

Storage & Data Transfer Service

dicos-support@twgrid.org

Academia Sinica Grid-computing Centre (ASGC)



Outline

- Storage space
- Introduction of Home Space
- Introduction of Group Working Space
- Subscribe to more Group Working Space
- Unsubscribe Space
- Working Space for Computing Jobs
- Using /tmp on the worker node as Intermediate Working Space
- Examples

Storage space

- Home Space (100GB for free, not extendable)

`/dicos_ui_home/{user_account}`

- Group Working Space (3TB free /each group, pay for extension)

`/ceph/work/{group}`

- [cryoEM group]

`/activeEM/data/{group}/{user}`

- Backup Space (will be available in the future)

- Please **back up to secure your data** in these spaces.

Storage space

- Check your usage on <https://dicos.grid.sinica.edu.tw/profile/info>

The screenshot shows the DICOS User Profile page. The navigation bar includes 'DICOS', 'Resources', 'Policy', 'About', 'Documentation', 'API', 'Apps', 'Contact', 'Live Chat', and 'Your account'. The 'Your account' dropdown menu is open, showing options: 'User Profile', 'Usage Summary', 'Group Member List', 'Subscribe Storage Space', 'Resource Usage', and 'Logout'. The profile information includes: Username: account_name, Group: group_name, First Name: your_first_name, Last Name: your_last_name, E-mail: your email, Last Login: time_joined, and Date Joined. A red box highlights the storage usage section, which shows two entries: '/dicos_ui_home/your_account' with 2.3GB/100GB usage and latest update 2023-11-14 00:50:02, and '/ceph/work/your_group' with 0.4TB/4.0TB usage and latest update 2023-11-14 00:30:03. At the bottom, there are buttons for 'Edit Profile' and 'Change Password'.

If /dicos_ui_home/ storage usage > 100 GB, user can't log in server.

Storage Usage	Usage	Latest Update
/dicos_ui_home/your_account	2.3GB/100GB	2023-11-14 00:50:02
/ceph/work/your_group	0.4TB/4.0TB	2023-11-14 00:30:03

Introduction of Home Space

`/dicos_ui_home/{user_account}`

- 100 GB for free, not extendable.
- Backup your important data in Home Space.
- `/dicos_ui_home/` is a temporary space, backup your data in the `/ceph/work/{group}` or transfer to your own data storage.
- Home space will be cleared once your account is expired and deleted.
- **Avoid using the Home Space as the job's working space.**

Introduction of Group Working Space

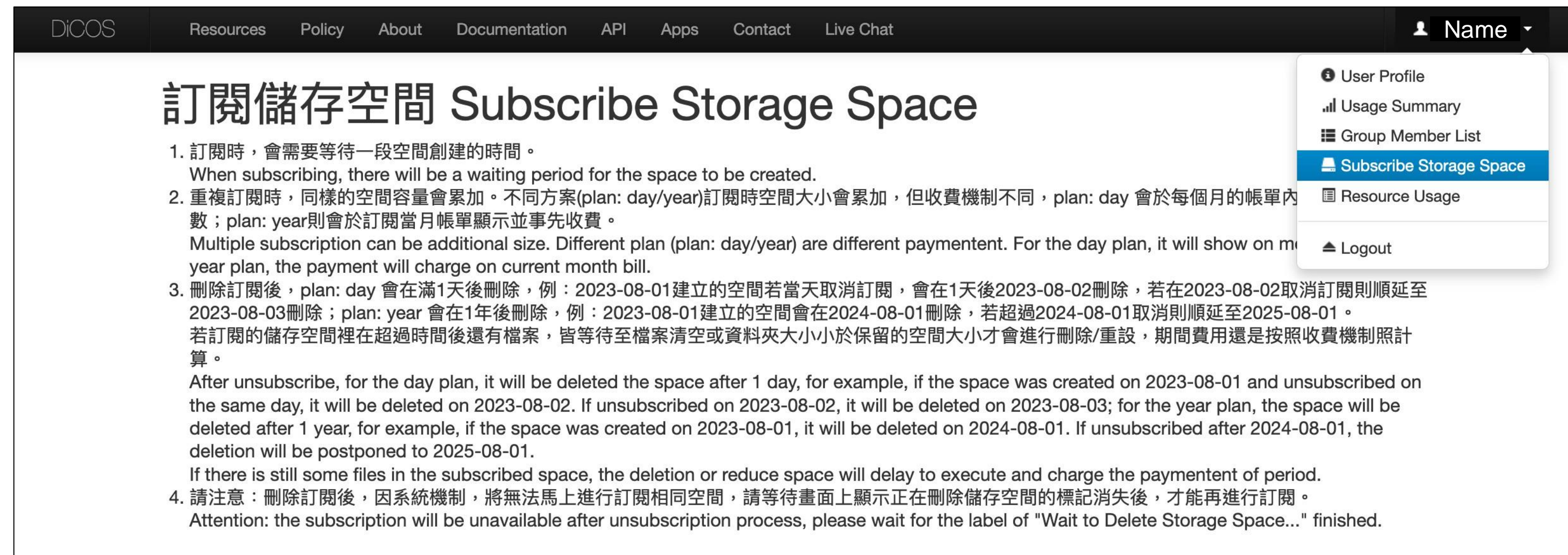
`/ceph/work/{group}`

- 3TB free space for each group.
- This space is shared among all members of the group.
- PI has full access and manage rights in this space.
- PI can subscribe & upgrade for more space via dicos.grid.sinica.edu.tw.
- Backup your important data in Group Working Space.

