トラブルシューティングと耐障害性の高いネットワーク作り

CSIネットワークマスター虎の穴 第10回

株式会社まほろば工房

代表取締役 近藤邦昭

2007/09/11

Agenda

- ネットワーク障害の分類と概要
  障害対応プロセスモデルによる障害対応の実際
- 障害に強いネットワーク構築とそのポイント

スケジュール

- 時間を調整しながら進めますが、おおむね下記の予定で進めます。
- セミナー1部　　13:15-14:15
  - ネットワーク障害の分類と概要
  - 障害対応プロセスモデルによる障害対応の実際
- セミナー2部　　14:30-15:30
  - 障害に強いネットワーク構築とそのポイント
- セミナー3部　　15:40-16:40
  - IPアドレスの扱い
  - 経路制御と冗長化プロトコル
  - 運用関連ツールTIPS
- 質疑応答　16:45-17:00

ネットワーク障害の分類と概要
障害対応プロセスモデルによる障害対応の実際

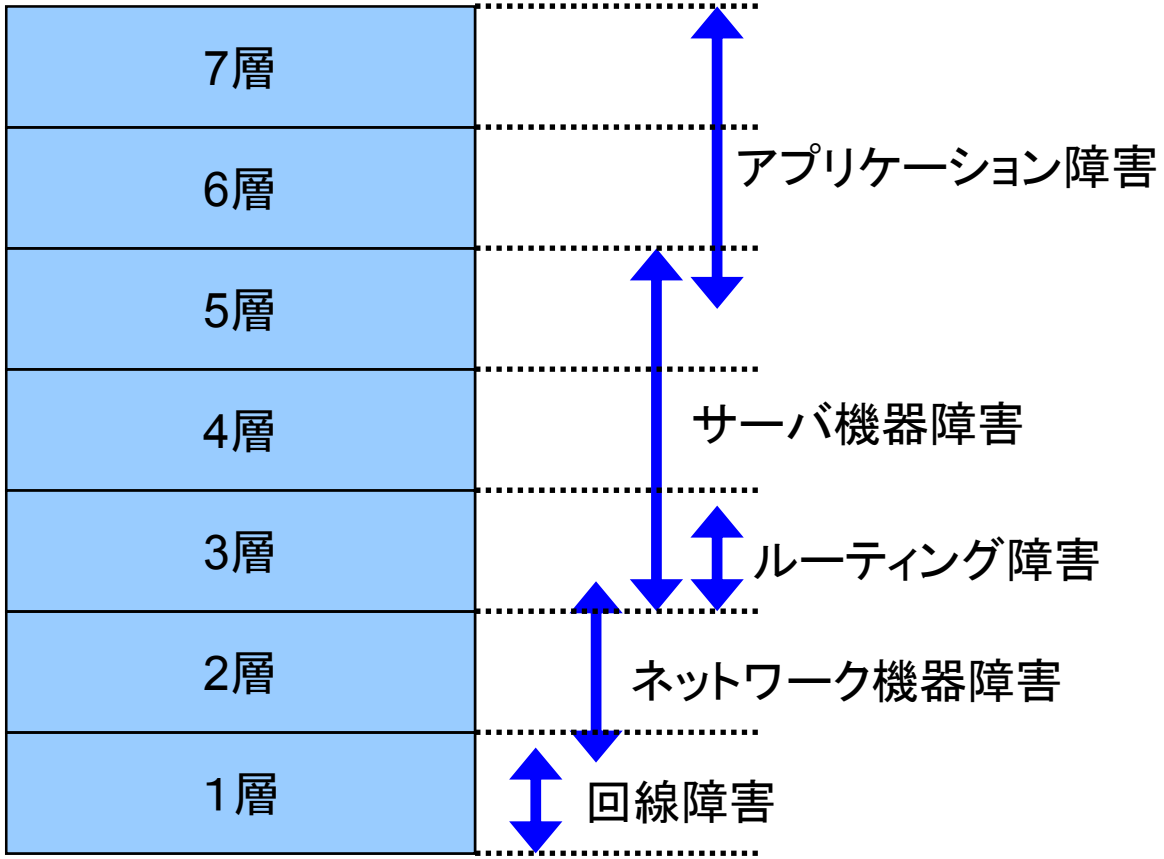
このセクションのAgenda

- 障害の種類の確認
- 個々の障害種別の大まかな概要
- 障害対応のプロセスモデル
- 障害の発見と障害の切り分け方法

障害の種類

- 回線障害　　レイヤ1
- ネットワーク機器障害　　レイヤ2
- ルーティング障害　　レイヤ3
- サーバ機器障害　　レイヤ3，４，５
- アプリケーション障害　　レイヤ5，６，７
- レイヤ8障害
  - 情報伝達ミスによる障害など

障害レイヤの概念図



障害の概要(回線障害)

- 専用線交換機の異常によるもの
- 回線提供業者の設定ミスによるもの
- 回線提供業者と回線利用者間の情報伝達ミスによるもの
- 回線利用者側の機器トラブルによるもの

⇨ 回線利用者がコントロールできる部分は非常に少ない

障害の概要（ネットワーク機器障害）

- スイッチ・ルータなどの故障による障害
- スイッチ・ルータなどの電源障害による障害
- 構内を結ぶ光ファイバやUTPケーブルの損傷による障害

⇨ ネットワークの構成によっては、ネットワーク全体の停止、または一部が分断される。

障害の概要（ルーティング障害）

- ルータソフトウエアのバグによる障害
- ルータの設定ミスによる障害
- 外部からの不正経路情報伝播による障害
- 外部からの不正アクセスによる障害

- パケットフォワーディングの全体、または一部に障害が発生する可能性がある。場合によっては、ルータが制御不能になる可能性も。

障害の概要(サーバ機器障害)

- ログファイルなどによるディスク容量あふれ
- ディスク故障によるサービス停止障害
- サーバの故障(ファン故障等)によるサービス停止障害
- サーバカーネル不具合(Panic)
- サーバ機器への不正アクセスによる障害

⇨ サーバ機器自体へのアクセスが不可能になるおそれがある。

障害の概要（アプリケーション障害）

- アプリケーションのバグによる障害
- アプリケーションの設定ミスによる障害
- アプリケーションの停止（バグ以外）による障害
- 外部からの不正アクセスによる障害

⇨ サーバには到達性があっても、目的のプロトコルによるアクセスが不能になるなどの障害が発生する

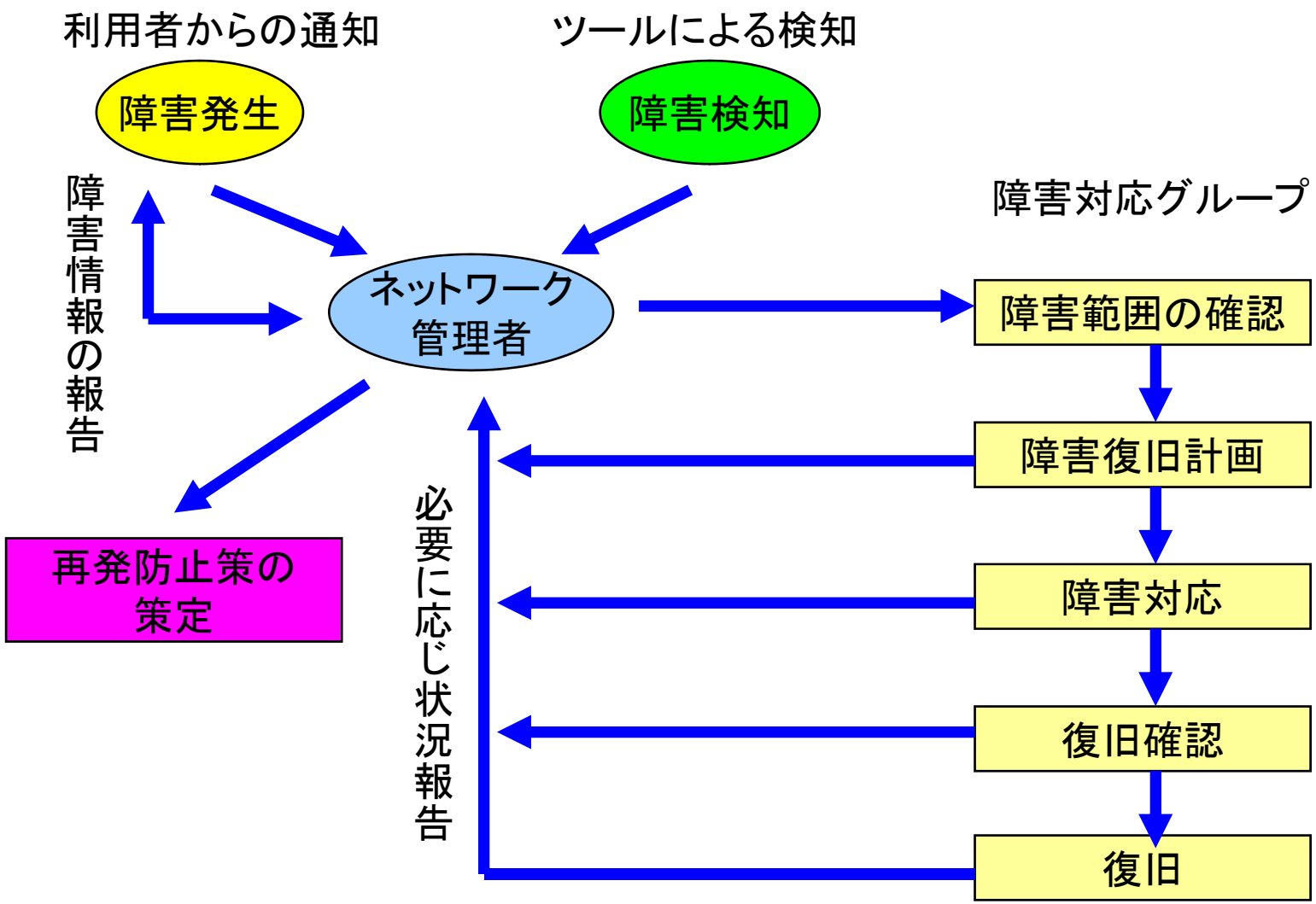
障害の概要のまとめ

- 障害の種類は様々
- また、障害によっても症状もまた様々。
- 障害に振る舞いは、ネットワーク階層に依存することが多いので、障害時の振る舞いとネットワーク階層をあわせて考えることで、障害ポイントを絞りやすくなる。

障害対応のプロセスモデル

- プロセスモデルとは
  - 障害発見から障害が完全に治るまでの流れ
  - また、障害が治った後の再発防止対策の策定なども含まれる
- プロセスのカテゴリ
  - 障害の発見とその確認
  - 障害の対応とその経過の報告
  - 復旧の報告と再発防止策の策定

障害対応プロセス概念図



障害の発見とその確認 -1

- 障害情報の取得
  - ネットワーク利用者から
    - ネットワーク監視ツールから
    - 取得する情報
      - 発生時刻
      - 影響範囲
      - ソースホストとディスティネーションホスト
      - 利用したプロトコル、また、障害と判断したプロトコル
      - 不具合が起きた時の詳しい状況
        - 他のIPアプリケーションが同一マシン上で動いていなかったか、など
    - 特に利用者からの情報の場合、技術的な通知はありえないので、どのようにこれらの情報を聞き出すかがポイントになる。

障害の発見とその確認 -2

- 影響範囲の確認
  - 再現性の確認
    - 通報のあった障害が、実際に起きているかの確認
      - 誤認による障害通報もありえる。
  - 障害時のネットワーク状況の確認
    - ネットワーク上で他のアプリケーションが動いていないか？
    - 他に関連する障害は起きていないか？
    - ネットワーク機器のログに関連するログは出ていないか？
  - 影響範囲はIPネットワークだけか？
    - ビルファシリティ、NTT回線などIPに関連する障害の影響をうけていないか？
  - 他のプロトコルも影響をうけているか？
    - 障害と見えているプロトコルだけに発生している障害か？

障害の発見とその確認 -3

- 障害レイヤの切り分け
  - 障害範囲の切り分けで得た情報をもとに、ネットワークレイヤのどの部分で障害が起きているかを推測
  - レイヤ3による場合分け
    - PingがOKであれば、レイヤ3以上が怪しい
    - そうでなければ、レイヤ3以下が怪しい
    - そうとも限らないことがあるので注意、pingはあくまで目安
    - telnet などで目的ホストの該当ポートにアクセスして、アプリケーションの稼動状況を確認る。

障害の対応とその経過の報告 -1

- 障害連絡
  - 実際に障害が発生していれば、その発生時刻、影響範囲などの詳細情報を利用者に連絡する。
  - 当然、障害が是正されていなければ復旧予定時刻も合わせてい調べる
  - 障害ではく、通常の動作であるならば、その旨を連絡する。

障害の対応とその経路の報告 -2

- 障害対応
  - ログなどにより電源障害のようなハードウエアトラブルと判定
    - 機器の交換によって復旧する可能性が高い
  - ログなどによる、特定のパケット特有の障害と判定
    - ファームウエアのアップグレート
    - バグの確認
    - ソフトウエアのバージョンアップ

障害の対応とその経路の報告 -3

- 障害対応
  - ネットワーク機器の追加、トラフィックの増加などが原因で物理的ネットワーク構成に起因する障害と判断
    - ネットワーク構成の変更
    - 該当回線の増速
    - 該当インタフェースの交換
      - これらは抜本的解決策だが、対応までの時間がかかる
    - 迂回経路への誘導
      - とりあえず、あいている回線へ迂回することで、短期的、暫定的に対応する方法。

障害の対応とその経路の報告 -4

- 障害復旧確認
  - 復旧対策後、少しの間は様子を見る
    - 熱問題によるトラブルなどは、再現までに時間がかかる。
  - 障害によって出力されたログはもうでていないか？
    - 同時に対応後に新しく出るようになったログはないか？
    - 障害対応したことで、新たな障害・トラブルを生むことは良くある。
  - 利用者に対する障害は依然発生しているかどうかの確認
    - 障害対応とは利用者が障害と感じなる時点で復旧といえる。

