UNIVERSIDADE DA CORUÑA

Guia docente					
Datos Identificativos					2021/22
Asignatura (*)	Procesamiento Paralelo			Código	614G02023
Titulación	Grao en Ciencia e Enxeñaría de Datos				
Descriptores					
Ciclo	Periodo	Curso		Tipo	Créditos
Grado	1º cuatrimestre	Tercero		Obligatoria	6
Idioma	Castellano				
Modalidad docente	Presencial				
Prerrequisitos					
Departamento	Enxeñaría de Computadores				
Coordinador/a	Enes Álvarez, Jonatan		Correo electrónico	jonatan.enes@udc.es	
Profesorado	Enes Álvarez, Jonatan Touriño Dominguez, Juan		Correo electrónico	jonatan.enes@udc.es juan.tourino@udc.es	
Web					
Descripción general	En esta asignatura el alumno aprenderá el papel que el uso del paralelismo tiene a la hora de acelerar la ejecución de programas en general y el procesamiento de datos en particular. El conocimiento teórico partirá de los conceptos más básicos de paralelismo, incluyendo su utilidad y aplicabilidad, el contexto técnico de la ejecución de programas paralelos, las diferentes técnicas existentes o las diferentes métricas para medir el paralelismo (Bloque I). A continuación se seguirá con un enfoque más dirigido al tratamiento de datos (Bloque II). Finalmente se acabará con una introducción al procesamiento paralelo de datos con las tecnologías más modernas basadas en el Big Data (Bloque III). En el aspecto práctico, el alumno realizará diversas sesiones con un enfoque incremental a fin de adquirir el conocimiento y la habilidad de programar y/o desplegar soluciones de procesamiento en paralelo de datos. Se empezará con prácticas dirigidas a adquirir competencias más técnicas y de 'bajo nivel', y se irá progresando hacia soluciones más completas y de más 'alto nivel'. Dichas prácticas se realizarán de forma coordinada con la explicación teórica, de tal forma que se usarán técnicas y tecnologías previamente explicadas. De igual forma, las prácticas serán autocontenidas y fuertemente enfocadas a la resolución de problemas o escenarios concretos con técnicas y tecnologías específicas. Esta asignatura tiene una fuerte dependencia con asignaturas previas como "Fundamentos de Programación I y II", principalmente por su requisito técnico para la programación y de "Diseño y Análisis de Algoritmos" por su conocimiento teórico del análisis de complejidad de los algoritmos. En menor medida se requieren conocimientos previos de la asignatura de "Fundamentos de Computadores" para comprender el comportamiento empírico y el rendimiento en general de algunos programas.				

Plan de contingencia	1. Modificaciones en los contenidos: + Sin modificaciones en los contenidos. 2. Metodologías: + Si fuese necesario, las sesiones magistrales y de prácticas pueden pasar a ser virtuales mediante el uso de Teams. De igual forma, si fuera preciso, la prueba objetiva (examen final) se realizaría de forma telemática mediante el uso de varios recursos de Moodle. 3. Mecanismos de atención personalizada al alumnado: + El alumnado podrá hacer uso de canales virtuales y telemáticos para interactuar con el docente, principalmente, videoconferencia por Teams y correo. 4. Modificaciones en la evaluación: + Se mantienen las ponderaciones para la evaluación. 5. Modificaciones de la bibliografía o webgrafía: + Ninguna

Competencias / Resultados del título	
Código	**Competencias / Resultados del título**
A12	CE12 - Capacidad de conocer y aplicar los principios fundamentales, principales paradigmas y técnicas de la programación paralela y distribuida al desarrollo de algoritmos para el procesamiento y análisis masivo de datos.
B2	CB2 - Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio
B3	CB3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética
B4	CB4 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado
B7	CG2 - Elaborar adecuadamente y con cierta originalidad composiciones escritas o argumentos motivados, redactar planes, proyectos de trabajo, artículos científicos y formular hipótesis razonables.
B8	CG3 - Ser capaz de mantener y extender planteamientos teóricos fundados para permitir la introducción y explotación de tecnologías nuevas y avanzadas en el campo.
B9	CG4 - Capacidad para abordar con éxito todas las etapas de un proyecto de análisis de datos: exploración previa de los datos, preprocesado, análisis, visualización y comunicación de resultados.
B10	CG5 - Ser capaz de trabajar en equipo, especialmente de carácter multidisciplinar, y ser hábiles en la gestión del tiempo, personas y toma de decisiones.
C1	CT1 - Utilizar las herramientas básicas de las tecnologías de la información y las comunicaciones (TIC) necesarias para el ejercicio de su profesión y para el aprendizaje a lo largo de su vida.
C4	CT4 - Valorar la importancia que tiene la investigación, la innovación y el desarrollo tecnológico en el avance socioeconómico y cultural de la sociedad.

