

Introducción

Un **sistema operativo** es un programa que administra el hardware de una computadora. También proporciona las bases para los programas de aplicación y actúa como un intermediario entre el usuario y el hardware de la computadora. Un aspecto sorprendente de los sistemas operativos es la gran variedad de formas en que llevan a cabo estas tareas. Los sistemas operativos para *main-frame* están diseñados principalmente para optimizar el uso del hardware. Los sistemas operativos de las computadoras personales (PC) soportan desde complejos juegos hasta aplicaciones de negocios. Los sistemas operativos para las computadoras de mano están diseñados para proporcionar un entorno en el que el usuario pueda interactuar fácilmente con la computadora para ejecutar programas. Por tanto, algunos sistemas operativos se diseñan para ser *prácticos*, otros para ser *eficientes* y otros para ser ambas cosas.

Antes de poder explorar los detalles de funcionamiento de una computadora, necesitamos saber algo acerca de la estructura del sistema. Comenzaremos la exposición con las funciones básicas del arranque, E/S y almacenamiento. También vamos a describir la arquitectura básica de una computadora que hace posible escribir un sistema operativo funcional.

Dado que un sistema operativo es un software grande y complejo, debe crearse pieza por pieza. Cada una de estas piezas deberá ser una parte bien diseñada del sistema, con entradas, salida y funciones cuidadosamente definidas. En este capítulo proporcionamos una introducción general a los principales componentes de un sistema operativo.

OBJETIVOS DEL CAPÍTULO

- Proporcionar una visión general de los principales componentes de los sistemas operativos.
- Proporcionar una panorámica sobre la organización básica de un sistema informático.

1.1 ¿Qué hace un sistema operativo?

Comenzamos nuestra exposición fijándonos en el papel del sistema operativo en un sistema informático global. Un sistema informático puede dividirse a grandes rasgos en cuatro componentes: el *hardware*, el *sistema operativo*, los *programas de aplicación* y los *usuarios* (Figura 1.1).

El **hardware**, la **unidad central de procesamiento**, UCP o CPU (central processing unit), la **memoria** y los **dispositivos de E/S** (entrada/salida), proporciona los recursos básicos de cómputo al sistema. Los **programas de aplicación**, como son los procesadores de texto, las hojas de cálculo, los compiladores y los exploradores web, definen las formas en que estos recursos se emplean para resolver los problemas informáticos de los usuarios. El sistema operativo controla y coordina el uso del hardware entre los diversos programas de aplicación por parte de los distintos usuarios.

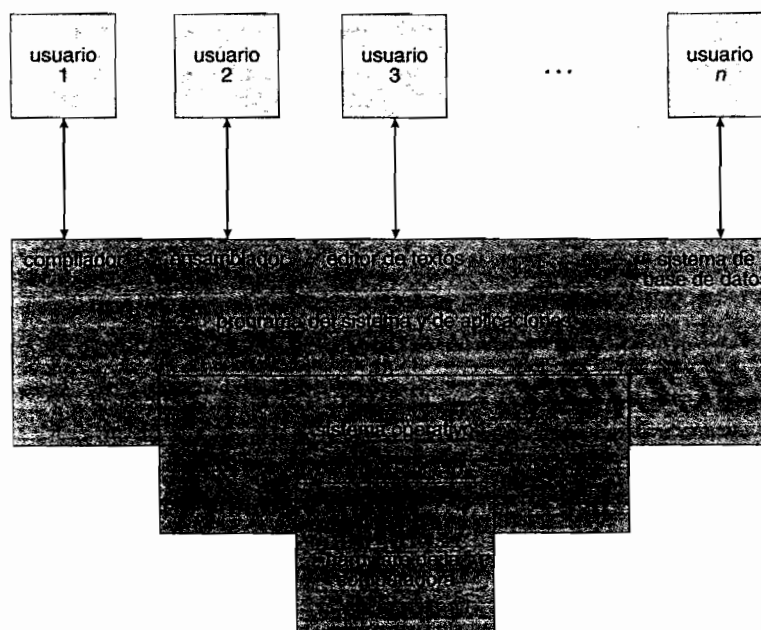


Figura 1.1 Esquema de los componentes de un sistema informático.

También podemos ver un sistema informático como hardware, software y datos. El sistema operativo proporciona los medios para hacer un uso adecuado de estos recursos durante el funcionamiento del sistema informático. Un sistema operativo es similar a un gobierno. Como un gobierno, no realiza ninguna función útil por sí mismo: simplemente proporciona un entorno en el que otros programas pueden llevar a cabo un trabajo útil.

Para comprender mejor el papel de un sistema operativo, a continuación vamos a abordar los sistemas operativos desde dos puntos de vista: el del usuario y el del sistema.

