

**University de Piura (UDEP)**  
Sillabus 2022-I

**1. COURSE**

CS3P2. Cloud Computing (Mandatory)

**2. GENERAL INFORMATION**

<b>2.1 Credits</b>	:	3
<b>2.2 Theory Hours</b>	:	1 (Weekly)
<b>2.3 Practice Hours</b>	:	2 (Weekly)
<b>2.4 Duration of the period</b>	:	16 weeks
<b>2.5 Type of course</b>	:	Mandatory
<b>2.6 Modality</b>	:	Face to face
<b>2.7 Prerequisites</b>	:	CS370. Big Data. (9 <sup>th</sup> Sem)

**3. PROFESSORS**

Meetings after coordination with the professor

**4. INTRODUCTION TO THE COURSE**

In order to understand the advanced computational techniques, the students must have a strong knowledge of the various discrete structures, structures that will be implemented and used in the laboratory in the programming language.

**5. GOALS**

- That the student is able to model computer science problems using graphs and trees related to data structures.
- That the student apply efficient travel strategies to be able to search data in an optimal way.

**6. COMPETENCES**

Nooutcomes

Nospecificoutcomes

**7. TOPICS**

<b>Unit 1: Distributed Systems (15)</b>	
<b>Competences Expected: a,b</b>	
<b>Topics</b>	<b>Learning Outcomes</b>
<ul style="list-style-type: none"> <li>• Faults (cross-reference OS/Fault Tolerance) <ul style="list-style-type: none"> <li>– Network-based (including partitions) and node-based failures</li> <li>– Impact on system-wide guarantees (e.g., availability)</li> </ul> </li> <li>• Distributed message sending <ul style="list-style-type: none"> <li>– Data conversion and transmission</li> <li>– Sockets</li> <li>– Message sequencing</li> <li>– Buffering, retrying, and dropping messages</li> </ul> </li> <li>• Distributed system design tradeoffs <ul style="list-style-type: none"> <li>– Latency versus throughput</li> <li>– Consistency, availability, partition tolerance</li> </ul> </li> <li>• Distributed service design <ul style="list-style-type: none"> <li>– Stateful versus stateless protocols and services</li> <li>– Session (connection-based) designs</li> <li>– Reactive (IO-triggered) and multithreaded designs</li> </ul> </li> <li>• Core distributed algorithms <ul style="list-style-type: none"> <li>– Election, discovery</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Distinguish network faults from other kinds of failures [Familiarity]</li> <li>• Explain why synchronization constructs such as simple locks are not useful in the presence of distributed faults [Familiarity]</li> <li>• Write a program that performs any required marshalling and conversion into message units, such as packets, to communicate interesting data between two hosts [Usage]</li> <li>• Measure the observed throughput and response latency across hosts in a given network [Usage]</li> <li>• Explain why no distributed system can be simultaneously consistent, available, and partition tolerant [Familiarity]</li> <li>• Implement a simple server – for example, a spell checking service [Usage]</li> <li>• Explain the tradeoffs among overhead, scalability, and fault tolerance when choosing a stateful v stateless design for a given service [Familiarity]</li> <li>• Describe the scalability challenges associated with a service growing to accommodate many clients, as well as those associated with a service only transiently having many clients [Familiarity]</li> <li>• Give examples of problems for which consensus algorithms such as leader election are required [Usage]</li> </ul>
<b>Readings :</b> [Cou+11]	

Unit 2: Cloud Computing (15)	
Competences Expected: a,b	
Topics	Learning Outcomes
<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>• Cloud services <ul style="list-style-type: none"> <li>– Infrastructure as a service <ul style="list-style-type: none"> <li>* Elasticity of resources</li> <li>* Platform APIs</li> </ul> </li> <li>– Software as a service</li> <li>– Security</li> <li>– Cost management</li> </ul> </li> <li>• Internet-Scale computing <ul style="list-style-type: none"> <li>– Task partitioning</li> <li>– Data access</li> <li>– Clusters, grids, and meshes</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Explicar el concepto de Cloud Computing. [Familiarity]</li> <li>• Listar algunas tecnologías relacionadas con Cloud Computing. [Familiarity]</li> <li>• Explain strategies to synchronize a common view of shared data across a collection of devices [Familiarity]</li> <li>• Discutir las ventajas y desventajas del paradigma de Cloud Computing. [Familiarity]</li> <li>• Expresar los beneficios económicos así como las características y riesgos del paradigma de Cloud para negocios y proveedores de cloud. [Familiarity]</li> <li>• Diferenciar entre los modelos de servicio. [Usage]</li> </ul>
Readings : [HDF11], [BVS13]	

