IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 11, Issue 3, Ver. I (May-Jun .2016), PP 100-110
www.iosrjournals.org

Impact of Cell Arrival Process on A Dynamic Call Admission Control-Based ATM Network

Hilary Ugo Ezea[1], Sunday Ojumo[2], Uchechukwu Rita Ezea[3]
[1]*(Electrical & Electronics Engineering Department, Federal University Oye-Ekiti, Nigeria)*
[2]*(Electrical & Electronics Engineering Department, Cross River State Univ. of Technology Calabar, Nigeria)*
[3]*(Electrical & Electronics Engineering Department, Federal University of Technology, Akure, Nigeria,)*

Abstract: *Network congestion control is comprised of connection admission control (CAC) and network policing control (NPC). Dynamic Call Admission Control (DCAC) is a class in CAC primarily intended to achieve network protection from congestion thereby enhancing network performance and utilization. ATM network is sensitive to varying traffic intensity and the accompanying QoS requirements. Therefore ATM based CAC system is chosen because of the distinct position it occupies in the present network technology. This work investigates the impact of variations in cell arrival process on the DCAC using cell burstiness as parameter. Equally, DCAC is investigated based on QoS parameters such as cell loss probability, cell delay and cell occupancy. Computer modeling and simulation techniques are employed. The simulation results obtained show that varying the cell arrival process brings about variation in the QoS metrics and so, the different cell arrival distribution models have different impacts on the Dynamic Call Admission Control (DCAC)*
Keywords: *Cell arrival, Network Congestion, CAC, DCAC QoS*

I. Introduction

Over four decades ago, precisely in 1968, CCITT established special study group D saddled with the responsibility of studying the use of digital technology in the telephone network. This study group, with the special assignment, established 4-year periods beginning with 1969. The first title of the group was “Planning of Digital Systems.” By 1977, the emphasis of the study group was on “Overall Aspects of Integrated Digital Networks” and “Integration of Services”. Then by 1989, the emphasis of the study shifted to “General Aspects of Integrated Services Digital Networks”. The concept of an “Integrated Services Digital Network, (ISDN)” was formulated in 1972 as one in which the same digital services and digital paths are used for different services such as telephony and data [1].

The first ISDN standard was published under the title “G.705 Integrated Services Digital Network (ISDN)”. Although this first document of an ISDN standard is in the series G Recommendations, most of the ISDN standards are in series I Recommendations, with some also in G, O, Q, and X series Recommendations [2].

The inefficiency of ISDN in achieving the desired unification of services became more pronounced with advancement in communication technology. Then in 1988, CCITT issued a set of Recommendations for ISDN, under the general name of “Broadband Aspects of ISDN” [2]. An interim set of Broadband ISDN (B-ISDN) Recommendations were first issued in 1990. The vision of B-ISDN involves the integration of voice, video, and data services end-to-end and with guaranteed Quality-of-Service (QoS). To support this B-ISDN, the Asynchronous Transfer Mode (ATM) network architecture was proposed by the ITU, as the transport mechanism of choice for the B-ISDN [3].

The traffic management is of paramount concern in the control of ATM traffics. In the traffic management, congestion control has been the focus area. Congestion control can be classified into two categories: Proactive and Reactive controls. In Proactive congestion control, schemes are set up to prevent the congestion condition from occurring. But in the Reactive congestion control, feedback information are collected and processed to control the network congestion level. Connection Admission Control (CAC) is an important part of the Proactive congestion control. The role of CAC is to control the number of connection flows into the network. A new connection request is progressed only when sufficient resources are available to establish and sustain the connection while guaranteeing that the QoS requirements of both the new connection and the already established connections within the network are not violated [4].

There is diversity in the nature and characteristics of traffics in an ATM network. These different traffics require different ways of handling so as to ensure that their respective QoS are guaranteed. Consequently, there is a necessity for CAC schemes to take into consideration the diverse traffic natures and their inherent characteristics such as the pattern by which the cells arriving from the traffic sources are distributed. Such arrival patterns should have impacts on QoS requirements such as the Cell Loss Probability

(CLP) constraints. The issues of interest in this study are determination of the impact of varying cell arrivals on Dynamic Call Admission Control (DCAC) based on cell-loss probability, cell delay and cell occupancy.

This study aims at showing how the average number of cells arriving at the entry node (admission point) of the network within a given interval of time varies from arrival process to arrival process. It also investigates the level of insensitivity or sensitivity of Dynamic Call Admission Control Scheme to variation in the cell arrival process with particular reference to Cell Loss Probability, Cell Delay and Cell occupancy.

II. Review of related Literature

A study of the different traffic classes and their features is imperative to the understanding of the data flow in ISDN networks, as well as, the call admission control in ATM network which is the basis of this research. Generally, there are three main classes of traffic in communication networks [5]. Voice and video are the representatives of the inherently real-time class I traffic [6]. Voice traffic has some level of tolerance to a certain amount of degradation and occasional blocking without losing any information. However, large transmission delays can disrupt a conversation if the necessary facilities are not put in place. Although the exact amount of subjectively acceptable delay is subject to argument, it seems generally agreed that the maximum allowable delay is in the approximate range between 100 and 500ms. Classes II and III traffic are collectively referred to as 'data' [6]. Class II traffic consists of person-to-machine (or machine-to-machine) interactive data. Although this class is not strictly real-time, it has certain delay limitations. Communication here takes the form of intermittent bursts of information separated by intervals of silence at unequal rates hence they could be characterized as being 'bursty'. Class II messages can tolerate short transmission delays but are not error-tolerant. Class III traffic consists of machine-to-machine "bulk data" which are characteristically unidirectional and relatively long. Not being real-time in nature, class III messages may be delayed substantially longer than class II messages and arrive in any random sequential order, provided they arrive without errors [6].

