COL334: Assignment 2
Distributed File Transfer

slowbrains
Sharad Kumar 2021CS10099
Rishabh Verma 2021CS10581
Bedanta Bhaumik 2021CS10097
Gauransh Sharma 2023VST9002

8 September 2023

1 Approach

We implemented the distributed file transfer program in Python utilising the `socket` and `threading` libraries of the language. We collaborated with different numbers of clients to download text data from a server efficiently. When a client successfully acquires a unique line, it performs two important actions:

1. The client stores the unique line within its individual hashmap for future reference. The key of this hashmap is the line number and content of that line is its value.
2. Additionally, the client shares the obtained unique line with the other clients in the group, enabling each of them to incorporate the line into their respective dictionaries.
3. To plot the graphs, we noted the number of unique lines on the client every 2 seconds.

This coordinated approach allows all the clients to work in synergy, resulting in a significant reduction in the overall time required to download the entire file when compared to a scenario where only one client is involved.

2 Average Completion Time Comparison

Note : The following datasets were obtained in the Ex-Hall(near CCD in Main building)

	Run 1	**Run 2**	**Run 3**	**Run 4**	**Run 5**	**Avg Completion Time**
4 clients	14.815	13.320	13.316	21.007	14.561	**15.404**
3 clients	21.998	23.482	20.199	19.756	14.101	**19.107**
2 clients	31.144	27.848	26.556	41.445	29.166	**31.032**
1 client	62.299	74.102	70.101	60.708	56.783	**64.198**

Table 1: Time taken (in seconds) to receive 1000 unique lines

3 Performance Analysis

3.1 Single Client

- Lines Received: 1000
- Time Taken: 62.30 seconds

In the first dataset with a single client, the download time is approximately 62.30 seconds. This scenario serves as a baseline for comparison.

