

HPC Workflow on Shaheen

(Chemistry, Physics & Materials Science)

Zhiyong Zhu

KAUST Supercomputing Core Lab

Outline

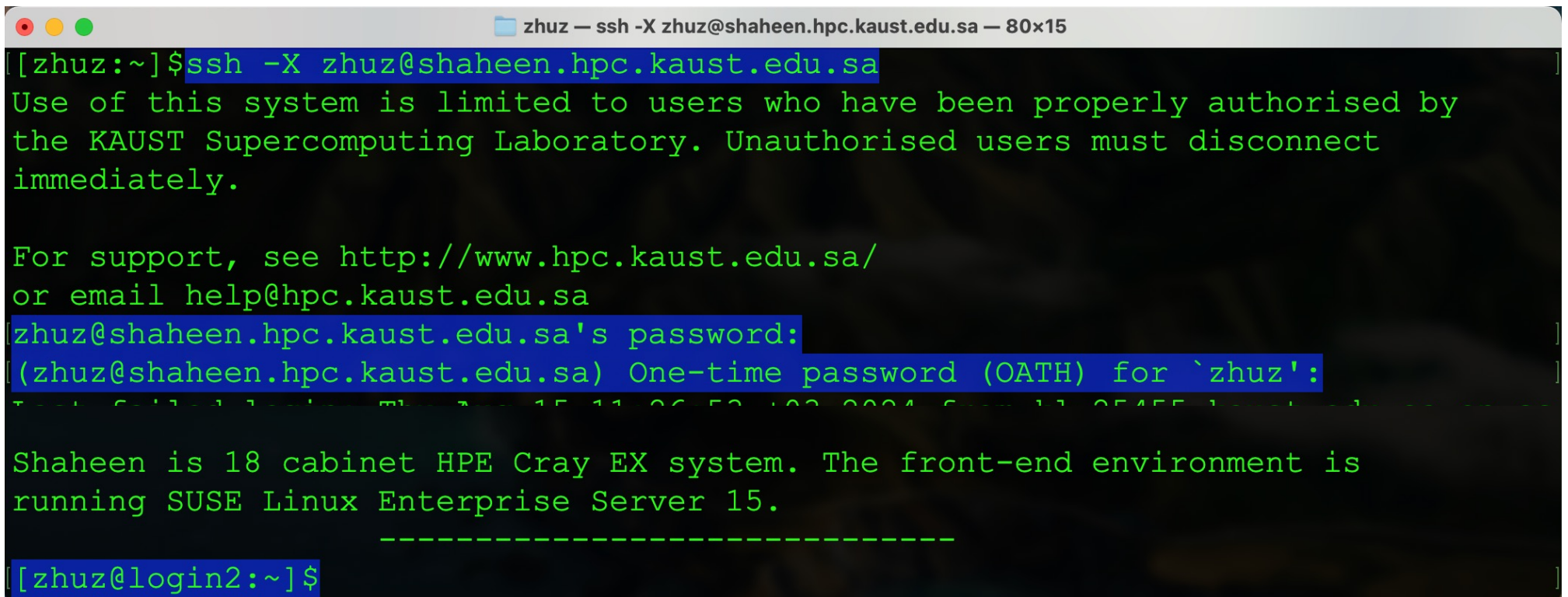
- Workflow
 - Use VASP as an example to show the steps to run
 - The Vienna Ab-initio Simulation Package (VASP) is a computer program for atomic scale materials modeling, e.g. electronic structure calculations and quantum-mechanical molecular dynamics from first-principles (<https://www.vasp.at>).

Workflow

- Login Shaheen
- Check Code Availability
- Working Directory
- Prepare Input Files for VASP
- Prepare Jobscript for Slurm Job Scheduler
- Job Submission using Slurm Commands
- Check Output Files

Login Shaheen

- Login
 - `ssh -X <UserName>@shaheen.hpc.kaust.edu.sa`



```
zhuz — ssh -X zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz:~]$ssh -X zhuz@shaheen.hpc.kaust.edu.sa
Use of this system is limited to users who have been properly authorised by
the KAUST Supercomputing Laboratory. Unauthorised users must disconnect
immediately.

For support, see http://www.hpc.kaust.edu.sa/
or email help@hpc.kaust.edu.sa
zhuz@shaheen.hpc.kaust.edu.sa's password:
(zhuz@shaheen.hpc.kaust.edu.sa) One-time password (OATH) for `zhuz':
-----
Shaheen is 18 cabinet HPE Cray EX system. The front-end environment is
running SUSE Linux Enterprise Server 15.

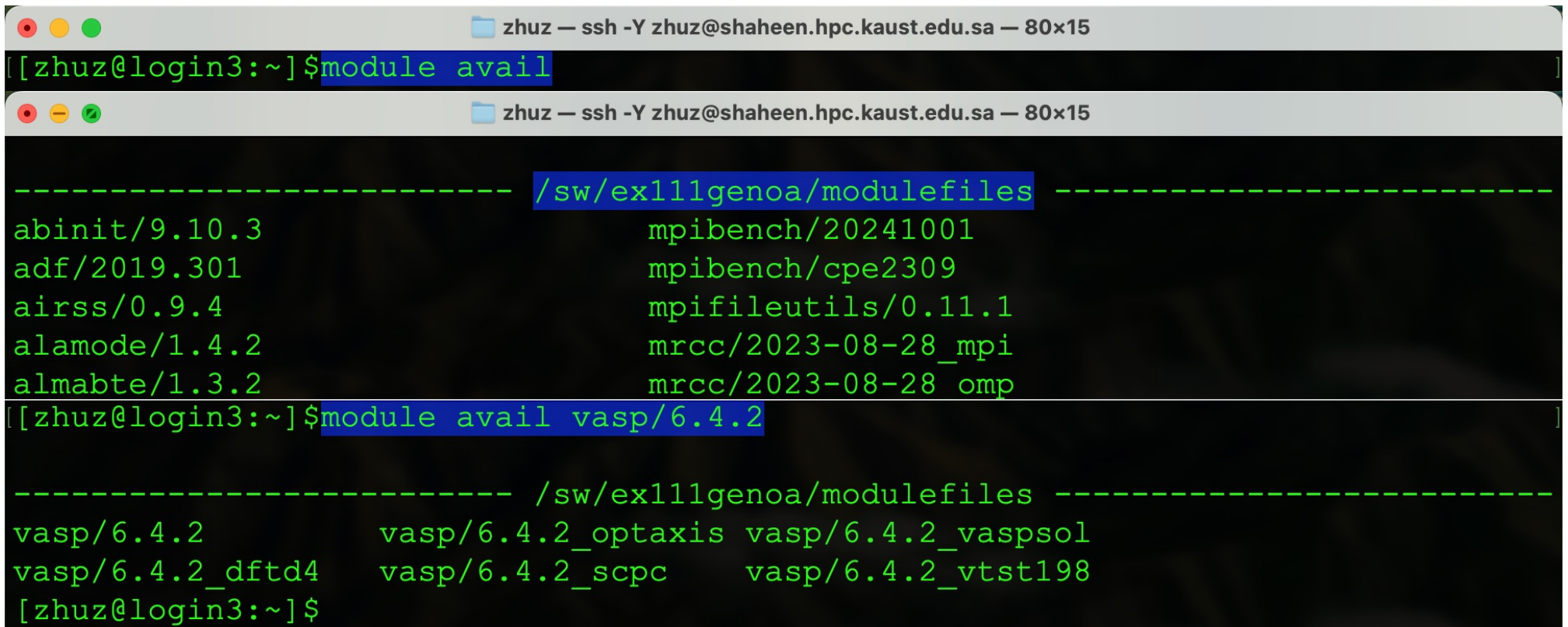
-----
[zhuz@login2:~]$
```

