Trust Based Resource Selection in Cloud Computing Using Hybrid Algorithm

V.Suresh Kumar
Research Scholar, M S University Tirunelveli, Tamilnadu, India
E-mail: sureshkumarv.clo@gmail.com

M. Aramudhan
Department of Information Technology, Perunthalaivar Kamarajar Engineering College, Pondicherry, India

Abstract—Cloud computing is experiencing rapid advancement in academia and industry. This technology offers distributed, virtualized and elastic resources as utilities for end users and can support full recognition of “computing as a utility” in the future. Scheduling distributes resources among parties which simultaneously and asynchronously seek it. Scheduling algorithms are meant for scheduling and they reduce resource starvation ensuring fairness among those using resources. Most Task-scheduling cloud computing procedures consider task resource requirements for CPU and memory, and not bandwidth. This study suggests optimizing scheduling with BAT-Harmony search hybrid algorithm.

Index Terms—Cloud Computing, Task Scheduling, BAT Algorithm, Harmony Search

I. INTRODUCTION

Cloud computing uses central remote servers and internet for data and applications maintenance. It permits consumers to use applications sans installation and access personal files at a computer having internet access. This ensures more efficient computing through centralizing memory, storage, processing and bandwidth [1]. Cloud computing is “a type of parallel and distributed system in a collection of inter-connected and virtualized computers dynamically provisioned and accessed as one or more unified computing resources established on Service Level Agreement (SLA), via negotiation between service provider and consumer”.

Due to benefits like elastic and scalable on-demand resources, Cloud service users explore ability of scalable platforms for efficient and cost-effective execution of applications like e-commerce, High Performance Computing (HPC), social network and web applications. For successful Cloud resources use, applications should adapt to the new environment and new scheduling solutions should be developed to ensure good performance. So also, Cloud providers should determine proper configurations and scheduling policies for proper use of computational, networks and storage resources like different applications that are executed both concurrently and in isolation [2].

To ensure that cloud services are effective as provider infrastructure, task scheduling algorithm is a major requirement. They are responsible to map jobs submitted to cloud environment to resources so that total response time in makespan, is reduced [3]. Task scheduling is allocating one or more time intervals to one or more resources. Scheduling is a problem of scheduling submitted tasks set from different users on computing resources to reduce a specific job’s completion time or system makespan. Many parameters are factors for scheduling problems like system throughput, load balancing, service cost, service reliability, and system use.

![Cloud Architecture](image)

Scheduling is a set of policies controlling order of work a computer system has to perform. Various scheduling algorithm exist in distributed computing, and job scheduling is one. The advantage of job scheduling algorithm is achieving high performance computing and best system throughput. Scheduling manages CPU memory availability. A good scheduling policy ensures maximum resource use [4]. Job grouping based scheduling is dynamic scheduling strategy maximizing resources use, processing capabilities and reducing overhead time and cost to execute jobs.

Job scheduling considers: (i) processing requirements for a job, (ii) jobs grouping mechanism, called job grouping, according to resources processing capabilities, and (iii) job grouping transmission to correct resource [5]. The scheduler orders jobs so that balance between improving quality of services and simultaneously maintaining jobs efficiency and fairness is maintained. So, evaluating scheduling algorithms performance is crucial.
to realizing large-scale distributed systems. Despite various scheduling algorithms for cloud environment, no comprehensive performance study providing a unified platform for comparing such algorithms exists [6].

As Cloud computing Systems users increased, tasks to be scheduled also increased proportionally. So, better algorithms to schedule tasks on systems are required. Algorithms to schedule tasks are service oriented and differ according to environments [7]. Cloud computing Task Scheduling algorithms minimize tasks makespan with minimum resources. Cloud computing, uses low-power hosts to ensure high usability. Cloud computing is a class of systems and applications using distributed resources to function in a decentralized way.

Available resources must be used efficiently without affecting service parameters of cloud. Cloud scheduling process service parameters are generalized in 3 stages namely [8],

1. Resource discovering and filtering - Datacenter
   Broker notices resources in network and gets their status information.

2. Resource selection - Target resource selection is based on task and resource parameters. This is a deciding stage

3. Task submissions - Tasks are submitted to selected resources

Task scheduling is a famous combinatorial optimization problem. The aim is determining a proper sequence during task execution while obeying some (transaction logic) constraints. As an optimization problem, getting an optimal packing (allocation) scheme needs huge computation time. After a resource calculation algorithm determines mapping scenarios to satisfy QoS requirements, an actual resource allocation algorithm maps task demands to available resources [9].

The target of scheduling is maximizing resource use and lowering tasks process time. An efficient job scheduling strategy must yield less response time [10] to ensure that jobs execution takes place in a stipulated time. Simultaneously there is an occurrence of in time resource reallocation. Due to this, jobs take place and more jobs are submitted to cloud by clients which results in accelerating cloud system’s business performance. Many scheduling algorithms are available and the most popular are discussed below:

*FCFS* is First Come First Serve algorithm where first data reaching queue first is executed first. The *round robin algorithm* has an advantage over FCFS algorithm. It allocates a process or task with a time slot and after this, the job is changed and the next job comes to execution. *Priority scheduling* algorithm discards FCFS and round robin algorithm’s disadvantages. In this algorithm each job’s priority is decided based on tasks properties [11]. *Hybrid algorithm* uses FCFS and priority algorithm where a job queue has only one algorithm to be done, when it goes for FCFS algorithm but when the next algorithm reaches the queue, it checks jobs priority and reschedules tasks accordingly. A hybrid approach is always better than a single approach.

Many-tasks workflow scheduling in a distributed computing platform is a NP-hard problem which is even more complex and challenging when virtualized clusters execute many tasks on a cloud computing platform. The difficulty is in satisfying multiple objectives that may be conflicting in nature. It is difficult to minimize many tasks makespan, when reducing resource cost and preserving fault tolerance and/or Quality of Service (QoS) simultaneously. Such requirements and goals are hard to optimize due to unknown runtime conditions, like resources availability and random workload distributions [12].

Usually, standard BA algorithm exploits search space, but sometime is trapped in local optima and so cannot perform global search well. For BA, search depends on random walks and hence fast convergence is not guaranteed. This study proposes a hybrid algorithm for BAT-Harmony search for scheduling optimization in cloud computing. The rest of the study is as follows: Section 2 reviews related works in literature. Section 3 explains methods used in the proposed work; Section 4 discusses experiments and results and Section 5 concludes the work.

**II. RELATED WORK**

A service flow scheduling with various QoS requirements in cloud computing was proposed by Liu et al., [13], adopting an Ant Colony Optimization (ACO) algorithm to improve service flow scheduling. In the new model, default rate described radio of cloud service provider breaking SLA, and introduced SLA module to monitor cloud services running state.

The relation between infrastructure components and power consumption of cloud computing environment was studied by Luo et al., [14], who discussed task types matching and component power adjustment methods. A Cloud Computing resource scheduling algorithm built on energy efficient optimization methods was presented. Experiments proved that jobs that used hardware environment had energy consumption reduced by the new algorithm.

An ACO based job scheduling algorithm, adapted to dynamic cloud computing characteristics that integrated specific ACO advantages in NP-hard problems was proposed by Song et al., [15]. It aimed to reduce job completion time based on pheromone. Experiments showed that ACO algorithm is promising for job scheduling in cloud computing environments.

System service cost and task scheduling algorithms of optimizing resource was studied by Wng et al., [16]. Besides, it also proposed a mathematical model of Dynamic Service-quality Cost (DSC) which in real-time synthetically measures service cost and node's service performance in cloud computing. Further, it also presented a complexity-aware optimization task scheduling algorithm. Finally, experiments proved the proposed algorithm's reasonability and effectiveness. It achieved not only reduced service cost but also increased service quality by reasonably assigning tasks to nodes for
calculation in cloud computing. And results were consistent with theoretical analysis.

A PSO based heuristic to schedule applications to cloud resources that considered computation cost and data transmission cost was presented by Pandey et al., [17]. A workflow application was experimented on varying computation and communication cost. Cost savings when using PSO was compared with Best Resource Selection (BRS) algorithm. Results revealed that PSO achieved: (a) three times more cost savings compared to BRS, and (b) also resulted in good workload distribution.

Use of working scheduling models to improve resource allocation in cloud computing was suggested by Chen and Tseng [18]. However, cloud computing and resources distribution is more responsive with time and resources optimal oriented. Authors analog users in accordance with demand work category parameters, and scheduling parameters to access Internet usage or take single time to work the parameter, by queuing theory for scheduling models like demand, scheduling parameters to model, generating reference data, through data analysis to compare various scheduling characteristics. Cloud computing systems future is a frame of reference to build own cloud computing models.

A new method to solve problem by applying stochastic integer programming for ideal resource scheduling in cloud computing was proposed by Li and Guo [19]. Applying Grobner bases theory to solve stochastic integer programming and experimental results of implementation are presented.

A Least Languages First Min-Min (LLFMM), algorithm based scheduling algorithm, suitable to multilingual information resources to improve information resources scheduling performance in cloud computing was proposed by Han and Luo [20]. Concurrent Scheduling Marking Graph (CSMG) of Colored Dynamic Timed Petri nets (CDTdPN) was defined and algorithm for constructing CSMG of CDTdPN was given. CSMG represented concurrent relation of transitions in Petri nets and reduced graph size to be constructed to some extent. Finally, example and simulation revealed that the new scheduling algorithm reduced makespan and that the new modeling and analyzing techniques were convenient to analyze scheduling system performance.

A QoS guided task scheduling model, composed of scheduling strategies and QoS guided scheduling Sufferage-min heuristic algorithm was presented by Han et al., [21]. Experiment revealed that make span value was successfully shortened.

A Multi Queue Scheduling (MQS) algorithm to reduce cost of reservation and on-demand plans using global scheduler was proposed by Karthick et al., [22]. Scheduling is an important and complex part in cloud computing. The new methodology depicted the clustering concept for jobs based on burst time. The new method overcame problems and reduced starvation in the process. The new MQS method gave more importance to job selection dynamically to achieve optimum cloud scheduling problem, and hence it used the unused free space economically.

A Period ACO (PACO) based scheduling algorithm, to solve task scheduling problems in cloud computing was proposed by Sun et al., [23]. PACO used ACO algorithm in cloud computing, with first proposed scheduling period strategy and improvement of pheromone intensity updated strategy. Experiment showed that PACO ensured good performance in makespan and load balance of the total cloud cluster.

A new framework built combining computing resources on Cloud computing and computing components in local systems was presented by Man and Huh [24]. The framework’s main component is cost with finish time-based scheduling algorithm that balanced application schedule performance and mandatory cost for use of Cloud resources. Experiments and comparison with the other scheduling approaches proved benefits of the new algorithm.

An architecture based on collaboration between thin-thick clients and clouds that provided effective task scheduling which boosted processing time in mobile cloud platform was proposed by Hung et al., [25] while considering network bandwidth and cost for cloud service usage. Experiments showed improved efficiency of task scheduling in ensuring desired processing time that corresponded to customer’s money.

An optimization model for task scheduling for reducing energy consumption in cloud-computing data centers was presented by Liu et al., [26]. The new approach was formulated as an integer programming problem to lower cloud-computing data center energy consumption by scheduling tasks to a minimum servers while keeping task response time constraints. Additionally, the most efficient server first task-scheduling scheme to minimize energy expenditure as a practical scheduling scheme was formulated. Simulation show that new task-scheduling scheme reduced server energy consumption on average over 70 times compared to energy consumed under a (not-optimized) random-based task-scheduling scheme. Energy savings were achieved by reducing allocated servers.

III. METHODOLOGY

This study proposed a hybrid method using BAT algorithm and Harmony search algorithm for scheduling optimization in cloud computing.

A. BAT algorithm

Xin-She Yang in 2010 [27] developed BAT algorithm which exploits the bats echolocation. Bats use sonar echoes to detect/avoid obstacles. Sound pulses are transformed to a frequency that reflects from obstacles. Bats use time delay from emission to reflection and use it to navigate. They emit short loud, sound impulses and the pulse rate is defined as 10 to 20 times /sec. After hitting and reflecting, bats transform own pulse to advantageous information to measure how far away the prey is. Bats
use wavelengths that vary from in range \([0.7, 17]\) mm or inbound frequencies \([20,500]\) kHz.

BAT algorithm is based on idealizing bats’ echolocation characteristics which follow approximate or idealized rules [28]:

1. Bats use echolocation to sense distance, and “know” surroundings in some magical way.
2. Bats fly randomly with velocity \(v_i\) and a fixed frequency \(f_{\text{min}}\) at position \(x_i\), varying wavelength \(\lambda\), and loudness \(A_0\) to hunt. They spontaneously accommodate wavelength (or frequency) of emitted pulses and adjust pulse rate emission \(r \in [0, 1]\), depending on target’s proximity.
3. Though loudness changes in different ways, it is supposed that loudness varies from a minimum constant (positive) \(A_{\text{min}}\) to a large \(A_0\).

### B. Harmony search

Harmony Search (HS) is a music-based meta-heuristic optimization algorithm inspired by observation that music aims to search for a perfect state of harmony. The harmony in music is analogous to finding optimality in optimization. Optimization search process is compared to a jazz musician’s improvisation process. On one hand, perfectly pleasing harmony is determined by audio aesthetic standards. A musician intends to produce music with perfect harmony. Alternatively, an optimal solution to an optimization problem is best solution to a problem under given objectives and limited by constraints. Both processes produce best or optimum [29].

HS algorithm initializes Harmony Memory (HM) with random generated solutions. The solutions stored in HM are defined by Harmony Memory Size (HMS). Then a new solution is created as follows iteratively. Every decision variable is generated on memory consideration and a possible additional modification, or through random selection. Parameters used in generation of a new solution are called Harmony Memory Considering Rate (HMCR) and Pitch Adjusting Rate (PAR). Each decision variable is set to value of corresponding variable of one solution in HM with a HMCR probability. Further modification of this value is done with a PAR probability [30]. Otherwise (with a probability of \(1-\text{HMCR}\), decision variable is set to random value. After a new solution is generated, it is evaluated and compared to worst HM solution. If its value is better than that of worst solution, it replaces worst solution in HM. This is repeated, till a termination criterion is fulfilled. The hybrid HS/BA meta-heuristic algorithm is as follows [31]:

**Begin**

**Step 1: Initialization.** Set the generation counter \(t = 1\);
initialize the population of \(NP\) bats \(P\) randomly and each bat corresponding to a potential solution to the given problem;
define loudness \(A\); set frequency \(Q\), the initial velocities \(V\), and pulse rate \(r\);
set the harmony memory consideration rate HMCR, the pitch adjustment rate PAR and bandwidth \(bw\);
set maximum of elite individuals retained KEEP.

**Step 2:** Evaluate the quality \(f\) for each bat in \(P\) determined by the objective function \(x\).

**Step 3:** While the termination criterion is not satisfied or \(t < \text{MaxGeneration}\) do
- Sort the population of bats \(P\) from best to worst by order of quality \(f\) for each bat.
- Store the KEEP best bats as KEEPBAT.

**Step 4:** for \(i = 1:NP\) (all bats) do
- \(v'_i = v_i^{-1} + (v'_i - x_i) \times Q\)
- \(x'_i = x_i^{-1} + v'_i\)
- if \((\text{rand} > r)\) then
  \(x'_i = x_i + \alpha e’\)
end if

**Step 5:** for \(j = 1:D\) (all elements) do //Mutate
- if \((\text{rand} < \text{HMCR})\) then
  \(r_i = [NP \times \text{rand}]\)
  \(x_i(j) = x_i(j) \text{ where } r_i \in \{1,2,...,\text{HMS}\}\)
else
  \(x_i(j) = x_{\text{min},j} + \text{rand} \times (x_{\text{max},j} - x_{\text{min},j})\)
end if

**Step 6:** Evaluate the fitness for the offsprings \(x'_u, x'_v, x'_w\).

**Step 7:** Select the offspring \(x'_k\) with the best fitness among the offsprings \(x'_u, x'_v, x'_w\).

**Step 8:** if \((\text{rand} < A)\) then
  \(x'_h = x'_k\)
end if

**Step 9:** Replace the KEEP worst bats with the KEEP best bats KEEPBAT stored.

**Step 10:** end while

**Step 11:** Post-processing the results and visualization;

**End**

Generally, Trust establishes/maintains relationship between 2 entities for long. Applying trust models to scheduling decreases failure ratio and reassigns cloud environments. Combining communication trust and data trust locates a component/resource/service’s overall trust when scheduling. Data trust decides resources list to separate trustful and untrusted nodes. Communication trust is calculated on client’s bandwidth availability and resource centers. Bayesian model is used for data fusion and communication trust. Reputation ratings are calculated by probability density functions based beta reputation given by [32].
\begin{equation}
f(p | \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1 - p)^{\beta-1}
\end{equation}

Where \( \alpha \) represents jobs completed and \( \beta \) represents unsuccessful jobs. \( R_{ij} \) is reputation for resource \( n_i \) observed from neighborhood resources \( n_j \).

\begin{equation}
R_{ij} = \text{Beta}(\alpha_j + 1, \beta_j + 1)
\end{equation}

Trust value is then calculated using expected reputation value.

\begin{equation}
T_j = E(R_j) = E\{\text{Beta}(\alpha_j + 1, \beta_j + 1)\} = \frac{(\alpha_j + 1)}{(\alpha_j + \beta_j + 2)}
\end{equation}

IV. EXPERIMENTAL RESULTS

The experiments are conducted with BAT and HS algorithms separately and with the hybrid algorithm. The proposed scheduling mechanism was implemented.

<table>
<thead>
<tr>
<th>Time required to complete all task</th>
<th>Time (second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random resource selection with BAT scheduling</td>
<td>391.14</td>
</tr>
<tr>
<td>Trust based resource selection with BAT scheduling</td>
<td>371.26</td>
</tr>
<tr>
<td>Random resource selection with Harmony Search scheduling</td>
<td>384.02</td>
</tr>
<tr>
<td>Trust based resource selection with Harmony Search scheduling</td>
<td>378.11</td>
</tr>
<tr>
<td>Random resource selection with BAT-HS scheduling</td>
<td>366.71</td>
</tr>
<tr>
<td>Trust based resource selection with BAT-HS scheduling</td>
<td>352.42</td>
</tr>
</tbody>
</table>

Trust based resource selection with BAT-HS scheduling decreases the time essential to complete all tasks by 3.9742% when compared with Random resource selection with BAT-HS scheduling, by 8.5818% when compared with Random resource selection with HS scheduling and by 10.4148% when compared with Random resource selection with BAT scheduling.

V. CONCLUSION

Cloud computing is a paradigm sharing computing and storage infrastructure over scalable resources network. In today’s world, data is scattered across different data centers and applications are stored in remote servers. Cloud technology brings scattered data and remote applications to user’s laptop virtually. The idea is to make computing and storage infrastructure available to cloud users regardless of time and location. Trust has an important role in commercial cloud environments and trust management is an integral part of cloud technology’s commercial aspects. This study proposed a trust based resource selection scheme using BAT-HS for scheduling problem in Cloud computing. The new Trust based resource selection with BAT-HS scheduling decreased time to complete tasks by 10.4148% compared to Random resource selection having BAT scheduling.

REFERENCES


Authors’ Profiles

V. Suresh Kumar has completed his BE Computer Science from Bharathidasan University, Tamilnadu, M.Tech in Computer Engineering from SASTRA, Tamilnadu. He is doing PhD in Computer Science at Manonmanium Sundaranar University, Tirunelveli. At present he is working as Dean of Dept.of Engineering, SNGIST, N.Paravoor His area of interest is Data mining, Networking, and Cloud computing.

Maramudhan has completed his undergraduate and postgraduate degrees in Computer Science and Engineering from Bharathidasan University. He obtained his PhD degree from Anna University, Chennai. At present, he is working as Associate Professor in Department of Information Technology, PKIET, Karaikal, Nedungadu, INDIA.