

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Arquitectura Big Data
Código	DTC-MBD-511
Título	Máster en Big Data. Tecnología y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics
Impartido en	Máster en Big Data. Tec. y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics [Primer Curso]
Nivel	Master
Cuatrimestre	Semestral
Créditos	3,0 ECTS
Carácter	Obligatoria
Departamento / Área	Departamento de Telemática y Computación

Datos del profesorado	
Profesor	
Nombre	David Contreras Bárcena
Departamento / Área	Departamento de Telemática y Computación
Despacho	Alberto Aguilera 25 [D-302]
Correo electrónico	davidcb@comillas.edu
Teléfono	4235

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
<p>Aportación al perfil profesional de la titulación</p> <p>This course takes you through the different technologies of Big Data, from the initial approach of data storage and processing, BI universe tools, to the dimensioning of a functional cluster.</p> <p>By the end of the course, students will:</p> <ul style="list-style-type: none"> • Understand the basic principles behind distributed systems. • Understand the pieces that make up the pillars of a Big Data architecture. • Have a practical experience sizing a cluster based on use cases. • Have well-formed criteria to choose how to deploy a Big Data cluster with the current technological tools and options.
<p>Prerequisitos</p> <p>Students willing to take this course should be familiar with basic Linux commands and knowledge of computer architecture</p>

Competencias - Objetivos

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Theory

1. Distributed systems

1.1 Distributed processing concepts: software clusters, multithread, multiprocess, high availability

1.2 Data replication

1.3. Cluster ICAI. Architecture and connections.

1.4. Big Data Projects. Cycle of life. Professionals profiles.

1.5. Tools. Virtualization. Containers.

2. Main Big Data Tools Architecture

2.1. HDFS

2.2. YARN and Map Reduce

2.3. Kafka

2.4. Main Non-Hadoop Tools

3. Commercial Big Data platforms & Distributions

3.1 Use of Distributions

3.1. On-premise

3.2. Cloud computing

3.3. Monitoring and administration

4. Design of Big Data architectures

4.1. Methodology of design and sizing of a cluster

Practice

The contents have been designed from a basic starting level, with the purpose of introducing participants to the technologies and tools most used in Big Data. Making an immersion to the most important concepts and demonstrations and practical exercises in each session.

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

GUÍA DOCENTE

2021 - 2022

Calificaciones

Grading

- Final exam will account for 70%.
- Lab will account for the remaining 30%.
- In order to pass the course, the mark of the final exam must be greater or equal to 5 out of 10 points.

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

Textbook

- Notes prepared by the lecturer (available in Moodle)
- Hadoop: The Definitive Guide, 4th Edition, ISBN-13: 978-1491901632
- Learning Spark: Lightning-Fast Big Data Analysis 1st Edition ISBN-13: 978-1491901632

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

<https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>