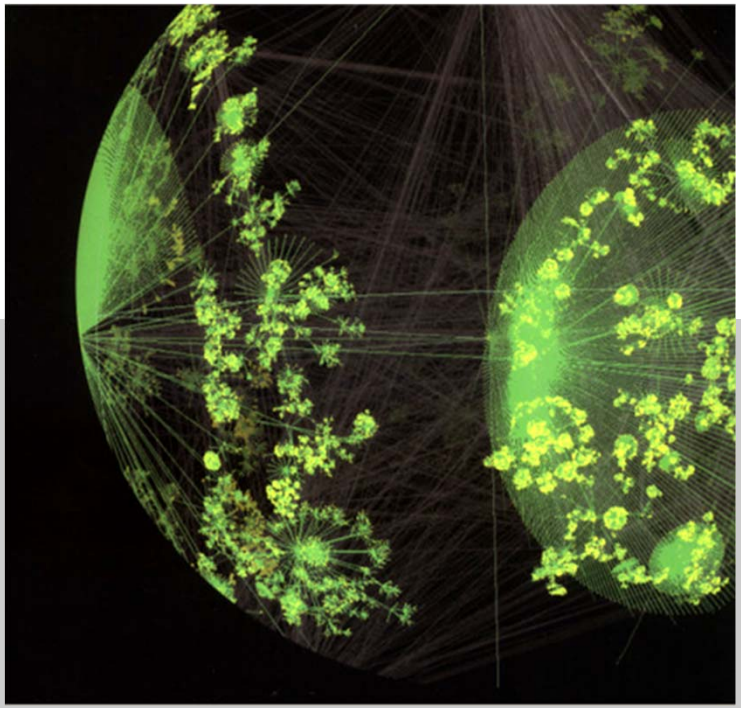


Chapter 8
The Web, DHCP, NTP and NAT

TCP/IP Essentials

A Lab-Based Approach

Spring 2017

World Wide Web (WWW)

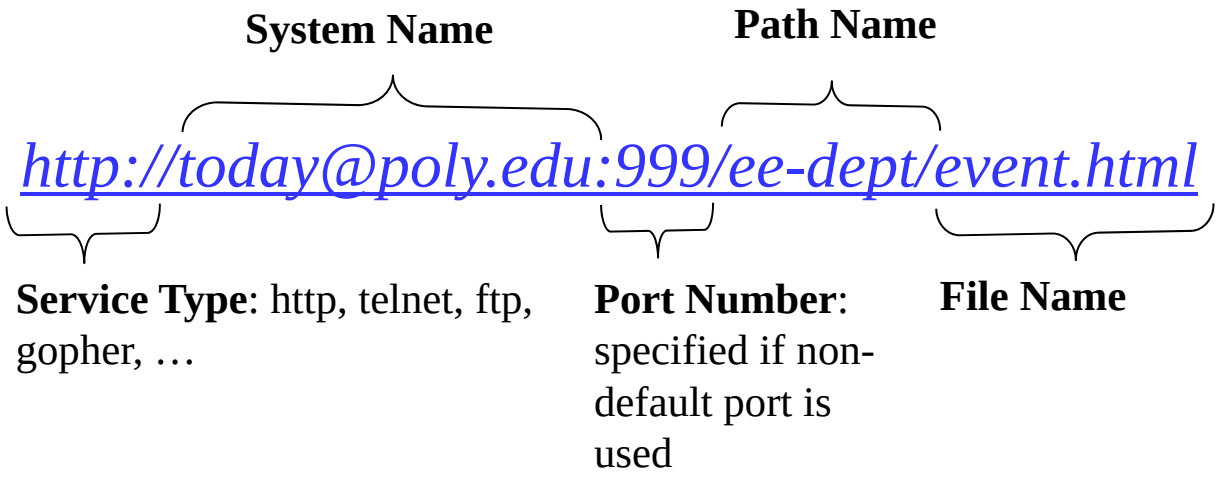
- The first WWW browser was written by Tim Berners-Lee in 1990.
- The popularity of WWW resulted in the exponential growth for the Internet.
- In WWW, information is typically provided as Hyper Text Markup Language (HTML) files (web pages).
- WWW resources are specified by Uniform Resource Locators (URL).

Uniform Resource Locator (URL)

A standard scheme for compactly locating any document on any Web server

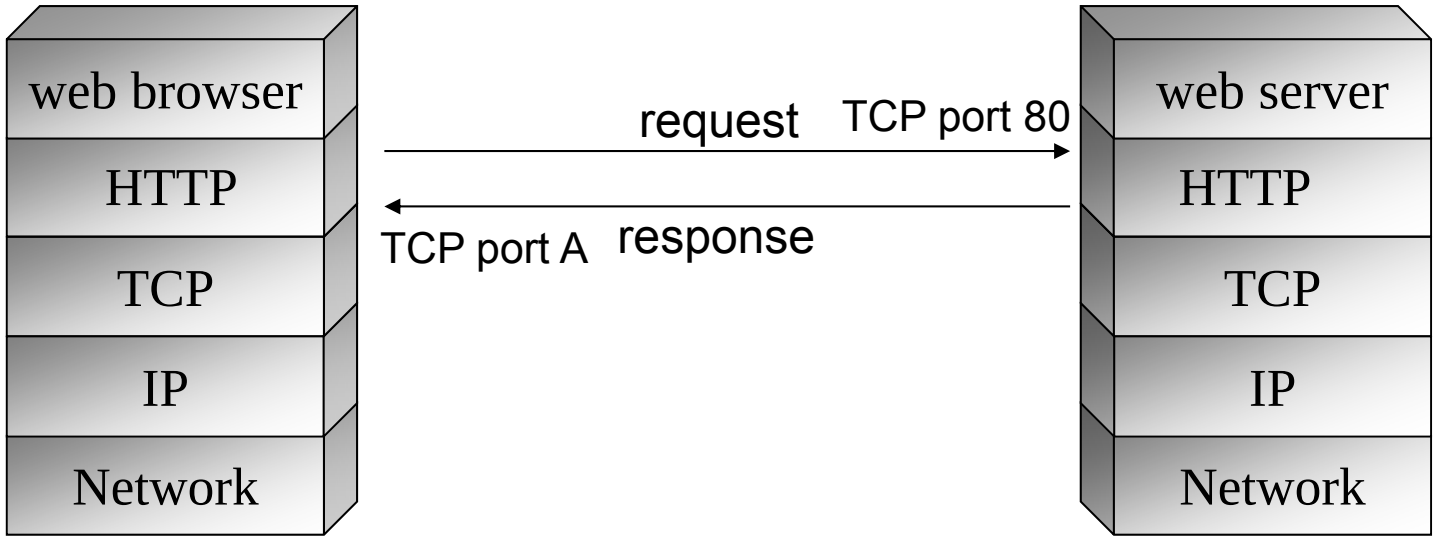
Components:

- A protocol name: http, rtp, rtsp
- ://
- A server domain name or server IP address
- A path to a resource (an HTML file or a CGI script)



HyperText Transfer Protocol (HTTP)

- Application layer protocol to distributes information in the WWW
- Based on the client/server architecture
  - HTTP client (web browser): sends a request to a server for a file
  - HTTP server (web server): well-known port number 80, responds with the requested file if it is available
  - A single TCP connection is used
- Stateless: server maintains no information about past client requests



HTTP Messages

- *English-based* and flexible, not *code-based* as lower layer protocols
- Components of an HTTP message:
  - A start-line
  - Optional headers, each has a header name and a value
  - A blank line (a "\r\n" only)
  - The requested file or other data in an HTTP response.

HTTP Request Message

Request Line:

- **Request Type**
- **URL**
- **HTTP version**

Optional Headers

- **Header name**
- **Value**

A blank line

```
GET /usage/try1.htm HTTP/1.1\r\n
Accept: image/gif, image/jpeg, */*\r\n
Accept-Language: en-us\r\n
Accept-Encoding: gzip, deflate\r\n
User-Agent: Mozilla/4.0 (Red Hat Linux)\r\n
Host: 128.238.42.129\r\n
Connection: Keep-Alive\r\n
\r\n
```

The Request Type defines methods in messages

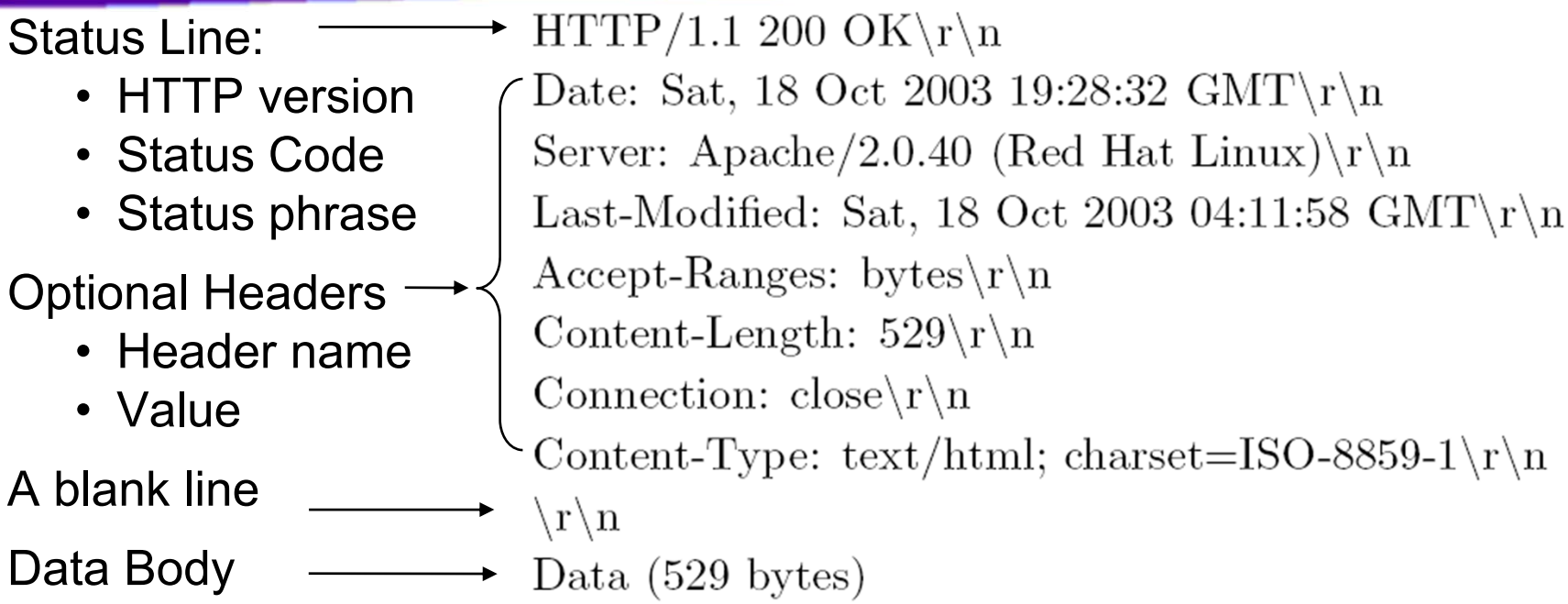
GET, HEAD – retrieve a full document or some info about a document from the server

PUT, PATCH – provide a new/replacement document or a list of difference to implement in an existing document to the server

COPY, MOVE, DELETE – copy, move, or delete a document

… …

HTTP Response Message



- The Status Code is similar to those in the FTP and the SMTP protocol with 3 digits
- The Status Phrase explains the status code such as continue, switching, OK, accepted, no content, multiple choices, bad request, unauthorized, forbidden, not found, internal server error, service unavailable, … …

HTTP TCP Connections

Client	Server
• First establishes a TCP connection to the server	• Always waits for TCP connection at port 80, and accepts connection
• Sends an HTTP request containing URL indicating the wanted object (an HTML file)	• Sends the HTTP response
• Processes the received HTML file to identify the embedded objects with respective URL	– May terminate the TCP connection after the response
– Send request for each embedded object if the TCP connection is terminated	

- In HTTP/1.0, the client establishes a TCP connection for each request, not efficient for a file with many embedded objects from the same URL
- In HTTP/1.1, Persistent Connections are supported
  - All embedded objects are sent through the same TCP connection established for the first request
  - Both the client and server have to enable the persistent connection feature

Internet Content Tsunami

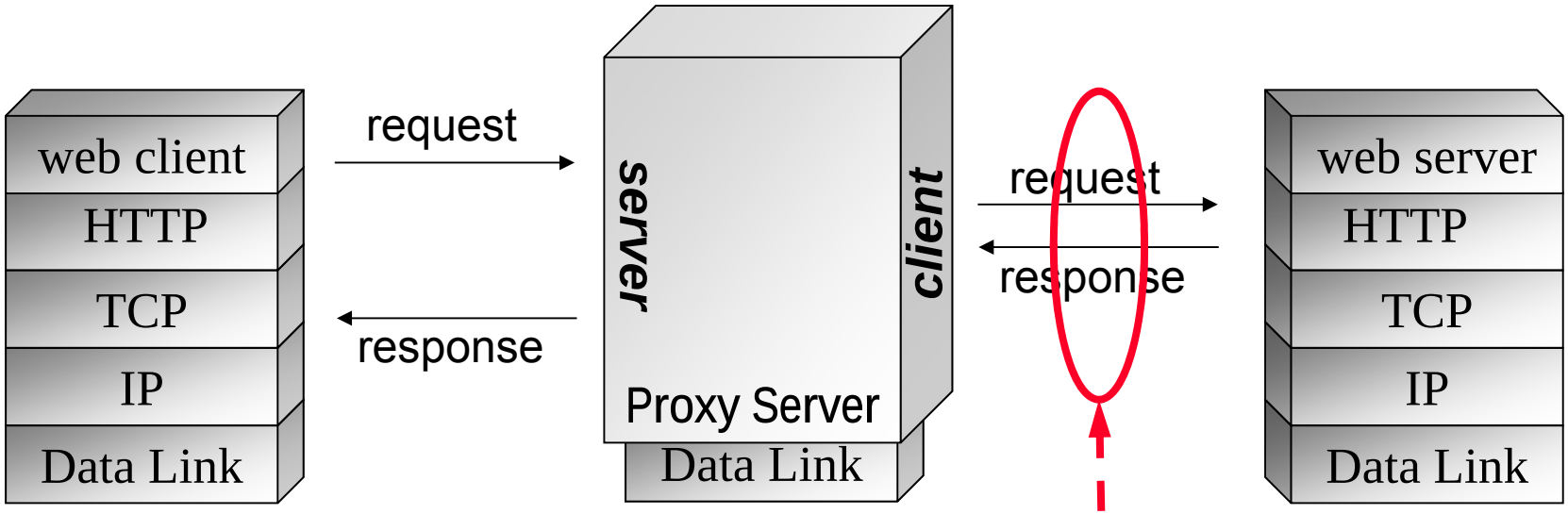
Internet content grows in many dimensions

