INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN 2278 – 0211 (Online)

The Analysis of the Route of a Queued Switch at an Arbitrary Time during Service

Boye Babalola
Department of Mathematics, Obafemi Awolowo University, Nigeria
Titi Obilade
Department of Mathematics, Obafemi Awolowo University, Nigeria

Abstract:
The random performance of any system is not hinged on the throughput of the system alone, it cuts across every interacting performance parameter in the system. Although the throughput gives a general idea of what to expect, the utility grants what to expect under certain laid down distributions at the arrival; a successive inter-arrival times of cells generated by a source, the service and the waiting, along this same line with a view to change certain function in time so as to improve the response and the quality of service delivery in the system. We consider the combined input cross-point queued switch and its performance during a service given a Markov Modulated Bernoulli process at the arrival and a geometric service time distribution.

Keywords: *Switching fabrics, virtual output queues, waiting time distribution, Markov modulated Bernoulli process, throughput*

1. Introduction

In a 2-MMBP, there is a geometrically distributed period of time during which an arrival occurs in Bernoulli fashion with a specific probability α; the system is said to be in state 1 during this period. The system may move from state 1 to another state, i.e. state 2, where arrivals will occur in a Bernoulli fashion with a different probability β; the system will then spend a geometrically distributed period of time in this state before moving to state 1. These periods alternate continuously. Given that the process is in state-1 (state-2) in slot n, it will remain in the same state in the next slot (n+1) with probability (w.p.) α(β), or will change to state-2 (state-1) with probability $1 - \alpha(1 - \beta)$. Furthermore, the MMBP is an ON/OFF process defined over a slotted axis. The ON/OFF periods are geometrically distributed, that is given that in slot i, the process is in an ON period and then in the next time slot i + 1, it would still be in the ON period w.p. α or it will change to the OFF period w.p. $1 - \alpha$. Likewise if it is in the OFF period in slot i, it will stay in the OFF period in the next time slot i + 1 w.p. β or will change to the ON period w.p. $1 - \beta$. If in the next time slot, it would be in the ON period, then a cell may arrive w. p. α. Upon this mechanism is our 2×3 Combined Input Cross-point Queued switch (CICQ) formed. To simplify this traffic model, we assume that each slot during the ON period contain a cell. That is p = 1. The traffic model can be completely defined if we know α and β. These probabilities can be further explained by using the Peak Cell Rate (PCR), Average Cell Rate (ACR) and the burstiness of the source as follows;

1.1. Peak Cell Rate

This is the maximum amount of traffic that can be submitted by a source to an ATM network, and it is expressed as ATM cells per second. Due to the fact that transmission speeds are expressed in bits per second, it is more convenient to talk about the peak bit rate of a source, i.e. the maximum number of bits per second submitted to an ATM connection, rather than its PCR. Thepeak bit rate can be translated to the peak cell rate, and vice versa, if we know which ATM adaptation layer is used [6].Preprint submitted to Performance Evaluation May 9, 2018

1.2. Maximum Burst Size

Depending upon the type of the source, cells maybe submitted to the ATM network in bursts. These bursts maybe fixed or may vary in size. For instance, in a file transfer, if the records retrieved from the disk are of fixed size, then each record result to a fixed number of ATM cells submitted to the network back-to-back. In an encoded video transfer, however, each coded image has a different size, which results to a variable number of cells submitted back-to-back. The Maximum Burst Size (MBS) is defined as the maximum number of cells that can be submitted by a source back-to-back at peak cell rate.

1.3. Burstiness

Typically a source is bursty if it transmits for a period of time and then becomes idle for another period of time. The longer the idle period the higher the arrival rate during the active period and the more bursty the source is. Burstiness of a source can significantly affect a cell loss in an ATM switch [6]. The arrival stream of an ATM cell can be seen as superposition of several different arrival streams coming from the input port of a switch (in this case the Virtual Output Queue), the cell that arrives for the buffer when it is occupied is lost.

From queueing theory, we know that as the arrival rate increases, the cell loss increases as well. What is interesting to observe is the similar behavior can be seen for the burstiness of a source. Burstiness of a source is defined by the Squared Coefficient of Variation (C2). This Quantity is defined as the variation of the inter-arrival times of cells divided by the mean inter- arrival times squared.

