TIPOS DE SISTEMAS OPERATIVO

Describiremos las características que clasifican a los sistemas operativos, básicamente se cubrirán tres clasificaciones: sistemas operativos por su estructura (visión interna), sistemas operativos por los servicios que ofrecen y finalmente sistemas operativos por la forma en que ofrecen sus servicios (visión externa).

Sistemas Operativos por su Estructura

Se deben observar dos tipos de requisitos cuando se construye un sistema operativo, los cuales son:

Requisitos de usuario: Sistema fácil de usar y de aprender, seguro, rápido y adecuado al uso al que se le quiere destinar.

Requisitos del software: Donde se engloban aspectos como el mantenimiento, forma de operación, restricciones de uso, eficiencia, tolerancia frente a los errores y flexibilidad.

A continuación se describen las distintas estructuras que presentan los actuales sistemas operativos para satisfacer las necesidades que de ellos se quieren obtener.

Estructura monolítica.

Es la estructura de los primeros sistemas operativos constituidos fundamentalmente por un solo programa compuesto de un conjunto de rutinas entrelazadas de tal forma que cada una puede llamar a cualquier otra (Ver Figura). Las características fundamentales de este tipo de estructura son:

Construcción del programa final a base de módulos compilados separadamente que se unen a través de los enlaces (links).

Buena definición de parámetros de enlace entre las distintas rutinas existentes, que puede provocar mucho acoplamiento.

Carecen de protecciones y privilegios al entrar a rutinas que manejan diferentes aspectos de los recursos de la computadora, como memoria, disco, etc.

Generalmente están hechos a medida, por lo que son eficientes y rápidos en su ejecución y gestión, pero por lo mismo carecen de flexibilidad para soportar diferentes ambientes de trabajo o tipos de aplicaciones.

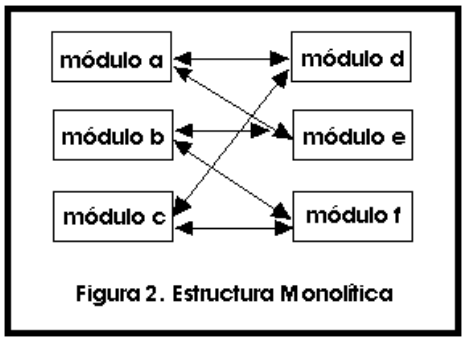

Figura 2. Estructura Monolítica

Estructura jerárquica.

A medida que fueron creciendo las necesidades de los usuarios y se perfeccionaron los sistemas, se hizo necesaria una mayor organización del software, del sistema operativo, donde una parte del sistema contenía subpartes y esto organizado en forma de niveles.

Se dividió el sistema operativo en pequeñas partes, de tal forma que cada una de ellas estuviera perfectamente definida y con un claro interface con el resto de elementos.

Se constituyó una estructura jerárquica o de niveles en los sistemas operativos, el primero de los cuales fue denominado THE (Technische Hogeschool, Eindhoven), de Dijkstra, que se utilizó con fines didácticos. Se puede pensar también en estos sistemas como si fueran `multicapa'. Multics y Unix caen en esa categoría.

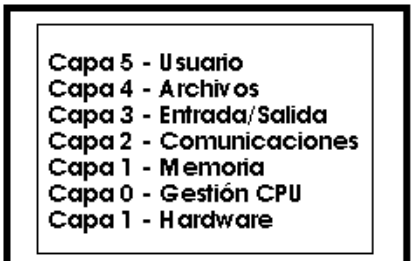

Figura 3. Sistema jerárquico THE

En la estructura anterior se basan prácticamente la mayoría de los sistemas operativos actuales. Otra forma de ver este tipo de sistema es la denominada de anillos concéntricos o "rings".

