Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

TQ-MR : Transient Queuing Theoretical Approach for the Performance Evaluation of MapReduce Model

Chandra Sekhar Darapaneni[1] , Bobba Basaveswara Rao[2] , V.N.Rama Devi [3] Neelima G[4]
1&4 Department of CSE, Acharya Nagarjuna University, Guntur 522510, India
2 University Computer Centre, Acharya Nagarjuna University, Guntur 522510, India
3 Department of Humanities and Sciences, GRIET, Hyderabad

ABSTRACT

In this paper a Transient queueing theoretical approach is identified, for the evaluation of performance of the MapReduce programming model (TQ-MR) followed by analysis of the same is also presented. The suggested approach is intended to analyze the behavior of the TQ-MR by considering different job arrival rates at mappers and given job completion times of both mappers and reducers for various values of mappers and reducers. This TQ-MR model, with two stages of services and no waiting in between these stages. The derived transient differential equations are used to finding the performance measures like Average queue length, waiting time and Blocking Probabilities of Mappers and Waiting Probability of Shuffling phase as time progresses. These measures are computed based on numerical simulation experiments conducted on MATLAB. Finally results are presented and also to portray the effect of various input parameters.

1. INTRODUCTION

Apache Hadoop is a framework intended to give the functionality of a Distributed system. MapReduce is a programming model written in Java developed in conjunction with Hadoop. The main objective of this combination is to process datasets with large size. It basically process data on computational nodes of clusters parallel in a well defined manner. MapReduce layer is programming model consisting of Mappers and Reducers that process data across the clusters. Mappers place the data in parallel clusters for computation and Reducers are used for assembling data together. Most of the search engines like Bing, Google Search, Yahoo, Yadex and Baidu etc do invariably use MapReducers on their data through Hadoop clusters.

Hadoop MapReducers are highly scalable which can execute on low-cost hardware commodities. Accordingly, it is also resilient and doesn’t require high coding and completely depends on Mappers and Reducers function to process data. As MapReducers process enormous of data, it is always required to define and fine tune parameters for jobbing the tasks. The decision for improving the performance always depends on defining the number of mappers and reducers to be used for process execution. With the increase in the number of reducers leads potentially may lead to some disadvantages like increase in the overhead of the framework coupled with load balancing. In this case, several administrators need to pay much attention in designing the MapReduce Hadoop framework for obtaining the better performance.

Over the past few years, with the growing demand of MapReduce framework, researchers are working on optimizing the performance in means of execution time, and processing speed through analytical, simulation and experimental models. When compared to other models the analytical models most of the time yield accurate and reliable estimates for performance metrics with a considerably low cost. Yang et al [1] designed an analytical model and validated with experimental evaluation for MapReduce performance evaluation Vianna et al [2] proposed an analytical model for Hadoop workloads, with combination of precedence graph and queueing network models. Khaled Ssalah et al [3] focussed on designing analytical model to run parallel jobs on cloud clusters using finite queuing model with minimum number of computing resources and also analyzed various performance metrics such as throughput, response time, and probability for blocking through simulation. Zuqiang Ke and Nohpill Park[4] developed an analytical model for availability analysis of MapReduce computing on a Hadoop platform and represented the same as a queueing model. They also optimized the availability of MapReduce computing resources supported by the derivation of suitable balancing equations. The same authors[5] also proposed an analytical model for Hadoop platform, using mapreduce model to evaluate the probability of availability of map-reduce computing at a instant of time. Based on these studies the TQ-MR model is developed and evaluated through numerical examples.

Despite the fact that various researchers proposed many models still there are certain limitations in suggesting best analytical MapReduce framework . These researchers do not consider the time as one of the factor that influence performance and they also confine to deterministic service time. This work, proposes analytical transient queuing model which evaluates the performance of MapReducer by considering jobs at different arrival times and completion times of jobs at Mappers including Reducers. The analytical transient queuing model define waiting queue with varying length by using queuing model with multi-server. 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 analytical transient queuing model for evaluating on the performance of Hadoop MapReduce with the configuration of m mappers and n reducers.
ii) The transient state diagram is drawn for all possible cases
iii) Based on the transient state diagram the transitional differential equations are derived to finding the performance measures.
iv) Numerical illustration is carried out 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 depicts how to arrive at the performance measures with this model. Finally section 5 discusses the conclusions and suggest the future scope of work.

2. RELATED WORK

For the past so many years, the performance evaluation of MapReduce frame work on Hadoop platform has been discussed by several researchers. The myriad of related works can be categorized in to three types of evolutionary models, they are analytical, simulation and experimental models. In view of the objective of this study, this section confirms only to review the literature belongs to the analytical models and specifically focus on queueing models that are used for computation of performance measures of MapReduce models.

