METHOD FOR CONDUCTING A COMBINED ANALYSIS OF GRID ENVIRONMENT'S FTA AND GWA THROUGH SESSION BASED MAPPING OF TRACE VARIABLES

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

[1]Lovely Professional University, Jalandhar, India

[2]Department of Computer Science, Punjabi University Patiala, India

ABSTRACT

Grid computing environment due to its scale and heterogeneous nature is more vulnerable to faults. To store and analyze fault and workload information, FTA Fault Trace Archive and GWA Grid Workload Archive are used. Previously researchers have analyzed FTA and GWA as separate research problems but in this research paper we have proposed a method for conducting a combined analysis of FTA and GWA based on session based mapping of trace file variables. This is the first attempt to conduct a combined analysis of these two trace files. Along with the step by step process of combining trace files we have also included do's and don'ts while conducting this analysis .Through this combined analysis we have established a correlation based relationship among number of node failures, number of failed jobs, failure duration and number of nodes. We have found that these variables are positively correlated with different correlation coefficients.

Keywords: *Grid Computing, FTA Fault Trace Archive, GWA Grid Workload Archive, Correlation*

1. Introduction

Grid computing is a concept which involves combining computing and data storage resources from different sources into virtual organizations (VO). These resources are of heterogeneous nature and are connected through dedicated network links or via internet. Due to its large size and lack of central control Grid resources are vulnerable to different types of faults. These faults affect the quality of service of Grid by either failing the submitted jobs or by delaying execution of these jobs. So researchers and developers need to analyze workload and fault information to study the Grid environment and to improve the overall Grid performance along with reliability. Trace files are used to collect data about the events taking place inside the Grid. These events include faults which are occurring on the Grid resources and events related to the jobs which are submitted for execution on Grid. Basically two types of trace files are used in Grid for collecting resource availability and job data which are FTA and GWA.

FTA stands for Fault Trace Archive [1] and GWA stands for Grid Workload Archive [2]. FTA collects information about the faults which are taking place on the Grid nodes and GWA collects

DOI: 10.5121/ijgca.2014.5101

information about the workload which is being submitted on the Grid. FTA contain information about nodes, platforms, sites, hardware configuration, availability and unavailability events, time of occurrence of these events and the corresponding step which was taken to deal with the failure situation. In the same way GWA contain information about job id, submission time, waiting time, processor requirement, estimated execution time, status and other information fields related to queues, users and groups. Status of a job can have different values e.g. completed, failed, cancelled or any other trace file specific value. Trace files are available in different formats e.g. raw, tabbed or mysql.FTA is a collection of different files which can be joined with each other with the help of the common fields, like we perform join operation among SQL tables. Due to space limitation we cannot include the whole format of trace files but details about the format and design of FTA can be found from FTA web page [3] and the detailed format information of GWA can be found from GWA web page [4][9].

These trace files can be analyzed using different tools e.g. Matlab, GridSim or mysql. GridSim is a good tool in case you want to run simulations on these trace files according to the constraints of a Grid environment. These trace files i.e. FTA and GWA have been individually analyzed by many researchers for studying Grid environment for different purposes, but no one have considered combining these trace files to study the relationship among two. In this research paper we are going to propose and implement a technique using which we can combine FTA and GWA. Both these trace files are collected from the same Grid platform simultaneously i.e. in parallel. Using this technique we can study the influence of the events of one trace file on the events of another trace file.

2. Related Work

Trace files are a good way of collecting data about the events taking place inside a system either distributed or centralized. Analysis of these trace files can reveal many interesting facts related to the system. Artur Andrzejak et. al [5]have worked on SETI@home host trace files, which consisted of 48000 hosts, to analyze host availability patterns. They proposed a model which can ensure that a certain number of hosts will be available for a certain amount of time either with replication or with over provisioning of resources. Similarly Bahman Javadi, et al., [6] 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 they have discovered that availability of around 34% of hosts is a truly random process but rest of these hosts can often be modeled in the form of different groups with few distinctions from one another. User oriented analysis reveals the facts about the workload patterns which are submitted by different users [11][12].

Bianca [7] have used trace file to analyze and predict the average life span of hard drives being used at a high performance cluster. Data sheets of these 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 to the actual field data show that the minimum disk replacement rate is 1% and usually it is up to 3-4%, and in some cases up to 13% replacement rate is there. Analysis show that actual data gathered from field may differ from the conceptual or ideal data. Nezih Yigitbasi et.al. have analyzed availability and unavailability of Grid nodes and have identified that there exist a predictable pattern in these events and this behavior can be modeled. They have also established a correlation based model for node failures [8].

