



Kelas Sistem Basis Data 02

KELOMPOK 1

# HADOOP

SISTEM BASIS DATA



Kelas Sistem Basis Data 02

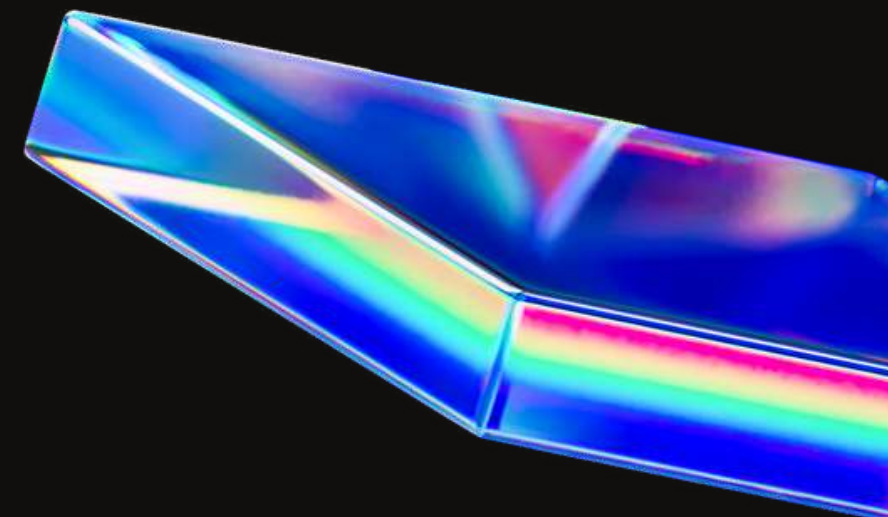
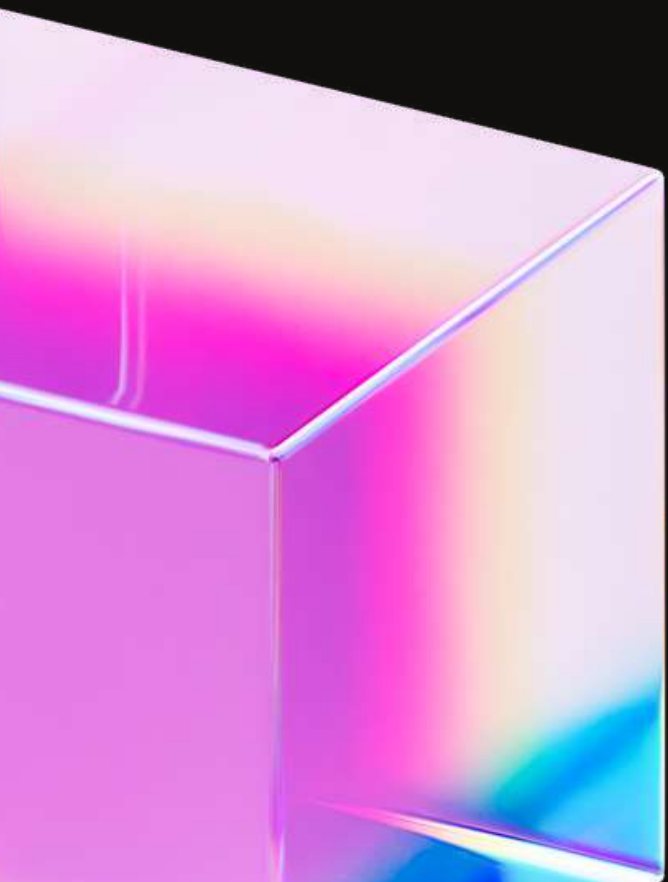
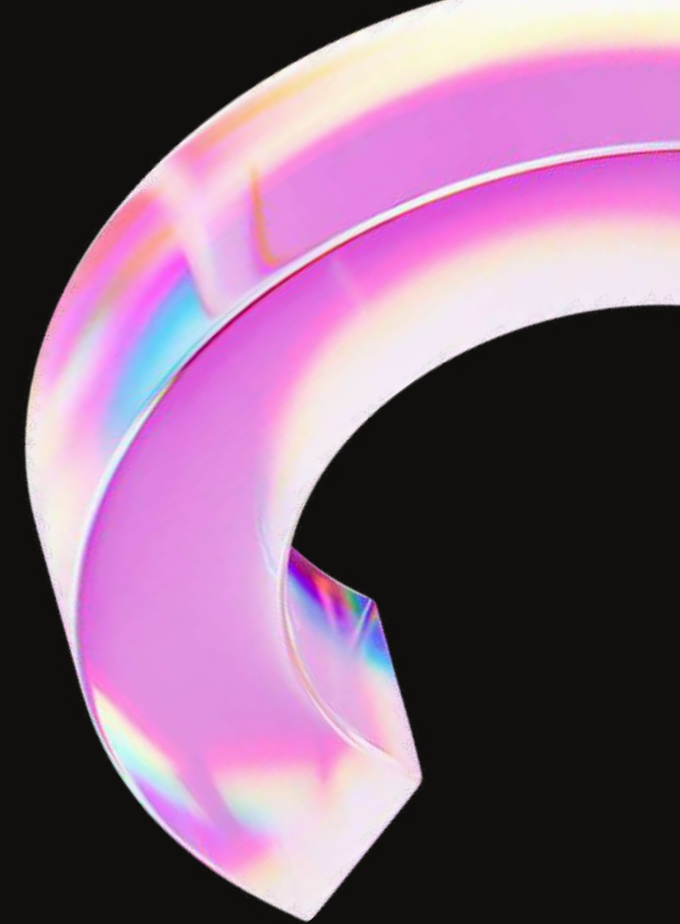
## KELOMPOK 1

**ROY OSWALDHA 2106731592**

**LEONARDO JEREMY PONGPARE 2106707914**

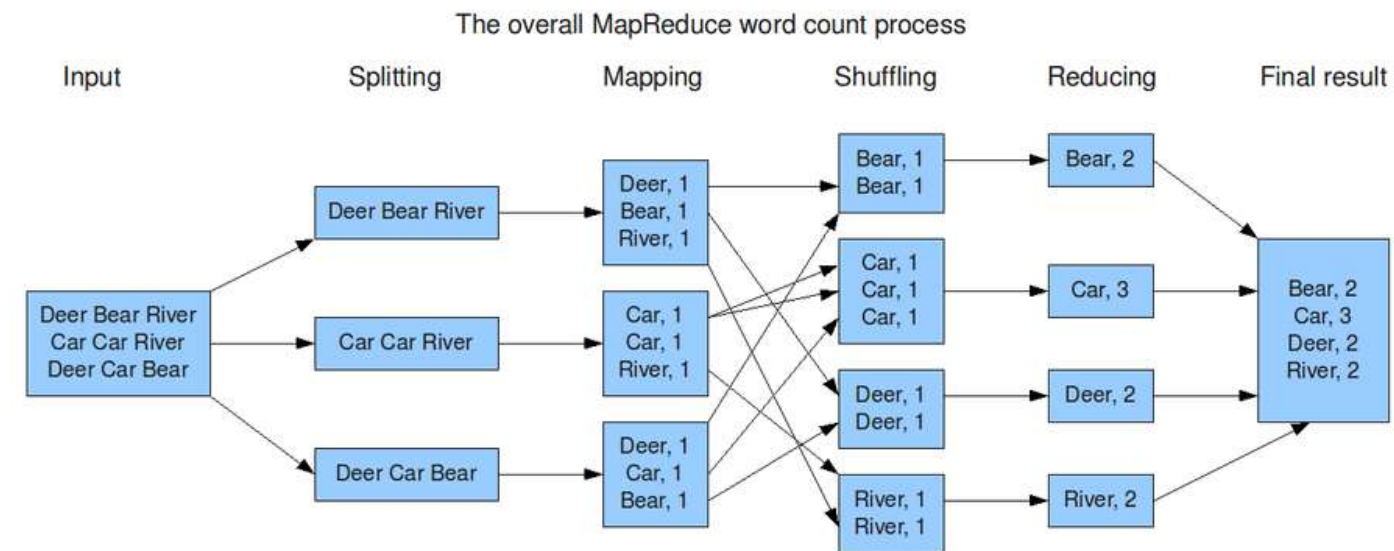
**LUTHFI MISBACHUL 2106706981**

**IVAN INDRAMATA 2106706981**





# MAP REDUCE



Source : <https://medium.com/pujanggateknologi/implementasi-sederhana-framework-mapreduce-8f80f22cc54f>

MapReduce digunakan untuk memproses dan menganalisis data besar secara terdistribusi.

Dalam pemrograman MapReduce dengan Java, data besar dibagi menjadi potongan-potongan kecil, kemudian setiap potongan ini diubah menjadi pasangan kunci-nilai dan diproses melalui tahap "Map". Selanjutnya, pasangan kunci-nilai tersebut dikelompokkan dan diproses melalui tahap "Reduce" untuk menghasilkan output akhir yang diinginkan. Hal ini memungkinkan pemrosesan data yang efisien di lingkungan komputasi terdistribusi seperti Apache Hadoop.

# WORDCOUNT

## 1. Tahap "Map":

- Dokumen atau kumpulan dokumen dibagi menjadi potongan-potongan kecil.
- Setiap potongan data diidentifikasi kata-katanya.
- Setiap kata diubah menjadi pasangan kunci-nilai, di mana kunci adalah kata dan nilai awalnya adalah 1.
- 

## 2. Tahap "Reduce":

- Pasangan kunci-nilai yang dihasilkan pada tahap "Map" dikelompokkan berdasarkan kata yang sama.
- Nilai-nilai yang terkait dengan setiap kata dijumlahkan.
- Pasangan kunci-nilai baru dihasilkan, di mana kunci adalah kata dan nilai adalah frekuensi kemunculan kata tersebut dalam dokumen atau kumpulan dokumen.
- 

## 3. Melalui proses MapReduce, WordCount dapat mencapai hasil akhir:

- Frekuensi kemunculan setiap kata dalam dokumen atau kumpulan dokumen dapat dihitung secara efisien.
- Proses tersebut dapat dilakukan dalam lingkungan pemrosesan terdistribusi seperti Apache Hadoop, yang memungkinkan pemrosesan data besar dengan efisiensi tinggi.



The image features four decorative, colorful, glossy swirls in the corners: top-left, top-right, bottom-left, and bottom-right. Each swirl is composed of multiple overlapping, translucent, rainbow-colored bands that create a vibrant, iridescent effect. The central text is positioned in the middle of the slide, between the top and bottom swirls.

Pada percobaan kali ini dilakukan perbandingan antara program wordcount yang menggunakan mapreduce dengan hadoop dengan program java tanpa mapreduce.

# WORDCOUNT

Using Java





# CODE IN JAVA

```
1 import java.io.BufferedReader;
2 import java.io.FileReader;
3 import java.io.IOException;
4 import java.util.HashMap;
5 import java.util.Map;
```

```
1 public class WordCount {
2     public static void main(String[] args) {
3         String filePath = "Z:\\My Projects\\Code\\Java\\WordCount\\500mb.txt";
4         int numIterations = 1;
5         long totalRuntime = 0;
6
7         for (int i = 0; i < numIterations; i++) {
8             long startTime = System.currentTimeMillis();
9
10            try (BufferedReader reader = new BufferedReader(new FileReader(filePath))) {
11                StringBuilder sb = new StringBuilder();
12                String line;
13
14                // Read text from the file
15                while ((line = reader.readLine()) != null) {
16                    sb.append(line).append("\n");
17                }
18
19                String text = sb.toString();
20
21                // Clean the text by removing non-alphanumeric characters and converting to lowercase
22                String cleanedText = cleanText(text);
23
24                // Split words into an array
25                String[] words = splitWords(cleanedText);
26
27                // Create a HashMap to count word frequencies
28                Map<String, Integer> wordCountMap = countWordFrequencies(words);
29
30                // Display the word count results
31                displayWordCountResults(wordCountMap);
32
33            } catch (IOException e) {
34                e.printStackTrace();
35            }
36        }
37    }
38 }
```

```
1         long endTime = System.currentTimeMillis();
2         long runtime = endTime - startTime;
3         totalRuntime += runtime;
4
5         System.out.println("Iteration " + (i + 1) + " Runtime: " + runtime + " milliseconds");
6     }
7
8     double averageRuntime = (double) totalRuntime / numIterations;
9     System.out.println("Average Runtime: " + averageRuntime + " milliseconds");
10 }
11
12 private static String cleanText(String text) {
13     return text.replaceAll("[^a-zA-Z0-9 ]", "").toLowerCase();
14 }
15
16 private static String[] splitWords(String text) {
17     return text.split(" ");
18 }
19
20 private static Map<String, Integer> countWordFrequencies(String[] words) {
21     Map<String, Integer> wordCountMap = new HashMap<>();
22
23     for (String word : words) {
24         wordCountMap.put(word, wordCountMap.getOrDefault(word, 0) + 1);
25     }
26
27     return wordCountMap;
28 }
29
30 private static void displayWordCountResults(Map<String, Integer> wordCountMap) {
31     for (Map.Entry<String, Integer> entry : wordCountMap.entrySet()) {
32         String word = entry.getKey();
33         int count = entry.getValue();
34         // System.out.println(word + ": " + count);
35     }
36 }
37 }
```



# RUNNING

Using Hadoop

[Back to Agenda](#)

```
require File.expand_path("../support/spec_helper.rb", __FILE__)
# Prevent database truncation if the environment is production
abort("The Rails environment is running in production mode!")
require 'spec_helper'
require 'rspec/rails'

require 'capybara/rspec'
require 'capybara/rails'

1
2 Capybara.javascript_driver = :webkit
13 Category.delete_all; Category.create
14 Shoulda::Matchers.configure do |config|
15   config.integrate do |with|
16     with.test_framework :rspec
17     with.library :rails
18   end
19 end
20
21 # Add additional requires below this line. Make sure they are not too early.
22 # Requires supporting ruby files with custom matchers (e.g. matchers.rb)
23 # spec/support/ and its subdirectories. These files will be loaded by default
24 # run as spec files by default. This will be overridden by the
25 # in _spec.rb will both be required and the files will be loaded.
26 # run twice. It is recommended that you do not require files in
27 # end with _spec.rb. You can configure the behavior by using
28 # option on the command line. For example:
29
30 No results found for 'mongoid'
31
32 mongoid
33
34 at buffer
```





