For all Tasks: Long outputs are shortend. Displayed by 2 lines with "...".

Lab Exercise: Linux Shell This exercise will familiarize you with the handling of the Linux operating system and some of its tools/services. Unless otherwise stated, use the command line interface (bash recommended). Complete the following tasks and write a concise report that will allow others to replicate your work Include the command used and a brief explanation (one-liner) of what the result looked like. If you use guides, cite them in the appendix.

Task 1:

Warm-up Answer the following questions:

• What is your OS kernel version?

```
andig91@labinf-ct01:~$ uname -r 5.4.143-1-pve
```

- Update your Linux apt repositories and upgrade your existing apps to the latest version
 - You might have to use root privileges for this task

```
andig91@labinf-ct01:~$ sudo apt update

OK:1 http://archive.ubuntu.com/ubuntu focal InRelease

OK:2 http://archive.ubuntu.com/ubuntu focal-updates InRelease

Holen:3 http://security.ubuntu.com/ubuntu focal-security InRelease [114 kB]

Holen:4 http://archive.ubuntu.com/ubuntu focal/main Translation-de [477 kB]

Holen:5 http://archive.ubuntu.com/ubuntu focal/restricted Translation-de [4.436 B]

Holen:6 http://archive.ubuntu.com/ubuntu focal/universe Translation-de [2.417 kB]

Holen:7 http://archive.ubuntu.com/ubuntu focal/multiverse Translation-de [109 kB]

Es wurden 3.121 kB in 1 s geholt (2.422 kB/s).

Paketlisten werden gelesen... Fertig

Abhängigkeitsbaum wird aufgebaut.

Statusinformationen werden eingelesen... Fertig

Alle Pakete sind aktuell.
```

• Using apt, download and install the "Unity Tweak Tool" and take a (voluntary) look at Ubuntu customization. Search for the exact package name of the tool using apt.

Optional: If you are not on Ubuntu or do not care about customization, install e.g. htop or the VLC media player instead

```
andig91@labinf-ct01:~$ sudo apt search Unity Tweak Tool
Sortierung... Fertig
Volltextsuche... Fertig
unity-tweak-tool/focal 0.0.7+-Oubuntu6 all
   configuration tool for the Unity desktop environment

youtube-dl/focal 2020.03.24-1 all
   Lädt Videos von YouTube und anderen Sites herunter

andig91@labinf-ct01:~$ sudo apt install unity-tweak-tool
```

```
Paketlisten werden gelesen... Fertig
Abhängigkeitsbaum wird aufgebaut.
Statusinformationen werden eingelesen.... Fertig
Die folgenden zusätzlichen Pakete werden installiert:
...
...
...
zeitgeist-core zeitgeist-datahub zenity zenity-common zip
0 aktualisiert, 865 neu installiert, 0 zu entfernen und 0 nicht aktualisiert.
Es müssen 325 MB an Archiven heruntergeladen werden.
Nach dieser Operation werden 1.646 MB Plattenplatz zusätzlich benutzt.
Möchten Sie fortfahren? [J/n] J
...
...
```

Unity Tweak Tool is not supported, we have the choice to use GNOME Tweak Tool or install an other application.

We have chosen to install VLC Media Player

```
student@student-virtual-machine:-$ sudo apt install vlc
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
...
...
0 upgraded, 46 newly installed, 0 to remove and 440 not upgraded.
Need to get 14,1 MB of archives.
After this operation, 73,0 MB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 http://at.archive.ubuntu.com/ubuntu focal/universe amd64 libaribb24-0 amd64
1.0.3-2 [26,9 kB]
...
student@student-virtual-machine:-$ vlc
```

• View the contents of the file /etc/passwd. What do you think it contains?

```
student@student-virtual-machine:~$ cat /etc/passwd
root:x:0:0:root:/root:/bin/bash
...
...
student:x:1000:1000:student,,,:/home/student:/bin/bash
systemd-coredump:x:999:999:systemd Core Dumper:/:/usr/sbin/nologin
sshd:x:126:65534::/run/sshd:/usr/sbin/nologin
```