Before the development of ATM networks, performance models of telecommunication systems were typically developed based on the assumption that arrival processes were distributed in Poisson fashion with the time between successive arrivals being exponentially distributed. In early ATM networks, arrival processes were also assumed to be Poisson distributed or Bernoulli distributed [7]. However, the performance models that were based on the Poisson and Bernoulli assumptions did not capture the bursty nature of ATM traffics. Consequently, there was a major shift that led to using the distributions of the ON/OFF type such as the interrupted Poisson process (IPP) and its discrete-time counterpart, the interrupted Bernoulli process (IBP) [7]. In IPP, the ON period is referred to as the ACTIVE period. This is the period during which arrivals occur in a Poisson fashion. On the other hand, the OFF period is referred to as the INACTIVE or IDLE period during which no arrivals occur.

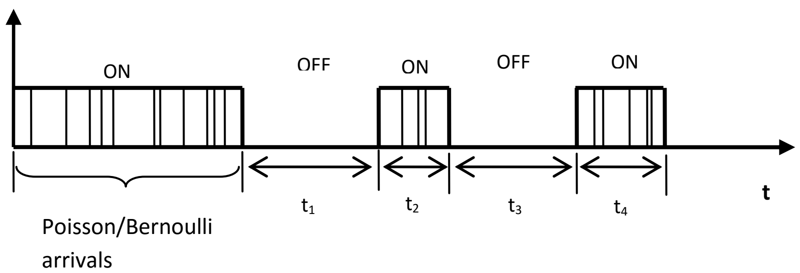

Fig 1: ON/OFF Traffic Model.

These two periods (t_1 and t_2) or (t_3 and t_4) are exponentially distributed, and they alternate continuously. In interrupted Bernoulli process (IBP), arrivals occur during the ON period in a Bernoulli fashion while no arrivals occur during the OFF period. The two periods, (t_1 and t_2) or (t_3 and t_4), are geometrically distributed. Another model known as interrupted fluid process (IFP) uses the fluid approach in which arrivals occur at a continuous rate during the ACTIVE period. Early characterization of ATM traffic showed that the inter-arrival times of cells from a specific source may well be correlated. But in contrast, an IPP or IBP does not capture the notion of correlation since successive inter-arrival times are independent of each other [7]. As a result, more complex distributions were introduced for modeling ATM traffic. These distributions are in the form of Markovian arrival processes such as the Markov modulated Poisson process (MMPP); the discrete-time process, Markov modulated Bernoulli process (MMBP); and the Markov modulated Fluid process (MMFP).

Unfortunately, the non-renewal of the aggregate arrival process poses a major challenge in modeling the superposition of several ON-OFF sources [8]. This non-renewal nature comes from variability of the instantaneous arrival rate due to the fluctuation of the number of sources in the ON state. The aggregate process differs from Poisson in the correlation between successive inter-arrival times. It is true this difference is small under low traffic intensity, it should be noted that at high traffic intensity, the long-term accumulation of these small correlations has a major impact on the system performance [8]. From the point of view of classical superposition limit theorems [9], one should expect that the superposition of many component stationary arrival processes would be nearly Poisson, but even with more than a hundred sources, the multiplexer under heavy loads experiences packet delays much greater than would occur with a Poisson arrival process. The deviation from Poisson behavior increases with increase in the number of streams. This brings in the notion of 'relevant time scale'. The notion of time scale appears indirectly in the classical superposition limit theorem [9] in the requirement that the individual stream gets sparse as the number of streams increases, so that the total average arrival rate remains unchanged. In [8], it is shown that the aggregate process does not behave like a Poisson process over relatively short time intervals, but under heavy loads the congestion in the multiplexer is determined by the behavior of the arrival process over much longer time intervals, where it does not behave like a Poisson process. From the point of view of the queue, as the number of sources increases and the traffic intensity tends to one, the traffic interaction in the queue spans over many intervals in the superposition arrival process. Thus, the long-term covariances between the interarrival times in the aggregate packet arrival process play a significant role, and the aggregate arrival process eventually looks substantially more variable than a Poisson process.

The Markov modulated Poisson process (MMPP) qualitatively models the time varying arrival rate and captures some of the important correlations between the interarrival times while still remaining analytically tractable [10], consequently, it has been extensively used to model the superposition of arrival processes. Considering a two-state Markov chain where the sojourn times in states 1 and 2 are r_1^{-1} and r_2^{-1} respectively.

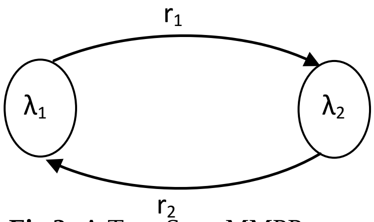

Fig 2: A Two-State MMPP

When the chain is in state k (k =1,2) the arrival process is Poisson with rate λ_k. Using the probability generating function of the number of arrivals in an interval [11],

$$g(z,t) = \pi exp\{[R + (z-1)\Lambda]\}e \qquad \text{........ .... (1)}$$

Where, for the two-state MMPP, the equilibrium probability vector π, is given by:

$$\pi = \frac{1}{r_1+r_2}(r_2, r_1) \qquad \text{............................... (2)}$$

$$e = (1,1)^T,\ R = \begin{bmatrix} -r_1 & r_1 \\ r_2 & -r_2 \end{bmatrix},\ \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

The zero values of the off-diagonal elements of Λ means that all inter-state transitions are accompanied by no arrivals. Arrivals are caused by only 'self-state' transitions.

If the number of arrivals of the stationary two-state MMPP over the interval [0,t] is denoted by N_t, then the average number of arrivals $\overline{N_t}$ is given by:

$$\overline{N_t} = E[N_t] = \frac{\lambda_1 r_2 + \lambda_2 r_2}{r_1 + r_2} t \qquad \text{.. (3)}$$

In Markov modulated Bernoulli process (MMBP), there are various states and in each state, arrivals occur in a Bernoulli fashion at a state-dependent rate. One underlying assumption of an MMBP is that the time the arrival process spends in each state is geometrically distributed. Let P_n denote the state of the underlying Markov chain at time n and let $a_j(k)$ denote the probability of k arrivals when the underlying Markov chain is in state j. Also, let $\{A_n;\ n=1,2,\ldots\}$ denote a sequence of non-negative integer random variables representing the number of arrivals. Then in [12],

$$a_j(k) = Pr\{A_n = k \mid P_n = j\}, \qquad j = 0,\ldots, m;\ k = 0,1,2,\ldots..$$

assumption: $a_j(0) = \begin{cases} 1, & j = 0 \\ 0, & otherwise \end{cases}$

This assumption implies that no cells arrive when the underlying Markov chain is in state 0, and at least one cell arrives in all other states. This assumption is satisfied by a discrete-time queue with a superposition of heterogeneous ON-OFF sources having geometrically distributed OFF-periods. A special type of MMBP is the two-state MMBP where the ON and OFF states alternate. The holding time on each state is assumed to be

DOI: 10.9790/7388-110301100110

geometrically distributed. The transition probabilities for the ON and OFF states are defined by P_{off} and P_{on} respectively.

