Statistical Modeling and Evaluation of Parallel Space-sharing Job Scheduling Algorithms for PC-cluster using Design of Experiments (DOE)

Amit Chhabra
Assistant Professor
Department of CSE, Guru Nanak Dev University,
Amritsar, India

Gurvinder Singh
Associate Professor
Department of CSE, Guru Nanak Dev University,
Amritsar, India

ABSTRACT
Parallel space-sharing job scheduling algorithms play an indispensible role in efficient allocation of processors of PC-cluster to the competing jobs to achieve one of the performance objective(s) viz. minimized mean response time (MRT), minimized average bounded slowdown or maximized throughput. Traditional performance modeling and evaluation studies of parallel space-sharing job scheduling algorithms are incompetent of predicting the combined or interaction effect on the response resulting due to simultaneous variation of two process variables. Present work is undertaken to predict and quantize the influence of main and interaction effects of the input scheduling process variables on the output MRT values using statistical approach of design of experiments (DOE). DOE based Response surface methodology (RSM) oriented experimental design is chosen to evaluate MRT values for two scheduling algorithms namely First Come First Serve (FCFS) and Fit Processors First Served (FPFS). Two empirical interaction models are suggested for both scheduling algorithms that predict MRT on the basis of multiple regression equations involving main and interaction effect terms of scheduling process variables. High value of adjusted coefficient of determination R^2 and insignificant lack of fit represent the goodness of fit of both the models to accurately predict the MRT values. Both the empirical interaction models are validated against additional experimental results. The comparative performance evaluation study on the basis of MRT reveals that the FPFS algorithm tends to outweigh the traditional FCFS policy.

General terms
PC-cluster, Space-sharing scheduling, Design of experiments, Response surface methodology, Mean response time and Empirical interaction model

Keywords
Statistical Modeling, First Come First Serve, Fit Processors First Served, DOE and RSM

1. INTRODUCTION

Incredible advances in the speed of microprocessors and networking technologies escort the way to the development of LAN based cluster of PCs[1] for high performance as well as high throughput computing activities. These cluster of PCs can be found in most of the educational institutions due to availability of requisite hardware (commodity desktop PCs and high speed local area network) and commonly available software (Windows Server 2003 and Windows XP). Cluster of PCs have a tendency to outperform large supercomputers in terms of extensibility capabilities and price/performance ratio. Job scheduling algorithms plays a great role in assigning the resources of PC-based cluster computing platform to the competing parallel jobs. These algorithms[2, 3] can be broadly classified into two categories; time-sharing and space-sharing. Time-sharing based scheduling algorithm shares the CPU time of PCs among multiple competing jobs. A space-sharing scheduling policy may allocate a distinct subset (partition) of processors (based on the job width of the job) of cluster's processor-pool to the selected job (job is selected on the basis of scheduling criteria). In this approach, no processor is concurrently assigned to more than one job. Parallel space-sharing job scheduling algorithm tends to play a double role; selecting a job from the set of competing jobs as well as allocating processors (out of available processors) to the job. Space-sharing algorithms [2, 3] are further categorized into two types; static and dynamic. In traditional static space-sharing, cluster is partitioned into equal sized partitions of processors and partition size is fixed for the whole life-span of the job. Contrarily in case of dynamic space-sharing, there can be a change in the subset size and the processors it contains. In program based machine partitioning technique, the partitions of processors are created for individual jobs based on their job sizes at the time of their servicing i.e. scheduling time.

FCFS and FPFS algorithms are mostly used for batch job scheduling[4] in space-shared clusters. In traditional FCFS algorithm [5, 6] jobs are considered for scheduling in the order of their arrival. The only job characteristic known to the scheduler when the job arrives, is the number of processors requested by the job i.e. job width or size. Rigid [7] class of data-parallel jobs is considered for scheduling in this research work. Such kind of parallel jobs will be selected for execution by the scheduler only if there are enough processors available to execute the job. In case the desired numbers of processors are not available, the first job and the other subsequent jobs in the job queue must wait for the availability of desired number of processors till they got freed from termination of some currently running jobs. This situation may lead towards an inefficient utilization of computing resources as processors sit idle waiting for their accumulation to fit to the job at the top of job queue. This also results into increase of mean response time of jobs due to increase in the wait times of individual jobs as the jobs are kept waiting in the queue for their turn to run. This shortcoming of FCFS led to the development of FPFS scheduling algorithm. In FPFS[6], if there are not enough processors available for the front job in the job queue, then job scheduler searches the job queue for the job which fits first to number of processors available and consequently that job is dispatched and processors are allocated to the job. This results into efficient utilization of processors as well as decrease in the mean response time of jobs. Traditional research on performance modeling and analysis studies[5, 6, 8] of job scheduling algorithms using experimental measurement, analytical/theoretical modeling and simulation is capable of showing the main effects corresponding to the variation of only one-factor-at-a-time (OFAT) on the observed output by keeping all other factors constant. These studies are not capable of predicting and quantizing the interaction i.e. combined effects on the output response resulting due to simultaneous variation of two process variables. An interaction between two input process variables occurs when effect of one variable on the observed output depends upon the level of another variable. The proposed work is helpful in investigating the relative importance of main as well as interaction effect of process variables with respect to the observed response with the help of statistical approach of design of experiments (DOE). This paper uses the DOE based approach of response surface methodology for performance modeling and analysis of job scheduling algorithms for PC-cluster computing environment with an emphasis on static space-sharing policy based on program specific partitioning. With the help of DOE based statistical techniques, empirical prediction models of performance metric MRT for both FCFS and FPFS polices are presented in terms of scheduling process variables.

