

Arquitectura de Computadores



Tema 4-2. Procesadores de Multihilos Simultáneos

Sumario

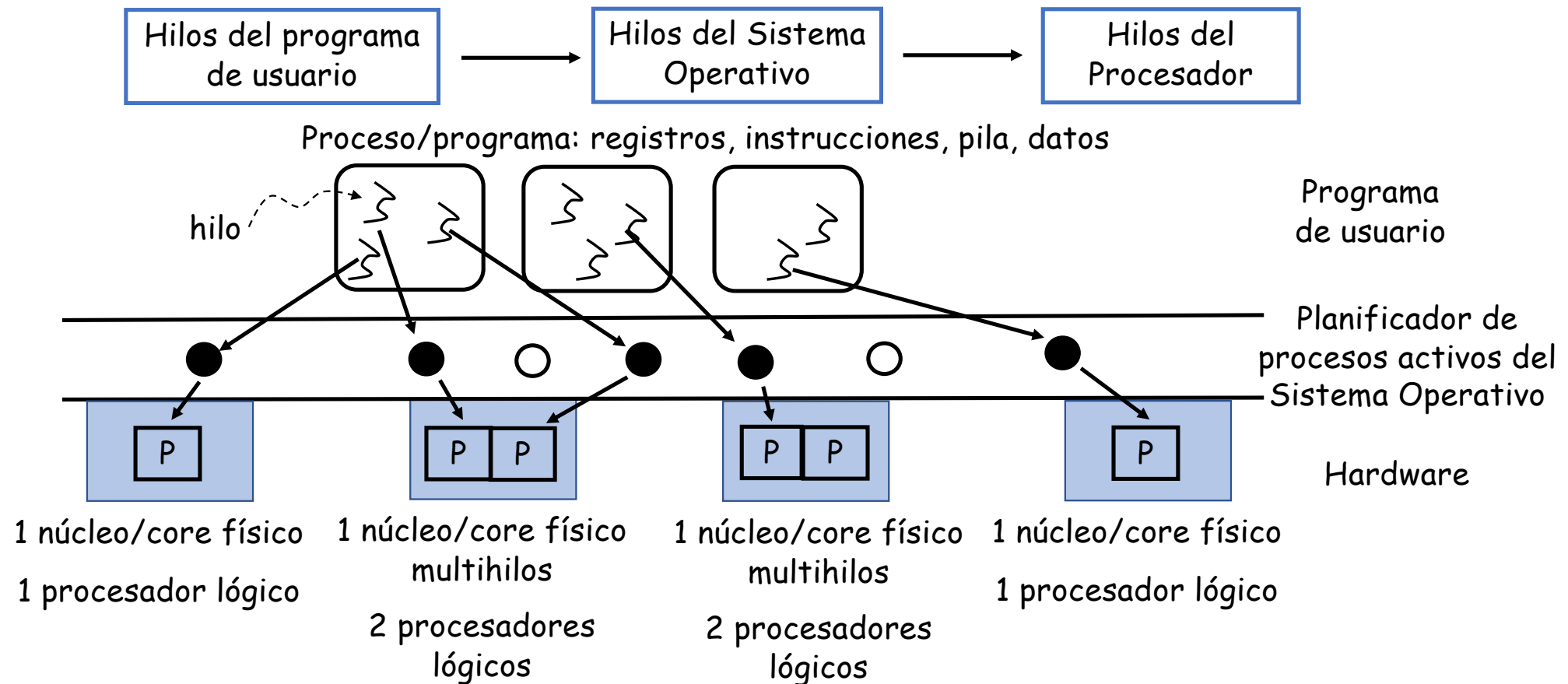
- Fundamentos
- Arquitectura multihilos de grano fino
- Arquitectura multihilos de grano grueso
- Arquitectura multihilos simultáneos
- Comparación entre arquitecturas multihilos
- Evaluación de prestaciones
- Procesadores comerciales

