Laboratorio Hilos y Paralelismo	01

Sebastián Cardona, Laura Gil, Zayra Gutiérrez

Ingeniero de Sistemas Javier Toquica Barrero

Universidad Escuela Colombiana de ingeniería Julio Garavito



# Contenido

Introducción	3
Desarrollo del Laboratorio	4
Parte I Hilos Java	4
Parte II Hilos Java	6
Parte III Evaluación de Desempeño	11
Conclusiones	23



## Introducción

El informe tiene como objetivo analizar y explorar el uso de hilos, paralelismo y concurrencia en Java, enfocándonos en la solución de problemas que requiere un alto rendimiento computacional. Se abordarán tres secciones importantes: la creación y manejo de hilos básicos, la implementación de algoritmos paralelos mediante el uso de hilos y la evaluación del desempeño del paralelismo en diferentes escenarios.

Se analizará el desempeño de las soluciones desarrolladas mediante la variación del número hilos utilizados, contrastando los resultados con la ley Amdahl para evaluar los límites del paralelismo en un entorno controlado.



#### Desarrollo del Laboratorio

### Parte I Hilos Java

1. De acuerdo con lo revisado en las lecturas, complete las clases CountThread, para que las mismas definan el ciclo de vida de un hilo que imprima por pantalla los números entre A y B.

Se creo la clase CounThread que extiende de Thread, tiene un constructor que recibe los números a y b y luego en el método run muestra los números de a hasta b, incluyendo el nombre del hilo actual.

- 2. Complete el método main de la clase CountMainThreads para que:
  - i. Se creo 3 hilos de tipo CountThread, asignándole al primero el intervalo [0..99], al segundo [99..199], y al tercero [200..299].

```
public class CountThreadsHain {

Run main | Debug main | Run | Debug public static void main(string a[]) ||

CountThread hijo2 - new CountThread(a:[90, b:99);

CountThread hijo2 - new CountThread(a:[90, b:99);

CountThread hijo2 - new CountThread(a:[90, b:99);

CountThread hijo2 - new CountThread(a:[90, b:299);

Activar Windows

Ve a Configuracion para activar Windows.

Ve a Configuracion para activar Windows.

Plava Ready SonarQube focus: overall code

Ln 9, Col 55 Spaces 4 UTF-8 CRLF (à Java P Go Live S C 400 pm.)

Three Ready SonarQube focus: overall code

Ln 9, Col 55 Spaces 4 UTF-8 CRLF (à Java P Go Live S C 400 pm.)

Addition of the count of
```

ii. Se inicio los tres hilos con 'start()'.



```
package edu.eci.arsw.threads;

public class CountThreaddMain {

Run main | Debug main | Run | Debug
public static void main(String a[]){}

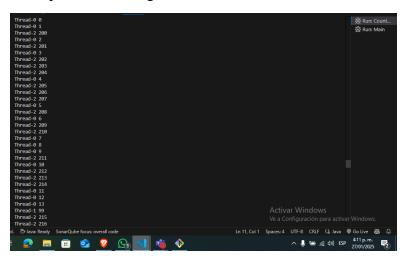
CountThread hllo1 = new CountThread(a:0, b:99);

CountThread hllo2 = new CountThread(a:280, b:299);

| hilo1.start();
| hilo2.start();
| hilo2.start();
| hilo3.start();
| hilo3.start();
| hilo3.start();
| hilo4.start();
| hilo5.start();
| hilo6.start();
| hilo6.start();
| hilo7.start();
| hilo8.start();
| hi
```

iii. Se ejecuto y reviso la salida por pantalla.

