

檔 號：
保存年限：

RND 0399

3

健行學校財團法人健行科技大學 函

地址：320桃園縣中壢市健行路
229號

傳 真：03-4598411

聯 絡 人：丁郁文

聯絡電話：03-4581196#3782

電子郵件

：winny_ding@cyu.edu.tw

受文者：國立暨南國際大學

發文日期：中華民國103年10月13日

發文字號：健推字第1030006350號

速別：普通件

密等及解密條件或保密期限：

附件：課程簡章（1030006350A-0.DOCX，共1個電子檔案）

主旨：函知本校推廣教育中心辦理「Big Data/Open Data 數據處理與分析實務」18小時雙認證研習課程，招生相關資訊，敬請 惠予公告周知並鼓勵所屬（轄）人員踴躍報名參加。

說明：

- 一、上課時間：103年11月13、14、15日共計3天。
- 二、上課地點：健行科技大學 商學院2樓 C206A電腦教室。
- 三、學員完成本課程18小時研習並通過考試者可取得兩張國際證照，包含HPATA: Serevr&Storage(伺服器與儲存管理)及MSCS: Big Data, Configuring (巨量資料分析技術師)。並可獲得18小時實務課程研習證明一張。
- 四、招生訊息及課程報名網頁請至本校推廣教育中心網站查閱、下載(<http://eec.uch.edu.tw/web3/index3-1-1.asp?cid=304138>)。
- 五、相關課程資訊詳如附件。

正本：公私立大專校院

副本：本校推廣教育中心

擬辦：

103/10/13
14:52:24

一、將來文上傳本校公文系統，
公告週知。

二、文陳閱後左。

許孟瑜

副教授兼研發處
施君興教授兼研發處
林佑昇

第1頁

共1頁

13

教授兼研發處
林佑昇代為
決行

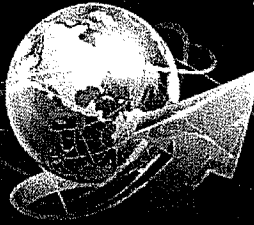
103年10月13日暨收文總字第(03001302)號



3352-4321



Big Data/Open Data 數據處理與分析實務課程 (雙認證)

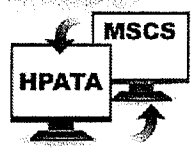


11/13~11/15 // 18小時專業授課!取得認證!

Big Data / Open Data

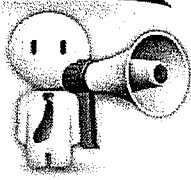
數據處理與分析實務課程

超值內容



- ▶ 兩張認證 (HPATA/MSCS)
- ▶ USB 隨身碟一支(含上課環境)
- ▶ 三天實務課程
- ▶ 18小時實務課程研習證明一張
- ▶ Big Data 系統授權書一張(限教育使用)

