**Pune Institute of Computer Technology**

Shape

Description automatically generated with medium confidence

**Department of Computer Engineering**

(2022- 2023)

# **“Implement merge sort and multithreaded merge sort.”**

**Submitted to the**

**Savitribai Phule Pune University**

**In partial fulfilment for the award of the Degree of**

**Bachelor of Engineering In Computer Engineering**

**By**

**Rohit James (41266)**

**Sufiya Sayyed (41278)**

**Aditya Wanjale (41282)**

Under the guidance of :

**Dr. Prajakta Khadkikar**

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sr. No** | **TITLE** | **Page no** |
| **1.** | Introduction | **3** |
| **2.** | Problem Statement | **4** |
| **3.** | Objectives | **4** |
| **4.** | Theory | **4** |
| **5.** | Conclusion | **15** |
| **6.** | References | **15** |

**Introduction**

## **Merge Sort**

Merge sort is a sorting technique which is based on divide and conquer technique where we divide the array into equal halves and then combine them in a sorted manner.

## **Algorithm to implement merge sort is**

* check if there is one element in the list then return the element.
* Else, Divide the data recursively into two halves until it can’t be divided further.
* Finally, merge the smaller lists into new lists in sorted order.

In merge sort, the problem is divided into two subproblems in every iteration.

Hence efficiency is increased drastically.

It follows the divide and conquer approach

Divide break the problem into 2 subproblem which continues until the problem set is left with one element only

Conquer basically merges the 2 sorted arrays into the original array

## **Multi-Threading**

In the operating system, **Threads** are the lightweight process which is responsible for executing the part of a task. Threads share common resources to execute the task concurrently.

**Multi-threading** is an implementation of multitasking where we can run multiple threads on a single processor to execute the tasks concurrently. It subdivides specific operations within a single application into individual threads. Each of the threads can run in parallel.

[Merge sort](https://www.prowaretech.com/articles/current/c-plus-plus/algorithms/merge-sort) is a good design for multi-threaded sorting because it allocates sub-arrays during the merge procedure thereby avoiding data collisions. This implementation breaks the array up into separate ranges and then runs its algorithm on each of them, but the data must be merged (sorted) in the end by the main thread. The more threads there are, the more unsorted the second to last array is thereby causing the final merge to take longer.

**Problem Statement**

Implement merge sort and multithreaded merge sort. Compare time required by both the algorithms. Also analyze the performance of each algorithm for the best case and the worst case.

**Objective**

Implement merge sort and multi-threaded merge sort. Compare their time complexities and analyze performance.

**Theory**

The **Merge Sort** algorithm is a sorting algorithm that is based on the **Divide and Conquer** paradigm. In this algorithm, the array is initially divided into two equal halves and then they are combined in a sorted manner.

## Merge Sort Working Process:

Think of it as a recursive algorithm continuously splits the array in half until it cannot be further divided. This means that if the array becomes empty or has only one element left, the dividing will stop, i.e. it is the base case to stop the recursion. If the array has multiple elements, split the array into halves and recursively invoke the merge sort on each of the halves. Finally, when both halves are sorted, the merge operation is applied. Merge operation is the process of taking two smaller sorted arrays and combining them to eventually make a larger one.

## Algorithm:

*step 1: start*

*step 2: declare array and left, right, mid variable*

*step 3: perform merge function.  
    if left > right  
        return  
    mid= (left+right)/2  
    mergesort(array, left, mid)  
    mergesort(array, mid+1, right)  
    merge(array, left, mid, right)*

*step 4: Stop*

*Multi-threaded Merge sort*

## *Multi-threading is way to improve parallelism by running the threads simultaneously in different cores of your processor. In this program, we’ll use 4 threads but you may change it according to the number of cores your processor has.*

## **For Example-:**

**In**−int arr[] = {3, 2, 1, 10, 8, 5, 7, 9, 4}

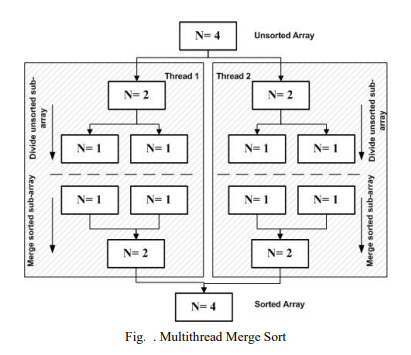
**Out**−Sorted array is: 1, 2, 3, 4, 5, 7, 8, 9, 10

**Explanation**−we are given an unsorted array with integer values. Now we will sort the array using merge sort with multithreading.

**In**−int arr[] = {5, 3, 1, 45, 32, 21, 50}

**Out**−Sorted array is: 1, 3, 5, 21, 32, 45, 50

**Explanation**−we are given an unsorted array with integer values. Now we will sort the array using merge sort with multithreading.



