A Quality of Experience (QOE) Based Estimate of Bandwidth Requirement for Computer Networks

Haruna Bege, Aminu Yusuf Zubairu

Department of Computer Engineering, Nuhu Bamalli Polytechnic, Kaduna State, Nigeria

Abstract: This paper centers on making proper estimate of the bandwidth requirement of a network (which has a specified user population) with the help of Monte Carlo simulation (MCS) - a statistical tool for analyzing stochastic processes. This was done to overcome the challenge of poor Quality of Experience (QoE) resulting from insufficient bandwidth. Three MCS scenarios: fully loaded scenario (FLS), irregular pattern scenario (IPS) and regular pattern scenario (RPS) which are synonymous to high, normal and low traffic periods respectively. A model based bandwidth estimation technique was adopted which involved observing network traffic, classification of the contents accessed on the network, determination of the minimum bandwidth requirement for accessing each class of content that guarantees good QoE is met, lastly obtaining the popularity factor (likelihood) of viewing each content. These served as input data for the simulations from which estimates were obtained.

Keywords; bandwidth, quality of experience (QoE), estimation, Monte Carlo simulation

I. INTRODUCTION

The increasing demand for high speed in data processing to provide good Quality of Experience (QoE) to network users in communication systems has led to the development of many network management techniques. This has also led to an insatiable demand for network resources that have effect on the speed of the network. Although other factors like latency, number of hops and physical distance between devices have impact on QoE, one of the major resources that determine the QoE of users on a network is the bandwidth. In digital systems, bandwidth can simply be defined as the data speed in bits per second (bps). While, in analogue systems bandwidth is defined in terms of the difference between the highestfrequency signal component and the lowest-frequency signal component of a communication channel.

As the number of users on a network increase, bandwidth challenges are bound to be encountered if such a number of users were not anticipated. Some of these challenges are usually evident in streaming multimedia and sometimes it gets as bad as inability to load a page. Proper network planning and dimensioning is one of the key challenges that network operators face. The importance of this issue is highlighted with the necessity of proper decision making regarding capital investments and the permanent increase in traffic volume. Under-investing in network resources leads to low network performances and hence dissatisfied customers, while overinvesting may cause over-dimensioned capacities which leads to lower resource utilization and opportunity losses [1]. This paper focuses on estimating bandwidth requirement for a campus network for good QoE.

II. LITERATURE REVIEW

The demand for high speed INTERNET performance to meet modern day demands like high definition multi-media streaming, online gaming and cloud computing has reinforced the need for studies on estimation of internet connection bandwidth as seen in the works of [2-5]

Bandwidth estimation is broadly classified into two areas which are mainly in terms of channel capacity estimation or available bandwidth estimation. Comparatively, most of the literatures on bandwidth estimation reviewed seem to centre more on available bandwidth estimation (ABE) [6-15]. This aspect of bandwidth estimation is concerned with determining the residual bandwidth during cross traffic session. ABE is usually employed for traffic management to determine if new flow can be admitted on a network while channel capacity estimation is concerned about determining the entire required size of a channel for a satisfactory quality of experience or quality of service as the case may be.

The techniques adopted for determining bandwidth estimates are active, passive and model based techniques [16]. Active techniques involve sending streams of probing packets to explore the entire network. This method generally has the drawback of generating extra traffic in the network. Examples of these tools include pathload, pathrate, pchar, SProbe, Treno, Nettimer and pathchar. Passive techniques rely on monitoring existing traffic between end-hosts to extract estimates. Examples include QoS Enabled routing in Mobile Ad hoc Networks (QoS-AODV) [17], Bandwidth Reservation under InTerferences influence (BRuIT) [18] and Adaptive Admission Control (AAC) [19]. Tools for passive measurement do not generate extra traffic. However, passive techniques are not adequate in estimating the available bandwidth (ABW) because they cannot predict result set after the admission of new flow in a network. They only reckon that on acceptance of new flow, network parameters get changed and affect the real ABW. So future result set cannot be predicted by active and passive bandwidth estimation techniques. Lastly, model based techniques have the unique ability to predict future result sets after admission of new flow in a network. They are analytical techniques which usually involve building a mathematical model and considering three main scenarios: saturated (high traffic), unsaturated (low

traffic) and semi-saturated (normal traffic) scenarios on a network [16]. This paper relies on model based techniques which are analytical techniques that are much more suited for network performance analysis.

