Carlos Ortega
CS385 HW3 Report

Summary

The purpose for this program is to try and find order in data files. The files that I used to test my program consisted of image files of different formats. Most of the image files however, were raw format. This was because jpeg and other similar image formats use compression techniques to lower the file size, meaning that for the purpose of the program so information may have been lost in these types of files. The program uses threads to analyze the file.

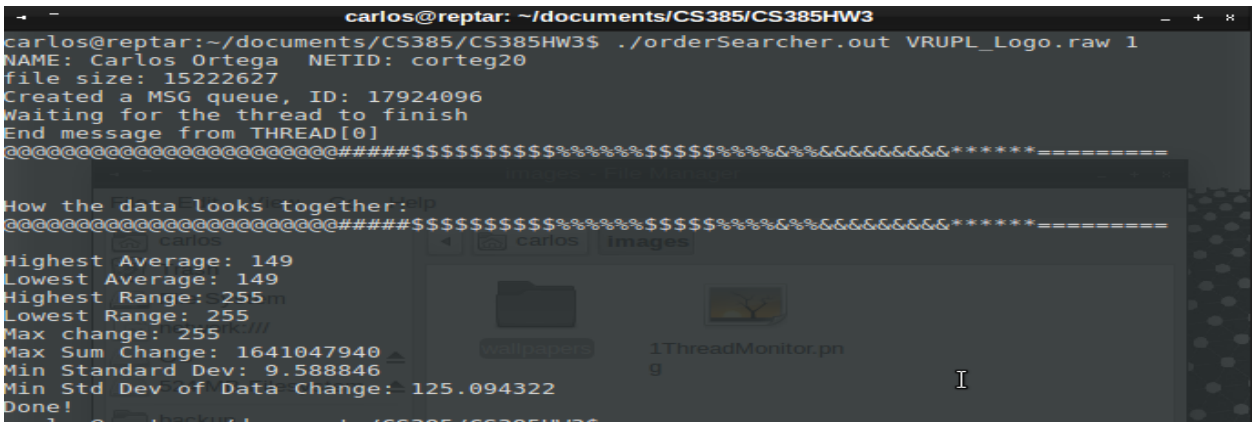
The file is read in as unsigned chars and then that data is divided as evenly as possible between the threads. This way each thread gets a “chunk” of data that it can then analyze. Using branch and bound techniques, each thread will find various parameters such a standard deviation of the data. Each thread also tries to visualize the data.

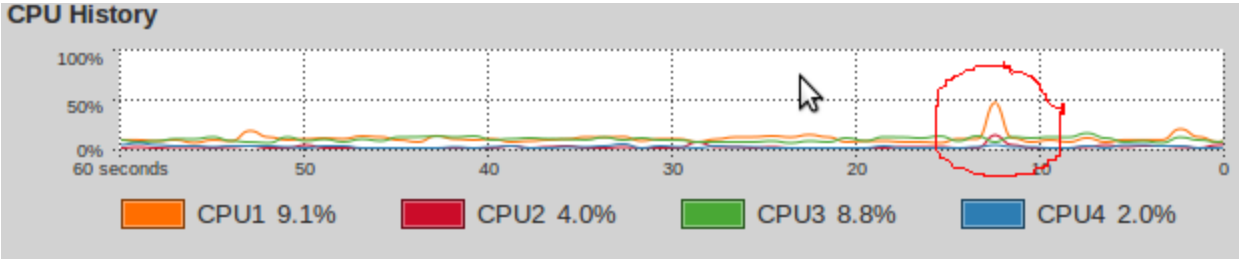
The way that visualization works in my program is that the data that each thread receives is divided by 80 to create smaller “chunks” of data. Then the average of each chunk is calculated. Based on that average a symbol is assigned to that chunk. Each individual thread then sends the symbols to a printing thread using a message queue. The printing thread then puts the symbol in order by thread the thread ID so that the first thread’s symbols appear first when they are printed.

