# Performance Analysis of Mixed Polling Schemes with Multiple Classes of Self-Similar Traffic Input to Build Comprehensive SLAs

Mohsin Iftikhar[1], Muhammad Imran[2], Ghazi Al-Naymat[3], Mohammad A. Alnuem[1], Hassan Mathkour[1]

[1]College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[2]Deanship of E-Transactions and Communications, King Saud University, Riyadh, Saudi Arabia
[3]College of Computer Science and Information Technology, University of Dammam, Dammam, Saudi Arabia
{miftikhar, cimran, malnuem}@ksu.edu.sa, ghalnaymat@ud.edu.sa

***Abstract*— Solitary polling technique is not a better choice to get rid of multiple queuing problems for getting enhanced performance against single server. Enhanced and reliable performance upon multi queued traffic can be achieved through the utilization of right selection of joint polling schemes. The complexity of polling design is directly proportional to its performance. Exhaustive polling scheme is outperformed by limited service polling policy but under average load of traffic both contain same mean transfer delay. Several polling methods are the part of literature but the question of their optimal utilization in joint manners against multiple queue buffers is still open. In this paper, we build a complete analytical framework for mixed exhaustive and limited service polling model based on G/M/1 queuing system. We build the Markov chain and extract the closed form expressions of packet delay for multiple classes of self-similar traffic. We further experimentally implement four queues model with different arrangements of mixed polling schemes to conclude the most optimal joint polling model for modern 4G wireless network through analyzing the co-relation between delay and Hurst paramete*r*.**

***Keywords***: *G/M/1, Self-Similar, Exhaustive, Gated,*

## I. INTRODUCTION

Polling Model is a multi-queue access management system that embodies several queues to a solitary server in a specific way. More often, polling techniques can accurately be analyzed if branching property is truly satisfied; which means each job in queue should always follow a unique (random) identifier during each server‘s call as reported in [1]. To achieve effectual and reliable performance in the presence of divergent load and congestion, an appropriate selection of most optimal polling scheme is necessary. Queue demonstration can be distinguished under three different polling routines [2]: (1) Exhaustive Method – in which a system (server) cannot leave the queue until it becomes vacant, (2) Gate Method - on each queue visit, the server serves only those number of packets from that queue which are present at the polling instant, (3) Limited Service Method – the server serves specific number of packets from each queue during each cycle. The polling model under exhaustive scheme [3] has been represented in Fig. 1.
Finite set of queues can be represented to single server in cyclic order with non-zero switching time through polling model [4] and its potential usages lie under transposition, telecommunication, reliability management of computer networks, manufacturing and health caring infrastructures etc. Recently, polling system has been actively implemented in resource sharing of wireless channels through Media Access Control (MAC) Protocol [5, 6]. Furthermore, polling model is actively being used in IEEE 802.11 cellular local area networks [7, 8], broadband cellular channels [9] and “Ethernet Passive Optical system” [10] because polling system is being considered a reliable access mechanism under peak load as compared to other connection oriented protocols [11].

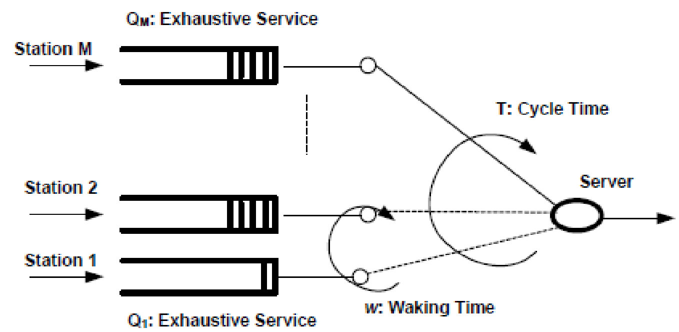


**Fig. 1:** Exhaustive Polling Scheme

In this paper, we build a complete analytical framework for mixed exhaustive and limited service polling model based on G/M/1 queuing system. We build the Markov chain and extract the closed form expressions for multiple classes of self-similar traffic input. To the best of our knowledge, this is the first time that closed form expressions has been presented for mixed exhaustive and limited service model with four distinct classes of self-similar traffic. We also experimentally implement the four queues model corresponding to different sorts of mixed polling mechanisms to conclude which combination provides more reliable results in the presence of asymmetric classes of traffic. We analyze the co-relation between delay and Hurst parameter for different combination of traditional polling schemes. The main objective of this work is to point out which combination of joint polling schemes provide guaranteed QoS to multiple classes of network traffic in today’s multiservice wireless internet.

## II. Related Work

Queuing methods facilitate the way to handle multidisciplinary traffic under the utilization of buffering technology. A lot of queuing techniques like Low Latency Queuing (LLQ), Priority Queuing (PQ), Class based Weighted Fair Queuing (CBWFQ) and Custom Queuing (CQ), have been introduced in the market [14]. Most of the prior work has analyzed these traditional queuing schemes with Poisson traffic. However, it has been shown clearly through high quality studies that multimedia traffic -- carried in modern fixed and wireless networks -- exhibits self-similar characteristics [15, 16]. Therefore, it is important to consider the bursty nature of traffic while conducting the analysis of any kind of queueing system. Further, in this study we also aim to investigate, which performance metric is responsible for degrading the efficiency of other performance metrics.

