ネットワーク構築 A to Z [I]

～基礎から始める最適ネットワーク設計～

2005年12月9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

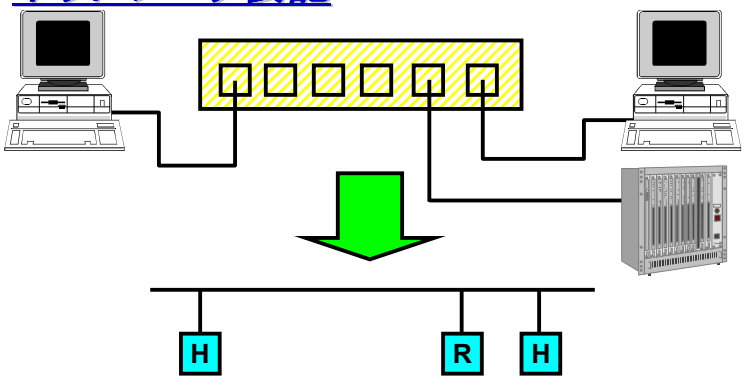
IIJ Internet Initiative Japan

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
- ネットワーク構築
- 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

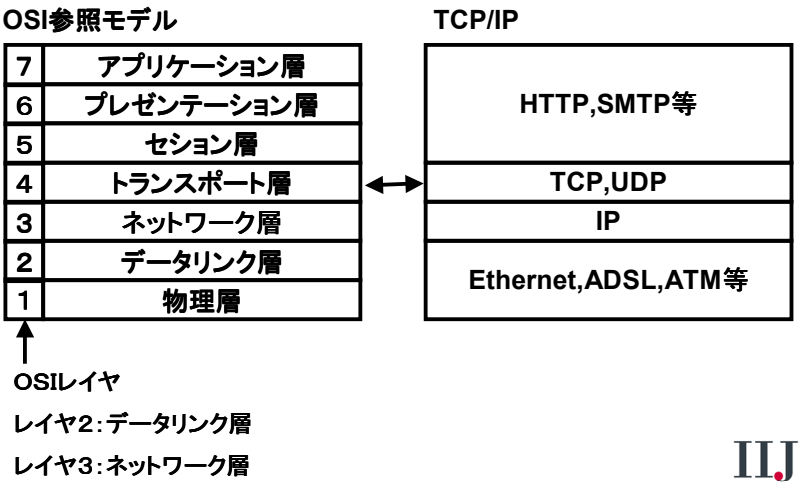
IIJ Internet Initiative Japan

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

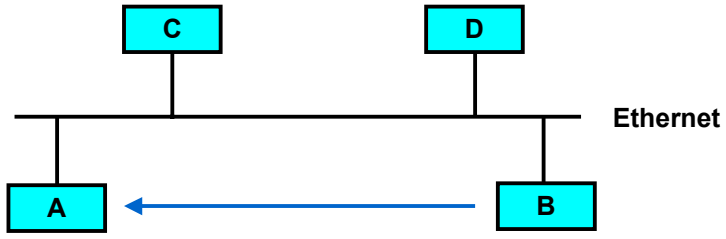
レイヤ3：ネットワーク層

IIJ Internet Initiative Japan

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

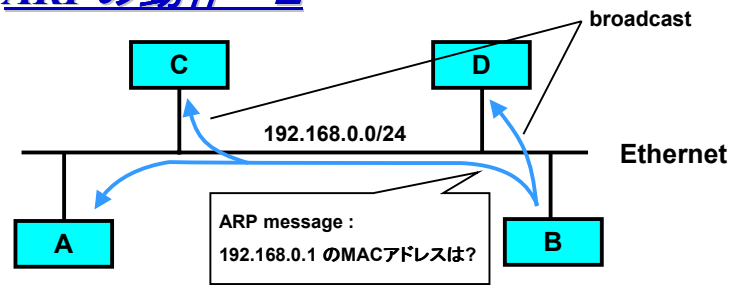
Host AのIP/MACアドレス対応表

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

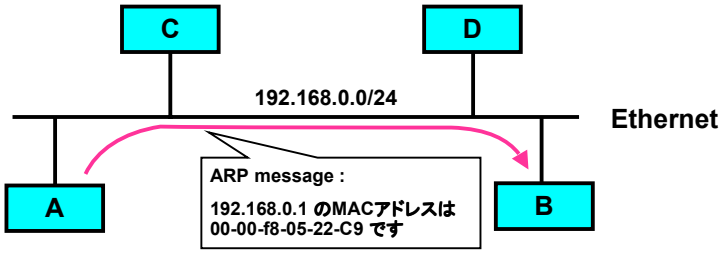
Host	IPアドレス	MACアドレス
A	192.168.0.1	不明
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

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

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

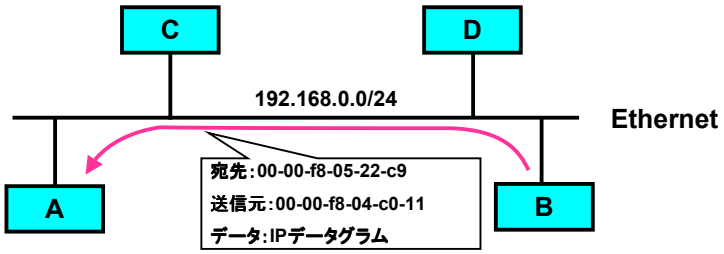
Host	IPアドレス	MACアドレス
A	192.168.0.1	00-00-f8-05-22-c9
B	192.168.0.2	00-00-f8-04-c0-11
C	192.168.0.3	不明
D	192.168.0.4	不明

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)ー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

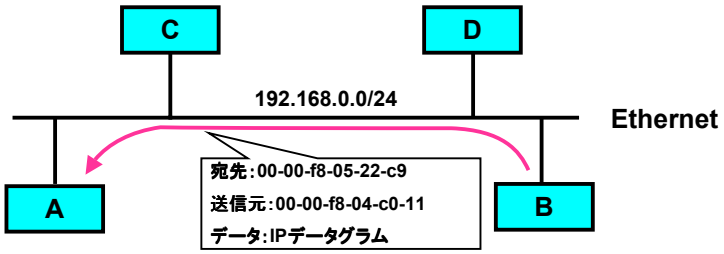
IIJ Internet Initiative Japan

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
      →詳しくは後述

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

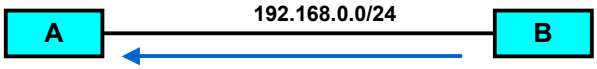
Host BのIP/MACアドレス対応表(固定)

- 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

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

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

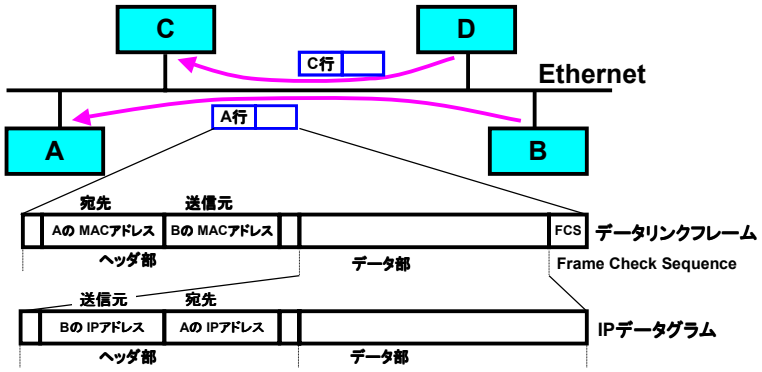
IIJ Internet Initiative Japan

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



IIJ Internet Initiative Japan

Connectedなネットワーク

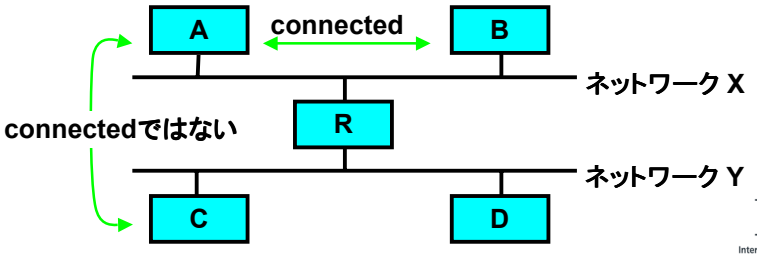
- A、Bは直接同じネットワークに接続している
  - MACアドレス、IPアドレスの対応表を ARP(address resolution protocol)などにより持っている