GWA trace files are used by researchers to study what kind of jobs are submitted on the Grid and how does the success and failure of a job is dependent on different parameters of job. Grid environments are either application specific or job specific. Application specific Grid can execute or process only a specific application related tasks while on the other hand job specific Grid can execute different types of jobs which may belong to different applications. Although individual analysis of trace files is very useful but in this research paper we are going to propose a method of combining these two trace files for a combined analysis.

3. Combining FTA and GWA

Following are the operations which need to be performed for conducting combined analysis.

3.1 Acquiring FTA and GWA Trace Files

For conducting a combined analysis of FTA and GWA very first condition is that both trace files should belong to the same Grid environment and should be collected at the same time i.e. events should be logged in parallel. Reason for this is that if these trace files belong to different platforms or are collected at different time intervals then there will be no benefit of conducting the combined analysis because then these events cannot be related to each other and the analysis will make no sense. In the absence of mysql format of trace files, we can use SPSS for importing data and converting it in required formats. From the excel format we can insert data in SQL or mysql table.

3.2 Trimming Trace Files

After acquiring both trace files we need to check event start time and event end time of both trace files. Reason for this is that collection process of these trace files may start and end at different times so we need to trim one or both the trace files at a common point of event start time and event end time. Let's consider an example where FTA trace file was started at 10:00:00 am, 10 Jan 2013 and it ends on 12:00:00 pm, 15 Jan 2014. Similarly GWA was started at 09:00:00 am, 31 Dec 2012 and it ended on 11:00:00 am, 1 March 2014. Although both these trace files are collected from the same environment and individually there is nothing wrong with these trace files but for combined analysis we need to trim either one or both from either beginning or end to synchronize the event start and event end time. So in case of above considered example we will trim GWA so that first event start time is 10:00:00 am, 10 Jan 2013 and last event end time will be 12:00:00 pm, 15 Jan 2014. In order to remove noise we will also have to remove the events which start before the trimming point but end after it, because such events may cause deviations in calculations. Same will be applicable to events which start before the trimming point but end after the trimming point. Although we can say that few such events will not make a difference when we are talking about millions of such events, but to minimize the errors this is the one precaution which should be considered in the beginning. So at the end of this step we will have two trace files which start and end at the same time.

3.3 Slicing and Dicing

Now that we have two trace files of equal time duration we can start with our analysis. Slicing means that we will divide trace files into sub parts. But the one thing to keep in mind is that these

slices should be of equal size, otherwise it will lead to inconsistent results. Slicing duration can vary from an hour to day or week and even a month depending on the type of analysis. Because the event time is epoch time so we would have to convert seconds to days, weeks or months and then add this value subsequently for slicing and dicing operations.

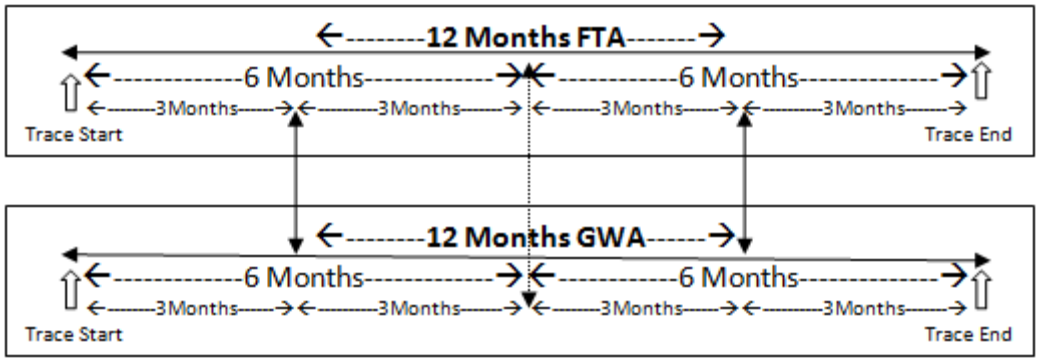

Figure 1: Mapping and Slicing of FTA and GWA

Figure 1 represents slicing operations in which FTA and GWA trace files of 12 months duration have been sliced in 6 month and 3 month durations. Vertical lines represent slicing operations. As we can see that synchronizing is very well considered here as both the trace files have been sliced at the same locations i.e. time.

