Trust Enabled CARE Resource Broker

Kumar Rangasamy and Thamarai Selvi Somasundaram

Abstract The objective of the grid system is to allow heterogeneous geographically distributed computational resources to be shared, coordinated and utilized for solving large scale problems. Grid users are able to submit and execute tasks on remotely available grid resources. Resource management is a vital part of a grid computing system. The notion of trust is used in resource management of grid computing system. In this work, we apply behavioral trust to the problem of resource selection. Through CARE Resource Broker (CRB), we demonstrate that our behavioral trust scheduling scheme significantly maximize the throughput while maintain the high success rate of task execution. We also analyze results of our behavioral trust based scheduling and rank based scheduling.

Keywords Grid system · Resource management · Scheduling · Behavioral trust

1 Introduction

A grid system is an infrastructure capable of managing the services and resources in a distributed and heterogeneous environment [1]. The computational grids enable the sharing, selection, and aggregation of a wide variety of geographically distributed computational resources and present them as a single, unified resource for solving large-scale compute and data intensive computing applications. A resource broker is an important component of computational grid systems and acts as a

K. Rangasamy (✉)
School of Computing Sciences and Engineering, Vellore Institute of Technology, Chennai Campus, Chennai, India
e-mail: rangasamykumarme@gmail.com

T.S. Somasundaram
Madras Institute of Technology, Anna University, Chennai, India

V. Vijayakumar and V. Neelanarayanan (eds.), *Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges (ISBCC – 16')*,
Smart Innovation, Systems and Technologies 49, DOI 10.1007/978-3-319-30348-2_2

bridge for the users to access grid resources. It leverages users in analyzing the choice of selecting the suitable resource provider for their job submission.

A grid user wants to have the transparent access and negotiates for the resources through the grid resource broker which identifies and allocates the suitable resource providers for job execution. The main function of a grid resource broker is to identify and exemplify the available resources and to select and allocate the most appropriate resources for a given job. The effective allocation of distributed resources requires not only the knowledge about the capabilities of the grid resources but also the assurance that the high availability and the requested capabilities can be fulfilled for successful completion of the job [2]. The resource discovery is one of the biggest challenges in the grid environment. Although, grid resource broker identifies the resource provider according to the suitability of the requested task and it is important to analyze the resource provider and rate them according to the quality of service rendered by them in their earlier performance. The quality of service may be determined based on the past behavior of the resource providers. The matching criteria of a resource provider in terms of credibility, computation power, memory size, resource performance, etc., for a given set of requirements also required for determining the good resource provider. The estimation of the belief on the behavior and reputation results in determining the trust values of the resource providers.

The trust values of the resource providers may be used for grading and ranking them to achieve better quality of service in the grid environment. In this paper, we propose a system that evaluates the trust of a resource provider by obtained by the trust metrics such as credibility and availability. The proposed Trust Management System is generic one and can be easily integrated with any metascheduler(s)/ resource broker(s). The proposed system is integrated with the CARE resource broker to evaluate and update the trust value of resource providers. In our pervious work [3], we have considered a simple averaging scheme for the calculation of trust value. In brief, the contributions of this research work are summarized below.

- We propose a trust model to evaluate the trustworthiness of resource provider in the computational grid infrastructure.
- We derive a novel mathematical model for credibility and availability as termed as Trust Scheduling Function (TSF). The credibility is the combinational value of direct and indirect satisfaction degree. The direct satisfaction degree obtained through the job success rate and the indirect satisfaction degree obtained by the feedbacks from the users. We have applied probabilistic approach based on Bayesian inferences. The availability is derived by resource performance and resource busy degree. The computed trust value can be used as input for the grid resource broker to select the good resource provider for the job execution.
- We propose a mechanism to verify the feedbacks from the users after the utilization of resource providers.
- The proposed trust model has been integrated with CARE resource broker. The impact of trust based scheduling versus rank based scheduling has been analyzed.

The rest of the paper is organized as follows: In Sect. 2 gives the definitions of trust. In Sect. 3, presents the related work. In Sect. 4, we discuss about various types of trust in brief. We derive a mathematical formula to compute overall trust value is explained in Sect. 5. The various components of Trust Management System are explained in Sect. 6. The experimental setup made in our research laboratory and the results are explained in Sect. 7. Finally, in Sect. 8, we conclude our research work and outline the future work.

2 Trust Definitions

Several researchers have proposed various definitions for trust. One of the earlier definitions of trust given by Gambetta [4] states “Trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently or his capacity ever to be able to monitor it) and in a context in which it affects his own action”. This definition reflects that trust is measure of a belief which paves the way for many researchers in exploring the computation of trust in a subjective manner. Castelfranchi and Falcone [5] extended Gambetta’s definition that the trustor should have a “theory of mind” of the trustee, which included the estimations of risk and a decision on whether to rely on the other based on the trustor’s risk acceptability.

Grandison and Sloman [6] surveyed various definitions of trust. Their definition states “the firm belief in the competence of an entity to act dependably, securely and reliably within a specified context”. They define that trust is a combined effect of various attributes which have been considered while defining the trust of a system.

Josang et al. [7] define trust as “the extent to which one party is willing to depend on something or somebody in a given situation with a feeling of a relative security, even though negative consequences are possible”. Their definition describes that the trust of an entity depends on the reliability of what other entities observe and also rely that the positive utility results in the good outcome and there is a possibility of risk occurrence by the relative previous entity.