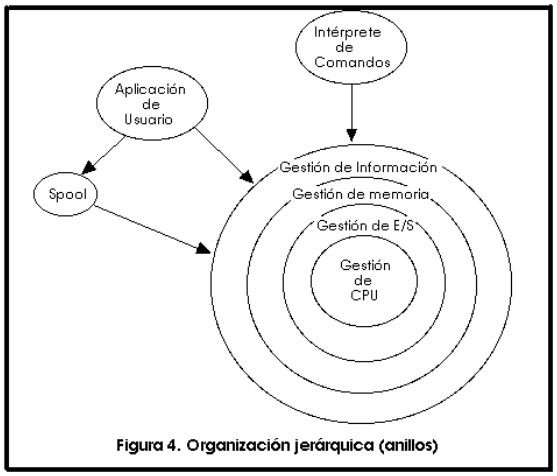

Figura 4. Organización jerárquica (anillos)

En el sistema de anillos, cada uno tiene una apertura, conocida como puerta o trampa (trap), por donde pueden entrar las llamadas de las capas inferiores. De esta forma, las zonas más internas del sistema operativo o núcleo del sistema estarán más protegidas de accesos indeseados desde las capas más externas. Las capas más internas serán, por tanto, más privilegiadas que las externas.

Máquina Virtual.

Se trata de un tipo de sistemas operativos que presentan una interface a cada proceso, mostrando una máquina que parece idéntica a la máquina real subyacente. Estos sistemas operativos separan dos conceptos que suelen estar unidos en el resto de sistemas: la multiprogramación y la máquina extendida. El objetivo de los sistemas operativos de máquina virtual es el de integrar distintos sistemas operativos dando la sensación de ser varias máquinas diferentes.

El núcleo de estos sistemas operativos se denomina monitor virtual y tiene como misión llevar a cabo la multiprogramación, presentando a los niveles superiores tantas máquinas virtuales como se soliciten.

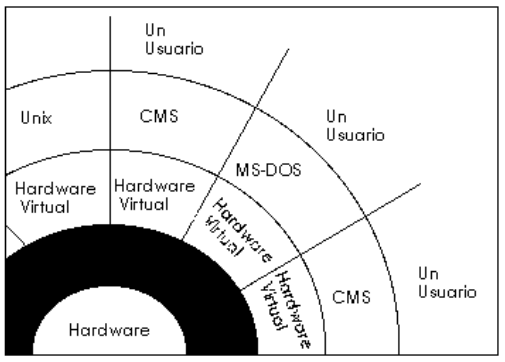

Figura 5. Máquina Virtual

Estas máquinas virtuales no son máquinas extendidas, sino una réplica de la máquina real, de manera que en cada una de ellas se pueda ejecutar un sistema operativo diferente, que será el que ofrezca la máquina extendida al usuario.

Cliente-servidor (Microkernel)

El tipo más reciente de sistemas operativos es el denominado Cliente-servidor, que puede ser ejecutado en la mayoría de las computadoras, ya sean grandes o pequeñas.

Este sistema sirve para toda clase de aplicaciones por tanto, es de propósito general y cumple con las mismas actividades que los sistemas operativos convencionales.

El núcleo tiene como misión establecer la comunicación entre los clientes y los servidores. Los procesos pueden ser tanto servidores como clientes. Por ejemplo, un programa de aplicación normal es un cliente que llama al servidor correspondiente para acceder a un archivo o realizar una operación de entrada/salida sobre un dispositivo concreto. A su vez, un proceso cliente puede actuar como servidor para otro". Este paradigma ofrece gran flexibilidad en cuanto a los servicios posibles en el sistema final, ya que el núcleo provee solamente funciones muy básicas de memoria, entrada/salida, archivos y procesos, dejando a los servidores proveer la mayoría que el usuario final o programador puede usar. Estos servidores deben tener mecanismos de seguridad y protección que, a su vez, serán filtrados por el núcleo que controla el hardware. Actualmente se está trabajando en una versión de UNIX que contempla en su diseño este paradigma.

