**CyberGIS and Big Data (GEOG6282/8282; INES 8090)**

**Title: Lab 2**

**Name: Tamim Adnan**

Question – 1:

Question 1: Use n CPUs (for example n = 5) to create 500 folders (you can modify the Shell scripts you made in lab1). Make sure each cpu only create 500/n folders. Create submission files and submit them one by one just like what we learned in Section 3.

* 1. **Shell Script (run.sh) for assigning 500/n folders to each cpus**

**(This script has been changed from lab 1, because assigning cpu needed local start and end, Also, create\_folder function was created).**

#!/bin/bash

# Define the number of CPUs to use

n=5

# Calculate the number of folders each CPU should create

num\_folders= 100 ## 500 / n folder in each CPU

# Define the function to create folders for a given range of indices

## start and end are two arguments.

## call the start and end variable as local variable with local.

## the current working directory has defined in pwd, mkdir for creating folder and t

create\_folders() {

local start=$1

local end=$2

for ((i=$start; i<=$end; i++))

do

echo "$(pwd)/$i"

mkdir -p "$i"

echo "$(pwd)" > "$i/text.txt"

done

}

# Loop over the CPUs and submit jobs to create folders

for ((i=1; i<=n; i++))

do

# Assign each CPUs numbers of folders to create

start=$((num\_folders\*(i-1) + 1))

end=$((num\_folders\*i))

# Submit the job to create folders for the current range of indices

create\_folders "$start" "$end" &

done

# Wait for all jobs to finish

wait

* 1. **Shell Script for (mkfiles.sh) the task of each CPU (0-100), (101 - 200), (201-300), (301-400) and (401-500)**

#! /bin/bash

for ((i=0;i<5;i++ ))

do

cat submit\_template.sh > submit$i.sh

echo "/users/tadnan/lab2/problems/problem\_1/run.sh $i" >> submit$i.sh

done

* 1. **Shell Script (submit\_template.sh) to submit the job.**

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=1

#SBATCH --ntasks-per-node=1

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================" echo ""

cd $SLURM\_SUBMIT\_DIR

echo "Hello World! I ran on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================" echo "End Time : $(date)"

echo "======================================================"

# Main program

* 1. **Shell Script (submitAll.sh)**

#!/bin/bash

for ((i=0;i<5;i++ ))