Resultados de aprendizaje			
Resultados de aprendizaje	**Competencias / Resultados del título**		
Conocer los requisitos técnicos y las tecnologías actuales que permiten el paralelismo.	A12	B8 B9	
Conocer las diferentes tecnologías de paralelismo existentes, su aplicación, límites, ventajas y desventajas.	A12	B2 B4 B8 B9	

Ser capaz de aplicar técnicas de paralelismo para transformar una solución existente y así permitir un procesamiento paralelo.	A12	B2 B4 B7 B8 B9 B10	C1
Ser capaz de analizar el rendimiento de una solución de procesamiento, existente sin paralelismo o nueva con paralelismo	A12	B2 B4 B7 B8 B9 B10	C1
Comprender la importancia del paralelismo a la hora de llevar a cabo procesos de análisis de datos críticos en el ámbito comercial o de investigación.	A12	B3 B4 B8 B10	C4

Contenidos	
Tema	**Subtema**
BLOQUE I - Conceptos básicos de Paralelismo	Tema 1 - Introducción y conceptos previos Tema 2 ? Paralelismo general
BLOQUE II - Tratamiento de Datos y Paralelismo	Tema 3 ? Numpy, Pandas y programación funcional
BLOQUE III - Paralelismo con Big Data	Tema 4 ? Paralelismo con MapReduce y Spark Tema 5 ? Otras tecnologías de procesamiento paralelo o auxiliares

Planificación				
Metodologías / pruebas	**Competencias / Resultados**	**Horas lectivas (presenciales y virtuales)**	**Horas trabajo autónomo**	**Horas totales**
Sesión magistral	A12 B3 B8 B9 C4	20	30	50
Prácticas de laboratorio	A12 B2 B4 B7 B9 B10 C1	20	70	90
Prueba objetiva	A12 B2 B4 B7 B9 C1 C4	3	1	4
Atención personalizada		6	0	6
(*)Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos				

Metodologías	
Metodologías	**Descripción**
Sesión magistral	* Las clases de sesión magistral introducirán la base teórica sobre la cual se basarán las prácticas que el alumnado realizará posteriormente, intentando en todo momento señalar esta conexión. * De igual forma, las sesiones magistrales explicarán y detallarán otros conceptos importantes relacionados con el paralelismo en general, ya sean previos o fundamentales a fin de entender el funcionamiento de las tecnologías usadas en las sesiones de prácticas, o bien más avanzados y que explican su uso y papel crítico en procesos relevantes para la sociedad.

Prácticas de laboratorio	* Las prácticas de laboratorio serán autocontenidas y estarán enfocadas a escenarios muy concretos o a la resolución de problemas muy comunes, todo mediante el uso de técnicas, paradigmas y tecnologías previamente explicadas en las sesiones magistrales. * Cada práctica, que puede ocupar una o varias sesiones, se centrará en un problema para el cual el alumno dispondrá de una explicación previa, un código ya proporcionado a utilizar y una serie de preguntas sobre las cuales tendrá que desarrollar la práctica en su trabajo en clase y autónomo. Las preguntas pueden pedir realizar una extensión o modificación del código, hacer un estudio empírico del rendimiento de este con distintas configuraciones de paralelismo, describir su funcionamiento o comportamiento, u otras preguntas enfocadas en general a que el alumno comprenda el problema y la solución de paralelismo aplicada. * Es posible que para algunas prácticas se realice un cuestionario evaluable muy breve una vez que todos los alumnos hayan entregado la práctica. _____
Prueba objetiva	* Al final del cuatrimestre se realizará un examen de evaluación de toda la asignatura, donde se incluirán principalmente preguntas de desarrollo de los conceptos y materia en general explicada en las sesiones magistrales. * De igual forma también se pueden incluir en menor medida preguntas breves específicamente enfocadas a evaluar la adquisición de los conocimientos clave de las prácticas de laboratorio.