Subscribe to more Group Working Space

- PI can subscribe for upgrade storage space.

<https://dicos.grid.sinica.edu.tw/accounting/subscribe-storage>



DICOS Resources Policy About Documentation API Apps Contact Live Chat Name

訂閱儲存空間 Subscribe Storage Space

- 訂閱時，會需要等待一段空間創建的時間。
When subscribing, there will be a waiting period for the space to be created.
- 重複訂閱時，同樣的空間容量會累加。不同方案(plan: day/year)訂閱時空間大小會累加，但收費機制不同，plan: day 會於每個月的帳單內數；plan: year則會於訂閱當月帳單顯示並事先收費。
Multiple subscription can be additional size. Different plan (plan: day/year) are different paymentent. For the day plan, it will show on m year plan, the payment will charge on current month bill.
- 刪除訂閱後，plan: day 會在滿1天後刪除，例：2023-08-01建立的空間若當天取消訂閱，會在1天後2023-08-02刪除，若在2023-08-02取消訂閱則順延至2023-08-03刪除；plan: year 會在1年後刪除，例：2023-08-01建立的空間會在2024-08-01刪除，若超過2024-08-01取消則順延至2025-08-01。
若訂閱的儲存空間裡在超過時間後還有檔案，皆等待至檔案清空或資料夾大小小於保留的空間大小才會進行刪除/重設，期間費用還是按照收費機制照計算。
After unsubscribe, for the day plan, it will be deleted the space after 1 day, for example, if the space was created on 2023-08-01 and unsubscribed on the same day, it will be deleted on 2023-08-02. If unsubscribed on 2023-08-02, it will be deleted on 2023-08-03; for the year plan, the space will be deleted after 1 year, for example, if the space was created on 2023-08-01, it will be deleted on 2024-08-01. If unsubscribed after 2024-08-01, the deletion will be postponed to 2025-08-01.
If there is still some files in the subscribed space, the deletion or reduce space will delay to execute and charge the paymentent of period.
- 請注意：刪除訂閱後，因系統機制，將無法馬上進行訂閱相同空間，請等待畫面上顯示正在刪除儲存空間的標記消失後，才能再進行訂閱。
Attention: the subscription will be unavailable after unsubscription process, please wait for the label of "Wait to Delete Storage Space..." finished.

- User Profile
- Usage Summary
- Group Member List
- Subscribe Storage Space**
- Resource Usage
- Logout

1TB Storage: NT\$3/day or NT\$1000/year

More pricing information at <https://dicos.grid.sinica.edu.tw/resources>

Subscribe to more Group Working Space

加購工作短期儲存空間
Upgrade Work Space

1T = NT \$3/1 day
PATH: /ceph/work/[group]/

1 T

Subscribe

加購工作短期儲存空間：1年方案
Upgrade Work Space：1 year plan

1T = NT \$1000/1 year
PATH: /ceph/work/[group]/

1 T

Subscribe

Current Storage Space

工作短期儲存空間 Work Space
GROUP: ASGC
PLAN: day
PATH: /ceph/work/ASGC/
Latest Update: 2023-07-27 00:30:04

0.0/5

1 T Unsubscribe

工作短期儲存空間 Work Space
GROUP: ASGC
PLAN: year
PATH: /ceph/work/ASGC/
Latest Update: 2023-07-27 00:30:04

0.0/5

1 T Unsubscribe

Unsubscribe Space

- Clear your data and release your space before unsubscription.
- Unable to unsubscribe if the space is not released.
→ remove data until less or equal to target subscription.
- Unsubscription process will be blocked when space can not be released or cancel year plan.
- Contact info: dicos-support@twgrid.org



Working Space for Computing Space

- Group Working Space: `/ceph/work/{group}`



- If your job has high I/O wait output:

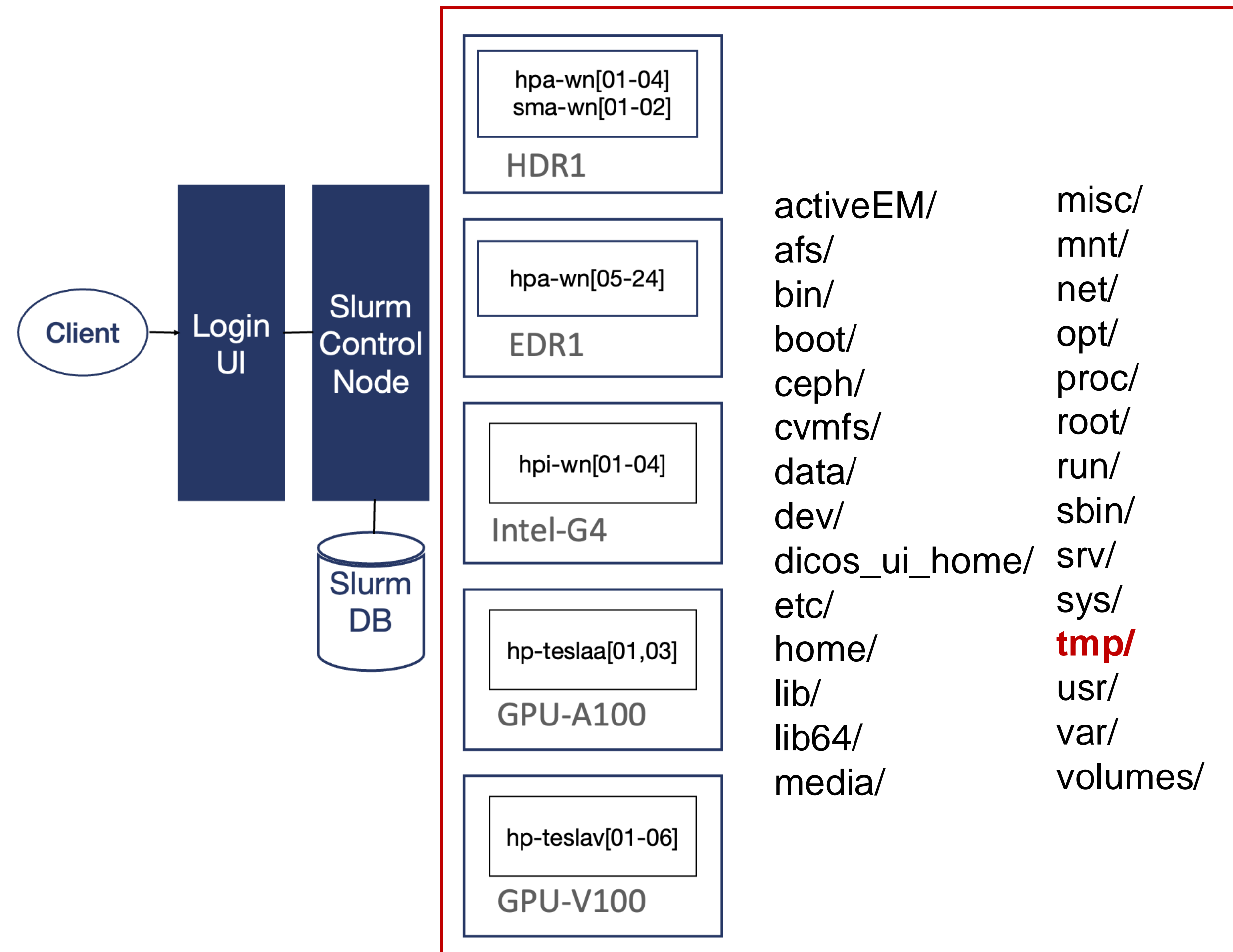
You can also use **/tmp on the worker node** as an intermediate job working space.

Copy intermediate data to `/ceph/work/{group}`.

Using /tmp on the worker node as Intermediate Working Space

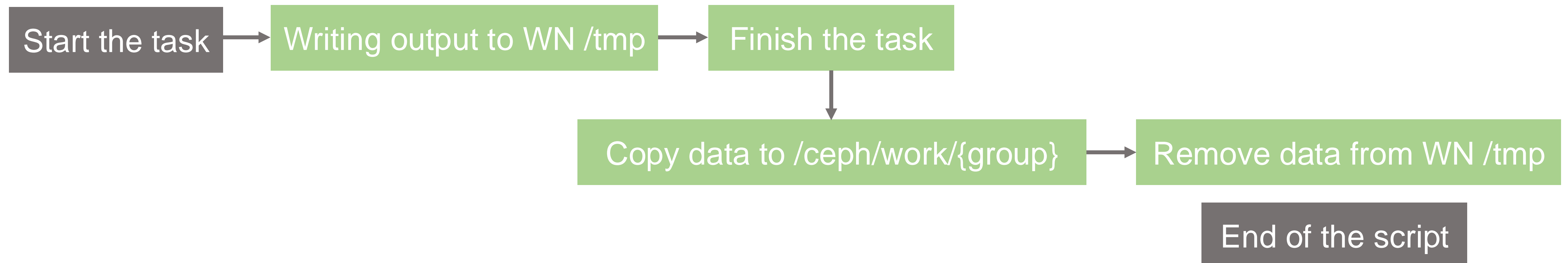
- Each worker node has a different spec (e.g., SSD) & capacity.

Cluster	Local Disk SSD
HDR1	1 TB
EDR1	2 TB
Intel-g4	2 TB



Using /tmp on the worker node as Intermediate Working Space

- Copy intermediate data to /ceph/work/{group}.
- Remove files from **/tmp on the worker node** when tasks finish.
- /tmp directory on the worker node will be cleaned anytime.



Examples

Linux Basic Command

Using /tmp on the worker node as Intermediate Working Space

Data Transfer

Linux Basic Command

`ssh username@slurm-ui.twgrid.org`

Login

`ssh username@dicos-sftp.twgrid.org`

`pwd`

Show current path

`mv {src} {dest}`

Move file

`rm filename`

Remove file

`rm -r dirname`

Remove directory

`cat filename`

View file content

Using /tmp on the worker node as Intermediate Working Space

```
# Create the temporary file
[username@slurm-ui01 ~]$ mktemp
/tmp/1mjAhAH3YM

[username@slurm-ui01 ~]$ mktemp /tmp/user-XXXXX
/tmp/user-WxWtv

# Create the temporary directory
[username@slurm-ui01 ~]$ mktemp -d
/tmp/tmp.1mjAhAH3YM

[username@slurm-ui01 ~]$ mktemp -d /tmp/user-XXXXX
/tmp/user-WxWtv
```