$\underline{\text{https://www.cyberciti.biz/faq/understanding-etcpasswd-file-format/}}$

Discriped on User student (delimited by ":"):

Text	Description
student	Username

Х	Password stored in /etc/shadow
1000	User ID
1000	Group ID
student,,,	User Beschreibung
/home/student	Pfad des Homeverzeichnis
/bin/bash	Command/Shell

- Create the following (empty) files (and directories, if needed) in your home directory:
 - /home//Pictures/bild01.png
 - /home//Documents/text01.md
 - /home//Secret/text02.txt

```
student@student-virtual-machine:~$ cd Pictures/
student@student-virtual-machine:~/Pictures$ touch bild01.png
student@student-virtual-machine:~/Pictures$ cd ../Documents/
student@student-virtual-machine:~/Documents$ touch tet01.md
student@student-virtual-machine:~/Documents$ cd ../
.cache/ .config/ Desktop/ Documents/ Downloads/ .gnupg/ .local/ Music/
Pictures/ Public/ .ssh/ Templates/ Videos/
student@student-virtual-machine:~/Documents$ mkdir ../Secret
student@student-virtual-machine:~/Documents$ cd ../Secret/
student@student-virtual-machine:~/Secret$ touch text02.txt
```

• Write the contents (using ls) of the directories /etc and /sbin to the "text02.txt" file

```
student@student-virtual-machine:~/Secret$ ls /etc/ > text02.txt
student@student-virtual-machine:~/Secret$ ls /sbin/ >> text02.txt
student@student-virtual-machine:~/Secret$ cat text02.txt
acpi
adduser.conf
alsa
...
...
xml
zsh_command_not_found
aa-remove-unknown
aa-status
...
...
xtables-monitor
xtables-nft-multi
zerofree
zic
zramctl
```

• Check the remaining free space on your disk

```
student@student-virtual-machine:~/Secret$ df -h
              Size Used Avail Use% Mounted on
Filesystem
udev
              953M
                      0 953M 0% /dev
              196M 1,6M 195M 1% /run
tmpfs
              29G 8,5G 19G 31% /
/dev/sda5
tmpfs
              980M
                   0 980M
                              0% /dev/shm
              5,0M 4,0K 5,0M 1% /run/lock
tmpfs
tmpfs
              980M 0 980M 0% /sys/fs/cgroup
              55M 55M 0 100% /snap/core18/1880
/dev/loop0
              30M 30M 0 100% /snap/snapd/8542
/dev/loop3
/dev/loop1
              63M 63M 0 100% /snap/gtk-common-themes/1506
/dev/loop2
             256M 256M 0 100% /snap/gnome-3-34-1804/36
                   50M 0 100% /snap/snap-store/467
/dev/loop4
              50M
/dev/sda1
              511M 4,0K 511M 1% /boot/efi
              44M 44M 0 100% /snap/snapd/14295
/dev/loop5
/dev/loop6
              56M 56M 0 100% /snap/core18/2253
              196M 40K 196M 1% /run/user/1000
tmpfs
/dev/loop7
              62M 62M 0 100% /snap/core20/1270
/dev/loop8
             128K 128K 0 100% /snap/bare/5
              55M 55M 0 100% /snap/snap-store/558
/dev/loop9
/dev/loop10
               66M
                   66M 0 100% /snap/gtk-common-themes/1519
              219M 219M 0 100% /snap/gnome-3-34-1804/77
/dev/loop11
                           0 100% /snap/gnome-3-38-2004/87
/dev/loop12
              248M 248M
```

• Create a new user "user01" on your machine

```
student@student-virtual-machine:~/Secret$ sudo adduser user01
[sudo] password for student:
Adding user `user01' ...
Adding new group `user01' (1001) ...
Adding new user `user01' (1001) with group `user01' ...
Creating home directory `/home/user01' ...
Copying files from `/etc/skel' ...
New password: *****
Retype new password: *****
passwd: password updated successfully
Changing the user information for user01
Enter the new value, or press ENTER for the default
    Full Name []: user
   Room Number []:
   Work Phone []:
   Home Phone []:
    Other []:
Is the information correct? [Y/n] Y
student@student-virtual-machine:~/Secret$
```

• Try to find out which access rights "user01" has on directory /etc You might need the commands id or groups

user01 has only read access, he is "everybody" in the permissions Ownership of /etc is group and user root.