Xio Yang and Jianling Sun [6] proposed an analytical model to improve the performance by minimizing the job execution time through modifying the Map split granularity and number of reducers without modifying the framework. Shouvik Bardhan and Daniel A.Menasce[7] build a model to predict completion time of the map phase of MapReduce Jobs using queueing network models. They also conducted experiments to validate the model. Khaled Salab et al.[8] presented an analytical queueing model to achieve elasticity for MapReduce jobs on cloud and to determine the minimum number of mappers and reducers required to satisfy the SLO response time under various workload conditions. Xiaolong Yu and Wei Li [9] investigated the analytical model with adoption of queueing theory for big data processing. The developed queueing model discovers the nature of the MapReduce programming model and also find the utilization and mean waiting time for the mappers as well as reducers. They have used this model to improve the system utilization by adjusting workload and tuning the system parameters. They suggested a simulation model for validating this model. F.Farhat et.al[10] analytically

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

investigated the stochastic behavior mapper nodes that impact on job completion time of the MapReduce job. They optimize the Mean Sojourn time with respect to task inter-arrival time to the mapper node. Their experimental results yields the performance and most required parameters of the various types of schedulers targeting MapReduce applications. Zuqiang Ke and Nohpill Park[11] proposes analytical model with representation of queueing model for availability of MapReduce computing on a Hadoop platform at an instant of time. The availability model considers the number of map and reduce tasks, number of nodes engaged, task arrival/exit rates and failure/repair rates to finding the availability. In this study the performance measures are derived from balance equations. The efficacy of the model is demonstrated by conducting parametric simulations and achieving the availability with respect to throughput. The M/G/1/K performance model with FCFS discipline is proposed by the Guzlan Miskeen[12] for finding the MapReduce performance. The performance measures like mean response time, loss probability and mean queue length are calculated with numerical investigation for various values of number of mappers, reducers and arrival rates of the jobs. The model is analyzed via discrete-event simulation. Kui Li et.al [13] developed elastic scaling algorithm based on finite multi server queueing model for balancing the average waiting time of tasks and total resource utilization rate of the cluster. Evaluation function and QoS constraints are the key factors of this model and Particle swarm optimization is used to search the feasible solution space determined by the constraints are applied in this algorithm.

3. PROPOSED TQ-MR MODEL

In his section a Transient Queueing model for MapReduce (TQ-MR) is proposed to analyse the performance of Hadoop MapReduce computing system by adopting M/M/1 model with two stages of services. In the first stage consists of mapper phase service and shuffle phase service with where C is the sum of the number of mappers and reducers are configured in the system. With help of the method, first order difference differential equations approach to explore the dynamic behaviour 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 in Khaled Salab et al.[8] and Guzlan Miskeen[12].

3.1 Assumptions

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

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

1.The TQ-MR is buffer less queueing model and no waiting in between the stages.

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

3. The jobs are accepted in FIFO manner.

4. The new jobs are blocked, when all the mappers are busy.

5. The failures of the system are not considered.

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

3.2 Derivation of performance metrics

The job arrival time and its processing time in MapReduce are fluctuated manner, because of the dynamic changes are occurred in Internet traffic, bandwidth consumption and user behaviour etc. In view of the fluctuation situation leads to be a complexity and the processing times are vary from time to time. In this scenario the analysis of TQ-MR behaviour is represented as a function of time. By observing all these issues that were faced by several researchers, this work has adopted transient queuing model design for TQ-MR. The corresponding state transient diagram is depicted in the 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 $P_{m,n}$ be the probability of 'm' mappers which are busy to provide service at the rate of μ_1 and 'n' represents the number of reducers that are busy with a service rate of μ_2

3.3 Derivation of Balanced Equations:

$$\frac{dP_{0,0}(t)}{dt} = -\lambda P_{0,0}(t) + \mu_2\, P_{0,1}(t) \quad (1)$$

$$\frac{dP_{i,0}(t)}{dt} = -(\lambda + \mu_1)P_{i,0}(t) + \mu_2\, P_{i,1}(t) + \lambda P_{i-1,0}(t), 0 < i < m, j = 0 \quad (2)$$

$$\frac{dP_{i,0}(t)}{dt} = -(\mu_1)P_{i,0}(t) + \mu_2\, P_{i,1}(t) + \lambda P_{i-1,0}(t), i = m, j = 0 \quad (3)$$

$$\frac{dP_{0,j}(t)}{dt} = -(\lambda + \mu_2)P_{0,j}(t) + \mu_1\, P_{1,j-1}(t) + \mu_2\, P_{0,j+1}(t), i = 0, 0 < j < n \quad (4)$$

$$\frac{dP_{0,j}(t)}{dt} = -(\lambda + \mu_2)P_{0,j}(t) + \mu_1\, P_{1,j-1}(t), i = 0, j = n \quad (5)$$

