Modeling of load balanced scheduling and reliability evaluation for on-demand computing based transaction processing system

Dharmendra Prasad Mahato
Computer Science and Engineering
BIT Sindri, Dhanbad-828123
Jharkhand, India
dpmahato.rs.cse13@iitbhu.ac.in

Jasminder Kaur Sandhu
Computer Science and Engineering
Thapar University, Patiala-147001
Punjab, India
jasminder.kaur@thapar.edu

***Abstract*—The load scheduling and reliability modeling in on-demand computing based transaction processing are complex tasks. This paper presents the CPNs (Coloured Petri Nets) based modeling for load balanced scheduling and reliability analysis for on-demand computing based transaction processing system.**

***Index Terms*—On-demand computing, Load balanced scheduling, Reliability analysis, CPN Tools.**

I. Introduction

Transaction processing is a key application used for an enormous range of economic activities, from travel reservation, ticketing systems and electronic banking, to financial transactions and e-commerce. It certainly has become a critical component of the worlds economic engine which is responsible for handling trillions of transactions every year. Because, the never-ending growth of the global economy along with e-business and Web-based commerce are placing more demands on transaction processing systems.

When choosing on-demand computing based transaction processing strategy, corporations must rely on solutions that ensure data availability, reliability, performance, and scalability. On-demand computing based transaction processing e.g. grid transaction processing is a largely invisible back-end business computing function [1]. New trends and emerging technologies, such as the increasing use of mobile messaging technology and the employment of Radio Frequency Identification (RFID) tracking, are generating more transaction traffic. Therefore, the workload volume managed by on-demand computing based transaction processing systems is enormous.

Thus, modeling of load balanced scheduling along with reliability evaluation for on-demand computing based transaction processing system have become important research in distributed system [2].

This paper uses CPNs to model the load balanced scheduling and reliability in on-demand computing system. We analyze the instance of the proposed model for on-demand computing system using CPN Tools. Our main contributions are two-fold. First, a load balanced scheduling is designed to represent transaction oriented grid architecture. Second, reliability of on-demand computing system is modeled and is analyzed using CPN Tools.

The rest of the paper is organized as follows. Section II presents the load balanced scheduling model and the reliability model. Section III presents the CPNs based model of on-demand based transaction processing system. Finally, section IV concludes the paper.

II. Load Balanced Scheduling with Reliability Evaluation

Load balanced transaction scheduling has become important task for the enhancement of throughput, makespan, resource utilization, and reliability of the system. In literature, we find a lot of works in the field of scheduling and reliability evaluation [3]–[9]. But we find rare examples of load balanced scheduling with reliability evaluation.

When load balanced scheduling for transaction processing is required, the deadline constraint in the system is also needed to model.

Let us consider the set T of m transactions T_i $(i = 1, 2, ..., m)$ to be processed within their given deadline, on the set of n nodes N_j $(j = 1, 2, ..., n)$.

Therefore, the expected queue length Q_j at node N_j is given by the expression:

$$Q_j = \sum_{i=1}^{m} T_{ij}.x_{ij}, \forall i = 1, \dots, m \text{ and } \forall j = 1, \dots, n. \quad (1)$$

where

$$x_{ij} = \begin{cases} 1 & \text{if } T_i \text{ is assigned to node } N_j \\ 0 & \text{otherwise} \end{cases}$$

and T the total number of transaction arrivals in the system is given as $\sum_{i=1}^{m} T_i$.

The transaction arrivals T is Poisson distributed within any interval of length τ, when transactions arrive with the rate λ. The transaction arrivals can be given as

$$T = e^{-\lambda\tau} \frac{(\lambda\tau)^m}{m!}, \; m = 0, 1, ... \quad (2)$$

978-1-5386-9156-4/18/$31.00 ©2018 IEEE
DOI 10.1109/eScience.2018.00114

The completion time of transactions are defined as $C_{i,j}$ ($i \in \{1, 2, ..., m\}$, $j \in \{1, 2, ..., n\}$) that the transaction T_i takes to finish the execution at node N_j. Then, makespan of the system is calculated as

$$C_{max} = max\Big\{\sum C_j\Big\} \tag{3}$$

where $\sum C_j$ represents the time that each node N_j finishes all the transactions scheduled for itself.

The system reliability that there is no transaction deadline-miss fault in addition to nodes and communication links are operational during the elapsed time for the execution can be computed as

$$R_{j,jb,DM}(X) = \left[\prod_{j=1}^{n} R_j(X)\right] \cdot \left[\prod_{j=1}^{n-1} \prod_{b>j} R_{kb}(X)\right] \cdot \left[\prod_{i=1}^{m} R_{DM}(X)\right] \tag{4}$$

where the node reliability $R_j(X)$ of a node N_j during a time interval t has been computed as $e^{-\gamma_j \sum_{i=1}^{m} x_{ij} e_{ij}}$ when $\sum_{i=1}^{m} x_{ij} e_{ij}$ is the total elapsed time t for executing the transactions assigned to N_j. While the communication link reliability $R_{jb}(X)$ at a time interval t has been computed as $e^{-\sigma_{jb} \sum_{i=1}^{m} \sum_{g \neq i} x_{ij} x_{gb} (cost_{ig}/w_{jb})}$ where the total elapsed time for transmitting the transaction communication via l_{jb} is $\sum_{i=1}^{m} \sum_{g=1}^{m} x_{ij} x_{gb} (cost_{ig}/w_{jb})$. Here transactions are steady state and follow queuing system model $M/M/c$. While, $R_{DM}(X)$, the probability that there is no deadline-miss with rate ψ_D when transaction T_i is scheduled on N_j can be computed by using the Markov model as

$$R_{DM}(X) = e^{-\psi_D \cdot \left[\frac{1}{\mu} + q_0 \cdot \frac{\rho (c\rho)^c}{c!(c\mu - \lambda)(1-\rho)}\right]}, \quad \forall c \in N \tag{5}$$

where q_0 is given by

$$q_0 = \left[\sum_{i=0}^{c-1} \frac{(c\rho)^i}{i!} + \frac{(c\rho)^c}{c!(1-\rho)}\right]^{-1} \tag{6}$$

where $\rho = \frac{\lambda}{c.\mu} < 1$.

III. The CPN model

The model is a hierarchical CPNs based model (as shown in Figure 1). It consists of various hierarchical nets to model the system properly. The users' requests for the transactional service are generated in the DATAGEN (Data Generation) net using exponential distribution. The CLIENT net receives these requests, abiding by the rules of the system. Then, the requests are scheduled in the SCHEDULER net which decides the order of arrival of the requests and schedules the job. The GLB (Gloabal Load Balancing) net models the global resource management system where the transactional tasks are subdivided into multiple subtasks and those subtasks are sent to the resources for execution. The LocalRM is the local level resource management net. After that the transaction management is performed in the TRANSACTION MANAGER net. Then, MONITOR net monitors these tasks. It decides where to send the received tasks either to the users or to the SCHEDULER net. If the subtasks are executed unsuccessfully, they are again rescheduled either on the local node (where the subtask has faced failure or on the replicated nodes (any other nodes rather than the local one). In the case of replicated node, the replication method is used.

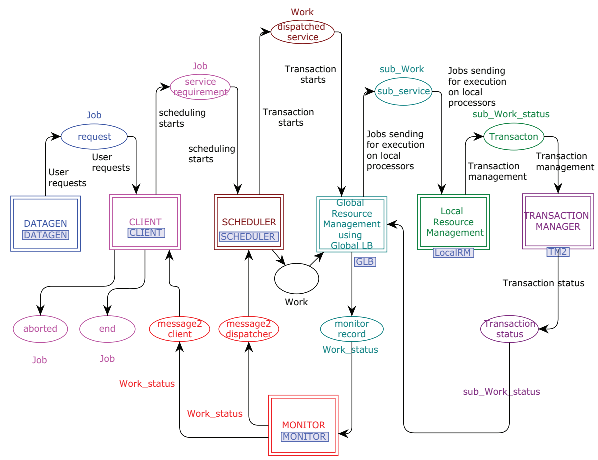

Fig. 1. CPN model of GTP

IV. Conclusion

This paper presents the modeling of the on-demand computing based transaction processing using CPNs based modeling.

References

[1] D. P. Mahato and R. S. Singh, "Balanced task allocation in the on-demand computing-based transaction processing system using social spider optimization," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 18, p. e4214, 2017.

[2] D. P. Mahato, "Cpns based reliability modeling for on-demand computing based transaction processing," in *Proceedings of the 47th International Conference on Parallel Processing Companion*, ser. ICPP '18. New York, NY, USA: ACM, 2018, pp. 24:1–24:4. [Online]. Available: http://doi.acm.org/10.1145/3229710.3229746

[3] D. P. Vidyarthi and A. K. Tripathi, "Maximizing reliability of distributed computing system with task allocation using simple genetic algorithm," *Journal of Systems Architecture*, vol. 47, no. 6, pp. 549 – 554, 2001.

[4] S. Kartik and C. S. R. Murthy, "Improved task-allocation algorithms to maximize reliability of redundant distributed computing systems," *IEEE Transactions on Reliability*, vol. 44, no. 4, pp. 575–586, Dec 1995.

[5] K. S. Trivedi, J. K. Muppala, S. P. Woolet, and B. R. Haverkort, "Composite performance and dependability analysis," *Performance Evaluation*, vol. 14, no. 3, pp. 197–215, 1992.

[6] D. P. Mahato and R. S. Singh, "On maximizing reliability of grid transaction processing system considering balanced task allocation using social spider optimization," *Swarm and Evolutionary Computation*, vol. 38, pp. 202–217, 2018.

[7] D. P. Mahato, R. S. Singh, A. K. Tripathi, and A. K. Maurya, "On scheduling transactions in a grid processing system considering load through ant colony optimization," *Applied Soft Computing*, vol. 61, pp. 875–891, 2017.

[8] D. P. Mahato, "Cuckoo search-ant colony optimization based scheduling in grid computing," in *Proceedings of the 47th International Conference on Parallel Processing Companion*, ser. ICPP '18. New York, NY, USA: ACM, 2018, pp. 39:1–39:10. [Online]. Available: http://doi.acm.org/10.1145/3229710.3229750

[9] D. P. Mahato and R. S. Singh, "Load balanced transaction scheduling using honey bee optimization considering performability in on-demand computing system," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 21, p. e4253, 2017.