

Proyecto 1

El objetivo de este proyecto es comparar el rendimiento de realizar una tarea colaborativa utilizando múltiples hilos. Para ello se deberá desarrollar una versión multihilos del programa *copy* que permita copiar el contenido de un directorio completo.

En este proyecto se deben realizar múltiples pruebas para determinar la cantidad óptima de hilos para ejecutar este tipo de tarea sobre un único directorio muy grande.

Uso del programa

El programa *copy* recibe el nombre del directorio origen y el directorio destino:

```
copy directorio_origen directorio_destino
```

Por omisión el programa debe mostrar el nombre de cada archivo que se copia y su cantidad de bytes.

Lectura de directorios

El hilo principal deberá leer el directorio indicado, crear un pool de hilos y asignar a cada hilo la copia de uno de los archivos en dicho directorio. La lectura de directorios y archivos se puede realizar utilizando como base los tutoriales respectivos. Se debe procesar el directorio y todos los subdirectorios bajo él.

Manejo de múltiples hilos

El programa debe contar con múltiples hilos que se encarguen de copiar los archivos. Debido a que la copia se realiza sobre múltiples archivos no sería eficiente crear un único hilo. Es por ello que es necesario crear un conjunto compartido de hilos (pool de hilos) que incluya solo una cantidad limitada de hilos, y en donde cada hilo trabajará sobre un archivo específico. Como parte del proyecto se deben realizar diferentes pruebas (con diferentes directorios grandes) para determinar un tamaño "óptimo" del pool de hilos. Se utilizará el tiempo de ejecución para determinar dicho tamaño óptimo.

En este caso se utilizará un pool de hilos estático, es decir, el programa creará la totalidad de los hilos al inicio del programa y únicamente utilizará estos para realizar todo el trabajo. Tal como usted debe suponer, al hilo que termine de copiar un archivo deberá asignarse otro archivo para copiar hasta terminar con todo el directorio.

Comunicación entre hilos

El hilo principal (coordinador) debe comunicarse con los hilos copiadores (esclavos) mediante variables (arreglos) compartidas. Las variables contendrán el nombre del archivo a copiar, y el hilo copiadador debe escribir en una variable (arreglo) de vuelta un mensaje indicando que terminó la copia. Note que se deben utilizar semáforos mutex para sincronizar el acceso de los hilos a las variables compartidas.

Al finalizar de copiar todos los archivos el hilo principal deberá comunicar la terminación a cada uno de los hilos esclavos. Hasta que cada uno de estos hilos no haya terminado, no podrá terminar el hilo principal.

Análisis de resultados

Se deberán realizar varias pruebas cambiando la cantidad de hilos en el pool. Una vez que se calcule el tiempo de ejecución final, se deben crear tablas y gráficas en donde se muestre cuál cantidad de hilos brinda el mejor rendimiento para realizar la copia de archivos en paralelo.

Cree un archivo de bitácora (logfile) desde el hilo principal que escriba una entrada por cada archivo que se copia, indicando el nombre del archivo, el subhilo que lo copió, y el tiempo que se duró. Utilice prefereriblemente un formato CSV para poder realizar el análisis de resultados desde Excel o R.

Consideraciones generales

- Todo el desarrollo del proyecto debe realizarse en lenguaje C y sobre ambiente Unix.
- Se deben utilizar hilos de ejecución.
- Se deberá generar una documentación formal, en formato pdf, en donde se describan las diferentes etapas del desarrollo del proyecto, las decisiones de diseño que se tomaron, los mecanismos de programación utilizados, y los resultados de las diferentes pruebas al programa. Dicha documentación deberá incluir al menos las siguientes secciones:
 - Introducción
 - Descripción del problema (este enunciado)
 - Definición de estructuras de datos
 - Descripción detallada y explicación de los componentes principales del programa
 - Mecanismo de creación y comunicación de hilos
 - Pruebas de rendimiento
 - Conclusiones

- El proyecto puede realizarse en grupos de dos estudiantes.
- No se permite la copia de código entre grupos de estudiantes, o entregar código realizado por estudiantes los semestres anteriores, tampoco es permitido utilizar librerías adicionales (desarrolladas por terceros).