Se logra evidenciar como los tres hilos ocurren de manera concurrente y se imprimen en pantalla los rangos de los tres hilos a la vez



iv. Se cambio el inicio con 'start()' por 'run()'. ¿Cómo cambia la salida?, por qué?



```
package edu.eci.arsw.threads;

public class CountThreadsMain {

Run main | Debug main | Run | Debug public static void main(string a[]) {

CountThread hilo1 - new CountThread(a:0, b:99);

CountThread hilo2 - new CountThread(a:200, b:299);

| CountThread hilo3 - new CountThread(a:200, b:299);

| bijol.stant(); hilo2.stant(); hilo2.stant(); hilo3.stant();

| hilo2.stant(); hilo3.run(); hilo3.run(); hilo3.run();

| bijol.stant(); hilo3.run(); hil
```

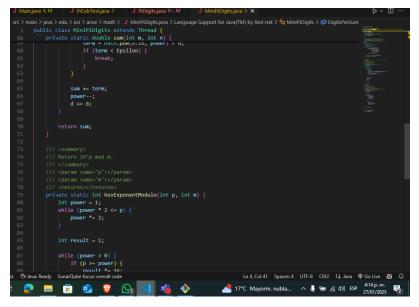
Con el método start(), se crea un nuevo hilo en ejecución. Este método realiza una llamada interna al método run(), permitiendo que la tarea se ejecute de forma paralela al hilo principal u otros hilos existentes.

Por otro lado, al llamar directamente al método run(), no se crea un nuevo hilo. En su lugar, la ejecución ocurre en el hilo actual, de forma secuencial, como cualquier otro método común de Java.

### Parte II Hilos Java

1. Cree una clase de tipo Thread que represente el ciclo de vida de un hilo que calcule una parte de los dígitos requeridos.

La clase MiniPiDigits se encargará de contener la lógica de calcular los count dígitos de pi desde start, solo que esta vez esta clase hereda de Thread





2. Haga que la función PiDigits.getDigits() reciba como parámetro adicional un valor N, correspondiente al número de hilos entre los que se va a paralelizar la solución. Haga que dicha función espere hasta que los N hilos terminen de resolver el problema para combinar las respuestas y entonces retornar el resultado. Para esto, revise el método join del API de concurrencia de Java.

En el método getDigits() efectivamente se agrega el parámetro N para primero dividir el rango de dígitos a calcular por cada hilo, y en el caso de hacer una división no exacta, guardar los números del rango que faltan, luego se crea una lista de los N hilos, donde se establece el inicio de cada hilo y su rango.

Después, todos los hilos pasan a ejecución con el método start() de cada hilo y antes de juntar la respuesta, usamos el método join para esperar que todos los hilos se ejecutaran y unirlos con arrayCopy.

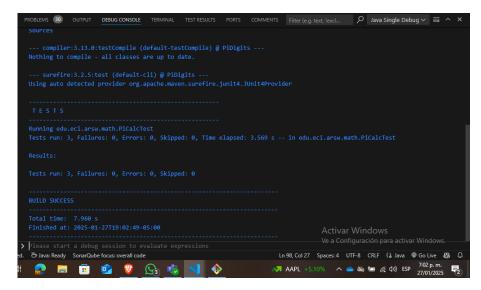


3. Ajuste las pruebas de JUnit, considerando los casos de usar 1, 2 o 3 hilos (este último para considerar un número impar de hilos)

Se crearon tres pruebas cada uno con 1, 2 y tres hilos respectivamente, todas las pruebas pasaron exitosamente

```
int numThreads = 1;
                   for (int start = 0; start < expected.length; start++) {
   for (int count = 0; count < expected.length - start; count++) {
      byte[] digits = PiDigits.egetbigists(start, count, numThreads);
      assertEquals(count, digits.length);</pre>
                            for (int i = 0; i < digits.length; i++) {
   assertEquals(expected[start + i], digits[i]);</pre>
                   Ln 98, Col 27 Spaces: 4 UTF-8 CRLF ( ) Java 🌳 Go Live 🔠 🚨
      📀 🔚 🗓 💁 🦁 🚫 🗞
                                                                                      int numThreads = 2:
                   for (int start = 0; start < expected.length; start++) {
    for (int count = 0; count < expected.length - start; count++) {
        byte[] digits = PiDigits.getDigits(start, count, numThreads);
        assertEquals(count, digits.length);
    }
}</pre>
                            for (int i = 0; i < digits.length; i++) {
    assertEquals(expected[start + i], digits[i]);</pre>
                   Ln 98, Col 27 Spaces: 4 UTF-8 CRLF ( ) Java @ Go Live 🔠 🚨
      📀 🔚 🗓 💁 🦁 🚱 🖏
                                                                                    public void piGenTest2() throws Exception [
                  for (int start = 0; start < expected.length; start++) {
    for (int count = 0; count < expected.length - start; count++) {
        byte[] digits = ploigits.getDigits(start, count, numThreads);
        assertEquals(count, digits.length);</pre>
                           for (int i = 0; i < digits.length; i++) {
    assertEquals(expected[start + i], digits[i]);</pre>
                                                                                      Ln 98, Col 27 Spaces: 4 UTF-8 CRLF ( ) Java @ Go Live 🔠 🚨
🕸 🕵 🕫 🙃 💁
```