復旧の報告と再発防止対策の策定 -1

- 障害復旧報告
  - 障害のあった時間帯、箇所、機器名、障害時の細かい状態を記録
  - 障害が復旧したのであれば、どのような対応で復旧したのかを記録
  - 復旧していないのであれば、どこまで対応し、今後どのように対応しなくてはならないのかを記録

復旧の報告と再発防止対策の策定 -3

- 障害再発防止対策
  - 原因を明確にし、再発しないように対策を講じることが目的。
  - あくまで現実的に、かつ、具体的に書くことがポイントとなる。
    - 悪い例
      - 「～の様に善処する。」と策定する例があるが、これはいわゆる「がんばれ！」といっているに過ぎないので、具体性に欠ける。

障害の発見方法 -1

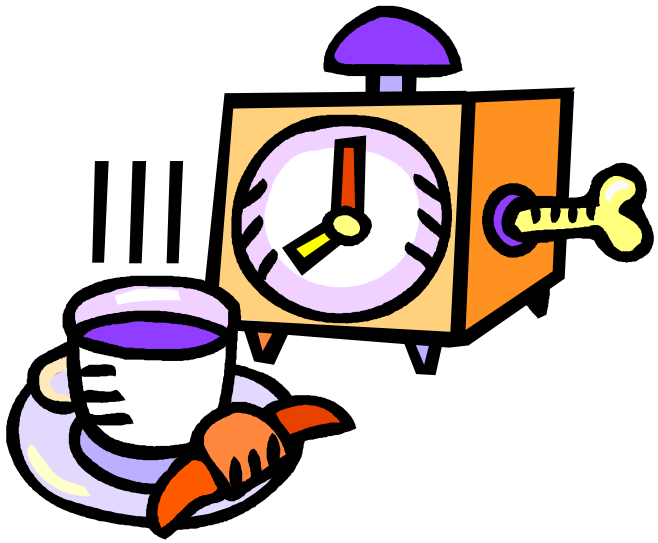
- ISPの場合
  - 管理ツールなどによる定常監視での障害検出
  - 顧客からの通信不具合の連絡
  - 他のISPからの通信不具合の連絡

障害の発見方法 -2

- 企業ネットワークの場合
  - 管理ツールなどによる定常監視での障害検出
  - ユーザ（利用者・社員）からの通信不具合の連絡
  - 通信相手の企業（VPN接続先等）のネットワーク管理者からの通信不具合の連絡

障害ポイントの切り分け

- 通信状態の確認
  - 障害通報者からの情報が非常に重要
  - 過去の障害履歴などから同様なものを検索
    - 同じ問題が多発するケースも少なくない
- 障害レイヤの特定
  - レイヤ3を境目に上下で対応部署が異なる場合が多い
    - IP網を管理している担当者と、サーバを管理している担当者、または、ビルインフラを管理しているところも違う。
    - これらはすべて、レイヤで分断することが可能。
- 障害箇所の特定
  - Ping, traceroute, telnetなど、ごく基本的なコマンドを利用して障害箇所を特定する
  - ネットワーク機器が残しているログから、障害発生機器を特定する



TAKE A BREAK

障害に強いネットワーク構築とそのポイント

障害に強いネットワーク構築とそのポイント

- 電源／ケーブリング
- 熱対策
- LAN
- WAN
- アドレッシング
- ルーティング
- ネットワーク障害監視

電源

- 電源容量の計算の仕方
- 電源の取り方の注意
- アースの必要性

電源

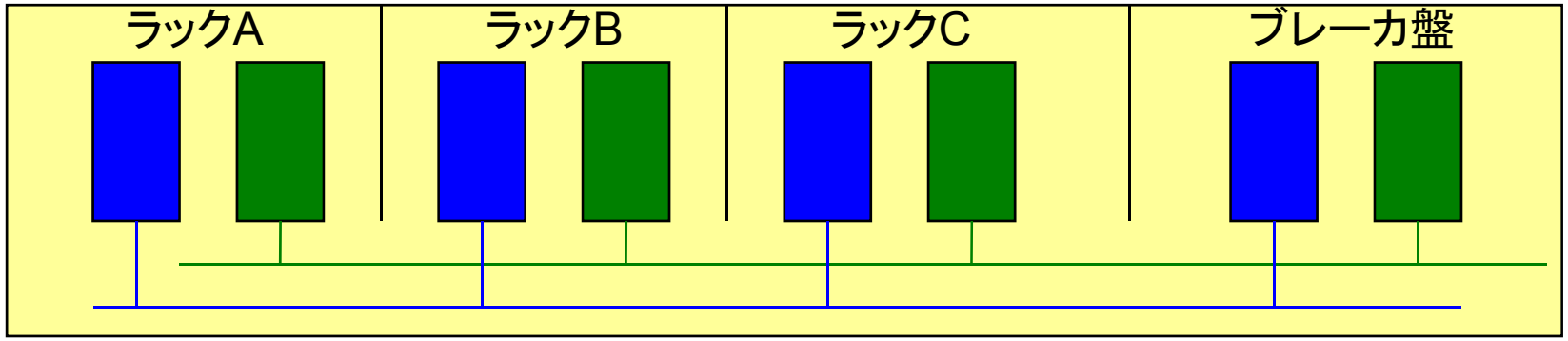
- 電源容量の計算の仕方
  - 電源容量の表示の仕方には、WとVAがある。
  - W=VAではない。（Wは力率を掛ける）
    - W = V × A × cosθ
      - θ = 30°～60°くらい：機器によって力率は異なる
      - θ = 0°（直流抵抗）　→ W=VA
      - θ = 30°のとき　　→ W=0.87VA
  - 機器によって表示が異なる場合がある。
    - UPSなどを使っている場合は、UPSの表記に合わせて計算をするのがわかりやすい。
    - W<VAなので、VAですべてを計算すると電力が足りなくなることは避けられる。

電源

- 電源容量の計算の仕方
  - 電源は、機器投入直後に急激に消費される。
    - モーターは回転を始めるときに最大電流が流れる。
      - ハードディスク、冷却用ファンなど
    - 設計は、起動時の電力で行う。
    - 通常時の電力で計算していると、全機器が同時に起動されるとオーバーフローする。
      - とはいえ、それでは現実的でない場合が多いので、電源を一斉に投入しないような仕組みを導入するなどして節約することも必要。
      - 機器の立ち上げは順次行っていく必要がある。
      - 特に停電時からの復旧は、トラブルを起こしやすい。
  - 電源ユニットを二つ以上持っている機器の場合
    - 通常時 ―ユニット1つあたりの機器の消費電力の1／2
    - 障害時 ―機器の消費電力のすべて（通常時の2倍）

電源

- 電源の取り方の注意
  - 電源ユニットを2つ以上持っている機器
    - それぞれのユニットごとにブレーカが違うコンセントからとる。
  - 同一機器でバックアップ関係のある機器（サーバなど）
    - それぞれの機器ごとにブレーカが違うコンセントから電源を取る
  - ラックに電源コンセントが2列ついている場合には、それぞれ違うブレーカからとる
    - 容量が1ラックで満たない場合には、複数ラックで共用する



電源

- アースの必要性
  - コンピュータやネットワーク機器は、スイッチング電源を使用しているため、筐体自体をアースに落とす必要がある。
    - アースを共通化しておかないと、個々の筐体ごとに電位が変わる
      - 最悪の場合、機器の破損につながる
    - アースのあるケーブルで、ケーブリングを行う機器同士のアースは共通にしておいたほうが好ましい。
      - シリアル、パラレル、CRTケーブルなど
      - UTPではできない・・・。
    - 2Pアース付のケーブルがついている機器は、アースなしに変換するアダプタ(通称:豚の鼻)を使わないようにする。
      - コンセントは、できるだけ2極アース付の抜け止めタイプ(ツイストロック)を使用する。
    - とはいえ、最近の聞きはアースを必要としないものも多い。機器の仕様にしたがって正しくつかうこと。

熱対策

- 最近のインターネットの普及で1ラックあたりの電力密度が高くなってきている。
  - 電力はそのまま熱になるので、電力密度が高いっていうことは、熱量も高いということ。
  - データセンターなどの冷却能力やエアフローをきちんと把握しなくては、ラック内の温度が急激にあがって、システムダウンにつながる
- 昨今では、熱対策は必須条件

熱量の計算

- データセンターの熱源
  - IT機器、USP、配電システム、空調ユニット、照明、人員
  - 全体としての熱量設計はデータセンターが行うので、設置側はラック、つまり、IT機器の部分の熱量と、ラックあたりの冷却能力について正確に把握する必要がある。
- 熱量の計算
  - 消費電力がそのまま熱に変換されると考えてよい。
    - ただし、端末への給電を行う装置の場合、消費電力の30%程度しか熱へと変換させず、その他は端末で消費されることがわかっている。
  - 100V 10A = 1000VA
  - VA * 0.67(力率) = W = 670W

熱量の単位

元になる単位	係数	変換後の単位
1時間あたりのカロリー	0.86	(W)ワット
1時間あたりのBTU	0.293	(W)ワット
(W)ワット	3.41	1時間あたりのBTU
トン	3520	(W)ワット
(W)ワット	0.000283	トン

※1:BTU : British Thermal Unit
※2:トン ：氷の冷却能力を示すもので、1870～1930年代の名残
※3:標準的単位：近年はワットを標準でつかうという動きが盛ん

ラックの熱管理

- 最高効率の場合は、気にする必要がない。
  - 多くのデータセンターでは、ラックに供給している電力分の冷却能力は想定しているので、20Aのラックであれば、100 * 20 * 0.67 = 1.3KW程度の冷却能力があると思ってよい。
  - しかし・・・・・。
- 外的要因による冷却能力の低下
  - 隣接するラックの温度による影響
  - ラック内のエアフローが悪いために起こる悪影響
  - ラック配置の不備による悪いエアフローによる影響
  - データセンター自体の冷却能力低下による悪影響
  - ラック内の機器の設置の仕方によるエアフローの低下
  - ラック内のケーブルがエアフローを邪魔することによる悪影響

  - 特に、前後左右のラックの排熱による影響は受けやすいので注意が必要

無理なものは無理。

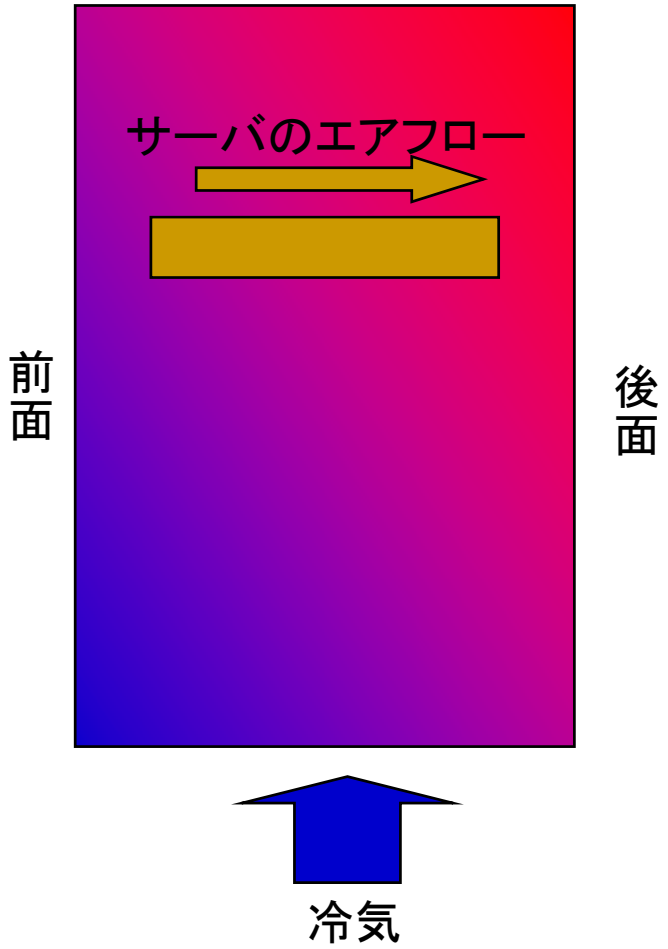
- ラックのスペースがいているからと言って・・・。
  - 電源容量が無いが、ラックスペースがあいているので、電源増やして機材を設置するとろくなことはありません。
    - 結局ラック単位の想定冷却能力のバランスが崩れるので、機械を設置できたからといって運用できるとは限りません。

設置限界

- 1ラックに果たしてどれくらい機材が設置できるか？
  - 定格電力は当てにならない。
    - 定格電力は安全のための電力で、実質の3割り増し程度になっているのが通常のようです。つまり、定格電力の3割減で突入電力を考えることができます。
    - 定常状態の電力はさらに低くなります。
      - 冷却能力的には、全体の70%程度で考えることができます。
  - 上記により設置限界が計算できる
    - 20Aのラックの場合、20*1.3 = 26A, 2600VA*0.67= 1740Wの冷却限界と考える。
    - 10Aの機器の場合、670Wと考えられるので、3台も設置できればいいところ。
    - もっと正しくは、電力計などを用いて逐次現状を把握することが望ましい。