In our prior work, we have figured out that how load and congestion affects the performance of remote services [24]. The mutual relationship between hop counts and congestion has also been found in [24].

Chernova et al. [2] implemented different queuing logics against an individual server. The authors have considered two different scenarios. Case 1: if queue is empty upon a visit then it will not be visited upon upcoming visit and Case 2: with limited service polling policy the stable and unstable condition of different queuing schemes have been investigated. For the analysis of polling models, a framework has been designed by Al-Hanbali et al. [1]. Their main contribution is that; the presented framework is able to manage those jobs that did not meet branching property.

The exhaustive and gated policies are dependent on the priority of each job that bounds the server execution to follow the indicated preference of job. On the other side, k- Limited polling strategy is preferred rather to these policies (gated, exhaustive) due to its facility of defining the job execution limits [1]. Recently Time Autonomous-Server polling strategy similar to K (time)- Limited polling model that control the server to reside at queue for specific time even in the condition of empty queue has been discussed in studies [12, 13]. The performance of polling model is highly affected under the following factors as reported in study [1].

1. Complexity of Polling Scheme that causes more load
2. Increment in cycle time with the linear addition of more queues
3. Increments in cycle time under overloaded situation

Many studies related to polling models have been conducted in recent years. The analysis of exhaustive and gated polling schemes has been conducted in [17] by considering the switch over time equivalent to zero. The analysis of exhaustive polling method with arbitrary number of queues has been reported in [18]. Lee et al. [19] have analyzed two queues in which one queue is being served by exhaustive service and the other one though limited service policy. We further refer the readers to [20-22] to have an overview of the related work being conducted in the area of polling models.

The outperformed polling design can be achieved through average transferring rate with small queue buffer and lesser possible server walk through time upon each queue. Larger visiting time and number of nodes cause higher cycle time and as a result waiting time is increased that badly affects the performance of any queuing model. Exhaustive polling strategy seems to be more efficient as compared to limited service polling policy but under average load of traffic. On the other hand both exhibit same performance in terms of mean transfer delay [3]. The interesting finding is that, in case of overload condition, the limited service polling technique is better in terms of lower delay but this performance is directly correlated to the value of k, because in case of higher value of k it almost performs similar to exhaustive technique as reported by AlQahatani et al. [3]. Therefore, in this current study, our motive is to analyze thoroughly different combinations of polling strategies. The ultimate objective is to find out different kinds of suitable polling combinations for different kinds of networks according to the requirement of applications found in those networks.

## III. Traffic Model

We have adopted the traffic model presented in [25] for this work because it is highly accurate, parsimonious, analytical, flexible and implementable. The working of the model is like an on/off process, especially its variation N-Burst model [26] where packets are incorporated. However, the model in [26] only considers one type of traffic. Like real networks, the traffic model captures the dynamics of packet generation while accounting for the scaling properties [25]. This traffic model belongs to a particular class of self-similar traffic models called infinite source Poisson models [27]. A common feature of these models is a heavy-tailed distribution for the sessions that occur at the flow level and arrive according to a Poisson process. On the other hand, the local traffic injection process over each session is a distinguishing feature. The Hurst parameter is implicit in the distribution of the sessions and its estimation has been recently investigated in [28]. Our traffic model is long-range dependent and almost second-order self-similar as the auto-covariance function of its increments is equal to that of fractional Gaussian noise for sufficiently large time lags. The traffic can be approximated by FBM when the rate of packet arrivals tends to infinity. The traffic is found by aggregating the number of packets generated by several sources. In the framework of a Poisson point process, the model represents an infinite number of potential sources. Each source initiates a session with a heavy-tailed distribution, in particular a Pareto distribution whose density is given by $g(r) = \delta b^{\delta} r^{-\delta-1}$, $r > b$, where $\delta$ is related to the Hurst parameter by $H = (3-\delta)/2$. The sessions are assumed to arrive according to a Poisson process with rate $\lambda$. Locally, the packets generated by each source arrive according to a Poisson process with rate $\alpha$ throughout each session. The local packet generation process could be taken as a compound Poisson process which would then represent packet sizes as well [25, 29].

For a single class of traffic, the traffic *Y (t)* measured as the total number of packets injected in *[0, t]* can be written as

$$Y(t) = \sum_{S_i \le t} U_i (R_i \wedge (t - S_i))$$

where $U_i$ denotes the local Poisson process over session *i*, $R_i$ and $S_i$ denote the duration and the arrival time of session *i*, respectively, and the values of i denote an enumeration of the arriving sessions. Here, $R_i$ is positive, $S_i$ is real valued and $U_i$ which counts the number of packets of session *i* is integer valued. As a result, *Y (t)* corresponds to the sum of packets generated by all sessions initiated in *[0,t]* until the session expires if that happens before *t*, and until *t* if it does not. We consider the stationary version of this model based on an infinite past. The sessions have been arriving for a long time and hence the incremental traffic is stationary. The sessions are represented with horizontal line segments with their lengths equal to the ordinate of their starting points *(s, r)*. The starting points of the sessions are indicated with a diamond. The vertical segments represent the packets which are placed over each session at the time of their arrivals. The component u, which is not represented in Fig.1, is the number of packets over a session. The numbers *s, r* and *u* denote the realization of $S_i, R_i$ and $U_i$ for session *i*, respectively.

