

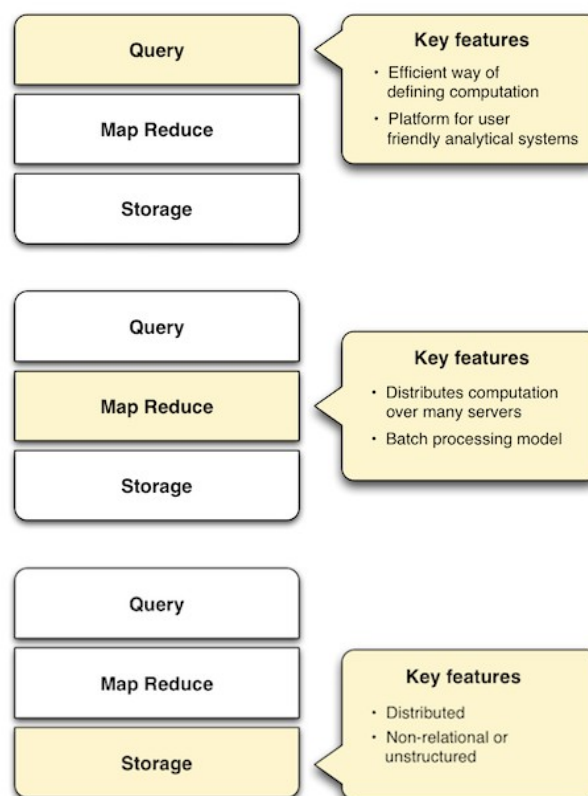
Hadoop 叢集佈建初探 - 以 hiCloud 為例

國家高速網路與計算中心 王耀聰 副研究員

用 Hadoop 打造您專屬的 SMAQ 資訊架構

隨著資訊科技的快速演進，我們每天產生數量眾多的數位文件、圖片、音樂跟影片等。根據 EMC 委託 IDC 對數位宇宙（Digital Universe）的研究報告指出，2007 年全球人類所產生的資料量已經遠超過全球所有儲存設施的總容量，意味著我們已經進入資料大爆炸（Data Explosion）的時代。2011 年數位宇宙的總容量更高達 1.8 ZB（Zetabyte, 10^9 TB），因此 2012 年可以稱為「巨量資料（Big Data）」的元年。

根據 Gartner CIO Agenda 2012 的調查，2012 年最重要的三大資訊科技分別為「商業智慧（Business Intelligence, BI）」、「行動科技（Mobile Technology）」與「雲端運算（Cloud Computing）」。事實上商業智慧早在 Gartner CIO Agenda 2009 就名列第一名，為何四年後捲土重來？那是因為要達成商業智慧之前，企業必須先具備處理巨量資料的資訊架構，也就是 2010 年 6 月 Edd Dumbill 於「The SMAQ stack for big data」一文中提出的 SMAQ 架構。



SMAQ 架構說明了處理巨量資料須具備儲存（**Storage**）、平行運算（**MapReduce**）與查詢（**Query**）三層資訊架構。其中，儲存層用來儲存分散而沒有關聯（Non-relational）的非結構化資料（unstructured data），平行運算層採用谷歌提出的 MapReduce 演算法，將龐大的非結構化資料轉化成結構化資料（資訊，Information），並且將結構化資料儲存到查詢層的資訊系統，如資料庫、資料倉儲等。

透過 SMAQ 資訊架構，企業還再次將資料倉儲的龐大資料經過 MapReduce 運算，運用統計的方法，找出重複出現的經驗法則（知識，Knowledge）。企業有了這些經驗法則，才

能夠進一步對充滿變數的未來進行預測，最後形成某種智慧（Wisdom）。十年前知識管理專家所提出的 DIKW 模型（Data、Information、Knowledge、Wisdom），如今終於在資訊系統中獲得了具體的實現。

Hadoop 這套自由軟體實現了類似 Google File System 的儲存層，提供給企業一個具備彈性擴充的儲存設施；其次它也實作了 MapReduce 演算法，並且將谷歌的在地運算（Data Locality）精神融入 Hadoop 的核心設計中，大幅降低了資料搬運的時間與頻寬成本。此外，建構於 Hadoop 之上，存在著 HBase 分散式資料庫與 Hive 資料倉儲等查詢層軟體。