Fluid models characterize the traffic as a continuous stream with a parameterized flow rate. The fluid traffic paradigm dispenses with individual traffic units such that a traffic count is replaced by a traffic volume which implies that the models are appropriate in case where individual units of traffic are numerous relative to chosen time scale, such that an individual unit is by itself of little significance [13]. Models other than fluid flow models that distinguish between the cells and consider the arrival of each cell as a standalone event (happening at a defined time instant), typically consume huge amounts of memory and processing time for the simulation. On the contrary, a fluid flow model that characterizes the incoming cells by a finite flow rate requires comparatively less resources [14]. This is because in a fluid flow model, an event is generated only when there is a change of flow rate, and such changes in flow rates are less frequent compared to the arrivals of cells. By capturing the rate of changes at the input, the models analyze the different events that occur in the network [15]. Typical fluid models assume that sources are bursty and of the ON-OFF type [16], [17].

2.1 Cell Arrival Process and Connection Admission Control

In B-ISDN ATM, burstiness is of great concern. This is because burstiness as well as correlation are two parameters that grossly affect QoS measures such as cell loss probability [18]. ATM forum standardized the following parameters: peak cell rate, average cell rate or sustainable cell rate, cell delay variation for the peak rate, and maximum burst length. One can effectively police the peak rate using the peak rate and cell delay variation. Similarly, using the maximum burst length, one can estimate a cell delay variation that can be used to police the average rate. But these parameters are fairly inadequate when it comes to bandwidth allocation, since it can easily be shown that there are different distributions (arrival patterns or traffic models) with the same peak rate, average rate, and maximum burst length, but with different burstiness and inter-arrival correlations [18]. In [19], it was shown that different arrival processes affect burstiness.

Several CAC schemes exist. The non-statistical CAC scheme uses only the knowledge of the Peak cell rate (PCR) parameter to compare against the network available bandwidth and decide whether to accept the new connection request or not [4]. The new connection is accepted if the sum of the peak rates of all the existing connections plus the peak rate of the new connection is less than the capacity of the output link. The disadvantage of peak allocation is that unless connections transmit at peak rates, the output port link will be grossly underutilized [7]. In statistical CAC, bandwidth for a new connection is not allocated on the basis of peak cell rates; rather, the allocated bandwidth is between the peak cell rate and the sustained cell rate [7], [4]. The consequence of this is that the sum of all the admitted connections peak cell rates may be greater than the outgoing link capacity. Statistical allocation makes economic sense when dealing with bursty sources, but it is difficult to characterize an arrival process here [7]. Many works have been done by a lot of researchers on the derivation of the equivalent capacity (one of the statistical allocation techniques) and the necessary approximations employed. [20] showed that effective bandwidth exists for a GI/G/1 system with a constraint on tail probability and an M/G/1 system with constraints on mean workload. [21] considered effective bandwidth for buffered network resources, while [22] considered the bufferless case. In [23], [24], and [25], results for the dominant negative exponential tails of certain non-Gaussian queues are obtained. Also, [21] considered the effective bandwidth based on the stochastic fluid flow model, while [23] is based on the batch Poisson arrival process. [26] discussed effective bandwidth for on/off sources with dependent and general distributions, while [27] considered on/off sources with different priorities. In [28], [29], equivalent bandwidth method is also used to allocate capacity based on source declarations and policing mechanisms. An extension of this method with some form of dependence on the current state of the network is looked at by [30], [31], [32]. In [33], the case of network node using a priority-service discipline to support multiple classes of service considered. In [34], the author proposed a measurement-based admissions control procedure based on the policed peak rates of the admitted connections as well as measurements of the aggregate arrival rate. Some effective bandwidth schemes may fail in specific situations highlighted by [35] [36]. In particular, it fails when the probability that the traffic load exceeds the link capacity is assumed to be close to one for a bufferless system having the same input. In the effective bandwidth scheme by [37], the effective bandwidth is computed from the combination of two different approaches. The first one relies on a fluid-flow model [38],[39],[40],[41], in which, the bit rate generated by a number of connections multiplexed together is represented as a continuous flow of bits with intensity varying in accordance with the state of an underlying continuous-time Markov chain. The second approximation focuses on the distribution of the stationary bit rate on a link. Due to the exclusive roles of each of the approximations, they are combined in other to complement each other in such a way that the first approximation estimates the equivalent capacity when the impact of individual connection characteristics is critical while the second approximation represents bandwidth requirements when the effect of statistical multiplexing is of significance. However, the CAC scheme may overestimate the aggregate bandwidth requirement when connections have short bursts because these bursts are normally smoothed out by the output buffer.

Admission control schemes based fully or partially on real-time resource measurements will estimate the resource usage more accurately. Basically, these schemes attempt to predict whether the QoS objectives can be achieved if a new connection is admitted, based on the real-time measurements of certain resources. Therefore, measurement-based CAC schemes came up in a bid to overcome the drawbacks of model-based CAC [42]. The measurement-based CAC algorithms use network measurement to estimate current load of existing traffic, instead of computing the traffic characteristics out of the user specified connection's parameters [4]. So the strategy does not make any assumption about the offered traffic, instead, the network monitors and measures the incoming traffic statistics, and makes decision based on the measured statistics [42]. In [43], a Chernoff bound measurement based admissions control procedure that is based on the equivalent capacity was derived. The Chernoff bound gives a measurement based admissions control procedure based on measurements of the aggregate arrival rate and on a single burstiness parameter for all of the admitted connections. Admission decisions are made based on the current load being less than a pre-calculated threshold. However, CAC strategy based on a decision-theoretic framework is highly vulnerable to errors in the specification of the burstiness parameter.

