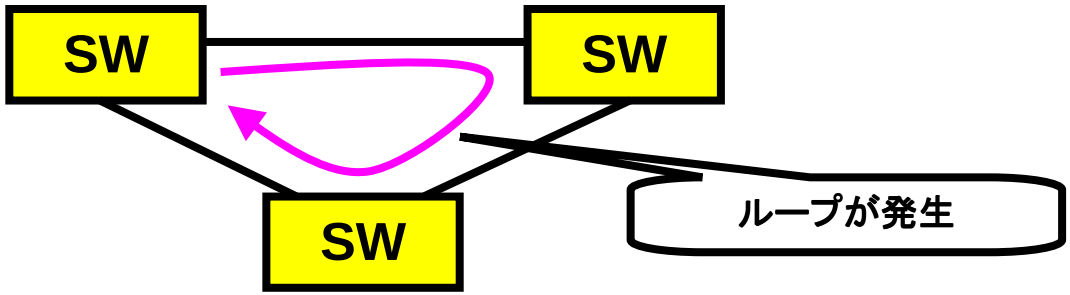
IIJ
Internet Initiative Japan

HSRP/VRRPまとめ

- 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
- HSRP
  - Hot Standby Router Protocol
  - RFC2281 (Informational)
  - Cisco社のパテント
- VRRP
  - Virtual Router Redundancy Protocol
  - RFC2338 (Proposed Standard)
  - ルータやファイアウォールなどに実装されている
- MHSRP
  - 1つのネットワークに複数のHSRPを同時利用し、不可分散することが可能

IIJ
Internet Initiative Japan

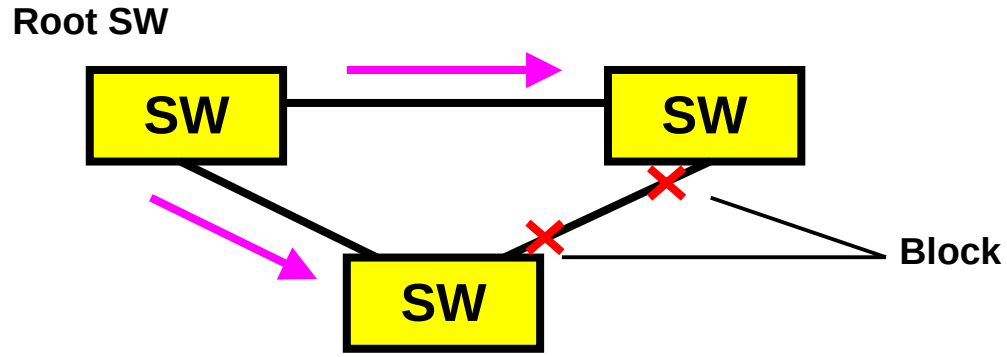
STPの動作-1



- STPを利用しない場合
  - STPを利用せずにスイッチの冗長化するとループが発生する
- ループの発生により様々な問題が発生
  - 各スイッチのアドレステーブルに混乱が生じる
  - 同じフレームが二重に届き、上位層の異常な動作につながる
  - Broadcast floodが発生する

