



國網中心大型臺灣杉1號與TWCC介紹

Linux 使用介紹與文字編輯器實作

國網中心 工程師

Y.Y Chen

2020.08.17

Outline

- 國網中心服務介紹
 - 台灣杉 HPC
 - TWCC
 - iService
- Linux 使用介紹及實作
 - 何謂 Linux
 - 使用 Linux (TWCC 虛擬運算服務/台灣杉一號主機)
 - <https://hackmd.io/@yychen1218/B1eclvbZP>
 - 在 Linux 上編輯文字檔案

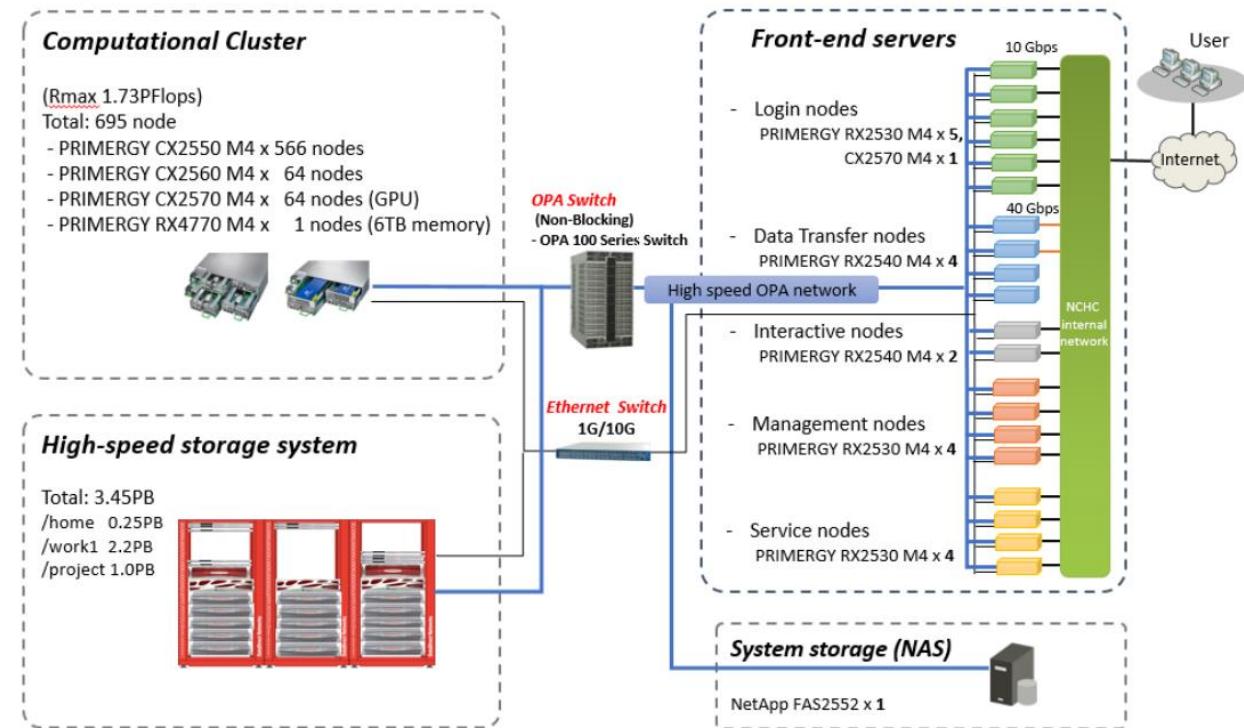
國網中心服務介紹

台灣杉(TAIWANIA) HPC

- 高速計算叢集
 - 107年5月上線服務，7月1日開始收費
 - 630台純CPU節點，64台GPU節點，總記憶體容量達157TB。
 - 3PB以上巨量平行檔案系統
- 服務模式
 - 適合大規模(跨節點)平行計算，如物理、化學、工程、生物。
 - 適合批次(batch)作業，大頻寬平行讀寫。
 - 提供網頁介面操作(SimPlatform)。
 - 沒有虛擬化機制，不適合網站、資料庫等一般的網路服務應用。
- 以透過排程系統申告佔用(allocation)之計算資源數量與時間乘積計費，如CPU核心小時或GPU小時。

台灣杉一號：系統概觀

- 組成
 1. 計算叢集 (Computational Cluster)
 2. 前端伺服器 (Front-end Servers)
 3. 高速儲存系統
 4. 系統儲存 (網路附加儲存) (System Storage (NAS))
- 由乙太網路與 Intel Omni-Path 高速網路互相串連

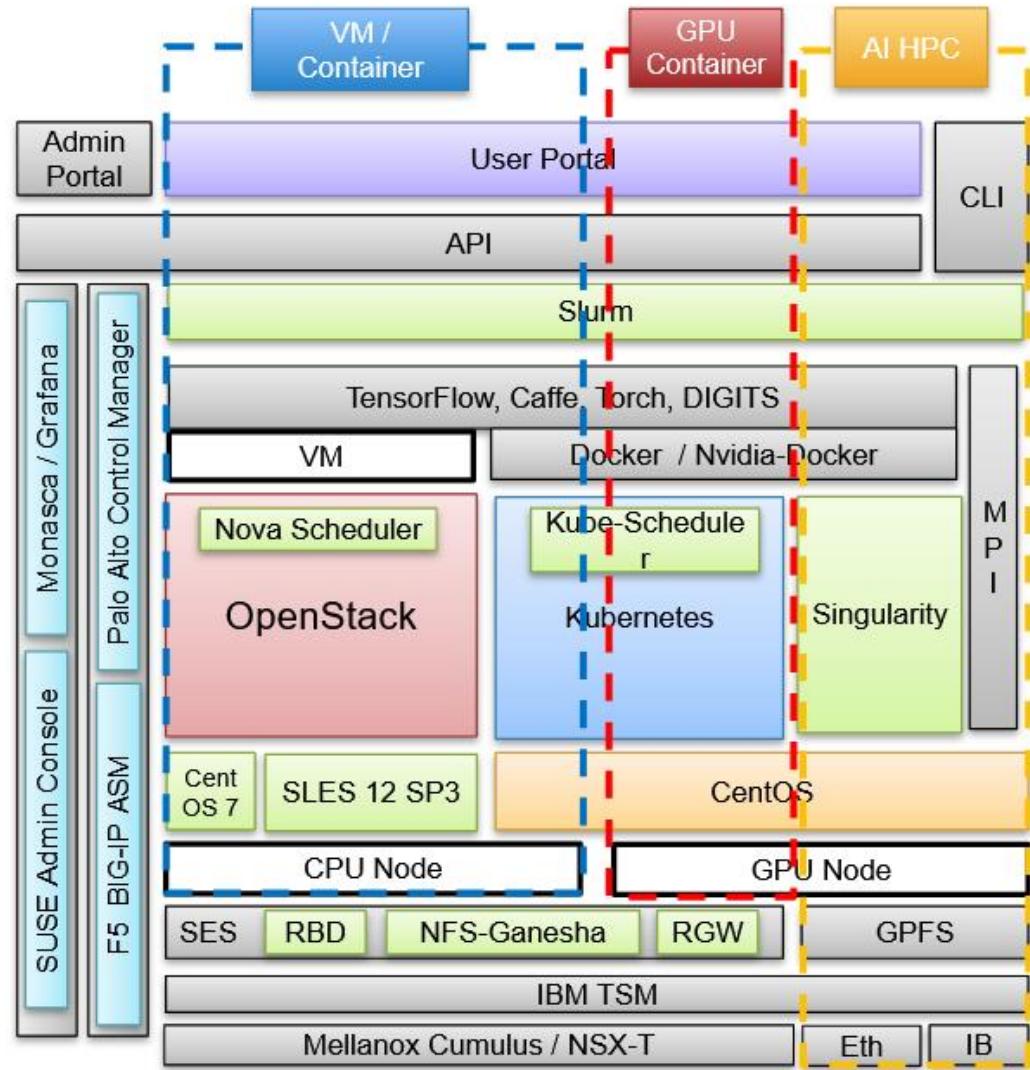


台灣杉一號：使用

- 一般命令列節點
 - 140.110.148.11 clogin1.twnia.nchc.org.tw
 - 140.110.148.12 clogin2.twnia.nchc.org.tw
- 互動式節點 - 可使用圖形化桌面，請勿進行計算工作
 - 140.110.148.17 intact1.nchc.org.tw
 - 140.110.148.18 intact2.nchc.org.tw
- 資料傳輸節點 - 以 40Gbps HCA 介面卡連接外部網路，支援 scp/sftp 傳輸，無 shell
 - 140.110.148.21 xdata1.twnia.nchc.org.tw
 - 140.110.148.22 xdata2.twnia.nchc.org.tw
- 生醫資訊節點
 - 140.110.148.14

*以上節點可由台灣 IP 直接連線進入，若不在台灣則無法連到主機

TWCC



- 建立整合式統一使用者入口網站
- 提供計算服務網站，提供用戶 AI 計算、大資料計算及儲存管理服務
 - 支援Docker、Singularity與Shifter等容器技術
 - 透過OpenStack, Slurm 與Kubernetes 進行資源管理
 - 採用VMWare NSX-T SDN技術
 - 提供統一的自助式入口網站提供基礎架構、應用、資料服務服務定義和可重複自動化工作流的標準化，保證資源控制和法規遵循及合規性，減少由於錯誤而導致的時間浪費

TWCC 虛擬運算服務

- 部署
 - 從網頁通過簡單的圖形化界面即可快速建立虛擬機
 - 支援多種作業系統、客製化 CPU 數量、記憶體等硬體資源
- 備份
 - 快照
- 安全性
 - 自行設定網路安全性群組
 - 預設使用鑰匙對登入

