

---

# COMP1680 - Clouds, Grids and Virtualisation Coursework Report

---

Azhar Muhammed - 001364857 Word Count:

## Part I

# Parallel Processing using Cloud Computing

In the contemporary landscape of computational technologies, small consultancy firms face critical decisions regarding their computational infrastructure. The choice between adopting cloud computing services or investing in an onsite High-Performance Computing (HPC) system significantly impacts operational efficiency, scalability, and financial expenditure. This report evaluates these options for a consultancy company comprising 30 consultants utilizing approximately 1,400 CPU hours each month. The analysis focuses on the advantages and disadvantages of leading cloud service providers such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) when compared to an onsite HPC system for multicore workloads using batch processing. A comprehensive financial cost comparison is provided, culminating in a recommendation that aligns with the company's strategic objectives.

## 1. Analysis of Cloud Providers vs. Onsite HPC for Multicore Workloads

Cloud computing has revolutionized the way businesses handle computational workloads, offering scalable resources and flexible pricing models. In contrast, onsite HPC systems provide dedicated resources but require substantial initial investments and ongoing maintenance.

Table 1. Onsite HPC Costs Breakdown

Cost Item	Minimum (\$)	Maximum (\$)
<b>Initial Setup Costs</b>		
HPC Hardware	80,000	120,000
Cooling Systems	10,000	20,000
Power Supplies and UPS	5,000	10,000
Networking Equipment	5,000	10,000
<b>Total Initial Setup Cost</b>	<b>100,000</b>	<b>160,000</b>
<b>Annual Operating Costs</b>		
Maintenance and Support Contracts	10,000	15,000
Energy Consumption	756	756
IT Personnel Salary	50,000	70,000
<b>Total Annual Operating Costs</b>	<b>60,756</b>	<b>85,756</b>
<b>Hardware Upgrade Costs (Every 4–5 Years)</b>	50,000	112,000
<b>Total 5-Year Cost</b>	<b>453,780</b>	<b>700,780</b>

### 1.1. Advantages of Cloud Computing Over Onsite HPC

Cloud computing offers several advantages over onsite HPC systems, particularly for small consultancies:

- **Scalability and Flexibility:** Cloud services allow dynamic scaling of computational resources to match workload

demands. This elasticity is essential for batch processing tasks that may have variable computational needs (Armbrust et al. 2010).

- **Cost Efficiency:** With pay-as-you-go models, companies avoid the substantial upfront capital expenditure associated with purchasing HPC hardware. Operational costs are aligned with actual usage, enhancing financial predictability (Li, O'Brien, and Zhang 2019).
- **Maintenance and Operational Overhead:** Cloud providers handle hardware maintenance, software updates, and security patches, reducing the burden on internal IT staff (Marinescu 2013).
- **Access to Advanced Technologies:** Cloud platforms frequently update their offerings, providing access to the latest hardware and software innovations without additional costs (Dillon, Wu, and Chang 2010).

## 1.2. Disadvantages of Cloud Computing

Despite these advantages, cloud computing presents certain challenges:

- **Data Security and Compliance:** Relying on third-party providers raises concerns about data sovereignty and compliance with regulations such as GDPR (Hashem et al. 2015).
- **Performance Variability:** Multi-tenant cloud environments can lead to inconsistent performance due to resource contention (Leitner and Cito 2016).
- **Potential for Unexpected Costs:** Without proper management, costs can escalate due to factors like data egress fees and underutilized resources (Rehman, Wah, and Hussain 2018).

## 2. Comparison of Leading Cloud Providers

**Amazon Web Services (AWS):** AWS offers a broad range of services suitable for batch processing, such as AWS Batch and EC2 compute-optimized instances. It provides robust scalability and a variety of pricing models, including on-demand, reserved, and spot instances (Amazon Web Services n.d.). AWS's global infrastructure ensures high availability and low latency.

**Microsoft Azure:** Azure provides services like Azure Batch and compute-optimized virtual machines suitable for parallel processing tasks. It integrates seamlessly with other Microsoft products, which can be advantageous for companies already utilizing Microsoft ecosystems (Microsoft Azure n.d.). Azure offers competitive pricing and various discount options, such as reserved instances.

**Google Cloud Platform (GCP):** GCP's Compute Engine and preemptible virtual machines offer cost-effective solutions for batch processing workloads. GCP emphasizes sustained use discounts, automatically reducing costs for consistent usage without long-term commitments (Google Cloud Platform n.d.). Google's expertise in containerization and Kubernetes can benefit firms looking to adopt modern deployment practices.

