Universitat Autònoma de Barcelona

2020/2021

Management of Data-Processing Infrastructures

Code: 104352
ECTS Credits: 6

Degree	Type	Year	Semester
2503758 Data Engineering	OB	3	1

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities.

Contact

Name: Remo Suppi Boldrito
Email: Remo.Suppi@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: Yes
Some groups entirely in Spanish: No

Prerequisites

Recommendation: to passed the subjects of Fundamentals of Computer Science, Fundamentals of Programming, Computer Networks and the Internet.

Objectives and Contextualisation

The objective of this subject is to present the main characteristics of the infrastructures for the Big Data processing from three point of view: computation, storage and communication. The concepts and use cases of the main technologies used for massive data infraestructures will be presented, as well as some aspects related to the structure and location (IoT, Edge, Cloud, Distributed)

cost, energy consumption, redundancy, scalability and availability.

Competences

- Conceive, design and implement efficient and secure data storage systems.
- Search, select and manage information and knowledge responsibly.
- Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Learning Outcomes

1. Deploy and manage data-processing infrastructures and the services associated with them.
2. Search, select and manage information and knowledge responsibly.
3. Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Content

1. Introduction to IT infrastructures for Big Data: Computing. Storage, Networks, Architecture (IoT, Edge, Cloud/Cluster/Distributed). Energy consumption and Cost.
2. Computing: Virtualization of processors and operating system. Hypervisors (KVM, HyperV, VWmare, VBox, ...), Containers (Docker and LXC) and serverless computing (Firecracker).
3. Storage: redundancy, scalability, high availability. NAS & DAS. Deployment and management
4. Networks: High throughput & very low latency networks. Software-defined networking. Linux Bridges/VXLAN/802.1Q. Deployment and Management.
5. Cluster and Cloud Infrastructures: Deployment (Manual, Terraform & Ansible), Administration and Management (services, security, monitoring, and tuning).

Methodology

The subject contains three sections where each will have an appropiated methodology to the type of teaching:

Conceptual classroom: theoric and conceptual aspects of the contents will be explained.

Applied concepts: short works will be done on virtualized infrastructure in the covered topics that will be developed during the course.

Practices: group sessions of 2 students to resolve problems within the practices sessions.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Applied Concepts	11.5	0.46	2, 1
Conceptual Classroom	22.1	0.88	2, 1
Pràctices	11.5	0.46	1, 3
Type: Autonomous			
Individual study	100	4	2, 1

Assessment

a) Evaluation activities

The grade will be based on the evaluation of the different activities of the subject:

General concepts: individual quiz of general concepts developed in the subject.

Applied concepts: individual work submitted to the Virtual Campus.

Practices: collaborative work and personal work developed during the sessions.

Important: the activity of Practices & Applied Concepts are submitted during the course and are not recoverable.

b) Calendar of evaluation activities

The evaluation activities will be continued and the submitted through the Virtual Campus. The dates of continuous evaluation and work delivery will be published at Virtual Campus.

The student will be informed in the Virtual Campus about possible changes since this is the information interchange platform between teachers and students.

c) Recovery process

If the student don't pass the individual evaluation of general concepts and with (practice + applied concepts)/2 grade >= 5 points then an additional quiz will be programed.

The grades will compute the indicated percentage >= 5 points. Otherwise and after the recover quiz the student don't reach 5 points in each part, the student will not pass the subject and as a final grade will have the equivalent weighted grade if <=5 or 4.5 if the grade weighted this note is greater >=5.

d) Review of qualifications

For each evaluation activity, a place, date and time of revision will be fixed. The student can review the activity with the teacher. If the student does not apply for this review, this activity will not be reviewed later.

e) Qualifications

Honor Grade. Honor grade is the decision of the staff. The regulations of the UABindicatethat HG can only be granted to students who have obtained a final grade >= 9.00 and only can be assigned up to 5% of HG of the total number of students enrolled.

The award of GH is considered a merit and sign of excellence and is reserved for students who meet the requirements and not will be assigned automatically.

In the case of not attending any evaluation the student will have a Not Graded as the final grade of the subject.

f) Irregularities by the student, copy and plagiarism

Without prejudice to other disciplinary measures and in accordance with current academic regulations, irregularities committed by a student in an evaluable activity will have a grade = zero (0).

These evaluation activities qualified with zero (0) will not be recoverable. If this activity is necessary to pass compulsory evaluation activities, this subject will be suspended directly (without the opportunity to recover it in the same course).

These irregularities include, among others:

- the total or partial copy of a work, report, or any other evaluation activity;
- let copy;
- present a group work not done entirely by the members of the group (applied to all members, not only those who have not worked);
- present as own materials prepared by a third party, even if they are translations or adaptations, and in general works with non original and exclusive contents of the student;
- the utilization of communication devices (such as mobile phones, smart watches, camera pens, etc.) during individual evaluation quiz;
- talk with other people during individual quiz;
- copy or attempt to copy from other students during the evaluation quiz;
- use or attempt to use writings related to the subject during the individual quiz.

In future editions of this subject, students with this irregular actions can't compesate activities from previous year.

In summary: the plagiarism (or try) in any of the evaluation activities is equivalent to don't pass the subject & this action invalidate compensatory activities in subsequent courses.

h) Students that don't pass the subject in previous year.

These students, with a practices grade >= 5 can compensate the practices of the curent year.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Applied Concepts	35%	1.7	0.07	2, 1
General concepts	30%	1.5	0.06	2, 1
Labs	35%	1.7	0.07	1, 3

Bibliography

IT Infrastructure Architecture - Infrastructure Building Blocks and Concepts. Sjaak Laan.

IT Architecture For Dummies 1st Edition, Kalani Kirk Hausman & Susan L. Cook.

Building a Future-Proof Cloud Infrastructure: A Unified Architecture for Network, Security, and Storage Services. Silvano Gai

Learn Ansible: Automate cloud, security, and network infrastructure using Ansible 2.x. Russ McKendrick.

Encyclopedia of Cloud Computing. San Murugesan, Irena Bojanova

Cloud Computing: Principles and Paradigms. Rajkumar Buyya, James Broberg, Andrzej Goscinski

Fog and Edge Computing: Principles and Paradigms. Rajkumar Buyya, Satish Narayana Srirama

Scheduling of Large-Scale Virtualized Infrastructures: Toward Cooperative Management. Flavien Quesnel

Cloud Computing. Web-Based Dynamic IT Services. Christian Baun, Stefan Tai, Marcel Kunze, Jens Nimis

Big Data made easy. Michael Frampton.

Big Data 4D. Judith Hurwitz, Alan Nugent, Dr. Fern Halper, and Marcia Kaufman