In general, various trust definitions have been proposed by many researchers based on the context of the relationship which establishes among grid entities. In a computational grid environment, we focus on the establishment of trust of a resource provider. Hence, our definition of trust of resource provider states “Trust is a prediction of reliance on the ability and competence of a resource provider based on the commitment of quality of services measured within the specified context of computational grid”.

3 Related Work

Trust concept has been addressed at different levels by many researchers. Trust can be broadly classified into Identity trust and Behavioral trust, where Identity trust is concerned with the authentication and authorization of an entity and Behavioral trust deals with entity's trustworthiness based on good or bad performance. This work is focused on evaluating behavioral trust of resource providers in a computational grid.

The notion of "trust management" was introduced by Blaze et al. [8] in their seminal paper. In the computer science literature, Marsh [9] proposed a computational model for trust in the distributed artificial intelligence (DAI) community. Marsh proposes a trust model takes into account direction. It categorized into three types of trust such as Basic trust, General trust and Situation trust.

Abdul-Rahman and Hailes [10] proposed a model for computing the trust for an agent based on the experience and recommendation in a specific context. Trust values are categorized into very trustworthy, trustworthy, untrustworthy and very untrustworthy. Each agent also stores the recommender trust with respect to another agent. The recommender trust values are semantic distances applied for adjusting the recommendation in order to obtain a trust value. They propose a method for evaluating and combining recommendations and updating the trust value. As the model is based on the set theory, each agent has to store all history of past experiences and received recommendations. On a system with a lot of participants and frequent transactions, each agent should have a large storage with this respect. Regarding the network traffic, this is caused by the messages exchanged between agents in order to get reputation information. The authors provide an example of applying the reputation management scheme, but no computational analysis is provided. However, we have given enough computational analysis based on the proposed mathematical model. Moreover, their approach is entirely based on subjective analysis. The researchers [11–14] have attempted to explore the subjective nature of trust metrics. Further, they focus on the reputation or recommendation that depends entirely on the opinion of the entities involved.

Azzedin and Maheswaran [15], view the trust at two different levels viz., direct and overall trust. In this model, the overall trust represents the reputation value earned by the resource provider. This model rates resource provider through various levels A through F, with the assumption that A being the highest trust value and F as the lowest value. It doesn't provide any confined values for those levels. The overall trust value earned by an entity is subjective in nature, depends on the user's view or opinion. Further, this model lacks the truthful assurance of the user's feedback which may lead to decrease or increase the trust level of those resources which they had experienced. In our approach, we concentrate on object nature of trust metrics. They have applied discrete approach in calculating the value of trust. We have applied probabilistic density function to calculate the trust value of the resource providers.

Kamvar et al. [16] designed a Eigen algorithm for P2P system and it is based on the transitive trust. This model has the drawback of storing each other peers trust value locally. We have used trust database to store the trust values of each provider in the computing grid infrastructure. In this way, our system reduces the burden storing trust values locally. However, it does not suggest any method for obtaining these trust values. This model also lacks the dynamic collection of the feedback system. In our work, feedback aggregation is done after the verification of positive feedbacks and negative backs. During the feedback verification, if the feedbacks are found to be faked ones, our system automatically discards it. von Laszewski et al. [17] exploits the beneficial properties of EigenTrust. This work extends the model to allow its usage in grids. They integrate the Trust Management System as part of the QoS management framework, proposing to probabilistically pre-select the resources based on their likelihood to deliver the requested capability and capacity. The authors present the design of the system, but they don't present the experiments in order to prove efficiency of the approach.

Aberer and Despotovic [18] referred to P2P networks, because it fully employs the referral network as being a source for obtaining recommendations. They assume the existence of two interaction contexts: a direct relationship where a destination node performs a task and recommendations when the destination node acts as a recommender of other nodes in the network. Rather than considering the standard P2P architecture, the graph of nodes is built by linking peers who have one the above-mentioned relationships. The standard models usual weight a recommendation by the trustworthiness of the recommender. Instead, they model the ability of a peer to make recommendations, which is different from the peer trustworthiness.

Chen et al. [19] proposed a model for selection and allocation of grid resources based on Trust Management System scheme. The authors have considered trust metrics such as affordability, success rate and bandwidth for the calculation of overall trust value. There is no substantial mathematical derivation has given for the calculation of trust value. The authors have followed the simple averaging scheme for their computation of trust value.

Dessi et al. [20] introduced a concept of Virtual Breeding Environment (VBE). This system has three types of operative contexts namely user operative context, resource operative context and organization operative context. Under these contexts, many trust metrics have been proposed. However, their work lacks in mathematical derivation of each and every metrics. Aforementioned trust metrics are subjective in nature. For instance, the resources are grouped into five reputation levels. The quantification of trust by adopting subjective metrics does not give much accuracy compared to objective metrics that we have adopted in our research work. There is no classification of direct and indirect trust metrics are obtained from the resource providers and the users. We have classified the trust metrics into direct and indirect one. The direct trust metrics are success rate and failure rate of the resource providers. The indirect metrics such as positive feedbacks and negative feedbacks. Moreover, the feedbacks are cross checked by our system.

An effective voting based reputation system has proposed by Marti and Garcia-Molina [21] to facilitate the judicious selection of P2P resources. The major concern in such a system is to isolate the adverse effects of a malicious peer. The simple voting scheme was found to be quite effective towards achieving this goal. Similarly, Damiani et al. [22] also suggested a reputation sharing system to enable reliable usage of remote peer's resources. Their system works by using a distributed polling algorithm. Our system achieves an effective resource selection and allocation of job execution by considering the past behaviours of resource providers.

