

LỜI NÓI ĐẦU

Hệ thống phân tán được biết đến như một hệ thống với nhiều bộ xử lý được nối kết lại với nhau nhờ vào một hệ thống mạng nhằm cung cấp cho các ứng dụng một nền tảng tính toán có năng lực mạnh mẽ, tài nguyên dồi dào và độ tin cậy vững chắc. Các ứng dụng giờ đây không chỉ chạy và khai thác tài nguyên trên một máy tính đơn lẻ mà chúng có thể được phân rã thành nhiều thành phần để chạy trên nhiều máy tính khác nhau và khai thác tài nguyên của toàn hệ thống. Người ta gọi các ứng dụng vận hành trên một hệ thống phân tán là các ứng dụng phân tán. Internet và các ứng dụng trên mạng Internet là một ví dụ điển hình về hệ thống phân tán và các ứng dụng phân tán. Các ứng dụng phân tán có nhiều lợi thế hơn so với các ứng dụng vận hành trên các máy tính đơn lẻ. Tuy nhiên, phát triển các ứng dụng phân tán rất phức tạp, đòi hỏi người lập trình phải xử lý nhiều vấn đề phát sinh như truyền thông giữa các thành phần của ứng dụng, sự tranh chấp tài nguyên giữa các ứng dụng, sự không đồng nhất về phần cứng, phần mềm hệ thống trong một hệ thống phân tán,...

Nhiều nghiên cứu đã và đang được thực hiện nhằm vào việc khai thác tối đa thế mạnh của các hệ thống phân tán và làm cho việc phát triển các ứng dụng phân tán trở nên đơn giản và dễ dàng như việc phát triển các ứng dụng trên một máy tính đơn. Giáo trình này sẽ chọn lọc và giới thiệu đến người đọc những mô hình, giải pháp, giải thuật tiêu biểu từ những nghiên cứu trên. Trong chương 1, chúng tôi bắt đầu bằng việc giới thiệu tổng quan về hệ thống phân tán và chỉ ra những vấn đề liên quan đến việc xây dựng một hệ thống phân tán. Trong chương 2 chúng tôi sẽ trình bày vai trò của các hệ thống truyền thông điệp trong một hệ thống phân tán và một số mô hình xây dựng hệ thống truyền thông điệp phổ biến hiện nay. Trong một hệ thống phân tán, các ứng dụng cũng như các thành phần của một ứng dụng (thể hiện dưới dạng các quá trình) luôn có sự hợp tác với nhau để hoàn thành một công việc nào đó. Điều này đòi hỏi cần có sự thống nhất về mặt thời gian giữa các quá trình. Trong chương 3 chúng tôi sẽ trình bày các giải thuật liên quan đến việc đồng bộ hóa thời gian giữa các quá trình và các cơ chế phối hợp giữa chúng. Trong chương 4 chúng tôi sẽ trình bày các tiếp cận cho phép quản lý và khai thác các tài nguyên như CPU, đĩa cứng,.. trong một hệ thống phân tán một cách hiệu quả nhất. Việc quản lý và khai thác tài nguyên trên một hệ thống phân tán đôi khi cần thiết phải di dời một tiến trình đang chạy trên máy tính này sang một máy tính khác. Người ta gọi đây là cơ chế di dời tiến trình. Các cơ chế di dời quá trình tiêu biểu sẽ được chúng tôi giới thiệu trong chương 5. Trong một hệ thống phân tán, không những tài nguyên phục vụ cho việc tính toán như CPU được phân bố rải rác trên nhiều máy tính khác nhau, mà các tài nguyên dùng cho lưu trữ dữ liệu như đĩa cứng cũng bị phân tán. Trong chương cuối cùng chúng tôi sẽ trình bày những vấn đề cần quan tâm khi xây dựng một hệ thống tập tin phân tán trong một hệ thống phân tán.

Với mong muốn cung cấp cho độc giả một cái nhìn tổng quan về hệ thống phân tán và các vấn đề liên quan đến việc xây dựng hệ thống phân tán, nhóm biên soạn đã rất cố gắng trong việc chọn lọc các nội dung trình bày. Tuy nhiên giáo trình chắc chắn vẫn còn nhiều vấn đề thiếu sót. Nhóm biên soạn rất mong nhận được các góp ý từ phía độc giả về mặt nội dung cũng như hình thức trình bày của giáo trình.

Nhóm biên soạn xin chân thành cảm ơn độc giả đã đón nhận quyển giáo trình này. Hy vọng giáo trình sẽ là một trong những người bạn đồng hành của quý vị độc giả trong con đường chinh phục thế giới điện toán.