計費

- 台灣杉一號
 - 申請帳號時有 NTD\$200 試用額度
 - 以核心小時 (使用一顆 CPU/GPU 一小時) 為單位
- TWCC (<https://www.twcc.ai/doc?page=price>)
 - 計價項目分為運算、儲存、網路
 - 以秒為單位計價，無須擔心用不滿一小時

iService

- 使用國網中心計算資源需透過 iService 登入
 - 台灣杉一號
 - TWCC
 - 國網生科雲(LIONS)
 - etc
- 可透過 Facebook、Google 帳號快速申請及登入
- 申請帳號需驗證台灣手機號碼

iService 功能介紹

1. 管理主機帳號資訊及查看 OTP
2. 計畫申請及管理
 - 檢視計畫狀態
 - 管理成員(需管理員權限)
 - 購買/管理額度(需管理員權限)
 - 查看明細報表
3. 查看各服務目前使用狀況
4. 線上提問

iService 帳號申請

- https://iservice.nchc.org.tw/nchc_service/

- 點選右上角註冊
- 閱讀並同意服務條款、個資及權利義務說明
- 填寫帳號資料或使用其他帳號登入
- 填寫主機帳號密碼(用於登入Taiwania/TWCC主機)
- 收取 Email 認證信
- 收取手機驗證碼簡訊
- 註冊完成

The screenshot shows the iService account registration process. It includes:

- Step 1:** 告知會員資訊及權利義務聲明 (Information and rights and obligations statement).
- Step 2:** 填寫會員基本資料 (Enter basic member information). This step is highlighted with a blue bar.
- Step 3:** 取得認證信 (Obtain verification code).
- Step 4:** 驗證成功 (Verification successful).
- Basic Member Information Form:** Fields include:
 - 會員資料:** E-mail (必填), 主機帳號資料 (Optional: Facebook, Google, EduRoam quick login).
 - 會員帳號資料:** *請輸入您的E-mail，作為登入服務網的會員帳號(或是可以選擇既有Facebook、Google或EduRoam帳號快速登入) (Please enter your E-mail as the member account for logging into the service website, or choose to log in with existing Facebook, Google, or EduRoam accounts).
 - 會員密碼:** *會員密碼 (必填), *再次輸入會員密碼 (Optional: Enter and re-enter member password).
- Consent Page:** A red arrow points from the '註冊' button in the top right to the '我同意' (I agree) button on the consent page. The page also contains a note: '本人已詳閱以上告知內容，並同意以上之告知內容。' (I have read the above information in detail and agree to the above information.)

加入 TWCC 計畫

- 需加入有餘額的計畫才可使用 TWCC
- 請助教協助將各位加入計畫

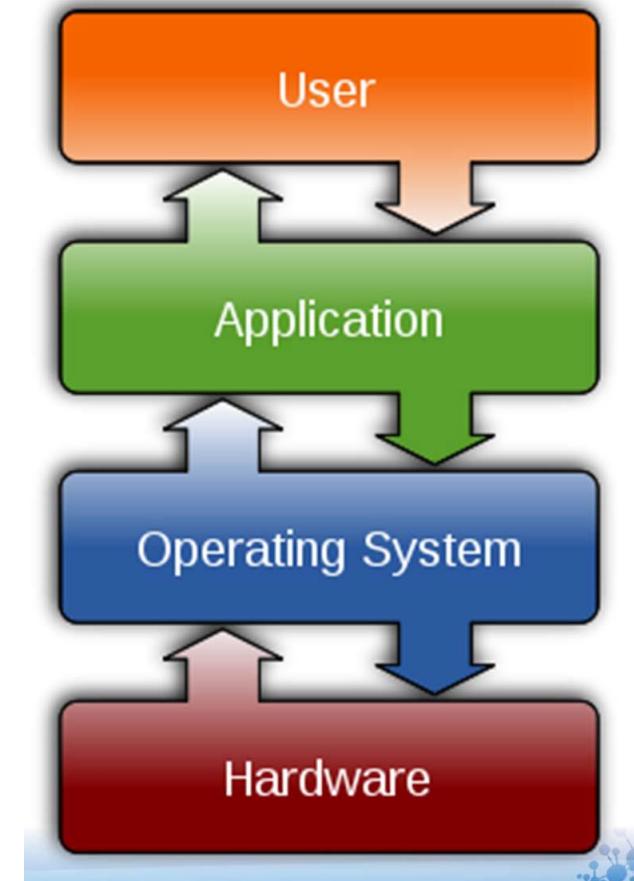
其他使用說明

- 使用上若有任何問題，可至 iService 操作說明頁面查看說明文件或影片
 - [台灣杉一號說明](#)
 - [TWCC 說明](#)
 - [iService 說明](#)

LINUX 介紹

何謂 Linux

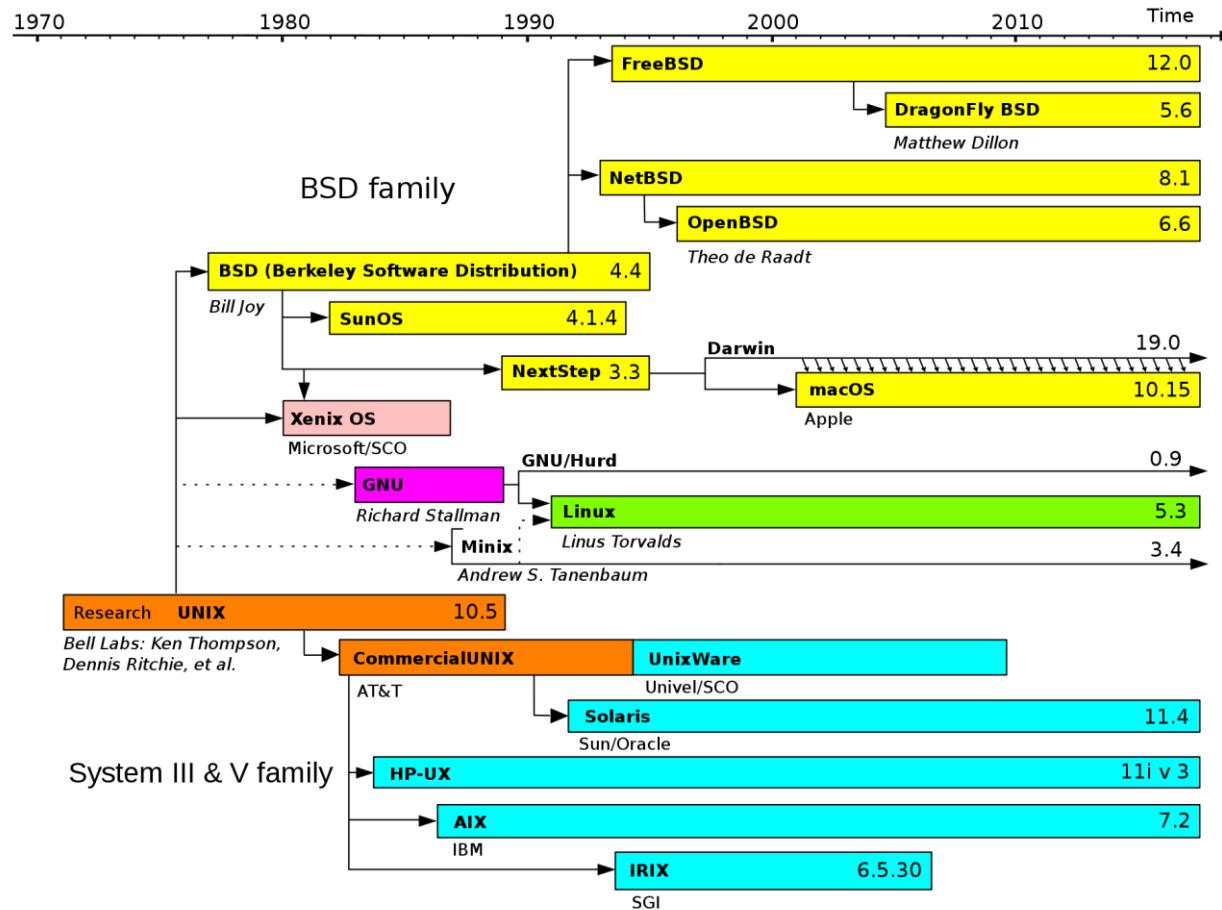
- 作業系統 (Operating System, OS) ，像是 Windows、MacOS、iOS、Android 等。管理軟硬體資源、提供軟體執行的平台。
- 開源
- Unix-Like OS



Unix

- Multitasking、Multiuser
- AT&T 貝爾實驗室
- 1969 發行第一版
- 1973 以 C 語言重寫
- 1991 芬蘭赫爾辛基大學 Linus Torvalds 發表 Linux

Unix & Unix-Like 簡史



Linux 特色

- 多人 (Multiuser)、多工 (Multitasking)
- 多執行緒 (Multiple Threading)
- 支援多 CPU
- 支援多種處理器 (Intel x86 相容 , AMD 64 bit, IA32/IA64, PowerPC, Sun Sparc, Dec Alpha,) 、硬體
- 套件 (發行版本)
- 常用於個人電腦、伺服器、嵌入式裝置、行動裝置等

Linux 發行版本

- 一些團體蒐集可使用於 Linux 的應用程式及管理工具等，製作成的完整作業系統
- 常見免費版本：Ubuntu、CentOS、Mint等
- 付費企業版：Red Hat Enterprise Linux、SuSE Linux Enterprise Server等

使用 Linux

- 終端機文字介面
 - 本地
 - 各種終端機工具
 - Windows : PuTTY、MobaXterm、Bitvise、PowerShell等
 - Mac : 終端機等
 - iOS/Android : Termius等
- 圖形介面
 - 本地
 - VNC、ThinLinc等

