Variable Size Packets Analysis in Load-balanced Switch with Finite Buffers

Yury Audzevich*, Levente Bodrog†, Yoram Ofek*, Miklós Telek† and Bülent Yener‡
* Department of Information Engineering and Computer Science,
University of Trento, Trento, Italy
Email: {audzevi,ofek}@disi.unitn.it
†Department of Telecommunications,
Budapest University of Technology and Economics, Budapest, Hungary
Email: {bodrog,telek}@hit.bme.hu
‡ Department of Computer Science,
Rensselaer Polytechnic Institute, Troy, New York, USA
Email: yener@cs.rpi.edu

***Abstract*—The Internet represents a complex asynchronous network operating with variable length packets which is strongly related to the application used. On the other hand, the requirements for the transmission and capacity characteristics of the Internet are rapidly increasing. Due to simple distributed control and high scalability Load-Balancing (LB) Birkhof-von Neumann switch appears to be a promising switch architecture. The previous research was focused on the assumption of unlimited amount of buffering and the transmission of fixed-size packets or cells. This paper analyzes packet and cell loss probabilities of the Load-Balanced switch operating with 1) *variable size packets* and 2) *finite central stage buffers*. In the course of the analysis we recognized a previously unpublished feature of the LB switch which is the asymmetry of different paths through the switch. This behavior has major impact on the fairness and loss inside the switch.**

I. Introduction

The single-stage load-balanced switch (LB switch) was recently presented in [1] and [2] and has remained one of the most relevant topics due to its attractive scalability properties. The first significant results in [1] and [2] was that under certain assumptions the switch can achieve high throughput and can be highly scalable, while keeping simple distributed control. However, LB switch have some problems of its own. One of the critical issues recently raised in the research appears to be mis-sequenced arrival of equal size packets (from now on we refer them simply as *cells*), while transmitting packets through the switch. Keeping correct sequence of packets through the system avoids unnecessary retransmissions of packets in the network protocol layer. The main efforts to resolve cell mis-sequencing were proposed in [2]. Each new solution for out-of-order packet arrival problem had increased the complexity of the LB switch. For example, the Mailbox switch [4] and the Contention and Reservation switch [8] have introduced a novel symmetric interconnection pattern for crossbar switches, which provides information feedback links with extra communication and computational overheads. Additional works resolving mis-sequencing problem were presented in [7], where matching algorithms and feedback between stages were used for appropriate packets exchange between the stages. Complexity and performance of this switch depends on the matching algorithm used. Finally, the Byte-Focal switch [5] does not require any feedback between the stages and has a simple controllable cell re-sequencer at the output stage. However, it uses large buffering space for resolution, which may not be scalable for large number of ports (N). Efficiency of mis-sequencing algorithm is important to take into account for LB switch with variable size packets, since inefficient algorithm can create extra delays and packet loss at the reassembly unit.

In most of the presented papers some strong assumptions are used. In particular, it is considered that the central stage buffers inside the switch are infinite, the packets coming through the system are of the same size (cells) [3], and that the traffic pattern obeys the admissibility conditions. The throughput analysis of the two stage load-balanced switch with finite buffers was presented in [9], in which the authors have shown that even under admissible traffic patterns the throughput cannot reach 100%. Unfortunately, in [9], only simulation results were presented. In the analysis presented in [6] there has been introduced the possibility of cell drop in the finite central stage buffers. It also provides solution for calculation of cell loss probability both for admissible and inadmissible traffic patterns. However, the analysis in [6] was done only for fixed size packets (cells), and there was not taken into account variable size packets.

The main goal of this paper is the analysis of packet (of variable size) loss in the internal LB switch. We also prove the significance of variable size packets consideration. We assume Markovian behavior to be able to use numerically efficient algorithms to solve Markov chains. In our case this means geometrically distributed packet and idle period lengths, which allows us to capture the mean of these distributions. Real internet traffic shows different packet size distributions [11], [12] and one can fit more parameters using other, more complex Markovian structures like discrete Phase Type (DPH) distributions or discrete Markovian arrival processes

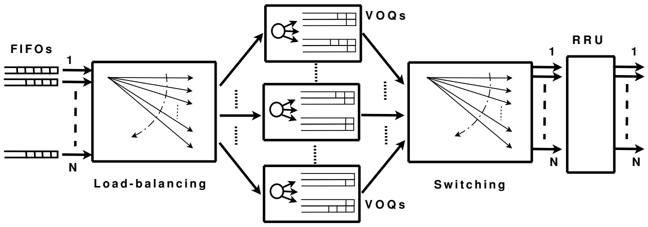


Fig. 1. The load-balanced switch considered for the analysis

(DMAPs). The number of fitted parameters can be increased at an arbitrary level, but it would greatly increase the complexity of the model as well and that would also hide the main contribution of our approach.

The single-stage buffering LB switch is equipped with First-In-First-Out (FIFO) buffers in the inputs and re-sequencing and reassembly units (RRU) in the output (see the illustration in Fig. 1). The implementation of RRU is not discussed in this paper, but it can be used as one among the proposed in research, like in [10]. In this analysis there is no feedback link between the switch stages and each stage is operating independently.

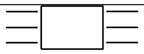
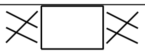
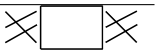
A. The main analysis assumptions

We assume finite FIFO buffers at the inputs which in fact are large enough to store as much cells as the loss probability remains under a predefined threshold. And in any case the dropped packets are not considered in the central stage which is the main focus of this paper. If a packet is dropped at the input buffer all of its cells are dropped and the whole packet is reset by the network layer and can be retransmitted by a network protocol. On the other hand if a single cell is dropped in the central stage, there is no possibility to drop all the remaining cells of this "broken" packet without sophisticated centralized controller (which is not the case of this paper). Such "broken" packets will definitely make impossible of RRU operation like is it introduced in [10]. Each variable size packet arrives with variable rate, this rate is always less than service rate of cells inside the switch – the switch is not overloaded. After the arrival of a variable size packet there is a possible idle period (measured in time slots). Cell transmissions inside the switch have a fixed rate. The destination outputs of the packets are chosen uniformly among all the available outputs (this is given in $\mathbf{T}$ parameter in our analysis). The analysis is done without any respect to cells mis-sequencing inside the switch and packets reassembly. The main goal of the presented analysis is to show the amount of packet/cell loss experienced by a single *central stage virtual output queue (VOQ).*

Moreover, as shown in our analytical results, the packet loss probability of a VOQ strongly depends on the specific traversing path of the traffic inside the switch, which is an interesting phenomenon described in Section II-A for the interconnection pattern applied. We analyze the least preferred path among the possibilities.