Table 2. Cloud Computing Costs for AWS, Azure, and GCP

Cost Item	AWS	Azure	GCP
<b>Instance Type</b>	c6a.large	F2s_v2	n2-standard-2
<b>vCPUs</b>	2	2	2
<b>Price per Hour (\$)</b>	0.085	0.085	0.084
<b>Monthly Compute Cost (\$)</b>	119.00	119.00	117.60
<b>Storage Cost per Month (\$)</b>	20.00	20.00	20.00
<b>Data Transfer Cost per Month (\$)</b>	10.00	10.00	10.00
<b>Total Monthly Cost (\$)</b>	149.00	149.00	147.60
<b>Annual Cost (\$)</b>	1,788.00	1,788.00	1,771.20
<b>5-Year Cost (\$)</b>	8,940.00	8,940.00	8,856.00

## 2.1. Comparison Between Cloud Providers

While all three providers offer similar core services, differences exist in pricing models, discount structures, and additional services:

- **Pricing Models:** AWS and Azure offer reserved instances for long-term commitments, while GCP provides sustained use discounts without upfront commitments.
- **Ecosystem Integration:** Azure may be preferable for companies using Microsoft tools, whereas AWS has a more extensive range of services and a mature ecosystem.
- **Specialized Services:** GCP excels in machine learning and big data analytics services, which could be beneficial depending on the consultancy's focus.

## 2.2. Advantages of Onsite HPC

Onsite HPC systems offer complete control over hardware and software configurations. They may provide performance benefits due to dedicated resources and reduced latency (Stergiou et al. 2018). Additionally, data remains on-premises, which can alleviate certain security and compliance concerns.

## 2.3. Disadvantages of Onsite HPC

The primary drawbacks of onsite HPC include:

- **High Initial Capital Expenditure:** Significant upfront investment is required for hardware, infrastructure, and installation.
- **Maintenance Costs:** Ongoing costs for maintenance, energy consumption, and staffing can be substantial.
- **Lack of Scalability:** Scaling up requires additional hardware purchases, leading to potential underutilization during periods of low demand.

## 2.4. Financial Cost Comparison

A detailed financial analysis over a five-year period highlights the cost implications of both options.

Table 3. Five-Year Total Cost Comparison

Option	Minimum Total Cost (\$)	Maximum Total Cost (\$)
Onsite HPC	453,780	700,780
Cloud Computing (AWS)	8,940	8,940
Cloud Computing (Azure)	8,940	8,940
Cloud Computing (GCP)	8,856	8,856
<b>Cost Savings with Cloud Computing</b>	<b>444,924</b>	<b>691,924</b>

## 2.5. Onsite HPC Costs

**Initial Setup Costs:** The purchase of HPC hardware, including high-performance servers, networking equipment, and storage systems, is estimated between \$80,000 and \$120,000. Infrastructure costs, encompassing cooling systems, power supplies, and additional networking equipment, add another \$20,000 to \$40,000. Thus, the total initial setup cost ranges from \$100,000 to \$160,000.

**Annual Operating Costs:** Maintenance and support contracts cost between \$10,000 and \$15,000 annually. Energy consumption is calculated based on an estimated 0.3 kWh per CPU hour at \$0.15 per kWh, resulting in an annual energy cost of \$756. Staffing requires a part-time HPC specialist, with salary expenses ranging from \$50,000 to \$70,000 per year. The total annual operating costs amount to \$60,756 to \$85,756.

**Hardware Upgrade Costs:** Upgrades are anticipated every 4-5 years, costing 50-70% of the initial hardware investment, equating to \$50,000 to \$112,000.

**Total Five-Year Cost:** Combining the initial setup, operating costs over five years, and hardware upgrades, the total expenditure ranges from \$453,780 to \$700,780.

## **2.6. Cloud Computing Costs**

**Amazon Web Services (AWS):** To fulfill the requirement of 1,400 CPU hours per month, AWS offers the c6a.large instance at \$0.085 per hour. The monthly compute cost is \$119, with additional storage and data transfer costs estimated at \$30 per month. The total monthly cost is \$149, leading to an annual cost of \$1,788 and a five-year cost of \$8,940. Potential discounts through reserved instances or spot instances can further reduce expenses.

**Microsoft Azure:** Azure's F2s\_v2 instance, also priced at approximately \$0.085 per hour, results in identical compute costs to AWS. Including storage and data transfer, the total monthly cost is \$149, with a five-year cost of \$8,940.

**Google Cloud Platform (GCP):** GCP's n2-standard-2 instance costs \$0.084 per hour, leading to a monthly compute cost of \$117.60. Including storage and data transfer, the total monthly cost is \$147.60. The annual cost is \$1,771.20, and the five-year cost is \$8,856. Sustained use discounts may reduce costs by up to 30%.

## **2.7. Total Cost Comparison**

Over five years, the total cost for cloud computing ranges from \$8,856 to \$8,940, significantly lower than the onsite HPC cost of \$453,780 to \$700,780. The cost savings with cloud computing amount to \$444,924 to \$691,924 over five years.

# **3. Recommendation**

Based on the analysis of operational and financial factors, it is recommended that the consultancy company adopts a cloud computing solution for its parallel processing needs.

**Significant Cost Savings:** The financial comparison demonstrates substantial cost savings with cloud computing, reducing expenditures by over \$440,000 over five years. This capital can be allocated to other strategic initiatives within the company.

**Operational Efficiency:** Cloud computing minimizes the need for in-house IT maintenance, allowing consultants to focus on core business activities. The responsibility for hardware maintenance and updates lies with the cloud provider, enhancing operational efficiency.

**Flexibility and Scalability:** Cloud services offer unparalleled flexibility, with the ability to adjust resources to match workload demands seamlessly. This scalability supports business growth without the need for additional infrastructure investment.

