

[Home](#)[網誌里程碑](#)[訪客留言](#)[關於作者](#)[旅行](#)

Search: type, hit enter



# Cash's Blog

Cash's report to the world!

## 在 Android 底下使用 CGROUP

Posted by cash on 2014-07-20

[Go to comments](#)[Leave a comment \(0\)](#)

CGROUP 是什麼呢？在 WIKI [1] 裡面簡直沒有幾行字，要看懂不容易。簡單地說，Linux 裡面有很多 process，我們為了分而治之，所以增加 group 的虛擬觀念。每個 group 的 process 能夠使用的資源會有不同的限制。比方說，某些 process 可以用雙核，某些 process 只能用單核。某些 process 可以用 2GB DDR，另一些 process 只能用 128MB。

我們可以在 console 下面 `cat /proc/cgroups`。就會看見，有 `cpuset`, `cpu`, `cpuacct`, `memory`, `device`, `freezer`, `blkio`, `net_prio` 等幾個 `subsys_name`。而且他們各自有 `hierachy`, `cgroups` 的數目，以及 `enable` 與否。因為 Linux 底下的排版跑掉了，所以除了第一列，其他每一列都縮排。

#subsys_name		hierachy		num_cgroups	enabled
cpuset	0	1	1		
cpu	3	3	1		
cpuacct	1	48	1		
memory	2	2	1		
devices	0	1	1		
freezer	0	1	1		
blkio	0	1	1		
net_prio		0	1	1	

以上面的 Android 系統為例 (Android 底下也是 Linux)，其中 `cpuacct` 這個 subsystem 就有 48 個 cgroup 之多。什麼是 `cpuacct` 呢？它是指 CPU accouting control [2]。假如 Linux 沒有生成 `cpuacct` 這個 cgroup，我們可以用 command line 產生。

```
mount -t cgroup -ocpuacct none /sys/fs/cgroup/cpuacct
```

我用原生的 Linux 3.10 只看得到 `memory` 在 cgroup 的目錄下。使用上面這個 `mount` 指令後，我們可以在 `/sys/fs/cgroup` 底下看見 `cpuacct`。如果這個 cgroup 已經成立，我們就可

可以看到一些統計資料, 例如 `cpuacct.stat`, `cpuacct.usage`, `cpuacct.usage_percpu...` 等等.

[Note 1]

loading

`cat /sys/fs/cgroup/cpuacct.stat` 可以看到 user 和 system 各自用了多少 cpu 時間, 例如:

user 100352

system 79817 (單位是 user\_hz [3])

`cat /sys/fs/cgroup/cpuacct.usage` 得到

2096699156712 (單位是 ns)

`cat /sys/fs/cgroup/cpuacct.usage_percpu` 得到兩顆 CPU 的數字

1145393229564 952319269088

兩者相加就是 `cpuacct.usage`, 不過不是同一個時間, 所以數字又變大了. 光看統計數字沒有太大的幫助, 我們要怎麼使喚它呢? 根據 [REF 4] 的說法, 至少有三種方法可以做到. 其中 command line 和 libcgroup 的方法都被它介紹過了, 還有一個 LXC 的方法, 在 [REF 5] 當中有說明.

我這邊就只整理 command line 的方法, 因為它最容易了解 cgroup 目錄階層的意義. 我們可以手動在 `/sys/fs/cgroup` 底下操作.

[CASE 1] 如果要限制的是可以用幾核的 CPU, 比方說雙核中, 誰可以用兩顆 CPU, 誰只能用 (某) 一顆, 那麼要先設定第一核與第二核的定義.

`cd /sys/fs/cgroup/cpuset`

`sudo mkdir first_core`

`echo 0 > /sys/fs/cgroup/cpuset/first_core/cpuset.cpus`

`sudo mkdir second_core`

`echo 1 > /sys/fs/cgroup/cpuset/second_core/cpuset.cpus`

接下來開始分配, 假設有兩個 process, pid 分別為 1185 和 1195.

```
echo 1185 > /sys/fs/cgroup/cpuset/first_core/tasks
```

```
echo 1195 > /sys/fs/cgroup/cpuset/first_core/tasks
```

```
echo 1195 > /sys/fs/cgroup/cpuset/second_core/tasks
```

結果使得 pid 1195 可以用雙核, 而 1185 只能用單核.

