Universidad Nacional Autónoma de México
Facultad de Ingeniería | Arquitectura Cliente Servidor
Ortega de la Paz Rafael | 420054085

Tarea 1: Investigar (Sistemas Operativos)

¿Qué es un proceso?

Un proceso es, en esencia, un programa en ejecución. A cada proceso se le asocia un espacio de direcciones (una lista de ubicaciones de memoria que va desde un mínimo, generalmente 0, hasta cierto valor máximo) donde el proceso puede leer y escribir información; ese espacio de direcciones contiene el programa ejecutable, los datos del programa y su pila. Además, a cada proceso se le asocia un conjunto de recursos: registros (entre ellos el contador de programa y el apuntador de pila), una lista de archivos abiertos, alarmas pendientes, listas de procesos relacionados y toda la demás información necesaria para ejecutar el programa. En pocas palabras, un proceso es un recipiente que guarda toda la información necesaria para ejecutar un programa (Tanenbaum, 2009, p. 38).

Es importante distinguir el proceso del programa: si un mismo programa se ejecuta por duplicado, esto cuenta como dos procesos distintos, aunque el sistema operativo pueda compartir el código entre ellos para que solo haya una copia en memoria (Tanenbaum, 2009, p. 86).

Para llevar el control de todos los procesos que existen en un momento dado, el sistema operativo mantiene una tabla de procesos —también llamada por algunos autores bloques de control de procesos (PCB)—, un arreglo con una entrada por cada proceso. Esa entrada guarda, entre otras cosas, el contador de programa, el apuntador de pila, la asignación de memoria, el estado de los archivos abiertos, la información de contabilidad y planificación, el identificador del proceso (PID), el proceso padre, señales pendientes, etc.; es decir, todo lo que se necesita para poder reanudar el proceso más adelante como si nunca se hubiera detenido (Tanenbaum, 2009, pp. 91-92).

¿Qué es la imagen de un proceso?

La imagen de un proceso es el conjunto completo de información que representa a ese proceso en un instante dado, y que el sistema operativo necesita conservar para poder suspenderlo y reanudarlo exactamente en el mismo estado. Tanenbaum (2009)

explica que un proceso suspendido consiste en dos partes: su espacio de direcciones —conocido comúnmente como imagen de núcleo (en honor a las memorias de núcleo magnético que se usaban antes)— y su entrada en la tabla de procesos, que guarda el contenido de sus registros y demás elementos necesarios para reiniciarlo posteriormente (p. 38). Esto se confirma más adelante cuando se describe la llamada al sistema fork: al crear un proceso hijo, "los dos procesos (padre e hijo) tienen la misma imagen de memoria, las mismas cadenas de entorno y los mismos archivos abiertos" (Tanenbaum, 2009, p. 87), y normalmente el hijo ejecuta después execve para reemplazar su imagen de memoria por un nuevo programa.

Desde una perspectiva más práctica, orientada a cómo se organiza esa imagen en la memoria de una máquina UNIX, Gay (2000) describe la composición de la imagen de memoria de un proceso (process memory image) en sistemas como FreeBSD y Linux. En la parte más alta de la memoria se ubican las variables de entorno; debajo está la pila (stack), que crece hacia direcciones más bajas, con una zona de holgura de memoria sin asignar en su base. El programa en sí —resultado de enlazar los módulos objeto del programa con los módulos de las bibliotecas estáticas— ocupa una región inferior de la memoria, y es precisamente esa colección de regiones (código, datos y las bibliotecas estáticas enlazadas) la que se almacena en el archivo ejecutable. Por debajo de esa zona se reserva un área para las bibliotecas dinámicas (compartidas) que se cargan en memoria cuando el programa inicia su ejecución (Gay, 2000, pp. 247-248).

En conjunto, ambas fuentes muestran que la imagen de un proceso está formada por:

- El segmento de texto (código ejecutable del programa).
- El segmento de datos (variables globales y estáticas).
- La pila (para variables locales, parámetros y direcciones de retorno).
- Las bibliotecas (estáticas, ya incluidas en el ejecutable, y dinámicas, cargadas en tiempo de ejecución).
- La información de control asociada en la tabla de procesos/PCB (registros, contador de programa, apuntador de pila, punteros a los segmentos de texto, datos y pila),

que permite al sistema operativo suspender y reanudar el proceso sin pérdida de información (Tanenbaum, 2009, pp. 91-92).

Referencias

Gay, W. W. (2000). Advanced UNIX programming. Sams Publishing.

Tanenbaum, A. S. (2009). Sistemas operativos modernos (3.ª ed.). Pearson Educación.