Sistemas Operativos por Servicios

Esta clasificación es la más comúnmente usada y conocida desde el punto de vista del usuario final. Esta clasificación se comprende fácilmente con el cuadro sinóptico que a continuación se muestra.

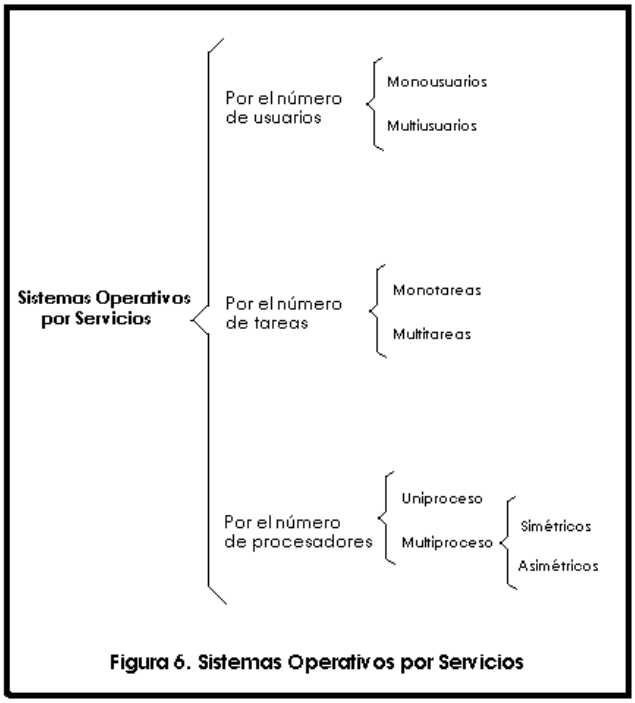

Figura 6. Sistemas Operativos por Servicios

Monousuarios

Los sistemas operativos monousuarios son aquéllos que soportan a un usuario a la vez, sin importar el número de procesadores que tenga la computadora o el número de procesos o tareas que el usuario pueda ejecutar en un mismo instante de tiempo. Las computadoras personales típicamente se han clasificado en este renglón.

Multiusuarios

Los sistemas operativos multiusuarios son capaces de dar servicio a más de un usuario a la vez, ya sea por medio de varias terminales conectadas a la computadora o por medio de sesiones remotas en una red de comunicaciones. No importa el número de procesadores en la máquina ni el número de procesos que cada usuario puede ejecutar simultáneamente.

Monotareas

Los sistemas monotarea son aquellos que sólo permiten una tarea a la vez por usuario.

Puede darse el caso de un sistema multiusuario y monotarea, en el cual se admiten varios usuarios al mismo tiempo pero cada uno de ellos puede estar haciendo solo una tarea a la vez.

Multitareas

Un sistema operativo multitarea es aquél que le permite al usuario estar realizando varias labores al mismo tiempo. Por ejemplo, puede estar editando el código fuente de un programa durante su depuración mientras compila otro programa, a la vez que está recibiendo correo electrónico en un proceso en background. Es común encontrar en ellos interfaces gráficas orientadas al uso de menús y el ratón, lo cual permite un rápido intercambio entre las tareas para el usuario, mejorando su productividad.

Uniproceso

Un sistema operativo uniproceso es aquél que es capaz de manejar solamente un procesador de la computadora, de manera que si la computadora tuviese más de uno le sería inútil. El ejemplo más típico de este tipo de sistemas es el DOS y MacOS.

Multiproceso

Un sistema operativo multiproceso se refiere al número de procesadores del sistema, que es más de uno y éste es capaz de usarlos todos para distribuir su carga de trabajo. Generalmente estos sistemas trabajan de dos formas: simétrica o asimétricamente. Cuando se trabaja de manera asimétrica, el sistema operativo selecciona a uno de los procesadores el cual jugará el papel de procesador maestro y servirá como pivote para distribuir la carga a los demás procesadores, que reciben el nombre de esclavos. Cuando se trabaja de manera simétrica, los procesos o partes de ellos (threads) son enviados indistintamente a cualesquiera de los procesadores disponibles, teniendo, teóricamente, una mejor distribución y equilibrio en la carga de trabajo bajo este esquema.

