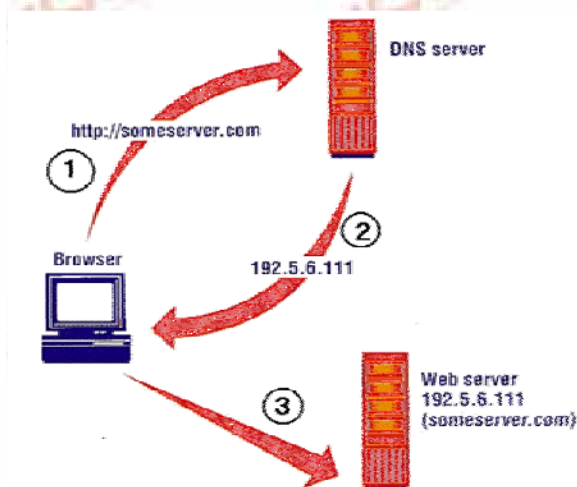


## BÀI 5: GIAO THỨC TCP/IP VÀ INTERNET



### Nội dung

- Các khái niệm về họ giao thức TCP/IP.
- Mạng Internet.
- Các ứng dụng trên TCP/IP.
- Dịch vụ tên miền (DNS).
- Electronic mail (SMTP, MINE, POP3, IMAP).
- World Wide Web (WWW).
- Truyền tập tin (FTP).
- Tài nguyên trên Internet.

### Hướng dẫn học

- Nghe giảng và đọc tài liệu để nắm bắt các nội dung chính.
- Làm bài tập và luyện thi trắc nghiệm theo yêu cầu của từng bài.
- Liên hệ và lấy các ví dụ trong thực tế để minh họa cho nội dung bài học.

### Thời lượng học

- 12 tiết.

### Mục tiêu

Sau khi học bài này, các bạn có thể:

- Trình bày được khái niệm về họ giao thức TCP/IP.
- Trình bày được một số ứng dụng của TCP/IP và tài nguyên trên Internet.

## TÌNH HUỐNG KHỞI ĐỘNG BÀI

### Tình huống dẫn nhập

Thực hiện các chức năng ở tầng 3 (Network) và 4 (Transport) trong mô hình OSI, mỗi hệ thống mạng sẽ có những giao thức riêng:

- UNIX: tầng 3 dùng giao thức IP, tầng 4 dùng giao thức TCP/UDP.
- Netware: tầng 3 dùng giao thức IPX, tầng 4 dùng giao thức SPX.
- Microsoft định nghĩa NETBEUI để thực hiện chức năng của cả tầng 3 và tầng 4.

Nếu chỉ dùng lại ở đây thì các máy tính UNIX, Netware, NT...sẽ không trao đổi thông tin được với nhau. Với sự lớn mạnh của mạng Internet, các máy tính cài đặt hệ điều hành khác nhau đòi hỏi phải giao tiếp được với nhau.



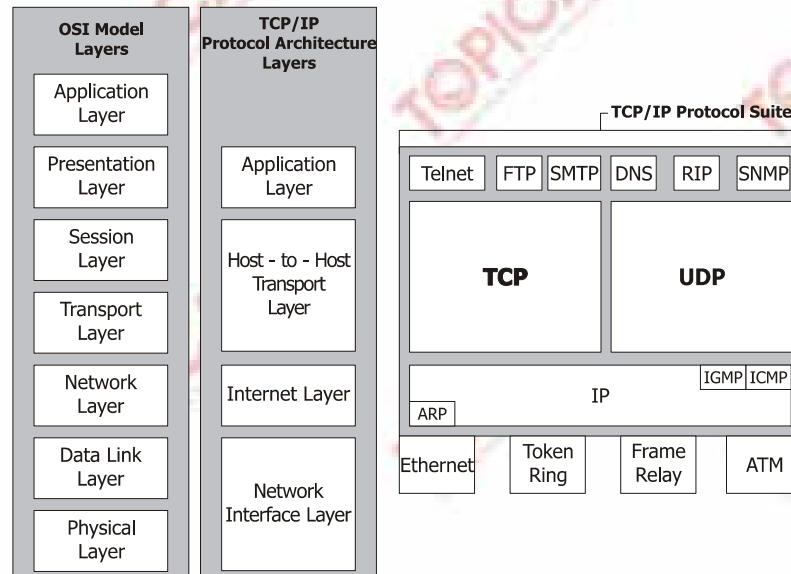
### Câu hỏi

1. Những giao thức nào đã được sử dụng để các máy tính khác nhau vẫn giao tiếp được với nhau?
2. Kiến trúc của bộ giao thức đó ra sao?
3. Những ứng dụng trên giao thức đó là gì?

## 5.1. Các khái niệm về họ giao thức TCP/IP

### 5.1.1. Khái niệm

Các tầng giao thức TCP/IP được chia làm bốn tầng chức năng được biết đến như là mô hình DARPA, được đặt tên sau khi chính phủ Mỹ bắt đầu phát triển TCP/IP. Bốn tầng của mô hình DARPA là: ứng dụng (Application), Giao vận (Transport), Internet, và Giao tiếp mạng (Network Interface). Mỗi tầng trong mô hình DARPA tương ứng với một hoặc nhiều tầng trong mô hình 7 tầng của mô hình OSI. Hình 5.1 biểu diễn kiến trúc giao thức TCP/IP.



**Hình 5.1: Kiến trúc giao thức TCP/IP**

#### 5.1.1.1. Tầng giao tiếp mạng (Network Interface Layer)

Tầng giao tiếp mạng (còn được gọi là tầng truy cập mạng) chịu trách nhiệm đặt các gói tin TCP/IP trên môi trường mạng và nhận các gói tin TCP/IP từ môi trường mạng. TCP/IP được thiết kế độc lập với phương pháp truy cập mạng, định dạng khung dữ liệu và môi trường mạng. Bằng cách này, TCP/IP có thể được sử dụng để kết nối các loại mạng khác nhau. Bao gồm các kỹ thuật mạng LAN như Ethernet hoặc Token Ring và các kỹ thuật mạng WAN như X.25 hay Frame Relay. Sự độc lập với bất kỳ kỹ thuật mạng nào cho phép TCP/IP có khả năng tương thích với các kỹ thuật mới như ATM (Asynchronous Transfer Mode).

Tầng giao tiếp mạng bao gồm tầng liên kết dữ liệu (Data Link) và tầng vật lý (Physical) của mô hình OSI. Chú ý rằng tầng Internet không có được các ưu điểm của các dịch vụ sắp xếp gói tin và thông báo có ở tầng Data Link. Một tầng giao tiếp mạng được giả thiết và việc truyền thông tin cậy qua các phiên thiết lập, sắp xếp và thông báo các gói tin thuộc về trách nhiệm của tầng giao vận.

#### 5.1.1.2. Tầng Internet

Tầng Internet chịu trách nhiệm địa chỉ hoá, đóng gói và dẫn đường. Lỗi các giao thức của tầng Internet là IP, ARP, ICMP và IGMP:

- Giao thức IP - Internet Protocol là một giao thức có khả năng dẫn đường cho các địa chỉ IP, phân chia và tập hợp lại các gói tin.

- Giao thức ARP - Address Resolution Protocol (giao thức phân giải địa chỉ) chịu trách nhiệm phân giải địa chỉ tầng Internet chuyển thành địa chỉ tầng giao tiếp mạng, như địa chỉ phần cứng.
- Giao thức ICMP - Internet Control Message Protocol chịu trách nhiệm đưa ra các chức năng chuẩn đoán và thông báo lỗi hay theo dõi các điều kiện lưu chuyển các gói tin IP.
- Giao thức IGMP - Internet Group Management Protocol chịu trách nhiệm quản lý các nhóm IP truyền multicast.
- Tầng Internet tương tự như tầng Network của mô hình OSI.

#### 5.1.1.3. Tầng giao vận

Tầng giao vận còn được gọi là tầng truyền Trạm – tới – Trạm (Host – to – Host Transport Layer) chịu trách nhiệm cung cấp cho tầng ứng dụng các dịch vụ tạo lập phiên và truyền dữ liệu. Các giao thức lõi của tầng giao vận là TCP và UDP (User Datagram Protocol).

- TCP cung cấp các dịch vụ truyền thông tin cậy một-một (One-to-One), hướng liên kết (Connection-Oriented). TCP chịu trách nhiệm thiết lập các kết nối TCP, gửi các gói tin có sắp xếp, thông báo và các gói tin phục hồi dữ liệu bị mất trong quá trình truyền.
- UDP cung cấp các dịch vụ truyền tin một-một, một-nhiều, không liên kết và không tin cậy. UDP được sử dụng khi lượng dữ liệu cần truyền nhỏ (ví dụ dữ liệu không điền hết một gói tin), khi việc thiết lập liên kết TCP là không cần thiết, hoặc khi các ứng dụng hoặc các giao thức tầng trên cung cấp dịch vụ đảm bảo trong khi truyền.

Tầng giao vận chịu trách nhiệm tầng giao vận trong mô hình OSI và một số nhiệm vụ của tầng phiên (Session) của OSI.

#### 5.1.1.4. Tầng ứng dụng

Tầng ứng dụng cung cấp các ứng dụng với khả năng truy cập các dịch vụ của các tầng khác và định nghĩa các giao thức mà các ứng dụng sử dụng để trao đổi dữ liệu. Có nhiều giao thức tầng ứng dụng và các giao thức mới luôn luôn được phát triển.

Các giao thức được ứng dụng rộng rãi nhất của tầng ứng dụng được sử dụng để trao đổi thông tin của người sử dụng là:

- Giao thức truyền tin siêu văn bản HTTP (HyperText Transfer Protocol) được sử dụng để truyền các tệp tạo nên trang Web của World Wide Web.
- Giao thức FTP - File Transfer Protocol được sử dụng để thực hiện truyền tệp.
- Giao thức SMTP - Simple Mail Transfer Protocol được sử dụng để truyền các thông điệp thư và các tệp đính kèm.
- Telnet, một giao thức mô phỏng trạm đầu cuối, được sử dụng để đăng nhập từ xa vào các máy trạm trên mạng.

Hơn nữa, các giao thức ứng dụng sau sẽ giúp dễ dàng sử dụng và quản lý mạng TCP/IP.

- Domain Name System (DNS) được sử dụng để chuyển từ tên trạm thành địa chỉ IP.
- Giao thức RIP - Routing Information Protocol là giao thức dẫn đường mà các router sử dụng để trao đổi các thông tin dẫn đường gói tin IP trong mạng.

- Giao thức SNMP - Simple Network Management Protocol được sử dụng giữa giao diện quản lý mạng và các thiết bị mạng (Router, Bridges và Hub thông minh) để thu thập và trao đổi thông tin quản lý mạng.

Ví dụ của tầng ứng dụng giao tiếp với các ứng dụng TCP/IP là Windows Sockets và NetBIOS. Windows Sockets cung cấp một chuẩn giao diện lập trình ứng dụng API (Application – Programming Interface) trên nền hệ điều hành Windows. NetBIOS là một chuẩn công nghiệp giao tiếp để truy cập các dịch vụ như dịch vụ phiên, truyền dữ liệu và phân giải tên. Thông tin chi tiết về NetBIOS được cung cấp ở cuối bài này.

### 5.1.2. Giao thức lõi TCP/IP

Thành phần giao thức TCP/IP được cài trong hệ điều hành mạng của bạn là một chuỗi các giao thức liên hệ với nhau được gọi là các giao thức lõi TCP/IP. Tất cả các ứng dụng khác và các giao thức khác trong bộ giao thức TCP/IP dựa trên các dịch vụ cơ sở cung cấp bởi các giao thức sau: IP, ARP, ICMP, IGMP, TCP và UDP.

#### 5.1.2.1. IP

##### Tổng quát

Nhiệm vụ chính của giao thức IP là cung cấp khả năng kết nối các mạng con thành liên kết mạng để truyền dữ liệu, vai trò của IP là vai trò của giao thức tầng mạng trong mô hình OSI. Giao thức IP là một giao thức kiểu không liên kết (Connectionless) có nghĩa là không cần có giai đoạn thiết lập liên kết trước khi truyền dữ liệu.

Sơ đồ địa chỉ hóa để định danh các trạm (Host) trong liên mạng được gọi là địa chỉ IP 32 bits (32 bit IP Address). Mỗi giao diện trong một máy có hỗ trợ giao thức IP đều phải được gán một địa chỉ IP (một máy tính có thể gắn với nhiều mạng do vậy có thể có nhiều địa chỉ IP). Địa chỉ IP gồm hai phần: địa chỉ mạng (netid) và địa chỉ máy (hostid). Mỗi địa chỉ IP có độ dài 32 bits được tách thành 4 vùng (mỗi vùng 1 byte), có thể biểu thị dưới dạng thập phân, bát phân, thập lục phân hay nhị phân. Cách viết phổ biến nhất là dùng ký pháp thập phân có dấu chấm (Dotted Decimal Notation) để tách các vùng. Mục đích của địa chỉ IP là để định danh duy nhất cho một máy tính bất kỳ trên liên mạng.

Do tổ chức và độ lớn của các mạng con (Subnet) của liên mạng có thể khác nhau, người ta chia các địa chỉ IP thành 5 lớp, ký hiệu là A, B, C, D và E. Trong lớp A, B, C chứa địa chỉ có thể gán được. Lớp D dành riêng cho lớp kỹ thuật *multicasting*. Lớp E được dành những ứng dụng trong tương lai.

Netid trong địa chỉ mạng dùng để nhận dạng từng mạng riêng biệt. Các mạng liên kết phải có địa chỉ mạng (netid) riêng cho mỗi mạng. Ở đây các bit đầu tiên của byte đầu tiên được dùng để định danh lớp địa chỉ (0 - lớp A, 10 - lớp B, 110 - lớp C, 1110 - lớp D và 11110 - lớp E).

Ở đây ta xét cấu trúc của các lớp địa chỉ có thể gán được là lớp A, lớp B, lớp C

Cấu trúc của các địa chỉ IP như sau:

- Mạng lớp A: địa chỉ mạng (netid) là 1 Byte và địa chỉ host (hostid) là 3 byte.
- Mạng lớp B: địa chỉ mạng (netid) là 2 Byte và địa chỉ host (hostid) là 2 byte.
- Mạng lớp C: địa chỉ mạng (netid) là 3 Byte và địa chỉ host (hostid) là 1 byte.



Lớp A cho phép định danh tới 126 mạng, với tối đa 16 triệu host trên mỗi mạng. Lớp này được dùng cho các mạng có số trạm cực lớn.

Lớp B cho phép định danh tới 16384 mạng, với tối đa 65534 host trên mỗi mạng.

Lớp C cho phép định danh tới 2 triệu mạng, với tối đa 254 host trên mỗi mạng. Lớp này được dùng cho các mạng có ít trạm.

	Netid			Hostid
Địa chỉ lớp A	0xxxxxxx	xxxxxxxx	xxxxxxxx	xxxxxxxx
Địa chỉ lớp B	10xxxxxx	xxxxxxxx	xxxxxxxx	xxxxxxxx
Địa chỉ lớp C	110xxxxx	xxxxxxxx	xxxxxxxx	xxxxxxxx

**Hình 5.2: Cấu trúc các lớp địa chỉ IP**

Một số địa chỉ có tính chất đặc biệt: Một địa chỉ có hostid = 0 được dùng để hướng tới mạng định danh bởi vùng netid. Ngược lại, một địa chỉ có vùng hostid gồm toàn số 1 được dùng để hướng tới tất cả các host nối vào mạng netid và nếu vùng netid cũng gồm toàn số 1 thì nó hướng tới tất cả các host trong liên mạng.

00001010	00001010	00001010	00001010	= 10.0.0.0 (lớp A) net id = 10
10000000	00000011	00000010	00000011	= 128.3.2.3 (lớp B) net id = 128.3 host id = 2.3
11000000	00000000	00000001	11111111	= 192.0.1.255 (lớp C) net id = 192.0.1 host id = 255 → hướng tới tất cả các host

**Hình 5.3: Ví dụ cấu trúc các lớp địa chỉ IP**

Cần lưu ý rằng các địa chỉ IP được dùng để định danh các host và mạng ở tầng mạng của mô hình OSI, và chúng không phải là các địa chỉ vật lý (hay địa chỉ MAC) của các trạm trên đó một mạng cục bộ (Ethernet, Token Ring.).

Trong nhiều trường hợp, một mạng có thể được chia thành nhiều mạng con (subnet), lúc đó có thể đưa thêm các vùng subnetid để định danh các mạng con. Vùng subnetid được lấy từ vùng hostid, cụ thể đối với lớp A, B, C như ví dụ sau:

Bit:	0	7	8	15	16	23	24	26	27	31	
	Net ID		Subnet ID		Host ID						(Lớp A)
	Net ID			Subnet ID			Host ID				(Lớp B)
	Net ID						Subnet ID		Host ID		(Lớp C)

**Hình 5.4: Ví dụ địa chỉ khi bổ sung vùng subnetid**

Đơn vị dữ liệu dùng trong IP được gọi là gói tin (Datagram), có khuôn dạng.

