

PRÁCTICA 3: Programación de altas prestaciones

Se dispone de un archivo de entrada con cotizaciones de valores con el siguiente formato:

dd-mm-aaa hh:mm:ss.ssssss "VVVVVV" precio

Por ejemplo

05-03-2016 11:03:15.034371 "GOOGLE" 804.60

05-03-2016 11:03:15.035123 "APPLE" 117.63

...

Como puede verse, el formato de la fecha usa dos dígitos para el día, dos para el mes y dos para el año. Por otra parte la hora, usa 2 dígitos para la hora (que puede variar entre 0 y 23), dos dígitos para los minutos (entre 0 y 59) y la representación de los segundos con precisión de microsegundo (entre 0 y 59.999999). El nombre del valor, irá delimitado por comillas y puede contener cualquier número de caracteres (incluyendo blancos y caracteres especiales). El valor de la cotización es un número real que utiliza precisión de céntimo.

Los datos en este fichero aparecen sin ningún tipo de orden. En particular, no están ordenados ni por fecha ni por nombre de sociedad.

Dado este fichero, se desea generar un fichero por cada valor con el siguiente formato:

dd-mm-aa hh:mm precio

El nombre de cada uno de estos ficheros debe ser el del valor con extensión txt (p. ej.: google.txt, apple.txt, ...). Para generar el nombre, convertirán previamente todos los caracteres a minúsculas y se eliminarán posibles caracteres especiales que puedan provocar problemas en el nombre de un fichero.

El fichero de salida, contendrá en cada línea como precio el valor promedio durante cada minuto. Si para un minuto determinado no hay cotizaciones no se generará la línea. Si para un minuto solamente hay una cotización, se generará una línea con dicho valor. Si para un minuto hay múltiples cotizaciones, se usará una función de interpolación lineal para determinar el precio medio (con pesos) durante ese minuto.

Los datos en cada fichero de precios se presentarán ordenados por los siguientes criterios: año, mes, día, hora, minuto.

El programa tomará los siguientes parámetros como opciones:

- -i input: Se leerán los datos del fichero input. Si no se pasa esta opción se leerá de la entrada estándar.
- -o path: Se escribirán los ficheros de salida en el directorio indicado por path. Si no se indica nada, se usará como ruta, el directorio actual.

Así, se considerará que para el programa ejecutable (cuyo nombre será p1), las siguientes invocaciones serán válidas:

p1

Lee los datos de la entrada estándar y genera los ficheros de salida en el directorio actual.

p1 -i datos.txt

Lee los datos del archivo datos.txt y genera los ficheros de salida en el directorio actual

p1 -o datos/prueba1

Lee los datos de la entrada estándar y genera los archivos de salida en el directorio datos/prueba1.

p1 -i datos.txt -o datos/prueba1

Lee los datos del archivo datos.txt y genera los archivos de salida en el directorio datos/prueba1.

Se deberán entregar dos versiones de la aplicación:

- a) Versión secuencial con menor tiempo de ejecución posible.
- b) Versión paralela, usando Intel TBB, con menor tiempo de ejecución posible.