IIJ
Internet Initiative Japan

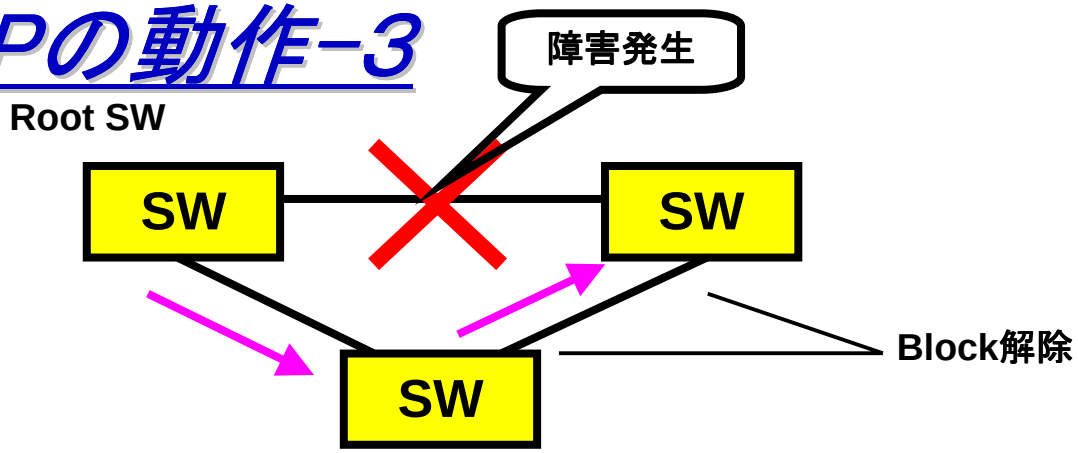
STPの動作-2



- STPを利用する
  - STPの利用によりRoot SWからツリーが構成され、冗長経路はブロッキングされる
  - ブロッキングによりループを防ぐ
- ブロッキングとは
  - 通信が止められている状態
  - ただし、STPのHelloは通信されている

