

High Performance Computing

Assignment 3

Theory Questions

1. Key Differences between SIMD and MIMD

SIMD	MIMD
Single Instructions Multiple Data Streams	Multiple Instructions Multiple Data Streams
Same instruction is executed simultaneously on all CPUs but on different data streams (as the name says lol)	Capable of executing multiple instructions on multiple data streams (again, as the name says)
Suitable for scientific computing because they involve lot of vector and matrix operations – the vector can be divided into multiple data sets and each data stream executes operations on one particular data set	Suitable for any kind of application because there are separate instructions and data streams
Processors work simultaneously	Processors work asynchronously
EXAMPLE: Thinking very simple, let's say I have a lot of potatoes to peel and cut. Using SIMD, there are different data streams, where each PE will be given a few potatoes and they will all peel and cut it at the same time.	EXAMPLE: With MIMD, I can have a few data streams that will peel my potatoes while some others will cut my potatoes. So I will divide the potatoes for peeling first and then as they are done they are sent to cut them
ANOTHER EXMAPLE: Like we did in lab, the vector matrix addition. SIMD would perform the addition operation on all the elements of matrix simultaneously with a single instruction	ANOTHER EXMAPLE: Using MIMD, we can divide the matrix into smaller subsets on which the operation is performed and then add up the results

2. OpenMP is a shared memory model (threads) whereas MPI is a distributed memory model (processes). If you wanted to have global variables in MPI how you create them?

MPI does not have shared memory so sharing 'global variables' like in OpenMP does not work the same way. When using MPI, data is unique to each processor. To transfer same data over all the processes, we use MPI existing functions. Like MPI_int, MPI_BCast, but these are for when lets say we are sharing the Vector Matrix or something at the beginning of the program. But global variable that can be shared across all processes? Not sure if that is possible. Each process is a new instance of a program, so the global variable that exists in that process is not exactly global anymore, it is instead just global to THAT process.

I guess, if we had to, we could send the variable back to process0 every time it was updated and BCast to all processes again and then take average of that value with the one they already had and override the value? This just seems overly complicated and useless honestly. It's almost like all processes trying to access the variable from the globally available memory but it would be overly complicated.

3. Discuss strategies for detecting deadlock, particularly in MPI

Deadlock would be caused if 2 processors or more are trying to access the same resource at the same time and the code does not take care of this. One way to avoid this would be how we take care of deadlock even in simple computer OS deadlock prevention.

We could prioritize. Now the factors used to prioritize may differ application to application, but even the ones where there is no such 'priority'. We can simply say that processor 3 gets more priority than processor 4.

We can do the Hold and Wait. Where we allocate the resource and wait till execution it finished by THAT process. This will definitely not be efficient because it many processes can be waiting for a long time.

Deadlock can be caused in MPI from MPI_Send and MPI_Recv. One way to solve this is to reverse the order of execution for one of the conditionals.

We need to have a resource/application taking care of deadlock like a banker's algorithm but for the MPI.

4. Compare the advantages and disadvantages of a different sparse matrix formats: Dictionary of Keys, List of lists, compressed sparse rows/columns. Give an example.

Sparse Matrix is when large number of elements are zero.

Considering the sparse matrix-

Ex – {[1,0,0,1], [0,2,0,0], [0,0,3,0], [4,0,0,5]}

Dictionary of Keys (DOK)

All non zero elements are stored in a dictionary where key represents the indices and value is the non-zero element. The format is not useful for accessing specific elements and/or iterating over in the sequential order

Ex – the above matrix will be stored as

(0,0) -> 1

(0,3) -> 1

(1,1) -> 2

(2,2) -> 3

(3,0) -> 4

(3,3) -> 5

List of lists (LIL)

As the name suggests, it is a List of lists, where each row is stored as a list and each entry in the list. The entries are stored as column index to be able to look up the values.

Ex – [[(0,1), (3,1)], [(1,2)], [(2,3)], [(0,4),(3,5)]]

Compressed sparse rows/columns

All the non zero elements are stored in 3 arrays – values, column index and row index.

The array of Values and Column Index are of length of the ‘number of non zero elements’ in the matrix.

The row index array stores the index where the first non zero element occurs.