do

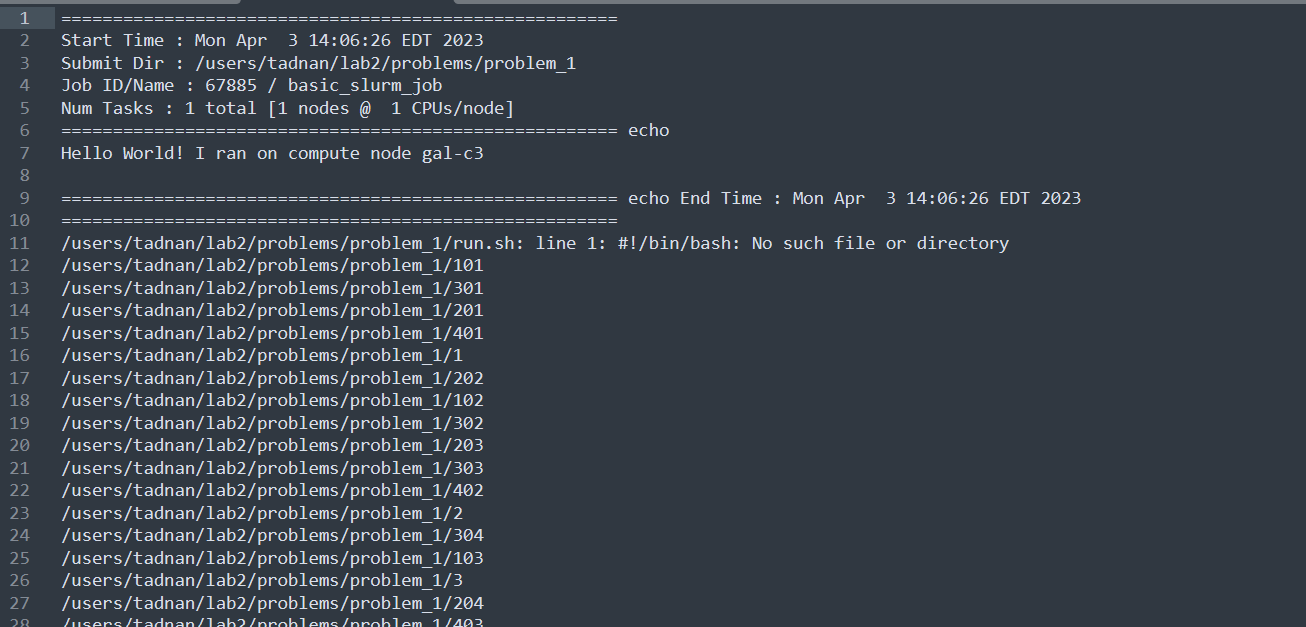
sbatch submit$i.sh

done

* 1. **Outputs**

Table

Description automatically generated



**Question-2: Write Shell scripts to solve Question 1, but use one single submission file to coordinate multiple CPUs.**

**2.1. Same run.sh file from question -1**

#!/bin/bash

# Define the number of CPUs to use

n=5

# Calculate the number of folders each CPU should create

num\_folders= 100 ## (500 / n) folder in each CPU

# Define the function to create folders for a given range of indices

## start and end are two arguments.

## call the start and end variable as local variable with local.

## the current working directory has defined in pwd, mkdir for creating folder and t

create\_folders() {

local start=$1

local end=$2

for ((i=$start; i<=$end; i++))

do

echo "$(pwd)/$i"

mkdir -p "$i"

echo "$(pwd)" > "$i/text.txt"

done

}

# Loop over the CPUs and submit jobs to create folders

for ((i=1; i<=n; i++))

do

# Assign each CPUs numbers of folders to create

start=$((num\_folders\*(i-1) + 1))

end=$((num\_folders\*i))

# Submit the job to create folders for the current range of indices

create\_folders "$start" "$end" &

done

# Wait for all jobs to finish

wait

**2.2 Create a shell script for submit.sh.**

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=5

#SBATCH --ntasks-per-node=1

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

### print some information about the job

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================"

echo ""

### Here the script changes submission directory and printing information about nodes

cd $SLURM\_SUBMIT\_DIR

echo "Hi, I'm running my script on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================"

echo "End Time : $(date)"

echo "======================================================"

# Main Program

srun -n 5 --time=1:00:00 --partition=Centaurus /users/tadnan/lab2/problems/problem\_2/run.sh 5

wait

**2.3. Output**

Text

Description automatically generated with medium confidence

**Question-3: Write Shell scripts to run the agent-based fire spread model that you learn in this lab *n* times (*n*>=1,000) using *m* CPUs (*m*>=10).**

**This question is followed by section 3 in lab 2 where each file was submitted one by one.**

**3. 1. Shell Script for run.sh**

#!/bin/bash

for i in $(seq 0 1 1000)

do

netlogo='/users/tadnan/lab2/problems/netlogo-6.2.0/app/netlogo-6.2.0.jar'

model='/users/tadnan/lab2/problems/problem\_3/Fire.nlogo'

output="/users/tadnan/lab2/problems/problem\_3/test\_$i.csv"

java -Xmx1024m -Dfile.encoding=UTF-8 -cp $netlogo org.nlogo.headless.Main --model $model --experiment experiment1 --table $output

done

**3.2 Script for mkfiles.sh**

#! /bin/bash

n1=10

n2=100

start=0

end=$n2

for((i=0;i<$n1;i++))

do

cat submit\_template.sh > ./submit$i.sh

s=($i+1)\*$n2

e=$s+$n2

for ((j=$s;j<$e;j++))

do

echo "/users/tadnan/lab2/problems/problem\_3/run.sh $j" >> ./submit$i.sh

done

done

**3.3. Create submit\_template.sh.**

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=3

#SBATCH --ntasks-per-node=10

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================" echo ""

cd $SLURM\_SUBMIT\_DIR

echo "Hello World! I ran on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================" echo "End Time : $(date)"

echo "======================================================"

# Main program

3.4 **Shell Script for submitAll.sh**

#!/bin/bash

for ((i=0;i<10;i++ ))

do

sbatch submit$i.sh

done

* 1. **Outputs**

**Graphical user interface, table

Description automatically generated**

**Question 4**

**In the Information Tab of the fire simulation model, it is mentioned that 59% of tree density is a threshold for fire spread. Regarding this, design an experiment with a parallel Monte Carlo approach (#Monte Carlo runs for each treatment: not less than 100) to investigate the relationship between fire spread (towards right edge of the landscape) and tree density and write an experiment report (single space; at least 1 page). The experiment report should include your hypothesis, method, experiment, discussion (model results and computing performance) and conclusion parts.**

**Hypothesis:**

As mentioned, 59% of the tree density is considered as the threshold for spreading fires. This report will present Monte Carlo simulation on several density and comparison through a graphical analysis.

