Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

BTQ-MR: Performance Evaluation of Map Reduce Programme Using Transient Queuing Model with Finite Buffer

Chandra Sekhar Darapaneni[1], Bobba Basaveswara Rao[2], S Mallikarjuna Rao[3] Neelima Guntupalli[4]
1,2,4 Department of Computer Science &Engineering, Acharya Nagarjuna University, Guntur 522510, India.
3 Department of IT, Vasireddy Venkatadri International Technological University,Nambur,India.
Corresponding Author E-mail::smallikarjun@vvit.net

ARTICLEINFO

Received: 29 Dec 2024

Revised: 15 Feb 2025

Accepted: 24 Feb 2025

ABSTRACT

This paper introduces a transient queueing model with a finite buffer (BTQ-MR) to evaluate the performance of the MapReduce programming model. While earlier research has explored analytical queueing models, many have missed two key aspects: (i) time-dependent workload variations and (ii) the impact of finite buffers for incoming user requests. Buffer size is crucial as it affects scheduling policies, system performance, and resource usage. The BTQ-MR model aims to address these aspects. The study examines the behavior of the BTQ-MR model under different conditions, including job arrival rates, scheduler allocation times, and mapper and reducer completion times. The model includes three service stages which are scheduling, mapper, shuffle, and reducer—with no waiting between them. Using transient differential equations, the study evaluates performance metrics such as average queue length, waiting time, blocking probabilities of mappers, and waiting probabilities in the shuffle phase. MATLAB simulations are used to analyze workload intensity, buffer size, and job completion times of various stages. Increasing buffer size reduces job blocking but may raise resource usage. Changes in arrival rates and completion times highlight the need for adaptive scheduling to ensure stability. The study emphasizes the importance of tuning parameters like the number of mappers and reducers to balance resource use and minimize completion times. The results provide insights into optimizing MapReduce systems to handle large-scale, dynamic workloads effectively. In summary, the BTQ-MR model addresses limitations of earlier approaches by considering finite buffers and time-dependent variations. It offers a practical method to analyze and improve MapReduce performance, laying the groundwork for further research and development.

Keywords : Transient queueing model, Finite buffer, MapReduce programming model, Performance metrics, Adaptive scheduling , MATLAB simulations

INTRODUCTION

The exponential growth of data in recent years has necessitated the development of robust distributed computing frameworks capable of efficient large-scale data processing. Apache Hadoop, with its MapReduce programming model, has become a pivotal solution in this domain, enabling the decomposition of complex computational tasks into manageable sub-tasks that can be processed in parallel across clusters [1]. This architecture not only enhances scalability but also ensures fault tolerance, making it indispensable for big data analytics [2].

Hadoop's MapReduce programming model plays a vital role in distributed computing environments, as it enables efficient processing by parallelizing workloads across multiple nodes in a cluster. Its ability to manage massive datasets effectively, through the divide-and-conquer approach, allows industries to handle diverse data types and volumes ranging from transactional logs to multimedia files. Furthermore, its open-source nature and compatibility with commodity hardware make it a cost-effective solution for large-scale data processing needs, democratizing access to powerful computing resources [3].

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Despite its widespread adoption, optimizing the performance of Hadoop MapReduce remains a critical area of research. Traditional performance evaluation methods have often relied on steady-state assumptions, which may not accurately capture the dynamic nature of real-world workloads. For example, systems often experience fluctuating job arrival rates and unpredictable resource availability, making steady-state models inadequate for comprehensive performance analysis [4]. Transient queueing models, on the other hand, offer the ability to analyze performance over time, capturing short-term behaviors that significantly affect the system's responsiveness and throughput [5]. This time-dependent perspective is particularly valuable in dynamic environments, such as cloud computing and data-intensive workflows, where demand spikes are common.

Moreover, the finite capacity of buffers in practical systems introduces additional complexities, such as blocking and waiting phenomena, which can significantly impact overall performance [6]. For instance, a saturated buffer can lead to job rejection or delays, negatively affecting the system's throughput and service-level objectives. These challenges are compounded in Hadoop MapReduce, where the shuffle phase, characterized by data movement between mappers and reducers, is particularly sensitive to resource constraints [7]. Effective buffer management and transient analysis are, therefore, critical to mitigating these challenges and ensuring smooth workflow execution.