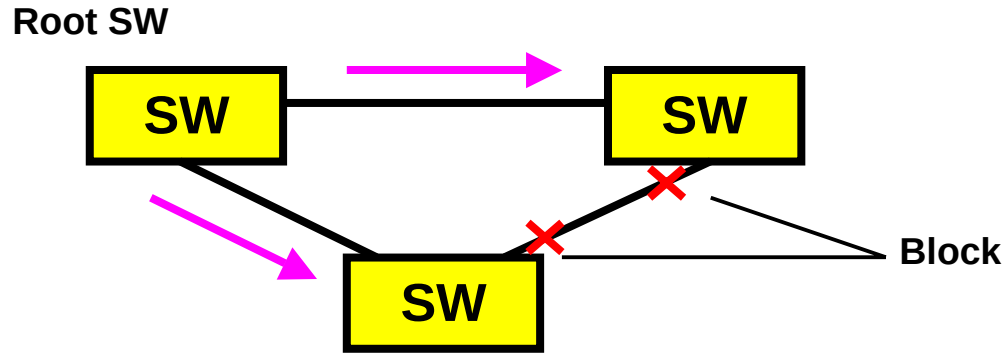
IIJ
Internet Initiative Japan

STPの動作-3



- 障害発生が発生した場合
  - 障害発生によりブロッキングが解除され、バックアップされる

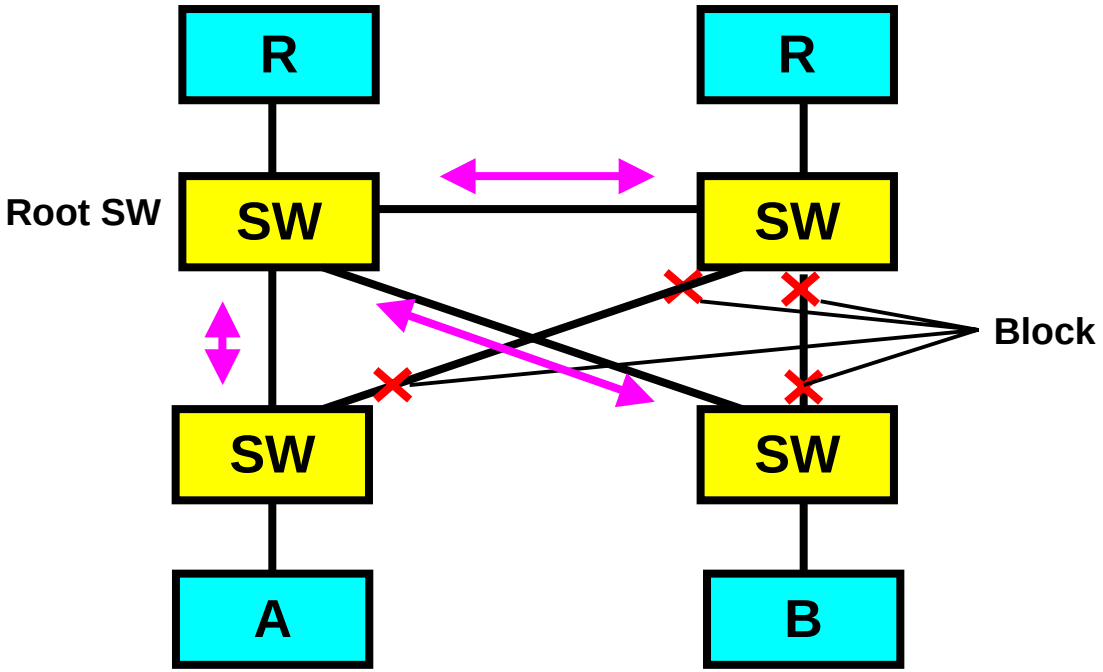
STPの動作-4



- 障害が復旧すると
  - 再びSTPによりRoot SWからツリーが構成され、冗長経路はブロッキングされる

IIJ
Internet Initiative Japan

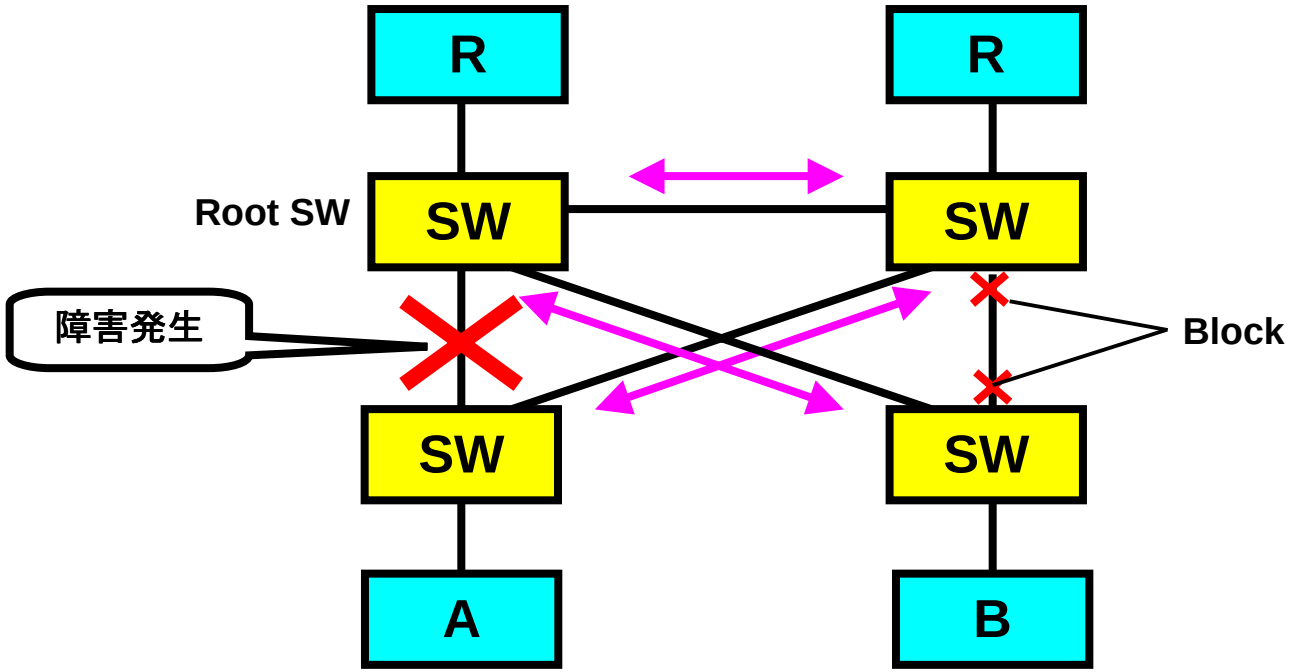
STP事例−1



- 全てのSWをRoot SWに最短となるように設計
- STPにより冗長化経路をブロッキングする

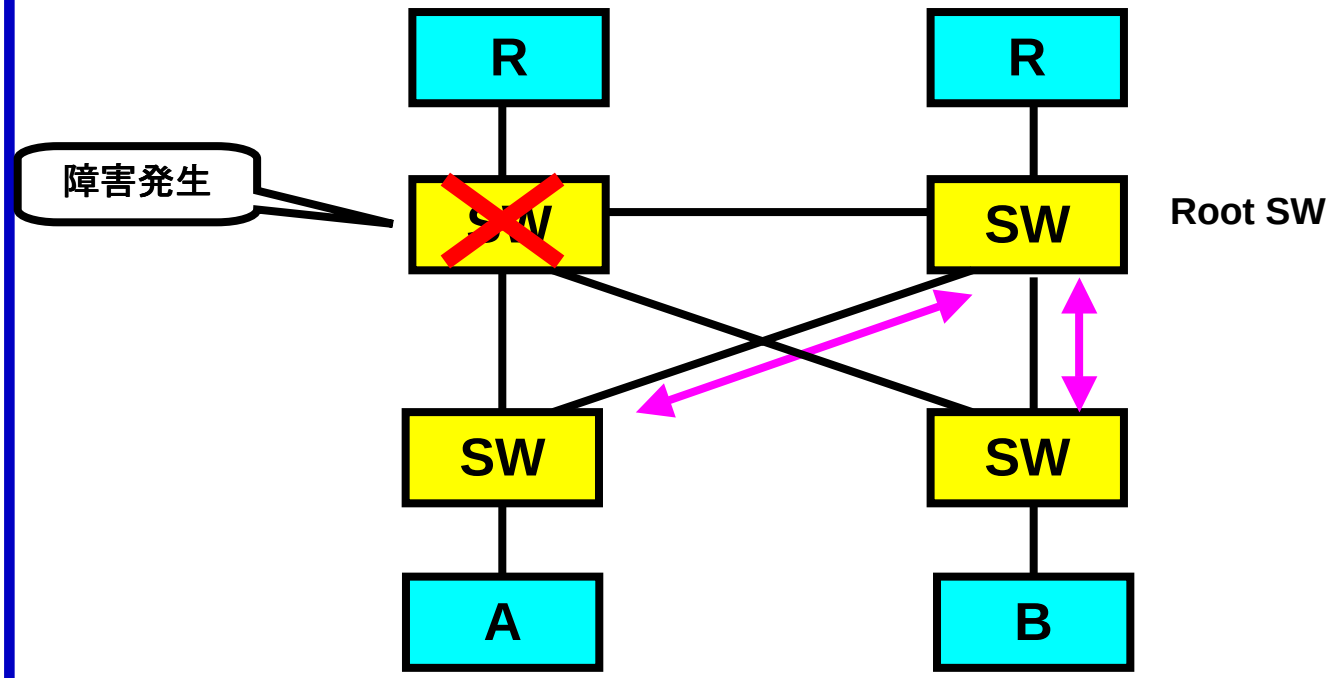