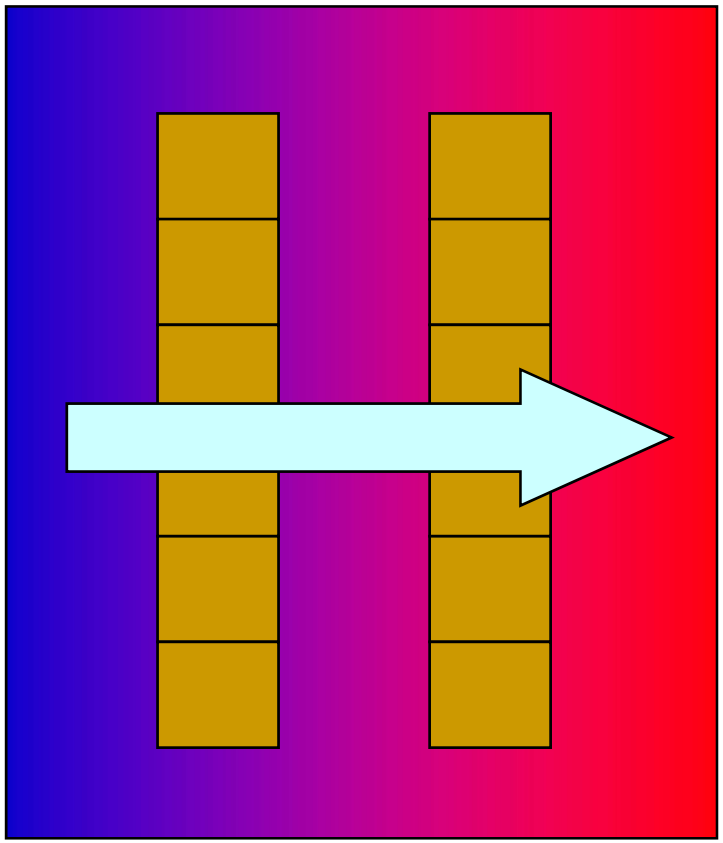
エアフローの考察 -1

- ラック内
  - 下から上に吹き上げるラック
    - 下は冷たいが、上部に行くほど温度が上がる。
    - サーバ機器のみの場合前面吸気、後面排気なので、前面に冷気を集中させる必要がある。
    - その他、巻き込みエアフローも考慮する。
      - 後面から排気された熱が、ラックの横を通過し前面冷気と混じることで、冷却能力を低下させる



エアフローの考察 -2

- データセンターのエアフローの設計
  - ホットアイルとコールドアイルに分けた設計をしているか？
  - していないと、前のラックの排気を後ろのラックが吸い込むため、ラックを通過するごとに空気は加熱される。



ラックの向きが同じだと、前列の排気が後列の吸気になる。

熱対策 – まとめ

- ラックに設置できる機材は、電力だけでなく、熱も考慮して設置しなくてはいけない。
- ラック内のエアフローを正しく把握しないと、熱暴走の危険が伴う。
- データセンターを選ぶ際には、データセンター全体として、熱対策をどのように考えているかを十分検討しなくてはならない。
- ラックを選ぶ際には、周りのラック、特に左右のラックで熱が過剰に排出されていないか確認したほうがよい。
  - 設置機器が少なくても、横面排気の危機が設置されている場合は、もろに自分のラックに排気される可能性がある。たとえ、間仕切りがされていても仕切り板が熱せられれば同じこと。

ケーブリング

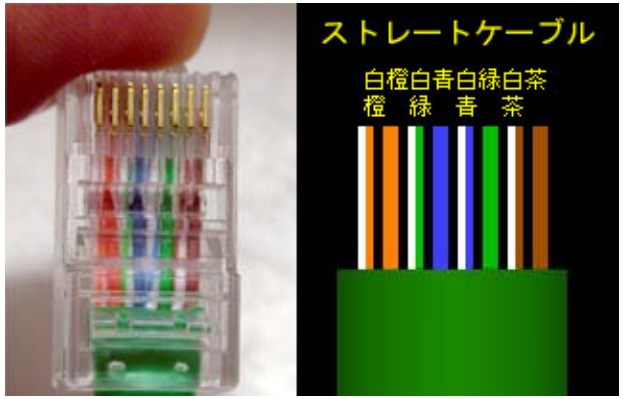
- ケーブルの種類
  - メタル（カッパ）ケーブル
    - ツイストペア
    - 同軸ケーブル
  - 光ファイバ
    - シングルモードファイバ
    - マルチモードファイバ
    - コネクタ形状様々
- ケーブリング時の注意事項

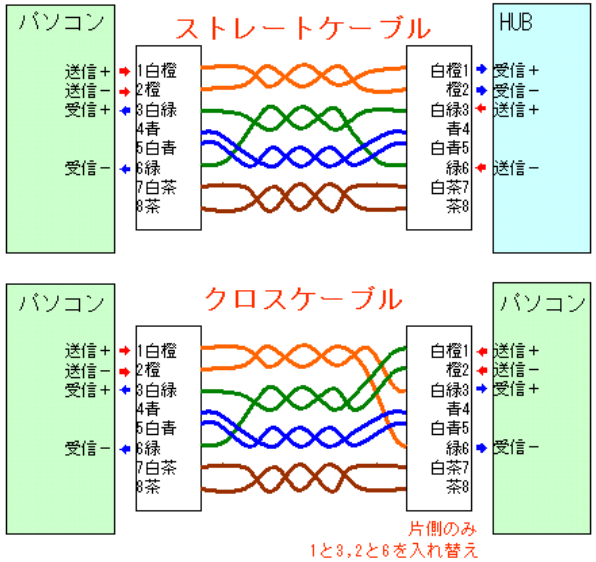
ケーブルの種類

- ツイストペアケーブル
  - より対線によりノイズの飛び込みを軽減している
  - クロストーク（漏話）とノイズ（雑音）に対する性能からクラス分けされている。
    - カテゴリ3　[CAT3]　(～10Mbps)
    - カテゴリ4　[CAT4]　(～20Mbps)
    - カテゴリ5　[CAT5]　(～100Mbps)
    - エンハンストカテゴリ5　[CAT5+]　(～1Gbps???)
    - カテゴリ6　[CAT6]　(～1Gbps)
  - カテゴリ5の規格では、コネクタにケーブルを差し込む時のより対部分のほぐす長さも決まっている。
    - 13mm以内
    - 電気的には相当厳しい規格である。

ケーブの種類

- ツイストペアケーブル
  - ツイストペア用コネクタには、主に次のものが用いられる。
    - RJ11　6線　　-電話用
    - RJ45　8線　　-LAN/ISDN用
    - RJ48　8線　　-ISDN新規格用
      - すでにISDNもほとんど使われないので、事実上RJ48は見られないが、RJ45の差込部分に凸部がある形状になっている。
    - ケーブル内の線は、色に応じてピン配置が決まっている。





ケーブルの種類

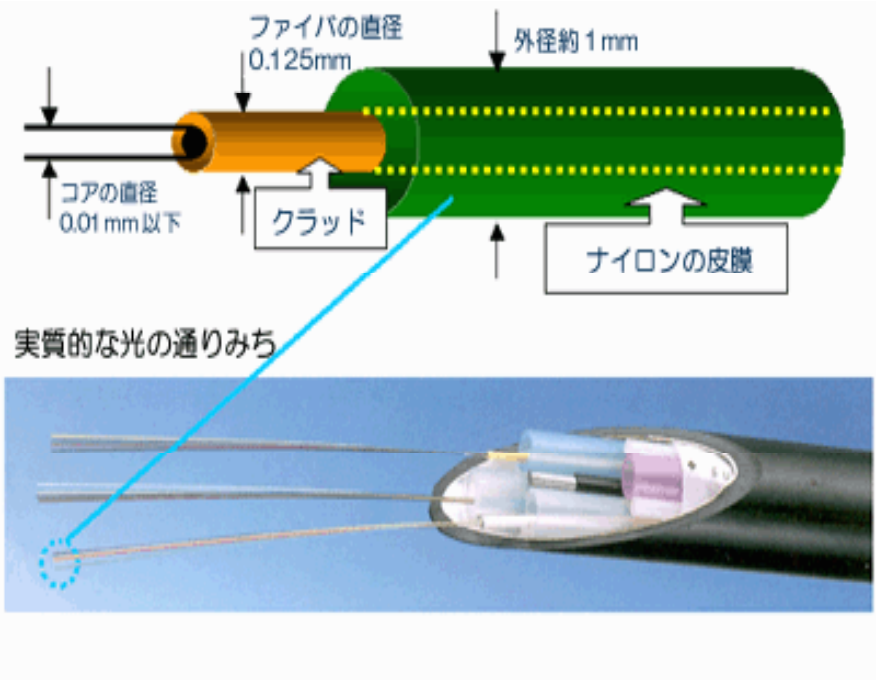
- ツイストペアケーブル
  - ツイストペアケーブルには、以下の2種類がある。
    - UTP(Un-shielded Twist Pair)ケーブル
      - より対線の外側の皮膜がそのまま（ビニールのみ）のもの
    - STP(Shielded Twist Pair)ケーブル
      - より対線と皮膜との間にシールド（同軸ケーブルのようにメッシュ上にあまれた伝導体の膜）がされているもの
  - 100Mbps以上のデータが流れるとツイストペアから雑音が出る。
    - 電子機器からの雑音の規制の厳しいドイツでは、STPしか使うことができなかった。
      - 今は大丈夫だと聞いている。
      - コネクターもシールドされていたらしい。
  - ケーブルには、単線ケーブルとより線ケーブルがある。
    - 工具で自作する場合には、単線ケーブルのほうが扱いが楽。
    - パッチケーブルに使うケーブルは、より線のほうがよい。
      - ケーブルがやわらかく、ねじっても癖がつきにくい。

ケーブルの種類

- 同軸ケーブル(ほとんど歴史のお話)
  - インピーダンスの違いで2種類ある。
    - インピーダンス50Ω
      - 主にLANケーブル(10Base-2)
        - 3D2V(JIS規格では、2文字目がDのものが50Ω)
        - RG-58A/U
    - インピーダンス75Ω
      - WANケーブル用(T3, DS3など)
        - 3C2V(二文字目Cが75Ωをあらわす)
        - RG-59A/U
  - コネクタは、BNCコネクタが利用される

ケーブルの種類

- 光ファイバ
  - 光ファイバの構造
    - 光ファイバは、コアとクラッドで構成されている
    - コアとクラッドは光の反射率が異なる素材で作られている
    - 入力された光はコアの中をクラッドとの境目で反射しながら進んでいく



ケーブルの種類

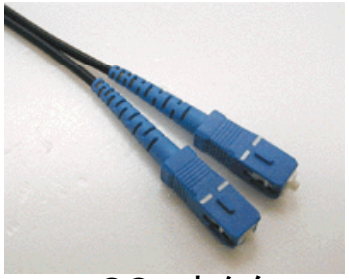
- 光ファイバ
  - ネットワークで使われる光ファイバ
    - クラッド径が125μm
    - コア径は、ケーブルの種類によって異なる
      - シングルモードファイバ
        - コア径が10μm以下（8.5μm, 9.5μm, 10μm)
        - 通過波長は、1310nm
        - コア径によって最適通過波長は異なり、減衰率が変わる。
      - マルチモードファイバー
        - コアの径が50μmのものと62.5μmのものがある。
    - そのほかクラッドとコアの屈折率の異なり具合から、2種類に分かれる
      - ステップインデックス（SI)
      - グレーデッドインデックス（GI)
        - 現在はほとんどがGIとなっている。違いについては割愛。

ケーブルの種類

- 光ファイバ
  - 光ファイバの特性を現すものとして、波長／伝送損失／伝送帯域などがある。
    - 最近は50μmのダブルウインドウが主流
      - 850nmと1300nmの両波長が使える
  - シングルモードファイバは、9μmとかが良く使われる。
    - JuniperのGigabitEtherのLHモジュールは、9μmを指定。
  - 使う機器の仕様にあわせて購入することが最善。
    - ただし、GigabitEtherだとか、インタフェースによって大体決まっているので、迷うことはあまりない。

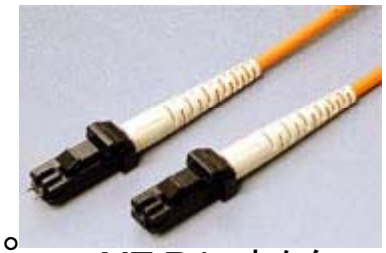
ケーブルの種類

- 光ファイバ
  - コネクタ
    - SC
      - プラスチックの角型のモールドタイプのもの
      - 二つが連結したSC-Dualというタイプもある
      - ATM/100Base-FX/1000Base-SX,LX,LHなど
    - ST
      - 一芯ごとにツイストロックするタイプのもの
      - ATM/100Base-FXなど
    - MT-RJ
      - 2芯1ペア構成のプラスチック製コネクタのもの
      - 主に1000Base-SX用のコネクタとして利用される。



SCコネクタ

STコネクタ



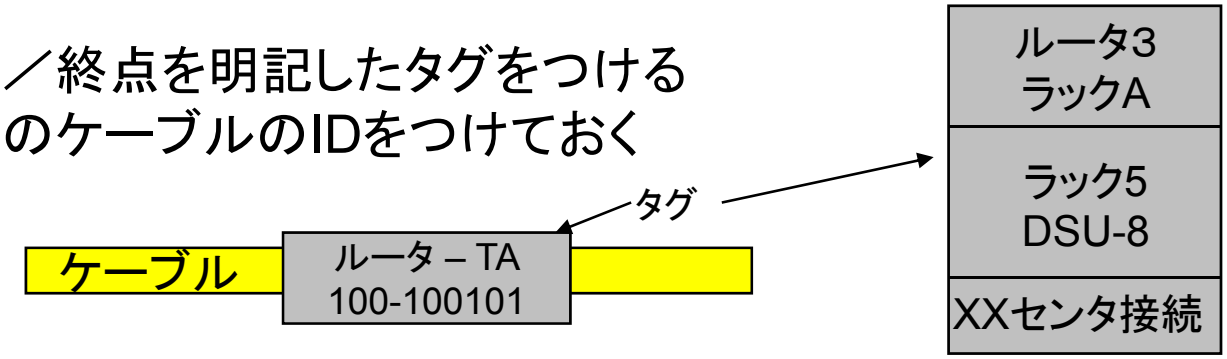
MT-RJコネクタ

ケーブルの種類

- 光ファイバ
  - コネクタ
    - GigabitEthernetなどでの光メディアの場合、多くはSCコネクタが用いられる。
    - ただし、スイッチなどポート密度に高密度が求められる場合は、MT-RJが多く使われる。
      - スイッチ製品など
    - このほかWDMなどのキャリア系光メディアはシングルモードファイバ＋SCコネクタという形で用いられるケースがおおい。
    - 最近では、FCコネクタやSTコネクタはほとんど見られない。

