ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

Sistemas distribuidos avanzados

CURSO ACADÉMICO - SEMESTRE

2016-17 - Segundo semestre

Datos Descriptivos

Nombre de la Asignatura	Sistemas distribuidos avanzados
Titulación	61AD - Master Universitario en Ciencias y Tecnologias de la Computacion
Centro responsable de la titulación	Escuela Tecnica Superior de Ingenieria de Sistemas Informaticos
Semestre/s de impartición	Segundo semestre
Módulos	Modulo 3a
Materias	Materia ciencias de la computacion
Carácter	Optativa
Código UPM	613000071
Nombre en inglés	Advanced distributed systems

Datos Generales

Créditos	6	**Curso**	1
Curso Académico	2016-17	**Período de impartición**	Febrero-Junio
Idioma de impartición	Castellano	**Otros idiomas de impartición**	

Requisitos Previos Obligatorios

Asignaturas Previas Requeridas

El plan de estudios Master Universitario en Ciencias y Tecnologias de la Computacion no tiene definidas asignaturas previas superadas para esta asignatura.

Otros Requisitos

El plan de estudios Master Universitario en Ciencias y Tecnologias de la Computacion no tiene definidos otros requisitos para esta asignatura.

Conocimientos Previos

Asignaturas Previas Recomendadas

El coordinador de la asignatura no ha definido asignaturas previas recomendadas.

Otros Conocimientos Previos Recomendados

Algorítmica

Redes de computadores

Sistemas Operativos

Programación

Competencias

E1 - Capacidad para aplicar las teorías, modelos y técnicas actuales en la identificación, análisis, diseño y documentación de soluciones informáticas

E4 - Capacidad para entender las restricciones temporales, de fallos y de comunicación que plantean los entornos de computación distribuida

E5 - Capacidad para desarrollar algoritmos y aplicaciones en entornos de computación distribuida propensos a fallos, con restricciones temporales en la computación y en las comunicaciones

G13 - Resolución de problemas

G6 - Trabajo en contextos internacionales

Resultados de Aprendizaje

RA26 - 1. Dominio de las técnicas de simulación, de los nuevos sistemas de computación y de las características de los sistemas distribuidos más recientes.

RA27 - 2. Capacidad para resolver problemas mediante aplicación de distintas técnicas algorítmicas y para estimar la complejidad de un problema.

Profesorado

Profesorado

Nombre	Despacho	e-mail	Tutorías
Jimenez Merino, Jose Ernesto		ernesto.jimenez.merino@upm.es	
Arevalo Viñuales, Sergio **(Coordinador/a)**		sergio.arevalo@upm.es	

Nota.- Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

Descripción de la Asignatura

Cada vez más nuestra sociedad depende de los computadores. En algunos casos esta dependencia es crítica ya que del buen funcionamiento del sistema de computación dependen vidas humanas o supone una gran cantidad de pérdidas económicas o de negocio. Esta dependencia se ve comprometida si fallan los elementos hardware del sistema. Ejemplos típicos de este tipo de depedencia crítica son: los sistemas de control de tráfico aéreo y de trenes, los sistemas de control de centrales nucleares, sistemas hospitalarios, las grandes web de sistemas de reservas, bancos, contratación en bolsa, etc? Este grado de dependencia hace que tengan que existir técnicas que garanticen un alto grado de fiabilidad y disponibilidad en los sistemas de computación. El problema se puede abordar desde el hardware y desde el software. En ambos casos se utiliza redundancia para conseguir tolerar fallos de hardware. Las técnicas hardware son en general muy caras debido al reducido nivel de producción que obliga a repercutir los costos de I+D en menos cantidad de productos. En cambio las técnicas software permiten una mayor versatilidad y adaptabilidad a diferentes hardware.

Este curso aborda el problema de cómo construir sistemas de computación que garanticen unos grados muy altos de fiabilidad y disponibilidad a las aplicaciones críticas usando técnicas software para tolerar fallos hardware. Para conseguir estas garantías estudiaremos como enmascarar fallos hardware (tolerancia a fallos). Veremos como utilizar la replicación de datos y de procesos. Usaremos como arquitectura de computación el sistema distribuido que nos garantiza redundancia intrinseca con probabilidad de fallos independiente en el funcionamiento de sus componentes. Para conseguir este enmascaramiento es necesario conocer los diferentes tipos de coherencia en el acceso a réplicas. Estudiaremos por tanto las diferentes semánticas de coherencia tanto en procesos como en datos. Extenderemos estos resultados a las transacciones estudiando las bases de datos replicadas.