Bit:	0	3	4	7	8	15	16	4	31
	VER	IHL		Type of service				Total Length	
	Identification					Flags	Fragment offset		
	Time to live			Protocol			Heder Checksum		
	Source Address								
	Destintion Address								
	Option + Padding								
	Data								

**Hình 5.5: Dạng thức của gói tin IP**

**Ý nghĩa của thông số như sau:**

- VER (4 bits): chỉ phiên bản hiện hành của giao thức IP hiện được cài đặt, việc có chỉ số phiên bản cho phép có các trao đổi giữa các hệ thống sử dụng phiên bản cũ và hệ thống sử dụng phiên bản mới.
- IHL (4 bits): chỉ độ dài phần đầu (Internet header Length) của gói tin, tính theo đơn vị từ ( 32 bits). Trường này bắt buộc phải có vì phần đầu IP có thể có độ dài thay đổi tùy ý. Độ dài tối thiểu là 5 từ (20 bytes), độ dài tối đa là 15 từ hay là 60 bytes.
- Type of service (8 bits): đặc tả các tham số về dịch vụ nhằm thông báo cho mạng biết dịch vụ nào mà gói tin muốn được sử dụng, chẳng hạn ưu tiên, thời hạn chậm trễ, năng suất truyền và độ tin cậy. Hình sau cho biết ý nghĩ của trường 8 bits này.

0	1	2	3	4	5	6	7
Precedence			D	T	R	Reserved	

**Hình 5.6: Ý nghĩa các bit trong Type Of Service**

- Precedence (3 bit): chỉ thị về quyền ưu tiên gửi gói tin, nó có giá trị từ 0 (gói tin bình thường) đến 7 (gói tin kiểm soát mạng).
- D (Delay) (1 bit): chỉ độ trễ yêu cầu trong đó:
  - D = 0 gói tin có độ trễ bình thường.
  - D = 1 gói tin độ trễ thấp.
- T (Throughput) (1bit): chỉ độ thông lượng yêu cầu sử dụng để truyền gói tin với lựa chọn truyền trên đường thông suất thấp hay đường thông suất cao:
  - T = 0 thông lượng bình thường.
  - T = 1 thông lượng cao.
- R (Reliability) (1 bit): chỉ độ tin cậy yêu cầu:
  - R = 0 độ tin cậy bình thường.
  - R = 1 độ tin cậy cao.
- Total Length (16 bits): chỉ độ dài toàn bộ gói tin, kể cả phần đầu tính theo đơn vị byte với chiều dài tối đa là 65535 bytes. Hiện nay giới hạn trên là rất lớn nhưng trong tương lai với những mạng Gigabit thì các gói tin có kích thước lớn là cần thiết.
- Identification (16 bits): cùng với các tham số khác (như Source Address và Destination Address) tham số này dùng để định danh duy nhất cho một gói tin trong khoảng thời gian nó vẫn còn trên liên mạng.

- **Flags (3 bits):** liên quan đến sự phân đoạn (Fragment) các gói tin. Các gói tin khi đi trên đường đi có thể bị phân thành nhiều gói tin nhỏ, trong trường hợp bị phân đoạn thì trường Flags được dùng điều khiển phân đoạn và tái lắp ghép bó dữ liệu. Tùy theo giá trị của Flags sẽ có ý nghĩa là gói tin sẽ không phân đoạn, có thể phân đoạn hay là gói tin phân đoạn cuối cùng. Trường Fragment Offset cho biết vị trí dữ liệu thuộc phân đoạn tương ứng với đoạn bắt đầu của gói dữ liệu gốc.
- Ý nghĩa cụ thể của trường Flags là:

0	1	2
<b>O</b>	<b>DF</b>	<b>MF</b>

**Hình 5.7: Ý nghĩa các bit của Flags**

- Bit 0: reserved - chưa sử dụng, luôn lấy giá trị 0.
- Bit 1: (DF) = 0 (May Fragment) = 1 (Don't Fragment)
- Bit 2: (MF) = 0 (Last Fragment) = 1 (More Fragments)
- **Fragment Offset (13 bits):** chỉ vị trí của đoạn (Fragment) ở trong gói tin tính theo đơn vị 8 bytes, có nghĩa là phần dữ liệu mỗi gói tin (trừ gói tin cuối cùng) phải chứa một vùng dữ liệu có độ dài là bội số của 8 bytes. Điều này có ý nghĩa là phải nhân giá trị của Fragment offset với 8 để tính ra độ lệch byte.
- **Time to Live (8 bits):** quy định thời gian tồn tại (tính bằng giây) của gói tin trong mạng để tránh tình trạng một gói tin bị quẩn trên mạng. Thời gian này được cho bởi trạm gửi và được giảm đi (thường quy ước là 1 đơn vị) khi gói tin đi qua mỗi router của liên mạng. Thời lượng này giảm xuống tại mỗi router với mục đích giới hạn thời gian tồn tại của các gói tin và kết thúc những lần lặp lại vô hạn trên mạng. Sau đây là một số điều cần lưu ý về trường Time To Live:
  - Nút trung gian của mạng không được gửi một gói tin mà trường này có giá trị bằng 0.
  - Một giao thức có thể ấn định Time To Live để thực hiện cuộc ra tìm tài nguyên trên mạng trong phạm vi mở rộng.
  - Một giá trị cố định tối thiểu phải đủ lớn cho mạng hoạt động tốt.
- **Protocol (8 bits):** chỉ giao thức tầng trên kế tiếp sẽ nhận vùng dữ liệu ở trạm đích (hiện tại thường là TCP hoặc UDP được cài đặt trên IP). Ví dụ: TCP có giá trị trường Protocol là 6, UDP có giá trị trường Protocol là 17.
- **Header Checksum (16 bits):** mã kiểm soát lỗi của header gói tin IP.
- **Source Address (32 bits):** địa chỉ của máy nguồn.
- **Destination Address (32 bits):** địa chỉ của máy đích.
- **Options (độ dài thay đổi):** khai báo các lựa chọn do người gửi yêu cầu (tùy theo từng chương trình).
- **Padding (độ dài thay đổi):** vùng đệm, được dùng để đảm bảo cho phần header luôn kết thúc ở một mốc 32 bits.
- **Data (độ dài thay đổi):** trên một mạng cục bộ như vậy, hai trạm chỉ có thể liên lạc với nhau nếu chúng biết địa chỉ vật lý của nhau. Như vậy vấn đề đặt ra là phải thực hiện ánh xạ giữa địa chỉ IP (32 bits) và địa chỉ vật lý (48 bits) của một trạm.



### Các giao thức trong mạng IP

Để mạng với giao thức IP hoạt động được tốt người ta cần một số giao thức bổ sung, các giao thức này đều không phải là bộ phận của giao thức IP và giao thức IP sẽ dùng đến chúng khi cần.

- Giao thức ARP (Address Resolution Protocol): ở đây cần lưu ý rằng các địa chỉ IP được dùng để định danh các host và mạng ở tầng mạng của mô hình OSI. Chúng không phải là các địa chỉ vật lý (hay địa chỉ MAC) của các trạm trên một mạng cục bộ (Ethernet, Token Ring). Trên một mạng cục bộ hai trạm chỉ có thể liên lạc với nhau nếu chúng biết địa chỉ vật lý của nhau. Như vậy vấn đề đặt ra là phải tìm được ánh xạ giữa địa chỉ IP (32 bits) và địa chỉ vật lý của một trạm. Giao thức ARP đã được xây dựng để tìm địa chỉ vật lý từ địa chỉ IP khi cần thiết.
- Giao thức RARP (Reverse Address Resolution Protocol): là giao thức ngược với giao thức ARP. Giao thức RARP được dùng để tìm địa chỉ IP từ địa chỉ vật lý.
- Giao thức ICMP (Internet Control Message Protocol): giao thức này thực hiện truyền các thông báo điều khiển (báo cáo các tình trạng các lỗi trên mạng) giữa các gateway hoặc một nút của liên mạng. Tình trạng lỗi có thể là: một gói tin IP không thể tới đích hoặc một router không đủ bộ nhớ đệm để lưu và chuyển một gói tin IP. Một thông báo ICMP được tạo chuyển cho IP. IP sẽ "bọc" (Encapsulate) thông báo đó với một IP header và truyền đến cho router hoặc trạm đích.

### Các bước hoạt động của giao thức IP

Khi giao thức IP được khởi động nó trở thành một thực thể tồn tại trong máy tính và bắt đầu thực hiện những chức năng của mình, lúc đó thực thể IP là cấu thành của tầng mạng, nhận yêu cầu từ các tầng trên nó và gửi yêu cầu xuống các tầng dưới nó.

Đối với thực thể IP ở máy nguồn, khi nhận được một yêu cầu gửi từ tầng trên, nó thực hiện các bước sau đây:

- Tạo một gói tin IP dựa trên tham số nhận được.
- Tính checksum và ghép vào header của gói tin.
- Ra quyết định chọn đường: hoặc là trạm đích nằm trên cùng mạng hoặc một gateway sẽ được chọn cho chặng tiếp theo.
- Chuyển gói tin xuống tầng dưới để truyền qua mạng.
- Đối với router, khi nhận được một gói tin đi qua, nó thực hiện các động tác sau:
  - Tính checksum, nếu sai thì loại bỏ gói tin.
  - Giảm giá trị tham số Time - to Live, nếu thời gian đã hết thì loại bỏ gói tin.
  - Ra quyết định chọn đường.
  - Phân đoạn gói tin, nếu cần.
  - Kiến tạo lại IP header, bao gồm giá trị mới của các vùng Time - to -Live, Fragmentation và Checksum.
  - Chuyển gói tin xuống tầng dưới để chuyển qua mạng.

Cuối cùng khi một gói tin nhận bởi một thực thể IP ở trạm đích, nó sẽ thực hiện bởi các công việc sau:

- Tính checksum. Nếu sai thì loại bỏ gói tin.
- Tập hợp các đoạn của gói tin (nếu có phân đoạn).
- Chuyển dữ liệu và các tham số điều khiển lên tầng trên.

### 5.1.2.2. ARP

Khi các gói tin IP được gửi trên môi trường chia sẻ truy cập, các công nghệ mạng dựa trên công nghệ truyền quảng bá như Ethernet hoặc Token Ring, Media Access Control (MAC) địa chỉ tương ứng với địa chỉ các gói tin IP phải được chuyển hoá. ARP sử dụng MAC mức truyền quảng bá để phân giải một địa chỉ IP gửi đi thành một địa chỉ MAC. ARP được định nghĩa trong RFC 826.

Để có thêm thông tin về ARP, xem phần “Phân giải địa chỉ vật lý” trong đoạn cuối của bài này.

### 5.1.2.3. ICMP

Giao thức điều khiển thông điệp Internet ICMP cung cấp tiện ích sửa chữa sự cố và thông báo lỗi cho các gói tin không truyền đi được. Ví dụ, nếu một gói tin IP không truyền tới trạm đích được, ICMP sẽ gửi thông điệp thông báo không tiếp cận đích được “Destination Unreachable” tới trạm nguồn. Bảng sau biểu diễn các thông điệp ICMP thông thường nhất.

**Bảng các thông điệp ICMP thông thường**

Thông điệp ICMP	Chức năng
Echo Request (yêu cầu phản hồi).	Thông điệp báo sự cố đơn giản được sử dụng để kiểm tra kết nối IP tới trạm mong muốn.
Echo Reply (trả lời phản hồi).	Dùng để trả lời cho Echo Request.
Redirect (định hướng lại).	Được gửi từ một router để khẳng định một trạm gửi tin dẫn đường tốt hơn tới một địa chỉ IP đích.
Source Quench (tắt nguồn).	Được gửi từ một router để khẳng định trạm gửi tin mà dữ liệu IP bị loại vì nghẽn tại router. Trạm gửi tin sẽ hạ thấp tỉ lệ truyền. Source Quench là một thông điệp ICMP không bắt buộc và thường không được cài đặt.
Destination Unreachable (không thể tiếp cận đích).	Được gửi bởi một router hoặc trạm đích để thông báo cho trạm gửi tin rằng gói tin không thể truyền được.

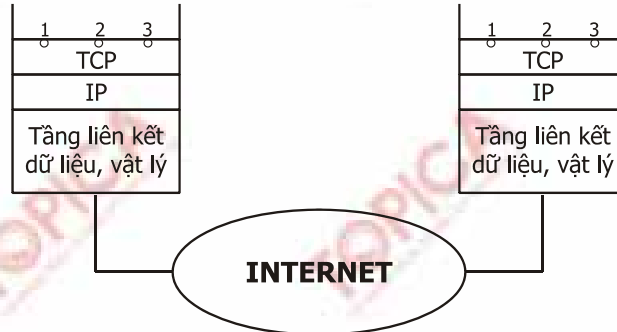
Để gửi một gói tin phản hồi yêu cầu ICMP và hiển thị thống kê trên các thông điệp trả lời trên máy sử dụng Windows NT chúng ta sử dụng tiện ích ping trong dấu nhắc lệnh. ICMP không làm cho giao thức IP trở nên tin cậy. ICMP cố gắng thông báo các lỗi và đưa ra phản hồi trong các điều kiện cụ thể. Các thông điệp ICMP được truyền đi như một gói tin IP không thông báo và bản thân chúng là không chắc chắn có tới được đích không. ICMP được định nghĩa trong RFC 792.

### 5.1.2.4. IGMP

Giao thức quản lý nhóm Internet IGMP (Internet Group Management Protocol) là một giao thức quản lý các trạm thành viên trong nhóm truyền IP multicast. Một nhóm IP multicast, được biết đến như một nhóm trạm. Gói tin IP multicast được truyền tới một địa chỉ MAC đơn nhưng được xử lý bởi nhiều trạm sử dụng giao thức IP. Một trạm cụ thể sẽ nghe một địa chỉ gói tin IP multicast cụ thể và nhận tất cả các gói tin từ địa chỉ đó.

### 5.1.2.5. TCP

TCP là một giao thức "có liên kết" (Connection - Oriented), nghĩa là cần phải thiết lập liên kết giữa hai thực thể TCP trước khi chúng trao đổi dữ liệu với nhau. Một tiến trình ứng dụng trong một máy tính truy nhập vào các dịch vụ của giao thức TCP thông qua một cổng (port) của TCP. Số hiệu cổng TCP được thể hiện bởi 2 bytes.



**Hình 5.8: Cổng truy nhập dịch vụ TCP**

Một cổng TCP kết hợp với địa chỉ IP tạo thành một đầu nối TCP/IP (socket) duy nhất trong liên mạng. Dịch vụ TCP được cung cấp nhờ một liên kết logic giữa một cặp đầu nối TCP/IP. Một đầu nối TCP/IP có thể tham gia nhiều liên kết với các đầu nối TCP/IP ở xa khác nhau. Trước khi truyền dữ liệu giữa hai trạm cần phải thiết lập một liên kết TCP giữa chúng và khi không còn nhu cầu truyền dữ liệu thì liên kết đó sẽ được giải phóng.

Các thực thể của tầng trên sử dụng giao thức TCP thông qua các hàm gọi (function calls), trong đó có các hàm để yêu cầu, để trả lời. Trong mỗi hàm còn có các tham số dành cho việc trao đổi dữ liệu.

Các bước thực hiện để thiết lập một liên kết TCP/IP: Thiết lập một liên kết mới có thể được mở theo một trong hai phương thức: chủ động (Active) hoặc bị động (Passive).

Phương thức bị động, người sử dụng yêu cầu TCP chờ đợi một yêu cầu liên kết gửi đến từ xa thông qua một đầu nối TCP/IP (tại chỗ). Người sử dụng dùng hàm passive Open có khai báo cổng TCP và các thông số khác (mức ưu tiên, mức an toàn)

Với phương thức chủ động, người sử dụng yêu cầu TCP mở một liên kết với một đầu nối TCP/IP ở xa. Liên kết sẽ được xác lập nếu có một hàm Passive Open tương ứng đã được thực hiện tại đầu nối TCP/IP ở đó.

**Bảng liệt kê một vài cổng TCP phổ biến**

Số hiệu cổng	Mô tả	Số hiệu cổng	Mô tả
0	Reserved	23	Telnet
5	Remote job entry	25	SMTP
7	Echo	37	Time
9	Discard	53	Name Server
11	Systat	102	ISO - TSAP
13	Daytime	103	X.400
15	Nestat	104	X.400 Sending
17	Quotd (Quote Odd Day)	111	Sun RPC
20	ftp-data	139	Net BIOS Session source
21	ftp (Control)	160 - 223	Reserved

Khi người sử dụng gửi đi một yêu cầu mở liên kết sẽ được nhận hai thông số trả lời từ TCP.

- Thông số Open ID được TCP trả lời ngay lập tức để gán cho một liên kết cục bộ (Local Connection Name) cho liên kết được yêu cầu. Thông số này về sau được dùng để tham chiếu tới liên kết đó. (Trong trường hợp nếu TCP không thể thiết lập được liên kết yêu cầu thì nó phải gửi tham số Open Failure để thông báo).
- Khi TCP thiết lập được liên kết yêu cầu nó gửi tham số Open Success được dùng để thông báo liên kết đã được thiết lập thành công. Thông báo này được chuyển đến trong cả hai trường hợp bị động và chủ động. Sau khi một liên kết được mở, việc truyền dữ liệu trên liên kết có thể được thực hiện.

Các bước thực hiện khi truyền và nhận dữ liệu: Sau khi xác lập được liên kết người sử dụng gửi và nhận dữ liệu. Việc gửi và nhận dữ liệu thông qua các hàm Send và Receive.