• Shut down your system with a delay/timer of 2 minutes. Make sure the command runs in the background so that it does not block your shell.

```
student@student-virtual-machine:~/Secret$ sudo shutdown -h 2
[sudo] password for student:
Shutdown scheduled for Sat 2021-12-18 11:41:14 CET, use 'shutdown -c' to cancel.
```

• Oh no! You reconsidered: Abort the shutdown using ps, kill,

```
student@student-virtual-machine:~/Secret$ sudo shutdown -c
```

Task 2

Now, download the 'linux-labs.tar.gz' ZIP archive from eCampus and extract the files to your home directory. If you do it from the host and not directly from inside the VM's browser, copy the file into your machine via the VM's "shared folders" functionality.

Part 1: Basic navigation

Follow below instructions and document the commands used. Useful commands for this exercise are tar, cd, ls, mkdir, grep, mv, cat, more, less

- Extract the archive with the tar command
- · Change to the directory with the extracted files and go to the directory "nav"
- Extract the archive within and list the now created directories
- Use grep to determine which files in the "output" subdirectories are about "caffeine". Remember: You can read entire files with cat, more, less or an editor of your choice
- There are only three "result*" directories. Create "result4" and "result5" and move "out_4.pdb" and "out_4.pdb" to those directories

```
student@student-virtual-machine:~$ mv Downloads/linux-labs.tar.gz ~
student@student-virtual-machine:~$ ls ~
```

```
Desktop Documents Downloads home linux-labs.tar.gz Music Pictures Public
Secret Templates Videos
student@student-virtual-machine:~$ tar -xf linux-labs.tar.gz
student@student-virtual-machine:~$ ls ~
Desktop Documents Downloads home linux-labs.tar.gz logs Music nav Pictures
Public Secret Templates Videos#
student@student-virtual-machine:~$ cd nav/
student@student-virtual-machine:~/nav$
student@student-virtual-machine:~/nav$ tar -xf directories.tar.gz
student@student-virtual-machine:~/nav$ ls ~/nav
directories.tar.gz inputs linux-labs.tar.gz logs nav outputs
student@student-virtual-machine:~/nav/outputs$ for file in */*; do echo $file; cat
$file | grep caffeine ; done
result1/out_1.pdb
result2/out_2.pdb
result3/out_3.pdb
TITLE
         caffeine
result3/out_4.pdb
result3/out_5.pdb
result3/out_6.pdb
result3/out_7.pdb
student@student-virtual-machine:~/nav/outputs$ mkdir result4
student@student-virtual-machine:~/nav/outputs$ mkdir result5
student@student-virtual-machine:~/nav/outputs$ mv result3/out_4.pdb result4/out_4.pdb
student@student-virtual-machine:~/nav/outputs$ mv result3/out_5.pdb result5/out_5.pdb
student@student-virtual-machine:~/nav/outputs$ ls -1 */*
result1/out_1.pdb
result2/out_2.pdb
result3/out_3.pdb
result3/out_6.pdb
result3/out_7.pdb
result4/out_4.pdb
result5/out_5.pdb
#We assumed, that we should move out_4 and 5 and not out_4 and 4#
```

Part 2: Manual pages and directory listings

Let's take a look at the man pages and the 1s command. Use man 1s to open the manual page for the 1s command. You can search the page by using "/" followed by the keyword and . Jump between results by pressing "n". "q" closes the man page.