CÁC TÁC GIẢ

MỤC LỤC

Chương 1. TỔNG QUAN VỀ HỆ THỐNG PHÂN TÁN	1
1.1 GIỚI THIỆU	1
1.2 HỆ THỐNG PHÂN TÁN LÀ GÌ?	1
1.3 NHỮNG LỢI ÍCH CỦA HỆ THỐNG PHÂN TÁN	3
1.4 NHỮNG MÔ HÌNH CỦA HỆ THỐNG PHÂN TÁN	3
1.4.1 Mô hình máy tính tâm trung	4
1.4.2 Mô hình trạm làm việc	4
1.4.3 Mô hình trạm làm việc - máy phục vụ	5
1.4.4 Mô hình vùng các bộ xử lý	6
1.4.5 Mô hình kết hợp	8
1.5 VAI TRÒ PHẦN MỀM TRONG HỆ THỐNG PHÂN TÁN	8
1.5.1 Hệ điều hành mạng	8
1.5.2 Phần mềm trung gian	9
1.5.3 Hệ điều hành phân tán	9
1.6 MỤC TIÊU HƯỚNG ĐẾN KHI THIẾT KẾ CÁC HỆ THỐNG PHÂN TÁN	10
1.6.1 Tính trong suốt	10
1.6.2 Tính mềm dẻo	13
1.6.3 Tính tin cậy	15
1.6.4 Hiệu năng của hệ thống	17
1.6.5 Tính năng mở rộng	19
1.6.6 Tính không đồng nhất	20
1.6.7 Tính bảo mật	20
1.6.8 Mô phỏng những hệ điều hành đã có	21
1.7 NHỮNG VẤN ĐỀ CẦN GIẢI QUYẾT KHI THIẾT KẾ HỆ THỐNG PHÂN TÁN	21
1.8 TÓM TẮT	22
CÂU HỎI ÔN TẬP	22
TÀI LIỆU THAM KHẢO	23
Chương 2. GIAO TIẾP LIÊN QUÁ TRÌNH	24
2.1 GIỚI THIỆU	24
2.2 QUÁ TRÌNH VÀ LUỒNG	25

2.3 GIAO TIẾP LIÊN QUÁ TRÌNH SỬ DỤNG SOCKET	25
2.3.1 Mô hình Client-Server	25
2.3.2 Socket	26
2.4 ĐẶT TÊN	28
2.4.1 Dịch vụ tên miền trên mạng Internet	30
2.4.2 Dịch vụ đặt tên cho các client di động	31
2.5 LỜI GỌI THỦ TỤC XA	32
2.5.1 Tiến trình thực hiện một lời gọi thủ tục xa	32
2.5.2 Những vấn đề cần quan tâm khi cài đặt cơ chế lời gọi thủ tục xa	34
2.5.3 Lập trình với lời gọi thủ tục xa	36
2.5.4 Các thể hệ của RPC	37
2.6 TÓM TẮT	50
CÂU HỎI ÔN TẬP	51
TÀI LIỆU THAM KHẢO	51
Chương 3. ĐỒNG BỘ HÓA THỜI GIAN VÀ PHỐI HỢP	52
3.1 GIỚI THIỆU	52
3.2 ĐỒNG BỘ HÓA ĐỒNG HỒ	52
3.2.1 Tầm nhìn nhất quán về thời gian	52
3.2.2 Đồng hồ luận lý	53
3.2.3 Đồng hồ vật lý	58
3.2.4 Bù đắp cho độ lệch drift	59
3.2.5 Thiết lập thời gian trên đồng hồ vật lý	60
3.2.6 Giải thuật của Cristian	61
3.2.7 Giải thuật Berkeley	61
3.2.8 Giao thức thời gian mạng	62
3.3 SỰ PHỐI HỢP VÀ THỐNG NHẤT	66
3.3.1 Giải thuật server tập trung	66
3.3.2 Loại trừ hồ tương phân tán của Ricart & Agrawala	67
3.3.3 Giải thuật Token Ring	70
3.3.4 So sánh ba giải thuật	71
3.3.5 Giải thuật bình chọn	72
3.4 TÓM TẮT	75
CÂU HỎI ÔN TẬP	75
TÀI LIỆU THAM KHẢO	76

