

Opleiding : Flink, toepassingen ontwikkelen voor Big Data

Praktijkcursus - 3d - 21u00 - Ref. FKB

Prijs : 2010 € V.B.

Apache Flink is een recent big data framework. Het vereenvoudigt de verwerking van grote real-time stromen en batchverwerking van enorme hoeveelheden gegevens (op Hadoop HDFS, Amazon S3, MongoDB, enz.). Deze cursus stelt je in staat om Flink te installeren en verschillende big data processen uit te voeren in Java.

Pedagogische doelstellingen

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

- ✓ De fundamentele concepten van Flink onder de knie krijgen
- ✓ Toepassingen ontwikkelen met de DataSet en DataStream API's
- ✓ Gedistribueerde gegevensverwerking met Flink en Hadoop
- ✓ Gegevens exploiteren met Tabel-API
- ✓ Een eerste benadering van machinaal leren

Doelgroep

Ontwikkelaars, architecten.

Voorafgaande vereisten

Goede kennis van Java.

Praktische modaliteiten

Praktisch werk

Praktische toepassing van de concepten die in de cursus worden behandeld met behulp van de taal Java.

Opleidingsprogramma

DEELNEMERS

Ontwikkelaars, architecten.

VOORAFGAANDE VEREISTEN

Goede kennis van Java.

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.

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.

1 Inleiding tot Apache Flink

- Geschiedenis van het kader.
- De verschillende versies van Flink.
- Vergelijking met de Apache Hadoop en Apache Spark omgevingen.
- De verschillende Flink-modules.

Praktisch werk

Flink installeren en configureren. Een eerste voorbeeld met woordentelling uitvoeren.

2 Gegevensverwerking met de DataStream API

- Runtime-omgeving en gegevensbronnen.
- Transformaties: Map, FlatMap, Filter, KeyBy, Reduce, enz.
- Bewerkingen op meerdere stromen: Union, Cogroup, Connect, Join, Iterate, etc.
- Vensterbewerkingen: Wereldwijd, Tuimelen, Schuiven, Sessie...
- Aangepaste fysieke partitionering, randomisatie, herbalancering en formaataanpassing.
- DataSink en connectoren: Kafka, X (voorheen Twitter), Elasticsearch...

Praktisch werk

Verbruik en verwerking van verschillende gegevensstromen.

3 Gegevens verwerken met de Batch API

- De verschillende soorten gegevensbronnen.
- Transformaties en aggregaties.
- Gegevens schrijven.
- DataSink en connectoren: HDFS, S3, Avro, MongoDB.

Praktisch werk

Manipuleer DataSets van meerdere gegevensbronnen.

4 Gegevensverwerking met behulp van de Table API

- Opgeslagen tabellen opslaan en lezen.
- Operatoren: selectie, filter, join, orderBy...
- Gebruik SQL op de gegevensstroom.
- Omgaan met complexe gebeurtenissen.

Praktisch werk

Een analyse opzetten met SQL op een gegevensstroom.

5 API Flink grafiek - Gelly

- Wat is een grafiek?
- De verschillende operaties.
- Grafieken maken.
- Grafiektransformaties.
- Presentatie van verschillende algoritmen.

Praktisch werk

Omgaan met de API aan de hand van verschillende voorbeelden.

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.

6 Flink implementeren

- Flink op YARN-configuraties.
- Een cluster starten en stoppen.
- Dien een opdracht in bij Flink.
- Flink op Google Cloud.
- Flink op AWS.

Praktisch werk

Een cluster met meerdere knooppunten configureren en een applicatie implementeren.