# SPEKIFIKASI YANG DIGUNAKAN

Using Hadoop

Ryzen 7 5700U 8 Core 16 Threads

23,4 GB Memory DDR4 3200Mhz

SSD Kingston 512 GB

Dijalankan di ubuntu menggunakan virtualbox

Base memory : 7206 MB

Processor : 4

[Back to Agenda](#)



# RUNNING WITH HADOOP

Menggunakan 1 Mb File txt

```
2023-06-08 04:28:32,528 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:28:38,743 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:28:45,896 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:28:46,955 INFO mapreduce.Job: Job job_1686172668550_0001 completed successfully
```

**15 Second**

```
2023-06-08 04:36:35,869 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:36:41,011 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:36:47,162 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:36:48,227 INFO mapreduce.Job: Job job_1686172668550_0003 completed successfully
```

**13 Second**

```
2023-06-08 04:41:39,184 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:41:45,404 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:41:52,531 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:41:53,598 INFO mapreduce.Job: Job job_1686172668550_0005 completed successfully
```

**14 Second**

```
2023-06-08 04:44:20,215 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:44:25,360 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:44:31,510 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:44:32,550 INFO mapreduce.Job: Job job_1686172668550_0007 completed successfully
```

**12 Second**

```
2023-06-08 04:46:13,294 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:46:19,456 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:46:24,572 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:46:26,632 INFO mapreduce.Job: Job job_1686172668550_0009 completed successfully
```

**13 Second**



# RUNNING WITH HADOOP

Menggunakan 10 Mb File txt

```
2023-06-08 04:33:33,540 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:33:40,751 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:33:47,891 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:33:48,953 INFO mapreduce.Job: Job job_1686172668550_0002 completed successfully
```

**15 Second**

```
2023-06-08 04:39:43,876 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:39:51,076 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:39:58,271 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:39:59,323 INFO mapreduce.Job: Job job_1686172668550_0004 completed successfully
```

**16 Second**

```
2023-06-08 04:43:25,838 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:43:34,135 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:43:40,268 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:43:41,318 INFO mapreduce.Job: Job job_1686172668550_0006 completed successfully
```

**16 Second**

```
2023-06-08 04:45:18,730 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:45:26,131 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:45:33,291 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:45:34,361 INFO mapreduce.Job: Job job_1686172668550_0008 completed successfully
```

**16 Second**

```
2023-06-08 04:47:06,715 INFO mapreduce.Job: map 0% reduce 0%
2023-06-08 04:47:14,944 INFO mapreduce.Job: map 100% reduce 0%
2023-06-08 04:47:21,043 INFO mapreduce.Job: map 100% reduce 100%
2023-06-08 04:47:22,098 INFO mapreduce.Job: Job job_1686172668550_0010 completed successfully
```

**16 Second**



# RUNNING WITH HADOOP

Menggunakan 100MB File txt

```
2023-06-09 20:02:41,036 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:03:03,351 INFO mapreduce.Job: map 14% reduce 0%
2023-06-09 20:03:09,543 INFO mapreduce.Job: map 25% reduce 0%
2023-06-09 20:03:15,706 INFO mapreduce.Job: map 26% reduce 0%
2023-06-09 20:03:22,185 INFO mapreduce.Job: map 38% reduce 0%
2023-06-09 20:03:28,530 INFO mapreduce.Job: map 50% reduce 0%
2023-06-09 20:03:35,301 INFO mapreduce.Job: map 62% reduce 0%
2023-06-09 20:03:40,435 INFO mapreduce.Job: map 67% reduce 0%
2023-06-09 20:03:46,531 INFO mapreduce.Job: map 76% reduce 0%
2023-06-09 20:03:52,648 INFO mapreduce.Job: map 95% reduce 0%
2023-06-09 20:03:54,694 INFO mapreduce.Job: map 100% reduce 0%
2023-06-09 20:04:15,600 INFO mapreduce.Job: map 100% reduce 100%
```

**1 Menit 36 Detik**

```
2023-06-22 18:01:32,430 INFO mapreduce.Job: map 0% reduce 0%
2023-06-22 18:01:48,750 INFO mapreduce.Job: map 28% reduce 0%
2023-06-22 18:01:55,040 INFO mapreduce.Job: map 50% reduce 0%
2023-06-22 18:02:01,140 INFO mapreduce.Job: map 62% reduce 0%
2023-06-22 18:02:07,176 INFO mapreduce.Job: map 75% reduce 0%
2023-06-22 18:02:13,330 INFO mapreduce.Job: map 98% reduce 0%
2023-06-22 18:02:14,355 INFO mapreduce.Job: map 100% reduce 0%
2023-06-22 18:02:26,623 INFO mapreduce.Job: map 100% reduce 100%
```

**54 Detik**

```
2023-06-22 18:03:17,974 INFO mapreduce.Job: map 0% reduce 0%
2023-06-22 18:03:34,290 INFO mapreduce.Job: map 36% reduce 0%
2023-06-22 18:03:40,400 INFO mapreduce.Job: map 50% reduce 0%
2023-06-22 18:03:46,492 INFO mapreduce.Job: map 62% reduce 0%
2023-06-22 18:03:52,616 INFO mapreduce.Job: map 76% reduce 0%
2023-06-22 18:03:57,697 INFO mapreduce.Job: map 100% reduce 0%
2023-06-22 18:04:11,957 INFO mapreduce.Job: map 100% reduce 100%
```

**54 Detik**

```
2023-06-22 18:05:58,525 INFO mapreduce.Job: map 0% reduce 0%
2023-06-22 18:06:15,247 INFO mapreduce.Job: map 37% reduce 0%
2023-06-22 18:06:22,081 INFO mapreduce.Job: map 50% reduce 0%
2023-06-22 18:06:27,170 INFO mapreduce.Job: map 62% reduce 0%
2023-06-22 18:06:33,339 INFO mapreduce.Job: map 81% reduce 0%
2023-06-22 18:06:37,445 INFO mapreduce.Job: map 100% reduce 0%
2023-06-22 18:06:49,701 INFO mapreduce.Job: map 100% reduce 100%
```

**51 Detik**

```
2023-06-22 18:07:34,416 INFO mapreduce.Job: map 0% reduce 0%
2023-06-22 18:07:50,764 INFO mapreduce.Job: map 29% reduce 0%
2023-06-22 18:07:56,884 INFO mapreduce.Job: map 48% reduce 0%
2023-06-22 18:08:03,063 INFO mapreduce.Job: map 62% reduce 0%
2023-06-22 18:08:09,203 INFO mapreduce.Job: map 67% reduce 0%
2023-06-22 18:08:15,296 INFO mapreduce.Job: map 100% reduce 0%
2023-06-22 18:08:28,554 INFO mapreduce.Job: map 100% reduce 100%
```

**54 Detik**



# RUNNING WITH HADOOP

Menggunakan 200MB File txt

**1 Menit 38 Detik**

```
2023-06-09 20:07:13,616 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:07:31,390 INFO mapreduce.Job: map 29% reduce 0%
2023-06-09 20:07:37,592 INFO mapreduce.Job: map 42% reduce 0%
2023-06-09 20:07:43,437 INFO mapreduce.Job: map 53% reduce 0%
2023-06-09 20:07:49,195 INFO mapreduce.Job: map 68% reduce 0%
2023-06-09 20:07:51,285 INFO mapreduce.Job: map 74% reduce 0%
2023-06-09 20:07:56,143 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:08:02,392 INFO mapreduce.Job: map 83% reduce 0%
2023-06-09 20:08:07,552 INFO mapreduce.Job: map 88% reduce 0%
2023-06-09 20:08:11,631 INFO mapreduce.Job: map 88% reduce 17%
2023-06-09 20:08:13,671 INFO mapreduce.Job: map 97% reduce 17%
2023-06-09 20:08:15,698 INFO mapreduce.Job: map 100% reduce 17%
2023-06-09 20:08:17,739 INFO mapreduce.Job: map 100% reduce 67%
2023-06-09 20:08:24,014 INFO mapreduce.Job: map 100% reduce 82%
2023-06-09 20:08:30,154 INFO mapreduce.Job: map 100% reduce 99%
2023-06-09 20:08:31,179 INFO mapreduce.Job: map 100% reduce 100%
```

```
2023-06-09 20:25:18,748 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:25:37,670 INFO mapreduce.Job: map 29% reduce 0%
2023-06-09 20:25:45,405 INFO mapreduce.Job: map 42% reduce 0%
2023-06-09 20:25:51,385 INFO mapreduce.Job: map 59% reduce 0%
2023-06-09 20:25:53,474 INFO mapreduce.Job: map 72% reduce 0%
2023-06-09 20:25:57,450 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:26:03,480 INFO mapreduce.Job: map 83% reduce 0%
2023-06-09 20:26:08,701 INFO mapreduce.Job: map 90% reduce 0%
2023-06-09 20:26:13,834 INFO mapreduce.Job: map 100% reduce 17%
2023-06-09 20:26:19,945 INFO mapreduce.Job: map 100% reduce 80%
2023-06-09 20:26:26,131 INFO mapreduce.Job: map 100% reduce 98%
2023-06-09 20:26:27,159 INFO mapreduce.Job: map 100% reduce 100%
```

**1 Menit 9 Detik**

**1 Menit 6 Detik**

```
2023-06-09 20:27:12,887 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:27:31,395 INFO mapreduce.Job: map 37% reduce 0%
2023-06-09 20:27:37,635 INFO mapreduce.Job: map 53% reduce 0%
2023-06-09 20:27:44,723 INFO mapreduce.Job: map 74% reduce 0%
2023-06-09 20:27:50,654 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:27:58,203 INFO mapreduce.Job: map 83% reduce 0%
2023-06-09 20:28:02,627 INFO mapreduce.Job: map 95% reduce 0%
2023-06-09 20:28:03,640 INFO mapreduce.Job: map 95% reduce 17%
2023-06-09 20:28:04,667 INFO mapreduce.Job: map 100% reduce 17%
2023-06-09 20:28:08,725 INFO mapreduce.Job: map 100% reduce 73%
2023-06-09 20:28:14,848 INFO mapreduce.Job: map 100% reduce 89%
2023-06-09 20:28:18,903 INFO mapreduce.Job: map 100% reduce 100%
```

```
2023-06-09 20:28:39,663 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:28:57,061 INFO mapreduce.Job: map 30% reduce 0%
2023-06-09 20:29:03,167 INFO mapreduce.Job: map 44% reduce 0%
2023-06-09 20:29:10,636 INFO mapreduce.Job: map 61% reduce 0%
2023-06-09 20:29:13,751 INFO mapreduce.Job: map 73% reduce 0%
2023-06-09 20:29:15,773 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:29:21,861 INFO mapreduce.Job: map 83% reduce 0%
2023-06-09 20:29:27,996 INFO mapreduce.Job: map 89% reduce 0%
2023-06-09 20:29:30,039 INFO mapreduce.Job: map 89% reduce 17%
2023-06-09 20:29:34,095 INFO mapreduce.Job: map 100% reduce 17%
2023-06-09 20:29:36,129 INFO mapreduce.Job: map 100% reduce 67%
2023-06-09 20:29:42,307 INFO mapreduce.Job: map 100% reduce 83%
2023-06-22 17:54:16,026 INFO mapreduce.Job: map 100% reduce 100%
```

**1 Menit 7 Detik**

**1 Menit 3 Detik**

```
2023-06-09 20:23:03,759 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:23:23,369 INFO mapreduce.Job: map 32% reduce 0%
2023-06-09 20:23:29,390 INFO mapreduce.Job: map 49% reduce 0%
2023-06-09 20:23:34,673 INFO mapreduce.Job: map 64% reduce 0%
2023-06-09 20:23:37,736 INFO mapreduce.Job: map 74% reduce 0%
2023-06-09 20:23:41,064 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:23:47,203 INFO mapreduce.Job: map 84% reduce 0%
2023-06-09 20:23:53,379 INFO mapreduce.Job: map 94% reduce 0%
2023-06-09 20:23:54,393 INFO mapreduce.Job: map 94% reduce 17%
2023-06-09 20:23:55,405 INFO mapreduce.Job: map 100% reduce 17%
2023-06-09 20:24:00,502 INFO mapreduce.Job: map 100% reduce 76%
2023-06-09 20:24:06,785 INFO mapreduce.Job: map 100% reduce 90%
2023-06-09 20:24:10,876 INFO mapreduce.Job: map 100% reduce 100%
```