Despite the fact that various researchers proposed many models still there are certain limitations in suggesting the best analytical MapReduce framework . These researchers do not consider the time as a factor influencing performance and they also confine their analysis to deterministic service times. This work, proposes an analytical transient queuing model with a finite buffer to evaluate the performance of MapReducer by considering jobs at different arrival times and their completion times at both mappers and reducers. The analytical transient queuing model analyzes jobs waiting in buffer of varying length using a multi-server queuing model. It may help in bringing out the optimality by minimizing the execution time of jobs, which leads to reduction in execution cost by proper parameter tuning.

The primary contributions towards the work are carried out as follows

i) The proposed BTQ-MR analytical transient queuing model evaluates the performance of Hadoop MapReduce with varying buffer sizes, k mappers and n reducers.
ii) A transient state diagram is drawn for all possible cases
iii) Based on the transient state diagram, the transitional differential equations are derived to determine performance measures.
iv) Numerical illustrations are performed using MATLAB with some examples and conclusions are drawn based on the results.

The rest of the paper is presented as follows. Section 2 describes the related work. Section 3 presents the analytical transient queueing model to capture the dynamism of the MapReduce computations. In section 4, the numerical illustration with examples depict how to compute performance measures with this model. Finally, section 5 discusses the conclusions and suggest the future scope of work.

The transient behavior of finite-buffer queuing models has been explored in various contexts, including telecommunication networks and production systems, providing insights into queue dynamics, waiting times, and throughput under varying load conditions [8]. However, research specifically addressing the transient analysis of MapReduce frameworks with finite buffer constraints remains limited. [8]. Addressing this gap is essential for improving performance and resource utilization, especially as workloads become increasingly complex and heterogeneous.

To address these challenges, this paper proposes a Buffer Transient Queueing Model for MapReduce (BTQ-MR) that incorporates finite buffer capacities and time-dependent workload variations. Using transient differential equations, the model captures the dynamic behavior of the MapReduce framework, providing insights into key performance metrics such as average queue length, waiting time, blocking probabilities, and waiting probabilities during the shuffling phase [9]. These performance metrics are critical for understanding the system's operational efficiency and identifying bottlenecks that could hinder job completion rates. Numerical simulations conducted

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

using MATLAB will illustrate the impact of various input parameters, including job arrival rates and service times of mappers and reducers, on system performance [10].

The proposed BTQ-MR model also contributes to the development of adaptive scheduling policies that dynamically adjust resource allocation based on workload characteristics. For example, by analyzing transient behavior, the system can prioritize high-priority jobs or scale resources during peak demand periods, reducing latency and improving overall throughput [11]. Such capabilities are essential in real-world scenarios, where workload patterns are often non-deterministic and influenced by external factors such as user behavior and network conditions.

The findings from this study are expected to contribute to the development of optimized scheduling policies and resource allocation strategies for Hadoop MapReduce environments, ultimately enhancing their efficiency and responsiveness in handling large-scale data processing tasks [12]. This work lays the foundation for further research into transient queueing models for distributed systems, highlighting the importance of time-dependent analysis in modern computing architectures.

LITERATURE SURVEY

The need for efficient data processing frameworks has driven significant research in distributed computing and performance optimization. MapReduce, a key programming model of Apache Hadoop, has garnered extensive attention for its scalability and fault tolerance. However, challenges related to workload dynamics and resource management continue to persist. This section reviews relevant studies that address these issues, focusing on queueing models, buffer constraints, and adaptive scheduling strategies for Hadoop MapReduce.

Recent studies have highlighted the importance of transient analysis for understanding system behavior under dynamic conditions. Darapaneni et al. [1] presented an analytical transient queueing model to evaluate MapReduce performance, focusing on job arrival rates and buffer constraints. Similarly, Kempa and Marjasz [2] explored transient queue-size distribution in finite-buffer systems, demonstrating the impact of batch arrivals and vacation policies on performance metrics. These studies underscore the necessity of incorporating transient analysis into distributed computing frameworks to capture time-dependent variations.

