

Practical Work 5: The Longest Path

Nguyen Thanh Dat
Student ID: 23BI14091

1 Introduction

This labwork implements a MapReduce-based project designed to find the longest file path within a provided dataset. This simulation mimics the processing of a large-scale file system index, similar to the output of a `find` / command, by distributing the computational load across multiple parallel processes.

2 Mapper and Reducer Design

The system utilizes a distributed approach where the task of string length comparison is parallelized across multiple MPI ranks to ensure efficiency.

- **Mapper:** Each rank processes a specific chunk of the total path list. It performs a local comparison to identify the longest string (by character count) within its assigned subset.
- **Reducer:** The master process (Rank 0) gathers the "champions" (the longest path from each mapper) and performs a final reduction step to determine the global longest path.

3 Implementation

I implemented this system using `mpi4py` on a Windows environment. For realistic testing, I used a set of local absolute paths from my project directory.

```
# Data sample used for testing
data = [
    "C:/Users/Admin/Documents/usth/b3/distributed_system/
        ds2026/Practical-Work-1/server.py",
    "C:/Users/Admin/Documents/usth/b3/distributed_system/
        ds2026/Practical-Work-2/RPC/client.py",
    "C:/Windows/System32/drivers/etc/hosts",
    "C:/Program-Files/Python310/python.exe",
    "C:/Users/Admin/AppData/Local/MiKTeX/miktex/log/latexmk.
        log"
]
```

```
# Mapper: Local max calculation
local_longest = max(my_paths, key=len)

# Reducer: Global max selection
final_winner = max(all_candidates, key=len)
```

4 Verification

The system was executed with 4 parallel processes. The following screenshot demonstrates the successful identification of the longest path and its character count.

```
--- MapReduce Longest Path Results ---
Longest Path Found: C:/Users/Admin/Documents/usth/b3/distributed_system/ds2026/Practical Work 2/RPC/client.py
Character Length: 89
```

Figure 1: MapReduce Output: Longest Path Found (89 characters)