In the present study, we exploit the traffic model to represent different classes of traffic streams. Each stream has its own parameters and is independent from the other(s). The packet sizes are assumed to be fixed because each queue or traffic class corresponds to a certain type of application where the packets have fixed size or at least fixed service time distribution.

## IV. ANALYTICAL FRAMEWORK

We considered a model of four queues based on G/M/1 queuing system. The queuing model is being fed by four distinct classes of self-similar traffic input. The first two queues are exhaustive queues; whereas; queue 3 and queue 4 are being treated in a limited service polling model fashion. The scheduler visits queue 1 first and serves all the packets, then visits queue 2 and serves all the packets. Furthermore, the limited service queues are being served in such a way that the scheduler serves two packets from queue 3 and one packet from queue 4 during each cycle. Let µ1, µ2, µ3 and µ4 denote the service time distribution for correspondingly type 1, 2, 3 and 4 packets, respectively. The observation of the queuing system at the time of arrival instants is the key factor on which the Markov chain formulation of G/M/1 is based on [30]. The embedded Markov chain at the time of arrival instants is denoted by $X_n : n \ge 0$. Moreover, the state space is defined as follows:

*S={ (i1,i2,i3,i4,a,s) : a*$\in$ *{a1,a2,a3,a4} , s*$\in$ *{s1,s2,s3,s4,* $I$ *}, i1,i2,i3,i4*$\in$ $Z$ + *}*

Where *a1, a2, a3* and *a4* represent the type of arriving packet, *s1, s2, s3* and *s4* denote the type of packet in a service. Moreover*, i1, i2, i3* and *i4* represent the number of packets in each queue plus a possible packet in a service if any, and $I$ indicates the idle condition in which no packet is in service or queued.

### A. Markov Chain Transition Probabilities

The transition probability matrix of the Markov chain can be built by taking each case scenario separately. Here, we are going to discuss only one transition in detail.

*Transition from* $(i_1, i_2, i_3, i_4, a_1, s_1) \longrightarrow (j_1, j_2, j_3, j_4, a_2, s_2)$

In this transition, a type 1 packet comes in the initial state followed by a type 2 packet in the next state. The arriving packet of type 1 in the initial state has viewed $i_1$ packets in queue 1 including a packet in service of type 1, $i_2$ packets of type 2, $i_3$ packets of type 3 and i4 packets of type 4 in the system. In the next state, a new arrival of type 2 packet has found $j_1$ packets of type 1, $j_2$ packets of type 2, $j_3$ packets of type 3 and $j_4$ packets of type 4 in the system.
Due to exhaustive serving logic for the first two queues, the new arriving packet of type 2 can find a type 2 packet in service only if all type 1 packets including the one that arrived in the previous state are exhausted during the inter-arrival period. Therefore, $j_1$ is equal to 0 and exactly $i_1 + 1$ packets of type 1 have been served. On the other hand, the no. of packets being served from queue 2 can be *k*, and *k* can have value between 0 and $i_2 - 1$ because at least one type 2 packet must remain in the system, the one being in service when the new arrival happens to fulfil this transition. Further, the scheduler has not got the chance to visit queue ¾ because, it is still busy with queue 2 on the arrival of new class 2 packet. This transition has been shown in Fig.1. The transition probability can be written as follow:

$$P\{X_{n+1} = (0, i_2-k, i_3, i_4, a_2, s_2) \mid X_n = (i_1, i_2, i_3, i_4, a_1, s_1)\}$$

$$= P\{i_1 + 1 \text{ served from queue 1, } k \text{ served from queue 2 and a type 2 packet remain in the service during } T_{12}\}$$

$$P\{X_{n+1} = (0, i_2-k, i_3, i_4, a_2, s_2) \mid X_n = (i_1, i_2, i_3, i_4, a_1, s_1)\}$$

$$= \int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_2}(s) f_{S_1^{i_1+1}+S_2^k}(x) f_{T_{12}}(t)\, ds\, dx\, dt$$

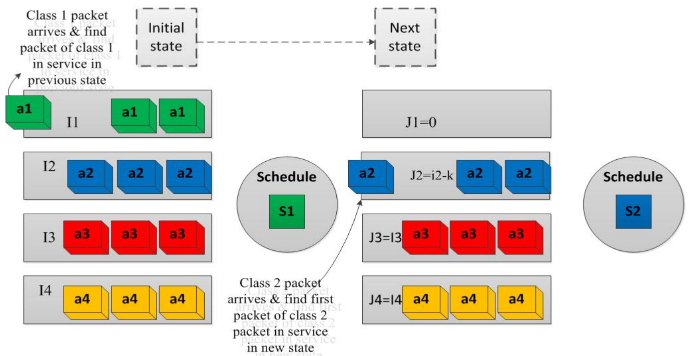


**Fig. 2**: A Markov Chain Transition for four queues of mixed exhaustive and limited service policy.

*($i_1$, $i_2$, $i_3$, $i_4$, $a_1$, $s_1$) ⟶ ($j_1$, $j_2$, $j_3$, $j_4$, $a_2$, $s_2$)*