The rest of the paper is organized as follows. We give the detailed analysis of a switch with $N = 3$ input and output ports in Section II, and right after we give the algorithmic description of the same process for the switch with arbitrary number (N) of ports in Section III. In Section IV we verify the result by comparing it with simulations. Finally Section V concludes the paper.

TABLE I
THE THREE POSSIBLE SETTINGS OF 3×3 LB SWITCH

$t \mod N$	0	1	2
switch state			

II. ANALYSIS OF 3×3 LOAD-BALANCED SWITCH

Throughout the paper we will use the notation $n \times m$ to denote a switch with n input and m output ports or simply refer them as a switch of size N if there are both N input and output ports.

In this section we give the detailed model of the switch of size 3 since it has all the important properties of the general switch of size N. In our analysis we consider VOQ_{00}, to which arrivals are possible from all three inputs directed to output 0.

The crossbar interconnections between input i, cental stage VOQ_{kj}, and output j obey the rules

$$\begin{aligned} k &= i + t \mod N \\ j &= k + t \mod N \end{aligned} \tag{1}$$

respectively. Due to this, the interconnection pattern of the switch has a periodic behavior with length N. In case of size 3 the possible interconnection settings are summarized in Table I.

In a time slot, firstly, the VOQs are connected to the outputs and then the inputs to the VOQs. This order of interconnections inhibits a cell from traverse the switch in a single time slot.

As the packets are segmented into fix-sized cells the arrival process to a VOQ can be described by a discrete time Markov chain (DTMC) on the cell level (see Section II-B). Extending the DTMC with two absorbing states we are able to model the system on the packet level (Section II-C). Having the packet level model we give its solution in Section II-D.

In Section II-A we will show our observations on the different behavior of the different paths through the switch.

A. Properties of the different paths

Using (1) the difference between the service of a given VOQ and the arrival to it can be expressed as

$$\Delta t = 2k - i - j \mod N. \tag{2}$$

This implies the difference between different paths as, in case of a given VOQ, it depends on the ordinal number of both the input and the output. The possible Δt values are $0, 1, \ldots, N-1$ time slots.

There are two important consequences of this observation.

a) Differences in cell loss: Depending on the value of Δt there can be N values of cell loss probabilities – $p_{c_l} \quad l \in [0, N[$, as well as N values of the packet loss probabilities – $p_{p_l} \quad l \in [0, N[$. The following set of inequalities holds for them

$$\begin{aligned} p_{c_0} = 0 \leq p_{c_1} \leq \ldots \leq p_{c_{N-1}}, \\ p_{p_0} = 0 \leq p_{p_1} \leq \ldots \leq p_{p_{N-1}}, \end{aligned} \tag{3}$$

which is explained by the fact that the higher the Δt value the higher the loss probability value.

In a time period, Δt inputs have the right to send a cell to the observed VOQ before the observed input. Consequently, the higher the Δt value the larger the chance that there are enough inputs sending cells before the observed input to fill up the queue.

b) Differences in other performance measures: The input – output pairs with higher Δt value will always suffer from higher delay as well as higher drop probability.

This way the interconnection patterns of the input and output crossbars determine the fairness of input – output pairs. Assuming the crossbars are set to provide a symmetric (fair) chance for the input – output pairs then the cells of an input – output pair suffers different loss at the different VOQs. This way the loss between an input – output pair is dominated by the VOQ where it has the maximal Δt value. I.e., the worst case dominates the loss. To the best of our knowledge this property has not been reported yet.

Due to this phenomenon it is not irrelevant which configuration is analyzed. We analyze the path with the maximal Δt value (worst case) in the following sections, and comment on the behaviour of the other cases in Section IV.

B. The cell level model

In the following we describe the model of the 3×3 switch. We have chosen to model VOQ_{00} – the first sub-queue of the first set of VOQs – or, more specifically, the path input $1 \rightarrow VOQ_{00} \rightarrow$ output 0. If one substitutes the ordinal number of these ports and virtual queue into (2) it will result in $2 \cdot 0 - 1 - 0 \mod N = 2$, which the highest loss values corresponds to.

VOQ_{00} is "fed" by three input processes with geometric distributed packet lengths. Each input can have packets destined to the different outputs.

First we model the operating mechanism on the cell level by building the appropriate DTMCs for the ith input (see Fig. 2). Each DTMC has four states denoted as follows

ij the states responsible for cell-arrivals from input i to output j and

$i\,id$ the state responsible for the idle period of input i.

The state transitions describe the beginning of either a new packet or the idle period or the continuation of an incomplete packet. The graph of the DTMC modeling the ith input is given in Fig. 2 and the transition probability matrix of the ith input is given in (4), i.e. for example the substitution of $i = 1$ will result in the graph together with the state transition probability matrix of the second, observed, input. The meaning of probabilities appearing in the DTMC according to the state transition probabilities are

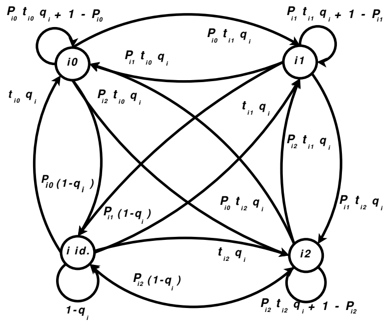


Fig. 2. The DTMC modeling input i

- $p_{ij} q_i t_{ij}$ is the probability that a packet from the ith input to the jth output arrives in the actual time slot,
- $1 - p_{ij}$ is the probability that the packet from the ith input to the jth output is still in progress,
- $p_{ij}\,(1 - q_i)$ is the probability that the packet from input i to output j is ended and the ith input changes to idle,
- input i remains in idle state with probability $1 - q_i$ in the actual time slot and
- $t_{ij} q_i$ is the probability that the next packet will be sent from input i to output j after the idle.