Chương 4. QUẢN LÝ TÀI NGUYÊN	77
4.1 GIỚI THIỆU	77
4.2 TIẾP CẬN PHÂN CHIA TÁC VỤ	78
4.2.1 Ý tưởng cơ bản	78
4.2.2 Tìm kiếm một sự phân chia tối ưu	78
4.3 TIẾP CẬN CÂN BẰNG TẢI	81
4.3.1 Chính sách ước lượng tải	81
4.3.2 Chính sách di dời quá trình	82
4.3.3 Chính sách định vị	84
4.3.4 Chính sách trao đổi thông tin trạng thái	86
4.3.5 Chính sách phân chia độ ưu tiên	87
4.3.6 Chính sách giới hạn việc di dời	88
4.4 TIẾP CẬN CHIA SẼ TẢI	88
4.4.1 Chính sách ước lượng tải	88
4.4.2 Chiến lược cho việc di dời quá trình	88
4.4.3 Các chính sách định vị	89
4.4.4 Chính sách trao đổi thông tin trạng thái	90
4.5 TÓM TẮT	90
CÂU HỎI ÔN TẬP	91
TÀI LIỆU THAM KHẢO	93
Chương 5. DI DỜI QUÁ TRÌNH	94
5.1 GIỚI THIỆU	94
5.2 DI DỜI QUÁ TRÌNH	94
5.2.1 Những cơ chế di dời quá trình	95
5.2.2 Những cơ chế di dời không gian địa chỉ	97
5.2.3 Những cơ chế chuyển tiếp thông điệp	99
5.3 LUỒNG	101
5.3.1 Những động lực cho việc sử dụng luồng	102
5.3.2 Những mô hình cho việc tổ chức các luồng	102
5.3.3 Cài đặt gói luồng	104
5.4 TÓM TẮT	106
CÂU HỎI ÔN TẬP	107
TÀI LIỆU THAM KHẢO	109

Chương 6. HỆ THỐNG TẬP TIN PHÂN TÁN	110
6.1 GIỚI THIỆU	110
6.2 NHỮNG MÔ HÌNH TẬP TIN	111
6.2.1 Những tập tin có cấu trúc và không có cấu trúc	111
6.2.2 Những tập tin có thể sửa đổi và không thể sửa đổi	112
6.3 NHỮNG MÔ HÌNH TRUY XUẤT TẬP TIN	113
6.3.1 Phương pháp truy xuất tập tin từ xa	113
6.3.2 Đơn vị truyền dữ liệu	114
6.4 NHỮNG NGỮ CẢNH CHIA SẺ TẬP TIN	115
6.4.1 Những ngữ cảnh UNIX	116
6.4.2 Ngữ cảnh phiên làm việc	117
6.4.3 Những ngữ cảnh chia sẻ tập tin dạng không thể sửa đổi	117
6.4.4 Ngữ cảnh tựa giao dịch	118
6.5 NHỮNG SƠ ĐỒ LƯU TRỮ ĐỆM TẬP TIN	118
6.5.1 Vị trí lưu trữ	119
6.5.2 Bộ nhớ chính của máy phục vụ	119
6.5.3 Đĩa của khách hàng	120
6.5.4 Bộ nhớ chính trên máy khách hàng	120
6.5.5 Truyền sự sửa đổi	121
6.5.6 Những sơ đồ cập nhật vùng lưu trữ đệm	123
6.6 NHÂN BẢN TẬP TIN	125
6.6.1 Sự khác nhau giữa nhân bản và lưu trữ đệm	125
6.6.2 Những ưu điểm của nhân bản	125
6.6.3 Trong suốt trong nhân bản	126
6.6.4 Vấn đề cập nhật nhiều bản sao	127
6.7 CHỊU LỖI	131
6.7.1 Lưu trữ bền vững	132
6.7.2 Hiệu quả của mô hình dịch vụ chịu đựng lỗi	133
6.8 NHỮNG GIAO DỊCH NGUYÊN TỬ	137
6.8.1 Nhu cầu hỗ trợ giao dịch trong một dịch vụ tập tin	137
6.8.2 Những phép toán trong một dịch vụ tập tin dựa trên giao dịch	137
6.8.3 Những kỹ thuật phục hồi	139
6.9 ĐIỀU KHIỂN ĐỒNG HÀNH	144

<i>Giáo trình Các hệ thống phân tán</i>	<i>Mục lục</i>
6.9.1 Khóa	145
6.9.2 Tối ưu hóa điều khiển đồng hành	145
6.9.3 Ghi tem thời gian	145
6.10 TÓM TẮT	147
CÂU HỎI ÔN TẬP	147
TÀI LIỆU THAM KHẢO	150

