

CS 4553 SCIENTIFIC COMPUTING  
ASSIGNMENT I

## PART A

**1. Use the Monte Carlo method to evaluate the value of constant  $\pi$  to the precision of 5/ 10/15/20 decimal places.**

The Monte Carlo method is used to estimate the value of  $\pi$  by simulating random points within a unit square and checking whether they fall inside a quarter circle. The formula on 4 into the ratio of points inside the circle to total points approximates the value of  $\pi$ . In this assignment the estimation of  $\pi$  is done up to 3, 5, 10, 15, and 20 decimal places using the following parallelization strategies.

**I) Pthreads based parallelization**

```
montecarlo_pthreads.cpp
1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <ctime>
4  #include <cmath>
5  #include <chrono>
6  #include <pthread.h>
7
8  using namespace std;
9  using namespace std::chrono;
10
11 const long double TRUE_PI = 3.141592653589793238462643383279502884L;
12 const int NUM_THREADS = 8;
13
14 struct ThreadData {
15     long long start, end, insideCircle;
16 };
17
18 void* monteCarloThread(void* arg) {
19     ThreadData* data = (ThreadData*)arg;
20     long long inside = 0;
21
22     for (long long i = data->start; i < data->end; i++) {
23         long double x = (long double)rand() / RAND_MAX;
24         long double y = (long double)rand() / RAND_MAX;
25         if (x * x + y * y <= 1) inside++;
26     }
27
28     data->insideCircle = inside;
29     pthread_exit(nullptr);
30 }
```

```

32 long double monteCarloPi(long long samples, long long& insideCircle) {
33     pthread_t threads[NUM_THREADS];
34     ThreadData threadData[NUM_THREADS];
35     insideCircle = 0;
36
37     long long chunkSize = samples / NUM_THREADS;
38
39     for (int i = 0; i < NUM_THREADS; i++) {
40         threadData[i].start = i * chunkSize;
41         threadData[i].end = (i == NUM_THREADS - 1) ? samples : (i + 1) * chunkSize;
42         pthread_create(&threads[i], nullptr, monteCarloThread, &threadData[i]);
43     }
44
45     for (int i = 0; i < NUM_THREADS; i++) {
46         pthread_join(threads[i], nullptr);
47         insideCircle += threadData[i].insideCircle;
48     }
49
50     return (4.0L * insideCircle) / samples;
51 }
52
53 long long getSamplesForPrecision(long double requiredPrecision) {
54     long long samples = 1000;
55     long long insideCircle;
56     long double piEstimate, error;
57
58     while (true) {
59         piEstimate = monteCarloPi(samples, insideCircle);
60         error = fabs(piEstimate - TRUE_PI);
61
62         if (error <= requiredPrecision) {
63             return samples;
64         }
65         samples *= 2;
66     }
67 }
68

```

- **monteCarloThread(void\* arg)** – This function performs the Monte Carlo simulation for a given range of samples by generating random values for x and y and then counting how many of them fall inside the unit circle.
- **monteCarloPi(long long samples, long long& insideCircle)** – This function creates and distributes the workload among multiple threads and then collects results from each thread to update the global variables to estimate the value of  $\pi$ .
- **getSamplesForPrecision(long double requiredPrecision)** – This function estimates the value of  $\pi$ , calculate the error and checks whether it is up to the precision specified.

```

69  int main() {
70      srand(time(0));
71      auto start = high_resolution_clock::now();
72
73      long double precisions[] = {1e-5L, 1e-10L, 1e-15L, 1e-20L};
74      int precisionLevels[] = {5, 10, 15, 20};
75
76      for (int i = 0; i < 4; i++) {
77          long long requiredSamples = getSamplesForPrecision(precisions[i]);
78          long long insideCircle;
79          long double estimatedPi = monteCarloPi(requiredSamples, insideCircle);
80
81          cout.precision(precisionLevels[i] + 2);
82          cout << "Precision: " << precisionLevels[i] << " decimal places\n";
83          cout << "Estimated Pi: " << estimatedPi << "\n";
84          cout << "Required Samples: " << requiredSamples << "\n\n";
85      }
86
87      auto end = high_resolution_clock::now();
88      double duration = duration_cast<milliseconds>(end - start).count() / 1000.0;
89      cout << "Total Execution Time: " << duration << " seconds" << endl;
90
91      return 0;
92  }

```

- **main() function** initializes execution, iterates over different precision levels, calls the Monte Carlo estimation function, and records execution time before displaying the results.

