Experimental Setup.

For Part 3, I implemented a simple MapReduce pipeline using ECS/Fargate tasks and S3. The pipeline consists of a Splitter service that splits an input file into 3 chunks and uploads them to S3, three Mapper services that process one chunk each in parallel and upload intermediate JSON word-count maps to S3, and a Reducer service that merges the three JSON outputs into a final result file in S3. To evaluate performance, I executed the full Split → Map → Reduce pipeline four times and recorded the latency of each phase (Splitter time, total Mapper phase wall-clock time, and Reducer time).

Results (4 runs).

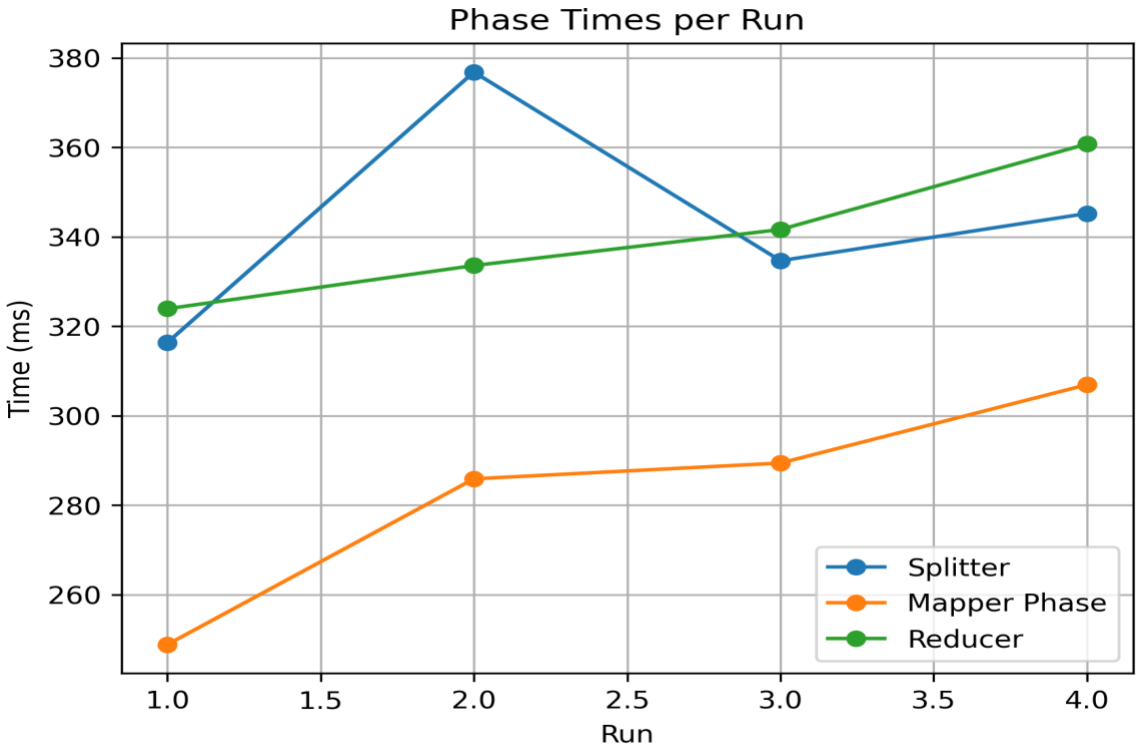
The end-to-end pipeline times (Splitter + Mapper phase + Reducer) were **889.008 ms, 996.224 ms, 965.642 ms**, and **1012.833 ms**. The **average total pipeline time** was **965.927 ms (~0.966 s)** with a standard deviation of **54.878 ms**, showing moderate run-to-run variability mainly due to network and service scheduling overhead.

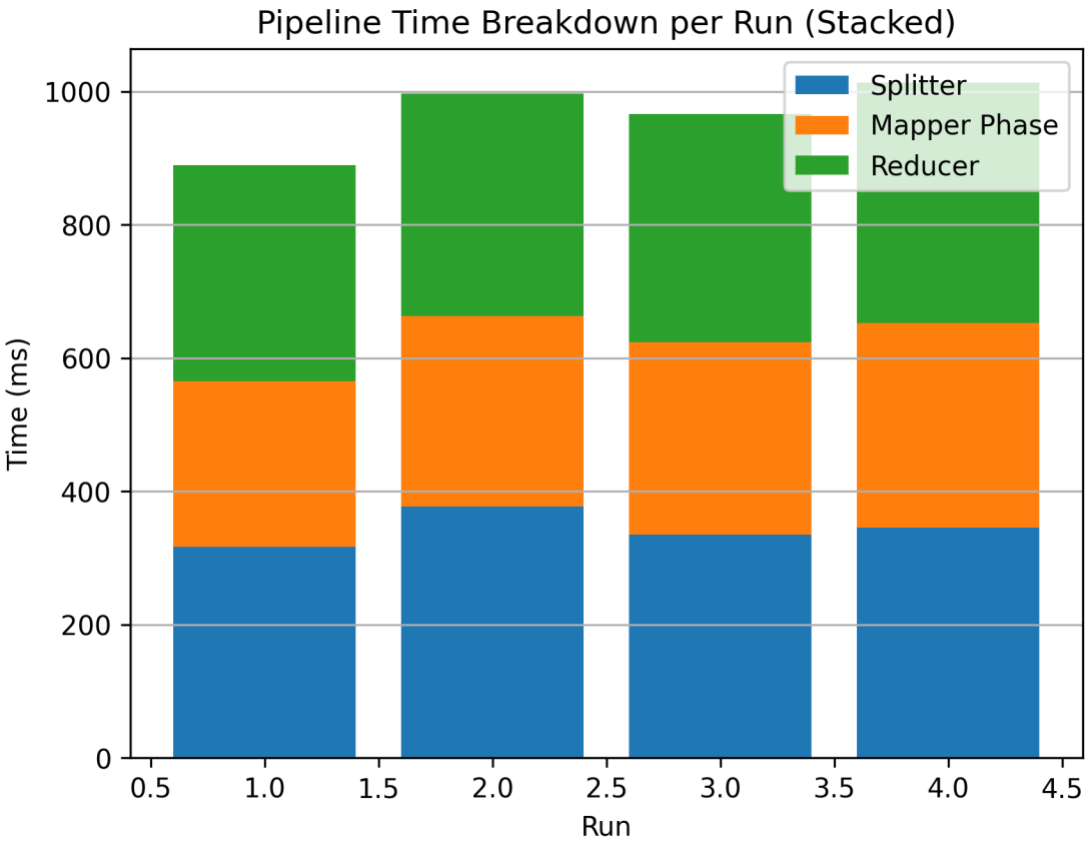
Phase timing observations.

