

一個在雲端應用服務環境支援 OLxP 交易和分析的商業連續性計劃

中興大學/資訊科學與工程學系

曾學文 老師

MOST-105-2221-E-005-058

(一) 計畫摘要

OLTP和OLAP的應用程式為了能夠有效率支援日益增長的資料儲存量與進行大規模的資料分析，帶動了雲端儲存需求。因為資料存取模式的差異，一般而言是分別應用在不同的系統上。隨著行動裝置的普遍，網路資料量急遽成長，透過分析這些儲存系統中的大量資料，能夠支援使用者進行決策。

為了減少從OLTP系統下載資料到OLAP系統花費的時間，必須在雲端儲存環境中將兩個系統進行整合。然而，由於OLTP及OLAP不同的工作負載特性，在OLxP整合性系統當中要能同時處理兩種應用程式流量。一方面要解決OLTP因請求數量眾多，致使儲存系統負載量不平衡所造成的效能延遲；另一方面，要能減少在系統中進行OLAP查詢時，節點間的資源競爭，改善整體計算效能的時間成本。

(二) 計畫架構



(三) 技術特色

- 針對NoSQL資料庫-Cassandra，提出可以自動化針對資料庫系統去偵測節點負載量，透過改變節點範圍的方法來改善各節點間的工作負載不平衡，自動進行調整節間的工作負擔。
- 改寫原先Hibench Benchmark延伸出能夠測試Spark-Cassandra開發環境的效能分析基準，可以對Spark-Cassandra開發環境進行完整的效能分析。
- 對Spark-Cassandra開發環境進行完整的效能分析，可提供給使用者參考，讓使用者在使用Spark-Cassandra前先評估自己的需求，協助使用者選擇合適的開發環境，降低成本的需求。

(四) 技術應用範圍

由於行動應用的普及化，越來越多消費者透過行動裝置購物，在網購平台工具選擇變多的情況下，商家無法僅靠數量規模及價格取勝，需深化客戶的信賴度，提供包括比價、導購、找商品等服務。

本計劃提出OLxP整合性系統的雲端應用，能夠儲存大量交易資料，並在不影響服務效能的狀況下，進行交易資料的分析以提供決策結果讓消費者參考。經過Benchmark測試不同Workloads在Spark的RDD計算時間與Spark寫入Cassandra所需要的時間；最後呈現出系統的效能可透過平衡Partition來進行優化，表示此OLxP整合性系統能彈性運用負載平衡的方式，讓不同Workloads的整體計算時間效能得到改善。

