THE BUSY PERIOD ANALYSIS OF A QUEUE WITH HETEROGENEOUS SERVERS AND THRESHOLD SERVICE POLICY

D. Efrosinin[1], V. Rykov[2]

[1] *Johannes Kepler University Linz, Institute for Stochastics*
[2] *Russian State Oil& Gas University*
[1] *Linz, Austria*
[2] *Moscow, Russia*

dmitry.efrosinin@jku.at, vladimir_rykov@mail.ru

In the present paper we consider a queueing system with heterogeneous exponential servers. The service policy is a threshold-based control one, i.e. the fastest server must be switched on whenever it is free and at least one customer is in the system, while the slower ones must be activated when the number of customers in a queue reaches some threshold level specified for a certain server.

We investigate algorithmically the busy period distribution by deriving expressions for the Laplace transforms and perform recursive calculation of the corresponding moments and the number of customers served during busy period. The optimal threshold levels for the mean busy period minimization are calculated.

Keywords: Multiserver queue; heterogeneous servers; busy period; threshold policy; number of customers served.

1. INTRODUCTION

The queueing systems with several heterogeneous servers take into account many practical aspects in modelling real systems, e.g. network nodes with servers supplied by different type of processors as a consequence of system updates, nodes in telecommunication network with links of different capacities, nodes in wireless systems serving different mobile users, etc. In this paper we consider such a system with one common queue and threshold-based service policy that can be described as follows. The fastest free server must be activated whenever at least one customer is in the queue and other servers must be switched on only if the queue length reaches some threshold levels defined for each server.

Threshold service policies have been applied for multiserver systems by many authors. As it was shown in [3] the optimal policy that minimizes the mean number of customers in the system with two heterogeneous servers and one common queue is of threshold type. This result was generalized in [4], where some convexity properties of the dynamic-programming value function were proved. In [2] the waiting time and sojourn time distributions of a three-server heterogeneous system under threshold policy are derived in form of Laplace transforms.

In the paper we provide algorithmic analysis of a length of busy period that seems to be a very important first-passage characteristic of any queueing model from the service provider's point of view. We obtain Laplace transforms and z-transforms assigning the length of busy period and the number of customers served and develop a recursive scheme for the computation of arbitrary moments.

2. MATHEMATICAL MODEL

We deal with a multi-server model consisting of K heterogeneous servers which serve customers according to the threshold control policy defined by the succession of the threshold levels, $1 = q_1^* \leq q_2^* \leq \cdots \leq q_K^* < \infty$. Denote by $Q(t)$, $D_k(t)$, $k = 1, 2, \ldots, K$ the number of customers in the queue and the states of servers. The random process $\{X(t)\} = \{Q(t), D_1(t), \ldots, D_K(t)\}_{t\geq 0}$ is a continuous-time Markov chain with state space defined as $E = \{x = (q, d_1, \ldots, d_K)|$

$$\begin{cases} d_j = \{0,1\}, & 1 \leq j \leq K, q = 0 \\ d_j = 1, d_i = \{0,1\}, & 1 \leq j \leq k, k+1 \leq i \leq K, q_k^* \leq q \leq q_{k+1}^*, 1 \leq k \leq K-1 \end{cases}\Big\}$$

and with threshold dependent infinitesimal matrix $\Lambda = [a_{ij}]$. The details of the model can be found in [1]. To limit the size of the formulas we only present the case of three servers, but the given method can be applied to any number of servers. The method can also be extended to some other models, e.g. where the inter-arrival and service times have phase type distribution.

3. BUSY PERIOD ANALYSIS

The busy period of the system is a duration L that starts by arrival of a new customer to the empty queue and ends when the system visits zero state after a service completion. First the following notations are defined:

L_x- the first-passage time to zero level given the initial state is $x \in E$,

$\varphi_x(t)-$ density distribution function of the conditional first-passage time,

$\bar{\varphi}_x(s) = \mathbb{E}[e^{-sL_x}] = \int_0^\infty e^{-st}\varphi_x(t)dt, Re[s] \geq 0-$ corresponding Laplace transform.

