



Présentation

Code interne : EIN9-ISC12

Description

L'objectif de ce cours est d'apprendre aux étudiants les concepts de base nécessaires pour traiter des larges quantités de données de façon distribuée.

Pour cela, les points suivants seront discutés :

1. Introduction au big data
2. Le framework Hadoop et le système de fichiers distribué HDFS
3. Le paradigme MapReduce
4. La création d'applications avec Spark
5. Introduction aux bases de données NoSQL

Heures d'enseignement

CI	Cours Intégrés	16h
----	----------------	-----

Syllabus

Management de masse de données:

Présentation sur les big data: Problématiques et défis

Stockage

Traitement et interrogation (NoSQL)

Infrastructures:

Virtualisation

Infra. de type cloud

Technologies: Hadoop

Visualisation:

Représentation, navigation, corrélation

Analyse et extraction de données

Modalités de contrôle des connaissances

Évaluation initiale / Session principale - Épreuves

Type d'évaluation	Nature de l'épreuve	Durée (en minutes)	Nombre d'épreuves	Coefficient de l'épreuve	Note éliminatoire de l'épreuve	Remarques
Contrôle Continu Intégral	Contrôle Continu			1		Evaluation sur moodle
Projet	Soutenance			1		

Infos pratiques

Contacts

Francieli Zanon boito

✉ Francieli.Zanon_Boito@bordeaux-inp.fr