学習情報ネットワーク論

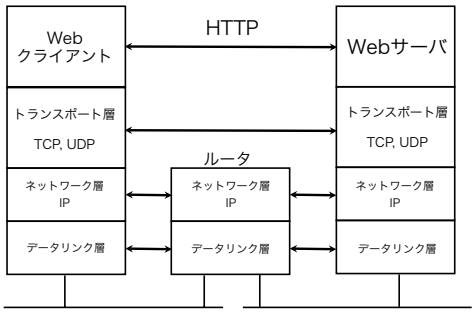
（第5回）

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（2008年1月31日）

- トランスポート層
  - ポート番号
  - UDP
  - TCP

インターネットプロトコル階層



トランスポート層の役割

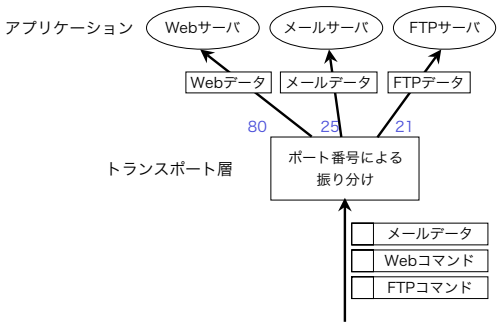
- End-to-Endの通信を担当
  - 上位アプリーケーションからのデータを，宛先のアプリケーションへ配送（プロセス間の通信）
  - ネットワーク層を利用してサービスを実現
- アプリケーション識別子：ポート番号
- 上位層からのデータ
  - 宛先IPアドレス，宛先ポート番号
  - データ

宛先IPアドレスを
取得するのは
上位層の仕事

ポート番号

- トランスポート層での識別子
  - アプリケーションごとに割り当て
  - 送信元と宛先で独立
  - 16ビットの正の整数
  - well-knownポート番号
    - http:80、電子メール(smtp):25、telnet:23、ftp:21

ポート番号による振り分け



コネクションの識別

- コネクション（通信）の識別
  - 送信元(IPアドレス,ポート番号)＋宛先(IPアドレス,ポート番号)
  - IPアドレスとポート番号の組：ソケット

アプリケーションの動作

- Webブラウザの例
  1. ユーザからURIの入力を得る
     - プロトコル（スキーム）：http
     - サーバのドメイン名：www.titech.ac.jp
  2. IPアドレスを取得
     - DNSへAレコードの問い合わせ：131.112.125.60
  3. ソケットを作成してコネクションを張る
     - TCP，IPアドレス：131.112.125.60，ポート番号：80
  4. HTTPリクエストフォーマットを作成（具体的フォーマットは省略）
  5. ソケットに送る
  6. サーバからの返信を得る
  7. 4〜6を繰り返しながら，ブラウザに表示
  8. コネクションを切断

アプリケーションが必要とするサービス

- 高信頼データ転送
  - 1bitも間違わないで転送
- 帯域幅
  - 転送量の保証
- 遅延
  - 高速性
  - 遅延保証

インターネットのトランスポートプロトコル

- ２つのトランスポートプロトコル
  - UDP (User Datagram Protocol)
  - TCP (Transmission Control Protocol)

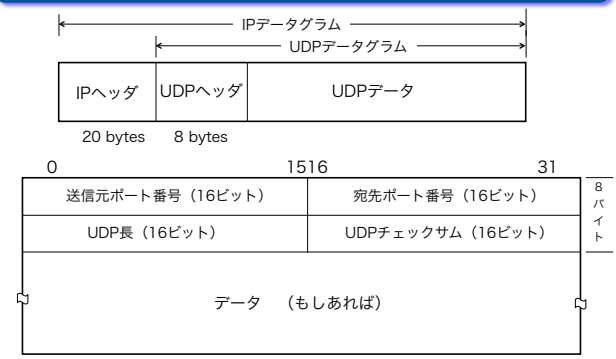
TCPとUDP

	TCP	UDP
高信頼性	○	×
帯域幅保証	×	×
遅延保証	×	×
高速性	×	○

UDP

- 仕様書（RFC768）はたった3ページ!!
- データ伝送用の簡単なメカニズムのみを提供
  - アプリケーションへのデータ分配のみ
  - オーバーヘッドが小さく高速処理
- 信頼性を提供しない
  - アプリケーションデータが宛先に届くことを保証しない

UDPデータグラム



UDPのアプリケーション例

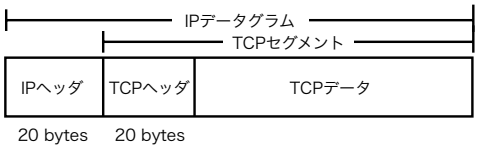
- パケット損失率が少ない安定した通信路を仮定
  - NFS (Network File System)
- 即時性，実時間性重視
  - DNS (Domain Name System)
  - 音声/動画ストリーミングアプリケーション
- 一対多通信
  - ブロードキャスト/マルチキャストアプリケーション

