International Journal of Grid and Distributed Computing
Vol.7, No.2 (2014), pp.53-68
http://dx.doi.org/10.14257/ ijgdc.2014.7.2.06

Correlation Based Combined Analysis Of FTA And GWA For Reliability And Performance Study Of Grid Environment

R. K. Bawa[1] and Ramandeep Singh[2]

[1]Department of Computer Science, Punjabi University, Patiala, Punjab, India
[2]Department of Computer Science, Lovely Professional University, Jalandhar, India
[1]rajesh_k_bawa@yahoo.com, [2]ramankhosa@gmail.com

Abstract

In this age of data and knowledge, Cloud, Grid and P2P systems are becoming common and advanced. Due to heterogeneous and distributed nature, Grid becomes more vulnerable to faults. Trace files are great way of storing and collecting fault and workload information from the system. FTA (Fault Trace Archive) and GWA (Grid Workload Archive) are two such trace files. Previously FTA and GWA have been individually analyzed by researchers, but in this research paper for the first time, we have analyzed the combination of FTA and GWA as a single research problem. Trace files have been joined based on the event timestamp values. Both the trace files have been analyzed to establish a correlation based model among node failures, failed jobs, number of nodes and failure duration. We have discovered that these factors are positively correlated with each other but to a different extent. Along with node failure frequency, failure resume time and node dedication factor, we have found that interactive jobs have a higher failure probability as compared to batch jobs.

Keywords: *FTA Fault Trace Archive, GWA Grid Workload Archive, Fault Tolerance, QoS Quality of Service*

1. Introduction

Grid computing which simply originated as a small idea of distributed computing has taken many remarkable turns in the form of data grids, computational grids and most importantly cloud computing [1, 2]. Grid is a set of heterogeneous resources that are connected via networks and used in different areas of science, research and industry. As any other distributed environment Grid is also prone to faults but its heterogeneous nature makes it more vulnerable. These faults can be in the form of hardware, software, network or exceptions, which leads to job failure or cancellation. Due to these faults, system performance degrades and it may not be able to meet the user quality of service requirements [3, 4].

Performance of Grid depends on resource management, job scheduling, monitoring and load balancing strategies. But careful planning and implementation of these strategies require an inside knowledge of what actually happens in the Grid environment. In this paper, we have used the combination of FTA (Fault Trace Archive) [5] and GWA [6] (Grid Workload Archive) to analyze resource availability, job submission and execution patterns. These trace files are collected from Grid5000 a Grid computing platform consisting of 17 laboratories in France. FTA is a trace file that stores information about node failures and GWA provides information about the workload that is submitted for execution on the Grid. Trace files are analyzed by researchers and developers to establish a model for system behavior according to the resource availability and job submission trends. Resource availability is critical for the reliability and responsiveness of Grid

ISSN: 2005-4262 IJGDC

services[7]. Available groups of resources are often required to execute tightly coupled distributed and parallel algorithm based workloads.

In the next section, we have discussed work related to trace file analysis. Section 3 contains FTA analysis and Section 4 contains GWA analysis. Combined analysis of FTA and GWA has been conducted in Section 5. Section 6 contains conclusions and future work have been discussed in Section 7.

2. Related Work

Trace files are used to collect information about events taking place inside a system either centralized or distributed. Trace file with different format serves different purposes due to the type of information included in a particular trace. FTA was proposed by Dick Epma, *et al.*, [8] as a standard format to collect data about fault events in distributed environment. FTA contains information about the events in the form of event type, event id, node or host on which this event took place, event start time, event end time, event end reason, platform info of the host *etc.*, Bahman Javadi, *et al.*, [9] have collected and analyzed SETI @home trace files for discovering subset of hosts, which share similar kind of statistical availability patterns. Out of 230000 hosts, it was discovered that availability of around 34% of hosts is a truly random process while rest of these hosts can often be modeled in the form of different groups, with few distinctions from one another. This distribution can be considered while designing job scheduling policies to meet user requirements.

Artur Andrzejak, *et al.*, [10] have also worked on SETI@home host trace files which consisted of 48000 hosts to analyze host availability patterns. They proposed a model which ensures that a certain number of hosts will be available for a certain amount of time, either by replication or by over provisioning of resources. So in this way trace files can be a very good way of knowing about the inside events of the system.

Similarly GWA (Grid Workload Archive) provides the information about the events related to the workload *i.e.*, jobs that are being submitted for execution. GWA contains data about the various factors related to the job in detail, a few of these are job type, job submission time, job runtime requirements, number of processors required, waiting time, job memory requirements, job queue and job status. Job status shows that whether the specific job was completed, failed or cancelled. It may also contain any system specific reason of job status based on the trace file attributes. Not only researchers but also the administrators as well as developers can also learn a lot from these trace files for making changes in the architecture and policies to minimize the fault recovery time and also reducing the fault occurrence probability [11, 12].

Careful analysis of GWA [6]can reveal that how a particular type of job is treated by Grid environment. Mainly a compute Grid environment is either application specific or job specific. Application specific architecture is designed to support the execution of a specific application *e.g.* Video encoding. On the other hand, a job based environment can execute more than one type of job simultaneously. Jobs are mainly categorized in two categories *i.e.* Batch jobs and Interactive jobs [6]. Researcher or system administrator can know from the analysis that how the job success or failure rate is affected by the job type and other factors like execution time, waiting time *etc*. Trace files also include the information about the users who have submitted the job for execution. Although grid user populations range from tens to hundreds of individuals, but only a few users dominate each grid’s workload both in terms of consumed resources and the number of jobs submitted to the system [13].

Other than jobs and resources, trace files can be generated and used for analyzing the performance of individual components of a host. Bianca [14] have used trace file to analyze and predict the average life span of hard drives being used at a high performance cluster. Data sheets hard drives show that MTTF (Mean Time to Failure) is around

1,000,000 to 15,000,000 hours, suggesting an annual failure rate around 0.88%. But analysis of field data shows that the minimum disk replacement rate is 1% which may go up to 3-4% and in some cases it can reach upto 13%. Analysis show that actual data gathered from the field may differ from the conceptual or ideal data. Main motive behind this research is to analyze Grid workload in relation to host failures. FTA and GWA from Grid5000 which are collected over the period of one and half year have been used in this analysis. Earlier researchers have analyzed FTA and GWA individually, but till now no one has combined both as a single research problem. Before performing combined analysis of FTA and GWA, we have analyzed each trace file separately.