**CODE :**

**Merge Sort**

#include <iostream>

using namespace std;

void merge(int array[], int const left, int const mid,

int const right)

{

auto const subArrayOne = mid - left + 1;

auto const subArrayTwo = right - mid;

auto \*leftArray = new int[subArrayOne],

\*rightArray = new int[subArrayTwo];

for (auto i = 0; i < subArrayOne; i++)

leftArray[i] = array[left + i];

for (auto j = 0; j < subArrayTwo; j++)

rightArray[j] = array[mid + 1 + j];

auto indexOfSubArrayOne

= 0, // Initial index of first sub-array

indexOfSubArrayTwo

= 0; // Initial index of second sub-array

int indexOfMergedArray

= left; // Initial index of merged array

while (indexOfSubArrayOne < subArrayOne

&& indexOfSubArrayTwo < subArrayTwo) {

if (leftArray[indexOfSubArrayOne]

<= rightArray[indexOfSubArrayTwo]) {

array[indexOfMergedArray]

= leftArray[indexOfSubArrayOne];

indexOfSubArrayOne++;

}

else {

array[indexOfMergedArray]

= rightArray[indexOfSubArrayTwo];

indexOfSubArrayTwo++;

}

indexOfMergedArray++;

}

while (indexOfSubArrayOne < subArrayOne) {

array[indexOfMergedArray]

= leftArray[indexOfSubArrayOne];

indexOfSubArrayOne++;

indexOfMergedArray++;

}

while (indexOfSubArrayTwo < subArrayTwo) {

array[indexOfMergedArray]

= rightArray[indexOfSubArrayTwo];

indexOfSubArrayTwo++;

indexOfMergedArray++;

}

delete[] leftArray;

delete[] rightArray;

}

void mergeSort(int array[], int const begin, int const end)

{

if (begin >= end)

return; // Returns recursively

auto mid = begin + (end - begin) / 2;

mergeSort(array, begin, mid);

mergeSort(array, mid + 1, end);

merge(array, begin, mid, end);

}

void printArray(int A[], int size)

{

for (auto i = 0; i < size; i++)

cout << A[i] << " ";

}

int main()

{

int arr[] = { 12, 11, 13, 5, 6, 7 };

auto arr\_size = sizeof(arr) / sizeof(arr[0]);

cout << "Given array is \n";

printArray(arr, arr\_size);

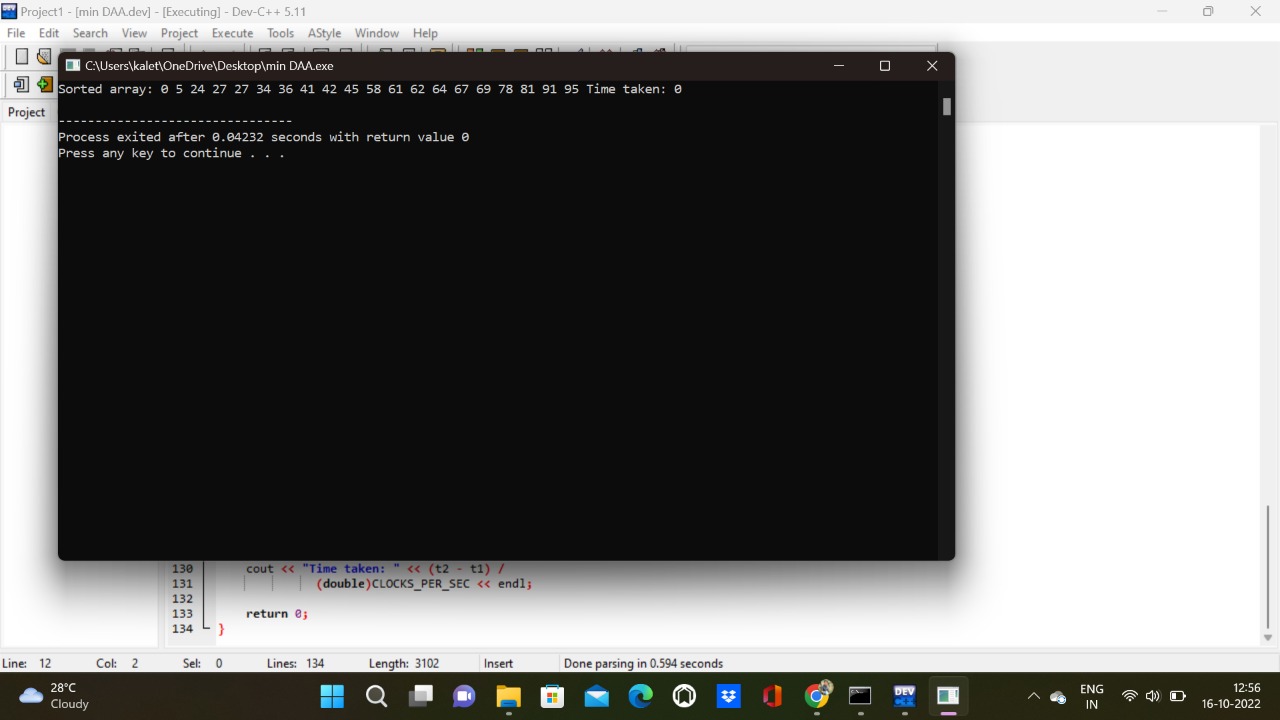
mergeSort(arr, 0, arr\_size - 1);

cout << "\nSorted array is \n";

printArray(arr, arr\_size);

return 0; }

OUTPUT :