Using /tmp on the worker node as Intermediate Working Space

```
Job_TemporaryFile.sh

#!/bin/bash
#SBATCH --job-name=helloworld      # Job name
#SBATCH --partition=intel-g4-al9_short  # Partiotion name
#SBATCH --nodes=1                  # Numbers of nodes
#SBATCH --ntasks-per-node=1        # Number of tasks per node
#SBATCH --cpus-per-task=1          # Number of CPUs per task
#SBATCH --output=%j.out            # Standard output file (%j: Job ID)
#SBATCH --error=%j.err             # Standard error file (%j: Job ID)
#SBATCH --time=00-01:00:00         # Time limit
#SBATCH --mail-type=ALL            # Email notifications = BEGIN, END, FAIL, ALL
#SBATCH --mail-user=jennifer.chen@twgrid.org # Email address to send notifications

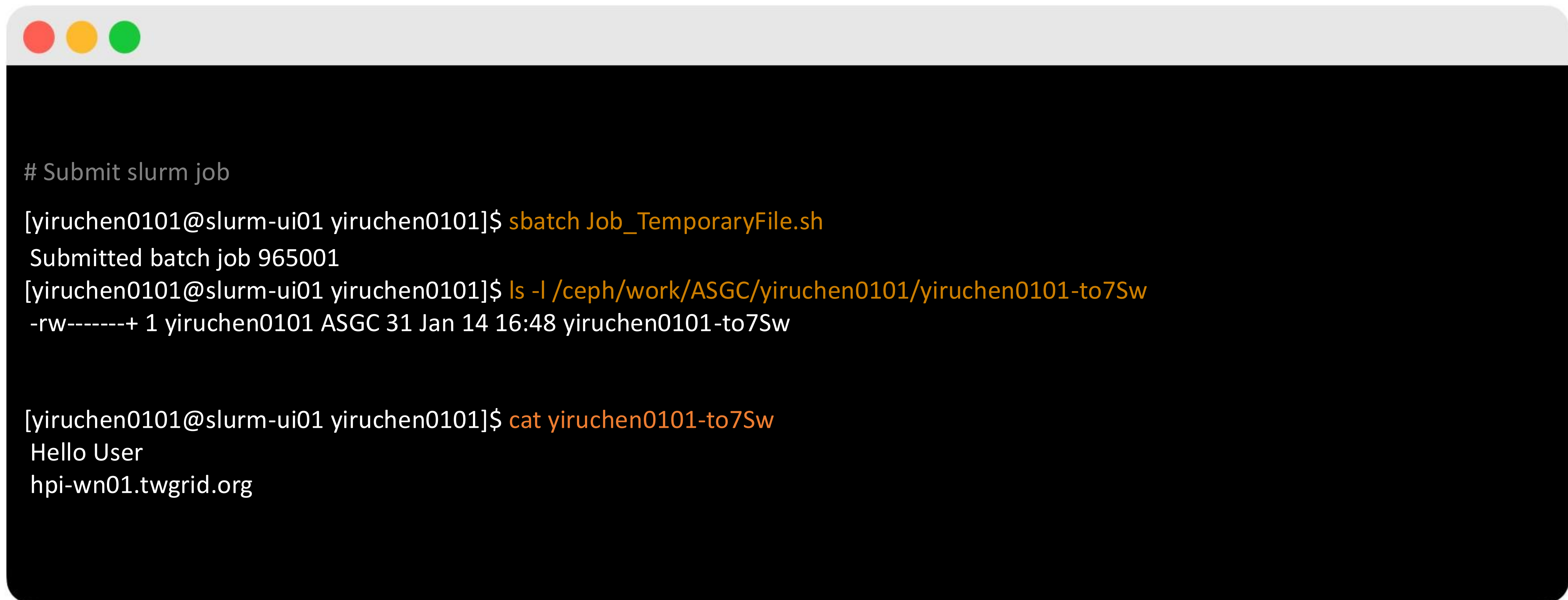
# Create a temporary file with a unique name under /tmp/ directory
TMPFILE=$(mktemp /tmp/yiruchen0101-XXXXX)

# Get the filename (without the path) of the temporary file
filename=$(basename "$TMPFILE")

#your computing program
echo "Hello User" >> $TMPFILE
hostname >> $TMPFILE

# Copy the temporary file to the specified destination directory (/ceph/work/{Group}/{your directory})
cp $TMPFILE /ceph/work/ASGC/yiruchen0101/$filename
# Remove the temporary file from /tmp/ directory to clean up
rm $TMPFILE
```

Using /tmp on the worker node as Intermediate Working Space



```
# Submit slurm job
[yiruchen0101@slurm-ui01 yiruchen0101]$ sbatch Job_TemporaryFile.sh
Submitted batch job 965001
[yiruchen0101@slurm-ui01 yiruchen0101]$ ls -l /ceph/work/ASGC/yiruchen0101/yiruchen0101-to7Sw
-rw-----+ 1 yiruchen0101 ASGC 31 Jan 14 16:48 yiruchen0101-to7Sw

[yiruchen0101@slurm-ui01 yiruchen0101]$ cat yiruchen0101-to7Sw
Hello User
hpi-wn01.twgrid.org
```