- More content producers
- More content consumers
- More content centric applications
- Higher per-content data volume

Solutions

- Server Farms – But does not address end-to-end latency problem
- Forward Caching (put content close to consumers)
  - Large percentage of HTTP objects are not cacheable
  - Content providers can't track how content data get distributed
- Content Distribution Network

HTTP Proxies & Caching



- **Proxy server** acts as both a client and server
  - receiving client's initial requests, translating requests, passing requests to other servers
- Proxies can be used with firewalls to block undesired traffic
- **Caching** feature of a Web proxy server reduces network traffic by saving recently viewed pages on the disk driver

Content Delivery Network (CDN)

Advantages:

- Reduced root server load and end-to-end latency
- Traffic loads distributed and routed around congested networks

A CDN uses an overlay network to provide a “rendezvous” between end-user and content providers

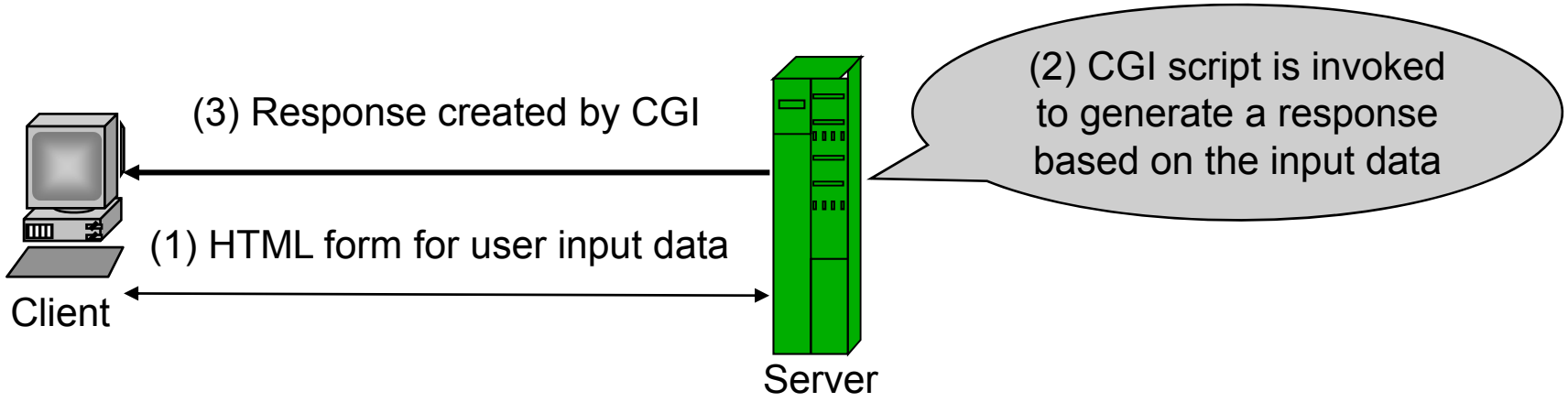
- Management of content requests/responses
- Store content to the edge of Internet, close to content consumers
  - The most popular content gets moved to ‘edge’ servers closer to the subscriber to reducing delay and transmissions costs
- Content networking between 1000’s of caches and load balancers
  - Modified DNS’, request routing systems, IP backbone facilities
  - CDN specific algorithms to move content based on user behavior, network conditions and sometimes anticipated demand.

Common Gateway Interface (CGI)

- CGI is a technology that creates and handles dynamic documents
- Dynamic web pages provide a two-way communication between web clients and servers: on-line opinion poll, e-commerce
- CGI uses two files
  - An HTML form, for a user to input data
  - A CGI script,
    - > processing user input data and generating a response dynamically
    - > Any program that can read input from the STandarD INput (STDIN) and write output to the STandarD OUTput (STDOUT) can be used as CGI script

CGI Operations

- A user downloads the HTML form (consisting of text inputs, checklists, and buttons), inputs data, and submit the data
- A web server invokes the CGI script
- The web server returns the CGI response to the client

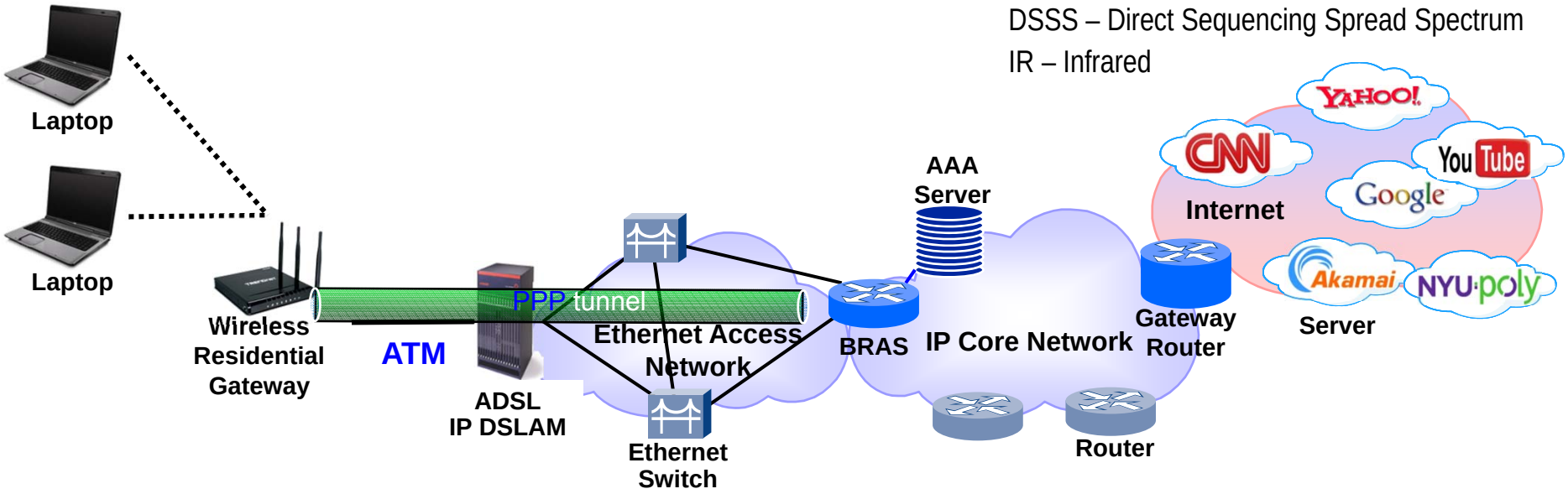


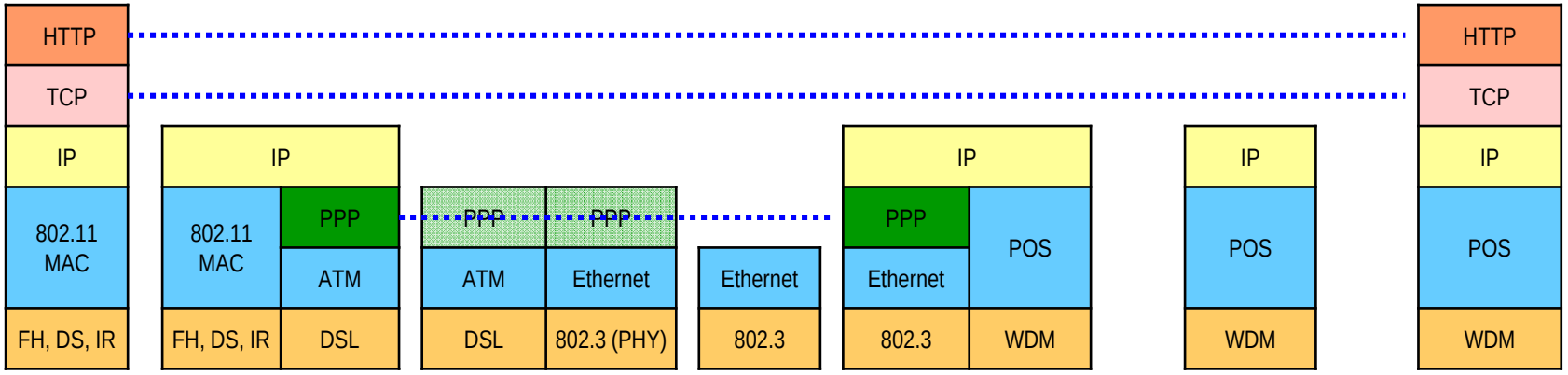
Apache Web Server

- The Apache server is the most popular web server in the Internet, according to the web server survey from Netcraft.
- Apache server is an open source software, included in both the Red Hat Linux 9 and Solaris 8 installation CDs.
- Apache is a process-based web server
  - Stable, other child processes won't be affected when one child process crashes
  - Scalable, allowing more client requests to be processed simultaneously

IP Networking Example
- High Speed Internet Browsing

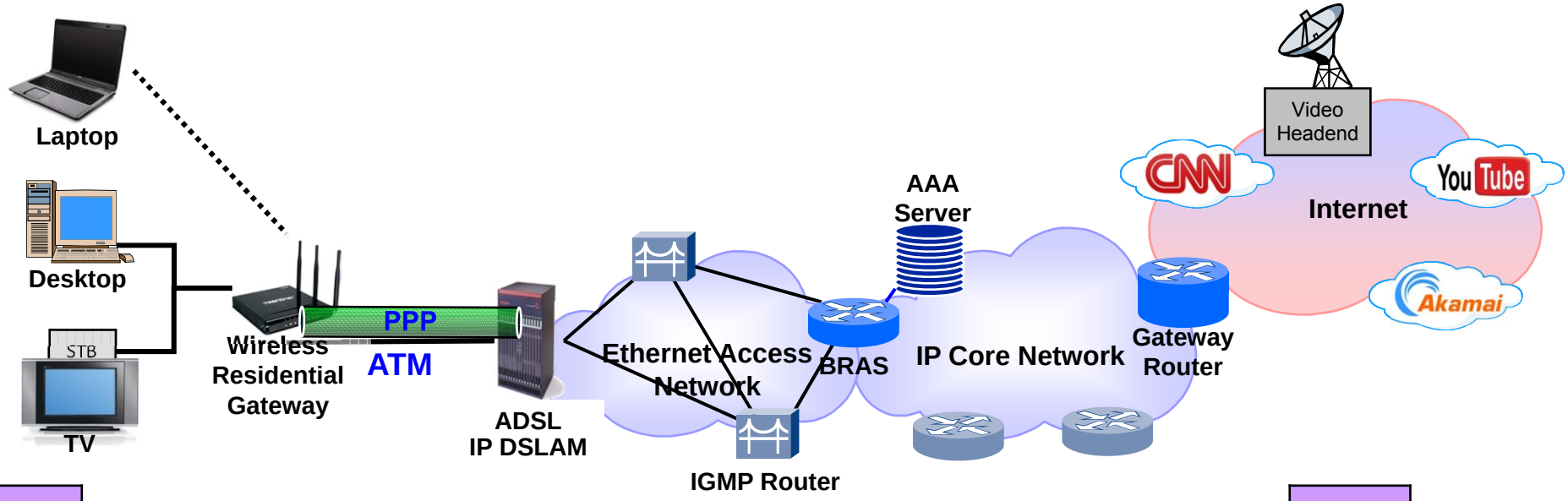
DSL – Digital Subscriber Line
DSLAM – DSL Access Multiplexer
BRAS – Broadband Remote Access Server
WDM – Wavelength Division Multiplexing
POS – Packet Over SDH/SONET
FHSS – Frequency Hopping Spread Spectrum
DSSS – Direct Sequencing Spread Spectrum
IR – Infrared

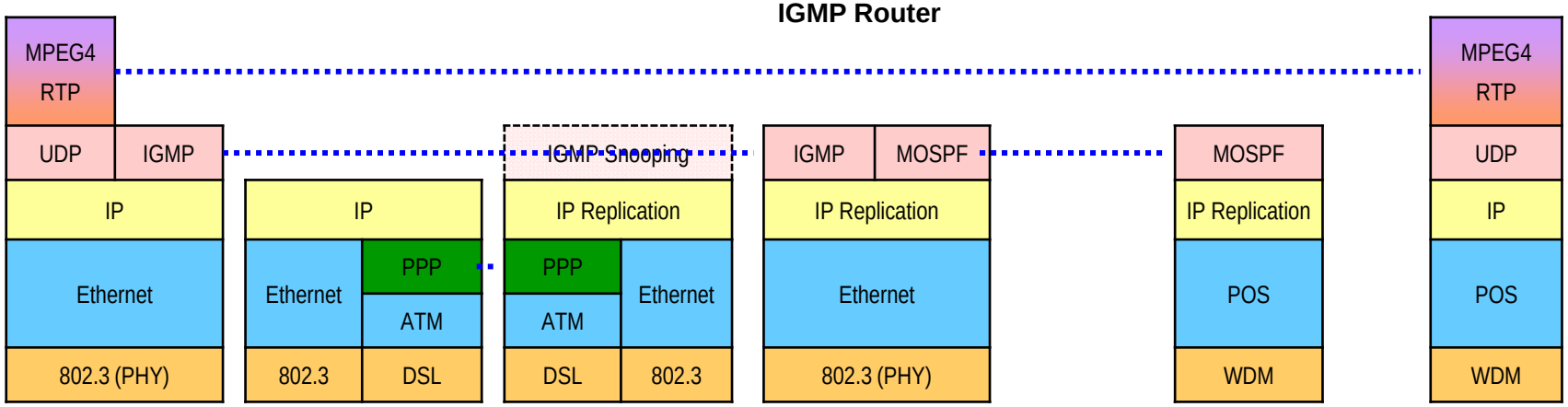




IP Networking Example
- IPTV Multicasting

DSL – Digital Subscriber Line
DSLAM – DSL Access Multiplexer
WDM – Wavelength Division Multiplexing
POS – Packet Over SDH/SONET
STB – Set Top Box