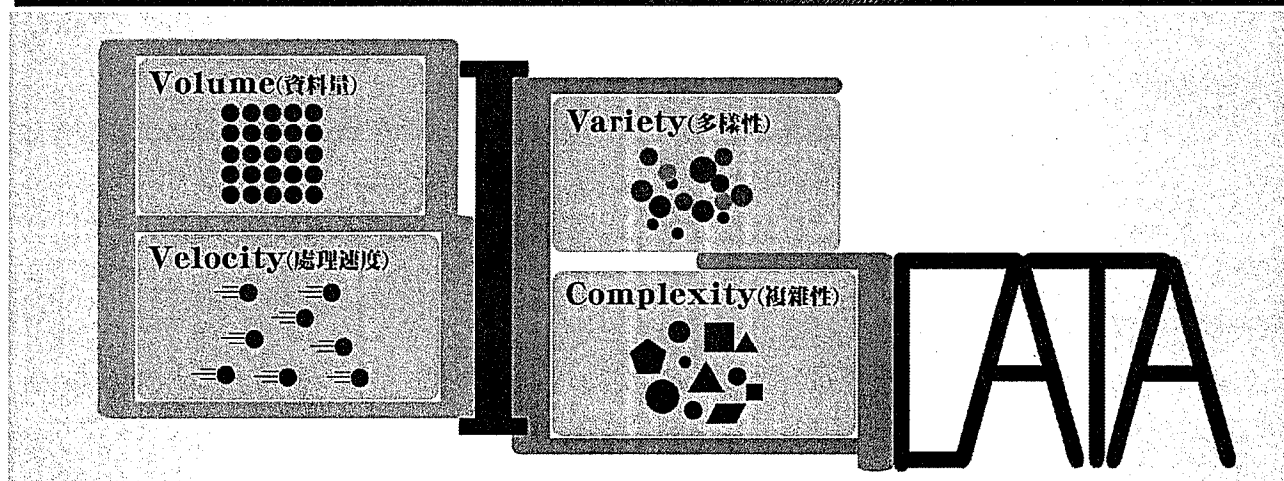
優惠對象



本校師生 → **NT\$8,000**

校外師生 → **NT\$9,500**

商業人士 → **NT\$11,500**



課程名稱：**Big Data/Open Data 數據處理與分析實務課程 (雙認證)**

上課天數：**三天**

上課日期：**11/13,14,15 (四五六)**

上課時間：**9:30-16:30 (11/14 及 11/15 16:30-17:20 考試)**

上課地點：**健行科技大學 商學院 C206A 教室**

地址：**桃園縣中壢市健行路 229 號**



1. 課程介紹

"何謂 Big Data ?"

我們每天建立幾百 G 的資料 — 數量之大，光是過去兩年所建立的資料就佔當今世界總量的 90%。這些資料來源廣泛，像是用於搜集氣候資訊的感應器(物聯網)、社交媒體網站的貼文、數位圖片與影像、採購交易記錄以及行動電話 GPS 訊號等，不一而足，這類的資料就是 Big Data。

Big Data 包括三種層面：巨量、即時性及多樣性。

1. 巨量 (Volume) - 海量資料的特色就在於：龐大。企業資料包羅萬端，很容易便達到數兆位元組，甚至千兆位元組之譜。
2. 即時性 (Velocity) - 海量資料通常具有時效性，一旦串流至企業便須立即使用，方能發揮其最大價值。
3. 多樣性 (Variety) - 海量資料的範疇不僅止於結構化資料，還包含各類非結構化的資料：諸如文字、音訊、視訊、點擊串流 (click stream)、日誌檔等等。

這 3 個資料特性，已經是現在式，而不是未來式。然而該如何解決日漸緊迫的巨量資料處理問題呢？像 Facebook、Twitter 這樣面臨資料量大爆炸的網路公司，開始用 Hadoop、NoSQL 等新興技術來解決問題。

"挑戰還是機會？"

Big Data 不只是一項挑戰，更是絕佳的機會，讓您能夠洞悉新興的資料類型、使企業運作更加靈敏並為過往所無法企及的問題提供解答。但在此之前，這種機會並無實際方法可以掌握。今天，Big Data 平台採用 Hadoop 等技術，能為充滿各種可能性的世界開啟一扇大門。

"Hadoop 技能有高度需求"

由於成長型企業在 Hadoop 和相關技術方面的興趣，同時也推動了企業在巨量資料(big data)方面的專業人力需求，而資料分析師、資料科學家和資料管理師將獲得大鳴大放的機會。

在紐約舉辦的 Hadoop World 會議上，分析師和 IT 經理人一再指出，採用 Hadoop 技術的公司面臨的主要挑戰之一就是技術可用性(skills availability)，並表示具有 Hadoop 技能的人同時也掌握了更多收入。

摩根大通集團(JP Morgan Chase)和 EBay 的 IT 主管利用主題演講的機會，大肆招攬與會學員加入 Hadoop 的行列，這種情況似乎也顯示了這樣的技能在市場供給上多麼有限。eBay 負責搜尋平台的資深副總裁 Hugh Williams 除了在演講中提到 eBay 正在招募 Hadoop 專才以外，同時更邀請現場有興趣的人直接找機會與他對談。摩根大通總經理 Larry Feinsmith 也接著 Williams 的講談內容，更開玩笑地告訴聽眾，摩根大通也雇用 Hadoop 技術專才，同時還願



意付比 eBay 高 10% 的酬勞！Forrester 分析師 James Kobielus 則說，Hadoop 是一種新的資料倉儲型態，是企業中資料的新來源。企業對那些具備足夠 Hadoop 技能的人提供一份優渥的薪資，期望能利用這樣的技術為公司在市場上取得優勢。