Chương 1

TỔNG QUAN VỀ HỆ THỐNG PHÂN TÁN

1.1 GIỚI THIỆU

Sự phát triển nhanh chóng của công nghệ vi điện tử và công nghệ truyền thông đã làm cho cách thức chế tạo và khai thác máy tính không ngừng thay đổi. Ở những năm đầu của thập niên 50, với việc sử dụng đèn điện tử và transistor làm linh kiện chế tạo chính, các máy tính thời đó rất cồng kềnh và đắt tiền. Chúng thường được biết đến với tên gọi là máy tính lớn (Mainframe computer). Chỉ những trung tâm nghiên cứu lớn hay các trường đại học hàng đầu mới có đủ khả năng về tài chính và năng lực vận hành các máy tính này. Mặc dù còn hạn chế về tốc độ xử lý, dung lượng lưu trữ, độ bền, các máy tính thời bấy giờ đã sớm khẳng định được tính hiệu quả và những khả năng mới mà chúng có thể mang đến cho con người trong các hoạt động nghiên cứu khoa học, quân sự, quản lý,.. vì thế nhu cầu sử dụng máy tính ngày càng một tăng lên. Người ta phải suy nghĩ để tìm ra cách thức giúp nhiều người có thể tiếp cận được các máy tính Mainframe đồ sộ và đắt tiền này. Chính trong ngữ cảnh đó, mạng Internet đã được ra đời vào thập niên 70 nhằm vào việc kết nối các máy tính lớn, máy tính tầm trung ở các trung tâm nghiên cứu, các trường đại học Hoa kỳ lại với nhau bằng các đường truyền viễn thông (chủ yếu là đường điện thoại) để chia sẻ phần cứng, phần mềm và dữ liệu giữa các Viện, Trường. Nhờ vào Internet mà một nhà khoa học đang ngồi làm việc trên máy tính ở một trường đại học này có thể truy cập vào một máy tính ở một trường đại học khác cách xa vài ngàn km để thực hiện một chương trình tính toán hoặc tham khảo tài liệu trên máy tính ở xa đấy. Kết quả các phép toán hay dữ liệu mà nhà khoa học muốn truy cập ở máy tính xa có thể sẽ được chuyển về trên máy tính mà ông đang ngồi. Mạng Internet là một thí dụ rất điển hình cho một mô hình sử dụng máy tính mới mà trong đó một yêu cầu công việc của người dùng không được xử lý bởi một máy tính cục bộ tại nơi người dùng đang tọa lạc mà yêu cầu của người dùng có thể được xử lý/đáp ứng bởi nhiều máy tính khác nhau, phân tán rải rác khắp nơi về mặt địa lý thông qua một hệ thống mạng. Internet là một hệ thống tiêu biểu cho một “Hệ thống phân tán” mà chúng tôi giới thiệu trong giáo trình này.

1.2 HỆ THỐNG PHÂN TÁN LÀ GÌ?

Có nhiều định nghĩa khác nhau về hệ thống phân tán. Tanenbaum [Tanenbaum 1995] định nghĩa: “*Một hệ thống phân tán là tập hợp các máy tính độc lập nhau mà chúng xuất hiện trước người dùng hệ thống như là một máy tính đơn*”.

Định nghĩa này đề cập đến 3 khía cạnh quan trọng trong một hệ thống phân tán, đó là: 1) Máy tính trong một hệ thống phân tán là độc lập; 2) Có nhiều máy tính trong một hệ thống phân tán và 3) Hệ thống như là một máy tính đơn.

Khía cạnh thứ nhất đề cập đến *tính tự chủ* của một máy tính trong một hệ thống phân tán. Tính tự chủ thể hiện ở sự độc lập, không ràng buộc lẫn nhau về mặt vật lý và địa lý của các máy tính. Các máy tính trong một hệ thống phân tán có thể được sản xuất bằng những công nghệ khác nhau và có thể được bố trí ở những vị trí cách xa nhau về mặt địa lý. Khái niệm máy tính được sử dụng ở đây phải hiểu theo nghĩa rộng hơn cách mà chúng ta nghĩ thông thường. Một máy tính trong định nghĩa này có thể là các máy tính truyền thống như máy tính lớn, máy tính tầm trung, máy tính để bàn, máy tính xách tay, máy tính mạng (netbook). Thêm vào đó, các thiết bị tính toán mới không dây di động cũng được xem như là các máy tính, thí dụ như điện thoại di động, điện thoại thông minh, các thiết bị nhúng, các cảm biến (sensor), ...

Khía cạnh thứ hai liên quan đến phần cứng của một hệ thống phân tán. Nếu xem một bộ xử lý trung tâm (CPU-Central Processing Unit) là một máy tính thì có 2 loại hệ thống phần cứng có thể thỏa mãn tiêu chí "*bao gồm một tập các máy tính*" đó là máy tính song song và hệ thống mạng. Tuy nhiên, chúng ta sớm loại bỏ các máy tính song song ra khỏi hệ thống phân tán vì các CPU của một máy tính song song không có tính tự chủ (chúng phụ thuộc nhau về mặt vật lý và không thể đặt xa nhau). Người ta thường dùng khái niệm hệ thống **kết nối chặt** (Tightly Coupled System) để chỉ các máy tính song song và hệ thống **kết nối lỏng** (Loosely Coupled System) để chỉ hệ thống mạng của các máy tính độc lập. Như vậy khía cạnh thứ hai cho thấy rằng, hệ thống kết nối lỏng chính là hạ tầng phần cứng của một hệ thống phân tán.

