Research Journal of Applied Sciences, Engineering and Technology 8(16): 1833-1837, 2014
DOI:10.19026/rjaset.8.1171
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: July 24, 2014 Accepted: September 13, 2014 Published: October 25, 2014

Research Article

Trust Worthiness on Load Balance in Grids Using Standard Cauchy Distribution

[1]S. Tamilselvi, [2]G. Manoharan and [3]K. Nirmala
[1]Department of Computer Science, Bharathiar University, Coimbatore 641046,
[2]Department of Computer Science, Manonmanium Sundaranar University, Tirunelveli,
[3]Department of Computer Science, Quaid-E-Millah Govt. College for Women, Chennai, India

Abstract: Computational tasks are split into many modules and the resources of participating nodes of Grid are shared with to solve intensive jobs. Since the Grid participants are open in many cases, the respective providers do not maintain reputation in providing their promised services. Therefore 'Trust' worthiness of nodes of clusters of Grid is determined through various techniques with several parameters. Computational efficiency in Grid is another major issue that needs to be addressed. 'Load Balancing' in Grids could be a major remedial measure for tackling this issue. Literature on these two components in isolation of Grid technology is seen to be aplenty. But a combined study for relational effectiveness is worth an attempt. Standard Cauchy distribution is a continuous probability distribution, which can be applied so as to determine the acceptance levels of Cluster functionalities of the Grid. In view of this, the study attempts to make use of standard Cauchy distribution technique for validating Load balancing, which is performed after the determination of the trust worthiness of the participating nodes of a cluster. The study argues that the Grid cluster that is trust worthy alone can be performed with the load balancing and the conclusion is thus derived. GridSim 5.0 has been used for the proposed experimental studies.

Keywords: CPU traces of nodes, grid technology, load balancing, trust computation

INTRODUCTION

The advancements in computing and networking technologies have made it possible to share resources that are distributed geographically over the network throughout the world, known as 'Grid Technology'. It is possible to submit computational tasks to remote resources of Grids for execution through clusters of nodes offered by the Grid middleware service providers. However, these resources may be unreliable and there is a risk involved in submitting tasks that might fail or could cost more than expected, as the participating clusters in the Grid may be from individual owners. Thus 'trust' and 'reputation' of Grid participants play an important role in Grids. Trust is a property of relationships rather than individuals and accordingly, it is usually specified as a relationship factor between a trusted, such as the supplier of remote resources and a trustee, such as the user awarding jobs. Trust is a process and it can only be influenced by the final product of accumulated experience through a long-term assessment. Literature has shown that certain empirical weighting on certain parameters namely, credibility and availability of the Grid suppliers and resources respectively have been adapted.

It has been observed from literature that computational efficiency used to be one of the major issues that need to be addressed. Literature also points out that 'Load Balancing' in Grids could be a major remedial measure to tackle this issue. Load balancing is a method applied in computer networking for distributing workloads across multiple computers or say in a Grid of clusters (resources). Successful load balancing not only will optimize the resource usages, but also will maximize the throughput. This method will minimize the response time, so as to avoid any overload. Determining the trust worthiness of participating Grid clusters, for deciding performance of load balancing with several components, for the sake of reliability needs to be researched upon.

Standard Cauchy distribution is a continuous probability distribution represented through a graph, which can be used to view the acceptance levels of any functional distribution. Literature shows that this Cauchy distribution has been successfully adopted for load balance computations in Grids. Instead of the probability values, the Cauchy dense function values are themselves used for the distribution, as the study attempts to obtain only a clue for the trust worthy behaviour of the participating cluster of the Grid.

LITERATURE REVIEW

GridSim toolkit to simulate Grid environment have been successfully adapted (Buyya and Murshed, 2002) for research studies. Grid environment containing

Corresponding Author: S. Tamilselvi, Department of Computer Science, Bharathiar University, Coimbatore 641046, India

multiple resources and user entities with different requirements have been experimented and tested. It has been demonstrated that the user and broker entities could be extended via the GridSim class. In such cases, all the users can create experiments which may contain application specification of Gridlets that represent jobs with different processing and quality of service requirements, such as deadline and budget constraints with optimization strategy. GridSim 5.0 has been adopted for the experiments demonstrated in this study. CPU running time may be very important task particularly interactive application based grid, (Zhang *et al*., 2006). A new evaluation method has been presented to predict the running time of tasks in a grid. It was concluded that the prediction of task running time is based on a novel CPU load prediction method and was calculated from predictions of CPU load. For the purpose of overcoming message overload, load balancing might be added to the system (Elenin and Kitakami, 2011). The relevant algorithms could be either static or dynamic. Four algorithms for static mode had been compared and merits/demerits demonstrated. Round Robin, Random, Central Manager and Threshold algorithms were for static load balancing, while Central Queue and Local Queue algorithms were considered for dynamic load balancing. In static load balancing, the resource performance was determined at the beginning of the execution. Hence a simple model for Grid monitoring had been proposed. Response time versus message sizes in linear increase for the four algorithms had been presented and compared. Variations between the algorithms were seen large in larger message sizes. It was concluded that overloading was a big problem in Grids so balancing could be added. While static method had been proposed in this study, dynamic value had been tried out by using ant colony analogy and demonstrated by a separate work.