Sabater and Sierra [23] proposed a model, named REGRET that considers three dimensions of the reputation models such as the individual dimension: which is the direct trust obtained by previous experience with another agent and the social dimension which refers to the trust of an agent in relation with a group and the ontological dimension which reflects the subjective particularities of an individual. In [24] review some works regarding reputation as a method for creating trust from the agent related perspective. They do not categorize the described models, but they try to find how those models are related with some theoretical requirement properties for reputation. The aforementioned works are related to subjective in nature. The quantification of overall trust value is more complex.

Karaoglanoglou and Karatza [25] presented a trust-aware resource discovery mechanism that guarantees satisfying requests with a high value of trustworthiness and in the minimum distance of hops in a Grid system. It is not clear that how trust is calculated for each Virtual Organization (VO) in the Grid system. The authors simply assume random values for all VOs. There is no substantial mathematical background for the calculation of overall trust of VOs. Further, the proposed model is not integrated any metascheduler/resource broker. In [26], the authors have proposed the social-network based reputation ranking algorithm for the peer to peer environment. It is capable of inferring while the calculation of indirect trust ranks of the peers more accurately. Further, the effective measures are still missing in their design and implementation. In [3], the authors mainly considered three trust metrics such as affordability, success rate and bandwidth of the resource providers. The overall trust value can be computed by the simple approach.

In this research work, we propose a trust model that aimed to explore the possibility of obtaining the trust value in an objective fashion and the same has been integrated with CARE resource broker. We compute the reputation of resource providers by means of resource performance and quality of service. We consider the user's opinion in terms of positive and negative feedbacks. The feedbacks are collected from the user after the usage of a particular resource provider can be quantified into a numerical value. We propose a mathematical model to evaluate the trust value of the resource provider. The trust values for each resource provider will be stored in the database for the future use. Each resource provider has its unique resource identity and associated trust value. We have considered two main factors, one reflecting the entity's past experience and the other reflecting the capabilities of resource provider.

4 Types of Trust for Computational Grid

The resource broker is responsible for identifying the suitable resource provider meeting the customer's requirement. It is also responsible for mediating the resource provider and users during grid transactions. Trust in the resource provider's competence, honesty, availability, success rate and reputation will influence the consumer's decision for accessing them. The trust models are categorized into four types such as Service provision trust, Behavioral trust, Identity trust and Reputation trust.

4.1 Service Provision Trust

Service provision trust describes the relying party's trust in service. The trustor trusts the trustee to provide a service that does not involve access to the trustor's resources. This type of trust reflects the resource provider's capability in terms of its computational and connectivity power. The grid resources with a higher computational power are expected to process the task with a reduced amount of execution time with respect to others. The trust metrics are considered in service provision trust such as CPU speed, size of memory, network bandwidth, latency and utilization of CPU, etc.

4.2 Behavioral Trust

Behavioral trust is a measurable trust by the resource broker by its experience and interactions over the resource providers in the grid. It measures the consistency of any grid entity over a period of time in the grid environment. It also helps in the determination of trust acquired by an entity which helps to predict the behavior of that entity in near future. A few examples of behavioral trust metrics are availability, success rate and network speed.

4.3 Identity Trust

Identity trust represents the entity's trust depending on their form of identification to the grid environment. This trust focuses more on the authentication and authorization system which has been adopted in the grid environment. This trust helps to classify the entity's security level. A few examples of identity trust metrics are certificate authority, authentication mechanism, authorization policies.

4.4 Reputation Trust

Reputation trust reflects the trust of a grid entity over a certain period of time based on the remarks made by the grid users. This trust helps to identify the experience of a grid entity directly or indirectly. Reputation based trust is subjective in nature and the value of trust vary from individual to individual. A few examples of reputation trust metrics are feedback, recommendation, grading based on the varieties of jobs handled, etc.

5 Mathematical Model for the Proposed Trust Management System

In this section, we propose a novel trust model of estimating the trustworthiness of the resource provider in the grid infrastructure. Our model has advantages by considering the objective nature of the trust in the computation of trust value. There are various possibilities of failures in the grid environment. The success rate, failure rate and system availability of resource providers are taken into consideration while calculating the overall trust value. Hence, the trustworthiness of the resource provider is the combined effect of the behavioral of current status and the historical data meaning that how it was performed in the past and present. In the following section, we explain the formation of a mathematical model for the proposed system.

Let us take the following scenario. Consider R is a problem, all the possible solutions are representing a set S, set S is called the solution set of the problem. Let us consider $S = (T, -T)$ where T represents a grid node to another grid node of a resource provider is not only credible but also available and $-T$ represents a node set which a grid node to another grid node is not entirely credible and available, is a set of grid nodes in line with the requirements of solutions. For instance, a grid node has very high credibility, but not available, or a node has high availability, but it is not credible, such nodes do not meet the requirement nodes. For the credibility and availability of the grid node, we formulate a trust scheduling function $TSF(i)$ to evaluate the trustworthiness of a resource provider. This scheduling function is evaluated via two functions namely credibility function $C(i)$ and availability function $AVL(i)$. $TSF(i)$ is defined as formula (1).

$$TSF(i) = \alpha \times C(i) + \beta \times AVL(i) \tag{1}$$

Here α and β are the weights of credibility and availability, sum of α and β is 1. We have assigned equal weights for credibility and availability. If there is a low requirement level for credibility, then we assign the lower value for α and higher value for β. If there is higher requirement of the credibility is needed, then we assign the bigger value for α and lower value for β.

