B03102082
陳孝思

SA HW2

1. Playing With LVM

1.1 Storage For NASA Course

(a). create 1 volume group **storage-vg** with these two disks

```
sudo pvcreate /dev/sdb
sudo pvcreate /dev/sdc
sudo vgcreate storage-vg /dev/sdb /dev/sdc
```

(b). create 1 logic volume of size 150G named **student** with ext4 filesystem

```
sudo lvcreate –L 150G –n student storage-vg
sudo mkfs –t ext4 /dev/storage-vg/student
```

(c). create 1 logic volume of size 350G named **ta** with ext4 filesystem

```
sudo lvcreate –L 350G –n ta storage-vg
sudo mkfs –t ext4 /dev/storage-vg/ta
```

(d). create 1 logic volume of size 300G named **hsinmu** with ext4 filesystem

```
sudo lvcreate -l 100%FREE -n hsinmu storage-vg
sudo mkfs –t ext4 /dev/storage-vg/hsinmu
```

(100%FREE 的用法來自 Ref)

Ref:https://www.centos.org/docs/5/html/Cluster_Logical_Volume_Manager/LV_create.html

1.2 Need More Space

```
sudo pvcreate /dev/sdd
sudo vgextend storage-vg /dev/sdd
sudo lvextend /dev/storage-vg/hsinmu /dev/sdd -r
sudo lvresize -L -150G /dev/storage-vg/ta -r
sudo lvresize -L +150G /dev/storage-vg/hsinmu –r
```

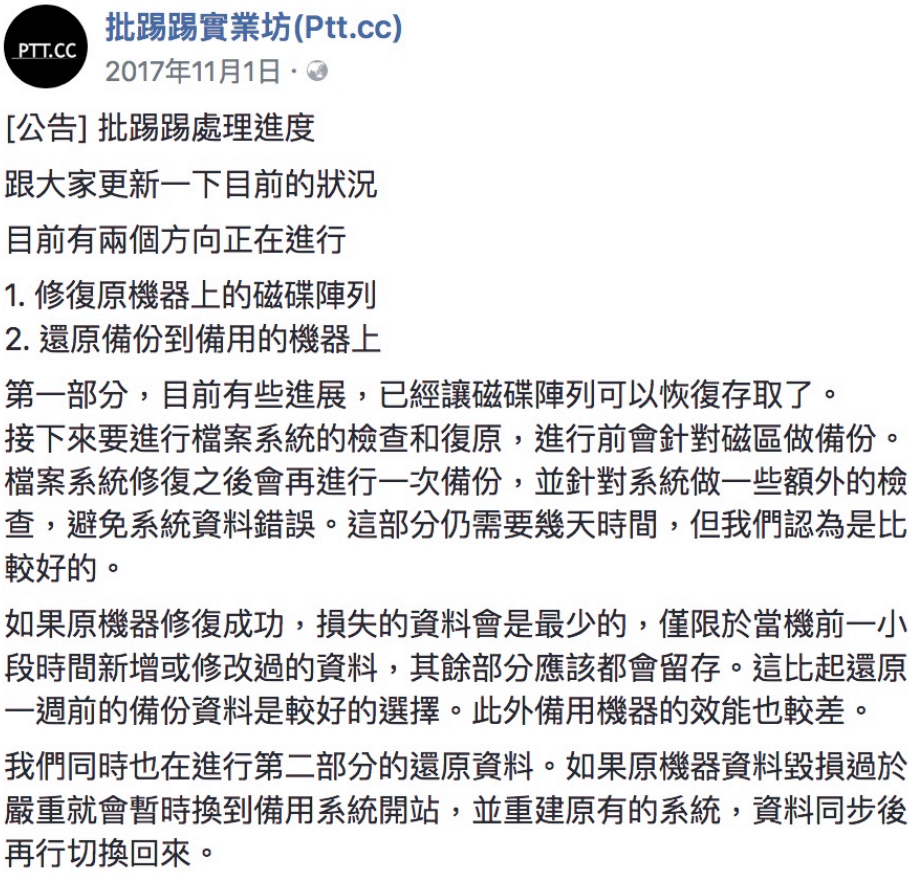
(-r 這個參數代表 resize，也是來自 Ref，如果不加-r，改用 resize2fs，那要先 resize2fs ta 才能 lvresize hsinmu)

Ref: https://www.rootusers.com/how-to-increase-the-size-of-a-linux-lvm-by-adding-a-new-disk/

2. PTT Alert

PTT 官方先公告說：「磁碟陣列掛了」

隨後又說：



推測原因：

詳細的原因官方沒有太多公告，僅能得出：伺服器使用的是硬碟陣列裡的硬碟壞了，硬碟陣列（RAID）是一種將多顆硬碟組合在一起，令系統視為一顆硬碟的技術，主要的目的有增加效能與增加資料安全性，希望能在部分硬碟損壞的情況下完整修復損壞的內容。

目前尚未得知 PTT 使用的是哪一種 RAID 方式，如果是 RAID 5 的話，RAID 5 要三顆以上的硬碟才能建立，在寫入資料時會計算 parity，並存在與寫入資料不同的硬碟裡面，目的是單一硬碟損壞的時候，可以透過另外兩顆硬碟上的

parity 重建資料，寫入替換壞掉硬碟的新硬碟中。且 RAID 5 只容許一顆硬碟損壞，壞兩顆就沒辦法救了。如果是 RAID 6 的話，類似 RAID 5，但要用四顆以上硬碟組成，但容許壞兩顆硬碟。

但是從故障到系統恢復上線花了好幾天，這篇公告是 11/1 號發的，11/1 晚間又延後開機時間到 11/3（來自維基百科），實際上正式修復上線是 11/4，加上 SYSOP 版上公告：「沒有備份到的資料基本上沒救」，與這篇貼文（未證實）提到：「這次先壞一顆，檢查時又爆了一顆，然後換了新的開始跑備份時又炸了一顆，然後就……」，故可以從起始公告跟最後處理方式來推測事情經過：一開始公告說先重建，代表可以 rebuild，但是因為未公開的原因導致 rebuild 失敗，可能原因從推文得知是 rebuild 時同個陣列上本來沒壞的硬碟也壞了，最後只好用備份資料恢復上線。

更好的處理辦法：

有錢的話直接放上 AWS 或 Google Cloud。

沒那麼多錢的話：

當下能做的事除了站方公告的之外，我也不太知道能做什麼，我覺得站方已經做到在預算人力限制之下最好的解決辦法了，事後諸葛來看的話就只有果斷用備份與舊機器上線。

至於之後為了避免跟這次一樣的失誤，可以

(1). 監控上線人數並準備備用伺服器，不只是單純備份，而是在主要伺服器壞掉的時候能直接上線的，平常處於待機狀態，一旦同時上線人數達到一定程度，備用機就可以在一旁待命

(2). 硬碟陣列能插滿就插滿，讓更多硬碟去分擔讀寫

(3). 增加 Raid Card Cache Memory 減輕硬碟負擔

(4). 同一陣列使用不同品牌且不同出廠年份的硬碟，以免同陣列裡的硬碟壽命同時到期一起掛掉

(5). Remote Backup，由不同於 PTT 團隊的網管管理（可能 NASA 團隊？）

Ref: http://linux.vbird.org/linux_basic/0420quota.php

https://zh.wikipedia.org/zh-tw/RAID

http://bytepile.com/raid_class.php

https://www.facebook.com/PttTW/posts/10154777910716364

https://linux.cn/article-2504-1.html