**Risk Mitigation:** Cloud computing eliminates risks associated with hardware failures and obsolescence. Providers offer high availability and disaster recovery solutions, ensuring business continuity.

## **3.1. Platform Selection**

While all three major cloud providers offer services that meet the company's needs, slight distinctions may influence the decision:

- **AWS:** With its extensive range of services and mature ecosystem, AWS is suitable for companies seeking a broad array of tools and global infrastructure.
- **Azure:** For firms already utilizing Microsoft products, Azure offers seamless integration and may provide operational efficiencies.
- **GCP:** Companies interested in advanced analytics and machine learning capabilities may prefer GCP due to its strengths in these areas.

Given the consultancy's size and requirements, any of these providers could be appropriate. A detailed assessment of

existing partnerships, staff expertise, and specific service offerings should guide the final selection. The consultancy should also consider any existing vendor relationships or preferred ecosystems when making the final decision.

### 3.2. Implementation Considerations

To maximize the benefits of cloud computing, the company should take the following steps during the implementation phase:

- **Leverage Cost Savings Opportunities:** Utilize reserved instances or savings plans for predictable workloads and explore spot instances for non-critical tasks to reduce costs.
- **Implement Cost Management Practices:** Use cloud monitoring tools to track resource utilization, set alerts for unexpected cost spikes, and establish budgets to manage cloud spending effectively.
- **Ensure Security and Compliance:** Establish robust cloud security policies, implement encryption for sensitive data, and leverage the cloud provider's compliance certifications to meet regulatory requirements.
- **Provide Staff Training:** Equip consultants with the necessary skills to use cloud services effectively, ensuring a smooth transition and minimizing resistance to the new system.

Table 4. Comparison of Cloud Computing and Onsite HPC

Criteria	Cloud Computing	Onsite HPC
<b>Scalability and Flexibility</b>	High; resources can be scaled up or down easily	Limited; scaling requires hardware investment
<b>Initial Capital Expenditure</b>	Low; pay-as-you-go model	High; significant upfront costs
<b>Maintenance Overhead</b>	Low; provider handles maintenance	High; requires dedicated IT staff
<b>Risk of Obsolescence</b>	Low; access to latest technology	High; hardware can become outdated
<b>Energy Consumption</b>	Lower; optimized data centers	Higher; increased operational costs
<b>Data Control</b>	Less control; data stored off-premises	Full control; data stored on-site
<b>Security and Compliance</b>	High; robust security measures and certifications	Dependent on in-house capabilities
<b>Total Cost over 5 Years</b>	\$8,856 – \$8,940	\$453,780 – \$700,780

### 3.3. Conclusion

Transitioning to cloud computing is a strategic decision that aligns with the consultancy company's operational needs and financial objectives. The substantial cost savings, coupled with enhanced scalability and reduced maintenance burdens, make cloud services the optimal choice over an onsite HPC system. By adopting cloud computing, the company positions itself to leverage advanced technologies, respond agilely to market demands, and focus its resources on delivering value to clients.

## Part II

# Parallel Programming Exercise

### 4. Step 1: Implementation and Optimization of Sequential Code

The C program was adapted to simulate the temperature distribution in a rectangular 2D grid with specific boundary conditions: the top boundary at 15°C, the bottom at 60°C, the left at 47°C, and the right at 100°C. The updated code uses a global preprocessor directive, `#define` keyword that creates a constant macro. These macros allow the program to use meaningful names throughout the code instead of magic numbers, making the code more readable and maintainable. When the code is compiled, the preprocessor replaces every occurrence of these macro names with their corresponding numerical values. These conditions were implemented by setting the corresponding edges of the grid to the specified temperatures.

In the code, after initializing the entire grid to an initial temperature (e.g., 30°C), the boundary conditions are applied:

```

1 // Set boundary conditions
2 for (int i = 1; i <= m; i++) {
3     t[i][0] = LEFT_TEMP; // Left boundary set to 47C
4     t[i][n + 1] = RIGHT_TEMP; // Right boundary set to 100C
5 }
6 for (int j = 1; j <= n; j++) {
7     t[0][j] = TOP_TEMP; // Top boundary set to 15C
8     t[m + 1][j] = BOTTOM_TEMP; // Bottom boundary set to 60C
9 }
10
11
12 // Set corner temperatures as average of adjacent boundaries
13 t[0][0] = (TOP_TEMP + LEFT_TEMP) / 2.0; // Top-left corner
14 t[0][n + 1] = (TOP_TEMP + RIGHT_TEMP) / 2.0; // Top-right corner
15 t[m + 1][0] = (BOTTOM_TEMP + LEFT_TEMP) / 2.0; // Bottom-left corner
16 t[m + 1][n + 1] = (BOTTOM_TEMP + RIGHT_TEMP) / 2.0; // Bottom-right corner

```

These code snippets ensure that the boundary temperatures are correctly set according to the problem's specifications. The corners are handled separately by averaging the temperatures of the adjacent boundaries to avoid anomalies at those points.