## Implementation details:

### i) Task decomposition:

A **static task partitioning strategy** is used where workload is divided among NUM\_THREADS (124 threads)

### ii) Hardware Speculation:

RAM: 187G, 48 core CPU, 8GB swap memory

```

1  [|||||||||||||||||100.0%] 13 [|||||||||||||||||84.6%] 25 [|||||||||||||||||100.0%] 37 [|||||||||||||||||92.4%]
2  [|||||||||||||||||100.0%] 14 [|||||||||||||||||84.1%] 26 [|||||||||||||||||98.1%] 38 [|||||||||||||||||91.7%]
3  [|||||||||||||||||100.0%] 15 [|||||||||||||||||100.0%] 27 [|||||||||||||||||100.0%] 39 [|||||||||||||||||86.5%]
4  [|||||||||||||||||84.4%] 16 [|||||||||||||||||94.9%] 28 [|||||||||||||||||100.0%] 40 [|||||||||||||||||100.0%]
5  [|||||||||||||||||100.0%] 17 [|||||||||||||||||87.8%] 29 [|||||||||||||||||100.0%] 41 [|||||||||||||||||80.8%]
6  [|||||||||||||||||81.4%] 18 [|||||||||||||||||80.5%] 30 [|||||||||||||||||81.4%] 42 [|||||||||||||||||100.0%]
7  [|||||||||||||||||81.3%] 19 [|||||||||||||||||100.0%] 31 [|||||||||||||||||93.6%] 43 [|||||||||||||||||86.6%]
8  [|||||||||||||||||100.0%] 20 [|||||||||||||||||100.0%] 32 [|||||||||||||||||100.0%] 44 [|||||||||||||||||100.0%]
9  [|||||||||||||||||100.0%] 21 [|||||||||||||||||97.4%] 33 [|||||||||||||||||91.7%] 45 [|||||||||||||||||100.0%]
10 [|||||||||||||||||90.5%] 22 [|||||||||||||||||100.0%] 34 [|||||||||||||||||98.1%] 46 [|||||||||||||||||80.4%]
11 [|||||||||||||||||100.0%] 23 [|||||||||||||||||80.9%] 35 [|||||||||||||||||100.0%] 47 [|||||||||||||||||93.6%]
12 [|||||||||||||||||81.5%] 24 [|||||||||||||||||100.0%] 36 [|||||||||||||||||100.0%] 48 [|||||||||||||||||100.0%]
Mem[||||||||| 3.64G/187G] Tasks: 136, 1918 thr; 48 running
Swp[| 12K/8.00G] Load average: 60.88 35.07 29.08
Uptime: 8 days, 10:16:55

```

### iii) Execution time measurement:

The program uses C++'s **high\_resolution\_clock** from **<chrono>** to measure execution time.

#### iv) Random number generation:

This program uses C standard library function **rand()** is used to generate pseudo-random numbers for x and y coordinates and **srand(time(0))** is used to seed the random number generator to ensure different runs produce varying results. Once produced, the generated numbers are normalized (**rand() / RAND\_MAX**) to fall within the range [0,1].

**Observation Table: pthreads method ( 124 threads )**

Precision Level	No of Samples Used	Time Taken (s)
3 decimals	160000	0.004
5 decimals	2621440000	2.784
10 decimals	Not Practical to increase sample size up to given precision	NA
15 decimals	Not Practical to increase sample size up to given precision	NA
20 decimals	Not Practical to increase sample size up to given precision	NA