Design of experiments (DOE) is a set of powerful and systematic statistical techniques[9,14] used for planning, designing and analyzing the experiments in a way to achieve authentic and objective conclusions effectively and proficiently. Response surface methodology (RSM) is a meta-modeling approach[10] of DOE aimed to be used in modeling, establishing and analyzing the relationships existing between process variables and the observed response using polynomial mathematical equations. RSM based experimental designs tend to minimize the number of experiments required for performance modeling and analysis of the observed response.

The rest of the paper is organized as follows: Section 2 deals with the development of user friendly GUI based PC-cluster computing environment and resource management system (RMS) for job scheduling activities. This section also deals with discussion on experimental procedure for implementation of FCFS and FPFS policies in the PC-cluster. Scheduling process experimental design based on the RSM oriented D-optimal coordinated exchange and mathematical performance models of FCFS and FPFS are also presented in this section. In section 3, results of the scheduling experiments are presented and analyzed. Finally comparative performance evaluation study of FCFS and FPFS is discussed.

2. MATERIALS AND METHODS

PC-cluster[1] is a group of interconnected stand-alone PCs working jointly as a single integrated computing resource with the help of single system image (SSI) functionality residing at cluster middleware abstraction layer. The SSI [11] represents the view of cluster's parallel and distributed system as a single unified computing resource to the user. It hides the hardware and software complexities of the PC-cluster's parallel and distributed computing environment from the user hence leading towards a convenient single unified environment to work with. The SSI of the cluster is realized with the help of a cluster distributed RMS. RMS[1, 11] is developed with an aim to manage cluster functionalities related to job scheduling such as job submission, job scheduling, processor allocation, job execution and other resource management activities. The generic architecture of the cluster distributed RMS system is shown in figure1.

2.1 Experimental set-up and procedure

PC-cluster is constituted by connecting 25 networked computers available in the three departmental computer laboratories with the help of layer 3 Virtual Local Area Network (VLAN) based switch (make Cisco 3750 series). PCs are connected within a computer lab with the help of layer 2 edge switch (make Cisco 2650 series). One of the nodes in the VLAN is designated as master node (Pentium Core 2 duo with 1 GB RAM, running on Windows Server 2003 Enterprises Edition) and others 24 PCs are accredited as slave or compute nodes (configured with Windows XP based Pentium IV, 3.0 GHz and 512 MB RAM). Number of rigid non-interactive

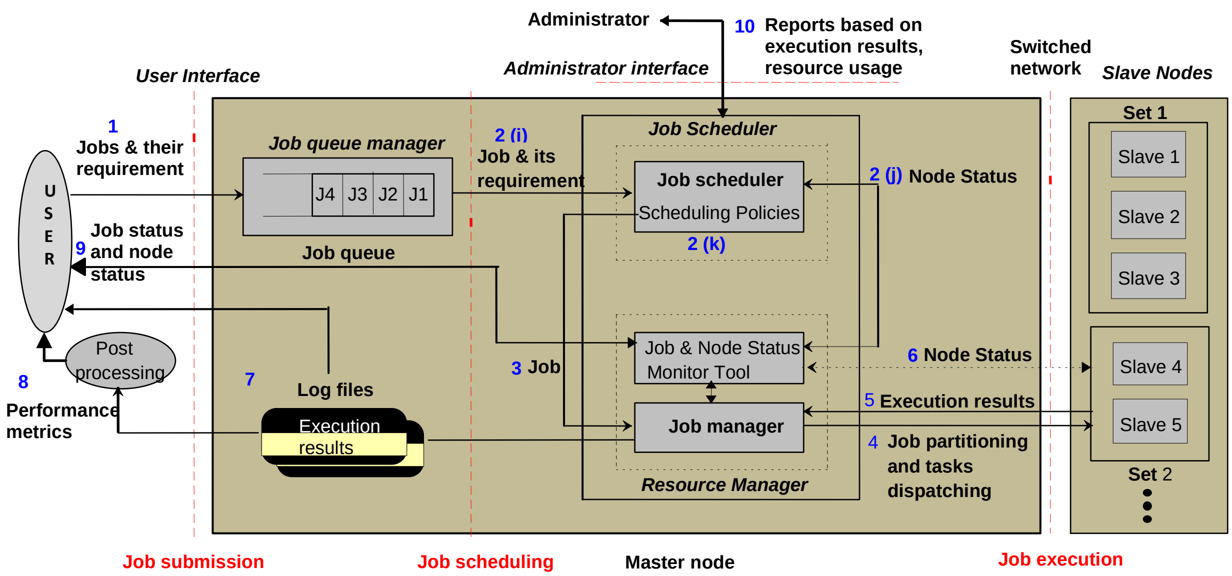


Fig. 1: Distributed Resource Management System and scheduling procedure

Table 1. Independent process variables and their range

Process variables(Factors)	Symbols	Category	Type	Levels	Levels (actual values)
ScheduleSize	SS	Quantitative	Discrete	4	66,100,132,168
ClusterSize	CS	Quantitative	Continuous	9	16 -24

Table 2. RSM based experimental design for FCFS and FPFS with experimental and model predictive response