- Hàm Send: Dữ liệu được gửi xuống TCP theo các khối (Block). Khi nhận được một khối dữ liệu, TCP sẽ lưu trữ trong bộ đệm (Buffer). Nếu cờ PUSH được đánh dấu thì toàn bộ dữ liệu trong bộ đệm được gửi, kể cả khối dữ liệu mới đến sẽ được gửi đi. Ngược lại cờ PUSH không được đánh dấu thì dữ liệu được giữ lại trong bộ đệm và sẽ gửi đi khi có cơ hội thích hợp (chẳng hạn chờ thêm dữ liệu nữa để gửi đi sẽ hiệu quả hơn).
- Hàm receive: Ở trạm đích dữ liệu sẽ được TCP lưu trong bộ đệm gắn với mỗi liên kết. Nếu dữ liệu được đánh dấu với một cờ PUSH thì toàn bộ dữ liệu trong bộ đệm (kể cả các dữ liệu được lưu từ trước) sẽ được chuyển lên cho người sử dụng. Còn nếu dữ liệu đến không được đánh dấu với cờ PUSH thì TCP chờ tới khi thích hợp mới chuyển dữ liệu với mục tiêu tăng hiệu quả hệ thống.

Nói chung việc nhận và giao dữ liệu cho người sử dụng đích của TCP phụ thuộc vào việc cài đặt cụ thể. Trường hợp cần chuyển gấp dữ liệu cho người sử dụng thì có thể dùng cờ URGENT và đánh dấu dữ liệu bằng bit URG để báo cho người sử dụng cần phải xử lý khẩn cấp dữ liệu đó.

- Các bước thực hiện khi đóng một liên kết: Việc đóng một liên kết khi không cần thiết được thực hiện theo một trong hai cách: dùng hàm Close hoặc dùng hàm Abort.
  - Hàm Close: yêu cầu đóng liên kết một cách bình thường. Có nghĩa là việc truyền dữ liệu trên liên kết đó đã hoàn tất. Khi nhận được một hàm Close TCP sẽ truyền đi tất cả dữ liệu còn trong bộ đệm thông báo rằng nó đóng liên kết.  
*Lưu ý:* Khi một người sử dụng đã gửi đi một hàm Close thì nó vẫn phải tiếp tục nhận dữ liệu đến trên liên kết đó cho đến khi TCP đã báo cho phía bên kia biết về việc đóng liên kết và chuyển giao hết tất cả dữ liệu cho người sử dụng của mình.
  - Hàm Abort: người sử dụng có thể đóng một liên kết bất kỳ và sẽ không chấp nhận dữ liệu qua liên kết đó nữa. Do vậy dữ liệu có thể bị mất đi khi đang được truyền đi. TCP báo cho TCP ở xa biết rằng liên kết đã được hủy bỏ và TCP ở xa sẽ thông báo cho người sử dụng của mình.
- Một số hàm khác của TCP:
  - Hàm Status: cho phép người sử dụng yêu cầu cho biết trạng thái của một liên kết cụ thể, khi đó TCP cung cấp thông tin cho người sử dụng.
  - Hàm Error: thông báo cho người sử dụng TCP các yêu cầu dịch vụ bất hợp lệ liên quan đến một liên kết có tên cho trước hoặc về các lỗi liên quan đến môi trường.



- Đơn vị dữ liệu sử dụng trong TCP được gọi là segment (đoạn dữ liệu), có các tham số với ý nghĩa như sau:

Bit 0	15				16				31
Source Port					Destination Port				
Sequence Number									
Acknowledgment Number									
Data Offset	Reserved	U G R	A C K	P S H	R S T	S I N	F I N	Window	
Checksum				Urgent Poiter					
Options				Padding					
TCP data									

**Hình 5.9: Dạng thức của TCP Segment**

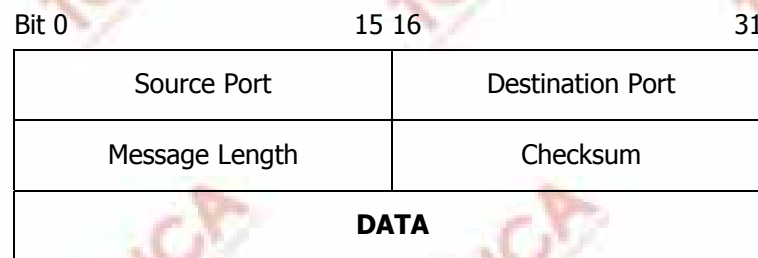
- Source Port (16 bits): Số hiệu cổng TCP của trạm nguồn.
- Destination Port (16 bit): Số hiệu cổng TCP của trạm đích.
- Sequence Number (32 bit): số hiệu của byte đầu tiên của segment trừ khi bit SYN được thiết lập. Nếu bit SYN được thiết lập thì Sequence Number là số hiệu tuần tự khởi đầu (ISN) và byte dữ liệu đầu tiên là ISN+1.
- Acknowledgment Number (32 bit): số hiệu của segment tiếp theo mà trạm nguồn đang chờ để nhận. Ngầm ý báo nhận tốt (các) segment mà trạm đích đã gửi cho trạm nguồn.
- Data offset (4 bit): số lượng bội của 32 bit (32 bit words) trong TCP header (tham số này chỉ ra vị trí bắt đầu của nguồn dữ liệu).
- Reserved (6 bit): dành để dùng trong tương lai.
- Control bit (các bit điều khiển):
  - URG: Vùng con trỏ khẩn (Urgent Pointer) có hiệu lực.
  - ACK: Vùng báo nhận (ACK number) có hiệu lực.
  - PSH: Chức năng PUSH.
  - RST: Khởi động lại (Reset) liên kết.
  - SYN: Đồng bộ hóa số hiệu tuần tự (Sequence Number).
  - FIN: Không còn dữ liệu từ trạm nguồn.
- Window (16 bit): cấp phát Credit để kiểm soát nguồn dữ liệu (cơ chế cửa sổ). Đây chính là số lượng các byte dữ liệu, bắt đầu từ byte được chỉ ra trong vùng ACK number, mà trạm nguồn đã sẵn sàng để nhận.
- Checksum (16 bit): mã kiểm soát lỗi cho toàn bộ segment (header + data).
- Urgent Pointer (16 bit): con trỏ này trỏ tới số hiệu tuần tự của byte đi theo sau dữ liệu khẩn. Vùng này chỉ có hiệu lực khi bit URG được thiết lập.
- Options (độ dài thay đổi): khai báo các tùy chọn của TCP, trong đó có độ dài tối đa của vùng TCP data trong một segment.
- Padding (độ dài thay đổi): phần chèn thêm vào header để đảm bảo phần header luôn kết thúc ở một mốc 32 bit. Phần thêm này gồm toàn số 0.
- TCP data (độ dài thay đổi): chứa dữ liệu của tầng trên, có độ dài tối đa ngầm định là 536 byte. Giá trị này có thể điều chỉnh bằng cách khai báo trong vùng options.



### 5.1.2.6. Giao thức UDP (User Datagram Protocol)

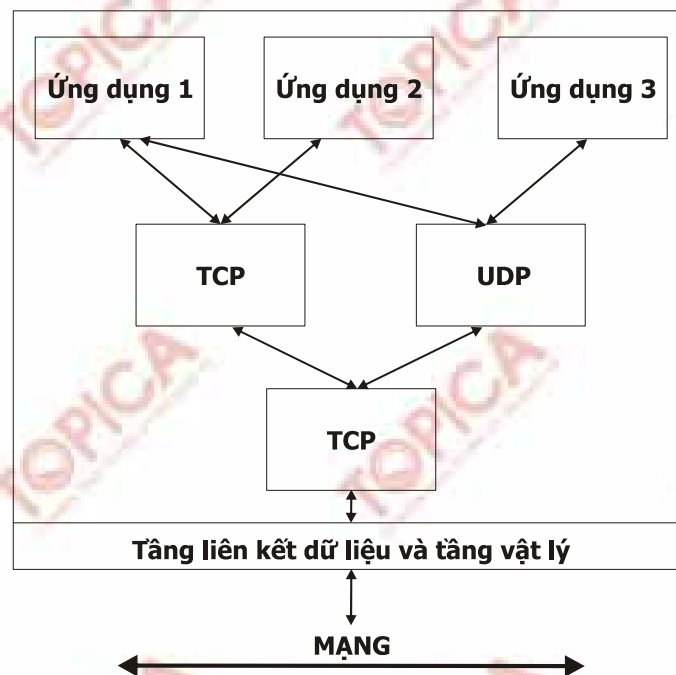
UDP (User Datagram Protocol) là giao thức theo phương thức không liên kết, được sử dụng thay thế cho TCP ở trên IP theo yêu cầu của từng ứng dụng. Khác với TCP, UDP không có các chức năng thiết lập và kết thúc liên kết. Tương tự như IP, nó cũng không cung cấp cơ chế báo nhận (Acknowledgment), không sắp xếp tuần tự các gói tin đến và có thể dẫn đến tình trạng mất hoặc trùng dữ liệu mà không có cơ chế thông báo lỗi cho người gửi. Qua đó ta thấy UDP cung cấp các dịch vụ vận chuyển không tin cậy như trong TCP.

Khuôn dạng UDP datagram được mô tả với các vùng tham số đơn giản hơn nhiều so với TCP segment.



**Hình 5.10: Dạng thức của gói tin UDP**

UDP cũng cung cấp cơ chế gán và quản lý các số hiệu cổng (Port Number) để định danh duy nhất cho các ứng dụng chạy trên một trạm của mạng. Do ít chức năng phức tạp nên UDP thường có xu thế hoạt động nhanh hơn so với TCP. Nó thường được dùng cho các ứng dụng không đòi hỏi độ tin cậy cao trong tầng giao vận.



**Hình 5.11: Mô hình quan hệ hệ giao thức TCP/IP**

## 5.2. Mạng Internet

Cùng với sự phát triển của NFSNET và ARPANET nhất là khi giao thức TCP/IP đã trở thành giao thức chính thức duy nhất trên các mạng trên thì số lượng các mạng, nút muốn tham gia kết nối vào hai mạng trên đã tăng lên rất nhanh. Rất nhiều các mạng vùng được kết nối với nhau và còn liên kết với các mạng ở Canada, châu Âu.

Vào khoảng giữa những năm 1980 người ta bắt đầu thấy được sự hình thành của một hệ thống liên mạng lớn mà sau này được gọi là Internet. Sự phát triển của Internet được tính theo cấp số nhân, nếu như năm 1990 có khoảng 200.000 máy tính với 3.000 mạng con thì năm 1992 đã có khoảng 1.000.000 máy tính được kết nối, đến năm 1995 đã có hàng trăm mạng cấp vùng, chục ngàn mạng con và nhiều triệu máy tính. Rất nhiều mạng lớn đang hoạt động cũng đã được kết nối vào Internet như các mạng SPAN, NASA, HEPNET, BITNET, IBM, EARN. Việc liên kết các mạng được thực hiện thông qua rất nhiều đường nối có tốc độ rất cao.

Hiện nay một máy tính được gọi là thành viên của Internet nếu máy tính đó có giao thức truyền dữ liệu TCP/IP, có một địa chỉ IP trên mạng và nó có thể gửi các gói tin IP đến tất cả các máy tính khác trên mạng Internet.

Tuy nhiên trong nhiều trường hợp thông qua một nhà cung cấp dịch vụ Internet người sử dụng kết nối máy tính của mình với máy chủ của nhà cung cấp dịch vụ và được cung cấp một địa chỉ tạm thời trước khi khai thác các tài nguyên của Internet. Máy tính của người đó có thể gửi các gói tin cho các máy khác bằng địa chỉ tạm thời đó và địa chỉ đó sẽ trả lại cho nhà cung cấp khi kết thúc liên lạc. Vì máy tính của người đó sử dụng trong thời gian liên kết với Internet cũng có một địa chỉ IP nên người ta vẫn coi máy tính đó là thành viên của Internet.

Vào năm 1992 cộng đồng Internet đã ra đời nhằm thúc đẩy sự phát triển của Internet và điều hành nó. Hiện nay Internet có 5 dịch vụ chính:

- **Thư điện tử (Thư điện tử):** đây là dịch vụ đã có từ khi mạng ARPANET mới được thiết lập, nó cho phép gửi và nhận thư điện tử cho mọi thành viên khác trong mạng.
- **Thông tin mới (News):** các vấn đề thời sự được chuyển thành các diễn đàn cho phép mọi người quan tâm có thể trao đổi các thông tin cho nhau, hiện nay có hàng nghìn diễn đàn về mọi mặt trên Internet.
- **Đăng nhập từ xa (Remote Login):** bằng các chương trình như Telnet, Rlogin người sử dụng có thể từ một trạm của Internet đăng nhập (Logon) vào một trạm khác nếu như người đó được đăng ký trên máy tính kia.
- **Chuyển tệp (File transfer):** bằng chương trình FTP người sử dụng có thể chép các tệp từ một máy tính trên mạng Internet tới một máy tính khác. Người ta có thể chép nhiều phần mềm, cơ sở dữ liệu, bài báo bằng cách trên.
- **Dịch vụ WWW (World Wide Web):** WWW là một dịch vụ đặc biệt cung cấp thông tin từ xa trên mạng Internet. Các tập tin siêu văn bản được lưu trữ trên máy chủ sẽ cung cấp các thông tin và dẫn đường trên mạng cho phép người sử dụng dễ dàng truy cập các tập tin văn bản, đồ họa, âm thanh.



**Hình 5.12: Ví dụ một trang Web cho phép dễ dàng khai thác các trang Web khác**

Người sử dụng nhận được thông tin dưới dạng các trang văn bản, một trang là một đơn thể nằm trong máy chủ. Đây là dịch vụ đang mang lại sức thu hút to lớn cho mạng Internet, chúng ta có thể xây dựng các trang Web bằng ngôn ngữ HTML (Hypertext Markup Language) với nhiều dạng phong phú như văn bản, hình vẽ, video, âm thanh và có thể có các liên kết với các trang Web khác. Khi các trang đó được đặt trên các máy chủ Web thì thông qua Internet người ta có thể xem được sự thể hiện của các trang Web đó và có thể xem các trang Web khác mà nó chỉ đến.

Các phần mềm thông dụng được sử dụng hiện nay để xây dựng và duyệt các trang Web là Mosaic, Navigator của Netscape, Internet Explorer của Microsoft, Web Access của Novell.

### 5.3. Các ứng dụng trên TCP/IP

Phần này sẽ tìm hiểu một số ứng dụng mạng phổ biến hiện nay, chủ yếu tập trung vào giao thức hoạt động của chúng.

Ví dụ đầu tiên được xem xét là dịch vụ tên miền phân tán, dịch vụ đầu tiên được cài đặt trong một mạng máy tính. Về thực chất dịch vụ tên miền là cái mà các ứng dụng khác phải phụ thuộc vào. Một máy chủ phục vụ tên miền thường được sử dụng bởi các ứng dụng khác hơn là bởi con người.

Sau đó, các ứng dụng mạng truyền thống và phổ biến sẽ được giới thiệu, bao gồm các dịch vụ MAIL, WEB và FTP.

Cũng cần nói trước rằng, những dịch vụ mạng vừa nói sẽ dựa trên hai giao thức vận chuyển đã được đề cập trong chương 6 là TCP và UDP.

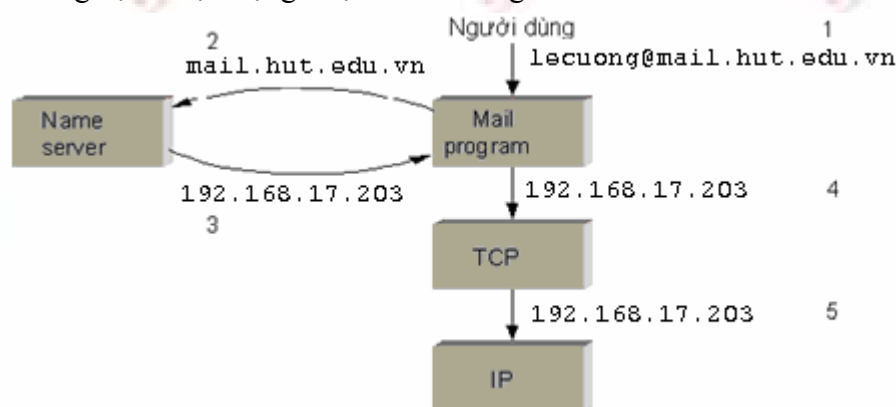
#### 5.3.1. Dịch vụ tên miền (DNS)

Cho đến bây giờ, chúng ta vẫn dùng địa chỉ để định danh các host. Trong khi rất thuận tiện cho việc xử lý của các router, các địa chỉ số không thân thiện với người dùng lắm. Vì lý do này, các host thường được gán cho một cái tên thân thiện và dịch vụ tên được

sử dụng để ánh xạ từ cái tên thân thiện với người dùng này sang địa chỉ số vốn rất thân thiện với các router. Dịch vụ như vậy thường là ứng dụng đầu tiên được cài đặt trong một mạng máy tính do nó cho phép các ứng dụng khác tự do định danh các host bằng tên thay vì bằng địa chỉ. Dịch vụ tên thường được gọi là phần trung gian (middleware) vì nó lấp đầy khoảng cách giữa các ứng dụng khác và lớp mạng phía dưới. Tên host và địa chỉ host khác nhau ở hai điểm quan trọng. Thứ nhất, tên host thường có độ dài thay đổi và dễ gọi nhớ, vì thế nó giúp người dùng dễ nhớ hơn. Thứ hai, tên thường không chứa thông tin gì để giúp mạng định vị (chuyển các gói tin đến) host. Ngược lại địa chỉ hàm chứa thông tin vạch đường trong đó.

