学習情報ネットワーク論

（第6回）

室田真男
大学院社会理工学研究科 人間行動システム専攻
murota@hum.titech.ac.jp

今日の内容

- ネットワーク層
  - IP
    - IPアドレス，ネットワーク部，ホスト部
    - ルーティング
  - ICMP
    - ping，traceroute
- インターネットのキーワード
- レポート課題
- 授業評価

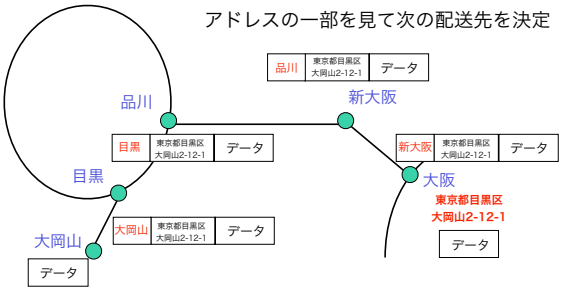
ネットワーク層

- 役割
  - データを宛先まで届ける
  - 上位層から（ネットワーク層における）宛先のアドレスを受け取り，配送する．
- インターネットにおけるネットワーク層は，IP (Internet Protocol)を利用
  - 宛先アドレス：IPアドレス
  - 配送の仕組み：ルーティングプロトコル

IPのアナロジー(1)

バケツリレー方式

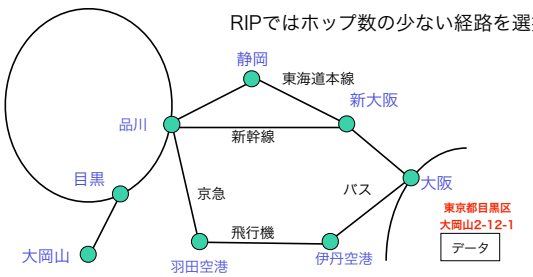
アドレスの一部を見て次の配送先を決定



IPのアナロジー(2)

経路の選択

RIPではホップ数の少ない経路を選択



IPの特徴

- シンプルな通信プロトコル
  - パケットと呼ばれるIPデータグラム
  - ルータ（中継ホスト）での単純な処理を規定
- 信頼性を保証しない
  - IPデータグラムが配送先に届くことを保証しない
  - ベストエフォートサービス
- コネクションレス
  - IPデータグラムの到着順序性や経路同一性に関与しない
  - それぞれのデータグラムは独立に扱われる

IPアドレス

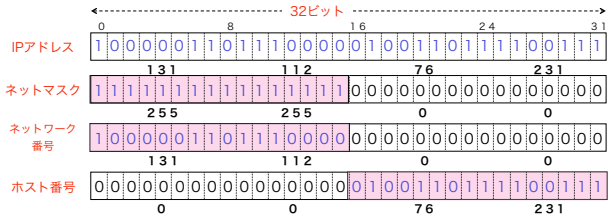
- IPv4
  - 現在主に利用されているIPアドレス
  - 32ビット
    - 約43億　(4,294,967,296)
- IPv6
  - 徐々に利用が進んできている次世代のIPアドレス
  - 128ビット
    - 地上の1cm²あたりに1020個のコンピュータがあっても対応可
    - 世界中の人が1028台コンピュータを持っていても対応可
- 授業で単にIPと書くときは、IPv4をさす

IPアドレス

- IPv4は32ビットのアドレス空間
- 表記法：8ビットを10進数で表記し、ドット(.)で接続
  - 131.112.76.231
- ネットワーク部とホスト部に分けられる
  - 目的：配送や管理を容易にするため
  - ネットマスク表記（ビットマスク表記）
    - 131.112.87.230 netmask 255.255.0.0
  - プレフィックス長表記
    - 131.112.87.230/16

ネットマスク表記

- ネットワーク部をネットマスクで表記する
- 131.112.76.231 netmask 255.255.0.0



プレフィックス長表記

- ネットワーク部の長さを “プレフィックス長”で表す
- 先頭から何ビットがネットワーク部を表すか
- 表記法：“IPアドレス/プレフィックス長”