III. METHODOLOGY

The steps employed for conducting the research presented in this paper involved design of a campus network, generation and observation of traffic from the designed network, development of a model of which data obtained from the observed traffic served as input to perform Monte Carlo simulations. Estimate of the bandwidth requirement for 900 users was obtained from the MCS results. These steps are shown in more details as follows;

A. Network design

Campus network designs are usually hierarchical. The design presents three layers which are the core layer, the distribution layer and the access layer. The distribution layer switches and higher layers are directly connected to the internet, while access layer switches are directly connected to the end users (computer or smart devices). In order to collect data, a network was set up from higher layers to lower layers in the hierarchies of networks. A dedicated database server was responsible for the collection of network traffic. As depicted in Figure 1, traffic data was collected on a private network by configuring different user profiles. This is a typical scenario of a campus network where each student or staff user has a unique username and password for accessing the internet.



Figure 1: Topology for Data Collection

B. Traffic Generation

Approximately, 50 computers (users) were accessing and surfing various websites on the network at the same time 24/7 in order to generate diversified traffic within a particular *C*. interval of time. The captured traffic from different LANs and D

WANs were monitored and stored in a database server. This was done constantly without any interruption and downtime.

As a network setup test bed, it was subjected to rigorous usage within a month to collect traffic data as presented in Figures 2-4





"Weekly" Graph (30 Minute Average)



"Monthly" Graph (2 Hour Average)



Figure 4: Monthly Traffic Data

C. Traffic Observation

In this study, the In-Out traffic on the Router's ports was monitor for all the incoming and outgoing traffic with the aid of Wireshark software. The traffic was recorded for varying times in hours and minutes as shown in Figure 5.



Figure 5: Sample of Traffic Observed

The extracted sample captured data is presented in Table 1.

			1	-
S/N	Days	Week Day	Time (Hours)	Payload (kilobytes)
1	Day 1	Thursday	24	1184110.82
2	Day 2	Friday	24	4910853.06
3	Day 3	Saturday	24	5978301.75
4	Day 4	Sunday	24	4644118.73
5	Day 5	Monday	24	8899905.00
6	Day 6	Tuesday	24	5476324.33
7	Day 7	Wednesday	24	1234567.67
8	Day 8	Thursday	24	3214561.65
9	Day 9	Friday	24	4567891.76
10	Day10	Saturday	24	1289085.76
11	Day 11	Sunday	24	5965301.75
12	Day 12	Monday	24	4710835.06
13	Day 13	Tuesday	24	3116451.14
14	Day 14	Wednesday	24	3011344.68
15	Day 15	Thursday	24	1156789.10
16	Day 16	Friday	24	9867543.10
17	Day 17	Saturday	24	1235471.12
18	Day 18	Sunday	24	1117854.19
19	Day 19	Monday	24	1213450.78
20	Day 20	Tuesday	24	4348721.11
21	Day 21	Wednesday	24	3216783.10
22	Day 22	Thursday	24	2315676.10
23	Day 23	Friday	24	7653293.10
24	Day 24	Saturday	24	2569791.76
25	Day 25	Sunday	24	6215643.1

Table 1: Extracted Sample Data

26	Day 26	Monday	24	1247321.88
27	Day 27	Tuesday	24	4126543.00
28	Day 28	Wednesday	24	3354678.01
29	Day 29	Thursday	24	1123863.11
30	Day 30	Friday	24	4590345.72

As presented in the captured sample data in Table 1, the bandwidth payload captured for the different days of the week were stored. The payload comprises of traffic generated from research, social media and multi-media streaming sites. The contents accessed were classified into video, audio, image and text based data as described in [20]. Based on a single user experience, the data obtained was analyzed from which the minimum bandwidth (MB) requirement and popularity factor (PF) (based on frequency of access) was obtained and presented in Table 2.