Se dice que un thread es la parte activa en memoria y corriendo de un proceso, lo cual puede consistir de un área de memoria, un conjunto de registros con valores específicos, la pila y otros valores de contexto. Un aspecto importante a considerar en estos sistemas es la forma de crear aplicaciones para aprovechar los varios procesadores. Existen aplicaciones que fueron hechas para correr en sistemas monoproceso que no toman ninguna ventaja a menos que el sistema operativo o el compilador detecte secciones de código paralelizable, los cuales son ejecutados al mismo tiempo en procesadores diferentes. Por otro lado, el programador puede modificar sus algoritmos y aprovechar por sí mismo esta facilidad, pero esta última opción las más de las veces es costosa en horas hombre y muy tediosa, obligando al programador a ocupar tanto o más tiempo a la paralelización que a elaborar el algoritmo inicial.

Sistemas Operativos por la Forma de Ofrecer sus Servicios

Esta clasificación también se refiere a una visión externa, que en este caso se refiere a la del usuario, el cómo accesa los servicios. Bajo esta clasificación se pueden detectar dos tipos principales: sistemas operativos de red y sistemas operativos distribuidos.

Sistemas Operativos de Red

Los sistemas operativos de red se definen como aquellos que tiene la capacidad de interactuar con sistemas operativos en otras computadoras por medio de un medio de transmisión con el objeto de intercambiar información, transferir archivos, ejecutar comandos remotos y un sin fin de otras actividades. El punto crucial de estos sistemas es que el usuario debe saber la sintaxis de un conjunto de comandos o llamadas al sistema para ejecutar estas operaciones, además de la ubicación de los recursos que desee accesar. Por ejemplo, si un usuario en la computadora hidalgo necesita el archivo matriz.pas que se localiza en el directorio /software/código en la computadora modelos bajo el sistema operativo UNIX, dicho usuario podría copiarlo a través de la red con los comandos siguientes: hidalgo% hidalgo% rcp modelos:/software/codigo/matriz.pas . hidalgo% En este caso, el comando rcp que significa "remote copy" trae el archivo indicado de la computadora morelos y lo coloca en el directorio donde se ejecutó el mencionado comando. Lo importante es hacer ver que el usuario puede accesar y compartir muchos recursos.

Sistemas Operativos Distribuidos

Los sistemas operativos distribuidos abarcan los servicios de los de red, logrando integrar recursos (impresoras, unidades de respaldo, memoria, procesos, unidades centrales de proceso) en una sola máquina virtual que el usuario accesa en forma transparente. Es decir, ahora el usuario ya no necesita saber la ubicación de los recursos, sino que los conoce por nombre y simplemente los usa como si todos ellos fuesen locales a su lugar de trabajo habitual. Todo lo anterior es el marco teórico de lo que se desearía tener como sistema operativo distribuido, pero en la realidad no se ha conseguido crear uno del todo, por la complejidad que suponen: distribuir los procesos en las varias unidades de procesamiento, reintegrar sub-resultados, resolver problemas de concurrencia y paralelismo, recuperarse de fallas de algunos recursos distribuidos y consolidar la protección y seguridad entre los diferentes componentes del sistema y los usuarios. Los avances tecnológicos en las redes de área local y la creación de microprocesadores de 32 y 64 bits lograron que computadoras mas o menos baratas tuvieran el suficiente poder en forma autónoma para desafiar en cierto grado a los mainframes, y a la vez se dio la posibilidad de intercomunicarlas, sugiriendo la oportunidad de partir procesos muy pesados en cálculo en unidades más pequeñas y distribuirlas en los varios microprocesadores para luego reunir los subresultados, creando así una máquina virtual en la red que exceda

