Memory Estimation of Internet Server Using Queuing Theory: Comparative Study between M/G/1, G/M/1 & G/G/1 Queuing Model

(Name : Abhishek.K.B , Roll-No : 04010203 , Term-Paper)

Abstract

How to effectively allocate system resource to process the Client request by Gateway servers is a challenging problem. In this paper, we propose an improved scheme for autonomous performance of Gateway servers under highly dynamic traffic loads. A methodology to calculate Queue Length and Waiting Time is devised utilizing Gateway Server information to reduce response time variance in presence of bursty traffic. The most widespread contemplation is performance, because Gateway Servers must offer cost-effective and high-availability services in the elongated period, thus they have to be scaled to meet the expected load. Performance measurements can be the base for performance modeling and prediction. With the help of performance models, the performance metrics (like buffer estimation, waiting time) can be determined at the development process. This paper describes the possible queue models those can be applied in the estimation of queue length to estimate the final value of the memory size. Both simulation and experimental studies using synthesized workloads and analysis of real-world Gateway Servers demonstrate the effectiveness of the proposed system

INTRODUCTION

The architecture of Gateway Server is in general referred to as "Web services". The term "Web Services" describes specific business functionality through Internet connections,for the purpose of providing a way for another entity to use the services it provides . Web services are the building blocks for the future generation of applications and solutions on the Internet. The services provided through Internet Server(s) are colossal. It transfers information in form of text, images and voice. The whiz of information technology is because of consistent and steady functioning of Internet Server. However, Web traffic is highly dynamic and volatile. The data arrives and departs from different nodes randomly. Thus, we can envisage that, "*number of channels*" for arrival and "*number of channels*" for departing must be identical. The incoming data can be stochastically treated as a "*process*"and so will be the case of departing from the memory of Web Servers. These situations make the working of Memory of Web Servers - a typical case of - "Queuing Process". During the last 40 years, research has shown that queuing models serve as a fundamental tool to model computing systems .In fact, queuing models have been successfully applied to areas such as capacity planning and performance analysis etc.The layout of Gateway Server is depicted in Figure 1, in which data arriving and data departing at a node has been revealed:
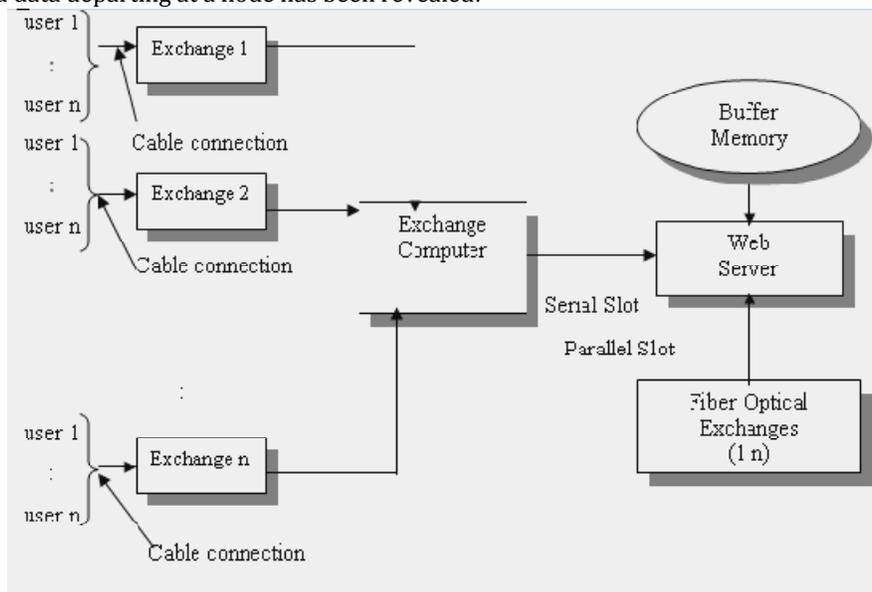
It has been implicated that memory passes through diverse situations of Queue Models, i.e., M/M/1, M/G/1, G/M/1 and G/G/1. In all these four models, M/M/1 model is most disciplined and can be analyzed analytically to estimate the queuing parameters, i.e., "Queue Length" and "Waiting time". These are derived as given in equations (1) and (2):

$$\bar{Q}(L) = \frac{\lambda/\mu}{(1-\lambda/\mu)} \qquad (1)$$

$$\bar{W}(t) = \frac{1}{\mu}\left\{\frac{\lambda/\mu}{(1-\lambda/\mu)}\right\} \qquad (2)$$

The other models (M/G/1, G/M/1, G/G/1) emerges to be most undisciplined and provides situations for calculating queuing parameter's analytically incredibly tough. One can relatively envision that this will be case, which requires bigger value of memory size, thereby, making task a challenging application. However, to get all cases verified, we propose to carry out the study for all the four cases employing stimulation techniques on workstation. Stimulation technique is a viable method when problem arises to computer analytically. Stimulation techniques for queuing models is possible if arrival instances and departing instances are generated having pre-defined rate of arrivals and rate of departures for a given distribution. After gazing the models, it is mandatory to delineate two types of strings, for the given average rate of arrivals and departures.

