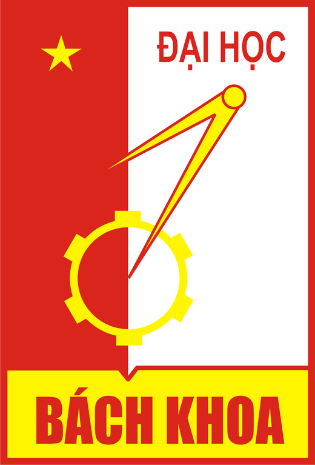
**ĐẠI HỌC BÁCH KHOA HÀ NỘI  
TRƯỜNG ĐIỆN – ĐIỆN TỬ  
\*\*\*\*\* □&□ \*\*\*\*\***



**BÁO CÁO BÀI TẬP LỚN:  
*MẠNG MÁY TÍNH***

**ĐỀ TÀI: LẬP TRÌNH SOCKET ỨNG DỤNG CHIA SẺ FILE DÙNG CƠ CHẾ CLIENT – SERVER**

|  |  |
| --- | --- |
| **GVHD**: PGS. TS. Nguyễn Quang Vinh  **Mã lớp**: | |
| **Họ và tên** | **MSSV** |
| Nguyễn Thị Minh Châu | 20192716 |
|  |  |
|  |  |
|  |  |

