Determination of the End-To-End Delays of any Switched Local Area Network

M. O.EYINAGHO‡
Department of Electrical and Information Engineering,
Covenant University, Ota, Nigeria
eyimon@yahoo.com

ABSTRACT
In designing Switched Local Area Networks (Switched LANs), there is the inherent need of being able to compute all the end-to-end delays and hence, the average end-to-end delays of such LANs. The widely held notion in the literature is that, the ability to perform this computation is contingent on the formulation of an origin-destination traffic matrix of the network with respect to the hosts that are attached to it. We have shown in a previous paper that this notion does not seem to be correct. In this paper, we explain a methodology for enumerating all the end-to-end delays of any switched LAN and hence, of determining the average (or average of maximum) end-to-end delay of any LAN.

Keywords
End-To-End Delay, Switched Local Area Networks

IJCIR Reference Format:

M. O.Eyinagho. Determination of the End-To-End Delays of any Switched Local Area Network. International Journal of Computing and ICT Research, Vol. 6, Issue 1. pp 24-30.
http://www.ijcir.org/volume6-numbe1/article 3.pdf

1. INTRODUCTION

In a previous paper (Eyinagho et al., 2011), it was shown that the widely held notion in the literature (for example, Kanem et al., 1999, Torab and Kanem, 1999) that to determine the average network delay of any switched Local Area Network (LAN), the origin-destination traffic matrix for all the hosts that are attached to the LAN has to be enumerated, does not seem to be correct. For example, in (Torab and Kanem, 1999), it was stated that, let T_{ij} denote the information flow in packets per unit time between end nodes i and j of the network; we define the traffic matrix of the network to be the n×n matrix T= (T_{ij}), where n is the number of end nodes. Moreover, Kanem et al. (1999) defined the average end-to-end delay of a switched Ethernet local area network, as the weighted combination of all end-to-end delay times; Elbaum and Sidi (1996) on the other hand, defined minimum average network delay as the average delay between all pairs of users in the network. A theory was, therefore, proposed in (Eyinagho et al., 2011) that, for any switched LAN, the hosts that are attached to the LAN cannot simply be enumerated in order to calculate all origin-destination end-to-end delays. The origin-destination end-to-end delays are usually needed for the purpose of designing upper bounded delay switched LANs, in efforts at solving the delay problem of this class of networks. How then should the end-to-end delays of any switched LAN be enumerated? This is the subject of this paper.

‡ Author's Address: M. O.Eyinagho. Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria.
eyimon@yahoo.com

International Journal of Computing and ICT Research, ISSN 1818-1139 (Print), ISSN 1996-1065 (Online), Vol. 6, Issue 1, pp.24-30, December 2012.

2. DETERMINING THE END-TO-END DELAYS OF ANY SWITCHED LOCAL AREA NETWORK

Shown in Figure 1 is a switched local area network having 3 switches and 8 hosts. For host H_1 to communicate with host H_8, the data packets must go through switches S_2 and S_1. The same thing happens if either hosts H_2 or H_3 want to communicate with host H_8. And hosts H_1, H_2, and H_3 cannot simultaneously communicate with host H_8; packets must wait for their turn to be sent to host H_8. On the other hand, an host that is attached to switch S_2 can be communicating with H_8 simultaneously as another host that is attached to switch S_2 which is communicating with an host that is attached to switch S_3.This means that, no matter the number of hosts that are attached to switch S_2 or the number of hosts that are attached to switch S_1 for example, we only have one end-to-end delay (and hence, one maximum end-to-end delay) between the hosts that are attached to switch S_2 and the hosts that are attached to switch S_1. We can therefore, aggregate all the hosts that are attached to any of the switches in a switched LAN as shown in Figure 2. Then if all the hosts that are attached to any switch in a switched LAN simultaneously have packets that are destined for another (the same) host in the LAN, then we have a maximum switch delay situation (as the offered load to a network or network device is the aggregate sum of the data packet rates presented to the network or network device (Gerd, 1989, p.203)). This maximum switch delay situation can be seen as one of the reasons for the difficulty in uploading/downloading to and from the Internet at certain times of the day when many users that are attached to the LAN are trying to upload and download at the same time. We will use these ideas to develop a methodology for designing upper-bounded delay switched local area networks.

