

Universidad Católica San Pablo
Facultad de Ingeniería y Computación
Escuela Profesional de
Ciencia de la Computación
SILABO



CS331. Cloud Computing (Obligatorio)

2010-1

1. DATOS GENERALES

1.1 CARRERA PROFESIONAL	:	Ciencia de la Computación
1.2 ASIGNATURA	:	CS331. Cloud Computing
1.3 SEMESTRE ACADÉMICO	:	10 ^{mo} Semestre.
1.4 PREREQUISITO(S)	:	CS230W. Computación Centrada en Redes. (8 ^{vo} Sem) , CS314. Algoritmos Paralelos. (7 ^{mo} Sem)
1.5 CARÁCTER	:	Obligatorio
1.6 HORAS	:	1 HT; 2 HP; 2 HL;
1.7 CRÉDITOS	:	3

2. DOCENTE

3. FUNDAMENTACIÓN DEL CURSO

La capacidad de procesamiento de una sola máquina es limitada y la Ley de Moore se ha encontrado con barreras antes de lo previsto, a pesar de esto la necesidad de mayor poder computacional es creciente.

El uso de las computadoras como elementos conectados entre sí es cada vez más común y cada vez en mayor escala, la capacidad de comunicación entre dispositivos (computadoras, celulares, pdas, etc.), abre las puertas a la existencia de una única plataforma donde la información de los usuarios esté disponible siempre, sin importar el medio de acceso a esta (*Cloud computing*).

La computación en la nube de internet o un grupo de computadores permite conseguir ambos objetivos, traspasando la barrera de una sola máquina para poder integrar las capacidades de distintos dispositivos y permitirles interactuar en un entorno que el usuario perciba como unificado; además, al conectarlos, el tope de desempeño del sistema ya no es la capacidad de un sólo elemento (e.g. CPU) sino la cantidad de participantes en este, por lo cual existe una escalabilidad del poder computacional muchísimo mayor.

4. SUMILLA

1. Introducción a cloud computing 2. Temas de investigación en cloud computing 3. Cloud data management 4. Data-intensive applications 5. Programando para Cloud Computing

5. OBJETIVO GENERAL

- Comprender los conceptos básicos de la computación en nube, incluyendo definiciones, historia, pros y cons de la misma, comparaciones con tecnologías relacionadas, tales como grid computing, o utility computing.
- Conocer la tecnología que soporta a la computación en nube.
- Comprender la relación entre data-intensive applications y cloud computing, y
- Evaluar el nuevo modelo de computación para conocer las tendencias de esta área emergente.

6. CONTRIBUCIÓN A LA FORMACIÓN PROFESIONAL Y FORMACIÓN GENERAL

Esta disciplina contribuye al logro de los siguientes resultados de la carrera:

- a) Aplicar conocimientos de computación y de matemáticas apropiadas para la disciplina. [Nivel Bloom: 3]
- b) Analizar problemas e identificar y definir los requerimientos computacionales apropiados para su solución. [Nivel Bloom: 4]
- c) Diseñar, implementar y evaluar un sistema, proceso, componente o programa computacional para alcanzar las necesidades deseadas. [Nivel Bloom: 4]
- d) Trabajar efectivamente en equipos para cumplir con un objetivo común. [Nivel Bloom: 3]
- i) Utilizar técnicas y herramientas actuales necesarias para la práctica de la computación. [Nivel Bloom: 3]
- j) Aplicar la base matemática, principios de algoritmos y la teoría de la Ciencia de la Computación en el modelamiento y diseño de sistemas computacionales de tal manera que demuestre comprensión de los puntos de equilibrio involucrados en la opción escogida. [Nivel Bloom: 4]
- k) Aplicar los principios de desarrollo y diseño en la construcción de sistemas de software de complejidad variable. [Nivel Bloom: 4]

7. CONTENIDOS

UNIDAD 1: Introducción a cloud computing (7 horas)

Nivel Bloom: 2

OBJETIVO GENERAL

CONTENIDO

- Describir tecnologías emergentes y el área de computación centradas en redes así como evaluar las capacidades y limitaciones actuales y su potencial a corto plazo.
- Identificar y discutir diferentes sistemas especializados.
- Sintetizar los temas técnicos centrales asociados con la implementación del crecimiento de sistemas especializados..
- Comprender como aparecio el paradigma de computación en nube.