Chakravarthy [3] extended the application of transient models by analyzing M/M/1 queueing systems with multiple differentiated vacations. This work provides insights into the effect of customer impatience and waiting servers on performance metrics, which are relevant for systems with dynamic workloads such as MapReduce. Earlier, Kempa [4] investigated finite-capacity queueing models with balking and repair periods, emphasizing the role of transient analysis in evaluating system reliability and responsiveness.

Buffer constraints significantly influence the performance of distributed systems. Effective management of buffer size is essential to minimize blocking probabilities and ensure smooth job execution. White [5] highlighted the role of buffer management in Hadoop's performance optimization, emphasizing the need for tailored scheduling policies. Fortes and Squillante [6] applied queueing models to evaluate resource allocation strategies, providing a foundation for understanding the interplay between system capacity and workload demands.

Scheduling policies play a critical role in optimizing MapReduce performance. Ananthanarayanan et al. [7] addressed outliers in MapReduce clusters using adaptive mechanisms to balance load and minimize job delays. This approach is particularly relevant for systems with varying job arrival rates. Zaharia et al. [8] further improved MapReduce performance in heterogeneous environments by incorporating dynamic scheduling policies that adapt to workload characteristics.

Early research by Dean and Ghemawat [9] introduced the MapReduce programming model, laying the groundwork for subsequent performance optimization studies. Their work highlighted the potential of MapReduce for handling large-scale datasets efficiently. Yang et al. [10] expanded on this by predicting the performance of parallel applications across platforms, demonstrating the scalability of the MapReduce framework.

Recent advancements in distributed computing focus on integrating queueing models with machine learning techniques to predict workload patterns and optimize resource allocation. Emerging research emphasizes the role of automation and intelligent scheduling in improving system efficiency. These studies point toward a future where

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

adaptive mechanisms driven by real-time data further enhance the performance of distributed computing systems like Hadoop MapReduce.

3. PROPOSED BTQ-MR MODEL

In this section a Transient Queueing model for MapReduce is proposed to analyse the performance of Hadoop MapReduce computing system by adopting M/M/(m+k) model with three phase of services, where m is the number of mappers and reducers configured in the system and K is the size of the buffer available at the entry of the map reducer model. With the help of this method, first-order differential equation approach is applied to explore the dynamic behavior of the time dependent MapReduce system and find the various performance measures like Average queue length, waiting time, and blocking probabilities of the mappers and reducers. For this analysis, the basic architecture of the Hadoop MapReduce computing system is considered, as defined by Khaled Salab et al. [8] and Guzlan Miskeen [12].

3.1 Assumptions

The BTQ-MR model is finite buffer transient queueing model with m mappers and n reducers, so the model depicted as a standard queueing notation i.e M/M/(m+K) with three stages.. The following assumptions are made for computing the performance measures through the transitional differential equations.

1. The BTQ-MR is a buffered queueing model and no waiting between these phases.

2. The job arrival rate follows Poisson distribution and the job completion times of both mappers and reducers follow an exponential distribution.

3. The jobs are accepted in a FIFO manner.

4. New jobs are blocked, when the buffer is fully occupied.

5. System failures are not considered.

6. The job completion time of the mapping phase also includes the shuffling phase.

$$\lambda P_{000} = \mu_3 P_{001} \tag{1}$$

$$(\lambda + \mu_1) P_{100} = \lambda P_{000} \tag{2}$$

$$(\lambda + \mu_2) P_{0i0} = \lambda P_{0,i-1,0} + \mu_3 P_{0i1} \;\; ; 1 \leq i \leq m-1 \tag{3}$$

$$(\mu_2) P_{0i0} = \lambda P_{0,i-1,0} \;+ \mu_1 P_{1,i-1,0} + \mu_3 P_{0i1} \;;\; i = m \tag{4}$$