In [44] CAC scheme is proposed using artificial neural network. This scheme is based on the complex relationship between the offered traffic and QoS requirements during stochastic multiplexing. In [45], a global fuzzy traffic controller that handles, via separate mechanisms, both admission control and congestion control, is presented. [46] proposed a CAC scheme based on fuzzy-logic in which Cell Loss Ratio (CLR) fuzzy estimation based on number of calls per class is performed. [47] proposed a neurocomputing CAC which employs neural networks (NN's) to calculate the bandwidth required to support multimedia traffic with multiple QoS requirements. [48] proposed a neural fuzzy admission control algorithm that combines the best of both neural and fuzzy methods. In [49], the authors considered a neural network-based CAC mechanism that estimates the cell delay and cell loss experienced by each class of traffic in a heterogeneous stream. In [50], an alternative CAC scheme is proposed using a delay and loss based algorithm called quasi-linear dual class correlation, which conservatively estimates the cell delay and cell loss per traffic class using pre-computed vectors derived from the results of three dual arrival queuing models.

In [42], a CAC that uses a variety of information, consisting of user declared UPC parameters and real time on-line traffic measurements was proposed. Using a queuing theoretic approach, the CAC then determines the required capacity or bandwidth to ensure the QoS objectives of all the classes are met.

$$C_{dyn}^{eqv} = sC_{upc}^{eqv} + (1-s)C_{mes}^{eqv} \qquad \text{…………………………………} (4)$$

where s is the smoothing factor that combines the declared UPC characterization and the on-line measured traffic statistics. This smoothing factor s is dynamically chosen using the fuzzy logic control approach based on on-line monitoring of QoS and link utilization.

A conventional CAC scheme makes use of traffic parameters specified by the user and stored in some lookup table [51]. Due to the difficulties in managing the lookup table classifying the wide range of traffic characteristics of the ATM network into a reasonable number of class groups and verifying the fitness of the existing model to other arrival processes, there was the need for a CAC scheme, the performance of which is independent of the classification of calls, and the arrival process modeling. In dynamic CAC, bandwidth allocation to a connection assumes a dynamic adjustment every fixed time period [51]. The scheme estimates the distribution of the number of arriving cells with a renewal mechanism called exponential forecasting [4]. This distribution is estimated from the measured number of cells arriving at the output buffer during the fixed interval and traffic parameters specified by users. Call acceptance (connection admission) is decided on the basis of on-line evaluation of the upper bound of cell loss probability, derived from the estimated distribution of the number of cells arriving. QoS standards can be guaranteed using this control when there is no estimation error. This control can achieve higher utilization than a control that uses only traffic parameters, and can cope with excessive flow from an individual source [51].

In summary, the non-statistical CAC algorithm, as earlier stated, though is easy to implement sequel to the lone requirement of knowing the peak cell rate of the new connection, does not support statistical multiplexing and therefore does not have multiplexing gain. A measure of multiplexing gain being the ratio of the number of sources that can be handled with statistical QoS to the number of sources that can be handled if each is provided its peak rate [52]. Actually, this gain is as a result of the fact that information to be transmitted usually varies in time with peaks and troughs. Consequently, if different calls are fitted together (multiplexed) in such a way that peaks do not coincide, then more calls can be carried as against the peak allocation. Hence, the statistical CAC algorithms in which bandwidth allocation is a function of not only the peak cell rate but also some other statistical parameters such as the sustainable (average) cell rate, prove to be a better option as they support statistical multiplexing in varying degrees. Narrowing down to the statistical CAC schemes, it is very pertinent to note that every bandwidth allocation procedure (including the CAC schemes) uses either a static or dynamic strategy. In static allocation, a reference model is used to determine the bandwidth allocation. This reference model, such as the source traffic descriptor used for CAC, is given apriori. The implication of this being that when the traffic conditions change, the static allocation scheme will fail as there is often no provision for modification. On the other hand, there is a monitoring of the actual traffic conditions in dynamic allocation schemes. This also implies that network engineers do not have to worry about detailed resource allocation at each moment. Instead, the network itself continuously adjusts the network resources assigned to each traffic class and each source-end pair [53]. To increase statistical multiplexing, parameters other than the peak cell rate must be specified in the source traffic descriptor. However, since for many services it is difficult to specify precise values for these additional parameters, users must specify large values for them. As a result, the resulting reference models become insufficient for static allocation [53]. Also, cell delay variation tolerance (CDVT) greatly impacts bandwidth determination but is difficult to specify accurately. Similarly cell delay variation (CDV) generated in the network may affect the amount of bandwidth needed [54]. The existence of CDV makes it difficult to derive an accurate reference model. These problems can be taken care of by employing a CAC algorithm that has dynamic characteristics. One of the salient features of the dynamic CAC that is of interest remains that the scheme incorporates on-line evaluation of cell loss probability. This feature makes it measurement–based and as such, the scheme enjoys the benefits of measurement–based CACs, most importantly, reliability and flexibility. Consequently, in the next sections of this paper, sensitivity of DCAC to variation in arrival process and also the impact of such variations on QoS parameters- the Cell Loss Probability, cell delay and cell occupancy, would be investigated.

III. Traffic Model

In this work, the traffic models considered are all of the ON-OFF type. However, the distributions vary. The cell arrival pattern is therefore bursty. All the traffic source models here have a 0.5 probability of being in the ON state. This also implies that the probability of being in the OFF state is 0.5. Each of the sources generates traffic at the respective rates of 100cells/s, 120cells/s, and 150cells/s.

These models include the Interrupted Poisson Process (IPP) in which the distribution of cells within each burst assumes a Poisson fashion (each arrival instant can only have an arrival and the instants are random); the Interrupted Bernoulli Process (IBP) in which distribution is Bernoulli fashioned (simultaneous arrivals can be accommodated within an arrival instant); the ON-OFF (constant distribution) model in which the distribution of cells within a burst assumes a constant trend.