TCP

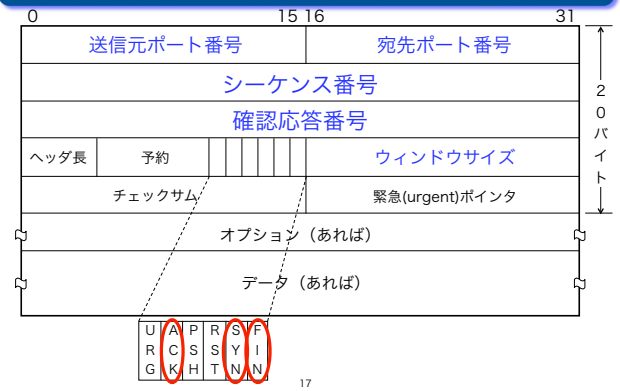
- 特徴
  - 非構造化ストリーム
  - コネクション型
  - 双方向通信
  - 信頼性のあるデータ転送サービス
  - フロー制御、輻輳制御

非構造化ストリーム

- データの内容/構造には関知しない
- アプリケーションからのデータをビット列と見なす
- 送信者からのビット列をそのまま宛先に渡す
- TCPでは，バイト単位に転送



TCPヘッダ

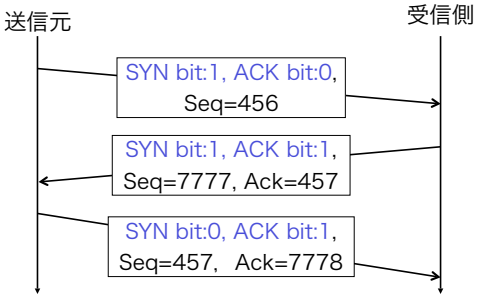


コネクション指向、双方向通信

- データ送信の前に送信側と受信側が「交渉」して、仮想コネクションを張る
  - バーチャルサーキット
- コネクションの確立手法
  - スリーウェイ・ハンドシェイク
- 双方向通信 (全二重化通信)
  - 通信する両者が同時にデータを送受信できる

コネクションの確立

- スリーウェイ・ハンドシェイク



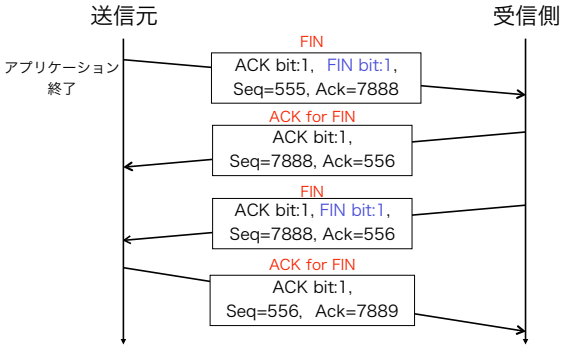
コネクション確立時の交渉

- シーケンス番号の初期値の設定
  - 双方向でそれぞれ独立に0でない値が設定される
  - 第三者に推定されないことが重要
- データ受信能力を相手へ通知
  - 受信側が最大のウィンドウサイズを通知
  - 送信側は受信側の受信バッファサイズを知る

コネクションの終了

- 双方向コネクションを片方向ずつ閉じる
  - データ送信が終了したらFINを送る（FIN bitを1）
  - FINを受信するとFINに対するACKを返す
  - 上記を双方向ともに行う
- FIN受信は相手側のデータ送信が終了したという意味
  - FIN受信後もデータ送信は可能
  - 片方向だけのコネクションが閉じた状態：ハーフクローズ (half-close)

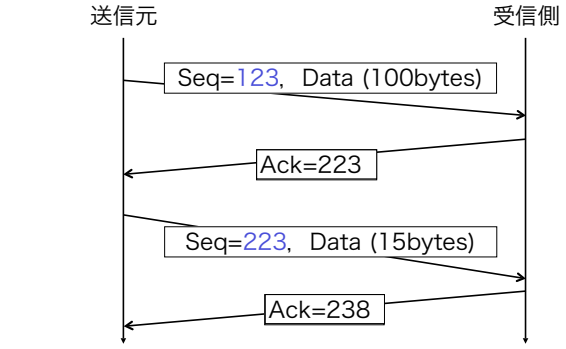
コネクション終了の様子



信頼性のあるデータ転送

- アプリケーションから見ると，TCPにデータを渡すと，必ず順序通りに届く
- 仕組み
  - シーケンス番号による送達確認
    - 送信データにバイト単位にシーケンス番号を割当
    - 受信側は，次に受信したいシーケンス番号を確認応答番号 (Acknowledge Number, ACK)として返す
    - ACKを受信したら次のデータを送る
  - データ再送機能
    - 失われた（と思われる）データを再送