$$\left(\lambda + \mu_3 P_{00j}\right) = \lambda P_{00ji-1} ++ \mu_3 P_{00j+1}; 1 \leq j \leq n-1 \tag{5}$$

$$\mu_3 P_{00j} \;\; = \lambda P_{00j-1} + \mu_2 P_{0,1,j-1}; \text{J=n} \tag{6}$$

$$(\lambda + \mu_2 + \mu_3) P_{0ij} = \lambda P_{0,i-1,j} + \lambda P_{0,i,j-1} + \mu_3 P_{0,i,j+1} + \mu_1 P_{1,,i-1,,j}; 1 \leq i \leq m-1, \quad 1 \leq j \leq \; n-1 \tag{7}$$

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

$$(\lambda + \mu_2 + \mu_3)P_{0ij} = \mu_1 P_{1,i-1,j} + \mu_3 P_{0,i\ j+1};\ i = m, 1 \le j \le n-1 \tag{8}$$

$$(\lambda + \mu_2 + \mu_3)P_{0i,j} + \mu_1, P_{i-1,j} + \mu_2 P_{0,i+1,j-1}; j = n, 1 \le i \le m-1 \tag{9}$$

$$(\lambda + \mu_2 + \mu_3)P_{0ij} = \mu_1 P_{1\ i-1\ j}; i = m\ , j = n \tag{10}$$

$$(\lambda + \mu_1 + \mu_2 + \mu_3)P_{1ij} = \lambda P_{0,i,j} + \mu_3 P_{1,i\ j+1}(t) + \mu_1 P_{2\ i-1,j}\ ; 1 \le i \le m-1\ , \quad 1 \le j \le n-1 \tag{11}$$

$$(\lambda + \mu_1 + \mu_2 + \mu_3)P_{1ij} = \mu_3 P_{1\ \ i\ j+1}(t)\ ; i = m, 1 \le j \le n-1 \tag{12}$$

$$(\lambda + \mu 1 + \mu_2 + \mu_3)P_{1ij}(t)\ \ + \mu_1 P_{2,i-1,j} + \mu_2 P_{1,i+1\ ,j-1}(t)\ ; \quad i \le m-1, \mathrm{j = n} \tag{13}$$

$$(\lambda + \mu_1 + \mu_2 + \mu_3)P_{1,i,j} = \mu_1 P_{2i-1\ \ j-1};\ i = m, \mathrm{J=n} \tag{14}$$

$$(\lambda + \mu_1 + \mu_2 + \mu_3)P_{b\ i,j} = \lambda P_{b-1,i,j} + \mu_3 P_{b,i,j+1}\ ; \mathrm{i=m}\ , 1 \le i \le n-1 \tag{15}$$

$$(\lambda + \mu_1 + \mu_2 + \mu_3)P_{bij} = \lambda P_{b-1,i,j}; \mathrm{i=m\ ,j=n} \tag{16}$$