Code Availability

- On Shaheen login node:
 - module avail
- In /sw/ex111genoa software stack:
 - ls /sw/ex111genoa
- From our website:
 - <https://docs.hpc.kaust.edu.sa>

Code Availability

- module avail
 - module avail
 - module avail <code>/<version>



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:~]$ module avail
----- /sw/ex111genoa/modulefiles -----
abinit/9.10.3                mpibench/20241001
adf/2019.301                 mpibench/cpe2309
airss/0.9.4                  mpifileutils/0.11.1
alamode/1.4.2                mrcc/2023-08-28_mpi
almabte/1.3.2                mrcc/2023-08-28_omp
[zhuz@login3:~]$ module avail vasp/6.4.2
----- /sw/ex111genoa/modulefiles -----
vasp/6.4.2                   vasp/6.4.2_optaxis vasp/6.4.2_vaspsol
vasp/6.4.2_dftd4             vasp/6.4.2_scpc    vasp/6.4.2_vtst198
[zhuz@login3:~]$
```

Code Availability

- `ls /sw/ex11genoa`

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:~]$ls /sw/ex11genoa
abinit      crystal14  jmol       openfoam   sod
adf         cuby4      koopmans   openmolcas softbv
airss       dftbplus  kwant      openmx     spack
alamode     dftd4     lammps     orca       sumo
almabte     dlpoly    lev0       osu-microbenchmarks tbmodels
amber       dpcode    libxc      ovito      tdep
amd         dssp      libxml2    p4vasp     thirdorder
ams         eddp      lobster    pacchem    totalview
amset       edmftf    materstudio packmol     towhee
ansys       egsnrc    matlab     periodic_nbo turbomole
arm-forge   eigen     milo       perturbo   uspex
ase         elk       mkl        phono3py   vampire
atk         elpa     modulefiles phonopy     vasp
atompaw     espresso  mohid      plumed     vaspkit
```


Code Availability

- <https://docs.hpc.kaust.edu.sa>
 - [Software ecosystem -> Software environment -> Applications catalogue -> Shaheen III](#)

The screenshot shows a web browser displaying the KAUST Supercomputing Laboratory website. The URL in the address bar is docs.hpc.kaust.edu.sa/apps_catalogue/shaheen3.html#. The page features a purple navigation bar with a search icon and a link to "Checkout, Frequently Asked Questions!". The main content area is titled "Shaheen III Compilers" and contains a table with the following data:

System Build	App	Version	Compiler
ex111genoa	python	3.10.13	lib64
ex111genoa	python	3.10.13	Python-3.10.13
ex111genoa	cmake	3.18.2	gcc7.5.0
ex111genoa	cmake	3.28.3	gcc7.5.0
ex111genoa	cmake	3.30.5	gcc7.5.0
ex109genoa	python	3.10.13	sles15sp4

