



Guía Docente				
Datos Identificativos				2013/14
Asignatura (*)	Computación en Sistemas Distribuídos		Código	614473009
Titulación				
Descriptor				
Ciclo	Período	Curso	Tipo	Créditos
Mestrado Oficial	1º cuatrimestre	Primeiro	Obrigatoria	3
Idioma	CastelánGalegoInglés			
Prerrequisitos				
Departamento	Electrónica e Sistemas			
Coordinación	Pardo Martínez, Xoán Carlos		Correo electrónico	xoan.pardo@udc.es
Profesorado	Pardo Martínez, Xoán Carlos		Correo electrónico	xoan.pardo@udc.es
Web				
Descrición xeral	Desde hai varios anos, o uso de arquitecturas de computación paralelas foi un aspecto fundamental que permitiu o desenvolvemento de importantes áreas en múltiples campos da ciencia básica e aplicada. Con todo, o elevado custo dos sistemas paralelos tradicionais limitou o seu uso practicamente a grandes industrias e centros de investigación. Hai tempo que o uso de redes de computadores de baixo custo representa unha alternativa práctica e barata aos grandes sistemas. Como outras alternativas, as infraestruturas Grid, e máis recentemente os sistemas Cloud, aparecen como paradigmas de computación distribuída que cambian o modo no que usamos os computadores, permitindo o acceso transparente, seguro e barato a enormes recursos computacionais desde calquera lugar do mundo. O obxectivo principal desta materia é dar a coñecer estes dous novos paradigmas de computación distribuída, isto é, Grid e Cloud Computing, e instruír ao alumno na súa utilización no ámbito da computación de altas prestacións.			

Competencias da titulación	
Código	Competencias da titulación

Resultados da aprendizaxe				
Competencias de materia (Resultados de aprendizaxe)		Competencias da titulación		
Habilidade para a procura, selección e manexo de recursos (bibliografía, software, etc.) relacionados coa computación Grid e Cloud	AI5	BI4	CM1	
	AI7	BI7	CM2	
	AI12	BI10	CM3 CM6	
Coñecer e saber utilizar as ferramentas básicas a nivel de usuario dispoñibles en contornas Grid e Cloud	AI7	BI1		
	AI8	BI7		
		BI9		
Coñecer e saber utilizar algunha das ferramentas dispoñibles para preparar e executar aplicacións científicas en contornas Grid e Cloud	AI6	BI7		
	AI7	BI9		
	AI8			
Xestionar a seguridade nunha contorna Grid	AI8	BI1		
		BI9		
Instalar e configurar unha contorna Grid.	AI5	BI1		
	AI8	BI9		
Instalar e configurar unha contorna de virtualización	AI5	BI1	CM3	
	AI7	BI9		
	AI8			

Contidos	
Temas	Subtemas



1. Introducción ás arquitecturas distribuídas	a. Sistemas distribuídos b. Computación a grande escala: Grid e Cloud Computing
2. Computación Grid	i. Introducción * Descrición e compoñentes dun Grid * Exemplos de aplicacións e infraestruturas Grid: ESG, NeesGrid, LCG * Middleware, organizacións e estándares Grid: OGF, WSI, OGSA ii. Middleware Grid: Globus Toolkit * Conceptos básicos sobre Globus * Conceptos básicos de seguridade en contorna Grid: PKI, certificados X.509, proxies, etc. * Compoñentes de Globus: seguridade, execución, información, etc.
3. Computación Cloud	i. Introducción * Conceptos xerais * Modelos de servizo: SaaS, PaaS, IaaS * Modelos de despregue: Público, Privado, Híbrido * Beneficios, riscos e oportunidades * Casos de estudo ii. IaaS (Infraestrutura como servizo) * Oferta de provedores públicos * Tecnoloxías de virtualización * Estándares: OVF, OCCl, CDMI, ... * Middleware Open-Source para Clouds privados: OpenNebula, OpenStack, CloudStack, Eucalyptus, ... iii. Exemplo de IaaS * Servizos básicos proporcionados * Servizos específicos: clusters virtuais, servizos para HPC,...
4. Procesamento de grandes datos (Big Data) na nube	i. Concepto de Big Data ii. Desafíos do Big Data * Almacenamento do Big Data: bases de datos NoSQL * Procesamento do Big Data ii. O Big Data e a nube
5. Xestión do Big Data	i. Introducción ao modelo de programación MapReduce * Exemplos de aplicacións * Execución e optimizacións de aplicacións MapReduce * MapReduce na nube ii. Implementacións MapReduce: Hadoop * Introducción a Hadoop * Hadoop Distributed Filesystem * Aplicacións MapReduce con Hadoop * Instalación de un cluster Hadoop iii. Ecosistema Hadoop * Introducción ao ecosistema Apache Hadoop * Programación MapReduce de alto nivel: Pig, Hive * Base de datos sobre HDFS: HBase * Outras tecnoloxías: Zookeeper, HCatalog, Oozie, ...