3.1 Node Model

Bearing in mind that with the knowledge of buffer capacity, effective service process, and adequate characterization of the cell arrival processes, the behavior of a queue taken in isolation is approximately the same as it is within the network, an isolated node is modeled.

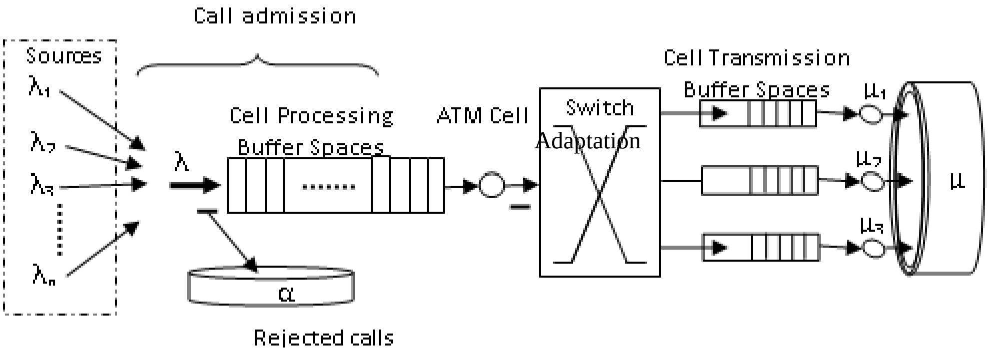

Figure 3: ATM network node model

Connection requests from various sources attached to the node arrive at the rate λ, as expressed in equation (5). The intensity of traffic accessing the node is specified through the arrival rate.

$$\lambda = \sum_{i-1}^{n} \lambda_i \, . \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(5)}$$

This node is composed of the following:

- A buffer, k, of capacity 540cells. The service discipline here is First-In-First-Out and it is assumed that each buffer space can accommodate only a cell.
- An output transmission link, μ. This is a single server with an exponential service rate of 480cells/s.
- The Call admission control (CAC) facility. In this facility, the status of the nodal resources- the buffer and transmission link, is monitored on-line and available resources evaluated. A call is accepted only if its

acceptance would not affect the QoS desired- the cell loss probability and cell delay requirements. If on the other hand, there are no resources to accommodate the call, it is rejected and lost.

The Markov chain, birth and death process is used in modeling this node in which the buffer has a capacity N while the transmission link is a single server.

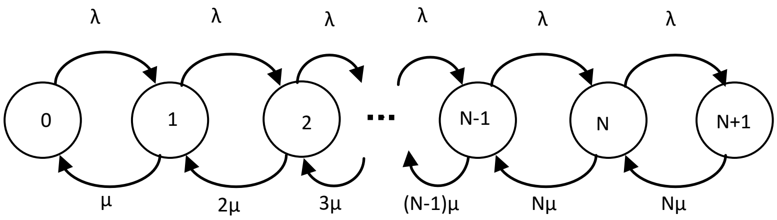

Fig 4: An N+1 state Markov Chain

Let the various states of the Markov chain denote the number of busy resources. Then using the fluid-flow concept, the equilibrium equation for state 0 is given as

$$\lambda P_0 = \mu P_1$$

$$\Rightarrow P_1 = \frac{\lambda}{\mu} P_0 \qquad \text{..} (6)$$

For state 1, $\lambda P_0 + 2\mu P_2 = \lambda P_1 + \mu P_1 \Rightarrow 2\mu P_2 = \lambda P_1$

Substituting for P_1,

$$2\mu P_2 = \frac{\lambda^2}{\mu} P_0$$

$$\therefore P_2 = \left(\frac{\lambda}{\mu}\right)^2 \frac{P_0}{2} \qquad \text{..}(7)$$

For state 2,

$$\lambda P_1 + 3\mu P_3 = \lambda P_2 + 2\mu P_2 \Rightarrow 3\mu P_3 = \lambda P_2$$

Substituting for P_2,

$$3\mu P_3 = \lambda \left(\frac{\lambda}{\mu}\right)^2 \frac{P_0}{2} \Rightarrow P_3 = \left(\frac{\lambda}{\mu}\right)^3 \frac{P_0}{3 \times 2 \times 1}$$

$$\therefore P_3 = \left(\frac{\lambda}{\mu}\right)^3 \frac{P_0}{3!} \qquad \text{...... ....} (8)$$

Comparing equations 6, 7, and 8, one can infer that

$$P_i = \left(\frac{\lambda}{\mu}\right)^i \frac{P_0}{i!} \qquad \text{..} (9)$$

The implication of equation 9 is that,

$$P_{(N-1)} = \left(\frac{\lambda}{\mu}\right)^{(N-1)} \frac{P_0}{(N-1)!} \qquad \text{..} (10)$$

Then for state (N-1), $\lambda P_{(N-2)} + N\mu P_N = \lambda P_{N-1} + (N-1)\mu P_{N-1} \Rightarrow N\mu P_N = \lambda P_{N-1}$

Substituting for P_{N-1}, $N\mu P_N = \lambda \left(\frac{\lambda}{\mu}\right)^{N-1} \frac{P_0}{(N-1)!}$

$$\Rightarrow P_N = \left(\frac{\lambda}{\mu}\right)^N \frac{P_0}{N!} \qquad \text{..} (11)$$

Then considering state N, $\lambda P_{N-1} + N\mu P_{N+1} = \lambda P_N + N\mu P_N \Rightarrow N\mu P_{N+1} = \lambda P_N$

Substituting for P_N we have, $N\mu P_{N+1} = \lambda\left(\frac{\lambda}{\mu}\right)^N \frac{P_0}{N!}$

$$\therefore P_{N+1} = \left(\frac{\lambda}{\mu}\right)^{N+1} \frac{P_0}{N \times N!} \quad \text{……………………………………………}(12)$$