因此，透過導入 Hadoop 生態系相關軟體元件，能提供貴單位符合 SMAQ 的資訊架構。像 Yahoo、Facebook、Tweeter、LinkedIn、Joost 等均已採用 Hadoop 當背後的支援火力。連 Oracle、Microsoft SQL Server 都提供 Hadoop 的資料庫連結元件，足見業界對於 Hadoop 的重視。

Hadoop 經過七年的開發，阿帕契軟體基金會（Apache Software Foundation）終於在 2012 年 3 月釋出 1.0.1 版本，其重點在於安全性的改良以及對 HBase 有更好的支援。此外 1.0 版也算是 Hadoop 成熟到能支持企業營運任務的重大里程碑。即便如此，**目前 Hadoop 最令人望之怯步的關鍵點在於「不夠友善」**。筆者自 2009 年起便在台灣推動 Hadoop 使用者社群(<http://www.hadoop.tw>)，經營 Hadoop 論壇(<http://forum.hadoop.tw>)，但初學者往往第一個要面對的問題是缺乏佈建 Hadoop 叢集所需的背景知識，如 Linux、Java 語言等。因此，撰寫本文讓更多資訊從業人員得以從實例操作中，學習如何使用 Hadoop 生態系。

Hadoop 單機版初探 - Windows 版

根據 hadoop 官方文件的建議，Hadoop 最佳營運案例主要是在 Linux 作業系統底下運行。但在台灣，多數資訊從業人員仍以 Windows 為主要的作業系統。筆者為了降低各位學習 Hadoop 的門檻，撰寫了一套名為 Hadoop4Win 的懶人包安裝程式。目前在 Windows XP 與 Windows 2000 環境下測試正常，至於 Windows 7 環境下目前因為安全性權限較嚴格，不易安裝成功，因此這次我們以 Windows XP 環境來做示範。

首先，請到 <http://hadoop4win.nchc.org.tw> 或 <http://hadoop4win.org> 下載 0.1.5 版本的 hadoop4win 完整安裝檔。Hadoop4win 0.1.5 目前包含五大軟體組成：(1) Cygwin；(2) JDK 1.6.0 update 18；(3) Hadoop 0.20.2；(4) HBase 0.20.6；(5) Ant 1.8.2。因此整個完整版懶人包大約有 189 MB，請**點擊 hadoop4win-setup-full_0.1.5.exe 執行檔**開始進行安裝的動作。

備註：由於 hadoop4win 會在安裝資料夾寫入數量眾多的檔案，假如您有安裝防毒軟體，安裝時間可能會比較長（因為每個檔案都被掃描過一次）。倘若您希望縮短安裝時間，可暫時性關閉防毒軟體。



