DASC 7606 Deep Learning

Tutorial 1

Tianshuo Yang

yangtianshuo@hku.hk

This Tutorial includes:

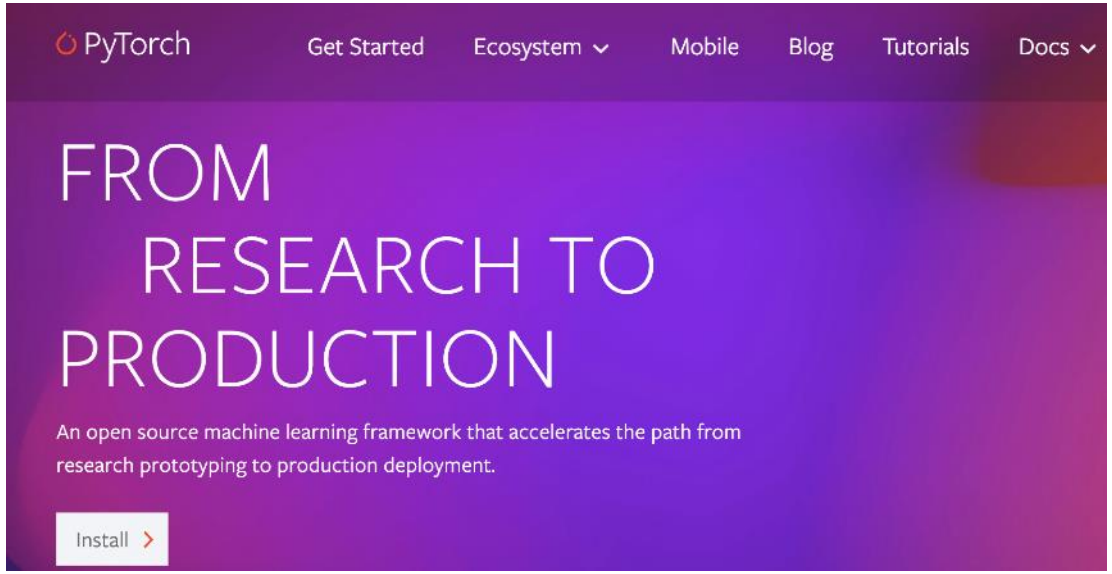
- Python and PyTorch resources for you to self-learn
- HKU GPU Farm
  - Apply for accounts
  - Accessing HKU GPU Farm using SSH and Basic Linux Commands
  - Connect to HKU GPU Farm Gateway Node and GPU Node
- Exercise using the GPU

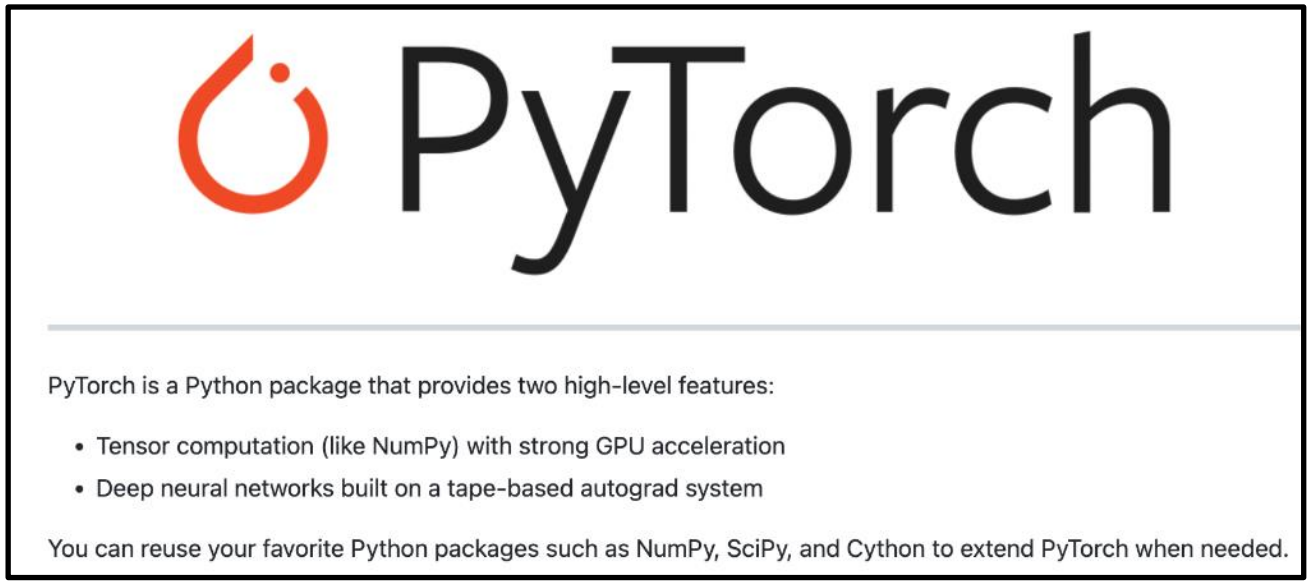
Python

- Install Python from Anaconda:
  - https://www.anaconda.com/products/distribution
- Or install Python directly:
  - https://www.python.org/downloads/
- Python Tutorial:
  - https://docs.python.org/3/tutorial/
  - https://www.w3schools.com/python/default.asp

PyTorch – Python package for deep learning

https://pytorch.org





PyTorch

- PyTorch quickstart tutorial:
  - https://pytorch.org/tutorials/beginner/basics/quickstart_tutorial.html

- PyTorch documentation:
  - https://pytorch.org/docs/stable/index.html

HKU GPU Farm

https://www.cs.hku.hk/gpu-farm/home

THE UNIVERSITY OF HONG KONG
DEPARTMENT OF
COMPUTER SCIENCE

HOME QUICK START POLICIES AND GUIDELINES

HKU CS GPU Farm

Department of Computer Science of HKU currently has two GPU farms.

