Asynchronous working vacations in the $M/M/2$ queue

Tuan Hai Nguyen[a], **Tien Van Do**[b]

[a]Budapest University of Technology and Economics
nthaitrx@gmail.com

[b]Budapest University of Technology and Economics
do@hit.bme.com

Abstract

Queues with vacations can be used to model computing systems where mechanisms are applied to save the energy consumption of servers in computing clusters. In the term of queueing theory, server vacations represents an event that physical servers go into a sleeping state if no processing needs can be found in the system.

In this paper, we consider a computing cluster with two servers that change states independently. Such a cluster configuration is commonly established in enterprise environment and small companies. To enhance the reliability two servers are configured with a resource manager (e.g., Pacemaker) that allows the configuration and the operation of the cluster. It supports both active-passive and active-active configurations. In an active-passive setup services run on only one computer and upon failure another computer takes over. In case of an active-active scenario services run on all computer nodes and the incoming requests are distributed through load balancing. After the server finishes its job if it cannot find a customer waiting for service it goes into a vacation period regardless of whether the other server is still serving a request or not. The duration of the vacation, the interarrival times and the service times in the normal states and the vacation states are exponentially distributed.

To analyze such a cluster with two asynchronous servers, we apply the $M/M/2$ queue with working vacations. We provide the closed form solution using the spectral expansion method. Numerical results show that in case of low traffic the power consumption is higher and the average queue length is lower if servers asynchronously change their states. At high traffic though the difference of the performance measures is negligible, and therefore the working vacation mode does not reduce the energy consumption of the cluster with two servers.

Keywords: multiserver, queueing, working vacation

MSC: 60K25 68M20