STEP 1：首先請選擇語系
「繁體中文，Chinese (Traditional)」



STEP 2：本軟體採用 Apache 2.0 授權，
請按「我接受」



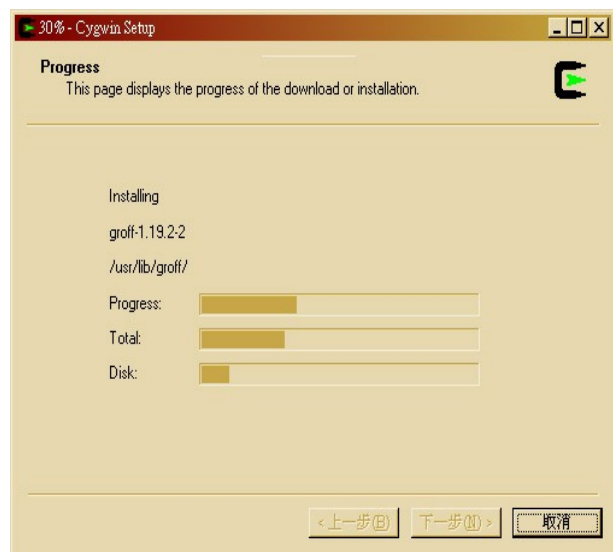
STEP 3：如果您不需要 HBase 跟 Ant，
可以取消選取



STEP 4：選擇安裝目標資料夾，
預設為 C:\hadoop4win



STEP 5：接著安裝程式會開始解壓縮



STEP 6：接著會跳出 Cygwin 的安裝畫面



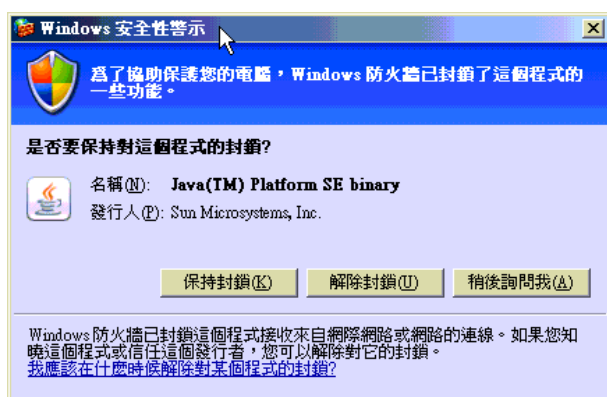
STEP 7：安裝完成後，請按關閉結束



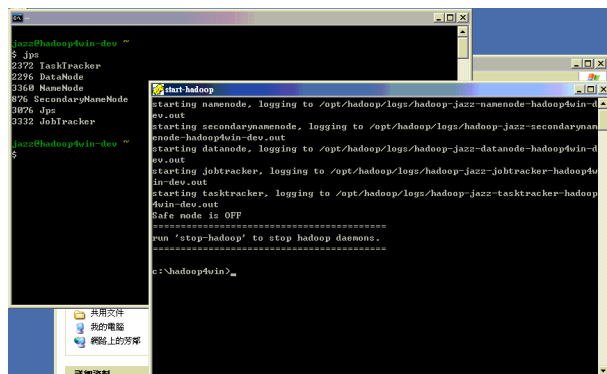
STEP 8：您可以從「開始」→「程式集」→「hadoop4win」點選 start-hadoop 啟動 Hadoop 服務於背景

當您完成 hadoop4win 的安裝後，有兩種方式可以在 Windows 上啟動 Hadoop。其一是從「開始」選單選取 start-hadoop 捷徑。其二是先點選 hadoop4win 捷徑，進入 Cygwin 指令操作的介面，才下指令 start-hadoop 來啟動 Hadoop 的五個背景服務（亦及 NameNode、DataNode、JobTracker、TaskTracker 與 SecondaryNameNode）。

備註：安裝過程中，Windows 防火牆會跳出警告說有 Java 程式嘗試開啟網路連線埠，這是因為 Hadoop 背景服務至少會開啟五個網路連線埠來進行底層的 RPC 通訊，請選擇「解除封鎖」。



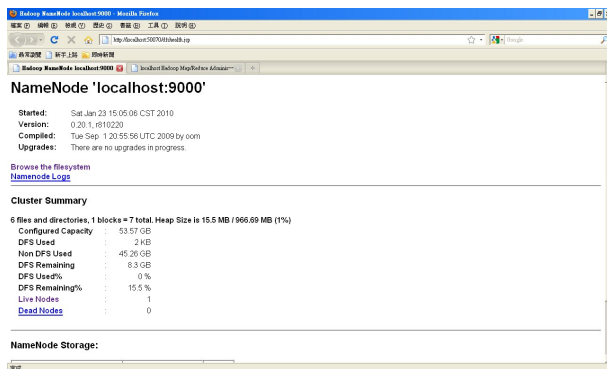
備註：請選「解除封鎖」



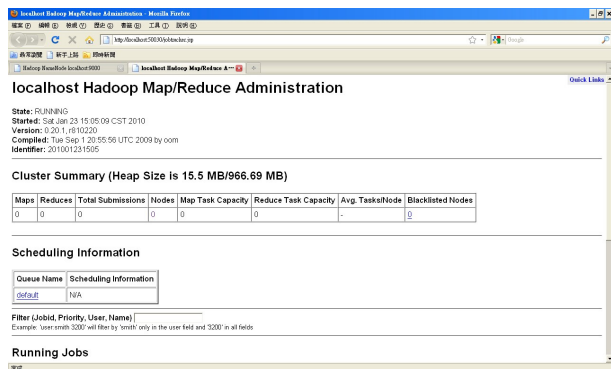
點選 start-hadoop，並等待五個背景程序啟動。可於 Cygwin 視窗中下 jps 指令檢查。

接著，您可以根據 hadoop4win 網頁的說明，學習基本的 Hadoop 指令操作。礙於版面篇幅，這裡我們僅列舉三個指令操作來進行說明。首先，第一個指令叫 jps，這個指令顧名思義就是列出所有 java process。單機版的 Hadoop4Win，基本上應該要有五個 Java 程序。因此，倘若您無法看到如圖的五個 Java 程序名稱（前面的數字為 PID，每次執行會不同），那代表您的電腦可能記憶體不足（最低需求為 512MB，建議至少有 1GB 記憶體）。倘若您在執行上遇到任何問題，歡迎到 forum.hadoop.tw 提問。