IIJ
Internet Initiative Japan

STP事例-2



- 障害が発生するとSTPによりバックアップ経路に切り替わる

STP事例-2



- SWに障害が発生した時もSTPによりバックアップされる
- Root SWに障害が発生した場合には切り替えに時間がかかる

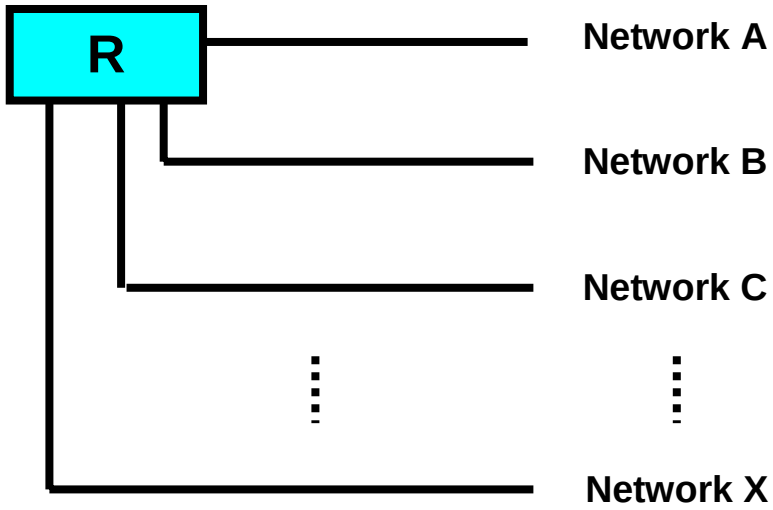
IIJ
Internet Initiative Japan

STPまとめ

- STP（Spanning Tree Protocol）
- IEEE 802．1Dの中で定義されている
- データリンク層（L2）プロトコル
- ルートSWからツリーを構成する
- ブロッキングによりループを防止する
- 遠隔地への伝送時などに有効に利用される
- ルートSWはMACなどにより決定されるが、設定することも可能

