

## Internet Week 2001 B5

### プロトコル詳説

～クリックしてからホームページが表示されるまで～

ネットワンシステムズ(株) 白橋明弘

## 本講座のねらい

- ◆「クリックしてからホームページが表示されるまで」を題材として
- ◆レイヤー2からレイヤー7までのプロトコルの動作を実際に即して解説する
- ◆ルータ、スイッチの世界とDNS、Mailなどアプリケーションの世界のギャップを埋める

## 前提とする知識

- ◆TCP/IP の基本的概念(IPアドレスやポート番号)を知っている
- ◆Web や DNS/Mail のしくみについて概念的な理解ができている

## 対象として想定する環境

- ◆ダイヤルアップまたはADSL常時接続の環境で
- ◆プロバイダへの接続が確立して
- ◆パソコン上のWebブラウザでインターネット上のWebサイトを閲覧するまで



## OSI参照モデル(1)

- ① 物理層 (Physical Layer)  
物理的な接続上でビット列をやり取りする
- ② データリンク層 (Data Link Layer)  
データをリンク上でパケットというまとまりでやり取りする
- ③ ネットワーク層 (Network Layer)  
パケットを複数のリンクをまたがって宛先に送り届ける

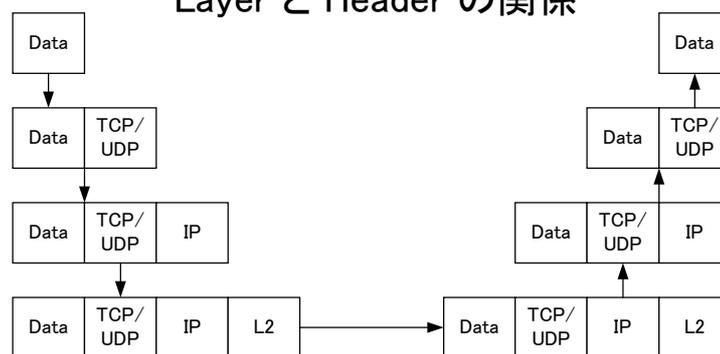
## OSI参照モデル(2)

- ④ トランスポート層 (Transport Layer)  
信頼性のある2点間の通信路を提供する
- ⑤ セッション層 (Session Layer)  
トランスポート層に付加機能を提供する
- ⑥ プレゼンテーション層 (Presentation Layer)  
アプリケーションのデータを通信に適する形式にエンコードする
- ⑦ アプリケーション層 (Application Layer)  
ファイル転送や電子メールといったアプリケーション

## OSI参照モデルとTCP/IPの対応

- ◆ 物理層・データリンク層 → データリンク層:L2
- ◆ ネットワーク層 → ネットワーク層 (IP):L3
- ◆ トランスポート層 → トランスポート層 (TCP):L4
- ◆ セッション層・プレゼンテーション層・アプリケーション層 → アプリケーション層:L7