Concepto de HILO en Sistemas Operativos

- Definición de **Proceso** en S.O.
 - Un Proceso es un programa en ejecución. Uno o más hilos se ejecutan en el contexto de un proceso.
- Definición de **Hilo** en S.O.
 - Un Hilo es la unidad básica a la que el Sistema Operativo asigna tiempo de ejecución. Un hilo puede ejecutar cualquier parte del código de un proceso, incluyendo partes actualmente siendo ejecutadas por otro hilo.
- Referencia (gugleando!!): [Processes and Threads - Win32 apps | Microsoft Learn](https://learn.microsoft.com/en-us/windows/win32/procthread/processes-and-threads)
 - <https://learn.microsoft.com/en-us/windows/win32/procthread/processes-and-threads>

Fundamentos de la Tecnología Multihilos

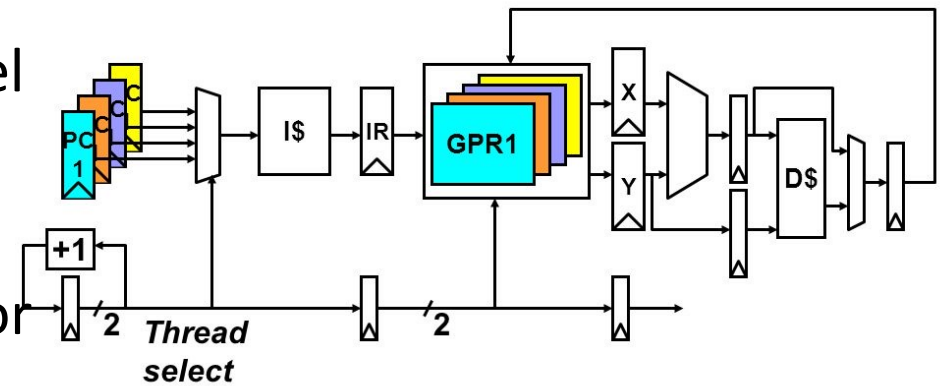
Hilo (thread) **en AC**: conjunto de instrucciones **independientes** que comparten un mismo espacio de direccionamiento (Tema 4-1).



Fundamentos de la Tecnología Multihilos

- **Procesador Multihilos (multithreading):** el núcleo procesador permite ejecutar varios **hilos de instrucciones** de forma paralela en el mismo hardware. Conmutación rápida entre hilos. El S.O. y las aplicaciones multihilo “pueden percibir” que el computador dispone de varios **procesadores “lógicos”**. Por ejemplo: Intel Core i7 960 (integra dos procesadores lógicos en cada núcleo físico).
- Aplica **procesamiento paralelo de tipo funcional** a la ejecución de los programas.
- Aprovecha la **implementación del paralelismo de instrucciones (ILP) y del paralelismo de hilos (TLP)** simultáneamente.