Khía cạnh thứ ba liên quan đến *hình ảnh hệ thống* mà người sử dụng cảm nhận được khi họ sử dụng một hệ thống phân tán. Về căn bản, hệ thống phân tán phải được thiết kế để mang đến cho người dùng cảm tưởng như họ đang sử dụng *một máy tính đơn*. Hệ thống phải che đậy đi sự phức tạp của việc kết nối các máy tính riêng lẻ trong hệ thống phân tán. Muốn thế cần phải có một lớp phần mềm nền tảng phủ lên hạ tầng phần cứng của hệ thống phân tán để cung cấp cho người sử dụng hình ảnh của một hệ thống phân tán như là một máy tính đơn.

Theo định nghĩa trên, các hệ thống sau có thể được xem là các hệ thống phân tán

- Mạng Internet: Nối kết máy tính và chia sẻ tài nguyên toàn cầu
- Mạng Intranet: là mạng cá nhân của một công ty
- Dịch vụ World Wide Web cung cấp thông tin dưới dạng các siêu văn bản
- Các hệ thống thương mại điện tử, đặt hàng trực tuyến, ...
- Các hệ thống mạng di động không dây
- Mạng của các thiết bị nhúng trong một tòa nhà thông minh,
- Hệ thống mạng cảm biến thu thập thông tin môi trường, quan trắc, ...

1.3 NHỮNG LỢI ÍCH CỦA HỆ THỐNG PHÂN TÁN

Ngày nay hệ thống phân tán được sử dụng rộng khắp và đang trở thành xu thế chính định hướng việc sản xuất và khai thác máy tính. Câu hỏi đặt ra là hệ thống phân tán mang lại những lợi ích gì cho chúng ta. Phần này sẽ trình bày những lợi ích và những bất lợi của hệ thống phân tán.

Về mặt lợi ích, hệ thống phân tán mang đến cho chúng ta những lợi ích sau:

- **Tính kinh tế:** các bộ vi xử lý cung cấp tỉ lệ giữa năng lực và giá tốt hơn máy tính Mainframe.
- **Tốc độ:** một hệ thống phân tán có khả năng tính toán tổng thể tốt hơn máy tính Mainframe.
- **Tính phân tán vốn có:** vài ứng dụng đòi hỏi không gian tính toán trên nhiều máy riêng biệt.
- **Khả năng tin cậy:** nếu một máy tính có sự cố, toàn bộ hệ thống vẫn tiếp tục hoạt động.
- **Phát triển nhanh:** năng lực tính toán có thể được bổ sung trong sự gia tăng nhỏ.
- **Chia sẻ dữ liệu:** cho phép nhiều người dùng truy xuất tới cơ sở dữ liệu chung
- **Chia sẻ thiết bị:** cho phép nhiều người dùng chia sẻ các thiết bị đắt tiền: máy in, thiết bị lưu trữ,..
- **Truyền thông:** giúp giao tiếp giữa người với người dễ dàng hơn, thí dụ bằng email
- **Linh hoạt hơn:** trải tải công việc qua những máy tính sẵn dùng một cách hiệu quả nhất về chi phí.

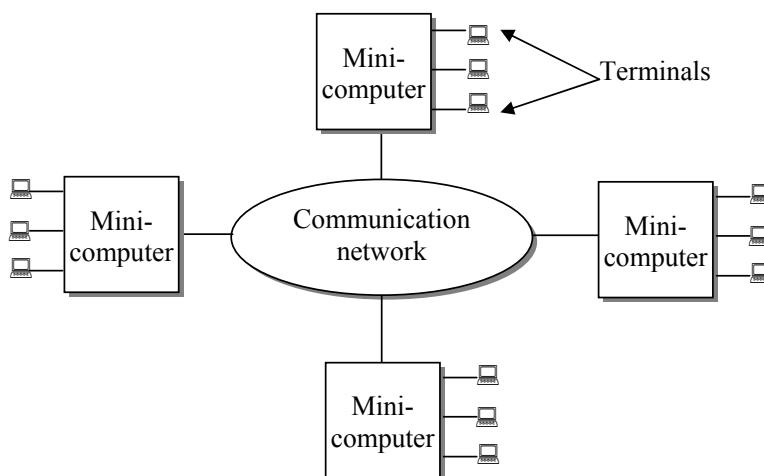
Bên cạnh đó, hệ thống phân tán cũng có một số bất lợi như:

- **Phần mềm:** ít những phần mềm hiện có vào thời điểm hiện tại cho hệ thống phân tán.
- **Mạng:** mạng có thể bão hòa hay gây ra các vấn đề khác.
- **Bảo mật:** truy xuất dễ dàng cũng áp dụng cho các dữ liệu bảo mật.