ケーブリング時の注意事項

- 全てのケーブル共通
  - 障害時に問題のあるネットワークのケーブルが特定できるように
    - 起点／終点を明記したタグをつける
    - 個別のケーブルのIDをつけておく



  - 巻かれているケーブルを延ばすときは、ねじりが出ないようにすること。
    - そのまま伸ばしたのでは、必ずねじりが発生する。
    - ケーブル自体を回転させながら延ばしてゆく。
      - ケーブルのねじりは、特にUTPの場合には、通信障害に通じる

ケーブリング時の注意事項

- ツイストペアケーブル
  - 電源ケーブルなどと並行してケーブルを敷設しない
    - 電源のラインからノイズが飛び込む
    - 特にフリーアクセスの下などでの工事時には注意
  - ケーブルを折り曲げると、伝送距離は短くなり、エラーレートは高くなる。
    - CAT5で100BASE-TXで、100mの規格いっぱいいっぱいで使うと最悪15%～20%程度のエラーが発生する。
    - 最低折り曲げ半径の10cm程度を保つ
    - ケーブルをねじっても同じくノイズが発生する。
      - UTPのより対線のよりが解けて電気的なバランスが崩れるため。

ケーブリング時の注意事項

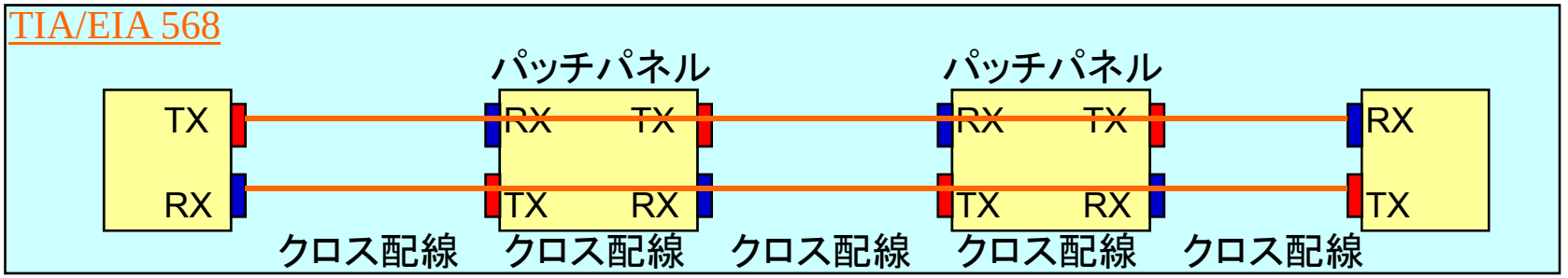
- 同軸ケーブル
  - 機器にあったインピーダンスのケーブルを用いる。
    - LAN用50Ω
    - WAN用75Ωなど
  - 起点から終点まで同じインーピーダンスのケーブルを使う。
    - インピーダンスの異なる同軸ケーブルを用いると、インピーダンスの変わるところで、信号の反射が起こり、波形が乱れてエラーとなる可能性がある。（特に、パッチ使用時に注意）
  - コネクタ類（プラグ、ジョイント、パッチ）にも、インピーダンスがあります。
    - パッチ等を設置する場合は十分に調べてから購入するようにしたい。

ケーブリング時の注意事項

- 光ファイバ
  - 最小曲げ半径は、10cm程度とること。(最小60mm)
    - 光ファイバは意外と折れやすい。
      - ファイバの外側の皮膜が保護しているに過ぎない。
    - ケーブルを小さい半径で曲げると内側と外側で光の反射率が変わってくる。
      - 光の全反射範囲が狭くなったりするため、減衰が大きくなる。
    - プラスチックファイバを使うと、曲げ半径は小さくできる。
      - 通常は石英ファイバなので、折れやすい。
  - マルチモードファイバのとき、ケーブルの混用に気をつける(特にジョイントして延長する場合)
    - 同じようなファイバでも、50μmと62.5μmがある。
    - 混用すると、反射がおきトラブルの元となる。

ケーブリング時の注意事項

- 光ファイバ
  - フリーアクセスなどの床下に光ファイバを入れる場合、事故をさけるために
    - ケプラーコートされた折れにくい光ファイバを使用する。
    - リボンケーブルなどの折れやすいケーブルを使う場合は、
      - 保護用パイプの中を通す。
      - スパイラルチューブを巻くようにする。
  - 光ケーブルは送信受信が双方の機器で入れ替わる必要があるので、全てをクロス配線するとよい



LAN

- ネットワークインタフェースの種類と特徴
  - イーサネット系
    - 10Base-2,5,FL,T
    - 100Base-TX,TX
    - 1000Base-SX,LX, T
  - xDDI系
    - FDDI
    - CDDI
  - その他
    - Token Ring
    - ATM
    - Fiber Channel

LAN

- よく見かけるトラブルの例
  - 10Base-5(Thick Ether)
    - LANを早くから導入したところに現在でも残っている場合がある。
    - トランシーバを同軸ケーブル(イエローケーブル)にタップして接続する
      - このため、経年変化により、タップの部分の接合が悪くなって障害が発生するケースがある。
    - 10Base-2(Thin Ether)
      - ある端末は接続できるが、別の端末からは接続できない。
        - 相次ぐ増設などにより、全長が200mを超えてしまうネットワークがある。
        - 経年変化によってコネクタ部分の接点不良もある。
        - 問題となっている端末と無関係なところに原因がある場合も多い。

LAN

- よく見かけるトラブルの例
  - 10Base-Xネットワークに共通
    - HUBが混入することでのトラブル。
      - 現在、多くの場合はSWでネットワークが構成されるが、古いHUBを使った場合、ブロードキャストドメインが広がるために、トラブルを起こす。
      - HUBは4段までしか接続してはいけない。

LAN

- よく見かけるトラブルの例
  - 100Base-TX
    - 10/100Mbpsの自動認識を信じてはいけない
      - Half/Full Duplexの自動選択も同じ
        - 条件がわかっている場合には、できるだけ固定の設定を行うこと
  - 1000Base-T(GigabitEthernet)
    - ケーブルの減衰が規定値以内にないため、1Gbpsで接続しない。
    - 片方が1Gで接続し、もう片方が100Mbpsにフォールバックしているケースがある。
      - オートネゴを信用したトラブルは多い。
    - Gigabit EtherにもHalf Duplexモードがあることを知らない人がいる。
      - オートネゴで、Half Duplexになっているケースも考慮する。

LAN

- GigabitEthernet
  - マルチモードファイバでもコア径が異なるものがある。
    - コア径によって伝送距離が異なるので注意が必要
  - パケットフレームのエンコーディング手法の差で接続できない場合があった。
    - 現在はほとんどない。

LAN

- よく見かけるトラブルの例
  - ARP忘れ問題
    - 同一アドレスで機器の交換をしたとき、ARPテーブルのキャッシュ情報を更新しないとうまく通信できない場合がある。
    - SWの場合、ポートに接続しているMACアドレスを学習しているので、注意が必要
  - ルータのインタフェースに設定しておいたほうがよい(かもしれない)項目
    - No ip redirect
    - No ip proxy-arp
    - No ip directed-broadcast

シェアードネットワークと スイッチドネットワーク

- ネットワークの規模によって違うが、時代の変遷によってLANの設計は変わってきている。
  - 第一期：～1992年
    - 10Base-5/2がバックボーンのネットワーク
      - トランシーバーからAUIケーブルで各機器に接続
      - 使ってもブリッジで、ルータはほとんど用いられなかった。
  - 第二期：1992～1993年
    - 10Base-Tの登場
      - フロア内で端末の接続をHUBとツイストペアケーブルで行う。
    - フロアが変われば、ブリッジやルータで接続する。
      - ブリッジやルータが非常に高価で多くのポートを持ったものは準備できなかった。(Cisco CGS/AGS)

シェアードネットワークと スイッチドネットワーク

- 第三期：1993年～1995年
  - ルータのポート単価が安くなってきた時期
    - 各フロアにルータを置き、同一フロア内でも、部課ごとに1つづつポートを分けてセグメントを設ける。
    - バックボーンが10Mbpsで足りない場合は、FDDIで各ルータ間を接続
    - ルータがそれなりに使われるようになった時期
      - Cisco 25xx / Cisco 4000 / Cisco 7000

シェアードネットワークと スイッチドネットワーク

- 第四期：1995年～1997年
  - 100Base-TXとスイッチの登場
    - 10M-HUBで足りなくなってきたポートに対して、10M-Switchに帰る事で、トポロジーをそのままで、高速化対応ができた。
      - リピータの4段制限の問題も解決
    - LAN間接続は、100Mbpsネットワーク
      - 当時は、CDDI vs 100Base-TX vs 100M VG-AnyLANが競っていた。
      - ルータ間の接続はFDDIが主流
      - 一部ルータレスで、スイッチとHUBだけで構成するネットワークも出てきた

シェアードネットワークと スイッチドネットワーク

- 第五期：1997年～2000年
  - スイッチ全盛
    - バックボーンは100Baseを使って高速なネットワークを組む
    - エッジは、10MSwitchか、10/100M自動認識のSwitch
    - ルータに置き換わって、Layer-3 Switchを使いながら、論理的なネットワークをVLANなどの技術を使って重ねていく
    - 100Mbpsで足りなくなった場合は、100Mbpsを束ねて使う技術（EtherChannel)や、GigabitEthernetを利用

シェアードネットワークと スイッチドネットワーク

- 第六期：2000年～
  - GigabitEthernet全盛
    - GigabitEthernetのSwitchの価格がどんどんやすくなり、100Base-TのSwitchに迫る。
    - LAN用メディアとして光ファイバが多く使われるようになる。
    - LANメディアをWAN用インタフェースとして使えるようになる。
      - 広域Ethernetサービスの出現
    - バックボーンでは1000Base-SXを主に使い足り名場合は、10GEtherを使い始める。
      - それでも足りない場合はLinkAggregationを使う。

- このように、かつてのSharedネットワークは、すでになくなりつつあり、全てがSwitchedネットワークに移行か完了しつつある。また、WAN/LANも区別のないネットワーク構成が可能になってきている。

WAN

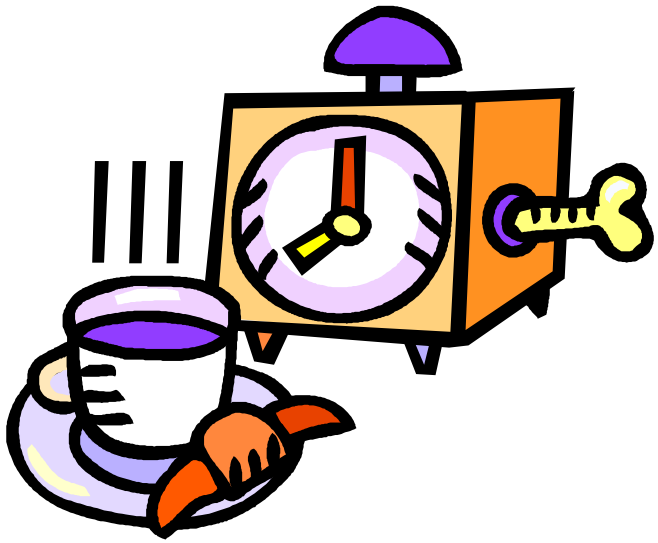
- NTTがサービスしている回線
  - 専用線
    - HSD(ハイスーパーディジタル）専用線
    - DA（ディジタルアクセス）専用線
    - DR（ディジタルリーチ)専用線
    - ATMメガリンク
    - 音声帯域専用線（3.4KHz）
  - 準専用線
    - スーパーリレーFR
    - スーパーリレーCR
  - ISDN
    - INS-64
    - INS-1500

WAN

- それ以外の回線
  - 構内自営線
    - 構内モデムを用いた回線
      - HDSLを用いて4線（2対）ケーブルで、最高2Mbps程度が出る。
      - 距離に応じて速度は反比例する。
    - 最近では、光ファイバを用いる。
  - 衛星回線
  - CATV
  - SONET
    - NTT以外のキャリアがUIとして出してくる場合が多い。
    - NTTも出している。
    - NTT規格よりもこちらのほうが標準的

WAN

- 回線の種類
  - 速度によって呼び方が異なる
    - T1　　1.5Mbps
    - T3　　45Mbps
    - OC-3　　155Mbps
    - OC-12　　622Mbps
    - OC-48　　2.4Gbps
    - OC-96　　4.9Gbps
    - OC-192　　10Gbps
    - OC-384　　20Gbps
    - OC-768　　40Gbps
    - OC-3072　　160Gbps



