

PGS.TS. HÀ QUANG THỤY – TS. NGUYỄN TRÍ THÀNH

Giáo trình

HỆ ĐIỀU HÀNH UNIX - LINUX

NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM

Chương 1

GIỚI THIỆU CHUNG VỀ HỆ ĐIỀU HÀNH LINUX

1.1. Giới thiệu về hệ điều hành Linux

1.1.1. Sơ bộ về sự phát triển của hệ điều hành

Khái niệm hệ điều hành (Operating System)

- Hệ điều hành là một bộ chương trình (và dữ liệu - tham số hệ thống) đã được cài đặt sẵn (dưới dạng các file trên các thiết bị lưu trữ như: băng từ, đĩa từ,...) có chức năng điều phối (lựa chọn) và phân phối các công việc cho các bộ phận chức năng trong máy tính điện tử để sử dụng hiệu quả nhất hệ thống của máy tính.
- Theo góc độ nhìn nhận từ chức năng trên, thì hệ điều hành được quan niệm như là một *hệ thống quản trị tài nguyên*, nhằm thoả mãn ở mức cao nhất yêu cầu đa dạng của người dùng.
- Theo góc độ nhìn nhận từ chức năng đáp ứng nhu cầu người dùng, hệ điều hành được quan niệm như một *máy tính mở rộng* (hoặc *máy tính ảo*).

Thực chất trong các thế hệ máy tính đầu tiên (thế hệ 1, 2) là chưa có hệ điều hành, việc lựa chọn công việc thực hiện thông qua con người (thao tác viên - operator). Theo thời gian, năng lực của máy tính được nâng cao về tốc độ xử lý của CPU, về dung lượng bộ nhớ, về hệ thống thiết bị, về phần mềm hệ thống, về số lượng và khả năng của người sử dụng và vì vậy tạo ra điều kiện cần thiết cho việc xuất hiện các hệ điều hành ở dạng đơn giản.

Ban đầu xuất hiện hệ điều hành đơn chương trình, mỗi chương trình của người dùng được xếp hàng lần lượt được đưa vào bộ nhớ trong để thực hiện. Một chương trình sau khi được nạp từ dòng đợi vào bộ nhớ trong, sẽ được hệ thống (cùng toàn bộ tài nguyên) phục vụ từ khi bắt đầu chạy cho đến lúc kết thúc chương trình. Một chương trình được nạp vào bộ nhớ như vậy có thể được thực hiện với nhiều bộ dữ liệu. Chỉ khi chương trình này được kết thúc, thì mới nạp tiếp chương trình khác trong dòng đợi vào bộ nhớ trong.

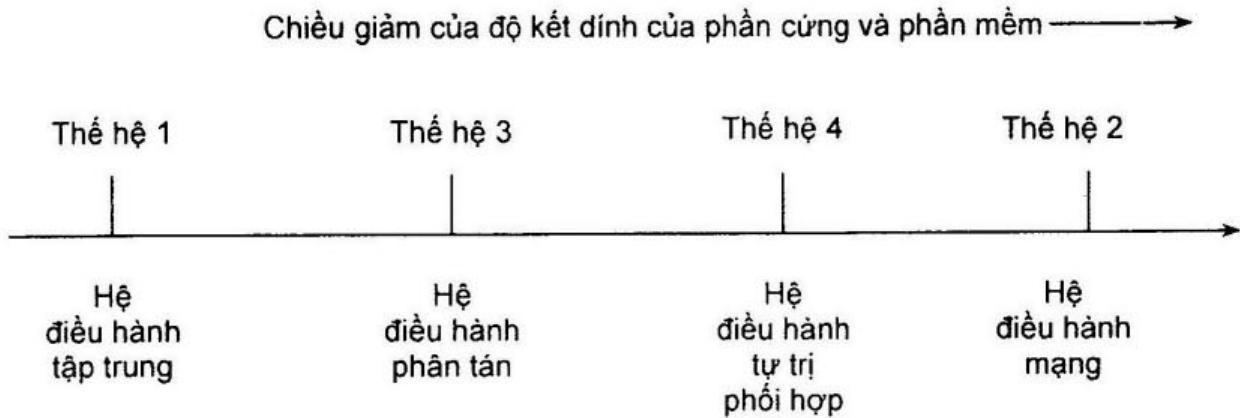
Tuy nhiên, việc nạp chương trình và dữ liệu vào bộ nhớ trong làm việc lại liên quan đến thiết bị vào - ra rất đa dạng. Trong thời kỳ đầu tiên, phổ biến là dữ liệu vào từ bìa đục lỗ và ra máy in. Cho đến một thời điểm, khi tốc độ làm việc của máy tính đã tăng nhanh và tốc độ nạp bìa và tốc độ in ra không theo kịp, đòi hỏi cần cải tiến để tăng hiệu quả. Một trong những cải tiến là hệ điều hành đơn chương trình hoạt động theo chế độ SPOOLING, mà trong chế độ đó, tất cả việc vào - ra đối với hệ điều hành là làm việc với đĩa cứng, còn vào - ra từ đĩa cứng với các vật mang tin khác được đảm bảo bằng những cơ chế riêng. Vì vậy tốc độ của toàn bộ hệ thống đã được tăng lên đáng kể. Chế độ SPOOLING còn được sử dụng trong những hệ điều hành đa chương trình xuất hiện sau này.