Using /tmp on the worker node as Intermediate Working Space

```
Job_TemporaryDir.sh

#!/bin/bash
#SBATCH --job-name=helloworld      # Job name
#SBATCH --partition=intel-g4-al9_short  # Partiotion name
#SBATCH --nodes=1                 # Numbers of nodes
#SBATCH --ntasks-per-node=1       # Number of tasks per node
#SBATCH --cpus-per-task=1         # Number of CPUs per task
#SBATCH --output=%j.out           # Standard output file (%j: Job ID)
#SBATCH --error=%j.err           # Standard error file (%j: Job ID)
#SBATCH --time=00-01:00:00        # Time limit
#SBATCH --mail-type=ALL           # Email notifications = BEGIN, END, FAIL, ALL
#SBATCH --mail-user=jennifer.chen@twgrid.org # Email address to send notifications

# Create a temporary directory with a unique name under /tmp/ directory
TMPDIR=$(mktemp -d /tmp/yiruchen0101-XXXXX)

# Get the directory name (without the path) of the temporary direcotry
filename=$(basename "$TMPDIR")

#your computing program
echo "Hello User" >> $TMPDIR/TMPD.txt
hostname >> $TMPDIR/TMPD.txt

# Recursively copy the entire temporary directory to the specified destination directory (/ceph/work/{Group}/{your directory})
cp -r $TMPDIR /ceph/work/ASGC/yiruchen0101/$filename
# Remove the temporary directory and its contents from /tmp/ directory to clean up
rm -r $TMPDIR
```

Using /tmp on the worker node as Intermediate Working Space

```
# Submit slurm job
[yiruchen0101@slurm-ui01 yiruchen0101]$ sbatch Job_TemporaryDir.sh
Submitted batch job 409355
[yiruchen0101@slurm-ui01 yiruchen0101]$ ls -l /ceph/work/ASGC/yiruchen0101/yiruchen0101-HeycZ
drwx-----+ 2 yiruchen0101 ASGC    1 Jan 14 17:32 yiruchen0101-HeycZ

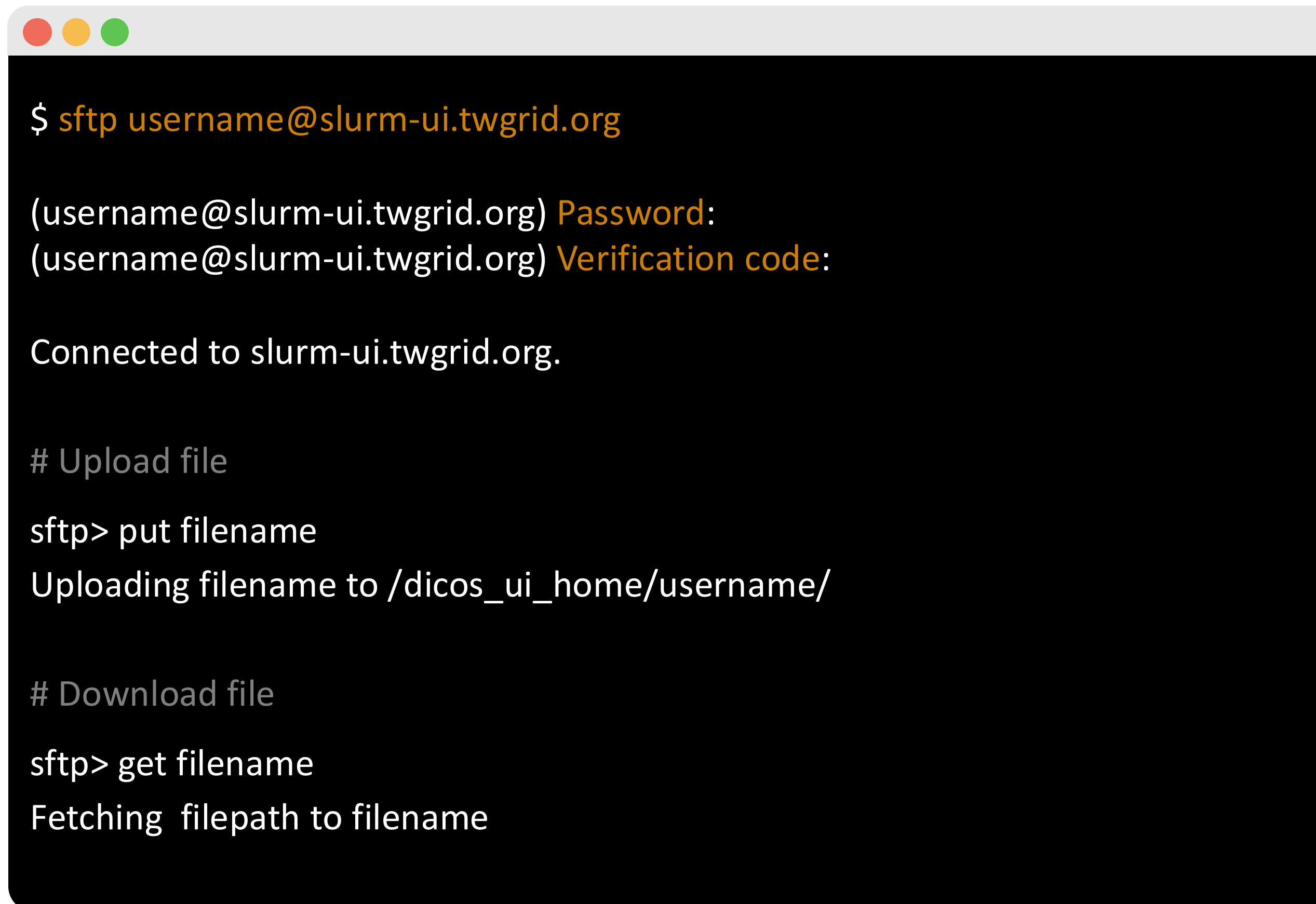
[yiruchen0101@slurm-ui01 yiruchen0101]$ cat yiruchen0101-HeycZ/TMPD.txt
Hello User
hpi-wn01.twgrid.org
```