en poder a un mainframe. El sistema integrador de los microprocesadores que hacer ver a las varias memorias, procesadores, y todos los demás recursos como una sola entidad en forma transparente se le llama sistema operativo distribuido. Las razones para crear o adoptar sistemas distribuidos se dan por dos razones principales: por necesidad (debido a que los problemas a resolver son inherentemente distribuidos) o porque se desea tener más confiabilidad y disponibilidad de recursos. En el primer caso tenemos, por ejemplo, el control de los cajeros automáticos en diferentes estados de la república. Ahí no es posible ni eficiente mantener un control centralizado, es más, no existe capacidad de cómputo y de entrada/salida para dar servicio a los millones de operaciones por minuto. En el segundo caso, supóngase que se tienen en una gran empresa varios grupos de trabajo, cada uno necesita almacenar grandes cantidades de información en disco duro con una alta confiabilidad y disponibilidad. La solución puede ser que para cada grupo de trabajo se asigne una partición de disco duro en servidores diferentes, de manera que si uno de los servidores falla, no se deje dar el servicio a todos, sino sólo a unos cuantos y, más aún, se podría tener un sistema con discos en espejo (mirror) a través de la red de manera que si un servidor se cae, el servidor en espejo continúa trabajando y el usuario ni cuenta se da de estas fallas, es decir, obtiene acceso a recursos en forma transparente.

Ventajas de los Sistemas Distribuidos

En general, los sistemas distribuidos (no solamente los sistemas operativos) exhiben algunas ventajas sobre los sistemas centralizados que se describen enseguida.

- Economía: El cociente precio/desempeño de la suma del poder de los procesadores separados contra el poder de uno solo centralizado es mejor cuando están distribuidos.
- Velocidad: Relacionado con el punto anterior, la velocidad sumada es muy superior.
- Confiabilidad: Si una sola máquina falla, el sistema total sigue funcionando.
- Crecimiento: El poder total del sistema puede irse incrementando al añadir pequeños sistemas, lo cual es mucho más difícil en un sistema centralizado y caro.
- Distribución: Algunas aplicaciones requieren de por sí una distribución física.

Por otro lado, los sistemas distribuidos también exhiben algunas ventajas sobre sistemas aislados. Estas ventajas son:

- Compartir datos: Un sistema distribuido permite compartir datos más fácilmente que los sistemas aislados, que tendrían que duplicarlos en cada nodo para lograrlo.
- Compartir dispositivos: Un sistema distribuido permite accesar dispositivos desde cualquier nodo en forma transparente, lo cual es imposible con los sistemas aislados. El sistema distribuido logra un efecto sinergético.
- Comunicaciones: La comunicación persona a persona es factible en los sistemas distribuidos, en los sistemas aislados no. _ Flexibilidad: La distribución de las cargas de trabajo es factible en el sistema distribuidos, se puede incrementar el poder de cómputo.

Desventajas de los Sistemas Distribuidos

Así como los sistemas distribuidos exhiben grandes ventajas, también se pueden identificar algunas desventajas, algunas de ellas tan serias que han frenado la producción comercial de sistemas operativos en la actualidad. El problema más importante en la creación de sistemas distribuidos es el software: los problemas de compartir datos y recursos es tan complejo que los mecanismos de solución generan mucha sobrecarga al sistema haciéndolo ineficiente. El evaluar, por ejemplo, quiénes tienen acceso a algunos recursos y quiénes no, el aplicar los mecanismos de protección y registro de permisos consume demasiados recursos. En general, las soluciones presentes para estos problemas están aún en pañales.

