**sDSCC 201/401 Homework Assignment #3**

**Due: September 29, 2021 at 9 a.m. EDT**

**Answers to these questions should be submitted via Blackboard. Please upload a file**

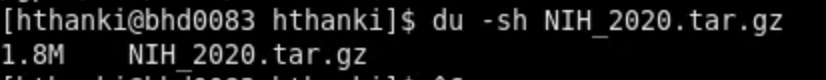
**containing your answers and explanations to Blackboard (Homework #3: Linux Environment) as a Word document (docx), text file (txt), or PDF. All input and output should be plain text (cut and paste from a terminal window) consolidated into one file, with your name provided at the top. It is very important to include the command you used to give the answer to the question. Credit will only be given if the command to answer the question and the output have been provided. For each of the following questions 1-8 provide the input command and the output**

**Generated.**

**1. Download the file NIH\_2020.tar.gz from Blackboard. Upload the file to your scratch directory on BlueHive. What is the total amount of disk space used by the NIH\_2020.tar.gz file that you copied to BlueHive? Show the command how you determined the size of the file.**

Command: $ du -sh NIH\_2020.tar.gz

Output: 1.8M NIH\_2020.tar.gz



**2. Submit an interactive job to the debug partition on BlueHive to obtain a compute node**

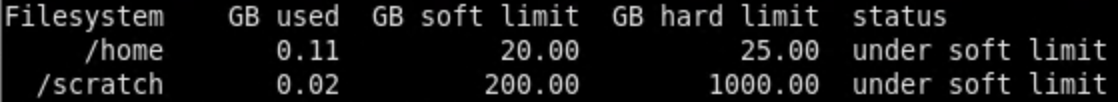
**with 1 core and 2 GB of RAM for 1 hour using a terminal session. What command did**

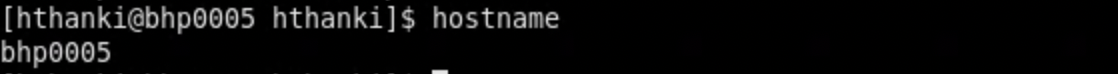
**you type to obtain these resources? What is the hostname of the compute node you are**