5.1 Credibility

The credibility is an assessment of grid node status and behaviors characterized by the credibility value. This is to make each others satisfaction degree evaluation of the trusted status after interacting information between nodes. We define a double $C(i) = (D(i), I(i))$ to indicate the credibility of the node. Here $D(i)$ is the direct satisfaction degree and $I(i)$ is the indirect satisfaction degree. $C(i)$ is defined as formula (2).

$$C(i) = \{w_1 \times D(i) + w_2 \times I(i)\} \tag{2}$$

Here, w_1 and w_2 are the weights of direct satisfaction degree and indirect satisfaction degree, the sum of w_1 and w_2 is 1. In general, we have more confidence for the direct satisfaction degree and less confidence for the indirect satisfaction degree. Therefore, we usually set the value of w_1 is larger and set the smaller value for w_2.

5.2 Direct Trust Satisfaction

It is the direct evaluation of trusted status of the other node after interacting with each other and comes from the historical record of information. This information is obtained from the job success rate of the resource provider. Direct satisfaction is a kind of reliable information in the derivation of trustworthiness value. The calculation of direct satisfaction degree is done by Bayesian inferences. Two metrics are used in beta distribution to represent the observations are chosen n_s as the number of previous satisfying interactions and n_u as the number of unsatisfying interactions. While computing the values of n_s and n_u, it is assumed that the desired type future interaction is identical to that of previous interactions. By setting $x = n_s + 1$ and $y = n_u + 1$ the estimated value of D(i) is obtained by the expected value of the probability distribution function of the beta distribution.

$$D(i) = E(f(a; x, y) = \frac{x}{x+y} = \frac{n_s + 1}{(n_s + 1) + (n_u + 1)} = \frac{n_s + 1}{n_s + n_u + 2} \tag{3}$$

The idea of adding 1 each to n_s and n_u (thus 2 to $n_s + n_u$) in formula (3) is that it follows the Laplace's *famous rule of succession* for applying probability to decision making from the history of values.

5.3 Indirect Trust Satisfaction

It is an indirect evaluation of trusted status of the grid node is performed by collecting the feedback from others. The feedback from others could help to find out the quality of the site even without direct interactions. However, feedback is not reliable source of information. By getting feedbacks from the grid users, we compute the indirect trust value of each site. The feedback evaluation is done by the following mechanism. After the usage of the grid resource, the feedback can be collected from the users through web forms. The feedbacks are classified by the user QoS parameter such as a deadline and the recommendation for the future.

The user can give an answer to the question as follows after usage of each grid node.

- Whether the job has been finished within the deadline or not?
- Do you recommend to use this node in the future for the other users?

It is obvious that the users could provide both positive feedbacks and negative feedbacks about the grid resource providers. These feedbacks are cross checked and then aggregated. Here P denotes the positive feedbacks and N denotes the negative feedbacks. The positive feedbacks and negative feedbacks are checked by the following scenario. The grid job is through the job submission template to the various resource providers available in the grid infrastructure. The job template consists of job requirements such as CPU, memory, disk size, deadline and etc. After every job execution, the user has to provide the feedback about the resources. Through the user, feedbacks are compared with the scheduler generated execution time of the job.

For instance, the user request 30 min as a deadline parameter to run the job on the resource. If the job is finished on or before 30 min, but the user gives the negative feedback about the resource. We consider this feedback is faked one. So we consider this as a negative feedback. Likewise, user has to provide the feedbacks for recommendation metric. It is possible that the resource provider performs well, but the user can give negative feedback. The grid metascheduler has logs at the end of every job execution. Through the system generated logs our trust model could verify the recommendation parameter and aggregate the negative and positive feedbacks. The mechanism for verifying the feedbacks are hidden to the users of the grid system. After the feedback verification, we categorize the feedbacks and then aggregate using the following formula for the calculation of overall trust.

By setting $x = n_P + 1$ and $y = n_N + 1$ the estimated value of I(i) is obtained by the expected value of the probability distribution function of the beta distribution.

$$I(i) = E(f(a; x, y) = \frac{x}{x+y} = \frac{n_P + 1}{(n_P + 1) + (n_N + 1)} = \frac{n_P + 1}{n_P + n_N + 2} \quad (4)$$

5.4 Resource Availability

Resource availability function $AVL(i)$ is defined as the characteristics of resource performance and resource busy degree characterization. It is decided by the resource performance RP(i), resource busy degree RB(i).

The resource performance $RP(i)$ is defined as a formula (5).

$$RP(i) = \sum_{j=1}^{n} (R_{ij} \times p_j) \quad p_j \in (0, 1) \tag{5}$$

Here $RP(i)$ is the resource performance number, i is the grid node, R_{ij} is the number, j is the resource attribute number, i grid node, such as a number of CPUs per core, CPU frequency, memory size, hard disk size, network speed, etc., p_j is the weight number, j is the resource attribute. It is a static value and there is no need to update periodically.

The degree of a resource busy is defined as a formula (6) below.

$$RB(i) = \sum_{j=1}^{n} ((1 - \frac{U_{ij}}{T_{ij}}) \times q_j) \quad q_j \in (0, 1) \tag{6}$$

Here, $RB(i)$ is the busy degree of a number, U_{ij} is the usage number, j is the category resource number, i grid node, such as a CPU, memory, hard disk size, network usage, etc., T_{ij} is the total amount number, q_j is the weight of the resource usage number. The value of $B(i)$ is dynamic in nature. So, this value can be updated periodically to the resource broker in order to take the scheduling decision. Therefore, the resource availability function $AVL(i)$ is the combination of resource performance value and resource busy value.

$$AVL(i) = RP(i) + RB(i) \tag{7}$$

6 Trust Management System

In our earlier work [3, 27], we have proposed a generic life cycle of Trust Management System and its various phases involved in establishing trust across the grid resources. Figure 1, present the proposed Trust Management System has been integrated with CARE resource broker.