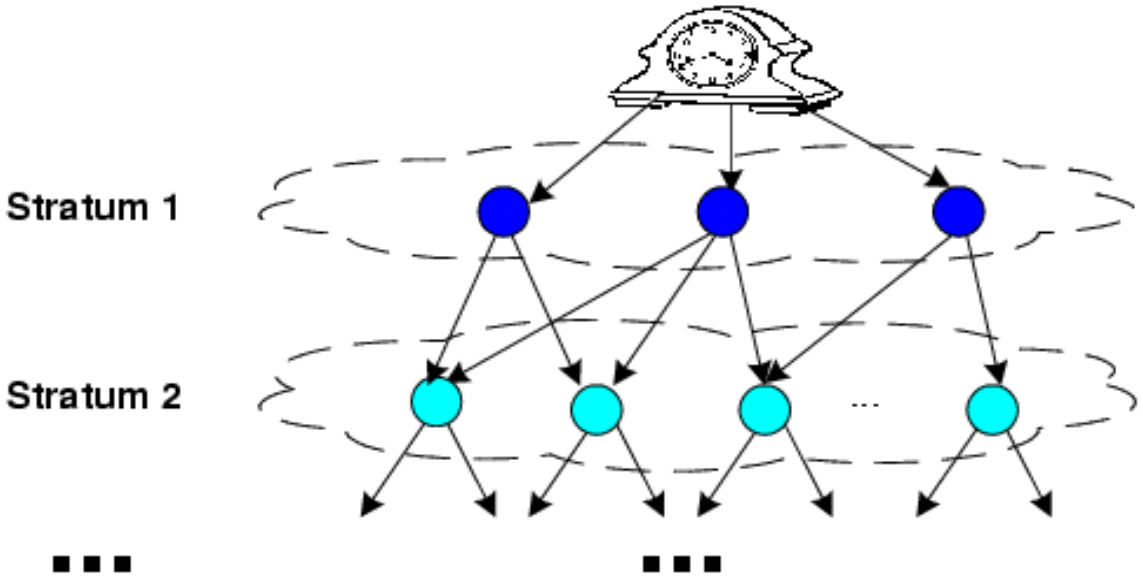
Network Time Protocol (NTP)

- Accurate timing is important in network design, management, security, and diagnosis.
- NTP is an application layer protocol, with UDP or TCP port 123, used to
  - Provide accurate timing in the network
  - Synchronize routers, hosts, and other network devices

NTP Timing Service

NTP timing service uses a hierarchical architecture organized into 16 stratums

- An NTP primary server, or stratum-1, is synchronized with a high precision clock
  - Over 300 valid stratum-1 servers
- About 175,000 hosts running NTP in the Internet, Each server chooses one or more higher stratum servers and synchronizes with them



NTP Operation Modes

Clients and servers can operate in the multicast or broadcast mode.

- Timing information is broadcast or multicast by the servers.
- A client can proactively poll the servers for timing information.

NTP client synchronize with a server in two ways

- Query time information from and synchronize to a remote NTP server, use *rdate* or *ntpdate*
- Synchronize with a remote server continuously and automatically, use *ntpd*

Dynamic Host Configuration Protocol (DHCP)

- DHCP, RFC 2131, is designed to dynamically configure TCP/IP hosts in a centralized manner from DHCP server.
- DHCP server maintains a collection of configuration parameters, such as IP addresses, subnet mask, default gateway IP address, to make a configured host work in the network.
- A DHCP client queries the server for the configuration parameters.
- The DHCP server returns configuration parameters to the client.
- Often use assigned UDP port numbers for BOOTP (BootStrap Protocol): 67 for DHCP server and 68 for DHCP client

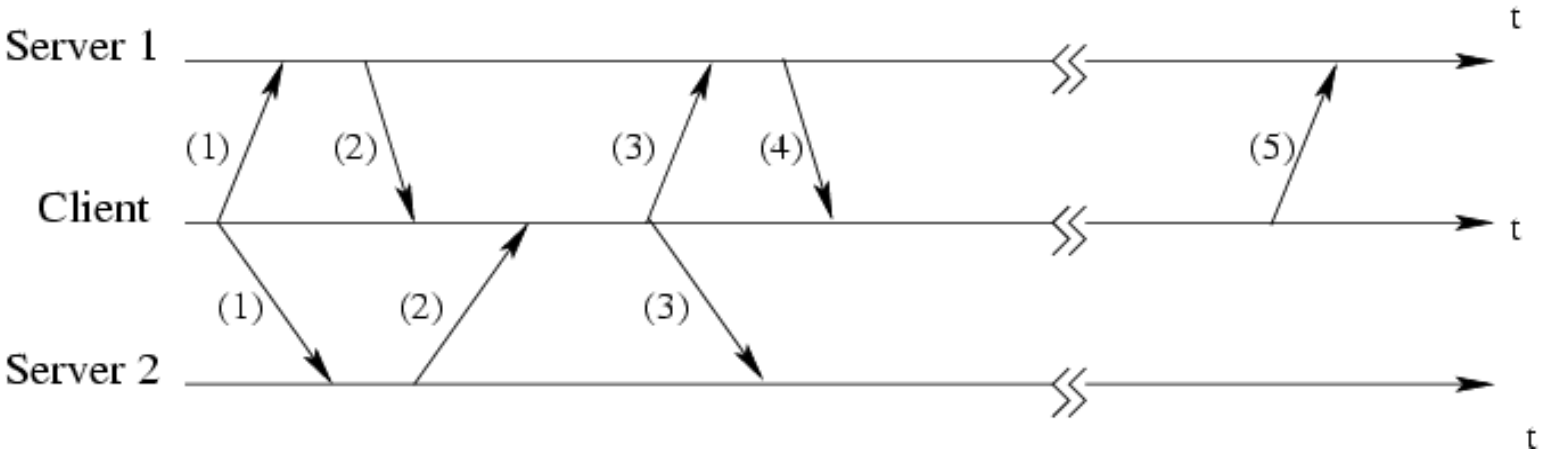
DHCP Network Parameters Assignment

- DHCP can provide persistent storage of network parameters for the clients
  - A client can be assigned with same set of parameters whenever it bootstraps, or is moved to another subnet
  - The DHCP server keeps a key-value entry for each client and uses the entries to match queries from the clients
  - The entry could be a combination of a subnet address and the MAC address (or domain name, host name, …) of a client
- DHCP can also assign configuration parameters dynamically
  - The DHCP server maintains a pool of parameters and assigns an unused set of parameters to a querying client
  - A DHCP client leases an IP address for a period of time. When the lease expires, the client may renew the lease, or the IP address is put back to the pool for future assignments

DHCP Operations (1/4)

When two DHCP servers are used

1) A client first broadcasts a DHCP DISCOVERY message on its local physical network during bootstrapping.
   - The message has 0.0.0.0 as the source IP address.
   - The message may be forwarded by relay agents to servers in other physical networks.
2) Each server may respond with a DHCPOFFER message with an available network address in the Your IP Address field.

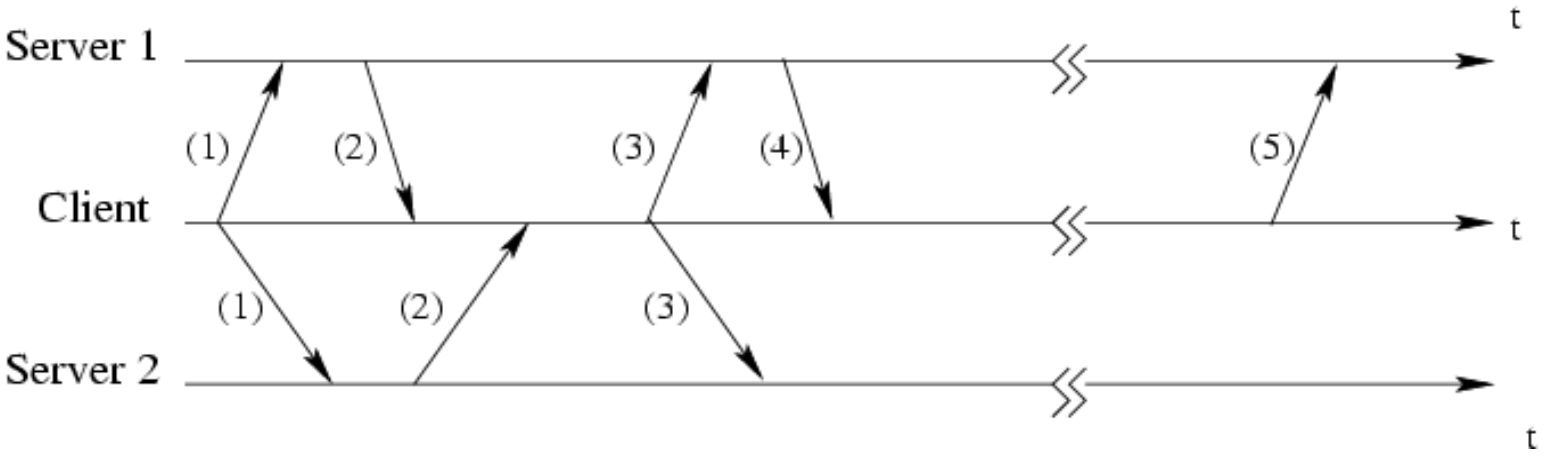


DHCP Operations (2/4)

When two DHCP servers are used

3) The client may receives more than one DHCPOFFER messages.

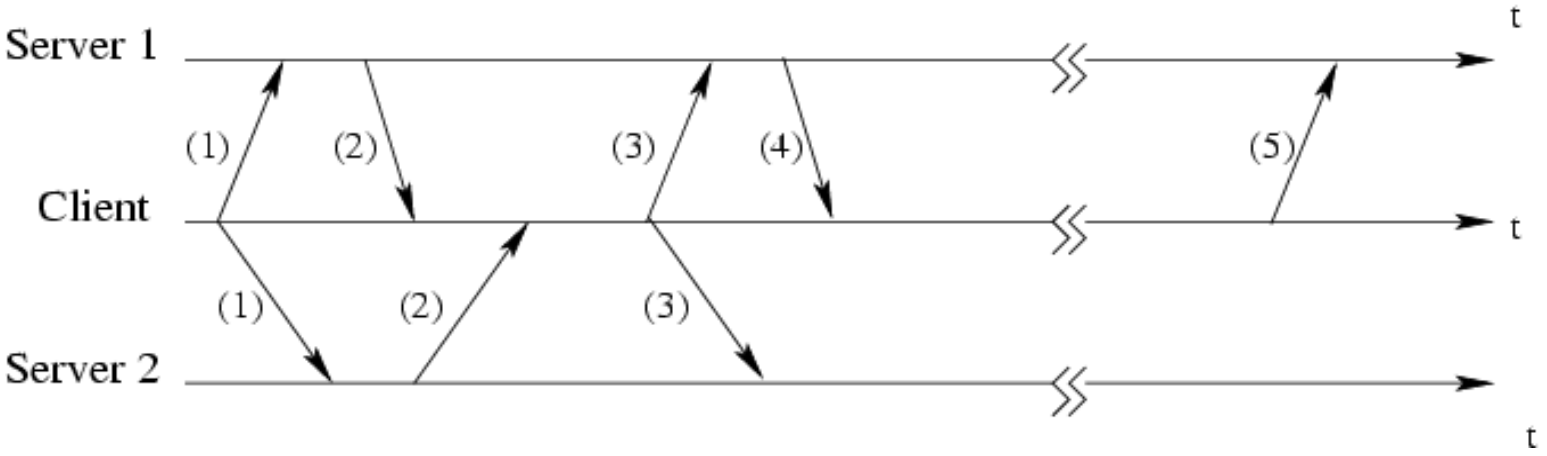
- It chooses one server from all responding servers based on the offered IP address and the lease duration.
- The client then broadcasts a DHCPREQUEST message with the Server Identifier option to indicated the selected server.