Content	PF	MB (Mbps)
Video	0.35	3.9
Audio	0.08	2.8
Image Text	0.34 0.23	1.8 0.005

IV. SIMULATION MODEL

The simulation was performed by considering three scenarios of the population (P) using the network. These three scenarios considered for the simulation process are the fully loaded scenario, regular pattern scenario and the irregular pattern scenario which corresponds with the saturated, unsaturated and semi-saturated scenarios mentioned earlier.



Figure 6: Parametric model of Monte Carlo technique [21]

Considering the parametric model shown in Fig.6, let the dependent variable Y represent the set of outputs (bandwidths) generated from the MC simulation, X represent the minimum bandwidth (MB) associated with each content on the network and the independent variable P represent the population using the network at every instance.

As the population on a network increases, so also the contents accessed as well as the bandwidth required to meet the

requests of the network users. Therefore the estimated bandwidth *Y* is proportional to *P* and *X*.

Mathematically,

 $Y = kPX \dots \dots \dots \dots \dots (2)$

Where *k* is a constant in equation 2.

Assuming k to be unity (k = 1), the general mathematical relation governing the bandwidth estimation simulation process is hereby given by equation 3.

Where

Y = Estimated bandwidth

P = Population on the network

X = Minimum bandwidth of network contents

It should be noted that in equation 3, *m* represents the number of iterations performed; *n* represents the total population considered. For further clarity equation three could be considered as a matrix with Y_j being a column matrix with elements $y_1 y_2 y_3 \dots y_m$, *X* a column matrix with elements, x_0 , x_1, x_2, x_3 and x_4 corresponding to [*nil* (C_0), video (C_1), audio (C_2), image (C_3), text (C_4)] and representing the MB of the network contents given as [0, 3.9, 2.8, 1.8, 0.0048] Mbps respectively. P_k is a row matrix with elements P_0, P_1, P_2, P_3 and P_4 corresponding to the population viewing contents C_0, C_1, C_2, C_3 and C_4 respectively. $P_0, P_1,$ P_2, P_3 and P_4 must always sum up to *n*. With this view, equation 3 can be represented by equation 4. It should also be noted that C_0 was introduced as dummy (0) content assigned to P_0 to represent those absent from the entire population

$\begin{bmatrix} y_1 \end{bmatrix}$						$\begin{bmatrix} x_0 \end{bmatrix}$	
<i>y</i> ₂						$ x_1 $	
	$= [p_0]$	p_1	p_2	p_3	$p_4]$.	$ x_2 $	
						<i>x</i> ₃	
y_m						$\begin{bmatrix} x_4 \end{bmatrix}$	(4)

being considered on the network for each iteration.

The simulation was performed using MATLAB R2015a (8.5.0.197613), 64 bits and Microsoft Excel 2013. For the three scenarios, little variations were made on equation 3 to fit the peculiarity of each case. Fig.7. shows the flowchart of the simulation process.



Fig.7: Flowchart of the simulation process

Generally, in theory, as the number of users on a network increases, the bandwidth consumed by the users also increases. This implies that bandwidth requirement will continue to increase infinitely as the users of the network increase. Practically bandwidth on a network is never infinite as it is constrained by the limitations of the network hardware. Nevertheless, the relationship between the number of users of the network and the amount of bandwidth consumed was verified by MCS for 900 users using the simulation parameters in table 2. This is shown in Fig 8.



Fig 8: Bandwidth/Network user relationship from MC simulation

a. Fully Loaded Scenario

Considering the population of 900 distinct users with an average of 526 users logged on at every instant, with the possibility of having all users needing access to the network simultaneously. This constitutes a scenario where the highest demand on the network in terms of bandwidth and other resources is expected. Therefore, the Fully Loaded Scenario (FLS) is a scenario that assumes all users of the network are logged on and are active. In this case each user was considered to be accessing at least one of the available contents on the network. The variables (parameters) for simulation in this case are shown in Table 3.

Table 3: Parameters for simulation in FLS

Content	Video	Audio randomly ac	Image ross 900 users	Text
PF	0.35	0.08	0.34	0.23
MB	3.9Mbps	2.8Mbps	1.8Mbps	0.005Mbps

For the FLS, equation 3 transforms into equation 5.