Hereinafter $\mathbf{T} = (\,t_{ij}\,)$ denotes the probabilities that a packet arrives from input i to output j. $\mathbf{P} = (\,p_{ij}\,)$ are the parameters of the geometric distributed packet length arrivals from input i to output j. $\mathbf{q} = (\,q_i\,)$ is the vector containing the parameters of the geometric distributed idle period of input i.

Now we split the state transition probability matrix of the ith input into two terms

$$\mathbf{P}_i = \mathbf{A}_i + \mathbf{K}_i \tag{5}$$

The first term includes state transitions responsible for cell arrival to output 0 – its first row equals to the first row of $\mathbf{P}_i$ and it is 0 otherwise. The second term includes the other cases, which has zeros in the first row and equals to $\mathbf{P}_i$ otherwise.

The system behavior in three consecutive time slots is described by a DTMC embedded right before the service of the VOQ at every $t \mod N = 0$th time instance. Each input (during the three time slots long period) is modeled by the third power of its state transition probability matrix. The joint behavior of the three inputs in the period is described by the Kronecker product of the third power of the state transition probability matrices of each input. It is

$$\boldsymbol{\mathcal{P}} = \mathbf{P}_0^3 \otimes \mathbf{P}_1^3 \otimes \mathbf{P}_2^3, \tag{6}$$

which is the phase process of a quasi birth-deathlike (QBD-like) process describing the queue length of the observed VOQ.

In addition to (6)

$$\boldsymbol{\mathcal{P}} = \mathbf{B} + \mathbf{L} + \mathbf{F}_1 + \mathbf{F}_2 \tag{7}$$

also holds for the phase process of the same QBD-like. Here $\mathbf{B}$ is the backward, $\mathbf{L}$ is the local and $\mathbf{F}_k$s are the set of forward level transition matrices (in our 3×3 case $k = 1, 2$).

$$\mathbf{P}_i = \begin{pmatrix} (1-p_{i0})+p_{i0}q_it_{i0} & p_{i0}q_it_{i1} & p_{i0}q_it_{i2} & p_{i0}\left(1-q_i\right) \\ p_{i1}q_it_{i0} & (1-p_{i1})+p_{i1}q_it_{i1} & p_{i1}q_it_{i2} & p_{i1}\left(1-q_i\right) \\ p_{i2}q_it_{i0} & p_{i2}q_it_{i1} & (1-p_{i2})+p_{i2}q_it_{i2} & p_{i2}\left(1-q_i\right) \\ q_it_{i0} & q_it_{i1} & q_it_{i2} & 1-q_i \end{pmatrix} \tag{4}$$

In the next step we substitute (5) into (6), expand it, identify the terms corresponding to 0, 1, 2 and 3 cell arrivals to VOQ_{00} and we match its subexpressions to the terms of (7). All these are given in (8).

In detail, one factor of the third powers in (6) is substituted by (5). Namely, in case of input 0 it is the first factor because it can send a cell to the observed central stage queue in the first time slot of the cycle (according to Table I). In case of input 1(2) the third(second) factor is substituted. Once the substitution is done and the expansion is executed the terms are collected based on "the number of $\mathbf{A}$s appearing in it" which equals to the number of cell arrivals to the observed VOQ.

After all manipulations we obtain (8) in which we also indicate the meaning of the terms. There is one cell served at the beginning of a period which results in one level transition backward in case of 0 cell arrivals, stay on the same level in case of 1 cell arrivals and 1(2) level transition(s) forward in case of 2(3) cell arrivals.

Finally we give the irregular levels of the QBD-like process. In the first irregular level, when the central stage queue is empty, the DTMC can have $0, 1, 2$ and 3 level transitions according to $\mathbf{B}, \mathbf{L}, \mathbf{F}_1,$ and $\mathbf{F}_2$ respectively. In case of a full buffer the level process remains in the bth level instead of level transitions forward. According to this the forward level transition matrix in the level before the last one is $\mathbf{F}_1' = \mathbf{F}_1 + \mathbf{F}_2$ and the local state transition in the last level is $\mathbf{L}' = \mathbf{L} + \mathbf{F}_1 + \mathbf{F}_2$.

Then the state transition probability matrix of the QBD-like process on the block level is

$$\mathbf{P} = \begin{pmatrix} \mathbf{B} & \mathbf{L} & \mathbf{F}_1 & \mathbf{F}_2 & 0 & \dots & 0 \\ \mathbf{B} & \mathbf{L} & \mathbf{F}_1 & \mathbf{F}_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \mathbf{B} & \mathbf{L} & \mathbf{F}_1 & \mathbf{F}_2 \\ 0 & 0 & \dots & 0 & \mathbf{B} & \mathbf{L} & \mathbf{F}_1' \\ 0 & 0 & 0 & \dots & 0 & \mathbf{B} & \mathbf{L}' \end{pmatrix}. \tag{9}$$

Its steady state solution is the solution of the linear equation system

$$\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi} \qquad \mathbf{P}\mathbf{h} = \mathbf{h}, \tag{10}$$

where $\mathbf{h}$ is the appropriate size column vector of ones.

C. The packet level model

We model the system on the packet level by a DTMC describing the life-cycle of a packet. Such an absorbing DTMC is described by its initial distribution and state transition probability matrix. In this section we give these properties.

This is obtained by the introduction of two absorbing states appended to the modified QBD-like model of the virtual queue. As these absorbing states correspond to the two possible endings of the packet transmission – either the packet is lost or transmitted successfully – this new transient DTMC will describe the life cycle of the packet.

The Markov model of the system in Fig. 3 consists of these parts, one is the revised QBD-like model of the virtual output queue and there are two absorbing states appended to the QBD-like part. The absorbing state ST corresponds to the successful packet transmission in the observed path of the LB switch and CL to the first cell loss and also to the packet loss. In the next sections we will discuss these parts of the model into details.

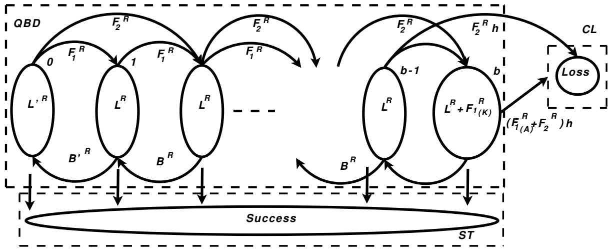


Fig. 3. The transient DTMC for the packet level

1) Modifications to the QBD-like process: Here we introduce the notation superscript $*^{\mathcal{R}}$ which denotes properties of the transient DTMC introduced in this section.