Process variables for FCFS and FPFS			Response for FCFS		Response for FPFS	
Actual values (coded values)			Mean Response Time(sec.)		Mean Response Time(sec.)	
Exp. No.	SS ScheduleSize	CS ClusterSize	Exp. values	Model predicted	Exp. values	Model predicted
1.	100 (-0.333)	16 (-1.000)	27.39	28.53	15.63	15.41
2.	100 (-0.333)	20 (0.000)	24.56	24.69	13.75	13.29
3.	168 (1.000)	16 (-1.000)	42.86	42.49	23.43	23.30
4.	168 (1.000)	16 (-1.000)	42.24	42.49	23.82	23.30
5.	132 (0.294)	17 (-0.650)	33.83	33.32	17.95	18.21
6.	66 (-1.000)	18 (-0.460)	20.23	20.18	10.66	10.61
7.	66 (-1.000)	18 (-0.460)	20.85	20.18	10.72	10.61
8.	168 (1.000)	19 (-0.220)	37.59	37.45	20.27	20.81
9.	132 (0.294)	20 (0.100)	29.46	29.51	15.28	16.24
10.	66 (-1.000)	21 (0.280)	18.85	18.31	9.04	9.44
11.	100 (-0.333)	22 (0.540)	22.34	22.61	11.91	12.15
12.	132 (0.294)	24 (0.900)	24.31	25.45	13.39	14.14
13.	66 (-1.000)	24 (1.000)	16.87	16.49	8.59	8.31
14.	66 (-1.000)	24 (1.000)	16.12	16.49	8.67	8.31
15.	168 (1.000)	24 (1.000)	30.05	29.55	17.02	16.92
16.	168 (1.000)	24 (1.000)	29.75	29.55	17.85	16.92

data-parallel jobs viz. matrix-matrix multiplication, matrix-vector multiplication, calculation of pi value, run-length image compression and finding prime numbers in a list with varying input sizes has been developed in accordance with power-of-two workload model(more details shown in Appendix A). The set of jobs and their job size will be acting as a workload to be submitted to the job scheduler for scheduling. In power-of-two workload model, the entire job sizes are of the type 2^n where n is a user specific integer within the range of [1, 4] and size of cluster falls in integer continuous range of [16, 24]. Based on the job size characteristics, rigid parallel jobs are classified as small (number of processors required by job varies from 1-4) and large (number of processors required varies from 5-16). Workload submitted by the user to the job queue at time zero for scheduling consists of roughly 50% small and 50% large jobs. Master node with the support of cluster RMS system helps the user to submit, schedule and execute jobs. These scheduling and other resource management activities are performed with the help of key components of RMS; user interface & queue manager, job scheduler and resource manager. Slave nodes are only responsible for execution of the partitioned tasks of jobs dispatched by the job manger of the master node as well as communicating the task execution results back to the master node.

The overall procedure for three major job scheduling activities viz. job submission, job scheduling and job execution is shown in figure 1 using labeled numbers from step 1 to 10. At step 1, user submits the jobs along with their job sizes to job queue manager with the help of user interface to the RMS at the master node. Based on the triplicate information obtained from step 2(i) (job size details), node availability information obtained from job & node status monitoring tool of the resource manager in step 2(j) and scheduling policy at step 2(k), a scheduling decision to select a job is taken by the scheduler and set of slave nodes are selected for allocation to job. The selected job is dispatched by the job manager in step 3. In step 4, job manager partitions the job into parallel tasks based on the number of slave nodes allotted to the whole job and dispatches these partitioned parallel tasks to the allocated slave nodes for execution. Task execution results are sent back to the job manger module after the tasks are executed by the slave nodes in step 5. Job manager is also responsible for merging of the partial results collected from various slaves to form the final result. Final result and various real-time parameters related to job submission times, job completion times and job waiting times are stored in the text based log files. Job and node status is updated at step 5 and step 6 with the help of resource manager. This scheduling procedure from step 1 to 6 continues till the job queue is empty. In step 7 users can access the log files at master node console with the help of cluster user interface. Results in terms of performance metric (MRT) also known as average turnaround time can be

Table 3. ANNOVA table and model diagnostics for FCFS & FPFS

	Mean Response Time(seconds) Interaction model for FCFS						Mean Response Time(seconds) Interaction model for FPFS			
Source FCFS model	Sum of squares	df	Mean square	F-value	p-value* (Prob. > F)	Source FPFS model	Sum of squares	df	Mean square	F-value	p-value* (Prob. > F)
Model terms	1065.41	3	355.14	952.42	< 0.0001	*Model terms*	361.58	3	120.53	389.78	< 0.0001
ScheduleSize [m]	731.64	1	731.64	1962.15	< 0.0001	ScheduleSize [m]	264.72	1	264.72	856.10	< 0.0001
ClusterSize [m]	170.16	1	170.16	456.36	< 0.0001	ClusterSize [m]	47.77	1	47.77	154.50	< 0.0001
ScheduleSize x ClusterSize [i]	24.33	1	24.33	65.24	< 0.0001	ScheduleSize x ClusterSize [i]	4.10	1	4.10	13.25	0.0034
Residual	4.47	12	0.37			Residual	3.71	12	0.31		
Lack of fit	3.76	8	0.47	2.65	0.1813##	Lack of fit	3.29	8	0.41	3.86	0.1039##
Pure error	0.71	4	0.18			Pure error	0.43	4	0.11		
Corr. Total	1069.88	15				Corr. Total	365.29	15			
Model statistics for FCFS:						*Model statistics for FPFS:*					
S.D: 0.611	C.V. %: 2.234		R^2: 0.996			S.D: 0.556	C.V. %: 3.739		R^2: 0.990		
Adjusted R^2: 0.995	Predicted R^2: 0.992					Adjusted R^2 :0.987	Predicted R^2: 0.982				

*significant at p≤0.05 ## not significant at p≤0.05 [m] main effect [i] interaction effect

obtained by doing standalone post-processing exercise on the data collected from log files as per (1) at step 8. Job and node status can be collected from resource manager module by the user as well as the administrator of the cluster at step 9 and 10 respectively. Mean response time (MRT) is chosen as performance metric to analyze the performance of scheduling algorithms viz. FCFS and FPFS. MRT is calculated using (1).

$$\text{MRT} = \frac{1}{N}\sum_{i=1}^{N}\left(\text{Job\_EndTime(i)} - \text{Job\_SubmitTime(i)}\right) \quad (1)$$

where N is the number of jobs with known job width characteristics, Job_SubmitTime(i) indicates the time when i^{th} job is submitted to the job queuing system and Job_EndTime(i) denotes the time when i^{th} job gets terminated. MRT being user specific metric, indicates an average completion time of the submitted job using a certain scheduling algorithm.