3. Fault Trace Archive Analysis

FTA (Fault Trace Archive) was introduced by[8] to standardize the format for collecting fault traces in distributed environments. These traces have been pre-processed for removal of any noise in the data and which can lead to prediction or analysis errors. In the trace file, nodes have been assigned node ids that range from 0 to 1287 as there are a total of 1288 nodes. For better understanding of cluster behavior we have mapped these node ids to the corresponding sites and clusters. This will also help in establishing and understanding the pattern from the inter cluster and intra cluster point of view. As we can see in Table1 that ids assigned to the nodes are not in a sequence but still there are five sites (colored) where nodes have been assigned ids sequentially and we can study graphs for these id ranges to understand the behavior of nodes, which belong to the same site or cluster.

3.1. Node Unavailability Time

In this section, we will analyze the trace events related to availability and unavailability of the Grid nodes. In FTA, 0 represents unavailability and 1 represents availability. Node unavailability duration represents the total time for which a node was available, which is plotted in Figure 1. We can observe from the graph that most of the nodes lie in the unavailability interval below 2000 hours. We have shown only those intervals as unavailability intervals which appeared after the node joined the Grid for the first time.

3.2. Number of Failures

Node failure can be defined as a state of the node when it is no longer able to serve its purpose as part of the Grid. So if we analyze the failure incidents which a node has gone through during the trace file collection period, then it looks like the pattern shown in Figure 2. If we observe the nodes that are part of the site ids 1, 3, 11, 12, 13 (Table1), we can see that the nodes of the same site have a common pattern of unavailability. The reason for this might be the homogeneous nature or any type of periodic effects or calendar effects.

Table 1. Node ID and Corresponding Node Site

Site Id	Site Name	Beginning Node ID	Last Node ID	Number of Nodes
1	G1/site1/c1	0	63	64
2	G1/site1/c2	320	477	64
3	G1/site1/c3	741	839	99
4	G1/site1/c4	406	889	32
5	G1/site2/c1	670	1287	95

6	G1/site3/c1	64	888	57
7	G1/site4/c1	125	669	106
8	G1/site4/c2	1018	1250	57
9	G1/site5/c1	490	886	56
10	G1/site5/c2	1141	1233	69
11	G1/site6/c1	542	647	103
12	G1/site6/c2	1208	1231	24
13	G1/site7/c1	840	885	46
14	G1/site8/c1	85	1017	342
15	G1/site9/c1	891	1286	74

3.3. Failure Prone Nodes

Failure prone nodes are those nodes which have more than average number of failures. Average number of failures is calculated by the total number of failure events divided by the total number of nodes. The result is shown in Figure 3. Nodes which are more prone to failures have a higher frequency of failures. Nodes with high failure frequency are also the one with short TTR (Time to Resume). Such nodes are a better option for executing small jobs whose execution time is in seconds or minutes. In other words, we can say that jobs will be executed between consecutive failures. Scheduling policy may be modified to refrain from submitting loner jobs on failure prone nodes for QoS requirements.