1.4 NHỮNG MÔ HÌNH CỦA HỆ THỐNG PHÂN TÁN

Qua nhiều giai đoạn phát triển, đến nay người ta thấy những hệ thống phân tán thường được thiết kế dựa theo một trong năm mô hình sau:

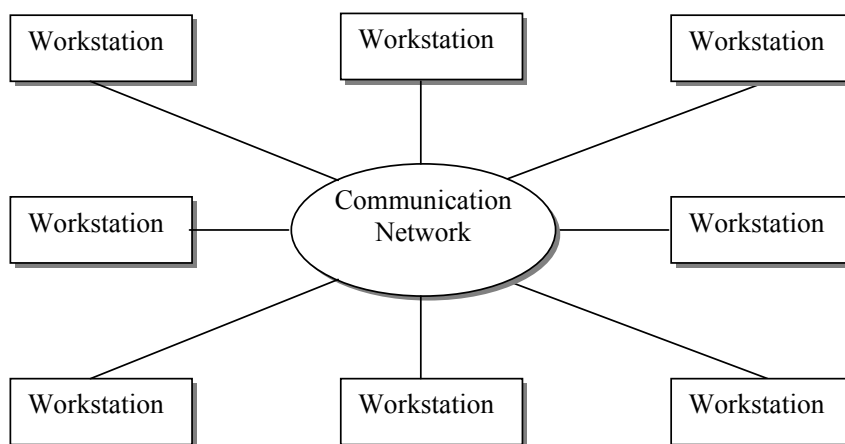
1.4.1 Mô hình máy tính tầm trung



Hình 1.1: Hệ thống phân tán dựa trên mô hình máy tính tầm trung

Mô hình máy tính tầm trung (minicomputer model) là sự mở rộng đơn giản của máy tính Mainframe. Nó bao gồm một số máy tầm trung (hoặc lớn hơn) liên kết với nhau bởi một hệ thống mạng, mỗi máy tính có nhiều người dùng đăng nhập (login) vào một cách đồng thời. Mạng máy tính cho phép một người dùng truy xuất đến một máy tính khác ở xa và sử dụng những tài nguyên trên máy ở xa đó. Trước đây, ARPANet là mạng thuộc dạng này.

1.4.2 Mô hình trạm làm việc

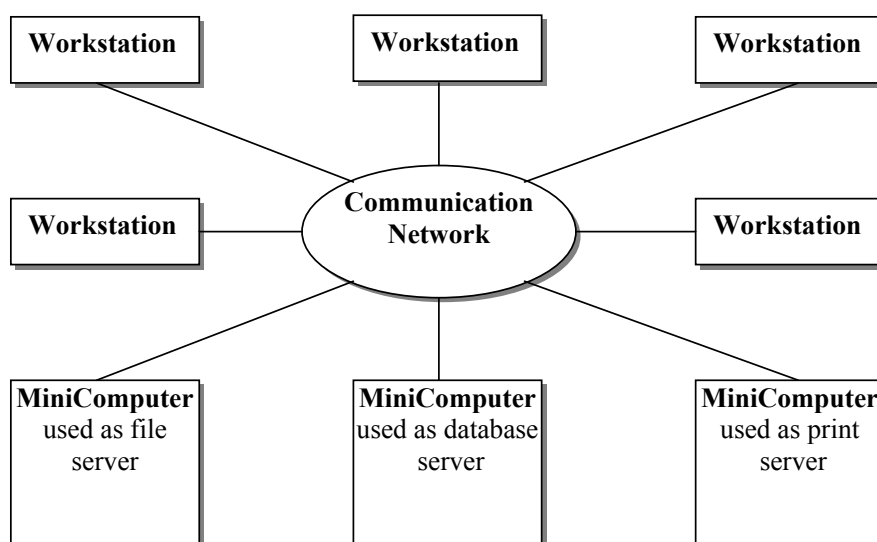


Hình 1.2: Hệ thống phân tán dựa trên mô hình trạm làm việc

Hệ thống phân tán dựa trên mô hình trạm làm việc (workstation model) bao gồm một số trạm làm việc được liên kết lại với nhau. Mỗi trạm trang bị đĩa cứng và phục vụ như một máy tính cho một người dùng. Trong mô hình này, có sự lãng phí CPU khi một số máy tính rảnh. Để tận dụng tất cả các CPU, người ta nối kết các trạm lại với nhau bằng đường mạng tốc độ cao. Nhờ đó khi một trạm quá tải nó có thể nhờ các trạm đang rảnh thực hiện hộ công việc của mình.