**Method:**

Net logo fire simulation is an efficient tool to show the relationship between the fire burning with respect to the tree density. The experiment was conducted with 40%, 59%, 65%, 80%, and 95% tree densities. With the behavior space tool in the Net logo, five experiments were created. Then each of them was loaded to the URC Cluster to run each experiment with slurm files. Each of the experiments has been ran 10 times and the test.csv file was overwritten to store the results. The purpose of the overwritten to find the best output of the experiment.

**Experiment**

There are some graphs that have been plotted from the generated excel sheet for each experiment.

Observation at 40% shows that with this density, fire will reach the right edge, but the lowest number doesn’t make any sense because it goes below zero. Then the values were set to 59%, which showed the tree density burned with a regular fluctuation and the fire reached the right edge of the forest.

In addition, tree density for 65%, 80%and 95% are almost similar that suddenly descended to the ground at the right edge of the graph even though 65% has more fluctuation in the middle.

**Discussion:**

It is clear that all of the simulation curves 65%, 80% and 95% don’t make any sense while 40% denotes values with regular ups and downs. However, 59% showed values with regular ups and downs. It means it that several points can be considered as threshold for different tasks.

**Shell Scripts for each density**

**For 40% Tree Density**

1. Create a run.sh file.

#!/bin/bash

netlogo='/users/tadnan/lab2/problems/netlogo-6.2.0/app/netlogo-6.2.0.jar'

model='/users/tadnan/lab2/problems/problem\_4/40/Fire.nlogo'

output='/users/tadnan/lab2/problems/problem\_4/40/test.csv'

java -Xmx1024m -Dfile.encoding=UTF-8 -cp $netlogo org.nlogo.headless.Main --model $model --experiment experiment1 --table $output

1. Create mkfiles.sh file

#! /bin/bash

for ((i=0;i<10;i++ ))

do

cat submit\_template.sh > submit$i.sh

echo "/users/tadnan/lab2/problems/problem\_4/40/run.sh $i" >> submit$i.sh

done

1. Create a Submit\_template.sh file

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=3

#SBATCH --ntasks-per-node=10

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================" echo ""

cd $SLURM\_SUBMIT\_DIR

echo "Hello World! I ran on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================" echo "End Time : $(date)"

echo "======================================================"

# Main program

1. Create submitAll.sh file

#!/bin/bash

for ((i=0;i<10;i++ ))

do

sbatch submit$i.sh

done

**For 59% Tree Density**

1. Create a run.sh file.

#!/bin/bash

netlogo='/users/tadnan/lab2/problems/netlogo-6.2.0/app/netlogo-6.2.0.jar'

model='/users/tadnan/lab2/problems/problem\_4/59/Fire.nlogo'

output='/users/tadnan/lab2/problems/problem\_4/59/test.csv'

java -Xmx1024m -Dfile.encoding=UTF-8 -cp $netlogo org.nlogo.headless.Main --model $model --experiment experiment1 --table $output

1. Create mkfiles.sh file

#! /bin/bash

for ((i=0;i<10;i++ ))

do

cat submit\_template.sh > submit$i.sh

echo "/users/tadnan/lab2/problems/problem\_4/59/run.sh $i" >> submit$i.sh

done

1. Create a Submit\_template.sh file

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=3

#SBATCH --ntasks-per-node=10

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================" echo ""

cd $SLURM\_SUBMIT\_DIR

echo "Hello World! I ran on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================" echo "End Time : $(date)"

echo "======================================================"