Demonstration

The demo here is meant to show how the use of threads makes the completion of the program both faster, but also more reliable and accurate. For this demo I ran the program with numerous threads to analyze the VRUPL_Logo.raw file. Later I also used a python script that runs the program 20 times for each number of threads. For the script I also used the VRUPL_Logo.raw file.
These are the results:

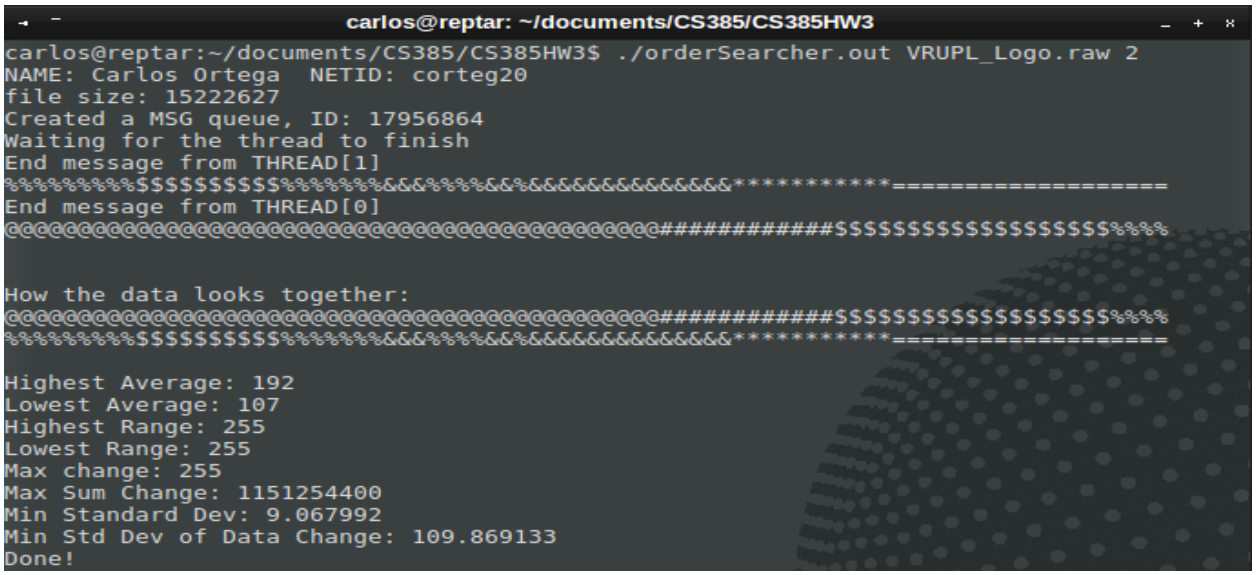
1 Thread

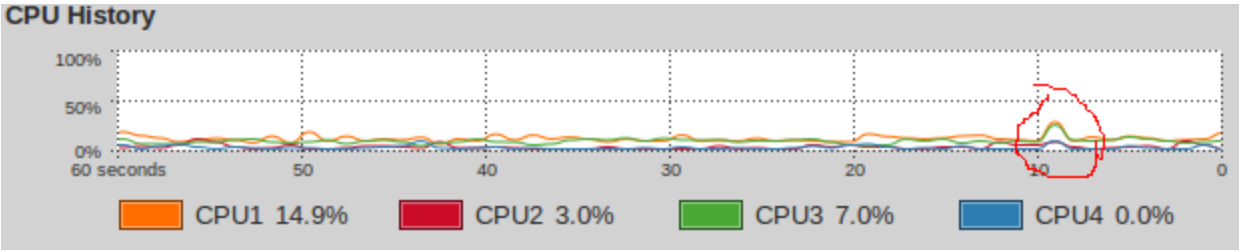




As you can see there is a significant spike in CPU1 when the program executes

2 Threads



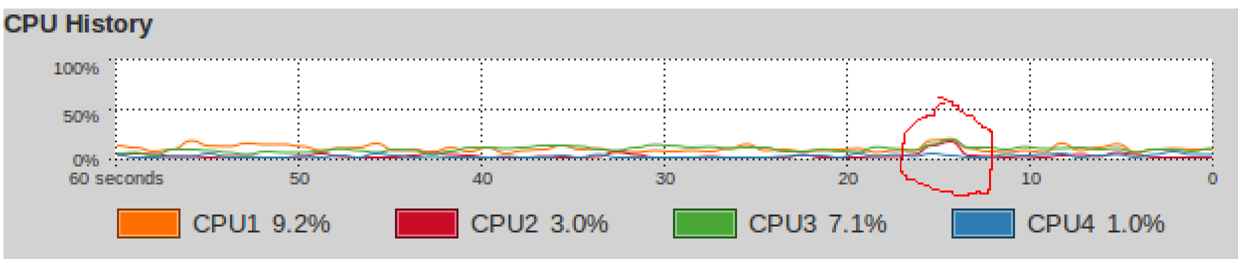


