老師 各位同學大家好 我們報告的是應用資料壓縮實現更具效率且經濟的雲端儲存系統

這是我們的outline首先我們先介紹甚麼是公共雲、私有雲及混合雲，以及資料壓縮演算法，在介紹本篇論文設定的架構及實驗方式，之後由偉祺來說明實驗結果及做一個總結。

首先雲有分三種，公共雲是由第三方提供使用，能够為用戶提供無限的存儲空間，且還負責數據可用性、數據可靠性等，使用公共雲我們無需購買、管理和維護存儲硬體，這些都由提供商負責，僅依據使用量多寡來收費。Amazon、Google是提供商

因為公共雲無法訂製及有安全性的考量就需考慮使用私有雲，私有雲就是自行購買設備架設。

混合雲將公共雲、私有雲或其他雲組合在一起。是個別獨立，僅通過網路相互傳遞資料。

由於私有雲具有更好的效能因此使用私有雲當作緩存空間放置常用資料，而公共雲作為備份或是放置不常用資料。

為了有效地利用私有雲有限的存儲空間及降低公共雲存儲的開銷，我們在系統中應用了資料壓縮。

使用資料壓縮優點:壓縮可以減少文件大小，降地公共雲儲存成本，也能使得私有雲放置更多文件從而提高效能減少從公共雲存取檔案的傳輸時間。

然而使用資料壓縮的缺點是需考量壓縮及解壓縮時間，例如一些檔案經常被使用，一些檔案則不常被使用，如果不區分檔案特性只使用一種壓縮演算法，結果可能比甚麼都不做來得更糟糕，因此在之後會說明如何選擇壓縮算法。

接下來可以看到論文設計的架構，該系統有兩部分左邊為私有雲當作緩存空間，右邊為公共雲當作備份。一旦接收到請求，緩存管理器就換檢查檔案是否存在本身儲存空間中，若有則直接訪問它，若沒有則去公共雲儲存空間做拿取，但在將檔案下載至私有雲之前需先判斷該檔案是否需要壓縮，如果需要壓縮需判斷要用哪種壓縮法。

從表中我們可以看到五種壓縮演算法的壓縮比、壓縮時間、解壓縮時間，我們可以看到說霍夫曼及LZ77具有較低的壓縮及解壓縮時間，如果一個檔案時常被讀取我們可以選擇LZ77當作演算法，因為其解壓縮時間短，而一個檔案要時常寫入則可以使用霍夫曼當作演算法因其壓縮時間短。而LZMA及LZW壓縮比小於huffman及LZ77，但C/D較高，但考慮到儲存空間較少的關係及LZ77壓縮及解壓縮時間過差異性太大，因此選擇LZMA當作讀取壓縮演算法其壓縮時間短，LZW當作寫入壓縮演算法，CM則作為備份壓縮演算法，但我們必須確保該檔案室幾乎不會被使用才採取CM。

接下來我們來看論文的實驗方式:該論文主要是探討如何讓混合雲的使用更加具有效能，因此效能是本文探討的重點，效能定義為每個請求的平均回應時間，包刮:檔案遺失時從雲端訪問的時間、從緩存訪問的時間及壓縮及解壓縮時間。

壓縮時間的考量:如果即將來到的檔案很大，但私有雲的儲存空間不夠勢必刪除一些檔案或是針對該檔案進行壓縮，但必須考量到刪除檔案是否會造成後續的讀取該檔案花費更多時間，比如要刪除的檔案時常被使用，把它刪除下次使用就必須再從公共雲上拿出來，造成效能不佳，如果即將到來的檔案很小不管是否壓縮，不會影響其他檔案，那們為了節省時間就不進行壓縮。

接下來我們計算即將到來的檔案壓縮與否是否會節省時間，Vc0:為如果即將來的檔案太大私有雲空間無法放下因此將刪除原有的檔案，而VC0則是代表如果不刪除該檔案可以節省的時間，避免下次使用所花費時間，因為將檔案從私有雲中刪除下次使用就會從公共雲取出造成效能增加，節省了從公共雲下載的時間，而VCI則是比較即將下載的檔案沒有壓縮和壓縮所花費的時間，如果VCI>0則代表壓縮該檔案可節省回應時間，而V則是將VCO-VCI v>0則表示可以壓縮該檔案可以節省時間。假設VCO為5秒而VC