Data Transfer

- sftp
- scp
- FileZilla

Data Transfer using sftp

Node for Data Transfer: **slurm-ui.twgrid.org** or **dicos-sftp.twgrid.org**



```
$ sftp username@slurm-ui.twgrid.org
(username@slurm-ui.twgrid.org) Password:
(username@slurm-ui.twgrid.org) Verification code:

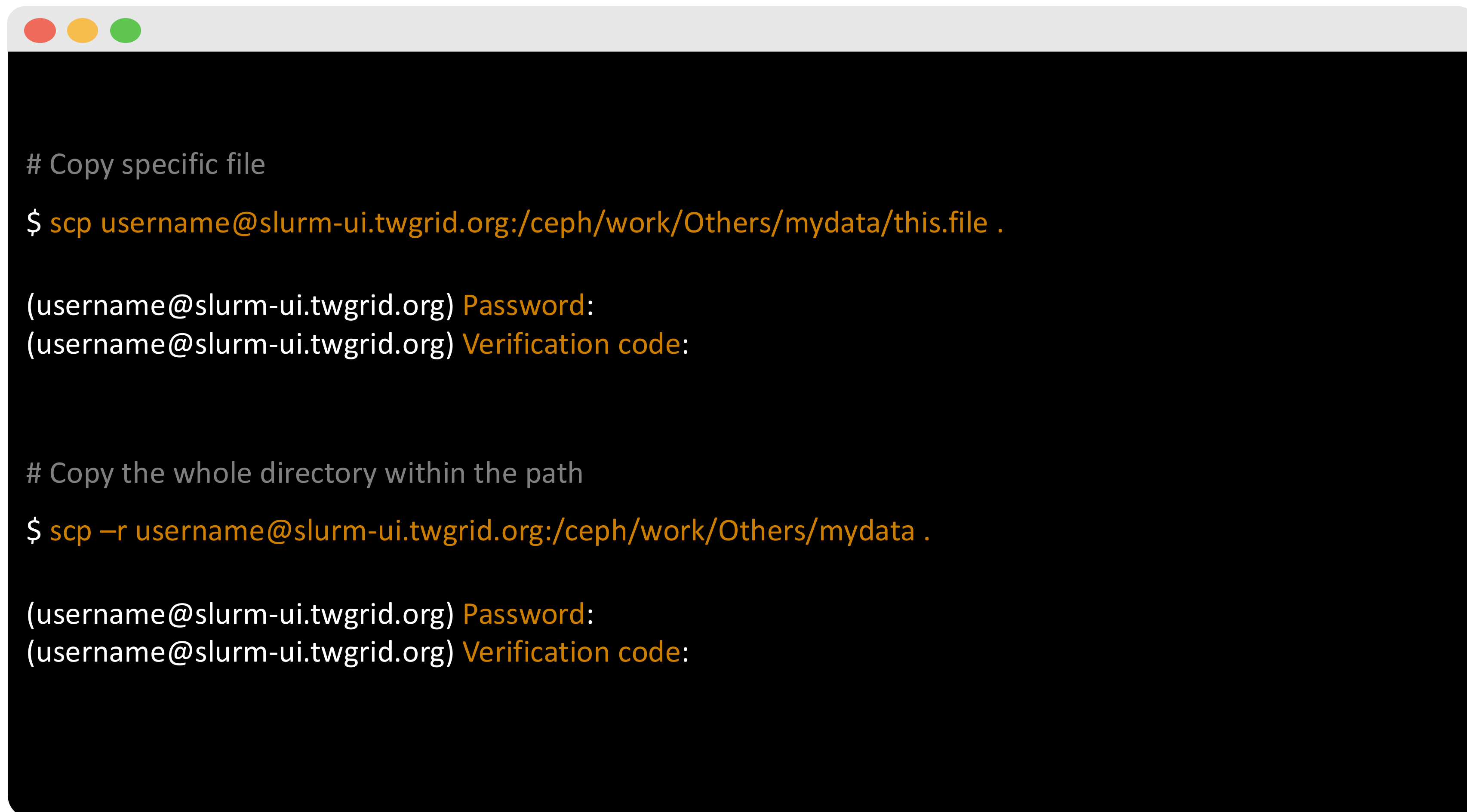
Connected to slurm-ui.twgrid.org.

# Upload file
sftp> put filename
Uploading filename to /dicos_ui_home/username/

# Download file
sftp> get filename
Fetching filepath to filename
```