TAKE A BREAK

IPアドレスの扱い
経路制御と冗長化プロトコル
運用関連ツールTIPS

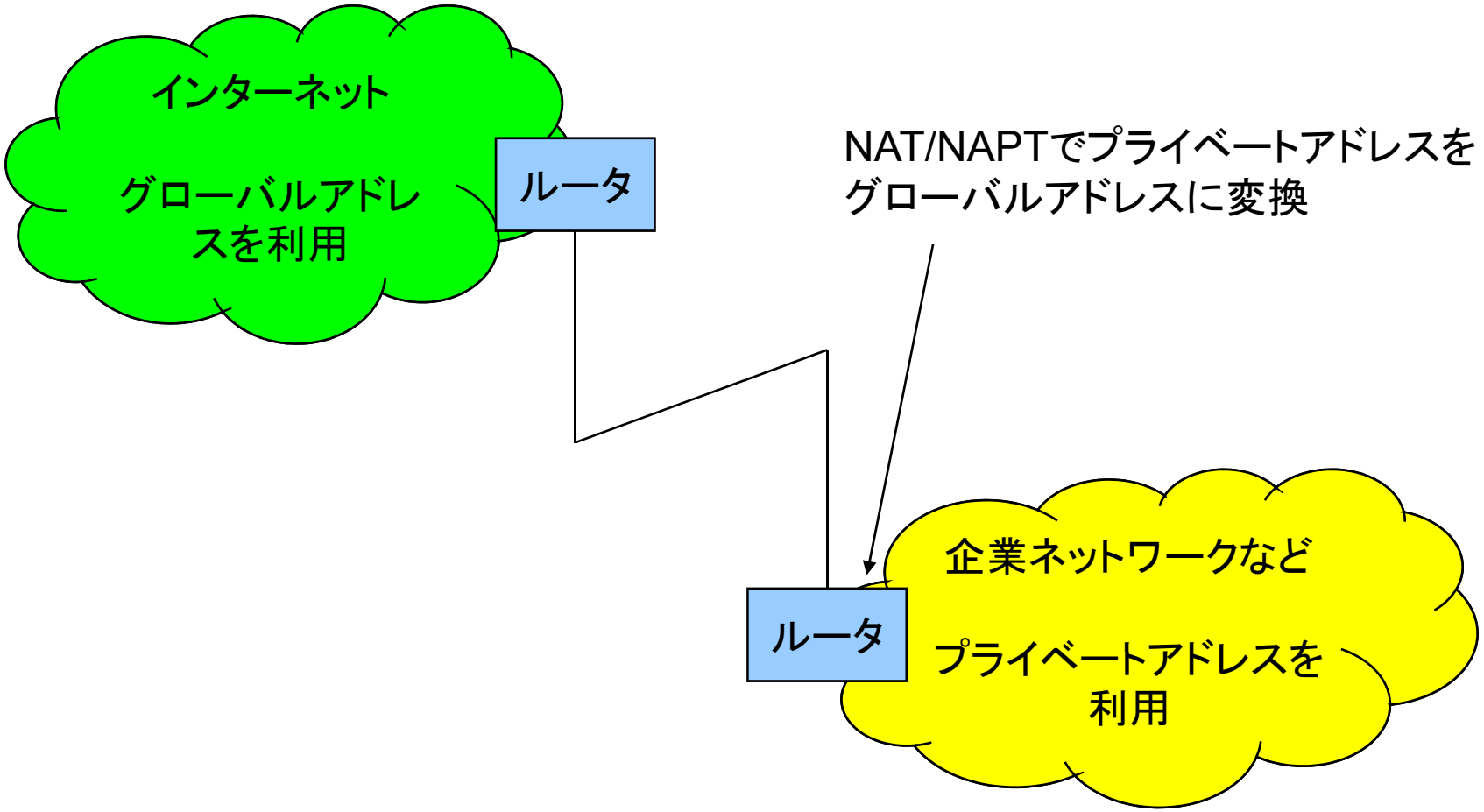
アドレッシング

- グローバルアドレスとプライベートアドレス
- アドレス変換の仕組み
  - NAT/NAPT
- 最適なアドレス採番とは
  - 障害を発見しやすく、メンテナンスをしやすくするアドレス採番方法

グローバルアドレスと プライベートアドレス

- グローバルアドレスとは
  - 一般にインターネットで使われるアドレス
  - 基本的に世界中で一意に決定できる番号
- プライベートアドレスとは
  - イントラネットなどの閉ざされたネットワーク空間で利用されるアドレス
  - グローバルインターネットには流出してはいけないアドレス

グローバルアドレスと プライベートアドレスの関係



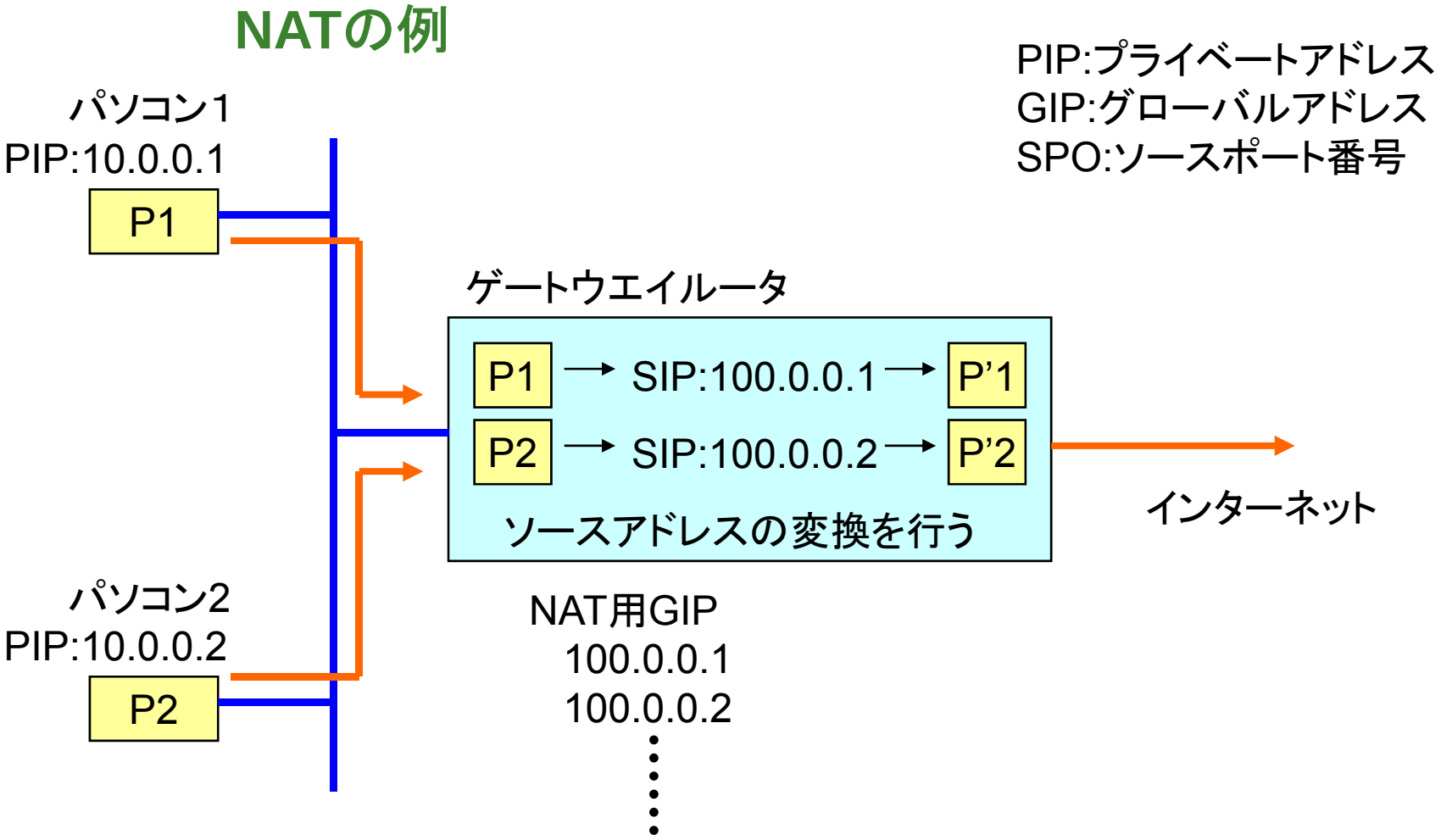
アドレス変換の仕組み

- NAT/NAPT(Masquerade)
  - 少ないグローバルアドレスを効率よく利用する仕組み
  - 1つ以上のグローバルアドレスをそれ以上のプライベートアドレスが振られた端末で共有する仕組み
  - 参考
    - IPv4アドレスは2009年くらいに枯渇するというレポートもある。
      - http://www.nic.ad.jp/ja/research/ipv4exhaustion/

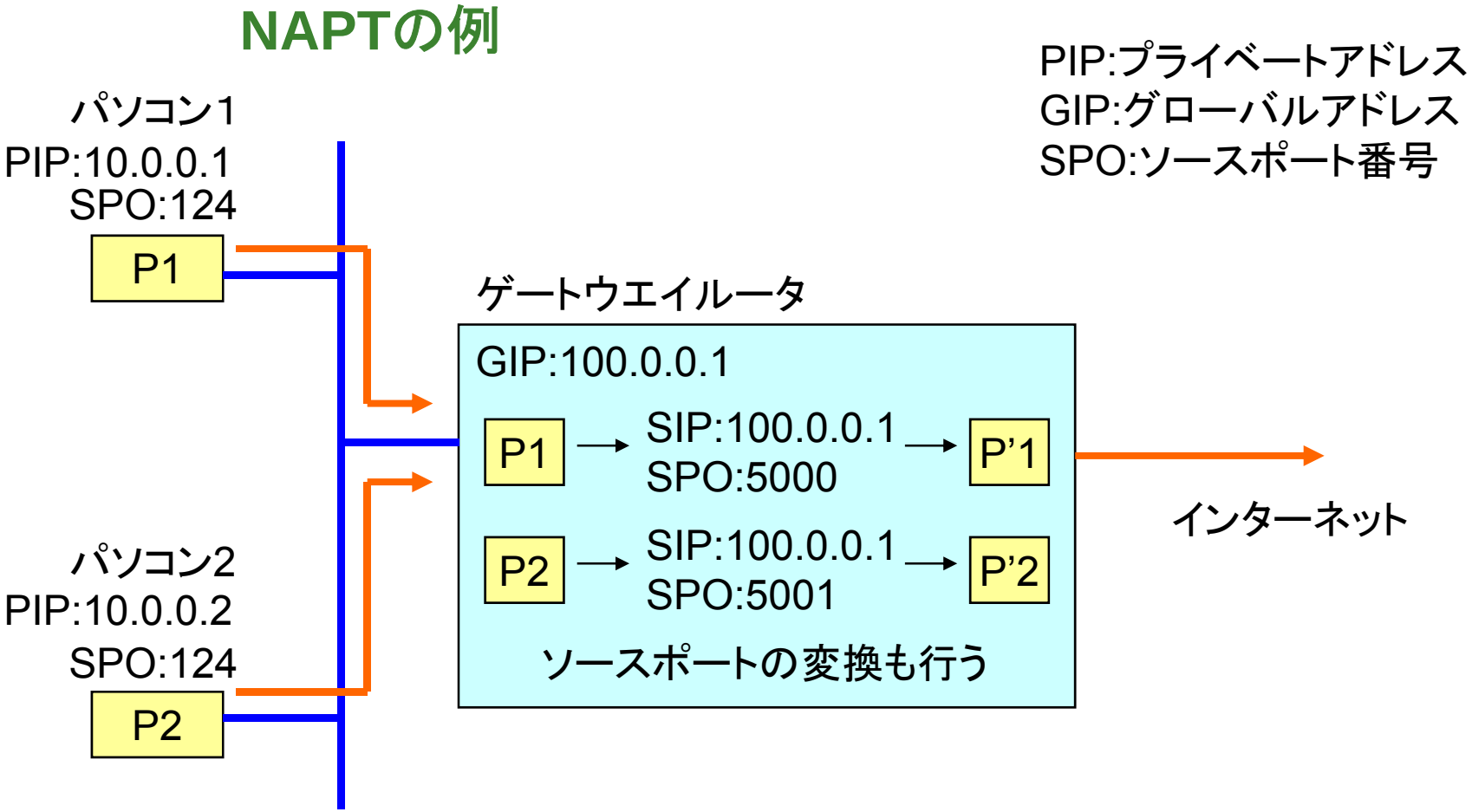
アドレス変換の仕組み

- NATとNAPTの違い
  - NATはソースポートを変えずにグローバルインターネットにパケットを送り出す。
  - NATは1つのグローバルアドレスに1つのプライベートアドレスが割り当てられる。
  - NAPTは、ソースポートを適当に変換する。このため、複数台の端末が1つのグローバルアドレスを利用して、インターネットにアクセスすることが可能

アドレス変換の仕組み



アドレス変換の仕組み



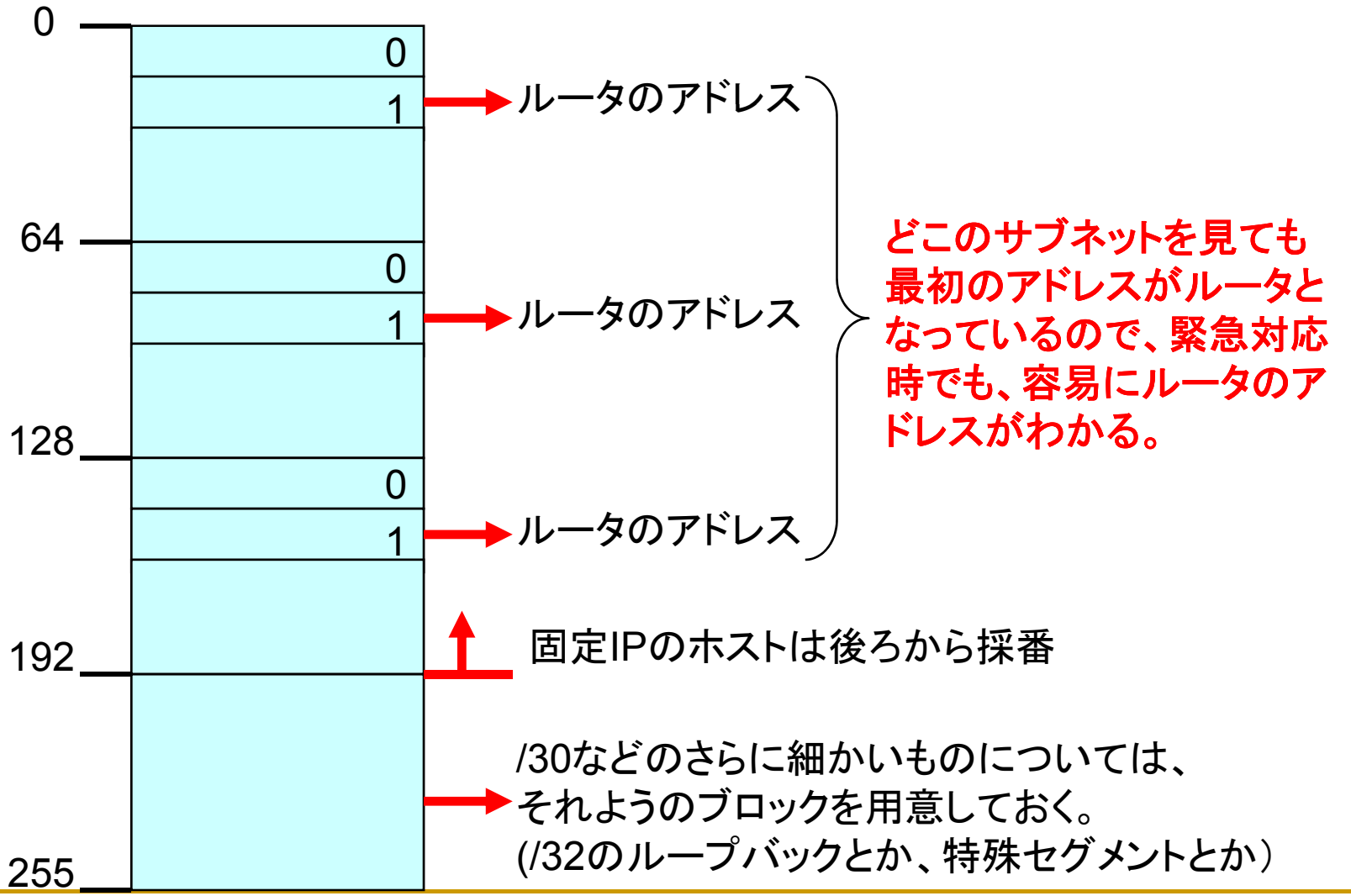
最適なアドレス採番とは

- 障害が起きたとき、その箇所が容易に特定可能な採番方法をとる
  - アドレスブロックでエリアを特定できるなど
  - グローバルアドレスの場合、余裕を持った採番が許されないので、実施は難しいが、台帳管理などでカバーしておく必要がある
- 採番されているアドレスがわからなくてもルータなどのアドレスが容易に推測可能であること
  - ルータや重要なサーバは、セグメントの先頭に置くなど