如果您可以正常看到 jps 出現的五個 Java 程序，代表您的 Hadoop 環境已經在背景正常執行了。此時您可以點選「開始」選單的「NameNode Web UI」與「JobTracker Web UI」，或者直接用瀏覽器開啟 <http://localhost:50070> 與 <http://localhost:50030> 的網址。



開啟 http://localhost:50070 頁面



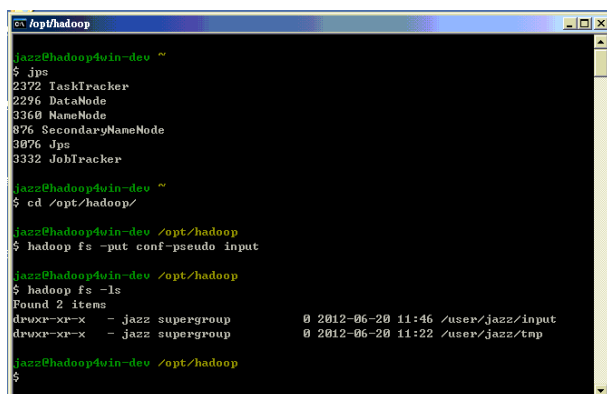
開啟 http://localhost:50030 頁面

第二個指令叫作 `hadoop fs`，這個指令是用來驗證 Hadoop 分散式檔案系統（HDFS，Hadoop Distributed File System）的功能是否正常。請在 Cygwin 指令視窗鍵入以下指令，將位於安裝目錄（`c:\hadoop4win`）的 `opt\hadoop\conf-pseudo` 子目錄，上傳到 HDFS 的抽象檔案系統：

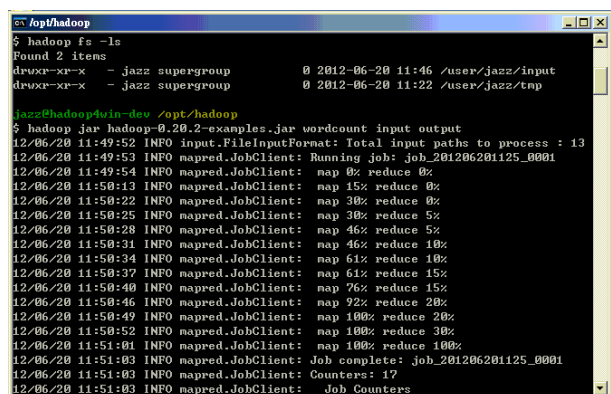
```
$ cd /opt/hadoop
$ hadoop fs -put conf-pseudo input ; hadoop fs -ls
```

最後一個指令叫作 `hadoop jar`，這個指令是用來驗證 Hadoop 的 MapReduce 運算環境功能是否正常。請在 Cygwin 指令視窗中鍵入以下指令，將剛剛上傳到 HDFS 的 `input` 目錄，共計 13 個檔案，進行「字數統計（Word Count）」的 MapReduce 範例運算。

```
$ hadoop jar hadoop-0.20.2-examples.jar wordcount input output
```



用 `hadoop fs` 指令做 HDFS 功能驗證



用 `hadoop jar` 指令做 MapReduce 功能驗證

Hadoop 叢集版初探 – hiCloud Ubuntu Linux 版

上個小節我們介紹了如何在 Windows 上建立一個 Hadoop 的實驗環境，然而 Windows 版本存在著許多小問題，像是 Cygwin 的路徑與 Java Runtime 認知的 Windows 路徑不同，因此許多時候，您在下指令的時候，還必須要透過 `cygpath` 指令將 Cygwin 路徑轉換成 Windows 路徑，相當不便。其次，假設您想要把多台 `hadoop4win` 串起來，有許多先決條件，像是必須要有不同的電腦名稱、要有相同的使用者名稱跟檔案存取權限。因此，一般我們在提供商業運轉的 Hadoop 環境時，仍舊會採用 Linux 版本。