Here you can see that CPU1 and CPU3 have spiked in activity

Carlos Ortega
CS385 HW3 Report

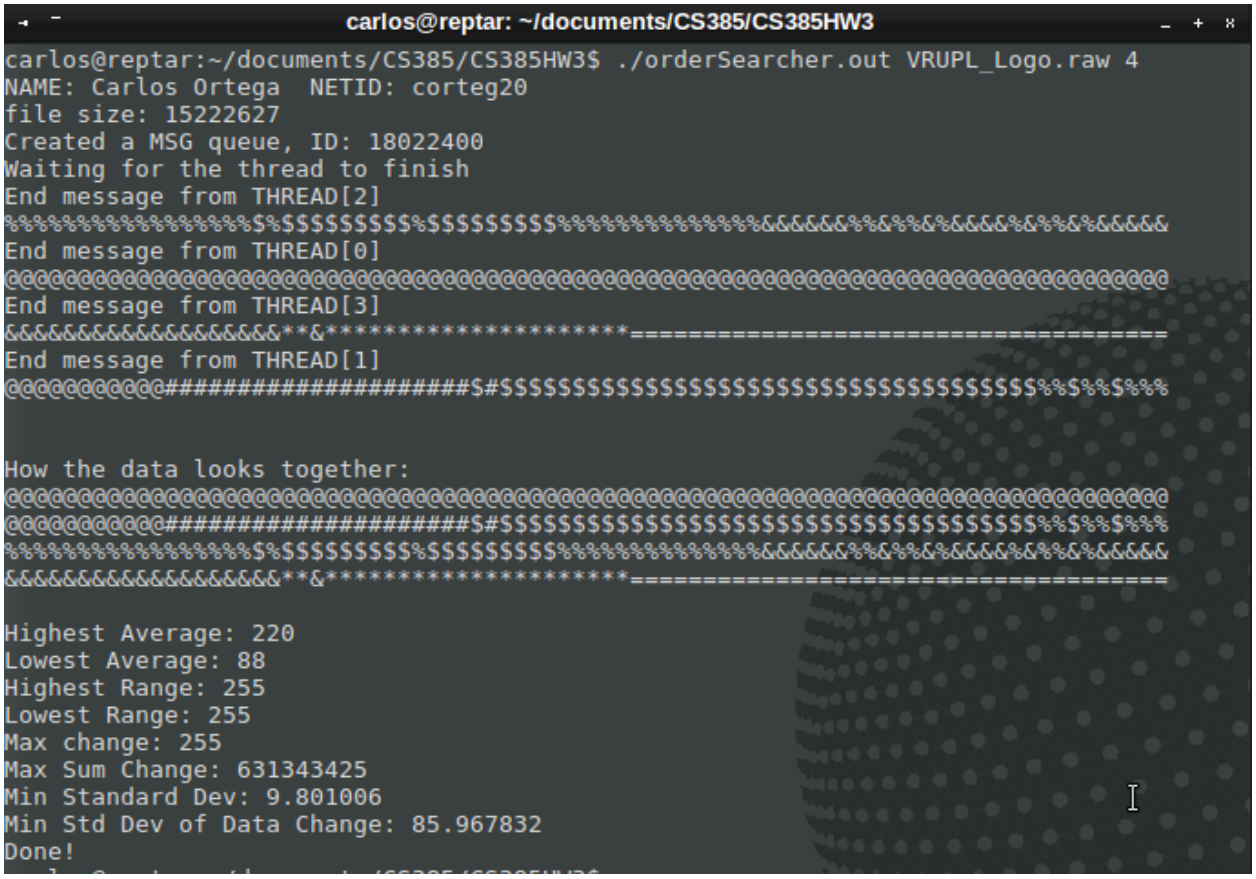
3 Threads

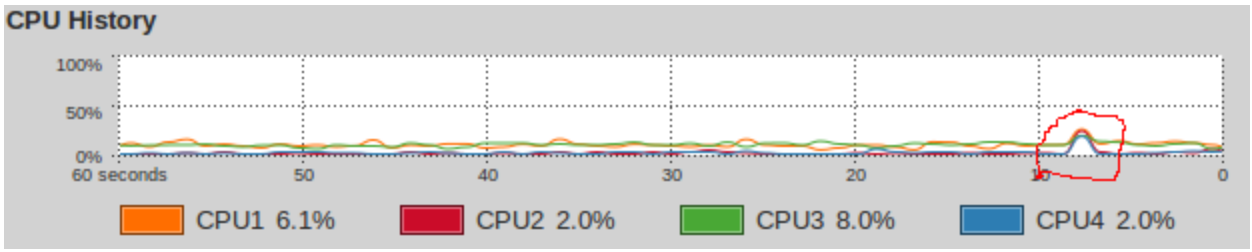
```
carlos@reptar: ~/documents/CS385/CS385HW3

carlos@reptar:~/documents/CS385/CS385HW3$ ./orderSearcher.out VRUPL_Logo.raw 3
NAME: Carlos Ortega  NETID: corteg20
file size: 15222627
Created a MSG queue, ID: 17989632
Waiting for the thread to finish
End message from THREAD[1]
#####$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$%%%%%%%%%%%%%%%%%%$$$$$$$$$$$$$$%%%%%%%%%%%%&
End message from THREAD[0]
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@#############
End message from THREAD[2]
&&&%%%%&%&&&%%&&&&&&&&&&&&&&&&&&&&&*****************============================


How the data looks together:
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@#############
#####$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$%%%%%%%%%%%%%%%%%%$$$$$$$$$$$$$$%%%%%%%%%%%%&
&&&%%%%&%&&&%%&&&&&&&&&&&&&&&&&&&&&*****************============================

Highest Average: 210
Lowest Average: 92
Highest Range: 255
Lowest Range: 255
Max change: 255
Max Sum Change: 832050685
Min Standard Dev: 255.000000
Min Std Dev of Data Change: 96.076185
Done!
```