Phase 1

The first GPU farm is a computer cluster with 84 NVIDIA GeForce GTX 1080Ti and 16 RTX 3090 GPU cards installed in 27 compute nodes. This computer cluster has 6 file servers providing 576 TB usable disk storage in total.

Phase 2

The second GPU farm has similar configurations to the first one with 100 NIVIDA GeForce RTX 2080Ti GPU cards in 25 compute nodes. This computer cluster also has 6 file servers providing another 576 TB usable disk storage.

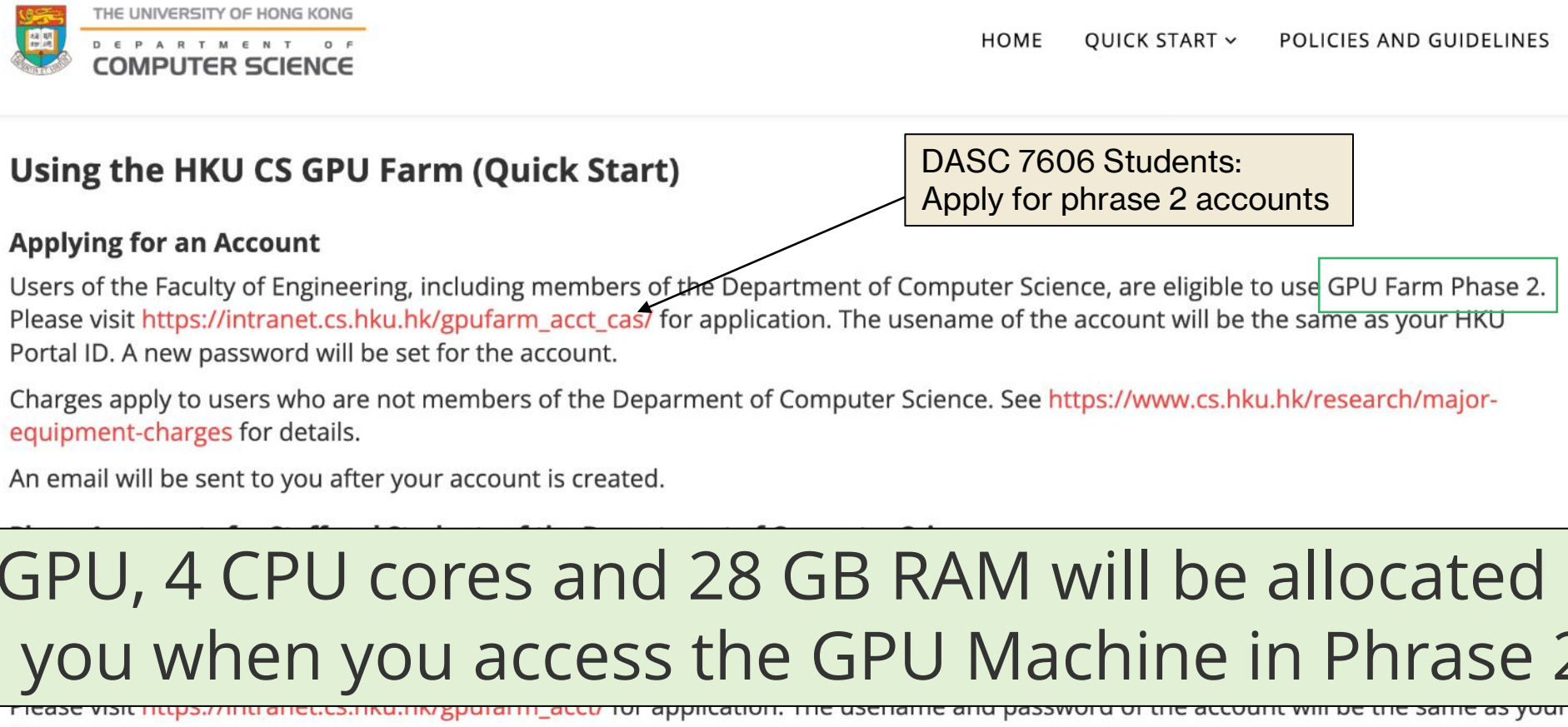
Both computer clusters are managed by the SLURM cluster management and job scheduling system. Compute- or data-intensive software, such as TensorFlow and PyTorch for deep learning, can be run on this platform.

- Charges and Payment (for non-CS department users)
- Applying for an Account and Using the HKU CS GPU Farm
- GPU Farm Usage Policies and Guidelines

If you have any feedback or query please click here.