2.2 Experimental design and scheduling performance modeling

The workload to the scheduling system consists of information about number of space-sharing rigid data-parallel jobs to be scheduled along with their job size characteristics. The input parameter for the scheduling system is the sum of job sizes of the total number of jobs in the workload and is known as schedule size (denoted as ScheduleSize(SS)). Another input variable chosen is the number of processors in the cluster known as cluster size (denoted as ClusterSize(CS)). The chosen independent process variables or parameters and observed output (known as factors and response respectively in terms of DOE terminology) along with their levels (variations) for modeling of observed response MRT values are shown in table 1. Based on RSM D-optimal coordinate exchange design, 16 experimental runs (table 2) in random order were conducted with various combinations of SS and CS for FCFS policy. Experiments for FPFS (table 2) were also carried out with the same design and combinations of SS and CS as is done in the case of FCFS. This RSM based experimental design for both scheduling algorithms helps to minimize the number of experiments required to model their performance. Number of experiments required for modeling purpose using RSM design are 32(16 for each scheduling policy) as compared to 72(4x9=36 for each scheduling policy) in case of OFAT approach[12].

Some of the experiments in both FCFS and FPFS were the replicated to check the variation in the computer based physical experimentation process due to uncontrolled experimental factors like variation in network load on the interconnection switches. This variation became the source for calculating the term mean square pure error. Experimental data of MRT values (table 2) of both the scheduling policies were fitted against the two independent interaction models (one for FCFS and the other for FPFS) with the presupposition that during the process of scheduling, interaction between any two process variables might occur.

2.1.1 ANNOVA analysis

ANNOVA results of the interaction models of MRT (for FCFS and FPFS) are helpful in determining the significance of models as well as their model terms. Insignificant terms in the models with p-value greater than 0.05 can be omitted to improve the models. Interaction model fitting, ANNOVA statistical analyses, coefficient estimation and visual result analyses using model diagnostic and other plots were carried out with the help of Design-Expert 8.0 software (StatEase Inc. USA)[13].

Goodness of fit of the each interaction model was observed [9,13] using high values of coefficient of determination R^2, adjusted R^2, predictive R^2 and low value of coefficient of variation(CV%) and insignificant lack of fit. Lack of fit compares the residual error with the pure error obtained from replicated model points and it is not desirable. Significant lack of fit[13] implies that the variation of the replicates about their mean values is less than the variation of the design points about their predicted values. Signal to noise ratio was observed from adequate precision value with ratio > 4 desirable for the model to navigate the design space.

2.1.2 Model adequacy checking

In each of the interaction model, model adequacy checking of the residuals was performed using various diagnostic plots [9,13]. Normal probability plot of studentized residuals was checked to see the normality of residuals. Plot of studentized residuals versus predicted values were studied to check the constant error. Plot of externally studentized residuals was checked to see the presence of outliers i.e. influential values. Box-Cox plot was investigated to look for power transformations suggestions to improve the model. Power transformations were required in those cases when the max to min ratio of response is greater than 10 and/or presence of non-normality in the residual data.

2.1.3 Model fitting

Response MRT can be related to independent scheduling process variables using mathematical interaction model equation. Empirical interaction models of MRT for both scheduling policies were described both in terms of coded factors and the actual factors with the help of least squares multiple regression equation given in (2).

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{1 \le i \le j}^{k} \beta_{ij} x_i x_j + \varepsilon \qquad (2)$$

where y is the predicted response, x_i and x_j are independent variables or factors, k is the number of independent factors. β_o, β_i and β_{ij} are the regression coefficients of intercept, first-order and interaction term respectively and ε is statistical random error.

The coded equation [13] is useful for understanding the relationship between independent input variables and the output response. It also helps in identifying the relative significance of the model factors in terms of their absolute effect on the model response by comparing the factor coefficients. This coefficient comparison cannot be made with the actual equation because the coefficients are scaled to accommodate the units of each factor. In coded equation, every factor is uniformly scaled between -1 and +1; hence it provides the unitless regression coefficients to estimate the relative importance of the model factors. Finally predicted values of both the interaction models (for FCFS and FPFS) are validated against the additional actual experimentation results.