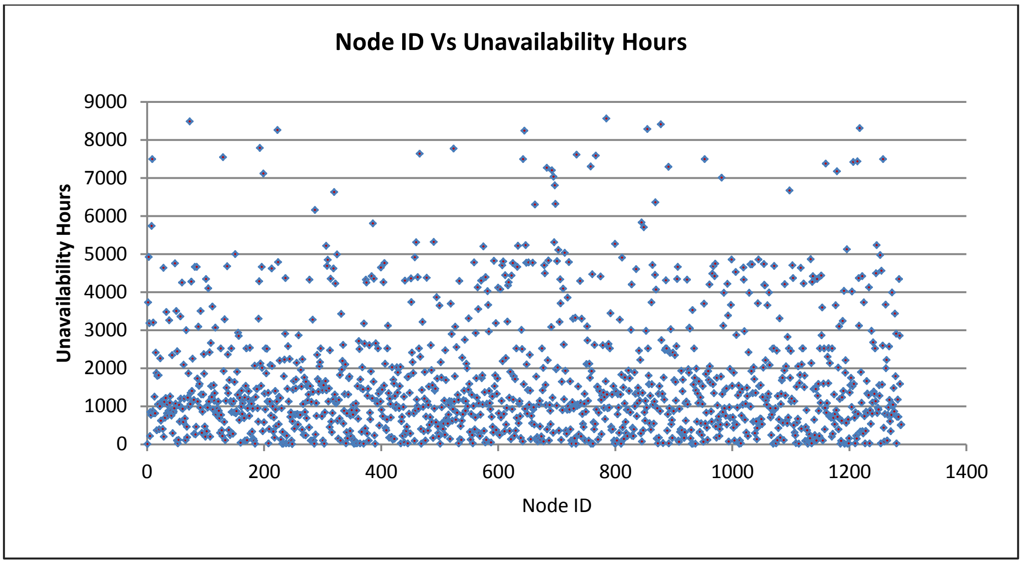


Figure 1. Node ID Vs Unavailability Hours

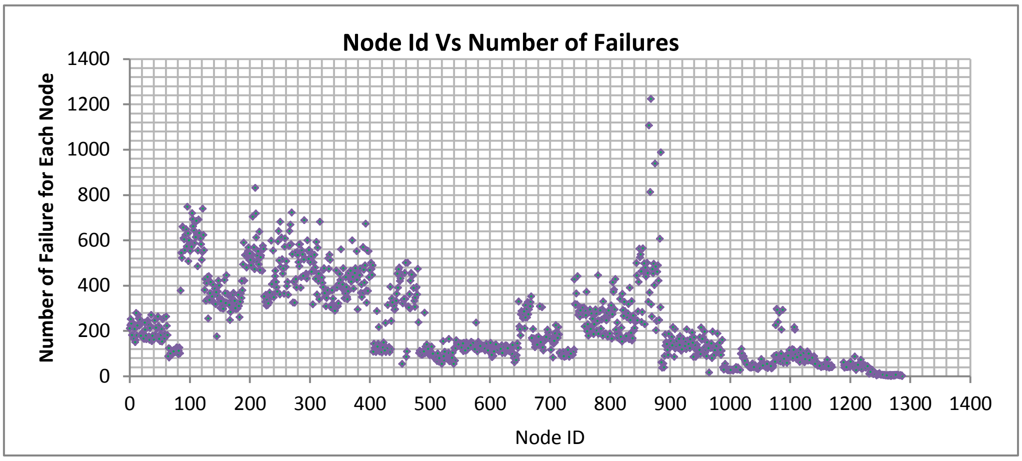


Figure 2. Node ID Vs Number of Failures

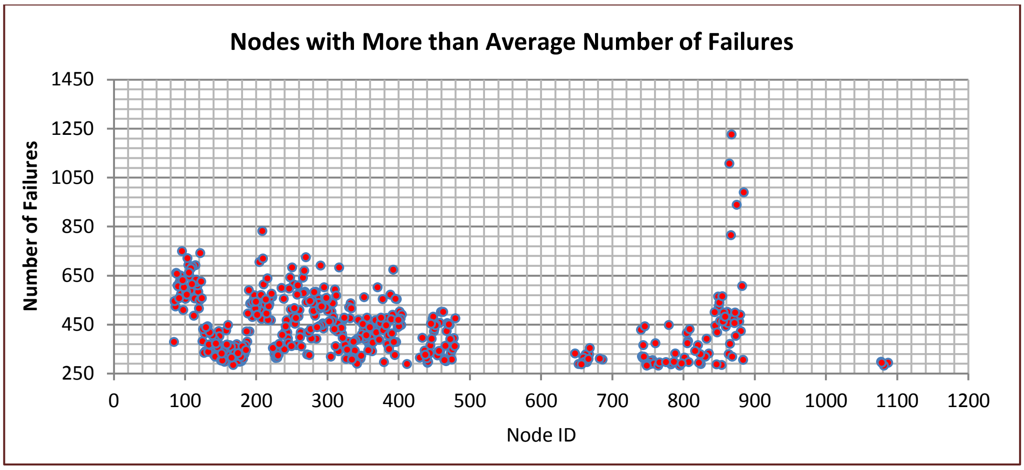


Figure 3. Failure Prone Nodes

3.4. Node Dedication

Node dedication factor is introduced as a new concept which represents the node participation on the Grid. Node dedication is calculated as the total availability duration divided by the duration for which we are calculating node dedication *e.g.*, one week, one month or year. NDF for Grid5000 is shown in Figure 4. For simplicity purpose, we have taken the dedication duration as the full duration of the trace file. Node dedication is a measurement of the time that a node spends on the Grid despite of the failures. We can combine dedication and number of failure values to design a scheduling metric. This combination will help us to create a profile of the node for its usage purpose. E.g. If a node has a high dedication factor, but also has high number of failures, then we can classify the node as a node with short duration failure. This type of information cannot be retrieved from the number of failures alone.

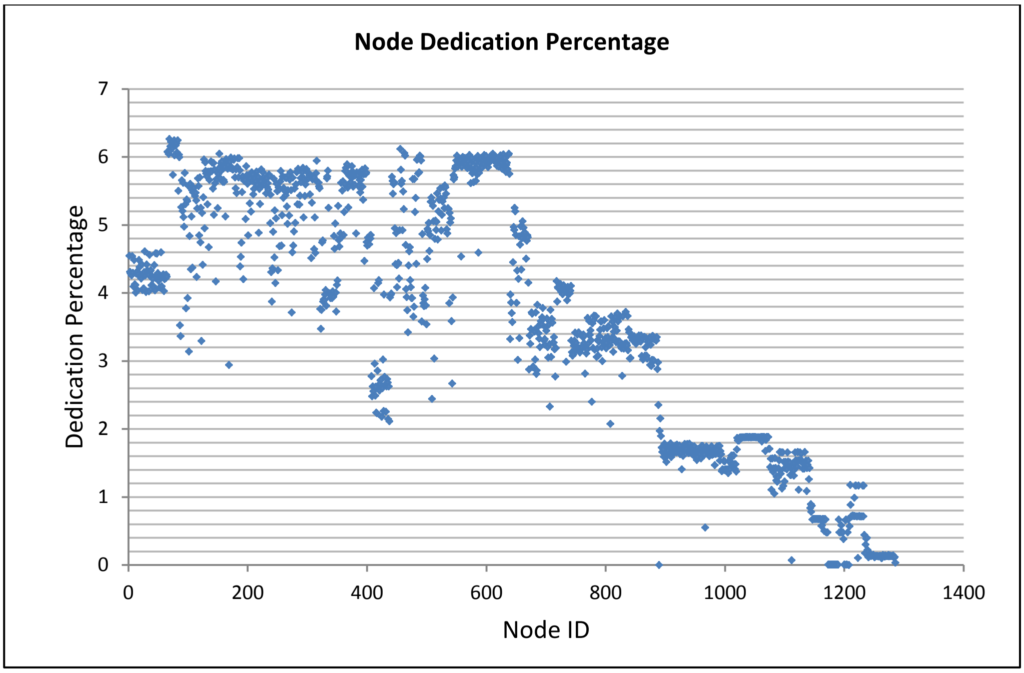


Figure 4. Node Dedication Factor

3.5. Average Time to Resume after Failure

ATR (Average Time to Resume) is the average time taken by a node to recover from failure. Average time to resume after failure will depend on the type of failure. Nodes with minor faults may recover quickly as compared to hardware or operating system related faults. This factor may be considered as one of the important factor while classifying nodes based on reliability and availability. It may also be helpful in better resource management and also for designing new job scheduling policies. We can also use these nodes to execute jobs with checkpointing capabilities. Result of the analysis are shown in Figure 5.