## Layer と Header の関係



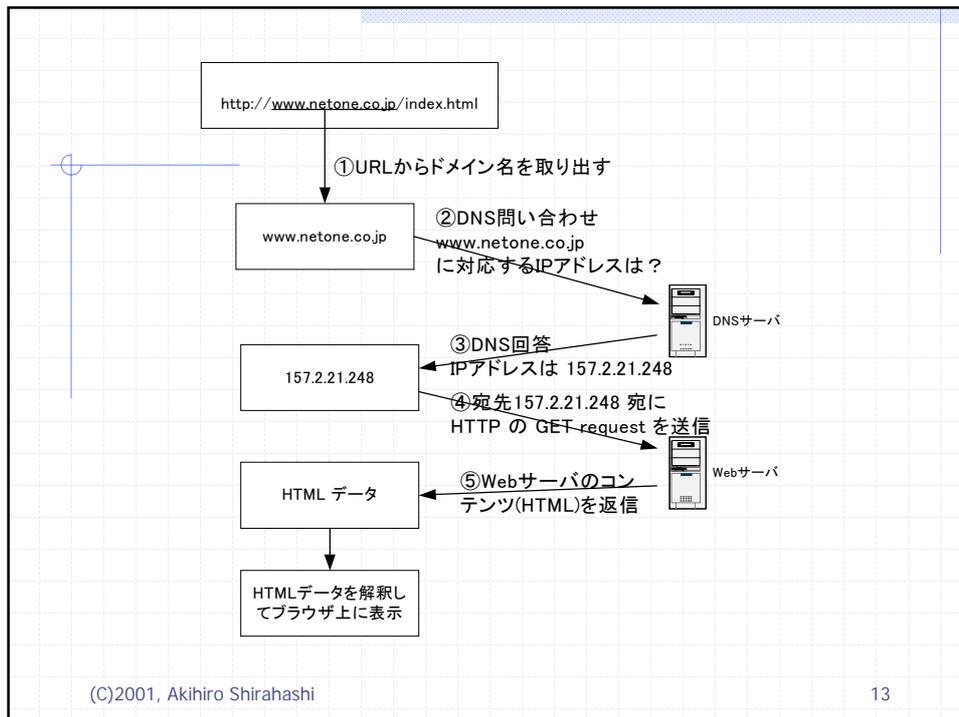
## 登場するプロトコル

- ◆ データリンク層  
Ethernet, ARP, PPP, PPPoE
- ◆ ネットワーク層  
IP
- ◆ トランスポート層  
TCP, UDP
- ◆ アプリケーション層  
DNS, HTTP

## クリックしてから表示されるまで

最も上位層で見ると、こうなる:

- ① URLをクリック (またはURLを入力)
- ② ドメイン名をDNSでIPアドレスに変換する
- ③ そのIPアドレスにHTTPのGETリクエストを送り、レスポンスとしてHTMLのデータを受け取る
- ④ WebブラウザがHTMLを解釈して表示する



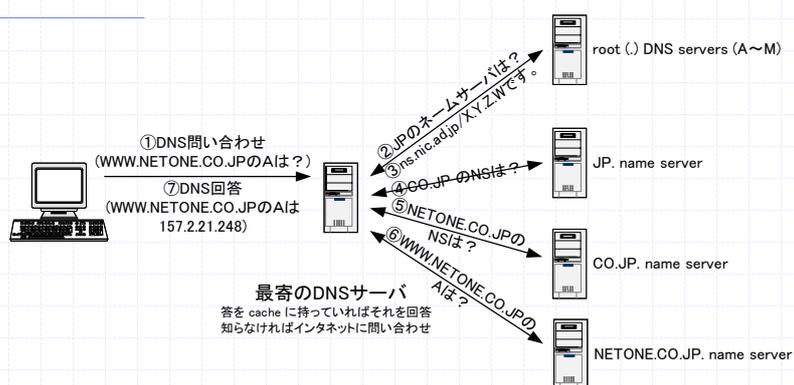
## ドメイン名をIPアドレスに変換

これを詳しく見ると:

- ① パソコンに設定されたDNSサーバのIPアドレスを宛先として、DNSのqueryのパケットを送る
- ② DNSサーバが、このパケットを受け取り、その中に含まれているドメイン名をIPアドレスに変換して、答をパソコンに返す

## DNSサーバが回答できるわけ

- ① 自分が知っている(当該ドメインについて authorityを持っているか、以前の答えをcacheしている)ものは、それを答える
- ② 自分が知らないものは、そのドメイン名について最も知っていそうなDNSサーバに問い合わせる
- ③ そのドメイン名について何も知らなければ、root server に問い合わせ . → jp → co → netone と降りてくる



(注) この説明はよく目にしますが、ちょっと嘘があります

## DNS に関する疑問(1)

クライアントPCはDNSの問い合わせを送るべき最寄のDNSサーバのアドレスをどうやって知るか

- 静的に設定
- DHCPで割り当て
- IPCPで割り当て(PPP)
- mode-configで割り当て(IPsec VPN)

## Windows 2000 ipconfig/all 出力

```
Ethernet adapter ローカル エリア接続:  
    Connection-specific DNS Suffix . . : iw2001.internetweek.jp  
    Description . . . . . : Intel(R) PRO/100 SP Mobile Combo Adapter  
    Physical Address. . . . . : 00-03-47-6E-XX-XX  
    DHCP Enabled. . . . . : Yes  
    Autoconfiguration Enabled . . . . : Yes  
    IP Address. . . . . : 61.215.25.135  
    Subnet Mask . . . . . : 255.255.255.0  
    Default Gateway . . . . . : 61.215.25.1  
    DHCP Server . . . . . : 61.215.25.1  
    DNS Servers . . . . . : 61.215.25.2  
    Lease Obtained. . . . . : 2001年12月7日 13:13:44  
    Lease Expires . . . . . : 2001年12月7日 13:28:44
```