雲端運算的五大基礎特徵中，第一點是「隨需自助服務」，其背後隱藏著「標準化」與「自動化」的精神。本文為了讓每位讀者能有標準化的實作環境，免除不同 Linux 版本造成的操作問題，此次採用的是中華電信 hiCloud 的 Ubuntu 10.04 虛擬機器當作示範架設 Hadoop 叢集版的環境。底下將示範如何在十台 Ubuntu 10.04 的虛擬機器上，建置一個 Hadoop 叢集。由於 Hadoop 安裝的指令繁瑣，筆者這裡提供一個自動化的安裝腳本程式，各位可至 <http://github.com/jazzwang/hicloud-hadoop> 或 <http://hidooop.sf.net> 下載原始碼。以下我們用圖解的方式，說明如何使用 hicloud 開啟十台虛擬機器，並設定成 Hadoop 叢集。



STEP 1：請連線至 hicloud.hinet.net 並點選右上角「客戶登入」



STEP 2：若您不是中華電信用戶，請先加入會員，然後用取得帳號登入



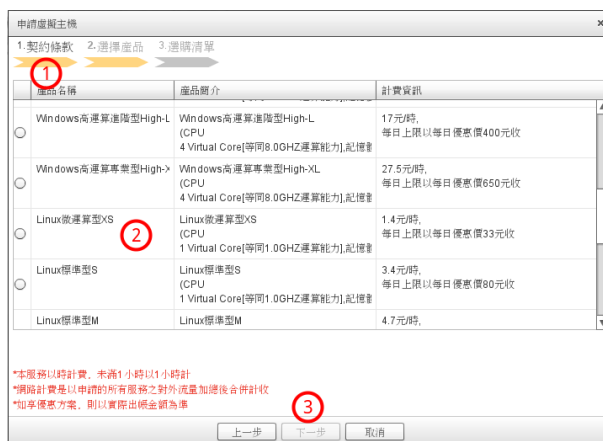
STEP 3：點選上方「雲運算」



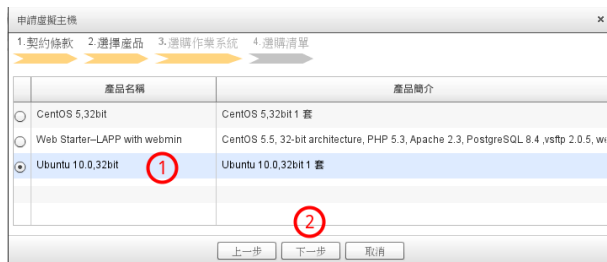
STEP 4：點選左方「虛擬主機」



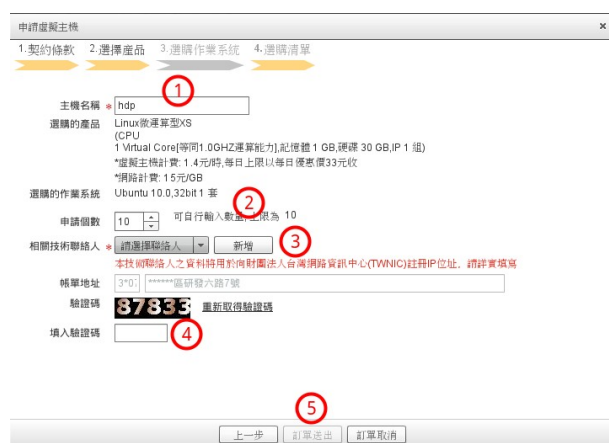
STEP 5：點選左上角「申請」



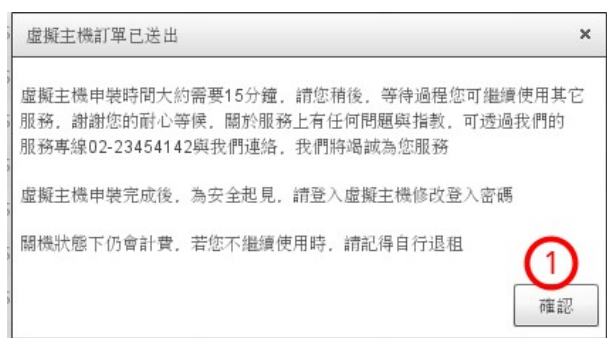
STEP 6：確認同意中華電信 hicloud 租約選擇「Linux 微運算型 XS」，按「下一步」



