

SLURM EXERCICES

Laia Barcenilla Mañá

Username: biohpc-31

NIA: 107694

laia.barcenilla@alum.esci.
upf.edu

Jan Carreras Boada

Username: biohpc-32

NIA: 107765

jan.carreras@alum.esci.upf
.edu

15/05/2024

2023/2024

Third term

2nd Year Bioinformatics

1. Write a SLURM job that solves the following problem:

We have several text files names_1.txt, names_2.txt,... For each file we want to generate a new file in upper case letters (with a different extension). (Hint: you can use the tr command to perform this action: i.e. tr [:lower:][:upper:]).

```
#!/bin/bash
#SBATCH --job-name=convert_to_upper
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=1
#SBATCH --partition=nodo.q

for file in names_*.txt; do
    newname=$(basename "$file" .txt)
    upername=$(echo "$newname" | tr '[:lower:]' '[:upper:]')

    tr '[:lower:]' '[:upper:]' < "$file" > "${uxername}.txt"
done
```

This code iterates over the filenames. For each one, it gets its name and converts it to uppercase. After that, it translates the content of the file and stores it in a file named after the variable previously created.

NAMES_0.txt	NAMES_1.txt	NAMES_2.txt	NAMES_3.txt	NAMES_4.txt
SERENITY AUTUMN ADELINE HAILEY GIANNA VALENTINA ISLA ELIANA QUINN NEVAEH IVY SADIE PIPER LYDIA ALEXA JOSEPHINE EMERY JULIA DELILAH ARIANNA VIVIAN KAYLEE SOPHIE BRIELLE MADELINE	EMMA OLIVIA AVA ISABELLA SOPHIA CHARLOTTE MIA AMELIA HARPER EVELYN ABIGAIL EMILY ELIZABETH MILA ELLA AVERY SOFIA CAMILA ARIA SCARLETT VICTORIA MADISON LUNA GRACE CHLOE	PENELOPE LAYLA RILEY ZOEY NORA LILY ELEANOR HANNAH LILLIAN ADDISON AUBREY ELLIE STELLA NATALIE ZOE LEAH HAZEL VIOLET AURORA SAVANNAH AUDREY BROOKLYN BELLA CLAIRE SKYLAR	LUCY PAISLEY EVERLY ANNA CAROLINE NOVA GENESIS EMILIA KENNEDY SAMANTHA MAYA WILLOW KINSLEY NAOMI AALIYAH ELENA SARAH ARIANA ALLISON GABRIELLA ALICE MADELYN CORA RUBY EVA	AUBREY OSCAR KAJOL SOFIA HARPER AMRITPAL ALESSIO JASMINE ZAID KOBI LEONA TEDDIE DELLA KAVAN LETICIA LAIBA SAIRA HADI LAWSON RAIFE TIANNA MICAH BLADE OLIVIA HAYLEY

2. Write a script (either in bash or Python) that submits a workflow that carries out the following operation with the names_*.txt:

Generates upper case versions of each file (as in the previous exercise).

All files (original and upper case) are sorted alphabetically (Hint: use the sort command from Linux).

A summary file with the first and last name of each file is generated using the script summary.sh (available at Escritorio/alumnos/SLURM/). Take a look at it and run it at the command line as a stand-alone command to understand how it works.

summary.sh file_1 file_2file_N

```
#!/bin/bash
#SBATCH --job-name=array-job
#SBATCH --nodes=1
#SBATCH --ntasks-per-node=1
#SBATCH --partition=nodo.q
#SBATCH --array=0-4

# GET THE CORRESPONDENT FILE
Taskid=$SLURM_ARRAY_TASK_ID
file=$(ls names_*.txt | sed -n "$((Taskid + 1))p")

# TRANSFORM THE NAME AND THE CONTENT OF THE FILE TO UPPERCASE
newname=$(basename "$file" .txt)
uppername=$(echo "$newname" | tr '[:lower:]' '[:upper:]')
tr '[:lower:]' '[:upper:]' < "$file" > "${uppername}.txt"

# SORT THE CONTENT OF THE FILE
sort "${uppername}.txt" -o "${uppername}.txt"

# MAKE A SUMMARY
./summary.sh "${uppername}.txt" > "summary_${Taskid}.txt"

# MERGE SUMMARIES
cat "summary_${Taskid}.txt" >> summary.txt

# REMOVE UNNECESSARY FILES
rm "summary_${Taskid}.txt"q

file: NAMES_0.txt
ADELINE
VIVIAN
*****
file: NAMES_1.txt
ABIGAIL
VICTORIA
*****
file: NAMES_3.txt
AALIYAH
WILLOW
*****
file: NAMES_2.txt
ADDISON
ZOEY
*****
file: NAMES_4.txt
ALESSIO
ZAID
*****
```

In this code, we are processing each file from 0 to 4, so each task will handle a file. We create a variable `newname` that contains the name of the file and transforms it to uppercase. Then, the content of the file is also transformed. After that, we sort the file and generate the summary file. All these summaries are stored in a new variable (which will be removed later). Finally, we merge all the summaries into a file named `summary.txt`.

At the right we can see the summary.txt file, this code also creates files named NAMES_*.txt sorted alphabetically in uppercase letters as shown.

```
ADELINE
ALEXA
ARIANNA
AUTUMN
BRIELLE
DELILAH
ELIANA
EMERY
GIANNA
HAILEY
ISLA
IVY
JOSEPHINE
JULIA
KAYLEE
LYDIA
MADELINE
NEVAEH
PIPER
QUINN
SADIE
SERENITY
SOPHIE
VALENTINA
```