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

$$\frac{dP_{i,j}(t)}{dt} = -(\lambda + \mu_1 + \mu_2)P_{i,j}(t) + \lambda P_{i-1,j}(t) + \mu_1 P_{i+1,,j}(t) + \mu_2 P_{i,j+1}(t), i < m, j < n \quad (6)$$

$$\frac{dP_{i,j}(t)}{dt} = -(\mu_1 + \mu_2)P_{i,j}(t) + \lambda P_{i-1,j}(t) + \mu_2 P_{i,j+1}(t), i = m, j < n \quad (7)$$

$$\frac{dP_{i,j}(t)}{dt} = -(\lambda + \mu_1 + \mu_2)P_{i,j}(t) + \lambda P_{i-1,j}(t) + \mu_1 P_{i+1,,j+1}(t), i < m, j = n \quad (8)$$

$$\frac{dP_{m,n}(t)}{dt} = -(\mu_1 + \mu_2)P_{m,n}(t) + \lambda P_{m-1,n}(t), i = m, j = n \quad (9)$$

In this study Average Queue Length, Blocking Probability of mappers and reducers, and Waiting time , performance measures are computed based on the above transitional differential equations.

3.4 Performance Measures

1. Expected length of the system =$(L_S^{(t)}) = \sum_{i=1}^{m}\sum_{j=1}^{n}(m+n) * p_{i,j}^{(t)}$ (10)
2. Mean waiting time $(W_S^{(t)}) = \frac{L_S^{(t)}}{\lambda'}$ $, where\ \lambda' = 1 - p_{0,0}$ (11)
3. $Blocking\ probability\ of\ Mappers = \sum_{j=1}^{n} p_{m,j}^{(t)}$ (12)
4. $Waiting\ probability\ of\ shuffle\ phase = \sum_{i=1}^{m} p_{i,n}^{(t)}$ (13)

4. NUMERICAL ILLUSTRATION

In this section the numerical evaluation is carried out to investigate the effect of performance of TQ-MR for various combinations of input parameters λ, m and n where µ1and µ2 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 the sum of the mapper phase service time and reducer phase service time. The mean service time of mapper phse (i.e 1/β) is 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 to execute on single reducer. So the single mapper and single reducer mean service times can be computed as 1/ β = 500/m ms and 1/r =100/n ms respectively. The service time x is calculated using the below formula.

$$x = \frac{1}{\beta}\sum_{i=1}^{m}\frac{1}{i} + \frac{1}{r}\sum_{i=1}^{n}\frac{1}{i} \quad (14)$$

In this study the number of mappers and reducers are taken as witha a ratio of m:n i.e 2:1. The transient values of ‘t’ 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 programme to solve the

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

system of differential equations. The influence of various parameters on system constants is studied and they are shown in Tables 1 and 2 which are presented in Annexure I and Annexure II.

(i) Effect of λ: For the numerical illustrations to show the effect of λ, the model parameters are considered with the values $\lambda = 0.25$, $\mu_1 = 0.0059, \mu_2 = 0.0192, m = 8$ and $n = 4$.The time instances are taken as $t_1 = 0.5, t_2 = 1.0, t_3 = 1.5, t_4 = 2.0$.

