

Data Warehousing on AWS

AWS 課堂培訓

課程說明

Data Warehousing on AWS 介紹使用 Amazon Redshift (AWS 中 PB 規模的資料倉儲) 設計雲端資料倉儲解決方案的概念、策略和最佳實務。本課程示範如何使用 Amazon DynamoDB、Amazon EMR、Amazon Kinesis 和 Amazon S3 等 AWS 服務，收集、存放和準備資料倉儲資料。此外，本課程還示範如何使用 Amazon QuickSight 執行資料分析。

- 課程等級：中級
- 持續時間：3 天

活動

本課程包括簡報、小組練習和實作實驗室。

課程目標

在本課程中，您將：

- 討論資料倉儲的核心概念，以及資料倉儲與大數據解決方案之間的交集
- 啟動 Amazon Redshift 叢集，並使用元件、特點和功能在雲端實作資料倉儲
- 使用 Amazon DynamoDB、Amazon EMR、Amazon Kinesis 和 Amazon S3 等其他 AWS 資料和分析服務，使資料倉儲解決方案更加完善
- 架構資料倉儲
- 識別效能問題、最佳化查詢，以及調校資料庫以獲得更佳效能
- 使用 Amazon Redshift Spectrum 直接從 Amazon S3 儲存貯體分析資料
- 使用 Amazon QuickSight 針對資料倉儲執行資料分析和視覺化任務

目標對象

本課程適用於：

- 資料庫架構師
- 資料庫管理員
- 資料庫開發人員
- 資料分析師
- 資料科學家

Data Warehousing on AWS

AWS 課堂培訓

先決條件

我們建議參加本課程的人員需具備下列先決條件：

- [AWS Technical Essentials](#) (或同等的 AWS 經驗)
- 熟悉關聯式資料庫和資料庫設計概念

立即註冊

立即瀏覽 aws.training 以尋找課程。

Data Warehousing on AWS

AWS 課堂培訓

課程大綱

第 1 天

單元 1：資料倉儲簡介

- 關聯式資料庫
- 資料倉儲概念
- 資料倉儲與大數據的交集
- AWS 中的資料管理概觀
- 實作實驗室 1：Introduction to Amazon Redshift

單元 2：Amazon Redshift 簡介

- 概念性概觀
- 實際應用案例
- 實作實驗室 2：啟動 Amazon Redshift 叢集

單元 3：啟動叢集

- 建置叢集
- 連接至叢集
- 控制存取
- 資料庫安全
- 載入資料
- 實作實驗室 3：最佳化資料庫結構描述

Data Warehousing on AWS

AWS 課堂培訓

第 2 天

單元 4：設計資料庫結構描述

- 結構描述和資料類型
- 單欄式壓縮
- 資料分佈樣式
- 資料排序方法

單元 5：識別資料來源

- 資料來源概觀
- Amazon S3
- Amazon DynamoDB
- Amazon EMR
- Amazon Kinesis Data Firehose
- 適用於 Amazon Redshift 的 AWS Lambda 資料庫載入器
- 實作實驗室 4：將即時資料載入 Amazon Redshift 資料庫

單元 6：載入資料

- 準備資料
- 使用 COPY 載入資料
- 維護資料表
- 並行寫入操作
- 對載入問題進行故障診斷
- 實作實驗室 5：使用 COPY 命令載入資料

Data Warehousing on AWS

AWS 課堂培訓

第 3 天

單元 7：撰寫查詢和調校效能

- Amazon Redshift SQL
- 使用者定義的函數 (UDF)
- 影響查詢效能的因素
- EXPLAIN 命令和查詢計劃
- 工作負載管理 (WLM)
- 實作實驗室 6：設定工作負載管理

單元 8：Amazon Redshift Spectrum

- Amazon Redshift Spectrum
- 設定 Amazon Redshift Spectrum 的資料
- Amazon Redshift Spectrum 查詢
- 實作實驗室 7：使用 Amazon Redshift Spectrum

單元 9：維護叢集

- 稽核記錄
- 效能監控
- 事件和通知
- 實驗室 8：稽核和監控叢集
- 調整叢集大小
- 備份與還原叢集
- 資源標記和限制與約束
- 實作實驗室 9：備份、還原和調整叢集大小

單元 10：分析和視覺化資料

- 視覺化的強大功能
- 建置儀表板
- Amazon QuickSight 版本與功能