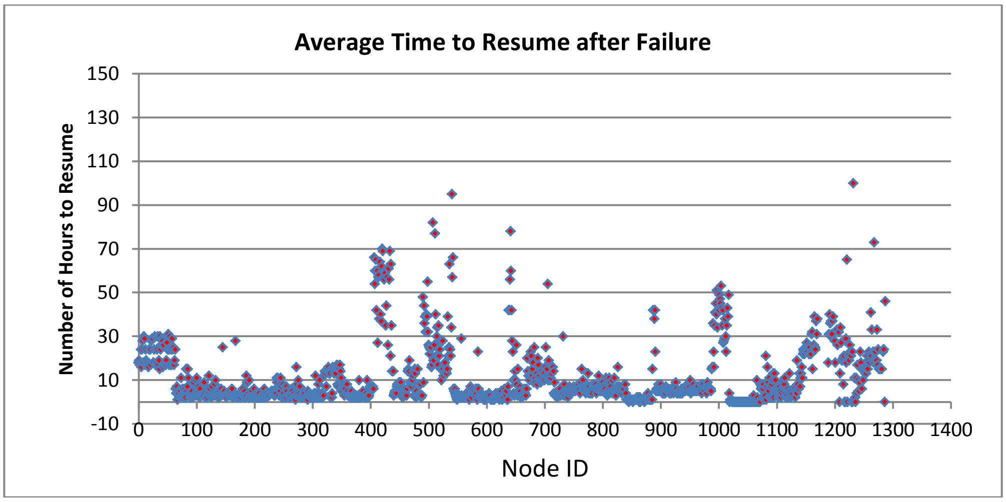


Figure 5. Average Time to Resume After Failure

4. Grid Workload Archive Analysis

FTA analysis has given an overview about how nodes actually behave inside a Grid. Graphs give us a clear idea about failures, failure duration, failure frequency, node dedication and resume time. These factors have a direct effect on the efficiency and reliability of Grid. Failures lead to the delayed or failed job execution. GWA trace file contains information about the type of job *e.g.*, batch job or interactive job. We have conducted an analysis of GWA to look into the trace file and to see how the Grid environment handle these jobs and also analyze the failure rate of these jobs. GWA analysis has been divided into following categories.

4.1. Job Execution Status

Before starting with a detailed analysis, we analyzed jobs based on the job status. Total 10, 20,151 number of jobs were submitted on the Grid for execution. Results of job execution status are shown in Figure 6. Considering workload consisting of batch and interactive jobs, we have found that around 4.93% of the total jobs are interactive jobs. Interactive jobs have special interaction requirements from the user side as compared to the batch jobs. If we compare the job execution status of batch and interactive jobs then, we can clearly see that interactive jobs have a higher failure rate as compared to batch jobs in Figure 7.

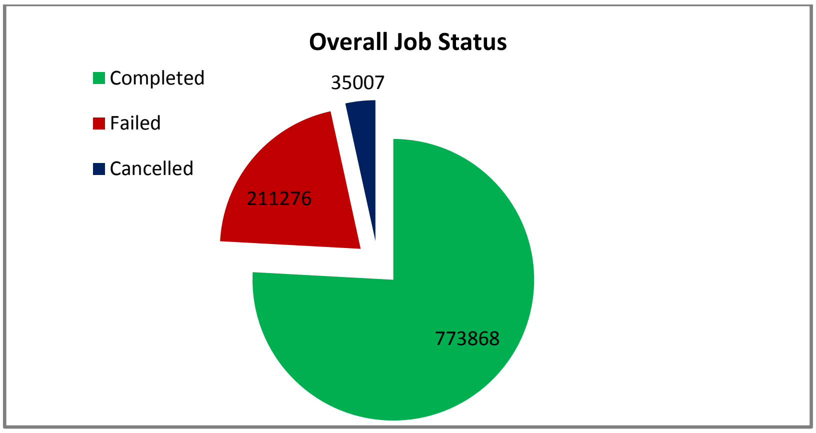


Figure 6. Job Execution Status

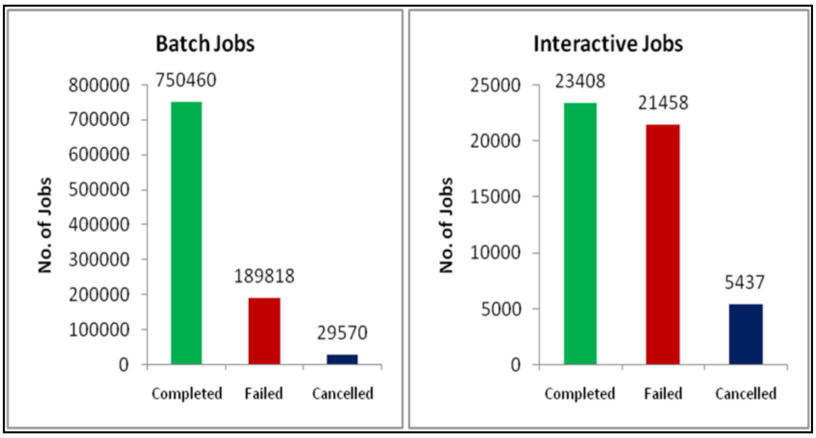


Figure 7. Batch Jobs Vs Interactive Jobs

4.2. Failure Stage of Jobs

Moving further, we have studied that how costly a job failure can be in terms of total processing power consumed by a job before failure. Trace file provides information about job submission time, requested CPU time and runtime of the job. To measure the failure stage of the job, we have used requested time as the target and run time as the failure point. Although a job can be completed before the requested time but to standardize the calculation we have assumed that the job shall be completed at the completion of the requested time. We have considered only those jobs as a part of analysis whose runtime is positive and less than the requested time. Job Failure analysis is beneficial while modifying or designing the scheduling policies because if most of the jobs are failing after completing a significant amount of processing then it will affect the system quality of service and throughput. Along with other solutions checkpointing and over provisioning of resources may be the possible solutions to minimize the job failure cost. Results of this analysis are shown in figure 8 given below. From this graph, we can observe that as we move from 0% to 100% completion the number of jobs in the next range are less than the previous range (except for 80-90). So based on our assumption we can say that maximum number of jobs fail in the initial stage of the execution.