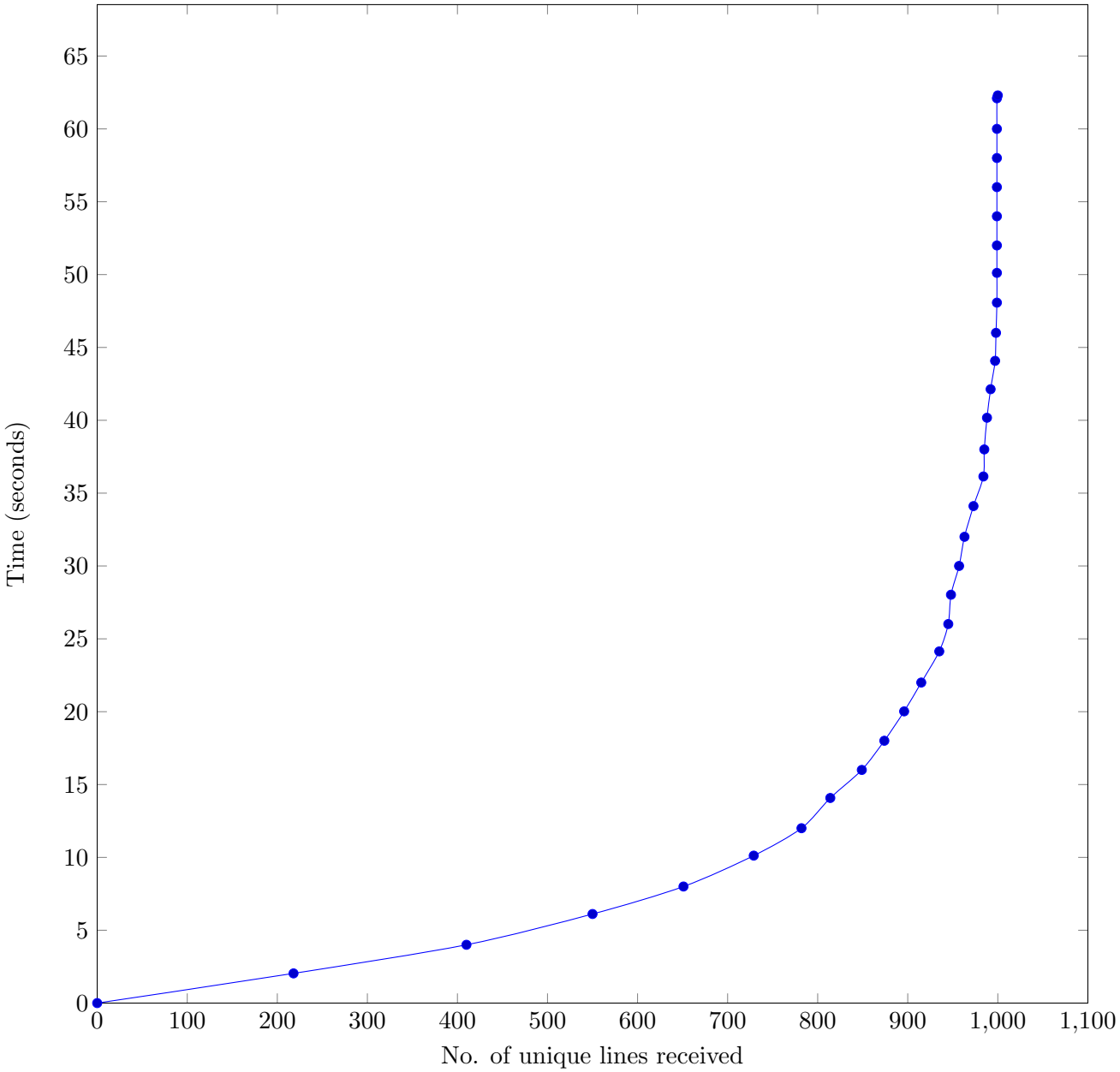

Figure 1: Downloading with One Client

3.2 Two Clients

- Lines Received: 1000
- Time Taken: 29.17 seconds

With two clients working collaboratively in the second dataset, the download time reduces to approximately 29.17 seconds. The addition of a second client approximately cut the completion time in half.