Hadoop 允許企業儲存和管理的結構化或非結構化資料的資料量，遠遠大於今日的 RDBMS（關連式資料庫管理系統）所能處理的量，而且經濟實惠。

在這方面技術有需求的企業數量一直不斷成長，為了儲存和分析像 weblogs 的 peta 級資料量、點擊串流資料和社交媒體內容，以對他們的顧客和生意有更深入的了解。由於越來越多企業的投入，也使得高階分析人才需求度不斷增加，這包含在以下方面具有專長的人：像是多變量分析、資料探勘、預測模型、自然語言處理，內容分析，文字分析與和社交網路分析，Kobielus 說。「就廣義的巨量資料（尤其是 Hadoop）來說，這些企業應用的需求同時也驅動了企業對 IT 專才的需求，企業更需要能夠在使用進階分析方法上（像是為了預測和統計模型的 MapReduce 和 R）有經驗的人才，」Kobielus 說，因為這些資料分析師或資料科學家們可以利用 Hadoop 的環境，來處理結構化或非結構化資料，以提商務智慧和洞察力。

企業對 Hadoop 平台的興趣也創造了 Hadoop 平台管理專才的需求，Kobielus 說。這些專才的工作是以 Hadoop 叢集、安全、管理與最佳化來回應企業需求，並確保叢集可用性。在每一個執行 Hadoop 應用程式的背後，他們就是能建立 Hadoop 平台並予以最佳化的關鍵。

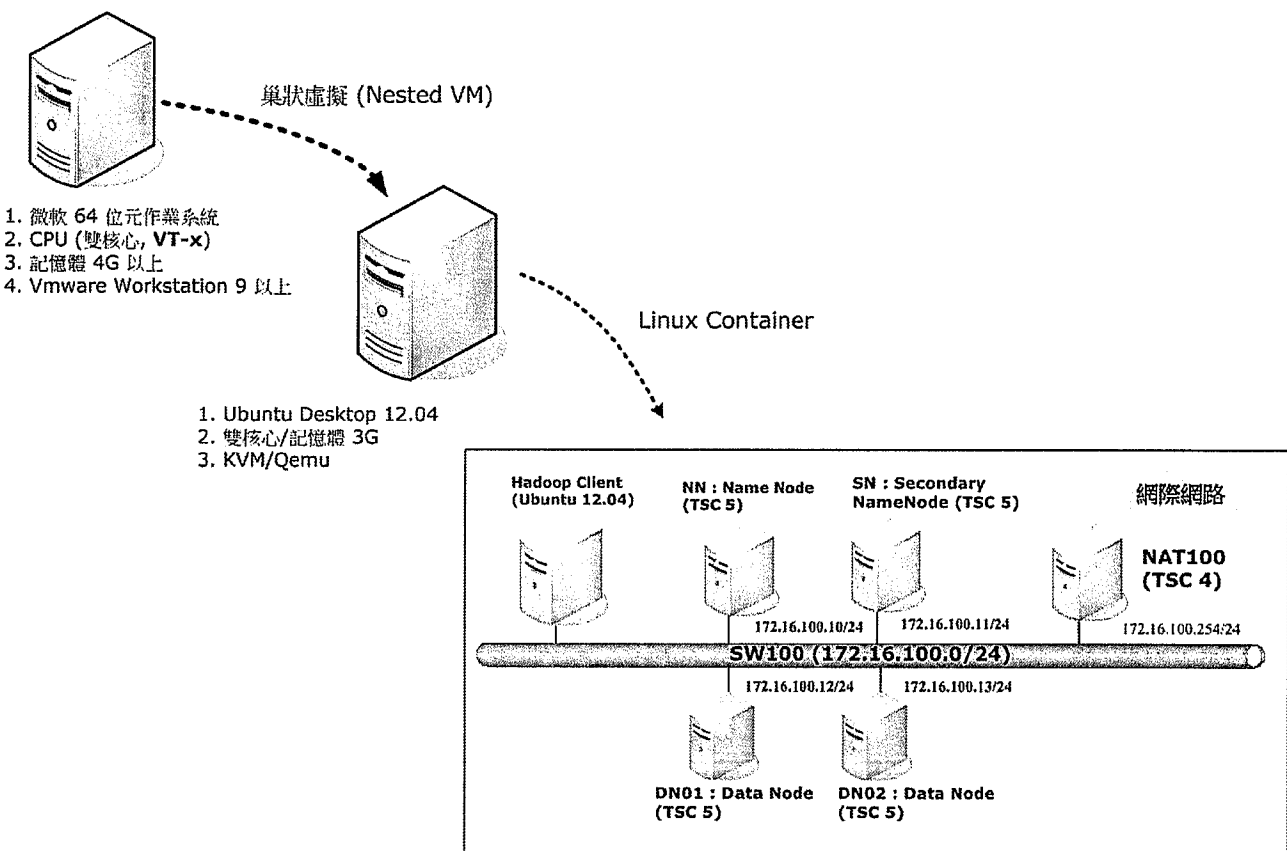
這些專才已經意識到全新世界的來臨，他們和負責管理 Teradata 或 Oracle Exadata 的資料庫管理員是同一批人，而他們正開始重新定義自己，將自己視為 Hadoop 叢集管理員的角色。另一方面，企業不僅期待可以找到儲存管理的專才，也希望可以找到能在關連式資料庫中進行 Hadoop 環境整合的專才。