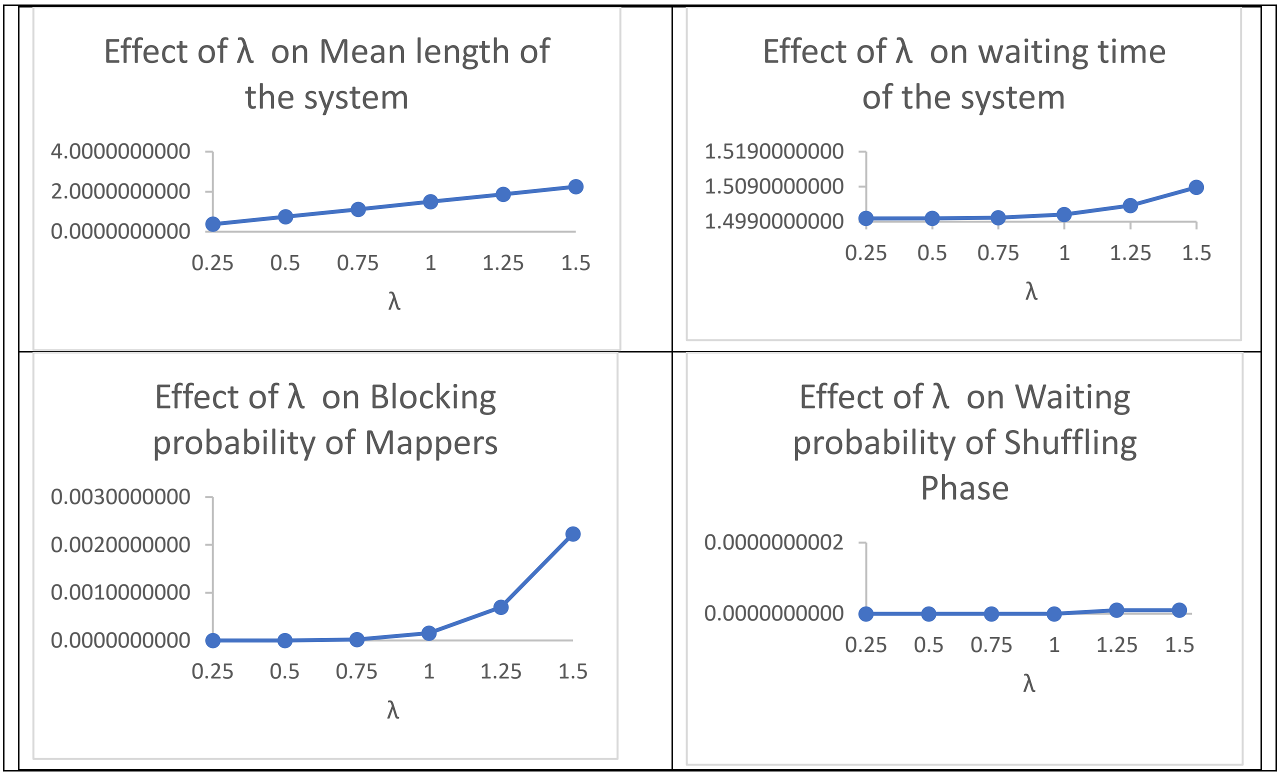


Figure 1: Effect of λ

(ii) Effect of m, n: For the numerical illustrations to show the effect of m, n, the model parameter λ is considered as "1" and μ_1 *and* μ_2 values are changed with the changes in m,n The time instances are taken as follows: $t_1 = 0.5, t_2 = 1.0, t_3 = 1.5, t_4 = 2.0$