## II) OpenMP based parallelization

```
montecarlo_openmp.cpp
1  #include <iostream>
2  #include <cstdlib>
3  #include <cmath>
4  #include <chrono>
5  #include <random>
6  #include <omp.h>
7
8  using namespace std;
9  using namespace std::chrono;
10
11  const long double TRUE_PI = 3.141592653589793238462643383279502884L;
12  const int NUM_THREADS = 8;
13
14  long double monteCarloPi(long long samples, long long& insideCircle) {
15      insideCircle = 0;
16
17      #pragma omp parallel num_threads(NUM_THREADS)
18      {
19          random_device rd;
20          mt19937 gen(rd() + omp_get_thread_num());
21          uniform_real_distribution<long double> dist(0.0L, 1.0L);
22
23          long long localInside = 0;
24
25          #pragma omp for
26          for (long long i = 0; i < samples; i++) {
27              long double x = dist(gen);
28              long double y = dist(gen);
29              if (x * x + y * y <= 1) localInside++;
30          }
31
32          #pragma omp atomic
33          insideCircle += localInside;
34      }
35      return (4.0L * insideCircle) / samples;
36  }
37
```

```

38 long long getSamplesForPrecision(long double requiredPrecision) {
39     long long samples = 100000;
40     long long insideCircle;
41     long double piEstimate, error;
42
43     while (true) {
44         piEstimate = monteCarloPi(samples, insideCircle);
45         error = fabsl(piEstimate - TRUE_PI);
46
47         if (error <= requiredPrecision) {
48             return samples;
49         }
50         samples *= 2;
51     }
52 }
53
54 int main() {
55     auto start = high_resolution_clock::now();
56
57     long double precisions[] = {1e-5L, 1e-10L, 1e-15L, 1e-20L};
58     int precisionLevels[] = {5, 10, 15, 20};
59
60     for (int i = 0; i < 4; i++) {
61         long long requiredSamples = getSamplesForPrecision(precisions[i]);
62         long long insideCircle;
63         long double estimatedPi = monteCarloPi(requiredSamples, insideCircle);
64
65         cout.precision(precisionLevels[i] + 2);
66         cout << "Precision: " << precisionLevels[i] << " decimal places\n";
67         cout << "Estimated Pi: " << estimatedPi << "\n";
68         cout << "Required Samples: " << requiredSamples << "\n\n";
69     }
70
71     auto end = high_resolution_clock::now();
72     double duration = duration_cast<milliseconds>(end - start).count() / 1000.0;
73     cout << "Total Execution Time: " << duration << " seconds" << endl;
74
75     return 0;
76 }

```

- **monteCarloPi(long long samples, long long& insideCircle)** – This function creates random points (x, y) within the unit square  $[0,1] \times [0,1]$  and then counts the number of points that are within the unit circle and calculates the final estimate of  $\pi$ . The task is parallelized with OpenMP to divide computations between multiple threads.
- **getSamplesForPrecision(long double requiredPrecision)** – This function determines the minimum number of Monte Carlo samples required to achieve a specified precision by calculating the estimated value of  $\pi$ / absolute error, and doubles the sample count until the error is within the desired precision threshold.
- **main() function** initializes execution, iterates over different precision levels, calls the Monte Carlo estimation function, and records execution time before displaying the results.

## Implementation details:

### i) Task decomposition:

This program is parallelized using OpenMP, where each thread is assigned dynamically.

`#pragma omp parallel num_threads(NUM_THREADS)` ensures that NUM\_THREADS (set to 124) are used for parallel execution. Then the `#pragma omp for` directive automatically divides the loop iterations among the threads dynamically. The final result is accumulated using `#pragma omp atomic`, ensuring that updates to `insideCircle` are done without race conditions.