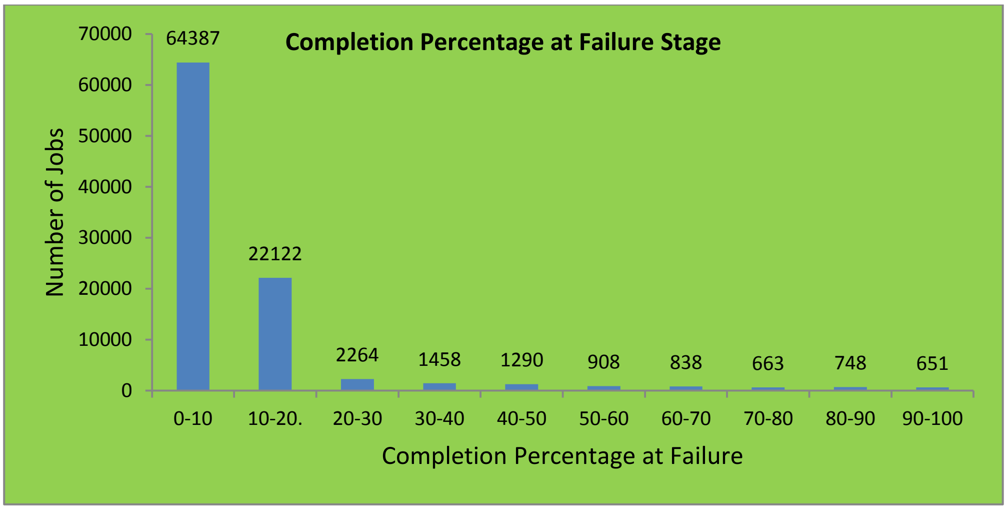


Figure 8. Job Completion Percentage

4.3. Completed, Failed and Cancelled Jobs

If we compare the average number of processors and the average processing time requirement of jobs, then we can learn from the data that the failed jobs have a higher value as compared to the completed jobs. Because the job which requests more processors or processor time is more vulnerable to faults and failures as more time a job spend on the Grid more is the probability that some fault will occur on the nodes executing the job which will lead to failure or cancellation of the job. The graphical representation of this is shown in Figure 9. As it is obvious from the dot plot chart that per job processing time request is maximum in case of cancelled jobs and it decreases when we move towards failed and then completed. Almost a similar pattern can be observed in figure number 10 as well, which represents the number of processors required for execution of the job. Detail of plot variablesis given in Table 2.

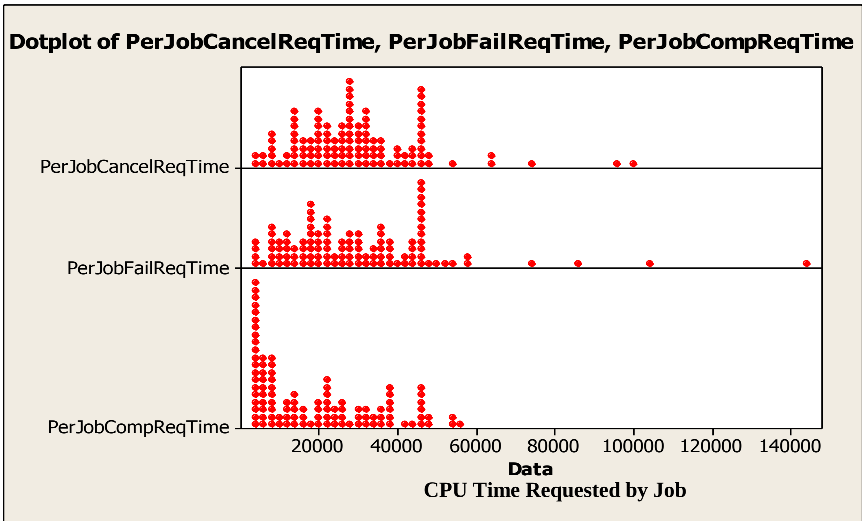


Figure 9. Average Processing Time Requested by Completed, Failed and Canceled Jobs

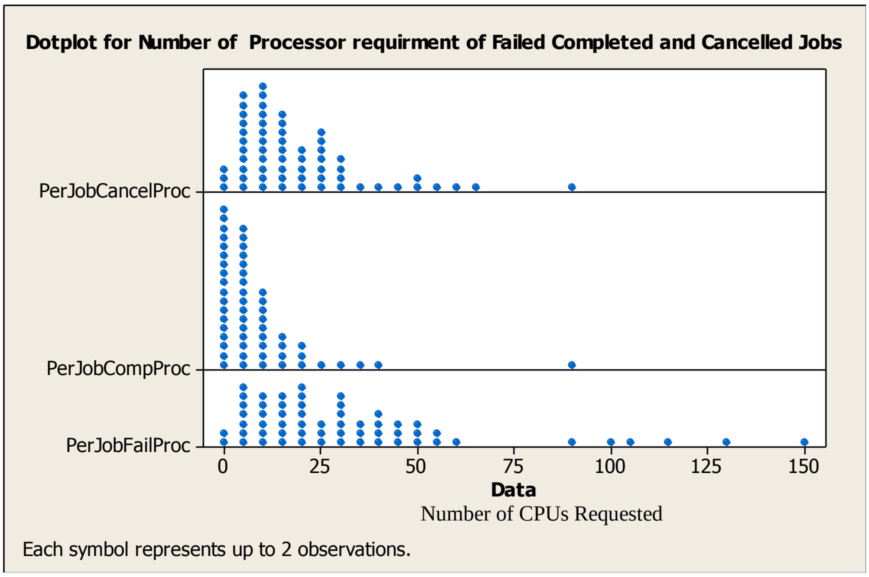


Figure 10. Average Number of Processor Requirement per Job for Completed, Failed and Cancelled Jobs

5. Combined Analysis of FTA and GWA

After the individual analysis of FTA and GWA, now we conducted a combined analysis of two to establish a relationship between nodes and jobs. Before going for correlation based analysis, we conducted a month wise analysis of nodes, failures, jobs from FTA and GWA. The results of this month wise analysis are shown in the graphs below. Figure 11 shows the number of nodes which have participated in each month. We can see that there is an increase in number of nodes in each month and the reason for this is that with time, more number clusters were added to the Grid. Number of failures which took place in each month are shown in Figure 12. Reason behind choosing this data set from 2005 and 2006 is that this is the only duration for which both FTA and GWA have been collected simultaneously. No other available data set satisfies this requirement.