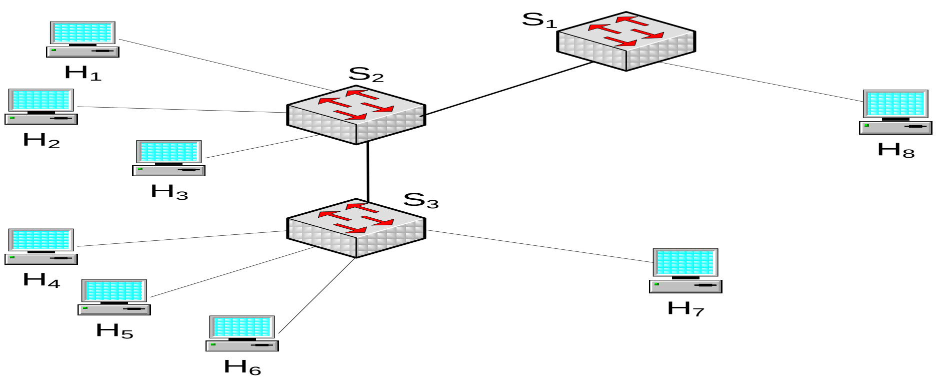

Figure 1 A three (3) switch, eight (8) hosts switched LAN

Referring to Figure 1, assuming switches S_1, S_2, and S_3, will be placed by the network designer (installer) as a result of the number of hosts and the locations of the hosts in a new LAN installation; we have been able to deduce that,

if m = the number of switches in the LAN,

p = the number of maximum end-to-end delays required for the design of an upper delay bounded LAN,

$$\text{then, } p = \sum_{x=0}^{m-1}(m-x) \qquad (1)$$

here, m =3,

$$\text{then, } p = \sum_{x=0}^{3-1}(3-x) = (3-0)+(3-1)+(3-2)$$

$$= 3+2+1 = 6 \text{ maximum end-to-end delays.}$$

Eq.(1) is called the Monday Eyinagho's equation for calculating the number of end-to-end delays in any switched local area network.

We use Figure 2 (Figure 2 is Figure 1 redrawn, with all the hosts that are attached to a switch aggregated as a single host) to illustrate how to enumerate the end-to-end delays (and hence, the maximum end-to-end delays) for any switched LAN. Our method (algorithm) for this enumeration which we call, 'right-most, pre-order transversal' (method of growing the spanning tree of the switches in any switched LAN) has two steps:

1. Process the root switch,
2. Transverse the right-most sub-tree pre-order, until all the switches have been processed.

Performing 'left-most, pre-order transversal' (the opposite of 'right-most, pre-order transversal') results in the same solution. The important thing is that, both methods cannot be mixed for the same LAN design.

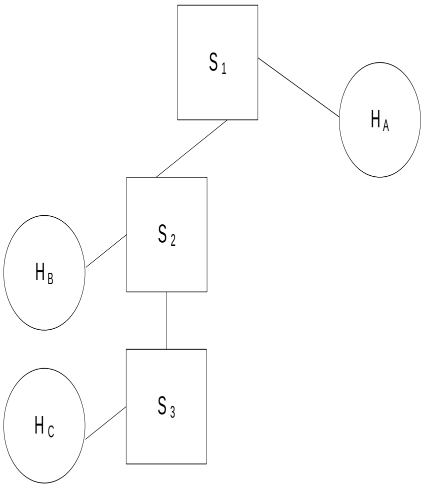

Figure 2 A Switched LAN, with the hosts that are attached to switches S_1, S_2, and S_3 aggregated as H_A, H_B, and H_C respectively.

Using the right-most, pre-order transversal, the required end-to-end delays are enumerated as follows (we take switch S_1 as the root switch).

1. Switch S_1 is placed as shown in Figure 3

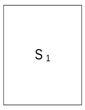

Figure 3 Switch S_1 is placed

2. Next, place switch S_2 and connect S_1 to S_2. This is shown in Figure 4.

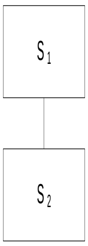

Figure 4 Switches S_1 and S_2 are placed and connected together

Now, for any host that is attached to switch S_1 to communicate with any host that is attached to switch S_2, data packets will experience delay in switch S_1 and switch S_2 and vice versa; this is 1 end-to-end delay.

(delay in S_1 + delay in S_2) = 1 end-to-end delay

3. Next, place switch S_3 and connect to already placed switches as shown in Figure 5.

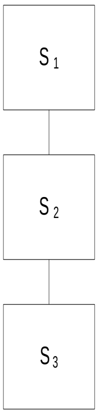

Figure 5 Switches S_1, S_2, and S_3 are placed and connected together

Again, for any host that is attached to switch S_1 to communicate with any host that is attached to switch S_3, the data packets will experience delays in switch S_1, switch S_2, and switch S_3; this is 1 end-to-end delay.