STEP 7：選擇「Ubuntu 10.0,32bit」



STEP 8：用 hdp 命名，產生 10 台虛擬機器



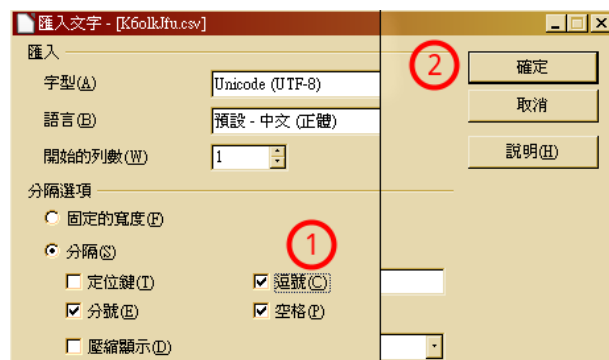
STEP 9：等待申裝狀態變為「已申裝」



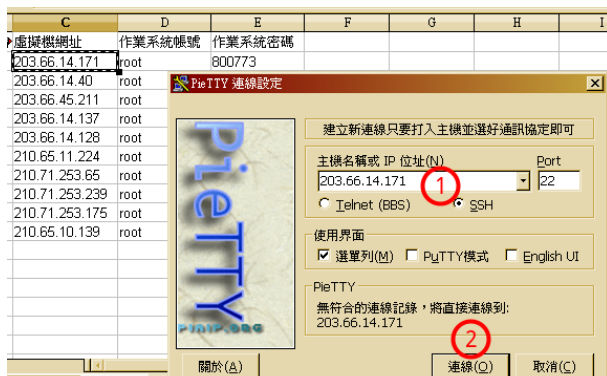
STEP 10：請確認供裝狀態變為「已申裝」
點左上角「全選」，並選「開機」



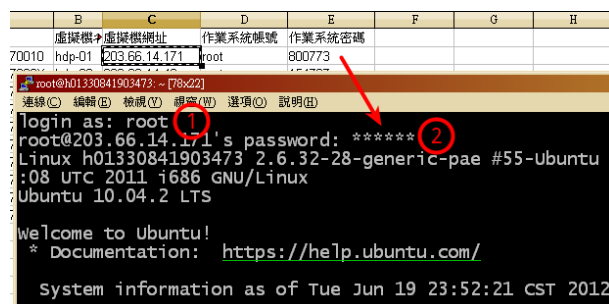
STEP 11：點選右上角「預設密碼」，將 password.zip 儲存到桌面



STEP 12：用預設密碼解壓縮 password.zip
並匯入 CSV 格式的密碼檔案



STEP 13：我們從密碼表中挑選第一台
用 PieTTY 連線到該台虛擬機器



STEP 14：預設請使用 root 帳號登入，
密碼請根據您的密碼表輸入六碼預設密碼

做到這邊，我們已經把十台虛擬機器開機。請從 hiCloud 介面上確認每一台虛擬機器都已經進入「已開機」的狀態，接著使用您慣用的 SSH 連線軟體，這裡我們使用的是 PieTTY，可以自 <http://ntu.csie.org/~piaip/pietty/> 下載。由於 hicloud-hadoop 自動安裝腳本是針對 Linux 環境撰寫的，因此我們直接挑選其中一台來進行安裝。您也可以在您的 Linux 環境上執行以下的步驟。在執行 hicloud-hadoop 自動安裝腳本前，有一些前置作業：(1)產生 SSH 連線用的金鑰；(2)安裝 Perl 的 Expect.pm 函式庫，因為我們會使用 Expect 來幫忙您自動填入預設密碼，登入另外九台虛擬機器，並進行 Hadoop 安裝作業。現在請您在登入第一台虛擬機器的 PieTTY 視窗內輸入以下兩個指令：

```
~# ssh-keygen -t rsa -P ""
~# apt-get install libexpect-perl
```

接下來，讓我們下載 hicloud-hadoop 自動安裝腳本，並且解壓縮安裝檔。然後在 hicloud-hadoop 目錄底下建一個 conf 子目錄。

```
~# wget http://nchc.dl.sourceforge.net/project/hadoop/0.0.1/hicloud-hadoop-0.0.1.tar.gz
~# tar zxvf hicloud-hadoop-0.0.1.tar.gz
~# mkdir -p hicloud-hadoop/conf
```

接著，使用記事本先編輯兩個檔案，一個叫 master，一個叫 slave。可用 PieTTY 或 WinSCP 上傳這兩個檔案到 hicloud-hadoop/conf 底下，或參考底下的作法在 conf 底下建立兩個設定檔。

```
~# cat > hicloud-hadoop/conf/master << EOF
203.66.14.171 root 800773
EOF
~# cat > hicloud-hadoop/conf/slave << EOF
203.66.14.40 root 154737
.....
210.65.10.139 root 367578
EOF
```