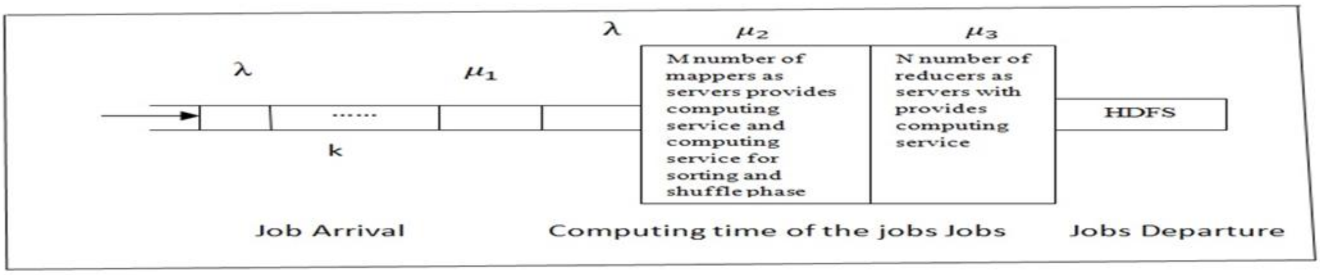


Fig.1 Hadoop mapreduce distributed process model with M/M/(C+K) queueing model

3.2 Derivation of performance metrics

The job arrival time and its processing time in MapReduce fluctuate due to dynamic changes in internet traffic, bandwidth consumption, and heterogeneous user demands. This fluctuation situation leads to be a complexity and the processing times vary over time. In this scenario the analysis of BTQ-MR behavior is represented as a function of time. To address these issues identified by several researchers, this work has adopted transient queuing model design for BTQ-MR. The corresponding state transient diagram is depicted in Annexure III for transient queuing model. Based on the state transient diagram the following transitional differential equations are derived for all possible cases at the time interval 't'. Let $P_{m,n,k}$ be the probability of 'm' mappers which are busy to provide service

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

at the rate of μ_2 , 'n' represents the number of reducers that are busy with a service rate of μ_3 and the scheduling service time in buffer is μ_1.

3.3 Derivation of Balanced Equations:

3.4 Performance Measures

1. Expected length of the system $=(L_S^{(t)}) = \sum_{i=1}^{m}\sum_{j=1}^{n}(m+n) * p_{i,j}^{(t)}$ (17)
2. Mean waiting time $(W_S^{(t)}) = \frac{L_S^{(t)}}{\lambda'}$ $, where\ \lambda' = 1 - p_{0,0}$ (18)
3. $Blocking\ probability\ of\ Mappers = \sum_{j=1}^{n} p_{m,j}^{(t)}$ (19)
4. $Waiting\ probability\ of\ shuffle\ phase = \sum_{i=1}^{m} p_{i,n}^{(t)}$ (20)

4. NUMERICAL ILLUSTRATION

In this section the numerical evaluation is carried out to investigate the effect of BTQ-MR for various combinations of input parameters λ, m, n, μ2 and K where μ2 and μ3 depend on the sizes of m,n. On the lines of Khaled Ssalah et al [3],the service time x for executing a single MapReduce job is considered as the sum of the mapper phase service time and reducer phase service time. The mean service time of mapper phse (i.e 1/β) depends on the speed of the slave node and the number of splits created for that MapReduce job. Let us assume 500 ms to execute the MapReduce job on single mapper and 100 ms on a single reducer. The mean service times for a single mapper and reducer can be computed as

$$\frac{1}{\beta} = \frac{500}{m}\,\text{ms}, \quad \frac{1}{r} = \frac{100}{n}\,\text{ms}$$

The service time x is calculated using the below formula.

$$x = \frac{1}{\beta}\sum_{i=1}^{m}\frac{1}{i} + \frac{1}{r}\sum_{i=1}^{n}\frac{1}{i} \quad (21)$$

In this study, the number of mappers and reducers are assumed as within a ratio of m:n =2:1. The transient values of 't' are considered in intervals from 0.5 to 2(number of intervals = 4). With the help of the MATLAB numeric computing platform, the above performance measures are calculated. MATLAB software is used to explore various probabilities and constants by developing a computational program to solve the system of differential equations. The influence of various parameters on system constants is studied, they are summarized in Tables 1 and 2, presented in Annexure I and Annexure II.

Results and Analysis

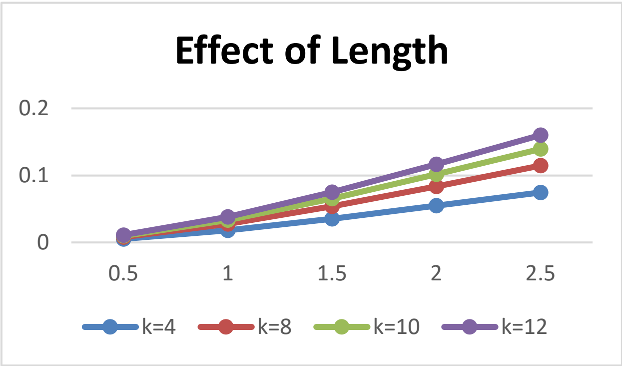
The earlier work [1] is carried out without buffer to find the effects of various performance metrics. It is also observed that there is no impact of number of reducers as well as job completion time. In this paper, the buffer, number of mappers and number of mappers and reducers effects are considered. The results analysis is presented as buffer increasing scenario and mappers, reducers increasing scenario.