# RUNNING WITH HADOOP

Menggunakan 500MB File txt

```
2023-06-09 20:09:48,669 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:10:10,422 INFO mapreduce.Job: map 4% reduce 0%
2023-06-09 20:10:12,274 INFO mapreduce.Job: map 14% reduce 0%
2023-06-09 20:10:16,076 INFO mapreduce.Job: map 18% reduce 0%
2023-06-09 20:10:18,080 INFO mapreduce.Job: map 22% reduce 0%
2023-06-09 20:10:23,790 INFO mapreduce.Job: map 32% reduce 0%
2023-06-09 20:10:29,246 INFO mapreduce.Job: map 34% reduce 0%
2023-06-09 20:10:30,425 INFO mapreduce.Job: map 41% reduce 0%
2023-06-09 20:10:36,124 INFO mapreduce.Job: map 48% reduce 0%
2023-06-09 20:10:37,206 INFO mapreduce.Job: map 49% reduce 0%
2023-06-09 20:10:41,464 INFO mapreduce.Job: map 52% reduce 0%
2023-06-09 20:10:42,481 INFO mapreduce.Job: map 56% reduce 0%
2023-06-09 20:10:47,691 INFO mapreduce.Job: map 57% reduce 0%
2023-06-09 20:10:48,765 INFO mapreduce.Job: map 61% reduce 0%
2023-06-09 20:10:54,280 INFO mapreduce.Job: map 67% reduce 0%
2023-06-09 20:10:56,472 INFO mapreduce.Job: map 69% reduce 0%
2023-06-09 20:11:00,712 INFO mapreduce.Job: map 73% reduce 0%
2023-06-09 20:11:06,265 INFO mapreduce.Job: map 76% reduce 0%
2023-06-09 20:11:07,321 INFO mapreduce.Job: map 78% reduce 0%
2023-06-09 20:11:08,347 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:11:12,538 INFO mapreduce.Job: map 83% reduce 0%
2023-06-09 20:11:13,568 INFO mapreduce.Job: map 85% reduce 0%
2023-06-09 20:11:14,636 INFO mapreduce.Job: map 86% reduce 0%
2023-06-09 20:11:15,771 INFO mapreduce.Job: map 87% reduce 0%
2023-06-09 20:11:18,920 INFO mapreduce.Job: map 89% reduce 0%
2023-06-09 20:11:19,989 INFO mapreduce.Job: map 91% reduce 0%
2023-06-09 20:11:21,071 INFO mapreduce.Job: map 93% reduce 0%
2023-06-09 20:11:25,212 INFO mapreduce.Job: map 96% reduce 0%
2023-06-09 20:11:27,234 INFO mapreduce.Job: map 99% reduce 0%
2023-06-09 20:11:29,263 INFO mapreduce.Job: map 100% reduce 0%
2023-06-09 20:11:52,952 INFO mapreduce.Job: map 100% reduce 8%
2023-06-09 20:11:59,102 INFO mapreduce.Job: map 100% reduce 33%
2023-06-09 20:12:05,195 INFO mapreduce.Job: map 100% reduce 70%
2023-06-09 20:12:11,411 INFO mapreduce.Job: map 100% reduce 75%
2023-06-09 20:12:17,552 INFO mapreduce.Job: map 100% reduce 82%
2023-06-09 20:12:23,626 INFO mapreduce.Job: map 100% reduce 88%
2023-06-09 20:12:29,732 INFO mapreduce.Job: map 100% reduce 94%
2023-06-09 20:12:35,825 INFO mapreduce.Job: map 100% reduce 100%
```

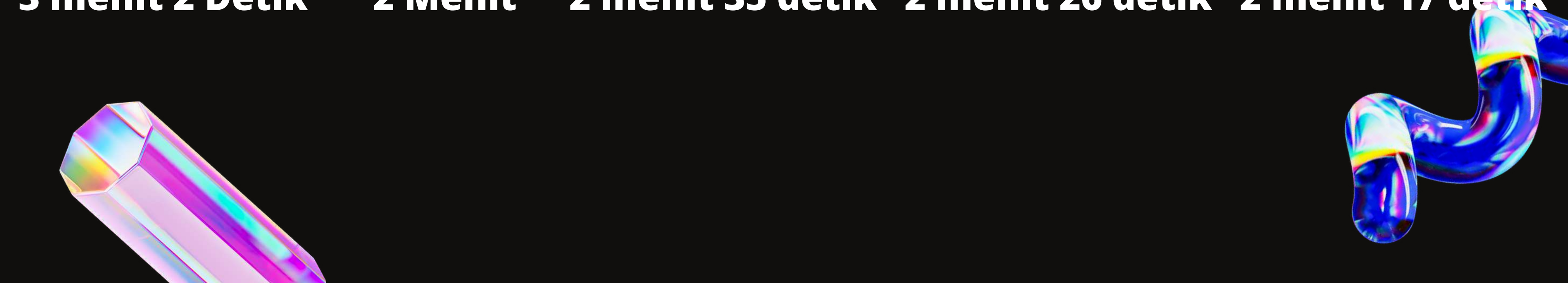
```
2023-06-22 17:58:07,204 INFO mapreduce.Job: map 0% reduce 0%
2023-06-22 17:58:27,174 INFO mapreduce.Job: map 19% reduce 0%
2023-06-22 17:58:34,127 INFO mapreduce.Job: map 27% reduce 0%
2023-06-22 17:58:39,192 INFO mapreduce.Job: map 31% reduce 0%
2023-06-22 17:58:41,181 INFO mapreduce.Job: map 37% reduce 0%
2023-06-22 17:58:46,218 INFO mapreduce.Job: map 45% reduce 0%
2023-06-22 17:58:48,181 INFO mapreduce.Job: map 48% reduce 0%
2023-06-22 17:58:52,026 INFO mapreduce.Job: map 55% reduce 0%
2023-06-22 17:58:54,173 INFO mapreduce.Job: map 58% reduce 0%
2023-06-22 17:58:58,315 INFO mapreduce.Job: map 60% reduce 0%
2023-06-22 17:58:59,388 INFO mapreduce.Job: map 65% reduce 0%
2023-06-22 17:59:00,430 INFO mapreduce.Job: map 67% reduce 0%
2023-06-22 17:59:04,507 INFO mapreduce.Job: map 74% reduce 0%
2023-06-22 17:59:05,522 INFO mapreduce.Job: map 80% reduce 0%
2023-06-22 17:59:10,590 INFO mapreduce.Job: map 91% reduce 0%
2023-06-22 17:59:11,687 INFO mapreduce.Job: map 94% reduce 0%
2023-06-22 17:59:12,756 INFO mapreduce.Job: map 96% reduce 0%
2023-06-22 17:59:14,853 INFO mapreduce.Job: map 98% reduce 0%
2023-06-22 17:59:16,904 INFO mapreduce.Job: map 100% reduce 0%
2023-06-22 17:59:37,591 INFO mapreduce.Job: map 100% reduce 67%
2023-06-22 17:59:43,683 INFO mapreduce.Job: map 100% reduce 72%
2023-06-22 17:59:49,785 INFO mapreduce.Job: map 100% reduce 79%
2023-06-22 17:59:55,858 INFO mapreduce.Job: map 100% reduce 87%
2023-06-22 18:00:01,943 INFO mapreduce.Job: map 100% reduce 94%
2023-06-22 18:00:07,007 INFO mapreduce.Job: map 100% reduce 100%
```

```
2023-06-09 20:16:21,022 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:16:42,583 INFO mapreduce.Job: map 10% reduce 0%
2023-06-09 20:16:43,610 INFO mapreduce.Job: map 13% reduce 0%
2023-06-09 20:16:47,940 INFO mapreduce.Job: map 15% reduce 0%
2023-06-09 20:16:48,982 INFO mapreduce.Job: map 16% reduce 0%
2023-06-09 20:16:50,133 INFO mapreduce.Job: map 21% reduce 0%
2023-06-09 20:16:54,639 INFO mapreduce.Job: map 26% reduce 0%
2023-06-09 20:16:55,788 INFO mapreduce.Job: map 28% reduce 0%
2023-06-09 20:16:56,838 INFO mapreduce.Job: map 31% reduce 0%
2023-06-09 20:17:01,995 INFO mapreduce.Job: map 35% reduce 0%
2023-06-09 20:17:03,425 INFO mapreduce.Job: map 40% reduce 0%
2023-06-09 20:17:06,995 INFO mapreduce.Job: map 45% reduce 0%
2023-06-09 20:17:10,461 INFO mapreduce.Job: map 50% reduce 0%
2023-06-09 20:17:13,866 INFO mapreduce.Job: map 54% reduce 0%
2023-06-09 20:17:15,988 INFO mapreduce.Job: map 56% reduce 0%
2023-06-09 20:17:17,009 INFO mapreduce.Job: map 58% reduce 0%
2023-06-09 20:17:19,118 INFO mapreduce.Job: map 60% reduce 0%
2023-06-09 20:17:20,225 INFO mapreduce.Job: map 61% reduce 0%
2023-06-09 20:17:23,421 INFO mapreduce.Job: map 62% reduce 0%
2023-06-09 20:17:24,610 INFO mapreduce.Job: map 65% reduce 0%
2023-06-09 20:17:25,677 INFO mapreduce.Job: map 67% reduce 0%
2023-06-09 20:17:29,768 INFO mapreduce.Job: map 69% reduce 0%
2023-06-09 20:17:31,850 INFO mapreduce.Job: map 74% reduce 0%
2023-06-09 20:17:35,977 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:17:38,041 INFO mapreduce.Job: map 83% reduce 0%
2023-06-09 20:17:42,110 INFO mapreduce.Job: map 88% reduce 0%
2023-06-09 20:17:44,359 INFO mapreduce.Job: map 92% reduce 0%
2023-06-09 20:17:46,680 INFO mapreduce.Job: map 93% reduce 0%
2023-06-09 20:17:47,696 INFO mapreduce.Job: map 95% reduce 0%
2023-06-09 20:17:48,715 INFO mapreduce.Job: map 97% reduce 0%
2023-06-09 20:17:49,729 INFO mapreduce.Job: map 100% reduce 25%
2023-06-09 20:18:12,600 INFO mapreduce.Job: map 100% reduce 25%
2023-06-09 20:18:17,767 INFO mapreduce.Job: map 100% reduce 67%
2023-06-09 20:18:23,863 INFO mapreduce.Job: map 100% reduce 74%
2023-06-09 20:18:29,987 INFO mapreduce.Job: map 100% reduce 81%
2023-06-09 20:18:36,089 INFO mapreduce.Job: map 100% reduce 88%
2023-06-09 20:18:42,309 INFO mapreduce.Job: map 100% reduce 94%
2023-06-09 20:18:47,672 INFO mapreduce.Job: map 100% reduce 100%
```

```
2023-06-09 20:19:21,333 INFO mapreduce.Job: map 0% reduce 0%
2023-06-09 20:19:40,332 INFO mapreduce.Job: map 9% reduce 0%
2023-06-09 20:19:41,469 INFO mapreduce.Job: map 18% reduce 0%
2023-06-09 20:19:46,373 INFO mapreduce.Job: map 20% reduce 0%
2023-06-09 20:19:47,452 INFO mapreduce.Job: map 25% reduce 0%
2023-06-09 20:19:53,266 INFO mapreduce.Job: map 31% reduce 0%
2023-06-09 20:20:00,382 INFO mapreduce.Job: map 40% reduce 0%
2023-06-09 20:20:09,037 INFO mapreduce.Job: map 50% reduce 0%
2023-06-09 20:20:13,996 INFO mapreduce.Job: map 59% reduce 0%
2023-06-09 20:20:19,407 INFO mapreduce.Job: map 60% reduce 0%
2023-06-09 20:20:20,551 INFO mapreduce.Job: map 63% reduce 0%
2023-06-09 20:20:26,044 INFO mapreduce.Job: map 68% reduce 0%
2023-06-09 20:20:32,444 INFO mapreduce.Job: map 74% reduce 0%
2023-06-09 20:20:33,503 INFO mapreduce.Job: map 77% reduce 0%
2023-06-09 20:20:37,802 INFO mapreduce.Job: map 79% reduce 0%
2023-06-09 20:20:38,842 INFO mapreduce.Job: map 87% reduce 0%
2023-06-09 20:20:42,053 INFO mapreduce.Job: map 88% reduce 0%
2023-06-09 20:20:44,085 INFO mapreduce.Job: map 90% reduce 0%
2023-06-09 20:20:45,117 INFO mapreduce.Job: map 94% reduce 0%
2023-06-09 20:20:48,332 INFO mapreduce.Job: map 98% reduce 0%
2023-06-09 20:20:49,354 INFO mapreduce.Job: map 100% reduce 0%
2023-06-09 20:21:13,184 INFO mapreduce.Job: map 100% reduce 25%
2023-06-09 20:21:19,309 INFO mapreduce.Job: map 100% reduce 69%
2023-06-09 20:21:25,471 INFO mapreduce.Job: map 100% reduce 75%
2023-06-09 20:21:31,571 INFO mapreduce.Job: map 100% reduce 82%
2023-06-09 20:21:37,768 INFO mapreduce.Job: map 100% reduce 88%
2023-06-09 20:21:43,910 INFO mapreduce.Job: map 100% reduce 96%
2023-06-09 20:21:47,954 INFO mapreduce.Job: map 100% reduce 100%
```

```
2023-06-22 17:54:43,402 INFO mapreduce.Job: map 0% reduce 0%
2023-06-22 17:55:02,949 INFO mapreduce.Job: map 15% reduce 0%
2023-06-22 17:55:04,191 INFO mapreduce.Job: map 20% reduce 0%
2023-06-22 17:55:09,355 INFO mapreduce.Job: map 27% reduce 0%
2023-06-22 17:55:10,661 INFO mapreduce.Job: map 30% reduce 0%
2023-06-22 17:55:15,451 INFO mapreduce.Job: map 37% reduce 0%
2023-06-22 17:55:16,900 INFO mapreduce.Job: map 39% reduce 0%
2023-06-22 17:55:21,578 INFO mapreduce.Job: map 46% reduce 0%
2023-06-22 17:55:22,603 INFO mapreduce.Job: map 49% reduce 0%
2023-06-22 17:55:28,039 INFO mapreduce.Job: map 55% reduce 0%
2023-06-22 17:55:29,177 INFO mapreduce.Job: map 58% reduce 0%
2023-06-22 17:55:35,261 INFO mapreduce.Job: map 67% reduce 0%
2023-06-22 17:55:40,415 INFO mapreduce.Job: map 69% reduce 0%
2023-06-22 17:55:41,459 INFO mapreduce.Job: map 75% reduce 0%
2023-06-22 17:55:46,591 INFO mapreduce.Job: map 80% reduce 0%
2023-06-22 17:55:47,603 INFO mapreduce.Job: map 85% reduce 0%
2023-06-22 17:55:52,782 INFO mapreduce.Job: map 91% reduce 0%
2023-06-22 17:55:53,811 INFO mapreduce.Job: map 93% reduce 0%
2023-06-22 17:55:56,040 INFO mapreduce.Job: map 95% reduce 0%
2023-06-22 17:55:58,102 INFO mapreduce.Job: map 96% reduce 0%
2023-06-22 17:55:59,132 INFO mapreduce.Job: map 99% reduce 0%
2023-06-22 17:56:01,182 INFO mapreduce.Job: map 100% reduce 0%
2023-06-22 17:56:19,888 INFO mapreduce.Job: map 100% reduce 17%
2023-06-22 17:56:32,368 INFO mapreduce.Job: map 100% reduce 67%
2023-06-22 17:56:38,501 INFO mapreduce.Job: map 100% reduce 73%
2023-06-22 17:56:44,590 INFO mapreduce.Job: map 100% reduce 80%
2023-06-22 17:56:50,676 INFO mapreduce.Job: map 100% reduce 87%
2023-06-22 17:56:57,245 INFO mapreduce.Job: map 100% reduce 96%
2023-06-22 17:57:00,281 INFO mapreduce.Job: map 100% reduce 100%
```