**Multi-threaded Merge Sort**

#include <iostream>

#include <pthread.h>

#include <time.h>

#define MAX 20

#define THREAD\_MAX 4

using namespace std;

int a[MAX];

int part = 0;

void merge(int low, int mid, int high)

{

int\* left = new int[mid - low + 1];

int\* right = new int[high - mid];

int n1 = mid - low + 1, n2 = high - mid, i, j;

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

left[i] = a[i + low];

for (i = 0; i < n2; i++)

right[i] = a[i + mid + 1];

int k = low;

i = j = 0;

while (i < n1 && j < n2) {

if (left[i] <= right[j])

a[k++] = left[i++];

else

a[k++] = right[j++];

}

while (i < n1) {

a[k++] = left[i++];

}

while (j < n2) {

a[k++] = right[j++];

}

}

void merge\_sort(int low, int high)

{

int mid = low + (high - low) / 2;

if (low < high) {

merge\_sort(low, mid);

merge\_sort(mid + 1, high);

merge(low, mid, high);

}

}

void\* merge\_sort(void\* arg)

{

int thread\_part = part++;

int low = thread\_part \* (MAX / 4);

int high = (thread\_part + 1) \* (MAX / 4) - 1;

int mid = low + (high - low) / 2;

if (low < high) {

merge\_sort(low, mid);

merge\_sort(mid + 1, high);

merge(low, mid, high);

}

}

int main()

{

for (int i = 0; i < MAX; i++)

a[i] = rand() % 100;

clock\_t t1, t2;

t1 = clock();

pthread\_t threads[THREAD\_MAX];

for (int i = 0; i < THREAD\_MAX; i++)

pthread\_create(&threads[i], NULL, merge\_sort,

(void\*)NULL);

for (int i = 0; i < 4; i++)

pthread\_join(threads[i], NULL);

merge(0, (MAX / 2 - 1) / 2, MAX / 2 - 1);

merge(MAX / 2, MAX/2 + (MAX-1-MAX/2)/2, MAX - 1);

merge(0, (MAX - 1)/2, MAX - 1);

t2 = clock();

cout << "Sorted array: ";

for (int i = 0; i < MAX; i++)

cout << a[i] << " ";

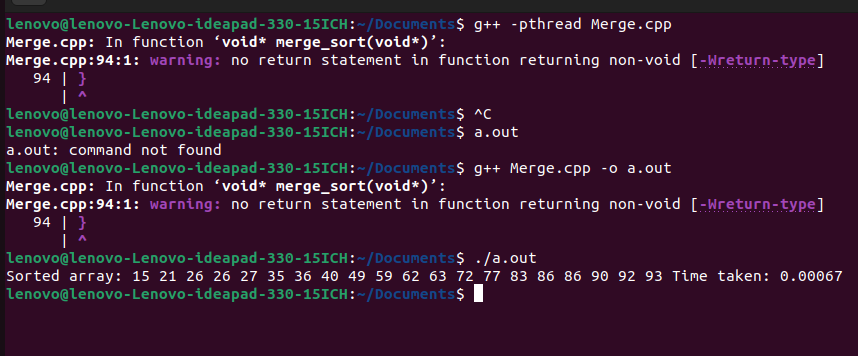
cout << "Time taken: " << (t2 - t1) /

(double)CLOCKS\_PER\_SEC << endl;

return 0;

}

OUTPUT :



**Time Complexity and Performance**

**Merge Sort**

Time Complexity: O(N log(N)),  Sorting arrays on different machines. Merge Sort is a recursive algorithm and time complexity can be expressed as following recurrence relation.

T(n) = 2T(n/2) + θ(n)

The above recurrence can be solved either using the Recurrence Tree method or the Master method. It falls in case II of the Master Method and the solution of the recurrence is θ(Nlog(N)). The time complexity of Merge Sort isθ(Nlog(N)) in all 3 cases (worst, average, and best) as merge sort always divides the array into two halves and takes linear time to merge two halves.

Auxiliary Space: O(n), In merge sort all elements are copied into an auxiliary array. So N auxiliary space is required for merge sort.

**Multi-threaded Merge Sort**

Multithread merge sort, creates thread recursively, and stops work when it reaches a certain size, with each thread locally sorting its data. Then threads merge their data by joining threads into a sorted main list. The multithread merge sort that have array of 4 elements to be sorted. Merge sort in multithread is based on the fact that the recursive calls run in parallel, so there is only one n/2 term with the time complexity (2): T(n) = Θ log(n) + Θ(n) = Θ(n)

**Conclusion**

Thus, We have implemented and compared time complexity and analysed performance of the Merge Sort and Multi-threaded Merge Sort.

**References**

* <https://www.tutorialspoint.com/merge-sort-using-multithreading-in-cplusplus>
* https://www.geeksforgeeks.org/merge-sort-using-multi-threading/
* <https://www.geeksforgeeks.org/merge-sort/>
* <https://en.wikipedia.org/wiki/Merge_sort>
* https://www.javatpoint.com/merge-sort