



## 1. CURSO

CS3P2. Cloud Computing (Obligatorio)

## 2. INFORMACIÓN GENERAL

2.1 Créditos	:	3
2.2 Horas de teoría	:	1 (Semanal)
2.3 Horas de práctica	:	2 (Semanal)
2.4 Duración del periodo	:	16 semanas
2.5 Condición	:	Obligatorio
2.6 Modalidad	:	Presencial
2.7 Prerrequisitos	:	CS370. Big Data. (9 <sup>no</sup> Sem)

## 3. PROFESORES

Atención previa coordinación con el profesor

## 4. INTRODUCCIÓN AL CURSO

Para entender las técnicas computacionales avanzadas, los estudiantes deberán tener un fuerte conocimiento de las diversas estructuras discretas, estructuras que serán implementadas y usadas en laboratorio en el lenguaje de programación.

## 5. OBJETIVOS

- Que el alumno sea capaz de modelar problemas de ciencia de la computación usando grafos y árboles relacionados con estructuras de datos.
- Que el alumno aplicar eficientemente estrategias de recorrido para poder buscar datos de una manera óptima.

## 6. COMPETENCIAS

- a) Aplicar conocimientos de computación y de matemáticas apropiadas para la disciplina. (**Usar**)
- b) Analizar problemas e identificar y definir los requerimientos computacionales apropiados para su solución. (**Usar**)
- c) Diseñar, implementar y evaluar un sistema, proceso, componente o programa computacional para alcanzar las necesidades deseadas. (**Evaluar**)
- g) Analizar el impacto local y global de la computación sobre los individuos, organizaciones y sociedad. (**Usar**)
- i) Utilizar técnicas y herramientas actuales necesarias para la práctica de la computación. (**Usar**)

## 7. COMPETENCIAS ESPECÍFICAS

- a5) Aplicar técnicas eficientes de resolución de problemas computacionales en ambientes paralelos y distribuidos.
- b13) Modelar base de datos a través de modelos ER, MR, optimización, transacciones y recuperación de la información.
- c4) Diseñar e implementar arquitecturas de software escalables debilmente acopladas, flexibles, seguras y de fácil mantenimiento en distintas plataformas.
- g10) Analizar el impacto de la computación en la nube en las organizaciones.
- i11) Utilizar y administración contenedores de microservicios para crear aplicaciones escalables.

## 8. TEMAS

Unidad 1: Sistemas distribuidos (15)	
Competencias esperadas: a,b	
Temas	Objetivos de Aprendizaje
<ul style="list-style-type: none"> <li>• Fallos: <ul style="list-style-type: none"> <li>– Fallos basados en red (incluyendo particiones) y fallos basados en nodos</li> <li>– Impacto en garantías a nivel de sistema (p.e., disponibilidad)</li> </ul> </li> <li>• Envío de mensajes distribuido: <ul style="list-style-type: none"> <li>– Conversión y transmisión de datos</li> <li>– Sockets</li> <li>– Secuenciamiento de mensajes</li> <li>– Almacenando <i>Buffering</i>, reenviando y desechando mensajes</li> </ul> </li> <li>• Compensaciones de diseño para Sistemas Distribuidos: <ul style="list-style-type: none"> <li>– Latencia versus rendimiento</li> <li>– Consistencia, disponibilidad, tolerancia de particiones</li> </ul> </li> <li>• Diseño de Servicio Distribuido: <ul style="list-style-type: none"> <li>– Protocolos y servicios Stateful versus stateless</li> <li>– Diseños de Sesión (basados en la conexión)</li> <li>– Diseños reactivos (provocados por E/S) y diseños de múltiples hilos</li> </ul> </li> <li>• Algoritmos de Distribución de Núcleos: <ul style="list-style-type: none"> <li>– Elección, descubrimiento</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Distinguir las fallas de red de otros tipos de fallas [Familiarizarse]</li> <li>• Explicar por qué estructuras de sincronización como cerraduras simples (<i>locks</i>) no son útiles en la presencia de fallas distribuidas [Familiarizarse]</li> <li>• Escribir un programa que realiza cualquier proceso de <i>marshalling</i> requerido y la conversión en unidades de mensajes, tales como paquetes, para comunicar datos importantes entre dos <i>hosts</i> [Usar]</li> <li>• Medir el rendimiento observado y la latencia de la respuesta a través de los <i>hosts</i> en una red dada [Usar]</li> <li>• Explicar por qué un sistema distribuido no puede ser simultáneamente Consistente (<i>Consistent</i>), Disponible (<i>Available</i>) y Tolerante a fallas (<i>Partition tolerant</i>). [Familiarizarse]</li> <li>• Implementar un servidor sencillo - por ejemplo, un servicio de corrección ortográfica [Usar]</li> <li>• Explicar las ventajas y desventajas entre: <i>overhead</i>, escalabilidad y tolerancia a fallas entre escoger un diseño sin estado (<i>stateless</i>) y un diseño con estado (<i>stateful</i>) para un determinado servicio [Familiarizarse]</li> <li>• Describir los desafíos en la escalabilidad, asociados con un servicio creciente para soportar muchos clientes, así como los asociados con un servicio que tendrá transitoriamente muchos clientes [Familiarizarse]</li> <li>• Dar ejemplos de problemas donde algoritmos de consenso son requeridos, por ejemplo, la elección de líder [Usar]</li> </ul>
Lecturas : [Cou+11]	