#### 4.1. Execution Time

To measure the execution time of the program, the `time.h` library is included, and timing functions are used:

```

1 #include <time.h> // For measuring execution time
2
3 // Start timing the execution
4 clock_t start = clock();
5
6 // ... [Jacobi iteration loop] ...
7
8 // Calculate execution time
9 clock_t end = clock();
10 double cpu_time_used = ((double)(end - start)) / CLOCKS_PER_SEC;
11
12 // Print results
13 printf("Execution time: %f seconds\n", cpu_time_used);

```

*Listing 1. Execution Time Measurement*

The `clock()` function records the processor time consumed by the program. By capturing the time before and after the main computation loop and calculating the difference, the execution time is obtained. This measurement is crucial for comparing performance across different grid sizes and compiler optimization levels.

#### 4.2. Adjustments for Realistic Execution Time Measurements

To obtain accurate execution time measurements, it is important to prevent the program from printing large amounts of data, which can significantly affect performance. In the code, the grid values are only printed if the grid size is small (e.g., 10x10 or smaller):

```

1 // Print grid values for small grids (10x10 or smaller)
2 if (m <= 10 && n <= 10) {
3     for (int i = 0; i <= m + 1; i++) {
4         for (int j = 0; j <= n + 1; j++) {
5             printf("%3.5lf ", t[i][j]);
6         }
7         printf("\n");
8     }
9 }
```

*Listing 2. Printing Grid Values for Small Grids*

By restricting the output for larger grids, the program avoids the overhead associated with I/O operations, leading to more accurate timing of the computational aspects.

#### 4.3. Automation of Testing with Bash Script

The provided Bash script automates the compilation and execution of the program across different optimization levels (`-O0` to `-O3`) and grid sizes (150x150, 200x200, 250x250). The script performs the following tasks:

##### 1. Compiler Selection:

The script checks for the availability of `gcc-14`; if not found, it defaults to `gcc`:

```

1 if command -v gcc-14 &> /dev/null; then
2     COMPILER=gcc-14
3 else
4     COMPILER=gcc
5 fi
```

*Listing 3. Compiler Selection*

##### 2. Compilation with Different Optimization Levels:

The program is compiled multiple times with different optimization flags:

```

1 for optmlvl in 0 1 2 3; do
2     $COMPILER -std=c99 -O${optmlvl} jacobi2d-Step1.c -o jacobi2d-Step1-O${optmlvl}
3     # Error checking omitted for brevity
4 done
```

*Listing 4. Compilation with Different Optimization Levels*

##### 3. Execution Across Grid Sizes and Optimization Levels:

The script runs the compiled programs for each combination of grid size and optimization level, capturing the output:

```

1  for size in 150 200 250; do
2      for optmlvl in 0 1 2 3; do
3          echo "Testing grid size of ${size}x${size} with optimization level -O${optmlvl}" >> jacobi2d-Step1.txt
4          ./jacobi2d-Step1-O${optmlvl} $size $size 0.000100 >> jacobi2d-Step1.txt
5          echo "-----" >> jacobi2d-Step1.txt
6      done
7  done

```

*Listing 5.* Automated Testing Loop for Grid Sizes and Optimization Levels

This automated approach ensures consistency and efficiency in data collection, facilitating the analysis of how different optimization levels impact execution time for various grid sizes.

#### 4.4. Results and Observations

The output collected demonstrates how compiler optimization levels affect the program's performance. Generally, higher optimization levels (-O1, -O2, -O3) result in reduced execution times compared to no optimization (-O0). The data also shows that as the grid size increases, the execution time increases, which is expected due to the larger computational workload.

#### 4.5. Considerations Regarding Compiler Optimizations

It is noted that aggressive compiler optimizations can potentially alter the program's behavior. In this case, the program was tested across all optimization levels without any observed issues. However, caution is advised, and validation checks should be included to ensure that the optimizations do not introduce unintended side effects.

*Table 5.* Execution Times for Different Grid Sizes and Optimization Levels (-ON)

Grid Size	Iterations	Maximum Difference ( $diff_{max}$ )	-O0 (s)	-O1 (s)	-O2 (s)	-O3 (s)
150x150	20,763	$9.999041 \times 10^{-5}$	5.250	1.570	1.090	1.060
200x200	32,108	$9.999248 \times 10^{-5}$	11.840	4.290	2.610	2.840
250x250	44,398	$9.999424 \times 10^{-5}$	25.720	9.420	6.290	6.050

#### 4.6. Conclusion

The modifications to the program successfully implemented the specified boundary conditions and enabled the measurement of execution time. The use of a Bash script to automate testing across different grid sizes and optimization levels provided valuable insights into the program's performance characteristics. These results form a solid foundation for further analysis and optimization in subsequent steps of the coursework.

## 5. Step 2: OpenMP Parallelization

The Jacobi 2D grid problem was extended from its sequential implementation to a parallel version using OpenMP. The primary objective was to modify the existing code to incorporate OpenMP directives for parallel execution, compile it with OpenMP support, and test its correctness using multiple threads. The performance was compared across different thread counts to observe the scalability and correctness of the parallel implementation.

### 5.1. Code Modifications for Parallel Execution

The original sequential code was adapted to leverage OpenMP for parallel processing. Key modifications were made to parallelize the computation-intensive sections while ensuring thread safety and correctness.

