

# Infrastructure pour le Cloud et le Big Data

 **ECTS**  
3 crédits

 **Volume horaire**

## Présentation

### Description

Le but est d'introduire aux étudiants les technologies logicielles permettant le déploiement et l'exécution de calculs dans des infrastructures de type cluster (grappes de machines). De telles infrastructures sont très largement utilisées dans les domaines du cloud computing où des datacenters permettent l'hébergement externalisés de services, du big data et du machine learning pour le traitement et l'exploitation de grands volumes de données.

La première partie aborde les concepts et outils liés au cloud computing tels que la virtualisation (KVM), les conteneurs (Docker), les outils d'administration (OpenStack, Kubertenes) et les principaux services fournis par les opérateurs du cloud (AWS).

La seconde partie aborde les concepts et outils liés au big data tels que le traitement parallèle de données massives (Hadoop, Spark) et le traitement temps réel de données (Spark-streaming, Storm).

### Objectifs

A la fin de ce module, l'étudiant devra avoir compris et pourra expliquer (principaux concepts) :

- Concepts généraux des infrastructures de calcul du cloud et du big data

- Principes des infrastructures virtualisées
- Les services du cloud
- Les outils d'utilisation des infrastructures du cloud
- Principes des plate-formes de traitement big data (map-reduce, chaines de filtrages)
- Les environnements de traitement big data (Hadoop, Spark, Storm)

L'étudiant devra être capable de :

- Utiliser des plate-formes de virtualisation
- Utiliser des plate-formes de cloud
- Programmer des applications big data
- Exécuter des applications big data dans une infrastructure de calcul

### Pré-requis nécessaires

Algorithmique, programmation en Java, environnement Linux (commandes shell).

### Évaluation

L'évaluation des acquis d'apprentissage est réalisée en continu tout le long du semestre. En fonction des enseignements, elle peut prendre différentes formes : examen écrit, oral, compte-rendu, rapport écrit, évaluation par les pairs...

### Infos pratiques

## Lieu(x)

 Toulouse