Planificación			
Metodoloxías / probas	Horas presenciais	Horas non presenciais / traballo autónomo	Horas totais



Sesión maxistral	7	7	14
Traballos tutelados	1.5	20	21.5
Prácticas de laboratorio	15	22.5	37.5
Atención personalizada	2	0	2

*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado

Metodoloxías	
Metodoloxías	Descrición
Sesión maxistral	Nas que se exporá o contido teórico do temario incluíndo exemplos ilustrativos e co soporte de medios audiovisuais. O alumno disporá do material de apoio (apuntamentos, copias das transparencias, artigos, etc.) con anterioridade e o profesor promoverá unha actitude activa, recomendando a lectura previa dos puntos do temario a tratar en cada clase, así como realizando preguntas que permitan aclarar aspectos concretos e deixando cuestións abertas para a reflexión do alumno.
Traballos tutelados	Formulación de casos de estudo avanzados para a súa resolución individual por parte dos alumnos. Estes casos de estudo permitirán aos alumnos profundar en aspectos do temario que lles interesen especialmente e que non se puideron tratar en profundidade durante as sesións maxistrais.
Prácticas de laboratorio	Nas que o alumno verá o funcionamento na práctica dalgúns dos contidos teóricos vistos nas clases maxistrais. Nestas prácticas o alumno utilizará diferentes ferramentas propostas polo profesor que lle permitirán profundar e afianzar os seus coñecementos sobre diferentes aspectos da computación Grid e Cloud. As prácticas poderían realizarse en grupos en función da dispoñibilidade dos medios necesarios para a súa realización.

Atención personalizada	
Metodoloxías	Descrición
Prácticas de laboratorio Traballos tutelados Sesión maxistral	A atención personalizada durante as prácticas servirá para orientar e comprobar o traballo que os alumnos vaian realizando segundo as indicacións que se lles proporcionen, dependendo do traballo concreto. Para a realización dos traballos tutelados os profesores proporcionarán as indicacións iniciais necesarias, bibliografía para consulta e realizarán un seguimento dos avances que o alumno vaia realizando para ofrecer as orientacións pertinentes en cada caso, de modo que se asegure a calidade dos traballos de acordo aos criterios que se indiquen. Todos os profesores da materia proporán ademais un horario de titorías no que os alumnos poderán resolver calquera dúbida relacionada co desenvolvemento da mesma. Recomendarase aos alumnos a asistencia a titorías como parte fundamental do apoio á aprendizaxe.

Avaliación		
Metodoloxías	Descrición	Cualificación
Prácticas de laboratorio	Realizaranse prácticas individuais. As condicións para a súa realización e valoración serán as indicadas con anterioridade polo profesor e dependerán dos recursos dispoñibles.	60
Traballos tutelados	O alumno realizará un traballo individual seguindo as indicacións do profesor que o orientará sobre que aspectos se terán en conta na avaliación.	40

Observacións avaliación
Terase tamén en conta na avaliación o seguimento continuado das sesións maxistrais e actividades propostas así como a participación activa nas mesmas.

Fontes de información



Bibliografía básica	- Toby Velte, Anthony Velte, Robert C. Elsenpeter (2009). Cloud Computing, A Practical Approach. McGraw-Hill - Barry Wilkinson (2009). Grid Computing: Techniques and Applications. CRC Press - Tom White (2011). Hadoop: The Definitive Guide. MapReduce for the Cloud, 2º ed.. O'Reilly - Frederic Magoules, Jie Pan, Kiat-An Tan, Abhinav Kum (2009). Introduction to Grid Computing. Chapman & Hall
Bibliografía complementaria	- Borja Sotomayor, Lisa Childers (2006). Globus Toolkit 4: Programming Java Services. Morgan Kaufmann - Anirban Chakrabarti (2007). Grid Computing Security. Springer - Chuck Lam (2011). Hadoop in Action. Manning - L. George (2011). HBase: The Definitive Guide. O'Reilly - E. Capriolo, D. Wampler, J. Rutherglen (2012). Programming Hive. O'Reilly - Alan Gates (2011). Programming Pig. O'Reilly - Jeanna Matthews, Eli M. Dow, Todd Deshane, Wenjin Hu, Jeremy Bongio, Patrick F. Wilbur, Brendan John (2008). Running Xen: A Hands-On Guide to the Art of Virtualization. Pearson Education - Ian Foster, Carl Kesselman (2004). The Grid 2: Blueprint for a new Computing Infrastructure. Morgan Kaufmann - Bernard Golden (2008). Virtualization For Dummies. Wiley

Recomendacións

Materias que se recomenda ter cursado previamente

Arquitectura e Tecnoloxía de Computadores/614473001

Arquitectura e Tecnoloxía de Supercomputadores/614473002

Programación Paralela/614473004

Administración de Sistemas I/614473010

Materias que se recomenda cursar simultaneamente

Materias que continúan o temario

Observacións

Para os fundamentos básicos desta materia recoméndase utilizar o material recomendado polos profesores (textos, apuntamentos, artigos, etc.), co fin de que a atención do alumno non se disperse ante a gran cantidade de recursos bibliográficos existentes en Internet sobre o tema. Con todo, para os contidos avanzados da materia, así como para a realización do traballo tutelado, recoméndase a procura e selección de recursos en Internet, debido á rápida evolución dos contidos que a materia trata (non así os seus fundamentos básicos) e para dispor dunha maior riqueza nas fontes documentais.

(*)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 de acordo coa normativa vixente que establece o proceso de elaboración de guías