Simple Multithreaded Pipeline

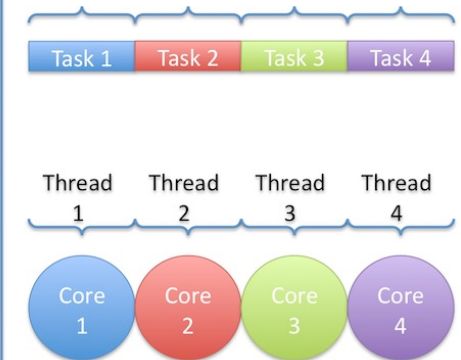


Single Threading



$$ExecutionTime \approx \sum_{i=1}^4 Task_i$$

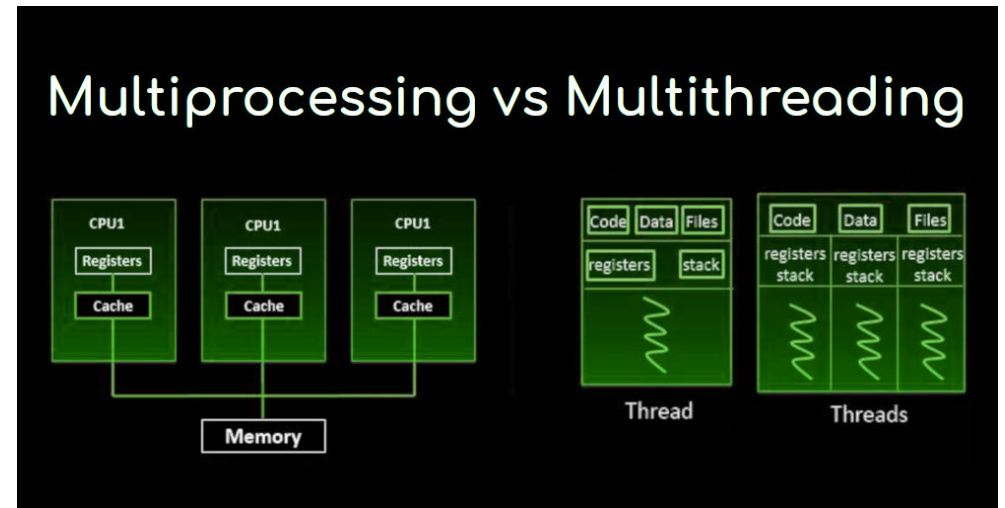
Thread-level Parallelism



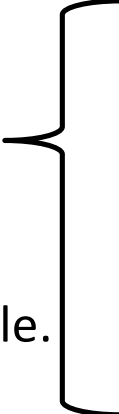
$$ExecutionTime \approx MAX_i (Task_i)$$

Fundamentos de la Tecnología Multihilos

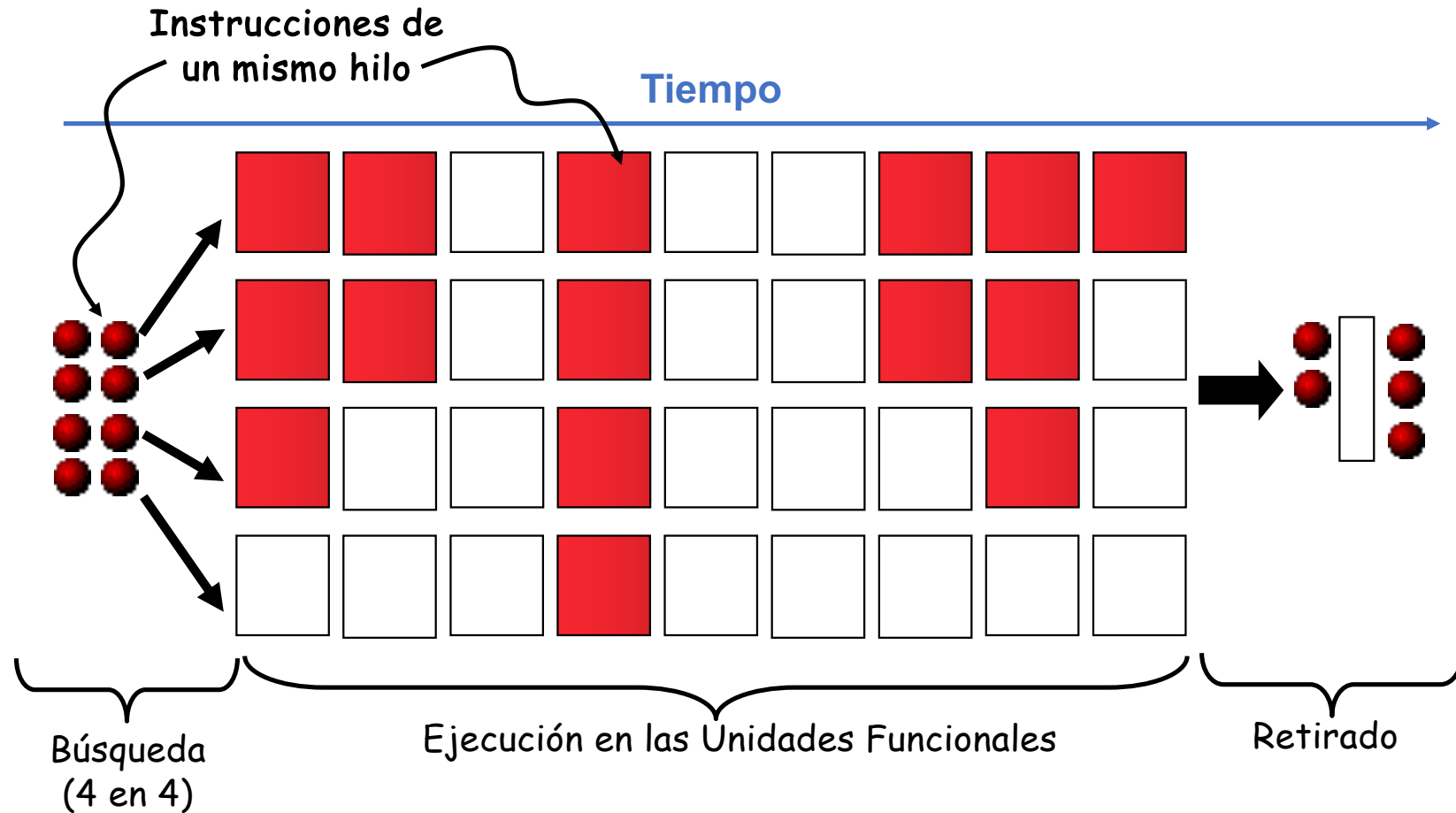
- **Recursos** que necesita un hilo para ejecutarse en el procesador lógico:
 - Contador de programa
 - Estado/contenido del banco de registros
 - Estado/contenido de la pila
- Utiliza los recursos no utilizados de los procesadores superescalares actuales.
- **Aumento de prestaciones** se basa en mejorar la tolerancia de la latencia de la memoria realizando trabajo útil en los periodos en los que un hilo está bloqueado por accesos a memoria.



