

# Systèmes de gestion de masses de données : Architectures, fonctionnalités et critères de choix

## Présentation

### Objectifs

En 2 jours, cette formation concernant les architectures de gestion de données massives, vous permettra de :

- Comprendre les concepts clés
- Acquérir les principales compétences
- Faire des choix appropriés et cohérents par rapport à l'évolution de l'offre des éditeurs

### Public visé et prérequis

Informaticiens, Data Scientist, Managers, Dirigeants.

Cette formation s'adresse aux personnes ayant des connaissances en systèmes de gestion de données.

**ATTENTION :** Merci de candidater auprès du Responsable Pédagogique de la Formation

### Compétences visées

- Mettre en pratique les méthodes de répartition de données des SGBD parallèles afin de minimiser les coûts de communication et assurer le passage à l'échelle des applications d'une organisation.
- Exploiter efficacement des systèmes de gestion de données en environnement parallèle et réparti dans une perspective d'aide à la décision au sein d'une organisation.
- Choisir de manière pertinent un système de gestion de données dans des multiples environnements (mono-processeur, parallèle, réparti) homogènes et hétérogènes pour un fonctionnement optimal des applications d'une organisation.

### Programme

#### Partie théorique :

- Introduction aux problèmes de gestion de données
  - Du SGF (Système de Gestion de Fichiers) au SGBD (Système de Gestion de Bases de Données)
- Traitement parallèle de masses de données
  - SGBD parallèles versus Systèmes de Gestion de Données Cloud
- Intégration de masses de données hétérogènes
  - Systèmes d'intégration virtuelle de données versus systèmes de gestion de données de type multi-store
- Critères clés de choix d'un système de gestion de données.

#### Partie pratique :

- Illustration du concept d'indépendance données-programmes
- Exemples de programmes via les interfaces de différents systèmes de gestion de données (e.g. SGBD relationnel, Hadoop MapReduce, Hive, Spark, ...)
- Etude de cas : choix d'un système de gestion de données approprié respectant un cahier des charges.

### Modalités d'évaluation

Un questionnaire d'évaluation est complété par chaque participant.

### Validation

Attestation de validation en fin de formation

### Responsable pédagogique

Abdelkader HAMEURLAIN

[abdelkader.hameurlain@irit.fr](mailto:abdelkader.hameurlain@irit.fr)

Composante de rattachement

Faculté des Sciences et d'Ingénierie

### Inscription Administrative

MISSION FORMATION CONTINUE ET APPRENTISSAGE

Sophie COBOURG

[sophie.cobourg@univ-tlse3.fr](mailto:sophie.cobourg@univ-tlse3.fr)

Tél : 05 61 55 87 12

### Prix

800 €

Tarif de groupe : nous consulter

### Déroulement de la formation

**Durée : 2 jours**

**Dates : 18 - 19 Avril 2019**

**27 - 28 Mai 2019**

**6 - 7 Juin 2019**

**Lieu : Univ. Paul Sabatier, MFCA  
Entreprise**

**Nombre de participants :**

Minimum : 5 personnes

Maximum : 10 personnes

*L'établissement se réserve le droit d'annuler la formation si le nombre minimal de participants n'est pas atteint.*

### Modalités d'enseignement

Présentiel

Les concepts introduits et les méthodes présentées sont illustrés par de nombreux exemples pratiques en s'appuyant sur des systèmes de gestion de données les plus représentatifs.

70% enseignement théorique avec des exemples pratiques.

30% enseignement pratique-démonstration.

**Intervenants :**

Abdelkader HAMEURLAIN

Franck MORVAN

Shaoyi YIN

Riad MOKADEM