

MACHINE LEARNING TA SESSION

Yu-Hsiu Chen, Chi-Long Tsai, Yu-Min Huang,
Yi-Chung Chiu, Po-Chien Hsu

National Chiao Tung University
Sep. 21, 2018

Outline

- Server Usage
- Python Tutorial
- Library Introduction
- Regression for Homework 1

資訊中心教學用GPU運算伺服器服務

伺服器環境

- 運算資源共有18張NVIDIA Tesla P100 16G, 採用container方式共享GPU
- 目前可提供的container環境為 (ubuntu 16.04, CUDA 9, tensorflow 1.4.1, python 2.7/3/3.5)
- 使用者透過openvpn連入內部網路後, 不經docker host, ssh直連container使用
- 每個container獨立一個private ip, 會創建一個學號的帳號, 以提供個別學生使用
- 本服務具有程式資源上限控制, 若同一個process執行過久(暫定1小時)或佔用GPU RAM過大(暫定1GB)會自動kill掉, 以確保資源共享, 並避免成為虛擬幣挖礦用途

環境建置

- 請先至 <https://openvpn.net/index.php/open-source/downloads.html> 下載安裝好 **openvpn**



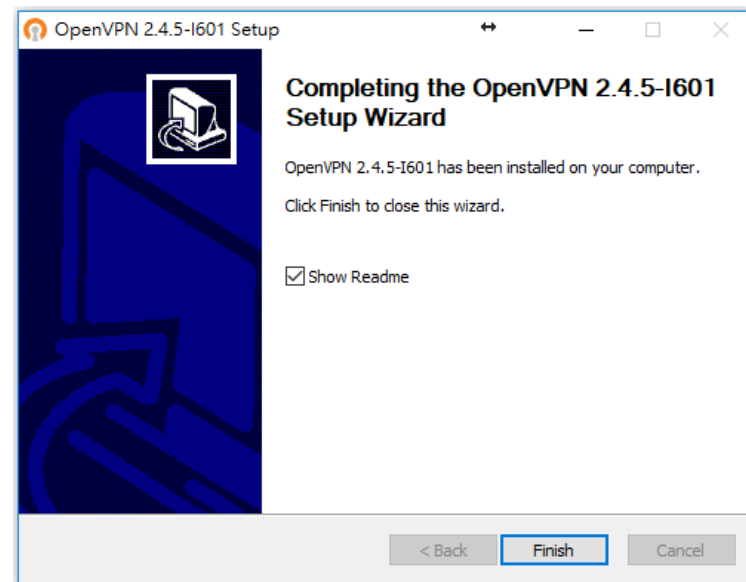
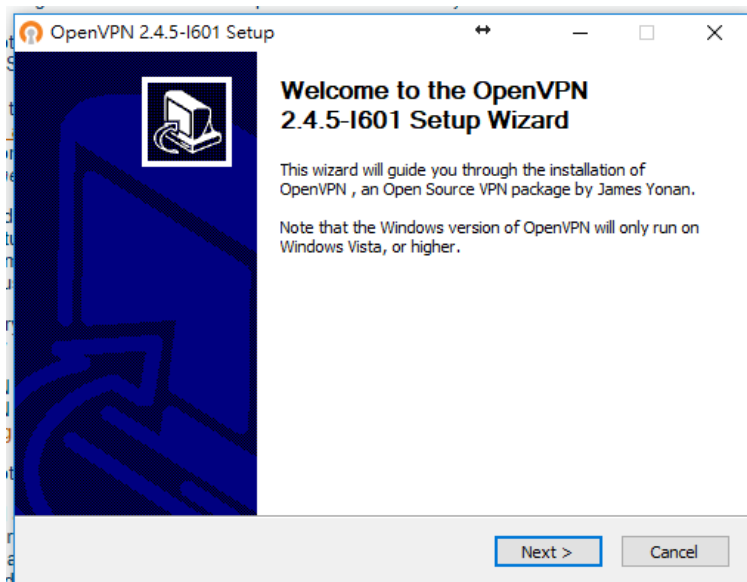
The screenshot shows the OpenVPN website's 'Downloads' section. The page has a navigation bar with links for Home, VPN Service, VPN Solution, Community, and Downloads. A left sidebar contains a table of contents with links to Overview, Downloads, Source Code, Documentation, HOWTO, Security Overview, Examples, Graphical User Interface, Manuals, Change Log, Installation Notes, Release Notes, Miscellaneous, Non-English, File Signatures, Articles, and FAQ. The main content area is titled 'Downloads' and features a section for 'OpenVPN 2.4.5 -- released on 2018.03.01 (Change Log)'. This section contains several paragraphs of text detailing the release, including notes about OpenSSL, LibreSSL, and Windows installers. Below the text is a table of download links for various formats, with the Windows installer link highlighted by a red box. At the bottom, there is a note about GPG signatures and a link to the official documentation.