↓

- これを「connected」な状態という

↓

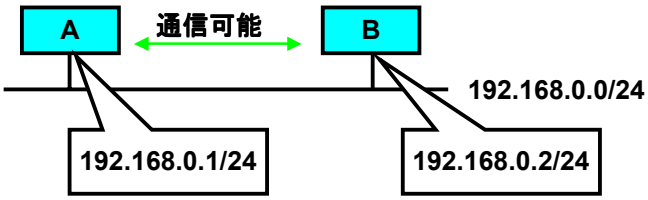
- ルーティング設定が不要で、ハブなどで接続すると通信できる



IIJ Internet Initiative Japan

ネットワーク層から見たConnectedなネットワーク－1

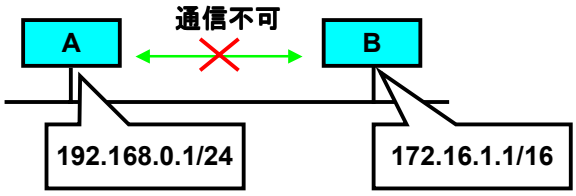
- Aのアドレス
  - 192.168.0.1/24
- Aから見たConnectedなアドレス空間
  - 192.168.0.0 ～ 192.168.0.255
- Bに192.168.0.2 ～ 192.168.0.254のアドレスを付ける
  - Bに192.168.0.2を付ける
  - A-B間の通信が可能



IIJ Internet Initiative Japan

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



IIJ Internet Initiative Japan

Connectedではないネットワーク－1

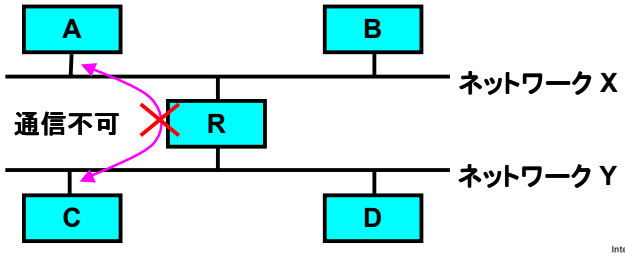
- A、Cはそれぞれ異なるネットワークに接続しているため connectedではない
- ルーティング*設定なし*ではA、C間の通信はできない

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



IIJ Internet Initiative Japan

Connectedではないネットワーク－2

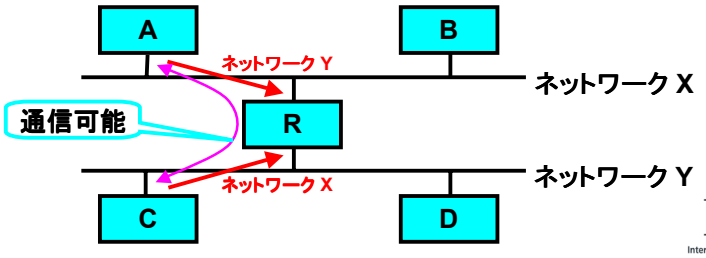
- ルーティング設定を行なう
  - A: ネットワークYを Rにルーティング
  - C: ネットワークXを Rにルーティング
- これにより、A⇔C間の相互通信が可能となる
  - Rは A,C共に connectedなため、アドレスを設定するだけで通信が可能

Aのルーティングテーブル

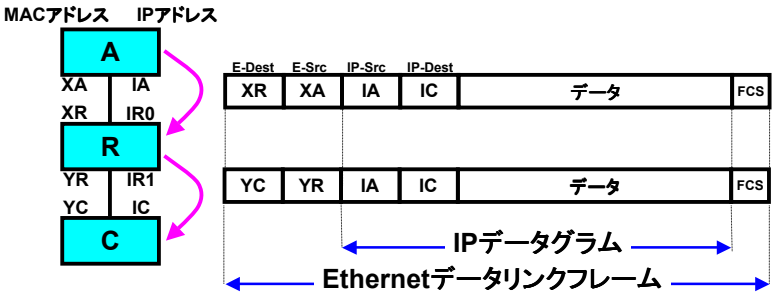
destination	Next Hop	到達性
X	Connected	到達可
Y	R	到達可

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」

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

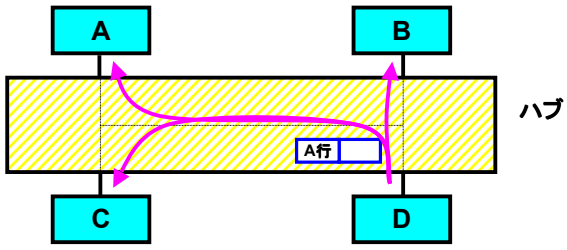
スイッチとルータの機能の違い

- 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
- 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

IIJ Internet Initiative Japan

ハブとスイッチの違い-1

ハブで構成した場合

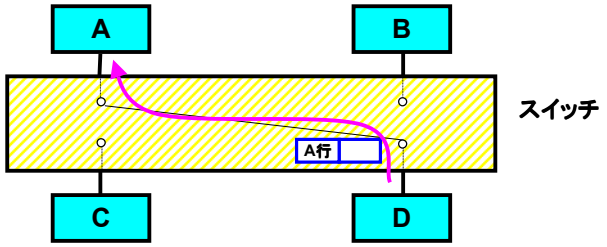


- ハブは全てのポートが常時接続された状態になっている
- このため異なるポート間の通信を、通信に関係の無い他のポートに伝播して、他の通信を妨げる

IIJ Internet Initiative Japan

ハブとスイッチの違い-2

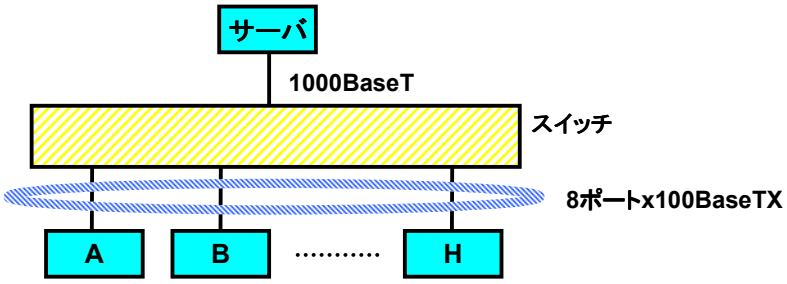
スイッチで構成した場合



- 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

IIJ Internet Initiative Japan

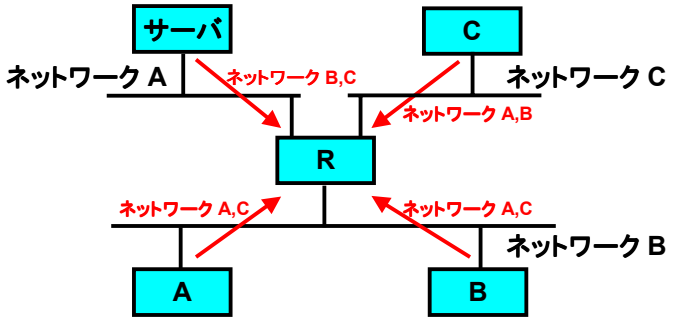
スイッチを有効に使うには



- 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

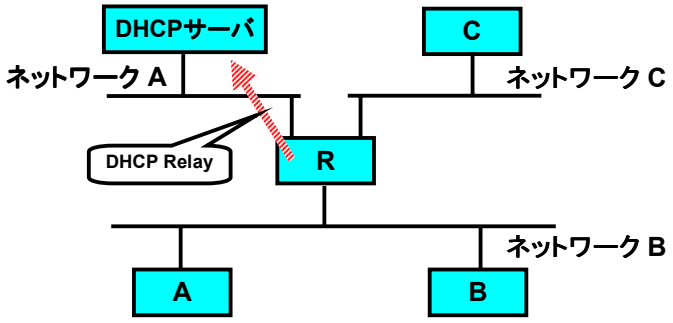
IIJ Internet Initiative Japan

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

ルータを利用するための設定－2



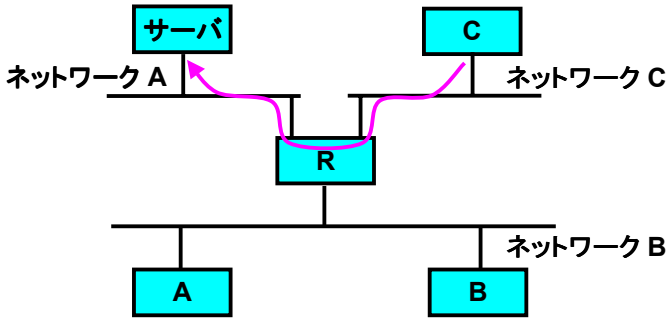
- DHCPサーバ設定
  - DHCPサーバは同一ネットワークに存在する必要がある
  - ルータのDHCP Relay設定により異なるネットワークでもDHCPを利用できる

