

# Opleiding : Hadoop, toepassingen ontwikkelen voor Big Data

*Praktijkcursus - 4d - 28u00 - Ref. APH*

*Prijs : 2520 € V.B.*

In deze praktische cursus leert u applicaties ontwikkelen waarmee u gedistribueerde gegevens in batchmodus kunt verwerken. Je zult gegevens in heterogene formaten verzamelen, opslaan en verwerken met Apache Hadoop om verwerkingsketens op te zetten die geïntegreerd zijn in je informatiesysteem.

## Pedagogische doelstellingen

Aan het einde van de training is de deelnemer in staat om:

- ✓ Een op MapReduce gebaseerd programma bouwen
- ✓ Hadoop HBase integreren in een bedrijfsworkflow
- ✓ Travailler avec Apache Hive et Pig depuis Hadoop Distributed File System (HDFS)
- ✓ Een taakgrafiek gebruiken met Hadoop

## Doelgroep

Concepteurs, développeurs.

## Voorafgaande vereisten

Goede ervaring met Java-ontwikkeling. Kennis van webarchitectuur is een pluspunt.

## Praktische modaliteiten

### Praktisch werk

Ontwikkeling van Big Data-toepassingen.

### Leer methodes

Hoorcolleges 30%, praktisch werk 70%.

## Opleidingsprogramma

### DEELNEMERS

Concepteurs, développeurs.

### VOORAFGAANDE VEREISTEN

Goede ervaring met Java-ontwikkeling. Kennis van webarchitectuur is een pluspunt.

### VAARDIGHEDEN VAN DE CURSUSLEIDER

De deskundigen die de cursus leiden zijn specialisten op het betreffende vakgebied. Zij werden geselecteerd door onze pedagogische teams zowel om hun vakkennis als hun pedagogische vaardigheden voor elke cursus die zij geven. Zij hebben minstens vijf tot tien jaar ervaring in hun vakgebied en oefenen of oefenden verantwoordelijke bedrijfsfuncties uit.

### BEOORDELINGSMODALITEITEN

De cursusleider beoordeelt de pedagogische vooruitgang van de deelnemer gedurende de gehele cursus aan de hand van meerkeuzevragen, praktijksituaties, praktische opdrachten, ...  
De deelnemer legt ook van tevoren en naderhand een test af ter bevestiging van de verworven kennis.

## 1 Grote gegevens

- De reikwijdte van big data definiëren.
- De rol van het Hadoop-project.
- De basisconcepten van big data projecten.
- Inleiding tot cloud computing.
- Het verschil tussen private en publieke cloud computing.
- Big data-architecturen gebaseerd op het Hadoop-project.

### Demonstratie

Gebruik van Hadoop en GoogleApp.

## 2 Gegevens verzamelen en MapReduce toepassen

- Analyse van gegevensstromen binnen het bedrijf.
- Gestructureerde en ongestructureerde gegevens.
- De principes van semantische analyse van bedrijfsgegevens.
- Op MapReduce gebaseerde taakgrafiek.
- Granulariteit gegevensconsistentie.
- Gegevens overbrengen van een persistentiesysteem naar Hadoop.
- Gegevens overbrengen van een cloud naar Hadoop.

### Praktisch werk

Het verzamelen van klantgegevens beheren met MapReduce. De Yarn implementatie configureren. Een op MapReduce gebaseerde taak ontwikkelen.

## 3 Gegevensopslag met HBase

- Verschillende soorten XML-databases.
- Gebruikspatronen en hun toepassing op de cloud.
- Toepassing van Hadoop-database binnen een workflow.
- Gebruik van Hive/Pig projecten.
- Gebruik van het HCatalog project.
- De HBase Java API.

### Praktisch werk

Wijzigingen in een gegevenscatalogus van leveranciers beheren.

## 4 Gegevensopslag op HDFS

- Gebruikspatronen en hun toepassing op de cloud.
- Architectuur en installatie van een HDFS-systeem, journaal, NameNode, DataNode.
- Operaties, bestellingen en orderbeheer.
- De Java HDFS API.
- Gegevensanalyse met Apache Pig.
- De taal Pig Latin. Apache Pig gebruiken met Java.
- Query's uitvoeren met Apache Hive.
- GegevensrePLICatie. Het delen van gegevens op een HDFS-architectuur.

### Praktisch werk

Een gedeelde cliëntrepository op Hadoop beheren. De visualisatieconsole gebruiken.

## PEDAGOGISCHE EN TECHNISCHE MIDDELEN

- De gebruikte pedagogische middelen en cursusmethoden zijn voornamelijk: audiovisuele hulpmiddelen, documentatie en cursusmateriaal, praktische oefeningen en correcties van de oefeningen voor praktijkstages, casestudies of reële voorbeelden voor de seminars.
- Na afloop van de stages of seminars verstrekt ORSYS de deelnemers een evaluatievragenlijst over de cursus die vervolgens door onze pedagogische teams wordt geanalyseerd.
- Na afloop van de cursus wordt een presentielijst per halve dag verstrekt, evenals een verklaring van de afronding van de cursus indien de stagiair alle sessies heeft bijgewoond.

## TOEGANGSMODALITEITEN EN TERMIJNEN

De inschrijving dient 24 uur voor aanvang van de cursus plaatsgevonden te hebben.

## TOEGANKELIJKHEID VOOR MINDERVALIDEN

Is voor u speciale toegankelijkheid vereist? Neem contact op met mevr. FOSSE, contactpersoon voor mindervaliden, via het adres psh-accueil@ORSYS.fr om uw verzoek en de haalbaarheid daarvan zo goed mogelijk te bestuderen.

## 5 Springgegevens Hadoop

- Inleiding tot Spring en Spring Data.
- De Hadoop-namespace voor Spring.
- Spring gebruiken om Hadoop-configuratie te vereenvoudigen.
- Gedistribueerde cacheconfiguratie.
- Jobdefinitie en afhankelijkheid tussen jobs.
- Integratie van tools (Pig, Hive, enz.).

### Praktisch werk

Herontwerp van het beheer van de leveranciersgegevenscatalogus met behulp van Spring Data.

## Data en plaats

### KLAS OP AFSTAND

2026 : 26 mei, 6 okt.

### PARIS LA DÉFENSE

2026 : 26 mei, 6 okt.