建立並登入 TWCC 虛擬運算容器

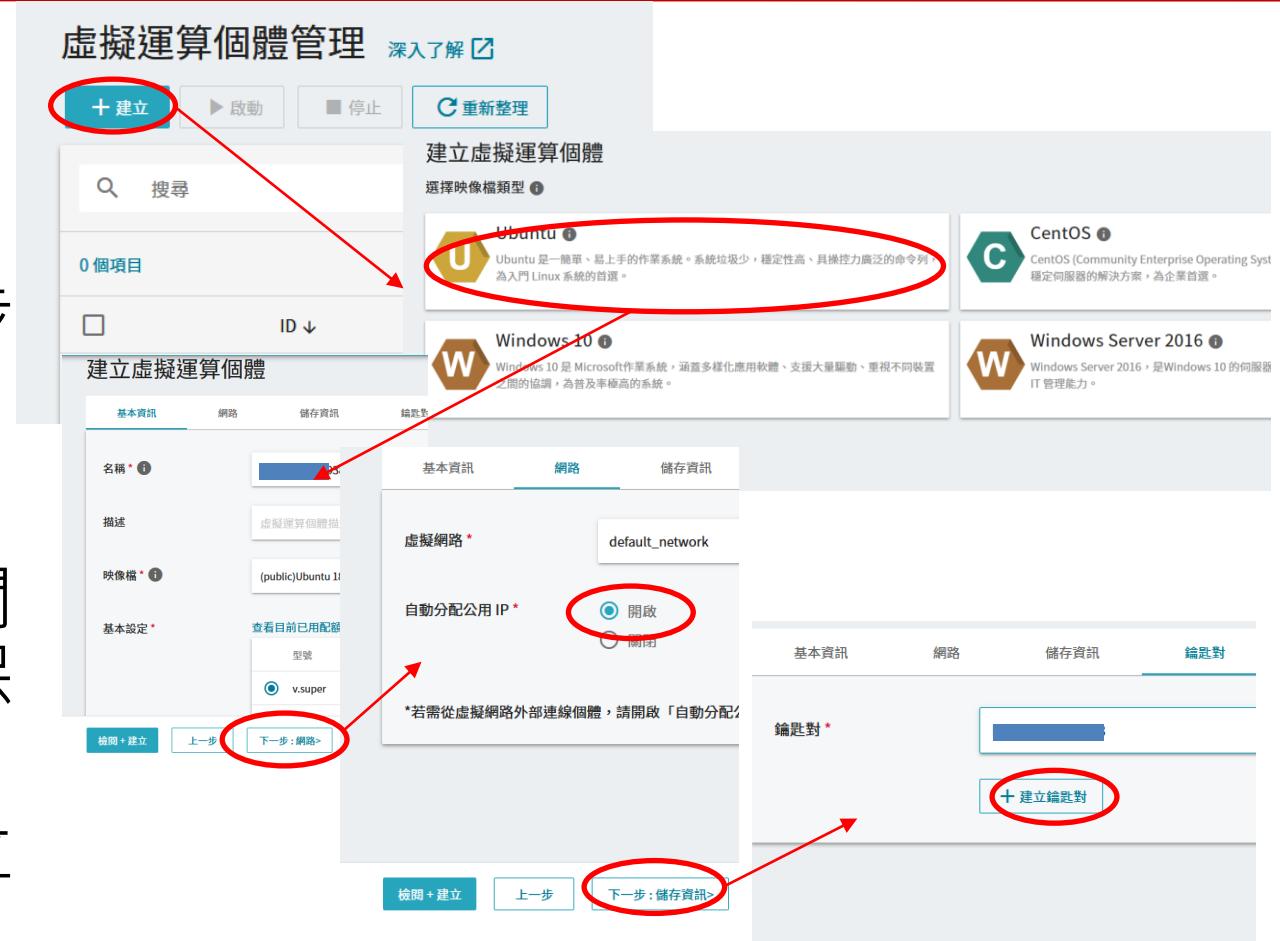
登入 TWCC

- <https://www.twcc.ai/>
 - 點選右上角登入，以 iService 帳號登入
 - 若是登入後顯示沒有計畫，請找助教將您加入教育訓練計畫
 - 登入後左上角選擇課程計畫，並確認可用額度非 0
 - 點選左下角「運算」的「虛擬運算」進入管理畫面



建立 TWCC 虛擬運算容器

1. 左上角「建立」
2. 左上角「Ubuntu」
3. 自行設定描述、名稱後下一步
4. 開啟自動分配公用 IP
5. 儲存資訊不改動下一步
6. 點選建立鑰匙對，不輸入公開金鑰直接確認，下載後妥善保管，請勿外流
7. 確認輸入資訊正確後點選建立
8. 等待狀態變成 Ready(綠燈)

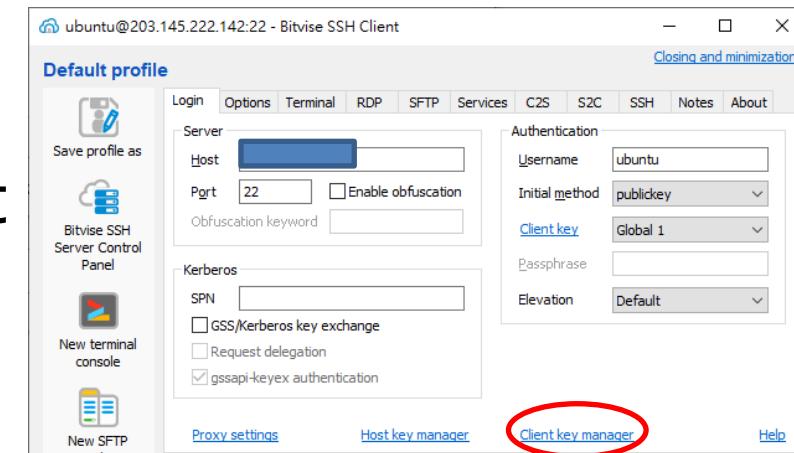
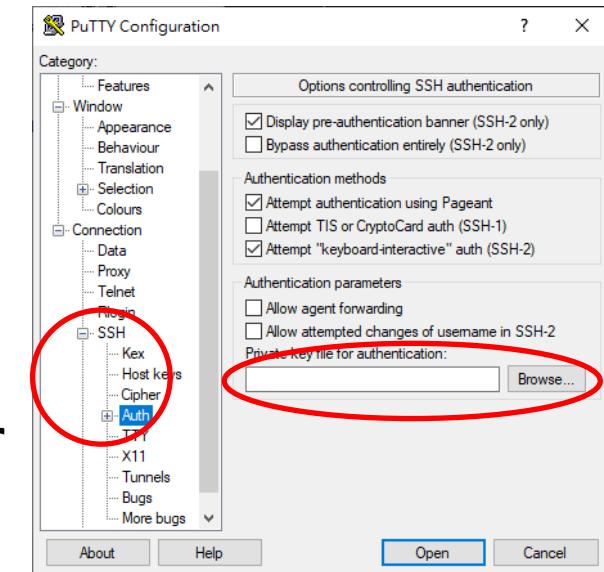


常用設定及操作

- 快照：快速備份，以便復原或建立相同內容的虛擬機
- 安全性群組：設定允許連線進/出虛擬機的 IP 網段、協定及 Port
 - ex：想要建立 http 伺服器，可以設定
 - 方向：ingress (連入)
 - 連接埠範圍：最小 80 最大 80
 - 協定：tcp
 - CIDR：0.0.0.0/0 (任何 IP)

首次登入

- 使用鑰匙對
- 預設帳號 : ubuntu
- SSH登入
 - PuTTY : SSH → AUTH → Private key file for authentication → login 輸入 ubuntu
 - Bitvise : Client key manager → Import → \$PATH (找不到選 All Files) → Import → Initial method 選 publickey
 - Terminal : \$ ssh -i \$PATH ubuntu@host

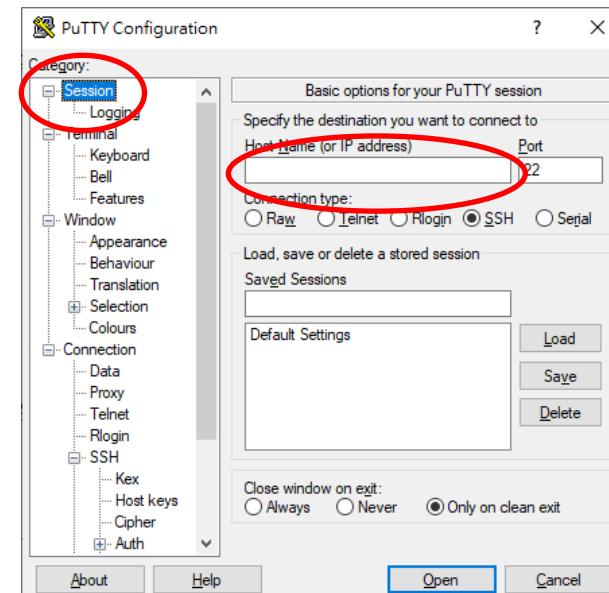
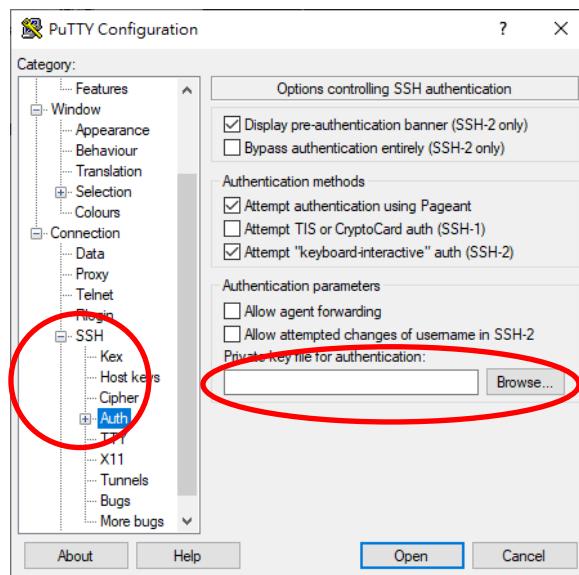


透過 SSH 遠端登入

- Host：
 1. 虛擬運算個體管理
 2. 剛剛建立的虛擬機
 3. 網路與連線 → 公用 IP
- Port : 22(大部分工具預設)
- Username : ubuntu
- Password : 無
- 鑰匙對 : 剛剛建立過程中下載的 pem 檔