(delay in S_1 + delay in S_2 + delay in S_3) = 1 end-to-end delay

Also, for any host that is attached to switch S_2 to communicate with any host that is attached to switch S_3, the data packets will experience delays in switch S_2 and switch S_3; this is one end-to-end delay.

(delay in S_2 + delay in S_3) = 1 end-to-end delay

We now have 3 end-to-end delays already. It can also be seen that, for any host that is attached to switch S_2 to communicate with another host that is attached to the same switch S_2, data packets will experience delay only in switch S_2; this is 1 end-to-end delay. This is illustrated in Figure 6.

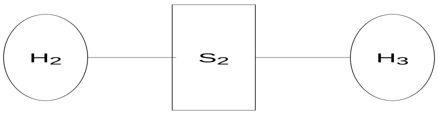

Figure 6 One (1) end-to-end delay through switch S_2

We have a similar situation if any host that is attached to switch S_3 wants to communicate with another host that is attached to the same switch S_3. This is similarly the case with hosts that are attached to switch S_1 (although Figure 1 shows that only one host (H_8) is attached to switch 1 – which is just for illustration purpose). These three situations gives us 3 end-to-end delays, which added to the previous end-to-end delays gives us a total of 6 end-to-end delays, which is what is obtained from Eq. (1).

While the network shown in Figure 1 is quite simple, our method of enumerating the number of end-to-end delays and of calculating these end-to-end delays can be applied to any switched LAN, no matter how complex it is. We now illustrate the methodology with the hypothetical switched LAN that is shown in Figure 7 without going into detailed explanation.

Basically, associated with each of the switches S_1, S_2, S_3, S_4, S_5, S_6, S_7, and S_8 is an end-to-end delay as a result of a host that is attached to a switch wanting to communicate with another host that is attached to the same switch. This gives us 8 end-to-end delays. We now apply our right-most, pre-order transversal method to the network.

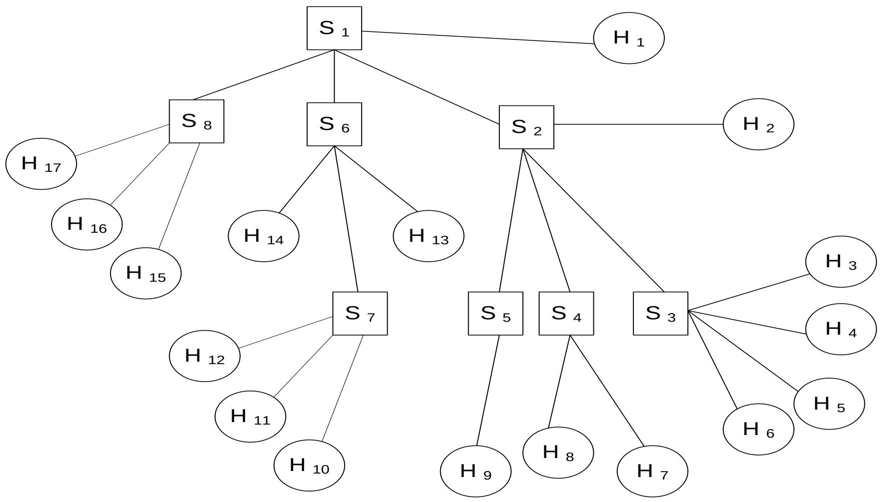

Figure7 An hypothetical Switched LAN having 8 switches (Si's) and 17 hosts (Hi's)

1. Switch S_1 is placed.

2. Switch S_2 is placed and connected to it; there is 1 end-to-end delay between switch S_1 and switch S_2.
(delay in S_1 + delay in S_2) = 1 end-to-end delay

3. Switch S_3 is now connected to switch S_2; there is 1 end-to-end delay between switch S_2 and switch S_3.
(delay in S_2 + delay in S_3) = 1 end-to-end delay

There is another end-to-end delay between switch S_1 and switch S_3.
(delay in S_1 + delay in S_2 + delay in S_3) = 1 end-to-end delay

4. .Switch S_4 is connected to switch S_2; there is an end-to-end delay between switches S_2 and S_4.
(delay in S_2 + delay in S_4) = 1 end-to-end delay

There is an end-to-end delay between switches S_3 and S_4.
(delay in S_3 + delay in S_2 + delay in S_4) = 1 end-to-end delay

There is another end-to-end delay between switches S_1 and S_4.
(delay in S_1 + delay in S_2 + delay in S_4) = 1 end-to-end delay

By similarly adding switches S_5, S_6, S_7 and S_8 in that order to the network, we have the following end-to-end delays.