- Find the flag of ls that sorts files by size
- Go to nav/inputs and list the files by their size

```
student@student-virtual-machine:~/nav/inputs$ ls -laS
total 48
```

```
      drwx-----
      2 student
      student
      4096
      Mai
      5
      2017
      .

      drwxrwxrwx
      4 student
      student
      4096
      Dez
      18
      12:03
      .

      -rw------
      1 student
      student
      2795
      Sep
      17
      2015
      inp_3.pdb

      -rw------
      1 student
      student
      1514
      Sep
      17
      2015
      inp_2.pdb

      -rw------
      1 student
      student
      1230
      Sep
      17
      2015
      inp_6.pdb

      -rw-------
      1 student
      student
      1173
      Sep
      17
      2015
      inp_5.pdb

      -rw-------
      1 student
      student
      1021
      Sep
      17
      2015
      inp_7.pdb

      -rw-------
      1 student
      student
      578
      Sep
      17
      2015
      inp_1.pdb

      -rwxrwxrwx
      1 student
      student
      13
      0kt
      30
      2018
      hallolo.pdb

      -rwxrwxrwx
      1 student
      student
      11
      0kt
      30
      2018
      1234567.xyz
```

• Repeat the same for reserve order (smallest file first)

• Change the output to use human-readable file sizes (KB, MB, GB)

```
student@student-virtual-machine:~/nav/inputs$ ls -laSr --block-size=K
total 48K
-rwxrwxrwx 1 student student 1K Okt 30 2018 1234567.xyz
-rwxrwxrwx 1 student student 1K Okt 30 2018 hallolo.pdb
-rwxrwxrwx 1 student student 1K Okt 30 2018 hallolo.txt
-rw----- 1 student student 1K Sep 17 2015 inp_1.pdb
-rw----- 1 student student 1K Sep 17 2015 inp_7.pdb
-rw----- 1 student student 2K Sep 17 2015 inp_5.pdb
-rw----- 1 student student 2K Sep 17 2015 inp_4.pdb
-rw----- 1 student student 2K Sep 17 2015 inp_6.pdb
-rw----- 1 student student 2K Sep 17 2015 inp_2.pdb
-rw----- 1 student student 3K Sep 17 2015 inp_2.pdb
-rw----- 1 student student 3K Sep 17 2015 inp_3.pdb
drwxrwxrwx 4 student student 4K Dez 18 12:03 ..
drwx----- 2 student student 4K Mai 5 2017 .
```

• Now, list all files in the inputs directory that contain an "a"

```
student@student-virtual-machine:~/nav/inputs$ ls -1 | grep a
hallolo.pdb
hallo.txt
```

• List files that have a name with exactly seven characters and end in ".pdb"

```
student@student-virtual-machine:~/nav/inputs$ ls -1 | grep ???????.pdb hallolo.pdb
```

• Go to the "outputs" directory and list all ".pdb" files in its subdirectories

```
student@student-virtual-machine:~/nav/outputs$ ls */* | grep .pdb
result1/out_1.pdb
result2/out_2.pdb
result3/out_3.pdb
result3/out_6.pdb
result3/out_7.pdb
result4/out_4.pdb
result5/out_5.pdb
```

• Now limit the list to only those out-files that have a number 2, 3, 4, or 5 in their name

```
student@student-virtual-machine:~/nav/outputs$ for file in */*; do if basename $file |
grep -qc "2\|3\|4\|5"; then echo $file; fi ; done
result2/out_2.pdb
result3/out_3.pdb
result4/out_4.pdb
result5/out_5.pdb
```

Part 3: Data extraction

Have a look at the /logs directory to see a list of files from chemistry computations. We will need grep, tar, cat, cut, more, less, awk, sed, wc, and sort in this exercise.