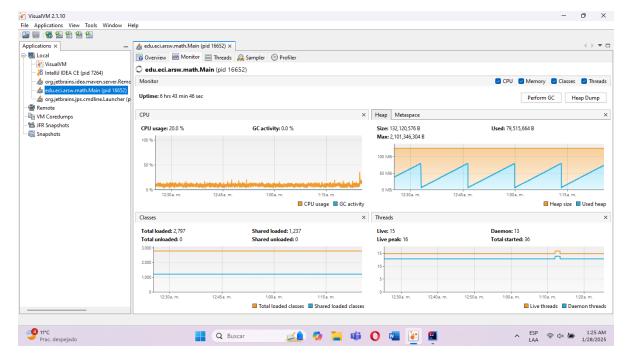


## Parte III Evaluación de Desempeño

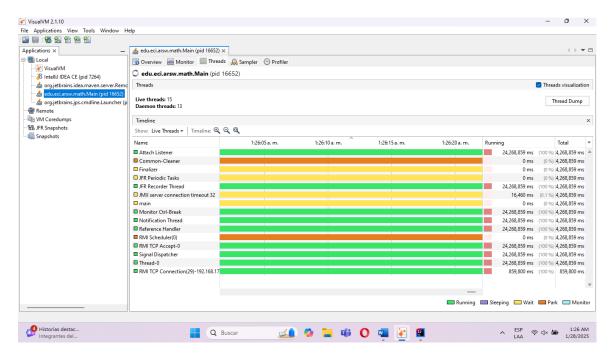
A partir de lo anterior, implemente la siguiente secuencia de experimentos para calcular el millon de dígitos (hex) de PI, tomando los tiempos de ejecución de los mismos (asegúrese de hacerlos en la misma máquina):