1.1.1 Punto de vista del usuario

La visión del usuario de la computadora varía de acuerdo con la interfaz que utilice. La mayoría de los usuarios que se sientan frente a un PC disponen de un monitor, un teclado, un ratón y una unidad de sistema. Un sistema así se diseña para que un usuario monopolice sus recursos. El objetivo es maximizar el trabajo (o el juego) que el usuario realice. En este caso, el sistema operativo se diseña principalmente para que sea de **fácil uso**, prestando cierta atención al rendimiento y ninguna a la **utilización de recursos** (el modo en que se comparten los recursos hardware y software). Por supuesto, el rendimiento es importante para el usuario, pero más que la utilización de recursos, estos sistemas se optimizan para el uso del mismo por un solo usuario.

En otros casos, un usuario se sienta frente a un terminal conectado a un **mainframe** o una **microcomputadora**. Otros usuarios acceden simultáneamente a través de otros terminales. Estos usuarios comparten recursos y pueden intercambiar información. En tales casos, el sistema operativo se diseña para maximizar la utilización de recursos, asegurar que todo el tiempo de CPU, memoria y E/S disponibles se usen de forma eficiente y que todo usuario disponga sólo de la parte equitativa que le corresponde.

En otros casos, los usuarios usan **estaciones de trabajo** conectadas a redes de otras estaciones de trabajo y **servidores**. Estos usuarios tienen recursos dedicados a su disposición, pero también tienen recursos compartidos como la red y los servidores (servidores de archivos, de cálculo y de impresión). Por tanto, su sistema operativo está diseñado para llegar a un compromiso entre la usabilidad individual y la utilización de recursos.

Recientemente, se han puesto de moda una gran variedad de computadoras de mano. La mayor parte de estos dispositivos son unidades autónomas para usuarios individuales. Algunas se conectan a redes directamente por cable o, más a menudo, a través de redes y modems inalám-

bricos. Debido a las limitaciones de alimentación, velocidad e interfaz, llevan a cabo relativamente pocas operaciones remotas. Sus sistemas operativos están diseñados principalmente en función de la usabilidad individual, aunque el rendimiento, medido según la duración de la batería, es también importante.

Algunas computadoras tienen poca o ninguna interacción con el usuario. Por ejemplo, las computadoras incorporadas en los electrodomésticos y en los automóviles pueden disponer de teclados numéricos e indicadores luminosos que se encienden y apagan para mostrar el estado, pero tanto estos equipos como sus sistemas operativos están diseñados fundamentalmente para funcionar sin intervención del usuario.

1.1.2 Vista del sistema

Desde el punto de vista de la computadora, el sistema operativo es el programa más íntimamente relacionado con el hardware. En este contexto, podemos ver un sistema operativo como un **asignador de recursos**. Un sistema informático tiene muchos recursos que pueden ser necesarios para solucionar un problema: tiempo de CPU, espacio de memoria, espacio de almacenamiento de archivos, dispositivos de E/S, etc. El sistema operativo actúa como el administrador de estos recursos. Al enfrentarse a numerosas y posiblemente conflictivas solicitudes de recursos, el sistema operativo debe decidir cómo asignarlos a programas y usuarios específicos, de modo que la computadora pueda operar de forma eficiente y equitativa. Como hemos visto, la asignación de recursos es especialmente importante cuando muchos usuarios acceden al mismo *mainframe* o minicomputadora.

Un punto de vista ligeramente diferente de un sistema operativo hace hincapié en la necesidad de controlar los distintos dispositivos de E/S y programas de usuario. Un sistema operativo es un programa de control. Como **programa de control**, gestiona la ejecución de los programas de usuario para evitar errores y mejorar el uso de la computadora. Tiene que ver especialmente con el funcionamiento y control de los dispositivos de E/S.

1.1.3 Definición de sistemas operativos

Hemos visto el papel de los sistemas operativos desde los puntos de vista del usuario y del sistema. Pero, ¿cómo podemos definir qué es un sistema operativo? En general, no disponemos de ninguna definición de sistema operativo que sea completamente adecuada. Los sistemas operativos existen porque ofrecen una forma razonable de resolver el problema de crear un sistema informático utilizable. El objetivo fundamental de las computadoras es ejecutar programas de usuario y resolver los problemas del mismo fácilmente. Con este objetivo se construye el hardware de la computadora. Debido a que el hardware por sí solo no es fácil de utilizar, se desarrollaron programas de aplicación. Estos programas requieren ciertas operaciones comunes, tales como las que controlan los dispositivos de E/S. Las operaciones habituales de control y asignación de recursos se incorporan en una misma pieza del software: el sistema operativo.