From basic concept of probability,

$$\sum_{n=0}^{N} P_n = 1 \quad \text{………………………………………………………………} (13)$$

$$\therefore P_0\left[\left(\sum_{n=0}^{N}\left(\frac{\lambda^n}{\mu^n} \times \frac{1}{n!}\right)\right) + \left(\frac{\lambda^{N+1}}{\mu^{N+1}} \times \frac{1}{N \times N!}\right)\right] = 1$$

$$\Rightarrow P_0 = \left[\left(\sum_{n=0}^{N}\left(\frac{\lambda^n}{\mu^n} \times \frac{1}{n!}\right)\right) + \left(\frac{\lambda^{N+1}}{\mu^{N+1}} \times \frac{1}{N \times N!}\right)\right]^{-1} \quad \text{……………………} (14)$$

$$\therefore P_0 = \left[\left(\sum_{n=0}^{N}\left(\left(\frac{\lambda}{\mu}\right)^n \times \frac{1}{n!}\right)\right) + \left(\left(\frac{\lambda}{\mu}\right)^{N+1} \times \frac{1}{N \times N!}\right)\right]^{-1} \quad \text{………………} (15)$$

Equation 15 is the probability of having a 100% resource availability. Therefore, at this probability, any cell arriving at the buffer would be conveniently accommodated for service. Of great importance in this work is equation 12. This is the probability of zero percent resource availability. That implies that at this probability, any cell arrival would not have any resource to accommodate it and subsequently, the cell is lost.

Using the buffer dimensioning for this model, the maximum admissible delay for any traffic model is given by [51]:

$$D_{MAX} = \frac{KL}{C_{LINK}} \quad \text{…………………………………………………} (16)$$

Where: K = buffer capacity, L = length of cell, C_{LINK} = capacity of the transmission link.

IV. Simulation Results and Analysis

The simulation was carried out using the models (i.e. traffic model and node model) presented in chapter. Each simulation was carried out using five traffic sources which are statistically identical but work independently. The duration for the simulation was chosen to be one hour in order to produce a stabilized result. Investigation of the impact of varying cell arrival process on DCAC is done by increasing the traffic intensity (number of connections) and observing the behaviors of the different traffic models. Here, the arrival rate is kept constant at 100cells/s while the number of connections is increased on a one-by-one basis. The results and analyses are shown below.

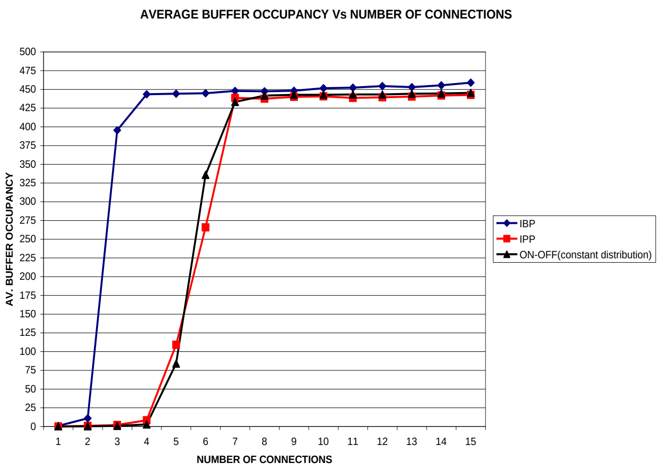

Fig 5: Average buffer occupancy versus traffic intensity (number of connections)

Fig 5 shows the behaviors of the traffic models with respect to buffer occupancy as the traffic intensity increases. Generally, as expected, as the number of connections increases, the buffer occupancy rate also increases. This increase is, however, not linear as the difference in values between successive connections, on the average, tend to be smaller at higher traffic intensities than at lower ones. Of the three models, the IBP attains a high occupancy ratio faster than either of the IPP and ON-OFF. This implies that the IBP would have the fastest cell loss rate and incur the longest delay. IPP and ON-OFF behave alike but a closer look at the table of values reveals that the upward trend in the values for the ON-OFF is not so for the IPP. This is explained by the probabilistic nature of the IPP.

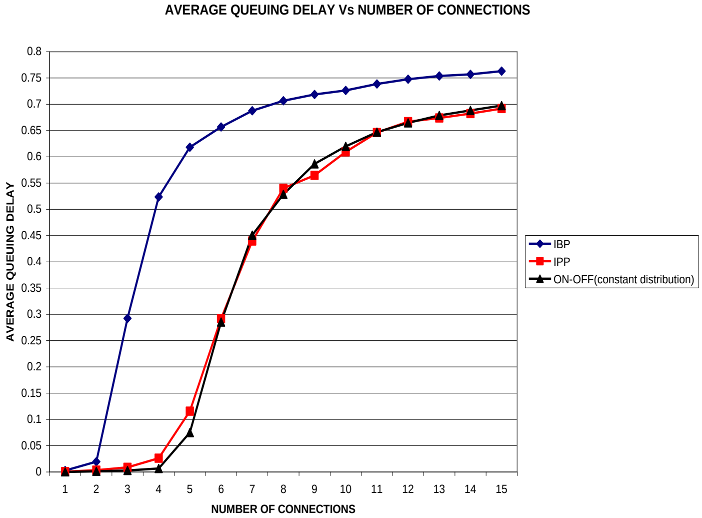

Fig 6: Average Queuing delay versus traffic intensity (number of connections)

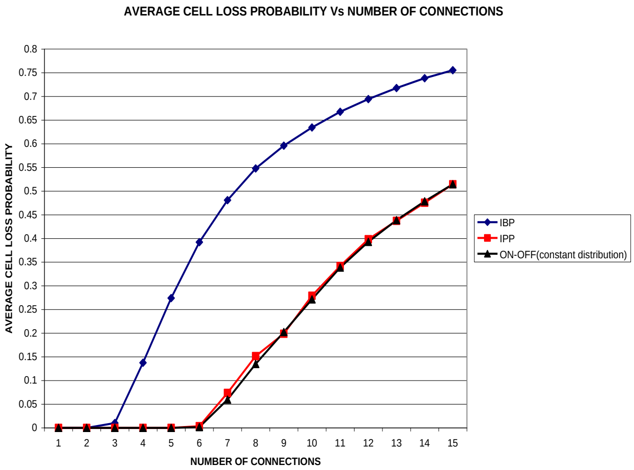

Fig 7: Average Cell loss probability versus traffic intensity (number of connections)