While equation 4 becomes equation 6



b. Regular Pattern Scenario

In comparison to the FLS, the regular case does not assume all users to be logged on and actively using the network. Rather it takes cognisance of the behaviour of a network. Considering the average of 526 users usually logged on and active, the Regular Pattern Scenario (RPS) assumes a user population of 526 and assigns the simulation parameters randomly across them while the remaining population are assigned a dummy PF of zero (0) to indicate their absence or are simply not considered in the simulation. Table 4 shows the variables for this simulation.

Table 4: Parameters for simulation in RPS

~	Video	Audio	o Imag	e Text		
Content	randomly across 526 users					
PF	0.35	0.08	0.34	0.23		
MB	3.9Mbps	2.8Mbps	1.8Mbps	0.005Mbps		

For the RPS, equation 3 transforms into equation 7.

While equation 4 becomes equation 8

$$\begin{bmatrix} y_1 \\ y_2 \\ \Box \\ \Box \\ y_m \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$
.......(8)

c. Irregular Pattern Scenario

Like the FLS, the Irregular Pattern Scenario (IPS) assumes all 900 users to have access to the network but not necessarily simultaneously. This is because users log on randomly. In this case a population of 900 is used but with the dummy PF of zero (0) to represent those not logged on at any instant. This case differs from the other cases in the sense that the active population continues to vary with every single iteration because the population absent also varies randomly with each iteration due to random assignment of 0 in the PF. the simulation parameters for this case are shown in Table 5.

Table 5: Parameters for IPS with dummy PF

Content	Video	Audio	Image	Text A	bsent	
	randomly across 900 users					
PF	0.35	0.08	0.34	0.23	0	
MB	3.9Mbps	2.8Mbps	1.8Mbps	0.005Mbps	0	

For the IPS, equation 3 transforms into equation 9.

While equation 4 becomes equation 10

$$\begin{bmatrix} y_{1} \\ y_{2} \\ \Box \\ \Box \\ y_{m} \end{bmatrix} = \begin{bmatrix} p_{0} & p_{1} & p_{2} & p_{3} & p_{4} \end{bmatrix} \cdot \begin{bmatrix} x_{0} \\ x_{1} \\ x_{2} \\ x_{3} \\ x_{4} \end{bmatrix} \dots \dots \dots (10)$$

Discussion of the results obtained from these simulations was made.

V. RESULTS

This section contains discussion of these results obtained from the MC simulations performed and was used to draw conclusions on the estimate of the bandwidth requirement

a. Fully Loaded Scenario

As explained in the preceding section, the FLS assumes all 900 users to be logged on and actively using the network simultaneously. Therefore, the MCS was performed with the variables specified in Table 2. The simulation was performed for 1000 iterations and the results obtained are shown in Fig 9 and Fig 10.



Fig. 9: MCS result for FLS showing its randomness



Fig 10: Range of bandwidth series generated for FLS MCS

MEAN =	1913.55 Mbps
MEDIAN =	1915.49 Mbps
MIN =	1777.23 Mbps
MAX =	2041.45 Mbps
STD. DEV=	43.54 Mbps

The result in Fig 9 shows the random variation of bandwidth consumed with each iteration by the 900 users assumed to be on the network. The Fig 10 shows a series of the bandwidths obtained from 1000 simulations with the minimum value at 1777.23 Mbps and the maximum value at 2041.45 Mbps. The variation of bandwidth observed with each iteration in Fig 9 is as a result of the randomness in assignment of the network contents across the 900 users, this depicts how users access content at will. The bandwidth estimated by this scenario is given by the mean value of the series which was obtained to be 1913.55 Mbps. This value gives the estimated minimum bandwidth requirement for good QoE when all 900 users are active on the network. This value can be considered to be the upper limit of bandwidth requirement for the 900 user. In reality this scenario is less likely to occur if users log on at will. The PF gives an idea of the bandwidth consumed by the various contents as shown in Table 6 and is applicable to the other scenarios.

Content	PF	% of estimate	consumption (Mbps)
Video	0.35	35%	669 74

Table 6: Content bandwidth consumption

Video	0.35	35%	669.74
Audio	0.08	8%	153.1
Image	0.34	34%	650.6
Text	0.23	23%	440.1

b. Regular Pattern Scenario

The MCS for the RPS was also performed using the parameters in table 3. Again the number of iteration I = 1000 was maintained. The results obtained from this simulation are shown in Fig 11 and Fig 12.