As the grid is dynamic in nature, heterogeneous nodes (resources) might join or leave the grid at any instance of time. In that respect it is inefficient to select a target resource that often leaves the grid (Naseera and Madhu Murthy, 2010) and hence the tasks may have to be rescheduled several times. Hence loads may have to be balanced. To perform such balancing, selection of trustworthy nodes may be essential. In grid technology, workload and resource management are two essential functions. Workloads have to be evenly scheduled for improving the throughput of the grid. Load balancing is performed to mitigate process loads from over loaded resources to under loaded resources. Benchmarks for load balancing processes have been proposed in the range levels from 1 to 10 from very untrustworthy to very trustworthy. It has been demonstrated that the reliability is inversely proportional to failure rate.

Grid analytical results deal a lot with statistical techniques. In statistical analysis, standard deviation and mean are not determined for Cauchy distributions (Jelasity *et al*., 2005). Predictions are performed in acceptable level, using Cauchy distributions. Load balancing can be performed using Cauchy and normal distribution methods. They even provide rough approximate but better solutions (Rajeswari, 2013). Literature on load balancing with 'K' means cluster selection and also on trust and reputation are seen aplenty. However, validating load balancing through Cauchy distribution in respect of 'Trust' of clusters is not to be seen in plenty. An attempt is made to validate load balancing by means of standard Cauchy distribution to testify 'Trust' of Grid clusters.

EXPERIMENTAL STUDY

The architecture of the chosen GridSim's compartment is third party developed package (GridSim5.0, 2013), which is available in public. The architecture is designed purposely for the output from different managers based on a middleware access driver. It is meant for interfacing with Grid execution services. Therefore this information manager module interfaces with the Grid information services. A Transfer Manager module interfaces with Grid data services. The scheduling process is decoupled from the Dispatch Manager through the use of an external and selectable scheduler module. This is a very important component, because to this selectable module, the researcher's input are provided. These inputs are related to Distributed Resource Management commands such as: submitting a job; killing a job; migrating a job; monitoring a job process; synchronizing jobs.

The experimental analysis consists of five jobs of capacities namely Job1: 243.34 KB; Job2: 305.45 KB; Job3: 103.56 KB; Job4: 72.3 KB; and Job5: 33.43 KB. The total job size works out to be 758.08 KB. The process time required for all the jobs put together in laboratory condition was tested and recorded to be 24361 ms. Applying the GridSim 5.0, the experiments determined the needed time for processing the overall jobs along with the actually performed processing time are presented in Table 1. Table 1 also presents the nodal details of jobs (sizes). Notations: N1, N2, N3.... denote node identities; C1, C2 and C3 represent cluster IDs. The intended experiments determined whether the clusters were trust worthy or not.

The objective of the experiments is to determine the trust worthiness of the clusters on their promised performances. The proposed process jobs (details of the jobs are presented in this study, as the study is part of a whole research work of the authors) would be fed inside the clusters of the GridSim. The difference between the actual processed time and the needed process time along with the promised resources of the Clusters would determine the worthiness. If there is a substantial mismatch between the promised and the actual, the values in percentage will be documented.

The objective of the research work is to establish the effectiveness of the performance of load balancing

Table 1: Comparison of process times of assigned jobs with Grid Sim

Processor speed performance based on trust					
Resource clusters	Jobs and required capacity of each job	Nodes and actual capacity awarded by GridSim 5.0	Needed total processor time from cluster in ms (input to GridSim 5.0)	Actual total processor time consumed by cluster in ms (Output from GridSim 5.0)	Performance in % (result will be –ve, if actual has taken more resource than expected)
C1	Job1:243.34 KB; Job2:305.45 KB; Job3:103.56 KB; Job4:72.3 KB; and Job5:33.43 KB.	N1: 548.4 KB; N2: 463.6 KB; N3: 257.9 KB; N4: 96.40 KB; and N5: 368.9 KB	24361	30420	-24.87 %
C2	Job1: 243.34 KB;Job2: 305.45 KB; Job3:103.56 KB; Job4:72.3 KB; and Job5:33.43 KB.	N6: 279.0 KB; N7: 738.0 KB; N8: 748.9 KB; N9: 46.00 KB; and N10: 36.80 KB	24361	21602	+11.32%
C3	Job1:243.34 KB; Job2:305.45 KB; Job3:103.56 KB; Job4:72.3 KB; and Job5:33.43 KB.	N11: 960.8 KB; N12: 978.2 KB; N13: 960.8 KB; N14: 960.8 KB; and N15: 928.8 KB	24361	24300	+0.25%