Fig 6 plots the average queuing delay against traffic intensity. The graph shows that the ON-OFF has the best delay performance at low traffic intensities. But as the traffic intensity increases, the difference in the performances of ON-OFF and IPP models becomes less pronounced. At high traffic intensities, the IPP performs better than the ON-OFF though the difference is very small. Conclusively on the average, the IBP has the worst delay performance while the ON-OFF has the best delay performance (incurring the least delay). Fig 7 plots the average cell loss probability against traffic intensity. IBP, as expected has the highest CLP values. Similarly, just as the delay performance, at lower traffic intensities, the ON-OFF has the best cell loss performance but at higher intensities, the IPP has the lowest CLP values. Furthermore, the ON-OFF, on the average, has the best performance.

The above results have shown that the ON-OFF (constant distribution) model has the best performance among the three models considered. It is also shown that the performances of both the IPP and the ON-OFF (constant distribution) models tend to become unified (i.e. the same) and as such become difficult to differentiate as the traffic intensity increases. The implication of this is that the IPP and ON-OFF models can substitute each other at high traffic intensity. The results also show that the IBP model has the worst performance among the three models with respect to the Quality of Service metrics (cell loss probability, cell queuing delay and buffer occupancy), as far as the dynamic call admission control (DCAC) is concerned.

V. Conclusion

The aim of the research work has been to investigate the impact of varying cell arrival process on dynamic call admission control using burstiness parameter, and with special focus on the quality of service metrics-cell delay, cell loss probability and buffer occupancy. The results of the simulation in chapter four have clearly shown the following:

- That varying the cell arrival process brings about variation in queuing delay experienced by the cell.
- That varying the cell arrival process also brings about variation in cell loss probability.
- That variation of the arrival process also causes a variation in buffer occupancy.
- Increasing the traffic intensity by increasing the arrival rate or by increasing the number of connections has an effect on the quality of service metrics with regards to the arrival process employed.

It is therefore concluded that the different cell arrival distribution models have different impacts on the dynamic CAC. Of the three traffic models compared, the interrupted Bernoulli process (IBP) has the poorest performance while the ON-OFF (with constant distribution) has the best performance.

References

[1] Stallings W. "ISDN and Broadband ISDN", Macmillan Publishing Company, New York, 1992.

[2] Ayanoglu E and Akar N, "B-ISDN (Broadband Integrated Services Digital Network)", University of California, Irvine. 2002.

[3] ITU-T, "Broadband Aspects of ISDN ITU-T Recommendation I.121", 1991.

[4] Tian Y, "A Survey on Connection Admission Control in ATM Networks," http://facweb.cs.depaul.edu/ctirs99/online/tian.html

[5] Ross M, "Circuit versus Packet Switching" in Fundamentals of digital switching. (J: McDonald, ed.) New-York, 1983

[6] Chen T.M and Messerschnitt D. G, " ISDN," IEEE communications Magazine, vol.26, No. 6, June 1988

[7] Perros H.G, "Call Admission Control Schemes: A Review," IEEE Communications Magazine, vol.34, no 11, Nov 1996, pp 82-91.

[8] Sriram K and Whitt W, "Characterizing superposition arrival processes in packet multiplexers for voice and data," IEEE Journal on Selected Areas in Communications, vol.SAC-4, No 6, Sept, 1986, pp 833-846.

[9] Cinlar E, "Superposition of Point Processes," In Stochastic Point Processes Statistical analysis, Theory and Applications, P. A.W. Lewis, Ed. New York; Wiley, 1972, pp 549-606

[10] Fisher W and Kathleen, "The MMPP Cookbook," October 1990

[11] Neuts M. F, "A Versatile Markovian Point Process," Journal on Applied Probability, vol 16. Pp 764-779, Dec 1979

[12] Fumio I and Tetsuya T, "Loss Probability in a finite discrete-time queue in terms of steady state distribution of an infinite queue," Queuing Systems Vol 31, 1999, pp 317-326

[13] Frost V and Melamed B, "Traffic Modeling for Telecommunication Networks," IEEE communications Magazine, vol.32, no 3, March 1994, pp 70-81

[14] Adas A, "Traffic Models in Broadband Networks," IEEE Communications Magazine, vol.35, no 7, July 1997

[15] Balakrishnan C, "Survey of Network Traffic Models," http://www.cse.wustl.edu/~jain/cse567-06/ftp/traffic_models3/index.html

[16] Anick D, Mitra D, and Sondhi M.M, "Stochastic Theory of a Data-Handling System with Multiple Sources," Bell System Tech. J., vol.61, No.8, pp 1871-1894, Oct 1982.

[17] Kobayashi H and Ren Q., "A Mathematical Theory for Transient Analysis of communication networks," IEICE Transactions on communication, vol. E75-B, No 12, 1992, pp 1266-1276

[18] Perros H.G, "Call Admission Control Schemes: A Review," IEEE Communications Magazine, vol.34, no 11, Nov 1996, pp 82-91.

[19] Ezea H. U. & Okwor C., "The Economic Implications of Burstiness of Cell Arrival Processes in Network Design", IOSR Journal of Electronics and Communication Engineering, Vol. 9, Issue 6, Nov-Dec. 2014

[20] Kelly F.P, "Effective Bandwidth at Multi-Class Queues." Queuing Systems, Vol. 9, pp 5-15, 1991

[21] Elwalid and Mitra "Effective Bandwidth of General Markovian Traffic Sources and Admission Control of High Speed Networks," IEEE/ACM Trans on Networking, Vol. 1, No 3, pp 329-343, June 1993

[22] Miyao Y, "A Call Admission Control Scheme in ATM networks," In Proceedings of International Conference On Communication (ICC)'91,Denver, USA; Jan 1991

[23] Kesidis G, Walrand J, and Chang C.S, "Effective Bandwidths for multi-class Markov Fluids and other ATM sources", IEEE/ACM Transactions on Networking, vol.1, No. 4, pp 424-428 Aug 1993.