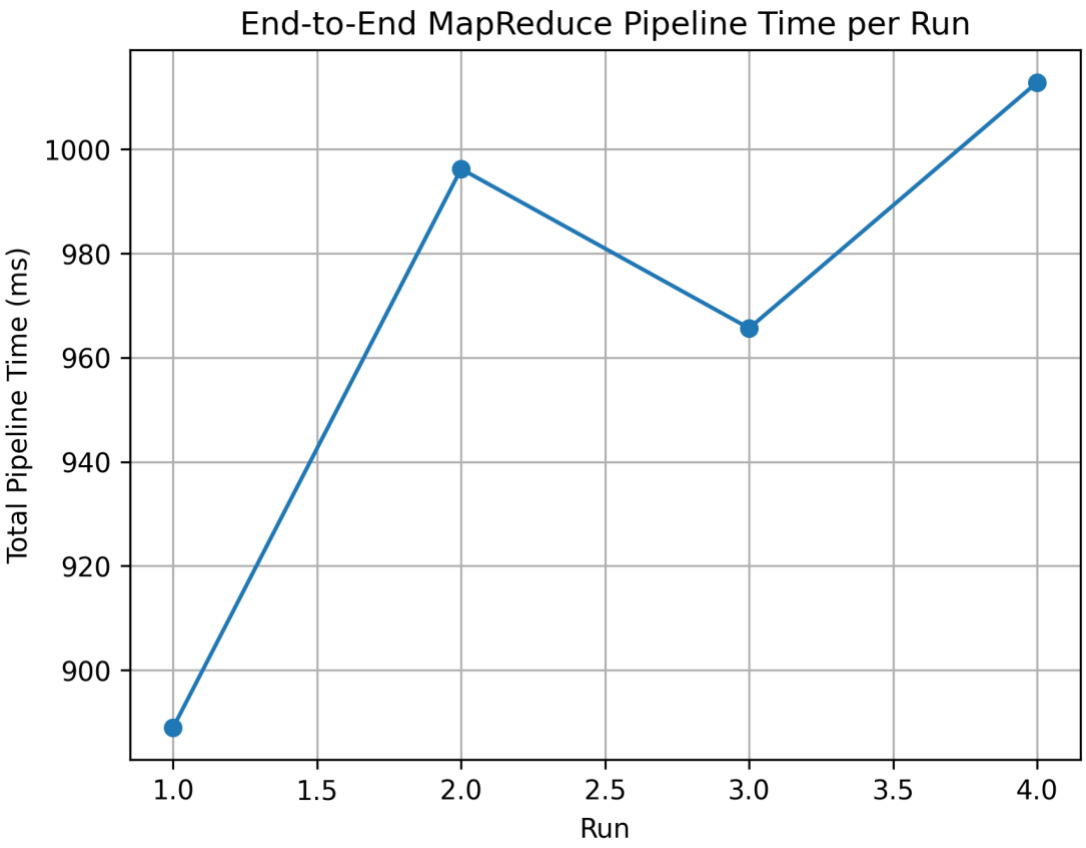
On average, the Splitter took **343.241 ms**, the Mapper phase took **282.736 ms**, and the Reducer took **339.949 ms**. The Mapper phase time represents the wall-clock time while running three mappers concurrently; therefore, it is approximately equal to the slowest mapper plus overhead, rather than the sum of all mapper times. This demonstrates parallel execution during the Map phase.

Discussion.

These measurements highlight that for smaller inputs, the total runtime is strongly influenced by overhead from HTTP requests, container networking, and S3 read/write operations. Even though the Map stage is parallelized across three tasks, the overall performance is limited by orchestration overhead and the sequential Split and Reduce stages. With larger input files (or more expensive per-chunk processing), the parallel Map phase would be expected to contribute more meaningful speedup relative to a single-machine baseline.


Phase Times per Run
Time (ms)
Run
Splitter
Mapper Phase
Reducer
380
360
340
320
300
280
260
1.0
1.5
2.0
2.5
3.0
3.5
4.0


Pipeline Time Breakdown per Run (Stacked)
Time (ms)
Run
Splitter
Mapper Phase
Reducer
1000
800
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5


End-to-End MapReduce Pipeline Time per Run
Total Pipeline Time (ms)
1000
980
960
940
920
900
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Run