#### 1. Including OpenMP Header

The OpenMP library was included by adding the header file:

```
1 #include <omp.h> // Added OpenMP library for parallel execution
```

*Listing 6. OpenMP Header Inclusion*

#### 2. Parallel Initialization

The initialization of the temperature grids was parallelized using OpenMP directives (Chapman, Jost, and Pas 2007). The `#pragma omp parallel` directive was used to define a parallel region, and loops were parallelized using `#pragma omp for` with appropriate scheduling:

```
1 #pragma omp parallel
2 {
3     // Initialize interior points to 30.0
4     #pragma omp for collapse(2) schedule(static) nowait
5     // ... [initializing interior points] ...
6
7     // Set boundary temperatures
8     #pragma omp for schedule(static) nowait
9     // ... [setting left and right boundary temperatures] ...
10
11    #pragma omp for schedule(static)
12    // ... [setting top and bottom boundary temperatures] ...
13 }
```

*Listing 7. Parallel Initialization*

- **Use of `collapse(2)`:** This clause was used to collapse the two nested loops into a single loop for parallel execution (*OpenMP Application Program Interface* n.d.), enhancing load balancing among threads.
- **Scheduling and `nowait`:** Static scheduling was chosen for simplicity, and the `nowait` clause was used to eliminate unnecessary barriers, improving performance.

#### 3. Parallel Main Iteration Loop:

The main computation loop, which performs the Jacobi iteration, was parallelized:

```
1 while (difmax > tol) {
2     iter++;
3     difmax = 0.0;
```

```

4      #pragma omp parallel
5      {
6          // Compute new temperature values
7          #pragma omp for private(j) schedule(static) nowait
8          // ... [Compute new temperature values] ...
9
10         // Update temperatures and compute maximum difference
11         #pragma omp for private(j, diff) reduction(max:difmax)
12         // ... [Update temperatures and compute maximum difference] ...
13     }
14 }
15

```

Listing 8. Parallel Main Iteration Loop

- **Private Variables:** Variables `j` and `diff` were declared as private to ensure thread safety.
- **Reduction Clause:** The `reduction(max:difmax)` clause was used to correctly compute the maximum difference across all threads (Chapman, Jost, and Pas 2007).
- **No Wait Clauses:** The `nowait` clause was employed to eliminate unnecessary synchronization barriers between loops.

#### 4. Timing with OpenMP Functions:

The execution time was measured using OpenMP's timing functions (*OpenMP Application Program Interface n.d.*):

```

1 double start_time = omp_get_wtime();
2 // ... [computation] ...
3 double exec_time = omp_get_wtime() - start_time;
4 printf("Execution time: %f seconds\n", exec_time);

```

Listing 9. Timing with OpenMP Functions

The use of `omp_get_wtime()` provides high-resolution wall-clock timing suitable for parallel regions.

#### 5. Memory Allocation Optimization:

The arrays were allocated using contiguous memory blocks to improve cache performance:

```

1 double *t_data = (double *)malloc((m + 2) * (n + 2) * sizeof(double));
2 double *tnew_data = (double *)malloc((m + 2) * (n + 2) * sizeof(double));
3 double **t = (double **)malloc((m + 2) * sizeof(double *));
4 double **tnew = (double **)malloc((m + 2) * sizeof(double *));
5
6 // Setup 2D array pointers
7 for (i = 0; i < m + 2; i++) {
8     t[i] = &t_data[i * (n + 2)];
9     tnew[i] = &tnew_data[i * (n + 2)];
10}

```

Listing 10. Memory Allocation Optimization

This approach reduces memory fragmentation and enhances data locality, which is beneficial for parallel execution.

## 5.2. Automation of Testing with Bash Script

To systematically test the parallel program across different thread counts, a Bash script was employed. The script automates the compilation and execution process, ensuring consistent and repeatable tests.

### 1. Execution Across Thread Counts:

The script runs the program for a grid size of 20x20 using 1, 2, 4, and 8 threads:

```

1  for size in 20; do
2      for threads in 1 2 4 8; do
3          echo "Testing grid size of ${size}x${size} with ${threads} threads" >>
jacobi2d-Step2.txt
4          OMP_NUM_THREADS=$threads ./jacobi2d-Step2 $size $size 0.000100 >> jacobi2d-
Step2.txt
5          echo "-----" >> jacobi2d-Step2.txt
6      done
7  done

```

*Listing 11. Execution Across Thread Counts*

- **Environment Variable OMP\_NUM\_THREADS:** This variable sets the number of threads for OpenMP at runtime.
- **Output Redirection:** The results are appended to `jacobi2d-Step2.txt` for later analysis.

## 5.3. Results and Observations

The output obtained from running the tests demonstrates the correctness and performance of the parallel implementation. For a grid size of 20x20, the program was executed using 1, 2, 4, and 8 threads. The iteration counts, maximum differences, and execution times were recorded.

*Table 6. Performance Comparison for Different Thread Counts*

Thread Count	Iterations	Maximum Difference ( $dif_{max}$ )	Execution Time (s)
1	752	$9.929954 \times 10^{-5}$	0.009949
2	754	$9.832363 \times 10^{-5}$	0.008693
4	740	$9.894563 \times 10^{-5}$	0.006277
8	736	$9.817051 \times 10^{-5}$	0.005612

**Correctness Across Threads:** The iteration counts and maximum differences are consistent across different thread counts, indicating that the parallel code operates correctly.