DHCP Operations (3/4)

When two DHCP servers are used

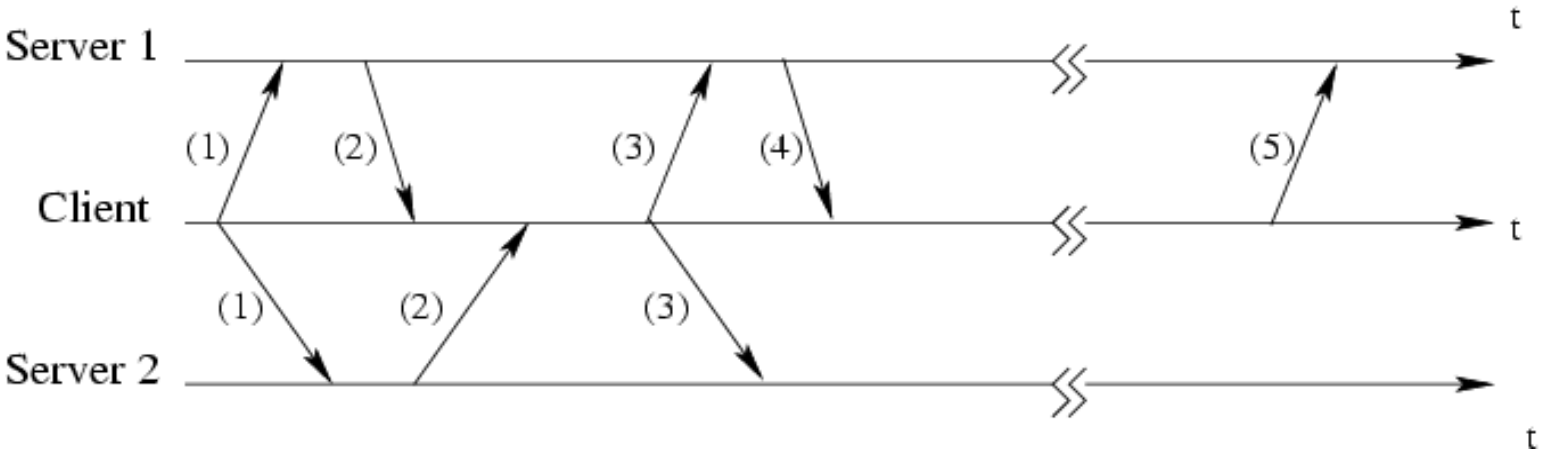
4) When the DHCPREQUEST message is received, only the chosen server responds with a DHCPACK message carrying a full set of configuration parameters to the client.
   - When the client receives, it checks the parameters and configures its TCP/IP modules using the parameters.
   - The message specifies the duration of the lease. When the lease expires, the client may ask the server to renew it. Otherwise, the address will be put back in the pool or assigned to other hosts.



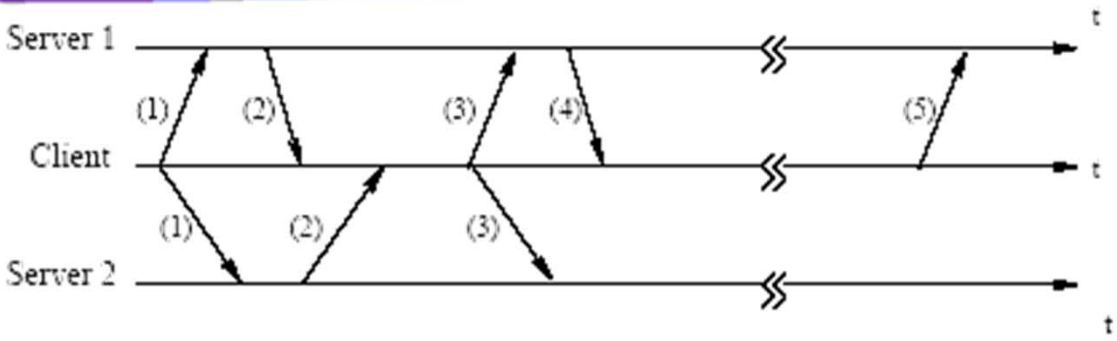
DHCP Operations (4/4)

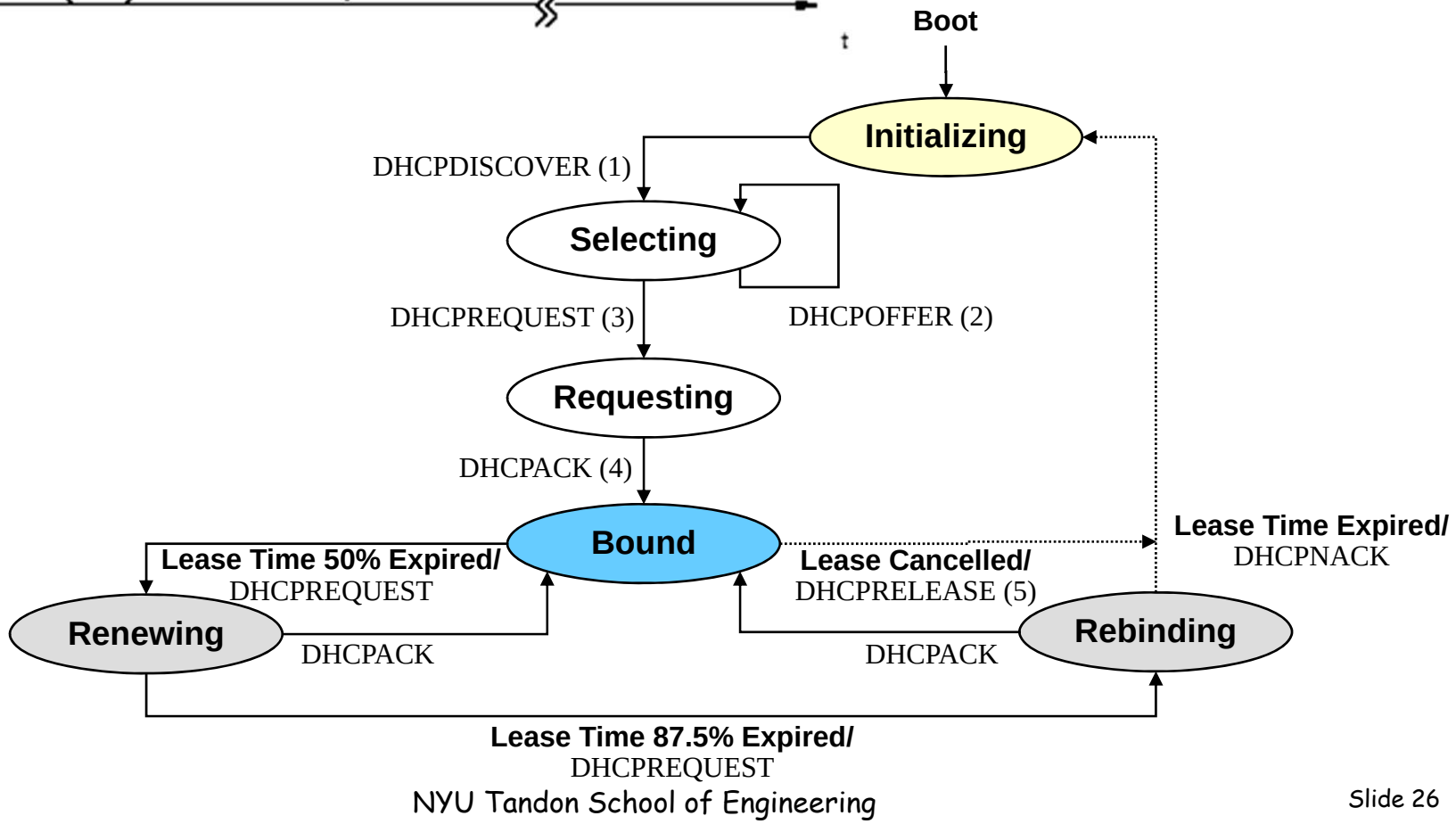
When two DHCP servers are used

5) The client may send a DHCPRELEASE message to the server to relinquish the lease on the network address.



DHCP Client Transition States

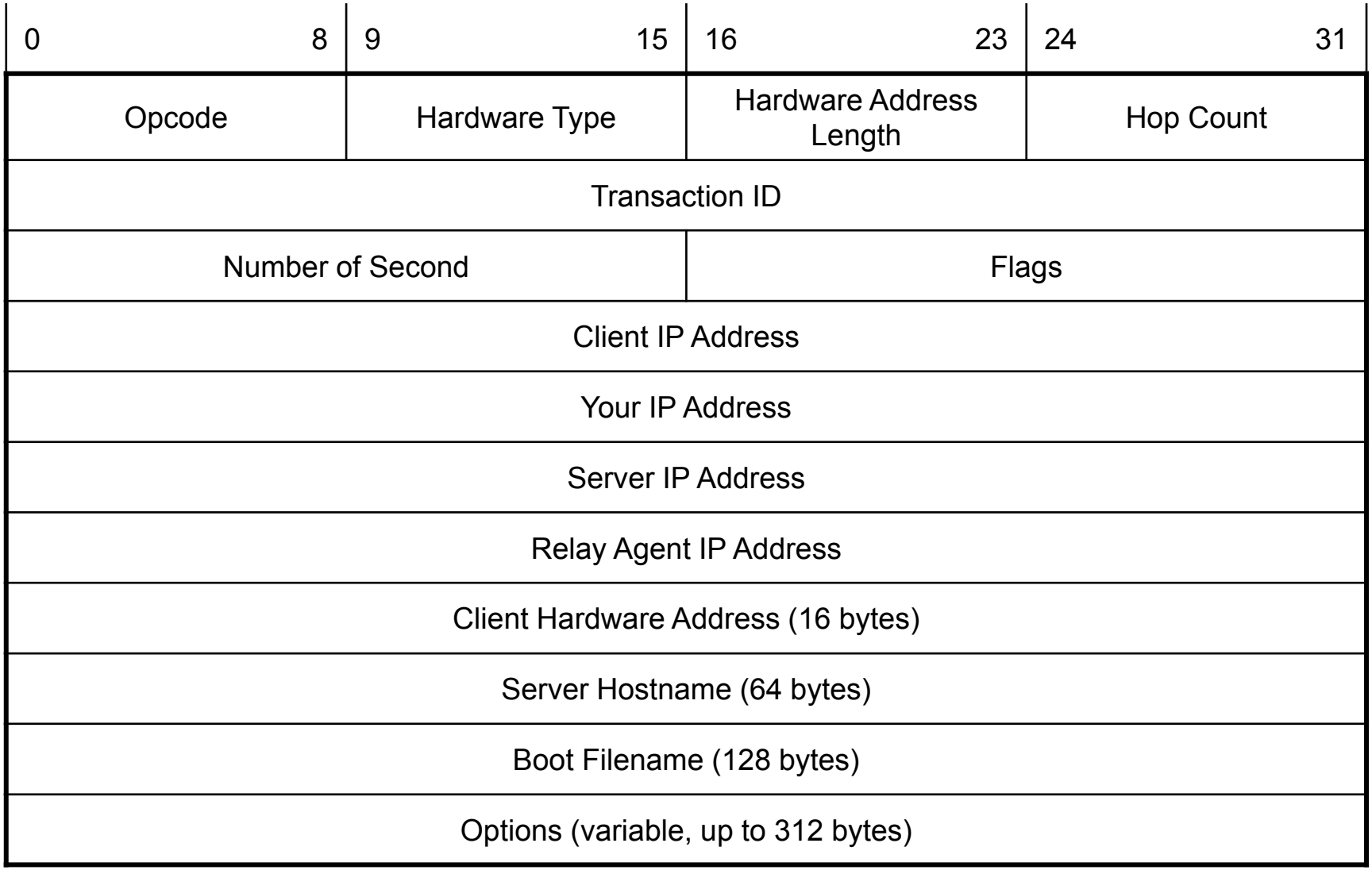




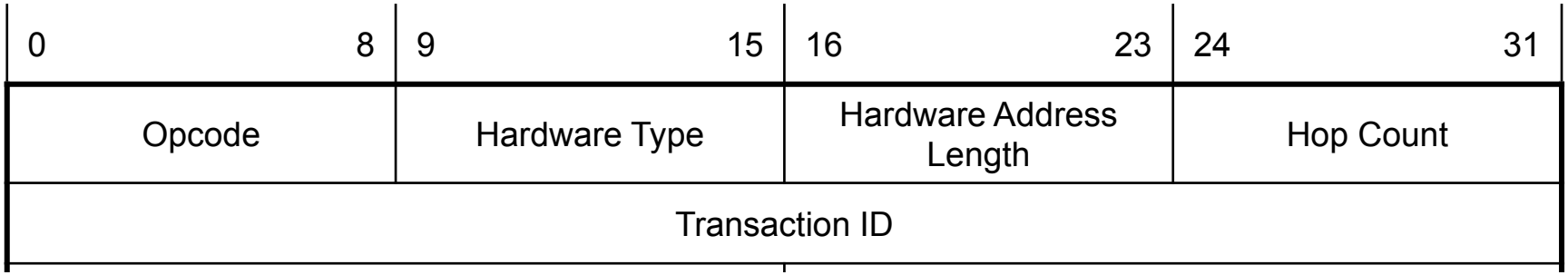
DHCP Server Address Allocation Mechanisms

- Three different address allocation mechanisms
  - Manual Allocation: DHCP communicates a pre-allocated IP address by an administrator to a specific single device
  - Automatic Allocation: DHCP automatically assigns an IP address permanently to a device selected from a pool of available addresses
  - Dynamic Allocation: DHCP assigns an IP address from a pool of addresses for a limited period of time chosen by the server, or until the client tells the DHCP server that it no longer needs the address

DHCP Message Format

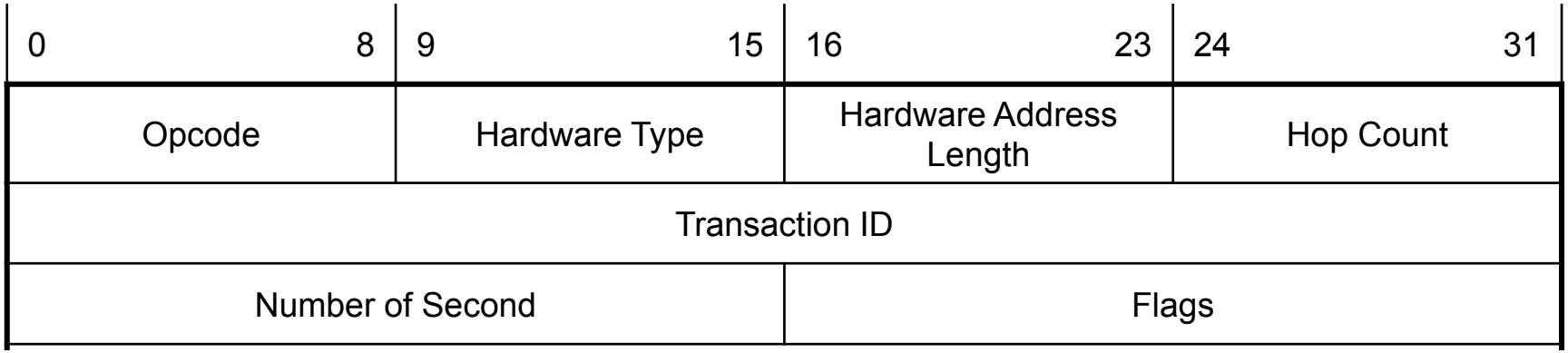


DHCP Message Fields (1/4)