#### 1. Un solo hilo.

## Tiempo Estimado: 6 horas, 43 min en la foto, pero en realidad duro más de 14 horas

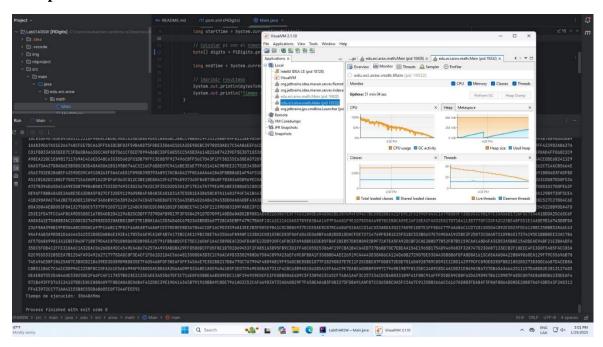




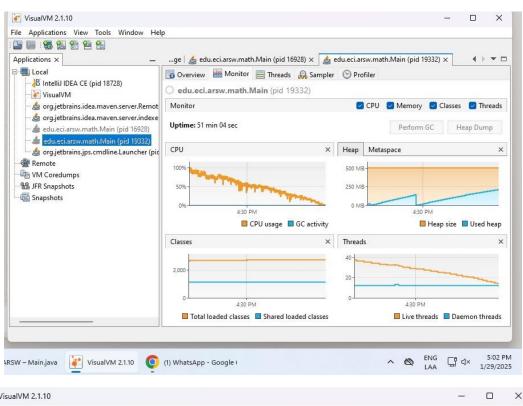


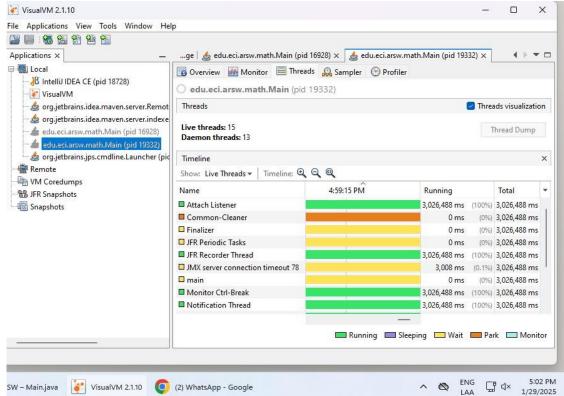
2. Tantos hilos como núcleos de procesamiento (haga que el programa determine esto haciendo uso del API Runtime).