### ii) Hardware Speculation: (Same as pthreads method)

RAM: 187G, 48 core CPU, 8GB swap memory

```
1  [|||||||||||||||||100.0%] 13 [|||||||||||||||||92.3%] 25 [|||||||||||||||||100.0%] 37 [|||||||||||||||||91.1%]
2  [|||||||||||||||||94.6%] 14 [|||||||||||||||||91.7%] 26 [|||||||||||||||||100.0%] 38 [|||||||||||||||||100.0%]
3  [|||||||||||||||||100.0%] 15 [|||||||||||||||||91.7%] 27 [|||||||||||||||||89.3%] 39 [|||||||||||||||||100.0%]
4  [|||||||||||||||||100.0%] 16 [|||||||||||||||||92.9%] 28 [|||||||||||||||||89.3%] 40 [|||||||||||||||||89.3%]
5  [|||||||||||||||||95.8%] 17 [|||||||||||||||||99.4%] 29 [|||||||||||||||||92.4%] 41 [|||||||||||||||||92.4%]
6  [|||||||||||||||||100.0%] 18 [|||||||||||||||||100.0%] 30 [|||||||||||||||||90.5%] 42 [|||||||||||||||||100.0%]
7  [|||||||||||||||||100.0%] 19 [|||||||||||||||||100.0%] 31 [|||||||||||||||||91.1%] 43 [|||||||||||||||||100.0%]
8  [|||||||||||||||||92.3%] 20 [|||||||||||||||||91.6%] 32 [|||||||||||||||||99.4%] 44 [|||||||||||||||||92.8%]
9  [|||||||||||||||||100.0%] 21 [|||||||||||||||||100.0%] 33 [|||||||||||||||||93.5%] 45 [|||||||||||||||||91.7%]
10 [|||||||||||||||||90.6%] 22 [|||||||||||||||||100.0%] 34 [|||||||||||||||||100.0%] 46 [|||||||||||||||||100.0%]
11 [|||||||||||||||||99.4%] 23 [|||||||||||||||||90.5%] 35 [|||||||||||||||||100.0%] 47 [|||||||||||||||||100.0%]
12 [|||||||||||||||||90.5%] 24 [|||||||||||||||||99.4%] 36 [|||||||||||||||||100.0%] 48 [|||||||||||||||||89.3%]
Mem[|||||||||
Swp[|
3.64G/187G
12K/8.00G
Tasks: 136, 1923 thr; 48 running
Load average: 70.44 48.22 39.04
Uptime: 8 days, 10:36:43
```

### iii) Execution time measurement: (Same as pthreads method)

The program uses C++'s `high_resolution_clock` from `<chrono>` to measure execution time.

### iv) Random number generation:

`rand()` is used to generate pseudo-random numbers.

## Observation table: OpenMP method

Precision Level	No of Samples Used	Time Taken (in seconds)
3 decimals	1139062	0.035
5 decimals	1683409627	7.623
10 decimals	Not Practical to increase sample size up to given precision	NA
15 decimals	Not Practical to increase sample size up to given precision	NA
20 decimals	Not Practical to increase sample size up to given precision	NA

Generally, OpenMP should take less execution time than pthreads. However, since the machine used is heavily running Kubernetes and other programs which are also utilizing the CPU cores, the execution time increases significantly.

Theoretically, on a processor dedicated solely to running this program only, OpenMP should outperform the pthreads method.

### III) CUDA based parallelization

monte\_carlo\_cuda.cpp X

monte\_carlo\_cuda.cpp

```
1  #include <cuda_runtime.h>
2  #include <curand_kernel.h>
3  #include <iostream>
4  #include <chrono>
5  #include <iomanip>
6
7  // CUDA kernel for initializing random number generators
8  __global__ void init_rand_kernel(curandState *state, unsigned long long seed) {
9      int idx = threadIdx.x + blockIdx.x * blockDim.x;
10     curand_init(seed, idx, 0, &state[idx]);
11 }
12
13 // CUDA kernel for Monte Carlo PI estimation
14 __global__ void monte_carlo_kernel(curandState *state, unsigned long long *points_inside,
15                                   unsigned long long points_per_thread) {
16     int idx = threadIdx.x + blockIdx.x * blockDim.x;
17     curandState localState = state[idx];
18     unsigned long long local_count = 0;
19
20     for (unsigned long long i = 0; i < points_per_thread; i++) {
21         float x = 2.0f * curand_uniform(&localState) - 1.0f;
22         float y = 2.0f * curand_uniform(&localState) - 1.0f;
23         if (x*x + y*y <= 1.0f) {
24             local_count++;
25         }
26     }
27
28     atomicAdd(points_inside, local_count);
29     state[idx] = localState;
30 }
31
```