Because of the markovity of the process $\{X(t)\}$, the residual first-passage time in state x consists of the time the system spend in state x until the next transition with density $\lambda_x e^{-\lambda_x t}$ plus the residual time in a new state y after possible transition from the state x, which take place with probability $\frac{\lambda_{xy}}{\lambda_x}$. Thus from the low of total probability for Laplace transforms we get

$$\bar{\varphi}_x(s) = \sum_{y \neq x} \frac{\lambda_{xy}}{s + \lambda_x} \bar{\varphi}_y(s), \; x \in E. \tag{1}$$

Now we partition the above Laplace transforms and define the following column-vectors:

$$\bar{\boldsymbol{\varphi}}_1(s) = (\bar{\varphi}_{(0,1,0,0)}(s), \bar{\varphi}_{(0,0,1,0)}(s), \bar{\varphi}_{(0,0,0,1)}(s))^t, \tag{2}$$

$$\bar{\varphi}_2(s) = (\bar{\varphi}_{(1,1,0,0)}(s), \bar{\varphi}_{(0,1,1,0)}(s), \bar{\varphi}_{(0,1,0,1)}(s), \bar{\varphi}_{(0,0,1,1)}(s))^t,$$

$$\bar{\varphi}_i(s) = \begin{cases} (\bar{\varphi}_{(i-1,1,0,0)}(s), \bar{\varphi}_{(i-2,1,1,0)}(s), \bar{\varphi}_{(i-2,1,0,1)}(s), \bar{\varphi}_{(i-3,1,1,1)}(s))^t, & 3 \le i \le q_2^* \\ (\bar{\varphi}_{(i-2,1,1,0)}(s), \bar{\varphi}_{(i-2,1,0,1)}(s), \bar{\varphi}_{(i-3,1,1,1)}(s))^t, & i = q_2^* + 1 \\ (\bar{\varphi}_{(i-2,1,1,0)}(s), \bar{\varphi}_{(i-3,1,1,1)}(s))^t, & q_2^* + 2 \le i \le q_3^* + 1 \\ \bar{\varphi}_{(i-3,1,1,1)}(s), & i = q_3^* + 2 \end{cases}$$

$$\bar{\varphi}(s) = (\bar{\varphi}_1(s), \bar{\varphi}_2(s), \dots, \bar{\varphi}_{q_3^*+2}(s))^t.$$

Theorem 1. *The vector $\bar{\varphi}(s)$ of the Laplace transforms $\bar{\varphi}_i(s)$, $1 \le i \le q_3^* + 2$ satisfy the following block tridiagonal system*

$$\bar{\varphi}(s) = \Lambda_L^{-1}(s)\bar{r}(s), \tag{3}$$

$\Lambda_L(s) = \Phi - sI_{2(q_2^*+q_3^*)+3}$, $\bar{r}(s) = -(A_0^t, 0, \dots, 0, \lambda\bar{\xi}(s)\bar{\varphi}_{q_3^*+2}(s))^t$, $\bar{\xi}(s) = \frac{s+\lambda+M-\sqrt{(s+\lambda+M)^2-4\lambda M}}{2\lambda}$. *Matrix Φ is of the form*

$$\Phi = \begin{pmatrix} -B_0 & C_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ A_1 & -B_1 & C_2 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & A_2 & -B_2 & C_2 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & & & & & & \ddots & \\ 0 & \dots & 0 & A_2 & -B_2 & C_3 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & A_3 & -B_3 & C_4 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & A_4 & -B_4 & C_5 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & A_5 & -B_4 & C_5 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & & & & & & \ddots & \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & A_5 & -B_4 & C_6 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_6 & -(\lambda+M) \end{pmatrix} \begin{matrix} \left.\right\} q_2^* \\ \left.\right\} q_3^* - q_2^* \end{matrix}$$

where $M = \mu_1 + \mu_2 + \mu_3$. The entries A_i represent the departures with elements depending on the servers which are active:

$$A_0 = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \quad A_1 = \begin{pmatrix} \mu_1 & 0 & 0 \\ \mu_2 & \mu_1 & 0 \\ \mu_3 & 0 & \mu_1 \\ 0 & \mu_3 & \mu_2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \mu_1 & 0 & 0 & 0 \\ \mu_2 & \mu_1 & 0 & 0 \\ \mu_3 & 0 & \mu_1 & 0 \\ 0 & \mu_3 & \mu_2 & \mu_1 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} \mu_2 & \mu_1 & 0 & 0 \\ \mu_3 & 0 & \mu_1 & 0 \\ 0 & \mu_3 & \mu_2 & \mu_1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} \mu_1+\mu_2 & 0 & 0 \\ \mu_3 & \mu_2 & \mu_1 \end{pmatrix}, \quad A_5 = \begin{pmatrix} \mu_1+\mu_2 & 0 \\ \mu_3 & \mu_1+\mu_2 \end{pmatrix},$$