## Tiempo Estimado: 51 min

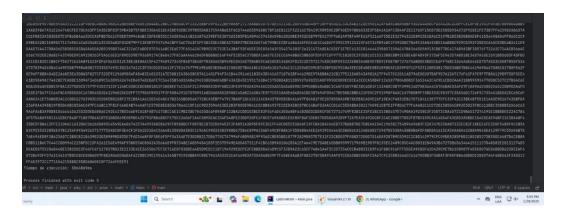






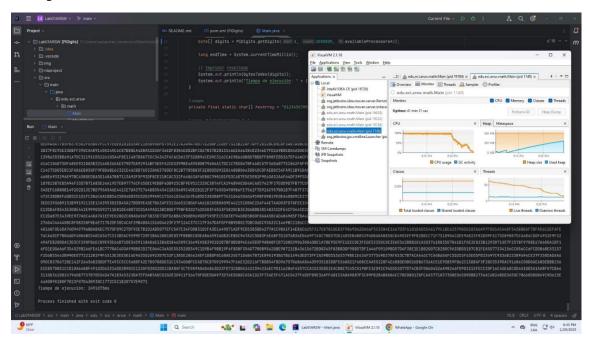




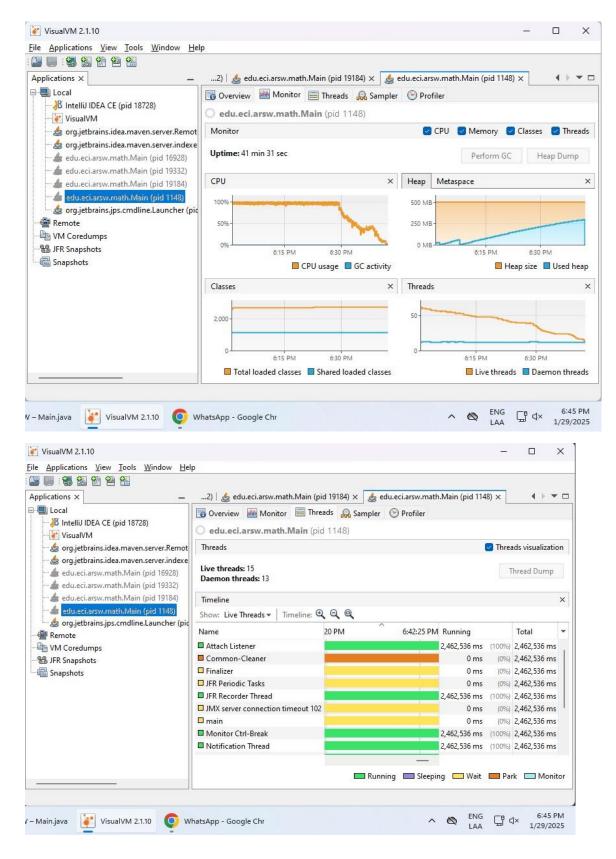


3. Tantos hilos como el doble de núcleos de procesamiento (24 hilos).

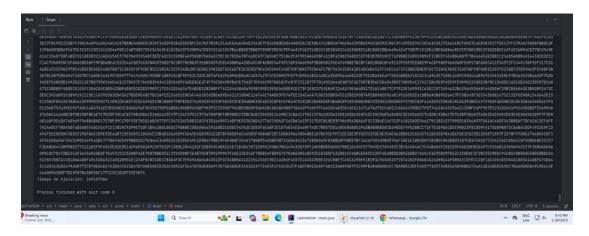
## Tiempo Estimado: 41 min





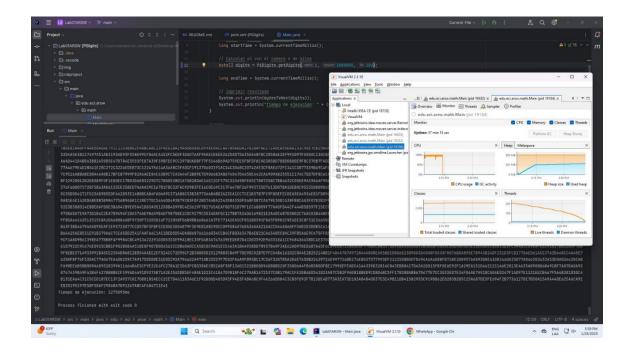




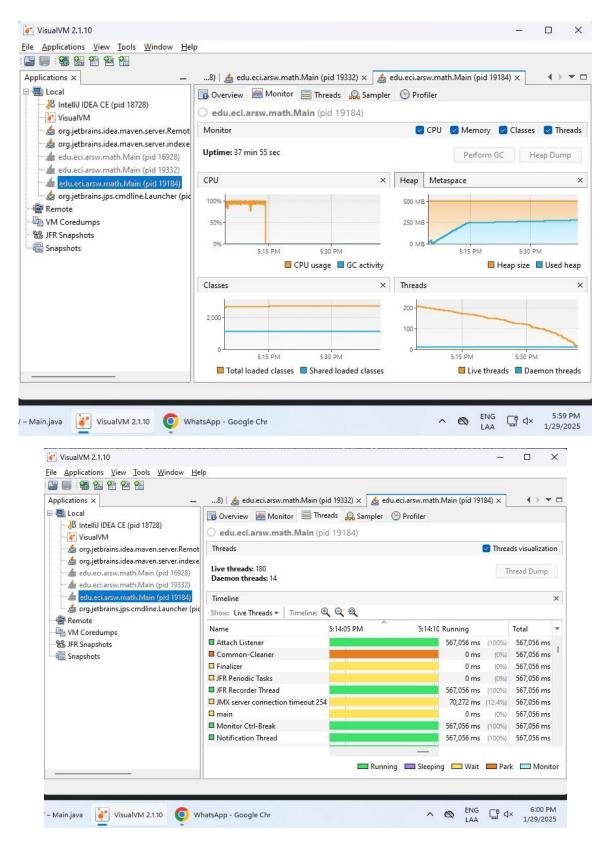


## 4. 200 hilos.

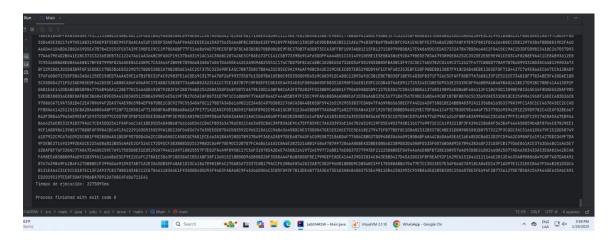
## Tiempo Estimado: 37 min









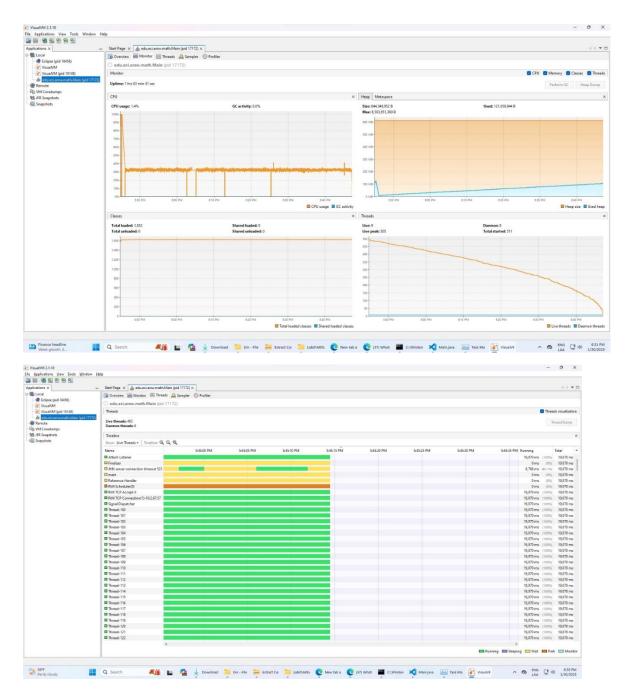


## 5. 500 hilos.

## Tiempo Estimado: 1 hora y 3 minutos

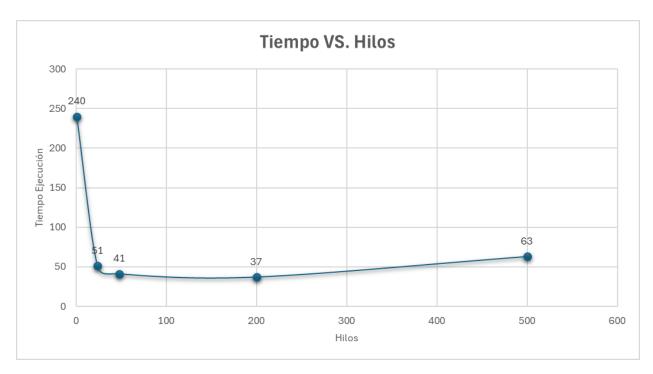
```
| Montoniple | Josephane | Josephane | Josephane | Montoniple | Montoniple | Josephane | J
```





Con lo anterior, y con los tiempos de ejecución dados, haga una gráfica de tiempo de solución vs. número de hilos. Analice y plantee hipótesis con su compañero para las siguientes preguntas (puede tener en cuenta lo reportado por jVisualVM):





Queremos mencionar que las características del equipo de cómputo fueron relevantes para el desarrollo de este laboratorio, pues, primero intentamos ejecutar el programa con un hilo desde un computador de 4 núcleos de procesamiento, luego con uno de 8, pero lamentablemente el tiempo de ejecución excedió las 14 horas, por lo que, toda esta parte se realizó desde una máquina del laboratorio de ingeniería de software cuya cantidad de núcleos de procesamiento son 24.