  131.112.76.231/16

特別なIPアドレス

- 自分自身を表すアドレス
  - 127.0.0.1　(127.0.0.1/32)
- ネットワークを表すアドレス
  - IPアドレスのホスト部のビットが全て0
  - マシンには割り当てない
  - ネットワークそのものを表し、経路制御などに用いられる
- ブロードキャストアドレス
  - IPアドレスのホスト部のビットが全て1
  - そのネットワークに接続されている全てのマシンを表す
  - ブロードキャストに使用される

プライベートアドレス

- 組織内のみで利用するIPアドレス
  - IPアドレスの枯渇対策
  - 外部とは限られた通信しかしない場合
- プライベートIPアドレス
  - 10.0.0.0/8 : 10.0.0.0 ～ 10.255.255.255
  - 172.16.0.0/12 : 172.16.0.0 ～ 172.31.255.255
  - 192.168.0.0/16: 192.168.0.0 ～ 192.168.255.255

東工大の例

- 東工大は，131.112を取得
  - 学外から見たら東工大のアドレスは 131.112.0.0/16
- 学内では，ネットワーク部を26ビットで運用
  - 学内で，10ビット=1024 (=2^10) 個のネットワーク
    - サブネットという
  - 各サブネットでは，62 (=64-2)個(台)のホスト接続が可能
  - サブネットマスク：255.255.255.192　/26
- 東工大の外では，東工大内でどのようにサブネットを利用しているかは関係ない

一般ホストの配送先決定法

- 宛先IPアドレスが自分と同じネットワークの場合
  - 送信先に直接IPデータグラムを送信
- 宛先IPアドレスが自分とは別のネットワークの場合
  - デフォルトルータにIPデータグラムを送信

ルーティングテーブル

- ルーティングテーブル
  - 宛先に応じた送付先を管理
  - 宛先、次の送り先 (ネクストホップ) 、インタフェース
  - 宛先：ネットワークアドレス
  - ネクストホップ：次のルータのアドレス
- 検索のルール
  - 最長一致

宛先	ネクストホップ	インタフェース
0.0.0.0/0	192.168.76.1	192.168.76.64
127.0.0.1	127.0.0.1	127.0.0.1
192.168.76/24	192.168.76.64	192.168.76.64
192.168.76.64	127.0.0.1	127.0.0.1

ルーティングプロトコル

- 前提
  - ルータは直接接続しているネットワーク(サブネット)の情報を持っている
- 各ルータは，
  - 隣接ルータとネットワーク情報を相互に交換
  - その情報から「経路のよさ」を計算し経路選択
  - ルーティングテーブルを更新
- ルーティングプロトコル
  - ネットワーク情報の交換方法
  - 経路のよさの計算方法

経路の選択

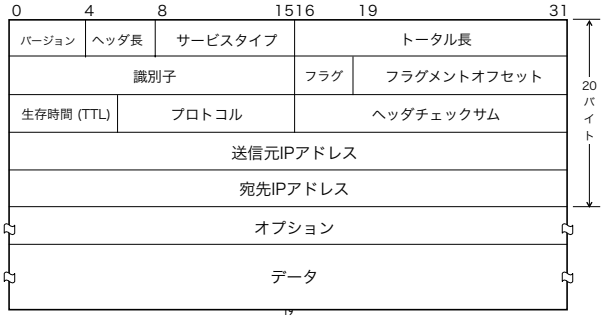
- 経路のよさを表現する正の整数値を使用
  - メトリック (metric) あるいはコスト (cost)
- データリンクに対して設定
  - 距離ベクトル型
  - リンク状態型
    - 方向により異なる可能性あり
- 宛先までの経路のよさの合計が最小の経路を選択

RIP

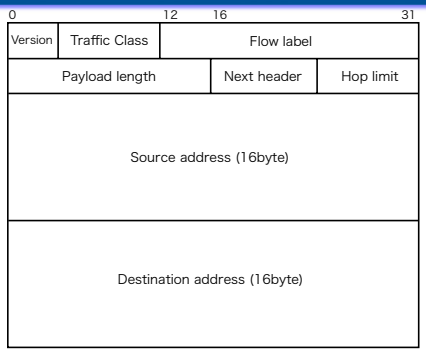
- Routing Information Protocol
- 距離ベクトル型
- メトリック
  - 送信元と宛先の間で通過するデータリンク数 (ホップ数)
- 最適ルート：ホップ数の一番小さいルート
- ルーティング情報を定期的（30秒）にネットワークにブロードキャスト