## B. Extraction of Packet Delay

Steady state distribution $\pi$ as seen by an arrival can be found by solving $\pi P = \pi$ using the transition matrix $P$ of the Markov chain analyzed above. In reality, each router has a limited capacity for each queue; hence the steady state distribution exists. Our investigation depends on the limiting distribution of the state of each queue at the arrival instances, which can be estimated by employing the analysis given above for our self-similar traffic model. In general, the following analysis is valid for any G/M/1 queuing system where the limiting distribution $\pi$ at the arrival instances can be enumerated. The expected waiting time for the first queue (exhaustive) packet can be found as:

$$E[W_1] = \sum_{j_1=1}^{J_1-1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\frac{j_1}{\mu_1}\pi(j_1,j_2,j_3,j_4,a_1,s_1) +$$
$$\sum_{j_1=0}^{J_1-1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{2}{\mu_3}+\frac{1}{\mu_4}\right)\pi(j_1,j_2,j_3,j_4,a_1,s_2) +$$
$$\sum_{j_1=0}^{J_1-1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{2}{\mu_2}+\frac{1}{\mu_4}\right)\pi(j_1,j_2,j_3,j_4,a_1,s_3^1) +$$
$$\sum_{j_1=0}^{J_1-1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{1}{\mu_3}+\frac{1}{\mu_4}\right)\pi(j_1,j_2,j_3,j_4,a_1,s_3^2) +$$
$$\sum_{j_1=0}^{J_1-1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=1}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{1}{\mu_4}\right)\pi(j_1,j_2,j_3,j_4,a_1,s_4) + \frac{c_2}{\mu_2}E[W_1]$$

where J1, J2, J3 and J4 are the respective capacities of each queue. This clearly invokes that a newly arrived highest priority packet will wait until all packets of same priority and the packet in the service are served. Through deliberating the category of the packet in service, we have constituent expressions in the sum.

The predictable waiting period of class 2 packets (queue 2) can be found as follows:

$$E[W_2] = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2-1}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_1) +$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2-1}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_2) +$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2-1}\sum_{j_3=1}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{1}{\mu_3}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_3) +$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2-1}\sum_{j_3=0}^{J_3}\sum_{j_4=1}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{1}{\mu_4}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_4) + \left(\frac{c_1}{\mu_1}\right)E[W_2]$$

On the other hand, we obtain the expected waiting time for the limited service queues by analyzing the events that constitute this delay. We consider two factors (the impact of high priority queues and the effect of round-robin service) to find out the expected waiting time of a packet (class 3 and class 4) arriving to non-priority queues (queue 3 and queue 4). The exact bounds on the expected waiting time for a class 3 packet can be computed as follows:

$$C_3 < E[W_3] < C_3'$$

$$C_3 = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=0}^{\lfloor j_3/2\rfloor}\pi(j_1,j_2,j_3,j_4,a_3,s_1)(j_1/\mu_1+j_2/\mu_2+j_3/\mu_3+j_4/\mu_4)$$
$$+\sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=\lceil j_3/2\rceil}^{J_4}\sum_{m=0}^{1}\pi(j_1,j_2,j_3,j_4,a_3,s_1)(j_1/\mu_1+j_2/\mu_2+j_3/\mu_3+(\lfloor j_3/2\rfloor+m)/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=0}^{\lfloor j_3/2\rfloor}\pi(j_1,j_2,j_3,j_4,a_3,s_2)(1/\mu_2+(j_2-1)/\mu_2+j_1/\mu_1+j_3/\mu_3+j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=\lceil j_3/2\rceil}^{J_4}\sum_{n=0}^{1}\pi(j_1,j_2,j_3,j_4,a_3,s_2)(1/\mu_2+(j_2-1)/\mu_2+j_1/\mu_1+j_3/\mu_3+(\lfloor j_3/2\rfloor+n)/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3-1}\sum_{j_4=0}^{\lfloor j_3/2\rfloor}\pi(j_1,j_2,j_3,j_4,a_3,s_3^1)(1/\mu_3+j_1/\mu_1+j_2/\mu_2+(j_3-1)/\mu_3+j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3-1}\sum_{j_4=\lceil j_3/2\rceil}^{J_4}\pi(j_1,j_2,j_3,j_4,a_3,s_3^1)(1/\mu_3+j_1/\mu_1+j_2/\mu_2+(j_3-1)/\mu_3+\lfloor j_3/2\rfloor/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3-1}\sum_{j_4=0}^{\lfloor j_3/2\rfloor}\pi(j_1,j_2,j_3,j_4,a_3,s_3^2)(1/\mu_3+j_1/\mu_1+j_2/\mu_2+(j_3-1)/\mu_3+j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3-1}\sum_{j_4=\lceil j_3/2\rceil}^{J_4}\pi(j_1,j_2,j_3,j_4,a_3,s_3^2)(1/\mu_3+j_1/\mu_1+j_2/\mu_2+(j_3-1)/\mu_3+\lceil j_3/2\rceil/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=1}^{\lfloor j_3/2\rfloor}\pi(j_1,j_2,j_3,j_4,a_3,s_4)(1/\mu_4+j_1/\mu_1+j_2/\mu_2+j_3/\mu_3+(j_4-1)/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=\lceil j_3/2\rceil}^{J_4}\pi(j_1,j_2,j_3,j_4,a_3,s_4)(1/\mu_4+j_1/\mu_1+j_2/\mu_2+j_3/\mu_3+\lfloor j_3/2\rfloor/\mu_4)$$
$$+\left(\frac{C_1}{\mu_1}+\frac{C_2}{\mu_2}\right)C_3$$