Unit 3: Centros de Procesamiento de Datos (10)	
Competences Expected: g,i	
Topics	Learning Outcomes
<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 Data Centers. [Familiarity]</li> <li>• Esbozar la arquitectura de un data center en detalle. [Familiarity]</li> <li>• Indicar consideraciones de diseño y discutir su impacto. [Familiarity]</li> </ul>
Readings : [HDF11], [BVS13]	

Unit 4: Cloud Computing (20)	
Competences Expected: i,j	
Topics	Learning Outcomes
<ul style="list-style-type: none"> <li>• Virtualization <ul style="list-style-type: none"> <li>– Shared resource management</li> <li>– Migration of processes</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>• Virtualization <ul style="list-style-type: none"> <li>– Shared resource management</li> <li>– Migration of processes</li> </ul> </li> <li>. [Familiarity]</li> <li>• Explain the advantages and disadvantages of using virtualized infrastructure. [Familiarity]</li> <li>• Identificar las razones por qué la virtualización está llegando a ser enorme útil, especialmente en la cloud. [Familiarity]</li> <li>• Explicar diferentes tipos de aislamiento como falla, recursos y seguridad proporcionados por la virtualización y utilizado por la cloud. [Familiarity]</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 cloud. [Familiarity]</li> <li>• Definir virtualización y identificar diferentes tipos de máquinas virtuales. [Familiarity]</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. [Familiarity]</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>. [Familiarity]</li> </ul>
<b>Readings :</b> [HDF11], [BVS13]	

Unit 5: Cloud Computing (12)	
Competences Expected: i,j	
Topics	Learning Outcomes
<ul style="list-style-type: none"> <li>• Cloud-based data storage               <ul style="list-style-type: none"> <li>– Shared access to weakly consistent data stores</li> <li>– Data synchronization</li> <li>– Data partitioning</li> <li>– Distributed file systems</li> <li>– Replication</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. [Familiarity]</li> <li>• Identificar los problemas de escalabilidad y administración de la big data. Discutir varias abstracciones en almacenamiento. [Familiarity]</li> <li>• Comparar y contrastar diferentes tipos de sistema de archivos. Comparar y contrastar el Sistema de Archivos Distribuido de Hadoop (HDFS) y el Sistema de Archivos Paralelo Virtual (PVFS). [Usage]</li> <li>• Comparar y contrastar diferentes tipos de bases de datos. Discutir las ventajas y desventajas sobre las bases de datos NoSQL. [Usage]</li> <li>• Discutir los conceptos de almacenamiento en la cloud. [Familiarity]</li> </ul>
Readings : [HDF11], [BVS13]	

Unit 6: Modelos de Programación (12)	
Competences Expected: g,j	
Topics	Learning Outcomes
<ul style="list-style-type: none"> <li>• Visión global de los modelos de programación basados en cloud computing.</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. [Familiarity]</li> <li>• Diferencias entre los modelos de programación: MapReduce, Pregel, GraphLab y Giraph. [Usage]</li> <li>• Explicar los principales conceptos en el modelo de programación MapReduce. [Usage]</li> </ul>
Readings : [HDF11], [BVS13], [Low+12], [Mal+10], [Bal+08]	

## 8. WORKPLAN

### 8.1 Methodology

Individual and team participation is encouraged to present their ideas, motivating them with additional points in the different stages of the course evaluation.

### 8.2 Theory Sessions

The theory sessions are held in master classes with activities including active learning and roleplay to allow students to internalize the concepts.

### 8.3 Practical Sessions

The practical sessions are held in class where a series of exercises and/or practical concepts are developed through problem solving, problem solving, specific exercises and/or in application contexts.

## 9. PLANNING

DATE	TIME	SESSION TYPE	PROFESSOR
See at EDU	See at EDU	See at EDU	See at EDU

## 10. EVALUATION SYSTEM

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

## 11. BASIC BIBLIOGRAPHY

- [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>.