In our traffic model, it is given by the following expression.

$$C^2 = \frac{(1-\alpha)(\alpha+\beta)}{2-\alpha-\beta} \quad (1)$$

1.4. Correlation

Consider a successive inter-arrival times of cells generated by a source. In an ATM environment, it is highly likely that the inter-arrival times are correlated either positively or negatively. Positive correlation means that if an inter-arrival time is large (or small), then it is highly likely that the next inter-arrival time will also be large (or small). Negative correlation implies the opposite. That is, if an inter-arrival time is large (or small), then it is highly likely that the next inter-arrival time will be small (or large). As in the case of burstiness, the correlation of the inter-arrival time of cells can significantly affect the cell loss probability in an ATM switch. The larger the value, the burstier the source. From [1], we considering transmission from an arbitrary queue J of the sub-network.

The sub-network will select Bi,j for service from the parent stream (say) SA with probability (w.p.) pi and w.p. 1 − pi selects Bi+1,j for service form the parent stream SB given that SA and SB are not empty. This in the time slot nth is Binomial distributed B(n, pi), considering that r went out from SA.

$$f(p) = \sum_{i=1}^{n} \binom{n}{r} p_i^r (1-p_i)^{n-r} \quad (2)$$

1.5. General Assumption for the Model

We list below the assumptions for developing our queueing model [5];

- The switch is symmetric N ×M and each input/output ports operates at the same speed.
- The size of each VOQ and crossbar (XB) are infinite and s respectively.
- The switch is operating synchronously over fixed size time slots, where each time slot is normalized as
- the time interval for transmitting a cell at input/output scheduling phase.
- Each time slot comprises two phases in sequence; the input scheduling and output scheduling phase.
- Both input and output arbitration uses the random selection policy, i.e., to select one randomly from
- all participating candidate of a contention.
- Cell losses in the switch are rare, which implies that the mean arrival and mean departure rates at
- any buffer in the switch are equal.
- The destination of arriving cells at the inputs are uniformly distributed, i.e., each arriving cell is
- destined to any output port with a probability of;
- The traffic port at each input is i.i.d., with a mean load NP
- At most one cell can arrive at each input port only at the beginning of a time slot
- At most one cell can depart at each output port at the end of a time slot

A state transition of the offered 2-MMBP traffic at each input port can occur only at the beginning of each time slot. A 2-state can generate cells in two cases

- From the ON state back to the ON state with a probability α
- From the OFF state to the ON state with a probability 1-β

It is important to note that each cell is uniformly destined for any output port w.p. 1 M . A 2-MMBP actually consists of two interleaving Bernoulli processes for the state ON and OFF, each with its own cell generation probability; α for the ON state and 1 − β for the OFF state respectively.

2. State of the System

It is desired to observe the system at the moment just before a departure from the output port. Among other variables, the departure at the end of each time slot can be categorized along the line of which VOQ

(A or B) can be associated with the release from three output ports.
Suppose that we define these categories as

$$i = \begin{cases} A_1 \text{if release from output 1 is from VOQ A,} \\ B_1 \text{if release from output 1 is from VOQ B} \end{cases}$$

$$j = \begin{cases} A_2 \text{if release from output 2 is from VOQ A,} \\ B_2 \text{if release from output 2 is from VOQ B} \end{cases}$$

$$k = \begin{cases} A_3 \text{if release from output 3 is from VOQ A,} \\ B_3 \text{if release from output 3 is from VOQ B} \end{cases}$$

Considering the probability of a consecutive departure (ON state to an ON state) from input i released out from output j, i.e., a departure at Output Queue j from input i at two consecutive time slots, given that in the last time slot there was a departure is
Pr{a departure at output j from stream A/the previously departed from stream A } = α
Pr{a departure at output j not from stream A/previously departed was from stream A} = 1 − α
Pr{a departure at output j from stream A/not previously departed was from stream A}= 1 − β
Pr{a departure at output j not from stream A/not previously departed was from stream A} = β
This is represented in a matrix shown below

$$P = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} \alpha_k & 1-\alpha_k \\ 1-\beta_k & \beta_k \end{bmatrix} \end{array} .$$

Thus for a release in each output at consecutive time slots. Where k signifies input queue 2 VOQ A and l
2 VOQ B

$$P = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} \alpha_1 & 1-\alpha_1 \\ 1-\beta_1 & \beta_1 \end{bmatrix} \end{array}$$

