



Guía Docente

Guía Docente				
Datos Identificativos				2022/23
Asignatura (*)	Procesamento Paralelo		Código	614G02023
Titulación	Grao en Ciencia e Enxeñaría de Datos			
Descriptor				
Ciclo	Período	Curso	Tipo	Créditos
Grao	1º cuatrimestre	Terceiro	Obrigatoria	6
Idioma	Castelán			
Modalidade docente	Presencial			
Prerrequisitos				
Departamento	Enxeñaría de Computadores			
Coordinación	Enes Álvarez, Jonatan	Correo electrónico	jonatan.enes@udc.es	
Profesorado	Enes Álvarez, Jonatan	Correo electrónico	jonatan.enes@udc.es	
Web				
Descrición xeral	Nesta materia o alumno estudará o papel que o uso do paralelismo ten á hora de acelerar a execución de programas en xeral, e o procesamento de datos en particular. O coñecemento teórico partirá dos conceptos técnicos máis básicos de paralelismo, incluíndo a súa utilidade, aplicabilidade e o contexto técnico da execución de programas paralelos (Tema 1). Despois expoñeranse as principais tecnoloxías hardware que existen actualmente para o procesamento paralelo, así como as súas capacidades técnicas subxacentes relacionadas e necesarias para explotar o paralelismo (Tema 2). A continuación, profundizarase no paralelismo con conceptos engadidos, clasificacións, posibles deseños para a súa implementación software e maneiras de analizar o seu rendemento (Tema 3). Finalmente, aplicarase todo o coñecemento previamente adquirido estudando as ferramentas e tecnoloxías modernas para o procesamento de datos masivos, é dicir, poñendo o foco no Big Data (Tema 4). No aspecto práctico, o alumno realizará diversas sesións cun enfoque incremental a fin de adquirir o coñecemento e a habilidade de programar e/ou despregar solucións de procesamento en paralelo de datos. Empezarase con prácticas dirixidas a adquirir competencias máis técnicas e de 'baixo nivel', e irase progresando cara a solucións máis completas e de máis 'alto nivel'. Estas prácticas realizaranse de forma coordinada coa explicación teórica, de tal forma que se usarán técnicas e tecnoloxías previamente explicadas. De igual forma, as prácticas serán autocontidas e fortemente enfocadas á resolución de problemas ou escenarios concretos. Esta materia ten unha forte dependencia con materias previas como "Fundamentos de Programación I e II", principalmente polo seu requisito técnico para a programación, e "Deseño e Análise de Algoritmos" polo seu coñecemento teórico da análise de complexidade dos algoritmos. En menor medida requírense coñecementos previos da materia de "Fundamentos de Computadores" para comprender o comportamento empírico e o rendemento en xeral dalgúns programas cando estes execútanse nun computador.			

Competencias do título

Código	Competencias do título
A12	CE12 - Capacidade de coñecer e aplicar os principios fundamentais, principais paradigmas e técnicas da programación paralela e distribuída ao desenvolvemento de algoritmos para o procesamento e análise masiva de datos.
B2	CB2 - Que os estudantes saiban aplicar os seus coñecementos ao seu traballo ou vocación dunha forma profesional e posúan as competencias que adoitan demostrarse por medio da elaboración e defensa de argumentos e a resolución de problemas dentro da súa área de estudo
B3	CB3 - Que os estudantes teñan a capacidade de reunir e interpretar datos relevantes (normalmente dentro da súa área de estudo) para emitir xuízos que inclúan unha reflexión sobre temas relevantes de índole social, científica ou ética
B4	CB4 - Que os estudantes poidan transmitir información, ideas, problemas e solucións a un público tanto especializado como non especializado
B7	CG2 - Elaborar adecuadamente e con certa orixinalidade composicións escritas ou argumentos motivados, redactar plans, proxectos de traballo, artigos científicos e formular hipóteses razoables.
B8	CG3 - Ser capaz de manter e estender formulacións teóricas fundadas para permitir a introdución e explotación de tecnoloxías novas e avanzadas no campo.



B9	CG4 - Capacidade para abordar con éxito todas as etapas dun proxecto de datos: exploración previa dos datos, preprocesado, análise, visualización e comunicación de resultados.
B10	CG5 - Ser capaz de traballar en equipo, especialmente de carácter multidisciplinar, e ser hábiles na xestión do tempo, persoas e toma de decisións.
C1	CT1 - Utilizar as ferramentas básicas das tecnoloxías da información e as comunicacións (TIC) necesarias para o exercicio da súa profesión e para a aprendizaxe ao longo da súa vida.
C4	CT4 - Valorar a importancia que ten a investigación, a innovación e o desenvolvemento tecnolóxico no avance socioeconómico e cultural da sociedade.