and

$$C_3' = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=0}^{J_4}\sum_{m=0}^{1}\pi(j_1,j_2,j_3,j_4,a_3,s_1)(j_1/\mu_1+j_2/\mu_2+j_3/\mu_3+(\lfloor j_3/2\rfloor+m)/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=0}^{J_4}\sum_{n=0}^{1}\pi(j_1,j_2,j_3,j_4,a_3,s_2)(1/\mu_2+(j_2-1)/\mu_2+j_1/\mu_1+j_3/\mu_3+(\lfloor j_3/2\rfloor+n)/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3-1}\sum_{j_4=0}^{J_4}\pi(j_1,j_2,j_3,j_4,a_3,s_3^1)(1/\mu_3+j_1/\mu_1+j_2/\mu_2+(j_3-1)/\mu_3+\lfloor j_3/2\rfloor/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3-1}\sum_{j_4=0}^{J_4}\pi(j_1,j_2,j_3,j_4,a_3,s_3^2)(1/\mu_3+j_1/\mu_1+j_2/\mu_2+(j_3-1)/\mu_3+\lceil j_3/2\rceil/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=1}^{J_4}\pi(j_1,j_2,j_3,j_4,a_3,s_4)(1/\mu_4+j_1/\mu_1+j_2/\mu_2+j_3/\mu_3+\lfloor j_3/2\rfloor/\mu_4)$$
$$+\left(\frac{C_1}{\mu_1}+\frac{C_2}{\mu_2}\right)C_3'$$

Similarly we can write down the exact bounds on expected waiting time of a class 4 packet (lowest priority queue) as follows:

$$C_4 < E[W_4] < C_4^{/}, \text{ where}$$

$$C_4 = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{2j_4}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_1)(j_1/\mu_1 + j_2/\mu_2 + j_4/\mu_4 + j_3/\mu_3)$$
$$+\sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=2j_4+1}^{J_3}\sum_{j_4=0}^{J_4-1}\sum_{m=0}^{2}\pi(j_1,j_2,j_3,j_4,a_4,s_1)(j_1/\mu_1 + j_2/\mu_2 + j_4/\mu_4 + (2j_4+m)/\mu_3)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{2j_4}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_2)(j_1/\mu_1 + j_2/\mu_2 + j_4/\mu_4 + j_3/\mu_3)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=2j_4+1}^{J_3}\sum_{j_4=0}^{J_4-1}\sum_{n=0}^{2}\pi(j_1,j_2,j_3,j_4,a_4,s_2)(j_1/\mu_1 + j_2/\mu_2 + j_4/\mu_4 + (2j_4+n)/\mu_3)$$
$$+$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{2j_4+2}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_3^1)(j_1/\mu_1 + j_2/\mu_2 + 1/\mu_3 + (j_3-1)/\mu_3 + j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=2j_4+3}^{J_3}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_3^1)(j_1/\mu_1 + j_2/\mu_2 + 1/\mu_3 + (2j_4+1)/\mu_3 + j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{2j_4+1}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_3^2)(j_1/\mu_1 + j_2/\mu_2 + 1/\mu_3 + (j_3-1)/\mu_3 + j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=2j_4+2}^{J_3}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_3^2)(1/\mu_3 + j_1/\mu_1 + j_2/\mu_2 + 2j_4/\mu_3 + j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{2j_4}\sum_{j_4=1}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_4)(1/\mu_4 + (j_4-1)/\mu_4 + j_1/\mu_1 + j_2/\mu_2 + j_3/\mu_3)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=2j_4+1}^{J_3}\sum_{j_4=1}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_4)(1/\mu_4 + (j_4-1)/\mu_4 + j_1/\mu_1 + j_2/\mu_2 + 2j_4/\mu_3)$$
$$+\left(\frac{C_1}{\mu_1} + \frac{C_2}{\mu_2}\right)C_4$$