HKU GPU Farm

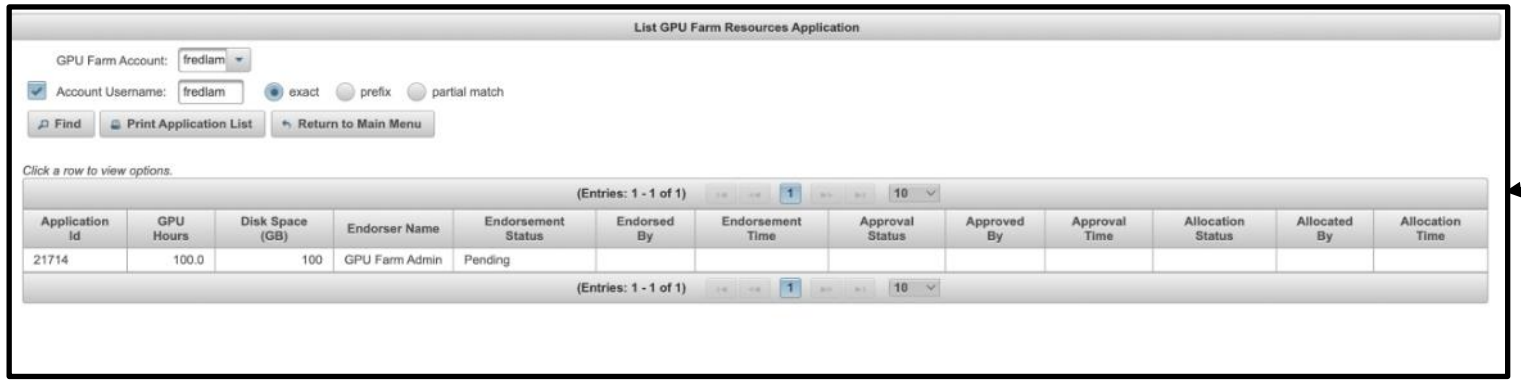
https://www.cs.hku.hk/gpu-farm/quickstart



HKU GPU Farm – Apply for phrase 2 accounts

https://intranet.cs.hku.hk/gpufarm_acct_cas

- Contact support@cs.hku.hk if you have any issue regarding the application for the account of HKU GPU Farm