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

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

IIJ Internet Initiative Japan

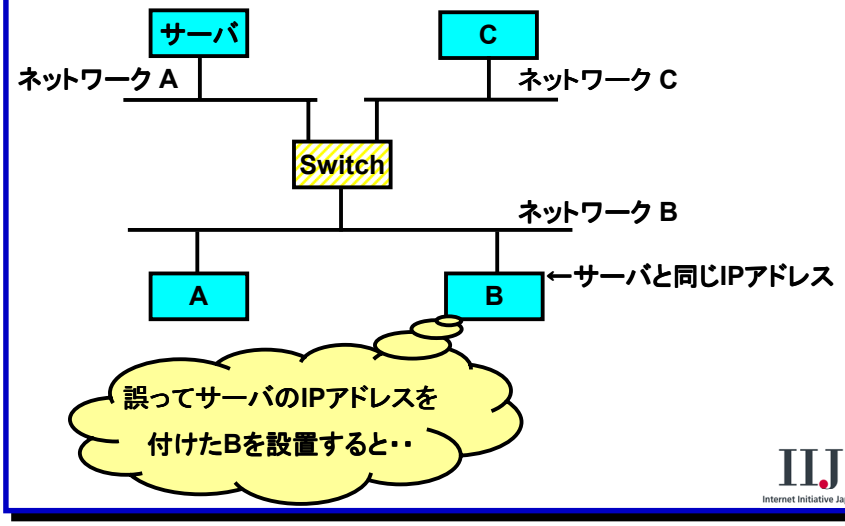
スイッチとルータの機能の違い

- 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

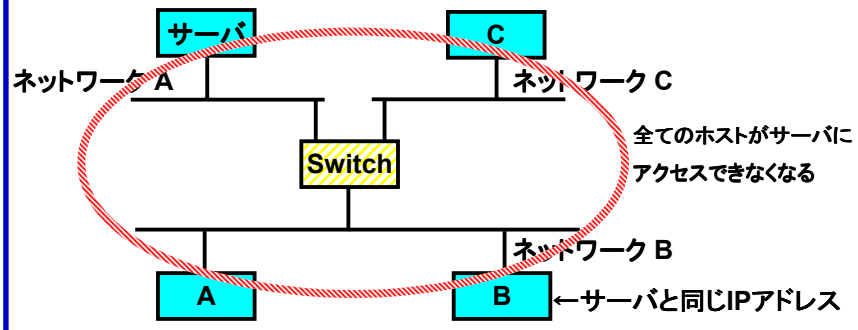
IIJ Internet Initiative Japan

スイッチの耐障害性-1



IIJ Internet Initiative Japan

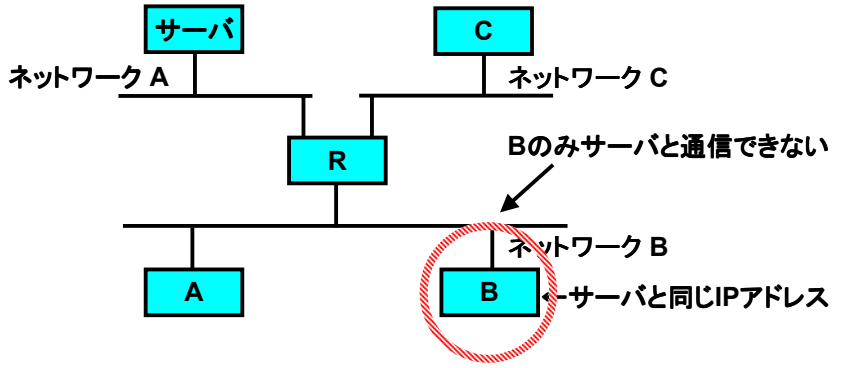
スイッチの耐障害性-2



- 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

IIJ Internet Initiative Japan

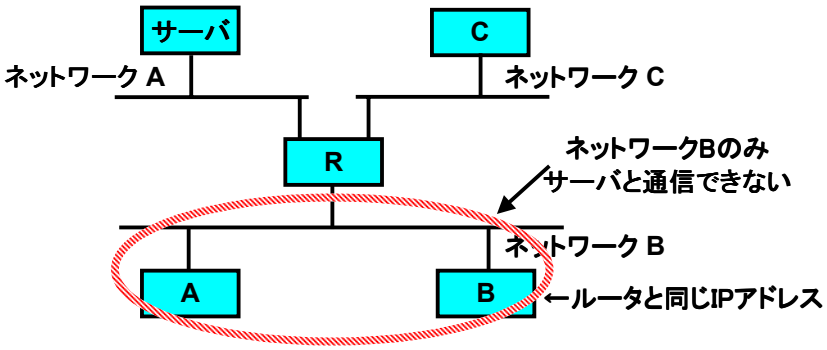
ルータの耐障害性-1



- 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

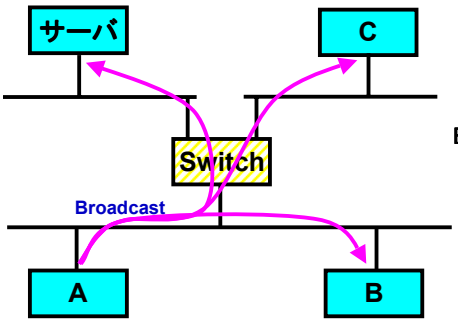
ルータの耐障害性-2



- 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

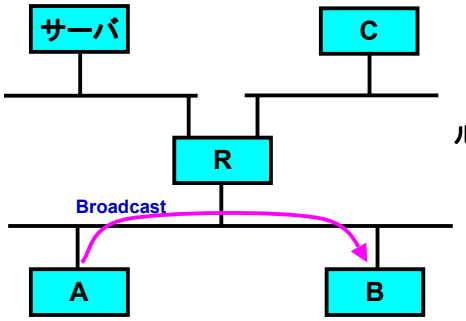


Broadcastパケットはスイッチの全てのポートに影響を与える

- 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

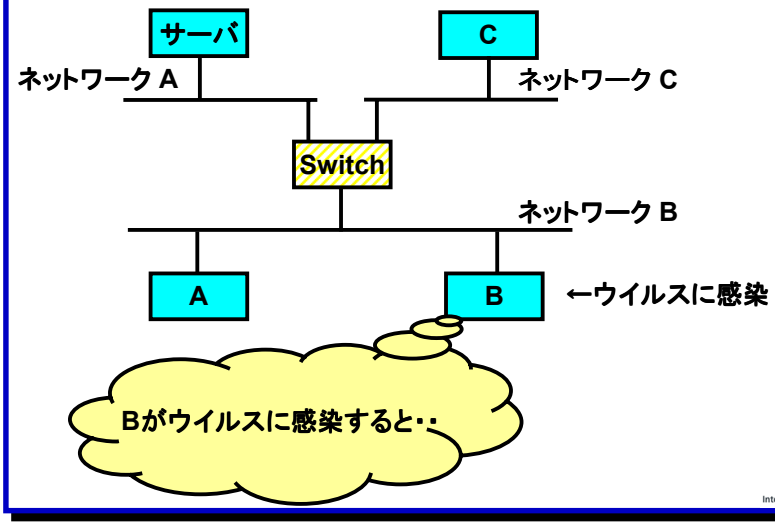


ルータは Broadcastパケットを
他のネットワークに通さない

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

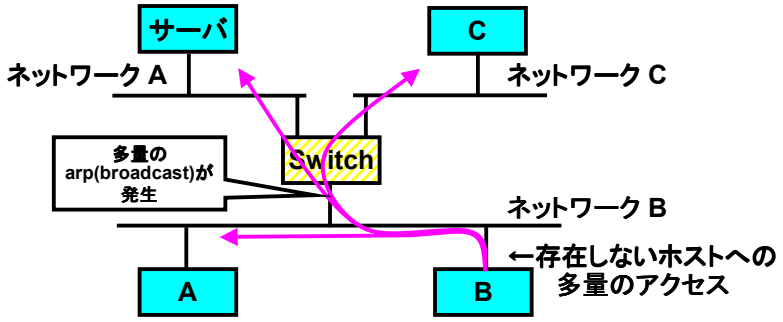