Trước khi đi vào chi tiết cách thức đặt tên cho các host trong mạng như thế nào, chúng ta đi định nghĩa một số thuật ngữ trước:

- Không gian tên (Name Space) định nghĩa tập các tên có thể có. Một không gian tên có thể là phẳng (Flat) – một tên không thể được chia thành các thành phần nhỏ hơn, hoặc phân cấp.
- Hệ thống tên duy trì một tập các ánh xạ (Collection of Bindings) từ tên sang giá trị. Giá trị có thể là bất cứ thứ gì chúng ta muốn hệ thống tên trả về khi ta cấp cho nó một tên để ánh xạ, trong nhiều trường hợp giá trị chính là địa chỉ host.
- Một cơ chế phân giải (Resolution Mechanism) là một thủ tục mà khi được gọi với tham số là một tên, sẽ trả về một giá trị tương ứng.
- Một máy chủ tên (Name Server) là một kết quả cài đặt của cơ chế phân giải luôn sẵn dùng trên mạng và có thể được truy vấn bằng cách gửi đến nó một thông điệp.
- Mạng Internet đã có sẵn một hệ thống đặt tên được phát triển tốt, gọi là hệ thống tên miền (Domain Name System – DNS). Vì thế chúng ta sẽ dùng DNS làm cơ sở để thảo luận về vấn đề đặt tên cho các host. Khi người dùng đưa một tên host đến một ứng dụng (có thể tên host đó là một phần của một tên hỗn hợp như địa chỉ thư điện tử chẳng hạn), ứng dụng này sẽ liên hệ với hệ thống tên để dịch tên host sang địa chỉ host. Sau đó ứng dụng liền tạo một nối kết đến host đó thông qua giao thức TCP chẳng hạn. Hiện trạng được mô tả trong hình 5.13.



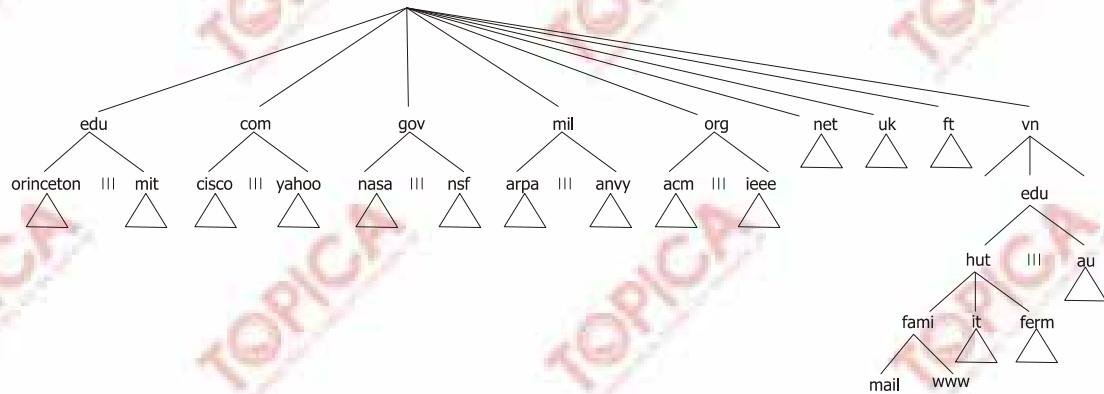
**Hình 5.13:** Tên máy được dịch sang địa chỉ, các số từ 1-5 thể hiện trình tự các bước xử lý

### 5.3.1.1. Miền phân cấp

DNS cài đặt không gian tên phân cấp dùng cho các đối tượng trên Internet. Các tên DNS được xử lý từ phải sang trái, sử dụng các dấu chấm (.) làm ký tự ngăn cách. (Mặc dù các tên DNS được xử lý từ phải qua trái, người dùng thường đọc chúng từ trái sang phải). Ví dụ tên miền của một host là mail.hut.edu.vn. Chú ý rằng các tên



miền được sử dụng để đặt tên các đối tượng trên Internet, không phải chỉ được dùng để đặt tên máy. Ta có thể tưởng tượng cấu trúc phân cấp của DNS giống như hình dáng cây. Hình 5.14 là một ví dụ.



**Hình 5.14: Cây phân cấp tên miền**

Có thể thấy rằng, cây phân cấp không quá rộng ở mức đầu tiên. Mỗi quốc gia có một tên miền, ngoài ra còn có 6 miền lớn khác gồm: edu, com, gov, mil, org và net. Sáu miền lớn này nằm ở Mỹ. Những tên miền không chỉ ra tên nước một cách tường minh thì mặc nhiên là nằm ở Mỹ.

### 5.3.1.2. Các máy chủ phục vụ tên

Một cấu trúc tên miền phân cấp hoàn chỉnh chỉ tồn tại trong ý niệm. Vậy thì trong thực tế cấu trúc phân cấp này được cài đặt như thế nào? Bước đầu tiên là chia cấu trúc này thành các cây con gọi là các vùng (Zone). Mỗi một vùng có thể được xem là đơn vị quản lý một bộ phận của toàn hệ thống phân cấp. Ví dụ, vùng cao nhất của hệ thống phân cấp được quản lý bởi NIC (Network Information Center), vùng HUT được quản lý bởi trường đại học.

Một vùng luôn có mối liên hệ đến các đơn vị cài đặt cơ bản trong DNS - các máy chủ tên. Thông tin chứa trong một vùng được thiết lập tại hai hoặc nhiều máy chủ tên. Mỗi máy chủ tên có thể truy xuất được qua mạng Internet. Máy trạm gửi yêu cầu đến máy chủ tên, máy chủ tên sẽ trả lời cho yêu cầu đó. Câu trả lời đôi khi chứa thông tin cuối cùng mà máy trạm cần, đôi khi lại chứa chỉ điểm đến một máy chủ tên khác mà máy trạm nên gửi câu hỏi đến đó. Vì thế, theo cách nhìn thiên về cài đặt, người ta có thể nghĩ về DNS được cài đặt bằng cấu trúc phân cấp các máy chủ tên hơn là bằng cấu trúc phân cấp tên miền. Để ý rằng mỗi vùng được cài đặt trong hai hoặc nhiều máy chủ tên với lý do dự phòng; nghĩa là nếu một máy chủ bị chết sẽ còn các máy chủ khác thay thế. Mặt khác, một máy chủ tên cũng có thể được dùng để cài đặt nhiều hơn một vùng.

Mỗi máy chủ tên quản lý thông tin về một vùng dưới dạng một tập các mẫu tin tài nguyên (Resource Record). Mỗi mẫu tin tài nguyên là một ánh xạ từ tên sang giá trị (Name to Value Binding), cụ thể hơn là một mẫu tin gồm 5 trường: (Tên, Giá trị, Kiểu, Lớp, TTL)

Các trường Tên và Giá trị là những gì chúng ta muốn có, ngoài tra trường Kiểu chỉ ra cách thức mà Giá trị được thông dịch. Chẳng hạn, trường Kiểu = A chỉ ra rằng Giá trị là một địa chỉ IP. Vì thế các mẫu tin kiểu A sẽ cài đặt kiểu ánh xạ từ tên miền sang địa chỉ IP:

- NS: Trường Giá trị chỉ ra tên miền của máy tính đang chạy dịch vụ tên và dịch vụ đó có khả năng thông dịch các tên trong một miền cụ thể.



- CNAME: Trường Giá trị chỉ ra một cái tên giả của một host nào đó. Kiểu này được dùng để đặt thêm bí danh cho các host trong miền.
- MX: Trường Giá trị chỉ ra tên miền của host đang chạy chương trình Mail Server mà máy chủ đó có khả năng tiếp nhận những thông điệp thuộc một miền cụ thể.
- Trường Lốp được sử dụng nhằm cho phép thêm vào những thực thể mạng không do NIC quản lý. Ngày nay, lốp được sử dụng rộng rãi nhất là loại được Internet sử dụng, nó được ký hiệu là IN.

Cuối cùng trường TTL chỉ ra mẫu tin tài nguyên này sẽ hợp lệ trong bao lâu. Trường này được sử dụng bởi những máy chủ đang trữ tạm các mẫu tin của máy chủ khác; khi trường TTL hết hạn, các mẫu tin chứa trường TTL hết hạn đó sẽ bị các máy chủ xóa khỏi Cache của mình.

Chú ý rằng trên lý thuyết các mẫu tin có thể được dùng để định nghĩa bất kỳ kiểu đối tượng nào, DNS lại thường được sử dụng để định danh các host và site. DNS không được dùng để định danh cá nhân con người hoặc các đối tượng khác như tập tin hay thư mục, việc định danh này được thực hiện trong các hệ thống phục vụ tên khác. Ví dụ X.500 là hệ thống định danh của ISO được dùng để định danh con người bằng cách cung cấp thông tin về tên, chức vụ, số điện thoại, địa chỉ,... X.500 đã chứng tỏ là quá phức tạp nên không được hỗ trợ bởi các máy tìm kiếm nổi tiếng hiện nay. Tuy nhiên nó lại là nguồn gốc phát sinh ra chuẩn LDAP (Lightweight Directory Access Protocol). LDAP vốn là thành phần con của X.500 được thiết kế để làm phần đầu cuối cho X.500. Ngày nay LDAP đang trở nên phổ biến nhất là ở cấp độ công ty, tổ chức lớn, đóng vai trò là hệ thống học và quản lý thông tin về người dùng của nó.

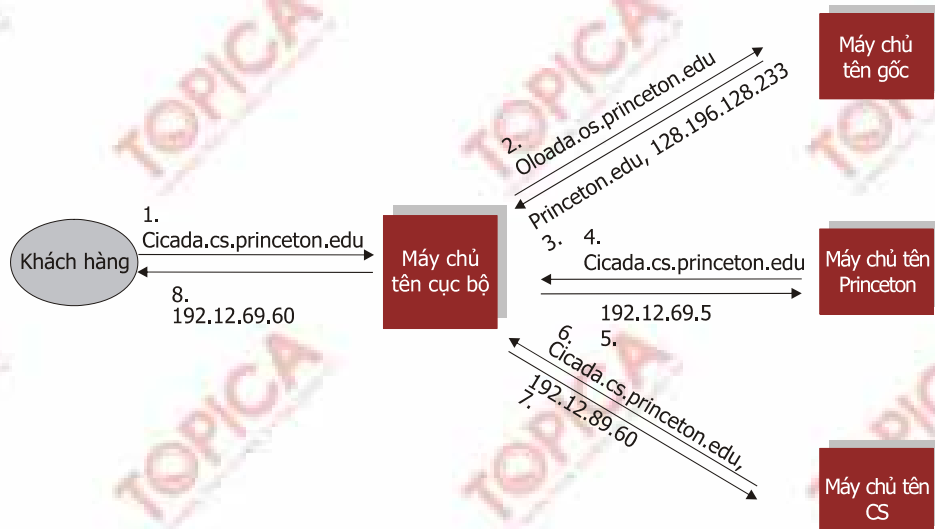
### 5.3.1.3. Phương pháp phân tích tên

Với một hệ thống phân cấp các máy chủ tên đã trình bày, bây giờ chúng ta đi tìm hiểu cách thức một khách hàng giao tiếp với các máy chủ này để phân tích cho được một tên miền thành địa chỉ. Giả sử một khách hàng muốn phân tích tên miền `www.hut.edu.vn`, đầu tiên khách hàng này sẽ gửi yêu cầu chứa tên này đến máy chủ tên gốc. Máy chủ gốc không thể so khớp tên theo yêu cầu với các tên mà nó chứa, liền trả lời cho khách hàng một mẫu tin kiểu NS chứa `edu.vn`. Máy chủ gốc cũng trả về tất cả các mẫu tin có liên quan đến mẫu tin NS trên, trong đó có mẫu tin kiểu A chứa địa chỉ của `dns1.vnnic.vnn.vn`. Khách hàng chưa có thông tin cuối cùng mà nó muốn, tiếp tục gửi yêu cầu đến máy chủ tên tại địa chỉ `192.168.17.203`. Máy chủ tên thứ hai này lại không thể so khớp tên theo yêu cầu với các tên mà nó chứa, tiếp tục trả lời cho khách hàng một mẫu tin loại NS chứa tên `hut.edu.vn` cùng với mẫu tin kiểu A tương ứng với tên máy chủ là `ns.hut.edu.vn`. Khách hàng lại tiếp tục gửi yêu cầu đến máy chủ tên tại địa chỉ `192.168.17.203` và lần này nhận được câu trả lời cuối cùng có kiểu A cho tên `www.hut.edu.vn`.

Ví dụ trên chắc chắn sẽ để lại nhiều câu hỏi về quá trình phân giải tên. Câu hỏi thường được đặt ra là: Lúc khởi đầu, làm sao khách hàng có thể định vị được máy chủ gốc? Đây là bài toán cơ bản đặt ra cho mọi hệ thống phục vụ tên và câu trả lời là: hệ thống phải tự thân vận động để có được thông tin về các máy chủ gốc! Trong tình huống của hệ thống DNS, ánh xạ từ tên sang địa chỉ của một hay nhiều máy chủ gốc được phổ

biến cho mọi người, nghĩa là ánh xạ đó được loan báo thông qua các phương tiện truyền thông khác nằm ngoài hệ thống tên.

Tuy nhiên, trong thực tế không phải tất cả khách hàng đều biết về các máy chủ gốc. Thay vào đó, chương trình khách hàng chạy trên mỗi host trong Internet được khởi động với các địa chỉ lấy từ máy chủ tên cục bộ. Từ đó, việc phân giải một tên miền bắt đầu từ câu truy vấn của khách hàng đến máy chủ cục bộ. Nếu máy chủ cục bộ không có sẵn câu trả lời, nó sẽ gửi câu hỏi đến máy chủ từ xa giúp cho khách hàng. Chuỗi hành động trên có thể được mô tả trong hình sau:



Hình 5.15: Quá trình phân giải tên trong thực tế, các số 1 đến 8 chỉ ra trình tự thực hiện

### 5.3.2. Electronic Mail (SMTP, MIME, POP3, IMAP)

Thư điện tử là một trong những ứng dụng mạng lâu đời nhất và phổ dụng nhất. Thử nghĩ khi bạn muốn gửi thông điệp đến một người bạn ở đầu kia của thế giới, bạn muốn mang thư chạy bộ qua đó hay chỉ đơn giản lên máy tính gõ ít hàng và nhấn nút Send? Thật ra, những bậc tiền bối của mạng ARPANET đã không tiên đoán được thư điện tử sẽ là ứng dụng then chốt chạy trên mạng này, mục tiêu chính của họ là thiết kế hệ thống cho phép truy cập tài nguyên từ xa. Hệ thống thư điện tử ra đời không mấy nổi bật, để bây giờ lại được sử dụng hằng ngày bởi hàng triệu người trên thế giới.

Mục tiêu của phần này là chỉ ra những thành phần hoạt động trong hệ thống thư điện tử, vai trò của nó, giao thức mà họ sử dụng và khuôn dạng thông điệp mà cách thức trao đổi với nhau.

#### 5.3.2.1. Các thành phần của hệ thống thư điện tử

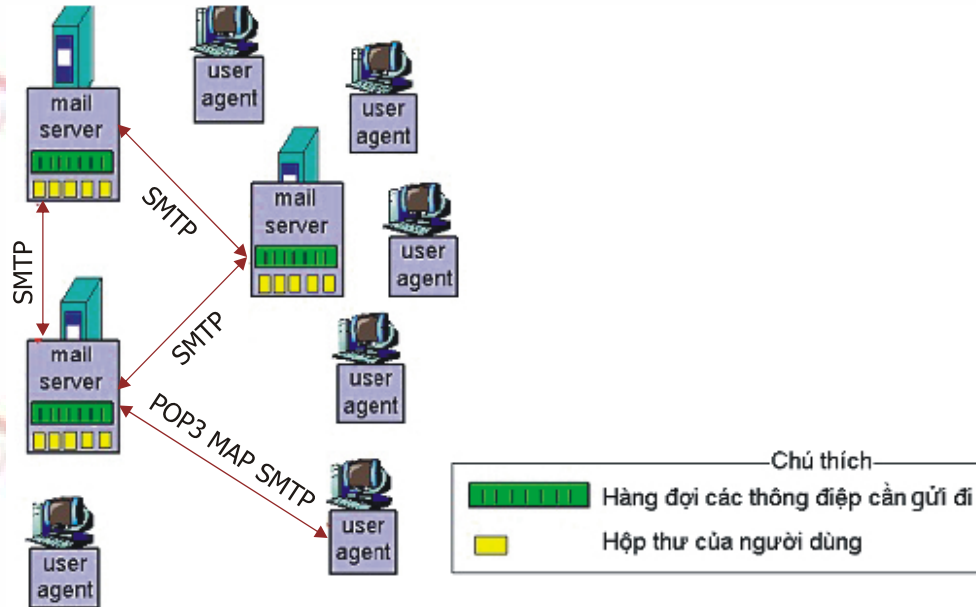
Một hệ thống thư điện tử thường có 3 thành phần chính: Bộ phận trợ giúp người dùng (User Agent), Mail Server và các giao thức mà các thành phần này dùng để giao tiếp với nhau.

Người ta phân loại các giao thức như sau:

- Giao thức giữa các Mail Server bao gồm:

SMTP (Simple Mail Transfer Protocol): được các máy chủ dùng để chuyển thư qua lại với nhau. Hình dung nó giống như cách thức mà các trạm bưu điện dùng để chuyển các thùng thư của khách hàng cho nhau. Thông tin chi tiết về giao thức này được mô tả trong tài liệu RFC 822.

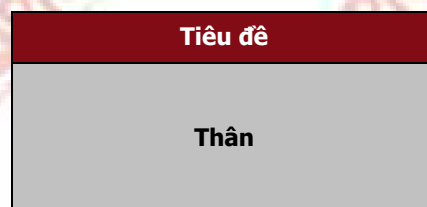
- Giao thức giữa *Mail Server* và *User Agent* bao gồm:
  - POP3 (Post Office Protocol version 3 [RFC 1939]): được user agent sử dụng để lấy thư về từ hộp thư của nó trên máy chủ.
  - SMTP: được user agent sử dụng để gửi thư ra máy chủ.
  - IMAP: (Internet Mail Access Protocol [RFC 1730]): có nhiều tính năng vượt trội hơn POP3. Ngoài ra IMAP còn cho phép gửi thư điện tử.