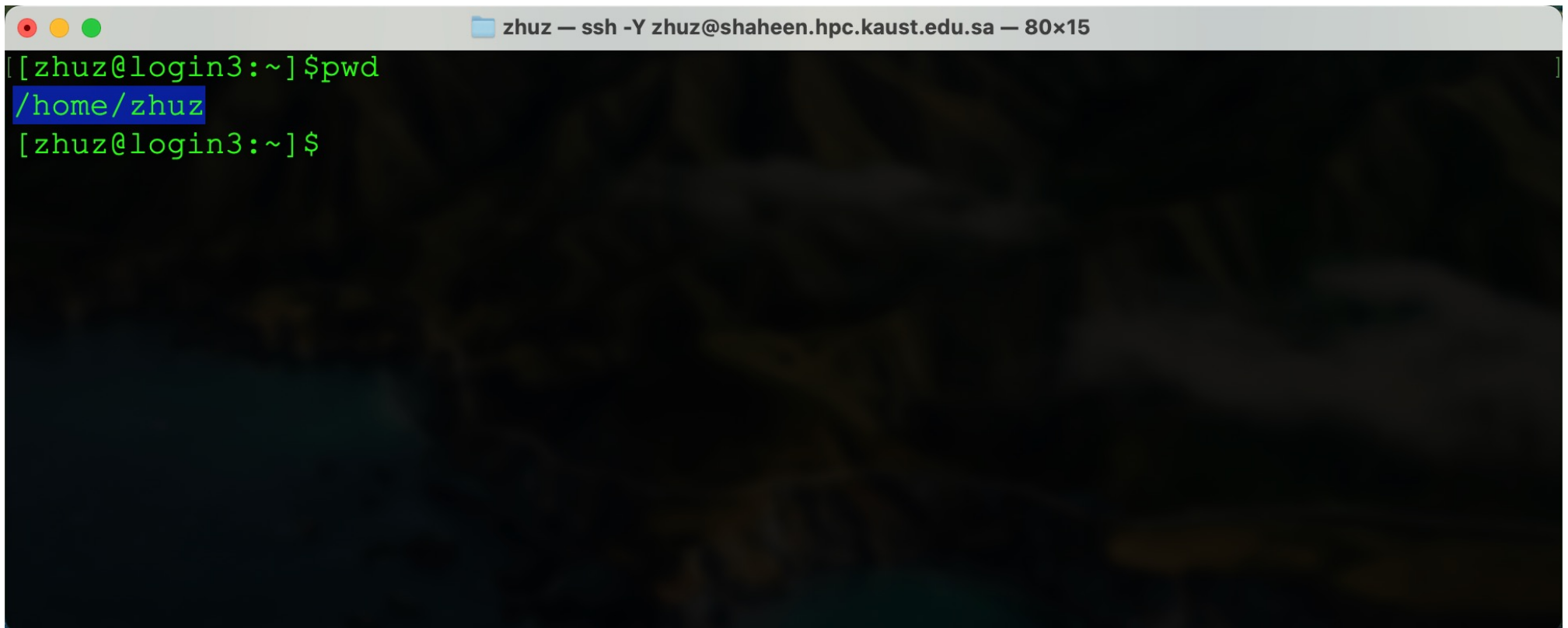
The left sidebar shows "Section Navigation" with "Applications catalogue" and "Shaheen III" highlighted. The right sidebar shows "On this page" with links to "Compilers", "Optimized Libraries", "Computational Chemistry", "Bioscience", "Computational Fluid Dynamics", "Data Science", and "Others".

Code Availability

- On Shaheen login node:
 - module avail
- In /sw/ex111genoa software stack:
 - ls -l /sw/ex111genoa
- From our website:
 - <https://docs.hpc.kaust.edu.sa>
- Not found?
 - help@hpc.kaust.edu.sa

3 Different Working Directories

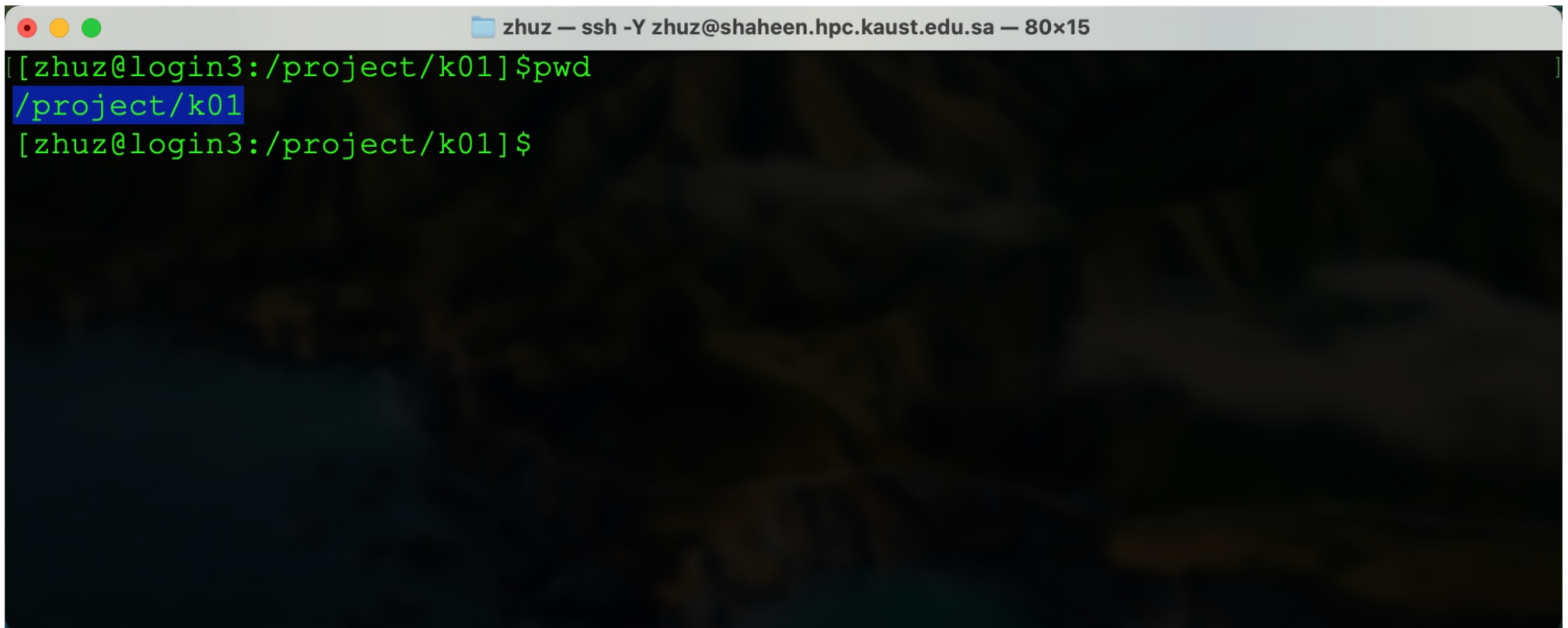
- /home
 - Very limited space; Not mounted on compute nodes (job submission will fail)