Data Transfer using scp

Node for Data Transfer: **slurm-ui.twgrid.org** or **dicos-sftp.twgrid.org**



```
# Copy specific file
$ scp username@slurm-ui.twgrid.org:/ceph/work/Others/mydata/this.file .

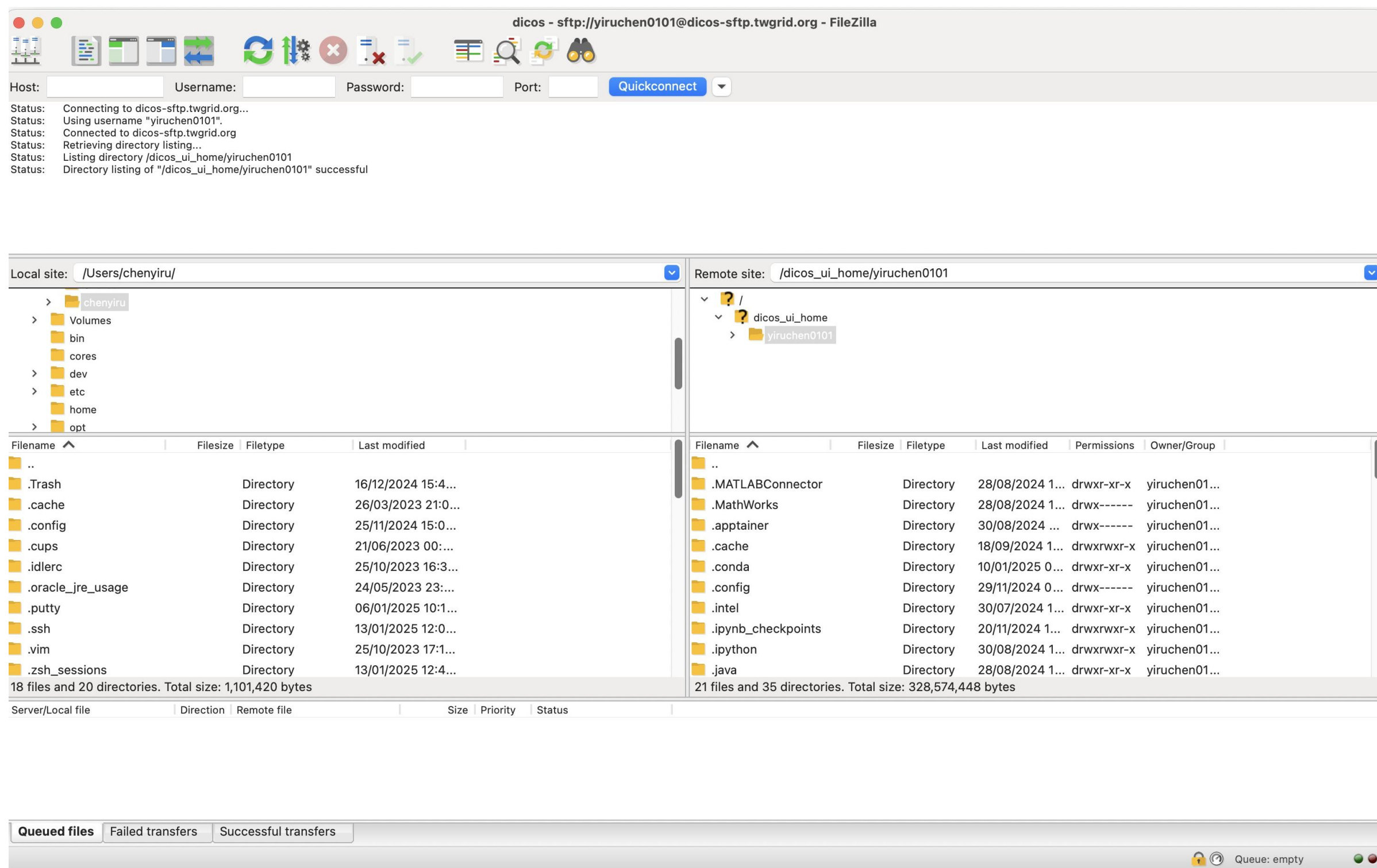
(username@slurm-ui.twgrid.org) Password:
(username@slurm-ui.twgrid.org) Verification code:

# Copy the whole directory within the path
$ scp -r username@slurm-ui.twgrid.org:/ceph/work/Others/mydata .

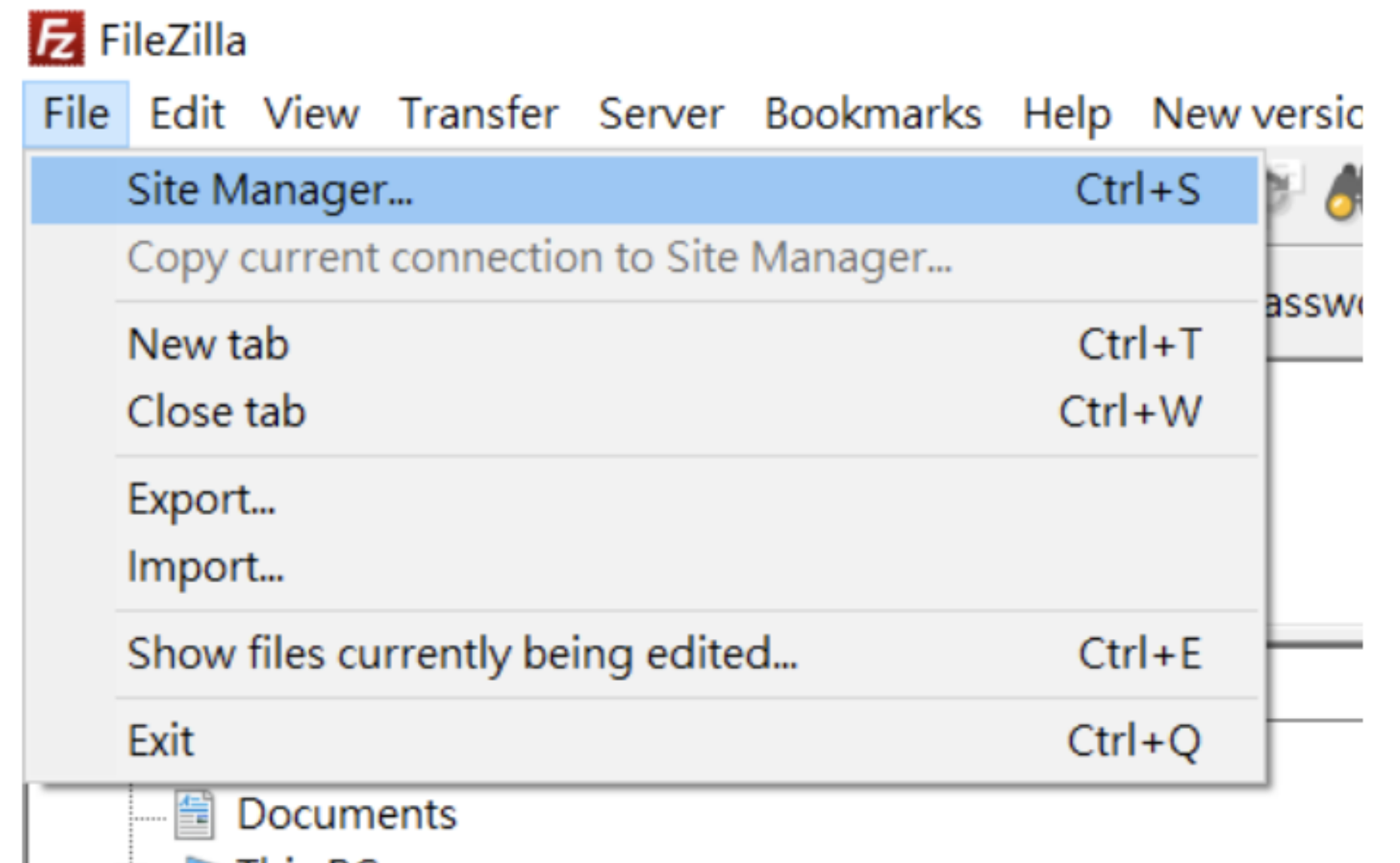
(username@slurm-ui.twgrid.org) Password:
(username@slurm-ui.twgrid.org) Verification code:
```