**Hình 5.16: Các thành phần của hệ thống thư điện tử**

### 5.3.2.2. Khuôn dạng của một thư điện tử

RFC 822 định nghĩa một thư điện tử gồm có hai phần: phần tiêu đề (header) và phần thân (body).



**Hình 5.17: Khuôn dạng của thư điện tử**

Cả hai phần đều được thể hiện dưới dạng ký tự ASCII. Lúc đầu, phần thân được quy định có khuôn dạng văn bản đơn giản. Sau này người ta đề nghị một chuẩn mới gọi là MIME, có thể cho phép phần thân của thư điện tử chứa bất kỳ loại dữ liệu nào.

Phần tiêu đề bao gồm nhiều dòng thông tin, mỗi dòng kết thúc bằng hai ký tự <CRLF>. Phần tiêu đề được chia khỏi phần thân bởi một hàng rỗng. Mỗi một hàng tiêu đề chứa một cặp “tên” và “giá trị”, cách nhau bởi dấu hai chấm (:). Người dùng có thể rất quen với nhiều hàng tiêu đề vì họ thường phải điền thông tin vào đấy. Ví dụ

Tên	Giá trị
From:	Địa chỉ người gửi
To:	Địa chỉ người nhận
Subject:	Chủ đề thư
Date:	Ngày gửi

RFC 822 được mở rộng năm 1993 (và được cập nhật lại năm 1996) để cho phép thư điện tử mang được nhiều loại dữ liệu: âm thanh, video, hình ảnh, tài liệu Word, ... MIME (Multipurpose Internet Mail Extensions) về cơ bản có ba phần. Phần đầu tiên là tập các dòng header dùng để bổ túc cho phần header cũ của RFC 822. Theo nhiều cách, những dòng header này mô tả dữ liệu chứa trong phần thân. Cụ thể như sau:

Tên	Giá trị
MIME – Version:	Phiên bản MIME đang sử dụng.
Content – Description:	Mô tả trong thư đang có dữ liệu gì.
Content – Type:	Mô tả kiểu dữ liệu đang nằm trong thư.
Content – Transfer – Encoding:	Mô tả cách thức mã hóa dữ liệu trong thư.

Phần thứ hai là các định nghĩa cho một tập các kiểu nội dung (và kiểu con nếu có). Ví dụ một số kiểu mà MIME định nghĩa:

Kiểu	Ý nghĩa
Image/gif	Ảnh dạng gif.
Image/jpeg	Ảnh dạng jpeg.
Text/plain	Văn bản đơn giản.
Text/richtext	Văn bản mở rộng (có đặt font chữ, được định dạng đậm, nghiêng hoặc gạch dưới ...).
Application	Dữ liệu trong thư được xuất ra từ một ứng dụng nào đó. Chẳng hạn: Application/postscript: Tài liệu Postscript (.ps). Application/msword: Tài liệu Microsoft Word (.doc).

MIME cũng định nghĩa kiểu multipart để chỉ ra cách mà phần thân của thư mang nhiều loại dữ liệu khác nhau như thế nào. Chỉ có một kiểu con của multipart là hỗn hợp (mixed) với ý nói rằng trong phần thân của thư có nhiều mảnh dữ liệu khác nhau, độc lập với nhau và được sắp xếp theo một trình tự cụ thể. Mỗi mảnh dữ liệu sẽ có phần tiêu đề riêng để mô tả kiểu dữ liệu của mảnh đó.

Phần thứ ba mô tả cách thức mã hóa các kiểu dữ liệu nói trên để có thể truyền chúng dưới dạng ASCII. Lý do để mọi bức thư phải chứa các ký tự ASCII là vì để đi được đến đích, bức thư đó có thể phải trung chuyển qua nhiều gateway, mà các gateway này đều coi mọi bức thư dưới dạng ASCII. Nếu trong thư chứa bất kỳ ký tự nào khác ASCII thì thư sẽ bị đứt gãy nội dung. MIME sử dụng phương pháp mã hóa trực tiếp dữ liệu nhị phân thành các ký tự nhị phân, gọi là base64. Ý tưởng của base64 là ánh xạ 3 bytes dữ liệu nhị phân nguyên thủy thành 4 ký tự ASCII. Giải thuật đơn giản như sau: tập hợp 3 bytes dữ liệu nhị phân lại thành 24 bits, sau đó chia 24 bits này thành 4 cụm, mỗi cụm 6 bits. Một cụm 6 bits được ánh xạ vào một trong 64 ký tự ASCII hợp lệ, ví dụ 0 ánh xạ thành A, 1 ánh xạ thành B... Nếu nhìn vào bức thư đã được mã hóa dạng base64, người dùng sẽ thấy chỉ có 52 chữ cái cả hoa lẫn thường, 10 chữ số từ 0 đến 9 và các ký tự đặc biệt + và /.

Đối với những người dùng chỉ sử dụng trình đọc thư hỗ trợ duy nhất kiểu ký tự thì việc đọc những bức thư có kiểu base64 sẽ rất là phức tạp. Vì lý do nhân đạo, MIME còn hỗ trợ kiểu mã hóa ký tự thường được gọi là 7-bit. 7-bit sẽ giữ nguyên dạng ký tự mà người ta nhập vào.



### 5.3.2.3. Chuyển thư

Kể đến, chúng ta sẽ xem xét giao thức SMTP – giao thức được dùng để chuyển thư từ máy này đến máy kia. Để đặt SMTP vào đúng ngữ cảnh, chúng ta nên nhắc lại các nhân vật then chốt trong hệ thống thư điện tử. Đầu tiên, người dùng tương tác với trình đọc thư (hay còn gọi là user agent) để soạn, lưu, tìm kiếm và đọc thư của họ. Hiện trên thị trường có nhiều phần mềm đọc thư, cũng giống như hiện cũng đang có nhiều loại trình duyệt Web vậy. Thứ hai, có trình xử lý thư (hay còn gọi là Mail Server) chạy trên một máy nào đó trong mạng nội bộ của người dùng. Có thể xem Mail Server như một bưu điện: Người dùng trao cho Mail Server các bức thư mà họ muốn gửi cho người dùng khác, Mail Server sử dụng giao thức SMTP trên TCP để chuyển bức các thư này đến Mail Server bên đích. Mail Server bên đích nhận các thư đến và đặt chúng vào hộp thư của người dùng bên đích. Do SMTP là giao thức mà rất nhiều người có thể tự cài đặt, vì thế sẽ có rất nhiều sản phẩm Mail Server hiện có trên thị trường. Sản phẩm Mail Server thường được sử dụng nhất là sendmail, ban đầu được cài đặt trong hệ điều hành Berkeley Unix.

Tất nhiên Mail Server bên máy gửi có thể kết nối SMTP/TCP trực tiếp tới Mail Server bên máy nhận, nhưng trong thực tế, một bức thư có thể đi ngang qua vài mail gateways trước khi đến đích. Cũng giống như máy đích, mỗi mail gateway cũng chạy một Mail Server. Không phải ngẫu nhiên mà các nút chuyển thư trung gian được gọi là mail gateway. Công việc của chúng cũng giống như các IP gateway là lưu tạm và chuyển phát tiếp các bức thư của người dùng. Điểm khác nhau duy nhất giữa chúng là, mail gateway trữ tạm các bức thư trong đĩa, trong khi các IP gateway trữ tạm các gói tin IP trong bộ nhớ.

Bạn có thể đặt câu hỏi: tại sao lại cần đến các mail gateways? Tại sao không dùng phương pháp nối kết SMTP/TCP trực tiếp từ bên gửi sang bên nhận? Lý do thứ nhất, người gửi không muốn kèm trong thư địa chỉ của máy đích. Ví dụ, riêng việc nhập vào trong thư địa chỉ đích `dhbk@mail.hut.edu.vn` đã rất mất công, không ai thấy thoải mái khi phải nhập thêm địa chỉ máy đích là `machine-of-dhbk.hut.edu.vn`. Thứ hai, không chắc lúc bên gửi thiết lập nối kết đến bên nhận, người dùng bên nhận đã bật sẵn máy! Thành thử chỉ cần địa chỉ thư bên nhận là đủ. Khi bức thư đến được mail gateway của Đại học Bách khoa Hà Nội, nếu người dùng dhbk đang mở máy, mail gateway sẽ chuyển thư cho anh ta ngay, nếu không mail gateway sẽ trữ tạm thư trên đĩa của nó đến khi dhbk bật máy lên và kiểm tra thư.

Dù có bao nhiêu mail gateways trung gian trên đường đến đích vẫn không đáng lo lắng, bởi vì mỗi mail gateway trung gian sẽ nỗ lực sử dụng một kết nối SMTP độc lập đến gateway kế tiếp trên đường đi nhằm chuyển thư càng ngày càng đến gần người nhận.

SMTP là một giao thức đơn giản dùng các ký tự ASCII. Sau khi thiết lập nối kết TCP đến cổng 25 của máy đích (được coi là máy chủ), máy nguồn (được coi là máy trạm) chờ nhận kết quả trả về từ máy chủ. Máy chủ khởi đầu cuộc đối thoại bằng cách gửi một dòng văn bản đến máy trạm thông báo danh tính của nó và khả năng tiếp nhận thư. Nếu máy chủ không có khả năng nhận thư tại thời điểm hiện tại, máy trạm sẽ hủy bỏ nối kết và thử thiết lập lại nối kết sau.

Nếu máy chủ sẵn sàng nhận thư, máy trạm sẽ thông báo lá thư đó từ đâu đến và ai sẽ là người nhận. Nếu người nhận đó tồn tại, máy chủ sẽ thông báo cho máy trạm tiếp tục



gửi thư. Sau đó máy trạm gửi thư và máy chủ báo nhận cho thư đó. Sau khi cả hai bên hoàn tất phiên truyền nhận, kết nối sẽ được đóng lại.

LỆNH CỦA MÁY TRẠM	
Lệnh	Ý nghĩa
HELLO	Câu chào và xưng danh của máy trạm.
MAIL FROM	Địa chỉ thư điện tử của người gửi.
RCPT TO	Địa chỉ thư điện tử của người nhận.
DATA	Bắt đầu truyền nội dung của thư.
QUIT	Hủy nối kết.
TRẢ LỜI CỦA MÁY CHỦ	
Lệnh	Ý nghĩa
250	Yêu cầu hợp lệ.
550	Yêu cầu không hợp lệ, không tồn tại hộp thư như máy trạm đã chỉ ra.
354	Cho phép bắt đầu nhập thư vào. Kết thúc bằng <CRLF>.<CRLF>.
221	Máy chủ đang đóng kết nối TCP.

Vẫn còn nhiều lệnh và mã trả lời chưa được trình bày, xin tham khảo tài liệu RFC 822 để có được đầy đủ thông tin.

#### 5.3.2.4. Phân phát thư

Như đã trình bày, khi đứng về góc độ người dùng thư, họ sẽ dùng user agent để gửi và nhận thư cho họ. User agent dùng giao thức SMTP để gửi thư đi, dùng giao thức POP3 hoặc IMAP để nhận thư về.

##### POP3

Một phiên làm việc theo giao thức POP3 bắt đầu tại user agent. User agent khởi động một nối kết TCP đến cổng 110 của Mail Server. Khi kết nối thực hiện xong, phiên làm việc POP3 sẽ trải qua theo thứ tự ba kỳ:

1. Chứng thực.
2. Giao dịch dữ liệu.
3. Cập nhật.

Kỳ chứng thực buộc người dùng thực hiện thủ tục đăng nhập bằng cách nhập vào hai lệnh sau:

Lệnh	Ý nghĩa
USER <tên người dùng>	Khai báo của người dùng là đúng.
PASS <mật khẩu>	Khai báo của người dùng là sai và lời giải thích.

Báo trả của Mail Server sẽ là một trong hai câu sau:

Lệnh	Ý nghĩa
+ OK <chú thích>	Khai báo của người dùng là đúng.
+ ERR <chú thích>	Khai báo của người dùng là sai và lời giải thích.

Trong kỳ giao dịch, người dùng có thể xem danh sách thư chưa nhận về, nhận thư về và xóa thư trong hộp thư của mình khi cần thiết. Các lệnh mà người dùng thường sử dụng để giao dịch với máy chủ là:

Lệnh	Ý nghĩa
LIST [<số thứ tự thư>]	Nếu dùng LIST không tham số, server sẽ trả về toàn bộ danh sách các thư chưa nhận. Nếu có tham số là số thứ tự thư cụ thể, server sẽ trả về thông tin của chỉ bức thư đó thôi.
RETR <số thứ tự thư>	Tải lá thư có số thứ tự <số thứ tự thư> về.
DELE <số thứ tự thư>	Xóa lá thư số <số thứ tự thư> khỏi hộp thư.
QUIT	Hoàn tất giai đoạn giao dịch và hủy kết nối TCP.

Các trả lời của máy chủ có thể là các số liệu mà máy trạm yêu cầu hoặc các thông báo +OK, -ERR như trong phần đăng nhập.

### IMAP

Với những người dùng có một tài khoản thư điện tử trên một ISP và người dùng này thường truy cập thư điện tử trên một PC thì giao thức POP3 hoạt động tốt. Tuy nhiên, một sự thật trong ngành công nghệ máy tính, khi một thứ gì đó đã hoạt động tốt, người ta lập tức đòi hỏi thêm nhiều tính năng mới (và tự chuốc lấy nhiều phiền nhiễu). Điều đó cũng xảy ra đối với hệ thống thư điện tử. Ví dụ, người ta chỉ có một tài khoản thư điện tử, nhưng họ lại muốn ngồi đâu cũng truy cập được nó. POP3 cũng làm được chuyện này bằng cách đơn giản tải hết các thư điện tử xuống máy PC mà người dùng này đang ngồi làm việc, và dĩ nhiên là thư từ của người dùng này nằm rải rác khắp nơi.

Sự bất tiện này khơi mào cho sự ra đời của giao thức phân phối thư mới, IMAP (Internet Message Access Protocol), được định nghĩa trong RFC 2060. Không giống như POP2, IMAP coi các thông điệp mặc nhiên nằm trên máy chủ vô hạn và trên nhiều hộp thư. IMAP còn đưa ra cơ chế cho phép đọc các thông điệp hoặc một phần của thông điệp, một tính năng hữu ích khi người dùng kết nối đến máy chủ bằng đường truyền tốc độ chậm như điện thoại nhưng lại đọc các thư điện tử có âm thanh, hình ảnh... Với quan niệm cho rằng người dùng không cần tải thư về lưu trên PC, IMAP cung cấp các cơ chế cho phép tạo, xóa và sửa đổi nhiều hộp thư trên máy chủ.

Cung cách làm việc của IMAP cũng giống như POP3, ngoài trừ trong IMAP có rất nhiều lệnh. IMAP máy chủ sẽ lắng nghe trên cổng 143. Cũng nên chú ý rằng, không phải mọi ISP đều hỗ trợ cả hai giao thức POP3 và IMAP.

**Bảng sau so sánh các tính năng của POP3 và IMAP**

Tính năng	POP3	IMAP
Giao thức được định nghĩa ở đâu?	RFC 1939	RFC 2060
Cổng TCP được dùng	110	143
Thư điện tử được lưu ở đâu?	PC của người dùng	Server
Thư điện tử được đọc ở đâu?	Off – line	On – line
Thời gian kết nối	Ít	Nhiều
Sử dụng tài nguyên của máy chủ	Tối thiểu	Nhiều hơn
Nhiều hộp thư	Không	Đúng

Tính năng	POP3	IMAP
AI lưu phòng hồ các hộp thư	Người dùng	ISP
Tốt cho người dùng di động	Không	Có
Kiểm soát của người dùng đối với việc tải thư về	Ít	Tốt
Tải một phần thư	Không	Có
Quota đĩa có là vấn đề không?	Không	Thỉnh thoảng
Dễ cài đặt	Có	Không
Được hỗ trợ rộng rãi	Có	Đang phát triển

### 5.3.3. World Wide Web (HTTP)

Ứng dụng Web đã rất thành công, giúp cho nhiều người có thể truy cập Internet đến nỗi Web được hiểu đồng nghĩa với Internet! Có thể hiểu Web như là một tập các máy trạm và máy chủ hợp tác với nhau và cùng nói chung một ngôn ngữ: HTTP (Hyper Text Transfer Protocol). Đa phần người dùng tiếp xúc với Web thông qua chương trình máy trạm có giao diện đồ họa, hay còn gọi là trình duyệt Web (Web Browser). Các trình duyệt Web thường được sử dụng nhất là Netscape Navigator (của Netscape) và Internet Explorer (của Microsoft).

Bất kỳ trình duyệt Web nào cũng có chức năng cho phép người dùng “mở một URL”. Các URL (Uniform Resource Locators) cung cấp thông tin về vị trí của các đối tượng trên Internet, chúng thường trông giống như sau: <http://www.hut.edu.vn/index.html>.