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:~]$ pwd
/home/zhuz
[zhuz@login3:~]$
```

3 Different Working Directories

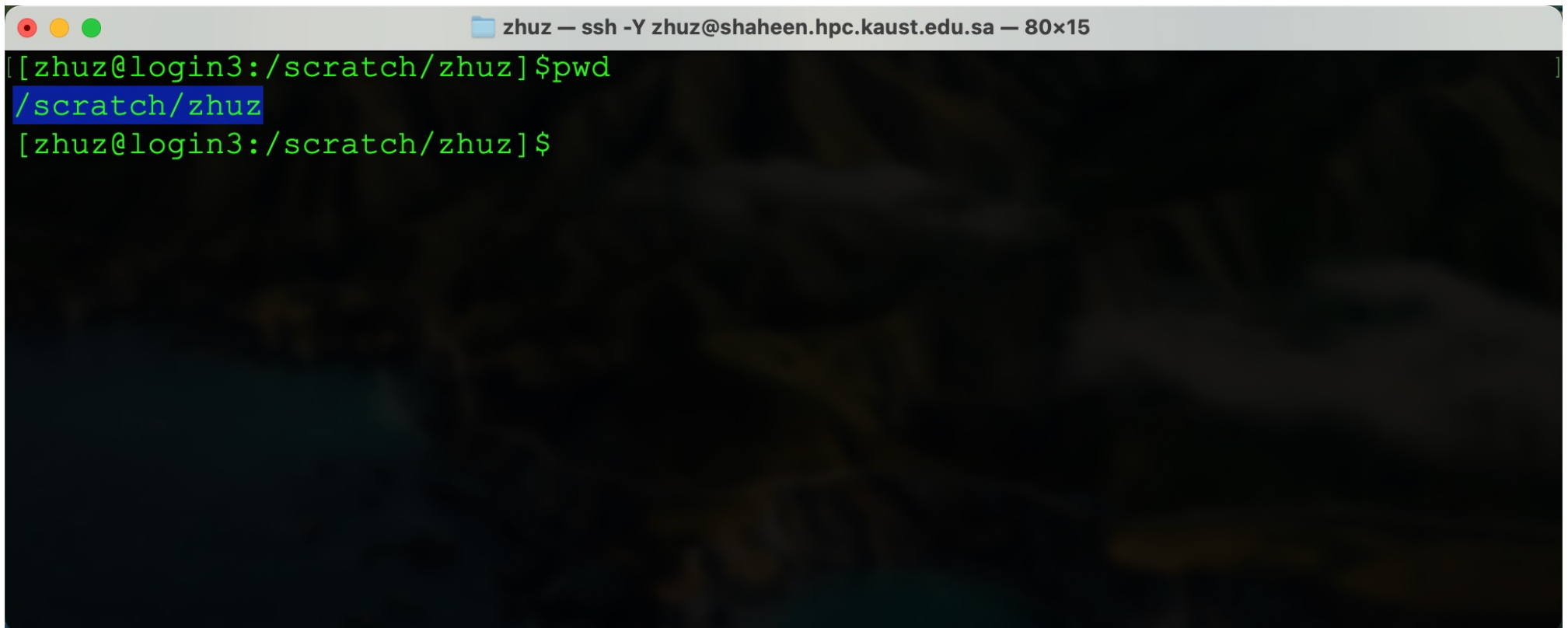
- /project/<projectname>
 - Read-only for compute nodes (job submission will fail); Used for data backup and data sharing



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/project/k01]$pwd
/project/k01
[zhuz@login3:/project/k01]$
```

3 Different Working Directories

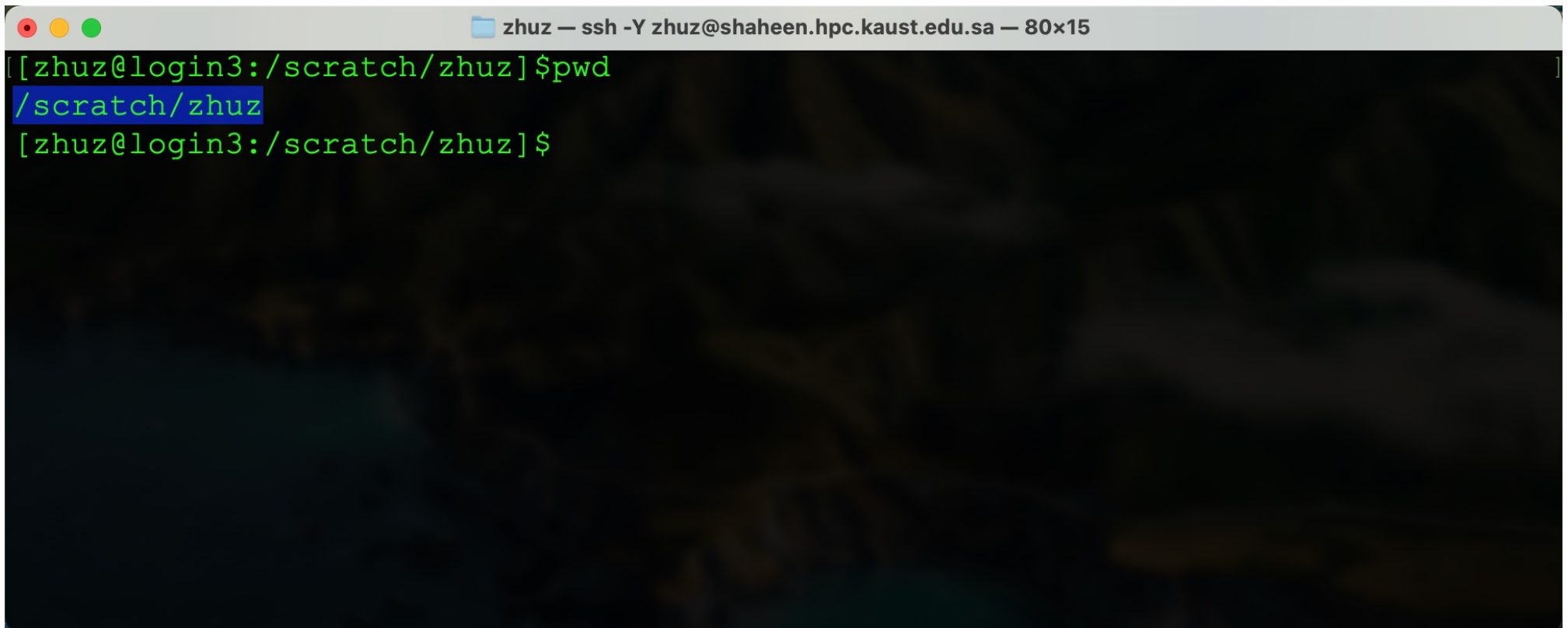
- /scratch/<username>
 - Almost unlimited space