Hệ thống Sprite [Ousterhout et al. 1988] và hệ thống thử nghiệm được phát triển tại Xerox PARC [Shoch and Hupp 1982] là hai thí dụ về hệ thống phân tán dựa trên mô hình trạm làm việc.

1.4.3 Mô hình trạm làm việc - máy phục vụ



Hình 1.3: Hệ thống phân tán dựa trên mô hình trạm làm việc-máy phục vụ

Hệ thống phân tán dựa trên mô hình trạm làm việc - máy phục vụ (workstation-server model) bao gồm một vài máy tính tầm trung và một vài trạm làm việc (phần lớn các trạm này đều không có đĩa cứng) được nối kết lại với nhau thông qua một mạng truyền thông.

Hệ thống tập tin được dùng bởi những trạm làm việc không có đĩa cứng được cài đặt trên một máy tính tầm trung hoặc trên các trạm làm việc có đĩa cứng, nhưng thông thường các máy tính tầm trung được dùng cho mục đích này. Một số máy khác được dùng cho các dịch vụ khác như phục vụ máy in. Trong hệ thống này, những máy tính đặc biệt được dùng để chạy những tiến trình phục vụ gọi là những máy phục vụ hay server làm nhiệm vụ quản lý và cung cấp việc truy xuất đến những tài nguyên được chia sẻ.

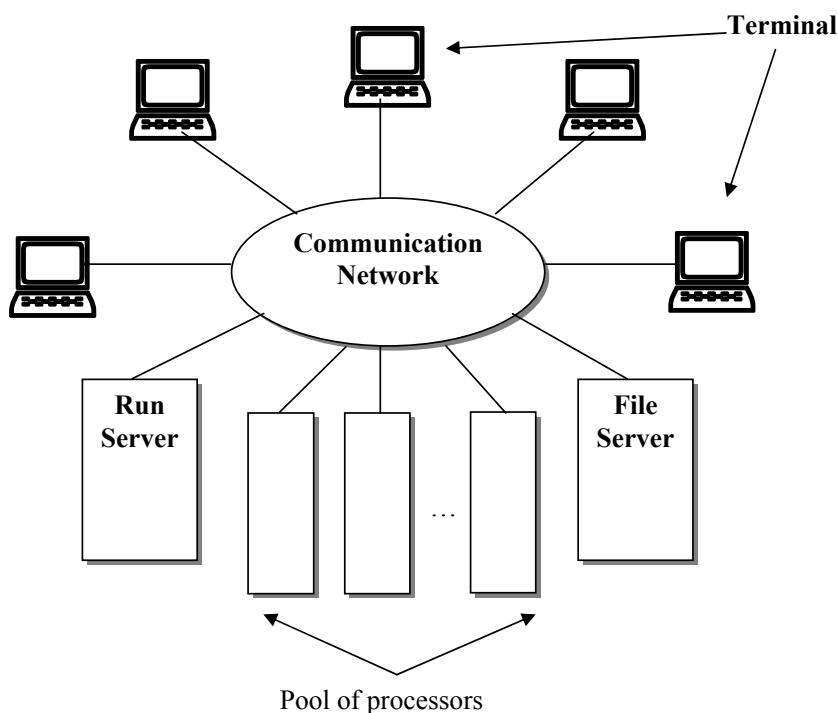
So với mô hình trạm làm việc, mô hình này có những ưu điểm sau:

1. Giá sẽ thấp hơn khi sử dụng một vài máy tính tầm trung được trang bị những ổ đĩa lớn so với việc dùng một số lượng lớn các trạm làm việc có đầy đủ ổ đĩa.
2. Dễ dàng trong việc lưu trữ và bảo trì.
3. Trong hệ thống trạm làm việc-máy phục vụ, người dùng có thể đăng nhập vào hệ thống từ bất kỳ trạm làm việc nào. Một điều khó thực hiện trong môi trường trạm làm việc.
4. Trong hệ thống trạm làm việc-máy phục vụ, giao thức yêu cầu-trả lời được sử dụng chủ yếu để truy xuất các dịch vụ của các máy phục vụ. Do đó, không giống như mô hình trạm làm việc, mô hình này không cần tiện ích di dời quá trình-một tiện ích rất khó cài đặt.
5. Một người dùng có thời gian đáp ứng được đảm bảo vì các trạm làm việc không được dùng để thực thi các quá trình ở xa. Tuy nhiên, mô hình này không tận dụng được khả năng xử lý của các trạm làm việc rảnh.

V-System [Cheriton 1988] là một thí dụ về hệ thống phân tán dựa trên mô hình trạm làm việc-máy phục vụ.