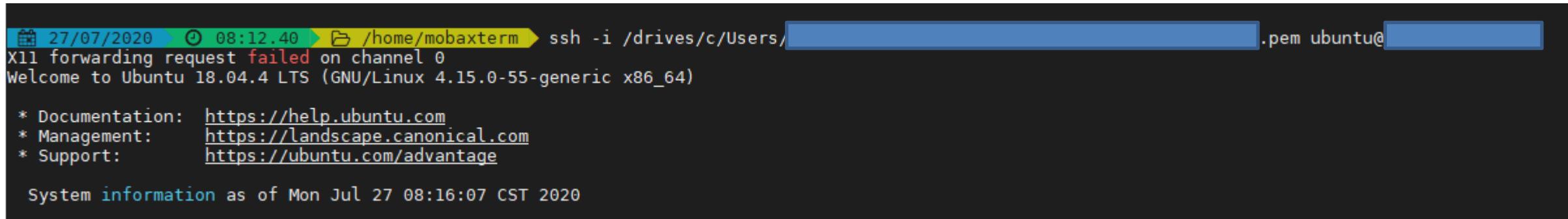
透過 SSH 遠端登入

- PuTTY：
 - 使用 PuTTYgen 將剛剛下載的 pem 檔轉為 ppk 檔
 - 在 SSH → AUTH → Private key file for authentication 選擇 ppk 檔
 - 回到 Session → 於 Host Name 輸入虛擬機的公用 IP 後按 Open
 - 出現警告視窗時按是
 - 在 login as: 後輸入 ubuntu 並按 Enter



透過 SSH 遠端登入

- MobaXterm：點選 Start Local Terminal 或開新分頁之後輸入
 - ssh -i 鑰匙對所在路徑 ubuntu@host
 - Windows 下的 MobaXterm 中以 /drives/c 表示 C 槽，其他槽以此類推



A screenshot of a terminal window showing an SSH session. The terminal title bar includes the date (27/07/2020), time (08:12.40), and current directory (/home/mobaxterm). The command entered is ssh -i /drives/c/Users/.ssh/id_rsa.pub ubuntu@host. The session output shows:

```
27/07/2020 08:12.40 /home/mobaxterm ssh -i /drives/c/Users/.ssh/id_rsa.pub ubuntu@host
X11 forwarding request failed on channel 0
Welcome to Ubuntu 18.04.4 LTS (GNU/Linux 4.15.0-55-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

System information as of Mon Jul 27 08:16:07 CST 2020
```

透過 SSH 遠端登入

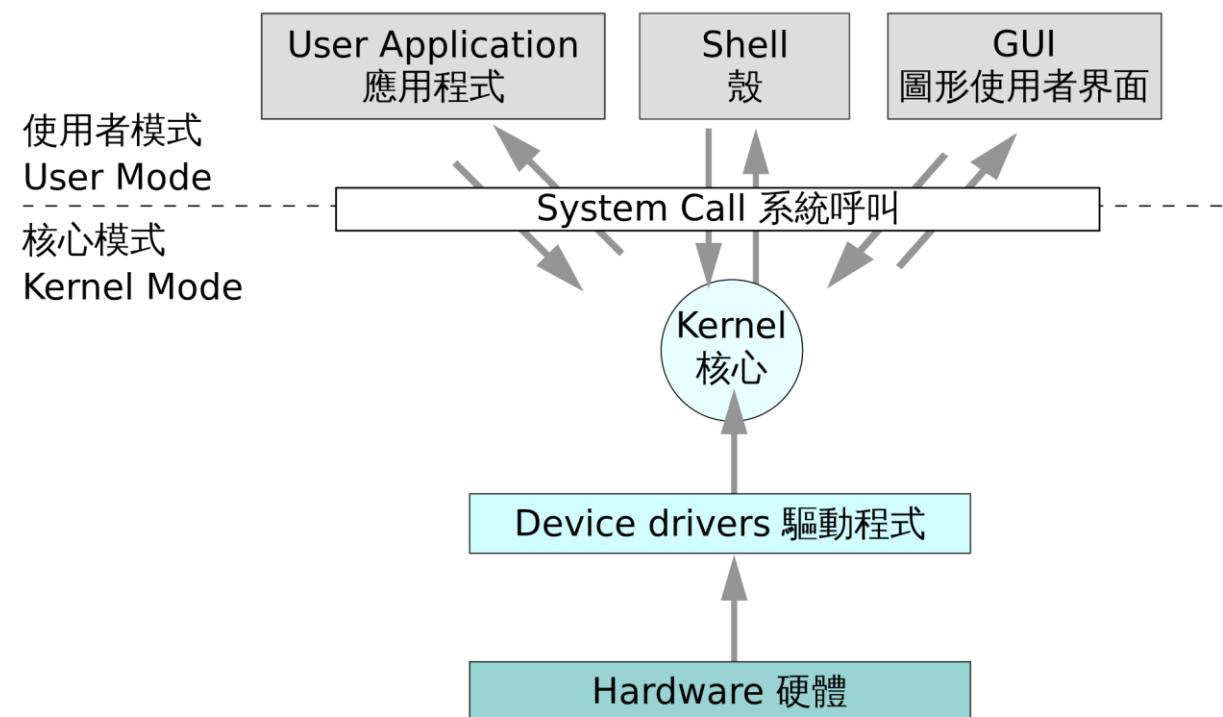
- 終端機工具(Mac 終端機、Windows PowerShell 等)
 - 輸入 ssh -i 鑰匙對所在路徑 ubuntu@host
 - 使用 PowerShell 需先將 pem 檔權限改成只有允許自己的帳戶可以使用 (內容 → 安全性 → 進階設定)

Shell

- 使用者用 SSH (Secure SHell) 連入系統時即啟動一個 Shell
- Shell 為使用者執行應用程式的一種介面的通稱，大部分 Linux 預設使用 bash，其他也有 zsh、csh等等
- 可以使用 chsh 指令切換使用的 shell

Shell

- 使用者透過使用者介面與 Kernel 溝通
- Kernel 透過驅動程式控制 硬體



環境變數

- 每個使用者登入啟動一次 shell 都會有一組環境變數(通常預設每個使用者登入時都相同)
- 使用者可以通過指令增加、修改、刪除環境變數，進而改變某些程式的行為
- env：列出所有環境變數
- export：增加/修改環境變數
- echo：印出東西，可以用來印出環境變數
 - ex : echo \$PATH

Prompt 命令提示字串



```
[yychen1218@login4 ~]$ | ubuntu@vm1595582893304-1090807-iaas:~$
```

- Prompt 為提示使用者現在可以輸入命令的字串
- 左圖以台灣杉一號主機為例，字串內容為
[username@主機 路徑]\$
- 右圖以 TWCC 虛擬運算服務的 Ubuntu 預設為例，字串內容為
username@主機:路徑\$
- 根據系統設定不同，常見的 prompt 結尾字元為 \$、#、>等
- 使用者可以修改設定檔來更改自己的 bash 的 prompt 的顯示資訊
- 於終端機執行的 shell 以外的程式有時也會使用這種提示

Linux 檔案系統

- 不同於 Windows，Linux 不以「槽」，而是以「根目錄」最為檔案系統的最上層
- 不同的硬碟會掛載在不同的目錄
- 路徑
 - 絕對路徑：以根目錄(/)開頭的路徑(ex : /home/username、/share)
 - 相對路徑：相對於使用者所在位置的路徑(ex : ../、mydir/)
- 可以使用 `pwd` 指令來得知自己所在的位置
- 家目錄(/home/username)常被縮寫為 ~，此設定可以在設定檔修改

Linux 檔案系統

```
total 264
drwx----- 33 yychen1218 TRI1091216 4096 Jul 17 11:33 .
drwxr-xr-x 331 root      root      20480 Jul 17 14:24 ..
-rw-r--r--  1 yychen1218 TRI1091216    12 Jul 16 17:57 123.txt
-rw-----  1 yychen1218 TRI1091216 21646 Jul 17 13:37 .bash_history
-rwx-----  1 yychen1218 TRI1091216    18 Jul 12 2016 .bash_logout
-rwx-----  1 yychen1218 TRI1091216   193 Jul 12 2016 .bash_profile
-rwx-----  1 yychen1218 TRI1091216   247 Nov 17 2017 .bashrc
```

檔案屬性及權限 檔案擁有者 檔案所屬群組 檔案大小 檔案修改日期 檔案名稱

- Linux 檔案屬性

- 類型(屬性第一碼)：一般檔案(-)、目錄(d)及連結檔案(l)
- 權限：
 - 分為擁有者權限、所屬群組權限及其他權限
 - 分為讀(r=4)、寫(w=2)、執行(x=1)、不可(-=0)



擁有者權限 所屬群組權限 其他權限

Linux 檔案系統

- 檔案權限可以三碼數字表達
 - 750 表示擁有者可讀可寫可執行($4+2+1=7$)、所屬群組可讀可執行($4+0+1=5$)、其他人無權限($0+0+0=0$)
- 檔案擁有者使用 chmod 可更改檔案的權限
 - ex : chmod 750 file
 - ex : chmod [ugo][+-=][rwx] file
 - 可用 chown 、 chgrp 更改檔案的擁有者及所屬群組
- 執行程式、讀取檔案時若出現 **permission denied**，可檢查相關檔案權限

Linux 指令

- 在 prompt 出現時可輸入指令，常見格式為

```
$ command [-options] parameter1 parameter2 ... # comment
```

- command 為主要指令
- options 為選項，有時也會稱為 flag，通常由 - 或 -- 開頭，有時可將幾個 flag 合在一起(ex: \$ ls -al = \$ ls -a -l)
- 以上各部分以空格區隔，無論幾個空格都視為一個，但用 "" 包起來的部分會視為同一個字串
- # 之後的部分為 comment (註解)，會被 shell 忽略