Với sự tiến bộ nhanh chóng của hệ thống máy tính làm cho dung lượng bộ nhớ tăng lên đáng kể (vượt xa dung lượng trung bình của các chương trình người dùng) và tốc độ xử lý cũng tăng nhanh không kém, thời gian chương trình chờ đợi đáp ứng nhu cầu vào - ra là quá lớn so với thời gian chạy chương trình để xử lý thông tin, nên chế độ đơn người dùng không còn thích hợp và chế độ hoạt động đa chương trình xuất hiện. Chế độ đa chương trình được đặt ra theo các hướng (chế độ mẻ, chế độ đa người dùng) được giới thiệu như dưới đây:

- Hướng độc lập người dùng (được gọi là chế độ mẻ - batch) có mục đích nhằm tăng số lượng chương trình được giải quyết trong một khoảng thời gian (như các hệ điều hành MFT, MVT). Trong các hệ điều hành này, phân phối bộ nhớ cho chương trình người dùng là liên tục; sự chuyển giao CPU giữa các chương trình người dùng xuất hiện khi chương trình đang thực hiện hướng tới khu vực ngoại vi.
- Hướng đa người dùng (multi-users, còn được gọi là hoạt động theo chế độ phân chia thời gian), có mục đích cho phép người dùng làm việc trực tiếp với máy tính thông qua trạm cuối. Trong hệ thống này, xuất hiện việc tổ chức bộ nhớ ảo, dùng để phân phối bộ nhớ cho chương trình là gián đoạn; sự chuyển giao CPU giữa các chương trình người dùng xuất hiện khi hết lượng từ thời gian.
- Hướng điều khiển tự động hóa được thể hiện thông qua các hệ điều hành thời gian thực (Real Time).

Hiện nay, để tăng cường khả năng tính toán, giải quyết các nhiệm vụ đặt ra, nên máy tính không còn được sử dụng một cách riêng lẻ, mà nhiều máy tính được kết nối thành một hệ thống chung; trong đó mỗi máy tính

đảm nhận một chức năng bộ phận. Theo quá trình đó, các loại hệ điều hành mạng, hệ điều hành phân tán và hệ điều hành tự trị phối hợp xuất hiện (hình 1.1). So với hệ điều hành tập trung, sự kết nối giữa phần cứng và phần mềm trong hệ thống máy tính trở nên mềm dẻo hơn, trong một số trường hợp (như hệ điều hành mạng) sự kết nối đó là chưa chặt chẽ.



Hình 1.1. Phân bố các hệ điều hành

Hệ điều hành mạng cho phép liên kết nhiều máy tính lại theo một cách thức không thực sự chặt chẽ. Tuy các máy tính trong mạng có thể trao đổi thông tin với nhau, truy nhập từ xa, song chưa có một hệ thống tổng thể điều phối các tài nguyên trong toàn bộ hệ thống mạng. Trong nhiều trường hợp, tài nguyên phần cứng và phần mềm của các máy tính thành phần hoàn toàn không có tác dụng đối với hoạt động xử lý thông tin của toàn bộ hệ thống.