IIJ Internet Initiative Japan

スイッチの耐ウイルス障害性-1



IIJ Internet Initiative Japan

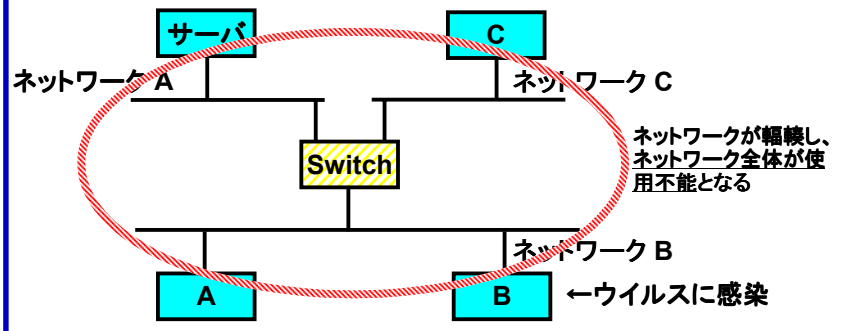
スイッチの耐ウイルス障害性-2



- 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

IIJ Internet Initiative Japan

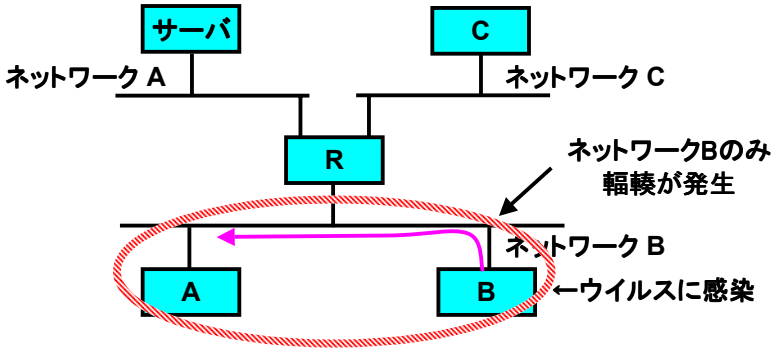
スイッチの耐ウイルス障害性-3



- 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

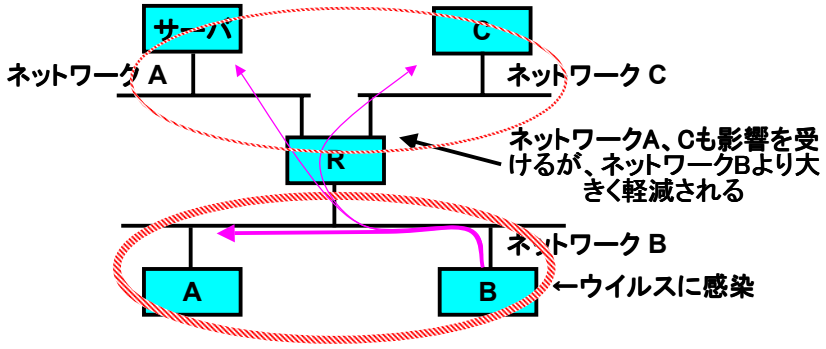
IIJ Internet Initiative Japan

ルータの耐ウイルス障害性-1



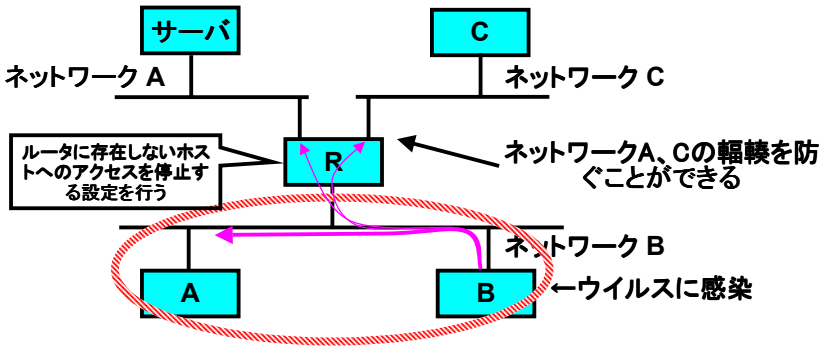
- ルータでは、ウイルス感染ホストによる多量の arp(broadcast)の影響は接続ネットワークに限られる

ルータの耐ウイルス障害性-2



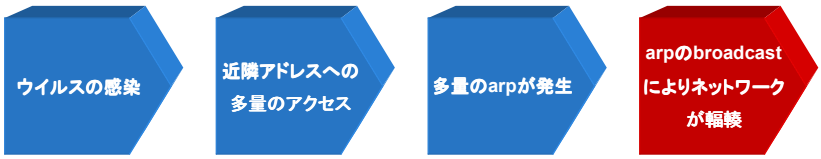
- ルータを越えるパケットも存在するが、ネットワークBに比べるとネットワークA、Cへの影響は小さくなる

ルータの耐ウイルス障害性-3



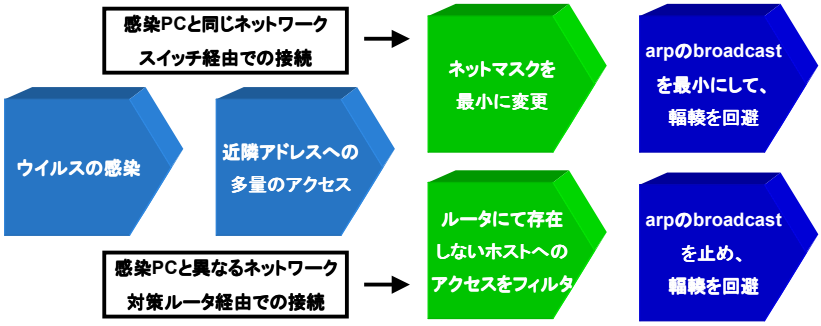
- 存在しないホストに対するフィルタをルータに行うことでネットワークA、Cに対する輻輳を防ぐことができる
- ネットマスクをできる限り小さくすることでフィルタと同様の効果を得ることができる

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

IIJ Internet Initiative Japan

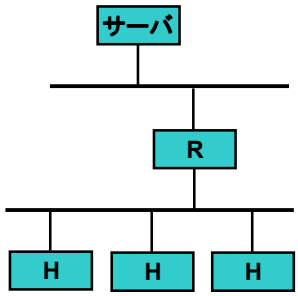
スイッチ VS ルータ

- 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
  - 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