and

$$C_4^{/} = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4-1}\sum_{m=0}^{2}\pi(j_1,j_2,j_{3,}j_4,a_4,s_1)(j_1/\mu_1 + j_2/\mu_2 + (2j_4+m)/\mu_3 + j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{J_3=0}^{J_3}\sum_{J_4=0}^{J_4-1}\sum_{n=0}^{2}\pi(j_1,j_2,j_3,j_4,a_4,s_2)(j_1/\mu_1 + j_2/\mu_2 + (2j_4+n)/\mu_3 + j_4/\mu_4)$$
$$+\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_3^1)(1/\mu_3 + j_1/\mu_1 + j_2/\mu_2 + (2j_4+1)/\mu_3 + j_4/\mu_4)$$
$$+$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3}\sum_{j_4=0}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_3^2)(1/\mu_3 + j_1/\mu_1 + j_2/\mu_2 + 2j_4/\mu_3 + j_4/\mu_4)$$
$$+$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=1}^{J_4-1}\pi(j_1,j_2,j_3,j_4,a_4,s_4)(1/\mu_4 + (j_4-1)/\mu_4 + j_1/\mu_1 + j_2/\mu_2 + 2j_4/\mu_3)$$
$$+\left(\frac{C_1}{\mu_1} + \frac{C_2}{\mu_2}\right)C_4^{/}$$

## V. SIMULATION RESULTS

We adopted experimental way to get the simulation results by using the discrete event simulator developed in C++ with different modern set of logics. The developed discrete event simulator is quite modular in design that provides the facility to implement different kinds of scheduling logic under multi disciplinary traffic input. In this study our logic revolves around five different set of polling strategies.

The developed discrete event simulator implements a special Scheduler class, where the design template has been used as discussed in [23]. A traffic generator was also written, which implements the bursty (self-similar) traffic input. This generator may also be over-ridden by another traffic model. A number of other associated classes like Simulation, RandomNumber and Packet, were also written to facilitate program function and accuracy. This section presents a comparison of simulation results of five different kinds of scheduling schemes. In all the cases, the capacity of each queue is 10 packets. The following values have been presented for the class 1 traffic: the session arrival rate is adjusted to $\lambda_1 = 6s^{-1}$, the in-session packet arrival rate is $\alpha_1 = 50s^{-1}$ (the property of VoIP traffic) and the service rate to $\mu_1 = 2500s^{-1}$. Also, the following values have been selected for queue 2, 3 4: the session arrival rate= $\lambda_2 = \lambda_3 = \lambda_4 = 50s^{-1}$, the in-session packet arrival rate= $\alpha_2 = \alpha_3 = \alpha_4 = 6s^{-1}$ and the service rate $\mu_2 = \mu_3 = \mu_4 = \mu_1$.

The effect of varying the Hurst parameter on mean delay has been studied for various newly developed mixed polling schemes. Fig. 3 shows the mean delay vs. Hurst parameter for mixed exhaustive and limited service model, whose analytical framework has been presented in section 4. Fig. 4 presents the mean delay vs. Hurst parameter for mixed exhaustive, gated and limited service model, whereas; Fig. 5 and 6 present the analysis for mixed exhaustive and gated service. In all the cases, we notice that with the increase in Hurst parameter, the queuing delay increases particularly for queue 3 and queue 3.