Otros problemas de los sistemas operativos distribuidos surgen debido a la concurrencia y al paralelismo. Tradicionalmente las aplicaciones son creadas para computadoras que ejecutan secuencialmente, de manera que el identificar secciones de código `paralelizable' es un trabajo arduo, pero necesario para dividir un proceso grande en sub-procesos y enviarlos a diferentes unidades de procesamiento para lograr la distribución. Con la concurrencia se deben implantar mecanismos para evitar las condiciones de competencia, las postergaciones indefinidas, el ocupar un recurso y estar esperando otro, las condiciones de espera circulares y finalmente, los "abrazos mortales" (deadlocks). Estos problemas de por sí se presentan en los sistemas operativos multiusuarios o multitareas, y su tratamiento en los sistemas distribuidos es aún más complejo, y por lo tanto, necesitará de algoritmos más complejos con la inherente sobrecarga esperada.

Por otro lado, en el tema de sistemas distribuidos existen varios conceptos importantes referentes al hardware que no se ven en este trabajo: multicomputadoras, multiprocesadores, sistemas acoplados débil y fuertemente, etc.

ARQUITECTURA DE WINDOWS NT
COMPONENTES DEL NT EXECUTIVE

El Administrador de Objetos (Object Manager).

Se encarga de crear, destruir y gestionar todos los objetos del Executive. Los siguientes son ejemplos de objetos en Windows NT: procesos, subprocesos, archivos, segmentos de memoria compartida, semáforos, mutex, sucesos, etc. Los subsistemas de entorno (Win32, OS/2 y POSIX) también tienen sus propios objetos. Por ejemplo, un objeto ventana es creado (con ayuda del administrador de objetos) y gestionado por el subsistema Win32. La razón de no incluir la gestión de ese objeto en el Executive es que una ventana sólo es innata de las aplicaciones Windows, y no de las aplicaciones UNIX o OS/2. Por tanto, el Executive no se encarga de administrar los objetos relacionados con el entorno de cada sistema operativo concreto, sino de los objetos comunes a los tres.

El Monitor Referente a la Seguridad (SRM).

Este componente da soporte en modo privilegiado al subsistema de seguridad, con el que interacciona frecuentemente. Su misión es actuar de alguna manera como supervisor de accesos, ya que comprueba si un proceso determinado tiene permisos para acceder a un objeto determinado, y monitoriza sus acciones sobre dicho objeto. De esta manera es capaz de generar los mensajes de auditorías. Soporta las validaciones de acceso que realiza el subsistema de seguridad local.

El Administrador de Procesos (Process Manager).

Se encarga (en colaboración con el administrador de objetos) de crear, destruir y gestionar los procesos y subprocesos. Una de sus funciones es la de repartir el tiempo de CPU entre los distintos subprocesos. Suministra sólo las relaciones más básicas entre procesos y subprocesos, dejando el resto de las interrelaciones entre ellos a cada subsistema protegido concreto. Por ejemplo, en el entorno POSIX existe una relación filial entre los procesos que no existe en Win32, de manera que se constituye una jerarquía de procesos. Como esto sólo es específico de ese subsistema, el administrador de objetos no se entromete en ese trabajo y lo deja en manos del subsistema.

La Facilidad de Llamada a Procedimiento Local (LPC Facility).

Este módulo se encarga de recibir y enviar las llamadas a procedimiento local entre las aplicaciones cliente y los subsistemas servidores.

El Administrador de Memoria Virtual (Virtual Memory Manager).

Windows NT y UNIX implementan un direccionamiento lineal de 32 bits y memoria virtual paginada bajo demanda. El VMM se encarga de todo lo relacionado con la política de gestión de la memoria. Determina los conjuntos de trabajo de cada proceso, mantiene un conjunto de páginas libres, elige páginas víctima, sube y baja páginas entre la memoria RAM y el archivo de intercambio en disco, etc.

El Administrador de Entrada/Salida (I/O Manager).

Consta de varios subcomponentes: el administrador del sistema de ficheros, el servidor de red, el redirector de red, los drivers de dispositivo del sistema y el administrador de cachés.