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

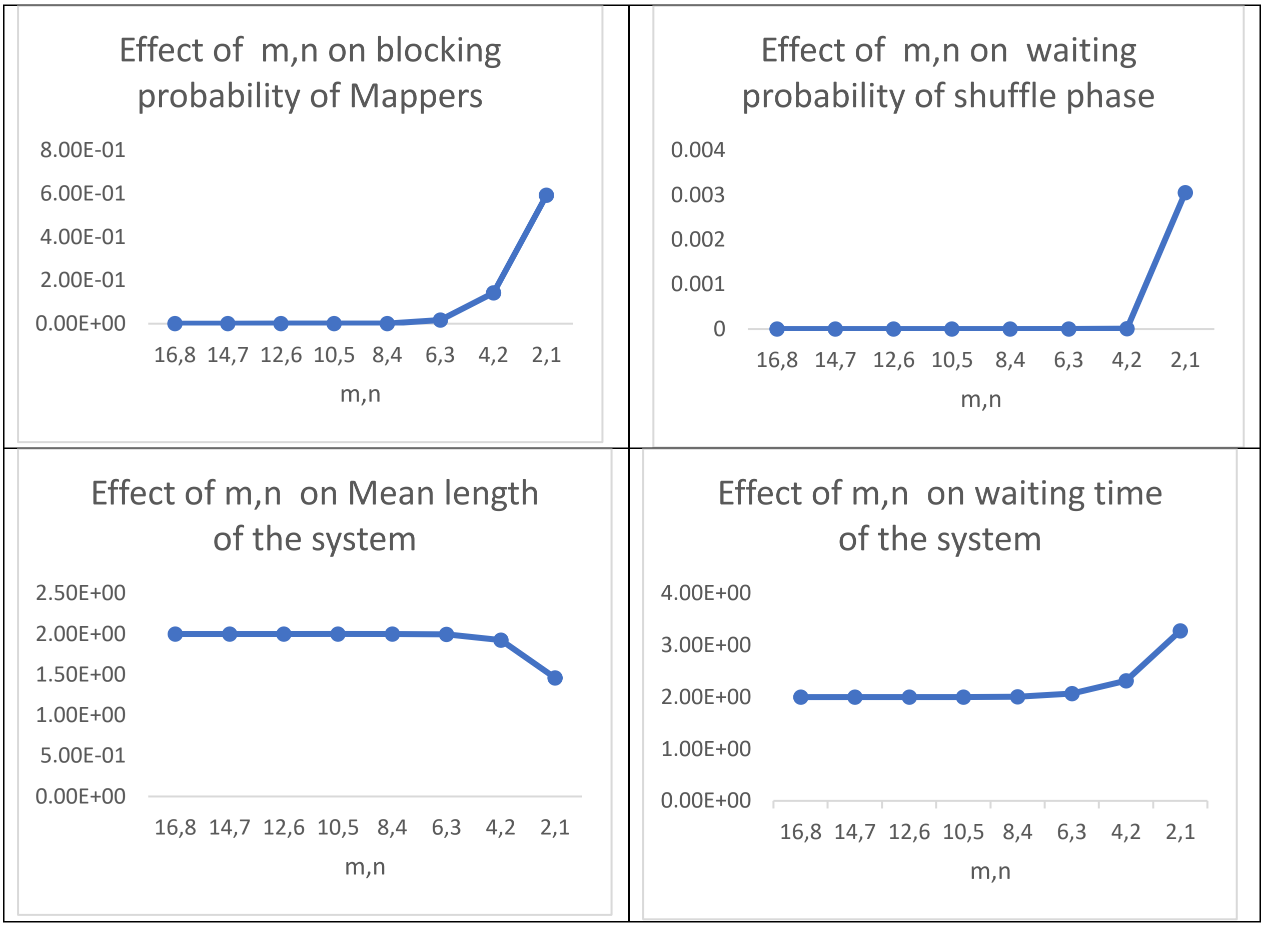


Figure 2: Effect of m, n

5.RESULTS AND DISCUSSION

The aforementioned section presents some results from the Matlab simulation setup described in previous sections. To explore the various performance measures for conducting numerical experiments the total completion time of the mapreduce job is assumed. It includes both mappers + shuffling and reducer phases completion times. This completion time is chosen from the SLA.

A. Mean Queue Length Analysis

To observe the behaviour of the mean queue length, to find out that the queue length is maintain stable average queue length or not for the peak arrival rates and the change rate is stable when the time is in progress. From the Table-I and Fig 1 it is observed that the mean queue length is increased when λ increases, this phenomena will be follows, when the time is in progress. The Table-II and Fig 2 shows that the queue length is all most all same for the increment of m,n ratio for fixed values of λ and job completion times, but when the time is progress from 0.5 to 2 the queue length is

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

significantly increased. For constant or minute variations of λ there is no need to scaling the mappers and reducers. Without aware of the job arrival pattern to scaleing the mappers and reducers it is a over burden for the system resource consumption like a memory ,CPU utilization and network bandwidth consumption etc.,. It leads to decrease of the total performance and the response time will be increased. Finally it is conclude that the queue length will not be effect for the almost all static values of λ , even though the increasing of mappers and reducers are not yield any further benefits.

B. Waiting Time

Waiting Time is key performance measure to identify the how much system response time will be effected. When the time is in progress it is observed in a two fold for satisfying the SLO (i) the various values of λ for the given configuration values m and n, (ii) the various values of m,n ratio for given λ. In order to obtain the results of Table II and Fig() the waiting time exhibits at any given point of time, little bit of variation when λ increases for constant values m,n. But the time is in progress (from 0.5 to 2), it will be significantly increased. If the m,n ratio will be increased the waiting time will not much effected for given λ to meet the required SLO and it yields opposite results i.e it increases noticeably when time is in progress. After keen observe the results of these scenarios, to find that the changes in waiting time is negligible for fixed point of time with the variations of λ and m,n ratio. The system operation time will be progressive the waiting time will be increased for fixed values of λ and m,n ratio. This will be true for various combination of fixed values of λ and m,n ratio to reach the SLO.

C. Blocking Probability of Mappers

Blocking Probability of Mappers is an important measure, that the configured number of mappers is enough for accepting the incoming job requests or not. It plays key role to take a decision for scaling the mappers. As per the Table I and II and graphs () the measure is low initially and gradually increased when the job arrival rates are increased for fixed values of mappers and reducers with a constant service rates of both mappers and reducers. This same line of behaviour exhibits when the time is in progress. The Blocking Probability of Mappers is always greater than Waiting Probability of Shuffle Phase for the static number of mappers and reducers with fixed service rates of mappers and reducers. For fixed job arrival rates and constant service rates with increasing of mappers and reducers the Blocking Probability of Mappers is decreases and it is move to zero. This will be happened because of two reasons i) the mappers are double in size of reducers or in some configuration it should be higher than number of reducers, ii) the tasks should be completed by the mappers first subsequently to be completed by the reducers. In view of this reasons the Blocking Probability of Mappers should be optimized more than Waiting Probabilities of Shuffle Phase.