3 menit 2 Detik      2 Menit      2 menit 35 detik      2 menit 26 detik      2 menit 17 detik





# RUNNING WITH HADOOP

Menggunakan 1GB File txt

**3 Menit 13 Detik**

```
2023-06-09 19:56:12,951 INFO mapreduce.Job: map 75% reduce 4%
2023-06-09 19:56:19,354 INFO mapreduce.Job: map 75% reduce 13%
2023-06-09 19:56:25,826 INFO mapreduce.Job: map 75% reduce 17%
2023-06-09 19:56:34,800 INFO mapreduce.Job: map 76% reduce 17%
2023-06-09 19:56:36,279 INFO mapreduce.Job: map 78% reduce 17%
2023-06-09 19:56:41,211 INFO mapreduce.Job: map 78% reduce 21%
2023-06-09 19:56:42,308 INFO mapreduce.Job: map 80% reduce 21%
2023-06-09 19:56:47,446 INFO mapreduce.Job: map 80% reduce 25%
2023-06-09 19:56:48,619 INFO mapreduce.Job: map 82% reduce 25%
2023-06-09 19:56:53,925 INFO mapreduce.Job: map 84% reduce 25%
2023-06-09 19:56:54,930 INFO mapreduce.Job: map 85% reduce 25%
2023-06-09 19:57:02,258 INFO mapreduce.Job: map 86% reduce 25%
2023-06-09 19:57:03,277 INFO mapreduce.Job: map 87% reduce 25%
2023-06-09 19:57:07,266 INFO mapreduce.Job: map 89% reduce 25%
2023-06-09 19:57:09,340 INFO mapreduce.Job: map 91% reduce 25%
2023-06-09 19:57:15,230 INFO mapreduce.Job: map 92% reduce 25%
2023-06-09 19:57:20,351 INFO mapreduce.Job: map 95% reduce 25%
2023-06-09 19:57:21,367 INFO mapreduce.Job: map 96% reduce 25%
2023-06-09 19:57:23,542 INFO mapreduce.Job: map 98% reduce 25%
2023-06-09 19:57:25,900 INFO mapreduce.Job: map 98% reduce 29%
2023-06-09 19:57:26,924 INFO mapreduce.Job: map 100% reduce 29%
2023-06-09 19:57:33,288 INFO mapreduce.Job: map 100% reduce 68%
2023-06-09 19:57:41,445 INFO mapreduce.Job: map 100% reduce 72%
2023-06-09 19:57:46,374 INFO mapreduce.Job: map 100% reduce 76%
2023-06-09 19:57:51,478 INFO mapreduce.Job: map 100% reduce 79%
2023-06-09 19:57:57,623 INFO mapreduce.Job: map 100% reduce 81%
2023-06-09 19:58:03,774 INFO mapreduce.Job: map 100% reduce 84%
2023-06-09 19:58:09,895 INFO mapreduce.Job: map 100% reduce 85%
2023-06-09 19:58:16,164 INFO mapreduce.Job: map 100% reduce 87%
2023-06-09 19:58:22,436 INFO mapreduce.Job: map 100% reduce 88%
2023-06-09 19:58:28,068 INFO mapreduce.Job: map 100% reduce 89%
2023-06-09 19:58:34,621 INFO mapreduce.Job: map 100% reduce 90%
2023-06-09 19:58:40,922 INFO mapreduce.Job: map 100% reduce 91%
2023-06-09 19:58:47,039 INFO mapreduce.Job: map 100% reduce 92%
2023-06-09 19:58:52,522 INFO mapreduce.Job: map 100% reduce 93%
2023-06-09 19:58:58,893 INFO mapreduce.Job: map 100% reduce 94%
2023-06-09 19:59:05,232 INFO mapreduce.Job: map 100% reduce 96%
2023-06-09 19:59:10,442 INFO mapreduce.Job: map 100% reduce 98%
2023-06-09 19:59:16,724 INFO mapreduce.Job: map 100% reduce 100%
2023-06-09 19:59:22,202 INFO mapreduce.Job: Job job_1686172668550_0011 completed successfully
2023-06-09 19:59:25,908 INFO mapreduce.Job: Counters: 57
```



# RUNNING WITH HADOOP

Menggunakan 1GB File txt

```
2023-06-22 20:21:15,030 INFO mapreduce.Job: map 0% reduce 0%
2023-06-22 20:21:38,757 INFO mapreduce.Job: map 8% reduce 0%
2023-06-22 20:21:46,950 INFO mapreduce.Job: map 14% reduce 0%
2023-06-22 20:21:51,468 INFO mapreduce.Job: map 16% reduce 0%
2023-06-22 20:21:52,546 INFO mapreduce.Job: map 18% reduce 0%
2023-06-22 20:21:58,720 INFO mapreduce.Job: map 21% reduce 0%
2023-06-22 20:21:59,734 INFO mapreduce.Job: map 23% reduce 0%
2023-06-22 20:22:05,352 INFO mapreduce.Job: map 27% reduce 0%
2023-06-22 20:22:06,382 INFO mapreduce.Job: map 28% reduce 0%
2023-06-22 20:22:11,834 INFO mapreduce.Job: map 30% reduce 0%
2023-06-22 20:22:12,844 INFO mapreduce.Job: map 31% reduce 0%
2023-06-22 20:22:18,800 INFO mapreduce.Job: map 35% reduce 0%
2023-06-22 20:22:19,967 INFO mapreduce.Job: map 36% reduce 0%
2023-06-22 20:22:25,372 INFO mapreduce.Job: map 41% reduce 0%
2023-06-22 20:22:31,854 INFO mapreduce.Job: map 44% reduce 0%
2023-06-22 20:22:33,109 INFO mapreduce.Job: map 48% reduce 0%
2023-06-22 20:22:37,484 INFO mapreduce.Job: map 49% reduce 0%
2023-06-22 20:22:43,926 INFO mapreduce.Job: map 51% reduce 0%
2023-06-22 20:22:45,021 INFO mapreduce.Job: map 54% reduce 0%
2023-06-22 20:22:50,167 INFO mapreduce.Job: map 57% reduce 0%
2023-06-22 20:22:51,204 INFO mapreduce.Job: map 60% reduce 0%
2023-06-22 20:22:56,434 INFO mapreduce.Job: map 62% reduce 0%
2023-06-22 20:22:57,651 INFO mapreduce.Job: map 65% reduce 0%
2023-06-22 20:23:01,923 INFO mapreduce.Job: map 66% reduce 0%
2023-06-22 20:23:03,028 INFO mapreduce.Job: map 68% reduce 0%
2023-06-22 20:23:04,202 INFO mapreduce.Job: map 70% reduce 0%
2023-06-22 20:23:08,157 INFO mapreduce.Job: map 71% reduce 0%
2023-06-22 20:23:09,738 INFO mapreduce.Job: map 74% reduce 0%
2023-06-22 20:23:14,586 INFO mapreduce.Job: map 75% reduce 0%
2023-06-22 20:23:40,868 INFO mapreduce.Job: map 80% reduce 17%
2023-06-22 20:23:46,892 INFO mapreduce.Job: map 82% reduce 25%
2023-06-22 20:23:54,281 INFO mapreduce.Job: map 85% reduce 25%
2023-06-22 20:23:59,863 INFO mapreduce.Job: map 87% reduce 25%
2023-06-22 20:24:06,225 INFO mapreduce.Job: map 90% reduce 25%
2023-06-22 20:24:12,864 INFO mapreduce.Job: map 92% reduce 25%
```

**2 Menit 57 Detik**



# RUNNING WITH HADOOP

Menggunakan 1 MB File txt

Uji	Waktu
1	15 second
2	13 second
3	14 second
4	12 second
5	13 second
Rata-rata	13,8 second

# RUNNING WITH HADOOP

Menggunakan 10 MB File txt

Uji	Waktu
1	15 second
2	16 second
3	16 second
4	16 second
5	16 second
Rata-rata	15,8 Second



# RUNNING WITH HADOOP

Menggunakan 100 MB File txt

Uji	Waktu
1	1 menit 36 detik
2	54 detik
3	54 detik
4	51 detik
5	54 detik
Rata-rata	1 menit 1,8 detik



# RUNNING WITH HADOOP

Menggunakan 200 MB File txt

Uji	Waktu
1	1 menit 38 detik
2	1 menit 6 detik
3	1 menit 7 detik
4	1 menit 9 detik
5	1 menit 3 detik
Rata-rata	1 menit 12 detik



# RUNNING WITH HADOOP

Menggunakan 500 MB File txt

Uji	Waktu
1	3 menit 2 detik
2	2 menit
3	2 menit 35 detik
4	2 menit 26 detik
5	2 menit 17 detik
Rata-rata	2 menit 28 detik



# RUNNING WITH HADOOP

Menggunakan 1 GB File txt

Uji	Waktu
1	3 menit 13 detik
2	2 menit 57 detik
Rata-rata	3 Menit 5 detik





# RUNNING WITH HADOOP

Running time rata-rata :

1 MB : 13,8 detik

10 MB : 15,8 detik

100 MB : 1 menit 1,8 detik

200 MB : 1 menit 12 detik

500 MB : 2 menit 28 detik

1 GB







# RUNNING

Without Hadoop

[Back to Agenda](#)

```
require File.expand_path("../support/spec_helper.rb", __FILE__)
# Prevent database truncation if the environment is production
abort("The Rails environment is running in production mode!")
require 'spec_helper'
require 'rspec/rails'

require 'capybara/rspec'
require 'capybara/rails'

1
2 Capybara.javascript_driver = :webkit
13 Category.delete_all; Category.create
14 Shoulda::Matchers.configure do |config|
15   config.integrate do |with|
16     with.test_framework :rspec
17     with.library :rails
18   end
19 end
20
21 # Add additional requires below this line. Make sure they're on top of the
22 # Requires supporting ruby files with custom matchers (e.g.
23 # spec/support/**/*.rb)
24 # run as spec files by default. This will run the matchers
25 # in _spec.rb will both be required when running specs
26 # run twice. It is recommended that you require the matchers
27 # end with _spec.rb. You can configure the require pattern
28 # option on the command line with --require.
29
30 No results found for 'mongoid'
31
32 mongoid
33
34 at buffer
```



# RUNNING WITHOUT HADOOP

Menggunakan 1 MB File txt

```
Iteration 1 Runtime: 156 milliseconds
Iteration 2 Runtime: 119 milliseconds
Iteration 3 Runtime: 88 milliseconds
Iteration 4 Runtime: 76 milliseconds
Iteration 5 Runtime: 73 milliseconds
Iteration 6 Runtime: 66 milliseconds
Iteration 7 Runtime: 86 milliseconds
Iteration 8 Runtime: 71 milliseconds
Iteration 9 Runtime: 96 milliseconds
Iteration 10 Runtime: 110 milliseconds
Iteration 11 Runtime: 87 milliseconds
Iteration 12 Runtime: 84 milliseconds
Iteration 13 Runtime: 102 milliseconds
Iteration 14 Runtime: 82 milliseconds
Iteration 15 Runtime: 77 milliseconds
Average Runtime: 91.53333333333333 milliseconds
```



# RUNNING WITHOUT HADOOP

Menggunakan 10 MB File txt

```
Iteration 1 Runtime: 923 milliseconds
Iteration 2 Runtime: 789 milliseconds
Iteration 3 Runtime: 699 milliseconds
Iteration 4 Runtime: 717 milliseconds
Iteration 5 Runtime: 716 milliseconds
Iteration 6 Runtime: 721 milliseconds
Iteration 7 Runtime: 878 milliseconds
Iteration 8 Runtime: 839 milliseconds
Iteration 9 Runtime: 1459 milliseconds
Iteration 10 Runtime: 1110 milliseconds
Iteration 11 Runtime: 1094 milliseconds
Iteration 12 Runtime: 1508 milliseconds
Iteration 13 Runtime: 870 milliseconds
Iteration 14 Runtime: 715 milliseconds
Iteration 15 Runtime: 810 milliseconds
Average Runtime: 923.2 milliseconds
```



# RUNNING WITHOUT HADOOP

Menggunakan 100 MB File txt

```
Iteration 1 Runtime: 10579 milliseconds
Iteration 2 Runtime: 12912 milliseconds
Iteration 3 Runtime: 10046 milliseconds
Iteration 4 Runtime: 9464 milliseconds
Iteration 5 Runtime: 9129 milliseconds
Iteration 6 Runtime: 9310 milliseconds
Iteration 7 Runtime: 9360 milliseconds
Iteration 8 Runtime: 9348 milliseconds
Iteration 9 Runtime: 9172 milliseconds
Iteration 10 Runtime: 9204 milliseconds
Iteration 11 Runtime: 9431 milliseconds
Iteration 12 Runtime: 9200 milliseconds
Iteration 13 Runtime: 9675 milliseconds
Iteration 14 Runtime: 9813 milliseconds
Iteration 15 Runtime: 9337 milliseconds
Average Runtime: 9732.0 milliseconds
```