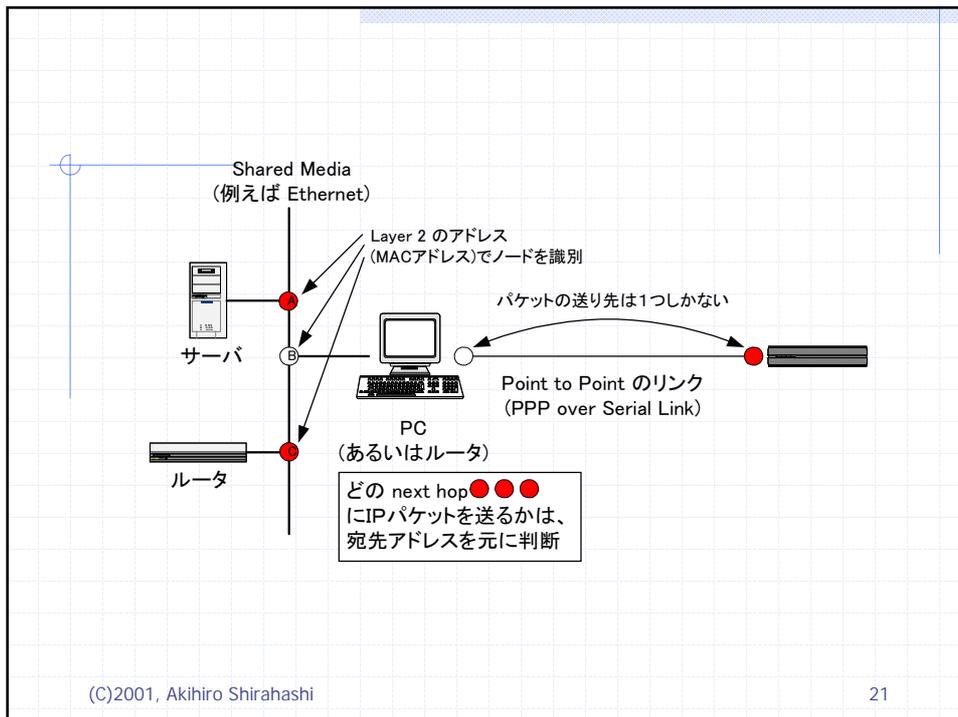
## DNSに関する疑問(2)

(最寄の)DNSサーバは、root servers(A～M)のIPアドレスをどうやって知るのか？

- DNSサーバの設定を最初から持っている
- それは Root cache と呼ばれるファイル
- Root servers はめったに変更にならないのでそれでも大丈夫