En conclusión, se observó una mejora absoluta entre correr el programa con un hilo a con los del procesador del computador, esto se ve, ya que cada hilo se reparte el trabajo equitativamente y lo hacen al tiempo, para luego unir el resultado, sin embargo, al momento de usar el doble de hilos del procesador, y de usar 200 hilos, hubo mejora, pero fue mínima, podemos suponer que esto sucedió, pues al usar demasiados hilos, puede reducir el rendimiento del computador esto sucede pues hay una sobrecarga de hilos, donde requiere trabajo pasar entre contexto de cada hilo y se consume mayor recursos de la memoria RAM. Es por esto que se evidencia un retroceso cuando pasamos de 200 a 500 hilos, pues se redujo el tiempo en 26 minutos

## 1. Según la ley de Amdahl:

$$S(n) = \frac{1}{(1-P) + \frac{P}{n}}$$

donde S(n) es el mejoramiento teórico del desempeño, P la fracción paralelizable del algoritmo, y n el número de hilos, a mayor n, mayor debería ser dicha mejora. Por qué el mejor desempeño no se logra con los 500 hilos?, cómo se compara este desempeño cuando se usan 200?.

• Tiempo con 200 hilos: 37 minutos.



• Tiempo con 500 hilos: 1 hora y 3 minutos (63 minutos).

Usamos los datos de 200 hilos y 500 hilos para estimar *P*.

1. Para 200 hilos:

$$37 = (1 - P) * T + (\frac{P * T}{200})$$

2. Para 500 hilos:

$$63 = (1 - P) * T + (\frac{P * T}{500})$$

Restamos las dos ecuaciones para eliminar T:

$$63 - 37 = (1 - P) * T + (\frac{P * T}{200})$$

$$P * T = -\left(\frac{26}{0.003}\right) = -8666.67$$

Aquí nos damos cuenta que el resultado no tiene sentido físico porque P y T deben ser positivos. Esto sugiere que los tiempos proporcionados no son consistentes con la ley de Amdahl o que existen otros factores, como el overhead, que no se están considerando.

