

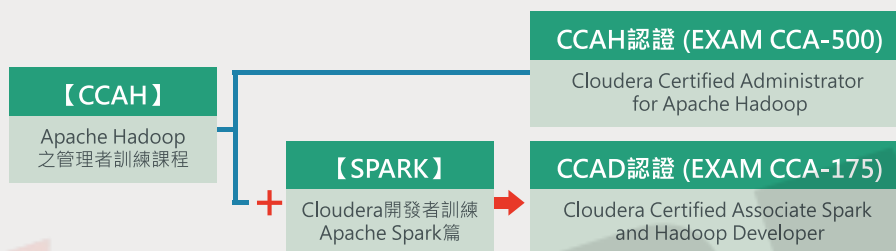
Cloudera

Cloudera認證介紹

大數據時代來臨，不論您的公司是否正在尋找相關技術或是您的客戶正在尋求解決方案，Cloudera所提供的技術及架構，都能為您帶來競爭優勢，現在就開始了解Cloudera並取得相關認證，成為一個新技術的領導者，也是向您的客戶或您的雇主提出技術能力的最好證明。

Cloudera認證學習流程圖

學習路徑1 針對想成為Hadoop管理者及開發者所規劃



學習路徑2 針對Spark開發者或已具備Hadoop實作經驗者



CCAHA

Apache Hadoop之管理者訓練課程

Cloudera Administrator Training for Apache Hadoop-Configuring, Deploying, and Maintaining a Hadoop Cluster

時數：28 小時 | 費用：64,000 元 | 點數：19 點 | 教材：Cloudera 原廠教材



獨家課程



CCAHA認證

- 適合對象
1. 系統管理人員
 2. IT經理
 3. RDBMS資料庫管理人員

預備知識 基礎的Linux操作(不需要任何Hadoop經驗)

課程目標 本課程將教您關於管理、操作、維護Hadoop叢集所需必要步驟的廣泛介紹。從安裝、組態到負載平衡與效能調校，本課程讓您在面對管理正式環境下的Hadoop叢集時，可以從容面對各項挑戰並做好最佳準備。本課程結束後，您將具備考取Cloudera Certified Administrator for Apache Hadoop(CCAHA)認證的專業知識。

[實作練習Hadoop]

透過講師引導式討論、互動與實際操作練習，您將一覽Hadoop生態系全貌，並學習到下列內容：

- 為Hadoop叢集做出正確的硬體規格與基礎架構規劃
- 合適地組態與部署Hadoop叢集，讓Hadoop叢集可以與現存的資料中心整合
- 使用Flume將動態生成的檔案內容和使用Sqoop將RDBMS的資料載入Hadoop叢集
- 在多使用者環境下，組態FairScheduler讓Hadoop叢集提供服務層級協定
- 提供最佳實踐建議關於準備與維護正式環境下的Hadoop叢集
- 分析、偵錯、效能調校與解決Hadoop叢集所遇到的問題

備註事項 白天班及假日班課程含供餐服務



CDA

資料分析人員-使用Pig及Hive and Impala操作Hadoop叢集課程

Cloudera Data Analyst Training : Using Pig, Hive, and Impala with Hadoop

時數：28 小時 | 費用：64,000 元 | 點數：19 點 | 教材：Cloudera 原廠教材



適合對象

1. 資料分析人員
2. 商業分析人員
3. 程式開發人員
4. 系統管理人員

預備知識

SQL 語言
基本Unix/Linux操作
(不需要具備任何Apache Hadoop經驗)

課程目標 本課程將教您如何使用Apache Pig、Hive與Cloudera Impala，讓您可以在Big Data環境裡使用傳統的資料分析與商業智慧技術。Cloudera將展示如何使用類SQL語言與常見的Script語言，讓資料專家可以存取、操作、分析複雜的資料集。

[提升您的生態系技術]

Apache Hive讓資料分析人員、資料庫管理人員和其他沒有Java程式撰寫技術的人員存取多結構化資料。Apache Pig將常見的Script語言技術應用在Hadoop叢集。

Cloudera Impala讓儲存在Hadoop叢集的資料透過原生SQL環境提供即時、互動的分析能力。

[實作練習Hadoop]

- 透過講師引導式討論、互動與實際操作練習，您將一覽Hadoop生態系的全貌，並學習到下列內容：
- Apache Hadoop的基礎與使用Hadoop工具進行資料的ETL(擷取、轉換與載入)、擷取與處理
- 使用Apache Pig結合多個資料集與分析異質資料
- 使用Apache Hive組織資料到表格裡，進行轉換以及簡化複雜的查詢操作
- 使用Cloudera Impala與SQL對儲存在HDFS或HBase的大量資料進行即時與互動式分析
- 如何因應不同工作選擇最合適的Hadoop分析工具

備註事項 白天班及假日班課程含供餐服務



SPARK

Cloudera開發者訓練Apache Spark篇

Cloudera Developer Training for Apache Spark

時數：21 小時 | 費用：52,000 元 | 點數：15 點 | 教材：Cloudera 原廠教材



適合對象 大數據程式開發人員、大數據系統工程師、資料工程師、資料科學家

預備知識

Python或Scala語言基礎
基本Linux系統知識

課程目標 透過講師引導的討論、上課互動與實作練習，強化學員對Hadoop生態體系的專業能力。了解到Spark身為MapReduce的下一代後繼者，所能提供的新功能。Spark框架支援串流處理與複雜、迭代演算法，在最佳設計下可以比傳統的MapReduce程式快上100倍。

學員在本課程中將學習到下列內容：

- 使用Spark Shell進行互動式資料分析
- 了解Spark RDD(Resilient Distributed Datasets)的特色
- 了解Spark如何在Hadoop Cluster執行程式
- 了解Spark如何進行平行地執行工作
- 撰寫Spark應用程式
- 使用Spark處理串流資料

備註事項 白天班及假日班課程含供餐服務



Dev

Spark與Hadoop開發者訓練

Developer Training for Spark and Hadoop

時數：28 小時 | 費用：64,000 元 | 點數：19 點 | 教材：Cloudera 原廠教材



適合對象 使用Spark在Hadoop之上撰寫程式的開發人員
希望取得CCP: Data Engineer認證的相關人員



預備知識

1. 熟悉Python或Scala任一程式語言
2. 基本的Linux Command Line操作
3. 基本的SQL知識



課程目標 在這4天由講師引導的實作課程裡，學員將學習到在Hadoop Cluster裡資料如何分散地儲存與處理，如何使用Sqoop與Flume將資料載入到Hadoop Cluster，以及如何使用Spark處理分散在Hadoop Cluster裡的資料，同時也可以學習到如何使用Impala與Hive將Hadoop Cluster的資料以結構化方式存取。最後學員可以知道如何因應不同的使用樣式，選擇最適當的資料儲存格式。

本課程同時也是Cloudera公司所推出的CCAD認證與資料工程師認證(CCP: Data Engineer)的最佳準備課程之一。

備註事項 白天班及假日班課程含供餐服務