**Performance Improvement:** There is a noticeable reduction in execution time as the number of threads increases, demonstrating the benefit of parallelization.

**Diminishing Returns:** The speedup from 4 to 8 threads is less significant than from 2 to 4 threads, which may be due to overheads associated with thread management and the small problem size.

For grids of size 20x20 or smaller, the program prints the final temperature distribution. This way, the program can be used to visualize the temperature distribution for small grids while maintaining performance for larger grids.

```

1  if (m <= 20 && n <= 20) {
2      // ... [print grid] ...
3  }

```

*Listing 12. Conditional Printing of Grid Values*





## 6. Step 3: Performance Analysis of OpenMP Implementation

The objective of this step is to evaluate the performance of the OpenMP implementation of the Jacobi method for solving partial differential equations. The analysis focuses on measuring the execution times and speedup achieved by varying the number of threads, optimization levels, and problem sizes. The experiments were conducted on the University's High-Performance Computing (HPC) system using the SLURM workload manager. This report presents the testing methodology, performance analysis, key findings, and recommendations based on the results obtained.

### 6.1. Testing Methodology

To comprehensively assess the performance, a systematic testing approach was adopted. A Bash script was developed to automate the compilation and execution of the program across different configurations. The script varied the following parameters:

**Optimization Level:** The program was compiled with different optimization levels (-O0, -O1, -O2, -O3) to assess the impact of compiler optimizations on performance.

**Thread Counts:** The number of threads was varied from 1 to 8 to evaluate the scalability of the parallel implementation.

**Grid Sizes:** The program was executed on grids of varying sizes (10x10, 50x50, 100x100, 150x150) to evaluate the impact of problem size on performance.

**Execution Time Measurement:** The execution times were recorded for each configuration to analyze the performance metrics.

The OpenMP directives were used to parallelize the computational loops, and the `textttomp_get_wtime()` function was utilized to measure execution times accurately. Print statements were minimized to reduce I/O overhead, especially for larger grid sizes. Each test case was executed, and the iteration counts, maximum difference (`texttdifmax`), and execution times were recorded.

### 6.2. Performance Analysis

Here are the performance metrics for the different configurations:

Table 7. Performance Metrics for Grid Size 150x150

Optimization Level	Threads	Iterations	Maximum Difference ( <i>difmax</i> )	Time (s)
<b>-O0</b>	1	20,763	$9.999041 \times 10^{-5}$	6.834
	2	20,036	$9.999421 \times 10^{-5}$	3.642
	4	19,412	$9.990420 \times 10^{-5}$	1.974
	8	19,558	$9.998308 \times 10^{-5}$	1.173
<b>-O1</b>	1	20,763	$9.999041 \times 10^{-5}$	1.761
	2	20,034	$9.999776 \times 10^{-5}$	1.154
	4	19,423	$9.999817 \times 10^{-5}$	0.635
	8	19,404	$9.999061 \times 10^{-5}$	0.381
<b>-O2</b>	1	20,763	$9.999041 \times 10^{-5}$	0.926
	2	20,032	$9.999765 \times 10^{-5}$	0.668
	4	19,423	$9.999867 \times 10^{-5}$	0.357
	8	19,387	$9.996107 \times 10^{-5}$	0.226
<b>-O3</b>	1	20,763	$9.999041 \times 10^{-5}$	0.869
	2	20,033	$9.999972 \times 10^{-5}$	0.633
	4	19,424	$9.999856 \times 10^{-5}$	0.339
	8	19,730	$9.999910 \times 10^{-5}$	0.221

Table 8. Performance Metrics for Grid Size 200x200

<b>Optimization Level</b>	<b>Threads</b>	<b>Iterations</b>	<b>Maximum Difference (<math>dif_{max}</math>)</b>	<b>Time (s)</b>
<b>-O0</b>	1	32,108	$9.999248 \times 10^{-5}$	18.041
	2	30,807	$9.999068 \times 10^{-5}$	9.308
	4	29,749	$9.998986 \times 10^{-5}$	4.915
	8	30,392	$9.998702 \times 10^{-5}$	2.917
<b>-O1</b>	1	32,108	$9.999248 \times 10^{-5}$	4.719
	2	30,799	$9.999040 \times 10^{-5}$	2.571
	4	29,717	$9.999629 \times 10^{-5}$	1.529
	8	29,989	$9.998269 \times 10^{-5}$	0.903
<b>-O2</b>	1	32,108	$9.999248 \times 10^{-5}$	2.560
	2	30,802	$9.999834 \times 10^{-5}$	1.451
	4	29,705	$9.998981 \times 10^{-5}$	0.933
	8	31,783	$9.999760 \times 10^{-5}$	0.539
<b>-O3</b>	1	32,108	$9.999248 \times 10^{-5}$	2.423
	2	30,811	$9.998999 \times 10^{-5}$	1.367
	4	30,818	$9.992472 \times 10^{-5}$	0.829
	8	29,988	$9.998336 \times 10^{-5}$	0.484