In the following we describe the revised QBD-like model of VOQ_{00} when a packet is present in the system. The revision covers the determination of the state transition probabilities to the two absorbing states. First we remove state transition probabilities from $\mathbf{P}_1$ according to the successful packet transmission. Later on, these probabilities will be added as state transitions to the absorbing state ST in Section II-C4. In practice it means that the QBD-like model will be determined based on the revised DTMC description of input 1 given in Fig. 4 and in (11).

The DTMCs of the other two inputs remain the same as in Fig. 2 and in (4) since the observed path contains only input 1.

The determination of the state transition probabilities to absorbing state CL will be given later in Section II-C2.

Similar to Section II-B we consider the switch operation during three consecutive time slots. Now we split $\mathbf{P}_1^{\mathcal{R}}$ into the two similar terms as in (5)

$$\mathbf{P}_1^{\mathcal{R}} = \mathbf{A}_1^{\mathcal{R}} + \mathbf{K}_1^{\mathcal{R}}. \tag{12}$$

Similar to the cell level QBD-like model we calculate the state transition probability matrix of the phase process in two different ways and do the matching between the two expressions in (13).

$$
\begin{aligned}
\boldsymbol{\mathcal{P}} &= \mathbf{P}_0^3 \otimes \mathbf{P}_1^3 \otimes \mathbf{P}_2^3 = \left(\mathbf{A}_0 + \mathbf{K}_0\right)\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\left(\mathbf{A}_1 + \mathbf{K}_1\right) \otimes \mathbf{P}_2\left(\mathbf{A}_2 + \mathbf{K}_2\right)\mathbf{P}_2 = \\
&= \underbrace{\mathbf{A}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{K}_1 \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2 + \mathbf{K}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{A}_1 \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2 + \mathbf{K}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{K}_1 \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2}_{\text{1 arrival}} + \\
&+ \underbrace{\mathbf{K}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{A}_1 \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \mathbf{A}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{K}_1 \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \mathbf{A}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{A}_1 \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2}_{\text{2 arrivals}} + \\
&+ \underbrace{\mathbf{A}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{A}_1 \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2}_{\text{3 arrivals}} + \underbrace{\mathbf{K}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{K}_1 \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2}_{\text{no arrivals}} = \mathbf{L} + \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{B}
\end{aligned} \tag{8}
$$

$$
\mathbf{P}_1^{\mathcal{R}} = \begin{pmatrix}
(1 - p_{10}) & 0 & 0 & 0 \\
p_{11}q_1t_{10} & (1 - p_{11}) + p_{11}q_1t_{11} & p_{11}q_1t_{12} & p_{11}\left(1 - q_1\right) \\
p_{12}q_1t_{10} & p_{12}q_1t_{11} & (1 - p_{12}) + p_{12}q_1t_{12} & p_{12}\left(1 - q_1\right) \\
q_1t_{10} & q_1t_{11} & q_1t_{12} & 1 - q_1
\end{pmatrix}. \tag{11}
$$

$$
\begin{aligned}
\boldsymbol{\mathcal{P}} &= \mathbf{P}_0^3 \otimes {\mathbf{P}_1^{\mathcal{R}}}^3 \otimes \mathbf{P}_2^3 = \left(\mathbf{A}_0 + \mathbf{K}_0\right)\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\left(\mathbf{A}_1^{\mathcal{R}} + \mathbf{K}_1^{\mathcal{R}}\right) \otimes \mathbf{P}_2\left(\mathbf{A}_2 + \mathbf{K}_2\right)\mathbf{P}_2 = \\
&= \underbrace{\mathbf{A}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{K}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2 + \mathbf{K}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{A}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2 + \mathbf{K}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{K}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2}_{\text{1 arrival}} + \\
&+ \underbrace{\mathbf{K}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{A}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \mathbf{A}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{K}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \mathbf{A}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{A}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2}_{\text{2 arrivals}} + \\
&+ \underbrace{\mathbf{A}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{A}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2}_{\text{3 arrivals}} + \underbrace{\mathbf{K}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{K}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2}_{\text{no arrivals}} = \mathbf{L}^{\mathcal{R}} + \mathbf{F}_1^{\mathcal{R}} + \mathbf{F}_2^{\mathcal{R}} + \mathbf{B}^{\mathcal{R}}
\end{aligned} \tag{13}
$$

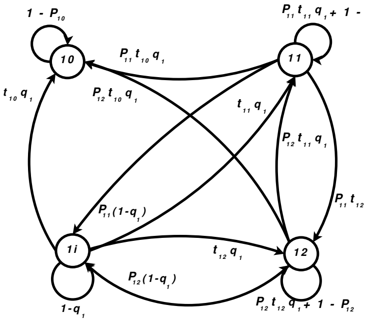


Fig. 4. The revised DTMC model of input 1

The first irregular level will be similar to that in (9). In the last level instead of one level transition forward there can be two cases. In the first case if there is no arrival from input 1 the system will stay on the same level, otherwise there will be a cell loss. According to these two cases $\mathbf{F}_1^{\mathcal{R}}$ is split into two terms

$$
\begin{aligned}
\mathbf{F}_1^{\mathcal{R}} = \Big(&\mathbf{K}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{A}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \\
&+ \mathbf{A}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{A}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2\Big) + \\
&+ \left(\mathbf{A}_0\mathbf{P}_0^2 \otimes {\mathbf{P}_1^{\mathcal{R}}}^2\mathbf{K}_1^{\mathcal{R}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2\right) = \mathbf{F}_1^{\mathcal{R}(\mathcal{A})} + \mathbf{F}_1^{\mathcal{R}(\mathcal{K})}.
\end{aligned} \tag{14}
$$

