

APRENDIZAJE BASADO EN PROYECTOS: INTRODUCCIÓN A LOS SISTEMAS DE CÓMPUTO PARALELO

1.- Pregunta motriz: ¿Para cuándo estará todo el patrimonio digitalizado?



Fotografías: Lazkaoko Beneditarren Fundazioa

2.- Escenario del proyecto

El Centro de Documentación de los Benedictinos de Lazkao, lleva varios años recopilando documentación de temática vasca y archivándola en el denominado Fondo de los Benedictinos. Parte de este patrimonio está digitalizado y se puede acceder desde Internet (<http://www.lbf.eus/es/>).

El trabajo de digitalización se está llevando a cabo durante los últimos años en orden cronológico inverso, y ahora toca digitalizar los documentos más antiguos de los que se dispone. El proceso consiste en escanear las páginas de los libros a digitalizar. Para mejorar la calidad de los documentos y reducir en parte el “ruido” que introduce el escáner, las páginas son filtradas, es decir, limpiadas y archivadas en una plataforma on-line para que puedan ser accedidas desde Internet. Los documentos se guardan cifrados para mantener la seguridad de la información. Para ello, antes de subirlas a la plataforma hay que cifrarlas.

El Centro de Documentación de los Benedictinos quiere finalizar con el proceso de digitalización en un tiempo limitado. Por ello, han comprado un multiprocesador de 48 procesadores para poder procesar la información en paralelo y así acabar antes. Es necesario por tanto programar un algoritmo paralelo que sea capaz de procesar las páginas de los documentos en el menor tiempo posible, o lo que es lo mismo, que pueda tratar la mayor cantidad de páginas por minuto. Se quieren procesar 10 millones de páginas en total, con un tamaño medio de 20 MB.

El Centro de Documentación te ha pedido que paralelices el proceso del modo más eficiente posible.

La Figura 1 (fuente: https://ca.wikipedia.org/wiki/Ferran_el_Cat%C3%B2lic) presenta un ejemplo del proceso que se quiere llevar a cabo. Página escaneada (Figura 1a), página filtrada (Figura 1b) y página cifrada (Figura 1c). Por otra parte, en la Tabla 1 se muestra el factor de aceleración que se logra en función del número de procesadores. El tamaño de la página escaneada es de 11 MB.

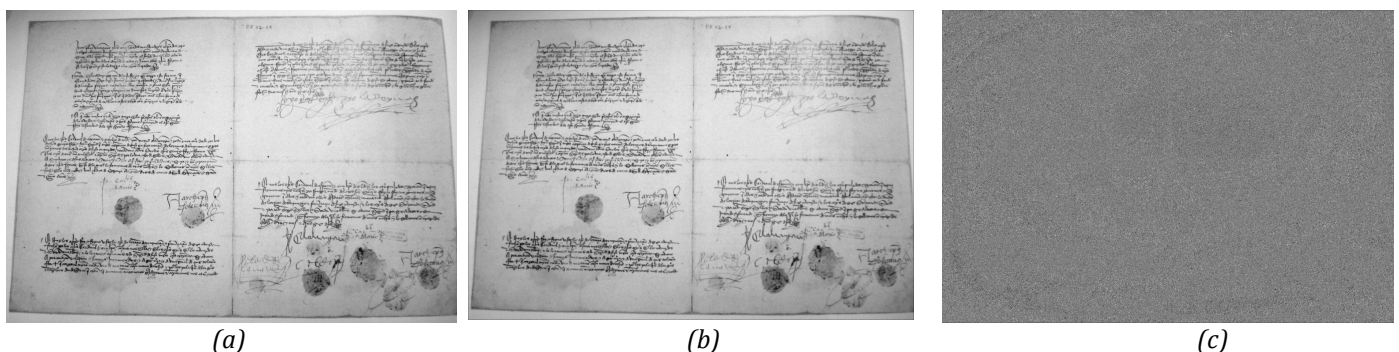


Figura 1. Procesado de una página de 11 MB: (a) pág. escaneada, (b) pág. filtrada y (c) pág. cifrada.

	Tiempo de ejecución (ms)	<i>speed-up</i>
1 proceso (serie)	7167	1
2 procesos	4450	1,6
4 procesos	2360	3,0
8 procesos	1302	5,5
16 procesos	795	9,0
32 procesos	574	12,5

Tabla 1. Factor de aceleración en función del número de procesadores.

3.- Resultados de aprendizaje

Al término de esta tarea el alumnado deberá ser capaz de:

1. Comprender los conceptos de paralelismo de bajo y alto nivel
2. Analizar las dependencias de una aplicación para establecer una estrategia de paralelización de la misma
3. Utilizar los mecanismos de sincronización necesarios para el correcto funcionamiento de los programas paralelos
4. Planificar un reparto de carga equilibrado entre los procesos
5. Analizar el rendimiento de la implementación paralela de una aplicación
6. Utilizar las directivas de OpenMP para programar una solución paralela

Así mismo, se trabajan las competencias generales C4, C8, C9 y las competencias RI1 y RI9 de la rama común informática. Estas competencias pueden consultarse en el documento:

<http://www.ehu.es/documents/340468/516505/Lista+de+competencias.pdf>

4.- Entregables del grupo de trabajo

- Entregable E1** Acta de constitución del grupo y documento de compromisos de los componentes del grupo, y las sucesivas actas de reuniones del grupo para el desarrollo del proyecto.
- Entregable E2** Bloque de entregables relacionados con el puzle a realizar al comienzo del proyecto: el informe que cada estudiante elabore del apartado que le corresponda (E2.1), su presentación (E2.2) y los ejercicios resueltos correspondientes (E2.3). En este entregable se incorporarán también aquellos materiales novedosos (no referenciados por el profesor o profesora) que el grupo haya utilizado para el desarrollo del puzle.
- Entregable E3** Desarrollo software realizado para la resolución del escenario planteado: informe técnico (E3.1) que explique la forma en la que se ha resuelto el escenario planteado y material utilizado para la presentación del desarrollo realizado (E3.2).