- Computación en redes y multimedia distribuida.
- Sistemas cliente-servidor.
- Sistemas distribuidos.
- Sistemas paralelos.
- Sistemas basados en web.

Lecturas: [Armbrust et al., 2009], [Mell and Grance, 2009]

UNIDAD 2: Temas de investigación en cloud computing (8 horas)

Nivel Bloom: 2

OBJETIVO GENERAL

CONTENIDO

- Entender la relación entre los diferentes tipos de investigación que procedieron a la computación en nube.
- Conocer distintas líneas de investigación de computación en nube.

- Data Center Network Architecture
- Network Management
- Resource and Performance Management
- Data management

Lecturas: [Vaquero et al., 2009], [Mei et al., 2008]

UNIDAD 3: Cloud data management (10 horas)	
Nivel Bloom: 3	
OBJETIVO GENERAL	CONTENIDO
<ul style="list-style-type: none"> ▪ Criticar y defender las aplicaciones de información de tamaño pequeño y mediano con respecto a la satisfacción de las necesidades reales del usuario. ▪ Explicar las medidas de eficiencia (estimación, tiempo de respuesta) y efectividad (<i>precision - recall</i>). ▪ Describir métodos para asegurar que los sistemas de información pueden escalar de lo individual a lo global. ▪ Identificar asuntos relacionados a la persistencia de datos en una organización. ▪ Evaluar estrategias simples para ejecutar una consulta distribuida para seleccionar la estrategia que minimice la cantidad de transferencia de datos. ▪ Conocer diferentes casos de objetos distribuidas. 	<ul style="list-style-type: none"> ▪ Almacenamiento y recuperación de información (IS&R). ▪ Búsqueda, recuperación, enlace, navegación. ▪ Escalabilidad, eficiencia y efectividad. ▪ Arquitectura de base de datos e independencia de datos. <ul style="list-style-type: none"> • Almacenamiento de datos distribuido. • Procesamiento de consultas distribuidas. • Modelo de transacción distribuido. • Control de concurrencia. • Soluciones heterogéneas y homogéneas. • Cliente-servidor. ▪ Big Data. ▪ Large small data. ▪ Bases de datos <i>NoSQL</i>.
Lecturas: [Stonebraker, 1986], [Stonebraker et al., 2007], [Agrawal et al., 2009]	

UNIDAD 4: Data-intensive applications (10 horas)	
Nivel Bloom: 3	
OBJETIVO GENERAL	CONTENIDO
<ul style="list-style-type: none"> ▪ Entender el modelo de programación MapReduce. ▪ Conocer diferentes modos de uso de MapReduce. ▪ Describir métodos para asegurar que los sistemas de información pueden escalar de lo individual a lo global. ▪ Identificar asuntos relacionados a la persistencia de datos en una organización. ▪ Evaluar estrategias simples para ejecutar una consulta distribuida para seleccionar la estrategia que minimice la cantidad de transferencia de datos. 	<ul style="list-style-type: none"> ▪ Modelo de programación MapReduce. ▪ Ejemplos de aplicaciones en la academia y en la industria. ▪ Aplicaciones usando MapReduce. ▪ Otros lenguajes de programación para Cloud Computing.
Lecturas: [Hey et al., 2009], [Bryant, 2007], [Dean and Ghemawat, 2008]	

UNIDAD 5: Programando para Cloud Computing (10 horas)	
Nivel Bloom: 3	
OBJETIVO GENERAL	CONTENIDO
<ul style="list-style-type: none"> ▪ Conocer los diferentes services de Amazon Web Services. ▪ Aplicar conocimientos de Cloud Computing para crear aplicaciones que usen otros servicios de Cloud Computing. ▪ Conocer los diferentes proveedores de servicios de Cloud Computing. ▪ Entender las similitudes y diferencias, ventajas y desventajas de los diferentes frameworks para crear <i>private clouds</i>. 	<ul style="list-style-type: none"> ▪ Usando Amazon Web Services. ▪ MapReduce en Amazon Web Services. ▪ Proveedores de Cloud Computing. ▪ Frameworks para crear servicios de Cloud Computing.
Lecturas: [Dean and Ghemawat, 2008], [Nurmi et al., 2009], [Services, 2010]	

8. METODOLOGÍA

El profesor del curso presentará clases teóricas de los temas señalados en el programa propiciando la intervención de los alumnos.