Fig 11: MCS result for RPS showing its randomness





Fig 12: Range of bandwidth series generated for RPS MCS

MEAN =	1118.86 Mbps
MEDIAN=	1118.81 Mbps
MAX =	1215.54 Mbps
MIN =	1023.08 Mbps
STD.DEV=	34.19 Mbps

The results in Fig 11 and Fig 12 show a variation of bandwidth between 1023.08 Mbps and 1215.54 Mbps. Like the case of the FLS, even though the number of users is fixed at 526, the bandwidth consumed varies randomly as a result of the random process of accessing various contents. The estimated bandwidth for this scenario is given by the mean value 1118.86 Mbps. This value is lower than the estimate of 1913.55 Mbps obtained in the FLS because the number of users is lower in the RPS. This is expected as shown by the user/bandwidth linear relationship. The bandwidth estimate obtained from this scenario is highly dependent on the behaviour of the network users as it makes use of the average population active on a network. As a result, this scenario is expected to give the most reliable estimate but can be misleading if the period of observation of the network is not sufficiently long to capture all activities on the network.

c. Irregular Pattern Scenario

Like the previous scenarios, the IPS simulation was performed with I = 1000. This scenario depicts a typical real life situation where all users log on randomly. It does not put into consideration the pattern or the average number of active users observed on a network. Therefore, in the simulation all PFs including the dummy (PF = 0) were distributed randomly across the 900 users for each iteration. In other words, the simulation variables in Table 5 were used and the results obtained are shown in Fig 11 and Fig 12.



iterations

1300 1250

Fig 12: Range of bandwidth series generated for IPS MCS

MEAN =	1531.12 Mbps
MEDIAN=	1532.88 Mbps
MAX =	1664.02 Mbps
MIN =	1397.44 Mbps
STD. DEV=	45.09 Mbps

The IPS result shows a random variation of bandwidth between 1397.44 Mbps and 1664.02 Mbps. The expected value of 1531.12 Mbps being the estimated bandwidth obtained and lies between 1118.86 Mbps and 1913.55 Mbps for RPS and FLS respectively. This scenario unlike the other two depicts reality by not considering a fixed number of users on the network. This reason makes the IPS unique as it can be used to make reasonable estimate for networks without necessarily having information about the behaviour of the population.

VI. CONCLUSION

A plot of the linear relationship, FLS, RPS and the IPS on a single graph is shown in Fig.13.



Fig.13: Single graphical result of FLS, IPS, RPS and LRBW

The Fig.13 shows a comparative view of the various scenarios. The FLS is at the peak and ranges between 1777.23 Mbps and 2041.45 Mbps. The LRBW for 900 users meets the FLS at exactly 1933.10 Mbps, this value specifies the bandwidth requirement when the network is fully loaded and is in agreement with that obtained in the bandwidth/network user relationship result in Fig 8. The LRBW flattens and remains constant at the meeting point with the FLS at exactly 900 user indicating that the LRBW simulation was done for only 900 users whereas the other curves were for 1000 simulations.

The IPS curve lies between the FLS and RPS curves showing that in reality, there will hardly be a time where all 900 users will be logged on and actively using the network when left to log on at will. It also shows that when left at will, the users will consume bandwidth greater than that estimated for the 526 average active users shown by the RPS result even though the average was drawn from observing a network with the users operating at will also.

Although the RPS estimate of 1118.86 Mbps was obtained based on the behavior of the population using a network, in a situation where the period of observation of the network is not sufficiently reliable, the best estimate to work with will be the IPS which gives an estimate of 1532.88 Mbps in this study. An example of such situation is a typical academic institution. If the period of observation does not cover all the different seasons experienced on the network, in this example (at least one academic year) which have different impacts on the behavior of network user, the IPS presents a better estimate as it is not constrained by the period of observation of the network.

ACKNOWLEDGMENT

The authors of this work are grateful to the Tertiary Education Trust Fund (TetFund) for funding this Institution Based Research.