IPヘッダ

- 伝送順序は、図の左から右、上から下



IPv6ヘッダ



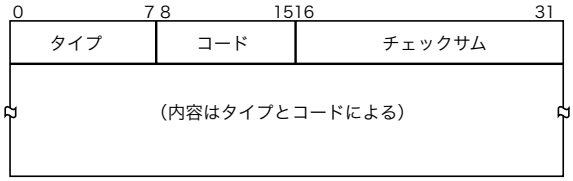
ICMP

- IPはベストエフォートサービス
  - IPデータグラムは相手に届かないかもしれない
  - しかし、黙って捨ててしまうのはあまりにも不親切
    - アプリケーションとしても困る
- 送信元に対してエラーの理由を報告する仕組み
  - ICMP (Internet Control Message Protocol)
  - 到達不能
    - ネットワークに到達不可
    - ホストに到達不可
    - 宛先ポートに到達不可
  - エコー応答　(エラー報告ではない)

ICMPメッセージ

- IPデータグラムにカプセル化される

- ICMPメッセージフォーマット



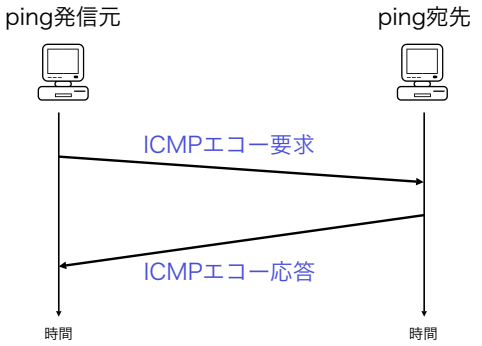
主なICMPメッセージタイプ

タイプ	コード	内容
0	0	エコー応答
3	0－15	宛先到達不可
	0	ネットワーク到達不可
	1	ホスト到達不可
	3	ポート到達不可
4	0	発信元規制
5	0－3	リダイレクト
	0	ネットワークへのリダイレクト
	1	ホストへのリダイレクト
8	0	エコー要求
11	0－1	時間超過

pingプログラム

- ping (Packet InterNet Groper)
- 機能
  - 他のホストにネットワーク的に到達可能かどうかをテスト
  - あるホストがpingに応答しなければ、そのホストはダウンしている可能性が高い
  - ホストまでの往復時間も測定
- 仕組み
  - ICMPエコー要求メッセージをホストに送る
  - ICMPエコー応答が返ってくるのを待つ

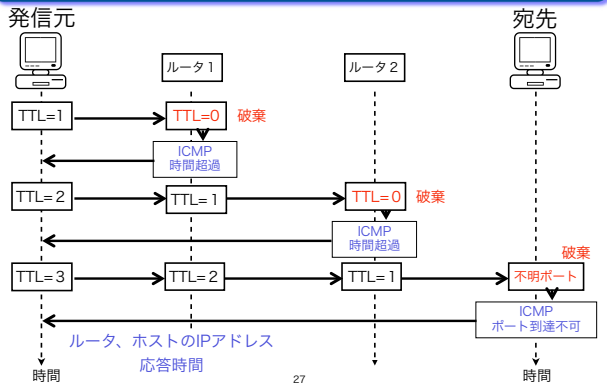
pingのモデル



tracerouteプログラム

- IPデータグラムの経路を調べることができる
  - 連続するIPデータグラムは，ほとんど同じ経路を通る
  - ルータは，TTLが0になったIPデータグラムは転送しないで破棄し，ICMP時間超過メッセージを送信元に送り返す
  - ICMPメッセージには，送信元ルータのIPアドレスが記載されている
- ルータの障害も調べることができる
- 何を利用しているか
  - IPヘッダのTTLフィールド
  - ICMPの時間超過メッセージ
  - ICMPの到達不能メッセージ

traceroutの動作



インターネットのキーワード

- 自律性（Autonomous）・分散性（Distributed）
- スケーラビリティ（Scalability）
- ベストエフォート（Best Effort）
- エンドシステム（End System, Edge System）
- セキュリティ（Security）
- 運用技術（Operation）