3. RESULTS AND DISCUSSION

Actual MRT values of obtained from the experimentation process for both FCFS and FPFS are shown in table 2. Interaction models to predict the main effects and interactions between two independent factors were fitted against the individual experimental data of FCFS and FPFS algorithms respectively. Statistical ANNOVA analysis (table 3) of the interaction models for MRT in both FCFS and FPFS showed that the model and all the model terms (main effect and interaction) are significant at $p \le 0.05$. Respective MRT model F-value of 952.42 and 389.78 indicates that interaction models are significant for both FCFS and FPFS policies and there is only 0.01% chance that such a high model F-value crop up due to the noise. Goodness of the fit of the interaction model for FCFS was determined by model statistics like high values of coefficient of determination R^2=0.996, adjusted R^2 =0.995, predicted R^2 =0.992 and low values of SD=0.611 and CV%=2.234(exceptionally good below 5).
Similar model statistics of coefficient of determination R^2=0.976, adjusted R^2 =0.970, predicted R^2 =0.955 and low values of SD=0.556 and CV %=3.739 conclude the goodness of fit for the interaction model of FPFS. Predicted R^2 and adjusted R^2 in both scheduling policies are in reasonable agreement (within 0.2) with each other. Adjusted R^2 values of 0.995 and 0.970 in case of FCFS and FPFS represent that respective models are true to explain 99.5% and 97% variation in the model process variables. Low F-value 2.65 and p-value 0.1813 entail that lack of fit is not significant relative to pure error. Similarly for FPFS policy, low F-value 3.86 and p-value 0.1039 conclude that lack of fit of the model is not significant. Adequate precision values 85.156 and 53.931 of both interaction models for FCFS and FPFS indicate an adequate signal to noise ratio to navigate the design space by the interaction model. Model diagnostics plots for both the models reveal that all the residuals of MRT are normally distributed as they fall on the linear line on the normal probability plot paper shown in figure 2(a) and 2(b). Predicted vs. actual fitted values are represented by a linear line passing from origin indicating the closeness of predicted and actual MRT values shown in figure 3(a) and 3(b). Box-Cox transformations were not required on the data in both the scheduling models due to the fact that max to min ratio of MRT is less than 10 and presence of normality of data in both cases as shown in figure 2(a) and 2(b). Predicted vs. actual fitted values are represented by a linear line passing from origin indicating the closeness of predicted and actual MRT values shown in figure 3(a) and 3(b). Box-Cox transformations were not required on the data in both the scheduling models due to the fact that max to min ratio of MRT is less than 10 and presence of normality of data in both cases as shown in figure 2(a) and 2(b).

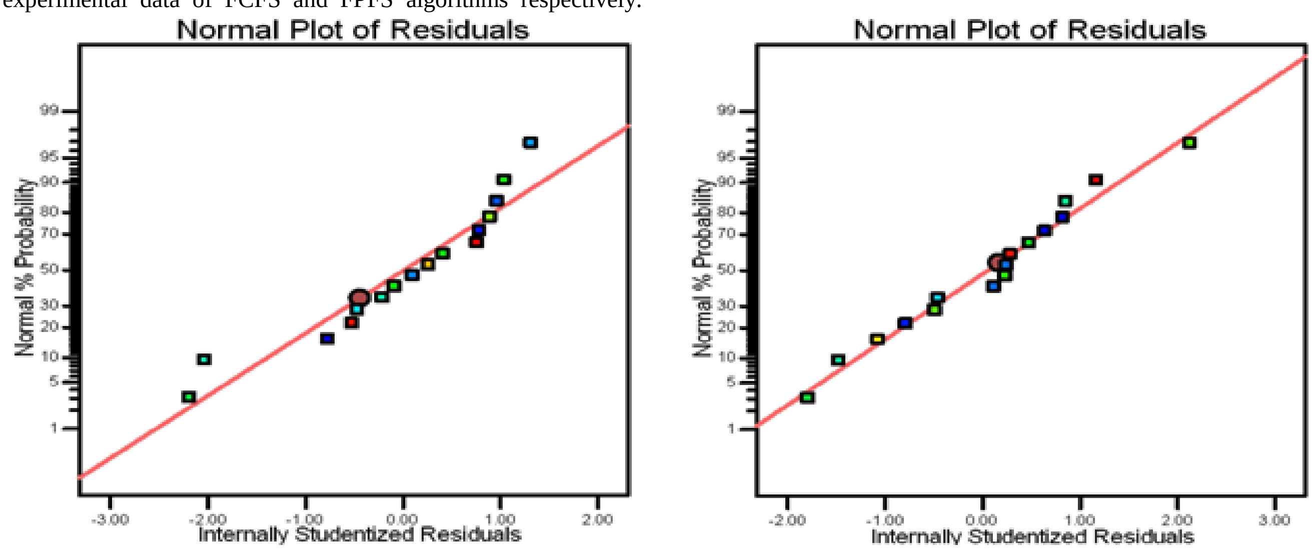


Fig. 2(a) and 2(b): Normal probability plot of MRT for FCFS (left) and FPFS (right)

