雲端運算與實務Final Project

0551287 張為舜 0551290 黃柏程

**Topic: Find a good solution for clustering Docker**

**Background**  
現在Docker之中並沒有一個眾所皆知的儲存系統，所以我們在Docker中建立3種分散式檔案系統，來進行檔案傳輸的測試，測試現在的分散式檔案系統中，哪一個比較適合在Docker中使用。

**Method**  
測試三種不同的分散式檔案系統在Docker下的處理速度。

在Docker安裝三種不同的分散式檔案系統，其中的不同的Container分別擔任Client端、Server端，由Client端下指令(Workload)，觀察各個DFS存取的速度差異、單個檔案大小的影響、檔案數量的影響。

**Experiment environment**   
1台8G ram 200G儲存空間 Ubuntu 16.04

2台8G ram 16G儲存空間 Ubuntu 16.04

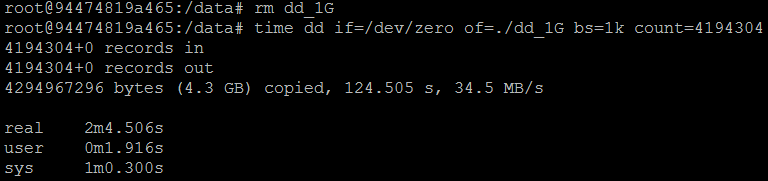
所使用的分散式檔案系統

* GlusterFS
* BTSYNC
* Network File System(NFS)

Workload  
使用dd指令從 /dev/zero 讀空白資料，寫到目前資料夾，對系統施加壓力，並得到傳輸所花的時間，以及傳輸的速率，例如：time dd if=/dev/zero of=./dd\_1G bs=1M count=1024

BS值代表單個檔案的大小，例如：bs=1M，代表每個檔案的大小是1MB；bs=1K，代表每個檔案大小是1KB。

Count值代表要傳輸幾個檔案，例如：count=1024，代表傳輸1024個檔案。



上圖為我們在container中執行dd指令對檔案系統施加壓力的範例圖，可以見到，我們傳輸了4194304個1KB的檔案到現在的資料夾。得到的結果是，總共傳了4.3GB，共花124.505秒，速率是34.5MB/s。

**Evaluation results**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 總傳輸量 | 每個檔案大小 | 總數量 | 傳輸速率 (MB/S) | | |
| GlusterFS | BTSYNC | NFS |
| 1GB | 1MB | 1024 | 74.1 | **91.4** | 71.8 |
| 4GB | 1MB | 4096 | 13.4 | 39.3 | **70.6** |
| 1GB | 1KB | 1048576 | 5.2 | **78.8** | 72.0 |
| 4GB | 1KB | 4194304 | 5.3 | **34.5** | 33.5 |
| 1GB | 1GB | 1 | **89.4** | 29 | 55.1 |
| 4GB | 2GB | 2 | 15 | 29.5 | **31.7** |

**Conclusion**

* GlusterFS在單一大型檔案傳輸方面(傳送1個1GB的檔案)，速度最快，其他方面則是表現較差。
* BTSYNC在小檔案傳輸上表現最佳(傳送大量1KB的檔案)，其他項目也有一定水準。
* NFS在各方面都有不錯的表現。
* 以我們小組實際架設檔案系統的情況來看，操作難易度為：NFS > GlusterFS > BTSYNC，在架設的過程中，NFS最為困難，再來是GlusterFS，BTSYNC則是比較容易的部分。