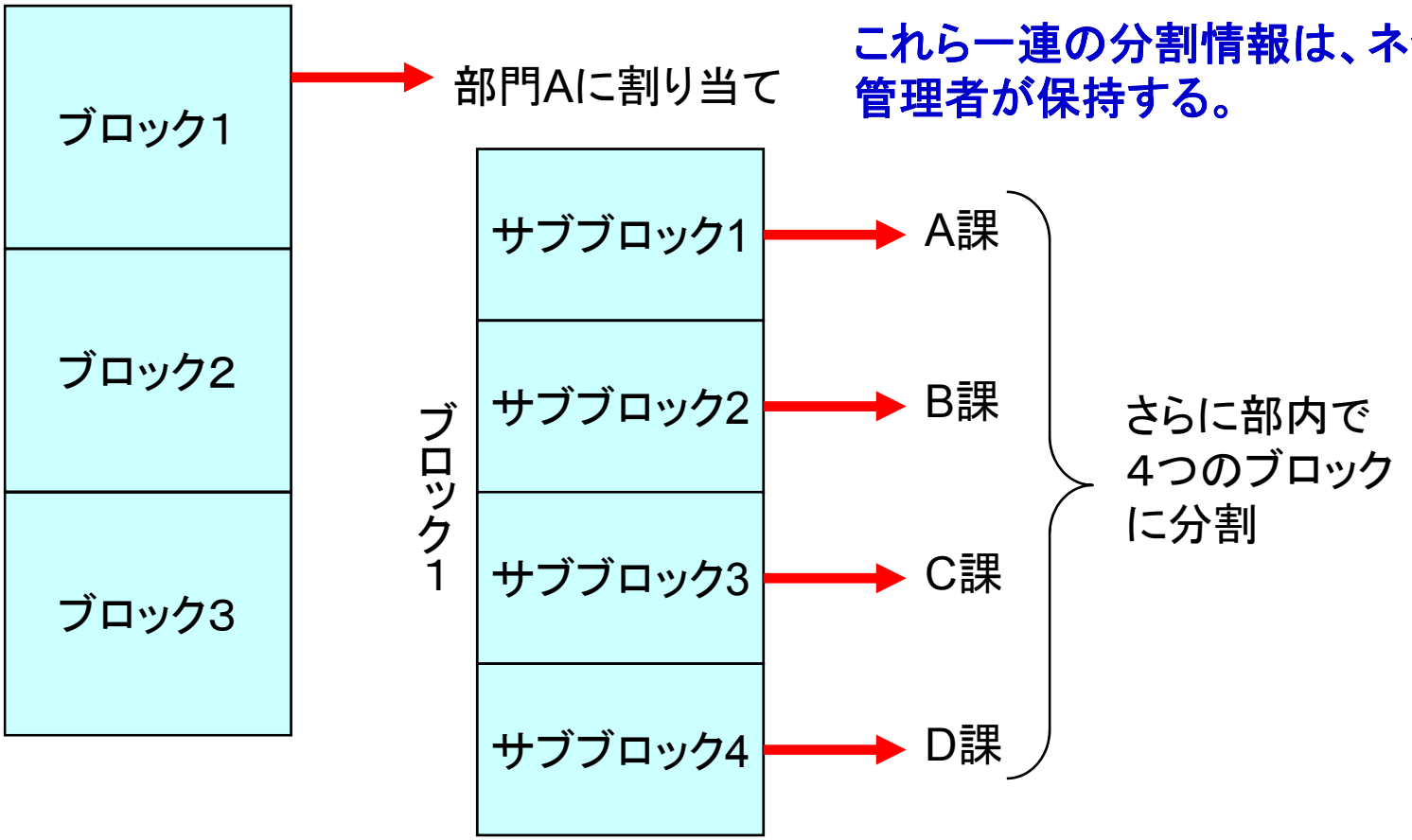
最適なアドレス採番の一例

- 10.0.0.0/24のネットワークなら
  - ルータは10.0.0.1
  - 固定IPアドレスのホストは、10.0.0.254から順に割り当てる
    - など
- /24が割り当てられたら概念的に/26に分割し、それぞれを部門別に分け、分けられたアドレスを部門内でサブネットに分割してりようする
  - など

最適なアドレス採番の一例



最適なアドレス採番の一例



障害があってもアドレスから細かい物理的位置が特定可能となる

ルーティング（経路制御）

- ルーティングプロトコル
  - RIP
    - Version 1,2,ngがある
    - 1はVLSMに対応していない
    - 2はVLSMに対応しているが、Supernetには対応していない。
    - RIPngはIPv6対応
    - 小規模なネットワークに今でもよく使われる
  - OSPF
    - Version2が標準。Version3でIPv6対応
    - VLSM/Supernetに対応し、中大規模なネットワークでよく使われる。
  - BGP
    - Version4, IPv6はMulti-Protocol Extensionで対応
    - プロバイダ間接続などで用いられる
    - インターネットにおけるISP間ルーティングプロトコルとして標準的に使われている。

RIPv2

- RIPv1プロトコルをそのままVLSMに対応させたもの
  - 実装は簡単で、安い機器にも実装されやすい
- でも…
  - 大規模ネットワークでスケールする技術ではない。
    - デフォルトでは、30秒に1回自分の持っている全てのルーティング情報を隣接するルータに配信する。（フラッディング）
    - ネットワークルーティングテーブルが大きくなると、フラッディングの負荷がばかにならない。
  - 障害時の即応性が低い
    - ネットワークダウンしても、デフォルトでは180秒たたないとルーティングテーブルから経路情報を削除しない。

OSPF

- OSPF
  - OSPFは、ある程度大規模なネットワークにも対応可能なルーティングプロトコルである。
    - ルーティングアップデートが起こらないと、原則経路情報を配らない。
      - 通常時は、10秒程度に1度のHelloパケットだけで、隣接の生存確認をする。
    - 40秒間Helloパケットが到着しなければ、その隣接ルータはダウンしたとみなされ、そのルータ配下の経路情報は削除され、その旨を他の隣接ルータに伝達する。
  - 設計上の注意点
    - エリア0(バックボーンエリア)を中心に、各エリアが接続されているというトポロジーで構成する必要がある。
    - LAN内のOSPF情報のやり取りはマルチキャストを使用する。
      - マルチキャストパケットをフィルタしていると、OSPFが正しく機能しない。

OSPF

- DR/BDR
  - OSPFでは、各セグメントごとにまずDRルータとBDRルータの選出を行う
  - DRルータやBDRルータは、自分が構築したルーティングデータベースを他のルータに配る。
  - DR/BDRになれるルータは限定しておいたほうがよい
    - Ospf priority 0 に設定すると、DR/BDRにならない。
    - DR/BDRなルータは以外と負荷がかかる。
  - 経路のベストパスの計算には、多くのCPUリソースを要求するので、速いCPUを持つルータがDR/BDRになったほうがよい。
    - Ospf priorityの数字をあげる。（デフォルトは1）(cisco)

OSPF

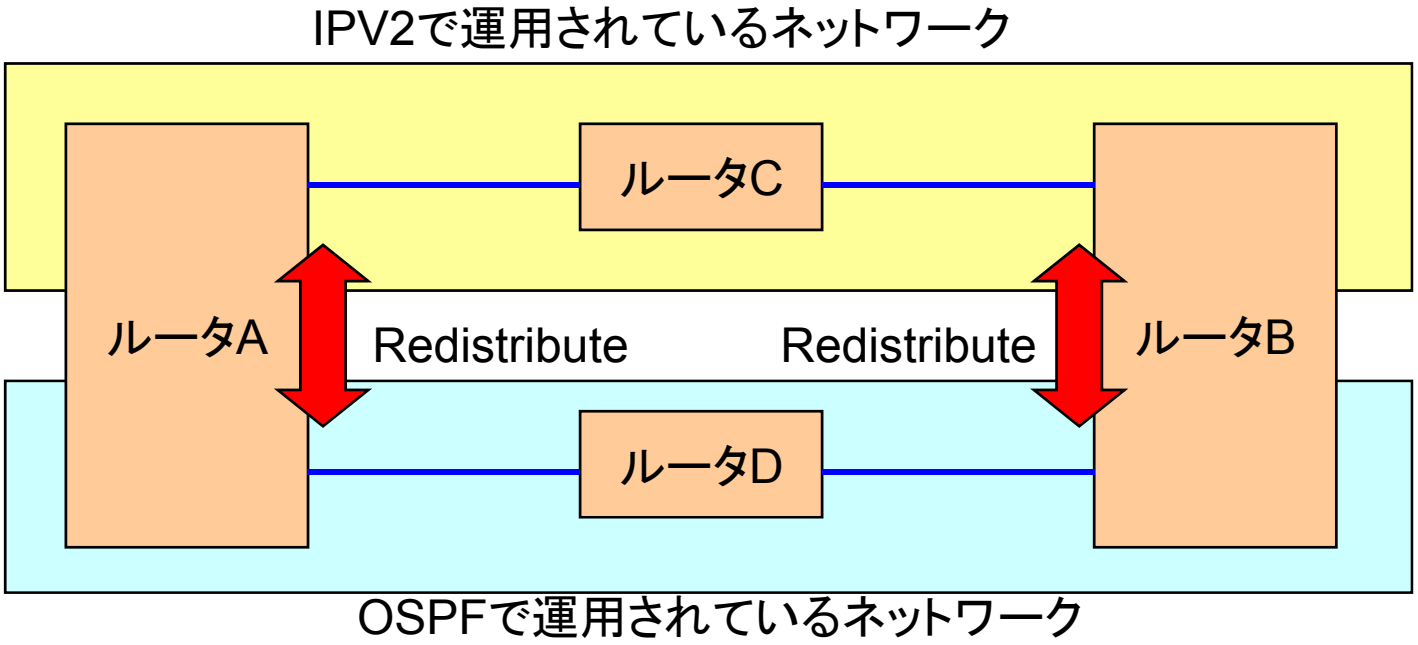
- OSPFは、複数のプロセスで独立したOSPFプロセスを複数同時に動かすことが可能（Ciscoなど）な機種tがある。
  - 経路が混じってほしくないネットワークで限定した経路だけを相互にやり取りしたい場合などに有効
    - Cisco IOSでは、機種によって起動できるプロセス数に制限がある。
- OSPFはトリガーがナイトupdateしないルーティングプロトコル
    - だからスケーラビリティが高いともいう。
  - 複数のOSPFプロセスを起動している場合、clear ip route をかけると、経路情報が他に流されなくなることがある・・・。
    - BUG?
    - 相当古い情報なので、今はなんとも無いかもしれない・・・。
    - 小さいネットワークの場合はRIP2がとり回しが楽です。

OSPF

- Redistributeの方法
  - Redistributeを行うときは、subnetsをつけよう
    - ルータに勝手に、ルーティング情報をまとめられては困る。
    - Subnetsはサブネット情報をそのまま扱う。
  - OSPFプロセス間でredistributeを行う時には、tagをつけておくと、show ip ospf databaseで、redistributeされた経路がわかるので、トラブルシュートには便利。
  - Redistributeには、
    - Connected
      - ルータに直接つながっているサブネットをOSPF経路情報に含める
    - Static
      - スタティックルートに設定された経路をOSPF経路情報に含める
    - そのほか、BGPなど他の経路テーブルから、OSPF経路情報に情報を移すというものがある。
  - デフォルトルートは、スタティックルートにあり、かつredistribute staticを行っても、redistributeされない。
    - Default-information-originateコマンドを使用します。

OSPF

- Redistributeの注意点
  - OSPF→RIPv2→OSPFなどとredistributeするときには、ルーティングループを起こす可能性が高い。



OSPF

- ローカルループバックアドレスのすすめ
(interface loopback 0を使う)
  - OSPFのルータIDは、アクティブなアドレスのなかで最も大きいアドレスを使う。
    - Ciscoの場合で、実装はルータによって異なる。
  - 機種やネットワークの切り替えの場合、WANインタフェースアドレスがRouter-IDになっていたりすると、トラブルの恐れがある。
    - ケーブル抜けやインタフェース障害でアドレスが無効になると、Router-IDが消える。
    - Shutdownしなくてはいけなくなる。
  - ループバックアドレスはを割り当てるとループバックアドレスがRouter-IDとなる実装がほとんどなので、アドレスが、他とぶつかることが無くなる。