D) Waiting Probabilities of Shuffle Phase

By observing Table I the Waiting Probabilities of Shuffle Phase values are gets zero, when λ increasing, with constant values of m,n and service rates of two stages. From the Table II and the Fig() the values are decreasing and tends to zero when the m,n ratio is increasing for the constant job arrival rates and service rates of stage 1 and stage 2. This means that the more number of reducers are configured the waiting probability of shuffle phase is tends to zero, this situation leads to job completion times are minimized. But more number reducers configured, will causes to more consumption of computational resources, This scenario degrade the performance of mapreduce model.

6.CONCLUSION

The presented analytical model based on transient queueing theoretical approach is used to investigate the performance of mapreduce model when time is in progress state. Performance measures are derived based on transitional differential equations. The measures Queue Length, Waiting Time, Blocking probability of Mappers and waiting probability of shuffle phase are computed for given input values of the job arrival rate, job completion time of the mappers and reducers with different combinations of configured mappers and reducers. The numerical experiments are conducted for various input parameters with time ‘t’. From the obtained results, key observations drawn are i) when the time is progress and the job arrival rate also increases the queue length will increases i.e the indication of number of jobs in process state in two stages of mapper , shuffle and reducer phases. ii) The waiting time will focus on how much time the jobs can wait in the system. The waiting time will not exhibit noticeable change while increasing job arrival rates, mappers and reducers ratios, but it will take opposite direction when time is in progress. iii) To identify how many minimum number of mappers are needed for accepting the incoming job requests with out rejecting due to overload. iv) To identify whether it is a suitable combination of mappers and reducers to meet the current workload conditions and also explores the number of reducers if they are busy or idle. Finally the authors have concluded that the presented TQ-MR model and the performance measures will be given birds-eye-view guidance for developing new job scheduling algorithms, performance optimization models and derive the auto tuning models for various mapreduce configurable parameters. The current work can be extended by enhancing the TQ-MR model with buffer queue for job arrivals and to suggest the suitable job scheduling algorithms that will help in reducing the costs for cloud service providers and also users by providing minimum resources with minimum job completion times.

Annexure I

Table 1: Effect of λ

Parameter(λ) Values	T	0.5	1	1.5	2
0.25	Length	0.124999	0.249996	0.374986	0.499967
	waiting time	0.499998	0.999982	1.499943	1.999873
	Blocking probability of Mappers	0	1.03E-10	5.47E-09	5.71E-08
	Waiting probability of Shuffling Phase	0	4.49E-16	2.51E-14	2.69E-13
0.5	Length	0.249999	0.499992	0.749974	0.99994
	waiting time	0.499998	0.999984	1.499967	2.000033
	Blocking probability of Mappers	0	2.65E-08	1.07E-06	9.70E-06
	Waiting probability of Shuffling Phase	0	7.19E-15	3.29E-13	3.20E-12
0.75	Length	0.374998	0.749988	1.124961	1.499891
	waiting time	0.499998	0.999994	1.500167	2.001369
	Blocking probability of Mappers	0	6.79E-07	2.08E-05	0.000164
	Waiting probability of Shuffling Phase	0	3.64E-14	1.37E-12	1.21E-11
1.0	Length	0.499998	0.999985	1.49993	1.999627
	waiting time	0.499998	1.00004	1.501059	2.006483
	Blocking probability of Mappers	0	6.78E-06	0.000156	0.001081
	Waiting probability of Shuffling Phase	0	1.15E-13	3.54E-12	2.85E-11
1.25	Length	0.624998	1.249982	1.874797	2.498452
	waiting time	0.499998	1.00018	1.503579	2.018574
	Blocking probability of Mappers	0	4.04E-05	0.000697	0.004236
	Waiting probability of Shuffling Phase	0	2.81E-13	7.07E-12	5.23E-11
1.5	Length	0.749997	1.49998	2.249334	2.994704
	waiting time	0.499998	1.000446	1.508818	2.039491
	Blocking probability of Mappers	0	0.000174	0.002232	0.011965
	Waiting probability of Shuffling Phase	0	5.82E-13	1.20E-11	8.22E-11