# RUNNING WITHOUT HADOOP

Menggunakan 200 MB File txt

```
Iteration 1 Runtime: 21902 milliseconds
Iteration 2 Runtime: 21828 milliseconds
Iteration 3 Runtime: 20253 milliseconds
Iteration 4 Runtime: 18430 milliseconds
Iteration 5 Runtime: 18558 milliseconds
Iteration 6 Runtime: 18256 milliseconds
Iteration 7 Runtime: 19300 milliseconds
Iteration 8 Runtime: 17979 milliseconds
Iteration 9 Runtime: 17560 milliseconds
Iteration 10 Runtime: 17775 milliseconds
Iteration 11 Runtime: 17713 milliseconds
Iteration 12 Runtime: 18039 milliseconds
Iteration 13 Runtime: 18551 milliseconds
Iteration 14 Runtime: 18170 milliseconds
Iteration 15 Runtime: 17865 milliseconds
Average Runtime: 18811.933333333334 milliseconds
```



# RUNNING WITHOUT HADOOP

Menggunakan 500 Mb File txt

```
Iteration 1 Runtime: 29127 milliseconds  
Iteration 2 Runtime: 24950 milliseconds  
Iteration 3 Runtime: 24615 milliseconds  
Iteration 4 Runtime: 24726 milliseconds  
Iteration 5 Runtime: 25422 milliseconds  
Iteration 6 Runtime: 24799 milliseconds  
Iteration 7 Runtime: 25520 milliseconds  
Iteration 8 Runtime: 24572 milliseconds  
Iteration 9 Runtime: 24736 milliseconds  
Iteration 10 Runtime: 24392 milliseconds  
Iteration 11 Runtime: 24981 milliseconds  
Iteration 12 Runtime: 25106 milliseconds  
Iteration 13 Runtime: 26645 milliseconds  
Iteration 14 Runtime: 26702 milliseconds  
Iteration 15 Runtime: 24524 milliseconds  
Average Runtime: 25387.8 milliseconds
```



# RUNNING WITHOUT HADOOP

Menggunakan 1 GB File txt

```
Iteration 1 Runtime: 61186 milliseconds
Iteration 2 Runtime: 50208 milliseconds
Iteration 3 Runtime: 50131 milliseconds
Iteration 4 Runtime: 50669 milliseconds
Iteration 5 Runtime: 49843 milliseconds
Iteration 6 Runtime: 51051 milliseconds
Iteration 7 Runtime: 51711 milliseconds
Iteration 8 Runtime: 51011 milliseconds
Iteration 9 Runtime: 51815 milliseconds
Iteration 10 Runtime: 50375 milliseconds
Iteration 11 Runtime: 51856 milliseconds
Iteration 12 Runtime: 53454 milliseconds
Iteration 13 Runtime: 56557 milliseconds
Iteration 14 Runtime: 55537 milliseconds
Iteration 15 Runtime: 54972 milliseconds
Average Runtime: 52691.73333333333 milliseconds
```





# ANALYZE

Hadoop vs Non Hadoop

Kelompok 1





# ANALYZE

On 1 MB

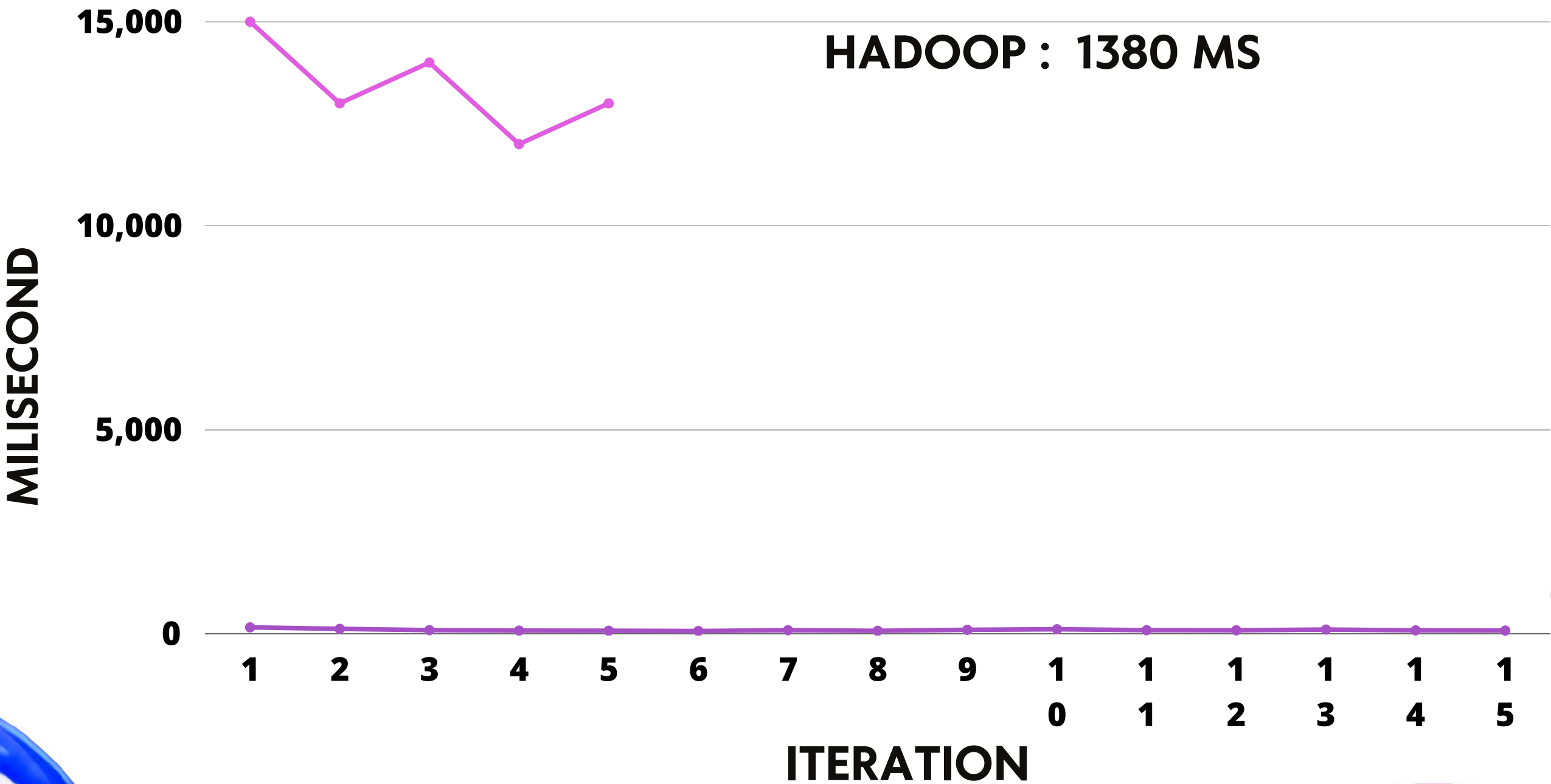
Kelompok 1



\*Perbandingan mungkin tidak sesuai karena dijalankan di PC yang berbeda.

# 1 MB

Compare 1 on 1



HADOOP : 1380 MS

JAVA : 91.53 MS



# ANALYZE

On 10 MB

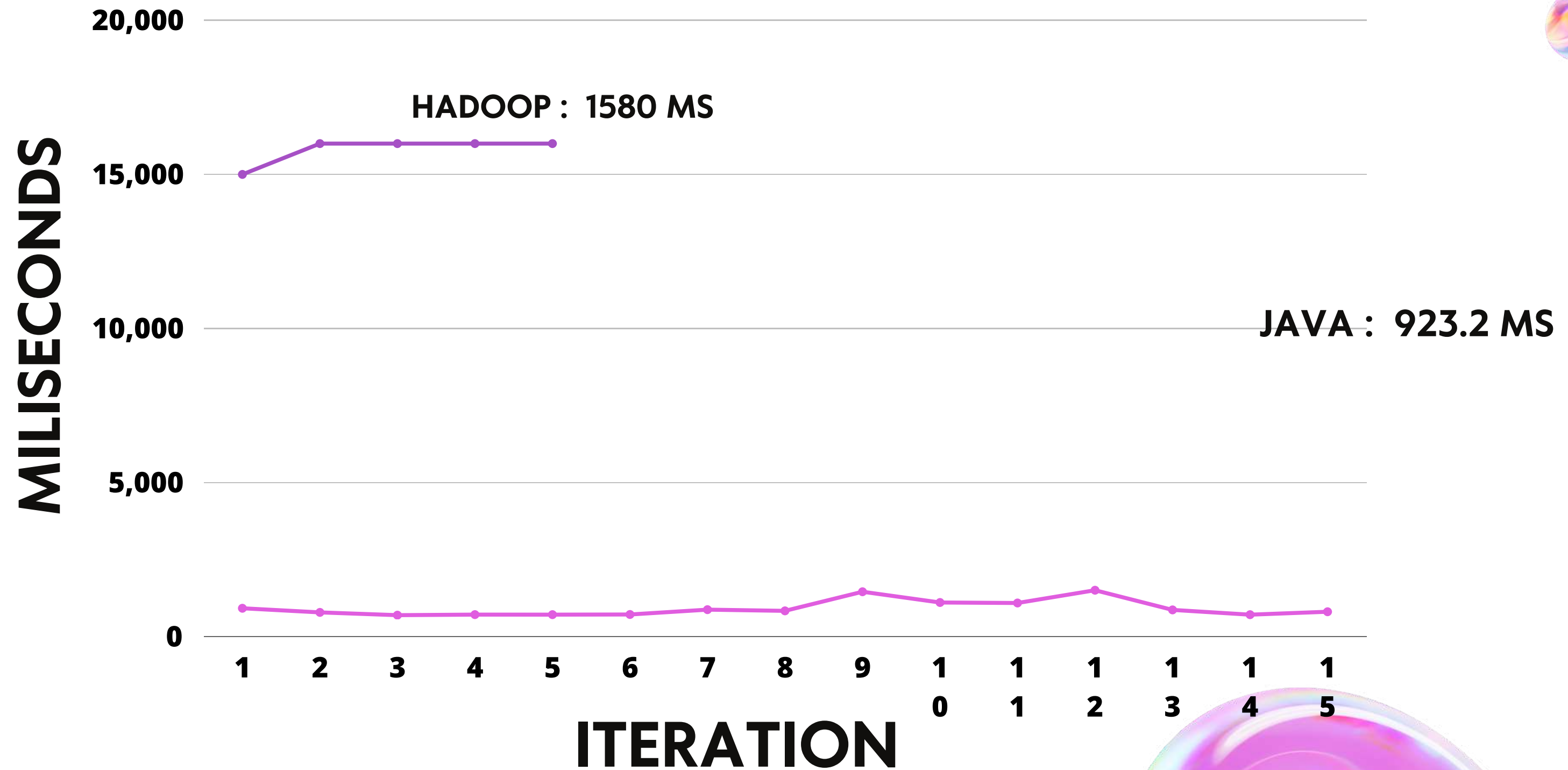
Kelompok 1



\*Perbandingan mungkin tidak sesuai karena dijalankan di PC yang berbeda.

# 10 MB

Compare 1 on 1





# ANALYZE

On 100 MB

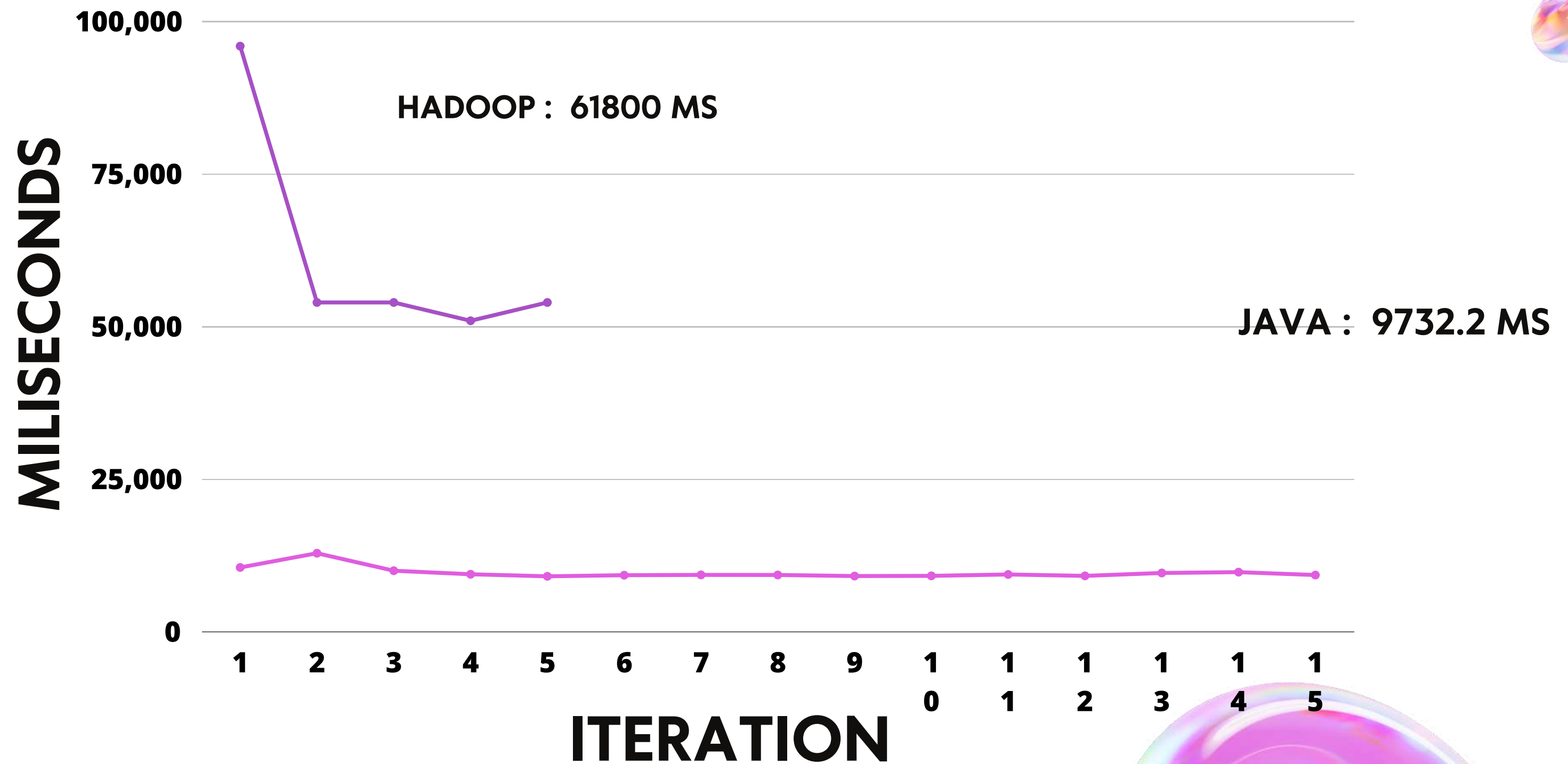
Kelompok 1



\*Perbandingan mungkin tidak sesuai karena dijalankan di PC yang berbeda.

