Information Systems

Academic Year: (2023 / 2024)   **Review date: 28-04-2023**

Department assigned to the subject: Telematic Engineering Department

Coordinating teacher: DIAZ SANCHEZ, DANIEL

Type: Electives ECTS Credits : 6.0

Year : 5 Semester :

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

System architectures (be familiar with the labs)
Telematic Applications (have knowledge about upper protocolo layers, specifically HTTP)

OBJECTIVES

The goal of this course is to enable the student to acquire knowledge relative to the fundamental concepts on modern distributed systems and enable him/her to use modern distributed systems that enable popular services used by millions of users all over the world, such as Dropbox, Spotify, Youtube, Google Search Engine

The philosophy of this course is to enable the practical learning of the distributed computing through the usage of interfaces from widely known and used services. In order to achieve this goal, the student should acquire a series of knowledges and capabilities.

Knowledge

To achieve this goal the student must acquire the following knowledge:

- To know the structure of a modern distributed systems, its internal characteristics and interfaces with other systems, such as mobile and desktop applications or others.
- To know the basic communication mechanisms of distributed systems, using different protocols used nowadays, such as HTTP (REST interfaces) or Web Services
- To know modern load balancing systems, from Enterprise Service Bus to the ones used by Map Reduce (used currently by Google, Amazon and others).
- To know real-world scenarios for web applications with high number of visits (Amazon, periódicos); cloud architectures for applications development and/or storing.
- To know the advantages and disadvantages of these systems and when to develop a specific application in a distributed way (resource scaling, data center economics). To know data sets that need distributed processing.

Specific capabilities

The specific capabilities that will be acquired by the student upon successful completion of the course will be:
- To use basic communication mechanisms in distributed systems.
- To know centralised and distributed applications through demonstrative labs.
- To model and deploy a distributed system and to use existing applications and clouds.

General capabilities

The general capabilities acquired by the students will be:

- Ability to use and apply knowledge of telecommunications technologies and engineering. The students will work on this capability in the lab sessions and in the exercise sessions.
- Ability to use the techniques and modern engineering tools necessary for the professional practice
- Ability to effectively communicate information in speech, presentation and in writing in english and spanish, through the development of the proposed activities in the subject (exercises, projects about

new trend technologies, etc)
- Recognition of the need of a life-long learning and the ability to obtain and apply the required and suitable information through the use of technical literature related to the subject field in spanish and in english
- Knowledge of contemporary issues and trends in the field of study

DESCRIPTION OF CONTENTS: PROGRAMME
The contents of the programme are the following:

1. Introduction to distributed computing
    1.1. From RPC/Corba to Main Frame, Grid and Cloud Computing
    1.2 Current state of the technologies
        1.2.a Legacy systems and evolution
        1.2.b Distributed systems
        1.2.c Cloud Computing
2. Distributed Systems and Cloud Computing
    2.1 What is it?
    2.2. Architectures (IaaS, PaaS, SaaS)
        2.2.a IaaS - Amazon EC2
        2.2.b SaaS - App Engine/Azure
    2.3 Distributed storing and distributed execution (for cloud systems as Google, Azure or Amazon EC2)
        2.3.b Security aspects in Cloud
    2.4 Use cases
        2.4.a Do I really need Cloud Computing for my project?
3. Communication protocols for distributed systems
    3.1 Legacy protocols
    3.2 Current protocols
        3.2.a REST / HTTP
        3.2.b Web Services
        3.2.c Messaging
    3.3 Data representation
        3.3.a Json vs XML - AJAX
    3.4 Modern Massive-use Web systems
    3.5 Mobile applications and REST systems
        3.5.a Why to use REST/HTTP?
4. Big Data: Mapreduce y otros sistemas
    4.1 Big Table from Google and Hadoop
    4.2 Introduction to Hadoop / Spark

LEARNING ACTIVITIES AND METHODOLOGY
The activities used to underpin the competences and the skills in the course are :

- During the first part of the course, the students should complete a set of guided labs, where they will modify code provided by the teaching staff in order to gradually acquire the needed skills to pass the course. These guided labs will be closely related with a real case study, presented at the beginning of the course to learn modern cloud systems as those provided by Google or Amazon.

- During a six-week period, students will be divided into pairs, and they are expected to complete a project entailing the design and implementation of a Cloud application (or an application that access a Cloud) that should use one or more of the technologies and paradigms covered during the course (PO a).

- Students are requested in several activities throughout the course to search for auxiliary documents to support the information studied in a topic. In their final report, they must acknowledge the information sources they used (PO: i).

- Use of the following tools: Virtual machines, IDEs, and version control in multiple laboratory sessions (PO: k).

- Exercises covering the topics of the course (PO a)

During these activities the teaching staff reviews the student work in the class, supervises the lab sessions, answers questions in course forum, maintains office hours and calls for plenary office hours

upon demand.

ASSESSMENT SYSTEM

The assessment will consist on the following activities:

Theoretical evaluation (2 points):

- Written exam that will assess the acquisition of the needed theoretical concepts and the maturity of the student when presenting them in a clear and proper way with a correct employ of the scientific terms used in the context of the subject (PO: a, k and g)

Lab assessment (8 points):

- Lab exercises (2 points) which structure will be designed with a twofold aim. On the one hand, their objective is to guide the students throughout their learning process and, on the other hand, they will be used to assess the actual acquisition of the necessary skills. Presentations regarding the application of the technologies to real use cases.

- Lab Project (in pairs) (6 points) where the students will show their mastering of the practical concepts of the subject. Evaluation will take place during the week for deliveries (at the end of the term) or the day of the exam.

% end-of-term-examination:	0
% of continuous assessment (assigments, laboratory, practicals…):	100

BASIC BIBLIOGRAPHY

 - Tom White Hadoop : the definitive guide, O'Reilly, 2009

 - George F. Coulouris Distributed systems : concepts and design, Addison-Wesley, 2005

BASIC ELECTRONIC RESOURCES

 - Chambers, Bill ; Zaharia, Matei . Spark: Big Data Processing Made Simple: https://learning.oreilly.com/library/view/spark-the-definitive/9781491912201/?ar
 - Mike Ryan, Federico Lucifredi . AWS System Administration: https://learning.oreilly.com/library/view/aws-system-administration/9781449342562/?ar
 - Tom White . Hadoop : the definitive guide: https://www.oreilly.com/library/view/hadoop-the-definitive/9780596521974/?ar