**Using?**

Command: $ interactive -p debug -t 1:00:00 -c 1 --mem-per-cpu=2GB

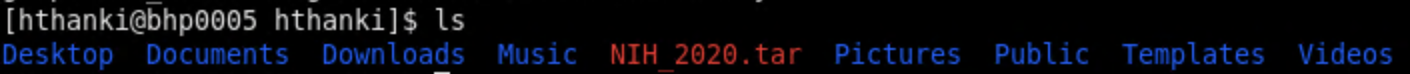
Hostname of compute node: bhp0005



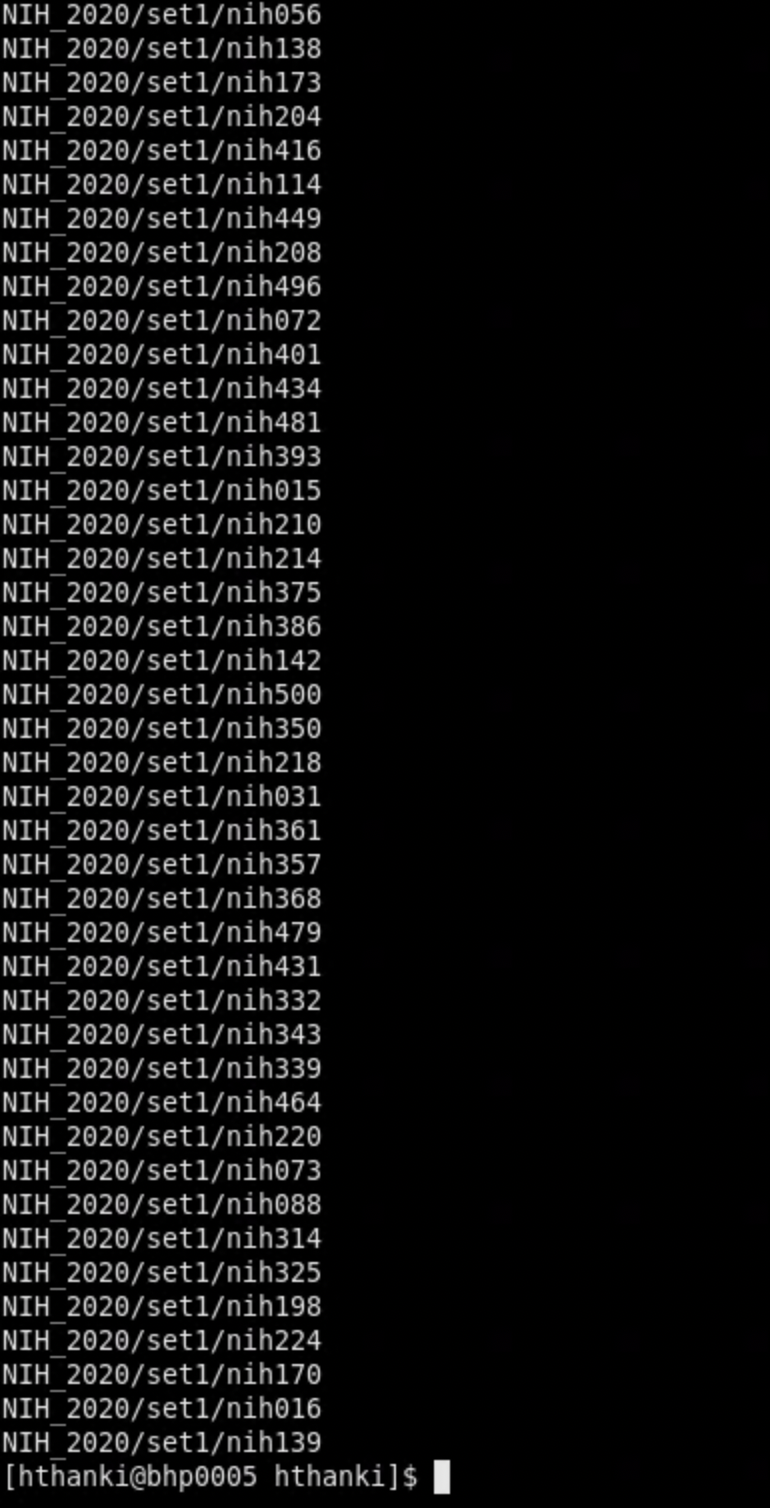


**3. Read the manual pages for the gunzip and tar commands. Uncompress the NIH\_2020.tar.gz file with the gunzip command and then subsequently unarchive the NIH\_2020.tar file while logged into the compute node. Show the commands you used to uncompress and unarchive the file.**

Command to uncompress the file: $ gunzip NIH\_2020.tar.gz



Command to unarchive the file: $ tar -xvf NIH\_2020.tar

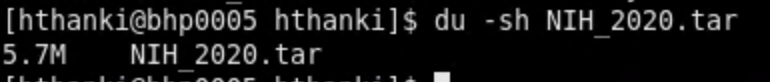


**4. What is the combined total size of all files (amount of storage on disk) in megabytes**

**located in the NIH\_2020 directory? Show the command you used to calculate this value.**

Command: $ du -sh NIH\_2020.tar

Output: 5.7M NIH\_2020.tar.gz

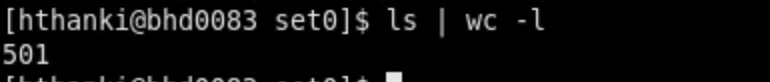


**5. How many files are located in the set1 directory? Show the command you used to**

**calculate this value.**

Command: $ ls | wc -l

Output: 501



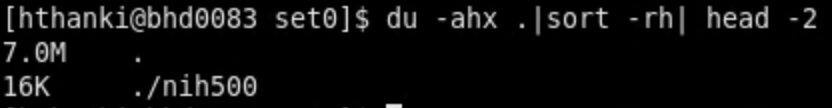
**6. Using the find command, what is the name of the single largest file located in the**

**set0 directory? What is the size of this file? Show the commands you used to determine the answer.**

Command: $ du -ahx .| sort -rh | head -2

Output: 7.0M .

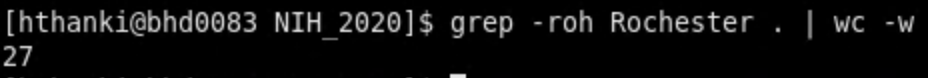
16K ./NIH500



**7. How many lines contain the string "Rochester" in all files in the NIH\_2020 directory and all of its subdirectories? Show the command you used to calculate this value.**

Command: grep -roh Rochester . | wc -w

Output: 27

****

**8. All files in the set0 and set1 subdirectories of the NIH\_2020 directory are CSV (comma-separated values) files. Write a bash script to execute in the NIH\_2020 directory that will add the '.csv' extension to all file names in the set0 and set1directories. For example, nih200 should become nih200.csv. (DSCC 401 Only: Convert lowercase nih file names to uppercase NIH file names. For example, nih200.csv should become NIH200.csv, while set0 and set1 remain**

**lowercase.)**

Command for converting filenames to uppercase: find .-type f-execdir rename ‘y/a-z/A-Z/‘{}\;

Command for adding .csv extensions to all files: -nv -- 's/$/.csv/' \*

**For questions 9, 10, and 11, create three separate files for each of the three Bash script**

**shown below. Run the scripts on BlueHive, examine the outputs, and answer the following questions:**

**#!/bin/bash**

**# script 1**

**pi=3.14159**

**two\_pi=2\*$pi**

**echo $two\_pi**

**#!/bin/bash**

**# script 2**

**pi=3.14159**

**two\_pi=$((2\*pi))**

**echo $two\_pi**

**#!/bin/bash**

**# script 3**

**pi=3**

**two\_pi=$((2\*pi))**

**echo $two\_pi**

**9. Show the outputs of script 1, script 2, and script 3.**

Output for script1: 2\*3.14159

Output for script2: syntax error: Invalid arithmetic operator (error token is “.14159”)

Output for script3: 6

**10. What is the purpose of the # in the Bash scripts?**

# is used for commenting in a Bash script. However, #! is used to tell your system that this file is a set of commands to be fed to the command interpreter indicated.

**11. What is the purpose of the double parentheses (( )) in the scripts?**

The (( )) construct permits arithmetic expansion and evaluation.