3.4 Data Extraction

After slicing down the trace files now we can retrieve data from these trace files for our combined analysis. Data selection and extraction is based on what kind of analysis we are performing and which fields of the trace file are required for the analysis. As we know that these two trace files do not share any common field except the event time variable so we cannot combine FTA and GWA directly. We have used a different approach which is based on aggregation functions.

a) FTA Data Extraction

Fault Trace Archive contains information about Grid nodes. If N_F represents number of failures and $N_F(i)$(s) represents number of failures for node i where $i \in J\{1,2,3,4,.....n\}$ .J represents collection of n nodes which are part of Grid environment and S represents time slot i.e. session. D represents total failure duration of all the nodes on the Grid whereas D(i) represents total failure duration for node i. F$(i)T$ represents failure frequency of node I for duration T.R(i) represents average resume time after failure for node i.

In the absence of common variable, we have used the method of aggregation functions (basically summation and average) to collect and retrieve data for analysis. $N_F(i)$, D(i) and F(i) can be directly calculated if we are considering the whole trace file as our possible data set. But we can also calculate these values for a specific duration S which is our sampling duration. Now the value of $N_F(i)$, D(i) and F(i) will be calculated from this duration S. S can be generated randomly

so that data should be retrieved form the whole trace file rather than one particular section of the trace file. This helps in getting more accurate results.

b) GWA Data Extraction

Similar to FTA now we need to extract data from GWA and map it with FTA variables. Once again we will use aggregation function and sample duration for retrieving data. J_{id} represents one job form the GWA with unique identifier *id* where $id \in W\{1,2,3,.....k\}$ here *W* represents set of jobs which were submitted on the Grid for execution. C_{id} represents number of processors which are requested by job J*id* for execution. WT_{id} represents waiting time for job J_{id} and R_{id} represents runtime requested by J_{id}. U_{id} represents user id of the user who submitted job $J_{id.}$ S_{id} represents status of the job J_{id}. Status as already discussed can be completed, failed, cancelled or any other trace file specific reason. $N_{completed}(s)$, $N_{failed}(s)$, $N_{cancelled}(s)$ represents number of completed, failed and canceled jobs respectively in duration *S*.

4. Combining Data

After the data extraction, now data of different variables should be mapped according to the time slots. In case the time slot shuffling is taking place in both the data sets simultaneously, then it will not lead to any inconsistencies, but if one data set's time slot have been shifted but not the other one, then it will lead to incorrect results. Graphical representation of this concept is shown in figure 2. In this figure S1, S2 and S3 represents the time slots for which aggregation functions have been used to retrieve data. In the figure *Fvar* and *Gvar* represents variables of FTA and GWA. Variable can be any variable which the analyst wants to include in the analysis and which have some relationship with the other variables of the trace file. This relation can be direct or indirect. For example we may consider that there exist a relationship between number of node failures and number of failed jobs. So on one side i.e. FTA we can have data about number of node failures and other relevant variables such as failure duration, resume time etc. and on the other side i.e. GWA we can have data related to number of job failures and the other relevant variables which are required for this analysis such as waiting time, number of processors requested by job, run time etc. It is not mandatory that the number of variables from FTA and GWA which are being used in analysis needs to be same in number. There can be a case where we may have only one or two variables from FTA but from GWA side there may be more than two. This is totally dependent on the analysis model.

