職訓局109年大數據與AI訓練班

關聯式與非關聯式資料庫 報告

學生 洪聖哲

指導老師 邱梓君老師

關聯式資料庫與非關聯式資料庫報告

1. 簡介關聯式/非關聯式資料庫:

關聯式資料庫最早是由庫德在1969年在IBM公司內部刊物提出的，目的在於取代更古早的階層式資料庫以及網路資料庫，由於是藉由欄位來進行連結，因此可以改進早期資料庫的問題，在關聯式資料庫普及之後，就一直在電腦世界中沿用至今天，直到大數據以及人工智慧的崛起，關聯式資料庫才漸漸無法達到人們的要求。

而人們發現關聯式資料庫已經無法達到人們對資料寫入及查詢的要求之後，又發展出非關聯式資料庫，也就是NoSQL的出現，與傳統的資料庫不同的是， NoSQL通常在實際部屬上會採取大量分散型的伺服器，散步在世界各地，所以如何讓資料變成一致性在 NoSQL 裡面變得很重要，但並不是每一份工作都需要非關聯式資料庫，通常是在有大量寫入、不間斷寫入、大量查詢需求的服務才需要，只要沒有上述需求傳統的關聯式資料庫已經十分足夠。

2. 詳細介紹任一非關聯式資料庫:

微軟的NOSQL資料庫: Azure Cosmos DB

因學習微軟的SQL Sever，那想微軟應該也有推出自己的NoSQL才是， Azure Cosmos DB是2017年微軟推出的新產品，是根據 Build 2017 改良而來的非關聯式資料庫，是微軟標榜"管理地球規模數據的資料庫"而來，微軟保證資料全球都是一樣的，而且可以讓全球的使用者不論在何時何地存取你的資料，都保證毫秒等級的速度。

因為微軟全球部屬資料中心，因此保證 Cosmos DB的高可用性，微軟也提供在某區域發生區域性災難損壞的狀況下，可以轉移用戶的程式與功能，讓服務不致中斷。

在某些情境中，因為某些意外的效果有可能在產生意外的大量連線或者大量需求，微軟提供透過程式介面就可以擴展對伺服器的連線請求，從全球範圍內數千個請求/秒擴展到數億個請求/秒，並且只需為有使用到的容量或流量付費。

在NoSQL的設計中，通常會有三大類的要求:資料一致性（Consistent）、可用性（Availability）和中斷容忍性（Partition Tolerance），是NoSQL在打造的時候開發團隊就必須顧及的三個方向，但在實際設計中非關聯型資料庫無法同時兼顧這三種特性，因此NoSQL資料庫通常會選擇其中兩種來設計，通常是選擇CP或AP，而 Cosmos DB中可以提供給程式設計師從強一致性到最終一致性的選擇權(共5種)，自己選擇自己要的模式。

微軟的 Cosmos DB會自己替資料編制索引，達成全球資料的一致性，微軟宣稱使用者不必再替資料建立索引，並可快速提供查詢服務。

3. 請比較關聯式/非關聯式資料庫優缺點

因為目前大數據的興起，導致傳統型資料庫在面對長時間讀寫、只讀不寫、巨量資料建立索引時表現出不好的效能，因此各家為了生存轉而投入非關聯型資歷庫的懷抱，因為非關聯式資料庫分散全球的關係，在資料一致性上會允許些許誤差，比較適合用來處理不那麼即時性的資料，比如FB的讚，如果一篇文章發布後少了幾個讚，使用者也不會發現，這是比較適合用NoSQL的地方，至於需要非常即時性的比如售票系統則不適合，因為平行控制難以處理。