El profesor del curso presentará demostraciones para fundamentar clases teóricas.

El profesor y los alumnos realizarán prácticas

Los alumnos deberán asistir a clase habiendo leído lo que el profesor va a presentar. De esta manera se facilitará la comprensión y los estudiantes estarán en mejores condiciones de hacer consultas en clase.

9. EVALUACIONES

Evaluación Permanente 1 : 20 %

Examen Parcial : 30 %

Evaluación Permanente 2 : 20 %

Examen Final : 30 %

Referencias

- [Agrawal et al., 2009] Agrawal, R., Ailamaki, A., Bernstein, P. A., Brewer, E. A., Carey, M. J., Chaudhuri, S., Doan, A., Florescu, D., Franklin, M. J., Garcia-Molina, H., Gehrke, J., Gruenwald, L., Haas, L. M., Halevy, A. Y., Hellerstein, J. M., Ioannidis, Y. E., Korth, H. F., Kossman, D., Madden, S., Magoulas, R., Ooi, B. C., O'Reilly, T., Ramakrishnan, R., Sarawagi, S., Stonebraker, M., Szalay, A. S., and Weikum, G. (2009). The Claremont report on database research. *Communication of ACM*, 52(6):56–65.
- [Armbrust et al., 2009] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R. H., Konwinski, A., Lee, G., Patterson, D. A., Rabkin, A., Stoica, I., and Zaharia, M. (2009). Above the clouds: A berkeley view of cloud computing. Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley.
- [Bryant, 2007] Bryant, R. E. (2007). Data-intensive supercomputing: The case for disc. Technical report, Carnegie Mellon University, School of Computer Science.
- [Dean and Ghemawat, 2008] Dean, J. and Ghemawat, S. (2008). Mapreduce: simplified data processing on large clusters. *Commun. ACM*, 51(1):107–113.

- [Hey et al., 2009] Hey, T., Tansley, S., and Tolle, K., editors (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, Redmond, Washington.
- [Mei et al., 2008] Mei, L., Chan, W., and Tse, T. (2008). A tale of clouds: Paradigm comparisons and some thoughts on research issues. *Asia-Pacific Conference on Services Computing. 2006 IEEE*, 0:464–469.
- [Mell and Grance, 2009] Mell, P. and Grance, T. (2009). The nist definition of cloud computing.
- [Nurmi et al., 2009] Nurmi, D., Wolski, R., Grzegorzcyk, C., Obertelli, G., Soman, S., Youseff, L., and Zagorodnov, D. (2009). The eucalyptus open-source cloud-computing system. In *Proceedings of the 2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid, CCGRID '09*, pages 124–131, Washington, DC, USA. IEEE Computer Society.
- [Services, 2010] Services, A. W. (2010). Amazon web services. <http://aws.amazon.com/>.
- [Stonebraker, 1986] Stonebraker, M. (1986). The case for shared nothing. *Database Engineering*, 9:4–9.
- [Stonebraker et al., 2007] Stonebraker, M., Madden, S., Abadi, D. J., Harizopoulos, S., Hachem, N., and Helland, P. (2007). The end of an architectural era: (it’s time for a complete rewrite). In *VLDB '07: Proceedings of the 33rd international conference on Very large data bases*, pages 1150–1160. VLDB Endowment.
- [Vaquero et al., 2009] Vaquero, L. M., Rodero-Merino, L., Caceres, J., and Lindner, M. (2009). A break in the clouds: towards a cloud definition. *SIGCOMM Comput. Commun. Rev.*, 39(1):50–55.