The first term stands for arrival and the other for no arrival. Using this $\mathbf{L}^{\mathcal{R}\prime} = \mathbf{L}^{\mathcal{R}} + \mathbf{F}_1^{\mathcal{R}(\mathcal{K})}$ and the state transition probability matrix of the QBD-like part is

$$
\hat{\mathbf{P}}^{\mathcal{R}} = \begin{pmatrix}
\mathbf{B}^{\mathcal{R}} & \mathbf{L}^{\mathcal{R}} & \mathbf{F}_1^{\mathcal{R}} & \mathbf{F}_2^{\mathcal{R}} & 0 & \dots & 0 \\
\mathbf{B}^{\mathcal{R}} & \mathbf{L}^{\mathcal{R}} & \mathbf{F}_1^{\mathcal{R}} & \mathbf{F}_2^{\mathcal{R}} & 0 & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & \dots & 0 & \mathbf{B}^{\mathcal{R}} & \mathbf{L}^{\mathcal{R}} & \mathbf{F}_1^{\mathcal{R}} & \mathbf{F}_2^{\mathcal{R}} \\
0 & 0 & \dots & 0 & \mathbf{B}^{\mathcal{R}} & \mathbf{L}^{\mathcal{R}} & \mathbf{F}_1^{\mathcal{R}} \\
0 & 0 & 0 & \dots & 0 & \mathbf{B}^{\mathcal{R}} & \mathbf{L}^{\mathcal{R}\prime}
\end{pmatrix}. \tag{15}
$$

2) The packet loss: There can be cell loss (or equivalently packet loss) in the system in two cases

- either if the queue length is $b - 1$ at the beginning of the cycle and there are three arrivals to VOQ_{00}
- or if the queue is full and there is arrival from input 1 to VOQ_{00}.

Appending the absorbing state CL to the QBD-like part and collecting the state transition probabilities to CL according to the two above cases we can build up the transpose of the state transition probability vector to CL as

$$
\mathbf{l}^{\mathsf{T}} = \left(0 \quad \dots \quad 0 \quad \left(\mathbf{F}_2^{\mathcal{R}}\mathbf{h}\right)^{\mathsf{T}} \quad \left(\left(\mathbf{F}_1^{\mathcal{R}(\mathcal{A})} + \mathbf{F}_2^{\mathcal{R}}\right)\mathbf{h}\right)^{\mathsf{T}}\right), \tag{16}
$$

where $\mathbf{h}$ is the appropriate size column vector of ones. Appending state CL to the QBD-like results in state transition probability matrix $\tilde{\mathbf{P}}^{\mathcal{R}} = \left(\begin{array}{c|c} \hat{\mathbf{P}}^{\mathcal{R}} & \mathbf{l} \\ \hline \multicolumn{2}{c}{\mathbf{e}_l^{\mathsf{T}}} \end{array}\right)$, where $\mathbf{e}_l^{\mathsf{T}} =$

$\left(\begin{smallmatrix}0 & \dots & 0 & 1\end{smallmatrix}\right)$ is the transpose of the last unit vector with appropriate size since CL is an absorbing state.

The introduced analytical approach is applicable for the analysis with different values of Δt (see (2)) with some modifications. The first is the modification of the block matrices in the last column of $\hat{\mathbf{P}}^{\mathcal{R}}$ – the irregular level matrices – in (15).

The second modification is the modification of the state transition probability vector to CL (l) as follows

- $\Delta t = 0$ no state transitions to CL and
- $\Delta t = 1$ state transitions to CL only from the last level.

There are other differences between analysis of the different types of paths mentioned in Section II-A (e.g. technically how the equations are built), but these two are the essential differences.

3) The cell loss: To calculate the cell loss of the system one should create $\mathbf{F}_1^{(\mathcal{A})}$ analogously to $\mathbf{F}_1^{\mathcal{R}(\mathcal{A})}$ by rearranging the term for $\mathbf{F}_1$ from (8) as

$$\begin{aligned}\mathbf{F}_1 = \Big(&\mathbf{K}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{A}_1 \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \\ &+ \mathbf{A}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{A}_1 \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2\Big) + \\ &+ \Big(\mathbf{A}_0\mathbf{P}_0^2 \otimes \mathbf{P}_1^2\mathbf{K}_1 \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2\Big) = \mathbf{F}_1^{(\mathcal{A})} + \mathbf{F}_1^{(\mathcal{K})}.\end{aligned}$$

It gives two terms, the first stands for the case when there is arrival from input 1 to VOQ_{00} and the second when there is no arrival.

Having $\mathbf{F}_1^{(\mathcal{A})}$ and using $\mathbf{F}_2$ from (8) the cell loss is

$$p_c = \boldsymbol{\pi}_{b-1}\mathbf{F}_2\mathbf{h} + \boldsymbol{\pi}_b\left(\mathbf{F}_1^{(\mathcal{A})} + \mathbf{F}_2\right)\mathbf{h}, \tag{17}$$

where b is the buffer size and $\boldsymbol{\pi}_l$ is the $l+1$st sub-vector – with length $N+1$ – of $\boldsymbol{\pi}$ given in (10) and $\mathbf{h}$ is the appropriate size column vector of ones.

4) The successful packet transmission: The DTMC absorbs in state ST if the last cell of a packet is transmitted successfully as well as all the other cells of it. The preceding parts of this model (the DTMC) do not contain the state transitions responsible for packet ending (see Fig. 4 and (11)). Accordingly the vector containing the probabilities to change state to ST, i.e. successful packet transmission, is calculated as

$$\mathbf{s} = \mathbf{h} - \tilde{\mathbf{P}}^{\mathcal{R}}\mathbf{h}. \tag{18}$$

By appending state ST to the DTMC we get

$$\mathbf{P}^{\mathcal{R}} = \left(\begin{array}{c|c|c} \hat{\mathbf{P}}^{\mathcal{R}} & \mathbf{l} & \mathbf{s} \\ \hline \multicolumn{2}{c|}{\mathbf{e}_l^{\mathsf{T}}} & 0 \\ \hline \multicolumn{3}{c}{\mathbf{e}_l'^{\mathsf{T}}} \end{array}\right) \tag{19}$$

since the state transition probability matrix of the DTMC contains two absorbing states (appearing in Fig. 3). Here $\mathbf{e}_l'^{\mathsf{T}} = \left(\begin{smallmatrix}0 & \dots & 0 & 1\end{smallmatrix}\right)$ is the transpose of the appropriate size last unit vector.