- Opcode: 1 – a boot request from client; 2 – a boot reply from server
- Hardware Address Type
  - The value is 1 for an Ethernet MAC address
  - The values are defined in the “Assigned Numbers” RFC
- HW address length
  - The length of the hardware address in byte
- Hop count
  - Optionally used by relay agents
  - A DHCP relay agent is a host or router that forwards DHCP messages between DHCP clients and servers

DHCP Message Fields (2/4)



- Transaction ID
  - Randomly assigned to link requests and replies between a client and a server
- Number of seconds
  - Elapsed time in seconds since the client began an address acquisition or renewal process
- Flags
  - Broadcast flag, the leftmost bit. Used when a client cannot receive a unicast IP datagram before its interface is configured
  - Remaining 15 bits must be 0 (reserved for future use)

DHCP Message Fields (3/4)

- Client IP address: use when the client is in BOUND, RENEW, and REBINDING state and can respond to ARP requests
- Your IP address: client's IP address from DHCP server
- Server IP address: the IP address of the next server to use in bootstrap
- Relay agent IP address: used when booting via a relay agent

Client IP Address
Your IP Address
Server IP Address
Relay Agent IP Address
Client Hardware Address (16 bytes)
Server Hostname (64 bytes)
Boot Filename (128 bytes)
Options (variable, up to 312 bytes)

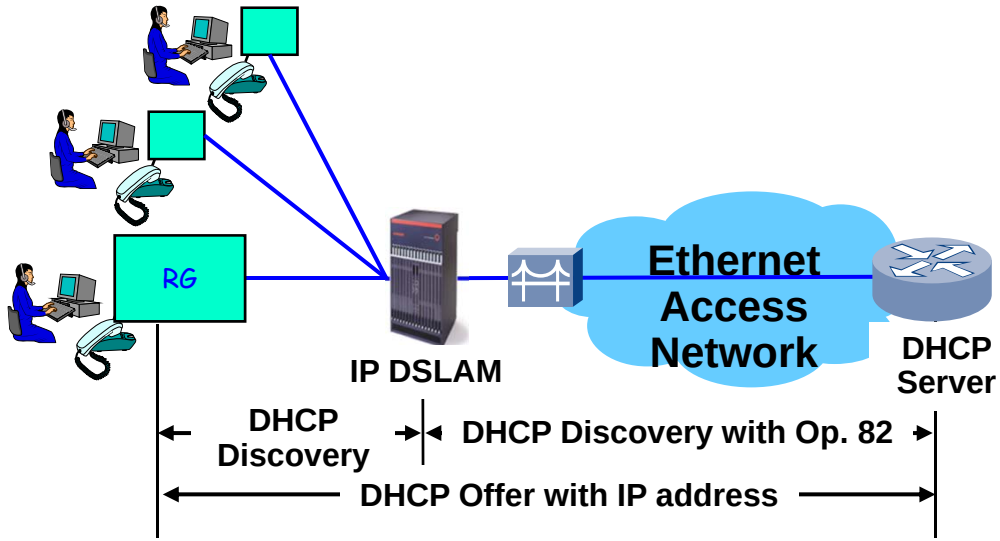
DHCP Message Fields (4/4)

- Client HW address
  - The hardware address of the client
  - For an Ethernet address, the first 6 bytes are filled and the remaining bytes are set to 0
- Server hostname: Hostname of the DHCP server
- Boot filename: Use in a DHCPOFFER message to specify the fully qualified, null terminated path name of a file to bootstrap from
- Options: optional vendor specific field

Client Hardware Address (16 bytes)
Server Hostname (64 bytes)
Boot Filename (128 bytes)
Options (variable, up to 312 bytes)

DHCP Relay Information Option (Option 82 per RFC3046)

- In broadband access service, how to assign IP address from a central DHCP server to each client device based on it's location?
- To get authenticated with a valid IP address, the DHCP client in a Residential Gateway (RG), ex. DSL/cable modem, will send a DHCP request to an ISP BRAS.
- A network access node, ex. an IP DSLAM, aggregates traffic from many users.
  - –Deployed closer to users with point-to-point connection to each user.
  - –Need to insert an identifier in all DHCP requests from end-user that allows the ISP to authenticate and control the rights for assigning IP addresses to the end-user.
- Network requirements
  - –Standard DHCP client in RG
  - –Switch/router (DSLAM) implementing Relay Agent with Option 82 support
  - –DHCP server (here BRAS) with Option 82 Support

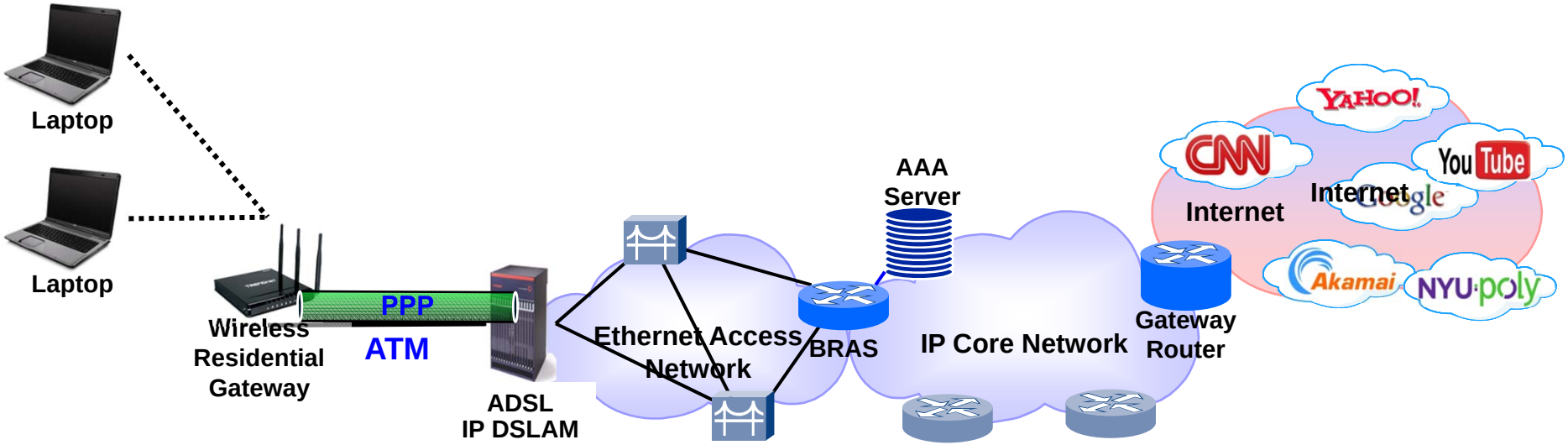


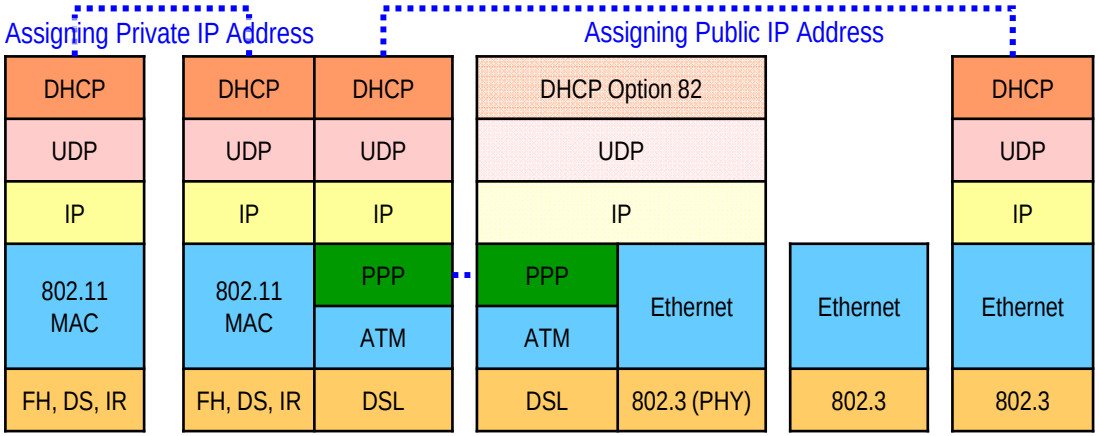
DHCP Configuration

An example of a DHCP server configuration file

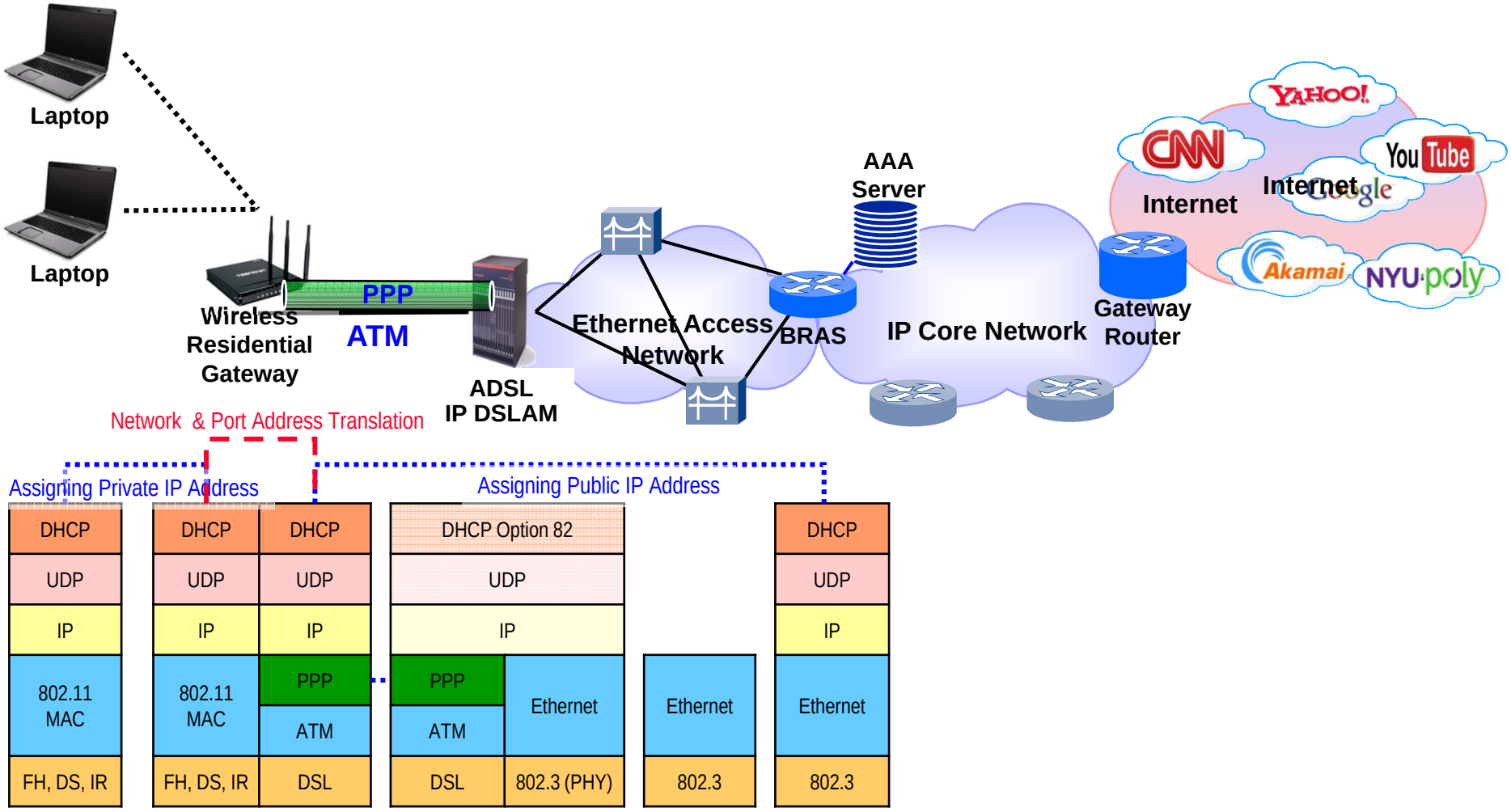
```
1   # Sample /etc/dhcpd.conf
2   default-lease-time 600;
3   max-lease-time 7200;
4   option subnet-mask 255.255.255.0;
5   option broadcast-address 128.238.66.255;
6   option routers 128.238.66.1;
7   #option domain-name-servers 128.238.2.38, 128.238.3.21;
8   #option domain-name "poly.edu";
9
10  subnet 128.238.66.0 netmask 255.255.255.0 {
11      range 128.238.66.111 128.238.66.112;
12  }
13
14  host apah {
15      hardware ethernet 08:00:20:79:e9:9f;
16      fixed-address 128.238.66.110;
17  }
```

IP Networking Example
- Dynamic Host Configuration Protocol





IP Networking Example
- Network Address Translation & Port Address Translation



Network Address Translator (NAT)

- Mapping one set of IP addresses (usually private) to another set of IP addresses (usually public).
  - Public Internet address are assigned by IANA.
  - Private address: 10/8, 172.16/12, and 192.168/16 are recommended to be used
- Used for
  - Security: internal network is invisible to the outside
  - Reuse IP address: Share a small pool of IP addresses with a large number of hosts