***Hà Nội, 12 /2022***

LỜI NÓI ĐẦU

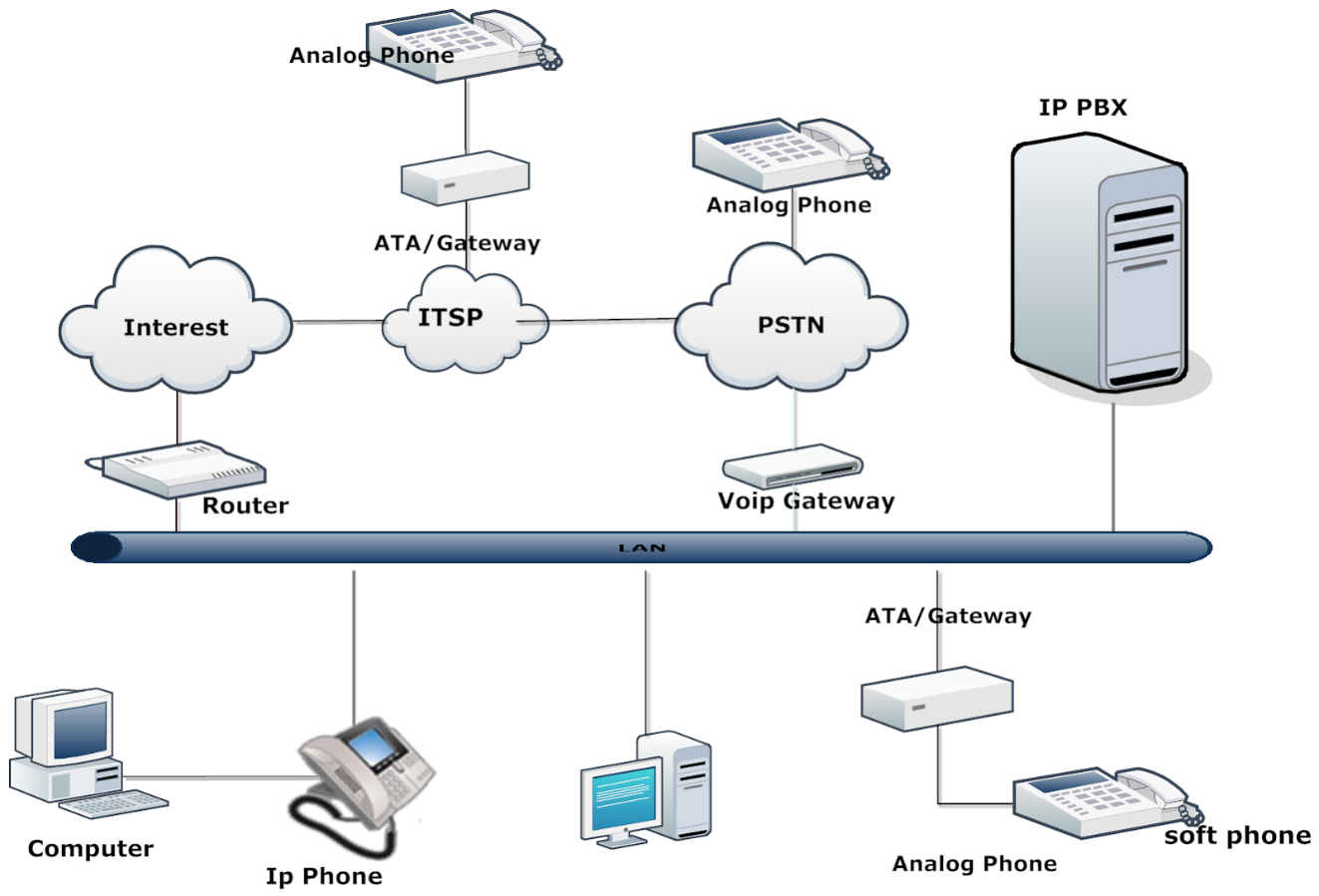
PHẦN 1: TỔNG QUAN

## Tổng quan mạng máy tính

### Mạng máy tính là gì?

Mạng máy tính là một tập hợp các máy tính được nối với nhau bởi đường truyền theo một cấu trúc nào đó và thông qua đó các máy tính trao đổi thông tin qua lại cho nhau.

Đường truyền là hệ thống các thiết bị truyền dẫn có dây hay không dây dùng để chuyển các tín hiệu điện tử từ máy tính này đến máy tính khác. Các tín hiệu điện tử đó biểu thị các giá trị dữ liệu dưới dạng các xung nhị phân (on - off). Tất cả các tín hiệu được truyền giữa các máy tính đều thuộc một dạng sóng điện từ. Tùy theo tần số của sóng điện từ có thể dùng các đường truyền vật lý khác nhau để truyền các tín hiệu. Ở đây đường truyền được kết nối có thể là dây cáp đồng trục, cáp xoắn, cáp quang, dây điện thoại, sóng vô tuyến ... Các đường truyền dữ liệu tạo nên cấu trúc của mạng. Hai khái niệm đường truyền và cấu trúc là những đặc trưng cơ bản của mạng máy tính.



### Tại sao cần mạng máy tính?

Ngày nay, khi máy tính được sử dụng một cách rộng rãi và số lượng máy tính trong một văn phòng hay cơ quan được tăng lên nhanh chóng thì việc kết nối chúng trở nên vô cùng cần thiết và sẽ mang lại nhiều hiệu quả cho người sử dụng.

Với một lượng lớn về thông tin, nhu cầu xử lý thông tin ngày càng cao, mạng máy tính đã trở nên quá quen thuộc đối với chúng ta trong mọi lĩnh vực như: khoa học, quân sự, quốc phòng, thương mại, dịch vụ, giáo dục... Người ta thấy được việc kết nối các máy tính thành mạng cho chúng ta những khả năng mới to lớn như:

- *Sử dụng chung tài nguyên*: những tài nguyên (như thiết bị, chương trình, dữ liệu) khi được trở thành các tài nguyên chung thì mọi thành viên của mạng đều có thể tiếp cận được mà không quan tâm tới những tài nguyên đó ở đâu.

- *Tăng độ tin cậy của hệ thống*: người ta có thể dễ dàng bảo trì máy móc, lưu trữ (backup) các dữ liệu chung và khi có trục trặc trong hệ thống thì chúng có thể được khôi phục nhanh chóng. Trong trường hợp có trục trặc trên một trạm làm việc thì người ta cũng có thể sử dụng những trạm khác thay thế.

- *Nâng cao chất lượng và hiệu quả khai thác thông tin*: khi thông tin có thể được sử dụng chung thì nó mang lại cho người sử dụng khả năng tổ chức lại các công việc với những thay đổi về chất như:

+ Ðáp ứng những nhu cầu của hệ thống ứng dụng kinh doanh hiện đại.

