

Infrastructure for BigData



Composante
École Nationale
Supérieure
d'Électrotechnique
d'Électronique
d'Informatique
d'Hydraulique
et des
Télécommunications

En bref

- **Code Ametys:** N9EN22D
- **Ouvert aux étudiants en échange:** Oui

Présentation

Objectifs

Partie big data (5C+5TP)

- C1) introduction au big data + Hadoop
- C2) Hadoop
- TP1) Hadoop
- C3) Spark
- C4) Spark
- TP2) Spark
- C5) Spark-streaming + Storm
- TP3) Spark-streaming
- TP4) passage à l'échelle
- TP5) passage à l'échelle + SparkML

Description

Le but est d'introduire aux étudiants les technologies logicielles permettant le déploiement et l'exécution dans des infrastructures de type cluster (grappes de machines) de calculs pour le traitement et l'exploitation de grands volumes de données (big data). Sont abordés les concepts et outils liés au big data tels que le traitement parallèle de données massives (Hadoop, Spark), le traitement temps réel de données (Spark-streaming, Storm), et le passage à l'échelle de ces traitements dans les infrastructures cluster.