Buffer Increasing Scenario

In this scenario, the metrics are calculated for increasing values of k and fixed values of , μ_1, μ_2, μ_3, m, n. The evaluated metrics are presented in the tables and the following figures.

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



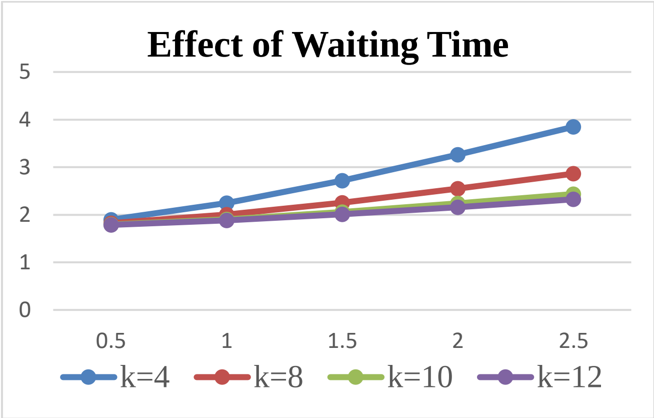
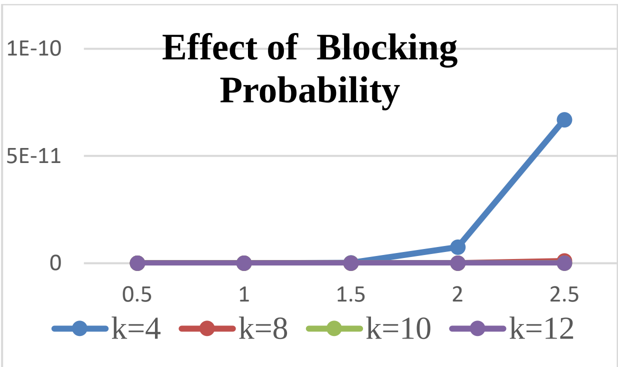


Fig.2 Impact of Length for various k values Fig.3 Impact of Waiting Time for various k values

The analysis of Table 8 is shown in the above graphs, the following trends are observed in this scenario:

- **Length:** Increases steadily with time (T) for all k values, indicating that the time progresses with respect to queue length.
- **Waiting Time:** Also exhibits a consistent increase with time, suggesting that as the system processes more tasks, the waiting time proportionally rises.
- **Blocking Probability of Mappers:** Remains close to zero initially but shows a slight increase at higher time intervals, implying that buffer saturation starts to impact mapper availability only at extended durations.
- **Waiting Probability of Shuffling Phase:** Follows a similar pattern to the blocking probability, indicating that delays in the shuffling phase are minimal until higher time intervals.

From the above buffer increasing scenario, it is concluded that the steady increase in length and waiting time suggests a need for adaptive scheduling to manage buffer saturation, especially at higher time intervals.



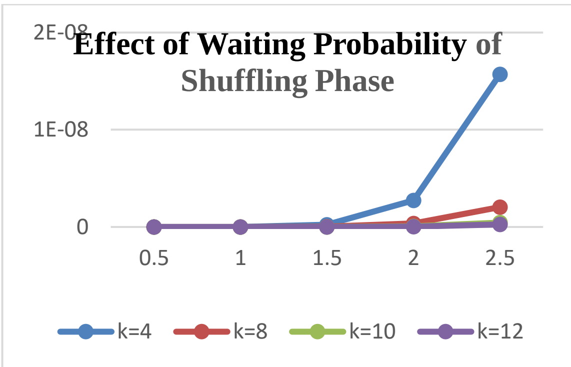


Fig.4 Impact of Blocking Probability of Mappers for various k values Fig.5 Impact of Waiting Probability of Mappers for various k values

Mappers Increasing Scenario

In this scenario, the metrics are calculated for increasing values of m and fixed values of , μ_1, μ_2, μ_3, k, n. The evaluated metrics are presented in the tables and the following figures.