Fundamentos de la Tecnología Multihilos

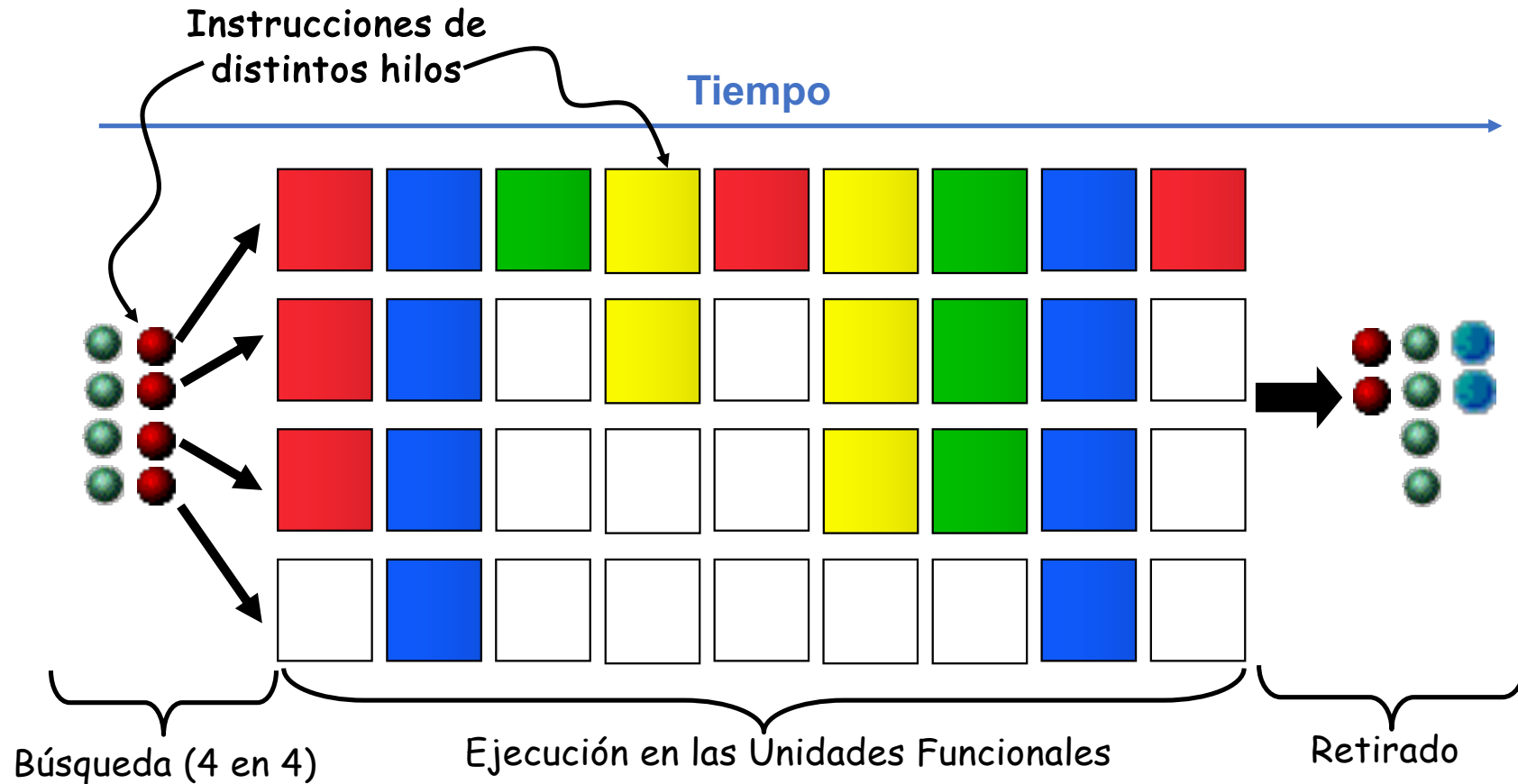
- Diferencia entre núcleo/procesador físico y procesador lógico:
 - En el **procesador físico**, el hilo dispone de todos los recursos hardware.
 - Mientras que en el **procesador lógico** los recursos están compartidos por los hilos que se ejecutan en paralelo.
 - Ventajas:
 - Aumentan las prestaciones del computador.
 - El código que ha sido escrito para computadores paralelos es compatible.
 - Puede ser escalable: el aumento de prestaciones puede ser proporcional al número de procesadores con tecnología multihilos (muy difícil de conseguir).
- 
- Permite **aumentar el número de usuarios** soportados.
 - **Aumenta el número de tareas** terminadas por unidad de tiempo.
 - **Aumenta el número de transacciones** que permite realizar el procesador.