Linux 指令

- 指令太長時可用 \ (反斜線)符號跳脫 enter，使按 enter 時不會直接執行，可以再輸入下一行
- 某些符號有特殊涵意，因此要表達該符號本身時需要在前面加一個 \
 - ex : echo "\\"
 - \"表示「"」，前後的"表示中間的內容是一個字串
- 不清楚指令的使用方式時
 - 以 \$ man 指令 來查看系統中指令的使用說明
 - 以 \$ 指令 **-help** 來查看指令自己的幫助說明
 - Google 或查詢來源文件

Linux 指令：範例

```
$ ls -a ~/
```

- ls : 列出目錄下的檔案
- -a : 要列出所有檔案，包含隱藏檔案、本目錄(.)及上層目錄(..)
- ~/ : 紿予命令的參數，此時表示要列出~(通常為家目錄)底下的檔案

```
$ ls -lh or $ ls -l -h
```

- -lh或-l -h : 顯示詳細資訊(l)，且大小以K/M/G為單位(h)
- 沒有輸入參數，預設列出所在目錄的檔案

Linux 指令：以 root 權限執行 & 安裝

- 為了系統安全，並非所有使用者都可以修改所有檔案，但是 super user (一般為 root) 可以進行所有操作
- sudo (Super User Do) 指令可以讓使用者以 root 權限進行一些操作
- 只有被寫入 sudoers 設定檔的使用者可以使用這個指令
- 若此使用者有設定密碼，使用 sudo 時需要輸入自己的密碼
- apt/apt-get 則是 Ubuntu 等系統下用來安裝 package 的指令(CentOS 等系統則是使用 yum)

```
$ apt-get install zsh # 應該會出現 permission denied  
$ sudo apt-get install zsh
```

*zsh 為另一款 shell 軟體，有興趣可以試著使用

Linux 指令：建立使用者

- 擁有權限的人可以在系統中新增使用者
- 使用者可以屬於一些 group
- 用下列指令新增 user (在前面加 sudo 以使用 root 權限)

```
$ adduser [--uid [指定uid] --ingroup [指定所屬 group] --  
disabled-password] username # 加入 --disabled-password  
可以暫時不設密碼
```

- 用下列指令修改密碼

```
$ passwd # 修改自己的密碼
```

```
$ passwd username # 修改其他使用者的密碼，需權限
```

*請勿在開啟密碼登入 ssh 的機器使用太簡單的 username/password，有被猜中入侵的風險，出現異常網路流量或 CPU 執行的虛擬機會被系統暫停運作

Linux 指令：以使用者身分執行

- 使用者可以使用 su 指令來以另一個使用者的身分登入
- su 不同於 sudo，是要輸入欲登入使用者的密碼
- 使用 - flag 可以設定新的變數環境(模擬重新登入 shell)
- 練習：

```
$ echo $HOME
```

```
$ su - username # 剛剛建立的 user
```

```
$ echo $HOME # 應該會跟以 ubuntu 執行時的值不同
```

啟用 ssh 密碼登入

- 沒有啟用 ssh 密碼登入時，除了 ubuntu 以外無法直接 ssh 登入
- 將 /etc/ssh/sshd_config 中的 PasswordAuthentication no 改成 PasswordAuthentication yes
- 重啟 ssh 服務
`$ sudo service ssh restart`

其他常用設定檔

- 所有使用者共用的設定檔通常位於 /etc/ 目錄
- 使用者自己的設定檔通常放在家目錄下，比如
 - bash : ~/.bashrc (修改後需用 source 指令重新讀取)
 - vim : ~/.vimrc (修改後直接套用新設定)

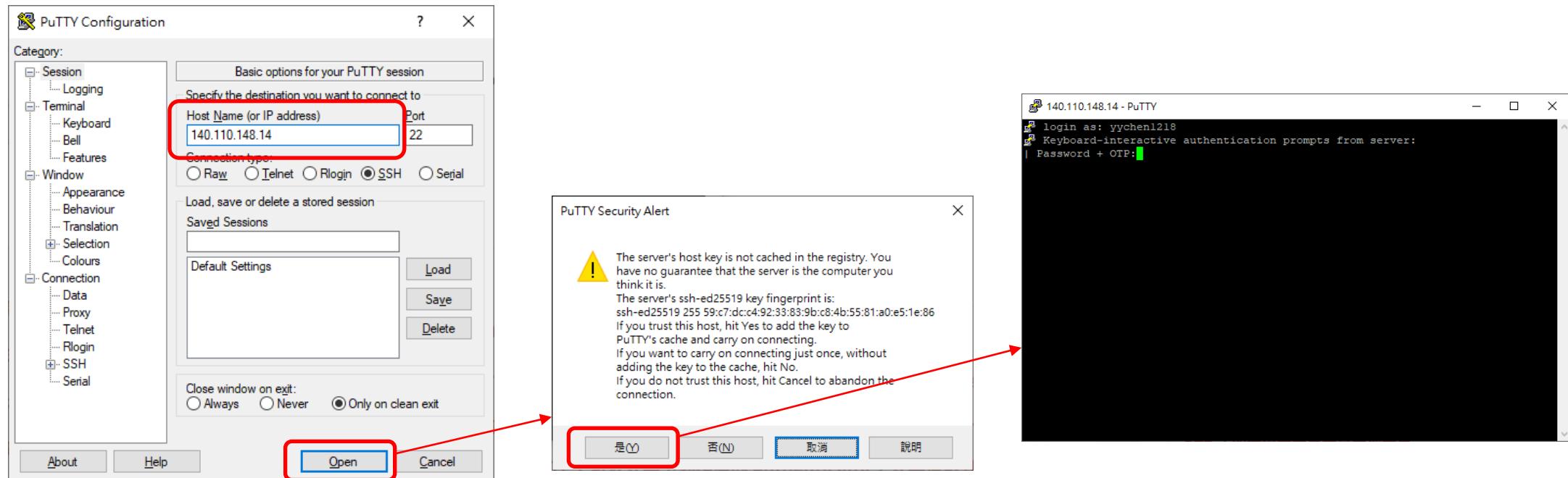
登入台灣杉一號主機

透過 SSH 遠端登入

- Host : 140.110.148.14 (生資節點)
- Port : 22(大部分工具預設)
- Username : 自行設定，可至 iService 主機帳號資訊查看
- Password : 自行設定，可至 iService 主機帳號資訊重設
- OTP : 至 iService 主機帳號資訊查看，或使用 Authenticator 搭配金鑰於其他裝置查看

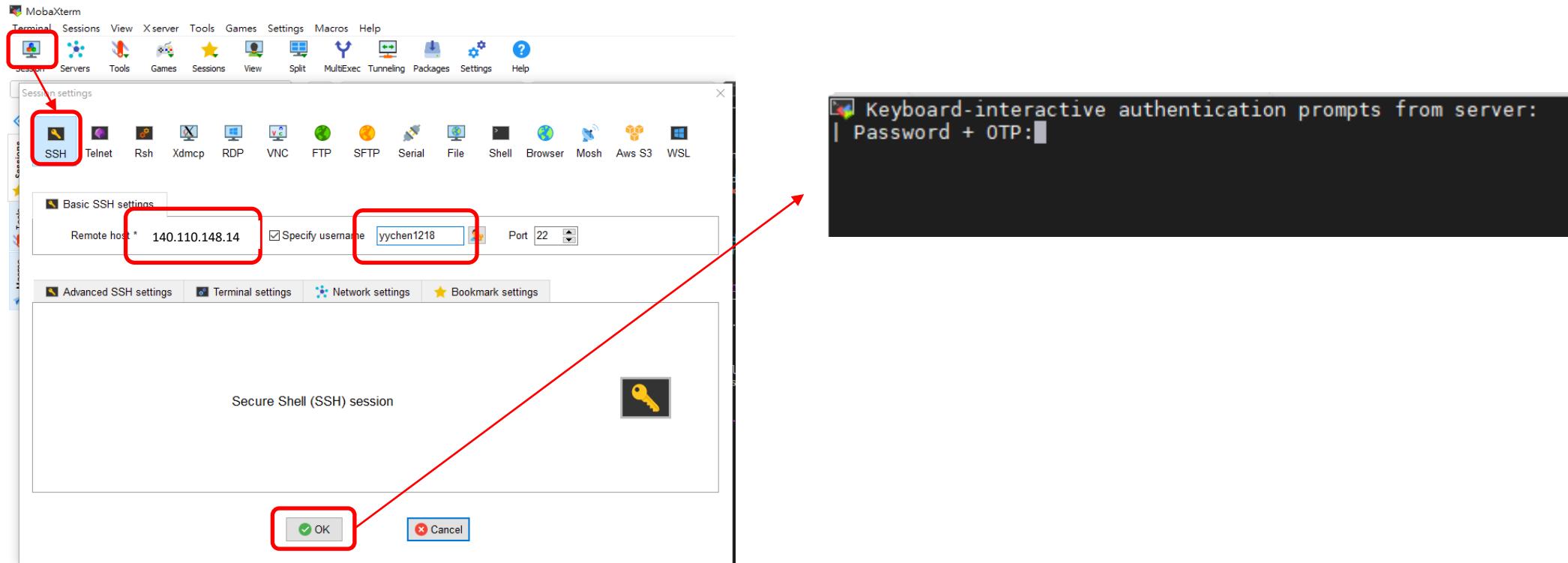
透過 SSH 遠端登入

- PuTTY：在 login as: 後輸入主機帳號、待出現 Password + OTP: 後輸入密碼+OTP



透過 SSH 遠端登入

- MobaXterm : 看到右圖畫面後輸入密碼+OTP



透過 SSH 遠端登入

- 終端機工具(Mac 終端機、Windows PowerShell 等)
 - 輸入 ssh username@host
 - 看到 Password + OTP: 後輸入密碼+OTP



A screenshot of a Windows PowerShell window titled "Windows PowerShell". The window displays the following text:
Windows PowerShell
Copyright (C) Microsoft Corporation. 著作權所有，並保留一切權利。
請嘗試新的跨平台 PowerShell <https://aka.ms/pscore6>
PS C:\Users\Y.Y Chen> **ssh** yychen1218@140.110.148.14
Password + OTP:

Linux 指令：基本資訊

- 查看自己是誰
 - whoami
 - id (會列出所屬的群組及 uid 等資訊，也可以查看別人的資訊)
- 查看路徑
 - pwd
- 切換目錄
 - cd [目標目錄]
- 查看目錄下的檔案
 - ls
- 新增目錄
 - mkdir

Linux 指令練習：基本資訊

- 登入 ssh 的終端機工具
- 查看自己是誰

```
$ whoami
```

- 查看自己所在的位置

```
$ pwd # 應該會在 /home/username
```

- 若不在自己的家目錄，切換到自己的家目錄

```
$ cd ~
```

或

```
$ cd /home/`whoami` # `命令` 會被替換成執行後的結果
```

Linux 指令練習：下載與解壓縮

```
$ mkdir nchc0817 # 新增叫做 nchc0817 的目錄
```

```
$ cd nchc0817
```

```
$ pwd # 應該會在 /home/username/nchc0817
```

- 下載網路上的檔案：

```
$ wget http://humem.nchc.org.tw/NGS/files/ngsData.zip
```

- 查看是否有下載成功：

```
$ ls # 應該出現 ngsData.zip
```

- 將 zip 檔解壓縮：

```
$ unzip ngsData.zip
```

Linux 指令：Pipe 及 Redirect

- 用 | 符號串起兩個指令，能將前面的指令輸出接到後面指令的輸入
- 在指令末尾加入 > 檔名，會將指令輸出寫入檔案
- 在指令末位加入 >> 檔名，會將指令輸出寫在檔案原本的內容後面

Linux 指令練習：Pipe 及 Redirect

- 查看解壓所後的資料夾有什麼，並且標示行號：

```
$ ls ngsData | nl # nl 指令為印出標準輸入或檔案內容，並加上行號標示
```

- 將以上的內容寫到一個叫做 result.txt 的檔案：

```
$ ls ngsData | nl > result.txt
```

- 在 result.txt 最末尾加入一行字

```
$ echo "last line" >> result.txt
```

Linux 指令練習：查看檔案內容

```
$ cat result.txt # 查看檔案內容
```

```
[yychen1218@clogin4 ncku0810]$ cat result.txt
1 alignment.sh
2 f010_raw_mirna.fastq
3 GCF_000013425.1_ASM1342v1_genomic.fna
4 gene_info.txt
5 SRR6491087_1.fastq.gz
6 SRR6491087_2.fastq.gz
7 SRR6491087.tar
last line
```

```
$ tac result.txt # 從最後一行看到第一行
```

```
$ nl result.txt # 查看檔案內容，同時加上行號
```

```
$ head -n 3 result.txt # 查看前三行
```

```
$ tail -n 3 result.txt # 查看最後三行
```

Linux 指令練習：查看檔案內容

- 當查看的檔案很長時，除了 `head` 及 `tail`，還可以用以下指令更方便查看

```
$ more ngsData/f010_raw_mirna.fastq # 查看檔案內容，可以一行一行  
或一頁一頁看
```

```
$ less ngsData/f010_raw_mirna.fastq # 查看檔案內容，可以一行一行  
或一頁一頁看，且可以往回看
```

```
$ less -NS ngsData/f010_raw_mirna.fastq # -N 顯示行號、-S 超過  
頁面寬度時不自動換行
```

- 以上兩個指令以 `enter` 鍵換行、空白鍵換頁，`less` 另可以上下鍵換行、`pgup/pgdn` 鍵換頁

Linux 指令：檔案操作

- 新增目錄
 - `mkdir [目錄名稱]`
- 移動檔案或目錄，目標是存在的目錄下不存在的檔案時，會將檔案移動並改名
 - `mv [要移動的檔案] [目標]`
- 複製檔案或目錄
 - `cp [要複製的檔案] [目標]` (複製整個目錄要加上 `-r`)
- 刪除檔案或目錄
 - `rm [檔案]` (刪除整個目錄要加上 `-r`)
 - `rmdir [目錄]` (用來刪除目錄)

Linux 指令練習：檔案操作

```
$ mkdir mydir # 新增一個目錄
```

```
$ mv result.txt mydir/ # 將 result.txt 移動到 mydir 中
```

```
$ ls mydir/
```

```
$ ls # 確認結果
```

```
[yychen1218@clogin4 ncku0810]$ mkdir mydir  
[yychen1218@clogin4 ncku0810]$ mv result.txt mydir/  
[yychen1218@clogin4 ncku0810]$ ls mydir/  
result.txt  
[yychen1218@clogin4 ncku0810]$ ls  
mydir ngsData ngsData.zip
```

```
$ cp mydir/result.txt . # 將 mydir 中的 result.txt 複製到  
所在目錄(.)
```

```
$ cp -r mydir mydir2 # 複製整個目錄並取名為 mydir2
```

```
$ ls mydir2/ # 確認結果
```

```
$ mv mydir2 mydir3 # 將 mydir2 改叫 mydir3
```

Linux 指令：字串處理

- 搜尋：

```
$ grep target file
```

- 插入(a)、取代(c/s)、刪除(d)：

```
$ sed -e 'script' [-e 'script' ...] file # 加入 -i flag
```

可直接將結果寫回原檔，加入 -n flag 可以只印出經過處理的行

- 以指定字元分割字串並處理：

```
$ awk -F 'delimiter' 'script' file
```

- 以上命令接支援 regex (正規表示法)

Regular Expression 正規表示法

- Regular Expression (常被簡稱為 Regex)是一種用來比對字串的 pattern，常以 /pattern/ 表示
- 常用表示：
 - .：任何字元
 - *：前一個字元重複 n 次 ($n \in$ 非負整數，aka 可以為 0)
 - []：只要符合裡面的字元就可以，其中 - 表示連續編碼的字元，如數字可用[0-9]，英文字母可用[a-zA-Z]表達
 - ^：放在一開始表示開頭，放在[]中表示不可以為後面接的字元
 - \：跳脫符號，想要表示上述符號時可以在前面加上\
 - \s：空白
 - \$：表示結尾
- 有的程式會再加上自己的獨特規則，也要注意各種語言 / IO 的不同表達

Linux 指令：字串處理

- 一些用法及範例：

```
$ grep -e 'pattern' file # 使用 regex
```

```
$ sed -e '1,5aline' -e '$d' -e 's/123/haha/g' file #  
a:於 1-5 行後各插入一行'line'、d:刪除第$行(最後一行)、s:  
將所有(g)'123'取代為'haha'
```

```
$ awk -F ',' 'match($1,/pattern/){print $2}' file #  
使用 regex 只有第 1 項符合 pattern 的行才印出第 2 項
```

```
$ sort -t ' ' -k2nr file # 以空格為分割，將第二欄(-k2)  
以數字(n)反向(r)排序後印出
```

Linux 指令練習：字符串處理

```
$ grep -e "[0-9]" result.txt # 抓出  
result.txt 中含有數字的行數
```

```
$ sed -i -e '$aadd a line' result.txt # 在  
第$行(最後一行)後面加入(a) 'add a line'
```

```
$ awk '{print $1" "$NF}' result.txt # 以空  
白為分割(預設)，印出第 1 項及第 NF 項(最後一項)
```

```
$ awk 'match($1,/^[0-9]*$/){print $2}'  
result.txt # 以空白分割，當第 1 項為數字時，  
印出第 2 項
```

```
$ sort -t ' ' -k1nr result.txt # 以空  
白為分割，第一欄以數值反向排序後印出
```

```
[yychen1218@clogin4 ncku0810]$ grep -e "[0-9]" result.txt  
1 alignment.sh  
2 f010_raw_mirna.fastq  
3 GCF_000013425.1_ASM1342v1_genomic.fna  
4 gene_info.txt  
5 SRR6491087_1.fastq.gz  
6 SRR6491087_2.fastq.gz  
7 SRR6491087.tar
```

```
[yychen1218@clogin4 ncku0810]$ awk '{print $1" "$NF}' result.txt  
1 alignment.sh  
2 f010_raw_mirna.fastq  
3 GCF_000013425.1_ASM1342v1_genomic.fna  
4 gene_info.txt  
5 SRR6491087_1.fastq.gz  
6 SRR6491087_2.fastq.gz  
7 SRR6491087.tar  
last line  
add line
```

```
[yychen1218@clogin4 ncku0810]$ sort -t ' ' -k1nr result.txt  
7 SRR6491087.tar  
6 SRR6491087_2.fastq.gz  
5 SRR6491087_1.fastq.gz  
4 gene_info.txt  
3 GCF_000013425.1_ASM1342v1_genomic.fna  
2 f010_raw_mirna.fastq  
1 alignment.sh  
add a line  
last line
```

Linux 指令：打包及壓縮

- tar 打包：將多個檔案包成一個檔案，大小沒有壓縮

```
$ tar -cf archive.tar file1 file2 # 將 file1 及 file2 打包成 archive.tar，  
也可直接打包一個目錄
```

```
$ tar -xf archive.tar # 解開打包
```

- zip 壓縮

```
$ zip -r archive.zip file1 file2 # 將 file1 及 file2 打包成 archive.tar，也  
可直接打包一個目錄
```

```
$ unzip arcive.zip # 解壓縮
```

- tar.gz 打包同時壓縮(Linux 系統常用格式，許多非 package 程式包會使用這種)

```
$ tar zcvf archive.tar file1 file2 # 將 file1 及 file2 打包成 archive.tar，  
也可直接打包一個目錄
```

```
$ tar zxvf archive.tar # 解壓縮
```

Linux 指令：常用快捷鍵

- **ctrl + c**：送出終止訊號，大部分時候可以打斷進行中的程式並結束
- **ctrl + z**：送出中斷訊號，大部分時候可以打斷進行中的程式並送到後台暫停，之後再用 `kill` 破掉或 `fg` 叫回繼續執行
- 上下鍵：回溯之前輸入過的指令
- **Tab 鍵**：自動補足指令或路徑(指令或路徑非唯一時根據 shell 軟體行為不同)