# 100 MB

Compare 1 on 1





# ANALYZE

On 200 MB

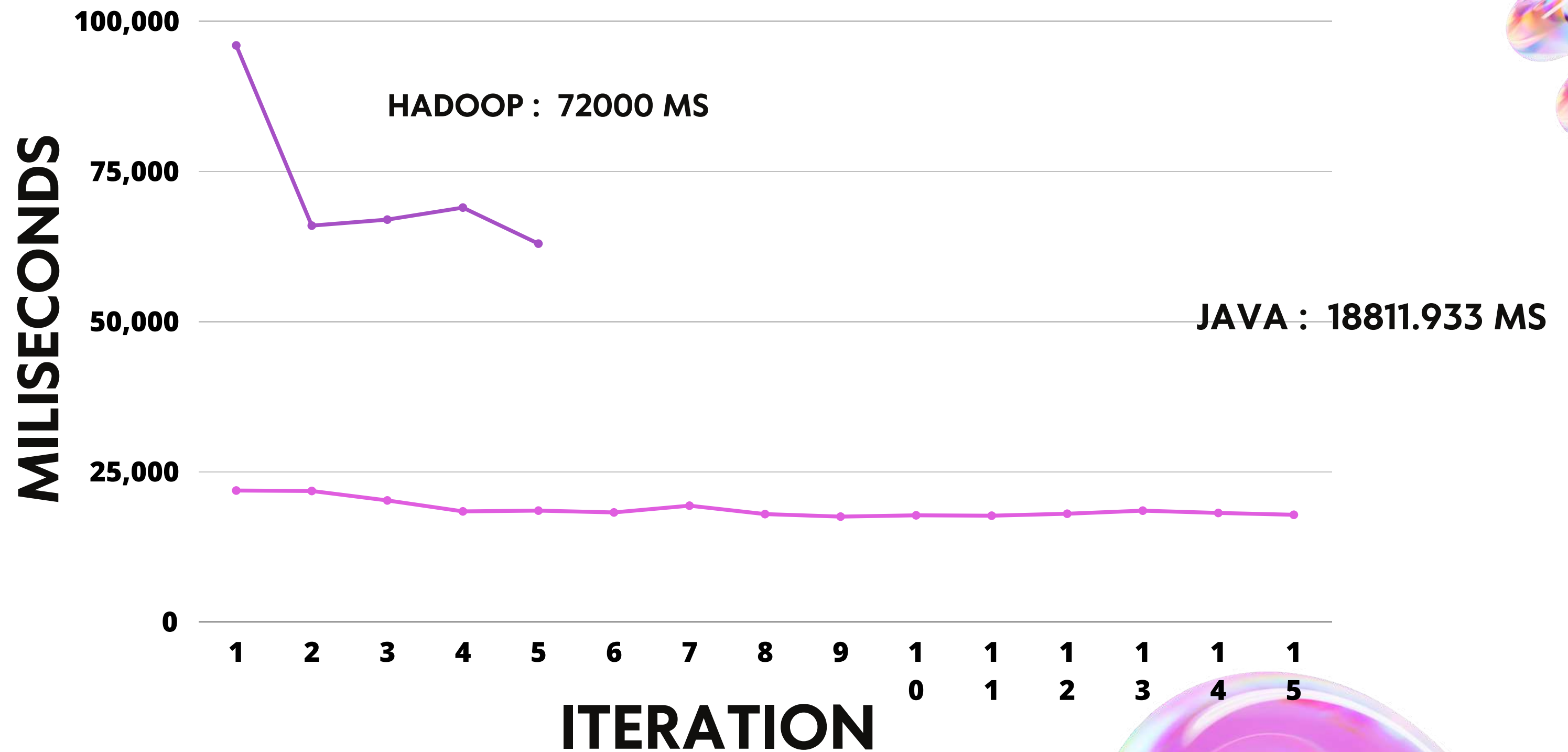
Kelompok 1



\*Perbandingan mungkin tidak sesuai karena dijalankan di PC yang berbeda.

# 200 MB

Compare 1 on 1





# ANALYZE

On 500 MB

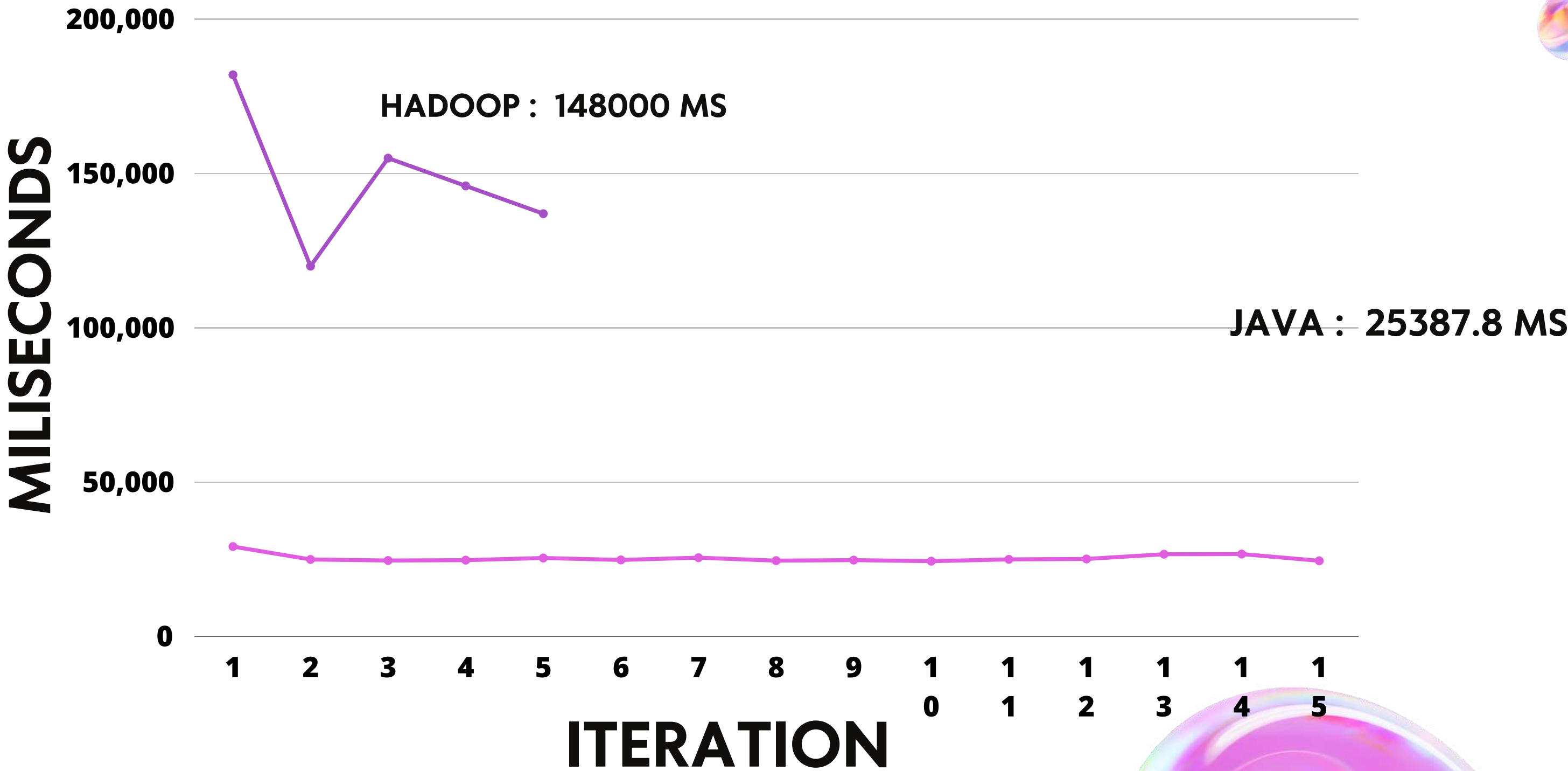
Kelompok 1



\*Perbandingan mungkin tidak sesuai karena dijalankan di PC yang berbeda.

# 500 MB

Compare 1 on 1





# ANALYZE

On 1000 MB

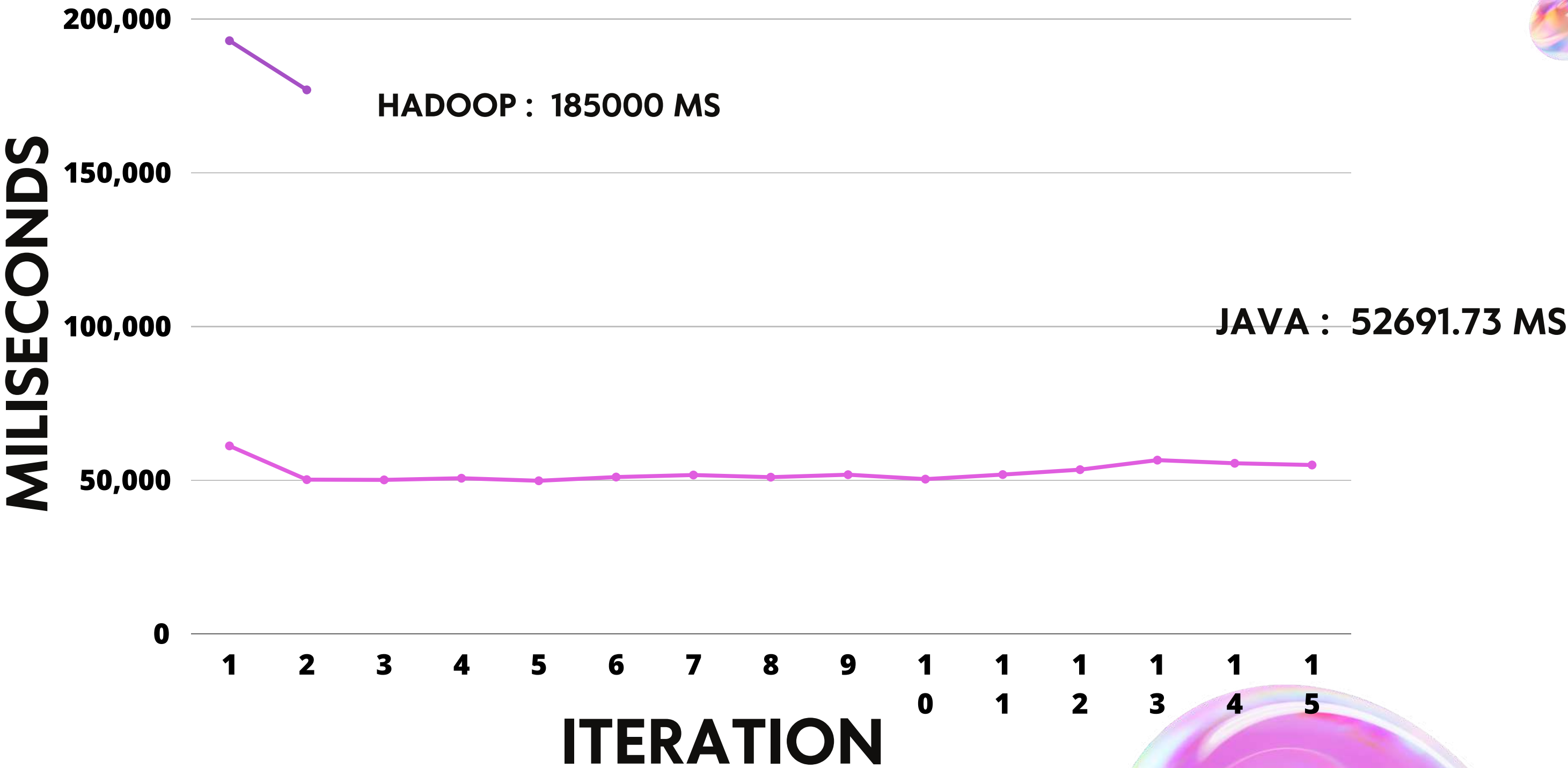
Kelompok 1



\*Perbandingan mungkin tidak sesuai karena dijalankan di PC yang berbeda.