Source Tarball (gzip)	Source Tarball (xz)	Source Zip	Installer, Windows Vista and later
openvpn-2.4.5.tar.gz	openvpn-2.4.5.tar.xz	openvpn-2.4.5.zip	openvpn-install-2.4.5-i601.exe
GnuPG Signature	GnuPG Signature	GnuPG Signature	GnuPG Signature

NOTE: the GPG key used to sign the release files has been changed since OpenVPN 2.4.0. Instructions for verifying the signatures, as well as the new GPG public key are available [here](#).

We also provide static URLs pointing to latest releases to ease automation. For a list of files look [here](#).

This release is also available in our non-oss source repository for Debian and Ubuntu: [Supported architectures are i386 and](#)

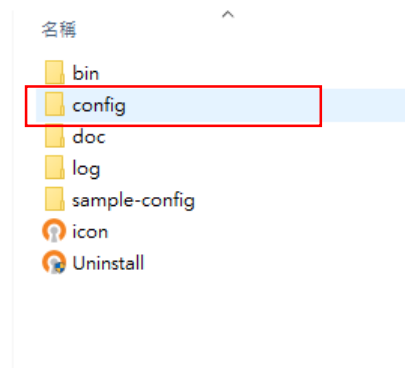
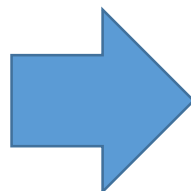
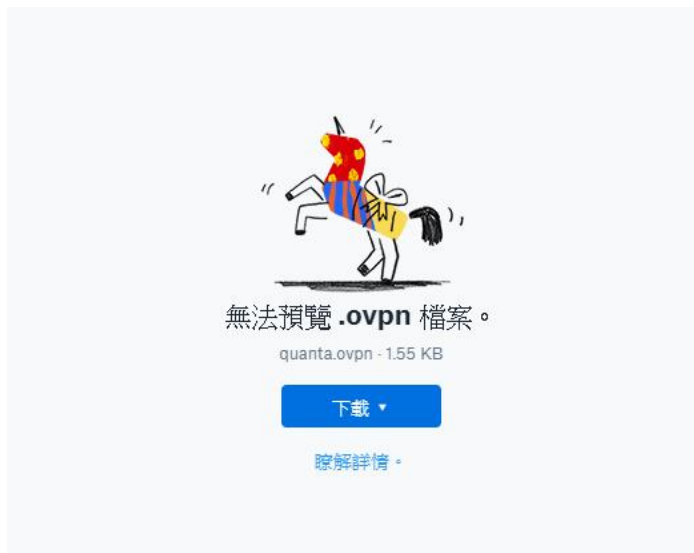