Dado que no podemos estimar P directamente, asumiremos un valor razonable para P, donde a ser P=0.95 (95% del algoritmo es paralelizable).

1. Para 200 hilos:

$$S(200) = \left(\frac{1}{(1 - 0.95) + \left(\frac{0.95}{200}\right)}\right) = 18.26$$

2. Para 500 hilos:

$$S(500) = \left(\frac{1}{(1 - 0.95) + \left(\frac{0.95}{500}\right)}\right) = 19.27$$

El mejor desempeño no se logra con 500 hilos debido al **overhead** y la **contención de recursos**. A más hilos, el sistema debe gestionar más contextos, lo que añade tiempo en creación, sincronización y comunicación. Además, los hilos compiten por recursos como la memoria caché y el ancho de banda de la RAM, reduciendo la eficiencia. La ley de Amdahl indica que la parte no paralelizable (1–*P*) limita la mejora, sin importar el número de hilos. Con 200 hilos, hay menos overhead y contención, resultando en un mejor tiempo (37 minutos) que con 500 hilos (63



minutos). Teóricamente, 500 hilos ( $S(500)\approx19.27$ ) superan ligeramente a 200 ( $S(200)\approx18.26$ ), pero en la práctica, el overhead y la contención empeoran el rendimiento.

2. ¿Cómo se comporta la solución usando tantos hilos de procesamiento como núcleos comparados con el resultado de usar el doble de éste?

Usar tantos hilos como núcleos suele ser óptimo porque cada hilo puede ejecutarse en un núcleo sin competir por recursos de CPU, minimizando el overhead y la contención. En cambio, usar el doble de hilos que núcleos obliga al sistema operativo a realizar **multiplexación** (cambio de contexto entre hilos), lo que introduce overhead y reduce el rendimiento debido a la competencia por recursos como la CPU y la memoria caché.

3. De acuerdo con lo anterior, si para este problema en lugar de 500 hilos en una sola CPU se pudiera usar 1 hilo en cada una de 500 máquinas hipotéticas, la ley de Amdahl se aplicaría mejor?. Si en lugar de esto se usaran c hilos en 500/c máquinas distribuidas (siendo c es el número de núcleos de dichas máquinas), se mejoraría? . Explique su respuesta.

La ley de Amdahl se aplica igual con 1 hilo en 500 máquinas o c hilos en 500/c máquinas, ya que el límite teórico depende de la fracción paralelizable P. Sin embargo, distribuir los hilos en múltiples máquinas reduce el overhead y la contención de recursos, mejorando el rendimiento práctico. Aunque no se supera el límite teórico, esta distribución es más eficiente que usar 500 hilos en una sola CPU, donde el overhead y la contención degradan el desempeño.



#### **Conclusiones**

Los hilos y el paralelismo son herramientas clave en programación para crear aplicaciones más rápidas. Los hilos comparten memoria dentro de un proceso, mientras que el paralelismo ejecuta tareas simultáneamente en diferentes núcleos. Usarlos mejora el rendimiento al dividir tareas, pero también introduce complejidad y posibles problemas de concurrencia.

Extender la clase Thread en Java permite crear hilos personalizados con comportamientos específicos. El método run() es donde se define el código que el hilo ejecutará, mientras que el método start() inicia la ejecución del hilo. En esencia, start() crea un nuevo hilo y luego llama al método run() dentro de ese nuevo hilo, permitiendo la ejecución concurrente del código.

El número de hilos no debe exceder la capacidad del hardware para evitar pérdidas de rendimiento. Aunque la ley de Amdahl sugiere que más hilos mejoran el desempeño, en la práctica, el overhead de gestión y la contención de recursos (como la memoria caché y el ancho de banda de la RAM) pueden degradar el rendimiento. Esto se evidencia al comparar 200 hilos (37 minutos) con 500 hilos (63 minutos), donde un mayor número de hilos empeoró el tiempo de ejecución debido a la saturación del sistema.