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

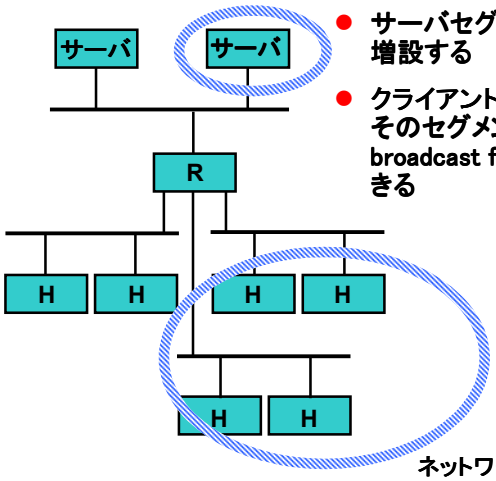


- 小規模であってもサーバのセグメントを分離する
  →サーバの安全性を確保する
- クライアントはDHCPによりアドレスの割り当てとdefault経路を得る
- Broadcast floodのサーバへの影響を防ぐ

IIJ Internet Initiative Japan

ネットワーク設計-2

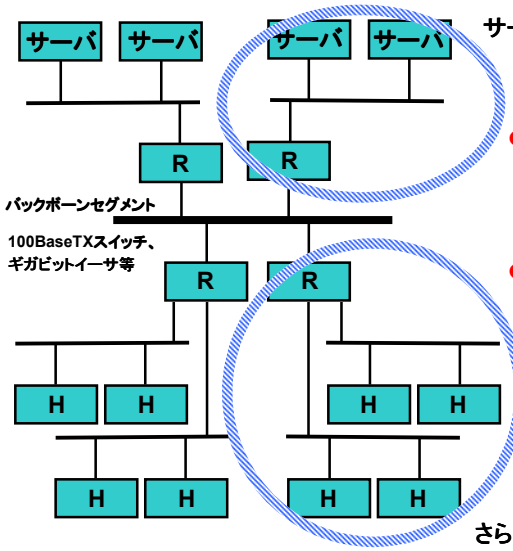
サーバの増設



- 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

IIJ Internet Initiative Japan

ネットワーク設計-3



- 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

IIJ Internet Initiative Japan

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
- 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。

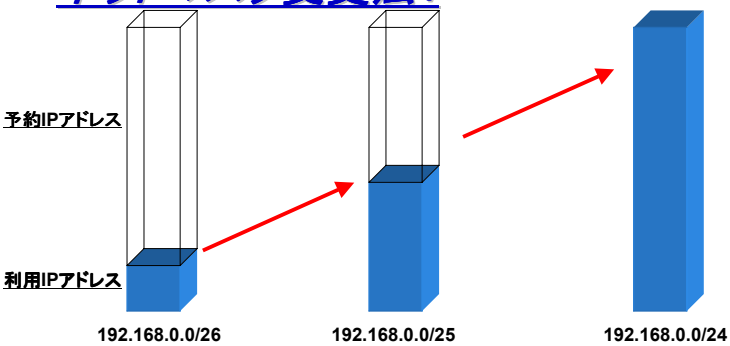
最適なネットマスクでの利用

- 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？

IIJ Internet Initiative Japan

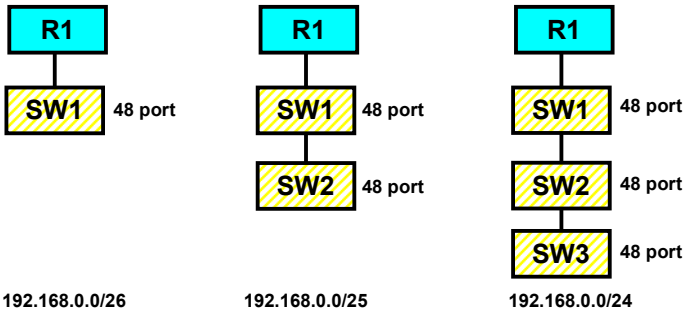
ネットマスク変更法1



- 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

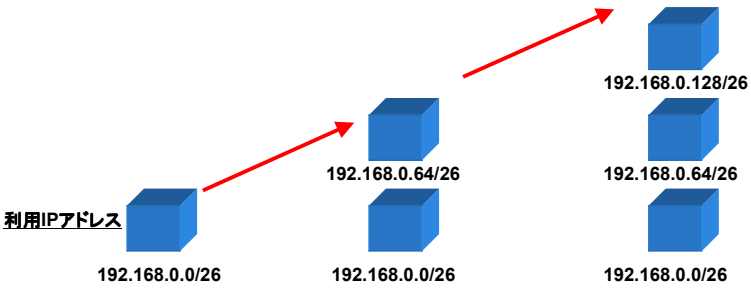
IIJ Internet Initiative Japan

ネットマスク変更法2



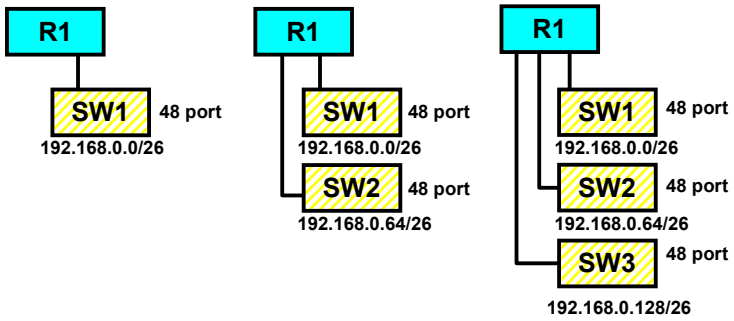
- ネットマスク変更法によるネットワーク拡張
  - マスク変更を前提とした割り当てはL2ネットワークを拡張していく必要があるため、カスケードされたL2ネットワークとなる場合が多い
  - カスケードされたL2ネットワークは障害箇所の特定などが困難となるだけでなく、broadcastドメインを大きくし、パフォーマンスを低下させる

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

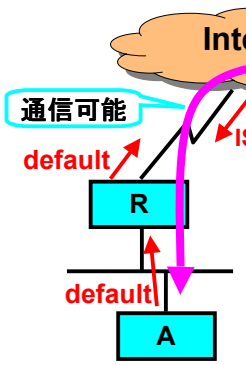


- サブネット追加法によるネットワーク拡張
  - サブネット追加による拡張では、スイッチの物理ポートとサブネットが対応し、スイッチの増設に合わせてサブネットも追加される
  - broadcastドメインの広さ(L2ネットワークの大きさ)を一定の大きさに保つことができるため、安定した拡張を行うことができる
  - 管理しやすくするために/24を予約して、/26だけ使っても良い
    - スペース重視: 192.168.0.0/26, 192.168.0.64/26, 192.168.0.128/26
    - 管理重視: 192.168.0.0/26, 192.168.1.0/26, 192.168.2.0/26