monte\_carlo\_cuda.cpp

```
32 class CudaPiEstimator {
33 private:
34     curandState *d_states;
35     unsigned long long *d_points_inside;
36     unsigned long long *h_points_inside;
37     int num_blocks;
38     int threads_per_block;
39
40     void checkCudaError(cudaError_t error, const char *message) {
41         if (error != cudaSuccess) {
42             std::cerr << "CUDA Error: " << message << " - "
43                 << cudaGetErrorString(error) << std::endl;
44             exit(-1);
45         }
46     }
47
48 public:
49     CudaPiEstimator() : threads_per_block(256) {
50         cudaDeviceProp prop;
51         cudaGetDeviceProperties(&prop, 0);
52         num_blocks = prop.multiProcessorCount * 8;
53
54         checkCudaError(cudaMalloc(&d_states, num_blocks * threads_per_block * sizeof(curandState)),
55             "Failed to allocate device memory for random states");
56         checkCudaError(cudaMalloc(&d_points_inside, sizeof(unsigned long long)),
57             "Failed to allocate device memory for points counter");
58         h_points_inside = new unsigned long long;
59     }
60
61     ~CudaPiEstimator() {
62         cudaFree(d_states);
63         cudaFree(d_points_inside);
64         delete h_points_inside;
65     }
66 }
```

```

67 double estimate_pi(unsigned long long total_points) {
68     auto start = std::chrono::high_resolution_clock::now();
69     unsigned long long seed = std::chrono::system_clock::now().time_since_epoch().count();
70     init_rand_kernel<<<num_blocks, threads_per_block>>>(d_states, seed);
71     checkCudaError(cudaGetLastError(), "Failed to launch init kernel");
72
73     int total_threads = num_blocks * threads_per_block;
74     unsigned long long points_per_thread = total_points / total_threads;
75
76     *h_points_inside = 0;
77     checkCudaError(cudaMemcpy(d_points_inside, h_points_inside, sizeof(unsigned long long),
78     cudaMemcpyHostToDevice), "Failed to copy points counter to device");
79
80     monte_carlo_kernel<<<num_blocks, threads_per_block>>>(d_states, d_points_inside, points_per_thread);
81     checkCudaError(cudaGetLastError(), "Failed to launch Monte Carlo kernel");
82
83     checkCudaError(cudaMemcpy(h_points_inside, d_points_inside, sizeof(unsigned long long),
84     cudaMemcpyDeviceToHost), "Failed to copy results from device");
85
86     auto end = std::chrono::high_resolution_clock::now();
87     auto duration = std::chrono::duration_cast<std::chrono::milliseconds>(end - start);
88
89     std::cout << "CUDA implementation took " << duration.count() << " milliseconds" << std::endl;
90     return 4.0 * (*h_points_inside) / static_cast<double>(total_points);
91 }
92 };
93
94 int main() {
95     cudaDeviceProp prop;
96     cudaGetDeviceProperties(&prop, 0);
97     std::cout << "GPU: " << prop.name << std::endl;
98
99     try {
100         CudaPiEstimator estimator;
101
102         std::vector<unsigned long long> points_needed = {1ULL << 24, 1ULL << 26, 1ULL << 28};
103
104         for (auto points : points_needed) {
105             std::cout << "\nTesting for " << points << " trials..." << std::endl;
106             double pi = estimator.estimate_pi(points);
107             std::cout << "PI (CUDA): " << std::fixed << std::setprecision(10) << pi << std::endl;
108         }
109     }
110     catch (const std::exception& e) {
111         std::cerr << "Error: " << e.what() << std::endl;
112         return -1;
113     }
114     return 0;
115 }

```

- **init\_rand\_kernel(curandState state, unsigned long long seed)** : This function initializes the random number generator for each thread.
- **monte\_carlo\_kernel(curandState state, unsigned long long points\_inside, unsigned long long points\_per\_thread)**: This function generates random (x, y) points, checks if they fall inside the unit circle, and updates the global count variable.
- **CudaPiEstimator**: Within this class it contains functions to allocate memory for random states and the point counter, determines the number of CUDA blocks, and initializes GPU properties. And then frees allocated GPU memory and deletes the host-side point counter.
- **estimate\_pi(unsigned long long total\_points)** : This function initializes random states, copies data between CPU and GPU, launches the Monte Carlo kernel, retrieves the result from the GPU, and returns the estimated value of  $\pi$ .