Table 9. Performance Metrics for Grid Size 250x250

<b>Optimization Level</b>	<b>Threads</b>	<b>Iterations</b>	<b>Maximum Difference (<math>dif_{max}</math>)</b>	<b>Time (s)</b>
<b>-O0</b>	1	44,398	$9.999424 \times 10^{-5}$	38.671
	2	42,381	$9.999290 \times 10^{-5}$	19.026
	4	42,321	$9.988293 \times 10^{-5}$	10.781
	8	42,684	$9.999864 \times 10^{-5}$	6.312
<b>-O1</b>	1	44,398	$9.999424 \times 10^{-5}$	10.183
	2	42,372	$9.999710 \times 10^{-5}$	5.149
	4	42,295	$9.999975 \times 10^{-5}$	3.203
	8	44,204	$9.999342 \times 10^{-5}$	1.934
<b>-O2</b>	1	44,398	$9.999424 \times 10^{-5}$	5.311
	2	42,383	$9.999860 \times 10^{-5}$	2.955
	4	42,328	$9.999459 \times 10^{-5}$	1.809
	8	41,784	$9.999386 \times 10^{-5}$	0.981
<b>-O3</b>	1	44,398	$9.999424 \times 10^{-5}$	5.051
	2	42,314	$9.999638 \times 10^{-5}$	2.718
	4	40,698	$9.999491 \times 10^{-5}$	1.722
	8	40,087	$9.999116 \times 10^{-5}$	0.883

#### 6.2.1. IMPACT OF PROBLEM SIZE

The execution times increased with the grid size due to the quadratic complexity of the Jacobi method. For single-threaded execution at optimization level -O0, the following base execution times were observed:

- $150 \times 150$  Grid: Approximately 6.83 seconds.
- $200 \times 200$  Grid: Approximately 18.04 seconds.
- $250 \times 250$  Grid: Approximately 38.67 seconds.

This increase aligns with the theoretical expectation that doubling the grid dimension results in a fourfold increase in computational work. The larger problem sizes provided more computational workload, which is beneficial for observing the effects of parallelization and optimization.

#### 6.2.2. IMPACT OF COMPILER OPTIMIZATION LEVELS

Compiler optimizations significantly affected performance. The optimization levels provided the following improvements over the baseline (-O0):

- -O1: Achieved a speedup of approximately 3 to 4 times compared to -O0.
- -O2: Further improved performance, yielding a speedup of approximately 5 to 7 times over -O0.
- -O3: Offered marginal improvements over -O2, typically around 5% to 10% faster.

The diminishing returns at higher optimization levels suggest that most beneficial optimizations occur at the -O2 level. The optimizations included loop unrolling, vectorization, and efficient memory access patterns, which are crucial for numerical computations.

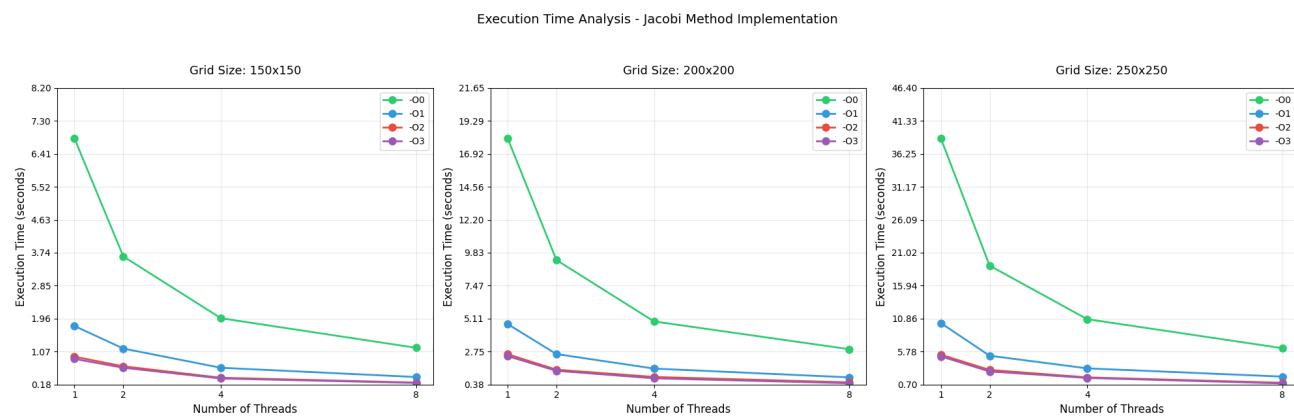


Figure 5. Execution Time Comparison for Different Grid Sizes and Optimization Levels

#### 6.2.3. PARALLEL SCALING ANALYSIS

The parallel performance was evaluated by measuring the speedup achieved with increasing thread counts. Speedup is defined as the ratio of the execution time with one thread to the execution time with multiple threads.

##### • 150×150 Grid:

- At -O0, an 8-thread execution achieved a speedup of approximately 5.8 times.
- At -O3, the speedup reduced to approximately 3.9 times with 8 threads.
- Diminishing returns were observed beyond 4 threads, likely due to insufficient workload per thread.

