

# 雲端運算環境中使用 MRU 演算法之 HBase 高效能資料分類 及處理

曾嘉影 楊希文 大同大學資訊工程系

## 摘 要

雲端運算環境中分散式資料庫系統管理為目前實用科技上的顯學，由多台電腦透過網路集中做平行大量資料處理可以加快運算時間。每一台電腦可以有著不同的專屬硬體環境，因此在此技術上，使用一般電腦發揮超級電腦效用，並非不可及。在如今效能不斷提升的硬體設備支援下，使用分散式運算技術是非常實際且確切能發揮其功效，系統的安裝也易於達成，但需要考量到系統設計與不同電腦運算速度的差異性。

本篇論文提出一種依照使用者常用檔案進行分類，利用分類結果使用 MRU 演算法來達到降低資料處理量與預測使用者接下來會使用的檔案進行預先處理，以達到效能的提升。在第一階段，我們會先對一些使用者進行常用檔案的次數與時間做紀錄，來當作演算法條件判斷的參考。在第二階段，我們將有使用 MRU 演算法和沒有使用演算法的 Hadoop 進行比對，觀察其效能提升的狀況並調整預測參數使其達到最佳的效能。

**關鍵詞：**HBase、資料分類、MRU 演算法