The sub components of the Trust Management System are described in this section. The Trust Metrics Identifier represents the trust metrics in a simpler form for the ease of trust value computation. The two main components of Trust Metrics Identifier are Trust Metric Information Collector and Trust Metric Indicator. The Trust Metrics Information Collector (TMIC) retrieves the basic information about trust metrics of every resource provider. It obtains the information from the grid resource broker or the grid middleware depending upon the trust metrics needed. The various metrics which is obtained by TMIC is availability status, job status, CPU and network information of the resource provider.

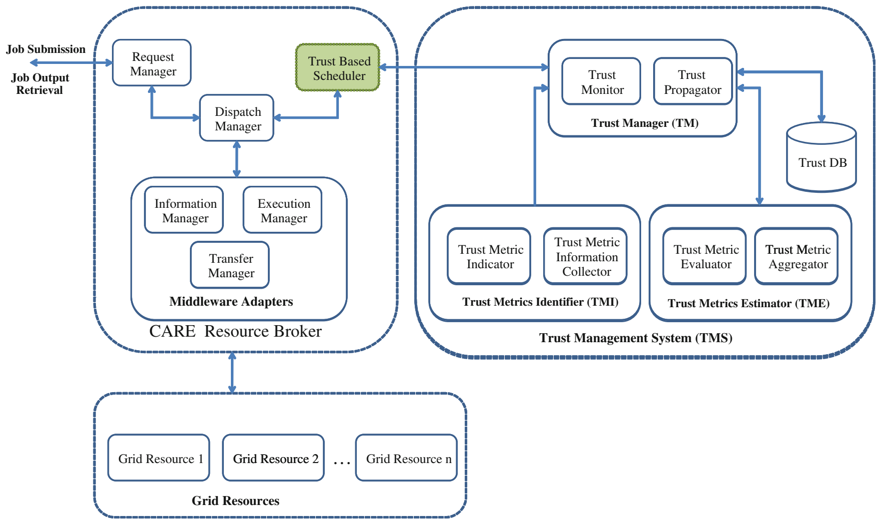

Fig. 1 Architecture of trust management system

Trust Metric Indicator provides the representation of the trust metrics which have been obtained through TMIC. It converts each trust metric into a suitable format for the computational aspect.

The Trust Estimator processes the trust metrics and applies a suitable mathematical formula for the computation of trust value of a resource provider. The two main components of Trust Estimator are Trust Metric Evaluator (TME) and Trust Metric Aggregator (TMA). The Trust Metric Evaluator computes the value of each trust metric considered for the resource provider. It provides the conversion of the trust metrics which is given by the TMI and represents the value of computing resources with respect to each trust metrics. The Trust Metrics Aggregator performs the computation of overall trust of computing resources. It processes each trust metrics, which is provided by the TME and applies a suitable mathematical model to compute the trust value.

The Trust Manager maintains the trust values computed by the Trust Estimator (TE). This maintenance of trust helps in evaluating the past experience of any resource provider in a grid environment. The two main components of Trust Manager are Trust Monitor and Trust Propagator. The Trust Monitor updates the trust value of any computing resources after each job execution. It also keeps logging information of every trust metrics computation. The Trust Propagator communicates with the grid metascheduler. The Trust Propagator retrieves the trust value of the resource provider from the Trust DB, if requested by the metascheduler. The metascheduler can utilize the trust value which is calculated by the Trust Management System and can schedule the jobs accordingly. The Trust DB holds the trust value of all resource providers that are available in the grid environment. It is also responsible for the Trust maintenance of every resource after each task/job completion.

6.1 CARE Resource Broker (CRB)

CARE resource broker [28] is a meta-scheduler that has been developed at the Center for Advanced computing Research Education Laboratory, Madras Institute of Technology. The motivation behind this work is that the application scheduling in grid is a complex task that often fails due to non availability of resources and required execution environment in the resources. The CARE resource broker addresses several scheduling scenarios using the concepts of virtualization. The Virtualization technology offers effective resource management mechanisms such as isolated, secure job scheduling, and utilization of computing resources to the possible extent. However, lack of protocols and services to support virtualization technology in high level grid architecture does not allow management of virtual machines and virtual clusters in grid environment. CRB proposes and implements necessary protocol and services to support creation and management of virtual resources in the physical hosts. Besides, CRB supports Semantic component, Resource leasing and Service Level Agreement (SLA). In addition to the aforementioned features of CRB, the proposed trust model that in assists to select good service provider for the reliable job execution. The current form of CARE resource broker provides more features but it lacks in selecting most reliable resource for the job execution. For this reason, we propose and model a novel trust scheme in addition to the aforementioned features.

The conventional CRB is working based on the ranking scheme. The computation of rank is done by considering the CPU and memory of the grid resources. Trust based scheduler makes scheduling decisions for jobs on the available grid resources. The grid jobs can be submitted to Care Resource Broker (CRB) through the job submission portal. The Request Manager component of CRB does the matchmaking process for the submitted jobs against the available resources in the specific interval. By default, the CRB is working based on the conventional (rank) scheme. The new feature is introduced in CRB is trust based scheme. The function of Dispatch Manager is invokes the rank or trust scheme is based on the configuration of CRB. If the trust scheme is configured, then the CRB calls the trust based scheduling function. The trust calculation is done by considering both direct and indirect experience. Trust based scheduler is working based on system capability, system performance, direct trust interaction and indirect trust interaction. The algorithm for trust based scheduling is given below.

Algorithm 1 Job submission and Trust based resource scheduling

Step 1 Start the CARE resource broker and Oracle 10g Database and initialize the Number of Jobs (NJ), Trust Scheduling Function (TSF), Credibility (C), Availability (AVL), Direct Trust (DT), In-Direct Trust (IDT), Resource Performance (RP), Resource Busy (RB).

Step 2 Submit 'N' numbers of job requests to CARE resource broker.

Step 3 The Dispatch Manager invokes the Trust Based Scheduler to select the trustworthy resources for job submission.

Step 4 The Trust Based Scheduler invokes the Trust Management System for trust computation.

Step 5 The Trust Management System computes the trust value based on the following trust metrics such as success rate, feedback and availability using the formula (1), (2) & (3) and computed Trust resource list has been sent to Trust Based Scheduler.

```
For (i=1; i<NJ.Size; i++) {
    For (j=1; j< Matchedresourcelist.Size; j++)
     {
        TSF(j)=C(j) +AVL(j);                          - Formula (1)
        C(j)= DT(j)+IDT(j);                           - Formula (2)
        AVL(j)=RP(j)+RB(j);                           - Formula (3)
       Trustresourcelistcount++;
       Trustresourcelist[j].add (Matchedresourcelist[j]);
     }
}
```

Step 6 The Trust Based Scheduler selects the most trustworthy resource from trust resource list and it is sent to Dispatch Manager.

Step 7 The Dispatch Manager invokes the Transfer Manager to transfer the executable and input files to selected trustworthy resource.

Step 8 Once the job execution is completed successfully, Dispatch Manager invokes the Transfer Manager to transfer the output files to broker.

Step 9 Then Dispatch Manager updates the job status, which is "Success" or "Failure" in Trust DB through Trust Management System.

Step 10 End

7 Results and Discussion

7.1 Experimental Setup

Figure 2 shows that the experimental testbed is done in our research laboratory for testing the proposed work in real world scenario. The test-bed consists of the Trust Resource Broker named Gridtrustbroker.mit.in and three cluster resources namely Xencluster.care.mit.in, Smscluster.care.mit.in and Centcluster.care.mit.in. This setup is managed by the grid middleware of Globus Toolkit (GT) 4.0.1, Torque-2.0.1 and Sun Grid Engine (SGE) 2.3.0 as Local Resource Manager (LRM), Ganglia 3.0.2 as resource monitoring tool, Network Weather Service (NWS) 2.13 as network monitoring tool and Oracle 10 g as a database for the trust repository. The purpose of using database is reduce the overhead and to store the historical information about the grid resources for a very large environment.

The Xen hypervisor is also installed in the environment because our CARE resource broker supports for the creation of virtual nodes, if there is a shortage of physical nodes. We have submitted and tested grid jobs in this testbed. The setup produces the fruitful results. We have tested 50 grid jobs in this environment initially. In the rank based scheme which is used in gridway metascheduler, the success ratio is 76 %. In trust based scheme, the success ratio is 85 %. Obviously, trust based scheduling is producing better results.

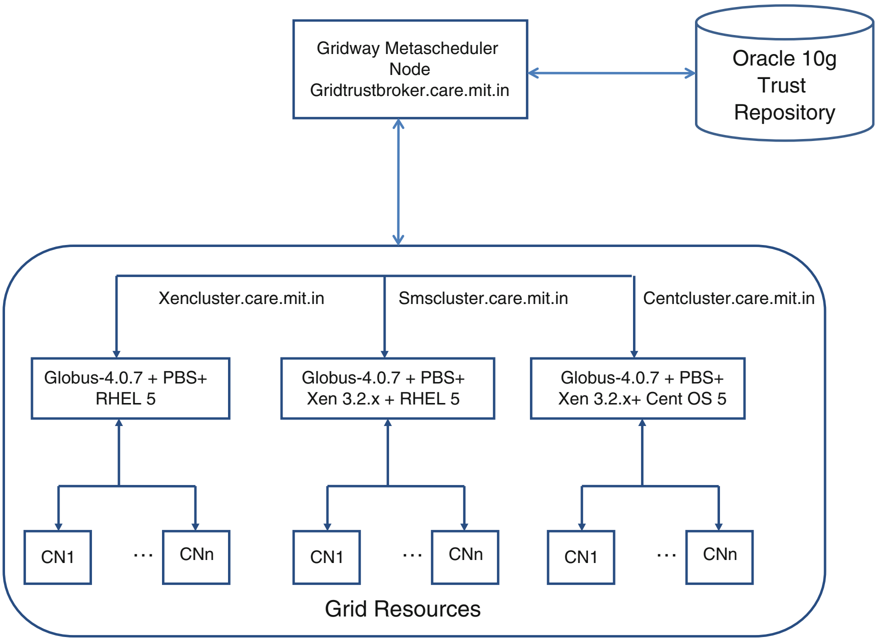

Fig. 2 Experimental setup

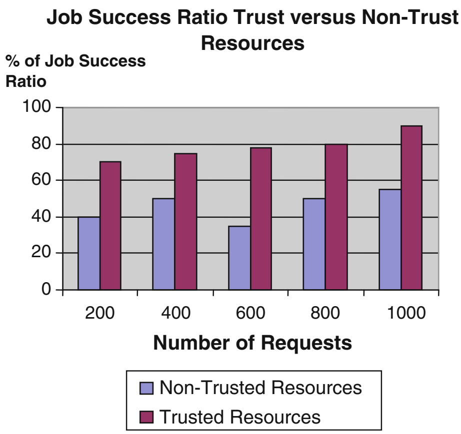

Fig. 3 Job success ratio

7.2 Simulation Experimental Results and Inferences

The proposed trust model has been simulated using the java based simulator code using a sample of 200, 400, 600, 800, 1000 requests and 200 nodes of grid resources and the requests (jobs) have been submitted to the trust enabled Gridway metascheduler node. The above requests have been tested both with trust and non-trust based model. The percentage of requests handled successfully with respect to the submitted requests is plotted as shown in Fig. 3. The proposed trust model increases the job success ratio, user satisfaction, and utilization of grid resources. The job success ratio in case of trusted resources is gradually increasing but in case of non-trusted resources, the job success ratio shows inconsistent.

The user's satisfaction level is plotted based on the feedback given by the user. The user's satisfaction increases for the trustworthy resources over a period of time whereas the satisfaction level is fluctuating and it is unpredictable for non-trusted resources as shown in Fig. 4.

The resource utilization of the trusted versus non-trusted resources is plotted as shown in Fig. 5. The X-axis represents the number of requests and Y-axis represents the utilization of the resources. The graph is plotted with the sample of 200, 400, 600, 800, 1000 and 1200 requests. The resource utilization of the trusted resources increases with the number of requests and the non-trusted resources utilized in a constant rate.

7.3 Trust Based Scheduling Versus Rank Based Scheduling Algorithm

The proposed Trust Management System has been integrated with the Gridway metascheduler for analyzing the impact of using trust in the selection of the

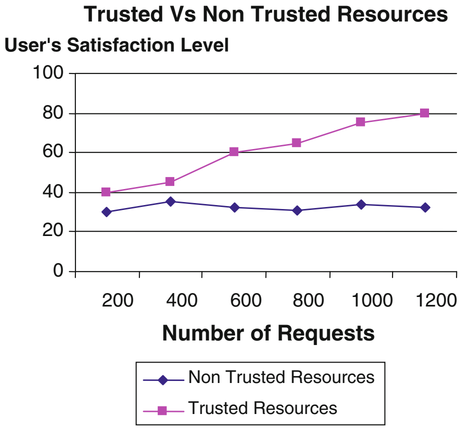

Fig. 4 Level of user's satisfaction

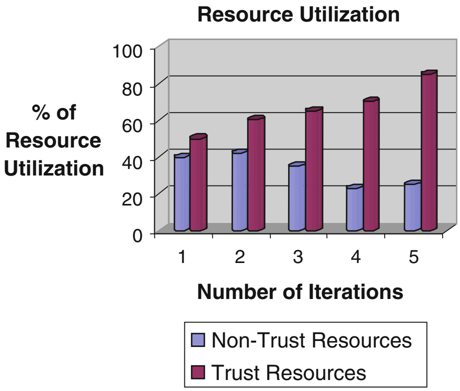

Fig. 5 Resource utilization

computational resource for the job execution. The Gridway metascheduler enables large scale reliable and efficient sharing of computing resources managed by different LRMS systems. However, Gridway approach follows the resource selection based on the rank mechanism.

The integration of our trust module with Gridway metascheduler is given in the following link: http://www.gridway.org/doku.php?id=ecosystem. In analyzing the impact of the resource selection in the grid environment, a simulator is developed to evaluate the performance of trust based scheduling algorithm with the rank based algorithm. In Fig. 6 represents the comparison of trust based scheduling and rank based scheduling. By default, Gridway uses the rank based scheduling algorithm. The rank based algorithm finds the ranking of computing resources by using the computational capability such as CPU and free RAM available.

The proposed trust model not only considers the computational capability and also takes the past behavior into an account while calculating overall trust. Our trust based scheduling algorithm has been tested using the same resources considered in rank based scheduling. Figure 6 shows that the trust based scheduling scheme

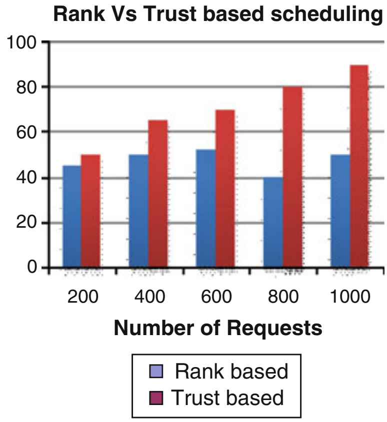

Fig. 6 Rank based scheduling versus trust based scheduling

outperforms the ranking mechanism used by the Gridway metascheduler. These results states that the resources are utilized more effectively compare to rank based scheduling. In this way, in a computational grid environment, the throughput of the resources can be maximized by using our approach. The comparison between these two schemes is represented below.

8 Conclusion

In this work, we present the trust enabled CARE by considering the metrics such as direct trust, indirect trust and system availability. The proposed trust model is proficient in choosing reliable, most trustworthy resources which are part of grid environment. The computation of trust value of each resource provider is done through the mathematical model and the measured value of trust is objective in nature. We have tested our trust model with CARE resource broker and observed that the trust based scheduling enables the identification of resource providers yielding increased throughput. We also analyze results of our behavioral trust based scheduling and rank based scheduling. Our trust model is a generic one and the same can be easily integrated with other grid metascheduler(s)/resource broker(s), thus improving the effective resource management of grid infrastructure.

References

1. Arenas, A., Wilson, M., Matthews, B.: On trust management in grids. In: Proceedings of the 1st International Conference on Autonomic Computing and Communication Systems (2007)
2. Zhao, S., Lo, V., Gauthier-Dickey, C.: Result verification and trust-based scheduling in peer-to-peer grids. In: Proceedings of the Fifth IEEE International Conference on Peer-to-Peer Computing (2005)

3. Thamarai Selvi, S., Balakrishnan, P., Kumar, R., Rajdendar, K.: Trust based grid scheduling algorithm for commercial grids, vol. 1, pp. 545–5581, ICCIMA (2007)
4. Gambetta, D.: Can we trust trust? In: Gambetta, D. (ed.) Trust: Making and Breaking Cooperative Relations, pp. 213–237 (1988)
5. Castelfranchi, C., Falcone, R.: Principles of trust for MAS: cognitive anatomy, social importance and quantification. In: Proceedings of the Third International Conference on Multi-agent Systems. IEEE C.S., Los Alamitos (1998)
6. Grandison, T., Sloman M.: Specifying and analysing trust for internet applications. In: The Second IFIP Conference on E-Commerce, E-Business, E-Government (I3E), Lisbon, Portugal (2002)
7. Josang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decis. Support Syst. **43**, 618–644 (2005)
8. Blaze, M., Feigenbaum, J., Lacy, J.: Decentralized trust management. In: IEEE Conference on Security and Privacy (1996)
9. Marsh, S.: Formalizing trust as a computational concept. PhD thesis, University of Stirling (1994)
10. Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: HICSS '00: Proceedings of the 33rd Hawaii International Conference on System Sciences, vol. 6, p. 6007, Washington, DC, USA, IEEE Computer Society (2000)
11. Huynh, D., Jennings, R., Shadbolt, R.: Developing an integrated trust and reputation model for open multi-agent system (2004)
12. Dingledine, R., Mathewson, N., Syverson, P.: Reputation in P2P anonymity systems. In: Proceedings of the First Workshop on Economics of P2P systems (2003)
13. Zhou, R., Hwang, K.: Trust overlay networks for global reputation aggregation in P2P grid computing. In: IEEE International Parallel and Distributed Processing Symposium (IPDPS-2006), Rhodes Island, Grace (2006)
14. Resnick, P., Zeckhauser, R., Friedman, E., Kuwabara, K.: Reputation systems. Commun. ACM **43**(12), 45–48 (2001)
15. Azzedin, F., Maheswaran, M.: Integrating trust in grid computing systems. In: Proceedings of the International Conference on Parallel Processing (2002)
16. Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H.: The Eigen Trust algorithm for reputation management in P2P networks (2004)
17. von Laszewski, G., Alunkal, B.E., Veljkovic, I.: Towards reputable grids. Scalable Comput. Pract. Experience **6**(3), 95–106 (2005)
18. Aberer, K., Despotovic, Z.: Managing trust in a peer-2-peer information system. In: CIKM '01: Proceedings of the Tenth International Conference on Information and Knowledge Management. pp. 310–317, New York, USA, ACM Press (2001)
19. Chen, C., Li-ze, G., Xin-xin, N., Yi-xian, Y.: An approach for resource selection and allocation in grid based on trust management system. In: First International Conference on Future Information Networks (2009)
20. Dessi, N., Pes, B., Fugini, M.G.: A distributed trust and reputation framework for scientific grids. In: Third International Conference on Research Challenges in Information Science (2009)
21. Marti, S., Garcia-Molina, H.: Limited reputation sharing in P2P systems. In: EC '04 Proceedings of the 5th ACM Conference on Electronic Commerce, New York, USA
22. Damiani, E., Vimercati, D.C., Paraboschi, S., Samarati, P., Violante, F.: A reputation-based approach for choosing reliable resources in peer-to-peer networks. In: 9th ACM Conference on Computer, ACM Press (2002)
23. Sabater, J., Sierra, C.: Regret: a reputation model for gregarious societies. In: Fourth Workshop on Deception, Fraud and Trust in Agent Societies, ACM Press (2001)
24. Sabater, J., Sierra, C.: Review on computational trust and reputation models. Artif. Intell. Revision **24**(1), 33–60 (2005)

25. Karaoglanoglou, K., Karatza, H.: Resource discovery in a grid system: directing requests to trustworthy virtual organizations based on global trust values. J. Syst. Softw. **84**(3), 465–478 (2011)
26. Wang, Y.F., Nakao, A.: Poisonedwater: an improved approach for accurate reputation ranking in P2P networks. Future Gener. Comput. Syst. **26**(8), 1317–1326 (2010)
27. Thamarai Selvi, S., Kumar, R., Balachandar, R.A., Balakrishnan, P., Rajendar, K., SwarnaPandian, J.S., Kannan, G., Rajiv, R. Prasath, C.A.: A framework for trust management system in computational grids. In: National Conference on Advanced Computing (NCAC) (2007)
28. Somasundaram, T.S., Amarnath, B.R., Kumar, R., Balakrishnan, P., Rajendar, K., Kannan, G., Rajiv, R., Mahendran, E., Rajesh, B.G., Madusudhanan, B.: CARE resource broker: a framework for scheduling and supporting virtual resource management. Future Gener. Comput. Syst. (2009)
29. Huedo, E., Montero, R.S., Llorente, I.M.: A modular meta-scheduling architecture for interfacing with pre-WS and WS Grid resource management services. Future Gener. Comput. Syst. **23**, 252–261 (2007)

http://www.springer.com/978-3-319-30347-5

Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges (ISBCC – 16')
Vijayakumar, V.; Neelanarayanan, V. (Eds.)
2016, XX, 520 p. 157 illus., 104 illus. in color.,
Hardcover
ISBN: 978-3-319-30347-5