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

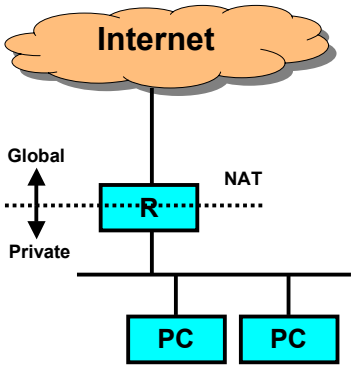
インターネットへの接続形態



- 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

IIJ Internet Initiative Japan

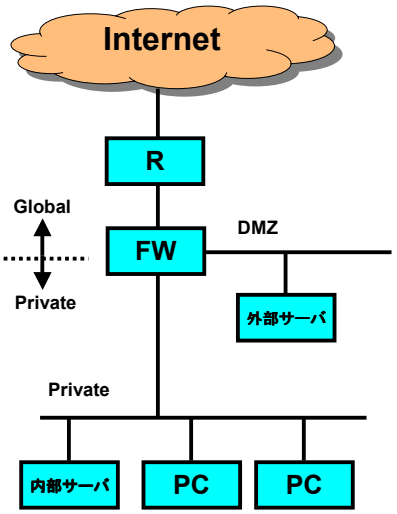
インターネット接続事例—A



- 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

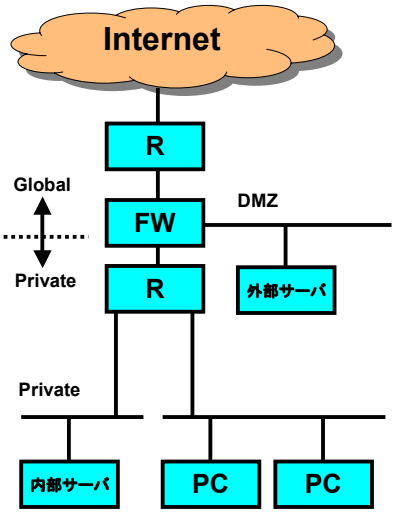
インターネット接続事例—B



- 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

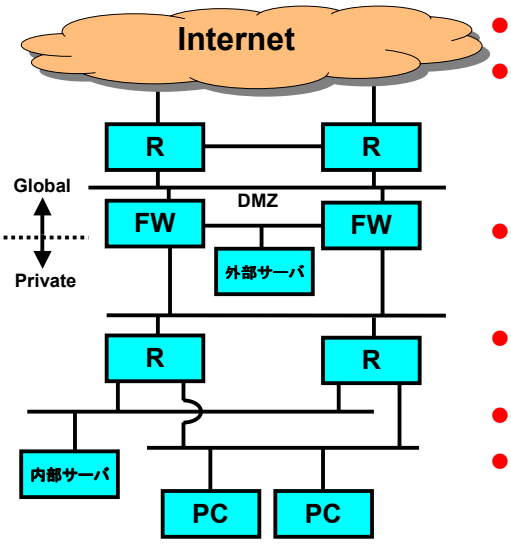
インターネット接続事例—C



- 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

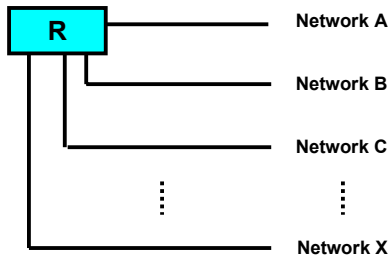
インターネット接続事例—D

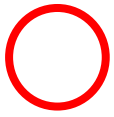


- 自社サーバ運用
- 外部サーバ
  - DMZ(Demilitarized Zone)に設置
  - 公開Web
  - 外部DNS
- 内部サーバ
  - Mail
  - 内部DNS
- セキュリティ
  - 標準的
- 内部サーバの保護
- 回線、機器の二重化
  - HSRP,VRRPの利用

IIJ Internet Initiative Japan

ネットワーク拡張(スター型)



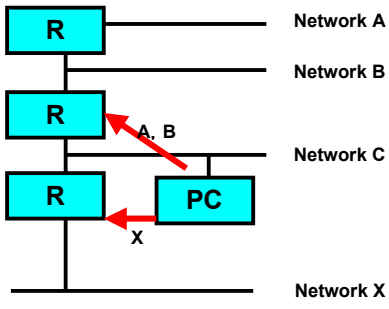


ネットワーク拡張の基本であり、特別な事情がない限り、まずこの形式を検討すべきである

- 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）

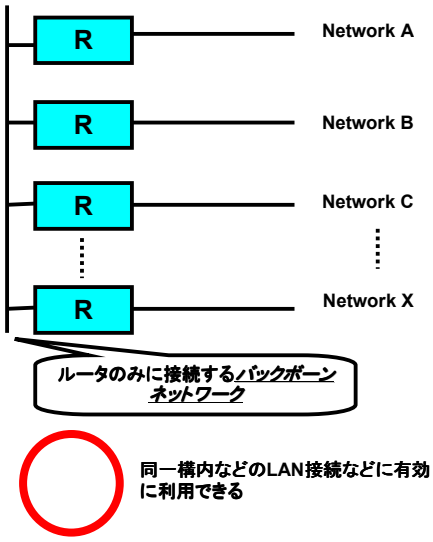


物理的にこの形式しか組めない場合を除いて避けるべき構成である

- 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）

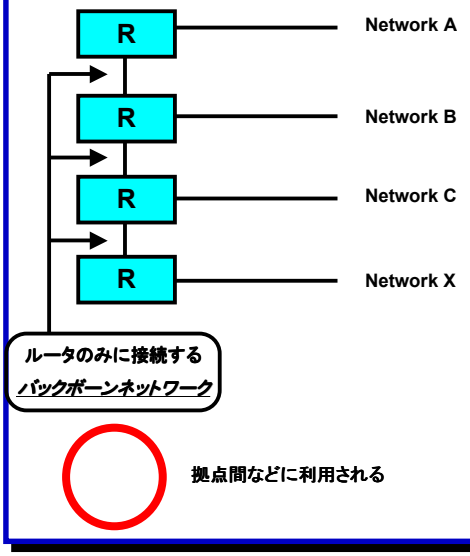


同一構内などのLAN接続などに有効に利用できる

- 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

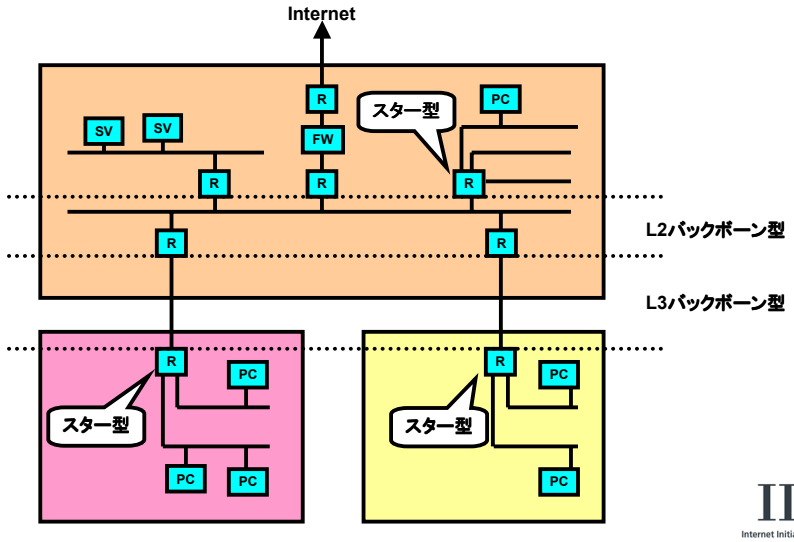
ネットワーク拡張（L3バックボーン型）



- 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



IIJ Internet Initiative Japan

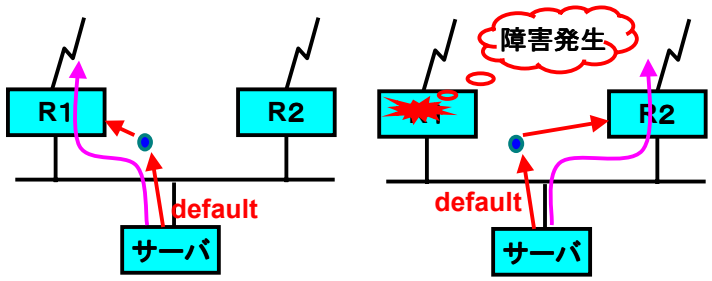
ネットワーク拡張のまとめ

- 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

ネットワーク構築に利用される冗長化の仕組み

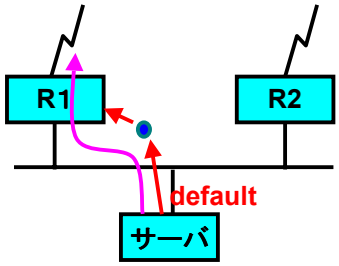
- STP(spanning tree protocol)
  - レイヤ2での冗長構成
  - 障害の発生から spanning tree変更までには10秒～60秒程度必要
- I/F downと static
  - I/Fの downを検出するとその i/flに向いているroutingが消えることを利用したbackup
  - Ethernet専用線等のdownしないI/Fでは利用できない。
- HSRP/VRRP
  - 一つの仮想的な MACアドレスを複数のルータで共有することで、サーバ等でダイナミックルーティングを利用せずに障害時の切り換えを行う