D. The probabilities of packet loss and successful packet transmission

The packet loss probability is given as the probability of absorbing in state CL

$$p_l = \boldsymbol{\pi}^{\mathcal{N}}\left(\mathbf{I} - \hat{\mathbf{P}}^{\mathcal{R}}\right)^{-1}\mathbf{l} \tag{20}$$

and the probability of successful packet transmission is the probability of absorbing in state ST

$$p_s = \boldsymbol{\pi}^{\mathcal{N}}\left(\mathbf{I} - \hat{\mathbf{P}}^{\mathcal{R}}\right)^{-1}\mathbf{s}, \tag{21}$$

where $\mathbf{I}$ is the appropriate size identity matrix. $\boldsymbol{\pi}^{\mathcal{N}}$ is the initial probability distribution of the system immediately after a new packet arrival from input 1.

1) The initial distribution of the system: The system is considered to be in the steady state when a new packet arrives. Then the initial distribution $\boldsymbol{\pi}^{\mathcal{N}}$ of the system is given as the probabilities being in each state right after a new packet arrival. Quantities with superscript $*^{\mathcal{N}}$ describes the system in this state.

In the $4 \le i \le b-1$st regular level the initial states are

$$\hat{\boldsymbol{\pi}}_i^{\mathcal{N}} = \boldsymbol{\pi}_{i-2}\mathbf{F}_2^{\mathcal{N}} + \boldsymbol{\pi}_{i-1}\mathbf{F}_1^{\mathcal{N}} + \boldsymbol{\pi}_i\mathbf{L}^{\mathcal{N}} + \boldsymbol{\pi}_{i+1}\mathbf{B}^{\mathcal{N}} \tag{22}$$

and in the irregular levels are

$$\hat{\boldsymbol{\pi}}_0^{\mathcal{N}} = \boldsymbol{\pi}_0\mathbf{B}^{\mathcal{N}} + \boldsymbol{\pi}_1\mathbf{B}^{\mathcal{N}} \tag{23}$$

$$\hat{\boldsymbol{\pi}}_1^{\mathcal{N}} = \boldsymbol{\pi}_0\mathbf{L}^{\mathcal{N}} + \boldsymbol{\pi}_1\mathbf{L}^{\mathcal{N}} + \boldsymbol{\pi}_2\mathbf{B}^{\mathcal{N}} \tag{24}$$

$$\hat{\boldsymbol{\pi}}_2^{\mathcal{N}} = \boldsymbol{\pi}_0\mathbf{F}_1^{\mathcal{N}} + \boldsymbol{\pi}_1\mathbf{F}_1^{\mathcal{N}} + \boldsymbol{\pi}_2\mathbf{L}^{\mathcal{N}} + \boldsymbol{\pi}_3\mathbf{B}^{\mathcal{N}} \tag{25}$$

$$\hat{\boldsymbol{\pi}}_3^{\mathcal{N}} = \boldsymbol{\pi}_0\mathbf{F}_2^{\mathcal{N}} + \boldsymbol{\pi}_1\mathbf{F}_2^{\mathcal{N}} + \boldsymbol{\pi}_2\mathbf{F}_1^{\mathcal{N}} + \boldsymbol{\pi}_3\mathbf{L}^{\mathcal{N}} + \boldsymbol{\pi}_4\mathbf{B}^{\mathcal{N}} \tag{26}$$

$$\begin{aligned}\hat{\boldsymbol{\pi}}_b^{\mathcal{N}} &= \boldsymbol{\pi}_{b-2}\mathbf{F}_2^{\mathcal{N}} + \boldsymbol{\pi}_{b-1}\Big(\mathbf{F}_1^{\mathcal{N}} + \mathbf{F}_2^{\mathcal{N}}\Big) + \\ &\quad + \boldsymbol{\pi}_b\Big(\mathbf{L}^{\mathcal{N}} + \mathbf{F}_1^{\mathcal{N}} + \mathbf{F}_2^{\mathcal{N}}\Big),\end{aligned} \tag{27}$$

where $\boldsymbol{\pi}$ is the steady state solution of the cell level model given in (10). $\mathbf{B}^{\mathcal{N}}$, $\mathbf{L}^{\mathcal{N}}$ and $\mathbf{F}_i^{\mathcal{N}}$ are the level transition matrices of a QBD-like model describing the system right after a new packet arrival.

This QBD-like model is built up in a similar way to that in Section II-C1. The difference is that the model of input 1 is containing only state transitions corresponding to a new packet arrival as it is shown in Fig. 5 and given as

$$\mathbf{P}_1^{\mathcal{N}} = \begin{pmatrix} p_{10}t_{10}q_1 & 0 & 0 & 0 \\ p_{11}q_1t_{10} & 0 & 0 & 0 \\ p_{12}q_1t_{10} & 0 & 0 & 0 \\ q_1t_{10} & 0 & 0 & 0 \end{pmatrix}. \tag{28}$$

The models of input 0 and 2 are the same as shown in Fig. 2.

$\mathbf{B}^{\mathcal{N}}$, $\mathbf{L}^{\mathcal{N}}$ and $\mathbf{F}_k^{\mathcal{N}}$ $\quad k = 1, 2$ are determined similar to the preceding cases. However, the behavior of input 1 needs some more considerations according to the new packet arrival.

Since the packet can arrive in all three time slots the state transition probability matrix of input 1 in a period is reconsidered based on Table II. Its notations are:

- $+$ denotes the arrival of a packet in a time slot,

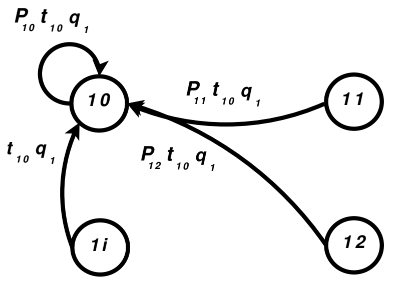


Fig. 5. The DTMC of input 1, a new packet arrival

TABLE II
THE POSSIBLE TIME EVOLUTION OF INPUT 1 WITH PACKET ARRIVAL

t mod $3 =$ 0	 1	 2	three consecutive time slots
$+$	$+$	$+$	$\mathbf{P}_1^2\mathbf{P}_1^{\mathcal{N}}$
$-$	$+$	$+$	
$+$	$-$	$+$	
$-$	$-$	$+$	
$+$	$+$	$-$	$\left(\mathbf{P}_1\mathbf{P}_1^{\mathcal{N}} + \mathbf{P}_1^{\mathcal{N}}\left(\mathbf{P}_1 - \mathbf{P}_1^{\mathcal{N}}\right)\right)\left(\mathbf{P}_1 - \mathbf{P}_1^{\mathcal{N}}\right)$
$-$	$+$	$-$	
$+$	$-$	$-$	

- the arrival instance(s) in a period is on the left hand side,
- the corresponding state transition probability matrix is on the right hand side.