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz]$pwd
/scratch/zhuz
[zhuz@login3:/scratch/zhuz]$
```

3 Different Working Directories

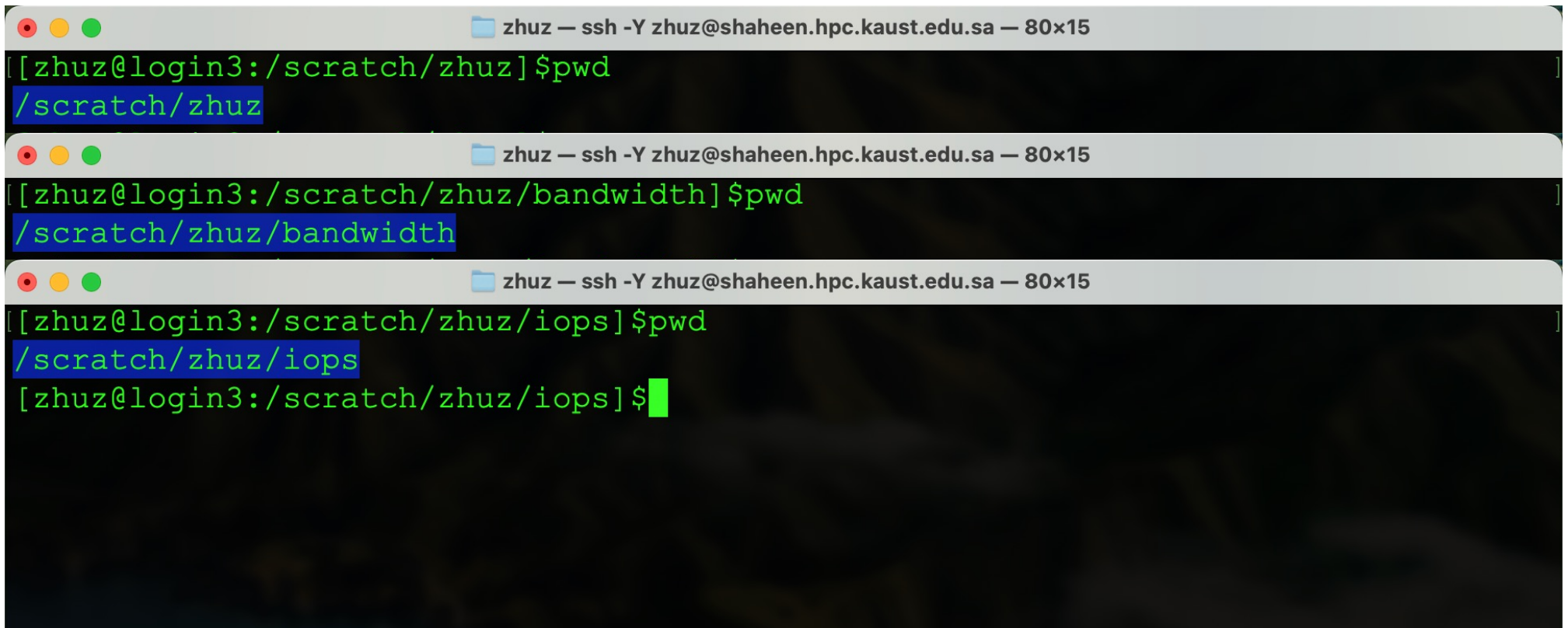
- Where to run? /scratch!
 - The only place to run
 - Remember to backup important data to /project



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz]$pwd
/scratch/zhuz
[zhuz@login3:/scratch/zhuz]$
```

3 Different Working Directories

- /scratch/<username> - 3 tiers
 - capacity, bandwidth, iops



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz]$pwd
/scratch/zhuz

zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz/bandwidth]$pwd
/scratch/zhuz/bandwidth

zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz/iops]$pwd
/scratch/zhuz/iops
[zhuz@login3:/scratch/zhuz/iops]$
```

3 Different Working Directories

- /scratch/<username>
 - Tier “capacity”
 - Large capacity (10T per user by default)
 - Low performance in terms of bandwidth and iops
 - Used for calculations that are not sensitive to IO performance
 - Data purged after 60 days of no access

3 Different Working Directories

- /scratch/<username>/bandwidth
 - tier “bandwidth”
 - Low capacity (1T per user by default)
 - High performance in terms of IO bandwidth
 - Used for calculations that read/write a large amount of data
 - Data purged after 60 days of no access

3 Different Working Directories

- /scratch/<username>/iops
 - Tier “iops”
 - Low capacity (50G per user by default)
 - High performance in terms of # IO operations
 - Used for calculations that read/write large number of files, and software installation (conda, python, etc)
 - No data purging

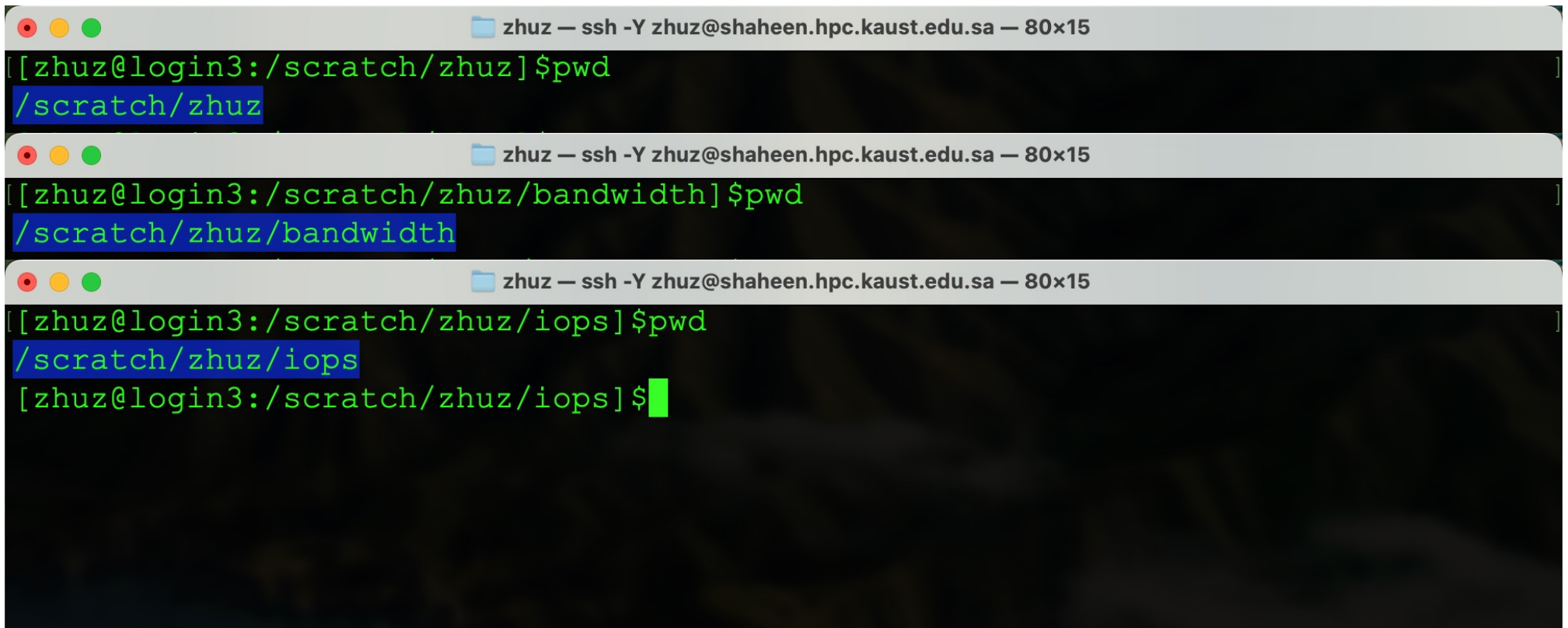
3 Different Working Directories

- Quota limits
 - kuq

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 100x20
[zhuz@login3:~]$ kuq
-----
Filesystem quota limits for user zhuz
Tier          Filesystem  used   quota  limit  grace  files  quota  limit  grace
-----
scratch       /scratch   678.3G  0k     11T    -      242181  0 1024000  -
  capacity    /scratch   667.1G  0k     10T    -      242181  0      0      -
  bandwidth   /scratch   458.2M  0k     1T     -      242181  0      0      -
  iops        /scratch   10.89G  0k     50G    -      242181  0      0      -
project       /project    3.491T  0k     0k     -      2211833  0 3000000  -
-----
[zhuz@login3:~]$
```

3 Different Working Directories

- Which tier to use? It depends!
 - Do your own tests



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz]$pwd
/scratch/zhuz

zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz/bandwidth]$pwd
/scratch/zhuz/bandwidth

zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz/iops]$pwd
/scratch/zhuz/iops
[zhuz@login3:/scratch/zhuz/iops]$
```

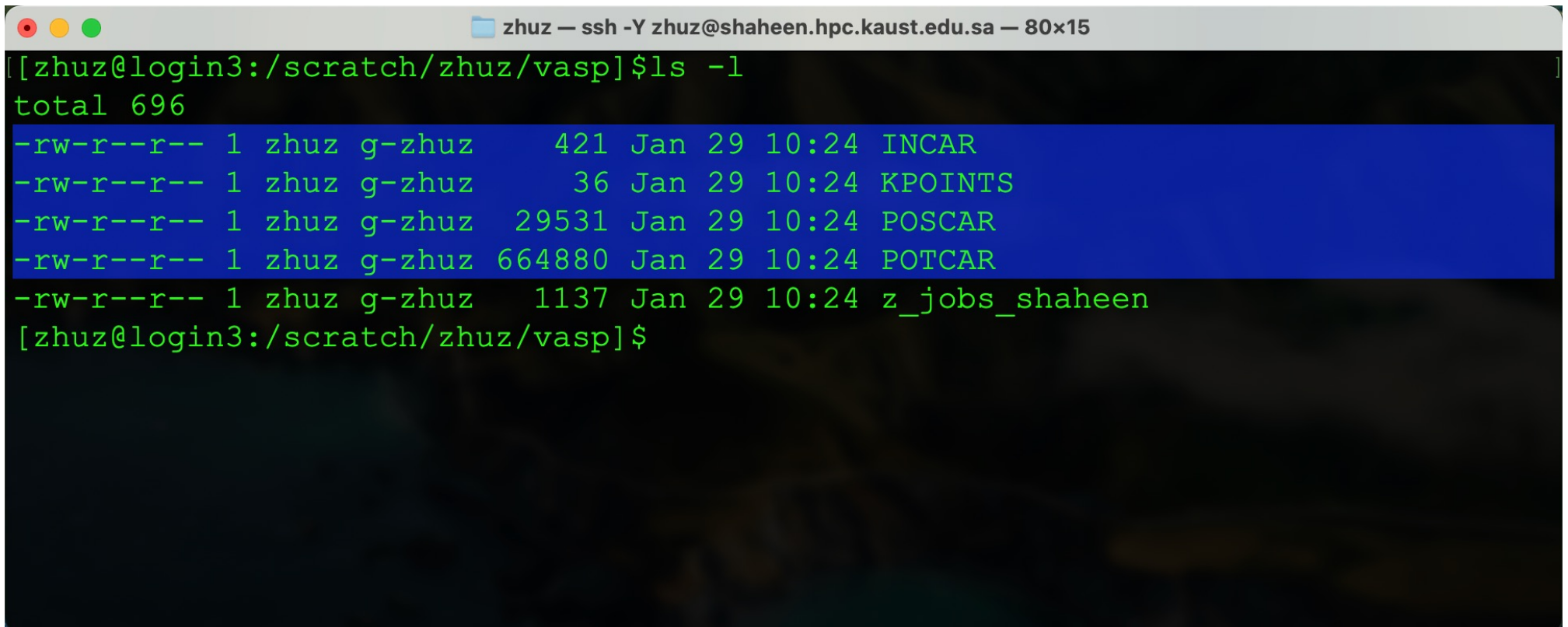
Prepare Input Files

- Examples under Installation Folder
 - /sw/ex111genoa/code/ver/compilation/example
 - Inputs for application; Jobscript for Slurm

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:~]$cd /scratch/zhuz
[zhuz@login3:/scratch/zhuz]$mkdir vasp
[zhuz@login3:/scratch/zhuz]$cd vasp
[zhuz@login3:/scratch/zhuz/vasp]$cp /sw/ex111genoa/vasp/6.4.2/intel19.0.5/example/02/* .
[zhuz@login3:/scratch/zhuz/vasp]$ls -l
total 696
-rw-r--r-- 1 zhuz g-zhuz    421 Jan 29 10:24 INCAR
-rw-r--r-- 1 zhuz g-zhuz     36 Jan 29 10:24 KPOINTS
-rw-r--r-- 1 zhuz g-zhuz  29531 Jan 29 10:24 POSCAR
-rw-r--r-- 1 zhuz g-zhuz 664880 Jan 29 10:24 POTCAR
-rw-r--r-- 1 zhuz g-zhuz   1137 Jan 29 10:24 z_jobs_shaheen
[zhuz@login3:/scratch/zhuz/vasp]$
```