- 匯入以下連結的設定檔

<https://www.dropbox.com/s/h0cg2oxoa54rwyc/quanta.ovpn?dl=0>

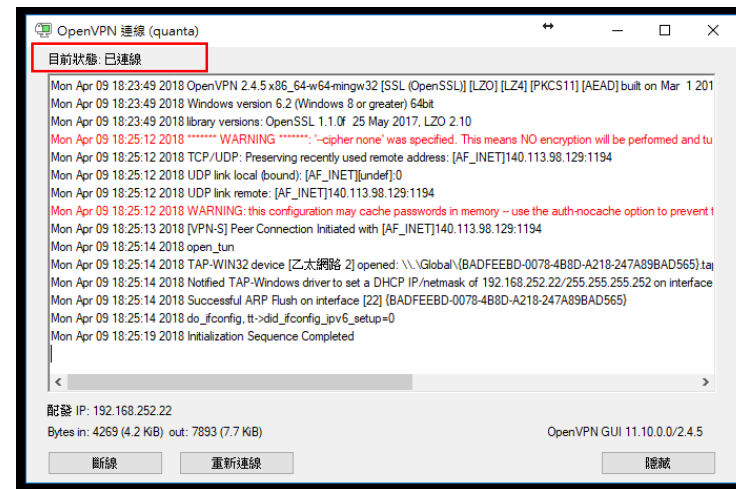
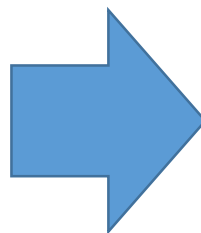
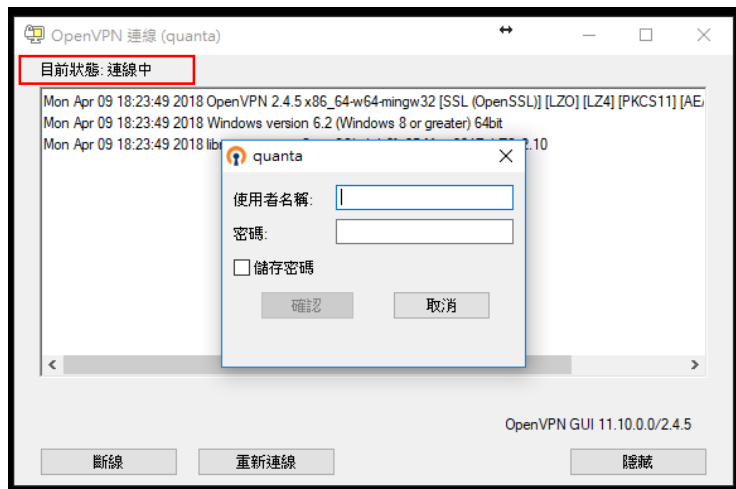
至openvpn的config目錄，

- 以win平台來說會在 %USERPROFILE%\OpenVPN\config

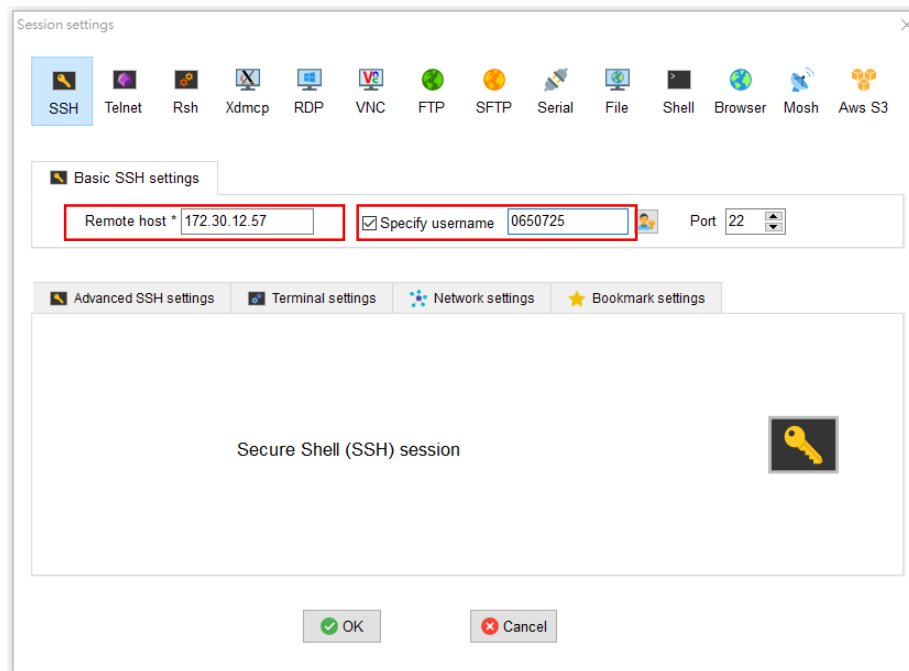


- 之後以 **管理權限** 啟動 openvpn 以進行 **連線**
- 帳號為 **class5084**, 密碼為 **ECM9042**

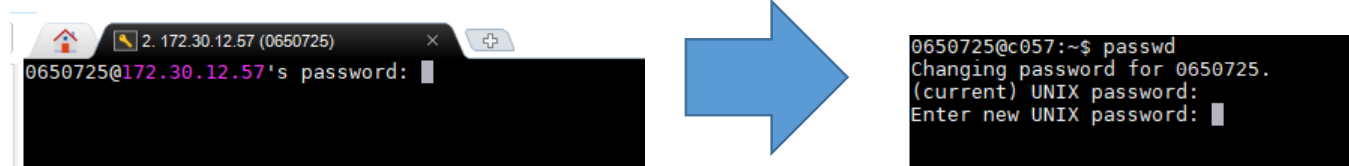




- openvpn連上後, 即可以ssh client連至所屬的主機, IP為172.30.12.XXX, 帳號為學號, 密碼為預設密碼欄位內容
- ssh client 推薦: PuTTY , MobaXterm