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

Annexure II

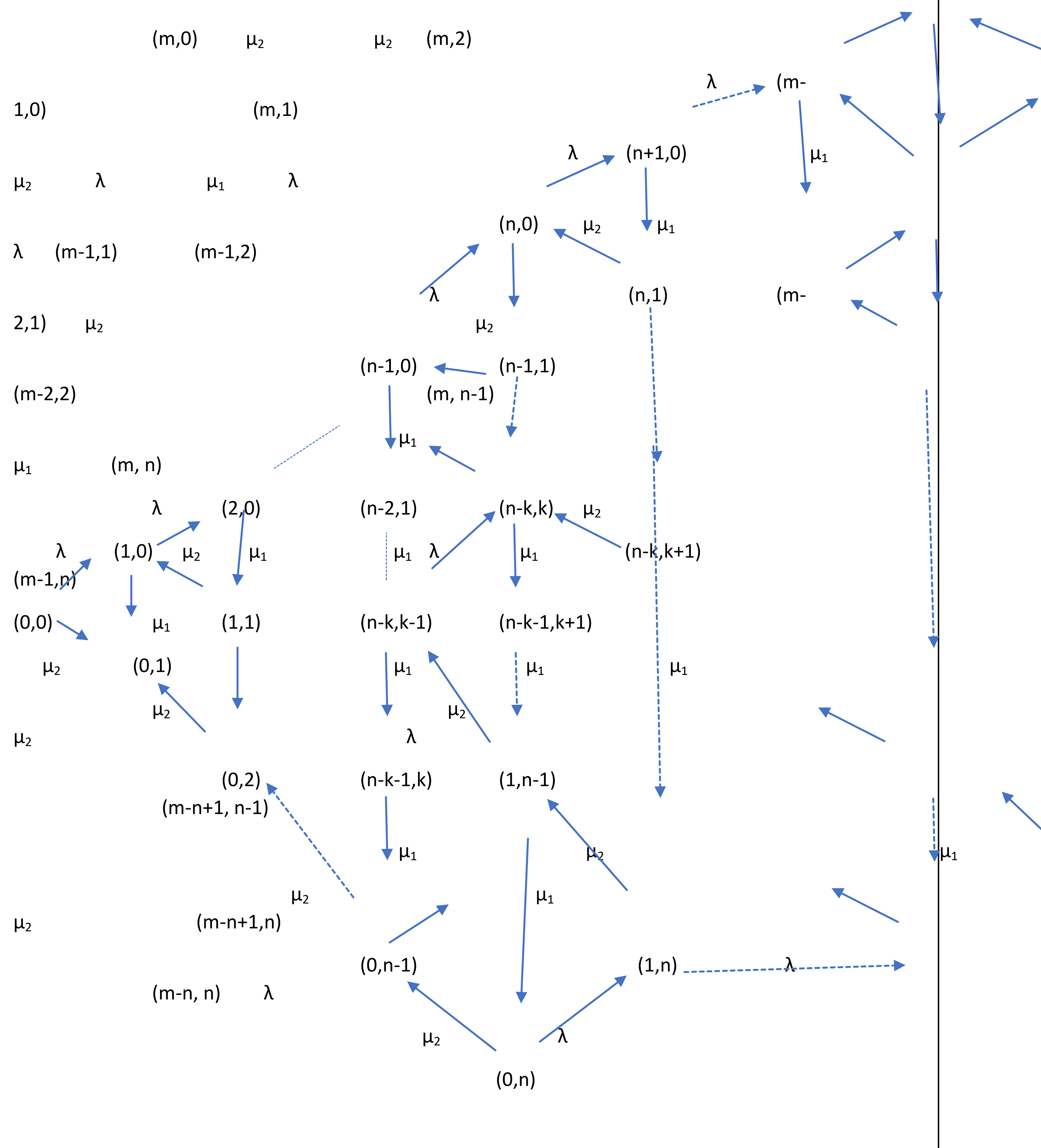
Parameters (m, n) Values	t	0.5	1.0	1.5	2.0
2,1	Length	0.484388888	0.897182242	1.219967749	1.459921479
	waiting time	0.63316842	1.299261181	2.08500815	3.278941059
	Blocking probability of Mappers	0.091082609	0.264863394	0.44209864	0.593155518
	Waiting probability of Shuffling Phase	0.000288012	0.000992007	0.001946547	0.003051114
4,2	Length	0.499999813	0.995586003	1.475728423	1.924989585
	waiting time	0.505160073	1.05522888	1.665811917	2.313336833
	Blocking probability of Mappers	0.002604167	0.019283804	0.065324213	0.142029142
	Waiting probability of Shuffling Phase	0.00000008	0.00000076	0.00000303	0.00000775
6,3 (1)	Length	0.499998548	0.999921912	1.498876208	1.9940355030
	waiting time	0.499998573	1.00005231	1.519983179	2.0654470550
	Blocking probability of Mappers	0	0.000574818	0.004493893	0.0165506760
	Waiting probability of Shuffling Phase	0.0000000000	0.0000000004	0.0000000039	0.0000000167
8,4 (2)	Length	0.499997942	0.999985165	1.499929515	1.999627119
	waiting time	0.499997978	1.000040353	1.501059445	2.006483299
	Blocking probability of Mappers	0	0.000007	0.000156329	0.001081158
	Waiting probability of Shuffling Phase	0.00E+00	1.15E-13	3.54E-12	2.85E-11
10,5 (3)	Length	0.499997296	0.999980516	1.499940629	1.999865521
	waiting time	0.499997344	0.999982389	1.499983384	2.000283519
	Blocking probability of Mappers	0	0	2.65E-06	4.16E-05
	Waiting probability of Shuffling Phase	0	0	1.93E-15	3.59E-14
12,6 (4)	Length	0.4999965770	1.0000000000	1.4999252050	1.9998396100
	waiting time	0.4999966380	1.0000000000	1.4999423600	1.9999184040
	Blocking probability of Mappers	0.0000000000	0.0000000000	0.0000000177	0.0000009250
	Waiting probability of Shuffling Phase	0.0000000000	0.0000000000	0.0000000000	0.0000000000
14,7 (5)	Length	0.5000000000	1.0000000000	1.4999080630	1.9998030910
	waiting time	0.5000000000	1.0000000000	1.4999289790	1.9998869460
	Blocking probability of Mappers	0.0000000000	0.0000000000	0.0000000000	0.0000000106
	Waiting probability of Shuffling Phase	0.0000000000	0.0000000000	0.0000000000	0.0000000000
16,8 (6)	Length	0.5000000000	1.0000000000	1.4999000000	1.9997825910
	waiting time	0.5000000000	1.0000000000	1.4999000000	1.9998844930
	Blocking probability of Mappers	0.0000000000	0.0000000000	0.0000000000	0.0000000000
	Waiting probability of Shuffling Phase	0.0000000000	0.0000000000	0.0000000000	0.0000000000