- Take a look at some of the files using cat, more, less, nano, vi, vim, or any other command line text editor
- Let's focus on "comp.log". The final energy per run is denoted in lines that look like this: "SCF Done: E(RPBE-PBE) = ..."
 - Extract all those lines with grep and write them directly into a new file

```
student@student-virtual-machine:~/logs$ cat comp.log | grep "SCF Done: E(RPBE-PBE) ="
> comp_filtered.log
student@student-virtual-machine:~/logs$ cat comp_filtered.log
SCF Done: E(RPBE-PBE) = -152.637052486 A.U. after 12 cycles
SCF Done: E(RPBE-PBE) = -152.649944913
                                          A.U. after 11 cycles
SCF Done: E(RPBE-PBE) = -152.654913859
                                          A.U. after 12 cycles
SCF Done: E(RPBE-PBE) = -152.656483798
                                         A.U. after 11 cycles
SCF Done: E(RPBE-PBE) = -152.657152494
                                         A.U. after 11 cycles
SCF Done: E(RPBE-PBE) = -152.657682699
                                         A.U. after 11 cycles
SCF Done: E(RPBE-PBE) = -152.658087670
                                          A.U. after 12 cycles
SCF Done: E(RPBE-PBE) = -152.658642047
                                        A.U. after 12 cycles
SCF Done: E(RPBE-PBE) = -152.659440588
                                         A.U. after 12 cycles
SCF Done: E(RPBE-PBE) = -152.660069288
                                          A.U. after 12 cycles
SCF Done: E(RPBE-PBE) = -152.660165007
                                          A.U. after 10 cycles
```

```
      SCF Done:
      E(RPBE-PBE) = -152.660185016
      A.U. after 10 cycles

      SCF Done:
      E(RPBE-PBE) = -152.660221654
      A.U. after 11 cycles

      SCF Done:
      E(RPBE-PBE) = -152.660243827
      A.U. after 10 cycles

      SCF Done:
      E(RPBE-PBE) = -152.660246491
      A.U. after 8 cycles

      SCF Done:
      E(RPBE-PBE) = -152.660246751
      A.U. after 7 cycles

      SCF Done:
      E(RPBE-PBE) = -152.660246843
      A.U. after 6 cycles
```

 Remove all text and leave only the number. You can use e.g. sed for that purpose

```
student@student-virtual-machine:~/logs$ cat comp.log | grep "SCF Done: E(RPBE-PBE) ="
-152.637052486 12
-152.649944913 11
-152.654913859 12
-152.656483798 11
-152.657152494 11
-152.657682699 11
-152.658087670 12
-152.658642047 12
-152.659440588 12
-152.660069288 12
-152.660165007 10
-152.660185016 10
-152.660221654 11
-152.660243827 10
-152.660246491 8
-152.660246751 7
-152.660246843 6
```

 \bullet We are interested in hydrogen atoms. Use grep and the wc command (together; think pipes) to count how many lines in the file "quickstep.report" start with "H"

```
student@student-virtual-machine:~/logs$ cat quickstep.report | grep " H" | wc 1024 4096 66560
```

• What is the smallest x-coordinate in "quickstep.report"? Use e.g. sort and head.

```
student@student-virtual-machine:~/logs$ cat quickstep.report | grep " H" | sort |
head -n 5
          -0.0458740000
                             21.3572900000
 Н
                                                  6.8759690000
                              4.2880610000
  Н
           0.0614840000
                                                  11.2611580000
 Н
          -0.0748410000
                               0.8127580000
                                                  0.0296070000
 Н
           0.1497510000
                               6.9271230000
                                                  21.0598550000
           0.2771940000
                              21.2327840000
                                                  20.7373030000
 Н
```

- There are temperatures in the file "quickstep.out".
 - Extract the lines into a new file and sort them by value (largest to smallest)

```
student@student-virtual-machine:~/logs$ cat quickstep.out | grep "TEMPERATURE\
extract_temp.txt
student@student-virtual-machine:~/logs$ cat extract_temp.txt
 TEMPERATURE [K]
                                              368.690
                                                                348.451
                                                                347.958
 TEMPERATURE [K]
                                              344.999
 TEMPERATURE [K]
                                              377.405
                                                                344.403
 TEMPERATURE [K]
                                              314.601
                                                                343.788
 TEMPERATURE [K]
                                              287.757
                                                                337.563
 TEMPERATURE [K]
                          =
                                              368.720
                                                                336.153
 TEMPERATURE [K]
                                              272.845
                                                                331.091
 TEMPERATURE [K]
                                              347.442
                                                                325.297
 TEMPERATURE [K]
                                              323.124
                                                                314.225
 TEMPERATURE [K]
                                              305.326
                                                                305.326
 INITIAL TEMPERATURE[K]
                                                                300.000
```