Linux 指令：其他常用指令

- clear：清除螢幕
- history：查看之前輸入過的指令
- ifconfig：查看網卡狀態及本機 IP
- scp：在機器間複製檔案
- who：查看有誰在線上
- jobstat：查看 job 的狀態
- ps：查看執行中的 process
- top：監控使執行中的 process 及各自使用的資源，按 q 離開
- diff：比較兩個檔案的差異
- git：常用版本控管程式

Linux 指令：其他常用指令

- df : 查看儲存空間容量
- du : 查看目錄容量
- nslookup : 解析 IP 與 domain name 的對應
- mail : 查看/寄出信件
- crontab : 排程執行
- gcc/g++ : 編譯 C/C++ 原始碼
- make : 根據寫好的 makefile 快速部署
- sort : 排序
- exit : 離開 shell

在 LINUX 上使用文字編輯器

Linux 下的文字編輯器

- 目標
 - 修改檔案
 - 撰寫程式
 - 建立文件
- 使用軟體
 - Vim
 - Nano
 - simPlatform 的 Editor

Vim

- Vim 是一款終端機常用文字編輯器
- 是從 Vi 編輯器發展而來的
- 具有自動縮排、highlight 部分程式語言關鍵字、復原/重作、分割窗口、比較文件.....等功能

Vim : 操作模式

Vim 分為多種模式，Normal 跟 Command-Line 以外的模式會標示在左下角

1. **Normal Mode**：一開始的狀態，此時無法輸入文字，可以移動游標、進行複製貼上刪除等命令。
2. **Insert Mode**：可以輸入文字的模式，使用方式類似一般記事本等常見輸入方式。
3. **Visual Mode**：可以移動游標來選擇文字以進行操作。又有 **Visual Line Mode**、**Visual Block Mode** 等類似模式。
4. **Command-Line Mode**：在 Normal Mode 輸入 : 或 / 或 ? 以輸入命令，命令會顯示在左下角，按 Enter 執行後會回到 Normal Mode
5. **Replace Mode**：可以將游標位置的字元取代成輸入字元。

Vim : 模式切換

一般以 Normal Mode 為基準

- 按 `a`、`I` 或 `o` 鍵切換為 Insert Mode
 - `a`：游標換到下一個字元
 - `i`：游標不動
 - `o`：在所在行下插入一個空白行，同時將游標換到該空白行行首
- 按 `r` 鍵切換為 Replace Mode
- 按 `v` 鍵切換為 Visual Mode
 - `v` (大寫) 切換為 Visual Line Mode
 - `ctrl + v` 切換為 Visual Block Mode
- 按 `:` 或 `/` 或 `?` 進入 Command-Line Mode

以上按 `Esc` 鍵可切換回 Normal Mode

Insert Mode 與 Replace Mode 可用 `insert` 鍵互相切換

三種 Visual Mode 可直接互相切換

Vim : 使用

於命令列輸入

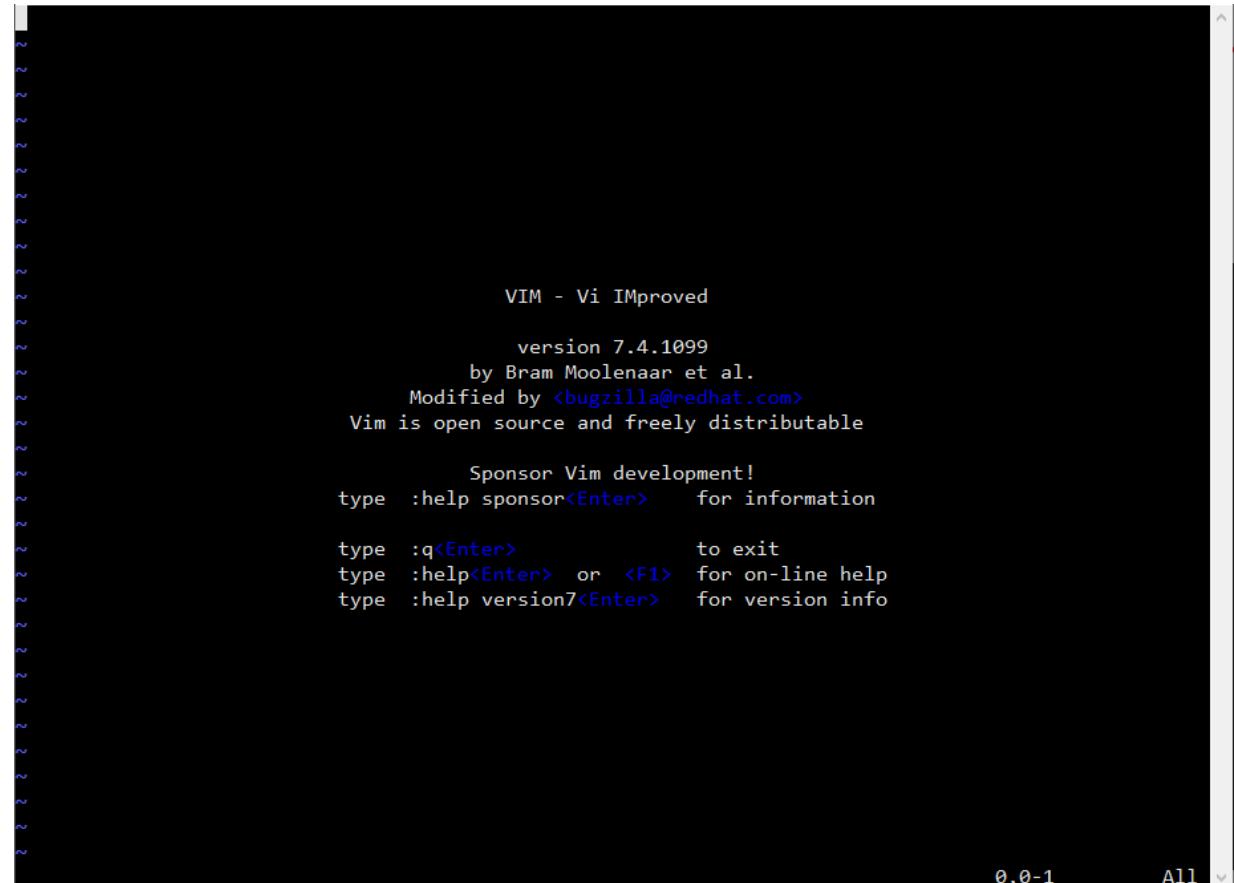
\$ vim

或

\$ vim file

啟動 Vim

若是 file 不存在，存檔時
會寫入新檔案



Vim : 練習

- 進入 Insert Mode , 輸入一些文字
- 按 Esc 鍵返回 Normal Mode
- 輸入 :w 寫入原檔或 :w path 寫入位於 path 的檔案
- 輸入 :q 離開 (有更動未存檔時會跳出警告 , 此時輸入 :q! 可不存檔離開)
- 可使用 :wq 同時寫入並離開
- 用 cat 檢查輸入的內容是否有被正確寫入檔案

Vim : 複製與刪除

- 複製(y)及刪除(d)的用法
 - 重複兩次/大寫以複製/刪除一行 ex : yy ` D
 - 數字+重複兩次/一次+數字+一次以複製/刪除數行 ex : 5yy ` d5d
 - 一次+物件簡稱 (Operator Pending Mappings)
 - w : 游標到下個 word 為止
 - aw : 游標所在 word
 - as : 游標所在 sentence
 - ap : 游標所在 paragraph
 - t+字元 : 到下個該字元為止, 不含該字元, 不算所在字元 ex : yt, ` dte
 - 一次+數字+物件簡稱 : 重複上述動作數次 ex : y5w 複製到下五個 word 為止
- 複製/刪除的內容會被放入剪貼簿

Vim : 快捷鍵

- Normal Mode 常用快捷鍵
 - p : 剪貼簿中為整行時，貼在下一行，否則貼在游標後一個字元處(Visual Mode 時取代選取的全部字元)
 - delete : 刪除一個字元並放在剪貼簿
 - u : 回復上一步
 - **ctrl+r** : 重作剛剛回復的那一步
 - . : 重複剛剛的動作(ex : 刪除行、貼上、輸入字串等)
 - gg : 跳到第一行
 - G : 跳到最後一行
 - % : 跳到對應的括號處
 - >> : 增加縮排(可用 >+數字+>)
 - << : 減少縮排(可用 <+數字+<)
 - 數字+Enter : 往下跳數行
 - gt : 切換到下個分頁
 - gT : 切換到上個分頁
 - 0/Home : 跳到行首
 - End : 跳到行尾
 - pgup/pgdn : 上下翻頁

Vim : Command-Line 常用命令

- :命令：
 - w/write：寫入檔案
 - e/edit：編輯檔案
 - q/quit：關閉這個分頁(只有一個分頁時會離開 Vim)
 - s/target/replace：搜尋並取代
 - tabe/tabedit：開新分頁編輯
 - n/new：開新檔案
 - help：查看說明
- /：向下搜尋字串，直接按 Enter 搜尋上一次搜尋的字串
- ?：向上搜尋字串，直接按 Enter 搜尋上一次搜尋的字串
- 可用上下鍵搜尋歷史命令

Nano

- 許多 Linux 發行版中內建的編輯器
- 被認為是文字模式的編輯器中較容易使用的
- 部分軟體如 git 會在需要啟動編輯器時，預設啟動 nano
(大部分可以在設定檔或用命令列更改)