Prepare Input Files

- VASP Input Files
 - Upload from your own personal workstations
 - Modifying existing input files



```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz/vasp]$ls -l
total 696
-rw-r--r-- 1 zhuz g-zhuz    421 Jan 29 10:24 INCAR
-rw-r--r-- 1 zhuz g-zhuz     36 Jan 29 10:24 KPOINTS
-rw-r--r-- 1 zhuz g-zhuz  29531 Jan 29 10:24 POSCAR
-rw-r--r-- 1 zhuz g-zhuz 664880 Jan 29 10:24 POTCAR
-rw-r--r-- 1 zhuz g-zhuz   1137 Jan 29 10:24 z_jobs_shaheen
[zhuz@login3:/scratch/zhuz/vasp]$
```


Prepare Input Files

- Slurm Jobscrip
– SLURM directives

```
zhu@zhu:~$ ssh -Y zhu@shaheen.hpc.kaust.edu.sa - 80x15
#!/bin/bash
#SBATCH --partition=workq
#SBATCH --job-name=vasp
#SBATCH --nodes=8
#SBATCH --time=4:00:00
#SBATCH --exclusive
#SBATCH --err=std.err
#SBATCH --output=std.out
#-----#
module switch PrgEnv-cray PrgEnv-intel
module switch intel intel/19.0.5.281
module load vasp/6.4.2
#module load vasp/6.4.4_dftd4 # https://github.com/dftd4/dftd4_vasp
#module load vasp/6.4.2_optaxis # https://github.com/Chengcheng-Xiao/VASP_OPT_AX
z_jobs_shaheen lines 1-14/25 42%
```


Prepare Input Files

- Slurm Jobscrip
– Environments settings

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
module switch PrgEnv-cray PrgEnv-intel
module switch intel intel/19.0.5.281
module load vasp/6.4.2
#module load vasp/6.4.4_dftd4 # https://github.com/dftd4/dftd4_vasp
#module load vasp/6.4.2_optaxis # https://github.com/Chengcheng-Xiao/VASP_OPT_AXIS (Fixing specific stress tensor element(s))
#module load vasp/6.4.2_scpc # https://github.com/aradi/SCPC-Method/tree/main
#module load vasp/6.4.2_vaspsol # https://github.com/henniggroup/VASPsol/tree/master
#module load vasp/6.4.2_vtst198 # http://theory.cm.utexas.edu/vtsttools
export FI_CXI_RX_MATCH_MODE=software
export MKL_DEBUG_CPU_TYPE=5
export MKL_CBWR=auto
export OMP_NUM_THREADS=1
z_jobs_shaheen lines 10-21/25 77%
```

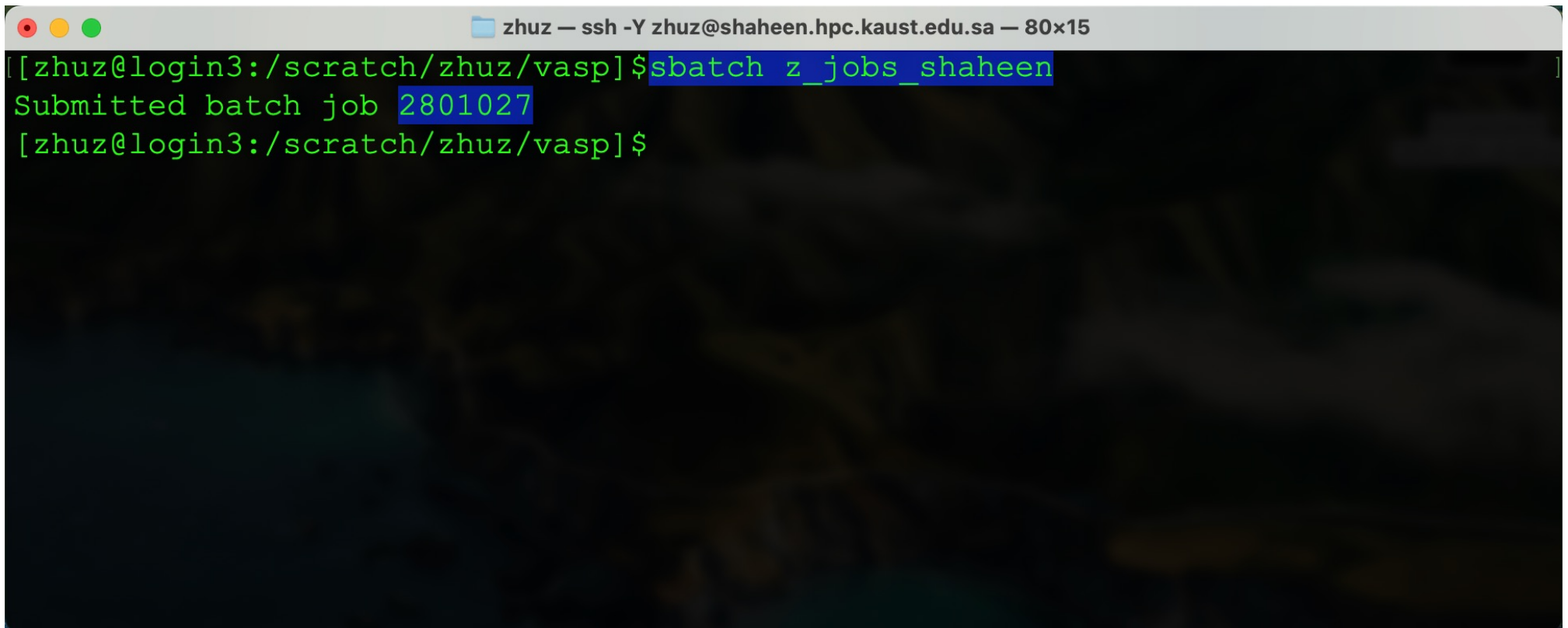
Prepare Input Files

- Slurm Jobscrip
– Commands to run

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
#module load vasp/6.4.2_optaxis # https://github.com/Chengcheng-Xiao/VASP_OPT_AXIS (Fixing specific stress tensor element(s))
#module load vasp/6.4.2_scpc # https://github.com/aradi/SCPC-Method/tree/main
#module load vasp/6.4.2_vaspsol # https://github.com/henniggroup/VASPsol/tree/master
#module load vasp/6.4.2_vtst198 # http://theory.cm.utexas.edu/vtsttools
export FI_CXI_RX_MATCH_MODE=software
export MKL_DEBUG_CPU_TYPE=5
export MKL_CBWR=auto
export OMP_NUM_THREADS=1
#-----#
echo "The job "${SLURM_JOB_ID}" is running on "${SLURM_JOB_NODELIST}"
#-----#
srun --ntasks=1536 --map-by=numa --hint=nomultithread ${VASP_HOME}/vasp_std
z_jobs_shaheen lines 14-25/25 (END)
```