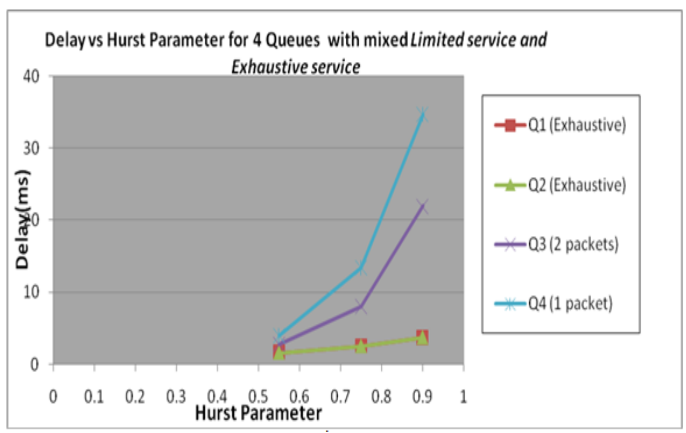


**Fig. 3:** Mean Delay vs. Hurst Parameter:
Simulation Results for 4-Queues with mixed exhaustive and limited service

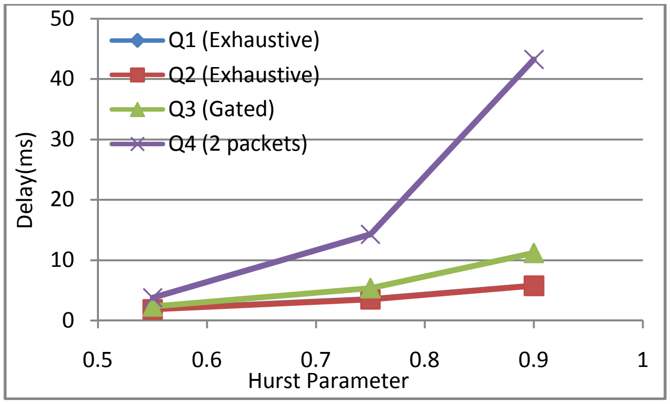


**Fig. 4:** Mean Delay vs. Hurst Parameter mixed Exhaustive, Gated and Limited service

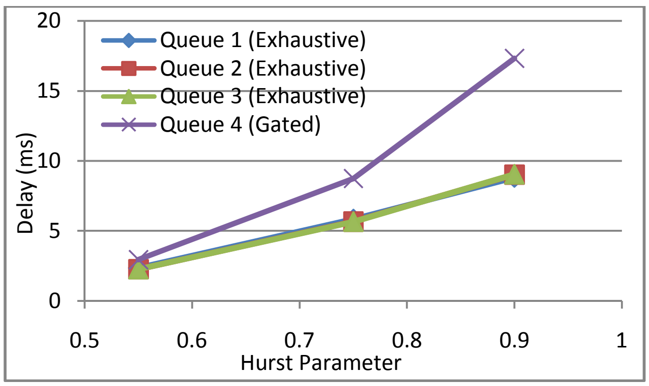


**Fig. 5:** Mean Delay vs. Hurst Parameter mixed Exhaustive and Gated service

Also, we can clearly observe that the mixed exhaustive and limited service model gives very little delay for exhaustive queues at the cost of high delay for limited service queues. As compared to that, the delay for queue 3 and queue 4 has been significantly reduced in mixed exhaustive and gated service models but at the same time, the delay for exhaustive queues has also been increased.

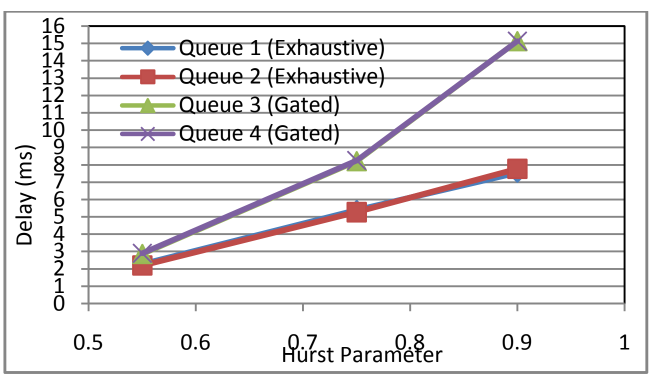


**Fig.6:** Mean Delay vs. Hurst Parameter mixed Exhaustive and Gated service

## VI. Conclusion And Future Work

In this paper, we have presented a detailed analysis for mixed polling schemes for variety of different combinations with multiple classes of self-similar traffic input. To the best of our knowledge, this is the first time that G/M/1 queuing model has been analyzed with four distinct classes of self-similar traffic input for mixed exhaustive and limited service scheme. The closed form expressions of packet delay have been presented for mixed exhaustive and limited service model for four different classes of traffic. Further, we simulated the behavior of multiple classes of self-similar traffic under variety of combinations of polling schemes. The effect of varying the Hurst parameter on mean delay has been studied. In our future work, we intend to analyze more sophisticated combinations of polling schemes with traditional queueing schemes such as priority, WFQ and CBWFQ etc. The main objective is to recommend most suitable combination for different kinds of wireless networks according to the application's requirement found in each network.

## Acknowledgement

The current work has been supported by research centre of college of computer and information sciences at King Saud University, Riyadh, Saudi Arabia.

## References

[1] Al-Hanbali A. Haan R, Boucherie RJ. et. al. (2011), Time-limited polling systems with batch arrivals and phase-type service times, Annal Operational Research, Springerlink.com, DOI 10.1007/s10479-011-0846-Y

[2] Chernova N, Foss S, Kim B. (2011), On the stability of a polling system with an adaptive service mechanism, Annal Operational Research, Springer Science + Business Media, LLC 2011.

[3] AlQahatani S. A. (2011), Performance Modeling and Analysis of Distributed-Based Polling Networks, International Journal of Computer Science Issues (IJCSI), Vol. 8, Issue 5, No 2, September 2011.

[4] Nandayapa LR, Foss S. (2011), Stability and performance of greedy server systems-A review and open problems, Queuing System (2011) 68: (221-227), Springer Science + Business Media, LLC 2011, DOI 10.1007/s11134-011-9235-0

[5] Li y, Zhu G. (2007) " Performance Analysis of Three Gated Service Polling Scheme on Wireless Network , 2nd International Symposium on Wireless Pervasive Computing, 2007.

[6] Vishnevsky V, Semenova O. (2008), Adaptive Dynamical Polling in Wireless Networks, Cybernetics and Information Technologies, vol 8, no 1, 2008, pp. 3-11.

[7] Law R. Y, Leung V. C, Chan H. C. (2006), Polling-based protocols for packet voice transport over IEEE 802.11wireless local area networks, IEEE Wireless Communications, Feb. 2006, pp.22-29.