Nano : 使用

於命令列輸入

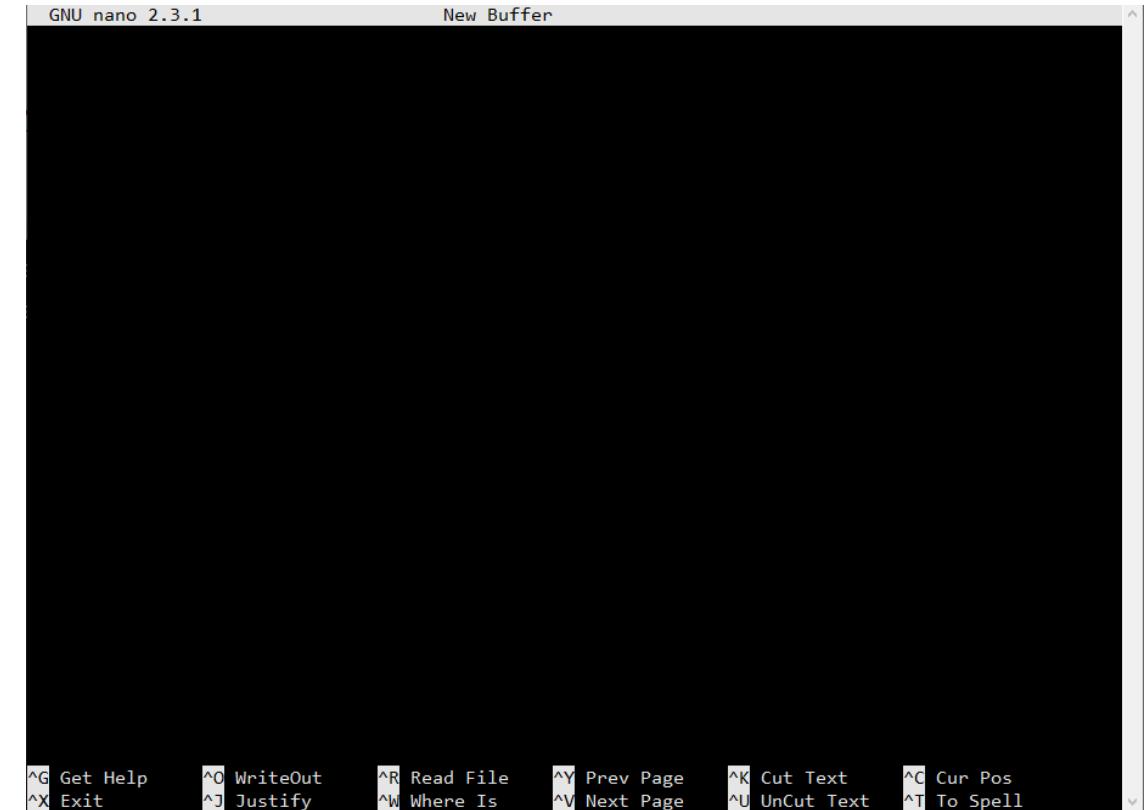
\$ nano

或

\$ nano file

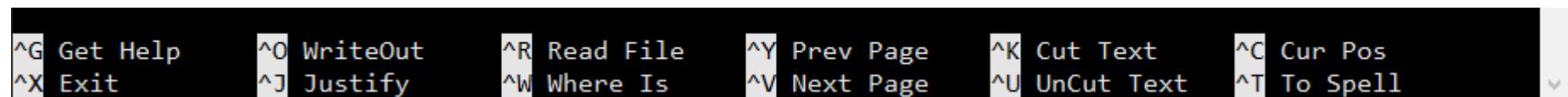
啟動 Vim

若是 file 不存在，存檔時
會寫入新檔案



Nano : 使用

- 可以直接輸入文字，大部分使用如記事本等常用編輯器
- ctrl + o : 寫入檔案
- ctrl + x : 離開 nano
- 其他指令可參考下方說明，其中 ^ 表示 ctrl，後面字母不分大小寫



Nano : 練習

- 使用 \$ nano 進入 Nano，輸入一些文字
- 輸入 ctrl + o 存檔
- 根據提示輸入欲儲存的檔名
- 輸入 ctrl + x 離開
- 使用 cat 指令確認輸入的內容是否有正確儲存

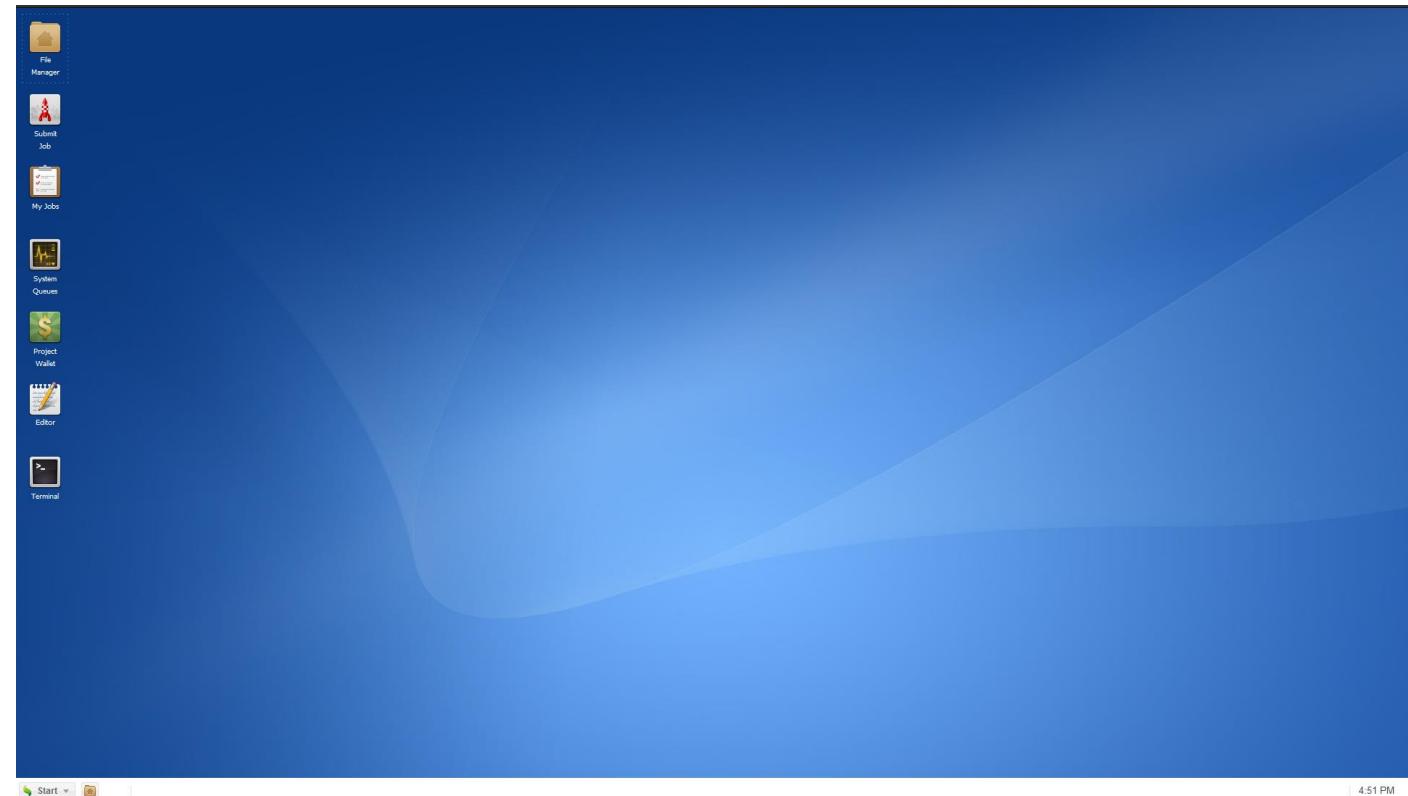
simPlatform

- 為了降低一般使用者使用大主機系統的門檻，
simPlatform 提供友善的圖形操作介面，使不熟悉指令操作的使用者也能簡單使用
- 只要透過瀏覽器就能登入，
不須額外下載工具



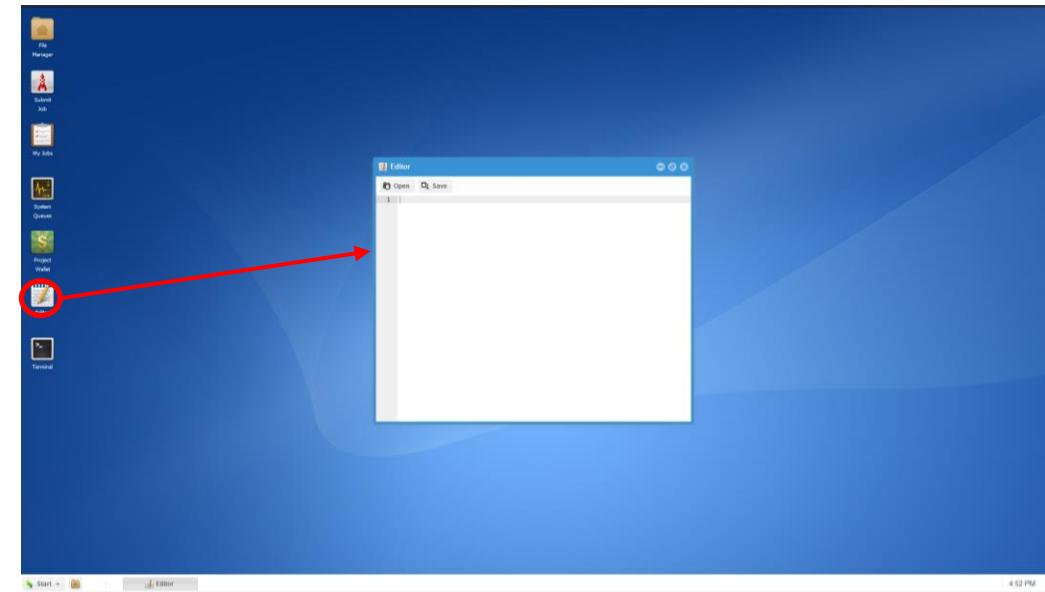
simPlatform

- 以瀏覽器打開網址
<https://simx.nchc.org.tw/Taiwania/>
- 於 Username 輸入主機帳號、
Password+OTP輸入密碼與OTP
- 選擇主機
- 登入
- 等待出現如右圖的畫面



simPlatform : Editor

- 點選 Editor 圖式，即可開啟編輯器
- 點選左上方 Open 可打開舊檔、Save 可儲存檔案
- 也可用 File Manager 點選欲編輯的文字檔案直接打開 Editor





自由練習時間



小測驗