$$A_6 = \begin{pmatrix} \mu_3 & \mu_1+\mu_2 \end{pmatrix}.$$

The entries C_i represent the arrivals with elements depending on whether the queue lengths are above or below threshold:

$$C_0 = \begin{pmatrix} \lambda & 0 & 0 \end{pmatrix}, C_1 = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \end{pmatrix}, C_2 = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}, C_3 = \begin{pmatrix} \lambda & 0 & 0 \\ \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix},$$

$$C_4 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \\ 0 & \lambda \end{pmatrix}, C_5 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, C_6 = \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}.$$

The diagonal blocks B_i represent the total outcome intensities of the certain state: $B_0 = \operatorname{diag}\{\lambda+\mu_1, \lambda+\mu_2, \lambda+\mu_3\}$, $B_1 = \operatorname{diag}\{\lambda+\mu_1, \lambda+\mu_1+\mu_2, \lambda+\mu_1+\mu_3, \lambda+\mu_2+\mu_3\}$, $B_2 = \operatorname{diag}\{\lambda+\mu_1, \lambda+\mu_1+\mu_2, \lambda+\mu_1+\mu_3, \lambda+M\}$, $B_4 = \operatorname{diag}\{\lambda+\mu_1+\mu_2, \lambda+M\}$, $B_3 = \operatorname{diag}\{\lambda+\mu_1+\mu_2, \lambda+\mu_1+\mu_3, \lambda+M\}$.

Proof. The inverse of the matrix $\Lambda_L(s)$ is well defined because its diagonal elements dominate in each column for any s, $Re[s] \geq 0$. The expression of the recurrent relations (1) in matrix form leads to the equality (3). □

Hence for the Laplace transform $\tilde{\varphi}(s)$ of the unconditional density function we get

$$\tilde{\varphi}(s) = \tilde{\varphi}_{(0,1,0,0)}(s) = \boldsymbol{e}_1^t(2(q_2^* + q_3^*) + 3)\tilde{\boldsymbol{\varphi}}(s). \tag{4}$$

The limit property of the Laplace transform allows us to get the value of the function $\varphi(t)$ at point $t = 0$:

$$\lim_{s\to\infty} s\,\tilde{\varphi}(s) = \begin{cases} \mu_1, & \text{if } (q, d_1, d_2, d_3) = (0, 1, 0, 0), \\ 0, & \text{otherwise.} \end{cases}$$

Therefore we have $\varphi(0) = \mu_1$.

Now we obtain the n-th moments of L_x which we denote by $\bar{L}_x(n) = \mathbb{E}[L_x^n]$. According to the introduced partition of the Laplace transforms we compose the following macro vectors

$$\bar{\boldsymbol{L}}(n) = (\bar{\boldsymbol{L}}_1(n), \bar{\boldsymbol{L}}_2(n), \ldots, \bar{\boldsymbol{L}}_{q_3^*+2}(n))^t. \tag{5}$$

Theorem 2. *The vectors of the n-th moments $\bar{\boldsymbol{L}}_i(n)$, $1 \leq i \leq q_3^* + 2$, $n \geq 0$ satisfy the following recurrent block tridiagonal system*

$$\begin{aligned} \bar{\boldsymbol{L}}(n) &= -n\Phi^{-1}\bar{\boldsymbol{L}}(n-1) + \Phi^{-1}\bar{\boldsymbol{R}}(n),\ n \geq 1 \\ \bar{\boldsymbol{L}}(0) &= \boldsymbol{e}(2(q_2^* + q_3^*) + 3), \end{aligned} \tag{6}$$

$\bar{\boldsymbol{R}}(n) = \lambda \sum_{m=0}^{n} \binom{n}{m} \bar{\boldsymbol{L}}_{q_3^*+2}(m)\bar{\Xi}(n-m)\boldsymbol{e}_{2(q_2^*+q_3^*)+3}$, $\bar{\Xi}(n) = (-1)^n \frac{d^n}{ds^n}\tilde{\xi}(s)\big|_{s=0}$.