[24] De Veciana G and Walrand J, "Effective Bandwidths: Call Admission, Traffic policing and Filtering for ATM Networks", University of California, Berkeley, Technical Report 1994.

[25] Whitt. W, "Tail Probabilities with statistical Multiplexing and effective bandwidth for multi-class queues," Telecommunications System, Vol.2, pp. 71-107, 1993

[26] Zhang Z and Acampora A. S, "Equivalent Bandwidth for Heterogeneous sources in ATM networks", In Proceedings of IEEE International Conference on Communication (ICC)'94, New Orleans, U.S.A., May 1994, 1025-1029.

[27] Kulkarani V.G, Gun L, and Chimento P.F, "Effective Bandwidth vector for two priority ATM traffic," In Proceedings Of Conference On Computer Communication (IEEE INFOCOM)'94,Toronto,Canada,Jan 1994, pp 1056-1064

[28] Dziong Z et al, "Admission Control and Routing in ATM Networks", In Proceedings of ITC Specialist Seminar, Adelaide, Australia, Sept 1989

[29] Woodruff G et al, "Control of ATM Statistical Multiplexing Performance", In Proceedings of ITC specialist seminar, Adelaide, Australia, Sept 1989

[30] Jain R, "Congestion Control and Traffic Management in ATM Networks: Recent Advances and A Survey," Computer Networks ISDN System, Jan 1995

[31] Liao K.Q and Mason LG, "A Congestion Control Frame Work for Broadband ISDN Using Selective Window Control," in proceedings of broadband communications'94, Paris, France.

[32] Mitra D, "A Sytomically Optimal Design of Congestion Control for High Speed Data Networks," IEEE Transactions on

communications, vol 40, pp 301-311, Feb 1992
[33] Berger A.W and Whitt W, "Effective Bandwidth with Priorities," IEEE/ACM Transactions on Networking. vol. 6, no 4, pp 447-460, Aug 1998.
[34] Floyd .S, "Comments on Measurement-based admissions control for controlled-load service", Technical report, Lawrence Berkeley Laboratory, Jul. 1996.
[35] Elsayed K and Perros H.G, "An Efficient Algorithm for Characterizing the Superposition of Multiple Heterogeneous Interrupted Bernoulli Process," Proc. 2nd Int'l Wksp. on numerical solution of large Markov chains, Jan 1995.
[36] Choudhury G.L, Lucantoni D.M, and Whitt W, "On the Effective Bandwidth for Admission Control in ATM Networks" In Proceedings of ITC 14, Jan 1994.
[37] Guerin R, Ahmadi H, and Naghshineh M., "Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks," IEEE JSAC, Vol.9, No.7, Sept.1991, pp 968-981
[38] Anick D, Mitra D, and Sondhi M.M, "Stochastic Theory of a Data-Handling System with Multiple Sources," Bell System Tech. J., vol.61, No.8, pp 1871-1894, Oct 1982.
[39] Kosten L, "Liquid Models For A Type of Information Storage Problems," Delft Prog. Rep.:Math. Eng., Math. and Inform. Eng., vol.11, pp 71-86, 1986.
[40] Mitra D, "Stochastic Theory of a Fluid Model of Producers and Consumers Coupled by a Buffer," Advanc. Appl. Prob., vol.20, pp 646-676, Sept 1988.
[41] Stern T. E, "Analysis of Separable Markov-Modulated Queuing Processes," Technical Report. CU/CTR/TR 113-8841, Columbia Univ., New York, 1988.
[42] Ren Q. and Ramamurthy G, "A Real-Time Dynamic Connection Admission Controller Based on Traffic Modeling, Measurement, and Fuzzy Logic Control," IEEE JSAC, Feb 2000.
[43] Gibbens R, Frank P, and Key P., "Decision-Theoretic Approach to Call Admission Control in ATM Networks," IEEE JSAC, Vol.13, No.6, Aug 1995, pp 1101-1114
[44] Hiramatsu A., "ATM Communications Network Control by Neural Networks", IEEE Transactions on Neural networks, vol.1, No1, pp 122-130, Mar 1990
[45] Cheng R. G, and Chang C. J, "Design Of A Fuzzy Traffic Controller for ATM Networks," IEEE/ACM Transactions on Networking, vol.4, no.3, June 1996, pp 460-469
[46] Uehara K. and Hirota K, "Fuzzy Connection Admission Control for ATM Networks Based on Possibility Distribution of Cell Loss Ratio," IEEE JSAC, Vol.15, No.2, Feb 1997, pp 179-190
[47] Youssef S.A, Habib I.W, and Saadawi T.N, "A Neurocomputing Controller for Bandwidth Allocation in ATM Networks," IEEE JSAC, Vol.15, No.2, Feb 1997, pp 191-199
[48] Cheng R.G, Chang C.J, and Lin L.F, "A QoS Provisioning Neural Fuzzy Connection Admission Controller For Multimedia High-Speed Networks", IEEE/ACM Transactions on networking, vol.7, no.1, pp 111-121, Feb 1999
[49] Hah J.M, Tien P.L, and Yuang M.C, "Neural Network Based Call Admission Control for ATM Networks with Heterogeneous Arrivals", Computer Communications Vol.20, No.9, pp 732-740, Sept 1997.
[50] Hah J.M and Yuang M.C, "Estimation-Based Call Admission Control with Delay and Loss Guarantees in ATM Networks", IEE Proceedings on Communications, vol.144, No.2, pp 85-92. April 1997.
[51] Saito H and Shiomoto K, "Dynamic Call Admission Control in ATM Networks," IEEE JSAC; Vol.9, pp 982-989, Sept 1991.
[52] Kumar A., Manjunath D., and Kuri J., "Communication Networking-An Analytical Approach", Morga Kaufmann, India, 2005.
[53] Hiroshi Saito, "Dynamic Resource Allocation in ATM Network", IEEE Communications Magazine, vol. 35, no 5, May 1997, pp 146-153
[54] Hiroshi Saito, "Teletraffic Technologies in ATM Networks", Boston: Artech House, 1994.