Hệ điều hành phân tán mới thực sự là một hệ điều hành quản lý tài nguyên máy tính trên phạm vi rộng. Các máy tính được kết nối trong hệ điều hành phân tán một cách chặt chẽ, hệ thống tài nguyên của mỗi máy tính đóng góp thực sự vào hệ thống tài nguyên chung thống nhất và tham gia vào việc giải quyết mỗi bài toán điều phối tiến trình, điều phối bộ nhớ, điều phối vào - ra,... Hệ điều hành phân tán, về logic là một hệ thống thống nhất, song về địa lý lại được "phân bố" chạy trên nhiều máy tính ở các vị trí khác nhau.

Hệ điều hành tự trị phối hợp cho một cách thức linh hoạt hơn so với hệ điều hành phân tán. Các máy tính thành viên vừa được phép tham gia kết nối vào toàn bộ hệ thống, lại vừa được phép chạy một cách độc lập. Khi tham gia vào hệ thống, tài nguyên của máy tính thành viên được toàn bộ hệ thống sử dụng, còn khi máy tính thành viên chạy độc lập, thì việc sử dụng tài nguyên là độc lập.

UNIX (và Linux) là *hệ điều hành đa người dùng* (multi-users). Hệ điều hành đa người dùng thuộc vào loại hệ điều hành đa chương trình định hướng "thân thiện với người dùng". Tại cùng một thời điểm, có nhiều người dùng cùng sử dụng máy tính và đối với mỗi người dùng như vậy đều có cảm giác như được sử dụng máy tính một cách "độc lập" vì họ được trực tiếp liên kết với chương trình của mình đang thực hiện trong máy tính. Điều này tương ứng với một chức năng của hệ điều hành là "hệ điều hành như một máy tính ảo" theo góc độ của người sử dụng.

Như vậy, trong máy tính đồng thời xuất hiện nhiều chương trình người dùng, các chương trình này chia nhau sử dụng các tài nguyên của hệ thống, trong đó có các tài nguyên quan trọng nhất là CPU, bộ nhớ trong và hệ thống file.

Mỗi người dùng hướng đến tài nguyên chung qua trạm cuối (terminal) của mình (các trạm cuối này được đặt tên và được hệ thống quản lý). Trong trường hợp đơn giản, trạm cuối chỉ bao gồm hai thiết bị là màn hình (để hiện thông tin cho người dùng) và bàn phím (để người dùng đưa yêu cầu đối với hệ điều hành). Trong nhiều trường hợp khác, có thể sử dụng một máy tính cá nhân đóng vai trò của một trạm cuối và như vậy mỗi người dùng vừa được phép sử dụng tài nguyên riêng vừa được phép sử dụng tài nguyên chung.

Điển hình nhất trong công việc phân chia tài nguyên của hệ thống máy tính trong hệ điều hành đa người dùng là việc phân chia CPU theo một chu kỳ thời gian, mà mỗi người dùng được sử dụng CPU trong một khoảng thời gian nhất định (được gọi là *lượng tử thời gian*) và sau khi mỗi người đã được phân chia CPU thì lại chuyển đến lượt phân chia tiếp theo. Như vậy, phân chia thời gian (Time shared system) là cách thức của hệ đa người dùng khi điều phối CPU.

Là một hệ điều hành đa người dùng, UNIX đã khá phổ biến trong lĩnh vực Công nghệ thông tin (CNTT), có thể được sử dụng từ máy vi tính cho tới máy tính mainframe. Nó đặc biệt thích hợp đối với các hệ Client-Server và mạng máy tính diện rộng.

1.1.2. Xuất xứ, sự phát triển và một số đặc trưng của hệ điều hành UNIX

Năm 1965, Viện công nghệ Massachusetts (MIT – Massachusetts Institute of Technology) và Phòng thí nghiệm Bell của hãng AT&T thực

hiện dự án xây dựng một hệ điều hành có tên gọi là Multics (MULTiplexed Information and Computing Service) với mục tiêu tạo lập được một hệ điều hành phủ trên vùng lãnh thổ rộng (hoạt động trên tập các máy tính được kết nối), đa người dùng, có năng lực cao về tính toán và lưu trữ. Dự án nói trên thành công ở mức độ hết sức khiêm tốn và người ta đã biết đến một số nhược điểm khó khắc phục của Multics.

