

SPM Project: Distributed External Memory Sorting

Marco Pampaloni
Department of Computer Science
m.pampaloni2@studenti.unipi.it
June 29, 2025

Contents

1. Introduction	2
1.1. Problem Statement	2
2. Implementation	2
2.1. Architecture	2
2.2. Shared-Memory Algorithm	2
2.2.1. FastFlow	2
2.2.2. OpenMP	2
2.3. Distributed Algorithm	2
2.3.1. MPI + OpenMP	2
3. Results	2
3.1. Performance Evaluation	2
3.2. Scaling	2
4. Cost Model Analysis	2
4.1. Bottlenecks	2
4.2. Main Optimizations	2
4.2.1. Buffered IO	2
4.2.2. Heap Arenas	2
4.2.3. Async IO	2
4.2.4. Computation-Communication Overlap	2
4.3. Challenges	2
5. Conclusions	2

1. Introduction

1.1. Problem Statement

2. Implementation

2.1. Architecture

2.2. Shared-Memory Algorithm

2.2.1. FastFlow

2.2.2. OpenMP

2.3. Distributed Algorithm

2.3.1. MPI + OpenMP

3. Results

3.1. Performance Evaluation

3.2. Scaling

4. Cost Model Analysis

4.1. Bottlenecks

4.2. Main Optimizations

4.2.1. Buffered IO

4.2.2. Heap Arenas

4.2.3. Async IO

4.2.4. Computation-Communication Overlap

4.3. Challenges

5. Conclusions