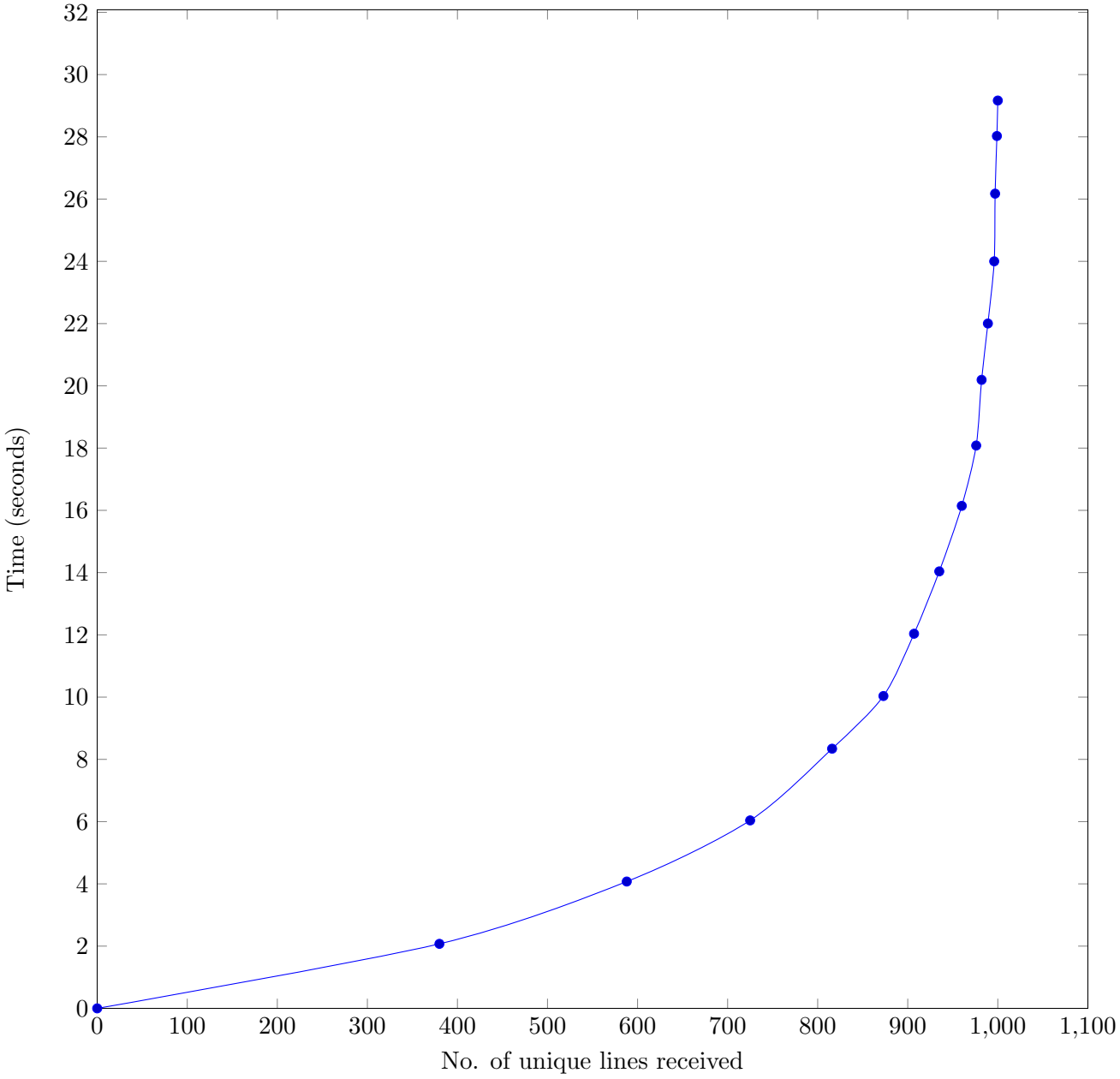

Figure 2: Downloading with Two Clients

3.3 Three Clients

- Lines Received: 1000
- Time Taken: 20.20 seconds

In the third dataset with three clients, the download time further decreases to approximately 20.20 seconds. The cooperative effort of three clients demonstrates a substantial improvement in efficiency.

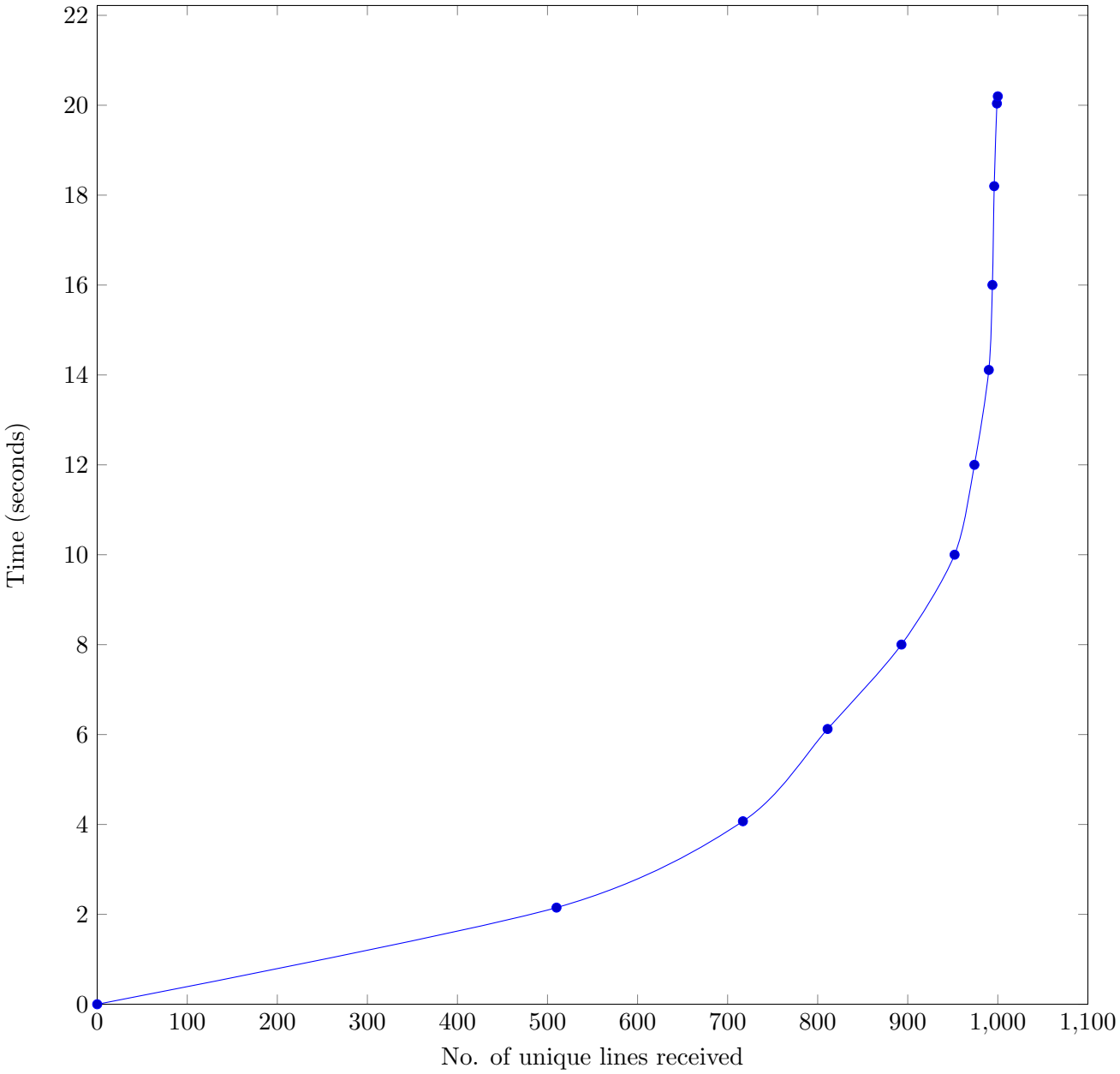

Figure 3: Downloading with Three Clients

3.4 Four Clients

- Lines Received: 1000
- Time Taken: 14.56 seconds

Finally, in the fourth dataset with four clients, the download time is approximately 14.56 seconds. This collaborative approach among four clients results in the most efficient performance, reducing the download time significantly compared to a single client.