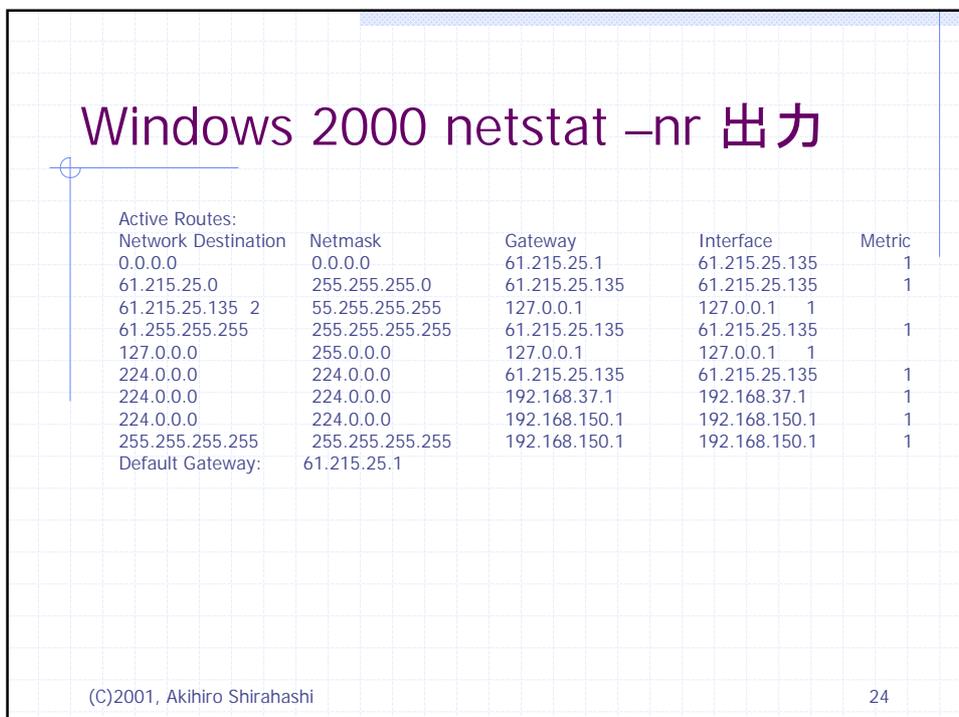
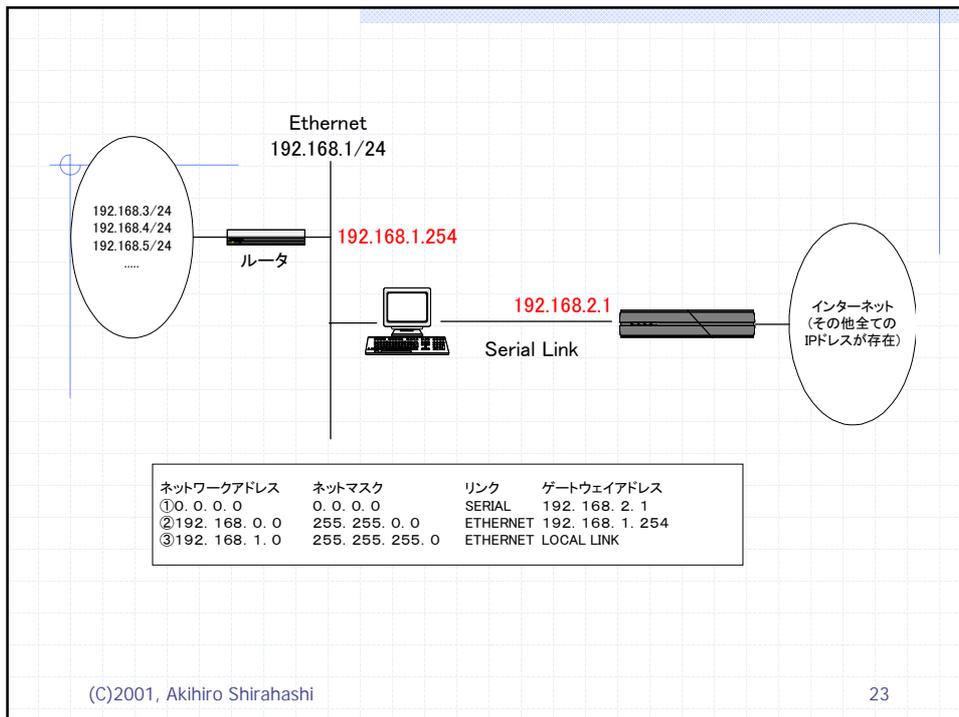
## DNSサーバへパケットが届く訳

- ◆ IPパケットは、どれかの(データリンク層の)リンクの next hop の宛先に対して送られる
- ◆ ダイヤルアップPPP接続の場合は、リンクは1つしかないし、next hop も point to point リンクなので選ぶ余地は無い
- ◆ 一般には、(複数ある)どのリンクを選ぶか、そして(Ethernet のようなリンク上に複数の宛先の存在する場合) next hop の選択は、宛先IPアドレスによって決まる



## 宛先IPアドレスから next hop を決める

- ◆ IPパケットをforwardする機器(ホスト、ルータ)は、  
経路情報テーブル(routing table)を持つ
- ◆ Routing table には、IPアドレス(A)/ネットマスク  
(M)と、それに対応するリンクとnext hop が書か  
れている
- ◆ (宛先IPアドレス and M) = A となる routing  
table のエントリのリンク,next hopが選択される
- ◆ 複数のエントリがマッチする場合は、ネットマスク  
が最長のものが選ばれる (longest match)