Private IP Address

- A Private Network is designed to use mainly inside an organization
  - Intranet is a private network (LAN) that its access is limited to the users inside the organization
  - Extranet is also a private network (LAN) like the intranet but it allows some users outside the organization to access the network
- A number of blocks in each class are assigned for private use
- Private IP addressed are not recognized globally
- Private IP addresses are used either in isolation or in connection with Network Address Translation (NAT) technique

Class	NetID	Block
A	10.0.0	1
B	172.16 to 172.31	16
C	192.168.0 to 192.168.255	256

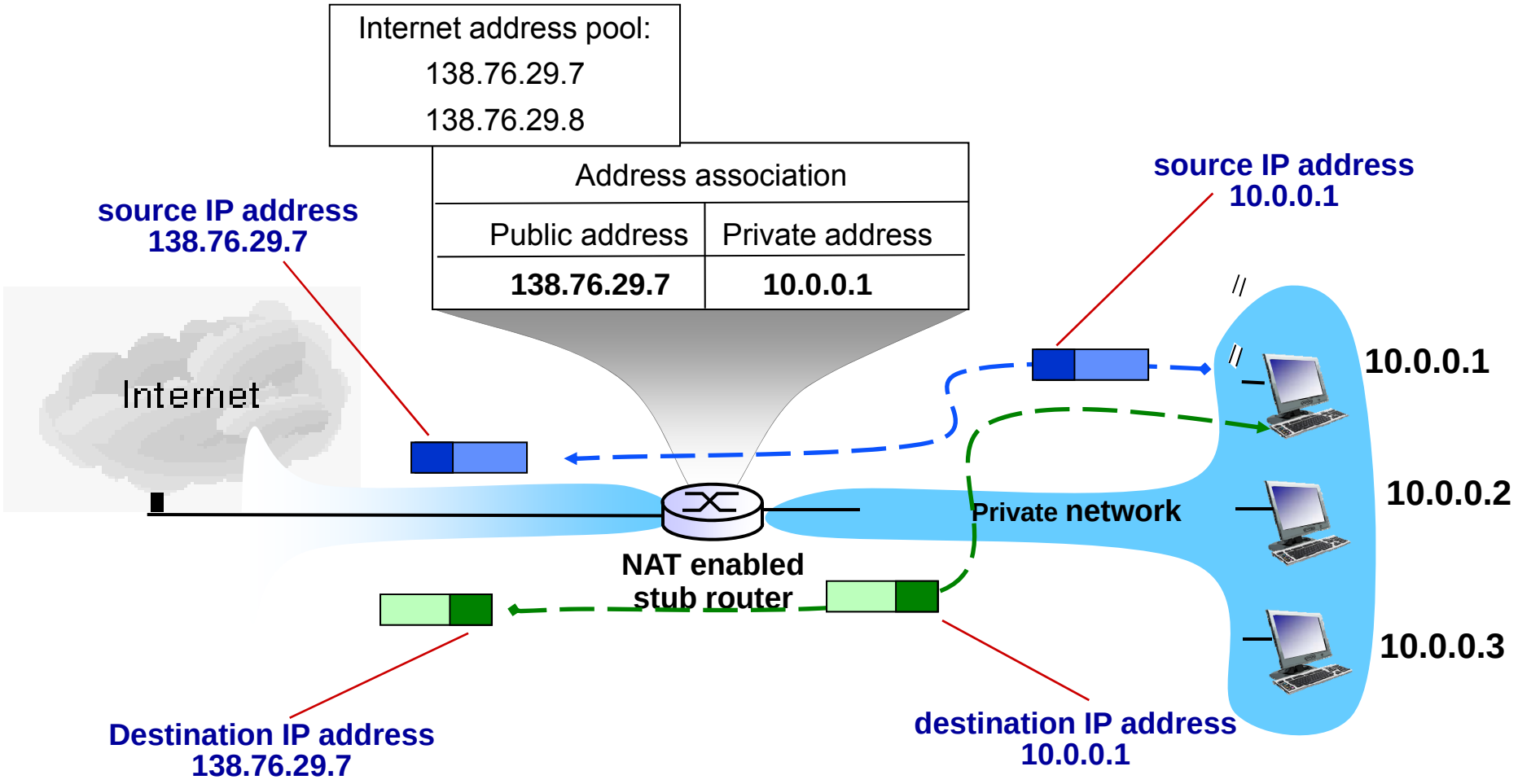
How NAT works

- Two type of networks:
  - Global network/External network: the Internet
  - Private network/Internal network
- NAT is carried between two networks: Internet and private network, or two private networks
- NAT is carried at stub routers:
  - Perform mapping/translation of the two set of address
  - ICMP message payload is also translated
- Two types of assignment:
  - Static address assignment: static mapping of public/private addresses
  - Dynamic address assignment: mapping is dynamic, based on requests and chosen from the available address pool

Three Step Operation

1. Address binding:
   - A private node IP address is associated with an external address.
   - The router maintains a table recording the associations
2. Address lookup and translation:
   - Outgoing packets: source IP address is changed to the corresponding external address
   - Incoming packets: destination IP address is changed to the corresponding private address
3. Address unbinding:
   - When the session is over, the entry is removed from the table and the external address is released for other connections to use

NAT: A Simple Example



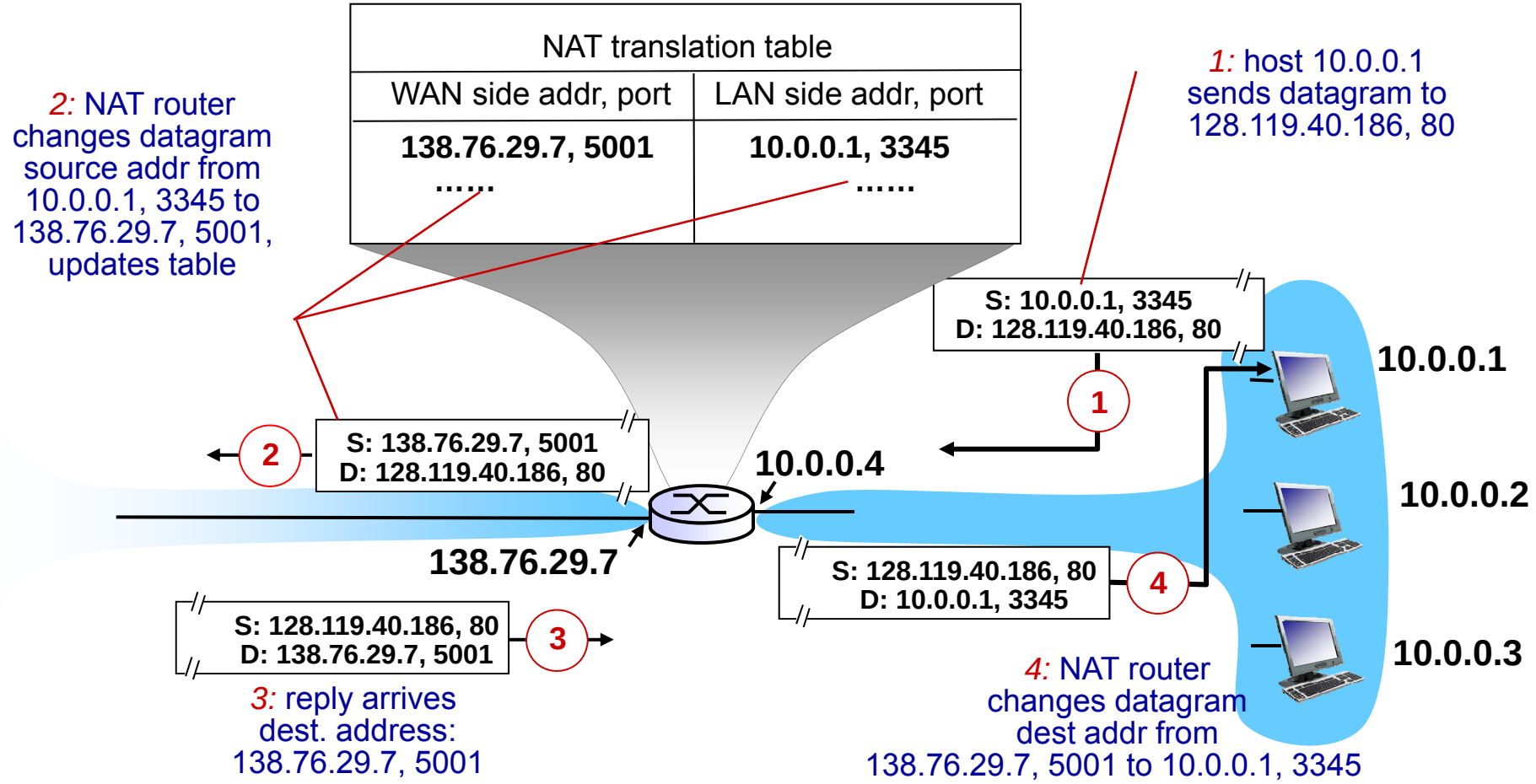
Port Address Translation

- Port Address Translation (PAT) extends the notion of translation one step further by also translating transport identifiers (e.g., TCP and UDP port numbers, ICMP query identifiers).
- Note that PAT can be combined with Basic NAT so that a pool of external addresses are used in conjunction with port translation.
  - PAT allows a set of hosts to share a single external address.
- The network devices that support both NAT and Pat are often referred as NAPTs (Network Address and Port Translators)

How PAT Works

- The association table: maintains mapping of IP address, port number, and ICMP query ID
- Outbound packets: translate the source IP address, source port number, ICMP query ID and related fields such as IP, TCP, UDP and ICMP header checksums.
- Inbound packets: the destination IP address, destination transport identifier and the IP and transport header checksums are translated.

NAT: An Example with Single External Address



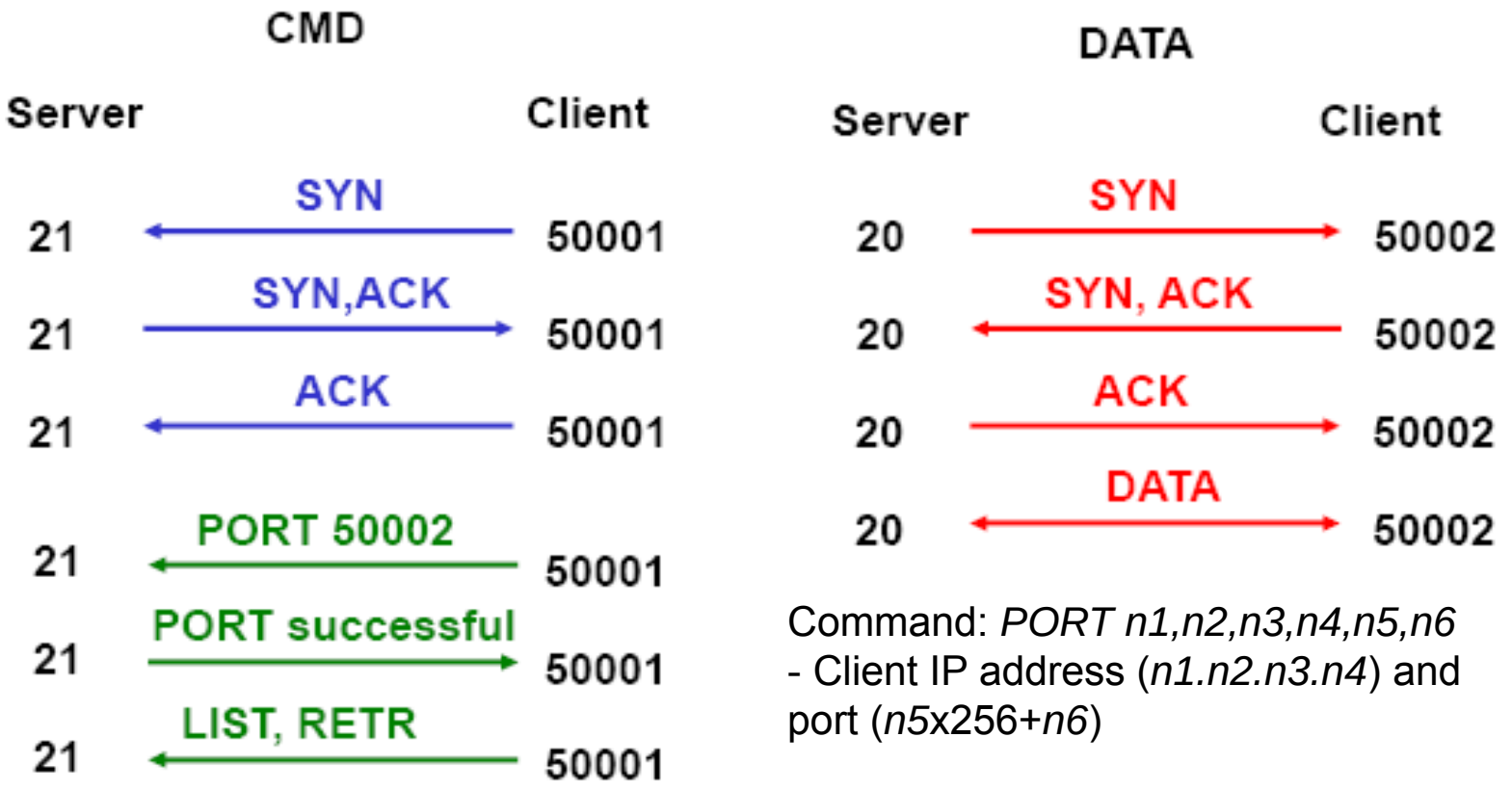
Other Issues for NAT

- Change the original IP packet embedded in the payload of the ICMP error message.
  - IP address and port numbers
  - Checksums
- All ICMP error messages (with the exception of Redirect message type) will need to be modified, when passed through NAT:
  - Destination-Unreachable,
  - Source-Quench,
  - Time-Exceeded,
  - Parameter-Problem.

