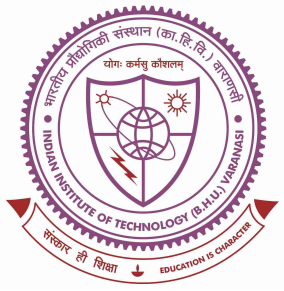
Soft Computing based Dependability Analysis for On-Demand Computing based Transaction Processing System



Thesis submitted in partial fulfillment

for the Award of Degree

Doctor of Philosophy

by

Dharmendra Prasad Mahato

COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY

(BANARAS HINDU UNIVERSITY)

VARANASI-221 005

Roll No: 13071009

January 27, 2018

Chapter 7

Conclusion and Future Directions

7.1 Concluding Discussion

Despite the popularity of on-demand computing based transaction processing system, not everything is positive. The real-time, heterogeneous, complex and dynamic nature of this environment make it challenging to provide unforeseen failures, load balancing, unexpected operating conditions, and malicious attacks. Therefore, dependability with load balancing and scheduling become important for on-demand computing system.

The balanced transaction scheduling in on-demand computing system is an NP-hard problem. To solve such problems, the meta-heuristic approaches are considerably best options. We proposed an ACO based balanced the grid transaction scheduling algorithm LBTS_ACO. We chose the concept and scenario of grid computing as a case study for applying this algorithm. The algorithm balanced the load of the system before the transactions were scheduled to the required nodes. It maximized the node utilization, minimized the makespan, increased the **throughput**, minimized the **miss ratio**, and maximized the **load balancing speedup** of the grid transaction processing system. The next step would be to model and analyze the availability, reliability, and dependability of the data grid system by meta-heuristic transaction scheduling algorithms.

Owing to the tremendous growth of transaction processing in on-demand computing system, the business enterprises needs maximum number of resources to execute the

transactions. The maximization of the **resource availability** becomes one of the prime factors for transaction scheduling in on-demand computing system. Thus, we formulated the problem with multi-objective functions; maximizing availability and minimizing makespan. we used ACO based transaction scheduling algorithm MATS_ACO. We compared our proposed algorithm with two meta-heuristic scheduling algorithms based on EO and GA. The experimental results showed that our proposed algorithm performed better than other two algorithms. We also carried out Wilcoxon statistical test for the validation of the results. The normality tests was carried out using Shapiro-Will tests. For the network simulation we followed NFSNet scenario.

For maximizing the **performability**, we proposed a load balanced scheduling technique for on-demand computing based transaction processing systems based on the behavior of honey bee foraging strategy LBTS_HBO. The algorithm balanced the load before scheduling the transactions to appropriate nodes in on-demand computing environment. We formulated the resource availability and performability considering load. We modified four scheduling algorithms and obtained transaction scheduling algorithms for the purpose of comparison with our proposed algorithm. The result showed that the algorithm enhanced the resource availability by decreasing the load. It also increased performability and reduced the miss ratio. This load balanced scheduling algorithm worked well for on-demand computing based transaction processing systems.

For the attribute **reliability**, we proposed the balanced task allocation algorithm for on-demand computing based transaction processing system using social spider optimization LBTA_SSO. The algorithm first balanced the load before it allocated the transaction to the appropriate node in the system. We also formulated the resource availability and reliability considering the load. We simulated the algorithm on two scenarios of on-demand computing system; grid and cloud. We compared the proposed algorithm with five modified algorithms. The result showed that the resource availability and reliability were maximized. It also reduced the miss ratio. The proposed algorithm worked well for a transaction in on-demand computing system.

7.2 Future Directions

The continued growth of business transactions, data sizes, advent of novel applications, and evolution of the infrastructure ensures that the area of on-demand based transaction processing system has many interesting research challenges. While some of these future research directions are direct extensions of the techniques presented in this dissertation, others are more radical.

Assignment and execution of transactions to the appropriate resources should also be within their prescribed deadline. This process in on-demand computing system will be profitable by employing efficient and deadline constrained task partitioning and scheduling strategies. More and more evidence show that the deadline constrained scheduling in such environment is highly critical to the performance of on-demand computing systems. The common objective of deadline constrained scheduling is to map tasks of a parallel application onto processors of system and order their executions, so that task precedence constraints are also satisfied and the minimum makespan is achieved. Scheduling of deadline constrained tasks in such distributed systems is a technique which requires that the tasks should be assigned to computing resources and the order of execution should be stipulated in such a manner that

- precedence constraints should not be violated
- all the tasks should be executed within their prescribed deadline
- all temporal constraints should be satisfied

Most of the scheduling approaches (either heuristic or meta-heuristic) in the literature are based on the deterministic version of the scheduling problem with deterministic computation and communication times. Such algorithms assume that parameters like task processing times and intertask communication times are fixed and deterministic, which are known in advance. But in real world scenarios, execution times of the tasks in on-demand computing based systems may not be fixed. These tasks may have different execution times for different inputs because they usually have conditional instructions and operations. Also the communication times among the tasks may fluctuate according to the network traffic. The deadline constrained tasks in this situation may frequently

face deadline-miss failure. The existing deterministic scheduling for these tasks becomes complex due to the nature of randomness and uncertainty of these tasks. Therefore, we have to consider stochastic task scheduling to tackle these problems so that we can interpret processing times and communication times as random variables.

We would like to model a stochastic differential equation (SDE) for dependability of on-demand computing based transaction processing system. The Wiener processes used in the SDE equations miss some important features of the transactional data. We can add jumps based on Gaussian jump diffusion process using normal distribution to the Wiener processes to make the estimation easy. Jump diffusion parameters can be estimated using meta-heuristic algorithms. Moreover, we can formulate a total expected software cost model for the environment of on-demand computing. The model would be estimated and validated on real-life datasets of transaction processing in literature to show its flexibility. Thus we obtain exquisite fit of the observed datasets with the stochastic differential equations. We would like to use the proposed method to construct SDEs-based models from different probability distribution also.