## 宛先アドレス 192.168.3.1 の場合

### ◆ ネットマスクと and を取る

- ①  $192.168.3.1 \text{ and } 0.0.0.0 = 0.0.0.0$
- ②  $192.168.3.1 \text{ and } 255.255.0.0 = 192.168.0.0$
- ③  $192.168.3.1 \text{ and } 255.255.255.0 = 192.168.3.0$

### ◆ ネットワークアドレスと比較する

- ①  $0.0.0.0 = 0.0.0.0 \rightarrow \text{OK}$
- ②  $192.168.0.0 = 192.168.0.0 \rightarrow \text{OK}$
- ③  $192.168.3.0 \neq 192.168.1.0 \rightarrow \text{NG}$

### ◆ ①と②のうちネットマスクの<長い>方を選択

- ①0bit/②16bitなので、②を選択  
→ next hop は 192.168.1.254

## 宛先アドレス 157.2.21.248 の場合

### ◆ ネットマスクと and を取る

- ①  $157.2.21.248 \text{ and } 0.0.0.0 = 0.0.0.0$
- ②  $157.2.21.248 \text{ and } 255.255.0.0 = 157.2.0.0$
- ③  $157.2.21.248 \text{ and } 255.255.255.0 = 157.2.21.0$

### ◆ ネットワークアドレスと比較する

- ①  $0.0.0.0 = 0.0.0.0 \rightarrow \text{OK}$
- ②  $157.2.0.0 \neq 192.168.0.0 \rightarrow \text{NG}$
- ③  $157.2.21.0 \neq 192.168.1.0 \rightarrow \text{NG}$

### ◆ ①を選択

0.0.0.0/0.0.0.0 の経路 = “default route”  
どのアドレスにもマッチする (が一番弱い)

# PPPの接続の確立

PPPのリンク確立は以下の手順で:

- ① 電話をかける
- ② LCPオプションの設定要求/確認  
LCP: Link Control Protocol
- ③ ユーザー認証 (PAP or CHAP)
- ④ NCP ICPオプションの設定要求/確認  
NCP: Network Control Protocol  
ICP: IP Control Protocol

## PPPの接続

クライアント

アクセスサーバ

LCPオプションの設定要求

LCPオプションの設定要求

LCPオプションの設定確認

LCPオプションの設定確認

ユーザー認証の要求

認証用チャレンジ

認証用レスポンス

認証結果通知

NCP/IPCPオプションの要求

NCP/IPCPオプションの要求

NCP/IPCPオプションの設定

NCP/IPCPオプションの設定

②LCP

③認証

④IPCP

## PPPoEの接続の確立

PPPoEのリンク確立は以下の手順で:

- ① クライアントがPPPoEのサーバを探す問い合わせをブロードキャスト
- ② PPPoEのサーバが、クライアントへ返答
- ③ クライアントがセッションの開始を要求
- ④ サーバがセッションIDを通知
- ⑤ 以下、PPPのネゴシエーションが始まる

## PPPoEの接続

クライアント

アクセスサーバ  
(Concentrator, BRAS)