• Bonus: Use grep and awk to calculate a cumulative sum of all the extracted temperatures.

```
student@student-virtual-machine:~/logs$ temperatur=0;for line in $(cat quickstep.out |
grep -n "TEMPERATURE\[K\]\|TEMPERATURE \[K\]" | sort | awk 'NF>1\{print $NF\}');do
temperatur=$(echo - | awk "{print $temperatur + $line}" | sed 's/,/./g' );echo
$temperatur; done; echo "Alle Durchschnittstemperaturen summiert sind "$temperatur"K."
300
605.326
919.551
1244.85
1581
1925.4
2273.85
2621.81
2965.6
3303.16
3634.25
Alle Durchschnittstemperaturen summiert sind 3634.25K.
```

Part 4: Data cleanup

Now for something else. Use the command line to download the file "ftp://mirbase.org/pub/mirbase/CURRENT/genomes/hsa.gff3" with wget. To clean up the file, do the following:

- Remove all comment lines (starting with "#")
- Remove all lines with the tag "miRNA_primary_transcript"
- Change chromosome names (1st column) from format "chr1"," chr2", ... to format "1", "2", ...
- Remove the second column with cut

New Link: https://www.mirbase.org/ftp/CURRENT/genomes/hsa.gff3

```
student@student-virtual-machine:~$ wget
https://www.mirbase.org/ftp/CURRENT/genomes/hsa.gff3
--2021-12-18 12:30:36-- https://www.mirbase.org/ftp/CURRENT/genomes/hsa.gff3
```

```
Resolving www.mirbase.org (www.mirbase.org)... 130.88.97.249
Connecting to www.mirbase.org (www.mirbase.org)|130.88.97.249|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 534796 (522K) [text/plain]
Saving to: 'hsa.gff3'
hsa.gff3
                                                                                                                                                       100%
522,26K 1,52MB/s
                                                                      in 0,3s
2021-12-18 12:30:37 (1,52 MB/s) - 'hsa.gff3' saved [534796/534796]
student@student-virtual-machine:~$ sed '/#/d' hsa.gff3 > hsa.gff3_without_#
student@student-virtual-machine:~$ sed '/'miRNA_primary_transcript'/d' hsa.gff3_
without_# > hsa-gff3_WITHOUTprimary
student@student-virtual-machine:~$ sed 's/chr//g' hsa.gff3_ WITHOUTprimary > hsa-
gff3_WITHOUTchr
student@student-virtual-machine: \verb|--| student| a.gff3\_ without\_withoutchr > hsa-like a.gff3\_ without\_withoutchr > hsa-like a.gff3\_ withoutchr > hsa-like 
qff3 WITHOUTcol2
```

Task 3: Bash scripting

Finally, write a short shell script that automates the steps in Task 4, Part 4. Supply the file to be downloaded as command line argument (you add the URL right after the script name when starting it, not in the code). Also include a check in the beginning that tests if the downloaded file is empty. If it is empty, print out the line "Error: File is empty!" instead of starting the cleanup process.

Note: Shell scripts for bash start with the shebang "#!/bin/bash" – this is a kernel convention and helps identify your code. Name the file "cleanup.sh". When starting it, use "./cleanup.sh". This tells the shell that you only want to look in the current directory, not the PATH

```
sed /#/d tmp.$$.download > tmp.$$.#
sed /miRNA_primary_transcript/d tmp.$$.# > tmp.$$.prim
sed s/chr//g tmp.$$.prim > tmp.$$.chr
cut -f 1,3- tmp.$$.chr > ~/hsa.gff3.Done
rm tmp.$$.# tmp.$$.prim tmp.$$.chr tmp.$$.download
```