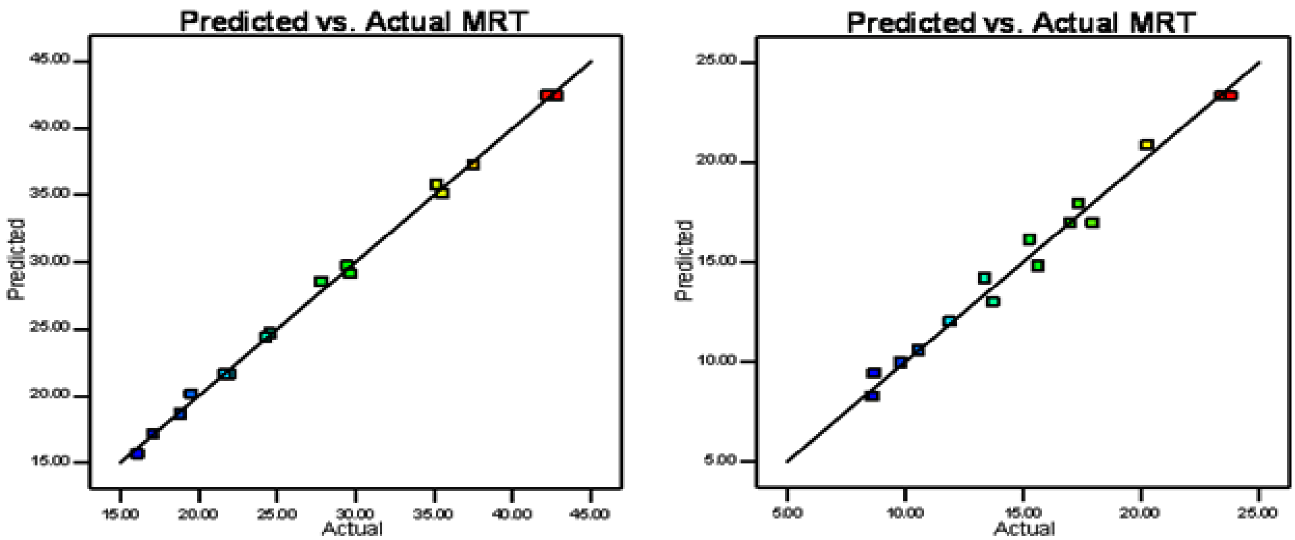


Fig. 3(a) and 3(b): Predicted vs. Actual MRT plot for FCFS (left) and FPFS (right)

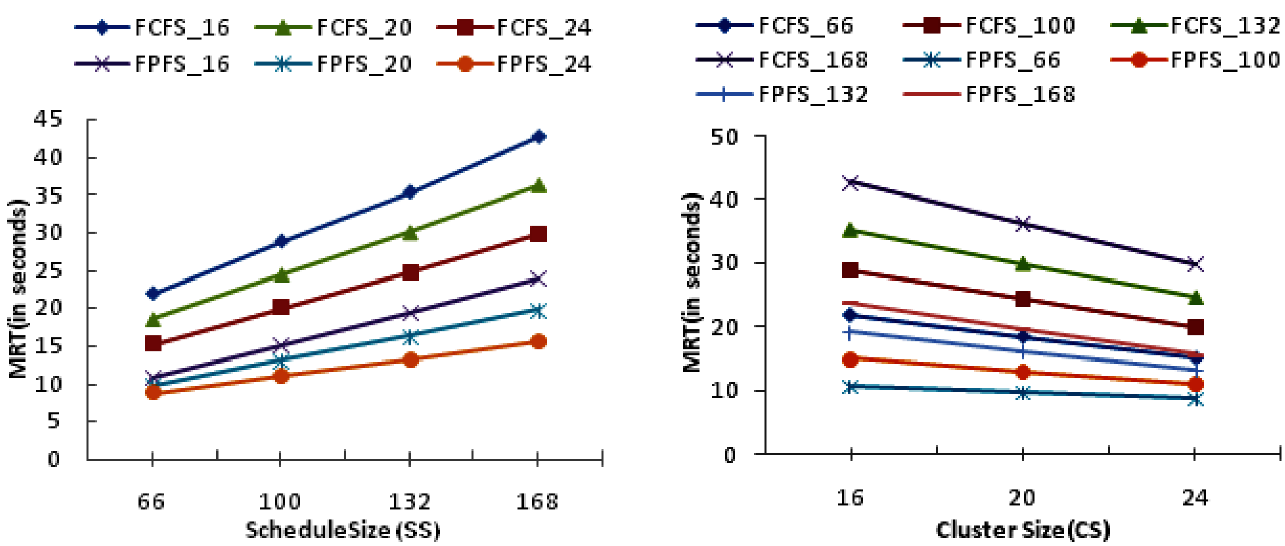


Fig. 4(a) and 4(b): Main effect (one variable) plots of SS vs. MRT and CS vs. MRT for FCFS and FPFS policy respectively

Box-Cox transformations were not required on the data in both the scheduling models due to the fact that max to min ratio of MRT is less than 10 and presence of normality of data in both cases as shown in figure 2(a) and 2(b).

Empirical interaction model equations of the MRT (in terms of coded values of process variables) for FCFS and FPFS policy are derived with the help of method of least squares based multiple regression models and are given in (3) and (4) respectively.

MRT for FCFS = 27.52 + 8.50 ScheduleSize – 4.50 ClusterSize – 1.97 ScheduleSize x ClusterSize (3)

MRT for FPFS = 15.00 + 5.11 ScheduleSize – 2.38 ClusterSize – 0.81 ScheduleSize x ClusterSize (4)

where ScheduleSize and ClusterSize represent the coded values of input variables schedule size and cluster size respectively. Equations (3) and (4) have terms ScheduleSize and ClusterSize that indicates the main effect of these process variables on the response MRT. Term ScheduleSize x ClusterSize indicates an interaction effect between schedule size and the cluster size. Positive and relatively higher regression coefficients(8.50 and 5.11) of term ScheduleSize in (3) and (4) respectively for FCFS and FPFS indicates that the process variable schedule size have higher relative impact on MRT values as compared to any other term in the equations. As the job sizes of the jobs contained in the workload increases, the MRT values are bound to increase in both policies though higher in case of FCFS as shown in main effect plot of ScheduleSize on the observed response MRT at the fixed levels of ClusterSize (16, 20 and 24) in figure 4(a). Regression coefficients of term ClusterSize with negative values of 4.50 and 2.38 indicates this variable have antagonistic effect on the MRT value and result into decrease in MRT value with the increase of processors in the cluster. Figure 4(b) of main effect of CS on the MRT values at fixed values of SS (66,100,132 and 168) witnesses that as the CS increases there is a sharp decrease in the MRT values in case of FCFS policy as compared to FPFS due to higher negative coefficient estimate of 4.50 of CS term in the equation of FCFS MRT model.