# Main program

1. Create submitAll.sh file

#!/bin/bash

for ((i=0;i<10;i++ ))

do

sbatch submit$i.sh

done

**For 65% Tree Density**

1. Create a run.sh file.

#!/bin/bash

netlogo='/users/tadnan/lab2/problems/netlogo-6.2.0/app/netlogo-6.2.0.jar'

model='/users/tadnan/lab2/problems/problem\_4/65/Fire.nlogo'

output='/users/tadnan/lab2/problems/problem\_4/65/test.csv'

java -Xmx1024m -Dfile.encoding=UTF-8 -cp $netlogo org.nlogo.headless.Main --model $model --experiment experiment1 --table $output

1. Create mkfiles.sh file

#! /bin/bash

for ((i=0;i<10;i++ ))

do

cat submit\_template.sh > submit$i.sh

echo "/users/tadnan/lab2/problems/problem\_4/65/run.sh $i" >> submit$i.sh

done

1. Create a Submit\_template.sh file

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=3

#SBATCH --ntasks-per-node=10

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================" echo ""

cd $SLURM\_SUBMIT\_DIR

echo "Hello World! I ran on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================" echo "End Time : $(date)"

echo "======================================================"

# Main program

1. Create submitAll.sh file

#!/bin/bash

for ((i=0;i<10;i++ ))

do

sbatch submit$i.sh

done

**For 80% Tree Density**

1. Create a run.sh file.

#!/bin/bash

netlogo='/users/tadnan/lab2/problems/netlogo-6.2.0/app/netlogo-6.2.0.jar'

model='/users/tadnan/lab2/problems/problem\_4/80/Fire.nlogo'

output='/users/tadnan/lab2/problems/problem\_4/80/test.csv'

java -Xmx1024m -Dfile.encoding=UTF-8 -cp $netlogo org.nlogo.headless.Main --model $model --experiment experiment1 --table $output

1. Create mkfiles.sh file

#! /bin/bash

for ((i=0;i<10;i++ ))

do

cat submit\_template.sh > submit$i.sh

echo "/users/tadnan/lab2/problems/problem\_4/80/run.sh $i" >> submit$i.sh

done

1. Create a Submit\_template.sh file

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=3

#SBATCH --ntasks-per-node=10

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================" echo ""

cd $SLURM\_SUBMIT\_DIR

echo "Hello World! I ran on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================" echo "End Time : $(date)"

echo "======================================================"

# Main program

1. Create submitAll.sh file

#!/bin/bash

for ((i=0;i<10;i++ ))

do

sbatch submit$i.sh

done

**For 95% Tree Density**

1. Create a run.sh file.

#!/bin/bash

netlogo='/users/tadnan/lab2/problems/netlogo-6.2.0/app/netlogo-6.2.0.jar'

model='/users/tadnan/lab2/problems/problem\_4/95/Fire.nlogo'

output='/users/tadnan/lab2/problems/problem\_4/95/test.csv'

java -Xmx1024m -Dfile.encoding=UTF-8 -cp $netlogo org.nlogo.headless.Main --model $model --experiment experiment1 --table $output

1. Create mkfiles.sh file

#! /bin/bash

for ((i=0;i<10;i++ ))

do

cat submit\_template.sh > submit$i.sh

echo "/users/tadnan/lab2/problems/problem\_4/95/run.sh $i" >> submit$i.sh

done

1. Create a Submit\_template.sh file

#! /bin/bash

#SBATCH --partition=Centaurus

#SBATCH --job-name=basic\_slurm\_job

#SBATCH --nodes=3

#SBATCH --ntasks-per-node=10

#SBATCH --time=1:00:00

##SBATCH --mem=30gb

##SBATCH --exclusive

##SBATCH --nodelist=str-c60

echo "======================================================"

echo "Start Time : $(date)"

echo "Submit Dir : $SLURM\_SUBMIT\_DIR"

echo "Job ID/Name : $SLURM\_JOBID / $SLURM\_JOB\_NAME"

echo "Num Tasks : $SLURM\_NTASKS total [$SLURM\_NNODES nodes @ $SLURM\_CPUS\_ON\_NODE CPUs/node]"

echo "======================================================" echo ""

cd $SLURM\_SUBMIT\_DIR

echo "Hello World! I ran on compute node $(/bin/hostname -s)"

echo ""

echo "======================================================" echo "End Time : $(date)"

echo "======================================================"

# Main program

1. Create submitAll.sh file

#!/bin/bash

for ((i=0;i<10;i++ ))

do

sbatch submit$i.sh

done