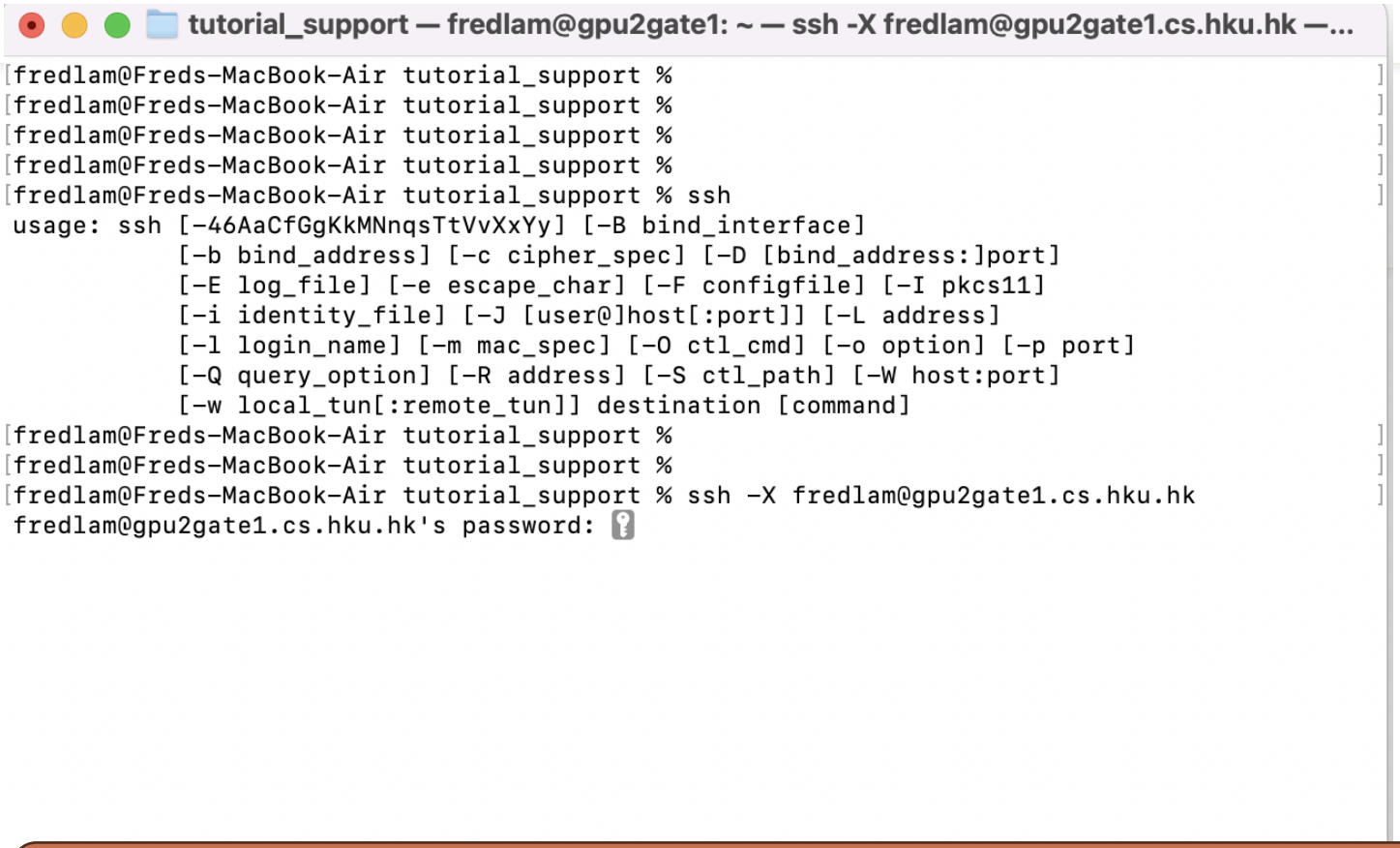


After applying for account, waiting for approval

Account allocated; you can now connect to HKU GPU Farm using SSH

Accessing GPU Farm using SSH

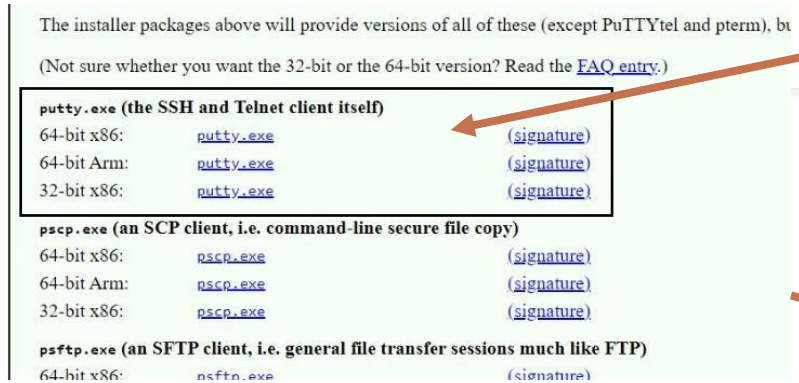
– Mac: Use Terminal *ssh -X <your_portal_id>@gpu2gate1.cs.hku.hk*