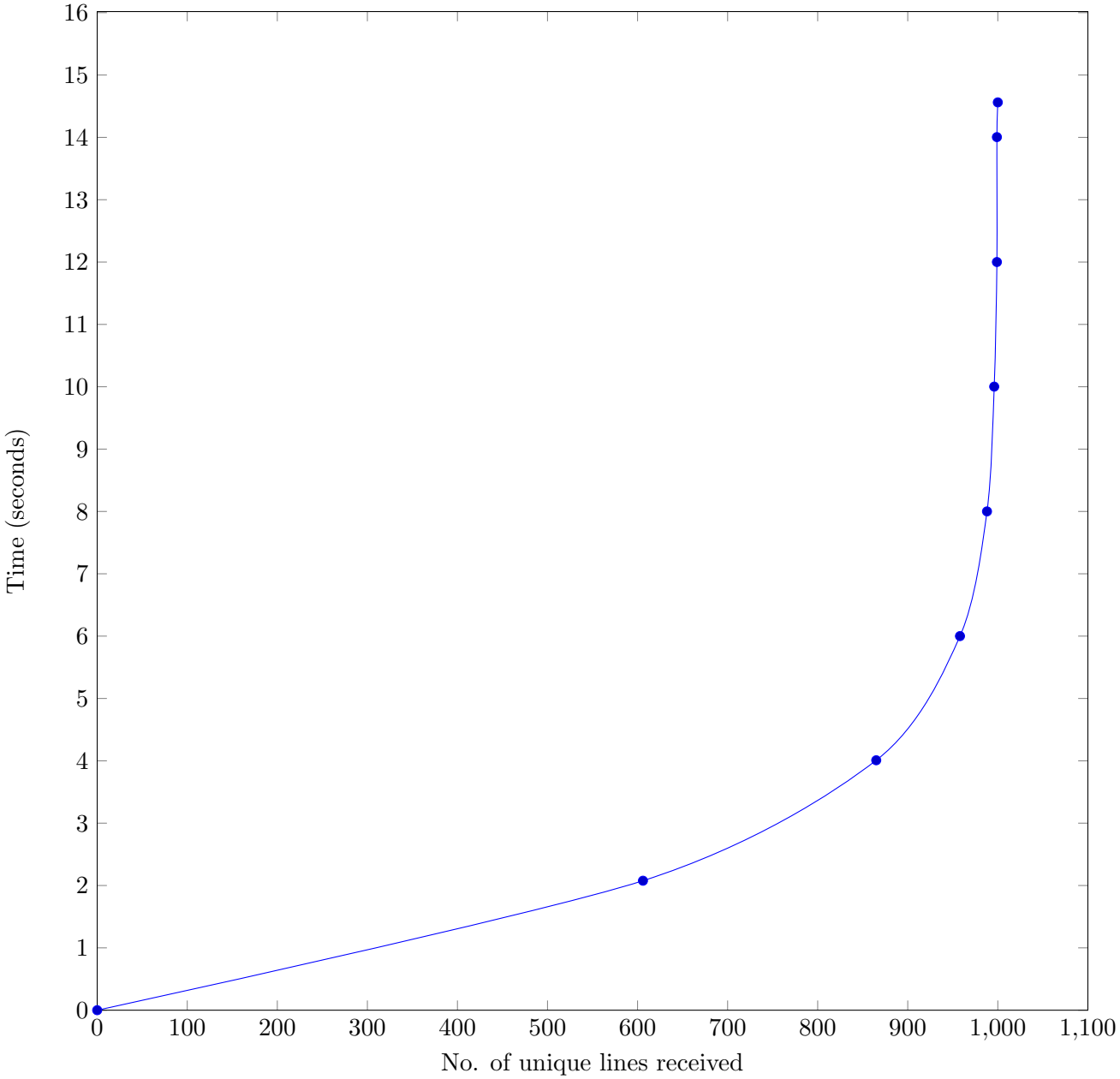

Figure 4: Downloading with Four Clients