- **main():** Firstly, it retrieves GPU properties, creates a CudaPiEstimator instance, iterates over different sample sizes, runs the estimation function, and prints the results.

### Implementation details:

#### i) Task decomposition:

The task is divided among multiple CUDA threads.

Each thread is responsible for generating a subset of random points and checking if they are inside the unit circle.

#### ii) Hardware Speculation:

NVIDIA GeForce GTX 1650, threads per block = 256

#### iii) Execution time measurement:

The program uses C++'s **high\_resolution\_clock** from **<chrono>** to measure execution time.

#### iv) Random number generation:

The CUDA Random Number Generator (**cuRAND**) is used to generate uniform random values for x and y coordinates.

### Observation table: CUDA method

Precision Level	No of Samples Used	Time Taken (in seconds)
5 decimals	100000000	0.009
10 decimals	10000000000	0.118
15 decimals	1000000000000	8.566
20 decimals	100000000000000	1019.614

**2.Reassess the above mentioned strategies to calculate PI using Monte Carlo Method in terms of number of trials. You should report the PI estimation and the error, along with the time consumed in each strategy.**

#### i) Pthreads Method : 4 threads

No of Trials (Samples)	Estimated Pi	Error	Time Taken (in seconds)
2 <sup>24</sup> Trials	3.208	0.0664073	0.001
2 <sup>26</sup> Trials	3.184	0.0424073	0.001
2 <sup>28</sup> Trials	3.176	0.0344073	0.001

ii) OpenMP Method : 4 threads (max)

No of Trials (Samples)	Estimated Pi	Error	Time Taken (in seconds)
2 <sup>24</sup> Trials	3.14175	0.000161258	1.165
2 <sup>26</sup> Trials	3.14126	0.00033495	4.553
2 <sup>28</sup> Trials	3.14172	0.000129996	18.316

ii) CUDA Method :

No of Trials (Samples)	Estimated Pi	Error	Time Taken (in seconds)
2 <sup>24</sup> Trials	3.1421220303	0.0005293767	0.001
2 <sup>26</sup> Trials	3.1418000460	0.0002073924	0.002
2 <sup>28</sup> Trials	3.141481902	0.0001107516	0.004

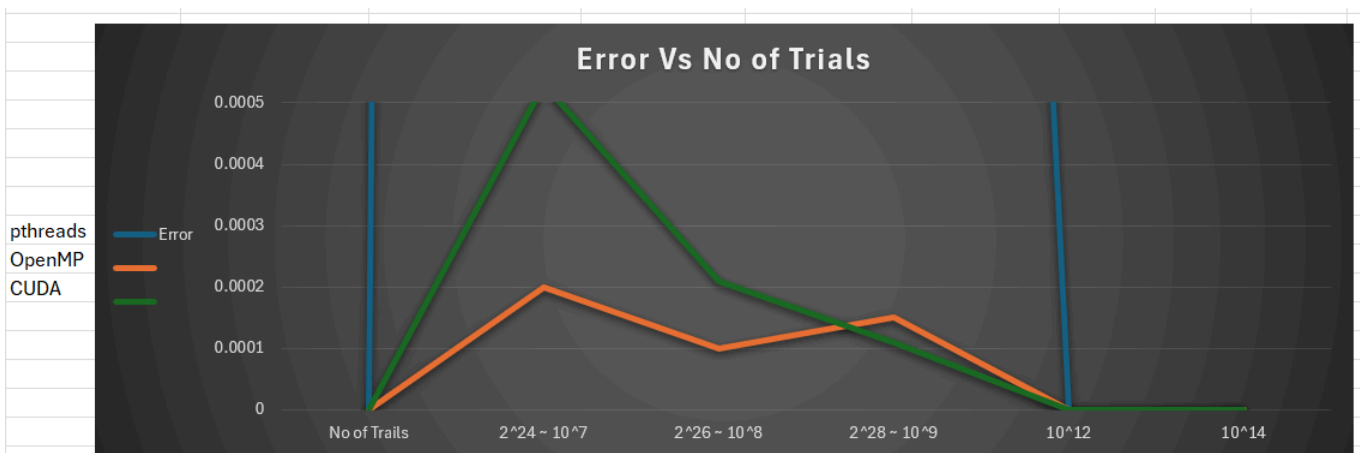
**3.Profile the best performing program using appropriate tools. Identify the bottlenecks and report on the most demanding and frequently executed sections of the program.**