5. When switch S_5 is connected to switch S_2.
(delay in S_2 + delay in S_5) = 1 end-to-end delay
(delay in S_4 + delay in S_2 + delay in S_5) = 1 end-to-end delay
(delay in S_3 + delay in S_2 + delay in S_5) = 1 end-to-end delay
(delay in S_1 + delay in S_2 + delay in S_5) = 1 end-to-end delay

6. When switch S_6 is connected to switch $S_{1.}$
(delay in S_1 + delay in S_6) = 1 end-to-end delay
(delay in S_2 + delay in S_1 + delay in S_6) = 1 end-to-end delay
(delay in S_3 + delay in S_2 + delay in S_1 + delay in S_6) = 1 end-to-end delay
(delay in S_4 + delay in S_2 + delay in S_1 + delay in S_6) = 1 end-to-end delay
(delay in S_5 + delay in S_2 + delay in S_1 + delay in S_6) = 1 end-to-end delay

7. When switch S_7 is connected to switch S_6.
(delay in S_6 + delay in S_7) = 1 end-to-end delay
(delay in S_1 + delay in S_6 + delay in S_7) = 1 end-to-end delay
(delay in S_2 + delay in S_1 + delay in S_6 + delay in S_7) = 1 end-to-end delay

(delay in S_3 + delay in S_2 + delay in S_1 + delay in S_6 + delay in S_7) = 1 end-to-end delay
(delay in S_4 + delay in S_2 + delay in S_1 + delay in S_6 + delay in S_7) = 1 end-to-end delay
(delay in S_5 + delay in S_2 + delay in S_1 + delay in S_6 + delay in S_7) = 1 end-to-end delay

8. When switch S_8 is connected to switch S_1
(delay in S_1 + delay in S_8) = 1 end-to-end delay
(delay in S_2 + delay in S_1 + delay in S_8) = 1 end-to-end delay
(delay in S_3 + delay in S_2 + delay in S_1 + delay in S_8) = 1 end-to-end delay
(delay in S_4 + delay in S_2 + delay in S_1 + delay in S_8) = 1 end-to-end delay
(delay in S_5 + delay in S_2 + delay in S_1 + delay in S_8) = 1 end-to-end delay
(delay in S_6 + delay in S_1 + delay in S_8) = 1 end-to-end delay
(delay in S_7 + delay in S_6 + delay in S_1 + delay in S_8) = 1 end-to-end delay

We have systematically enumerated all the end-to-end delays that are inherent in this LAN. The total end-to-end delays = 8+1+2+3+4+5+6+7 = 36, which is what is obtained by applying Eq. (1). Eq. (1) can therefore, be used as a check to ensure that all the end-to-end delays that are inherent in any switched local area network has been systematically enumerated using our right-most, pre-order transversal method.

With respect to Figure 2 therefore, if;

end-to-end delay 1-1 = time for a data packet to cross S_1 for any host that is attached to switch 1 that is communicating with another host that is attached to the same switch 1

end-to-end delay 2-2 = time for a data packet to cross S_2 for any host that is attached to switch 2 that is communicating with another host that is attached to the same switch 2.

end-to-end delay 3-3 = time for a data packet to cross S_3 for any host that is attached to switch 3 that is communicating with another host that is attached to the same switch 3.

end-to-end delay 1-2 = time for any data packet to cross S_1 + the time for the same data packet to cross S_2 for any host that is attached to switch 1 that is communicating with another host that is attached to switch 2.

end-to-end delay 2-3 = time for any data packet to cross S_2 + the time for the same data packet to cross S_3 for any host that is attached to switch 2 that is communicating with another host that is attached to switch 3.

end-to-end delay 1-2-3 = time for any data packet to cross S_1 + the time for the same data packet to cross S_2 + the time for the same data packet to cross S_3 for any host that is attached to switch 1 that is communicating with another host that is attached to switch 3.
Therefore, assume that the,
maximum time for any data packet to cross S_1 = x_1 seconds,
maximum time for any data packet to cross S_2 = x_2 seconds,
maximum time for any data packet to cross S_3 = x_3 seconds,

Then,
maximum end-to-end delay 1-1 = x_1 seconds, maximum end-to-end delay 2-2 = x_2 seconds,
maximum end-to-end delay 3-3 = x_3 seconds, maximum end-to-end delay 1-2 = (x_1+x_2) seconds,
maximum end-to-end delay 2-3 = (x_2+x_3) seconds, maximum end-to-end delay 1-2-3 = $(x_1+x_2+x_3)$ seconds,

the average of the maximum end-to-end delays of the whole network is therefore = $\frac{x_1+x_2+x_3+(x_1+x_2)+(x_2+x_3)+(x_1+x_2+x_3)}{6}$ seconds