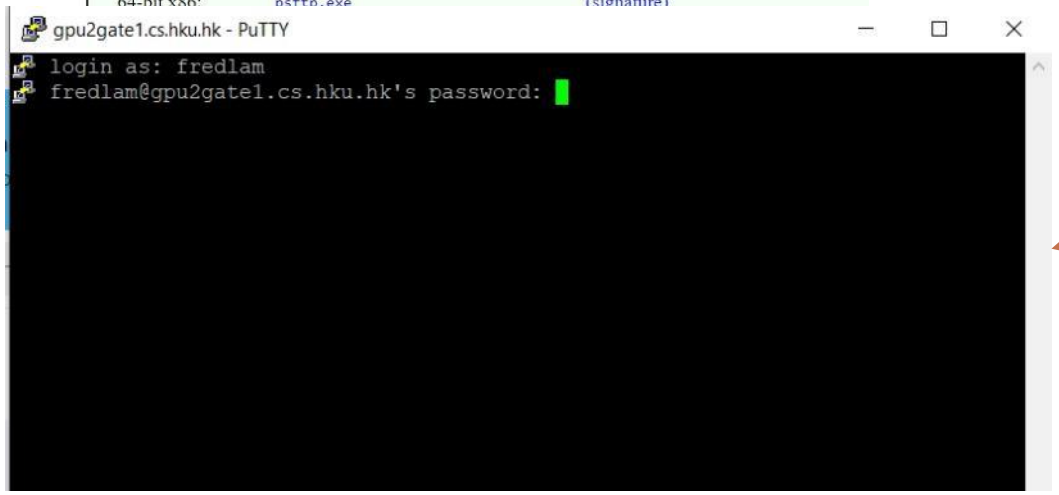


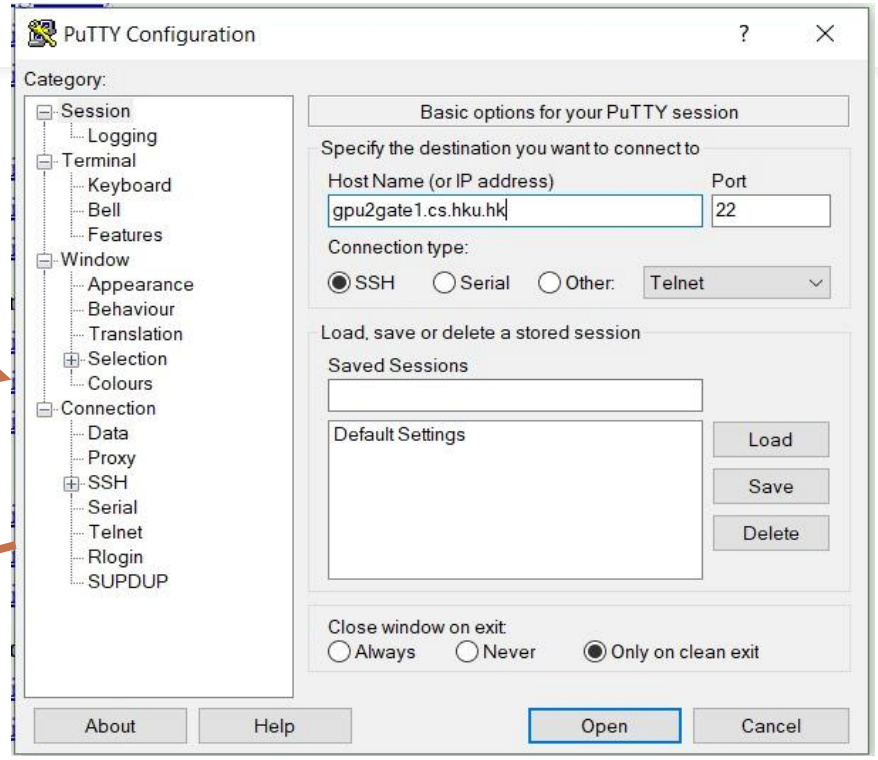
To access the GPU farm, you need to be connected to the HKU network (e.g., when you are using the wired network in CS laboratories and offices, or connected to HKU Wifi or HKUVPN). Use SSH to connect to one of the *gateway nodes*:

Accessing GPU Farm using SSH

– Windows: Use putty https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html







To access the GPU farm, you need to be connected to the HKU network (e.g., when you are using the wired network in CS laboratories and offices, or connected to HKU Wifi or HKUVPN). Use SSH to connect to one of the *gateway nodes*:

Useful Linux Commands

Linux Command	Purpose	
ls	display the contents of a directory	https://docs.rockylinux.org/books/admin_guide/03-commands/#ls-command
pwd	display the absolute path of the current directory	https://docs.rockylinux.org/books/admin_guide/03-commands/#pwd-command
cd	change the current directory	https://docs.rockylinux.org/books/admin_guide/03-commands/#cd-command
mkdir	create a directory or directory tree	https://docs.rockylinux.org/books/admin_guide/03-commands/#mkdir-command
mv	moves and renames a file	https://docs.rockylinux.org/books/admin_guide/03-commands/#mv-command
rm	delete a file or directory	https://docs.rockylinux.org/books/admin_guide/03-commands/#rm-command
cp	copy a file or directory	https://docs.rockylinux.org/books/admin_guide/03-commands/#cp-command
bash	execute commands read from the standard input or from a file	https://www.tutorialspoint.com/unix_commands/bash.htm
wget	download of files from the Web	https://www.tutorialspoint.com/unix_commands/wget.htm

Ubuntu Linux commands tutorial:
https://ubuntu.com/tutorials/command-line-for-beginners#1-overview

Accessing HKU GPU Farm

```
fredlam@Freds-MacBook-Air tutorial_support %
fredlam@Freds-MacBook-Air tutorial_support % ssh -X fredlam@gpu2gate1.cs.hku.hk
fredlam@gpu2gate1.cs.hku.hk's password:
Welcome to Ubuntu 18.04.6 LTS (GNU/Linux 4.15.0-175-generic x86_64)

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