Además, no hay ninguna definición universalmente aceptada sobre qué forma parte de un sistema operativo. Desde un punto de vista simple, incluye todo lo que un distribuidor suministra cuando se pide "el sistema operativo". Sin embargo, las características incluidas varían enormemente de un sistema a otro. Algunos sistemas ocupan menos de 1 megabyte de espacio y no proporcionan ni un editor a pantalla completa, mientras que otros necesitan gigabytes de espacio y están completamente basados en sistemas gráficos de ventanas. (Un kilobyte, Kb, es 1024 bytes; un megabyte, MB, es 1024^2 bytes y un gigabyte, GB, es 1024^3 bytes. Los fabricantes de computadoras a menudo redondean estas cifras y dicen que 1 megabyte es un millón de bytes y que un gigabyte es mil millones de bytes.) Una definición más común es que un sistema operativo es aquel programa que se ejecuta continuamente en la computadora (usualmente denominado *kernel*), siendo todo lo demás programas del sistema y programas de aplicación. Esta definición es la que generalmente seguiremos.

La cuestión acerca de qué constituye un sistema operativo está adquiriendo una importancia creciente. En 1998, el Departamento de Justicia de Estados Unidos entabló un pleito contra

Microsoft, aduciendo en esencia que Microsoft incluía demasiada funcionalidad en su sistema operativo, impidiendo a los vendedores de aplicaciones competir. Por ejemplo, un explorador web era una parte esencial del sistema operativo. Como resultado, Microsoft fue declarado culpable de usar su monopolio en los sistemas operativos para limitar la competencia.

1.2 Organización de una computadora

Antes de poder entender cómo funciona una computadora, necesitamos unos conocimientos generales sobre su estructura. En esta sección, veremos varias partes de esa estructura para completar nuestros conocimientos. La sección se ocupa principalmente de la organización de una computadora, así que puede hojearla o saltársela si ya está familiarizado con estos conceptos.

1.2.1 Funcionamiento de una computadora

Una computadora moderna de propósito general consta de una o más CPU y de una serie de controladoras de dispositivo conectadas a través de un bus común que proporciona acceso a la memoria compartida (Figura 1.2). Cada controladora de dispositivo se encarga de un tipo específico de dispositivo, por ejemplo, unidades de disco, dispositivos de audio y pantallas de vídeo. La CPU y las controladoras de dispositivos pueden funcionar de forma concurrente, compitiendo por los ciclos de memoria. Para asegurar el acceso de forma ordenada a la memoria compartida, se proporciona una controladora de memoria cuya función es sincronizar el acceso a la misma.

Para que una computadora comience a funcionar, por ejemplo cuando se enciende o se reinicia, es necesario que tenga un programa de inicio que ejecutar. Este programa de inicio, o **programa de arranque**, suele ser simple. Normalmente, se almacena en la memoria ROM (read only memory, memoria de sólo lectura) o en una memoria EEPROM (electrically erasable programmable read-only memory, memoria de sólo lectura programable y eléctricamente borrable), y se conoce con el término general **firmware**, dentro del hardware de la computadora. Se inicializan todos los aspectos del sistema, desde los registros de la CPU hasta las controladoras de dispositivos y el contenido de la memoria. El programa de arranque debe saber cómo cargar el sistema operativo e iniciar la ejecución de dicho sistema. Para conseguir este objetivo, el programa de arranque debe localizar y cargar en memoria el *kernel* (núcleo) del sistema operativo. Después, el sistema operativo comienza ejecutando el primer proceso, como por ejemplo “init”, y espera a que se produzca algún suceso.

La ocurrencia de un suceso normalmente se indica mediante una **interrupción** bien hardware o bien software. El hardware puede activar una interrupción en cualquier instante enviando una señal a la CPU, normalmente a través del bus del sistema. El software puede activar una interrupción ejecutando una operación especial denominada **llamada del sistema** (o también llamada de **monitor**).

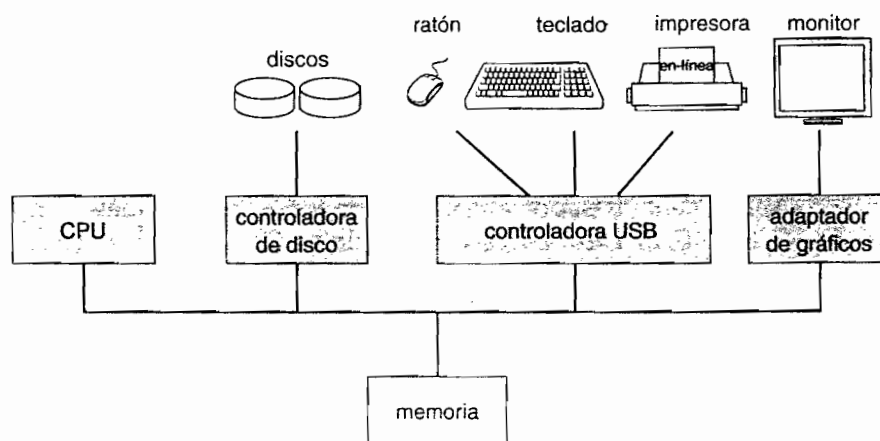


Figura 1.2 Una computadora moderna.