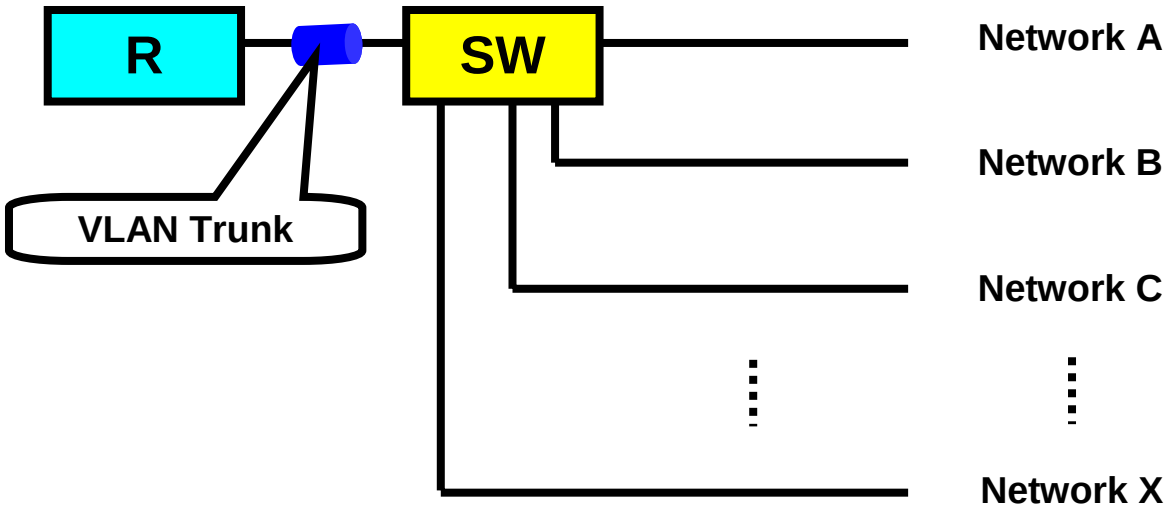
IIJ
Internet Initiative Japan

VLAN Trunk-1



- VLAN Trunkしない場合
  - 多くのネットワークを接続する場合、VLAN Trunkを利用しないとルータに多くのインターフェースを用意する必要があり、コストがかかる

VLAN Trunk-2



- VLAN Trunkを利用する
- VLAN Trunkによりルータのインターフェースが1つであっても仮想的に複数のインターフェースがあるように見せることができる
- ルータと比較して安価なスイッチのポートをルータのインターフェースとして見せることができる