5.- Sistema de evaluación

El peso del proyecto en la calificación final de la asignatura es **4 puntos**. Se evaluará, de la forma que indica la tabla, utilizando los siguientes mecanismos:

- En el proyecto se contemplan dos exposiciones individuales: la primera correspondiente a los distintos temas trabajados en el puzle inicial y la segunda correspondiente a la defensa final de la solución planteada en el proyecto. En la evaluación de las exposiciones se considerará la evaluación entre pares. Esta exposición la realizará uno de los miembros del grupo, elegido al azar, por lo que la nota que obtenga el estudiante que realice la exposición será trasladada a los componentes del grupo. Esta actividad tienen una valoración de **0,75 puntos**.
- Dos exámenes de conocimientos mínimos: un primer examen para evaluar los conocimientos del Lenguaje C (**0,5 puntos**) y un segundo examen para evaluar los conocimientos sobre paralelismo (**1 punto**). **En este segundo caso, el estudiante deberá obtener al menos un 30% de la valoración del examen.**
- El desarrollo software realizado, junto con el informe técnico que explica la solución adoptada. Esta actividad tiene una valoración de **1,75 puntos**.

Cada grupo deberá entregar además el informe del puzle (entregable E2), que será revisado por el profesor y devuelto al grupo de trabajo, con las indicaciones de mejora pertinentes. Esta actividad es de tipo FILTRO, es decir, obligatoria en el proyecto pero sin carácter sumatorio sobre la calificación final.

En el sistema de evaluación se contempla un sistema de “extra bonus”, esto es, independientemente de la actividad que se evalúe según el sistema de evaluación propuesto, se premiarán, con puntos adicionales sobre la calificación final, las actividades que de alguna forma sean sobresalientes.

	Evaluador		Nota
	Profesor	Estudiantes	
Individual	Examen conocimientos C	0,5 p.	1,5 p.
	Examen Paralelismo	1 p.	
Grupo	Presentaciones	0,4 p.	2,5 p.
	Informe técnico/desarrollo	1,75 p.	
	Informe puzle	FILTRO	
Nota	3,65 puntos		4 puntos

6. Planificación del trabajo del alumnado

Semana	Clase	Actividades presenciales	t/sesión	Actividades no presenciales (semana)	t/semana	Entregables	Evaluación
17/10	0.1	Prácticas de laboratorio: Lenguaje C	4 h	Estudio de programación en C	4 h		
31/10	0.2	Examen de conocimientos mínimos	45 m	Ejercicios de seguimiento de C			
5-9/11	1	Pregunta motriz + escenario Discusión ABP: póster + debate Exposición de conceptos básicos + planificación	15 m 45 m 30 m	Asimilación del planteamiento del proyecto y repaso de conceptos básicos	2 h	E1: acta de constitución del grupo	
	2	Prácticas de laboratorio: introducción a OpenMP	90 m	Estudio de la programación en OpenMP			
	3	Puzzle: presentación y reparto de tareas Puzzle: estudio individual de la tarea asignada	30 m 60 m	Estudio de la tarea asignada en el puzzle			
12-16/11	4	Puzzle: estudio individual de la tarea asignada	90 m	Estudio de la tarea asignada en el puzzle	8 h		
	5	Puzzle: reunión de expertos	90 m	Revisión del puzzle. Informe y presentación del puzzle			
	6	Puzzle: reunión de grupo	90 m	Elaboración del informe del puzzle y de su presentación			
26-30/11	7	Presentación de la aplicación (algoritmos) Trabajo en equipo: puzzle / desarrollo de la aplicación	90 m	Puzzle: informe, presentación y ejercicios	4 h		
	8	Puzzle: presentación Puzzle: debate aclaratorio	60 m 30 m				[0,75 p.]
	9	Puzzle: resolución de ejercicios	90 m	Aplicación: desarrollo, documentación, conocimiento		E2: informe puzzle	FILTRO
5/12	10	Trabajo en equipo: aplicación (desarrollo, documentación, adquisición de conocimiento)	90 m	Aplicación: desarrollo, documentación, conocimiento	4 h		
10-14/12	11	Trabajo en equipo: aplicación	90 m		6 h		
	12	Trabajo en equipo: aplicación	90 m	Aplicación: desarrollo, documentación, conocimiento			
	13	Trabajo en equipo: aplicación	90 m				
17-21/12		Semana de finalización de trabajos (entrega: 21/12)		Preparación de la presentación	6 h	E3.1: informe técnico + software	1,75 p.
8/1	14	Presentación del proyecto Examen de conocimientos mínimos	90 m 90 m		2 h	E3.2: material presentación	[0,75 p.] 1 p. (min.: 0,3)