HSRP/VRRP-1



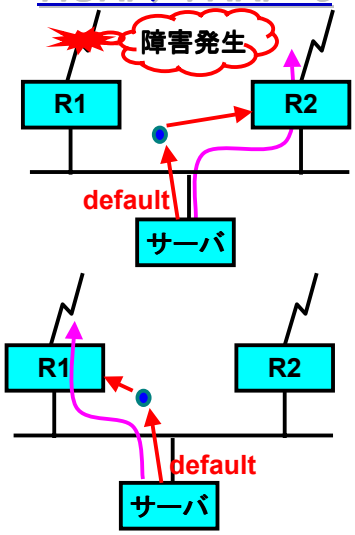
- 障害時には仮想MACアドレスがR1からR2に切り替わる
  - スイッチ等にルータを接続している場合には、ポート、MACアドレスの対応に食い違いが生じるため、さらに切り換えに時間を要する場合がある

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

HSRP/VRRP-3

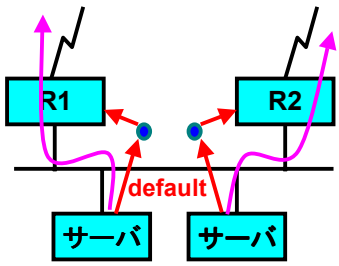


- HSRP+InterfaceTracking（障害発生時）
  - Interface Trackingにより切り替え
- HSRP+Interface Tracking（障害復旧時）
  - 復旧により切り戻しが発生

IIJ Internet Initiative Japan

MHSRP-1

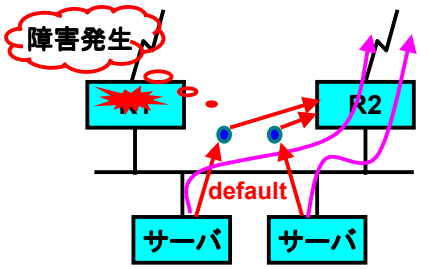
- マルチグループを用いて MHSRPを利用すれば、サーバ毎にトラフィックを分ける事ができる



- 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



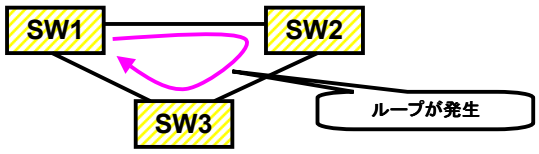
- MHSRP (障害発生時)

- 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

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

STPの動作-1

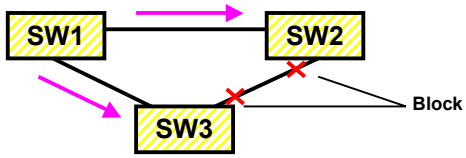


- STPを利用しない場合
  - STPを利用せずにスイッチの冗長化するとループが発生する
- ループの発生により様々な問題が発生
  - 各スイッチのアドレステーブルに混乱が生じる
  - 同じフレームが二重に届き、上位層の異常な動作につながる
  - Broadcast floodが発生する

IIJ Internet Initiative Japan

STPの動作-2

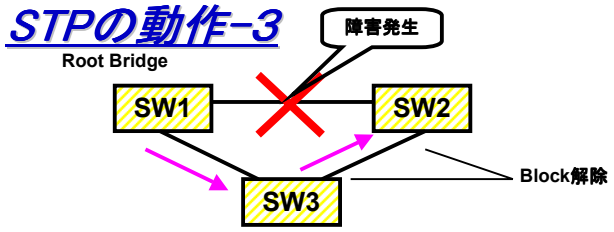
Root Bridge



- STPを利用する
  - STPの利用によりRoot Bridge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

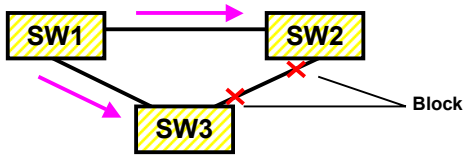
STPの動作-3



- 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

IIJ Internet Initiative Japan

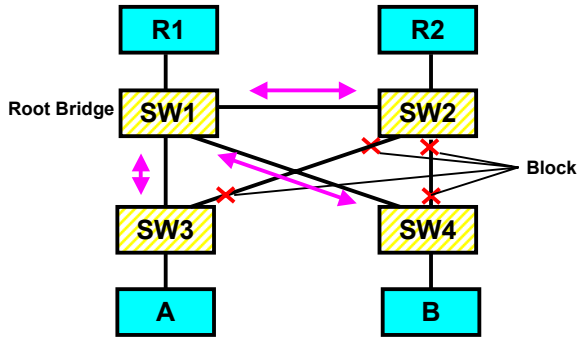
STPの動作-4



- 障害が復旧すると
  - 再びSTPによりRoot Bridge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

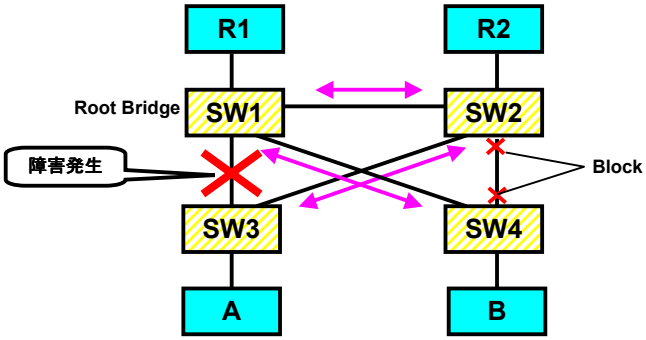
STP事例-1



- 全てのSWをRoot Bridge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

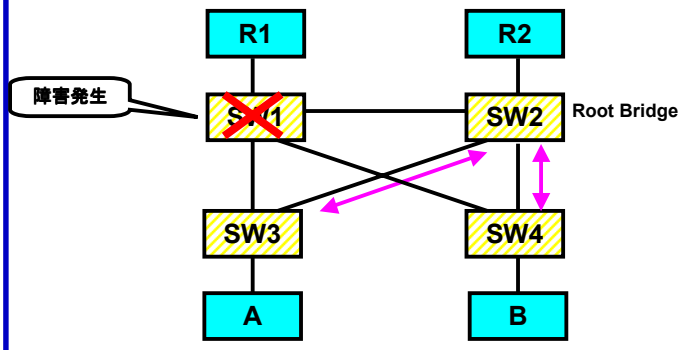
STP事例-2



- 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

IIJ Internet Initiative Japan

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 Bridgeに障害が発生した場合には切り替えに時間がかかる

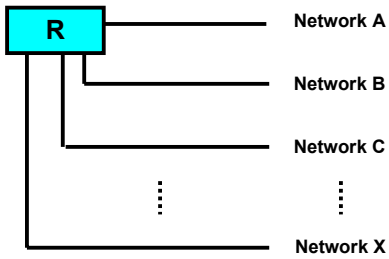
IIJ Internet Initiative Japan

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

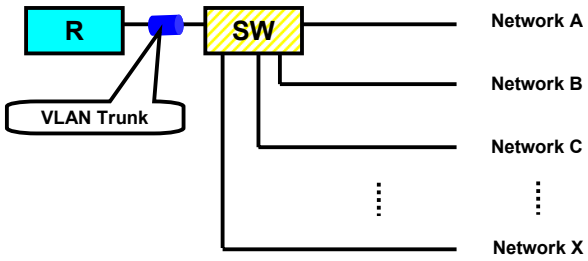
IIJ Internet Initiative Japan

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

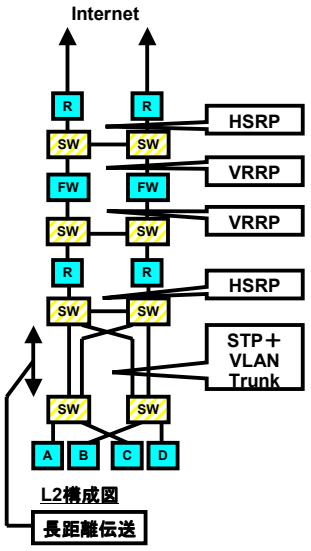


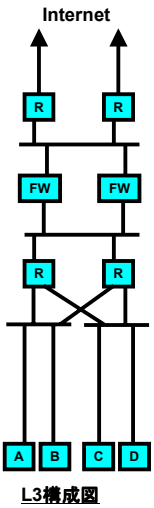
- VLAN Trunkを利用する
- VLAN Trunkによりルータのインターフェースが1つであっても仮想的に複数のインターフェースがあるように見せることができる
- ルータと比較して安価なスイッチのポートをルータのインターフェースとして見せることができる