Therefore, if y_1= maximum end-to-end delay 1, y_2= maximum end-to-end delay 2, y_3= maximum end-to-end delay 3,…,y_p= maximum end-to-end delay p respectively, and Davmaxnetdelay = average of the maximum end-to-end delays of the whole network, where p = no of end-to-end delays inherent in the network, then,

$$\text{Davmaxnetdelay} = \frac{y_1+y_2+y_3+\dots\dots+y_p}{p} \text{ seconds} \qquad (2)$$

Eq. (2) is an enhancement (with respect to a switched LAN) of the network delay performance measures defined in (Elbaum and Sidi, 1996), where it was stated that the minimum average network delay is the average delay between all pairs of users in the network. What we think is meant here is that, the minimum average network delay is the average of the minimum end-to-end delay between all pairs of users in the network. Correspondingly therefore, following from this statement, the maximum average network delay is the average of the maximum end-to-end delay between all pairs of users in the network. But we have previously shown in (Eyinagho et al., 2011) that, for a switched local area network, it does not seems to be correct when all the users (hosts) that are attached to the network are enumerated for the purpose of end-to-end (for example, average end-to-end) delay computation.

We are thus able to know (using Eq. (2)) whether this average of the maximum end-to-end delay of the whole network is less than the maximum end-to-end delay that can be tolerated by any application that is to be deployed on this network. If it is less than, then we have designed an efficient network; if it is more than, then we can iteratively choose switches of higher capacities (high switching fabric data transfer rates) and hence, low switching latency (delay).

This method of specifying the switches in a switched LAN finds agreement with the switched network design problem as formulated by Reiser (1982), which, stated in words, is 'design the switched network under maximum loading condition, such that the average network delay is less than or equal to a given delay bound'.

Another way of looking at this switched LAN design methodology is to state the network design problem in this way: can we specify the switches in a switched LAN such that the maximum of the maximum (max-max) end-to-end delays of the LAN is less than the maximum tolerable delay of the applications to be deployed in the network? For example, Georges et al. (2005) asserted with respect to the deployment of switched Ethernet networks in an industrial environment that, the network calculus is used to determine the upper-bounded end-to-end delays of each packet; if all the end-to-end delays are less than the time-cycle of the programmable controller, then the network organization (arrangement of switches) is in accordance with the application constraint.

3. CONCLUSION

Reiser (1982) has once averred that, the packet switched network design specification may be to obtain load values such that the mean network delay remains below a given bound. Relating this to switched local area networks, the design objective of a network will be to design the network such that the average maximum end-to-end delay is below an upper bound. This upper bound can be determined by the maximum delay constraints (requirements) of the applications to be deployed in the network. This paper has described a methodology for enumerating all the end-to-end delays of any switched local area network, and therefore, of calculating the average (or average of the maximum end-to-end) delays of such a network.

REFERENCES

ELBAUM, R.AND SIDI, M. 1996. Topological Design of Local Area Networks using Genetic Algorithms. IEEE/ACM Transactions on Networking, 4(5), 766-778.

EYINAGHO, M. O., FALAKI, S. O. AND, ATAYERO, A. A. 2011. Revisiting the notion of Origin-Destination Traffic Matrix of the hosts that are attached to a Switched Local Area Network. International Journal of Distributed and Parallel Systems (IJDPS), 2(6), 227-235.

GEORGES, J., DIVOUX, T. AND RONDEAU, E. 2005. Confronting the Performances of a Switched Ethernet Network with Industrial Constraints by using Network Calculus. International Journal of Communications Systems, 18, 877-903.

GERD, K. 1989. Local Area Networks. McGraw-Hill Book Co., New-York.

KANEM, E., TORAB, P. COOPER, K. AND CUSTODI, G. 1999. Design and Analysis of Packet Switched Networks for Control Systems. Proceedings, IEEE Conference on Decision and Control. Phoenix, AZ, 4460-4465.

REISER, M. 1982. Performance Evaluation of Data Communications Systems. Proceedings of the 1982 IEEE Conference, 70(2), 171-194.

TORAB, P. AND KANEM, E. 1999. Load Analysis of Packet Switched Networks in Control Systems. Proceedings 25th Annual Conference of the IEEE Industrial Electronics Society, San Jose, CA, 3, 1222 – 1227.