シーケンス番号による送達確認



データ再送機能

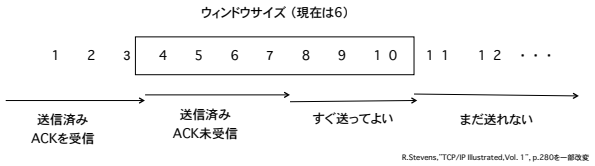
- 失われた（と思われる）データを再送
- 再送するタイミングは基本的に次の2種類
  - 再送タイマがタイムアウト
    - データ送信後，一定時間待っても確認応答が戻ってこない
  - 重複する確認応答を３つ以上受信
    - 同じ番号のACKが戻ってきた場合は，そのデータが失われた可能性が高い
    - 一定数以上受信した場合，タイムアウトを待たないでそのデータを再送する

データ転送の効率化

- 1バイトごとACK待ってから送るのは効率が悪い
- ACKが戻ってこなくても，ある程度先に送る
- どの程度先に送ってよいのか？

  → フロー制御
- スライディングウィンドウ方式
  - 相手の受信能力にあわせて送信
  - ネットワーク混雑状況にあわせて送信

スライディングウィンドウ方式



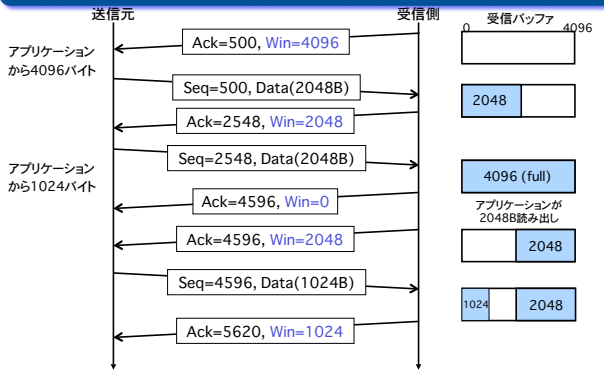
R.Stevens,"TCP/IP Illustrated,Vol. 1", p.280を一部改変

- ウィンドウは次の様に右に移動していく
  - ウィンドウの左端：ACKが返ってくることにより右に移動
  - ウィンドウの右端：ウィンドウサイズに応じて右に移動
- ウィンドウサイズを変化させてフロー制御を行う

2種類のウィンドウサイズ

- 受信側から通知されるウィンドウサイズ
  - 受信側が受信バッファのサイズを送信
  - TCPヘッダフィールド中に含まれている
  - 相手の受信能力にあわせて送信可
- 輻輳ウィンドウ(cwnd)サイズ
  - 送信側がネットワークの混雑状況を判断して決める
  - ネットワーク混雑状況にあわせて送信可
- 上記2つのウィンドウのうち，サイズの小さいウィンドウサイズを用いる

受信側からのウィンドウサイズ



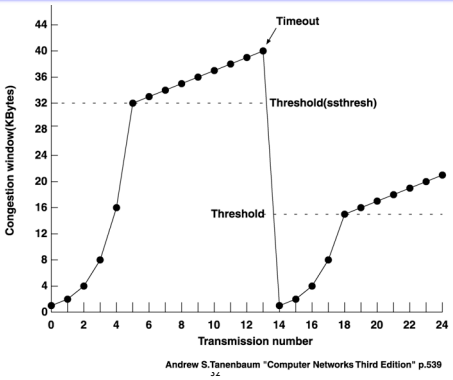
混雑状況の判断

- 基本的な考え方
  - ACK受信までの時間で、ネットワークの混雑状況を判断
  - RTT (Round Trip Time)
    - 返事が早い → 空いている
    - 返事が遅い(こない) → 混んでいる
  - 輻輳ウィンドウ(cwnd)のサイズ変化の原則
    - 空いているとき → たくさん送る
    - 混んでいるとき → 少しずつ送る

輻輳ウィンドウの動作(1)

- 2種類の増やし方
  - スロースタート段階
    - 新しい TCPコネクション確立時は、送信側のcwndはあるサイズに初期化される
    - ACKを受け取るたびに、cwndは倍に増やされる
    - 結果的にcwndは指数関数的に増加する
  - 輻輳回避段階
    - スロースタートしきい値(ssthresh)が定められる
    - cwndはしきい値に達すると、そこからは線形に増加する

輻輳ウィンドウの例



Andrew S.Tanenbaum "Computer Networks Third Edition" p.539

輻輳ウィンドウの動作(2)

- 2種類の減らし方
  - タイムアウトが起こった場合
    - 現在のウィンドウサイズの1/2がssthreshとされる
    - cwndは初期値にもどされる
    - スロースタート段階
  - 重複ACKによる再送
    - ACKが3つ重複するとすぐに再送する
    - cwnd、ssthreshともに現在のウィンドウサイズの1/2とされる
    - 輻輳回避段階

今日のキーワード

- トランスポート層
  - ポート番号，Well-knownポート番号，ソケット
  - TCP・UDP
  - スリーウェイハンドシェイク
  - シーケンス番号，確認応答番号
  - フロー制御，スライディングウィンドウ方式