Other Issues for NAT (cont'd)

- Computation intensive at the stub routers:
  - Search the headers, table lookup, change the headers, checksum computation, …
- Does not support applications with IP-address content: e.g., early SNMP versions
- Does not support applications with inter-dependent control and data sessions: e.g. H.323, RTP, FTP
  - FTP: the client uses PORT command in the control TCP connection to identify the IP address and port number for the data connection
- Solution: use special designed Application Level Gateways (ALG) with additional security benefits:
  - Restricts outside sessions' access to internal hosts
  - Can be used in conjunction with a firewall to filter unwanted traffic

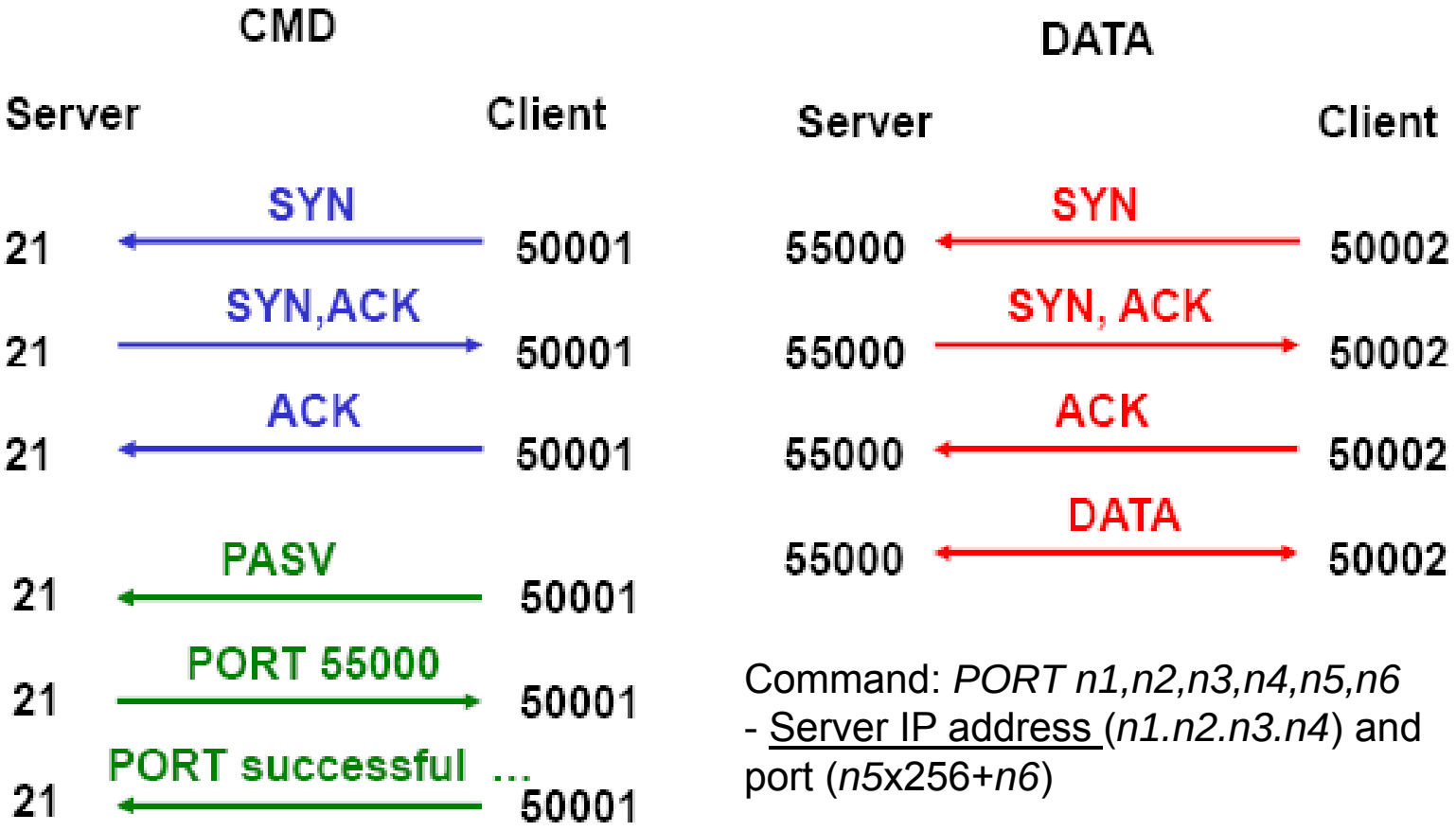
FTP Active Mode



Command: *PORT n1,n2,n3,n4,n5,n6*
- Client IP address (*n1.n2.n3.n4*) and port (*n5*x256+*n6*)

Active FTP won't work if client PORT command doesn't translated by NAT

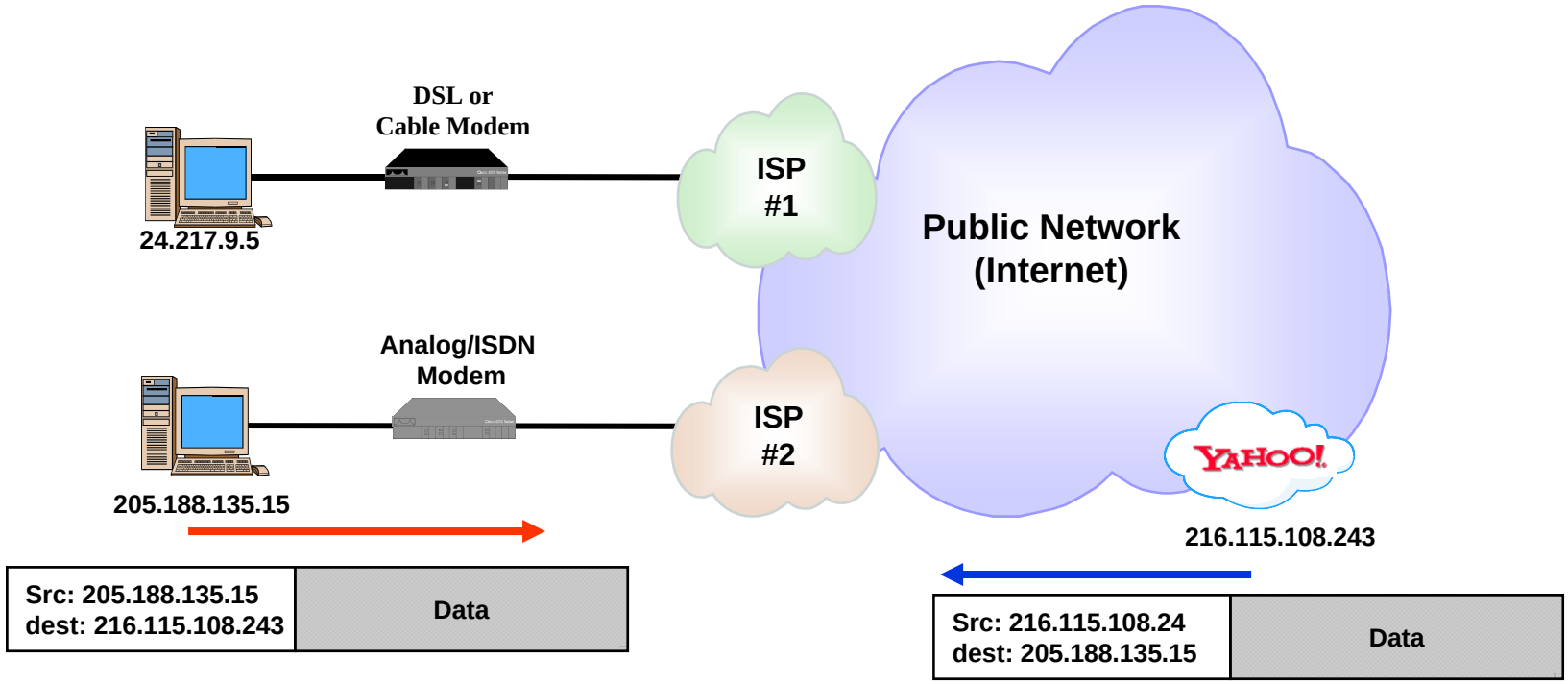
FTP Passive Mode



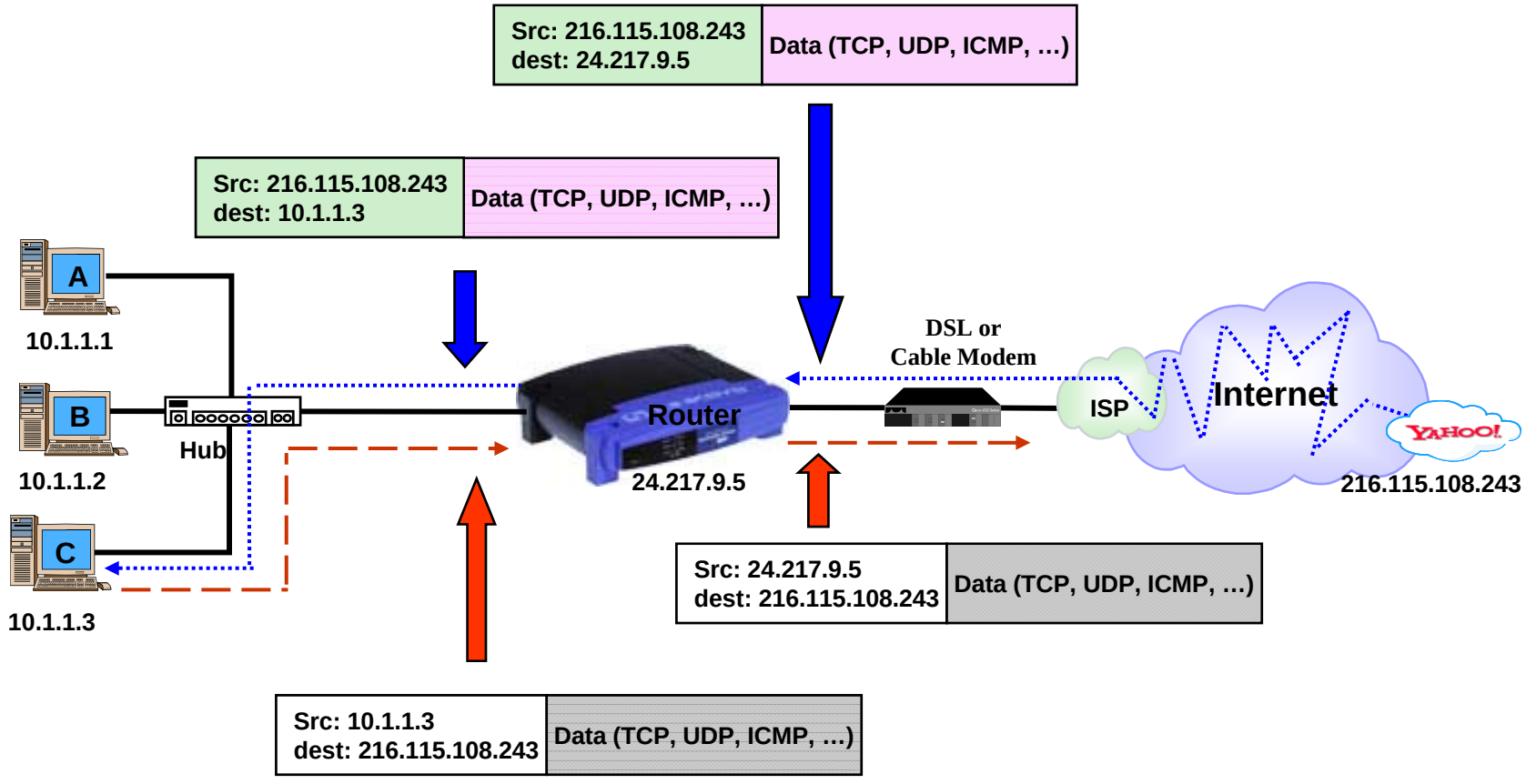
Command: *PORT n1,n2,n3,n4,n5,n6*
- Server IP address (*n1.n2.n3.n4*) and port (*n5*x256+*n6*)

Often used where the client is behind a firewall and unable to accept incoming TCP connections

Internet Access without NATing

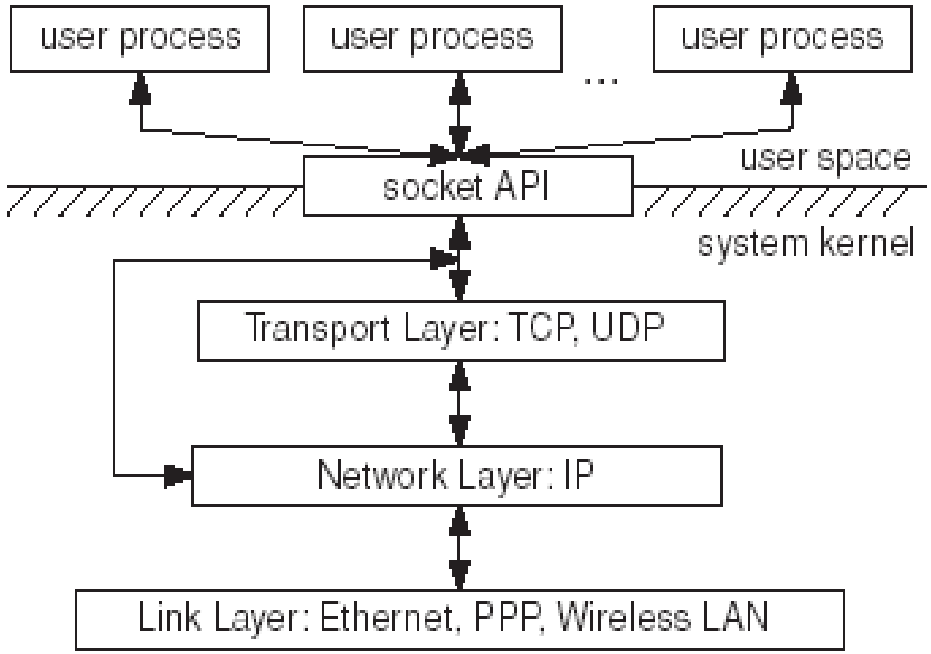


Home LAN Internet Access with NATing



Socket Programming in a Nutshell

- Most of the applications are implemented using the socket Application Programming Interface (API).
- TCP/IP protocols are implemented in the system kernel.
- User applications can use the TCP/IP service through the socket API.