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

STP+VLAN Trunk事例





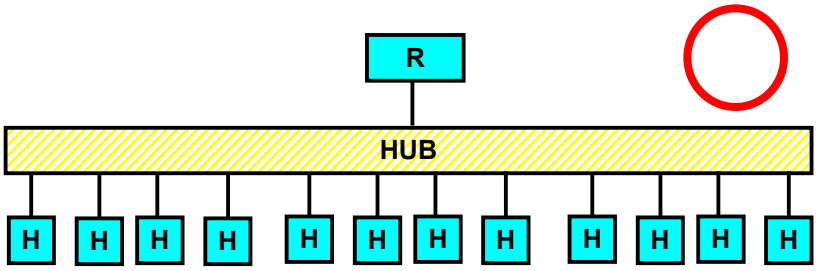
- 冗長化されたネットワーク
  - HSRP/VRRP
  - STP
- VLAN Trunkを利用
- 長距離伝送の冗長化とコスト削減を実現
- L3構成図とL2構成図の違いとポイント
  - STPはL2冗長化プロトコルであるため、L3構成図には表れない
  - VLAN Trunkに関してもL同様にL3構成図に表れない

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

IIJ Internet Initiative Japan

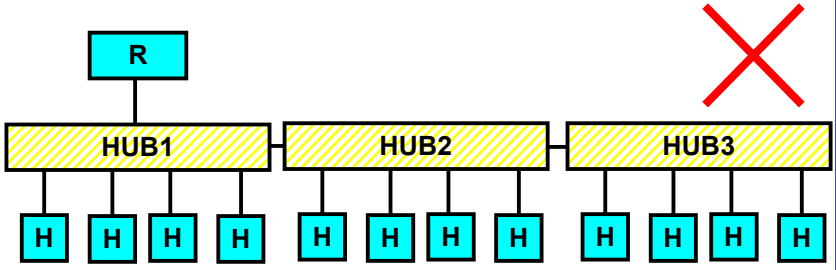
L2ネットワーク構成 カスケードなし



- 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

IIJ Internet Initiative Japan

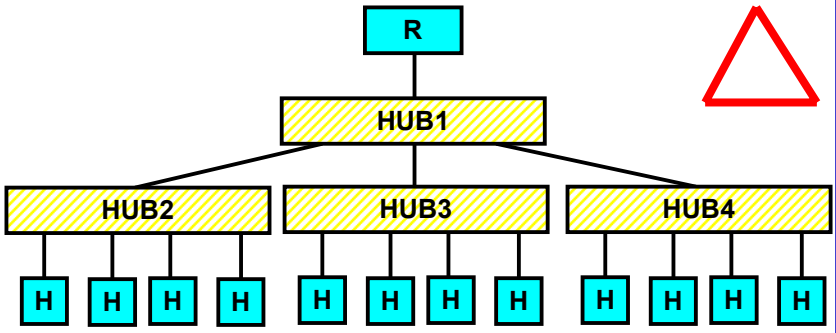
L2ネットワーク構成 横繋ぎカスケード



- 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 Japan

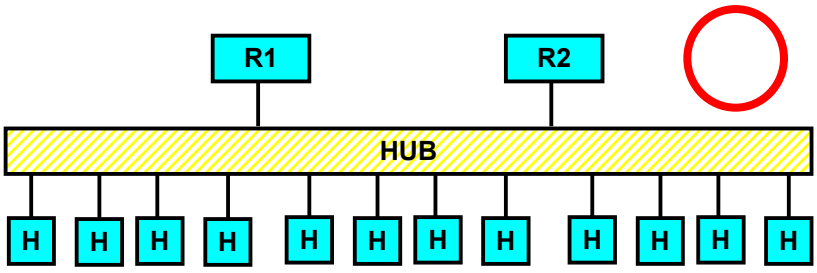
L2ネットワーク構成 ツリー型カスケード



- 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

IIJ Internet Initiative Japan

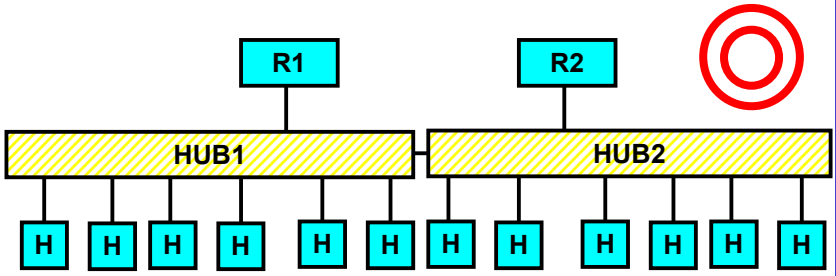
冗長化L2ネットワーク構成 カスケードなし



- 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

IIJ Internet Initiative Japan

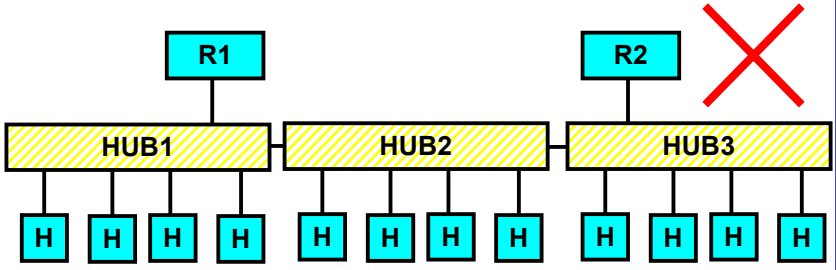
冗長化L2ネットワーク構成 2台カスケード



- 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

IIJ Internet Initiative Japan

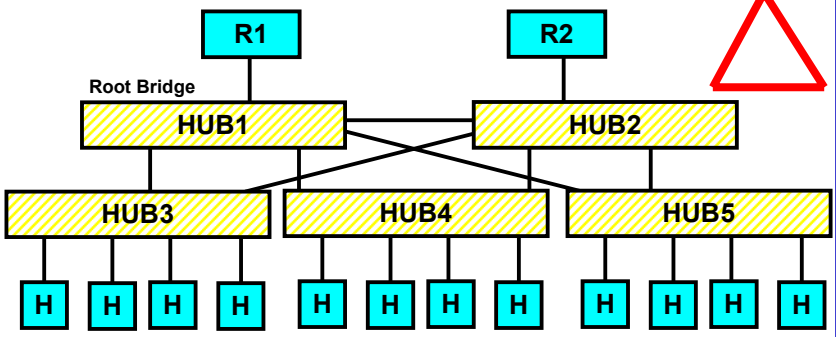
冗長化L2ネットワーク構成 3台以上カスケード



- 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 Japan

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 Japan

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

IIJ Internet Initiative Japan

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

IIJ Internet Initiative Japan

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

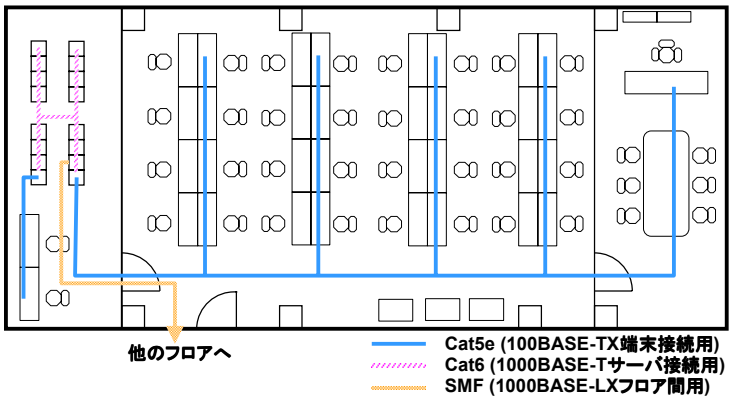
SMF: Single-mode Fiber

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 Japan

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

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

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

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
- 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

IIJ Internet Initiative Japan