Utilización de Recursos: Procesador Superescalar Monohilo



Los procesadores superescalares consiguen más prestaciones, pero disminuyen la eficiencia de utilización de los recursos del procesador.

Arquitectura Multihilo de Grano Fino

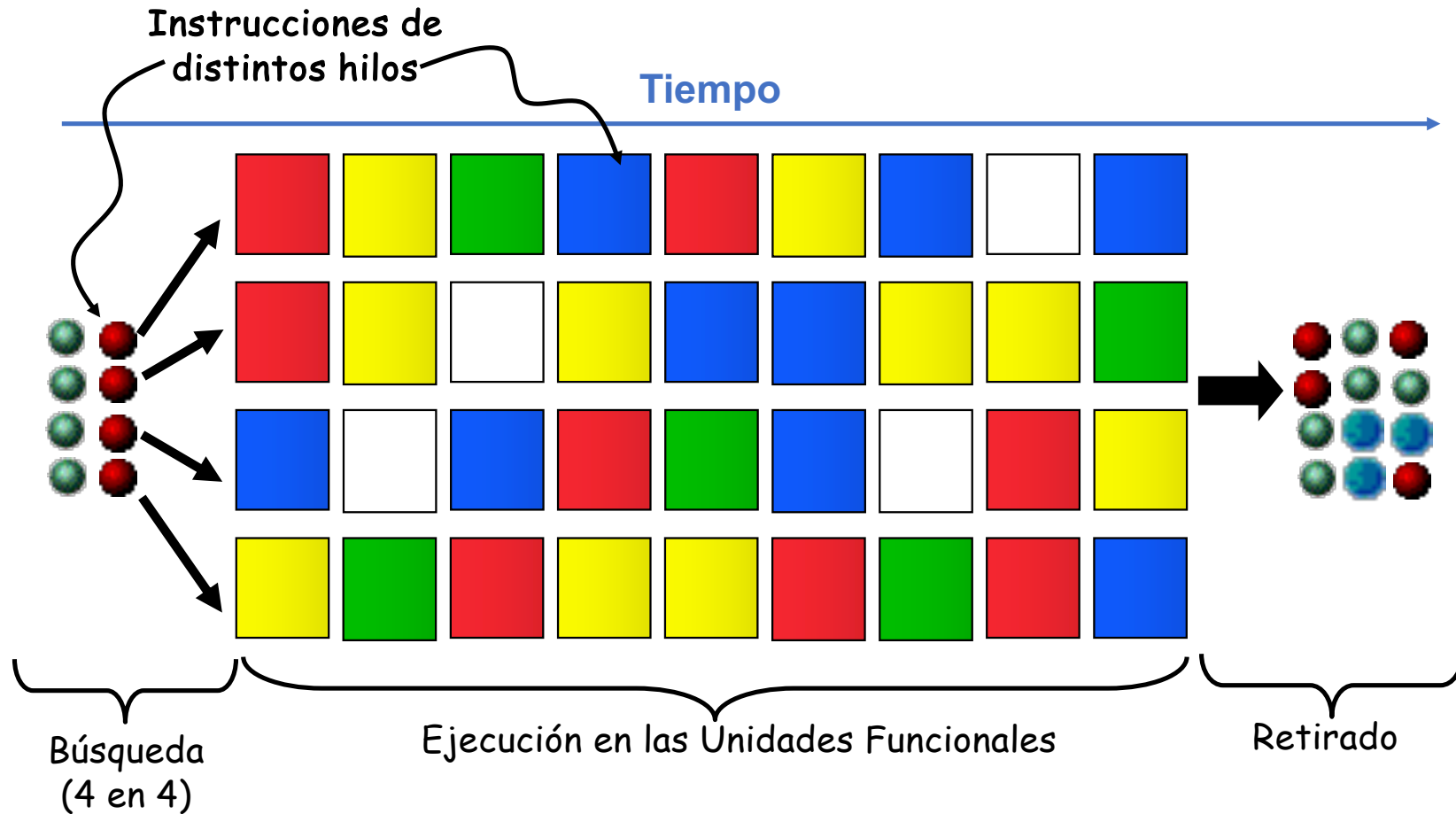


- En cada ciclo de reloj se pueden introducir instrucciones de un sólo hilo.
- Conmutación entre hilos ciclo a ciclo; ejecución entrelazada de hilos.
- Ideal cuando existen grandes penalizaciones en la ejecución de un determinado hilo.
- Desventaja: la ejecución de un hilo que no tiene penalizaciones se puede ralentizar.

Arquitectura Multihilo de Grano Grueso

- Alternativa al grano fino.
- Conmutación entre hilos solo después de eventos relevantes que ocasionen **grandes penalizaciones** tales como un fallo en el último nivel de la memoria cache.
- Ideal cuando existen grandes penalizaciones en la ejecución de un determinado hilo.
- Es menos probable que la ejecución de un hilo que no tiene penalizaciones se puede ralentizar.
- Desventaja: **no puede reaccionar tan rápido** como la arquitectura de grano fino en penalizaciones pequeñas y, por tanto, pierde la habilidad de aumentar el ritmo en el que se terminan los programas por unidad de tiempo.
- Desventaja: para realizar la conmutación entre hilos, se tiene que **vaciar el cauce de instrucciones** antes de introducir en el cauce las instrucciones del segundo hilo. Esto ocasiona penalizaciones en la ejecución de las instrucciones del segundo hilo.

Arquitectura Multihilos Simultáneos



- En cada ciclo de reloj se pueden introducir instrucciones de varios hilos.
- Se produce un aprovechamiento máximo de las unidades funcionales por parte de operaciones independientes.

Comparación entre arquitecturas multihilo en un procesador superescalar

En los 4 hilos superiores se muestra cómo cada hilo se ejecutaría en un **procesador superescalar convencional**. Existe un color distinto para las instrucciones de cada hilo.