Plots for interaction effects between ScheduleSize and ClusterSize on the response MRT for both policies are shown in figure 5(a) and 5(b) respectively. The interaction plots in both policies show that effect of ScheduleSize on response MRT is different for different levels of ClusterSize. Non-parallel lines of variable ClusterSize at different levels of ScheduleSize in both the interaction plots also validate the presence of interaction effect of ScheduleSize and ClusterSize on the response.

Table 4. Validation experiments for FCFS and FPFS

			FCFS		FPFS	
Exp. No.	SS	CS	Predicted MRT	Exp. MRT	Predicted MRT	Exp. MRT
1.	66	16	22.02	21.96	11.46	10.95
2.	100	24	22.11	22.33	11.18	10.88
3.	168	20	37.99	36.87	20.11	19.81

Approximative mathematical interaction model equations of the MRT (in terms of actual factors) for FCFS and FPFS using multiple regression models are given in (5) and (6).

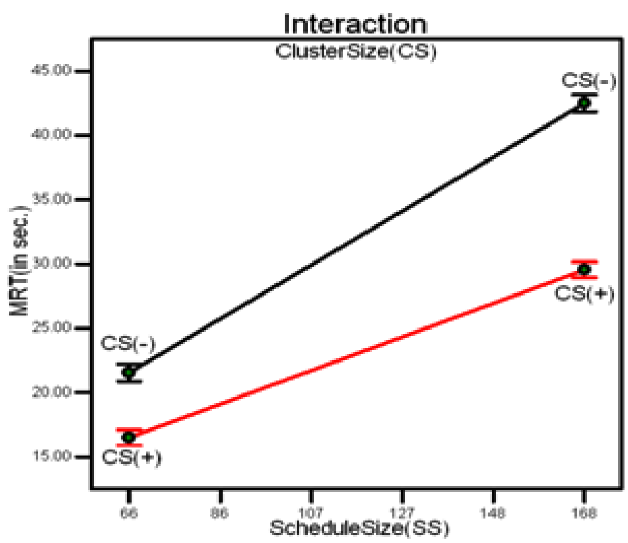
$$\text{MRT for FCFS} = 7.86721 + 0.36018\ SS - 0.00737\ CS - 0.0090\ SS \times CS \tag{5}$$

$$\text{MRT for FPFS} = 5.89207 + 0.17967\ SS - 0.13135\ CS - 0.00397\ SS \times CS \tag{6}$$

where SS and CS represent the actual values of the input variables schedule size and cluster size. Actual MRT values for FCFS and FPFS can be obtained from (5) and (6) respectively by fitting the actual values at all the levels of SS and CS. Individual interaction MRT model for both FCFS and FPFS is validated against the additional actual experimentation results shown in table 4.

4. CONCLUSIONS

Response surface methodology approach of DOE has been used for statistical performance modeling and analysis of program-based static space-sharing scheduling algorithms in PC-cluster computing environment. The mathematical interaction models for both FCFS and FPFS policies, expressed in terms of main and interaction effect terms of scheduling process variables viz. ScheduleSize and ClusterSize have been found to be remarkably statistically fit for predicting the process response MRT. Model term ScheduleSize have higher relative impact on the MRT values than any other term in both of the models. Goodness of the fit of the both interaction models was observed with the help of high values of adjusted R^2 and insignificant lack of fit. Respective empirical models of MRT for FCFS and FPFS are validated against additional actual experimental results. Performance analysis study showed that FCFS produced higher values of MRT at all the levels of SS and CS as compared to FPFS.



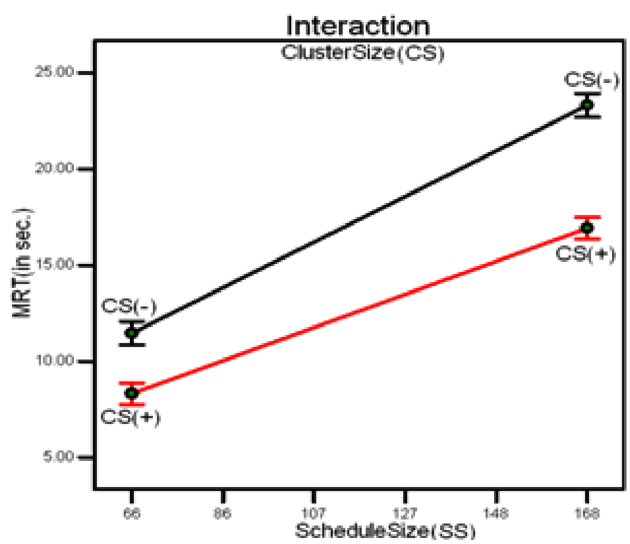


Fig. 5(a) and 5(b): Interaction effect plots of SS x CS for FCFS (left) and FPFS (right)