Temario

1. Replicación y coherencia
   1.1. Tipos de coherencia.
   1.2. Teorema CAP de imposibilidad.
2. Replicación de procesos
   2.1. Modelo
   2.2. Elección de lider
   2.3. Replicación de log.
   2.4. Coherencia de réplicas.
   2.5. Coherencia ante fallos
   2.6. Membresia de replicas
   2.7. Compactación de logs
   2.8. Interacción con procesos clientes

3. Transacciones distribuidas

   3.1. Recuperación distribuida.

   3.2. Protocolo de compromiso en dos fases.

4. Replicación de datos.

   4.1. Motivación

   4.2. Leer uno, escribir en todos.

   4.3. Taxonomia.

   4.4. Copia impaciente primaria

   4.5. Modificación impaciente cualquier copia

   4.6. Copia perezosa primaria

   4.7. Copia perezosa cualquier copia.

   4.8. Caso de estudio: spanner.

Cronograma

Horas totales: 56 horas

Horas presenciales: 56 horas (35.9%)

Peso total de actividades de evaluación continua: 100%

Peso total de actividades de evaluación sólo prueba final: 100%

Semana	Actividad Prensencial en Aula	Actividad Prensencial en Laboratorio	Otra Actividad Presencial	Actividades Evaluación
Semana 1	**Tema 1** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 2	**Tema 2.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 3	**Tema 2.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 4	**Tema 2.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 5	**Tema 2.** Duración: 03:00 LM: Actividad del tipo Lección Magistral			**Examen Parcial** Duración: 01:00 EX: Técnica del tipo Examen Escrito Evaluación continua Actividad presencial
Semana 6	**Tema 2.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 7	**Tema 2.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 8	**Tema 3.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 9	**Tema 3.** Duración: 03:00 LM: Actividad del tipo Lección Magistral			**Examen Parcial** Duración: 01:00 EX: Técnica del tipo Examen Escrito Evaluación continua Actividad presencial
Semana 10	**Tema 3.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 11	**Tema 3.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			

Semana 12	**Tema 4.** Duración: 04:00 LM: Actividad del tipo Lección Magistral			
Semana 13	**Tema 4.** Duración: 03:00 LM: Actividad del tipo Lección Magistral			**Examen Parcial** Duración: 01:00 EX: Técnica del tipo Examen Escrito Evaluación continua Actividad presencial
Semana 14				**Exposición Oral de artículos** Duración: 04:00 PI: Técnica del tipo Presentación Individual Evaluación continua Actividad presencial
Semana 15				
Semana 16				
Semana 17				**Exposición Oral de artículos** Duración: 01:00 PI: Técnica del tipo Presentación Individual Evaluación sólo prueba final Actividad presencial **Examen final** Duración: 02:00 EX: Técnica del tipo Examen Escrito Evaluación sólo prueba final Actividad no presencial

Nota.- El cronograma sigue una planificación teórica de la asignatura que puede sufrir modificaciones durante el curso.

Nota 2.- Para poder calcular correctamente la dedicación de un alumno, la duración de las actividades que se repiten en el tiempo (por ejemplo, subgrupos de prácticas") únicamente se indican la primera vez que se definen.

Actividades de Evaluación

Semana	Descripción	Duración	Tipo evaluación	Técnica evaluativa	Presencial	Peso	Nota mínima	Competencias evaluadas
5	Examen Parcial	01:00	Evaluación continua	EX: Técnica del tipo Examen Escrito	Sí	25%	3 / 10	G13, E1, E4, E5
9	Examen Parcial	01:00	Evaluación continua	EX: Técnica del tipo Examen Escrito	Sí	25%	3 / 10	G13, E1, E4, E5
13	Examen Parcial	01:00	Evaluación continua	EX: Técnica del tipo Examen Escrito	Sí	25%	3 / 10	G13, E1, E4, E5
14	Exposición Oral de artículos	04:00	Evaluación continua	PI: Técnica del tipo Presentación Individual	Sí	25%	4 / 10	G6, E1, E4, E5
17	Exposición Oral de artículos	01:00	Evaluación sólo prueba final	PI: Técnica del tipo Presentación Individual	Sí	25%	5 / 10	G6, E1, E4, E5
17	Examen final	02:00	Evaluación sólo prueba final	EX: Técnica del tipo Examen Escrito	No	75%	5 / 10	G13, E1, E4, E5

Criterios de Evaluación

En los examenes parciales se valoraran los conocimientos adquiridos por los alumnos y su capacidad para resolver problemas. En las presentaciones orales se valorarán tres aspectoas: 1) la calidad de las transparcias presentadas (claridad y concisión); 2) la calidad de la presentación oral realizada (calidad expositiva, conocimientos demostrados); 3) Comprensión técnica del artículo presentado.

Recursos Didácticos

Descripción	Tipo	Observaciones
Articulo sobre Stream Processing	Bibliografía	Making Sense of Stream Processing - Martin Kleppmann. O'Reilly 2016
Artículo de Raft	Bibliografía	D. Ongaro, J. Ousterhout. In Search of an Understandable Consensus Algorithm. USENIX Annual Technical Conference (ATC), Philadelphia, PA, 2014
Datos replicados	Bibliografía	Database Replication: A Tutorial. Bettina Kemme, Ricardo Jiménez-Peris, Marta Patiño-Martínez, and Gustavo Alonso B. Replication, LNCS 5959, pp. 219?252, 2010.
Transacciones	Bibliografía	Philip A. Bernstein, Vassos Hadzilacos and Nathan Goodman}, Concurrency Control and Recovery in Database Systems, Addison-Wesley, 1987.
Spanner	Bibliografía	Spanner: Google?s Globally-Distributed Database. James C. Corbett et al. Proceedings of OSDI 2012, pp. 1-14.
Moodle upm	Recursos web	