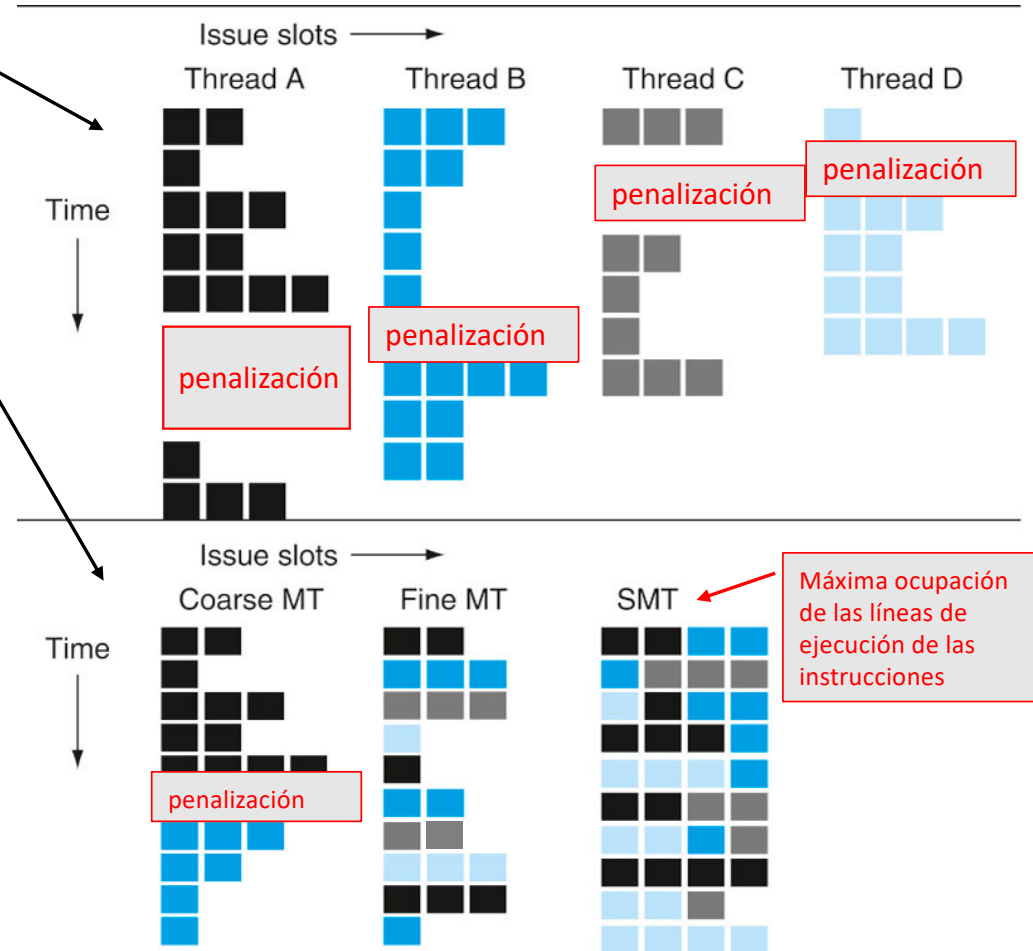
En los 3 ejemplos inferiores se muestra cómo se ejecutarían los 4 hilos en las **tres arquitecturas multihilo** presentadas anteriormente.

En el eje horizontal se representa la capacidad del procesador para enviar a ejecutar varias instrucciones en el mismo ciclo de reloj (superescalaridad).

En el eje vertical se representan los ciclos de reloj.

Un rectángulo blanco representa en un determinado espacio donde podría ir una instrucción pero que no está utilizada.

Para la arquitectura multihilo de grano grueso no se muestran los ciclos de vaciado y llenado del flujo de instrucciones.