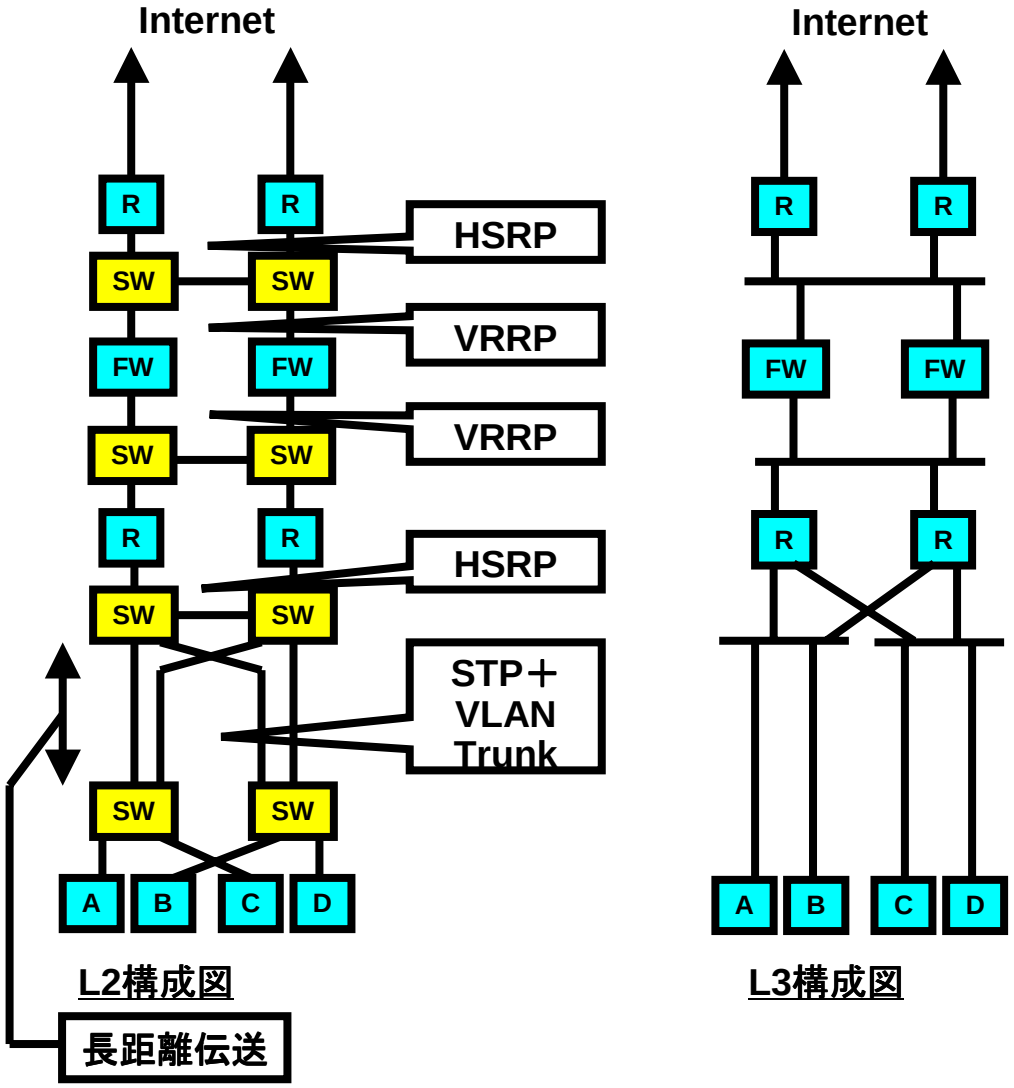
IIJ
Internet Initiative Japan

VLAN Trunkまとめ

- VLAN
  - Virtual LAN
  - 1つのスイッチ内の異なるLANの扱いをVLANと呼ぶ
  - VLAN Trunk、タグVLANのことを略してVLANと呼ぶこともある
- VLAN Trunk
  - 複数のVLANを1つのデータリンク層でまとめて通信する
  - 「タグ付VLAN」「タグVLAN」とも呼ばれる
  - IEEE 802.1Qで定義されている
  - メーカー独自のものも存在する
- VLAN Trunkによりルータのインターフェースが1つであっても仮想的に複数のインターフェースがあるように見せることができる
- ルータに比較して安価なスイッチのポートをルータのインターフェースとして見せることができる
- スイッチをカスケードして利用する場合にはスイッチのダウンを検知できなくなることがあるため注意が必要

IIJ
Internet Initiative Japan

STP+VLAN Trunk事例



- 冗長化されたネットワーク
  - HSRP/VRRP
  - STP
- VLAN Trunkを利用
- 長距離伝送の冗長化とコスト削減を実現
- L3構成図とL2構成図の違いとポイント
  - STPはL2冗長化プロトコルであるため、L3構成図には表れない
  - VLAN Trunkに関してもL同様にL3構成図に表れない

IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

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

IIJ
Internet Initiative Japan

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
- 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

IIJ
Internet Initiative Japan

まとめ－2

- 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
- 冗長化のためにSTP、HSRP/VRRPなどを利用する
- VLAN Trunkによりポート単価を下げることが可能
- 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

IIJ
Internet Initiative Japan