V = [1,1,2,3,4,5]

Col_index = [0,3,1,2,0,3]

Row_index = [0,1,2,3]

Programming Questions

Before implementing MPI.

```
● punyajamishra@LAPTOP-786BHK21:~/4350$ g++ A3_2.cpp -o A3_o2
● punyajamishra@LAPTOP-786BHK21:~/4350$ ./A3_o > A3_before
○ punyajamishra@LAPTOP-786BHK21:~/4350$
```

After implementing the MPI. I have attached the output file to the submission as well, it is called A3_console. I know I did mess up a bit. Its not exactly correct.

```
● punyajamishra@LAPTOP-786BHK21:~/4350$ mpicxx A3.cpp -o A3_o
● punyajamishra@LAPTOP-786BHK21:~/4350$ mpirun -np 4 ./A3_o > A3_console
○ punyajamishra@LAPTOP-786BHK21:~/4350$
```

First, let me tell what I wished to implement. I wanted to break the matrix into 4 squares and send each to a process, in which they do the calculations of the while loop. And then its all send back and gathered and implemented one last time (this would be needed just for the common edge of row and column) and that's it. Seems simple if you think about it right?

This is what I tried to implement. I split the matrix into small ‘chunks’, where the chunk size is SIZE/2 because at the end of the day a matrix size is row/2 and column/2. Since we are using an even square matrix, chunk_size=SIZE/2. I did call MPI_Abort for odd number of matrix because I don't think my code will execute for odd number of SIZE, it might send segmentation fault in the MPI.

As you will see in the output, the calculations are happening, but I think the way I tried to scatter the matrix was not exactly correct? Not all processes got the desired matrix I wanted.

Since I sent the matrix chunks in the beginning when the only cell with a value is in one part of the matrix, therefore, the values and all calculations happened only in that process. And all other processes had 0 values, therefore, no calculations per say happened.