Possible release from Output Queue (OQ)1 could either come from VOQ A (or VOQ B) given that it last release was from VOQ A with a probability of α1 (1 − α1).

$$P = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} \alpha_2 & 1-\alpha_2 \\ 1-\beta_2 & \beta_2 \end{bmatrix} \end{array}$$

Possible release from Output Queue (OQ) 2 could either come from VOQ A (or VOQ B) given that it last release was from VOQ A with a probability of α2 (1 − α2).

$$P = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} \alpha_3 & 1-\alpha_3 \\ 1-\beta_3 & \beta_3 \end{bmatrix} \end{array}$$

Possible release from Output Queue (OQ) 3 could either come from VOQ A (or VOQ B) given that it last release was from VOQ A with a probability of α3 (1 − α3).

Given that the transition probability matrix of the Markov Modulated Bernoulli Process at consecutive time slots behave in the manner thus

$$P = \begin{array}{c c} & \begin{array}{cc} ON & OFF \end{array} \\ \begin{array}{c} ON \\ OFF \end{array} & \begin{pmatrix} \alpha & (1-\alpha) \\ (1-\beta) & \beta \end{pmatrix} \end{array}$$

For steady state, we have

$$\Pi = \Pi.P$$

Where

$$\Pi = [\Pi_{ON} \quad \Pi_{OFF}]$$

It follows that

$$\Pi_{ON} = \alpha\Pi_{ON} + (1-\beta)\Pi_{OFF}$$

and since

$$\Pi_{ON} + \Pi_{OFF} = 1$$

we have that

$$\Pi_{ON} = \alpha\Pi_{ON} + (1-\beta)(1-\Pi_{ON})$$

rearranging, we finally have,

$$\Pi_{ON} = \frac{1-\beta}{2-\alpha-\beta}$$

We note that _j OFF (1−_j ON) is obtainable as the steady state for the Markov Modulated Bernoulli (arrival) Process MMBP.

Thus

$$\Pi^j_{ON} = \frac{1-\beta_j}{2-\alpha_j-\beta_j}$$

And

$$\Pi^j_{OFF} = \frac{1-\alpha_j}{2-\alpha_j-\beta_j}$$

From what follows, we have that αj(βj) is the respective probability that an ON (OFF) state is followed by an ON(OFF) state.

3. Throughput Performance

In the general situation when there is M input ports, see [1,2] is given as

$$\sum_{j=1}^{M} \frac{\lambda_{1,j}\beta_j + \lambda_{2,j}\alpha_j}{\lambda(\alpha_j+\beta_j)}(1-\overline{p_{0j}}) + \overline{p_{0j}}[1-\Pi^j_{OFF}\Pi^j_{ON}] \qquad (3)$$

In the event that p0j tends to zero as N tends to infinity, it is easy to see that SN ! 1. For this particular case, the probability of j occupancy in the next time slot given that i packets are in the system at present is Binomial (capacity is infinite in size) distributed and it is given as

$$p_{i,j} = C^{N-i}_{j+1-i}\left(\frac{1}{M}\right)^{j+1-i}\left(1-\frac{1}{M}\right)^{N-(j+1-i)-i} + \delta^i_j\left(1-\frac{1}{M}\right)^N \qquad (4)$$

Where

$i = 1, 2, ...N, \quad j = 1, 2, ...M$
$p_{i,j} = \begin{cases} \delta_j^i(1 - \frac{1}{M})^N & if j + 1 - i \geq 0 \\ 0 & otherwise \end{cases}$

$$p_{i,j} = \frac{(N-i)!}{(j+1-i)!(N-j-1)!} \frac{1}{M^{j+1-i}} (\frac{M-1}{M})^{-j-1} (\frac{M-1}{M})^N + \delta_j^i(1 - \frac{1}{M})^N$$

But
$(1 - \frac{1}{M})^N = e^{-1}$, for sufficiently large N
and N < M
It follows that

$$\frac{(N-i)!}{(j+1-i)!(N-j-1)!} \frac{1}{M^{j+1-i}} (\frac{M-1}{M})^{-j-1} e^{-1}$$

is reduced to

$$p_{ij} = \frac{e^{-1}}{(j+1-i)!} \quad (5)$$

Under the same condition,

$$p_{i,j} = \frac{e^{-1}}{(j+1-i)!} + \delta_j^i(1 - \frac{1}{M})^N \quad (6)$$

where δij is not equal to zero (5) yields (6).