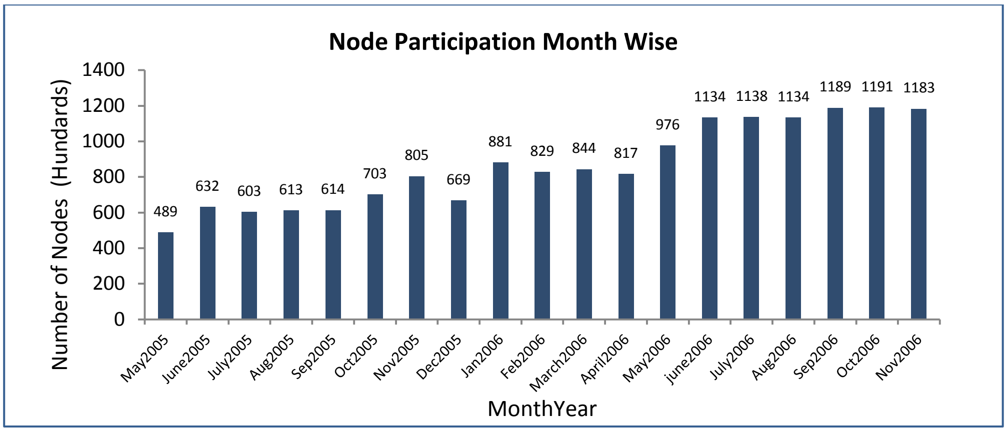


Figure 11. Node Participation Month Wise

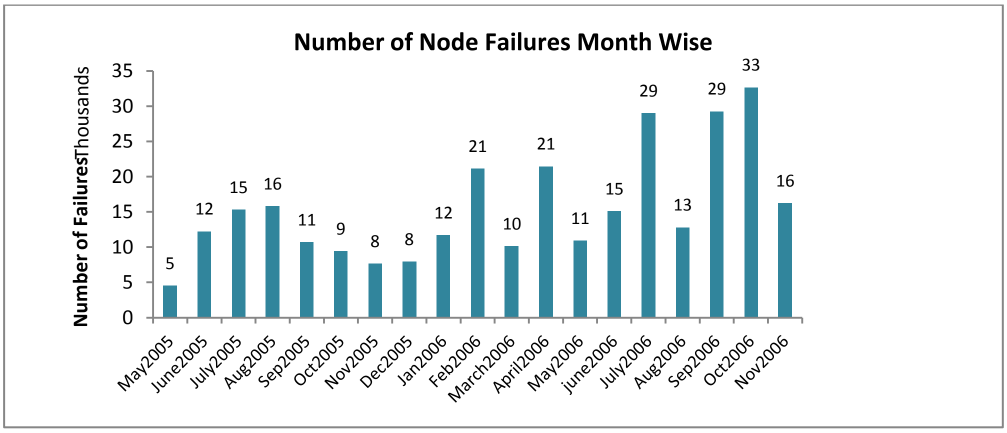


Figure 12. Month Wise Number of Node Failures

Total number of jobs submitted during these months is shown in Figure 13 below and the corresponding execution status of these jobs is shown in Figure 14.

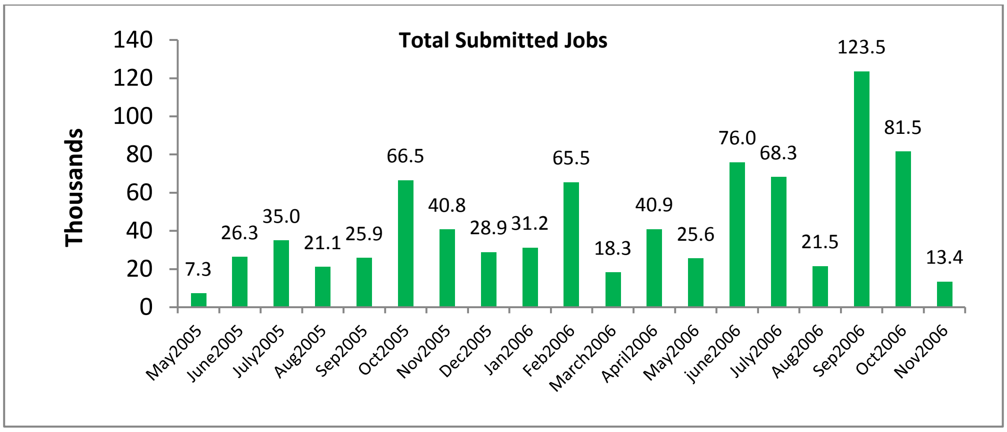


Figure 13. Month Wise Job Submissions

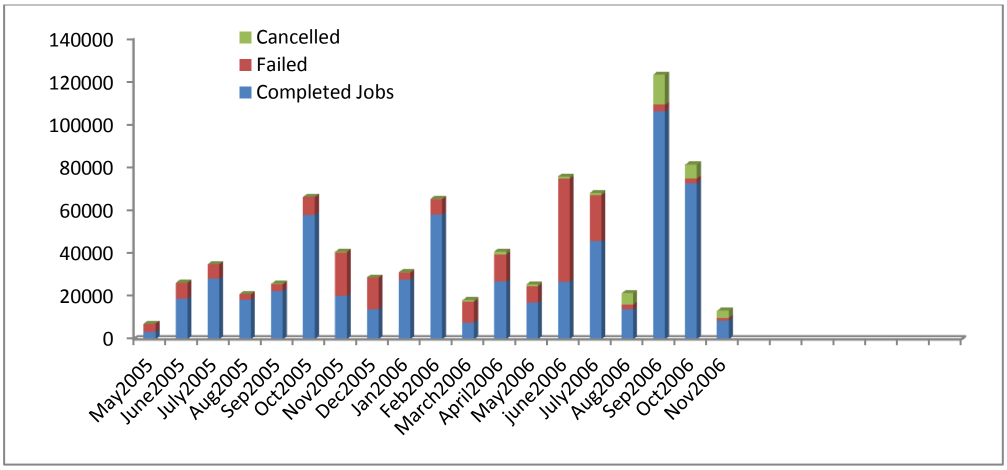


Figure 14. Month Wise Job Execution Status

5.1. Sample Size Selection

Sample data for the combined analysis of FTA and GWA were collected using random data selection technique. Time duration was randomly generated by the random function from a range of maximum and minimum values. Then the corresponding data were fetched from FTA and GWA for the randomly generated duration. With the help of aggregation functions total number of failures, total failure time, total number of submitted jobs, number of failed, completed and cancelled jobs were calculated. Correlation analysis was conducted on sample data. Different key elements of this sample set are shown in Table 2.