OSPF

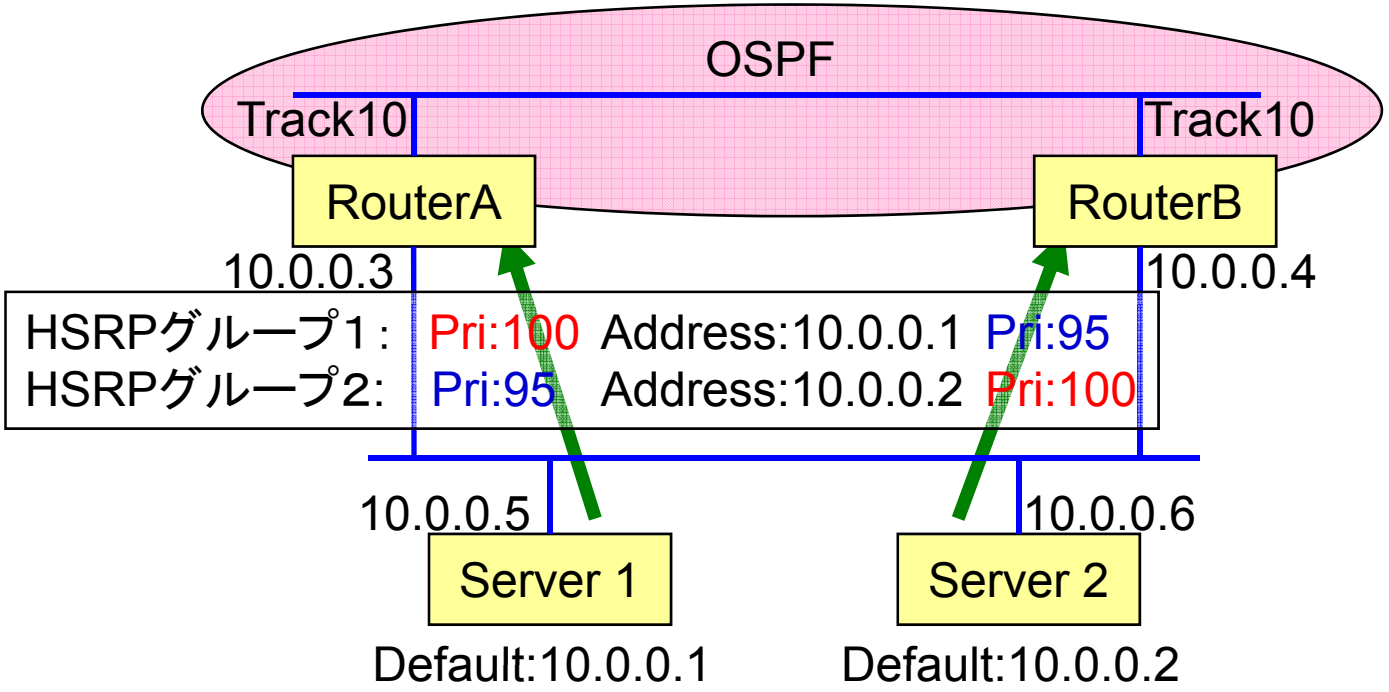
- ローカルループバックアドレスのすすめ
  - ループバックアドレスを使用する場合の効用
    - インタフェースを複数持っている場合に用いると、障害発生時にあるインタフェースがダウンしても、ループバックアドレスだけは生き続けるので、そのまま使用できる。
      - telnetの接続先アドレスとして使える
      - Syslogのソースアドレスとして利用できる
        - 障害解析が容易
      - BGPなのでピアリングするときも、同じアドレスが使える。
    - 問題点
      - /32(ホストルート）が経路表に載ってしまう。
        - 大きな問題ではない。

HSRP

- HSRPの活用
  - HSRP(Hot Standby Routing Protocol)とは
    - 一つの架空な実アドレスに対して、MACアドレスの割り当てを2台以上のルータから適宜変更することで、複数のルータの冗長構成を提供し、耐障害性をあげる技術
    - ダイナミックルーティングが使えない、機器の障害回避に有効
    - HSRPはCiscoの実装。
      - 標準的なのはVRRPという
      - このほか、各メーカで異なる名前の同様な仕組みを提供している。

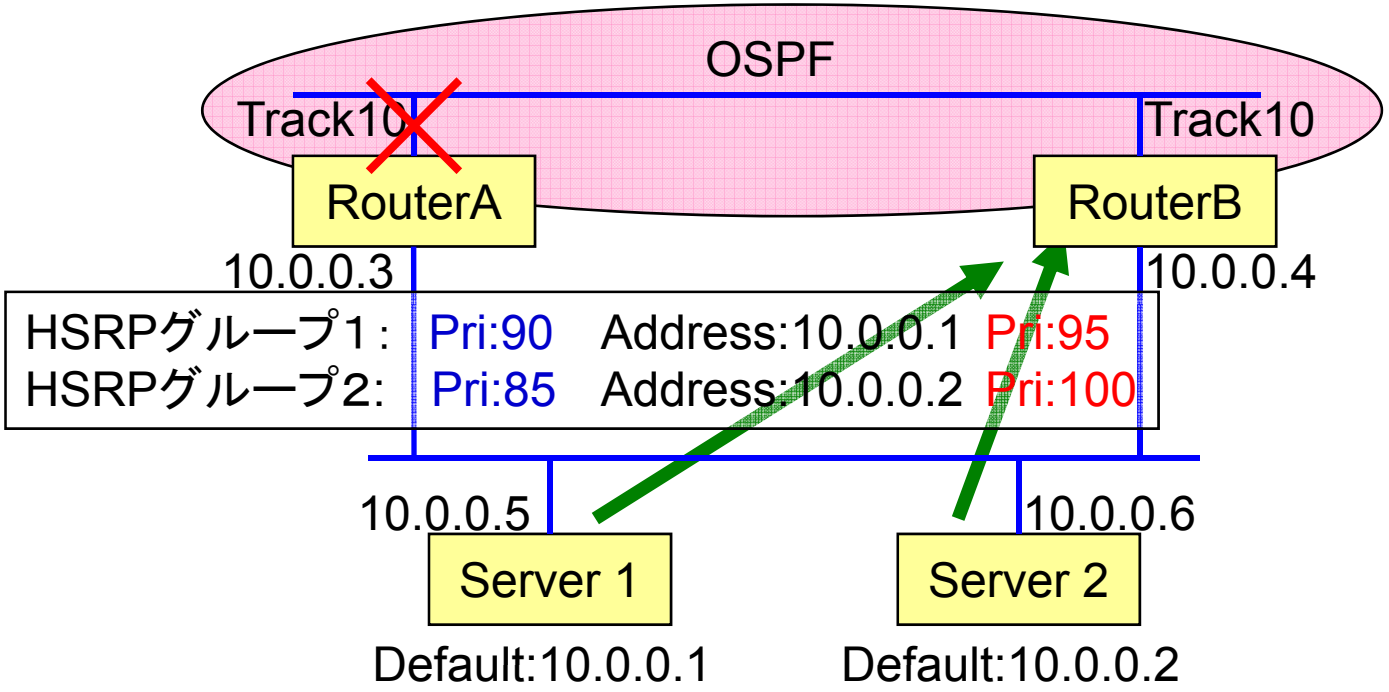
HSRP

- 通常時
  - 最も高いPriorityを持ったルータがActiveルータに、他のルータがStandbyルータになる。



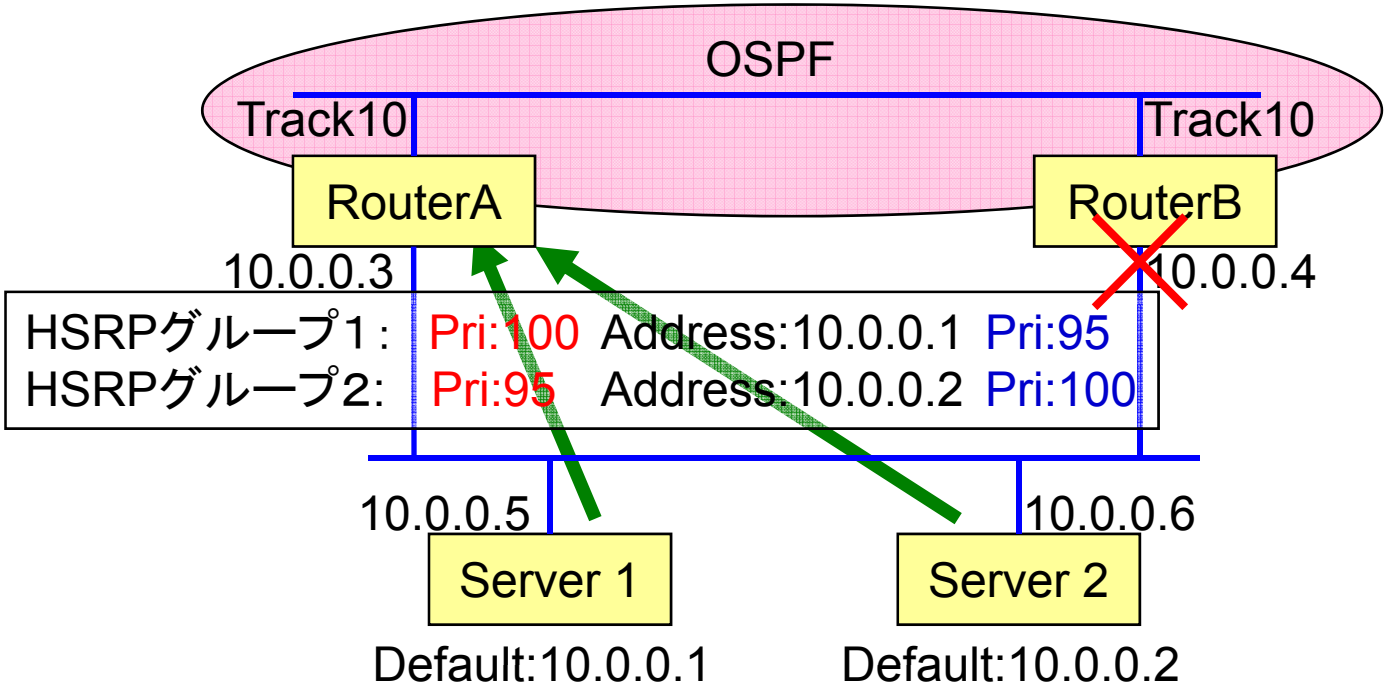
HSRP

- ルータAに障害発生
  - HSRPでは、Trackに指定したインタフェースがダウンするとTrackに指定していた数をPriorityから引いて、その結果のPriorityに従い、最も高いPriorityを持つルータがActiveになる。



HSRP

- ルータBに障害発生
  - HSRPでは、アクティブルータに対してKeepaliveパケットを出していて、Timeoutすると相手がダウンしているものとして、次にPriorityの高いルータがActiveとなる。



HSRP

- HSRPでは、複数のグループを同一のインタフェースで使用することが可能。
  - セグメント内で、デフォルトルートを分けることで負荷分散しつつ、冗長構成が作れる。
- ルータの機種によって、扱えるグループ数に制限がある。
- インタフェースのセカンダリアドレスと同時に使用することで、複数のセグメントを同一の物理ネットワーク上でHSRPを使って負荷分散と、障害時の迂回を同時に実現できる。
- HSRPを設定するインタフェースには、パケットリダイレクトが起こると問題がおきるので、no ip redirectの設定は必須となる。

ネットワーク障害監視

- ネットワークの障害監視の必要性
- 監視を行ううえでの留意点
- ネットワーク監視のためのツール
  - MRTG
  - RRDTool
  - Ping/Traceroute/Telnet
  - Sniffer
  - TTCP
  - Pathchar
  - Net-snmp
  - Looking Glass
  - PDAなどの活用

なぜネットワーク監視が必要なのか？

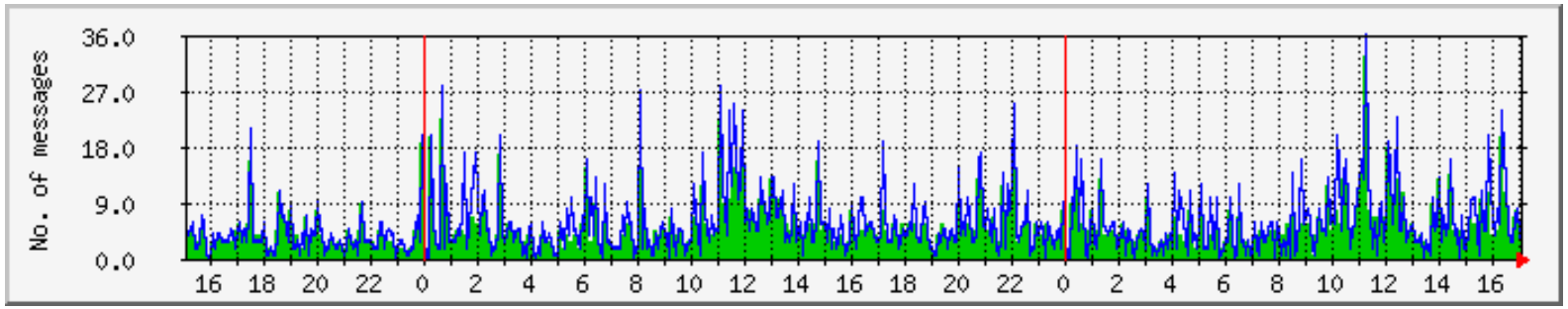
- トラブルは発生しないほうが良い
  - 発生させないためのネットワーク監視
    - 変な挙動をいち早く察知し、対応する。
- ネットワークの健康状態を知る
  - 不健康なら対策が必要、それによってトラブルは減る
- ネットワーク拡張の予測を立てる
  - 無計画なネットワーク拡張は、輻輳と無駄を発生させるだけ
- 自分のネットワークを守る
  - トラフィックの監視などで自分のネットワークへのアタックをもつけ、いち早く対処する。