Proof. Applying Leibnitz's formula to derive the n-th derivative of (3), we find that $\Lambda_L(s)\frac{d^n}{ds^n}\tilde{\boldsymbol{\varphi}}(s) - n\frac{d^{n-1}}{ds^{n-1}}\tilde{\boldsymbol{\varphi}}(s) = \frac{d^n}{ds^n}\tilde{\boldsymbol{r}}(s)$. Since for the Laplace transform the relations $\bar{\boldsymbol{L}}(n) = (-1)^n\frac{d^n}{ds^n}\tilde{\boldsymbol{\varphi}}(s)\big|_{s=0}$, $\bar{\boldsymbol{R}}(n) = (-1)^n\frac{d^n}{ds^n}\tilde{\boldsymbol{r}}(s)\big|_{s=0}$ hold and taking into account that $\Lambda_L(0) = \Phi$, we acquire the expression (6). □

4. THE NUMBER OF CUSTOMERS SERVED

In this section we study the number of customers N served during a busy period. Define: N_x – r.v. of the number of customers served given initial state x, $\psi_x(k) = \mathbb{P}[N_x = k]$ – probability density function of N_x, $\tilde{\psi}_x(z) = \mathbb{E}[z^{N_x}] = \sum_{k=0}^{\infty} \psi_x(k) z^k$, $|z| \le 1$ – corresponding z- transform, $\tilde{\boldsymbol{\psi}}(z) = (\tilde{\boldsymbol{\psi}}_1(z), \tilde{\boldsymbol{\psi}}_2(z), \dots, \tilde{\boldsymbol{\psi}}_{q_3^*+2}(z))^t$ – macro-vector of z-transforms $\bar{\boldsymbol{N}}(n) = (\bar{\boldsymbol{N}}_1(n), \bar{\boldsymbol{N}}_2(n), \dots, \bar{\boldsymbol{N}}_{q_3^*+2}(n))^t$ – macro-vector of the factorial moments $\bar{N}(n) = \mathbb{E}[N \dots (N-n+1)]$.

For the Markov process $\{X(t)\}$ using the low of total probability we get the relations for z-transforms $\tilde{\psi}_x(z)$ in form

$$\tilde{\psi}_x(z) = \frac{z a_{xy'}}{a_x} \tilde{\psi}_{y'}(z) + \sum_{y \ne x, y'} \frac{a_{xy}}{a_x} \tilde{\psi}_y(z). \tag{7}$$

where the first term represents the service completion and the second one corresponds to all other possible transitions that do not change the event under consideration.

Theorem 3. *The vectors of z-transforms $\tilde{\boldsymbol{\psi}}_i(z)$, $1 \le i \le q_3^* + 2$ satisfy the following block tridiagonal system*

$$\tilde{\boldsymbol{\psi}}(z) = \Lambda_N^{-1}(z) \tilde{\boldsymbol{q}}(z), \tag{8}$$

$\Lambda_N(z) = \Phi - A(z)$, $\tilde{\boldsymbol{q}}(z) = -(zA_0^t, \dots, \lambda \tilde{\zeta}(z) \tilde{\boldsymbol{\psi}}_{q_3^*+2}(z))^t$, $\tilde{\zeta}(z) = \frac{\lambda + M - \sqrt{(\lambda+M)^2 - 4\lambda M z}}{2\lambda}$,

$$A(z) = (1-z) \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & & & & & & & \\ 0 & \dots & 0 & A_2 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & A_3 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & A_4 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & A_5 & 0 & 0 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & & & & & & & \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & A_5 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_6 & 0 \end{pmatrix} \begin{matrix} \left.\right\} q_2^* \\ \left.\right\} q_3^* - q_2^* \end{matrix}$$

Theorem 4. *The vectors of n-th factorial moments $\bar{\boldsymbol{N}}(n)$, $1 \le i \le q_3^* + 2$ satisfy the following block tridiagonal recurrent system*

$$\bar{\boldsymbol{N}}(n) = -n\Phi^{-1}A(0)\bar{\boldsymbol{N}}(n-1) - \Phi^{-1}\bar{\boldsymbol{Q}}(n), n \ge 1, \qquad \bar{\boldsymbol{N}}(0) = \boldsymbol{e}(2(q_2^* + q_3^*) + 3), \tag{9}$$

$\bar{\boldsymbol{Q}}(n) = \delta_{n,1} A_0 \boldsymbol{e}_1 + \lambda \sum_{m=0}^{n} \binom{n}{m} \bar{\boldsymbol{N}}_{q_3^*+2}(m) \bar{\Psi}(n-m)$, $\bar{\Psi}(n) = \frac{d^n}{dz^n} \tilde{\zeta}(z)\big|_{z=1}$.