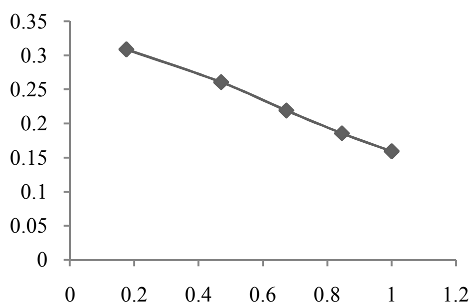

Fig. 1: Distribution of Cauchy values for cluster C1

Table 2: Distribution of Cauchy values of cluster C1 for head capacity of 548.4 KB

S. No	Node Id and capacity	x	$\frac{1}{\pi(1+x^2)}$
1	N1 548.4 KB	1.0000	0.1592
2	N2 463.6 KB	0.8000	0.1856
3	N3 257.9 KB	0.4703	0.2606
4	N4 96.40 KB	0.1758	0.3088
5	N5 368.9 KB	0.6727	0.2191

Table 3: Distribution of Cauchy values of cluster C2 for head capacity of 748.9 KB

S. No	Node Id and capacity	x	$\frac{1}{\pi(1+x^2)}$
1	N6q279.0 KB	0.3725	0.2795
2	N7 738.0 KB	0.9854	0.1615
3	N8 748.9 KB	1.0000	0.1592
4	N9 46.00 KB	0.0614	0.3171
5	N10 36.80 KB	0.0491	0.3175

only with that worthy Grid cluster. The success of the performance of load balancing with respect to the trust worthy cluster will hence prove to be the one most efficient.

Observation: It is clearly shown in Table 1 that clusters C1 and C2 have no reputation, as the performance and the promise did not tally to a great extent. While C1 was negative in its performance, cluster C2 even though performed better than what it did promised, as they did not tally well, C2 too is not trust worthy and lags reputation. As the difference is insignificant in the case of cluster C3 (only to the level of 0.25%), cluster C3 may be considered trust worthy and further experiments may be carried on with it by applying Cauchy distribution.

STANDARD CAUCHY DISTRIBUTION AND LOAD BALANCING

The Cauchy standard distribution formula is provided in Eq. (1):

$$f(x;0,1) = \frac{1}{\pi(1+x^2)} \tag{1}$$

where, 'x' represents the ratio of node(s) capacity with respect to the highest node capacity of a cluster. For example, when x becomes 0 and 1, as extreme cases, the distribution will become $1/\pi$ and $1/(2\pi)$ respectively. The experiment aims to determine only the standard distribution of the presence of node capacities in one cluster and not normal distribution. It is aimed to determine the performance based only on the distribution of Cauchy function Eq. (1). Table 2 presents the distribution of Standard Cauchy values for Cluster 1 of cluster head value of node 1, which is 548.4 KB.

The maximum cluster head has been considered for determining the Cauchy distribution according to the theory. The distribution is shown in Fig. 1. The distribution is somewhat shallow and not a straight line distribution, as can be seen from the figure. It is also seen from Table 1 that the trust value is also negative for cluster C1.

Table 3 presents the distribution of Standard Cauchy values for Cluster 2 of cluster head value of node 8, which is 748.9 KB.

Table 4: Initial value(s) distribution of Cauchy values of cluster C3 for head capacity of 978.2 KB

S. No	Node Id and capacity	x	$\frac{1}{\pi(1+x^2)}$
1	N11 960.8KB	0.9822	0.1620
2	N12 978.2 KB	1.0000	0.1592
3	N13 960.8 KB	0.9822	0.1620
4	N14 960.8 KB	0.9822	0.1620
5	N15 928.8 KB	0.9495	0.1674

Table 5: Initial value(s) distribution of Cauchy values of cluster C3 for head capacity of 878.0 KB

S. No	Node Id and Capacity	x	$\frac{1}{\pi(1+x^2)}$
1	N16 848.9 KB	0.9669	0.1645
2	N17 878.0 KB	1.0000	0.1592
3	N18 868.8 KB	0.9895	0.1608
4	N19 860.8 KB	0.9804	0.1623
5	N20 848.0 KB	0.9658	0.1647