Table 2. Variables and Description

Key Attribute Name	Description of Attributes
NumNodeFailures	Number of Failure Events
NumNodes	Total Number of nodes participated
FailureDuration	Duration for which Nodes were Unavailable
JobTotal	Total Jobs Submitted for Execution
NCompleted	Number of completed Jobs
NJobFailed	Number of Failed Jobs
NCanelled	Number of cancelled jobs
NProc	Number of Processors required for Execution
TotalReqTime	Total processing time requested by all the jobs
TotalNProc	Total Number of processors requested by all jobs
PerJob(Status)ReqTime	Average time requested by Failed, completed and cancelled jobs
PerJob(Status)Proc	Average number of processors requested by Failed, completed and cancelled jobs.

6. Correlation and Result Analysis

Correlation is a statistical measure that indicates the extent to which two or more variable values fluctuate together. A positive correlation indicates the extent to which two or more than two variables increase in parallel and a negative correlation indicates the extent to which one variable increase as the other decreases. The strength of the linear association between two variables is quantified by the correlation coefficient whose value can vary from -1 to +1.From the randomly collected data set we conducted a correlation based analysis of different variables. Results of this analysis are shown in the following Table 3 and Figure 15.

Table 3. Correlation Coefficient and Significance Values

	Correlation	NJobFailed	NumNodes	NumNode Failures	FailureDuration
NJobFailed	Correlation Coefficient	1.000	.181*	.551**	.498**
	Sig. (2-tailed)	.	.049	.000	.000
	N	520	518	520	516
NumNodes	Correlation Coefficient	.181*	1.000	.126	.041
	Sig. (2-tailed)	.049	.	.174	.665
	N	518	518	518	514
NumNodeFailures	Correlation Coefficient	.551**	.126	1.000	.660**
	Sig. (2-tailed)	.000	.174	.	.000
	N	520	518	520	516
FailureDuration	Correlation Coefficient	.498**	.041	.660**	1.000
	Sig. (2-tailed)	.000	.665	.000	.
	N	516	514	516	516

We can make the following conclusions from these results.

I. There exists a positive correlation between the number of node failures and the number of failed jobs. So we can say that with the increase in the number of node failures, number of failed jobs are also increasing and this explains the very basic behavior of the Grid environment. So number of node failures has a direct effect on the quality of service of the environment. Scheduling policy can use this information in order to make better scheduling decisions in a failure critical situation.

II. If we look at the correlation coefficient of the number of node failures and failure duration, we can see that there exists a considerable correlation of 0.660 between two. So with the increase in number of node failures, the failure duration can be predicted.

III. Correlation coefficient between failure duration and number of failed jobs is also positive and is equal to 0.498. Although it is not a strong correlation value but it supports our hypothesis that longer the nodes stay unavailable, more will be the failed or cancelled jobs.

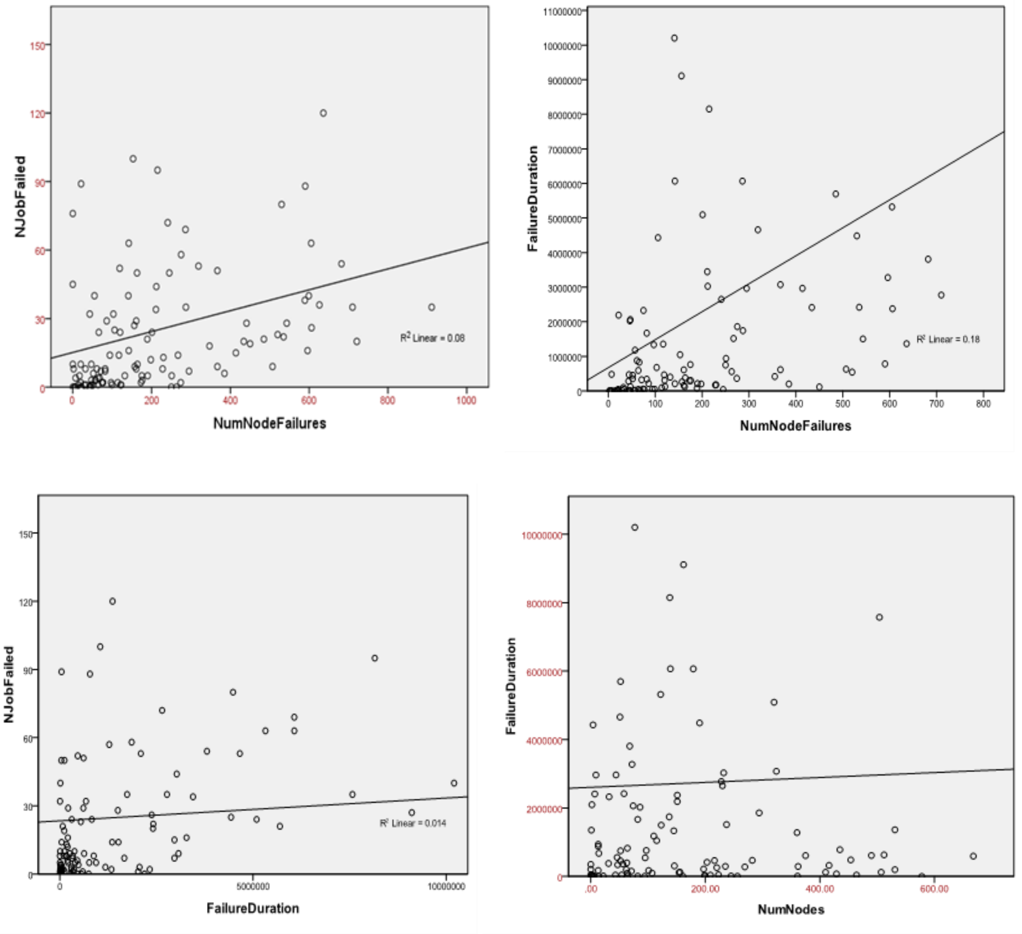


Figure 15. Correlation Plots of Different Variables

7. Conclusion and Future Work

From the combined analysis of FTA and GWA, we have found many interesting facts related to the Grid nodes and Grid jobs.