[CASE 2] 若是要進一步把 pid 1185 的使用權種降得更低, 我們就要定義出比重.

```
cd /sys/fs/cgroup/cpu
```

```
sudo mkdir high
```

```
echo 2048 > /sys/fs/cgroup/cpu/high/cpu.shares
```

```
sudo mkdir low
```

```
echo 512 > /sys/fs/cgroup/cpu/low/cpu.shares
```

此時高低權重就出來, 接下來同樣是把 pid 加入 task 當中.

```
echo 1185 > /sys/fs/cgroup/cpu/low/tasks
```

```
echo 1195 > /sys/fs/cgroup/cpu/high/tasks
```

這樣在同一顆 CPU 中, 1195 也可以拿到 1185 四倍的 CPU 時間.

[CASE 3] 如果要分配記憶體用量, 首先要設一個會限制記憶體的 group, 例如 limited\_memory.

```
cd /sys/fs/cgroup/memory
```

```
sudo mkdir limited_memory
```

將它的記憶體上限定為 128M Bytes

```
echo 128000000 > /sys/fs/cgroup/memory/limited_memory/memory.limit_in_bytes
```

然後將這個 pid 加到這個 group

```
echo 1185 > /sys/fs/cgroup/memory/limited_memory/tasks
```

如果它的記憶體用量硬是超過限制, 它會 swap 到 disk, 不會用 DDR. 不過看來這樣也會拖累效率就是了.

[ref]

## 1. 中文版 WIKI

cgroups(控制組)是Linux核心的一個功能, 用來限制報告和分離一個行程組的資源(CPU、記憶體、磁碟輸入輸出等)。這個工作是由Google的工程師(主要是Paul Menage和Rohit Seth)在2006年以「process containers(行程容器)」的名字開始的; [1] 在2007年的晚些時候被重新命名為控制組(由於在核心中「容器」這個名詞的歧義引起的混亂)並被合併到2.6.24版的核心中去。[2] 自那以後, 又添加了很多功能和控制器。

cgroups的一個設計標的是為不同的應用情況提供統一的介面, 從控制單一行程(像nice)到系統級虛擬化(像openVZ, Linux-VServer, LXC)。cgroups提供:

- 資源限制: 組可以被設定不超過設定的記憶體限制; 這也包括虛擬記憶體。[3] 原來的分頁機制是在Linux研討會的*Containers: Challenges with the memory resource controller and its performance*報告中提出的。[4]
- 優先化: 一些組可能會得到大量的CPU[5] 或磁碟輸入輸出通量。[6]
- 報告: 用來衡量系統確實把多少資源用到適合的目的上。[7]
- 分離: 為組分離名稱空間, 這樣一個組不會看到另一個組的行程、網路連線和檔案。[2]
- 控制: 凍結組或檢查點和重新開機動。[7]

## 2. CPU Accounting Controller

### 3. 3.3. cpuacct

## 4. Linux Control Group 介紹

## 5. lxc 建立和設定虛擬機器

[Note 1] 同理, 我也可以 `mount -t cgroup -ocpuset none /sys/fs/cgroup/cpuset`, 不過需要手動在 cgroup 底下建這些目錄才 mount 得到. 我用 Linux 3.10, 最初 cgroup 底下只有 memory 一個目錄而已.

建完新目錄之後, 底下就會自動繼承出一大堆子目錄. 例如 cpuset 底下的 first\_core 就會自動有 cpuset.cpus.

技術      cgroup

← 第十八屆交大高階經理人培訓班第七課

大陸網路與多媒體新聞小整理 →

Leave a comment ?

0 Comments.

Leave a Comment



NOTE - You can use these [HTML](#) tags and attributes:  
<a href="" title=""> <abbr title=""> <acronym title=""> <b> <blockquote cite=""> <cite> <code> <del  
datetime=""> <em> <i> <q cite=""> <s> <strike> <strong>

NAMEEMAILWebsite URL

SUBMIT

- ☐ 用電子郵件通知我後續的迴響。
- ☐ 新文章使用電子郵件通知我。

十月 2017

一	二	三	四	五	六	日
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29

« 九月

一	二	三	四	五	六	日
30	31					

« 九月

近期文章

- Webkit 歷史圖示
- ATSC 3.0 小整理
- 我讀《證券分析》第六版 – 證券分析的其他方面
- 區塊鏈小註解
- 我讀《證券分析》第六版 – 資產負債表 4
- 我讀《證券分析》第六版 – 資產負債表 3
- 我讀《證券分析》第六版 – 資產負債表 2
- 我讀《公司的品格》
- 中國 ETF 小整理
- Firstrade 小補充

近期迴響

- cash 在 神秘的 04 開頭電話
- cash 在 DiskPart 小註解
- 奏 在 神秘的 04 開頭電話
- Yugo 在 DiskPart 小註解
- cash 在 HDCP 相容性小註解