There is a spike in activity with the CPUs except for CPU4

4 Threads



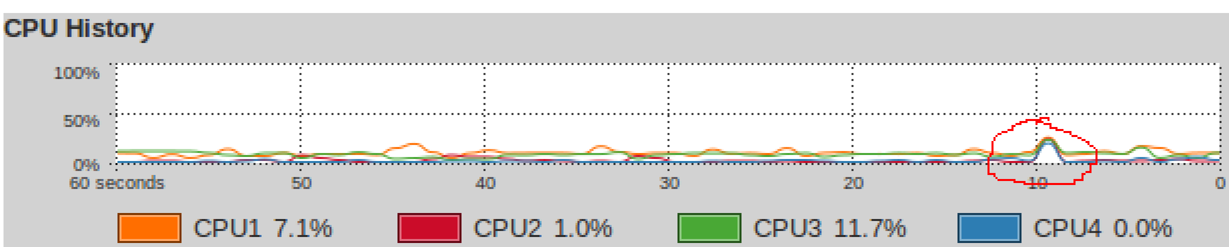


All the CPUs are working

Carlos Ortega
CS385 HW3 Report

8 Threads

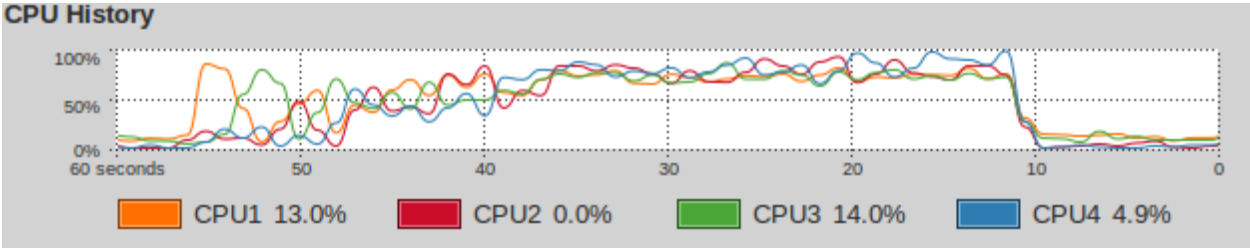
```
carlos@reptar: ~/documents/CS385/CS385HW3
carlos@reptar:~/documents/CS385/CS385HW3$ ./orderSearcher.out VRUPL_Logo.raw 8
NAME: Carlos Ortega  NETID: corteg20
file size: 15222627
Created a MSG queue, ID: 18055168
Waiting for the thread to finish
End message from THREAD[4]
%%%%%%%%%%%%%%%%%%%%%%$%%$%%%%%%%%%$%%$$$$$$$$$$$$$$$$$$%$$$$$$%$$$$$$$$$$$%%$%%
End message from THREAD[2]
@@@@@@@@@@@@@@@@@@@@#@###########################################$#$$$$$$$$$$$$$
End message from THREAD[0]
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
End message from THREAD[5]
%%%%&%%%%%%%%%%&%%%&%%&%%&&&&&&&&&%&%%%&%%%&%%&%%&&%&&%&&&&&&&%&%%&&%&&&%&&&*&&&
End message from THREAD[3]
$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$%$%%$$%%$%%$%%$%%%
End message from THREAD[7]
***=**===========*==============================================================
End message from THREAD[6]
&&&&&&&&&&&&&&&&&&&&&&*&&*&&*&&&&&&&&&*&&*&&**&*********************************
End message from THREAD[1]
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@


How the data looks together:
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@@@@@@@@@#@###########################################$#$$$$$$$$$$$$$
$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$%$%%$$%%$%%$%%$%%%
%%%%%%%%%%%%%%%%%%%%%%$%%$%%%%%%%%%$%%$$$$$$$$$$$$$$$$$$%$$$$$$%$$$$$$$$$$$%%$%%
%%%%&%%%%%%%%%%&%%%&%%&%%&&&&&&&&&%&%%%&%%%&%%&%%&&%&&%&&&&&&&%&%%&&%&&&%&&&*&&&
&&&&&&&&&&&&&&&&&&&&&&*&&*&&*&&&&&&&&&*&&*&&**&*********************************
***=**===========*==============================================================

Highest Average: 241
Lowest Average: 88
Highest Range: 255
Lowest Range: 255
Max change: 255
Max Sum Change: 315854813
Min Standard Dev: 2.972566
Min Std Dev of Data Change: 56.677707
Done!
carlos@reptar:~/documents/CS385/CS385HW3$
```