- 請用預設密碼連線後，執行passwd這個指令來變更使用者密碼



常見問題

- Error 1 : *“ImportError: No module named _tkinter, please install the python-tk package”*

Solution : `sudo apt-get install python-tk`

- Error 2 : *“Temporary failure resolving ...”*

Solution : `echo "nameserver 8.8.8.8" | sudo tee /etc/resolv.conf > /dev/null`

- Error 3 : *“_tkinter.TclError: no display name and no \$DISPLAY environment variable”*

Solution : Add this code to the start of your **script** (before importing pyplot)

*“import matplotlib
matplotlib.use('Agg') ”*

Python Tutorial

- Tutorial from Google

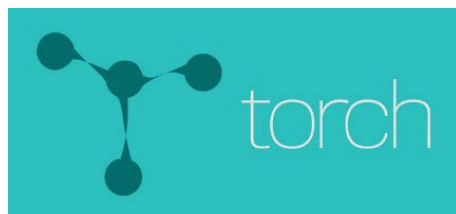
<https://developers.google.com/edu/python/introduction?hl=zh-CN>

- Colab

<https://colab.research.google.com/>

Deep Learning Library

theano



Theano

- Python framework developed by the MILA Lab run by Yoshua Bengio at the University of Montreal
- Computational graph
- There is a [long list of libraries built on top of Theano](#)
- Pros
 - Computational graph is nice abstraction
 - High level wrappers ease the pain
- Cons
 - Error messages can be unhelpful
 - Large models can have long compile times

theano

Tensorflow

- Google created TensorFlow to replace Theano
- TensorFlow runs dramatically slower than other frameworks such as CNTK and MxNet
- Pros
 - Computational graph abstraction, like Theano
 - Error messages can be helpful
 - supports distributed computations
- Cons
 - Not all operations in Tensorflow work as they do in Numpy
 - Slower than other frameworks
 - Not very toolable



MXnet

- MxNet involves Pedro Domingos and a team of researchers at the University of Washington
- Mxnet has been [adopted by Amazon Web Services](#)
- Pros
 - Less memory to use
 - Variety of interfaces, ex: Javascript, Python, R, Matlab, C++
 - Computational graph abstraction
 - supports distributed computations
- Cons
 - Upgrade slowly
 - Hard for finetuning existing networks



Torch & Pytorch

- [Torch](#) is a computational framework with an API written in Lua that supports machine-learning algorithm
- A Python API for Torch, known as [Pytorch](#), was open-sourced by Facebook in January 2017
- Some version of it is used by large tech companies such as Facebook and Twitter
- Pros
 - Lots of modular pieces that are easy to combine
 - Flexibility and fast
 - Faster compile times
- Cons
 - Spotty documentation



Caffe

- [Caffe](#) is a well-known and widely used **machine-vision library** that ported Matlab's implementation of fast convolutional nets to C and C++
- [Yangqing Jia](#) created the project during his PhD at UC Berkeley
- Pros
 - Good for feedforward networks and image processing
 - Good for finetuning existing networks
- Cons
 - Need to write C++ / CUDA for new GPU layers
 - Not good for recurrent networks
 - Not extensible, bit of a hairball



Library Comparision

Design Choice	Torch.nn	Theano-based	Caffe	autograd (NumPy, Torch)	Chainer	MXNet	Tensor-Flow
1.NN definition	Script (Lua)	Script* (Python)	Data (protobuf)	Script (Python, Lua)	Script (Python)	Script (many)	Script (Python)
2. Graph construction	Prebuild	Prebuild	Prebuild	Dynamic	Dynamic	Prebuild**	Prebuild
3. Backprop	Through graph	Extended graph	Through graph	Extended graph	Through graph	Through graph	Extended graph
4. Parameters	Hidden in operators	Separate nodes	Hidden in operators	Separate nodes	Separate nodes	Separate nodes	Separate nodes
5. Update formula	Outside of graphs	Part of graphs	Outside of graphs	Outside of graphs	Outside of graphs	Outside of graphs**	Part of graphs
6. Optimization	-	Advanced optimization	-	-	-	-	Simple optimization
57 Parallel computation	Multi GPU	Multi GPU (libgpuarray)	Multi GPU	Multi GPU (Torch)	Multi GPU	Multi node Multi GPU	Multi node Multi GPU