```
(base) fredlam@gpu2gate1:~$
(base) fredlam@gpu2gate1:~$ gpu-interactive
(Text mode only session.)
(base) fredlam@gpu2-comp-102:~$
(base) fredlam@gpu2-comp-102:~$ nvidia-smi
Sat Jan 28 16:52:21 2023
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 460.91.03    Driver Version: 460.91.03    CUDA Version: 11.2     |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|                               |                      |               MIG M. |
|===============================+======================+======================|
|   0  GeForce RTX 208...  On   | 00000000:05:00.0 Off |                  N/A |
| 18%   40C    P8    40W / 250W |      1MiB / 11019MiB |      0%      Default |
|                               |                      |                  N/A |
+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Processes:                                                                  |
|  GPU   GI   CI        PID   Type   Process name                  GPU Memory |
|        ID   ID                                                   Usage      |
|=============================================================================|
|  No running processes found                                                 |
+-----------------------------------------------------------------------------+
(base) fredlam@gpu2-comp-102:~$
```

Your Local Machine → HKU CS Gateway Node gpu2gate1.cs.hku.hk → HKU CS GPU Node gpu2-comp-xxx

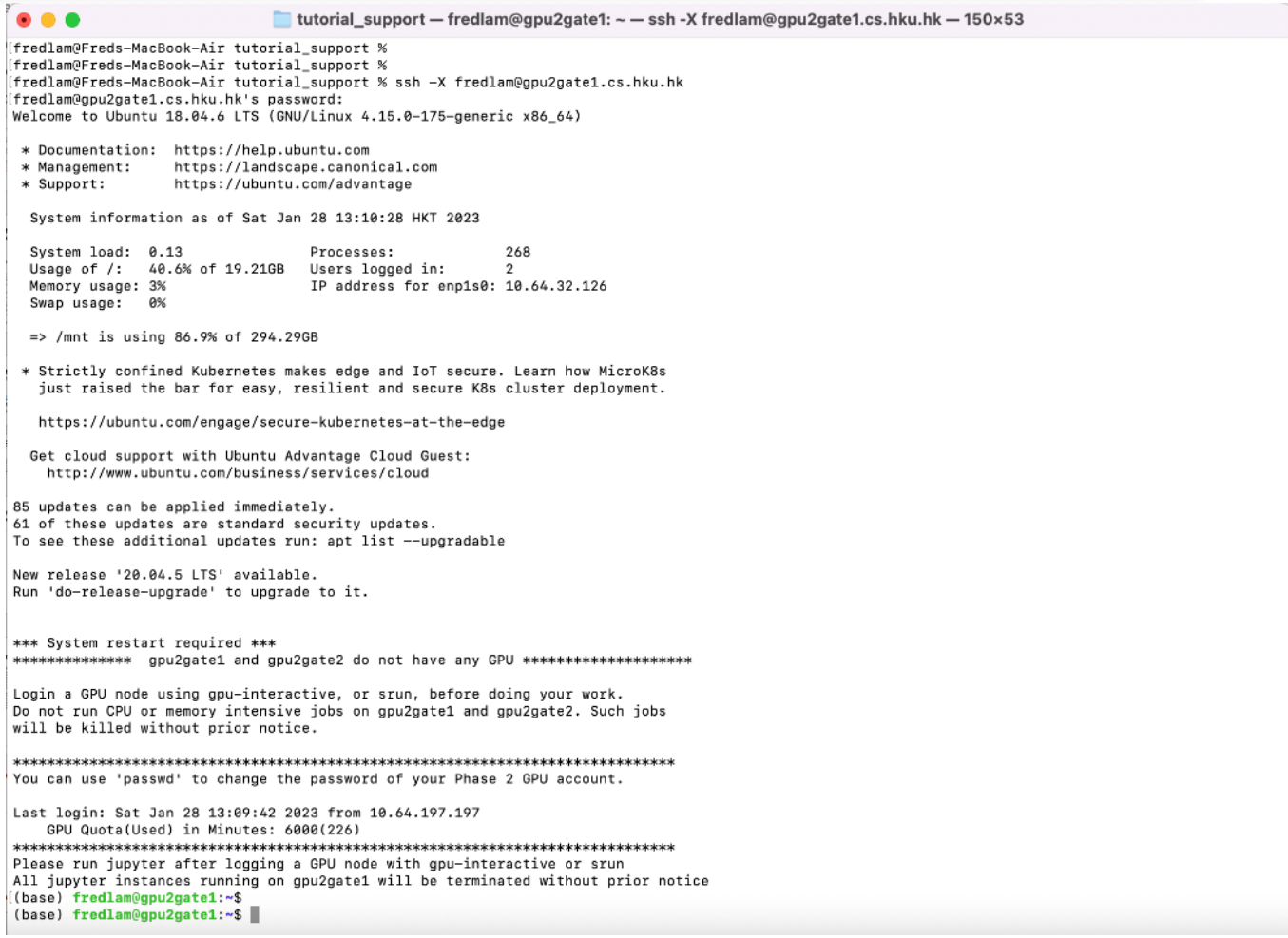
- All your processes running on the GPU node will be terminated when you exit from gpu-interactive command.
- You might be allocated another GPU node when you type ‘gpu-interactive’ again, but your files in the disk storage would remain the same.

```
(base) fredlam@gpu2-comp-102:~$
(base) fredlam@gpu2-comp-102:~$ hostname -I
10.64.32.112
(base) fredlam@gpu2-comp-102:~$ exit
exit
(base) fredlam@gpu2gate1:~$
(base) fredlam@gpu2gate1:~$
(base) fredlam@gpu2gate1:~$ hostname -I
10.64.32.126
(base) fredlam@gpu2gate1:~$ gpu-interactive
(Text mode only session.)
(base) fredlam@gpu2-comp-100:~$
(base) fredlam@gpu2-comp-100:~$
(base) fredlam@gpu2-comp-100:~$ hostname -I
10.64.32.110
(base) fredlam@gpu2-comp-100:~$
```

Accessing GPU Farm

SSH Connect to HKU GPU Farm Gateway node

- Connect to Gateway nodes for GPU Farm Phase 2:
  - ssh -X <your_portal_id>@gpu2gate1.cs.hku.hk
  - If my account is fredlam, then my command is:
  - ssh -X fredlam@gpu2gate1.cs.hku.hk
  - Then enter your password



Accessing GPU Farm

SSH Connect to Gateway node -> then connect to GPU node

Ref: https://www.cs.hku.hk/gpu-farm/quickstart#:~:text=Using%20GPUs%20in%20Interactive%20Mode

Using GPUs in Interactive Mode

After logging on a gateway node, you can now login to a server node with actual GPUs attached. To have an interactive session, use the gpu-interactive command to run a bash shell on a GPU node. An available GPU compute node will be selected and allocated to you, and you will be logged on the node automatically. The GPU compute nodes are named gpu-comp-x for Phase 1 and gpu2-comp-x for Phase 2. Note the change of host name in the command prompt when you actually log on to a GPU node, e.g.,