4. Analysis

From the works of Takagi (1991), we consider the queue size at an arbitrary time of the system in consideration with three input queues and three output queues with an initial regard for the systematic competition at the entry point such that the input port of all the VOQs compete for a space at the buffer which entertains a packet at a time into the switching fabrics.

We say that we let the number of packets in the system at time t be N(t) and the time with which the service of packets elapses. From the analytic point of view, the probability of a certain number say m at the said time x for a particular stream

$$P_m(t) \cong \Pr\{N(t) = m | N(0) = i\} \quad i, m = 0, 1, 2, ... \ t \geq 0 \quad (7)$$

Taking a double transformation; LST and z, (6) becomes

$$P^*(z; s) \cong \int_0^\infty e^{-st} P(z; t) dt \quad (8)$$

also that for a stable system, the number of messages in the system after the service commences is

$$\Pi_k(t)dx = \Pr\{N(t) = m, 0 < X(t) < dx | N(0) = i\} \quad (9)$$

X(t) is the elapsed service time of a message in service at time t.
It is upon this that we have

$$\Pi^*(z; s) \cong \int_0^\infty e^{-st} \overline{U}(z; t) dt = \sum_{m=1}^{\infty} \Pi^m(s) z^m \quad (10)$$

The interest of this research is in the fact that hinges on the arrival of packets before service commences, arrival during service, arrival during a residual time left for the time allotted for the completion of service on the basis of no arrival during a service.

The point to note are; if service is commenced and at a time t – x; whose service duration is x and if the service time of the first packet is t.

For the first,

$$\Pi_m(t) = \delta_c \lambda P_0(t) + \sum_{j=1}^{\infty} \int_0^t \Pi_j(t-x) A(x) b(x) e^{-\lambda x} dx \quad (11)$$

and later as

$$\Pi_m(t) = \delta_c \lambda P_0(t) + \sum_{j=1}^{\infty} \int_0^t \Pi_j(t-x) A(x) b(x) e^{-\lambda x} dx + A(t) b(t) e^{-\lambda t} \quad (12)$$

where the A(x) and b(x) are the arrival and the mean service time at time x respectively. Given a Bernoulli arrival at the tagged stream, we have

$$A(x) = \binom{n}{r} p^r q^{n-r} \lambda x \quad (13)$$

and the mean of the arrival process

$$E[p] = n\lambda p x \quad (14)$$

n is the number of streams the server chooses from, p is the probability arrival, λ is the rate of the arrival and x is the time of the arrival. For the service time distribution which is geometrically distributed.

$$E[s] = \frac{xp}{b(x)} \quad (15)$$

While the second moment of the service time distribution follows from

$$E[p^2] = t[\frac{1}{p} - \frac{2(1-p)}{p^2}] \quad (16)$$

We define i as L(θ) the number of messages in the system and X(t) as the elapsed service time of the message in service at time t.

$$P_k(t) = \Pr\{L(t) = k | L(0) = i\} \quad (17)$$

Is a time dependent process and L(0) is the number of messages at an instance time. It follows then that

$$\pi_k(t) dx = \Pr\{L(t) = k, 0 < X < dx | L(0) = i\}$$

Is the shortest time within which time is already spent into service? Hence, a double transformation yields

$$\Pi(z; s) = \sum_{k=1}^{\infty} \int_0^t \pi_k z^k e^{-st} dt \quad (18)$$

and as such, the underlying equation for the problem is as follows;

$$\pi(t) = \eta_{k,1} \lambda P_0(t) + \sum_{j=1}^{k+1} b(x) \frac{(\lambda x)^{k-j+1}}{(k-j+1)!} e^{\lambda x} dx + b(t) \frac{(\lambda x)^{k-j+1}}{(k-j+1)!} e^{\lambda t} \quad (19)$$

Taking the z−transforms of (7), we have

$$\Pi(z; t) = z\lambda P_0(t) + z^{i-1} b(t) e^{-\lambda(1-z)t} + z^{-1} \int_0^t \Pi(z; t-x) b(x) e^{-\lambda(1-z)x} dx - \int_0^t \Pi_1(t-x) e^{-\lambda x} dx \quad (20)$$