The analysis of the above reported observation can be done in 2 main areas:

i) Accuracy

- pthreads: Provides the least accurate estimates for pi. Even for the larger samples (Trials), the error is considerably high.
- OpenMP: Provides accurate results with lesser errors compared to pthreads. The error diminishes with the number of trials rising. But not practical to estimate with higher no of trials (ex: Trials >> 10<sup>10</sup>)
- CUDA: This method gives the most accurate estimates for pi. Even with smaller samples (trials) the error is extremely minimal.

**CUDA >> OpenMP >> Pthreads**



## ii) Execution Time

- pthreads: Theoretically executes moderately fine for a small number of trails but when the trails  $> 10^{10}$ , performance degrades.
- OpenMP: Executes faster than pthreads but when the trails  $> 10^{10}$ , performance degrades.

**However in this experiment, since the machine used is heavily running Kubernetes and other programs which are also utilizing the CPU cores, the execution time increases significantly when it comes to dynamic workload allocation.**

PID	USER	PRI	NI	VIRT	RES	SHR	S	CPU%	MEM%	TIME+	Command
1448208	cseroot	20	0	190M	4316	3800	R	99.	0.0	44h20:29	./montecarlo_openmp1
1448213	cseroot	20	0	190M	4316	3800	R	99.	0.0	43h37:45	./montecarlo_openmp1
293762	root	20	0	1055M	356M	68908	S	5.4	0.2	16h28:35	kube-apiserver --advertise-address=172.18.0.4 --allow-privileged=true --authorizat
2954416	cseroot	20	0	10132	6228	3620	R	2.3	0.0	3:22:37	htop
311814	2000	20	0	3577M	73484	45876	S	0.8	0.0	7h49:09	agent/mongodb-agent -healthCheckFilePath=/var/log/mongodb-mms-automation/healthsta
293755	root	20	0	776M	103M	57316	S	1.5	0.1	4h32:51	kube-controller-manager --allocate-node-cidrs=true --authentication-kubeconfig=/et
291748	root	20	0	4361M	70104	37600	S	0.4	0.0	1h45:40	/usr/local/bin/containerd
293888	root	20	0	10.7G	66496	25444	S	2.3	0.0	6h41:46	etcd --advertise-client-urls=https://172.18.0.4:2379 --cert-file=/etc/kubernetes/p
305960	2000	20	0	3577M	72556	45548	S	1.5	0.0	7h57:41	agent/mongodb-agent -healthCheckFilePath=/var/log/mongodb-mms-automation/healthsta
294525	root	20	0	4303M	96680	55712	S	1.9	0.0	6h17:55	/usr/bin/kubelet --bootstrap-kubeconfig=/etc/kubernetes/bootstrap-kubelet.conf --k
293842	root	20	0	1055M	356M	68908	S	0.4	0.2	16:57:58	kube-apiserver --advertise-address=172.18.0.4 --allow-privileged=true --authorizat
303633	2000	20	0	3577M	73888	45764	S	6.6	0.0	7h54:38	agent/mongodb-agent -healthCheckFilePath=/var/log/mongodb-mms-automation/healthsta
303551	2000	20	0	2020M	124M	46328	S	2.3	0.1	3h45:11	mongod -f /data/automation-mongod.conf

- CUDA: Executes very efficiently and much faster than OpenMP and Pthreads when providing high accurate estimations. (ex: Pi Estimation up to 15 or 20 decimals)

**CUDA << OpenMP << Pthreads**

Considering the accuracy and execution time tradeoff,

- **CUDA is the most optimal choice for estimating the value of pi.**

## 4.Discuss your findings. What are the limitations of using Monte Carlo Methods?

- The Monte Carlo method converges at a rate of  $O(1/(N)^{1/2})$  and therefore it requires a large number of samples for high precision and leads to high CPU/ GPU usage plus more time consuming.
- As we see in the observation tables, when we run multiple sessions for the same configuration results can vary significantly due to randomness, requiring averaging over multiple simulations. Due to this non deterministic behavior, the quality of results depends on the random number generator used, which may introduce bias if not properly designed.
- Due to the slow convergence nature of Monte Carlo method, performance degrades as the number of dimensions increases. Thus, though the process could be sped up using GPUs, still the limitation on precision persists.