Discovery パケットを Ethnet に broadcast

Confirmation パケットを返信

Session の開始を要求

Session ID を通知

以下はPPPの通信と同じ

LCP交渉

認証

NGP/IPCP交渉

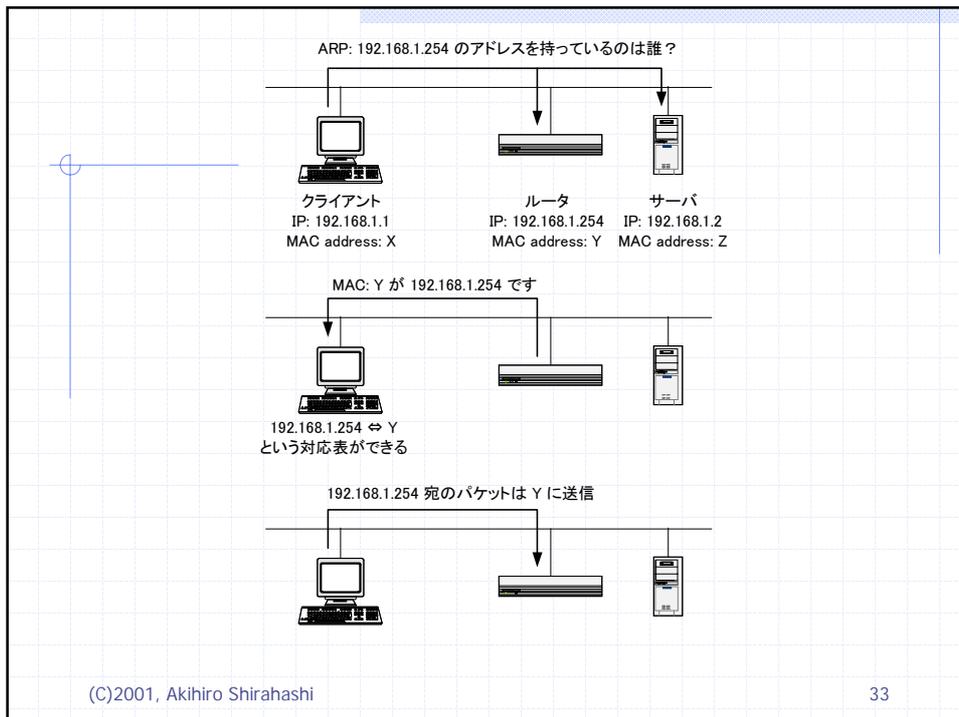
PPPの通信

## PPPoEは何故必要か

- ◆ データリンク層であるEthernetの上で、同じデータリンク層であるPPPをする意味は？
- ◆ PPPにはユーザ認証の仕組みがあるが、Ethernetにはユーザ認証の仕組みがない
- ◆ ユーザ認証の仕組みをEthernet上で使うことを主たる目的としてPPPoEが作られた

## Ethernet上での通信は

- ◆ Ethernet上での通信はMACアドレスに対してフレームを送ることで実現される
- ◆ IPアドレスでは直接通信できない
- ◆ 宛先IPアドレスからMACアドレスを知る必要がある → このためのプロトコルがARP (Address Resolution Protocol)
- ◆ IPアドレスをブロードキャストのEthernetフレームに入れて送信 → そのIPアドレスを持つ機器がMACアドレスを返信
- ◆ 得られたMACアドレスは一定時間cacheして利用



## Windows 2000 arp -a の出力

Interface: 0.0.0.0 on Interface 0x4

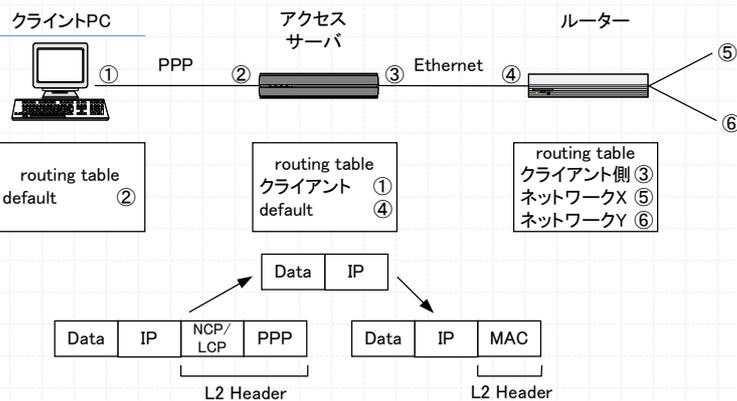
| Internet Address | Physical Address  | Type    |
|------------------|-------------------|---------|
| 10.32.10.238     | 00-00-66-61-xx-xx | dynamic |

Interface: 61.215.25.135 on Interface 0x1000006

| Internet Address | Physical Address  | Type    |
|------------------|-------------------|---------|
| 61.215.25.1      | 00-50-2a-84-8c-00 | dynamic |
| 61.215.25.2      | 00-50-b9-00-01-35 | dynamic |

# IPパケットの転送

- ◆ IPパケットは  
クライアントPC → (PPP/PPPoE) → アクセス  
サーバ → ルーター → ルーター → サーバ  
と転送されていく
- ◆ ルーターは、それぞれ自分の持つ routing  
table を参照して次の転送先を決める
- ◆ その routing table はどうやって決まる？



## Routing table は

- ◆ Routing Table は、静的(static)または動的(dynamic)で決まる
- ◆ 末端ホストや、単純なネットワークのルータは routing table 情報を管理者が設定して静的に与える (static routing)
- ◆ 複雑なネットワークでは、RIP/OSPF/BGPといった routing protocol で隣のルータと情報交換をして、routing table を自動的に構成する (dynamic routing)