Job Submission

- sbatch
 - Submit jobs

A terminal window with a grey title bar containing the text "zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15". The terminal content shows a green prompt "[zhuz@login3:/scratch/zhuz/vasp]\$" followed by the command "sbatch z_jobs_shaheen" which is highlighted in blue. The next line shows the output "Submitted batch job 2801027" where "2801027" is highlighted in blue. The final line shows the prompt "[zhuz@login3:/scratch/zhuz/vasp]\$" again.

```
[zhuz@login3:/scratch/zhuz/vasp]$sbatch z_jobs_shaheen
Submitted batch job 2801027
[zhuz@login3:/scratch/zhuz/vasp]$
```

Job Submission

- `squeue`
 - Check job status

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 120x24
[zhuz@login3:/scratch/zhuz/vasp]$ squeue --me
  JOBID      USER ACCOUNT      NAME  ST REASON   START_TIME          TIME  TIME_LEFT  NODES
  2795375    zhuz   k01      vasp  R None     2025-01-29T09:10:46  1:27:08  2:32:52    4
  2801027    zhuz   k01      vasp  R None     2025-01-29T10:36:57  0:57    3:59:03    8
[zhuz@login3:/scratch/zhuz/vasp]$
```


Job Submission

- scancel
 - Cancel jobs

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 120x24
[[zhuz@login3:/scratch/zhuz/vasp]$squeue --me
  JOBID      USER ACCOUNT      NAME  ST REASON      START_TIME          TIME  TIME_LEFT  NODES
   2795375   zhuz    k01      vasp  R None      2025-01-29T09:10:46  1:27:08  2:32:52    4
   2801027   zhuz    k01      vasp  R None      2025-01-29T10:36:57   0:57  3:59:03    8
[[zhuz@login3:/scratch/zhuz/vasp]$scancel 2801027
[[zhuz@login3:/scratch/zhuz/vasp]$squeue --me
  JOBID      USER ACCOUNT      NAME  ST REASON      START_TIME          TIME  TIME_LEFT  NODES
   2795375   zhuz    k01      vasp  R None      2025-01-29T09:10:46  1:28:26  2:31:34    4
[[zhuz@login3:/scratch/zhuz/vasp]$
```

Check Output Files

- Successful or Not? If yes, Analyze Results
 - Standard output/error files:std.out/std.err)
 - Application output files: OUTCAR

```
zhuz — ssh -Y zhuz@shaheen.hpc.kaust.edu.sa — 80x15
[zhuz@login3:/scratch/zhuz/vasp]$ls -l std.*
-rw-r--r-- 1 zhuz g-zhuz 132500 Jan 29 10:41 std.err
-rw-r--r-- 1 zhuz g-zhuz   5653 Jan 29 10:49 std.out
[zhuz@login3:/scratch/zhuz/vasp]$grep "free energy TOTEN" OUTCAR
  free energy TOTEN =      -3042.35709769 eV
[zhuz@login3:/scratch/zhuz/vasp]$
```

Tips

- Do not run directly on the login nodes
 - login nodes are shared
- It won't work to submit jobs from /home
 - /home is not seen on the compute nodes
- Backup important data from /scratch to /project (or /home, or your local computers)
 - Files in /scratch are not backed up, and are deleted automatically after 60 days
 - Do not confuse /scratch/project and /project

Tips

- Licensed software
 - Need license for commercial software
 - Different software have different license terms
 - VASP: you can use it on Shaheen, as long as you have your own license
 - Gaussian: you cannot use it on Shaheen, even if you have your own license
 - Contact us if you have any questions

Thank You!

help@hpc.kaust.edu.sa

Agenda

- 8:30am Welcome
- 8:35am Shaheen III Hardware Overview
- 8:55am How to apply on Shaheen III
- 9:05am Getting Started on Shaheen III
- 9:15am Software Environment
- 9:35am Job Scheduling
- 10:00am Coffee Break
- 10:15am Storage overview & Best practices
- 10:30am Applications software example: VASP workflow
- **10:50 am Applications software example: CFD applications**
- 11:10 am Applications software example: Bio informatics workflow
- 11:20-11.30am Q&A and Open Discussion



Shaheen III Survey