APPENDIX A: WORKLOAD INFORMATION

Workload	Workload description: with job arrival order Workload format (Job no - Job name - job width - problem size)	Input parameter ScheduleSize= $\sum_{i=1}^{N}$ JobSize(i)
Workload 1 No. of jobs =10	J1-Runlength Image Compression-2-303x239, J2-Matrix Vector Product.-8-3000x1, J3-Matrix Multiplication-4-256x256,J4-Matrix Multiplication-16-800x800,J5-Calculation of PI-2-100000,J6-Finding Total Prime No-4-10000,J7-Matrix Multiplication-16-800x800,J8-Matrix Multiplication-2-128x128, J9-Matrix Vector Product-8-2400x1,J10-Runlength Image Compression-4-303x239	66
Workload 2 No. of jobs =15	J1-Runlength Image Compression-2-303x239, J2-Matrix Vector Product-8-3000x1, J3-Matrix Multiplication-4-256x256,J4-Matrix Multiplication-16-800x800,J5-Calculation of PI-2-100000,J6-Finding Total Prime No-4-10000,J7-Matrix Multiplication-16-800x800,J8-Matrix Multiplication-2-128x128, J9-Matrix Vector Product-8-2400x1,J10-Runlength Image Compression-4-303x239,J11-Matrix Multiplication-4-256x256,J12-Matrix Vector Product-8-2000x1,J13-Finding Total Prime No-2-10000,J14-Matrix Multiplication-16-512x512,J15-Runlength Image Compression-4-303x239	100
Workload 3 No. of jobs =20	J1-Runlength Image Compression-2-303x239, J2-Matrix Vector Product-8-3000x1, J3-Matrix Multiplication-4-256x256,J4-Matrix Multiplication-16-800x800,J5-Calculation of PI-2-100000,J6-Finding Total Prime No-4-10000,J7-Matrix Multiplication-16-800x800,J8-Matrix Multiplication-2-128x128, J9-Matrix Vector Product-8-2400x1,J10-Runlength Image Compression-4-303x239, J11-Matrix Multiplication-4-256x256,J12-Matrix Vector Product-8-2000x1,J13-Finding Total Prime No-2-10000,J14-Matrix Multiplication-16-512x512,J15-Runlength Image Compression-4-303x239,J16-Calculation of PI-2-100000,J17-Matrix Multiplication-16-800,J18-Matrix Multiplication-4-256x256,J19-Matrix Vector Product-8-3000x1,J20-Runlength Image Compression-2-303x239	132
Workload 4 No. of jobs =25	J1-Runlength Image Compression-2-303x239, J2-Matrix Vector Product-8-3000x1, J3-Matrix Multiplication-4-256x256,J4-Matrix Multiplication-16-800x800,J5-Calculation of PI-2-100000,J6-Finding Total Prime No-4-10000,J7-Matrix Multiplication-16-800x800,J8-Matrix Multiplication-2-128x128, J9-Matrix Vector Product-8-2400x1,J10-Runlength Image Compression-4-303x239,J11-Matrix Multiplication-4-256x256,J12-Matrix Vector Product-8-2000x1,J13-Finding Total Prime No-2-10000,J14-Matrix Multiplication-16-512x512, J15-Runlength Image Compression-4-303x 239,J16-Calculation of PI-2-100000,J17-Matrix Multiplication-16-800x800,J18-Matrix Multiplication-4-256x256,J19-Matrix Vector Product-8-3000x1,J20-Runlength Image Compression-2-303x239,J21-Calculation of PI-4-1000000,J22-Matrix Multiplication-6-512x512,J23-Matrix Multiplication-8-512x512,J24-Matrix Multiplication-16-1000x1000,J25-Matrix Vector Product-2-1000x1	168

5. REFERENCES

[1] Yeo, C.S., Buyya, R. and Pourreza, H.2006 "Cluster Computing: High-Performance, High-Availability and High-Throughput Processing on a Network of Computers",*vol.29 (6),Springer Science + Business Media*Inc,New York,USA,pp.521-551.

[2] Ismail, I.M. 1995, "Space-sharing job scheduling policies for parallel computers", Ph.D thesis, Iowa state university, Iowa.

[3] Figueira, S.M. 2004, "Optimal partitioning of nodes to space-sharing parallel tasks", *Parallel Computing*, 32,pp 313-324.

[4] Iqbal,S., Gupta,R. and Fang,Y.C. 2005 "Planning considerations for job scheduling in HPC clusters", reprinted from Dell Power Solutions, pp 133-136.

[5] Schweigelshohn, U. and Yahyapour, R. 1998 "Analysis of first-come-first-serve parallel job scheduling", *In Proceedings of the ninth annual ACMSIAM symposium on discrete algorithms*, pages 629–638, Philadelphia, PA, Society for Industrial and Applied Mathematics.

[6] Aida, k. 2000, "Effect of job size characteristics on job scheduling performance" In Job Scheduling Strategies for Parallel Processing, Springer Verlag, Lect. Notes Computer Science vol. 1911, pp. 1—17.

[7] Lublin, U and Feitelson, D.G. 2003, "The Workload on Parallel Supercomputers: Modeling the Characteristics of Rigid Jobs", *Journal of Parallel and Distributed Computing*, 63(11):1105–1122.

[8] Aida,K., Kasahara, H. and Narita, S.1998. "Job Scheduling Scheme for Pure Space Sharing Among Rigid Jobs", *In fourth workshop on Job Scheduling Strategies for Parallel Processing, Lecture Notes in Computer Science*, vol. 1459, pages 98-121, Springer-verlag.

[9] Montgomery, D.C. 2009 Design and analysis of experiments (5th ed.). New York: Wiley & Sons, Inc. 672 pp.

[10] Myers, R.H., Montgomery, D.C. and Anderson-Cook, C. M., 2009. Response Surface Methodology: Process and product optimization using designed experiments (3rd ed.).New York: John Wiley and Sons, Inc. 728 pp.

[11] Buyya, R., Cortes, T. and Jin, H. 2001. "Single System Image (SSI)," *International Journal of High Performance Computing Applications*, vol. 15, no. 2, pp. 124-135.

[12] Anderson, M.J. and Whitcomb, P.J. 2000, *DOE Simplified: Practical Tools for Effective Experimentation*, Productivity press.

[13] Design Expert Software version 8.0 user's guide 2009.

[14] Antony, J. 2003. Design of experiments for engineers and scientists. Elsevier Science & Technology Books.