以上的 IP、帳號、密碼三個欄位，您可以直接從 CSV 貼到記事本，加入中間的空白，在貼到命令列。最後切換到 hicloud-hadoop 目錄，然後執行 install-ubuntu 自動安裝腳本。

```
~# cd hicloud-hadoop
hicloud-hadoop# ./install-ubuntu
```

	A	B	C	D	E
1	虛擬機ID	虛擬機	虛擬機網址	作業系統帳號	作業系統密碼
2	BV550103870010	hdp-01	203.66.14.171	root	800773
3	BV55010387000X	hdp-02	203.66.14.40	root	154737
4	BV55010387000Y	hdp-03	203.66.45.211	root	459715
5	BV55010387000Z	hdp-04	203.66.14.137	root	747421
6	BV550103870016	hdp-05	203.66.14.128	root	848416
7	BV550103870012	hdp-06	210.65.11.224	root	330179
8	BV550103870011	hdp-07	210.71.253.65	root	767701
9	BV550103870013	hdp-08	210.71.253.239	root	792269
10	BV550103870014	hdp-09	210.71.253.175	root	798999
11	BV550103870015	hdp-10	210.65.10.139	root	367578

從密碼表中挑選第一台的 IP、帳號、密碼貼到記事本，補上空格，然後轉貼到 master

	A	B	C	D	E
1	虛擬機ID	虛擬機	虛擬機網址	作業系統帳號	作業系統密碼
2	BV550103870010	hdp-01	203.66.14.171	root	800773
3	BV55010387000X	hdp-02	203.66.14.40	root	154737
4	BV55010387000Y	hdp-03	203.66.45.211	root	459715
5	BV55010387000Z	hdp-04	203.66.14.137	root	747421
6	BV550103870016	hdp-05	203.66.14.128	root	848416
7	BV550103870012	hdp-06	210.65.11.224	root	330179
8	BV550103870011	hdp-07	210.71.253.65	root	767701
9	BV550103870013	hdp-08	210.71.253.239	root	792269
10	BV550103870014	hdp-09	210.71.253.175	root	798999
11	BV550103870015	hdp-10	210.65.10.139	root	367578

從密碼表中挑選剩下的 IP、帳號、密碼貼到記事本，補上空格，然後轉貼到 slave

等自動安裝腳本執行完畢後，您會看到類似底下的訊息。此時，點選兩個網址，就會看到 Hadoop 的 HDFS NameNode 與 MapReduce JobTracker 網頁。

```
*****
SHUTDOWN_MSG: Shutting down NameNode at node9/210.65.10.139
*****
starting datanode, logging to /opt/hadoop/logs/hadoop-root-datanode-node9.out
starting tasktracker, logging to /opt/hadoop/logs/hadoop-root-tasktracker-node9.out
Please check http://203.66.14.171:50070 and http://203.66.14.171:50030 for the
root@h01330841903473:~/hadoop-hadoop#
```

安裝完畢的最後一行會顯示 50070 與 50030 的兩個網址

Hadoop NameNode hadoop-9000 - Mozilla Firefox

203.66.14.171:50070/dfshealth.jsp

NameNode 'hadoop:9000'

Started: Wed Jun 20 00:16:39 CST 2012
Version: 0.20.2, r911707
Compiled: Fri Feb 19 08:07:34 UTC 2010 by chrisdo
Upgrades: There are no upgrades in progress.

[Browse the filesystem](#)
[Namenode Logs](#)

Cluster Summary

7 files and directories, 1 blocks = 8 total. Heap Size is 15.19 MB / 966.69 MB (1%)

Configured Capacity	: 169.61 GB
DFS Used	: 144.01 KB
Non DFS Used	: 16.68 GB
DFS Remaining	: 152.92 GB
DFS Used%	: 0 %
DFS Remaining%	: 90.16 %
Live Nodes	: 6

50070 的 HDFS 頁面

Hadoop NameNode hadoop-9000 - Mozilla Firefox

203.66.14.171:50070/dfshealth.jsp?whatNode=LIVE

NameNode 'hadoop:9000'

Started: Wed Jun 20 00:16:39 CST 2012
Version: 0.20.2, r911707
Compiled: Fri Feb 19 08:07:34 UTC 2010 by chrisdo
Upgrades: There are no upgrades in progress.