- From the FTA analysis, we have found node failures, failure durations, failure frequency, failure prone errors, node dedication factor and node TTR (Time to Resume). This gives researchers an idea about node availability and unavailability patterns.
- While conducting GWA analysis, we have found that 20% of the jobs submitted on the Grid for execution have failed and approximately 3.43% of the jobs were cancelled.
- From the comparative analysis of interactive jobs Vs batch jobs, we have learnt that an interactive job has a higher probability of failure and cancellation as compared to the batch job.
- Through combined analysis of FTA and GWA we have also observed that there exists a positive correlation between number of node failures and number of failed jobs.
- Failure duration of nodes and the number of failures are also positively correlated with a correlation coefficient of 0.660.

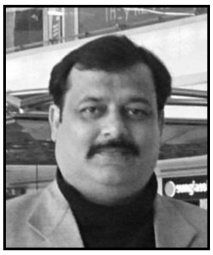
- We have also found out that with the increase in failure duration, number of failed jobs also increases because these two factors are also positively correlated.

Information which was missing in these trace files is the fault description and the corresponding corrective actions taken to recover from it and the memory requirements of the job. With this information in hand we can conduct a detailed analysis of how these factors influence system reliability and performance. We can also develop a new type of trace file which can record both FTA and GWA in a single log file because in that case we will be able to establish a relationship between a particular node and job which is being executed by it.

References

[1] D. Sharma and M. P. Mittal, "Grid Computing An Overview", International Journal of Science, Engineering and Technology Research, vol. 2, no. 3, **(2013)**, pp. 606-611.
[2] U. Schwiegelshohn, R. M. Badia, M. Bubak, M. Danelutto, S. Dustdar, F. Gagliardi, A. Geiger, L. Hluchy, D. Kranzlmüller and E. Laure, "Perspectives on grid computing", Future Generation Computer Systems, vol. 26, no. 8, **(2010)**, pp. 1104-1115.
[3] R. Garg and A. K. Singh, "Fault Tolerance in Grid Computing: State of the Art and open issues", International Journal of Computer Science & Engineering Survey (IJCSES), vol. 2, no. 1, **(2011)**.
[4] M. F. L. P. Z. Qiaoming and j. Qin, "Comparison Of Several Fault-Tolerance Policies On Grid Environment", Computer Applications and Software, vol. 1, **(2010)**, pp. 004
[5] http://fta.scem.uws.edu.au.
[6] C. D. Shanny Anoep, D. Epema, A. Iosup, M. Jan, H. Li and L. Wolters, "Grid Workload Archive".
[7] N. Yigitbasi, M. Gallet, D. Kondo, A. Iosup and D. Epema, "Analysis and modeling of time-correlated failures in large-scale distributed systems", Editor (Ed.)^(Eds.): 'Book Analysis and modeling of time-correlated failures in large-scale distributed systems' (IEEE, 2010, edn.), pp. 65-72.
[8] D. Kondo, B. Javadi, A. Iosup and D. Epema, "The failure trace archive: Enabling comparative analysis of failures in diverse distributed systems", Editor (Ed.)^(Eds.): 'Book The failure trace archive: Enabling comparative analysis of failures in diverse distributed systems' (IEEE, 2010, edn.), pp. 398-407.
[9] B. Javadi, D. Kondo, J.-M. Vincent and D. P. Anderson, "Mining for statistical models of availability in large-scale distributed systems: An empirical study of seti@ home", (Ed.)^(Eds.): 'Book Mining for statistical models of availability in large-scale distributed systems: An empirical study of seti@ home' (IEEE, 2009, edn.), pp. 1-10.
[10] A. Andrzejak, D. Kondo and D. P. Anderson, "Ensuring collective availability in volatile resource pools via forecasting", Managing Large-Scale Service Deployment' (Springer, 2008), pp. 149-161.
[11] A. Litke, D. Skoutas, K. Tserpes and T. Varvarigou, "Efficient task replication and management for adaptive fault tolerance in mobile grid environments", Future Generation Computer Systems, **(2007)**, vol. 23, no. 2, pp. 163-178.
[12] E. Huedo, R. S. Montero and I. M. Llorente, "Evaluating the reliability of computational grids from the end user's point of view'", Journal of Systems Architecture, vol. 52, no. 12, **(2006)**, pp. 727-736.
[13] B. Song, C. Ernemann and R. Yahyapour, "User group-based workload analysis and modelling", Editor: 'Book User group-based workload analysis and modelling' (IEEE, 2005, edn.), pp. 953-961.
[14] B. Schroeder and G. A. Gibson, "Disk failures in the real world: What does an MTTF of 1, 000, 000 hours mean to you?', in Editor: 'Book Disk failures in the real world: What does an MTTF of 1, 000, 000 hours mean to you?' (2007, edn.), pp. 1.

Authors



Dr. Rajesh K. Bawa holds Master's and Ph.D degree in the area of Numerical Computing from IIT Kanpur INDIA. Presently he is Associate Professor at Department of Computer science Punjabi University, Patiala, INDIA. His present area of interest is Parallel and Scientific Computation and has published many research papers in journals of reputed publishers and Also associated with review and editing work with many journals.

Ramandeep Singh received his B.Tech from Punjab Technical University and M.Tech from Punjabi University, Patiala in 2006 and 2009, respectively. At present he is pursuing his PhD in the field of fault tolerance in Grid computing from Punjabi University, Patiala. He is also working as Assistant Professor at department of computer science Lovely Professional University, Jalandhar, India. His research interests include Grid computing, Cloud and distributed computing. He is a PhD student at department of Computer Science at Punjabi University, Patiala.