[8] Vishnevsky V. M, Lyakhov A. I, Guzakov N. N, (2004), An Adaptive Polling Strategy for IEEE 802.11 PCF, In Proc. of 7th Int. Symp. on Wireless Personal Multimedia Communications (WPMC'04). Vol. 1. Abano Terme, Italy, September 12-15, 2004, pp. 87-91.

[9] Vishnevsky V. M, Lakontsev D. V,. Semenova O. V, et. al. (2006), Polling Model for Investigation of the Broadband Wireless Networks, Automation and Remote Control, Vol. 67, no 12, 2006, pp. 123-135.

[10] Geun C, Park B. et. al. (2004), Queuing Analysis of Gated Polling System for Dynamic Bandwidth Allocation Scheme in an EPON, J. Appl. Math. & Computing Vol. 16, No. 1 - 2, 2004, pp. 469 – 481.

[11] Vander Mei R. D. (2007), Towards a Unifying Theory on Branching-Type Polling Systems in Heavy Traffic , Queuing Systems, Vol. 57, No 1, 2007, pp. 29-46.

[12] Al-Hanbali, A., Haan D, Boucherie R , et. al. (2008a). A tandem queuing model for delay analysis in disconnected ad hoc net works. In LCNS:Vol. 5055. Proc. of ASMTA (pp. 189–205), Nicosia, Cyprus, June 2008.

[13] Haan D, Boucherie RJ, Ommeren V. (2009). A polling model with an autonomous server. Queuing Systems, 62(3), 279–308.

[14] Iftikhar M, Rahhal AM, Zuair M. (2011), On the Provisioning of Guaranteed QoS in Wireless Internet with Self-Similar Traffic Input based on G/M/1 Queueing System and Limited Service Polling Model, 2011 International Conference on Network and Electronics Engineering, IPCSIT vol.11 (2011) © (2011) IACSIT Press, Singapore.

[15] Chakravorty R, Cartwright J, Pratt I. (2002), Practical Experience with TCP over GPRS, in IEEE GlobeCom, Nov. 2002.

[16] Balachandran A, Voelker G. MBahl, P. et. al. (2002), Characterizing user behavior and network performance in a public Wireless LAN", Sigmetrics Performance Evaluation. Review, vol. 30. no. 1, 2002, pp. 195- 205.

[17] Cooper R. B. (1970), Queues Served in Cyclic Order : Waiting Times. The Bell Systems Technical Journal, 49:399–413, 1970.

[18] Bux W, Truong H. L.(1983), Mean-delay Approximation for Cyclic-Service Queuing Systems, Performance valuation, 3(3):187–196, 1983.

[19] Lee D.S. (1996), A two-queue model with exhaustive and limited service disciplines. Stochastic Models, 12(2):285–305, 1996.

[20] Hayat M. F, Khan F. Z, et. al. (2011), Performance Modeling of a Priority Scheduling System with Exhaustive Service, Finite Capacity and Switchover, 978-1-4577-0681-3/11, IEEE.

[21] Saffer Z, Telek M. (2011), Analysis of Globally Gated Markovian Limited Cyclic Polling Model and its Application to Uplink Traffic in the IEEE 802.16 Network, Journal of Industrial And Management Optimization Volume 7, Number 3, August 2011.

[22] Boon M, Onno J. et al. (2011), One Problem of Polling System, Queuing System (2011) 68:365–374, Springer, DOI 10.1007/s11134-011-9247-9.

[23] Gross D, Shortle J. F. et al. (2002), Difficulties in Simulating Queues with Pareto Service, In proc. of the 2002 Winter Simulation Conference, Vol. 1, Issue 8-11 Dec. 2002, PP.(405-415). IEEE.

[24] Ijaz A. Shoukat, Iftikhar M. (2011), Experimental Performance Evaluation of different Remote Servers to Analyze the Effect of Divergent Load and Congestion, International Journal of the Physical Sciences, Vol. 6(14), pp. (3354–3368), 18 July, 2011.

[25] M. Caglar, "A Long-Range Dependant Workload Model for Packet Data Traffic", Mathematics of Operations Research, 29, pp. 92-105, 2004

[26] H. P. Schwefel and L. Lipsky, "Impact of Aggregated Self-Similar On/Off Traffic on Delay in Stationary Queueing Models (Extended Version)", Performance Evaluation, 43, 2001, pp. 203-221

[27] I. Kaj, "Limiting Fractal Random Processes in Heavy Tailed Systems", In Fractals in Engineering, New Trends in Theory and Applications, Eds. J. Levy-Lehel, E. Lutton, Springer-Verlag London, 2005, pp. 199-218

[28] G. Fay, F. Roueff and P. Soulier, "Estimation of the memory parameter on the infinite-source poisson process", Bernoulli (13) 2007 473-491

[29] I. Kaj and M.S Taqqu, "Convergence to Fractional Brownian Motion and to the Telecom Process: the integral representation approach" Book Chapter in "In and Out Equilibrium 2" Book Series, Progress in Probability, volume 60, pages 383-427

[30] E. Cinlar, "Introduction to Stochastic processes" 1975