For unconditional z-transform and factorial moment of the n-th order we get, respectively

$$\tilde{\psi}(z) = \tilde{\psi}_{(0,1,0,0)}(z) = \boldsymbol{e}_1^t(2(q_2^* + q_3^*) + 3)\tilde{\boldsymbol{\psi}}(z), \tag{10}$$

$$\bar{N}(n) = \bar{N}_{(0,1,0,0)}(n) = \boldsymbol{e}_1^t(2(q_2^* + q_3^*) + 3)\bar{\boldsymbol{N}}(n), \; n \ge 0. \tag{11}$$

5. NUMERICAL EXAMPLES

In Figure 1 we apply the numerical inversion algorithms to obtain the distribution functions $\Phi(t)$ and $\Psi(k)$ under threshold policy that minimizes the mean length of busy period, figures labelled by (a) and (b), respectively. In our examples we fix the service rates $\mu_1 = 1.5$, $\mu_2 = 0.5$ and $\mu_3 = 0.3$. As expected we observe that when λ increases the distribution functions reveal a heavier tail, moreover the optimal threshold levels decreases, see Table 1.

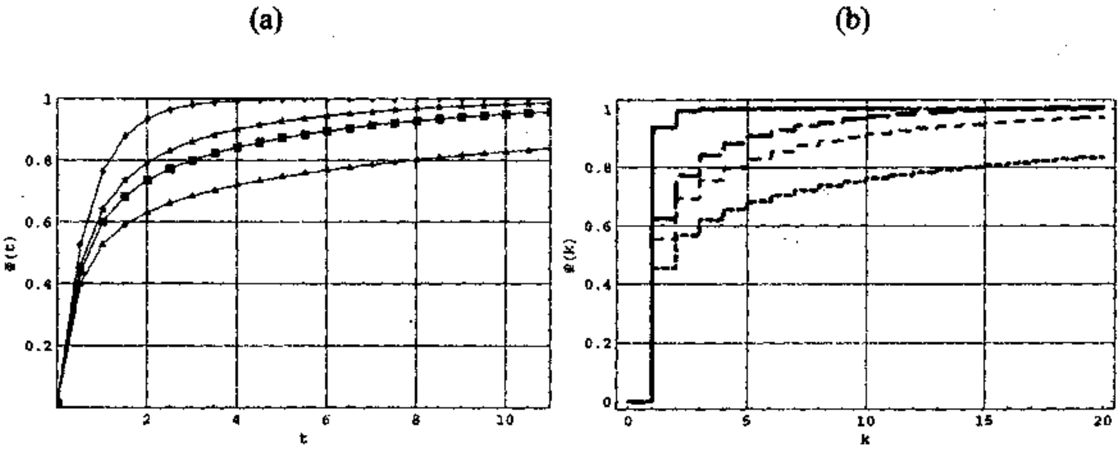

Fig. 1. The distributions $\Phi(t)$ and $\Psi(k)$ versus λ

Table 1 **Moments of L and T with corresponding optimal thresholds**

λ	$\mathbb{E}[L]$	$\mathbb{V}[L]$	$\mathbb{E}[N]$	$\mathbb{V}[N]$	q_2^*	q_3^*
0.1	0.7142	0.5831	1.0714	0.0875	4	11
0.9	1.5091	6.3237	2.5098	10.7093	3	6
1.2	2.2004	16.1579	3.6470	34.3611	2	5
1.8	6.3645	202.4630	8.0992	207.6810	2	4

REFERENCES

1. *Efrosinin D.* Controlled queueing systems with heterogeneous servers. VDM Verlag Dr. Müller, Saarbrücken. 2008.
2. *Efrosinin D., Rykov V.* On performance characteristics for queueing systems with heterogeneous servers // Autom. And Remote Control. 2008. № 1. P. 64–82.
3. *Lin W., Kumar P. R.* Optimal control of a queueing system with two heterogeneous servers // IEEE Trans. on Automatic Control, 1984. № 29. P. 696–703.
4. *Rykov V.* Monotone Control of Queueing Systems with Heterogeneous Servers // Queueing Systems. 2001. № 37. P. 391–403.