REFERENCES

- B. Mikavica, V. Radojičić, and A. KostićLjubisavljević, "estimation of optical access network bandwidth demand using Monte Carlo simulation," International Journal for Traffic and Transport Engineering, vol. 5, no. 4, pp. 384-399, 2015.
- www.rsisinternational.org

- [2]. K. Lee, Heung & Hall, Varrian & Hwan Yum, Ki & Kim, Kyoung & Kim, Eun. (2006) 'Bandwidth Estimation in Wireless LANS for Multimedia Streaming Services''. pp1181-1184
- [3]. Awodele & Akanni, Adeniyi. (2019). Estimation of required bandwidth for organisations, UNIASCIT, Vol 2 (2), 2012, 257-262 ISSN 2250-0987
- [4]. J. Jung Kim.(2017). "Bandwidth Estimation of Networks with Random Services". International Journal of Communication Technology for Social Networking Services. Vol 5. Pp 1-6.
- [5]. M. E. McLaughlin and J. Moran "Predicting Bandwidth Demand and Network Planning Implications on the Internet" March 2004
- [6]. M. Jain and C. Dovrolis, "End-to-end available bandwidth: measurement methodology, dynamics, and relation with TCP Journal of Computer Networks and Communications 9 throughput," IEEE/ACM Transactions on Networking, vol. 11, no. 4, pp. 537–549, 2003.
- [7]. K. Heni, L. Fatma, F. Mounir and Farouk "A QOS routing protocol based on available bandwidth estimation for wireless adhoc networks" International Journal of Computer Networks & Communications (IJCNC), Vol 3, No.1, January, 2011.
- [8]. E. Goldoni and M. Schivi. (2010) "End-to-end available bandwidth estimation tools; an experimental comparison Traffic monitoring and analysis," Lecture notes in Computer Science volume. 6003 p. pp.171–182.
- [9]. J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," in Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC '03), pp. 39– 44, October 2003
- [10]. M. Jain & C. Dovrolis, "Pathload: A measurement tool for end-toend available bandwidth", in Proceedings of passive and active measurements workshop, 2002, pp. 14–25.
- [11]. J. Sommers, P. Barford and W. Willinger (2006) "A Proposed Framework for Calibration of Available Bandwidth Estimation Tools". PP 709-718.
- [12]. M. Kassim, A Azmi, R Ab.Rahman, M. I. Yusof, R. Mohamad, A. Idris "Bandwidth Control Algorithm on YouTube Video Traffic in Broadband Network," vol. 10, no. 1-5, pp. 151-156, 2018
- [13]. Chitanana, L., & Govender, D. W. (2015). Bandwidth management in the era of brings your own device. The Electronic Journal of Information Systems in Developing Countries, 68(1), 1-14
- [14]. V. Ribeiro, R. Riedi, R. Baraniuk, J. Navratil, and L. Cottrell, "PathChirp: efficient available bandwidth estimation for network paths," in Proceedings of the Passive and Active Measurements Workshop, 2003.
- [15]. N. Rana, K. P. Bhandari, and S. J. Shrestha, "Network Bandwidth Utilization Prediction Based on Observed SNMP Data," vol. 13, no. 1, pp. 160168, 2017
- [16]. M. Airon and N. Gupta. (2017) "Bandwidth estimation tools and techniques: A Review". Mangalam university, Guregoan, India 2017.
- [17]. R.de Renesse, V. Friderikos and H.Aghvami, "QoS enabled routing in mobile ad hoc networks", in Fifth IEEE international conference on 3G mobile communications technologies, 2004, vol.1, no. 4, pp.678-682.
- [18]. Claude Chaudet, Isabelle Guerin Lassous, "BRuIT: Bandwidth Reservation under In Terferences influences", in Proceedings of European Wireless, 2002.
- [19]. de Renesse, R., Ghassemian, M., Friderikos, V., & Aghvami, A. H., "Adaptive admission control for ad hoc and sensor networks providing quality of service", in Technical Report, King College London, 20
- [20]. T. Alexandru, D. Marcelo, F. Serge and A. Panayotis. (2014) "A Survey on Predicting the Popularity of Web Content." Journal of Internet Services and Applications, Vol 5,
- [21]. J. Wittwer, "Monte Carlo Simulation Basics,". [Online].Available:https://www.vertex42.com/ExcelArticles/mc/M onteCarloSimulation.html