Here again we split $\mathbf{P}_1^{\mathcal{N}}$ into two parts

$$\mathbf{P}_1^{\mathcal{N}} = \mathbf{A}_1^{\mathcal{N}} + \mathbf{K}_1^{\mathcal{N}}, \tag{29}$$

containing two terms. First stands for the cell arrival and the second for no arrival.

Based on Table II, using (29) and (5) for $i = 1$, we give the matrices describing input 1 in a period if there is arrival – $\boldsymbol{\mathcal{A}}_1^{\mathcal{N}}$ – and if there are no arrivals – $\boldsymbol{\mathcal{K}}_1^{\mathcal{N}}$.

Now the state transition probability matrix of the phase process can be expressed in two different ways using (30) and (5) (for $i = 0, 2$). The matching of the terms is resulting in the level transition matrices of this QBD-like process. It is given in (31).

Similarly to (8), (31) makes mapping with $\mathbf{B}^{\mathcal{N}}$, $\mathbf{L}^{\mathcal{N}}$, $\mathbf{F}_1^{\mathcal{N}}$ and $\mathbf{F}_2^{\mathcal{N}}$ matrices. Substituting them into the expressions (22) through (27) and normalizing them by $c = \sum_{i=0}^{b} \hat{\boldsymbol{\pi}}_i^{\mathcal{N}}\mathbf{h}$ we obtain

$$\boldsymbol{\pi}_i^{\mathcal{N}} = \frac{1}{c}\hat{\boldsymbol{\pi}}_i^{\mathcal{N}}, \quad 0 \le i \le b, \tag{32}$$

the individual parts of the initial distribution of the transient DTMC in Fig. 3 modeling the system on the packet level. Combining the parts together we get $\boldsymbol{\pi}^{\mathcal{N}}$ – the initial distribution.

Finally substituting (15), (16), (18) and $\boldsymbol{\pi}^{\mathcal{N}}$ into (20) and (21) we get the packet loss probability and the probability of successful packet transmission.

The model presented in this Section describes a path with the highest loss probability according to Section II-A. And we just flash the results of the other two kind of paths in Section IV on Fig. 6.

TABLE III
THE PARAMETERS OF THE CONSIDERED SWITCH

variable	value
N	3
p_{ij}	0.2 (av. 5 cells)
q_i	0.9 (av. 1.1 cells)
t_{ij}	$\frac{1}{N}$

III. ANALYSIS OF $N \times N$ LOAD-BALANCED SWITCH

The analysis of the $N \times N$ switch can be done in an analogous way to the 3×3 case. Since we are short in space we only give here the short overview of the basic steps of it.

Step1. Based on the chosen path create the model of the switch in N time slots long time period for the *cell level* analogously as it is described in the Section II-B for 3×3 case;

Step2. build up the transient DTMC describing the system for the *packet level,* similarly to the procedure described in Section II-C;

Step3. based on the considered path of the cells determine the possible way of *cell/packet loss*, similarly to the derivations in Section II-D and determine the *initial probability* of the transient DTMC, as it is done in Section II-D1; and

Step4. solve the transient DTMC.

The above steps give the outline of the algorithm and based on Section II all the steps are well defined for its detailed program-automated implementation.

Unfortunately, even after proper description of the various steps of the algorithm in the general case (for arbitrary N) the state space increases exponentially with the size of the switch. This can lead to insolvable DTMCs even with the usage of the various sophisticated tools and numerical methods.

IV. COMPUTATIONAL STUDY

In this section we give the comparison of our models to the simulation results in case of switch size $N = 3$, depending on the buffer size – b.

As it is mentioned earlier in Section II-B, all packets arriving to the inputs contain geometric distributed number of cells and there are geometric distributed idle periods between them. If a packet goes to a specific output. These three set of parameters are

- $\mathbf{P} = (\, p_{ij} \,)$ the parameter of the geometric distributed packets length directed from input i to output j,
- $\mathbf{q} = (\, q_i \,)$ the parameter of the geometric distributed idle length (in cells) of input i and
- $\mathbf{T} = (\, t_{ij} \,)$ the probability that the packet is directed from input i to output j.

Using these parameters our modelling scenario is given in Table III.

In Section II we just present the model of path: input 1 → VOQ_{00} → output 0, but we also did it for the other two inputs. These three paths have three different kinds of loss

$$\begin{aligned}\mathcal{A}_1^{\mathcal{N}} &= \mathbf{P}_1^2\mathbf{A}_1^{\mathcal{N}} + \left(\mathbf{P}_1\mathbf{P}_1^{\mathcal{N}} + \mathbf{P}_1^{\mathcal{N}}\left(\mathbf{P}_1 - \mathbf{P}_1^{\mathcal{N}}\right)\right)\left(\mathbf{A}_1 - \mathbf{A}_1^{\mathcal{N}}\right)\\ \mathcal{K}_1^{\mathcal{N}} &= \mathbf{P}_1^2\mathbf{K}_1^{\mathcal{N}} + \left(\mathbf{P}_1\mathbf{P}_1^{\mathcal{N}} + \mathbf{P}_1^{\mathcal{N}}\left(\mathbf{P}_1 - \mathbf{P}_1^{\mathcal{N}}\right)\right)\left(\mathbf{K}_1 - \mathbf{K}_1^{\mathcal{N}}\right)\end{aligned} \tag{30}$$