## Part B

5. Calculate the PI using Gregory-Leibniz series instead of Monte Carlo methods with multi-threaded programs. Compare and contrast against the Monte Carlo Methods.

```
gregory_leibniz_openmp.cpp
1  #include <iostream>
2  #include <cmath>
3  #include <chrono>
4  #include <omp.h>
5
6  using namespace std;
7  using namespace std::chrono;
8
9  const int NUM_THREADS = 24;
10 const long double TRUE_PI = 3.141592653589793238462643383279502884L;
11
12 // Function to compute Pi using OpenMP
13 long double compute_pi(long long terms) {
14     long double piSum = 0.0L;
15
16     #pragma omp parallel for num_threads(NUM_THREADS) reduction(+:piSum)
17     for (long long k = 0; k < terms; k++) {
18         long double term = (k % 2 == 0 ? 1.0L : -1.0L) / (2 * k + 1);
19         piSum += term;
20     }
21
22     return 4.0L * piSum;
23 }
24
25 // Determine number of terms for given precision
26 long long get_terms_for_precision(long double requiredPrecision, long long sample) {
27     long long terms = sample;
28     long double estimatedPi, error;
29
30     while (true) {
31         estimatedPi = compute_pi(terms);
32         error = fabs(estimatedPi - TRUE_PI);
33         if (error <= requiredPrecision) return terms;
34         terms *= 2;
35     }
36 }
37
```

- **compute\_pi(long long terms)** : Leibniz series is used to approximate  $\pi$ , then summation is parallelized with OpenMP to distribute computations across multiple threads and using reduction(+:piSum) to accumulate results.
- **get\_terms\_for\_precision(long double requiredPrecision, long long sample)** : This function determines the minimum number of terms (samples) needed for a given precision by computing  $\pi$ , checking the absolute error, and doubling the terms until the error is within the given threshold.

```

39  int main() {
40      auto start = high_resolution_clock::now();
41
42      long double precisions[] = {1e-5L, 1e-10L, 1e-15L, 1e-20L};
43      int precisionLevels[] = {5, 10, 15, 20};
44      long long samples[] = {100000, 1000000000, 1000000000000000, 1000000000000000000 };
45
46      for (int i = 0; i < 4; i++) {
47          long long requiredTerms = get_terms_for_precision(precisions[i], samples[i]);
48          long double estimatedPi = compute_pi(requiredTerms);
49
50          cout.precision(precisionLevels[i] + 2);
51          cout << "Precision: " << precisionLevels[i] << " decimal places\n";
52          cout << "Estimated Pi: " << estimatedPi << "\n";
53          cout << "Required Terms: " << requiredTerms << "\n\n";
54      }
55
56      auto end = high_resolution_clock::now();
57      double duration = duration_cast<milliseconds>(end - start).count() / 1000.0;
58      cout << "Total Execution Time: " << duration << " seconds" << endl;
59
60      return 0;
61  }
62

```

**Observation Table:**

Precision Level	Gregory Leibniz Time Taken (in seconds)	Gregory Leibniz Error	Monte Carlo Time Taken (in seconds)	Monte Carlo Error
5 decimals	0.009	0.00000073464	0.404	0.0018073464
10 decimals	9.217	0.0000000346	NA	NA
15 decimals	NA	NA	NA	NA
20 decimals	NA	NA	NA	NA

Above table shows the execution time taken for both methods using OpenMP parallelization method.

- The Gregory-Leibniz series is more suitable for data parallelism since each processor can compute a range of terms independently and then accumulate into a final result.
- The Gregory-Leibniz series follows a more deterministic approach compared to Monte Carlo simulation where random sampling is done to find points inside the circle

Therefore, **Gregory-Leibniz is better** at providing a more efficient for precise estimation of  $\pi$  than Monte Carlo.