+ Cung cấp sự thống nhất giữa các dữ liệu.

+ Tăng cường năng lực xử lý nhờ kết hợp các bộ phận phân tán.

+ Tăng cường truy nhập tới các dịch vụ mạng khác nhau đang được cung cấp trên thế giới.

Với nhu cầu đòi hỏi ngày càng cao của xã hội nên vấn đề kỹ thuật trong mạng là mối quan tâm hàng đầu của các nhà tin học. Ví dụ như làm thế nào để truy xuất thông tin một cách nhanh chóng và tối ưu, trong khi việc xử lý thông tin trên mạng quá nhiều, đôi khi có thể làm tắc nghẽn và gây ra mất thông tin một cách đáng tiếc. Hiện nay, việc làm sao có được một hệ thống mạng chạy thật tốt, thật an toàn với lợi ích kinh tế cao đang rất được quan tâm.

### Phân loại mạng máy tính

Do hiện nay mạng máy tính được phát triển khắp nơi với những ứng dụng ngày càng đa dạng cho nên việc phân loại mạng máy tính là một việc rất phức tạp.

Dựa theo phạm vi phân bố của mạng ta có thể phân ra các loại mạng như sau:

- *Mạng cục bộ LAN (Local Area Network)*: Mạng LAN là một nhóm máy tính và các thiết bị truyền thông mạng được nối kết với nhau trong một khu vực nhỏ như một tòa nhà cao ốc, khuôn viên trường đại học, khu giải trí... Các mạng LAN thường có đặc điểm sau:

+ Băng thông lớn, có khả năng chạy các ứng dụng trực tuyến như xem phim, hội thảo qua mạng.

+ Kích thước mạng bị giới hạn bởi các thiết bị

+ Chi phí các thiết bị mạng LAN tương đối rẻ

+ Quản trị đơn giản

- *Mạng đô thị MAN (Metropolitan Area Network):* Mạng MAN gần giống như mạng LAN nhưng giới hạn của nó là một thành phố hay một quốc gia. Mạng MAN nối kết các mạng LAN lạ với nhau thông qua các phương tiện truyền dẫn khác nhau (cáp quang, cáp đồng, sóng....) và các phương thức truyền thông khác nhau.

Đặc điểm của mạng MAN:

+ Băng thông mức trung bình, đủ để phục vụ các ứng dụng cấp thành phố hay quốc gia như chính phủ điện tử, thương mại điện tử, các ứng dụng của các ngân hàng...

+ Do MAN nối kết nhiều LAN với nhau nên độ phức tạp cũng tăng đồng thời công tác quản trị sẽ khó khăn hơn.

+ Chi phí các thiết bị mạng MAN tương đối đắt tiền.

- *Mạng diện rộng WAN (Wide Area Network):* Mạng WAN bao phủ vùng địa lý rộng lớn có thể là một quốc gia, một lục địa hay toàn cầu. Mạng WAN thường là mạng của các công ty đa quốc gia hay toàn cầu, điển hình là mạng internet. Do phạm vi rộng lớn của mạng WAN nên thông thường mạng WAN là tập hợp các mạng LAN, WAN nối lại với nhau bằng các phương tiện như: vệ tinh (satellites), sóng biva (microwave), cáp quang, cáp điện thoại.

Đặc điểm của mạng WAN:

+ Băng thông thấp, dễ mất kết nối, thường chỉ phù hơp với các ứng dụng offine như e-mail, web, ftp...

+ Phạm vi hoạt động rộng lớn không giới hạn

+ Do kết nối của nhiều LAN, WAN lại với nhau nên mạng rất phức tạp và có tính toàn cầu nên thường là có tổ chức quốc tế đứng ra quản trị

+ Chi phí cho các thiết bị và các công nghệ mạng WAN rất đắt tiền

- *Mạng Internet:* Là trường hợp đặc biệt của mạng WAN, nó cung cấp các dịch vụ toàn cầu như mail, web, chat, ftp và phục vụ miễn phí cho mọi người.