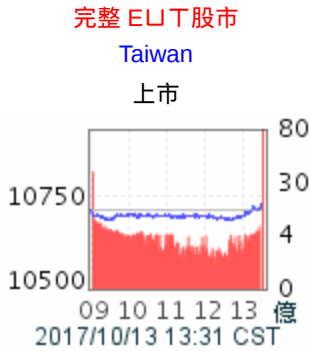
分類

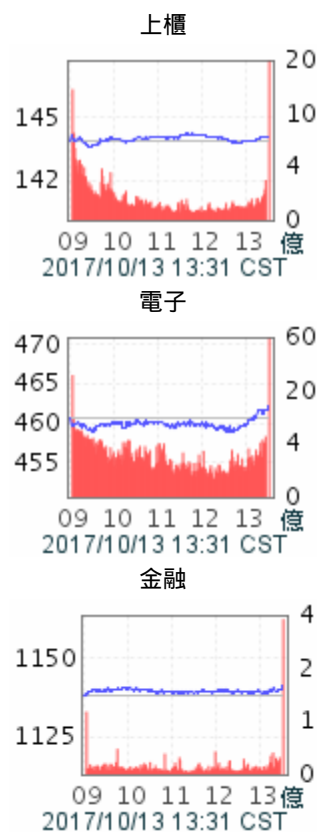
分類

選擇分類

台股速覽

資料來源: Yahoo 股市. 僅供參考





## 個人網站類書籤

4THINK | 為增進思考而閱讀  
RusRule Financial Broadcasting  
三分鐘熱度  
叛徒在火星的日子  
史蛋粒的網路樂園  
瘋小貓的華麗冒險:: 痞客邦PIXNET ::  
趙胤丞的One Piece學習部落格

## 參考鏈結

ColorHexa  
flyingV 群眾募資平台  
JUNKCALL.org 台灣  
My Travel Map  
NetSpeak  
SlideShare  
Time Zone Abbreviations – Worldwide List  
TOP SAFETY PICKs by years  
Worldtimebuddy  
全國地政電子謄本系統

## 技術類書籤

---

Android Cross Reference  
codecombat  
Free HEVC/H.265 bitstreams  
GFXBench  
GPU 的 Gflops  
How Browsers Work: Behind the scenes of modern web browsers  
IOH 開放個人經驗平台  
IP 位置追蹤  
IP下車位置  
Linux Cross Reference  
Myoops 線上課程  
Realtek WIKI  
wftp3.itu.int  
健忘的工程師筆記本  
常用u-boot命令詳解（全）  
經緯度轉換

## 理財專家個人網站

---

flying1229 (飛嫂之家)  
Mr. Market 市場先生 – 投資理財入門  
不敗教主~理財不 Buy  
兩座山  
共創價值投資  
定錨投資隨筆  
小樂的生存之道  
巴小智的粉絲團  
恩汎理財投資團隊  
樺誌的異想世界  
美股夢想家  
股魚-如何幫你輕鬆避開地雷股，長期投資必看  
華倫存股 穩中求勝  
談股論經---究天人之際，通古今之變，成一家之言  
麥克風 市場求生手冊  
麥樹仁投資社群網站(Makssin)

## 理財類書籤

---

105年殖利率扣抵稅率  
106年殖利率扣抵稅率



CMoney 大盤分析  
CMoney 選股網  
finance168財經學院  
Hi Stock 嗨投資  
MacroMicro 財經M平方  
MorningStar.com  
Stock Q. org  
stock-ai 投資級經濟指標  
TAURAS 財富守護星 – 選股軟體  
tivo168\_的投資理財\_Excel\_應用教學  
TWSE 台灣證券交易所  
USA Stock  
元大 ETF 收益分配表  
勞工個人退休金專戶試算表(勞退新制)  
勞工適用勞動基準法退休金制度年資之退休金  
試算表(勞退舊制)  
台灣股市資訊網 Stock Info  
基金試算專區  
強基金  
投信投顧公會  
政府線上繳稅  
淘股網  
淘股網 樂活五線譜  
熱血流成河  
神秘金字塔  
網龍大富翁  
線上掃雷系統  
老年年金給付  
股息 現金流 被動收入 理財的心路歷程  
財報狗  
集保戶股權分散表查詢

## 發票號碼查詢

---

中華電信  
台灣自來水公司  
台灣電力公司  
新竹瓦斯

## 網誌管理

---

登入  
文章 RSS 訂閱  
迴響 RSS 訂閱

WordPress 台灣正體中文