監視を行ううえでの留意点

- 既存の各種ツールの有効活用
  - あるものを使うことで、管理コストを下げる
- 現在のトラフィックパターンを周知
  - パターン異常は、よからぬことが起きていることを示唆している。
- 各ネットワークの管理担当者を明確に
  - トラブルが発生しているネットワークを特定できても、連絡先が無ければ意味がない。
- 不要な機器はネットワークに接続しない
  - 不正アクセスの進入経路になりかねない。
- 機器の試験などは、専用のセグメントで
  - 試験の時に、誤って変な経路、トラフィックをだすことで、動いているネットワークに影響を及ぼす可能性もある。
- 機器で取得可能なログはできる限り残す
  - どんな情報が役に立つかわからない。

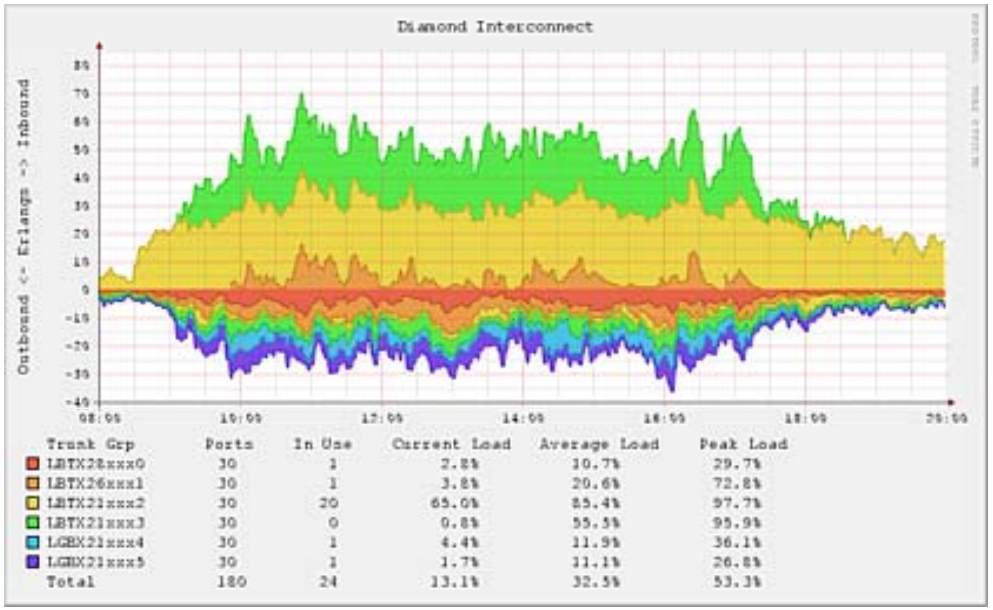
ネットワーク監視のためのツール

- MRTG
  - トラヒック計測、変動値のグラフ化が得意
  - http://oss.oetiker.ch/mrtg/



ネットワーク監視のためのツール

- RRDTool
  - MRTGと同じ開発者が作っているグラフ化ツール
  - MRTGをより発展させ、より細かいグラフ化が可能
  - http://oss.oetiker.ch/rrdtool/



ネットワーク監視のためのツール

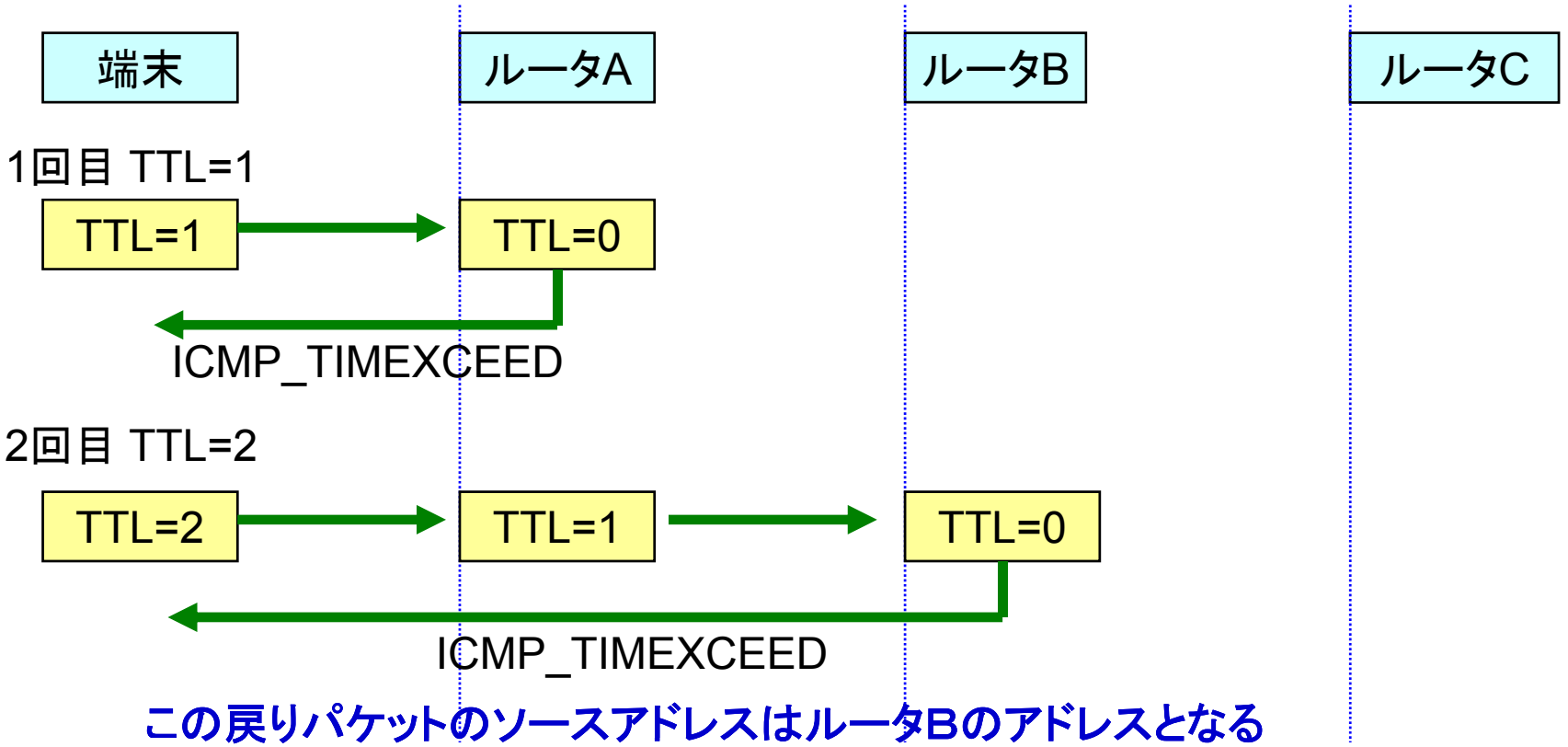
- Ping
  - ターゲットホストまでのRTTの参考になる。
  - ICMP_ECHOリクエストを利用したツール
  - Windows版とUNIX版でオプションが違います。

ネットワーク監視のためのツール

- Traceroute
  - UDPを用いるものとICMP_ECHOを用いるものの2通りがある。
  - 検査パケットを送る際、TTLを1から順に増やして行くことで、つながっているルータで順次ICMP_TIMEXCEEDを発生させることで、その帰ってきたルータのアドレスから、到達ホストまでの経路を検査する。
    - 行きの経路しか検査できない
    - 基本的に行きと帰りでパケットの通り道（経路）は異なる。
  - ICMPに対するフィルタを掛けているネットワークなどは、うまくtracerouteの結果が返ってこない。

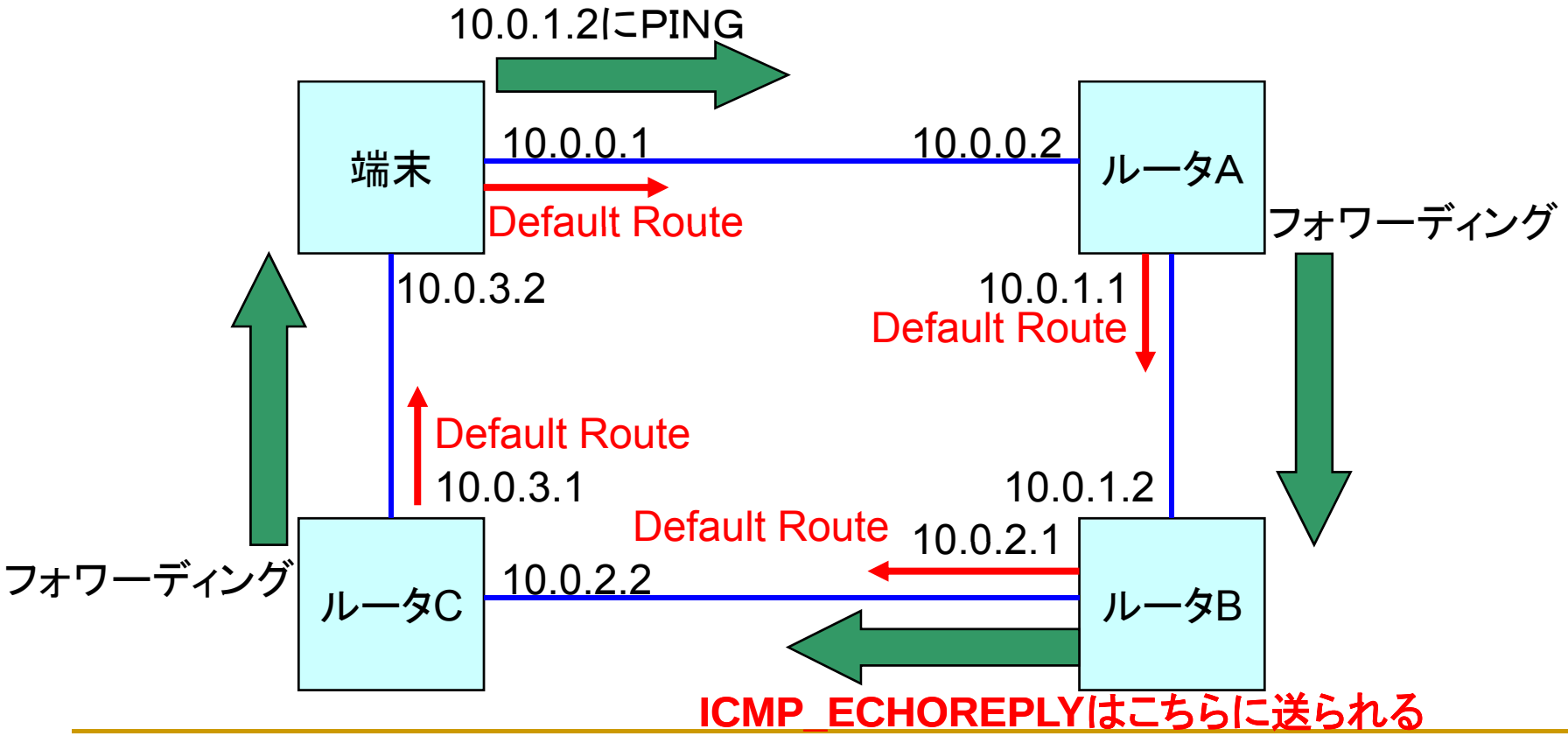
ネットワーク監視のためのツール

- Tracerouteの動き



ネットワーク監視のためのツール

- 非対称ルーティングの実態



ネットワーク監視のためのツール

- telnet
  - サーバがアプリケーションレベルで稼動しているかどうかを確認するために利用する。
  - telnet <host> <port>
  - httpであれば、 telnet <hostname> 80 とすれば、TCPセッションが確立でき、それによって、アプリケーションが正しく稼動しているかどうかを確認できる。
  - アプリケーションの動作がおかしいときの確認に有効

ネットワーク監視のためのツール

- Sniffer
  - LAN/WAN/ATM対応のアナライザ
  - OSI7層までのネットワーク障害をリアルタイムに検出が可能
  - OSI7層までのデータ解析が可能
  - 簡易アナライザとして、ソフト販売もされている。

  - 最近では、Etherealとかでも同じことができるので、そちらのほうがお勧めか？
    - http://www.ethereal.com/

ネットワーク監視のためのツール

- TTCP
  - 目的のサーバ間のTTCP同士でTCPパケットをバースト的に送信して、ネットワークのテストを行う。
  - ホスト間のパケットロス、伝達時間などを測定できる。
  - ネットワークにかなりの負荷をかける
  - http://www.pcausa.com/Utilities/pcattcp.htm

ネットワーク監視のためのツール

- Pathchar
  - ターゲットホストまでの回線の残容量を測定する
  - ICMPパケットを利用している
  - ネットワークにかなりの負荷をかける。
  - 最近似たようなものも結構あり、ネットワークの負荷を掛けないものもある（ようだ・・・）
  - http://www.caida.org/tools/utilities/others/pathchar/

ネットワーク監視のためのツール

- Net-snmp
  - SNMP Agentを含む様々なSNMPツールのパッケージ
  - コマンドによるため適用範囲が広い
  - 当然だがSNMPの知識が必要
  - http://net-snmp.sourceforge.net/

ネットワーク監視のためのツール

- ホームページからのping, tracerouteなども有効に利用できる。
  - http://www.nanog.org/lookingglass.html
  - http://neptune.dti.ad.jp/
    - など

お疲れ様でした

株式会社まほろば工房

http://www.d-commune.jp/

User Support/Collaboration Web Package

SPEED *http://www.ate-mahoroba.jp/speed/*

Emergency notification system

Network Operation System

http://www.ate-mahoroba.jp/iq3000/

Powered by MAHOROBA KOBO, Inc