### Một số topo mạng thông dụng

Theo định nghĩa về mạng máy tính, các máy tính được nối với nhau bởi các đường truyền vật lý theo một kiến trúc nào đó, các kiến trúc đó gọi là Topology. Thông thường mạng có ba loại kiến trúc đó là: mạng hình sao (Star Topology), mạng dạng tuyến (Bus Topology), mạng dạng vòng(Ring Topology).

- *Ring Topology:* Mạng được bố trí vòng tròn, đường dây cáp được thiết kế làm thành một vòng khép kín, tín hiệu chạy theo một chiều nào đó. Các nút truyền tín hiệu cho nhau tại một thời điểm được một nút mà thôi. Mạng dạng vòng có thuận lợi là có thể nới rộng ra xa nhưng đường dây phải khép kín, nếu bị ngắt ở một nơi nào đó thì toàn bộ hệ thống cũng bị ngưng.

- *Bus Topology:* Ở dạng Bus tất cả các nút được phân chia một đường truyền chính (bus). Đường truyền này được giới hạn hai đầu bởi một loại đầu nối đặc biệt gọi là Terminator. Khi một nút truyền dữ liệu, tín hiệu được quảng bá trên hai chiều của bus, mọi nút còn lại đều được nhận tín hiệu trực tiếp. Loại mạng này dùng dây cáp ít, dễ lắp đặt. Tuy vậy cũng có những bất lợi đó là sẽ có sự ùn tắc giao thông khi di chuyển với lưu lượng lớn và khi có sự hỏng hóc ở đoạn nào đó thì rất khó phát hiện, nếu một nút ngừng hoạt động sẽ ảnh hưởng tới toàn bộ hệ thống

- *Star Topology:* Mạng hình sao bao gồm một bộ tập trung và các nút thông tin. Các nút thông tin có thể là các trạm cuối, các máy tính hay các thiết bị khác của mạng. Mạng hoạt động theo nguyên lý nối song song nên nếu có một nút bị hỏng mạng vẫn hoạt động bình thường. Mạng có thể mở rộng hoặc thu hẹp tùy theo yêu cầu của người sử dụng, tuy nhiên mở rộng phụ thuộc và khả năng của trung tâm.

### Các giao thức mạng

Giao thức mạng là một tập các quy tắc, quy ước để trao đổi thông tin giữa hai hệ thống máy tính hoặc hai thiết bị máy tính với nhau. Nói một cách hình thức thì giao thức mạng là một ngôn ngữ được các máy tính trong mạng sử dụng để trao đổi dữ liệu với nhau. Có nhiều loại giao thức được sử dụng trong mạng máy tính như: Apple Talk, DLC, NetBEUI,... nhưng hiện nay giao thức được sử dụng phổ biến nhất trong mạng máy tính là giao thức TCP/IP.

#### Giao thức TCP

##### Tổng quan giao thức TCP

TCP (Transmission Control Protocol) là giao thức hướng kết nối, nó cung cấp một đường truyền dữ liệu tin cậy giữa hai máy tính. Tính tin cậy thể hiện ở việc nó đảm bảo dữ liệu được gửi sẽ đến được đích và theo đúng thứ tự như khi nó được gửi.

Khi hai ứng dụng muốn giao tiếp với nhau một cách tin cậy, chúng sẽ tạo ra đường kết nối giữa chúng và gửi dữ liệu thông qua đường này.