[Browse the filesystem](#)
[Namenode Logs](#)
[Go back to DFS home](#)

Live Datanodes : 6

Node	Last Contact	Admin State	Configured Capacity (GB)	Used (GB)	Non DFS Used (GB)	Remaining (GB)	Used (%)
node1	0	In Service	28.27	0	2.78	25.49	0
node2	2	In Service	28.27	0	2.78	25.49	0
node3	2	In Service	28.27	0	2.78	25.49	0
node4	0	In Service	28.27	0	2.78	25.49	0
node5	1	In Service	28.27	0	2.78	25.49	0
node6	1	In Service	28.27	0	2.78	25.49	0

50030 的 MapReduce 頁面

雲端精算師：好的設計可以省下更多錢

做到這邊，各位讀者應該可以繼續在 hiCloud 的十台虛擬機器上練習在 Hadoop4Win 的指令。不過這裡要提醒的一件很重要的事情：「與 Amazon EC2 有開機才計價，hiCloud 是從申裝開始計費，退租才停止計費，就像您租用手機門號一樣，不管有沒有開機都會計費」。因此建議您如果口袋不夠深的話，練習完麻煩一台一台把虛擬機器退租吧！一節課程大約花個兩三個小時，14 塊台幣/小時，跟買杯超商咖啡差不多，這就是雲端帶給各位的新學習環境。

hiCloud 控制台截圖，顯示主機列表及退租確認對話框。

確認退租？退租後此虛擬主機之資料將全部刪除

☒ 我已同意並確認將止租

是 否

練習完畢後，請務必退租，免得收到帳單會很痛！

服務項目	起始日期	起始時間	終止日期	終止時間	雲端服務編號
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:40:51	101/06/19	23:59:59	BV55010387000X
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:42:06	101/06/19	23:59:59	BV55010387000Z
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:42:25	101/06/19	23:59:59	BV550103870010
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:42:39	101/06/19	23:59:59	BV55010387000Y
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:42:51	101/06/19	23:59:59	BV550103870016
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:43:09	101/06/19	23:59:59	BV550103870015
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:43:21	101/06/19	23:59:59	BV550103870014
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:43:39	101/06/19	23:59:59	BV550103870011
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:43:52	101/06/19	23:59:59	BV550103870013
HICLD11 hicloud 微運算型-XS雲伺服器(Linux)	101/06/19	22:44:08	101/06/19	23:59:59	BV550103870012

服務項目	起始日期	起始時間	終止日期	終止時間	雲端服務編號	流量
HICLD21 hicloud 計量制流量	101/06/19	22:43:52	101/06/19	23:59:59	BV550103870013	10MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						10MB

花了 28 塊 (10 台*2 小時*1.4 元)
學 Hadoop 叢集佈署

倘若您手邊有兩台以上的實體機器，或者多核心效能強大的伺服器，不妨安裝 VirtualBox 虛擬化軟體，在一台電腦上建立兩至三台 Ubuntu 10.04 的 Linux 環境，各位聰明的讀者可以透過上述的步驟來安裝您專屬的 Hadoop 叢集，就算是一個小型的私有雲了。

在本文最後，筆者想提出幾個關於「雲端精算師」的軟體設計精神，與各位讀者分享。目前 hicloud-hadoop 0.0.1 版本是循序安裝，因此機器愈多，安裝時間愈長。未來將加入多執行緒的程式碼，讓安裝流程能同時進行。對 hiCloud 這種按時計費的公有雲而言，如何節省安裝時間變得很重要。未來 hiCloud 若能讓使用者建立自己專屬的虛擬機器範本，相信能有效降低各位讀者荷包的負擔。

其次，中華電信的計價模式還包含網路流量，因此如何**減少網路傳輸也是採用公有雲服務應該要思考的第二個重點**。目前的腳本每一台虛擬機器都會重複下載 Java 與 Hadoop 安裝檔。若能把 Hadoop 安裝檔先存在同一個公有雲提供的「儲存服務」，讓資料來自於「內部網路」而非「網際網路」，不納入流量計費，那也是對用戶有利的計價模式。

最後，若能夠把要處理的事情變成自動化，計算完就關閉虛擬機器，以上這些考量就是在雲端時代程式設計師應該具備的「雲端精算師」精神，**擁有節省成本的思考模式將是您未來在職場上的新競爭力！**