Further from (15),

$$\Pi^*(z;s) = \frac{z^2\lambda P_0(t) + z^i B^*(s+\lambda-\lambda z) - z\Pi_1^*(s+\lambda)}{z - B*(s+\lambda-\lambda z)} \tag{21}$$

for i = 1

$$\Pi^*(z;s) = \frac{z^2\lambda P_0(t) + z[B^*(s+\lambda-\lambda z) - \Pi_1^*(s+\lambda)]}{z - B*(s+\lambda-\lambda z)} \tag{22}$$

To find z, one solves quadratically, but the value of __(z; s) is not known. Hence, from Rouch´e's theorem, the argument is that the function z is equal to the function z − B _ (s + λ − λz) by |z| = 1 + ϙ and that both functions have the same number of zeros. In actual fact, the denominator of equation (15) is closer to z than further away as it may seem. for our model,

$$B*(s+\lambda-\lambda z) = \ln(t)e^{s+\lambda-\lambda z} \tag{23}$$

and, from Takagi (1991), we found our P0(t)

$$P_0(t) = \frac{1-\beta}{2-\alpha-\beta}\ln(t)e^{(2s+\lambda)t} \tag{24}$$

for a general arrival flow within the links that flows into the switching fabric. Eventually, (23) becomes

$$\Pi^*(z;s) = \frac{z^2\lambda\frac{1-\beta}{2-\alpha-\beta}\ln(t)e^{(2s+\lambda)t} + z[e^{s+\lambda-\lambda z} - \Pi_1^*(s+\lambda)]}{z - [\ln(t)e^{(s+\lambda-\lambda z)t}]} \tag{25}$$

To find the number of messages in the system at a service time, we make a numerical 2D inverse transforms of both Laplace and Z computation of (25) are considered along the line with the conditions from (22)-(24) and then at (19) where it all started.

For the 2D continuous-discrete system in (25), its 2D Laplace-Z transformation is defined by

$$\Pi(z,s) = \sum_{n=0}^{\infty} \pi(n,t)e^{-st}dtz^{-n} \tag{26}$$

Its inverse representation is

$$\pi(n,t) = \frac{1}{2\pi}\oint_{r}\int_{\sigma-\infty}^{\eta+\infty} \Pi(z,s)e^{-st}z^{n-1}dsdz \tag{27}$$

which is further expressed as

$$\pi(n,t) = (L_t Z_n)^{-1}[\Pi(z,s)] \tag{28}$$

see [8] It is not new then that for a large t and a more frequent influx at the VOQ end where λ is the rate at the arrival end, n increases and largely so, except for a packet loss policy wherewith some packets are unaccounted for

5. Concluding Remarks

This paper attempts to describe an adaptive scheme for a high traffic system whose arrival is not deter-ministic and with some uncertainty within the traffic route. With no consideration to packet loss, a side by side relationship is with the throughput and the queue size at arbitrary time is obtained based on the superposition of the MMBP traffic source. Although, no adaptive schemes were built than the ones already established in queueing theory texts.

6. References

i. Boye Babalola and Titi Obilade (2016) Analytic Queueing Performance of a Combined Input Cross-point Queued Switch
ii. (CICQ) under bursty traffic. International Journal of Scientific and Engineering Research vol.7 pp.1346-1364.
iii. Lin, M. and N. Mckeown (2005) The Throughput of a Buffered Crossbar Switch. IEEE Communications letters, vol.9,
iv. pp.465-467.

v. Medhi J. (2003) Stochastic Models in Queueing Theory. Academic Press, second edition, an imprint of Elsevier Science,
vi. Califonia, USA.
vii. Nabeshima, M. (2000) Performance Evaluation of a Combined Input-and Crosspoint Queued Switch. IEICE.
viii. Nong, G., N. Situ, and M. Hamdi (2007) Delay Analysis of Combined Input-Crosspoint Queueing Switches IEEE, 1-4244-
ix. 1251-X/07, pp. 912-917.
x. Perros, H. G. (2002) An Introduction to ATM Networks. John Wiley & sons, Ltd, Baffins lane, Chichester, West Susset
xi. PO19 IUD, England.
xii. Takagi, H.(1991) Queueing Theory Volume 1. Elsevier science publishers B. V. Amsterdam.
xiii. Xiao Y. and M. H. Lee (2006) 2-D Laplace-Z Transformation. IEICE Trans. Fundamentals E89-A.5.