2. 教學內容

本課程旨在建立雲端運算之大量資料(Big Data)處理、分析、應用的根基，讓參訓學員瞭解正確的觀念與方法，重點並不在講解程式設計的細節，而在於透過體驗式教學方式的實作，經由指令剪貼方式來體驗實際的操作方式，以從體驗中驗證課程所學。

由於 Hadoop 是採用 Java 語言撰寫，對於許多不熟悉 Java 語言的學員來說有相當大的入門障礙，因此本課程額外介紹可與其他語言介接的 Hadoop Streaming 方法，並以實際的案例說明如何讓 Hadoop 與現存的應用程式(如 Shell Script、C 語言撰寫的執行檔、PHP Script、Python Script 等)進行整合實作，期能讓學員學會如何將 Hadoop 這項技術與現存資訊架構進行整合。

上課方式採用巢狀虛擬 (Nested VM) 實作教學環境，使學生在各自的 VMware 虛擬系統中，啟動 Linux Container (LXC) 雲端虛擬系統，得以完全實做多節點 Hadoop 運算分析平台。如下圖：





3. 課程大綱

11/13 09:00~16:30

第一章 認識 Big Data

- 1.1 Big Data 資料特質與應用
 - 1.2 Big Data 關鍵技術
 - 1.3 Hadoop 運算分析系統架構
-

11/14 (上課)09:00~16:30 (考試)16:30~17:20 考試

第二章 Big Data 實作平台

- 2.1 Linux Container 雲端虛擬系統
- 2.2 HDFS 資料作業系統安裝與管理
- 2.3 HDFS 實務管理 (最佳化, 高可用性)

第三章 Hadoop Streaming

- 3.1 Hadoop Streaming 運作原理
 - 3.2 Bash Scripting 資料分析
-

11/15 (上課)09:00~16:30 (考試)16:30~17:20 考試

第四章 Big Data 資料分析

- 4.1 Apache Pig 運算分析 (人力資源失業率調查)
- 4.2 Apache Hive 運算分析 (各行業用電量分析)

第五章 HBase - Big Data 資料庫

- 5.1 安裝 HBase 海量資料庫
 - 5.2 設定 HBase 分散運算模式
 - 5.3 HBase 實務應用
-

4. 適合對象：

1. 欲親身體驗 Big Data 實作平台。
2. 資工/資管/資科/電通/企管 等相關學系學生。
3. IT 專案經理、系統架構師 或 技術決策人員。
4. 網路管理工程師 或 應用程式設計師。

5. 知識需求：

1. 需熟悉 Linux 桌面系統的 "目錄" 與 "檔案" 基本操作 (建立, 刪除, 壓縮, 複製...)
2. 具有 TCP/IP 網路基礎 (IP 位址, 遮罩, Default Gateway)

6. 上課環境：

1. 硬體需求: RAM 4GB 以上；建議最佳環境: **8GB RAM**。
2. Big Data 上課系統，由全域科技有限公司建置，含 USB 隨身碟內置上課環境。

7. 取得認證：

1. HPATA: Server&Storage(伺服器與儲存管理) 一張。
2. MSCS: Big Data, Configuring (巨量資料分析技術師) 一張。

8. 課程教材：

由全域科技有限公司之專業講師編輯 Big DATA 線上電子教材。

9. 學員可取得內容

1. 兩張認證 (HPATA/MSCS)。
2. USB 隨身碟一支 (含上課環境)。
3. 三天實務課程。
4. 實務課程研習證明一張。
5. Big Data 系統授權書一張 (限教育使用)。