Resultados da aprendizaxe			
Resultados de aprendizaxe		Competencias do título	
Coñecer os requisitos técnicos e as tecnoloxías actuais que permiten o paralelismo.		A12	B8 B9
Coñecer as diferentes tecnoloxías de paralelismo existentes, a súa aplicación, límites, vantaxes e desvantaxes.		A12	B2 B4 B8 B9
Ser capaz de aplicar técnicas de paralelismo para transformar unha solución existente e así permitir un procesamento paralelo.		A12	B2 B4 B7 B8 B9 B10
Ser capaz de analizar o rendemento dunha solución de procesamento, existente sen paralelismo ou nova con paralelismo		A12	B2 B4 B7 B8 B9 B10
Comprender a importancia do paralelismo á hora de levar a cabo procesos de análise de datos críticos no ámbito comercial ou de investigación.		A12	B3 B4 B8 B10

Contidos	
Temas	Subtemas
Tema 1 - Introducción e conceptos previos	* Programa secuencial e Proceso * Anatomía e ciclo de vida dun proceso * Fíos * Programa paralelo * Utilidades do paralelismo
Tema 2 - Paralelismo en Hardware, xerarquía do paralelismo	* Niveis de paralelismo * Paralelismo interno do procesador (agochado) * Funcionalidades do procesador (paralelismo baixo nivel) * Recursos accesibles do procesador (paralelismo alto nivel) * Conxunto de máquinas (Cluster e Supercomputador) * Redes distribuídas de Computación * Dispositivos específicos * Estado del arte de los procesadores



Tema 3 - Paralelismo en Software, diseño e implementación	* Clasificación do paralelismo (Taxonomía de Flynn) * Linguaxes e frameworks do paralelismo * Conceptos chave * Paradigmas do procesamento paralelo * Análise de programas paralelos * Deseño de programas paralelos
Tema 4 - Tecnoloxías Big Data	* Almacenamiento de datos * Xestión de recursos e a execución * Procesamento Batch * Procesamento Streaming

Planificación				
Metodoloxías / probas	Competencias	Horas presenciais	Horas non presenciais / traballo autónomo	Horas totais
Sesión maxistral	A12 B3 B8 B9 C4	20	30	50
Prácticas de laboratorio	A12 B2 B4 B7 B9 B10 C1	20	70	90
Proba obxectiva	A12 B2 B4 B7 B9 C1 C4	3	1	4
Atención personalizada		6	0	6
*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado				

Metodoloxías	
Metodoloxías	Descrición
Sesión maxistral	* As clases de sesión maxistral introducirán a base teórica sobre a cal se basearán as prácticas que o alumnado realizará posteriormente, tentando en todo momento sinalar esta conexión. * De igual forma, as sesións maxistrais explicarán e detallarán outros conceptos importantes relacionados co paralelismo en xeral, xa sexan previos ou fundamentais a fin de entender o funcionamento das tecnoloxías usadas nas sesións de prácticas, ou ben máis avanzados e que explican o seu uso e papel crítico en procesos relevantes para a sociedade.
Prácticas de laboratorio	* As prácticas de laboratorio serán autocontidas e estarán enfocadas a escenarios moi concretos ou á resolución de problemas moi comúns, todo mediante o uso de técnicas, paradigmas e tecnoloxías previamente explicadas nas sesións maxistrais. * Cada práctica, que pode ocupar unha ou varias sesións, centrarase nun problema para o cal o alumno dispoñerá dunha explicación previa, un código xa proporcionado a utilizar e unha serie de preguntas sobre as cales terá que desenvolver a práctica no seu traballo en clase e autónomo. As preguntas poden pedir realizar unha extensión ou modificación do código, facer un estudo empírico do rendemento con distintas configuracións de paralelismo, describir o seu funcionamento ou comportamento, ou outras preguntas enfocadas en xeral a que o alumno comprenda o problema e a solución de paralelismo aplicada. * É posible que para algunhas prácticas realícese un cuestionario avaliable moi breve unha vez que todos os alumnos entregasen a práctica.