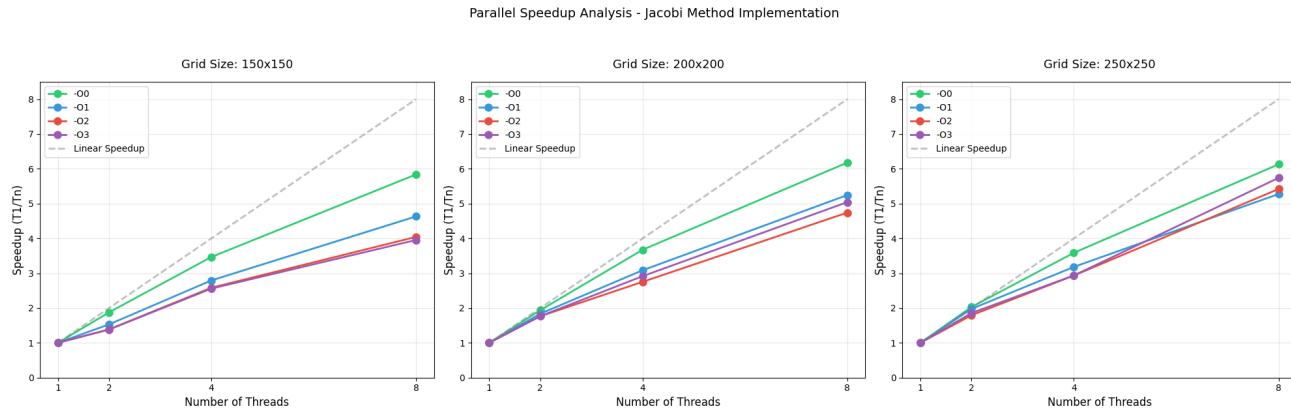
##### • 200×200 Grid:

- At -O0, an 8-thread execution achieved a speedup of approximately 6.2 times.
- At -O3, the speedup was approximately 5.0 times with 8 threads.
- Better scaling was observed compared to the smaller grid due to increased computational work.

##### • 250×250 Grid:

- At -O0, an 8-thread execution achieved a speedup of approximately 6.1 times.
- At -O3, the speedup increased to approximately 5.7 times with 8 threads.

- The highest parallel efficiency was noted for this grid size, attributed to a higher computation-to-communication ratio.



*Figure 6. Speedup Comparison for Different Grid Sizes and Optimization Levels*

## References

- Amazon Web Services (n.d.). *AWS Pricing*. Retrieved from <https://aws.amazon.com/pricing/>.
- Armbrust, Michael et al. (2010). “A View of Cloud Computing”. In: *Communications of the ACM* 53.4, pp. 50–58. DOI: [10.1145/1721654.1721672](https://doi.org/10.1145/1721654.1721672).
- Chapman, Barbara, Gabriele Jost, and Ruud Van Der Pas (2007). *Using OpenMP: Portable Shared Memory Parallel Programming*. MIT Press. ISBN: 978-0-262-03313-7.
- Dillon, Tharam, Chen Wu, and Elizabeth Chang (2010). “Cloud Computing: Issues and Challenges”. In: *Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications*. IEEE, pp. 27–33. DOI: [10.1109/AINA.2010.187](https://doi.org/10.1109/AINA.2010.187).
- Google Cloud Platform (n.d.). *Google Cloud Platform Pricing*. Retrieved from <https://cloud.google.com/pricing/>.
- Hashem, IAT et al. (2015). “The Rise of ”Big Data” on Cloud Computing: Review and Open Research Issues”. In: *Information Systems* 47, pp. 98–115. DOI: [10.1016/j.is.2014.07.006](https://doi.org/10.1016/j.is.2014.07.006).
- Leitner, Philipp and Juergen Cito (2016). “Patterns in the Chaos—A Study of Performance Variation and Predictability in Public IaaS Clouds”. In: *ACM Transactions on Internet Technology* 16.3, 15:1–15:23. DOI: [10.1145/2885497](https://doi.org/10.1145/2885497).
- Li, Zhenkun, Liam O’Brien, and He Zhang (2019). *Decision Support for Cloud Computing*. Springer. ISBN: 978-3-030-05645-2. DOI: [10.1007/978-3-030-05646-9](https://doi.org/10.1007/978-3-030-05646-9).
- Marinescu, Dan C. (2013). *Cloud Computing: Theory and Practice*. Morgan Kaufmann. ISBN: 978-0-12-404627-6.
- Microsoft Azure (n.d.). *Azure Pricing*. Retrieved from <https://azure.microsoft.com/en-us/pricing/>.
- OpenMP Application Program Interface* (n.d.). <https://www.openmp.org/>. Accessed: 2023-03-24.
- Rehman, Muhammad HU, Teh Ying Wah, and Farookh Khadeer Hussain (2018). “Towards Multi-Criteria Cloud Service Selection: A Systematic Literature Review”. In: *Journal of Network and Computer Applications* 104, pp. 1–13. DOI: [10.1016/j.jnca.2017.12.018](https://doi.org/10.1016/j.jnca.2017.12.018).
- Stergiou, Charalampos et al. (2018). “Secure Integration of IoT and Cloud Computing”. In: *Future Generation Computer Systems* 78, pp. 964–975. DOI: [10.1016/j.future.2016.11.031](https://doi.org/10.1016/j.future.2016.11.031).

*Due to space constraints, only relevant portions of the code have been included in the main report. The full code is provided in the attached files.*