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

Table 2: Effect of (m, n)

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

Annexure III



Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 6, June : 2024

References

1. Xiao Yang, Jianling Sun [XJ],'AN ANALYTICAL PERFORMANCE MODEL OF MAPREDUCE', Proceedings of IEEE CCIS2011
2. E.Vianna et.al. 'Analytical Performance Models for MapReduce Workloads' Int J Parallel Prog (2013) 41:495–525, DOI 10.1007/s10766-012-0227-4
3. Khaled Salah, Jose M Alcaraz Calero 'Achieving Elasticity for Cloud MapReduce Jobs', 2013 IEEE
4. Zuqiang Ke, Nohpill Park, 'Availability Modeling and Assurance of Map-Reduce Computing', 2017 IEEE
5. Zuqiang Ke, Nohpill Park, 'Availability Modeling and Assurance for a Big Data computing ', Int.J. big Data Intelligence, Vol 6, No 2,2019.
6. Xiao Yang, Jianling Sun [XJ],'AN ANALYTICAL PERFORMANCE MODEL OF MAPREDUCE', Proceedings of IEEE CCIS2011
7. Shouvik Bardhan, Daniel A. Menasce 'Queueing Network models to predict the Completion Time of the Map Phase of MapReduce Jobs' Proc 2012, International Computer measurement Group Conf Las Vegas NV,Dec 3-4 2012.
8. Khaled Salah,Jose,M. AlcarazCalero [KA], 'Achieving Elasticity for Cloud MapReduce Jobs', 2013 IEEE
9. Xiaolong yu and Wei Li [9] Performance Modelling and Analysis of MapReduce/Hadoop Workloads', 2015 IEEE
10. F.Farhat, D.Z. Tootaghai, A Sivasubramaniam, M.Kanandemir,C.R.Das 'Modeling and Optimization of Straggling Mappers'
11. Zuqiang Ke, Nohpill Park, 'Availability Modeling and Assurance of Map-Reduce Computing', 2017 IEEE
12. Guzlan Miskeen 'Modeling and Performance Evalution of MapReduce in Cloud Computing Systems Using Queueing Network Model'
13. Kui Li, **Yi-mu Ji**, Shang-dong Liu,Hai-chang Yao,Hang Li,Shuai You,and Si-si Shao 'ACEA: A Queueing Model-Based Elastic scaling Algorithm for Container Cluster , Wireless Communications and Mobile Computing Volume 2021.