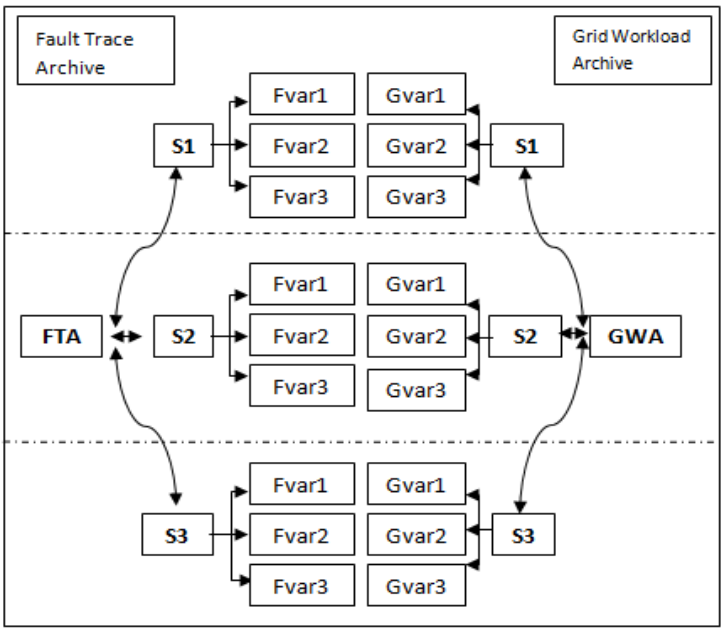

Figure 2: Combining FTA and GWA with Mapping Time Slots

5. Conducting Analysis

Based on the above proposed technique we have conducted a correlation based analysis by combining FTA and GWA. Through this analysis we have identified relationship among the variables of FTA and GWA. Variables which we have considered in our analysis are $N_F(s)$ (Number of node Failures in duration *S*), D(s) (Failure Duration in time S), N_N (Number of Nodes), $N_{failed}(s)$ (Number of Failed Jobs in duration S).

Serial No.	Variable	Variable Description	Variable Name for Plotting
1	$N_F(s)$	Number of node Failures in duration *S*	NumNodeFailures
2	D(s)	Failure Duration in time S	FailureDuration
3	N_N	Number of Nodes	NumNodes
4	$N_{failed}(s)$	Number of Failed Jobs in duration S	NJobFailed

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 increases as the other decreases. The strength of the linear association among two variables is quantified by the correlation coefficient whose value can vary from -1 to +1. We have used Spearman correlation for this analysis because of the absence of linear variation in data from both trace files. Reason being number of jobs submitted or failed and number of node failures are varying at a large scale from time to time .This technique uses a ranking based approach for calculating correlation coefficients. Following equation represents Spearman's correlation.

$$r_S = 1 - \frac{6 \times \sum d^2}{n(n^2 - 1)}$$

Here r_s represents Spearman's correlation coefficient, d represents the difference in the ranks of two variables whose correlation we are calculating and n represents the number of values which are being used for conducting this correlation. From the randomly collected data set we conducted a correlation based analysis of different variables. Result of this analysis is shown in the following table 1 and the corresponding plots are shown in figure 3

We can make the following conclusions from these results.

I. There exists a positive correlation between number of node failures and number of failed jobs. So we can say that with the increase in the number of node failures number of failed jobs is 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 number of node failures and failure duration then we can see that there exists a considerable correlation between two equal to 0.660. So based on the correlation with 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 equal to 0.498. Although it is not a strong correlation value but it supports out hypothesis that longer the nodes stay unavailable more will be the failed or cancelled jobs.