```
tmchan@gpugate1:~$ gpu-interactive
tmchan@gpu-comp-1:~$
```

You can verify that a GPU is allocated to you with the nvidia-smi command, e.g.:

```
tmchan@gpu-comp-1:~$ nvidia-smi

Sat Feb 16 17:22:06 2019
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 410.79       Driver Version: 410.79       CUDA Version: 10.0     |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|===============================+======================+======================|
|   0  GeForce GTX 108...  Off  | 00000000:06:00.0 Off |                  N/A |
| 28%   38C    P0    56W / 250W |      0MiB / 11178MiB |      0%      Default |
+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Processes:                                                       GPU Memory |
|  GPU       PID   Type   Process name                             Usage      |
|=============================================================================|
|  No running processes found                                                 |
+-----------------------------------------------------------------------------+
```

With the gpu-interactive command, 1 GPU, 4 CPU cores and 28 GB RAM are allocate to you.

You can now install and run software like using a normal Linux server.

Accessing the HKU GPU Farm

SSH Connect to Gateway node -> then connect to GPU node

With the gpu-interactive command, 1 GPU, 4 CPU cores and 28 GB RAM are allocate to you.

You can now install and run software like using a normal Linux server.

Note that you do not have sudo privileges. Do not use commands such as 'sudo pip' or 'sudo apt-get' to install software.

The time limits (quotas) for GPU and CPU time start counting once you have logged on to a GPU compute node, until you logout the GPU compute node:

```
tmchan@gpu-comp-1:~$ exit
tmchan@gpugate1:~$
```

All your processes running on the GPU node will be terminated when you exit from gpu-interactive command.

Accessing Your Session with Another Terminal

After you are allocated a GPU compute node with gpu-interactive, you may access the same node with another SSH session. What you need is the actual IP address of the GPU compute node you are in. Run 'hostname -I' on the GPU compute to node find out its IP address. The output will be an IP address 10.XXX.XXX.XXX, e.g.,