Similar CPU activity as with 4 threads

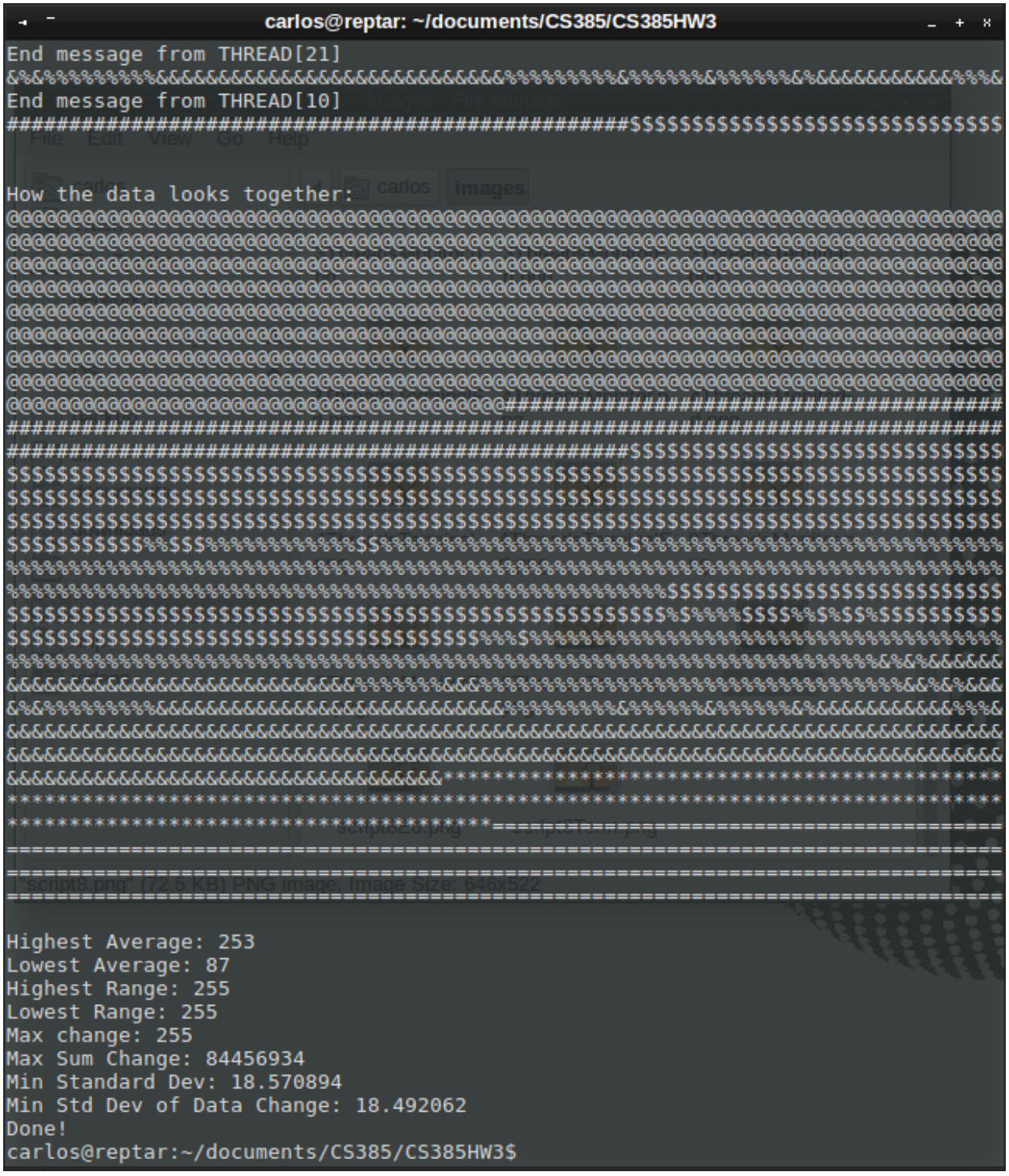
Running the script

The script call.py was executed with the argument 8. This means that it will execute the program orderSearcher.out 20 times with 1 thread, then with 2, all the way up to 8 threads.



As you can see when the script first begins orderSearcher.out is being run with only 1 thread. This is evident by the fact that only one CPU is active in the beginning. As the script keeps running, orderSearcher.out is executed with more threads, you can see that the CPU activity increases for all CPUs until the script is finished.

Example execution with 30 threads with the file VRUPL_Logo.raw



As you can see the program is able to show segments where the data is uniform. Comparing this output to the image it is clear that the program shows where there are differences in color.