Tính tin cậy của đường truyền được thể hiện ở hai điểm sau:

* Mọi gói tin cần gửi sẽ đến được đích. Để làm được điều này thì mỗi lần phía gửi gửi xong một gói tin nó sẽ chờ nhận một xác nhận từ bên nhận rằng đã nhận được gói tin. Nếu sau một khoảng thời gian mà phía gửi không nhận được thông tin xác nhận phản hồi thì nó sẽ phát lại gói tin. Việc phát lại sẽ được tiến hành cho đến khi việc truyền tin thành công, tuy nhiên sau một số lần phát lại max nào đó mà vẫn chưa thành công thì phía gửi có thể suy ra là không thể truyền tin được và sẽ dừng việc phát tin.
* Các gói tin sẽ được trình ứng dụng nhận được theo đúng thứ tự như chúng được gửi. Bởi các gói tin có thể được dẫn đi trên mạng theo nhiều đường khác nhau trước khi tới đích nên thứ tự khi tới đích của chúng có thể không giống như khi chúng được phát. Do đó để đảm bảo có thể sắp xếp lại gói tin ở phía nhận theo đúng thứ tự như khi chúng được gửi, giao thức TCP sẽ gắn vào mỗi gói tin một thông tin cho biết thứ tự của chúng trong cả khối tin chung được phát nhờ vậy bên nhận có thể sắp xếp lại các gói tin theo đúng thứ tự của chúng.

Như vậy có thể thấy TCP cung cấp cho chúng ta một kênh truyền thông điểm - điểm phục vụ cho các ứng dụng đòi hỏi giao tiếp tin cậy như HTTP (HyperText Transfer Protocol), FTP (File Tranfer Protocol), Telnet.... Các ứng dụng này đòi hỏi một kênh giao tiếp tin cậy bởi thứ tự của dữ liệu được gửi và nhận là yếu tố quyết định đến sự thành công hay thất bại của chúng. Hãy lấy ví dụ khi HTTP được sử dụng để đọc thông tin từ một địa chỉ URL, dữ liệu phải được nhận theo đúng thứ tự mà chúng được gửi nếu không thứ mà bạn nhận được có thể là một trang HTML với nội dung lộn xộn hoặc một file zip bị lỗi và không giải nén…

##### Cấu trúc gói tin TCP

#### Giao thức UDP

Định nghĩa: UDP (User Datagram Protocol) là giao thức không hướng kết nối, nó gửi các gói dữ liệu độc lập gọi là datagram từ máy tính này đến máy tính khác mà không đảm bảo việc dữ liệu sẽ tới đích. Ở phần trước chúng ta đã thấy trong giao thức TCP khi hai chương trình muốn giao tiếp với nhau qua mạng chúng tạo ra một kết nối liên kết hai ứng dụng và trao đổi dữ liệu qua kết nối đó. Trái lại ở giao thức UDP khi hai ứng dụng muốn giao tiếp với nhau chúng không tạo ra kết nối mà chỉ đơn thuần gửi các gói tin một cách độc lập từ máy này tới máy khác.Các gói tin như vậy gọi là các datagram. Việc gửi các gói tin như vậy tương tự như việc chúng ta gửi thư qua đường bưu điện. Các bức thư bạn gửi độc lập với nhau, thứ tự các thư là không quan trọng và không có gì đảm bảo là thư sẽ đến được đích. Trong truyền thông bằng UDP thì các datagram giống như các lá thư, chúng chứa thông tin cần gửi đi cùng thông tin về địa chỉ đích mà chúng phải đến, tuy nhiên chúng khác với các lá thư ở một điểm là nếu như trong việc gửi thư, nếu lá thư của bạn không đến được đích thì nó sẽ được gửi trả lại nơi gửi nêu trên lá thư đó bạn có đề địa chỉ gửi còn UDP sẽ không thông báo gì cho phía gửi về việc lá thư đó có tới được đích hay không.

Vậy nếu UDP là một giao thức không đảm bảo giao tiếp tin cậy thì tại sao người ta lại dùng chúng. Điều đó là bởi nếu như giao thức TCP đảm bảo một kết nối tin cậy giữa các ứng dụng thì chúng cũng đòi hỏi nhiều thời gian để truyền tin do chúng phải kiểm tra các header các gói tin để đảm bảo thứ tự các gói tin cũng như để phát lại các gói tin không đến được đích do đó trong một số trường hợp thì điều này không cần thiết. Dưới đây là một số trường hợp trong đó giao thức không hướng kết nối là thích hợp hơn so với giao thức hướng kết nối:

Khi chỉ một gói dữ liệu cần truyền đi và việc đó đến được đích hay không là không quan trọng, sử dụng giao thức UDP sẽ loại bỏ được các thủ tục tạo và hủy kết nối. So sánh một chút chúng ta sẽ thấy giao thức hướng kết nối TCP phải dùng đến bẩy gói tin để gửi một gói tin do nó cần phát và nhận các gói tin yêu cầu và chấp nhận kết nối cũng như các gói tin yêu cầu và xác nhận việc hủy kết nối, trong khi đó giao thức không kết nối UDP chỉ sử dụng duy nhất một gói tin chính là gói tin chứa dữ liệu cần chuyển đi.

Chúng ta lấy ví dụ về một server đồng hồ, nhiệm vụ của nó là gửi thời gian hiện tại của nó cho các ứng dụng trên client khi có yêu cầu. Nếu gói tin chứa thời gian bị thất lạc trên đường truyền và không tới được đích thì client cũng sẽ không đòi hỏi server phải gửi lại gói tin đó bởi khi gói tin đó được phát lại lần hai và tới được client thì thông tin thời gian chứa trong nó đã không còn đúng nữa. Nếu client tạo ra hai yêu cầu và nhận được các gói tin trả lời không theo đúng thứ tự mà server đã gửi thì client cũng không gặp phải vấn đề gì bởi nó hoàn toàn có thể suy ra được rằng các gói đã không được chuyển đến đúng thứ tự bằng cách tính thời gian được chứa trong các gói. Trong trường hợp này tính tin cậy của TCP là không cần thiết bởi nó làm giảm hiệu suất và có thể cản trở hoạt động của server.

Trường hợp thứ hai chúng ta xem xét việc sử dụng giao thức UDP là các ứng dụng đòi hỏi chặt chẽ về thời gian như các ứng dụng nghe audio thời gian thực. Trong trường hợp này việc hướng tới một kênh giao tiếp tin cậy không phải là ưu điểm mà ngược lại đó là một nhược điểm bởi nếu việc phải chờ cho khi một gói tin bị mất được nhận có thể gây ra những tác động dễ nhận thấy hoặc khiến chương trình phải tạm ngừng. Với các ứng dụng này giao thức không hướng kết nối đã được phát triển và chúng làm việc tốt hơn hẳn. Chúng ta có thể tham khảo ứng dụng RealAudio, trong đó người ta sử dụng một giao thức không hướng kết nối để truyền các dữ liệu âm thanh qua mạng.

Bảng sau so sánh sự khác biệt giữa hai chế độ giao tiếp hướng kết nối và không hướng kết nối.

|  |  |
| --- | --- |
| Chế độ có hướng kết nối (TCP) | Chế độ không hướng kết nối (UDP) |
| Tồn tại kênh giao tiếp ảo giữa hai bên giao tiếp | Không tồn tại kênh giao tiếp ảo giữa hai bên giao tiếp |
| Dữ liệu được gửi đi theo chế độ bảo đảm: Có kiểm tra lỗi truyền lại gói tin lỗi hay mất bảo đảm thứ tự đến của các gói tin... | Dữ liệu được gửi đi theo chế độ không bảo đảm: Không kiểm tra lỗi, không phát hiện không truyền lại gói tin bị lỗi hay mất, không bảo đảm thứ tự đến của các gói tin... |
| Dữ liệu chính xác, tốc độ truyền chậm | Dữ liệu không chính xác, tốc độ truyền nhanh. |
|  | Thích hợp cho các ứng dụng cần tốc độ, không cần chính xác cao: Truyền âm thanh, hình ảnh |

### Các mô hình hoạt động của mạng máy tính

#### Mô hình peer-to-peer

#### Mô hình client – server

## Tổng quan socket

### Khái niệm socket

#### Lịch sử hình thành

#### Nguyên lý hoạt động

### Các loại socket

### Cấu trúc socket

### Một số hàm cơ bản dùng trong lập trình socket

# PHẦN 2: XÂY DỰNG CHƯƠNG TRÌNH ỨNG DỤNG