Nếu người dùng mở URL trên, trình duyệt Web sẽ thiết lập một kết nối TCP đến Web Máy chủ tại địa chỉ [www.hut.edu.vn](http://www.hut.edu.vn) và tải tập tin [index.html](http://www.hut.edu.vn/index.html) về và thể hiện nó. Hầu hết các tập tin trên Web chứa văn bản và hình ảnh, một số còn chứa âm thanh và video. Chúng còn có thể chứa các liên kết đến các tập tin khác – được gọi là các liên kết siêu văn bản (Hypertext Links). Khi người dùng yêu cầu trình duyệt Web mở ra một liên kết siêu văn bản (bằng cách trỏ chuột và nhấp chuột lên liên kết đó), trình duyệt sẽ mở một nối kết mới, tải về và hiển thị một tập tin mới. Vì thế, rất dễ để duyệt từ máy chủ này đến máy chủ khác trên khắp thế giới để có được hết những thông tin mà người dùng cần.

Khi người dùng chọn xem một trang Web, trình duyệt Web sẽ nạp trang Web đó từ Web máy chủ về sử dụng giao thức HTTP chạy trên TCP. Giống như SMTP, HTTP là giao thức hướng ký tự. Về cốt lõi, một thông điệp HTTP có khuôn dạng tổng quát sau:

```
START_LINE <CRLF>
MESSAGE_HEADER <CRLF>
<CRLF>
MESSAGE_BODY <CRLF>
```

Hàng đầu tiên chỉ ra đây là thông điệp yêu cầu hay trả lời. Nó sẽ chỉ ra “thủ tục cần được thực hiện từ xa” (trong tình huống là thông điệp yêu cầu) hoặc là “trạng thái trả về” (trong tình huống là thông điệp trả lời). Tập hợp các hàng kế tiếp chỉ ra các tùy chọn hoặc tham số nhằm xác định cụ thể tính chất của yêu cầu hoặc trả lời. Phần MESSAGE\_HEADER có thể không có hoặc có một vài hàng tham số và được kết thúc bằng hàng trống.

HTTP định nghĩa nhiều kiểu header, một số liên quan đến các thông điệp yêu cầu, một số liên quan đến các thông điệp trả lời và một số lại liên quan đến phần dữ liệu trong thông điệp. Ở đây chỉ giới thiệu một số kiểu thường dùng. Cuối cùng, sau hàng trống là phần nội dung của thông điệp trả lời (MESSAGE\_BODY), phần này thường là rỗng trong thông điệp yêu cầu.

### 5.3.3.1. Các thông điệp yêu cầu

Hàng đầu tiên của một thông điệp yêu cầu HTTP sẽ chỉ ra ba thứ: thao tác cần được thực thi, trang Web mà thao tác đó sẽ áp lên và phiên bản HTTP được sử dụng. Bảng sau sẽ giới thiệu một số thao tác phổ biến.

Hành động	Mô tả
OPTIONS	Yêu cầu thông tin về các tùy chọn hiện có.
GET	Lấy về tài liệu được xác định trong đường dẫn.
HEAD	Lấy về thông tin thô về tài liệu được xác định trong đường dẫn.
POST	Cung cấp thông tin cho máy chủ.
PUT	Tải tài liệu lên server và đặt ở vị trí được xác định trong đường dẫn.
DELETE	Xóa tài liệu nằm ở vị trí đường dẫn trên máy chủ.
TRACE	Phản hồi lại thông điệp yêu cầu.
CONNECT	Được sử dụng bởi các proxy.

Hai thao tác thường được sử dụng nhiều nhất là GET (lấy một trang Web về) và HEAD (lấy về thông tin của một trang Web). GET thường được sử dụng khi trình duyệt muốn tải một trang Web về và hiển thị nó cho người dùng. HEAD thường được sử dụng để kiểm tra tính hợp lệ của một liên kết siêu văn bản hoặc để xem một trang nào đó có bị thay đổi gì không kể từ lần tải về trước đó.

**Ví dụ:** dòng START\_LINE

GET [http:// www.hut.edu.vn /index.html](http://www.hut.edu.vn/index.html) HTTP/1.1

nói rằng người dùng muốn tải về trên máy chủ [www.hut.edu.vn](http://www.hut.edu.vn) trang Web có tên index.html và hiển thị nó. Ví dụ trên dùng đường dẫn tuyệt đối. Ta cũng có thể sử dụng tương đối như sau:

GET /index.html HTTP/1.1 Host: [www.hut.edu.vn](http://www.hut.edu.vn) /index.html

Ở đây, Host là một trong các trường trong MESSAGE\_HEADER

### 5.3.3.2. Các thông điệp trả lời

Giống như các thông điệp yêu cầu, các thông điệp trả lời bắt đầu bằng một hàng START\_LINE. Trong trường hợp này, dòng START\_LINE sẽ chỉ ra phiên bản HTTP đang được sử dụng, một mã ba ký số xác định yêu cầu là thành công hay thất bại và một chuỗi ký tự chỉ ra lý do của câu trả lời này.

**Ví dụ:** dòng START\_LINE



HTTP/1.1 202 Accepted chỉ ra máy chủ đã có thể thỏa mãn yêu cầu của người dùng.

Còn dòng HTTP/1.1 404 Not Found chỉ ra rằng máy chủ đã không thể tìm thấy tài liệu như được yêu cầu.

Có năm loại mã trả lời tổng quát với ký số đầu tiên xác định loại mã.

Cũng giống như các thông điệp yêu cầu, các thông điệp trả lời có thể chứa một hoặc nhiều dòng trong phần MESSAGE\_HEADER. Những dòng này cung cấp thêm thông tin cho máy trạm. Ví dụ, dòng header Location chỉ ra rằng đường dẫn được yêu cầu đang có ở vị trí khác. Vì thế, nếu trang Web của Đại học Bách Khoa Hà Nội được di chuyển từ địa chỉ <http://www.hut.edu.vn/index.html> sang địa chỉ <http://www.hut.edu.vn/index.html/FAMI/index.html> mà người dùng lại truy cập vào đường dẫn cũ, thì Web máy chủ sẽ trả lời như sau: HTTP/1.1 301 Moved Permanently Location:

**<http://www.hut.edu.vn/index.html/FAMI/index.html>**

Trong tình huống chung nhất, thông điệp trả lời cũng sẽ mang theo nội dung trang Web được yêu cầu. Trang này là một tài liệu HTML, nhưng vì nó có thể chứa dữ liệu không phải dạng văn bản (ví dụ như ảnh GIF), dữ liệu này có thể được mã hóa theo dạng MIME. Một số hàng trong phần MESSAGE\_HEADER cung cấp thêm thông tin về nội dung của trang Web, bao gồm Content-Length (số bytes trong phần nội dung), Expires (thời điểm mà nội dung trang Web được xem như lỗi thời) và Last-Modified (thời điểm được sửa đổi lần cuối cùng).

### 5.3.3.3. Các kết nối TCP

Nguyên tắc chung của giao thức HTTP là máy trạm nối kết đến cổng TCP số 80 tại máy chủ, máy chủ luôn lắng nghe trên cổng này để sẵn sàng phục vụ máy trạm. Phiên bản đầu tiên (HTTP/1.0) sẽ thiết lập một nối kết riêng cho mỗi hạng mục dữ liệu cần tải về từ máy chủ. Không khó để thấy rằng đây là cơ chế không mấy hiệu quả: Các thông điệp dùng để thiết lập và giải phóng nối kết sẽ phải được trao đổi qua lại giữa máy trạm và máy chủ và khi mà tất cả máy trạm muốn lấy thông tin mới nhất của một trang Web, máy chủ sẽ bị quá tải.

Cải tiến quan trọng nhất trong phiên bản HTTP/1.1 là nó cho phép các kết nối lâu dài – máy trạm và máy chủ sẽ trao đổi nhiều thông điệp yêu cầu/trả lời trên cùng một kết nối TCP. Kết nối lâu dài có hai cái lợi: thứ nhất, nó làm giảm thiểu chi phí cho việc thiết lập/giải phóng nối kết. Thứ hai, do máy trạm gửi nhiều thông điệp yêu cầu qua một kết nối TCP, cơ chế điều khiển tắc nghẽn của TCP sẽ hoạt động hiệu quả hơn.

Tuy nhiên, kết nối lâu dài cũng có cái giá phải trả. Vấn đề phát sinh ở chỗ: không ai trong máy trạm và máy chủ biết được kết nối đó sẽ kéo dài bao lâu. Điều này thực sự gây khó khăn cho phía máy chủ bởi vì tại mỗi thời điểm, nó phải đảm bảo duy trì kết nối đến cả ngàn máy trạm. Giải pháp cho vấn đề này là: máy chủ sẽ mãn kỳ và cắt nối kết nếu nó không nhận được một yêu cầu cụ thể nào từ phía máy trạm trong một khoảng thời gian định trước. Ngoài ra, cả máy trạm và máy chủ phải theo dõi xem phía bên kia có chủ động cắt nối kết hay không và lấy đó làm cơ sở để tự cắt nối kết của mình (nhắc lại rằng, cả hai bên phải cắt nối kết thì nối kết TCP mới thực sự kết thúc).



#### 5.3.3.4. Trữ đệm

Một trong những lĩnh vực nghiên cứu tích cực nhất hiện nay về Internet là làm sao để trữ tạm các trang Web một cách hiệu quả. Việc trữ tạm mang lại nhiều lợi ích. Từ phía máy trạm, việc nạp và hiển thị một trang Web từ bộ đệm gần đây là nhanh hơn rất nhiều so với từ một máy chủ nào đó ở nửa vòng trái đất. Đối với máy chủ, có thêm một bộ đệm để can thiệp vào và phục vụ giúp yêu cầu của người dùng sẽ giảm bớt tải trên máy chủ.

Việc trữ đệm có thể được cài đặt tại nhiều nơi khác nhau. Ví dụ, trình duyệt Web có thể trữ tạm những trang Web mới được nạp về gần đây, để khi người dùng duyệt lại những trang Web đó, trình duyệt sẽ không phải nối kết ra Internet để lấy chúng về mà dùng bản trữ sẵn. Ví dụ khác, một khu vực làm việc (khu vực làm việc) có thể để cử một máy làm nhiệm vụ trữ tạm các trang Web, để những người dùng sau có thể sử dụng các bản trữ sẵn của những người dùng trước. Yêu cầu của hệ thống này là mọi người dùng trong khu vực làm việc phải biết địa chỉ của máy tính làm nhiệm vụ bộ trữ tạm, và họ chỉ đơn giản là liên hệ với máy tính này để tải các trang Web về theo yêu cầu. Người ta thường gọi máy tính làm nhiệm vụ trữ tạm các trang Web cho một khu vực làm việc là proxy. Vị trí trữ đệm có thể di chuyển gần hơn đến phần lõi của Internet là các ISP. Trong tình huống này, các khu vực làm việc nối kết tới ISP thường không hay biết gì về việc trữ tạm ở đây. Khi các yêu cầu HTTP từ các khu vực làm việc được chuyển phát đến router của ISP, router liền kiểm tra xem đường dẫn được yêu cầu có giống với các đường dẫn được trữ sẵn hay không. Nếu có, router sẽ trả lời ngay. Nếu không, router sẽ chuyển yêu cầu đến máy chủ thật sự và cũng không quên lưu vào bộ đệm của mình thông điệp trả lời từ phía máy chủ đó. Việc trữ tạm là đơn giản. Tuy nhiên bộ đệm phải đảm bảo những thông tin trữ đệm trong đó không quá cũ. Để làm được việc này, các Web máy chủ phải gán “ngày hết hạn” (tức là trường Expires trong header) cho mọi trang Web mà nó phục vụ cho máy trạm. Nhân đó, các bộ đệm cũng lưu lại thông tin này. Và từ đó, các bộ đệm sẽ không cần phải kiểm tra tính cập nhật của trang Web đó cho đến khi ngày hết hạn đến. Tại thời điểm một trang Web hết hạn, bộ đệm sẽ dùng lệnh HEAD hoặc lệnh GET có điều kiện (GET với trường If-Modified-Since trong phần header được đặt) để kiểm tra rằng nó có một phiên bản mới nhất của trang Web kia. Tổng quát hơn, cần phải có “các chỉ thị hướng dẫn” cho việc trữ đệm và các chỉ thị này phải được tuân thủ tại mọi bộ đệm. Các chỉ thị sẽ chỉ ra có nên trữ đệm một tài liệu hay không, trữ nó bao lâu, một tài liệu phải tươi như thế nào ...

#### 5.3.4. Truyền tập tin (FTP)

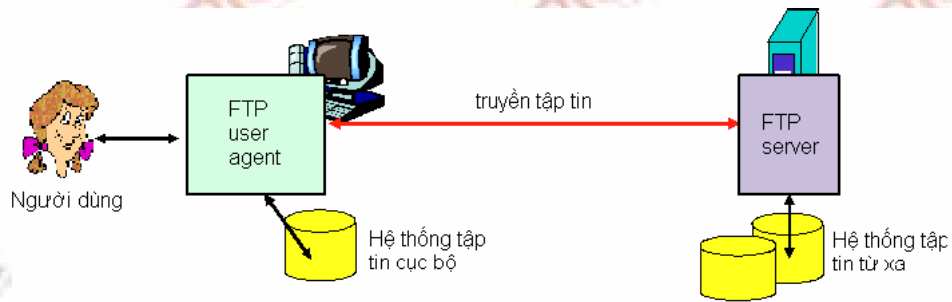
Thông qua dịch vụ FTP, người dùng tại một máy tính có thể đăng nhập và thao tác lên hệ thống tập tin được chia sẻ của một máy tính từ xa.

Mục tiêu của dịch vụ FTP là:

- 1) Đảm bảo việc chia sẻ tập tin (chương trình máy tính hoặc dữ liệu) trên mạng.
- 2) Khuyến khích việc sử dụng không trực tiếp (thông qua chương trình) tài nguyên trên các máy tính khác.
- 3) Người dùng không cần phải quan tâm đến sự khác nhau của các hệ thống tập tin trên mạng.
- 4) Truyền dữ liệu một cách tin cậy và hiệu quả.

#### 5.3.4.1. Mô hình dịch vụ FTP

Hình sau mô tả mô hình của dịch vụ FTP.

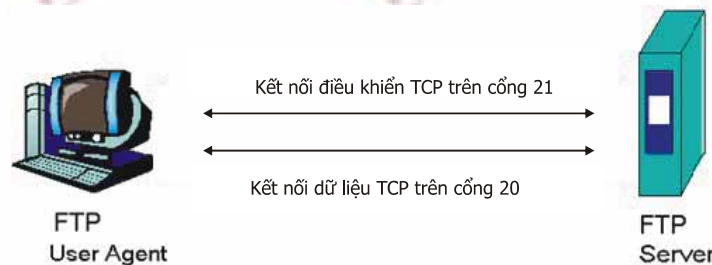


**Hình 5.18: Mô hình dịch vụ FTP**

Trong hệ thống này, người dùng sẽ ra lệnh cho FTP user agent. User agent sẽ nối kết tới FTP máy chủ để dàn xếp thủ tục làm việc, thực thi các tác vụ theo yêu cầu và trả kết quả về cho người dùng.

#### 5.3.4.2. Giao thức FTP

Đầu tiên, user agent thiết lập một kết nối điều khiển trên cổng 21 tới FTP máy chủ. Sau khi đã thỏa thuận các tham số truyền nhận, hai bên sẽ thiết lập một kênh dữ liệu chạy trên cổng 20. Dữ liệu của các tập tin được trao đổi qua lại giữa user agent và máy chủ sẽ chạy trên kênh dữ liệu này. Kênh dữ liệu là kênh hoạt động theo phương thức hai chiều và không nhất thiết phải luôn tồn tại.



**Hình 5.19: Giao tiếp giữa máy trạm và máy chủ trong giao thức FTP**

#### 5.3.4.3. Các lệnh cơ bản

Sau đây là các lệnh cơ bản mà người dùng có thể sử dụng để thao tác lên hệ thống FTP.

Lệnh	Tham số	Ý nghĩa
FPT	Host – name	Nối kết đến FTP server có địa chỉ Host – name.
USER	User – name	Cung cấp tên người dùng cho FTP server để thực hiện quá trình chứng thực.
ASCII		Chỉ định kiểu dữ liệu truyền nhận là ký tự.
BINARY		Chỉ định kiểu dữ liệu truyền nhận là nhị phân.
LS		Xem nội dung thư mục từ xa.
CD	Rremote – dir	Chuyển đến thư mục khác trong hệ thống tập tin từ xa.
GET	Remote – file, Local – file	Tải tập tin Remote – file trên FTP server về hệ thống tập tin cục bộ và đặt tên là Local – file.
PUT	Local – file, Remote – file	Nạp tập tin cục bộ Local – File lên server và đặt tên là Remote – file.
MKDIR	Dir – name	Tạo một thư mục có tên Dir – name trên hệ thống tập tin từ xa.
RMDIR	Dir – name	Xóa thư mục có tên Dir – name trên hệ thống tập tin từ xa.
QUIT		Đóng nối kết FTP và thoát khỏi chương trình FTP

#### 5.4. Tài nguyên trên Internet

Nói đến Internet là nói đến một kho tài nguyên vô cùng rộng lớn. Tuy nhiên trong khuôn khổ giáo trình này chúng ta chỉ liệt kê những tài nguyên cơ bản:

##### Dịch vụ Finger (Finger Service)

- Hầu hết các máy tính trên Internet đều cung cấp các công cụ cho phép bạn hỏi thông tin về cá nhân một người dùng nào đó. Dịch vụ này được biết tới với tên mô tả là Finger. Như bạn sẽ thấy, mọi người trên Internet thường được biết tới bằng userid của họ. Ví dụ, bạn có thể tìm ra rằng userid hardley được đăng ký cho Hardley Hahn.
- Chúng tôi dùng từ "Finger" không chỉ như một danh từ, mà còn như một động từ. Ví dụ, bạn có thể nghe ai đó hỏi: "Ai là userid TLN ?", một người thứ hai trả lời : "Tôi không biết. Sao chúng ta không finger anh ta và tìm ra?".
- Phụ thuộc vào dịch vụ finger được cài đặt như thế nào trên máy tính bạn liên hệ, bạn có thể tìm ra những thông tin khác về một cá nhân : số điện thoại, địa chỉ văn phòng...Thêm vào đó, một vài hệ thống finger cho bạn biết một người vào mạng lần cuối khi nào, và anh ta có thư chưa đọc không. Điều này thật thuận tiện, nếu bạn cần kiểm tra xem ai đó đã nhận một thông điệp quan trọng hay chưa.
- Cũng có một cách để bạn quy định theo ý mình một phần những gì người ta sẽ nhìn thấy khi họ finger tới userid của bạn. Bạn có thể chỉ ra những thông tin nhất định nào bạn muốn hiện lên. Ví dụ, một ông giáo sư có thể chỉ ra những giờ làm việc của ông ta. Một ai đó tổ chức một buổi dạ hội có thể đưa ra các hướng tới nhà cô ta. Bạn có thể làm hiện ra các thông tin này bất cứ lúc nào bạn muốn bằng cách finger vào userid của người đó.
- Bạn cũng có thể finger một máy tính, thay vì một userid. Trong trường hợp này, máy tính sẽ đáp lại bằng cách cho bạn thấy tóm tắt của tất cả các userid hiện thời đang thâm nhập vào (login).
- Cuối cùng, một vài hệ thống finger phục vụ cho những câu hỏi chung về những thông tin nhất định. Ví dụ, có một userid riêng và một máy tính tại Đại học Washington Seattle mà bạn có thể finger để hiện ra thông tin về những vụ động đất gần đây.