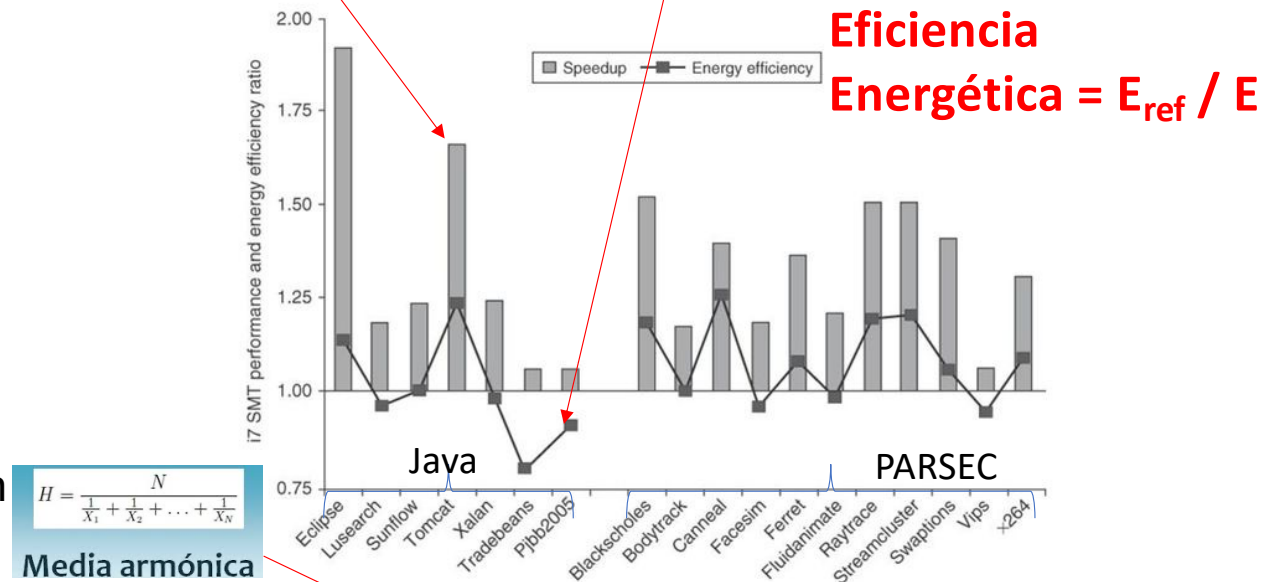


Ganancia de prestaciones de una arquitectura multihilo simultánea (SMT)

Datos obtenidos por Esmailzadeh et. al. [2011] usando la máquina virtual Java 1.6.0 y gcc v4.4.1.

- Sólo se activa **un núcleo físico** en un procesador **Intel Core i7 920 (4 cores, 2,66 GHz, 130 W, 45 nm, Q4'2008), 2 hilos hardware por core.**
- Benchmark con aplicaciones multihilos: aplicaciones científicas paralelas (PARSEC), además de programas Java multihilos obtenidas de los benchmarks DaCapo y SPEC Java.
- Sistema referencia: SMT desactivado
- Usando media armónica sin pesos, con SMT activado, el **speed-up es 1,31** para los programas benchmarks PARSEC y **1,28** para Java.
- En promedio, el aumento en la **eficiencia energética** (energíaSMToff/energíaSMTon) es **1,07** para PARSEC y **0,99** para Java.

Speedup versus Energy Efficiency



$$H = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_N}}$$

Media armónica

The speedup from using multithreading on one core on an i7 processor averages 1.28 for the Java benchmarks and 1.31 for the PARSEC benchmarks. The energy efficiency averages 0.99 and 1.07, respectively. Anything above 1.0 for energy efficiency indicates that the feature reduces execution time by more than it increases average power.

Intel Core i7-920 (2008)

CoreFreq (software de monitorización de procesadores):
<https://github.com/cyring/CoreFreq>

```
Processor [Intel(R) Core(TM) i7 CPU 920 @ 2.67GHz] 8/8 CPU Online
Architecture [Nehalem/Bloomfield] Caches L1 Inst=32 Data=32 KB
Base Clock ~ 152 502 700 Hz L2=256 L3=8192 KB
CPU Ratio -----12-----20-21-22

#0 |
#1 |
#2 |
#3 |
#4 |
#5 |
#6 |
#7 |

Frequency - Ratio - Turbo -- C0 --- C1 --- C3 --- C6 --- C7 --- Temps --
#0 81.57 MHz ( 0.53) 2.67% 4.38% 1.95% 13.46% 80.20% 0.00% 45 C
#1 55.05 MHz ( 0.36) 1.80% 2.84% 43.76% 0.97% 52.43% 0.00% 43 C
#2 16.97 MHz ( 0.11) 0.56% 0.91% 1.60% 1.97% 95.52% 0.00% 44 C
#3 16.00 MHz ( 0.10) 0.52% 0.86% 1.59% 1.55% 96.00% 0.00% 43 C
#4 29.17 MHz ( 0.19) 0.96% 1.58% 4.76% 13.46% 80.20% 0.00% 45 C
#5 12.68 MHz ( 0.08) 0.42% 0.67% 45.90% 1.00% 52.43% 0.00% 43 C
#6 9.16 MHz ( 0.06) 0.30% 0.49% 2.02% 1.97% 95.52% 0.00% 44 C
#7 14.61 MHz ( 0.10) 0.48% 0.79% 1.66% 1.55% 96.01% 0.00% 43 C
Averages [ 0.96% 1.57% 12.90% 4.49% 81.04% 0.00% ]
Tech [TSC-INV, HT, TURBO, EIST, C1E, PM3, C3A, C1A, C3U, C1U, TM1, TM2]
Linux [4.7.4-1-ARCH] Tasks [ 335] Mem [ 9828220/12298568 KB]
```

CPU Ratio: multiplicador (0..22) de la frecuencia base del procesador (Base clock: 152,5 MHz)