- **Length and Waiting Time:** Increase with a higher number of mappers due to more tasks being processed concurrently. However, this also leads to a slight rise in blocking probability, indicating that buffer limitations might become a concern at higher mapper counts.

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

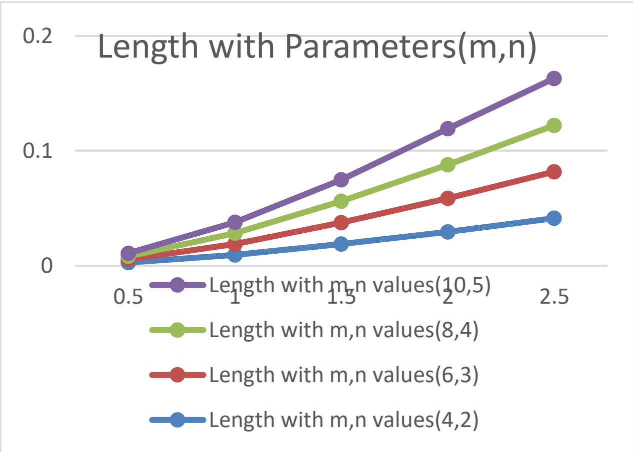
- **Blocking Probability:** Although low, shows a slight increase, suggesting that buffer management becomes critical with more mappers.
- **Waiting Probability of Shuffling Phase:**

It is thus observed that a balanced count of mappers with adequate buffer size is essential to prevent bottlenecks.

Mappers and Reducers Increasing Scenario

In this scenario, the metrics are calculated for increasing values of m, n and fixed values of , μ_1, μ_2, μ_3 and k. The evaluated metrics are presented in the tables and the following figures.

- **Length and Waiting Time:** Both metrics decrease as the number of mappers (m) and reducers (n) increases, demonstrating that higher parallelism in MapReduce tasks effectively reduces queue length and processing delays.
- **Blocking and Waiting Probabilities:** These probabilities remain low, reinforcing that a higher count of mappers and reducers alleviates bottlenecks in the system.



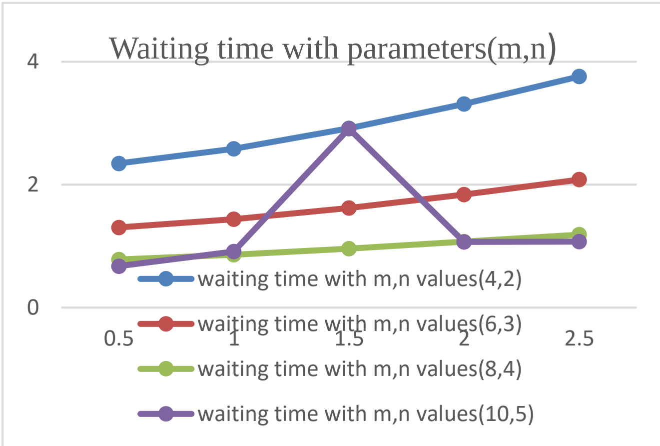
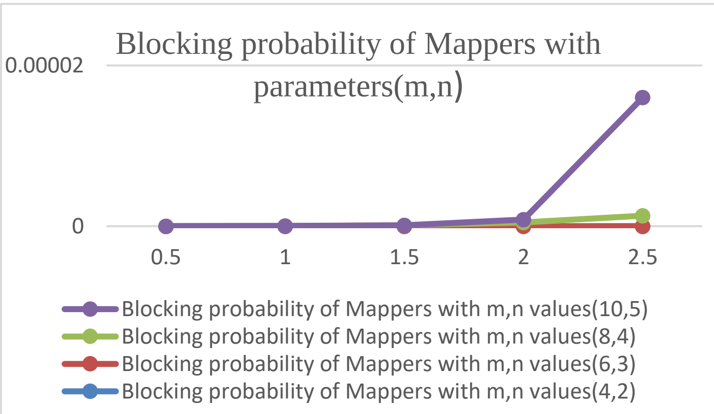