Unidad 2: Cloud Computing (15)	
Competencias esperadas: a,b	
Temas	Objetivos de Aprendizaje
<ul style="list-style-type: none"> <li>• Visión global de <i>Cloud Computing</i>.</li> <li>• Historia.</li> <li>• Visión global de las tecnologías que envuelve.</li> <li>• Beneficios, riesgos y aspectos económicos.</li> <li>• Servicios en la nube. <ul style="list-style-type: none"> <li>– Infraestructura como servicio <ul style="list-style-type: none"> <li>* Elasticidad de recursos</li> <li>* APIs de la Plataforma</li> </ul> </li> <li>– Software como servicio</li> <li>– Seguridad</li> <li>– Administración del Costo</li> </ul> </li> <li>• Computación a Escala de Internet: <ul style="list-style-type: none"> <li>– Particionamiento de Tareas</li> <li>– Acceso a datos</li> <li>– Clusters, grids y mallas</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Explicar el concepto de <i>Cloud Computing</i>. [Familiarizarse]</li> <li>• Listar algunas tecnologías relacionadas con <i>Cloud Computing</i>. [Familiarizarse]</li> <li>• Explicar las estrategias para sincronizar una vista común de datos compartidos a través de una colección de dispositivos [Familiarizarse]</li> <li>• Discutir las ventajas y desventajas del paradigma de <i>Cloud Computing</i>. [Familiarizarse]</li> <li>• Expresar los beneficios económicos así como las características y riesgos del paradigma de Cloud para negocios y proveedores de cloud. [Familiarizarse]</li> <li>• Diferenciar entre los modelos de servicio. [Usar]</li> </ul>
Lecturas : [HDF11], [BVS13]	

Unidad 3: Centros de Procesamiento de Datos (10)	
Competencias esperadas: g,i	
Temas	Objetivos de Aprendizaje
<ul style="list-style-type: none"> <li>• Visión global de un centro de procesamiento de datos.</li> <li>• Consideraciones en el diseño.</li> <li>• Comparación de actuales grandes centros de procesamiento de datos.</li> </ul>	<ul style="list-style-type: none"> <li>• Describir la evolución de los <i>Data Centers</i>. [Familiarizarse]</li> <li>• Esbozar la arquitectura de un data center en detalle. [Familiarizarse]</li> <li>• Indicar consideraciones de diseño y discutir su impacto. [Familiarizarse]</li> </ul>
Lecturas : [HDF11], [BVS13]	

Unidad 4: Cloud Computing (20)	
Competencias esperadas: i,j	
Temas	Objetivos de Aprendizaje
<ul style="list-style-type: none"> <li>• Virtualización. <ul style="list-style-type: none"> <li>– Gestión de recursos compartidos</li> <li>– Migración de procesos</li> </ul> </li> <li>• Seguridad, recursos y aislamiento de fallas.</li> <li>• Almacenamiento como servicio.</li> <li>• Elasticidad.</li> <li>• Xen y Wmware.</li> <li>• Amazon EC2.</li> </ul>	<ul style="list-style-type: none"> <li>• Virtualización. <ul style="list-style-type: none"> <li>– Gestión de recursos compartidos</li> <li>– Migración de procesos</li> </ul> </li> <li>. [Familiarizarse]</li> <li>• Explicar las ventajas y desventajas de usar una infraestructura virtualizada. [Familiarizarse]</li> <li>• Identificar las razones por qué la virtualización está llegando a ser enormemente útil, especialmente en la cloud. [Familiarizarse]</li> <li>• Explicar diferentes tipos de aislamiento como falla, recursos y seguridad proporcionados por la virtualización y utilizado por la nube. [Familiarizarse]</li> <li>• Explicar la complejidad que puede tener el administrar en términos de niveles de abstracción y interfaces bien definidas y su aplicabilidad para la virtualización en la nube. [Familiarizarse]</li> <li>• Definir virtualización y identificar diferentes tipos de máquinas virtuales. [Familiarizarse]</li> <li>• Identificar condiciones de virtualización de CPU, reconocer la diferencia entre <i>full virtualization</i> y <i>paravirtualization</i>, explicar emulación como mayor técnica para virtualización del CPU y examinar planificación virtual del CPU en Xen. [Familiarizarse]</li> <li>• Esbozar la diferencia entre la clásica memoria virtual del SO y la virtualización de memoria. Explicar los múltiples niveles de mapeamiento de páginas en oposición a la virtualización de la memoria. Definir memoria <i>over-commitment</i> e ilustrar sobre Wmware <i>memory ballooning</i> como técnica de reclamo para sistemas virtualizados con memoria <i>over-committed</i>. [Familiarizarse]</li> </ul>
Lecturas : [HDF11], [BVS13]	