$$\begin{aligned}\mathcal{P} &= \mathbf{P}_0^3 \otimes \left(\mathcal{A}_1^{\mathcal{N}} + \mathcal{K}_1^{\mathcal{N}}\right) \otimes \mathbf{P}_2^3 = \left(\mathbf{A}_0 + \mathbf{K}_0\right)\mathbf{P}_0^2 \otimes \left(\mathcal{A}_1^{\mathcal{N}} + \mathcal{K}_1^{\mathcal{N}}\right) \otimes \mathbf{P}_2\left(\mathbf{A}_2 + \mathbf{K}_2\right)\mathbf{P}_2 = \\ &= \underbrace{\mathbf{A}_0\mathbf{P}_0^2 \otimes \mathcal{K}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2 + \mathbf{K}_0\mathbf{P}_0^2 \otimes \mathcal{A}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2 + \mathbf{K}_0\mathbf{P}_0^2 \otimes \mathcal{K}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2}_{\text{1 arrival}} + \\ &+ \underbrace{\mathbf{K}_0\mathbf{P}_0^2 \otimes \mathcal{A}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \mathbf{A}_0\mathbf{P}_0^2 \otimes \mathcal{K}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2 + \mathbf{A}_0\mathbf{P}_0^2 \otimes \mathcal{A}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2}_{\text{2 arrivals}} + \\ &+ \underbrace{\mathbf{A}_0\mathbf{P}_0^2 \otimes \mathcal{A}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{A}_2\mathbf{P}_2}_{\text{3 arrivals}} + \underbrace{\mathbf{K}_0\mathbf{P}_0^2 \otimes \mathcal{K}_1^{\mathcal{N}} \otimes \mathbf{P}_2\mathbf{K}_2\mathbf{P}_2}_{\text{no arrivals}} = \mathbf{L}^{\mathcal{N}} + \mathbf{F}_1^{\mathcal{N}} + \mathbf{F}_2^{\mathcal{N}} + \mathbf{B}^{\mathcal{N}}\end{aligned} \tag{31}$$

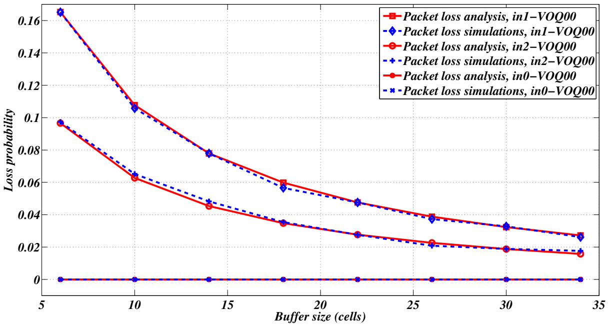


Fig. 6. The packet loss probability (p_p) in case of the analysis and simulation versus the buffer size (b)

probabilities given in Section II-A. The result with parameters given in Table III for all three path types are depicted in Fig. 6.

The results in Fig. 6 show good match in terms of both mathematical model and simulations. It is also captured that in case of larger buffer sizes the packet loss tends to be smaller, which is obvious in case of the same system with more capacity. There are three kind of loss types with relation $p_{p_0} = 0 \leq p_{p_1} \leq p_{p_2}$ as expected in Section II-A in (3).

V. Conclusions

In this paper we presented a combined analysis in order to calculate loss probabilities of a finite central stage buffer both for variable size packets and fixed size cells. In spite of the fact that our analysis does not make calculation according the packet size distribution of the real networks [12], it makes an attempt to present the analysis of LB switch operating with variable size packets contrary to the previous model in [6]. The results show that switch loss with variable size packets can be considerably greater than that for packets of fixed size.

Another important designing issue observed from the analysis is the difference in the packet loss probabilities depending on the traffic traversing path. This property is making complex evaluation of the loss probabilities for large switch sizes since it has strong dependence on the queue number and crossbar interconnection policy, i.e. how the LB switch actually operates. This is not mentioned elsewhere according to our best knowledge.

References

[1] C.S. Chang, D.S. Lee and Y.S. Jou, *Load-Balanced Birkhoff-von Neumann switches, Part I: One-Stage Buffering,* Computer Communications, Vol. 25, pp. 611-622, 2002.

[2] C.S. Chang, D.S. Lee and C.M. Lien, *Load-Balanced Birkhof-von Neumann switches, Part II: Multi-Stage Buffering,* Computer Communications, Vol. 25, pp. 623-634, 2002.

[3] I.Keslassy, S.T. Chuang, K.Yu, D. Miller, M. Horowitz, O. Solgaad and N. McKeown, *Scaling Internet Routers Using Optics,* ACM SIGCOMM'03, Karlsruhe, Germany, 2003.

[4] C.S. Chang, D. Lee, Y.J. Shih, *Mailbox Switch: A Scalable Two-Stage Switch Architecture for Conflict Resolution of Ordered Packets,* in Proceedings of IEEE INFOCOM'04, vol.3, pp. 1995 - 2006, Hong Kong, March 2004.

[5] Y. Shen,S. Jiang,S.S. Panwar, H.J. Chao, *Byte-Focal: A Practical Load-Balanced Switch* IEEE HPSR'05, pp. 6 - 12, Hong Kong, May 2005.

[6] Y.Audzevich, Y. Ofek, M. Telek and B. Yener, *Analysis of load-balanced switch with finite buffers*, IEEE Globecom '08, pp.1 - 6, New Orleans, USA, 2008.

[7] Bill Lin and Isaac Keslassy, *The Concurrent Matching Switch Architecture,* IEEE INFOCOM'06, pp. 1 - 12, Barcelona, Spain, April 2006.

[8] C. L. Yu, C.S. Chang, D.S. Lee, *CR Switch: A Load-Balanced Switch with Contention and Reservation,* IEEE INFOCOM'07, pp. 1361-1369, Anchorage, Alaska, May 2007.

[9] C.-Y. Tu, C.-S. Chang, D.-S. Lee, C.-T. Chiu *Design a Simple and High Performance Switch Using a Two-Stage Architecture,* IEEE GLOBECOM'05, vol. 2, pp. 6 - 11, St. Louis, USA, November 2005.

[10] Jonathan Turner. *Resilient Cell Resequencing in Terabit Routers,* Washington University, Department of Computer Science, Technical Report WUCS-03-48, June 2003.

[11] K. Thompson, G.J. Miller, R. Wilder, *Wide-area Internet traffic patterns and characteristics,* IEEE Network, vol. 11,pp. 10 - 23, Nov.-Dec. 1997.

[12] W.E. Leland, M.S. Taqqu, W. Willinger and D.V. Wilson, *On the Self-similar Nature of Ethernet Traffic,* IEEE/ACM Trans. Networking, vol.2, pp. 1-15, 1994.