##### Usenet

- Usenet - Tên này viết tắt của "User's Network" (Mạng của người sử dụng) - là một trong những lý do chính người ta sử dụng Internet. Usenet bản thân nó không phải là một mạng thực sự. Nó là một hệ thống của những nhóm thảo luận trong đó những bài riêng được phân phối trên toàn thế giới.
- Usenet có hàng ngàn nhóm thảo luận theo nghĩa đen.
- Tại mỗi vị trí trên Internet (Internet site), người quản trị hệ thống quyết định cho truyền tải hoặc không truyền tải các nhóm thảo luận Usenet. Vì vậy, Usenet không phải là dùng được ở khắp mọi nơi. Hơn nữa, thậm chí những vị trí cung cấp phương pháp này sẽ không truyền tải mọi nhóm thảo luận có thể thấy được.

##### Anonymous FTP

- Như chúng tôi đã đề cập trong phần trước, dịch vụ FTP cho phép bạn sao chép file từ máy tính này sang máy tính khác. Anonymous FTP là một hệ thống, trong đó

một tổ chức làm cho những tệp nào đó trở thành sử dụng công cộng được. Bạn có thể truy cập một máy tính như vậy bằng cách sử dụng userid là anonymous. Không mật mã đặc biệt nào cả.

- Anonymous FTP là một trong những dịch vụ Internet quan trọng nhất. Thực sự, mọi kiểu dữ liệu có thể được lưu trữ ở một vài nơi trên một vài máy tính, bạn đều có thể sử dụng miễn phí được. Ví dụ, rất nhiều chương trình được sử dụng trên Internet được tạo ra và bảo trì do những người khác nhau, phân phối những chương trình trên khắp thế giới thông qua Anonymous FTP. Bạn cũng có thể tìm thấy tạp chí điện tử, văn bản của các nhóm thảo luận Usenet, tài liệu kỹ thuật và nhiều hơn nữa. Như một người sử dụng Internet còn non nớt, bạn sẽ phụ thuộc rất nhiều vào Anonymous FTP.

### **Các Archie Máy chủ (Archie Máy chủ)**

- Có hàng ngàn các Anonymous FTP máy chủ trên toàn thế giới, cung cấp một số lượng khổng lồ các tệp. Vai trò của các Archie máy chủ là làm cho toàn bộ hệ thống trở nên quản lý được, bằng cách giúp đỡ bạn tìm kiếm những gì bạn muốn.
- Giả sử rằng bạn muốn một tệp xác định nào đó, chẳng hạn như một chương trình mà bạn đã nghe nói tới. Bạn có thể dùng một Archie máy chủ để nói cho bạn biết vị trí Anonymous FTP nào lưu trữ tệp đó. Khi bạn biết tên của vị trí này, thật là đơn giản để sử dụng FTP và nạp tệp đó xuống (download).
- Nếu bạn coi thế giới của Anonymous FTP như một thư viện toàn cầu khổng lồ, được thay đổi liên tục, thì bạn có thể nghĩ về các Archie máy chủ như là một danh mục. Thực tế, không có các Archie máy chủ, đa số các tài nguyên Anonymous FTP là không thể với tới.
- Cái tên "Archie" được đặt ra từ ý tưởng về một "archie máy chủ". Trong sự kính trọng với một khuynh hướng nổi tiếng của máy tính, người ta vương bạn trong những dấu hiệu cá nhân kỳ quặc, chúng tôi thường đề cập tới "Archie" như là một người. Ví dụ, bạn có thể hỏi ai đó : "Bạn có biết vị trí Anonymous FTP nào truyền tải tạp chí điện tử Unplastic News không ?". "Không, tôi không biết", bạn của bạn đáp, "Sao anh không hỏi Archie?"

### **Công cụ Talk (Talk Facility)**

- Công cụ Talk thiết lập một kết nối giữa máy tính của bạn với một máy khác. Bạn có thể dùng mỗi kết nối này để gõ thông điệp đến và đi (tới khi bạn chán thì thôi).
- Điều vĩ đại của Internet Talk là nó giữ cuộc đối thoại với những người nào đó, bất kể họ ở xa tới đâu. Người khác thấy những gì bạn gõ khi bạn gõ nó và các bạn có thể cùng gõ đồng thời, mà không trộn lẫn các thông điệp của các bạn với nhau.
- Internet Relay Chat (Cuộc trò chuyện tiếp sức Internet).
- Internet Relay Chat (IRC) giống như một công cụ Talk dành cho nhiều người đồng thời. Như bạn có thể hình dung, IRC được sử dụng thường xuyên và cung cấp nhiều thứ hơn là một cuộc đàm thoại đơn thuần.
- Bạn có thể chiếm một phần trong những cuộc đàm thoại công cộng với một số lớn người. Những cuộc đàm thoại này được tổ chức rời rạc với rất nhiều chủ đề hay ý tưởng. Một cách luân phiên, bạn có thể sử dụng IRC để sắp xếp các cuộc đàm thoại riêng với những người bạn chọn, giống như một hội nghị qua điện thoại.



## Gopher

- Gopher cung cấp một loạt những bảng chọn, từ đó bạn có thể truy cập thực sự vào mọi loại thông tin văn bản, kể cả nó được cung cấp bởi tài nguyên Internet khác. Có rất nhiều hệ thống Gopher trên Internet, mỗi cái được quản trị một cách cục bộ. Mỗi Gopher chứa bất kỳ thông tin nào mà những người dùng Gopher cục bộ quyết định chia sẻ.
- Chỉ có một vài Gopher là những hệ thống đơn độc, đa số Gopher được cài đặt để kết nối tới những Gopher khác. Ví dụ, bạn đang sử dụng một Gopher ở California. Với một sự chọn lựa đơn giản trên bảng chọn, bạn có thể kết nối với một Gopher khác ở Châu Phi hay ở Nam Mỹ. Điều làm cho những Gopher mạnh mẽ như vậy là bất kỳ bạn đang sử dụng Gopher nào và loại thông tin gì, giao diện luôn là một bảng chọn đơn giản giống nhau.

## Veronica và Jughead

- Không ai thực sự biết có bao nhiêu Gopher trên thế giới. Nói đầy đủ là có rất nhiều, tất cả đều có một loạt đề mục bảng chọn cung cấp thông tin và dịch vụ.
- Veronica là một công cụ theo dõi rất nhiều bảng chọn Gopher trên khắp thế giới. Bạn có thể sử dụng Veronica thực hiện việc tìm kiếm và quan sát tất cả các đề mục bảng chọn có chứa những từ khóa nhất định (những gì bạn chỉ ra). Jughead làm cùng một việc đó với một nhóm đặc biệt của bảng chọn Gopher.
- Kết quả tìm kiếm của Veronica hay Gopher là một bảng chọn custom, chứa những đề mục được tìm thấy. Việc lựa chọn bất kỳ đề mục nào từ bảng chọn này sẽ tự động kết nối bạn với Gopher thích hợp, nơi mà có thể nó ở đó. Thực ra trừ khi bạn hỏi một cách đặc biệt, bạn thậm chí không biết bạn đang sử dụng máy tính nào và nó ở nước nào. Thật là ngạc nhiên.

## Các Wais Servers (Wais Servers)

- Những Wais server cung cấp một phương pháp khác để tìm kiếm thông tin được trải ra trên khắp Internet. Wais có khả năng truy cập một số lớn các cơ sở dữ liệu (data base). Để khởi đầu, bạn nói cho Wais biết bạn muốn tìm cơ sở dữ liệu nào. Tiếp theo, bạn chỉ ra một hay nhiều từ khóa để tìm. Wais sẽ tìm tất cả các từ trong mọi mục trong tất cả các cơ sở dữ liệu mà bạn chỉ ra.
- Kết quả tìm kiếm của Wais là một danh sách các mục, được lựa chọn từ nhiều cơ sở dữ liệu, có vẻ giống cái mà bạn thích. Wais trình bày chúng như một bảng chọn, với những đề mục thích hợp trước. Từ danh sách này, bạn có thể yêu cầu Wais hiển thị bất kỳ mục nào hấp dẫn bạn.
- Tên "Wais" là do "Wide Area Information Service" (Dịch vụ thông tin diện rộng), và được phát âm là "Wayz".

## World Wide Web

- Nguồn gốc của World Wide Web:

World Wide Web ra đời từ một dự án nghiên cứu phát triển tại Trung tâm nghiên cứu hạt nhân châu Âu vào năm 1989. Nhóm nghiên cứu đã nghiên cứu và phát triển một giao thức truyền và nhận các tập tin siêu văn bản theo mô hình khách chủ được gọi tắt là HTTP (Hyper Text Transfer Protocol). Sau đó nhóm đã phát triển công bố rộng rãi thư viện chương trình nguồn của giao thức này cho các nhà phát

triển và các tổ chức phát triển trên Internet để họ phát triển các trình duyệt Web (Web Browser). Các tổ chức và tập đoàn đã thành lập W3 Consortium để tiếp tục chuẩn hoá và phát triển bộ giao thức. W3 Consortium đã phát triển thêm các tính năng mới của HTML và các mức cũng như các chuẩn để thực hiện các phần mềm đi kèm. Từ đó thuật ngữ World Wide Web ra đời.

- Cơ chế hoạt động
  - World Wide Web về cơ bản rất giống với những hệ thống văn bản trợ giúp cho những chương trình ứng dụng như các help online trong hệ điều hành Windows. Ý tưởng cơ bản của hệ thống này là thay thế việc tổ chức văn bản theo kiểu cấu trúc cứng nhắc, lần lượt văn bản sẽ được tổ chức theo kiểu liên kết phức tạp cho phép nhảy từ đề mục này sang đề mục khác bằng cách thao tác trên văn bản và tùy theo sự lựa chọn của người đọc.
  - Các liên kết cho phép người đọc điều khiển việc chuyển đổi từ văn bản này sang văn bản khác trong siêu văn bản. Như vậy mỗi một liên kết chính là một móc xích cho phép liên kết hai văn bản lại với nhau và các liên kết này rất phức tạp vì trong mỗi một văn bản có rất nhiều các điểm liên kết.
  - Cũng giống như các dịch vụ khác trên Internet, World Wide Web có thể hoạt động trên mọi hệ điều hành và giữa các hệ điều hành khác nhau. Sở dĩ như vậy là do World Wide Web đã sử dụng một ngôn ngữ giao tiếp chung HTTP. Ngoài ra World Wide Web cũng có thể hiểu được một số giao thức khác của Internet như FTP, TELNET...
  - WWW hoạt động theo mô hình khách chủ. Dữ liệu Web được lưu trữ trên các Web Máy chủ. Web Browser (trình duyệt Web) là chương trình để xem các tài liệu Web, được đặt trên máy trạm.
  - Có hai chương trình Web Browser hay được sử dụng nhất là Internet Explorer của Microsoft và Netscape Navigator của Netscape. Người dùng sử dụng các công cụ này để yêu cầu hiển thị các trang Web. Những nhà cung cấp dịch vụ Web có một Web Máy chủ (chương trình quản lý hệ thống dữ liệu Web), nhận các yêu cầu từ khách hàng (Web Browser) thông qua hệ thống mạng, sau đó gửi các văn bản tương ứng đến đúng nơi yêu cầu. Khi giao tiếp với Web Máy chủ, Web Browser sử dụng giao thức HTTP. Ngoài ra Web cũng sử dụng một số giao thức khác của Internet như: FTP, TELNET, GOPHER.
- Địa chỉ của một văn bản được gọi là đường dẫn (Universal Resource Locator), có cú pháp như sau: Protocol://Internet Address/path  
Trong đó:
  - Protocol: giao thức truy nhập và lấy thông tin, có thể là http, ftp, gopher.
  - Internet Address: địa chỉ Internet của trạm (site) chứa văn bản.
  - Path: đường dẫn tới nơi chứa thông tin.
  - Ví dụ: <http://www.microsoft.com/sitebuilder/intro.htm>
- Mỗi trang Web có một đường dẫn duy nhất. Hyperlink là các siêu liên kết giúp di chuyển giữa các trang Web, mỗi hyperlink trỏ tới đường dẫn của một trang Web. Trang Web được lưu trữ dưới dạng tệp có phần mở rộng là .htm, .html... HTML (HyperText Markup Language) là ngôn ngữ để tạo nên các trang Web.

## Các khái niệm cơ bản về trang Web

Có nhiều phần mềm hỗ trợ thiết kế trang Web theo một mẫu có sẵn như FontPage Express, PageMill, Web Hotdog, Netscape Composer, ...Bạn có thể sử dụng các phần mềm này để tạo ra trang Web của bạn một cách nhanh chóng. Bạn chỉ cần lựa chọn cái bạn muốn để đưa vào trang Web của bạn mà không cần biết ngữ pháp là gì. Nhờ các phần mềm này mà trang Web của bạn có thể rất đẹp, rất sinh động ...Nhưng trang Web đó vẫn chưa phải thực sự là tác phẩm của bạn ví như việc bạn xếp đồ hình này đồ hình kia để đưa vào bộ sưu tập sao cho đẹp mắt chứ đâu phải bạn là người đã tạo ra các đồ hình ấy. Bạn sẽ có cảm giác tê nhạt khi làm ra một trang Web mà không cần đến một trí lực gì đáng kể. Nếu như bạn chưa từng tự nấu lấy món ăn cho mình thì mời bạn hãy tiếp tục sử dụng đồ ăn sẵn. Còn nếu như bạn muốn tự mình làm nên tất cả thì xin mời bạn hãy tìm hiểu HTML. Khi bạn đã có HTML bạn có thể tự thiết kế trang Web mà không cần một hỗ trợ nào hết, miễn sao bạn phải có công cụ để soạn thảo trang Web và thử nghiệm trang Web, đó là một hệ soạn thảo văn bản dạng văn bản bất kỳ (NotePad chẳng hạn) và trình duyệt Web (Web Browser) chưa cần nối mạng (Internet Explorer).

## Vậy HTML là gì?

- HTML (HyperText Markup Language) không phải là ngôn ngữ lập trình như Pascal, C, Java... Nó không thể tạo ra các chương trình ứng dụng dùng trực tiếp ngôn ngữ máy. HTML là ngôn ngữ đánh dấu siêu văn bản để tạo ra các liên kết giữa các trang văn bản đa dạng với nhau và liên kết với các Multimedia như phim, ảnh số, âm thanh... Thuộc tính ngôn ngữ của HTML là ở chỗ nhờ có các kí hiệu quy ước đánh dấu trên văn bản mà trình duyệt Web biết phải làm gì với các kí hiệu đó.
- Giao thức (Protocol): Giao thức là một phương thức truy cập các trang Web của trình duyệt. Mỗi giao thức có một tên gọi xác định và tên đó có dạng TênGiaoThức://
- http:// là một giao thức hay dùng nhất để mở các tài liệu HTML trong trình duyệt. Đó là từ viết tắt của HyperLink Text Transfer Protocol nghĩa là giao thức truyền các văn bản siêu liên kết trên mạng.
- Web Site: Trang WEB là một tài liệu HTML dạng văn bản trong đó chứa các lệnh của ngôn ngữ HTML. Web Site là một số trang Web liên kết với nhau. Mỗi trang Web được lưu cất thành một tệp tin trên đĩa với tên đuôi phải là .HTM.
- Thẻ (Tag) định dạng: Các lệnh của HTML đơn giản chỉ là các mã đánh dấu định dạng. Mỗi mã đánh dấu định dạng gọi là một Tag nằm giữa hai dấu ngoặc nhọn trong đó có tên tag và các thuộc tính kèm theo của tag nếu có. Mỗi Tag chỉ dẫn cho trình duyệt biết định dạng của các đối tượng nằm sau mã theo một quy cách đã định. Tag không có dấu gạch chéo gọi là tag mở, Tag có dấu gạch chéo gọi là tag đóng. Tag bằng chữ in và chữ thường có tác dụng như nhau. Các tag nhập sai thì trình duyệt bỏ qua mà không báo lỗi vì trình duyệt sẽ coi đó là một phần của văn bản mà không coi đó là mã định dạng. Các tag có thể nhập liền nhau, cách nhau hoặc xuống dòng.