Năm 1969, Ken Thompson, một chuyên viên tại Phòng thí nghiệm Bell, người đã tham gia dự án Multics, cùng Dennis Ritchie viết lại hệ điều hành đa bài toán trên máy PDP-7 với tên là UNICS (UNiplexed Information and Computing Service) từ một câu gọi đùa của một đồng nghiệp. Trong hệ điều hành UNICS, một số khởi thảo đầu tiên về hệ thống file đã được Ken Thompson và Dennis Ritchie thực hiện. Đến năm 1970, hệ điều hành được viết trên ngôn ngữ Assembler cho máy PDP-11/20 và mang tên là UNIX.

Năm 1973, Ritchie và Thompson viết lại nhân của hệ điều hành UNIX trên ngôn ngữ C và hệ điều hành đã trở nên dễ dàng cài đặt trên các loại máy tính khác nhau; tính chất như thế được gọi là tính khả chuyển (portable) của UNIX. Trước đó, khoảng năm 1971, hệ điều hành được thể hiện trên ngôn ngữ B (mà dựa trên ngôn ngữ B, Ritchie đã phát triển thành ngôn ngữ C).

Hãng AT&T phổ biến chương trình nguồn UNIX tới các trường đại học, các công ty thương mại và chính phủ với giá không đáng kể.

Năm 1982, hệ thống UNIX-3 là bản UNIX thương mại đầu tiên của AT&T.

Năm 1983, AT&T giới thiệu Hệ thống UNIX-4 phiên bản thứ nhất, trong đó đã có trình soạn thảo vi, thư viện quản lý màn hình được xây dựng, phát triển bởi trường Đại học Tổng hợp California, Berkley.

Giai đoạn 1985-1987, UNIX-5 phiên bản 2 và 3 tương ứng được đưa ra vào các năm 1985 và 1987. Trong giai đoạn này, có khoảng 100000 bản UNIX đã được phổ biến trên thế giới (cài đặt từ máy vi tính đến các hệ thống lớn).

Đầu thập kỷ 90 của thế kỷ XX, UNIX-5 phiên bản 4 được đưa ra như là một chuẩn của UNIX. Đây là sự kết hợp của các phiên bản sau:

- AT&T UNIX-5 phiên bản 3.
- Berkley Software Distribution (BSD).
- XENIX của MicroSoft.
- SUN OS.

Trong thời gian gần đây (khoảng năm 1997) một số phiên bản mới của UNIX được giới thiệu và phổ biến trên Internet. Có thể tìm thấy các nội dung liên quan tại địa chỉ website <http://problem.rice.edu>.

Các nhóm nhà cung cấp khác nhau về UNIX đang hoạt động trong thời gian hiện nay được kể đến là:

- Unix International (viết tắt là UI): UI là một tổ chức gồm các nhà cung cấp thực hiện việc chuyển nhượng hệ thống UNIX-5 và cung cấp bản AT&T theo các nhu cầu và thông báo phát hành mới, chẳng hạn như điều chỉnh bản quyền. Giao diện đồ họa người dùng là Open Look.
- Open Software Foundation (OSF): OSF được hỗ trợ bởi IBM, DEC, HP... theo hướng phát triển một phiên bản của UNIX nhằm cạnh tranh với hệ thống UNIX-5 phiên bản 4. Phiên bản này có tên là OSF/1 với giao diện đồ họa người dùng được gọi là MOTIF.
- Free Software Foundation: là một tổ chức chủ trương phát hành một dòng của UNIX mã nguồn mở, miễn phí, đó là hệ điều hành Linux.

Bảng 1.1 liệt kê một số dòng UNIX khá phổ biến (thường thấy có chữ X ở cuối tên gọi của Hệ điều hành).

Bảng 1.1. Một số hệ UNIX phổ biến

<i>Tên hệ</i>	<i>Nhà cung cấp</i>	<i>Nền phát triển</i>
AIX	International Business Machines	AT&T System V
A/UX	Apple Computer	AT&T System V
Dynix	Sequent	BSD (Berkeley Software Distribution)
HP-UX	Hewlett-Packard	BSD
Irix	Silicon Graphics	AT&T System V
Linux	Free Software Foundation	
NextStep	Next	BSD
OSF/1	Digital Equipment Corporation	BSD
SCO UNIX	Santa Cruz Operation	AT&T System V
Solaris	Sun Microsystems	AT&T System V
SunOS	Sun Microsystems	BSD UNIX
Ultrix	Digital Equipment Corporation	BSD UNIX
Unicos	Cray	AT&T System V
UnixWare	Novell	AT&T System V
XENIX	Microsoft	AT&T System III-MS