## HTTP のリクエスト

- ◆ DNS の名前解決が出来て Web サーバの IPアドレスがわかると、次にはWebサーバに対してHTTPのリクエストが送られる
- ◆ HTTPは、パケットを送りっぱなしのUDPであるDNSと違いTCPである
- ◆ TCPでは、信頼性のある通信を保証するためTCPのセッションという手続きがある

# TCP のセッションの確立

TCPセッションは次のように確立される

- ① クライアントからサーバにTCP同期要求(SYN)パケットを送信
  - ② サーバから同期要求+応答(SYN+ACK)パケットを返信
  - ③ クライアントから応答(ACK)パケットを送信
- この 3way handshake でTCPのセッションが確立

# TCPの接続

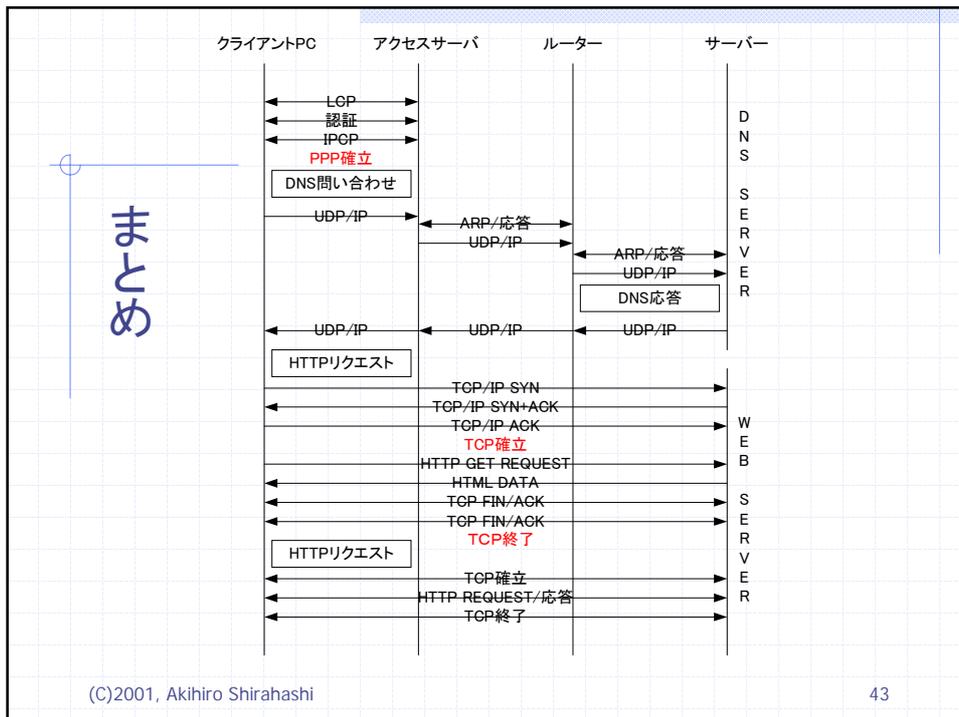


## TCP の役割

- ◆信頼性のある、クライアントとサーバ間の双方向データチャンネルを提供
  - コネクションの確立 (3way handshake)
  - 確認応答 (ACK)
  - 再送
  - 到着順序の保証
  - フロー制御 (輻輳制御)

## HTTPのリクエストと応答

- ◆サーバのHTTPのポート(標準は80)にリクエストが送られる  
GET index.html HTTP/1.1
- ◆サーバからレスポンスが返される  
HTTP/1.1 200 OK  
Content-Type: text/html  
<以下HTMLのコンテンツ>
- ◆基本は、1 HTTP request/1 TCP session



- ## 関連RFC一覧
- ◆ IP RFC791
  - ◆ UDP RFC768
  - ◆ TCP RFC793
  - ◆ DNS RFC1034,1035
  - ◆ HTTP RFC2616
  - ◆ PPP RFC1661
  - ◆ PPP/LCP RFC1570
  - ◆ PPP/IPCP RFC1332
  - ◆ PPPoE RFC2516
  - ◆ ARP RFC826
- (C)2001, Akihiro Shirahashi 44

## 参考文献

- ◆ 詳解TCP/IP Vol.1 プロトコル  
(ピアゾン・エデュケーション)
- ◆ インターネットRFC事典  
(アスキー)