Buena parte de su trabajo es la gestión de la comunicación entre los distintos drivers de dispositivo, para lo cual implementa una interfaz bien definida que permite el tratamiento de todos los drivers de una manera homogénea, sin preocuparse del funcionamiento específico de cada uno. Trabaja en conjunción con otros componentes del Executive, sobre todo con el VMM. Le proporciona la E/S síncrona y asíncrona, la E/S a archivos asignados en memoria y las caches de los archivos. El administrador de caches no se limita a gestionar unos cuantos buffers de tamaño fijo para cada archivo abierto, sino que es capaz de estudiar las estadísticas sobre la carga del sistema y variar dinámicamente esos tamaños de acuerdo con la carga. El VMM realiza algo parecido en su trabajo.

SUBSISTEMAS PROTEGIDOS

Subsistemas de entorno

El subsistema Win32.

Es el más importante, ya que atiende no sólo a las aplicaciones nativas de Windows NT, sino que para aquellos programas no Win32, reconoce su tipo y los lanza hacia el subsistema correspondiente. En el caso de que la aplicación sea MS-DOS o Windows de 16 bits (Windows 3.11 e inferiores), lo que hace es crear un nuevo subsistema protegido. Así, la aplicación DOS o Win16 se ejecutaría en el contexto de un proceso llamado VDM (Virtual DOS Machine, máquina virtual DOS), que no es más que un simulador de un ordenador funcionando bajo MS-DOS. Las llamadas al API Win16 serían correspondidas con las homónimas en API Win32. Microsoft llama a esto WOW (Windows On Win32). El subsistema soporta una buena parte del API Win32. Así, se encarga de todo lo relacionado con la interfaz gráfica con el usuario (GUI), controlando las

entradas del usuario y salidas de la aplicación. Por ejemplo, un buen número de funciones de las bibliotecas USER32 y GDI32 son atendidas por Win32, ayudándose del NT Executive cuando es necesario.

El subsistema POSIX.
Define un conjunto de llamadas al sistema en lenguaje C. El subsistema sirve las llamadas interaccionando con el NT Executive. Se encarga también de definir aspectos específicos del sistema operativo UNIX, como pueden ser las relaciones jerárquicas entre procesos padres e hijos (las cuales no existen en el subsistema Win32, por ejemplo, y que por consiguiente no aparecen implementadas directamente en el NT Executive).

El subsistema OS/2.
Igual que el subsistema POSIX proporciona un entorno para aplicaciones UNIX, este subsistema da soporte a las aplicaciones del sistema operativo OS/2. Proporciona la interfaz gráfica y las llamadas al sistema; las llamadas son servidas con ayuda del NT Executive.

Subsistemas integrales
El subsistema proceso de inicio.
El proceso de inicio (Logon Process) recibe las peticiones de conexión por parte de los usuarios. En realidad son dos procesos, cada uno encargándose de un tipo distinto de conexión: el proceso de inicio local, que gestiona la conexión de usuarios locales directamente a una máquina Windows NT; y el proceso de inicio remoto, el cual gestiona la conexión de usuarios remotos a procesos servidores de NT.

El subsistema de seguridad.
Este subsistema interacciona con el proceso de inicio y el llamado Monitor Referente a la seguridad. El subsistema de seguridad interacciona con el proceso de inicio, atendiendo las peticiones de acceso al sistema. Consta de dos subcomponentes: la autoridad de seguridad local (LSA) y el administrador de cuentas de seguridad (SAM).
El primero es el corazón del subsistema de seguridad. En general, gestiona la política de seguridad local; así, se encarga de generar los permisos de acceso, de comprobar que el usuario que solicita conexión tiene acceso al sistema, de verificar todos los accesos sobre los objetos (para lo cual se ayuda del Monitor Referente a la seguridad) y de controlar la política de auditorías, llevando la cuenta de los mensajes de auditoría generados por el monitor de referencias.
El administrador de cuentas mantiene una base de datos con las cuentas de todos los usuarios (login, claves, identificaciones, etc.). Proporciona los servicios de validación de usuarios requeridos por el subcomponente anterior.