A String of departure or arrival for pre-assigned average values having negative exponential distribution. The other set of strings having pre-assigned average rate of arrivals/average rate of departure having general distribution. General distribution is a case when there is no distribution for which the string belongs more than, say, 50. The computer for estimation of Queue Length and Waiting time can employ the two strings. For stable maneuver of the system, it is to be ensured that processing speed of workstation be larger than the arrival rates. This condition when convene, is called *"functioning of the system in Ergodic State".*

FORMATION OF ARRIVAL INSTANCES

For model(s) M/M/1, M/G/1, G/M/1, and G/G/1, We need two sets of Arrival Instances, each having same arrival rates but different distribution, namely:-
A. Negative-Exponential Distribution
B. General Distribution
General N-E distribution
The arrival process for N number of arrival instances is accumulated by generating N number of equiprobable distributions. Subsequently, these are transformed to Negative-Exponential distribution for pre-assigned rate of arrivals. Equation (3) is used for transformation of equiprobable data to Negative-Exponential sequence:

$$y = \frac{1}{k}\left(\log(1-x)\right) \qquad (3)$$

where,
y=Negative-Exponential data
x=Equiprobable random between 0 and 1
k= appropriate constant

Let these arrival instances are $_{1,2,\ldots\ldots\ldots N}a\ a\ a$. Then, the arrival instances which is nothing but Markov chain is general using equation (4):

$$A_N = a_{N-1} + a_n \qquad (4)$$

where, A_N for n=0 to N are time instances of the arrival of data. The second string for the arrival is having general distribution. The general distribution has been achieved in the study by merging 5 distributions having pre-assigned rate of arrivals. These distributions are:
1. Bernaulli Distribution
2. Geometric Distribution
3. Equiprobable Distribution
4. Gaussian Distribution
5. Negative-Exponential Distribution

The equiprobable method is provided with the system are used for generation of any distribution employing equal area criteria method using following formula:

A. *Bernaulli Distribution*
The equation for the transformation of equiprobable to Bernaulli distribution is given in equation (5):

$$f(y(i)) = a + by(i) \text{ for } i \in 1 to N \qquad (5)$$

Equiprobable Distribution = $x(i)$
Bernaulli Distribution = $y(i)$, then

$$y(i) = -a \pm \sqrt{a^2 + 2bx(i)/b} \qquad (6)$$

where,
$b = 1/\lambda$ in arrival process
$b = 1/\mu$ in departure process
$a = 1-b$

B. *Transformation of Equiprobable to N-E*
$$y = \frac{1}{k}\left(\log(1-x)\right) \qquad (7)$$

where,
y=Negative-Exponential data
x=Equiprobable random between 0 and 1
k= appropriate constant

C. *Geometric Distribution*
$$f(y(i)) = \frac{a}{(1-by(i))} \quad \text{for } i \in 1 to N \qquad (8)$$

Equiprobable Distribution= $x(i)$ for $i=1$☐1 *to N*
Geometric Distribution = $y(i)$ for $i=1$☐1 *to N* then,
$$y(i) = \frac{1}{b(1-e^{-\delta x(i)})} \qquad (9)$$

where,
$\delta = b/a$
$b=1-a$
$b\lambda = \mu$ for arrival/departure.

D. *Gaussian Distribution*
There is varied procedure of generating the Gaussian Distribution. The accepted practice is by intriguing mean of 12 Equiprobable numbers and arranges them in a String.Mathematical expression is specified beneath:
Let Equiprobable number are $_{1, 2, 12} xx\ldots\ldots\ldots x$ .Then, Gaussian distribution

$$y = \frac{1}{12}\left[\sum_{1}^{12} x(i)\right] \qquad (10)$$

E. *Equiprobable Distribution*
Equiprobable numbers are spawned between 0 and 1 for the identical range using command offered with the system.

FORMATION OF DEPARTURE INSTANCES

As epitomized in arrival process, Departure process will also have two elongated strings for departure instances are to be stimulated for:

Negative-Exponential Distribution
Average rate of departure, $\mu = \lambda + 1$

The identical approach is employed, as been depicted in thearrival process. The other string is made of, for $\mu = \lambda + 1$, having general distributions with combination of 5 distributions as conferred on arrival process. What is done actually, 5 sets of N/5 numbers are generated having distributions – Bernaulli, Gaussian, Negative-Exponential, Geometric, Equiprobable and are merged randomly.
From these two strings, two departure chains are computed using formulae:

$$D_n = D_{n-1} + d_n \qquad (11)$$

where, $n = 0$ to N and $d = n_{th}$ departure duration. To guarantee that departure doesn't take situate before arrival, the departure chain is shifted by 1 second.

QUEUE LENGTH AND WAITING TIME

In the segment, we have computed "Queue Length" and "Waiting Time" for 4 models, namely, M/M/1, M/G/1, G/M/1
and G/G/1, respectively.

A. M/M/1 Model
Two strings are chosen, both having Negative-Exponential Distribution. One should be of arrival process and other for the departure process. At each departing instance, Queue Length, is computed. Let these Queue Length be, $q_1, q_2, \ldots, q_n$. Then, Average Queue Length is

$$\overline{Q(L)} = y_n \sum_{i=1}^{n} Q(i) \qquad (12)$$

$$W(t) = \frac{1}{\mu}\big(Q(i)\big) \qquad (13)$$

B. M/G/1 Model
In this case, arrival process is selected having distribution Negative-Exponential. The departure process is selected having general distribution at the each instance of the departure, Queue Length is counted and using equation(s) 10 and 11 and the average queue length is computed.

C. G/M/1 Model
In this case, the arrival string is chosen having general distribution and the departure distribution is picked up for negative exponential distribution. The Queue Length and Waiting time is computed as described in the segment affirmed above.

D. G/G/1 Model
This is a emblematic case and offers largest Queue length and largest Waiting time. In this case, a string from arrival process is chosen having general distribution. Similarly, a string from departure process is also chosen having generaldistribution. The value of $Q(L)$ and $W(t)$ are computed as the technique portrayed previously.