Socket Programming for Applications

- Each participating process in the application should create a socket, containing the IP address and a unique port number.
- The application process can use the socket functions for sending or receiving data.
- Three types of sockets for applications to use:
  - TCP sockets, used to create a TCP connection
  - UDP sockets, used to provide the datagram service
  - Raw sockets, for applications to bypass the transport layer protocols and use the IP datagram service directly
- Client-server architecture is used in socket programming.

Network Management

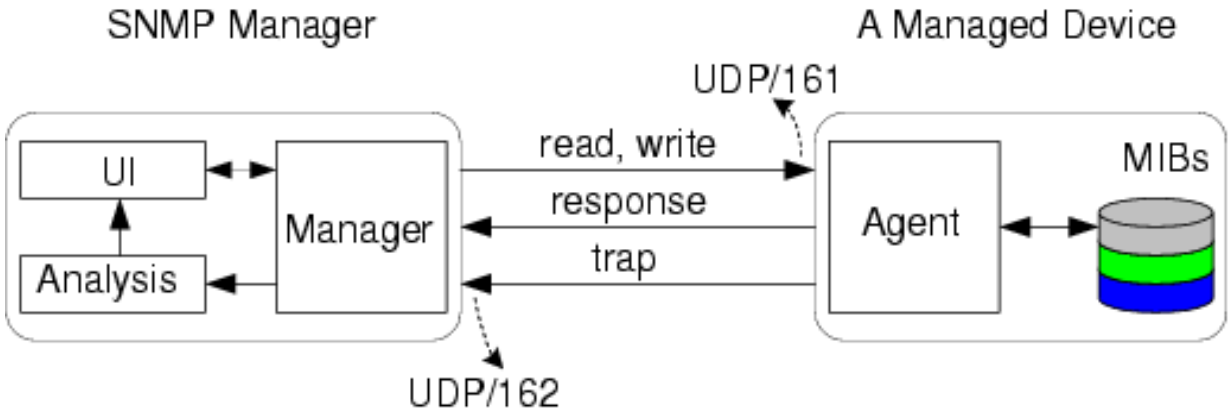
Network administrators need to

- Collect statistics from a device to see if it works properly (element management)
- Monitor network traffic load on routers to see if the load is appropriately distributed (traffic monitoring)
- Go through collected information to identify the cause when a network failure occurs (trouble shooting)

Simple Network Management Protocol (SNMP)

Simple Network Management Protocol (SNMP) is an application layer protocol for exchange management information between network devices

- Each Managed Device, a host or a router, maintains a number of Management Information Bases (MIBs)
- Each managed device has an SNMP Agent to provide interface between MIBs and an SNMP Manager
- An SNMP manager, usually implemented in Network Management System, can work with multiple SNMP agents
- Well-known UDP port number 161/162 at SNMP agent/manager

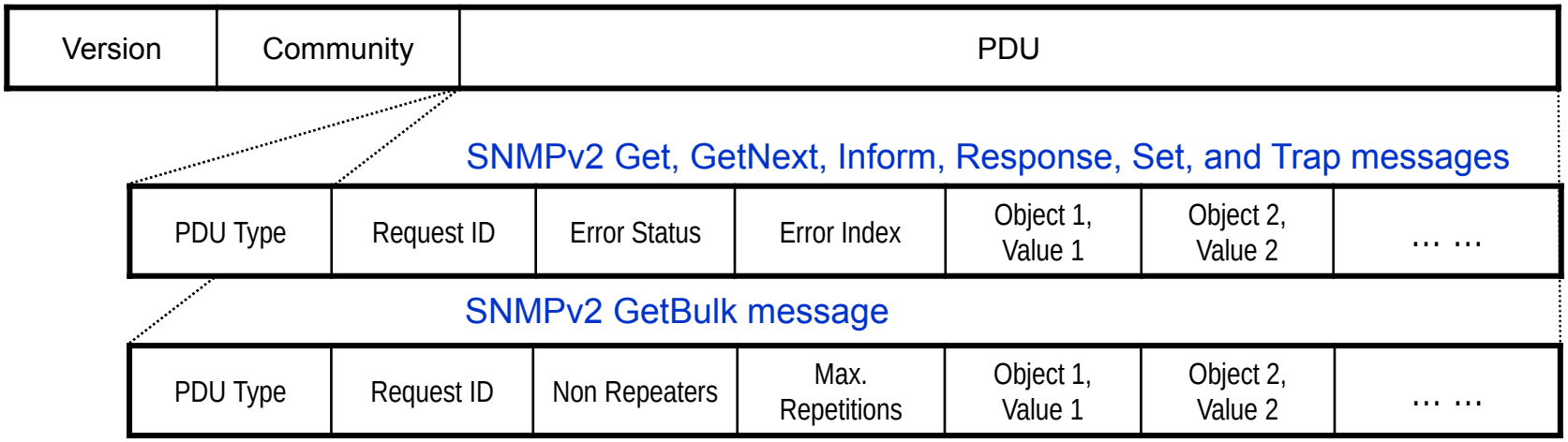


SNMP Messages

SNMP messages exchange information between an SNMP manager and an SNMP agent

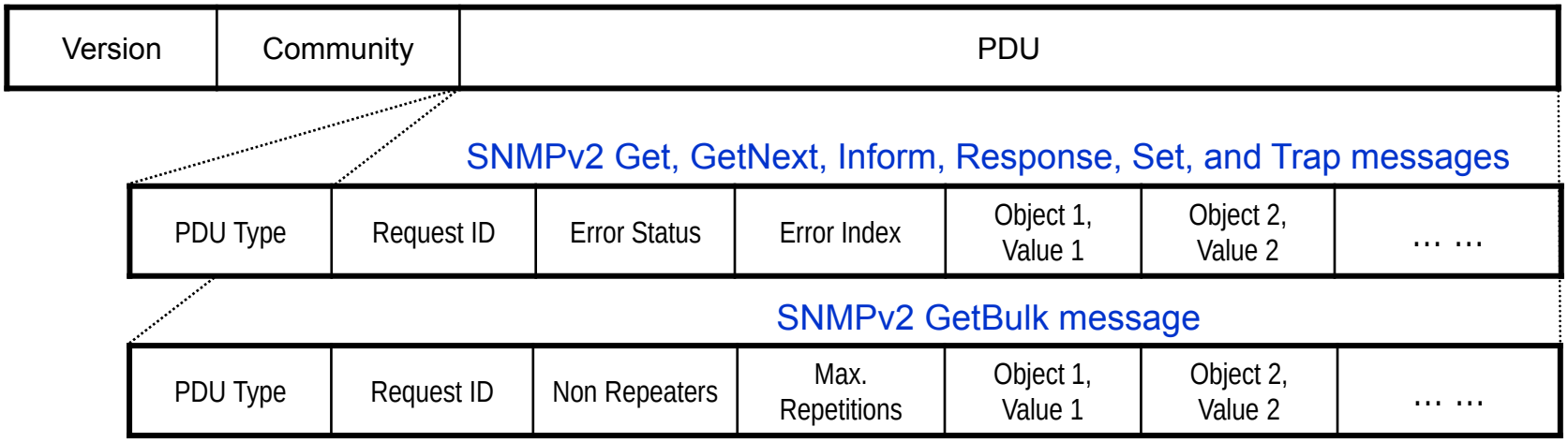
- Get: fetches the value of one or more objects
- GetNext: fetches the value of the next object after the specified object
- Set: sets the value of one or more objects
- Response: returns the value of one or more objects
- Trap: reports the occurrence of some significant events in a managed device.
- Inform: reports the occurrence of some significant events in a managed device and requests a response from the manager.
- GetBulk: allows exchanging of responses with a large amount of management information.

SNMP Message Formats



- Version Number
  - The version of SNMP: SNMPv1,SNMPv2, SNMPv3
  - SNMPv2 extends SNMPv1 by defining additional operations (GetBulk, Inform)
  - SNMPv3 extends SNMPv2 by adding security and remote configuration capabilities
- Community Name
  - Defines the access scope for SNMP managers and agents
  - An SNMP message carrying a different community name is discarded
- Protocol Data Unit (PDU) Type
  - Specifies the SNMP message type

SNMP Message Format (cont'd)



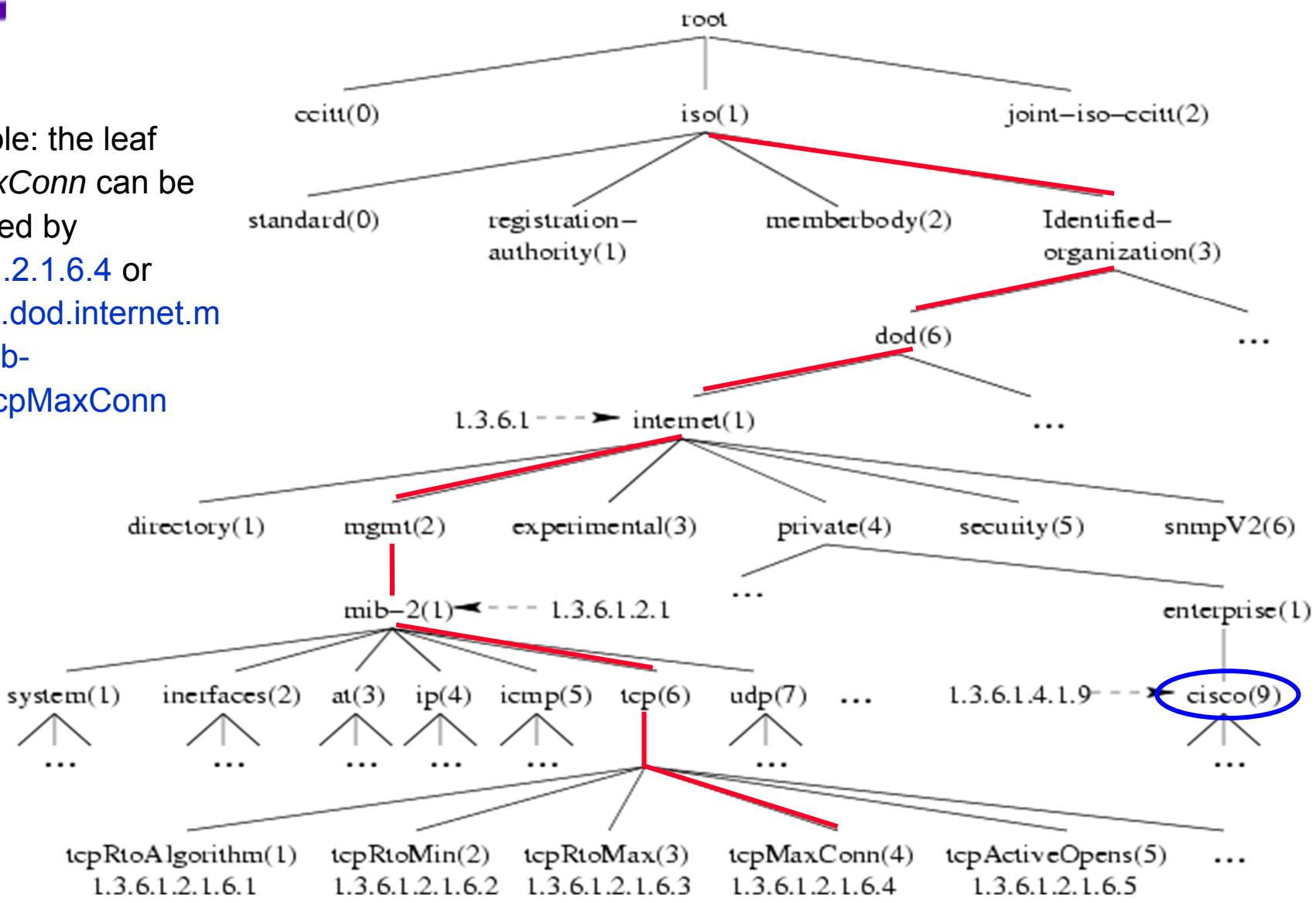
- Request ID
  - Used to match an SNMP request with the corresponding response
- Error Status
  - An integer specifying an error only set by an SNMP response
- Error Index
  - An integer offset specifying which object was in error only set by an SNMP response
- Objects and Values
  - A list of objects and their values

MIB Structure

- A managed device maintains a large number of SNMP objects to store management information
- The Structure of Management Information (SMI)
  - Defines the rules for describing management information and the data types used in SNMP
  - Data types: Integer, Octet String, Sequence
- MIB objects are organized as a tree
  - Each level of the tree consists of groups
  - Each group has its name and the associated numerical identifier
  - Leaves in the mib-2 subtree are MIB objects
  - Vendor-specific MIBs are located in the enterprise subtree
  - Each node (leaf) is identified by a concatenation of the names (or IDs) of all its predecessors starting from the root

MIB Tree Hierarchy

Example: the leaf *tcpMaxConn* can be identified by 1.3.6.1.2.1.6.4 or iso.org.dod.internet.mgmt.mib-2.tcp.tcpMaxConn



NET-SNMP

- Formerly known as UCD-SNMP
- A very popular public domain SNMP implementation
- Consists of
  - an extensible SNMP agent
  - a set of tools to request or set information from SNMP agents
  - a set of tools to generate and handle SNMP traps
  - an SNMP API library for writing SNMP related programs
- See Section 9.2.3 for details