Data Transfer using FileZilla



1. Choose "Site Manager"





Data Transfer using FileZilla

2. Click "New site" -> choose Protocol to "SFTP" -> enter Host "slurm-ui.twgrid.org" -> choose Logon Type to "Interactive" -> enter your DiCOS account to User field -> click "Connect"

Port: 22 >

Site Manager

select entry:

- My Sites
 - New site

General | Advanced | Transfer Settings | Charset

Protocol: SFTP - SSH File Transfer Protocol

Host: slurm-ui.twgrid.org Port: 22

Logon Type: Interactive

User: <Your DiCOS account>

Background color: None

Comments:

New site | New folder | New Bookmark | Rename | Delete | Duplicate

Connect | OK | Cancel

3. After connect, the first pop-up window please enter your DiCOS password

Enter password

Please enter a password for this server:

Name: New site

Host: slurm-ui.twgrid.org

User: <Your DiCOS account>

Challenge:

SSH server authentication

Password:

Password: ●●●●●●●●●● <Your DiCOS password>

OK | Cancel



Data Transfer using FileZilla

4. The second pop-up window please enter 6-digit code of OTP shows on Google authenticator

Enter password

Please enter a password for this server:

Name: New site

Host: slurm-ui.twgrid.org

User: <Your DiCOS account>

Challenge:

SSH server authentication

Verification code:

<6-digit code of one-time password>

Password: ●●●●●●

OK Cancel

5. After two authentication succeeded, you can see the right window list your DiCOS home directory

Remote site: /dicos_ui_home/<your DiCOS account>

dicos_ui_home

<your DiCOS account>

Filename	Filesize	Filetype	Last modifi...	Permissi...	Owner/G...
..					
.bash_history	173	BASH_H...	9/20/2024 ...	-rw-----	<your
.bash_logout	18	Bash Lo...	9/19/2024 ...	-rw-----	<your
.bash_profile	141	Bash Pr...	9/19/2024 ...	-rw-----	<your
.bashrc	492	Bash RC...	9/19/2024 ...	-rw-----	<your
.google_authent...	303	GOOGL...	9/20/2024 ...	-rw-r--r--	<your
.viminfo	914	VIMINF...	9/20/2024 ...	-rw-----	<your

Thanks for listening

Q&A

dicos-support@twgrid.org

Academia Sinica Grid-computing Centre (ASGC)