1.4.4 Mô hình vùng các bộ xử lý

Mô hình vùng các bộ xử lý (processors-pool model) được hình thành dựa trên quan điểm là: hầu hết những khoảng thời gian, người dùng không cần thiết đến năng lực tính toán của máy tính, nhưng đôi khi họ cần một năng lực tính toán rất lớn trong khoảng thời gian khác. Vì thế người ta nghĩ ra việc gộp chung tất cả các bộ xử lý lại, sau đó chia sẻ cho người dùng.



Hình 1.4: Hệ thống phân tán dựa trên mô hình vùng các bộ xử lý

Trong mô hình này, có một máy chủ đặc biệt gọi là *máy chủ chạy* (run server) quản lý và cấp phát bộ xử lý trong vùng cho những người dùng khác nhau theo yêu cầu. Khi người dùng có một yêu cầu tính toán, máy chủ chạy sẽ cấp tạm thời một số bộ xử lý cho công việc của người dùng đó. Thí dụ người dùng cần biên dịch một chương trình có n đoạn, mỗi đoạn có thể được biên dịch độc lập nhau. Khi đó n bộ xử lý trong vùng sẽ được cấp cho tác vụ này để biên dịch n đoạn song song. Khi tính toán đã xong những bộ xử lý được trả về cho vùng để sau đó có thể cấp phát cho những người dùng khác. Mô hình này tận dụng bộ xử lý tốt hơn mô hình khách hàng-người phục vụ. Tuy nhiên, trong thực tế mô hình này rất khó cài đặt.

Amoeba [Mullender et al 1990], Plan 9 [Pine et al. 1990] và Cambridge Distributed Computing System [Needham and Herbert 1982] là những thí dụ của các hệ thống phân tán dựa trên mô hình vùng các bộ xử lý.

1.4.5 Mô hình kết hợp

Trong bốn mô hình trên, mô hình trạm làm việc - máy phục vụ được sử dụng nhiều trong xây dựng những hệ thống phân tán. Điều này xảy ra bởi vì một số lớn những người dùng máy tính chỉ thực hiện những công việc đơn giản như soạn thảo văn bản, gửi thư điện tử hoặc thực thi những chương trình nhỏ. Tuy nhiên, trong những môi trường làm việc có số lượng người dùng lớn thì họ thường thực hiện những công việc đòi hỏi sự tính toán rất lớn. Mô hình vùng các bộ xử lý rất thích hợp cho trường hợp này. Mô hình kết hợp (Hybrid model) là một sự tổng hợp những ưu điểm của hai mô hình trên để xây dựng những hệ thống phân tán.

1.5 VAI TRÒ PHẦN MỀM TRONG HỆ THỐNG PHÂN TÁN

Việc hình thành các mô hình hệ thống phân tán khác nhau là kết quả của việc cài đặt một lớp phần mềm hệ thống lên trên hạ tầng phần cứng (hệ thống kết nối lỏng) của hệ thống phân tán. Cùng một hạ tầng phần cứng, nhưng cài đặt những lớp phần mềm hệ thống khác nhau sẽ tạo ra những hình ảnh hệ thống phân tán khác nhau đối với người sử dụng.

Chúng ta có thể kể ra 3 loại phần mềm hệ thống phổ biến giúp hình thành các hệ thống phân tán gồm:

- Hệ điều hành mạng
- Phần mềm trung gian
- Hệ điều hành phân tán

1.5.1 Hệ điều hành mạng

Hệ điều hành mạng (Network Operating System) là một lớp phần mềm được cài đặt bổ sung vào các hệ điều hành truyền thống của các máy tính trong một hệ thống kết nối lỏng nhằm giúp các máy tính trong hệ thống kết nối lỏng có thể trao đổi thông tin và chia sẻ tài nguyên qua lại với nhau. Giao thức, dịch vụ và giao diện là 3 thành phần quan trọng nhất trong một hệ điều hành mạng.

- **Giao thức:** qui định cách thức trao đổi thông tin giữa các máy tính trong một hệ thống kết nối lỏng.
- **Dịch vụ:** mô tả những gì mà một thành phần/một máy tính có thể cung cấp, phục vụ cho các thành phần/máy tính khác.
- **Giao diện:** mô tả cách thức để một thành phần/máy tính có thể sử dụng/truy cập được một dịch vụ do một thành phần/máy tính cung cấp.

Hệ điều hành mạng cung cấp những tiện ích, thí dụ như socket, để hình thành một hệ thống phân tán hoạt động theo mô hình Client-Server. Trong mô hình Client-Server, một ứng dụng được chia thành 2 phần có thể thực thi trên 2 máy tính khác nhau của hệ thống phân tán. Phần Client được dùng để tiếp nhận các yêu cầu của người dùng và chuyển yêu cầu đó đến server thông qua hệ thống