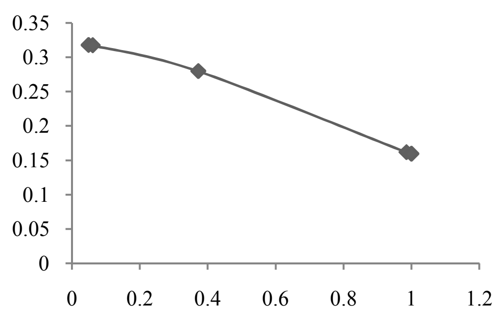

Fig. 2: Distribution of Cauchy values for cluster C2

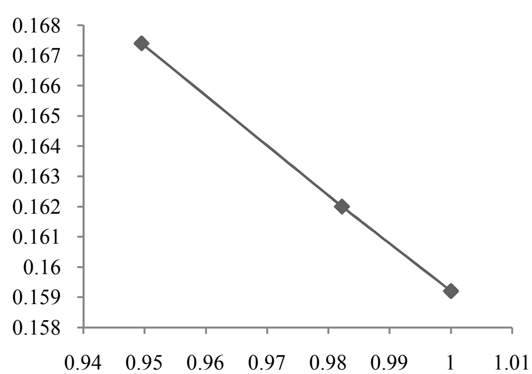

Fig. 3: Distribution of Cauchy values for cluster C3 from initial values

The maximum cluster head has been considered for determining the Cauchy distribution according to the theory. The distribution is shown in Fig. 2. This distribution is also shallow and the distribution is a curve, as can be seen from the figure. It is also seen from Table 1 that the trust value is high positive for cluster C2.

Table 4 presents the distribution of Standard Cauchy values for Cluster 3 of cluster head value of node 12, which is 978.2 KB.

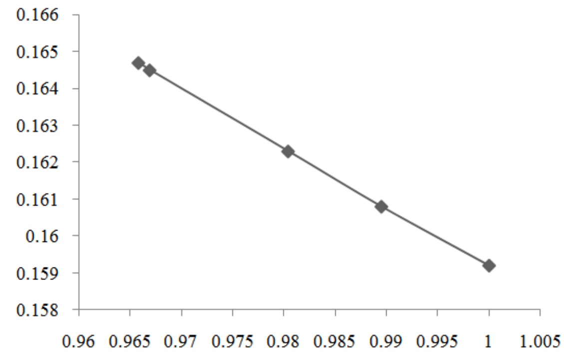

Fig. 4: Distribution of Cauchy values for cluster C3 after trust assertion

Once again the maximum cluster head has been considered for determining the Cauchy distribution according to the theory. The distribution is shown in Fig. 3. This distribution shows somewhat a steeper and the distribution is a straight line, as can be seen from the figure. It is clearly seen from Table 1 that the trust value is acceptable for cluster C3.

Comparing the three distributions, it is evidence that the trust computation is related to the standard Cauchy distribution of the third cluster. To validate the observation, GridSim 5.0 was once again executed only with the cluster 3 and results obtained and presented in Table 5.

Table 5 presents the distribution of Standard Cauchy values once again for Cluster 3 of cluster head value of node 17, which is 878.0 KB. As the second instance is different from the earlier one, new nodes namely N16 to N20 were obtained with different capacities.

A clear distinction in the style of distribution of Cauchy values in Fig. 1 along with 2.0 and the Fig. 3 is seen. While the clusters C1 and C2 are not trust worthy, Cluster C3 has been observed to be trust worthy as per Table 1. It is with this inference the load balance was performed with Cluster C3 and the results obtained and provided in Fig. 4. The reliability has been yet again tested with Fig. 3 and 4 in which both resemble the same style. It is clearly evidenced that the trust worthiness is related to the standard Cauchy distribution, after the revised results obtained from GridSim for the third cluster, as can be seen from Fig. 4.

CONCLUSION

It is clearly demonstrated that the trust computation would facilitate to reliably perform load balance computations on nodes of clusters that are trust worthy. Standard Cauchy distribution gives a clear picture to decide on the probability of distribution of performances.

REFERENCES

Buyya, R. and M. Murshed, 2002. Gridsim: A toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing. Concurr. Comp-Pract. E., 14(13-15): 1175-1220.

Elenin, S.A. and M. Kitakami, 2011. Performance analysis of static load balancing in grid. Int. J. Electr. Comput. Sci., 11(3): 57-63.

GridSim5.0, 2013. Retrieved from: www.GridWay.org.

Jelasity, M., A. Montresor and O. Babaoglu, 2005. Gossip-based aggregation in large dynamic networks. ACM T. Comput. Syst., 23(3): 219-252.

Naseera, S. and K.V. Madhu Murthy, 2010. Trust based load distribution in a data grid environment. Int. J. Comput. Sci. Eng., 2(01S): 24-29.

Rajeswari, S., 2013. A novel exhaustive criterion based load balancing algorithm for e-learning platform by data grid technologies. Int. J. Adv. Netw. Appl., 4(6): 1786-1792.

Zhang, Y., W. Sun and Y. Inoguchi, 2006. CPU load predictions on the computational Grid *. Proceeding of the 6th IEEE International Symposium on Cluster Computing and the Grid (CCGRID'2006). Singapore, May16-19, pp: 321-326.