Fig.10 Impact of Length for various m,n values Fig.11 Impact of Waiting Time for various m,n values



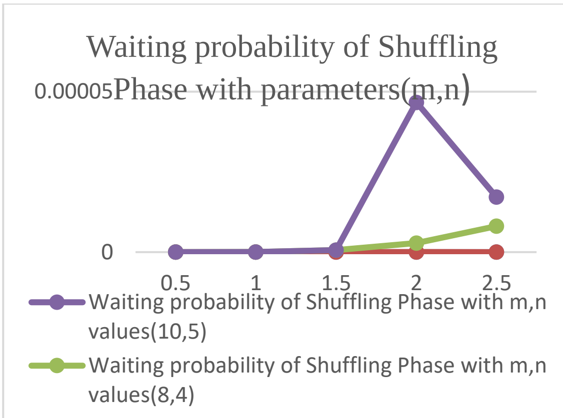


Fig.12 Impact of Blocking probability of Mappers for various m,n values Fig.13 Impact of Waiting Probability of Shuffling Phase for various m,n values

5. CONCLUSION AND FUTURE WORK

In this work, a transient analytical queueing model with finite buffers is proposed to address the problem of performance modeling in the Hadoop MapReduce framework. To achieve this objective, the BTQ-MR model is designed. This model effectively illustrates the impact of various parameters on performance metrics, providing valuable insights for optimizing MapReduce systems to handle large-scale, dynamic workloads efficiently. From

Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

this numerical study, it is concluded that increasing the buffer size significantly increases blocking probabilities and waiting time while also reducing the number of blocked jobs. These results may be useful for job scheduling and parameter tuning strategies. Increasing the number of mappers directly improves throughput and minimizes delays while maintaining a balanced buffer size to avoid bottlenecks. The model can be further enhanced to predict workloads in real-time and can also be applied to a wider range of buffer sizes.

REFERENCES

[1] Chandra Sekhar Darapaneni, Bobba Basaveswara Rao, Boggavarapu Bhanu Venkata Satya Vara Prasad, Suneetha Bulla, "An Analytical Performance Evaluation of MapReduce Model Using Transient Queuing Model," *Advances in Modelling and Analysis B*, vol. 64, no. 1-4, pp. 46-53, 2021.
[2] Wojciech M. Kempa, Rafał Marjasz, "Study on Transient Queue-Size Distribution in the Finite-Buffer Model with Batch Arrivals and Multiple Vacation Policy," *Entropy*, vol. 23, no. 11, 1410, 2021.
[3] S. R. Chakravarthy, "Transient Analysis of an M/M/1 Queueing System Subjected to Multiple Differentiated Vacations, Customer Impatience and a Waiting Server," *Proceedings of the Indian Academy of Sciences - Mathematical Sciences*, vol. 130, 2020.
[4] Wojciech M. Kempa, "Transient analysis of a finite-capacity queueing model with balking and repair periods," *AIP Conference Proceedings*, vol. 2048, 060008, 2018.
[5] T. White, *Hadoop: The Definitive Guide*, 4th ed., O'Reilly Media, 2015.
[6] J. A. B. Fortes, M. S. Squillante, "Queueing Models for Performance Evaluation of Computer Networks," in *Performance Evaluation of Computer and Communication Systems*, Springer, 2014.
[7] G. Ananthanarayanan, S. Kandula, A. Greenberg, I. Stoica, Y. Lu, B. Saha, E. Harris, "Reining in the Outliers in Map-Reduce Clusters Using Mantri," *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation*, 2010.
[8] M. Zaharia, A. Konwinski, A. D. Joseph, R. Katz, I. Stoica, "Improving MapReduce Performance in Heterogeneous Environments," *Proceedings of the 8th USENIX Conference on Operating Systems Design and Implementation*, 2008.
[9] 9.J. Dean, S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107-113, 2008.
[10] L. T. Yang, X. Ma, F. Mueller, "Cross-Platform Performance Prediction of Parallel Applications Using Partial Execution," *Proceedings of the 2005 ACM/IEEE Conference on Supercomputing*, 2005.