Atención personalizada	
Metodologías	**Descripción**
Sesión magistral Prácticas de laboratorio	* La atención personalizada se destinará para el apoyo de forma general al alumnado a lo largo de la asignatura y se extenderá sobre todo el temario que se exponga en esta. * Por un lado dicha atención estará disponible para la comprensión/ampliación de conceptos explicados en las sesiones magistrales de teoría, a fin de que ninguna alumna o alumno tenga dificultades en la adquisición del conocimiento más teórico y que será sujeto de evaluación en la prueba final. Esto es especialmente importante en las fases tempranas de aquel temario que sea incremental, o en los momentos de 'cambio' de temario (e.g., cambio del Bloque I al Bloque II). * Y por otro lado la atención personalizada también estará disponible al alumnado que requiera una atención más específica para la resolución de incidentes o dudas originadas en las prácticas de laboratorio, ya sean problemas técnicos para realizar dichas prácticas, como dificultades para comprender algún concepto clave de ellas. Esta atención estará disponible a lo largo de todo el curso, aunque se recomienda dirigir cualquier duda de una práctica a lo largo de la duración de esta o pocos días después de su finalización. _____ Los alumnos con dispensa académica pueden pedir, mediante atención personalizada, una explicación de una práctica en las mismas condiciones que en las sesiones de prácticas.

Evaluación			
Metodologías	**Competencias / Resultados**	**Descripción**	**Calificación**

Prácticas de laboratorio	A12 B2 B4 B7 B9 B10 C1	* Las prácticas serán sujetas a evaluación por parte del profesor. * Cada práctica será presentada y brevemente explicada por el profesor en una sesión de prácticas, donde además el alumnado iniciará su realización. * La entrega de la práctica se realizará en una fecha previamente acordada, hasta la cual se espera que el alumnado finalice la práctica en su tiempo de trabajo autónomo asociado. Cada grupo tendrá su fecha de entrega. * Es posible que para algunas prácticas la evaluación recaiga total o parcialmente en la contestación de un cuestionario una vez finalizada la práctica y en una fecha preacordada con el alumnado. _____	50
Prueba objetiva	A12 B2 B4 B7 B9 C1 C4	* Examen escrito realizado al final del cuatrimestre. * Se evaluarán principalmente conceptos teóricos explicados en las sesiones magistrales. * En menor medida también se incluirán preguntas sobre conceptos clave de las sesiones prácticas.	50

Observaciones evaluación
Para aprobar la asignatura será necesario obtener un mínimo del 40% de la nota máxima de la prueba objetiva (2 puntos de 5).La nota de prácticas NO ES RECUPERABLE para la segunda oportunidad.Los alumnos con matrícula a tiempo parcial podrán asistir (previo aviso) a un grupo de prácticas distinto al que tengan asignado.Los alumnos con matrícula a tiempo parcial o que dispongan de dispensa académica podrán realizar su entrega en un plazo más tardío, o en el caso de que se evalúe una práctica mediante cuestionario, podrán preacordar con el profesor una fecha distinta si es necesario.

Fuentes de información	
Básica	- Giancarlo Zaccone (2015). Python Parallel Programming Cookbook. Packt Publishing - Wes McKinney (2011). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly - Tomasz Drabas, Denny Lee (2017). Learning PySpark: Build data-intensive applications locally and deploy at scale using the combined powers of Python and Spark 2.0. Packt Publishing
Complementária	- Bertil Schmidt et al. (2017). Parallel programming : concepts and practic. Cambridge, MA : Morgan Kaufmann - Peter S.Pacheco (2011). An introduction to parallel programming. Burlington, MA : Morgan Kaufmann - Francisco Almeida et al. (2008). Introducción a la programación paralela. Madrid : Paraninfo Cengage Learning - Jesús Carretero Pérez et al. (2007). Sistemas operativos : una visión aplicada. Madrid : McGraw-Hill

Recomendaciones
Asignaturas que se recomienda haber cursado previamente
Diseño y Análisis de Algoritmos/614G02011 Fundamentos de Computadores/614G02005 Fundamentos de Programación II/614G02009 Fundamentos de Programación I/614G02004
Asignaturas que se recomienda cursar simultáneamente
Asignaturas que continúan el temario
Procesamiento Paralelo Avanzado/614G02034
Otros comentarios

Se recomienda tener unas competencias básicas/intermedias en la programación con el lenguaje Python dado que todas las prácticas se realizarán con este lenguaje.Se recomienda tener unos conocimientos básicos en el uso de un sistema Linux, gestión de procesos y conocimiento de ficheros y entrada/salida.

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías