
Performance prediction in time-sharing servers using simulation approach

Perera T.D.J.S. * and Rajapakse C.

Department of Industrial Management, University of Kelaniya, Dalugama, Kelaniya, Sri Lanka.

Although the concept of time-sharing was invented a long time ago to make better use of expensive processor time, it is still in use and applicable in the current era. To provide excellent customer service while minimizing operating and capital expenses, server performance must be managed efficiently and effectively. Consequently, it is necessary to assess the performance of a new time-sharing server to ensure that it performs as expected, prior to deployment. It is also crucial to analyze how an existing time-sharing server operates under various circumstances. As a result, it is vital to have a model that can predict server performance measures under varying load levels and server configurations. Identified parameters of the system are server parameters which are number of CPU cores, average processing time, maximum kernel-thread pool size, time quantum that causes a currently running thread to be preempted, context switch overhead, and load parameters which are number of concurrent users and their average think time. When measuring the performance (latency and throughput) of a time-sharing server, the previous work has not taken into account all of the parameters. By doing a thorough literature review, two approaches were identified (analytical modeling and simulation modeling) to create a model with above parameters, and discrete-event simulation (DES) was selected with its pros and relevancy for the study. As the initial experiments, we tested numerous scenarios by using the built DES model to review the effect of each parameter on the performance of time-sharing server and results were graphically represented.

Keywords: Time-sharing, Server-performance, Simulation

*Corresponding author: jithmisachinika@gmail.com