Unidad 5: Cloud Computing (12)	
Competencias esperadas: i,j	
Temas	Objetivos de Aprendizaje
<ul style="list-style-type: none"> <li>• Almacenamiento de datos en la nube: <ul style="list-style-type: none"> <li>– Acceso compartido a data stores de consistencia débil</li> <li>– Sincronización de datos</li> <li>– Particionamiento de datos</li> <li>– Sistemas de Archivos Distribuidos</li> <li>– Replicación</li> </ul> </li> <li>• Visión global sobre tecnologías de almacenamiento.</li> <li>• Conceptos fundamentales sobre almacenamiento en la cloud.</li> <li>• Amazon S3 y EBS.</li> <li>• Sistema de archivos distribuidos.</li> <li>• Sistema de bases de datos NoSQL.</li> </ul>	<ul style="list-style-type: none"> <li>• Describir la organización general de datos y almacenamiento. [Familiarizarse]</li> <li>• Identificar los problemas de escalabilidad y administración de la big data. Discutir varias abstracciones en almacenamiento. [Familiarizarse]</li> <li>• Comparar y contrastar diferentes tipos de sistema de archivos.</li> <li>• Comparar y contrastar el Sistema de Archivos Distribuido de Hadoop (HDFS) y el Sistema de Archivos Paralelo Virtual (PVFS). [Usar]</li> <li>• Comparar y contrastar diferentes tipos de bases de datos. Discutir las ventajas y desventajas sobre las bases de datos NoSQL. [Usar]</li> <li>• Discutir los conceptos de almacenamiento en la cloud. [Familiarizarse]</li> </ul>
Lecturas : [HDF11], [BVS13]	

Unidad 6: Modelos de Programación (12)	
Competencias esperadas: g,j	
Temas	Objetivos de Aprendizaje
<ul style="list-style-type: none"> <li>• Visión global de los modelos de programación basados en <i>cloud computing</i>.</li> <li>• Modelo de Programación MapReduce.</li> <li>• Modelo de programación para aplicaciones basadas en Grafos.</li> </ul>	<ul style="list-style-type: none"> <li>• Explicar los aspectos fundamentales de los modelos de programación paralela y distribuida. [Familiarizarse]</li> <li>• Diferencias entre los modelos de programación: MapReduce, Pregel, GraphLab y Giraph. [Usar]</li> <li>• Explicar los principales conceptos en el modelo de programación MapReduce. [Usar]</li> </ul>
Lecturas : [HDF11], [BVS13], [Low+12], [Mal+10], [Bal+08]	

## 9. PLAN DE TRABAJO

### 9.1 Metodología

Se fomenta la participación individual y en equipo para exponer sus ideas, motivándolos con puntos adicionales en las diferentes etapas de la evaluación del curso.

### 9.2 Sesiones Teóricas

Las sesiones de teoría se llevan a cabo en clases magistrales donde se realizarán actividades que propicien un aprendizaje activo, con dinámicas que permitan a los estudiantes interiorizar los conceptos.

### 9.3 Sesiones Prácticas

Las sesiones prácticas se llevan en clase donde se desarrollan una serie de ejercicios y/o conceptos prácticos mediante planteamiento de problemas, la resolución de problemas, ejercicios puntuales y/o en contextos aplicativos.

## 10. SISTEMA DE EVALUACIÓN

\*\*\*\*\* EVALUATION MISSING \*\*\*\*\*

## 11. BIBLIOGRAFÍA BÁSICA

- [Bal+08] Shumeet Baluja et al. “Video Suggestion and Discovery for Youtube: Taking Random Walks Through the View Graph”. In: *Proceedings of the 17th International Conference on World Wide Web*. WWW '08. Beijing, China: ACM, 2008, pp. 895–904. ISBN: 978-1-60558-085-2. DOI: 10.1145/1367497.1367618. URL: <http://doi.acm.org/10.1145/1367497.1367618>.
- [BVS13] Rajkumar Buyya, Christian Vecchiola, and S. Thamarai Selvi. *Mastering Cloud Computing: Foundations and Applications Programming*. 1st. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2013. ISBN: 9780124095397, 9780124114548.
- [Cou+11] George Coulouris et al. *Distributed Systems: Concepts and Design*. 5th. USA: Addison-Wesley Publishing Company, 2011. ISBN: 0132143011, 9780132143011.
- [HDF11] Kai Hwang, Jack Dongarra, and Geoffrey C. Fox. *Distributed and Cloud Computing: From Parallel Processing to the Internet of Things*. 1st. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011. ISBN: 0123858801, 9780123858801.
- [Low+12] Yucheng Low et al. “Distributed GraphLab: A Framework for Machine Learning and Data Mining in the Cloud”. In: *Proc. VLDB Endow*. 5.8 (Apr. 2012), pp. 716–727. ISSN: 2150-8097. DOI: 10.14778/2212351.2212354. URL: <http://dx.doi.org/10.14778/2212351.2212354>.
- [Mal+10] Grzegorz Malewicz et al. “Pregel: A System for Large-scale Graph Processing”. In: *Proc. ACM SIGMOD*. SIGMOD '10 (2010), pp. 135–146. DOI: 10.1145/1807167.1807184. URL: <http://doi.acm.org/10.1145/1807167.1807184>.