```
tmchan@gpugate1:~$ hostname -I 10.21.5.225
```

Then using another terminal on your local desktop/notebook, SSH to this IP address:

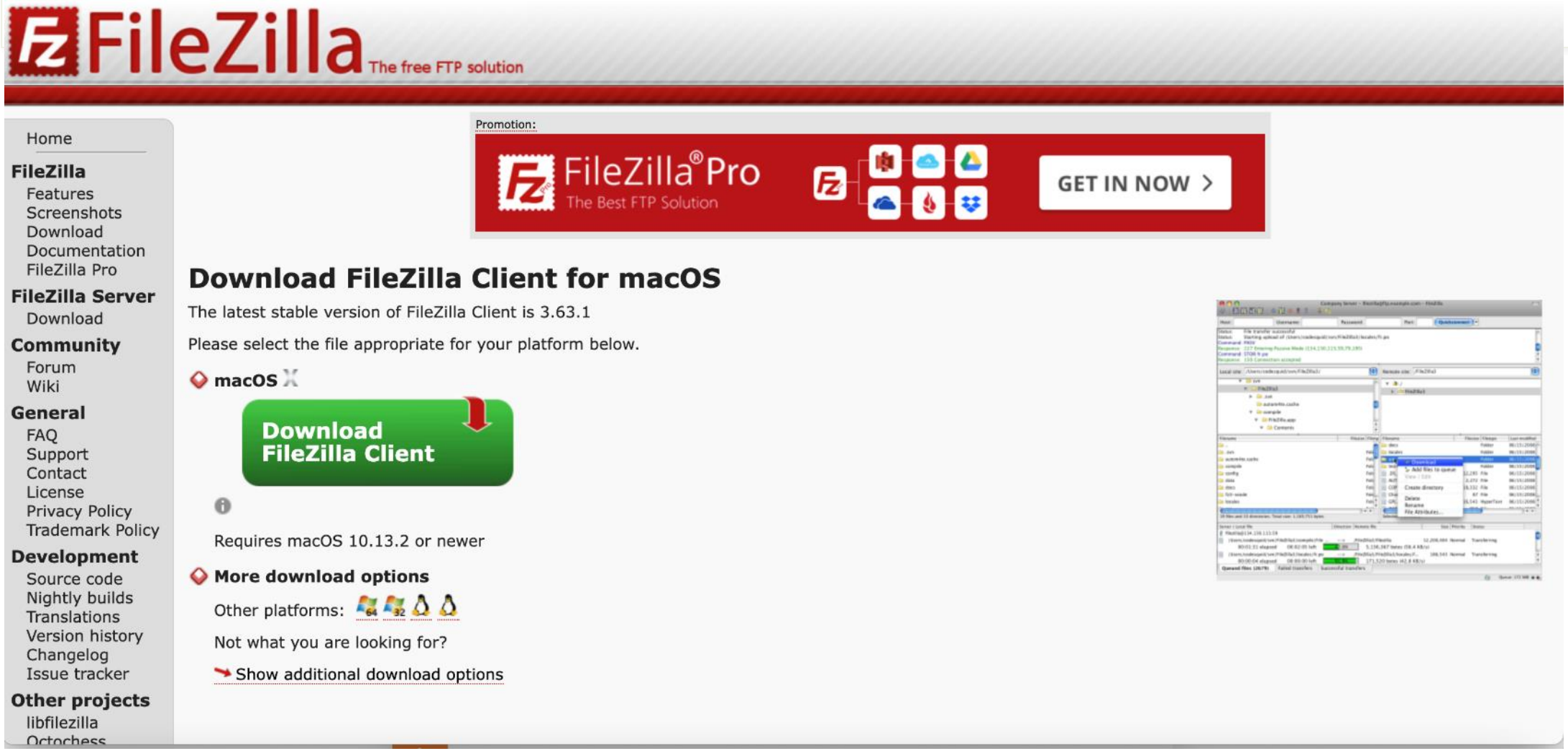
```
ssh -X <your_cs_username>@10.XXX.XXX.XXX
```

These additional SSH sessions will terminate when you exit the gpu-interactive command.

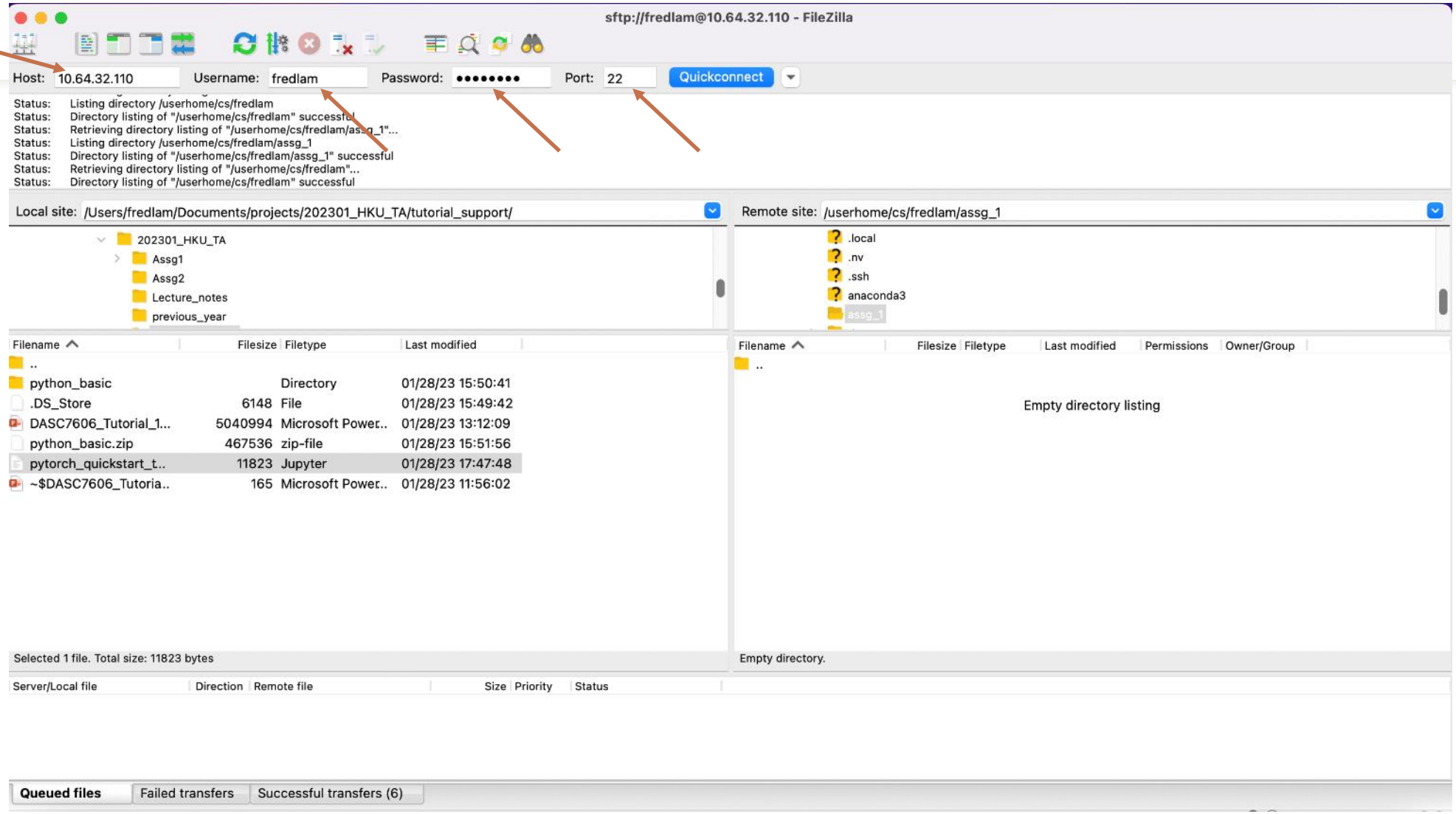
Note: Do not use more than one gpu-interactive (or srun) at the same time if you just want to access your current GPU session from a second terminal, since those commands will start a new independent session and allocate an additional GPU to you, i.e., your GPU time quota will be doubly deducted. Also, you cannot access the GPUs of your previous sessions.

HKU GPU Farm – Use Filzilla to Transfer files to GPU Node

https://filezilla-project.org/download.php



HKU GPU Farm – Use Filzilla to Transfer files to GPU Node



The End