Proba obxectiva	* Ao final do cuadrimestre realizarase un exame de avaliación de toda a materia, onde se incluírán principalmente preguntas de desenvolvemento dos conceptos e materia en xeral explicada nas sesións maxistrais. * De igual forma tamén se poden incluír en menor medida preguntas breves especificamente enfocadas a avaliar a adquisición dos coñecementos chave das prácticas de laboratorio.
-----------------	---

Atención personalizada	
Metodoloxías	Descrición
Sesión maxistral Prácticas de laboratorio	* A atención personalizada destinarase ao apoio de forma xeral do alumnado ao longo da materia e estenderase sobre todo o temario que se expoña. * Por unha banda dita atención estará dispoñible para a comprensión/ampliación de conceptos explicados nas sesións maxistrais de teoría, co fin de que ningunha alumna ou alumno teña dificultades na adquisición do coñecemento máis teórico e que será suxeito de avaliación na proba final. Isto é especialmente importante nas fases temperás daquel temario que sexa incremental, ou nos momentos de 'cambio' de temario (e. g., cambio do Bloque I ao Bloque II). * Por outra banda a atención personalizada tamén estará dispoñible de cara ao alumnado que requira unha atención máis específica para a resolución de incidentes ou dúbidas orixinadas nas prácticas de laboratorio, xa sexan problemas técnicos para realizar ditas prácticas, como dificultades para comprender algún concepto chave delas. Esta atención estará dispoñible ao longo de todo o curso, aínda que se recomenda dirixir calquera dúbida dunha práctica ao longo da duración desta ou poucos días despois da súa finalización. * Os alumnos con dispensa académica poden pedir, mediante atención personalizada, unha explicación dunha práctica nas mesmas condicións que nas sesións de prácticas ordinarias.

Avaliación			
Metodoloxías	Competencias	Descrición	Cualificación
Prácticas de laboratorio	A12 B2 B4 B7 B9 B10 C1	* Todas as prácticas estarán suxeitas á avaliación por parte do profesor. * Cada práctica será presentada e brevemente explicada polo profesor nunha sesión de prácticas, onde ademais o alumnado iniciará a súa realización. * A entrega da práctica realizarase nunha data acordada. Cada grupo terá a súa data de entrega. * É posible que para algunhas prácticas a avaliación recaia total ou parcialmente na contestación dun cuestionario unha vez entregada a práctica e nunha data previamente comunicada.	50
Proba obxectiva	A12 B2 B4 B7 B9 C1 C4	* Exame escrito realizado ao final do cuadrimestre. * Avaliaranse principalmente conceptos teóricos explicados nas sesións maxistrais. * En menor medida tamén se incluírán preguntas sobre conceptos chave das sesións prácticas.	50

Observacións avaliación



Para aprobar a materia será necesario obter un mínimo do 40% da nota máxima da proba obxectiva (2 puntos de 5). A nota de prácticas NON É RECUPERABLE para a segunda oportunidade. Os alumnos con matrícula a tempo parcial poderán asistir (previo aviso) a un grupo de prácticas distinto ao que teñan asignado. Os alumnos con matrícula a tempo parcial ou que dispoñan de dispensa académica poderán realizar a súa entrega máis tarde, ou no caso de que se avalíe unha práctica mediante cuestionario, poderán preacordar co profesor unha data distinta se é necesario.

Fontes de información

Bibliografía básica	- 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 - Giancarlo Zaccone (2015). Python Parallel Programming Cookbook. Packt Publishing - 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
Bibliografía complementaria	- 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 - Wes McKinney (2011). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly

Recomendacións

Materias que se recomenda ter cursado previamente

Deseño e Análise de Algoritmos/614G02011
Fundamentos de Computadores/614G02005
Fundamentos de Programación II/614G02009
Fundamentos de Programación I/614G02004

Materias que se recomenda cursar simultaneamente

Materias que continúan o temario

Procesamento Paralelo Avanzado/614G02034

Observacións

Recoméndase ter unhas competencias básicas/intermedias na programación coa linguaxe Python dado que todas as prácticas realizaranse con esta linguaxe. Recoméndase ter uns coñecementos básicos no uso dun sistema Linux, xestión de procesos e coñecemento de ficheiros e entrada/saída.

(*)A Guía docente é o documento onde se visualiza a proposta académica da UDC. Este documento é público e non se pode modificar, salvo casos excepcionais baixo a revisión do órgano competente dacordo coa normativa vixente que establece o proceso de elaboración de guías