[illegible]

```
Final Matrix:
0 0 0 0 1.288e-35 1.38224e-30 3.78093e-30 1.26882e-30 7.96475e-35 1.4269e-30 1.60375e-35 1.09801e-23 2.43444e-31 1.72679e-31 1.7549e-31 5.64472e-30 7.23971e-21 3.45786e-31 1.0051e-37
0 0 0 0 0 2.1469e-30 1.72155e-30 4.26102e-28 8.42538e-28 7.41167e-30 1.68365e-25 1.29085e-23 2.45561e-31 1.26518e-21 2.23494e-21 3.64435e-30 8.21464e-30 1.15518e-30 6.21907e-31 2.34158
Final Matrix:
0 0 0 0 4.0806e-35 5.26827e-32 2.2018e-30 7.70193e-34 9.43131e-28 8.00486e-28 9.02755e-26 1.66918e-25 1.18382e-23 1.51412e-23 9.22015e-22 1.54023e-22 5.67933e-20 8.48836e-20 1.36344e-18
0 0 0 1.288e-35 1.38224e-30 3.78093e-30 1.26882e-30 7.96475e-35 1.4269e-30 1.60375e-35 1.09801e-23 2.43444e-31 1.72679e-31 1.7549e-31 5.64472e-30 7.23971e-21 3.45786e-31 1.0051e-37
0 0 0 0 1.4668e-30 1.2156e-30 4.61602e-28 1.2538e-30 7.41167e-30 1.68365e-25 1.29085e-23 2.45561e-31 1.26518e-21 2.23494e-21 3.64435e-30 8.21464e-30 1.15518e-30 6.21907e-31 2.34158
0 0 0 4.3169e-35 1.89544e-32 2.64424e-30 3.58752e-30 8.42749e-30 8.11998e-28 5.50046e-26 8.76542e-26 9.82243e-24 4.73169e-22 5.68212e-22 3.3888e-20 4.78787e-20 2.37878e-18
0 0 0 0.89279e-35 8.12366e-33 1.89642e-32 2.17527e-30 3.74428e-30 4.26581e-28 7.83998e-28 8.40743e-26 8.17248e-26 6.96213e-24 5.3478e-24 2.14248e-22 2.78987e-21 2.83429e-20 2.44315e-18
1 0 0 0 4.0806e-35 5.2682e-32 2.2018e-30 7.70193e-34 9.43131e-28 8.00486e-28 9.02755e-26 1.66918e-25 1.18382e-23 1.51412e-23 9.22015e-22 1.54023e-22 5.67933e-20 8.48836e-20 1.36344e-18
0 0 0 4.3169e-35 1.89544e-32 2.64424e-30 3.58752e-30 8.42749e-30 8.11998e-28 5.50046e-26 8.76542e-26 9.82243e-24 4.73169e-22 5.68212e-22 3.3888e-20 4.78787e-20 2.37878e-18
0 0 0 1.1515e-35 7.26556e-33 1.60780e-32 1.71748e-31 3.01814e-30 3.05998e-28 7.90828e-28 4.69346e-26 7.66017e-26 2.7579e-24 4.23315e-24 3.27088e-22 3.57313e-22 9.3396e-21 2.34228e-20 3.9
1 0.16666e-35 7.38193e-33 1.38382e-32 7.05982e-33 1.41725e-30 2.20543e-30 2.42676e-28 4.6003e-28 3.89607e-26 2.8044e-26 1.92639e-24 4.24694e-24 1.5385e-22 1.71315e-22 5.6911e-21 1.6
0 0 0 1.1515e-35 7.26556e-33 1.60780e-32 1.71748e-31 3.01814e-30 3.05998e-28 7.90828e-28 4.69346e-26 7.66017e-26 2.7579e-24 4.23315e-24 3.27088e-22 3.57313e-22 9.3396e-21 2.34228e-20 3.9
0 0 0 4.96348e-30 2.58128e-35 4.82446e-33 5.38575e-33 1.01407e-31 1.44997e-30 1.32658e-28 1.71723e-28 1.56083e-26 1.1999e-26 1.04342e-24 1.612e-24 3.81042e-23 5.98155e-23 5.62682e-21 8.728
1 0.03198e-37 1.00281e-35 1.4758e-33 5.53275e-33 7.72717e-33 6.8449e-31 8.82371e-31 9.28363e-29 6.53399e-29 9.83553e-27 1.6061e-27 7.00829e-25 5.37406e-25 4.38166e-23 5.64781e-23 1.0033
1 0.16666e-35 7.58838e-35 5.38582e-33 7.05848e-33 1.41725e-30 2.20543e-30 2.42676e-28 4.60843e-28 3.89607e-26 2.80463e-26 1.92636e-24 4.24694e-24 1.5389e-22 1.71359e-22 5.69151e-21 1.6
0 0 4.96348e-30 2.58128e-35 4.82446e-33 5.38575e-33 1.01407e-31 1.44997e-30 1.32658e-28 1.71723e-28 1.56083e-26 1.1999e-26 1.04342e-24 1.612e-24 3.81042e-23 5.98155e-23 5.62682e-21 8.728
1 0.84896e-38 8.83161e-36 1.56595e-35 2.44222e-33 3.70626e-33 4.43534e-31 5.42929e-31 6.0033e-29 4.19678e-29 3.91915e-27 2.97632e-27 2.3777e-25 1.94726e-25 1.88915e-23 2.6722e-23 1.26066
0 0 0.84896e-38 8.83161e-36 1.56595e-35 2.44222e-33 3.70626e-33 4.43534e-31 5.42929e-31 6.0033e-29 4.19678e-29 3.91915e-27 2.97632e-27 2.3777e-25 1.94726e-25 1.88915e-23 2.6722e-23 1.26066
1 0.03198e-37 1.00281e-35 1.4758e-33 5.53275e-33 7.72717e-33 6.8449e-31 8.82371e-31 9.28363e-29 6.53399e-29 9.83553e-27 1.6061e-27 7.00829e-25 5.37406e-25 4.38166e-23 5.64781e-23 1.0033
5.64992e-41 2.37211e-38 4.63292e-36 7.69779e-36 1.04774e-31 1.26961e-32 2.45018e-31 1.72333e-31 1.18784e-29 4.40712e-29 1.06428e-27 2.67083e-27 9.67548e-26 1.67896e-25 3.34373e-24 1.76213
1 0.84896e-38 8.83161e-36 1.56595e-35 2.44222e-33 3.70626e-33 4.43534e-31 5.42929e-31 6.0033e-29 4.19678e-29 3.91915e-27 2.97632e-27 2.3777e-25 1.94726e-25 1.88915e-23 2.6722e-23 1.26066
1.37942e-38 1.38682e-38 1.38021e-36 3.49615e-36 6.08672e-34 1.43822e-33 4.33212e-33 1.26476e-31 1.02484e-31 1.00144e-29 1.5715e-29 6.00395e-28 1.55426e-27 2.29424e-26 1.06363e-25 3.29961e-25 2.46512
5.80184e-39 3.909
```

Please see the outputs properly in the file I provided