**Ví dụ 1:** <BODY background="Hoa.GIF">

trong ví dụ này BODY là tên của Tag, thuộc tính tag này là Background="Hoa.GIF".



### Ví dụ 2:

<b>Đoạn văn bản này được định dạng chữ béo </b>

<u>Đoạn văn bản này được định dạng gạch chân </u>

<i>Đoạn văn bản này được định dạng chữ nghiêng</i>

Trong ví dụ này thì tag <b> để mở đầu định dạng chữ béo, tag </b> để kết thúc định dạng chữ béo. Tag <u> mở đầu định dạng gạch chân, tag </u> để kết thúc gạch chân. Tag <i> mở đầu định dạng in nghiêng, tag </i> kết thúc định dạng in nghiêng.

- URL (Uniform Resource Locator): URL là địa chỉ của nguồn tài nguyên thống nhất của WEB. Bạn có thể tạo trang Web và khai thác tài nguyên ngay trên máy của bạn. Khi đó URL đơn giản chỉ là tên ổ đĩa và đường dẫn tới thư mục mà bạn cần khai thác. Chẳng hạn URL của tệp index.htm trong thư mục HTML của ổ đĩa C sẽ là "C:\HTML\INDEX.HTM"

### Tổ chức của một trang Web HTML

- Một trang Web HTML gồm 2 phần nằm giữa Tag <HTML> và </HTML>:
- Phần mở đầu nằm giữa tag mở <head> và tag đóng </head>, đó là phần khai báo các định dạng chung cho toàn bộ trang Web như khai báo dòng tiêu đề của thanh tiêu đề ở cửa sổ trình duyệt, định dạng phông chữ, nhạc nền...Phần này có thể không có.
- Phần sau là phần thân của trang nằm giữa tag mở <BODY> và tag đóng </BODY>. Phần này chứa các tag định dạng và các văn bản kèm theo của trang. Đây là phần chứa các nội dung mà trình duyệt cần đưa lên màn hình.

### Soạn thảo trang Web

- Bạn có thể dùng một trình soạn thảo NotePad (hoặc WordPad) để nhập tài liệu HTML theo cấu trúc trên sau đó bạn lưu vào tệp với tên đuôi là .HTM với kiểu tệp là Text document. Cần nhớ rằng tài liệu HTML chỉ được phép định dạng Text Document mà thôi, do đó nếu bạn soạn thảo trên các hệ soạn thảo khác (như MicroSoft Word chẳng hạn) thì bạn phải chọn kiểu định dạng Text Document để cất vào tệp. Trang chủ của bạn nên cất vào tệp có tên là INDEX.HTM vì đó là tên mặc định của trang được trình duyệt nạp vào đầu tiên khi kho tài nguyên của bạn được động tới.
- Khi nhập văn bản html bạn không nên dùng bất kỳ một định dạng nào cho văn bản này mà hãy soạn thảo bình thường như soạn thảo trong DOS. Trình duyệt sẽ bỏ qua mọi khoảng trắng và mọi kiểu trình bày ngay trong tài liệu html.

### Thử nghiệm trang Web

Đảm bảo chắc chắn là trang Web của bạn đã được lưu (save) vào thư mục của bạn ở trên đĩa của máy cá nhân với tệp kiểu .HTM. Để thử nghiệm, bạn hãy mở Windows explore rồi vào thư mục chứa trang Web của bạn, nháy chuột vào tên tệp để chạy thử. Bạn cũng có thể lôi tệp HTM của bạn ra Desktop thành một biểu tượng rồi kích hoạt biểu tượng đó để thử nghiệm trang Web của bạn. Nếu bạn thấy trang Web của bạn chưa thật ổn thì hãy trở về hệ soạn thảo để mở tệp và sửa đổi.

Bạn hãy nhập tài liệu HTML sau và chạy thử xem điều gì sẽ xảy ra:



```
<HTML>
<head>
<title> Trang Web của tôi </title>
</head>
<body>
<h1><marquee>Hello! Tôi đang làm trang Web</marquee></h1>
</body>
</HTML>
```

### Đưa trang Web lên mạng

Nếu bạn muốn tung trang Web của bạn lên mạng toàn cầu để giao lưu với thiên hạ thì bạn phải mất tiền cho nhà cung cấp dịch vụ internet (ISP=Internet Máy chủ Provider) để mua quyền truy cập và mua không gian đĩa trên máy chủ để đặt tài nguyên của bạn vào đó. Khi đã làm việc với ISP thì bạn sẽ được cấp một không gian đĩa và một địa chỉ URL để truy cập nó. Địa chỉ này thường bắt đầu bằng tên một giao thức và tên máy chủ của ISP tiếp theo là đường dẫn đến trang HTML của bạn trên máy chủ. Máy chủ của ISP (còn được gọi là Web Máy chủ) sẽ cung cấp cho bạn dịch vụ quay số và các dịch vụ khác. Bạn cũng sẽ được cung cấp một tài khoản riêng về mật khẩu, cách thức truy cập và cách thức nạp bổ sung các trang Web vào kho tài nguyên của mình trên Web Máy chủ bằng chương trình FTP hoặc Shell.

### Các thư mục White Pages

- Trong phạm vi tràn ngập khắp thế giới của Internet, không có gì quan trọng hơn các địa chỉ điện tử cá nhân. Chỉ khi bạn biết địa chỉ của ai đó, bạn mới có thể gửi thư, đàm thoại bằng Talk, hoặc thậm chí tìm hiểu thêm về người đó bằng cách dùng Finger.
- Bạn sẽ làm gì, khi mà bạn muốn liên hệ với một người, mà bạn không biết địa chỉ của họ? Bạn sử dụng một trong số các thư mục White Pages. Dĩ nhiên tên này nhắc chúng ta nhớ về danh bạ điện thoại.
- Tuy nhiên, bản sao điện tử này lại khác xa, chủ yếu vì không có một thư mục Internet đơn lẻ. Không có một tổ chức hay cá nhân nào chịu trách nhiệm, nên cũng không ngạc nhiên khi không có một nguồn trung tâm các tên và địa chỉ.
- Thay vào đó, có một số thư mục White Pages khác nhau - những máy chủ có mục đích đặc biệt - mà bạn có thể tìm một tên. Khi bạn đánh mất tờ ghi địa chỉ của ai đó trong kỳ nghỉ mới đây của bạn. Có cơ hội là White Pages có thể giúp bạn theo dõi địa chỉ điện tử của những người quen của bạn, trong khi anh ta hay cô ta là một người sử dụng Internet (tại sao bạn lại sẽ giao thiệp với một người không phải là người dùng Internet?).

### Các tạp chí điện tử (Electronic Magazines)

- Internet được làm chủ bởi rất nhiều tạp chí xuất bản điện tử. Đó là những bài được lưu trữ như các tệp văn bản, có thể truy cập từ một bài đến tất cả. Một vài tạp chí điện tử là những tờ báo uyên bác, chủ yếu hấp dẫn những nhà chuyên môn. Những tạp chí khác làm tất cả mọi người thích thú.
- Có hai cách phân phối các tạp chí điện tử. Một vài tạp chí chứa một danh sách địa chỉ thư điện tử (Mailing List - danh sách địa chỉ thư điện tử của người dùng mạng). Khi một sự phát sinh mới tới, nó được gửi tới bạn như một thông điệp mail. Các tạp chí khác được lưu trữ tại các vị trí anonymous FTP nổi tiếng. Bạn có thể chép xuống, kể cả những gì phát sinh trước đó, bất kỳ lúc nào bạn muốn.

### **Các danh sách địa chỉ thư điện tử (Mailing lists)**

- Một Mailing list là một hệ thống có tổ chức, trong đó một nhóm những người được gửi các thông điệp thuộc về những chủ đề riêng biệt. Những thông điệp này có thể là các bài báo, lời giải thích, hoặc bất kỳ cái gì thích hợp với chủ đề này.
- Tất cả các Mailing list - có hàng ngàn cái như vậy - có một người nào đó chịu trách nhiệm. Bạn có thể đặt trước hoặc thôi đặt trong danh sách này bằng cách gửi một thông điệp tới địa chỉ thích hợp. Rất nhiều mailing list được "điều tiết", nghĩa là có ai đó quyết định những tài liệu nào được nhận. Những danh sách khác sẽ chấp nhận và gửi đi những thông điệp từ bất kỳ ai.
- Nếu bạn cảm thấy cô đơn và bị bỏ rơi, việc đặt mua một năm các mailing list sẽ bảo đảm giữ cho hộp thư điện tử của bạn luôn luôn đầy tới mức giới hạn.

### **Các BBSs của Internet**

- Một BBS, hay Bulletin Board System, là một kho chứa các thông điệp và các tệp, thường công bố cho các chủ đề riêng biệt. Để sử dụng BBS, bạn kết nối tới nó, và lựa chọn những đề mục từ một loạt các bảng chọn.
- Một cách tiêu biểu, một BBS sẽ được bảo trì bởi một cá nhân hoặc một tổ chức. Có vô số các hệ thống BBS trên thế giới, đa số chúng được kết nối bằng dây điện thoại. Internet có nhiều BBS bạn có thể kết nối bằng cách dùng Telnet.

### **Các trò chơi**

- Nơi nào có máy tính mà lại không có trò chơi? Tất nhiên, có rất nhiều trò chơi trên máy tính bạn có thể chép xuống thông qua Anonymous FTP và chạy nó trên máy tính của bạn. Tuy nhiên, cũng có nhiều trò chơi đặc biệt trên mạng, sử dụng những lợi điểm của các công cụ Internet.
- Ví dụ bạn có thể chơi các trò Chess (Cờ Vua) và Go (Đi) với những người khác trên Internet. Hoặc bạn có thể chơi trò Thuật ngoại giao (Diplomacy) bằng thư điện tử. Có thể bạn sẽ thích một cuộc không chiến giả lập giữa các tàu vũ trụ, dựa trên cuộc du hành Ngôi sao. Có lẽ bạn cũng sẽ thích sử dụng kỹ năng của bạn như một lập trình viên để chơi trò Core War, một trò chơi trong đó những người chơi sử dụng những chương trình viết bằng hợp ngữ (Assembly Language Program) (cho một máy tính ảo) để thống trị và điều khiển bộ nhớ của máy tính này.
- Bất kỳ khuynh hướng nào của bạn, bạn đều có thể tìm thấy trên Internet. Nhưng đừng nói với ông chủ của bạn và những nhà cầm quyền điều đó. Vì bạn đã tiêu tốn thời gian và tiền bạc vào các trò chơi không mang lại lợi ích cho họ.

### **MUD**

- MUD, hay Multiple User Dimension, là một chương trình máy tính cung cấp thực tế ảo (Virtual Reality). Để tham dự vào MUD, bạn dùng Telnet để kết nối tới một MUD máy chủ, giữ một vai trò và khảo sát. Khi bạn làm thế, bạn tương tác với những người khác đang diễn vai trò của họ. Nói cách khác, MUD cho phép bạn luyện tập trí tưởng tượng và khả năng giả vờ của bạn, giống như chơi trò Dungeons and Dragons (Ngục tối và những con rồng) hoặc thăm một quán rượu của người độc thân.
- Điều bảo đảm còn lại là MUD phức tạp, mê hoặc và gây nghiện ngay lập tức. Khi bạn đi quanh các thực tế ảo của một MUD riêng biệt, bạn có thể nói chuyện với mọi người, giải trò xếp hình, khám phá những nơi xa lạ (như các hang động), và thậm chí tạo ra một vài thực tế của chính bản thân mình.

## TÓM LƯỢC CUỐI BÀI

- Các khái niệm về họ giao thức TCP/IP, các tầng giao thức TCP/IP:
  - Tầng giao tiếp mạng (Network Interface Layer).
  - Tầng Internet (Internet Layer).
  - Tầng giao vận (Transport Layer).
  - Tầng ứng dụng (Application Layer).
- Mạng Internet.
- Các ứng dụng trên TCP/IP như:
  - Dịch vụ tên miền (DNS).
  - Electronic Mail (SMTP, MIME, POP3, IMAP).
  - World Wide Web (WWW).
  - Truyền tập tin (FTP).
  - Tài nguyên trên Internet.

## **CÂU HỎI TỰ LUẬN**

- Câu 1.** Tên và chức năng chính của các tầng trong mô hình TCP/IP ?
- Câu 2.** So sánh các điểm giống nhau và khác nhau giữa hai mô hình OSI và TCP/IP ?
- Câu 3.** Trình bày nhiệm vụ và các giao thức lõi của tầng Internet trong mô hình TCP/IP
- Câu 4.** Trình bày nhiệm vụ chính của giao thức IP và các đặc điểm của nó .
- Câu 5.** Trình bày cấu trúc của một địa chỉ IP
- Câu 6.** Trình bày cấu trúc các lớp địa chỉ IP có thể gán được ?
- Câu 7.** Lấy ba ví dụ địa chỉ IP thuộc các lớp A, B, C
- Câu 8.** Xác định khoảng địa chỉ, số mạng và số host của các lớp A, B, C.
- Câu 9.** Đơn vị dữ liệu dùng trong IP là gì ? Vẽ khuôn dạng dữ liệu của nó.
- Câu 10.** Cho biết ý nghĩa của thông số Type of service và cấu tạo của thông số này .
- Câu 11.** Liệt kê các thông điệp ICMP thông thường nhất .
- Câu 12.** Nêu một số hàm cơ bản trong TCP và chức năng của chúng ?
- Câu 13.** Các dịch vụ chính của Internet
- Câu 14.** Trình bày các giao thức của hệ thống thư điện tử ?
- Câu 15.** Nêu mục tiêu của dịch vụ truyền tập tin FTP ?



## BÀI TẬP TRẮC NGHIỆM

- Giao thức nào sau đây không phải giao thức của tầng Internet:
  - Giao thức ARP.
  - Giao thức UDP
  - Giao thức ICMP
  - Giao thức IGMP
- Nhiệm vụ của giao thức IP:
  - Đánh địa chỉ IP.
  - Cho phép trao đổi tập tin qua mạng Internet.
  - Cung cấp khả năng kết nối các mạng con thành liên kết mạng để truyền dữ liệu.
  - Cả ba đáp án trên đều sai.
- Độ dài của một địa chỉ IP
  - 4 bytes.
  - 16 bits.
  - 3 bytes.
  - 8 bits.
- Mỗi địa chỉ IP được tách thành mấy vùng :
  - 5 vùng.
  - 4 vùng.
  - 3 vùng.
  - 2 vùng.
- Phát biểu nào sau đây là sai :
  - Mạng lớp A: địa chỉ mạng (netid) là 1 byte và địa chỉ host (hostid) là 3 byte
  - Địa chỉ IP được thành 5 lớp A, B, C, D, E
  - Mạng lớp B: địa chỉ mạng (netid) là 2 byte và địa chỉ host (hostid) là 2 byte
  - Các bit đầu tiên của byte đầu tiên của mạng lớp D có dạng: 110-

6. Địa chỉ IP sau thuộc lớp nào?

10000000	00000101	00000010	00010001
----------	----------	----------	----------

- Lớp A.
- Lớp B
- Lớp C
- Lớp D

7. Một địa chỉ IP có biểu diễn dưới dạng nhị phân như sau:

11000000	10101000	00000001	00010110
----------	----------	----------	----------

Biểu diễn dạng thập phân của địa chỉ này là:

- 192.160.1.36
- 192.168.1.36
- 192.168.1.38
- 128.168.1.38

8. Ý nghĩa của thông số IHL trong khuôn dạng của datagram là:

- Chỉ độ dài phần đầu của gói tin.
- Xác định version hiện tại của giao thức IP.
- Chỉ thị về quyền ưu tiên của gói tin.
- Cả ba đáp án trên đều sai.

9. Ý nghĩa của thông số khi R = 1:

- Gói tin có quyền ưu tiên khi gửi thấp.
- Độ thông lượng yêu cầu để truyền gói tin cao.
- Gói tin có độ tin cậy cao.
- Gói tin có độ trễ bình thấp.

10. Giao thức nào sau đây là giao thức trong mạng IP?

- MINE.
- WAP.
- UDP.
- ICMP.

11. Đơn vị dữ liệu sử dụng trong IP:

- Bit.
- Datagram.
- Byte.
- Frame.

12. Chức năng của giao thức ARP:
- Phân giải địa chỉ tầng Internet, chuyển thành địa chỉ tầng giao tiếp mạng.
  - Quản lý các nhóm IP truyền multicast.
  - Tìm địa chỉ IP từ địa chỉ vật lý.
  - Cả ba đáp án trên đều sai.
13. Khi người sử dụng gửi đi một yêu cầu mở liên kết sẽ được nhận hai thông số trả lời từ TCP, đó là:
- MessageID và OpenSuccess.
  - OpenID và OpenSuccess.
  - OpenID và MessageID.
  - Cả 3 đáp án trên đều sai.
14. Một máy tính được gọi là thành viên của Internet khi:
- Máy tính đó có giao thức truyền dữ liệu TCP/IP.
  - Có một địa chỉ IP trên mạng.
  - Có thể gửi các gói tin IP đến tất cả các máy tính khác trên mạng Internet.
  - Cả ba đáp án trên đều đúng.
15. Các dịch vụ chính của Internet:
- WWW (World Wide Web), News, Thư điện tử, File Receiver, HTTP
  - WWW (World Wide Web), File Transfer, File Receiver, Remote Login, HTTP
  - News, File Transfer, Remote Login, WWW, Thư điện tử
  - Remote Login, WWW, Thư điện tử, HTTP.