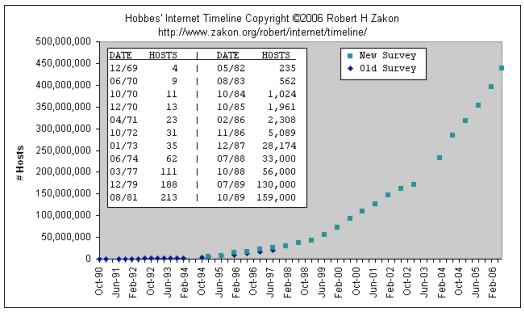
自律性・分散性

- インターネットは自律分散システムの典型的実装例
- ネットワークを運用している単位（組織など）はそれ自身で自律
  - 例えば研究室ネットワーク
  - インターネットからみた東工大ネットワーク
- 自律ネットワーク内ではどんな技術も利用可
- 外部ネットワークと接続するときの約束がプロトコル
- 全体を制御する組織や仕組みはなし
  - スケーラビリティ

スケーラビリティ

- これからインターネットに何人のユーザが繋がるか？
- 何台のコンピュータが繋がるか？
  - 誰にも予想できない
- 指数関数的に増加していくことが多い
  - 今後も続くと予想される
- インターネットに用いられる技術・アーキテクチャにはスケーラビリティが不可欠
- 指数関数的増加を予想したリソースの配分

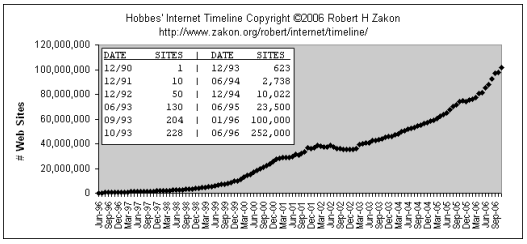
インターネットホスト数の増加



DATE	HOSTS	DATE	HOSTS
12/69	4	05/82	235
06/70	9	08/83	562
10/70	11	10/84	1,024
12/70	13	10/85	1,961
04/71	23	02/86	2,308
10/72	31	11/86	5,089
01/73	35	12/87	28,174
06/74	62	07/88	33,000
03/77	111	10/88	56,000
12/79	188	07/89	130,000
08/81	213	10/89	159,000

http://www.zakon.org/robert/internet/timeline/

WWWサーバ数の増加



DATE	SITES	DATE	SITES
12/90	1	12/93	623
12/91	10	06/94	2,738
12/92	50	12/94	10,022
06/93	130	06/95	23,500
09/93	204	01/96	100,000
10/93	228	06/96	252,000

http://www.zakon.org/robert/internet/timeline/

ベストエフォート

- 最善努力
- インターネットがデータを端末から端末まで運ぶときの根本的な考え方
- データを送るために最大限の努力はするが、もしかしたらデータが届かない可能性もある
- データ転送の保証をしない
- インフラストラクチャを軽く作れる
  - コスト、スケーラビリティ
- 今後はベストエフォートではない要素が入ってくる

エンドシステム

- エンドシステム、エッジシステム
  - 端末のコンピュータのこと
- インターネットはベストエフォートサービスであるが、最終的には信頼性のあるデータ転送が必要
- エンドシステムが信頼性を提供する
- 届かなかったら再送する
- 誤ったデータが届いたら、再送あるいは訂正する
- コストがかかること、難しいことは端末で
  - 分散システム
  - スケーラビリティ

セキュリティ

- 社会のインフラストラクチャとして経済活動に利用
  - オンラインショッピング
  - 税金の申告
  - インターネットバンク
  - 電子商取引
- ９１１、ウィルス、アタック、スパム
- 様々な対策が必要
  - 暗号化、認証、電子署名
  - インターネット全体の管理運用体制
  - 法整備
  - セキュリティ教育、モラル教育

運用技術

- 忘れてはならない技術（ノウハウ）
- 365日24時間稼働する必要がある
- 瞬間的にのみ優れた技術ではダメ