Table 1 : Correlation Analysis Results

	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

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

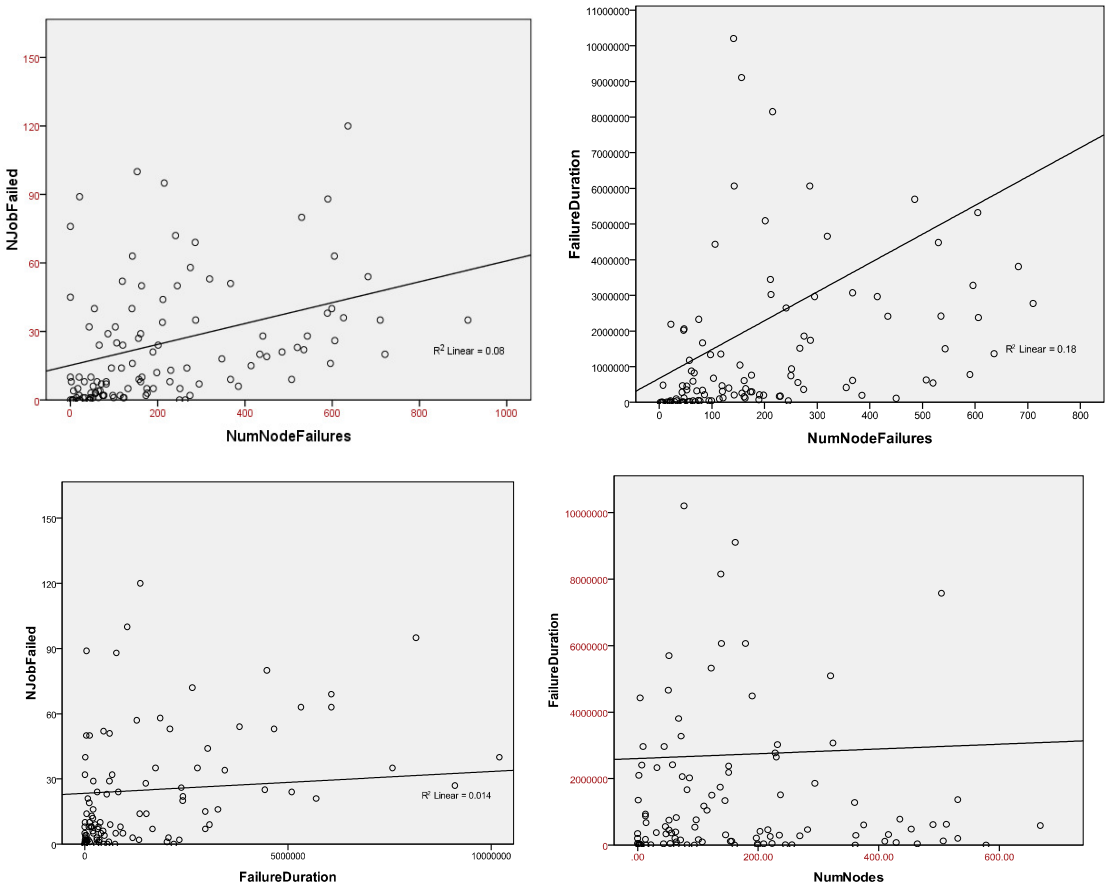

Figure 3: Correlation Analysis Plots

6. Conclusion and Future Work

In this research paper we have proposed first technique of combining Fault Trace Archive and Grid Workload Archive as a single research problem. We have discussed the step by step approach of combining FTA and GWA. We have also identified and discussed that what type of mistakes can be made while conducting the analysis and what kind of impact these mistakes can have on the results. Finally with the help of correlation based combined analyses we have found out that there are positive correlations among FTA and GWA variables and have also identified the coefficient of these correlations. In future we can establish a regression based model of different variables and can predict system behavior in response to variation of events.

References

[1] Fault Trace Archieve. Fault Trace Archieve. [Online]. http://fta.scem.uws.edu.au

[2] Hui Li, Mathieu jan, Shanny Anoep Alexandru Iosup, "The Grid Workload Archieve,".

[3] Alexandru Iosup,Matthieu Gallet,Emmanuel Jeannot,Derrick Kondo,Bahman Javadi,Artur Andrzejak Dick Epema. (2009) Fault Trace Archive: For improving the reliability of distributed systems. [Online]. HYPERLINK http://fta.scem.uws.edu.au/index.php?n=Main.FTAFormat

[4] Catalin Dumitrescu, Dick Epema, Alexandru Iosup, Mathieu Jan, Hui Li, Lex Wolters Shanny Anoep. (2007, January) The Grid Workloads Archive. [Online]. HYPERLINK http://gwa.ewi.tudelft.nl

[5] Derrick Kondo, David P.Anderson Attur Andrzejak, "Ensuring Collective Availability in Volatile Resource Pools via Forecasting,".

[6] Derrick Kondo, Jean-Marc Vincent, David P.Anderson Bahman Javedi, "Mining for Statistical Models of Availability in Large-Scale Distributed Systems: An Emperical Study of SETI@home,".

[7] Garth A,Gibson Bianca Schroeder, "Disk Failures in Real World: What does and MTTF of 1,000,000 hours mean to you?," in 5th USENIX Conference on File and Storage Technologies, San Hose, CA, Feb 14-16,2007.

[8] Matthieu Gallet, Derrick Kondo, Alexandru Iosup, Dick Epema Nezih Yigitbasi, "Analysis and Modeling of Time Correlated Failures in Large Scale Distributed Systems," 8th IEEE/ACM International Conference on Grid Computing, PDS-2010-004, 2010.

[9] Bahman Javadi, Alexandru Iosup, Dick Epema Derrick Kondo, "The Failure Trace Archive : Enabling Comparative Analysis of Failures,".

[10] Dick Epema Alexandru Iosup, "Grid Computing Workloads," IEEE Transactions on Internet Computng, vol. 15, no. 2, pp. 19-26, 2011.

[11] Dick Epema Alexandru Iosup, "Grid Computing Workloads: Bags of Tasks, Workflows, Pilots, and Others," , Netherlands, 2011.

[12] Carsten Ememann, and Ramin Yahyapour Baiyi Song, "User Group-based Workload Analysis and Modelling," in 2005 IEEE International Symposium on Cluster Computing and the Grid, 2005, pp. 953-961.