# 1 GB

Compare 1 on 1





# ANALYZE

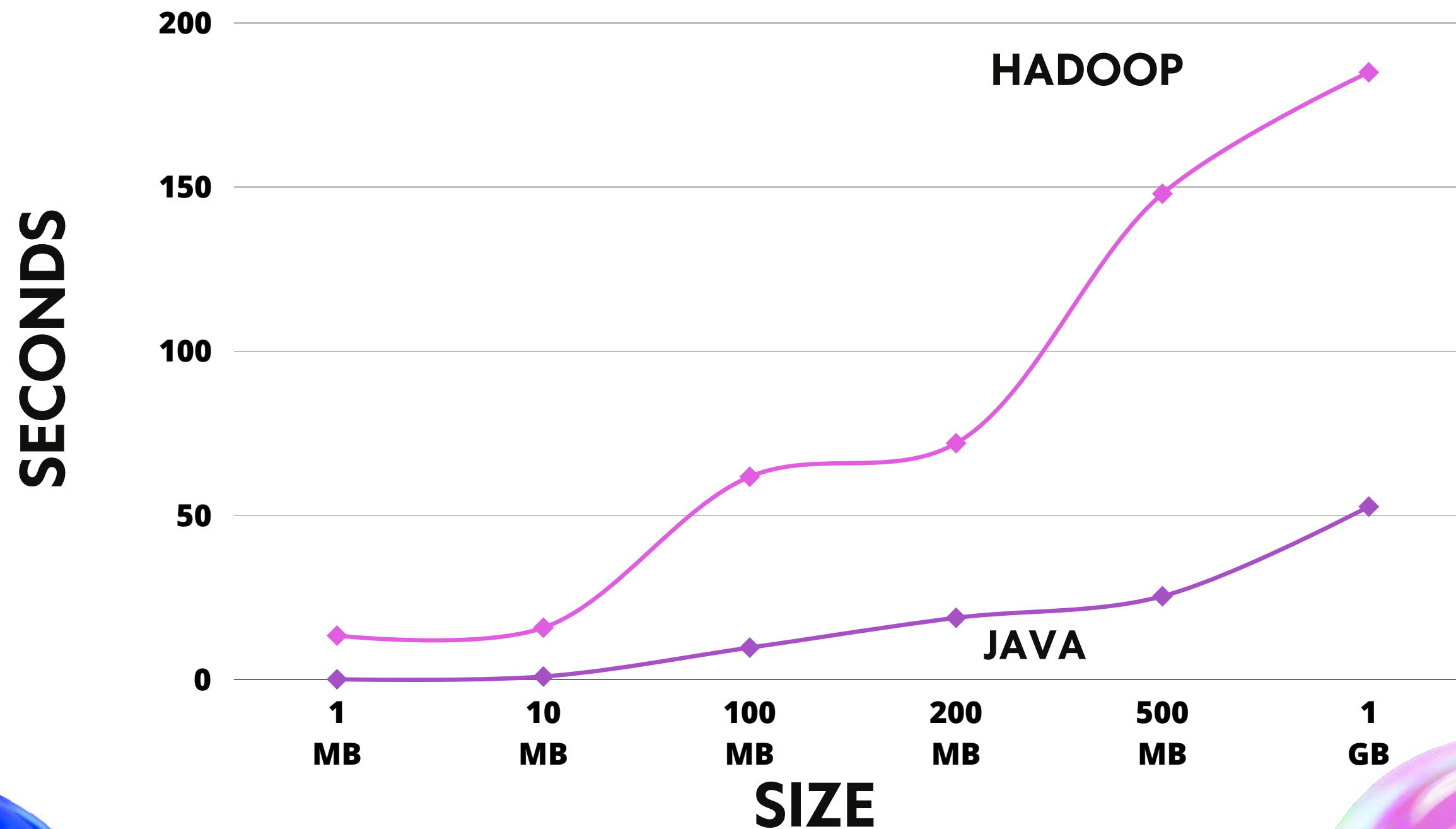
At All

Kelompok 1



\*Perbandingan mungkin tidak sesuai karena dijalankan di PC yang berbeda.

# COMPARE ALL

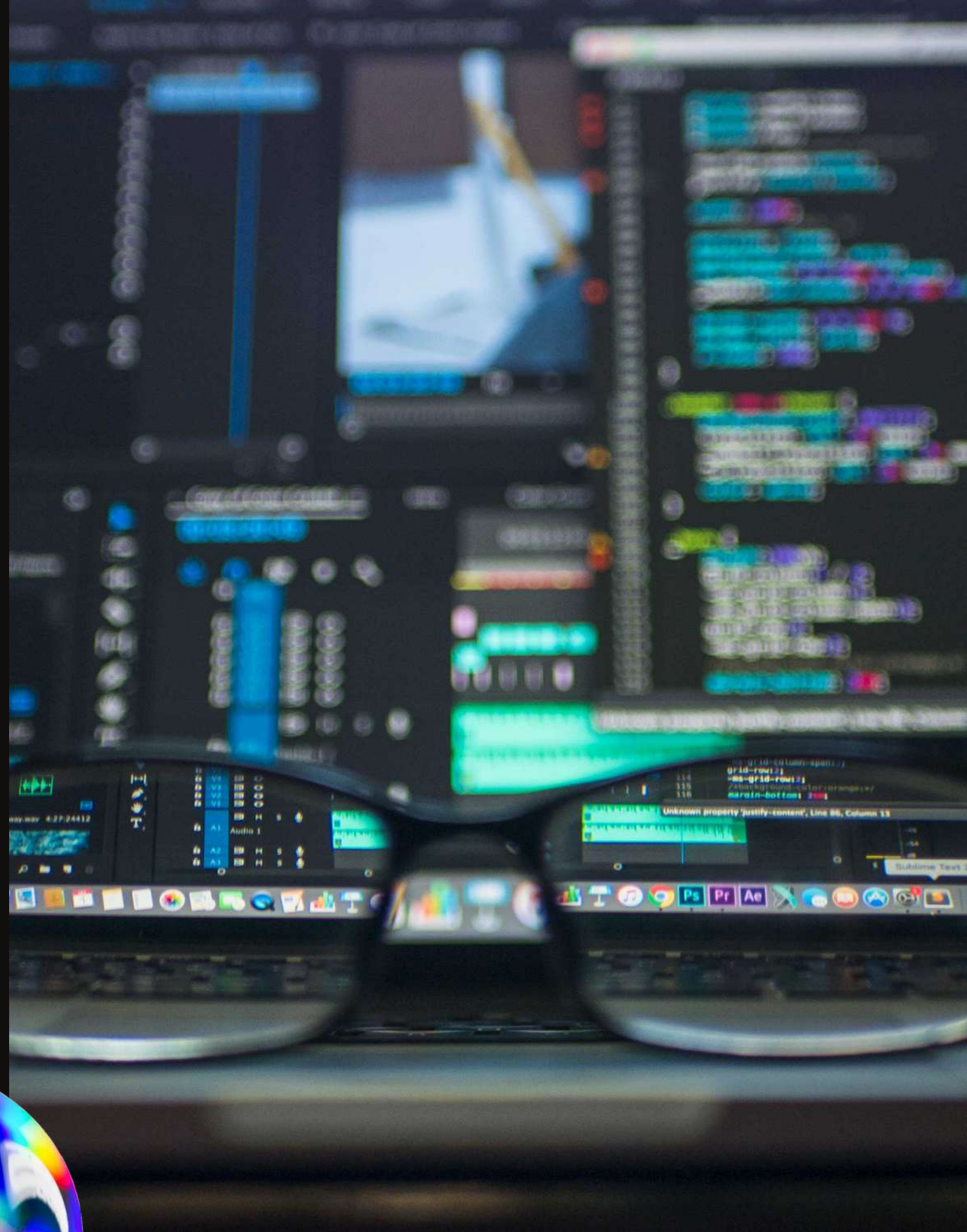




# ANALISIS

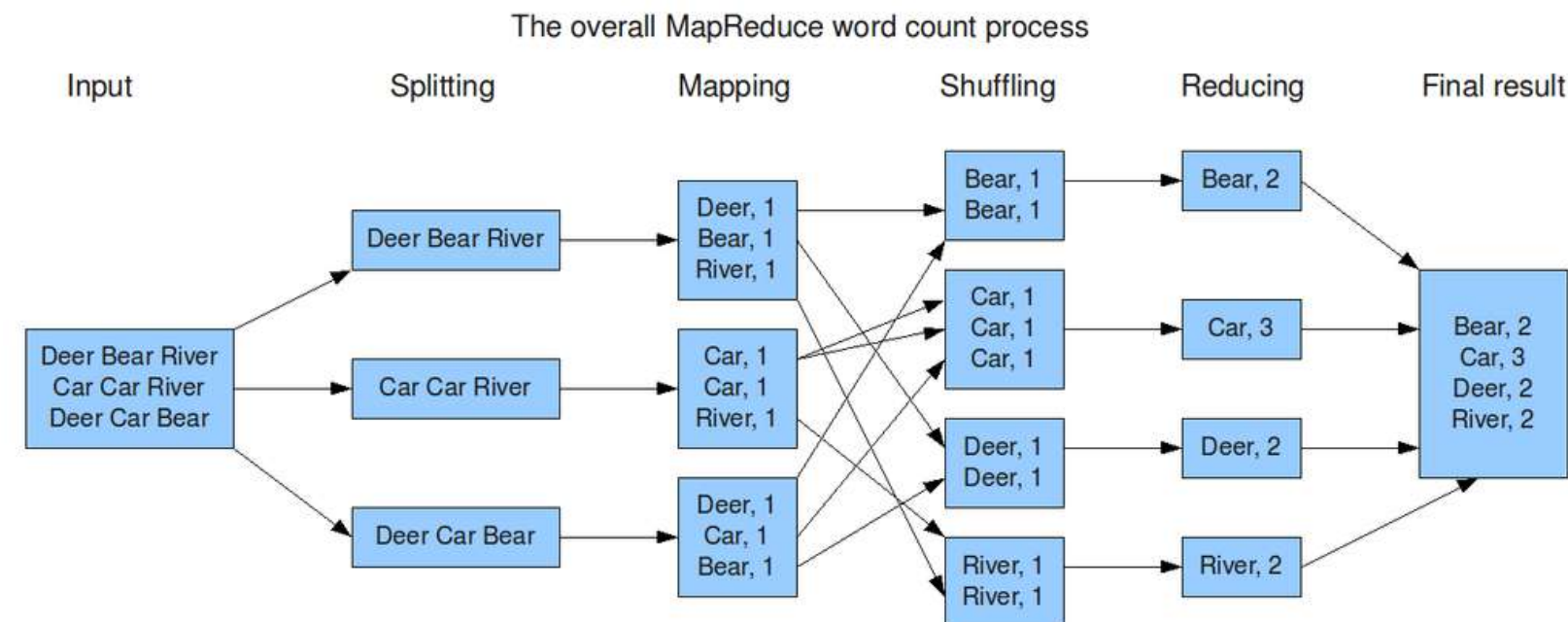
Perbandingan tanpa Hadoop dan dengan Hadoop

Kelompok 1





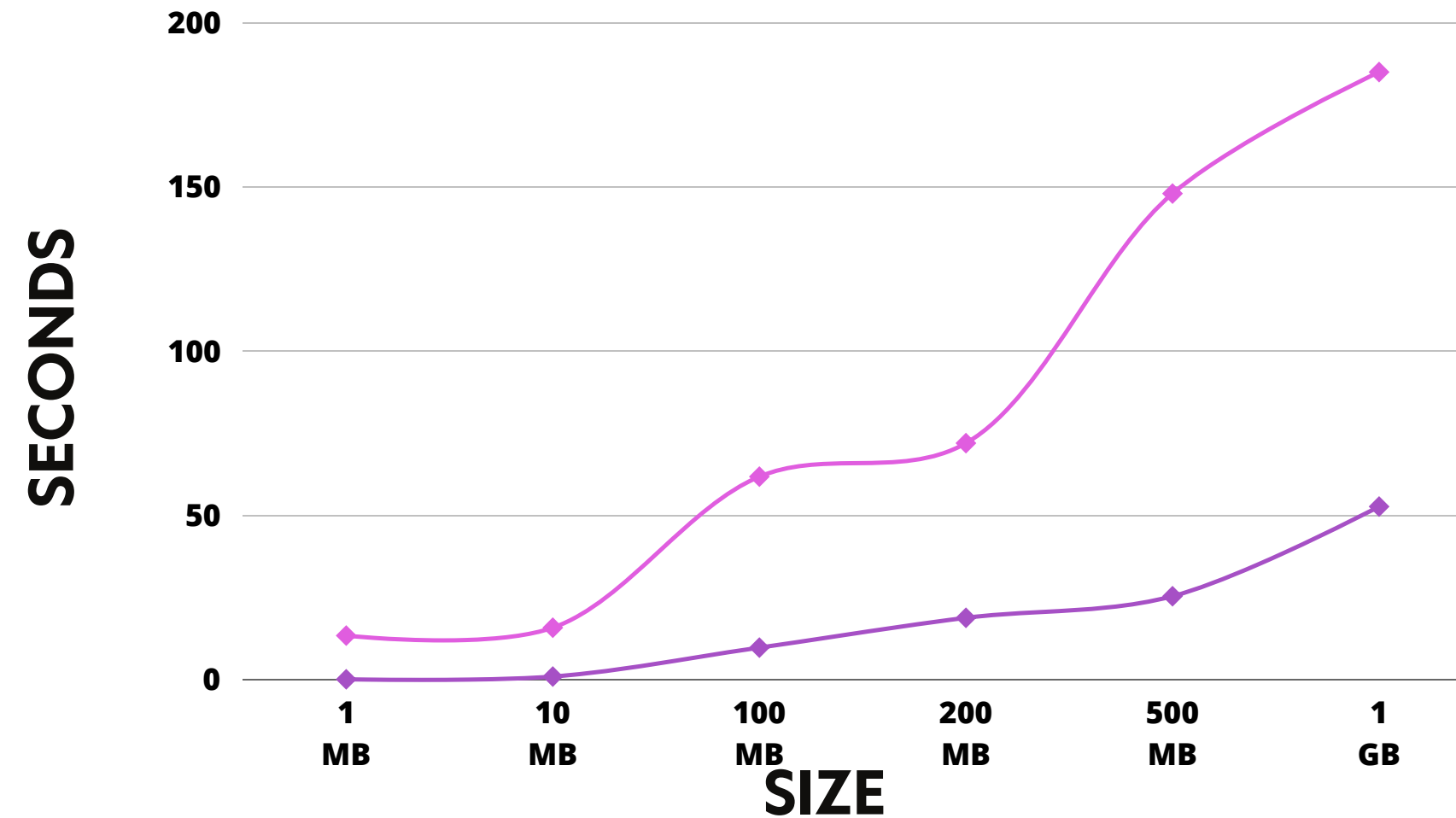
# WAKTU EKSEKUSI



Source : <https://medium.com/pujanggateknologi/implementasi-sederhana-framework-mapreduce-8f80f22cc54f>

Hadoop dengan MapReduce memiliki waktu eksekusi yang lebih lama daripada Java tanpa MapReduce untuk setiap ukuran data. Hal ini dapat disebabkan oleh overhead yang terlibat dalam proses distribusi, pengiriman data antara node, serta pengaturan dan sinkronisasi tugas MapReduce.

# WAKTU EKSEKUSI



Meskipun Java tanpa MapReduce memiliki waktu eksekusi yang lebih cepat daripada Hadoop dengan MapReduce, perbedaannya semakin menonjol seiring dengan pertambahan ukuran data. Hal ini menunjukkan bahwa Java tanpa MapReduce mungkin lebih cocok untuk pemrosesan data kecil hingga menengah, sedangkan Hadoop dengan MapReduce lebih cocok untuk pemrosesan data besar yang dapat dipecah menjadi tugas yang dapat disebarluaskan di beberapa node.





# PERBEDAAN SPESIFIKASI YANG DIGUNAKAN

Perbedaan spesifikasi antara laptop dan virtual machine dapat mempengaruhi perbandingan waktu eksekusi Wordcount. Dalam hal ini, laptop dengan spesifikasi Ryzen 7 5600U yang memiliki 8 core dan 16 thread serta RAM sebesar 23,4 GB mungkin memiliki kinerja yang lebih baik daripada virtual machine yang hanya menggunakan 4 core dan 7,2 GB RAM.

Dengan memiliki lebih banyak core dan thread, laptop dapat melakukan pemrosesan yang lebih cepat. Selain itu, RAM yang lebih besar pada laptop juga memungkinkan untuk menyimpan dan memanipulasi jumlah data yang lebih besar secara efisien, yang dapat mempercepat eksekusi Wordcount.

Di sisi lain, virtual machine dengan spesifikasi yang lebih rendah mungkin mengalami pembatasan dalam hal pemrosesan paralel dan kapasitas memori. Hal ini dapat mempengaruhi kinerja Hadoop saat menjalankan Wordcount melalui virtual machine. Waktu eksekusi yang lebih lama mungkin terjadi karena keterbatasan sumber daya yang tersedia dalam virtual machine tersebut.

Jadi, perbedaan spesifikasi antara laptop dan virtual machine dapat mempengaruhi waktu eksekusi Wordcount, dengan laptop mungkin memiliki kinerja yang lebih baik karena memiliki spesifikasi yang lebih tinggi.






# **WORDCOUNT DENGAN JAVA TANPA HADOOP MEMERLUKAN MEMORI YANG BESAR**

Wordcount dengan Java tanpa MapReduce memiliki kecenderungan untuk menggunakan memori yang besar dan rentan terhadap kesalahan OutOfMemoryError. Hal ini disebabkan oleh pendekatan pemrosesan data secara serial, di mana seluruh teks atau file dibaca dan dimuat ke dalam memori secara keseluruhan sebelum dilakukan penghitungan frekuensi kata.

Ketika memproses volume data yang besar, penggunaan memori dapat meningkat secara signifikan. Jika memori yang tersedia tidak mencukupi, program Wordcount dapat mengalami kesalahan OutOfMemoryError, yang menunjukkan bahwa memori yang dialokasikan telah habis.

Untuk mengatasi masalah OutOfMemoryError dapat menggunakan perintah "-Xms" untuk mengatur ukuran awal memori yang dialokasikan untuk program. Misalnya, "-Xms16g" akan mengalokasikan 16 gigabyte (16g) memori awal.








# FITUR TAMBAHAN HADOOP

Hadoop memiliki kemampuan untuk berjalan pada mode multinode, yaitu dapat dijalankan di beberapa node atau server yang saling terhubung dalam sebuah kluster. Namun, pada percobaan ini, implementasinya menggunakan mode singlenode, yang berarti hanya menjalankan Hadoop pada satu mesin atau server.

**Jika percobaan ini menggunakan mode multinode** dengan kluster Hadoop yang terdiri dari beberapa mesin atau server, **kemungkinan besar waktu pemrosesan akan jauh lebih cepat dibandingkan dengan mode singlenode.**

Dalam mode multinode, tugas pemrosesan MapReduce dapat didistribusikan di antara beberapa mesin atau server dalam kluster. Setiap mesin atau server dapat menjalankan tugas pemrosesan pada subset data yang berbeda secara paralel. Dengan demikian, lebih banyak sumber daya komputasi yang tersedia secara bersamaan, yang menghasilkan peningkatan kecepatan pemrosesan secara signifikan.



# KELEBIHAN KEKURANGAN WORDCOUNT MENGGUNAKAN JAVA TANPA MAPREDUCE

## KELEBIHAN

### **Kesederhanaan**

Dapat menulis kode Java langsung untuk membaca dan memproses file teks, menghitung frekuensi kata, dan menghasilkan output yang diinginkan.

### **Kecepatan**

Dalam beberapa kasus, Wordcount dengan Java tanpa MapReduce dapat lebih cepat daripada menggunakan MapReduce. Terutama saat memproses volume data yang relatif kecil.

### **Ketersediaan sumber daya**

Pendekatan Wordcount dengan Java tanpa MapReduce tidak memerlukan penggunaan virtual machine atau cluster Hadoop.

## KEKURANGAN

### **Skalabilitas terbatas**

Saat menghadapi volume data yang sangat besar, memprosesnya secara serial dengan menggunakan Java tanpa MapReduce dapat memakan waktu yang lama dan membutuhkan sumber daya yang lebih besar.

### **Keterbatasan pemrosesan paralel**

Dalam pendekatan ini, pemrosesan paralel harus diimplementasikan secara manual jika ingin memanfaatkan potensi pemrosesan paralel pada mesin yang memiliki banyak core atau thread. Ini dapat menjadi lebih rumit dan memerlukan pengetahuan teknis yang lebih dalam dalam pemrograman paralel.

### **Toleransi kesalahan terbatas**

Ketika terjadi kegagalan pada salah satu bagian pemrosesan, seperti bacaan file atau pemrosesan data, perlu menangani penanganan kesalahan secara manual.





# **KELEBIHAN KEKURANGAN WORDCOUNT MENGUNAKAN HADOOP DENGAN MAPREDUCE**

## **KELEBIHAN**

### **Skalabilitas**

Hadoop dapat membagi tugas Wordcount menjadi bagian-bagian yang dapat dijalankan secara paralel di berbagai mesin dalam cluster Hadoop. Ini memungkinkan pemrosesan data yang lebih cepat dan efisien.

### **Error Handling**

Hadoop memiliki mekanisme built-in untuk mendeteksi dan menangani kegagalan pada level perangkat keras maupun perangkat lunak. Jika terjadi kegagalan pada salah satu mesin dalam cluster, tugas Wordcount dapat dilanjutkan pada mesin lain tanpa kehilangan data atau harus mengulang proses secara keseluruhan.

## **KEKURANGAN**

### **Kompleksitas**

User harus memahami arsitektur Hadoop, mengonfigurasi cluster, dan menulis kode MapReduce yang sesuai. Ini dapat memerlukan tingkat pemahaman yang lebih dalam tentang teknologi terkait.

### **Latensi**

Terdapat overhead tambahan dalam distribusi dan pengoordinasian tugas pada cluster Hadoop, waktu respons yang dibutuhkan untuk menyelesaikan tugas Wordcount mungkin lebih lama dibandingkan dengan pendekatan non-distribusi seperti menggunakan Java tanpa MapReduce.



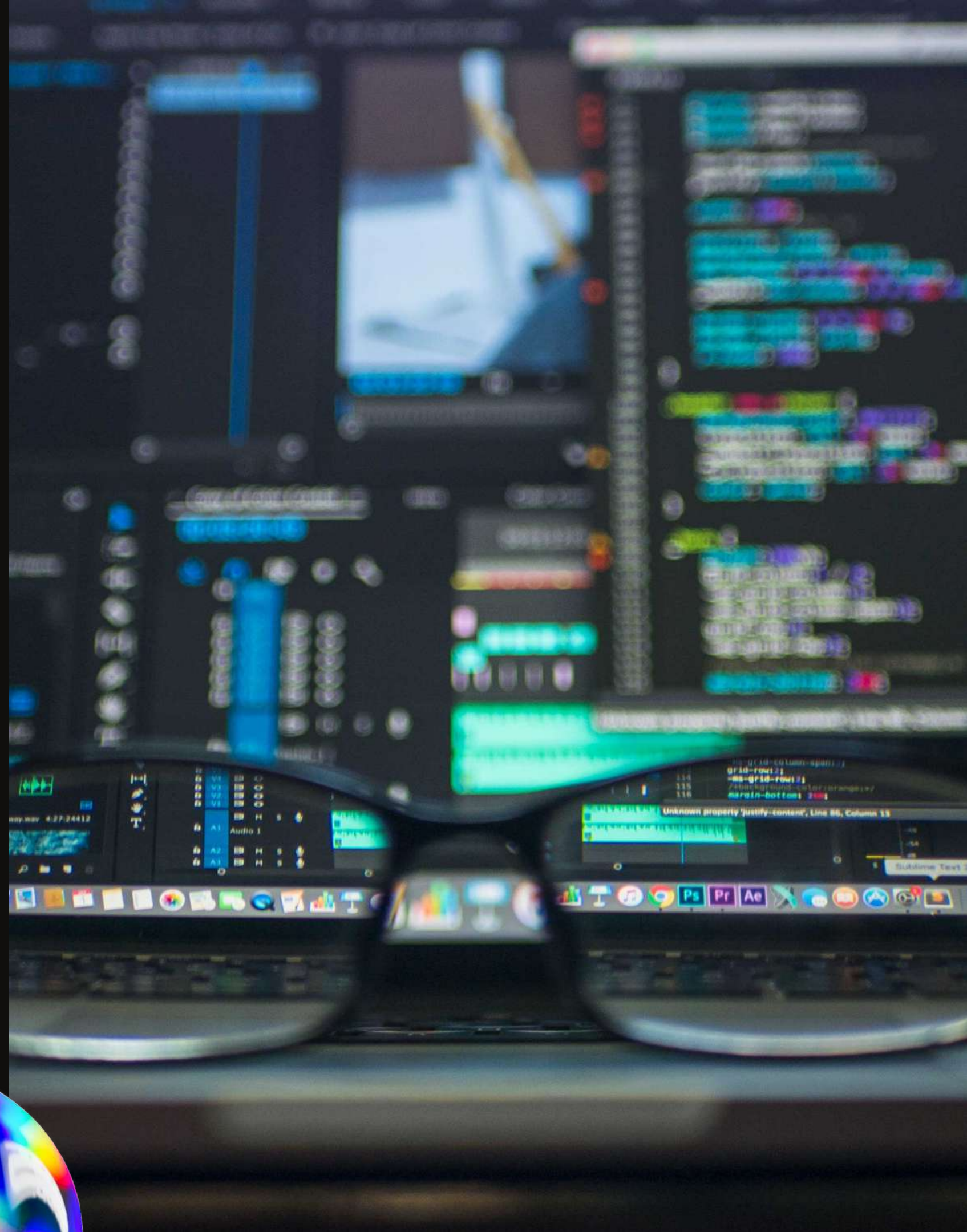




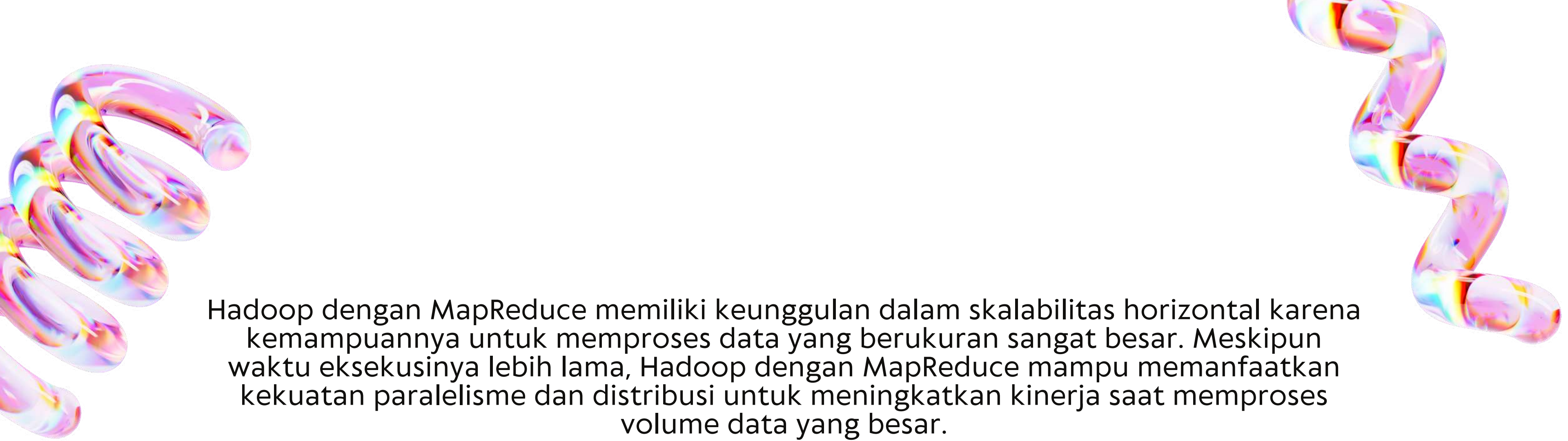
# KESIMPULAN

Perbandingan tanpa Hadoop dan dengan Hadoop

Kelompok 1





The top corners of the slide are decorated with vibrant, multi-colored wavy lines that resemble liquid or smoke, featuring shades of purple, blue, and yellow.

Hadoop dengan MapReduce memiliki keunggulan dalam skalabilitas horizontal karena kemampuannya untuk memproses data yang berukuran sangat besar. Meskipun waktu eksekusinya lebih lama, Hadoop dengan MapReduce mampu memanfaatkan kekuatan paralelisme dan distribusi untuk meningkatkan kinerja saat memproses volume data yang besar.

Java tanpa MapReduce dapat menjadi pilihan yang lebih cepat dan sederhana untuk pemrosesan data kecil dengan memanfaatkan kemampuan pemrosesan lokal pada satu mesin atau server.

The bottom corners of the slide are decorated with vibrant, multi-colored wavy lines that resemble liquid or smoke, featuring shades of purple, blue, and yellow.