A PERFORMANCE EVALUATION OF COMPUTER NETWORK USING MULTI-LAYERED PROTOCOL

Tadao SAITO*, Hitoshi AIDA*, Atsushi SHIRAHATA*, Masahiro HAMAOGI*, Wataru TAKAHASHI*, and Hisashi ISHIHARA**

* Department of Electrical Engineering, University of Tokyo, Tokyo Japan
** CSK, Tokyo Japan

OSI reference model has seven layers of which six layers are logical layers having their own header. The computers connected by the network or the intermediate switching nodes must process the header to handle the data unit to be routed through the network. The standardization and layering of the protocol improves the connectability of the network in logical sense, however, to maintain the efficiency of the network, detailed understandings of the performance of the network and protocol parameters are needed. This paper treats the lower three logical layers of the OSI model, i.e., the transport, network and data link layers from the viewpoint of performance evaluation.

1. INTRODUCTION

Protocol parameters in the layered model are defined layer by layer. The transmission speed, and propagation delay are the basic transmission media parameter. The data unit size and the window size of each layers are the basic protocol parameters. The effects of these parameters on network performance in the data link layer have been studied extensively(1,2), however, very few are known on the performance of network having multi-layered protocol(3,4).
Fig. 1 shows an example of sequence of data units in three lower logical layers in a network in which two communication nodes are connected through a data link. If a transport command $T(C)$ is transmitted from the DTE A to the DTE B, a network command $N(C)$ is transmitted in the network layer. If one response is to be used for each command in the layer, a network response $N(R)$ will be needed. To transmit network layer command, again, a command response pair of data link layer is needed. Therefore, to transmit a transport command and to receive the response back, 8 data unit of data link layer will be transmitted in the data link.

At the transmitting and the receiving computer, some processing time will be needed for send and receive data units of each layer and some delay is incorporated with the transmission media. As an example, Fig. 2 is a sequence diagram of communication processing in a case that one computer is used for data link, network and transport layer processing. Considering the fact that the transmitter which transmitted information data units confirm correct data transmission by receiving supervisory data unit (ack), if the reception of supervisory data unit is delayed, the transmission efficiency of the network will be degraded. In the chart shown in Fig. 2, some processing time is needed for data transmission of each layer. This is data send processing delay. Then after the reception of data some data reception processing delay is needed. Then acknowledge send processing delay is caused in each layer. After the reception of acknowledgement receive processing delay and post processing delay are needed.

The relation between the network parameters and the network performance also very much effected by the network configuration. If a number of data links are connected serially, the effect of transmission delay on the performance is further serious.

Generally, performance of network can be measured in terms of throughput and response time. Throughput may be measured assuming infinite size of data at one end to be continuously supplied to the network. The response time is defined as the time duration between transport data unit transmission and the corresponding response reception in this paper.
2. Network Configuration

Computer networks can be configured in various way depending on the application requirement. In the simples form, a network just includes two computers inter-connected through a data link. In more complicated form, a number of packet switching nodes are connected by data links to form a packet switching network, through which two ends nodes are connected.

In Fig.3, physical configuration of various computer networks are shown schematically in the left hand side column. End nodes, i.e., DTEs are indicated by a square block and intermediate nodes having buffering capability are indicated by a small circle. If the intermediate nodes are packet switches, the packet switching network is indicated by circles noted by PSDN.

The indications in the right hand column of Fig.3 are the corresponding diagram to describe the layered protocol structure used for the network. If two DTEs are directly connected by a data link (a), just one layer protocol i.e. data link layer protocol may be used (a-1). Or end to end transport layer protocol can be used (a-2). Three layer protocols i.e. data link layer, network layer and transport layer protocols may be used (a-3).

Similarly, if two DTEs are interconnected through a intermediate node as is shown by (b), transport layer end to end protocol is used in addition to link by link layer protocol. This is shown by graph (b-1) in the right hand side column. In the case (b), network layer protocol may be used as is shown by (b-2). The case (b-2) may also be applied for the case where packet switched network is used as the case (c). In the case (c) if network layer protocol of local significance is used, the corresponding logical diagram may be shown as c-2. The case two or more intermediate nodes are used, more complex networks and the corresponding logical diagrams are obtained. The networks shown in Fig.3 (d) and (e) and the corresponding logical diagrams may be understood similarly.

The time chart shown in Figs.1 and 2 corresponds to the network having the logical structure of Fig.3 (a-3). For more complex networks, much more number of lower layer data units must be exchanged to a pair of transport layer command and responses. Therefore, general tendency that throughput is degraded for more complex networks can be observed.
3. Evaluation Model

The timing diagram as is shown in Fig.2 can be obtained for various types of networks. Network throughput can be evaluated by measuring the number of bits in transport receive data units in a fixed time duration and the efficiency is evaluated by dividing the throughput by the transmission speed to the line.

In the evaluation model used in this paper, only a pair of DTEs are connected through the network. An information source of infinite capacity is provided in one DTE and the information is transmitted to the peer DTE continuously. Throughput is evaluated by the number of transmitted bits and the response time is measured as the time duration between the time when a transport service data unit is given and the valid response to the data unit is returned.

Some processing time is needed for protocol processing of each node. The processing time is assumed to be

- data send : 2 msec
- data receive : 3 msec
- ack send : 1 msec
- ack receive : 2 msec
- ack post processing : 1 msec

The transmission delay is assumed to be 5 msec which roughly corresponds to the surface distance of 1000 km assuming the speed of 200 m/μsec. If satellite transmission system is used, the transmission delay will be about 270 msec.

4. Throughput and response time in a point to point network

Fig.4 shows the graph indicating the transmission efficiency versus transmission speed in the case of network shown in Fig.3 (c). The size to transport service data unit is assumed to be 128 bytes. In this figure, the line indicated (L) shows the efficiency of the network having only one layer and the line indicated by (L+N) shows the efficiency of network having 2 layered protocol and the line indicated by (L+N+T) shows that in the network having 3 layered protocol. In each of the protocol, the efficiency will be lowered as the transmission speed increases, because the comparative time needed for processing at the end node and transmission delay increases compared with pure transmission time.

The efficiency of the network will be improved if an supervisory acknowledgment data unit is used for a number of information data units. The transmission efficiency degradation is partly caused by processing delay for protocol

---

Fig. 4 Transmission efficiency versus transmission rate in multi-layered protocol (all window size = 1)

Fig. 5 Response time versus transmission rate in multi-layered protocol (all window size = 1)
processing in each layer and partly by transmission delay. For example in the case of the transmission speed of 2400 b/s, the efficiency of one layer system is 88.4%, and 77.8% and 59.9% for two and three layered system respectively. If protocol processing machines are provided for each layer, protocol processing delay is improved, however, even if protocol processing delay is reduced to zero the efficiency of 59.9% can be improved just 2.1%.

The response time values for the cases of Fig.3 (a) are shown by the graph Fig. 5. Response time is of course improved by increasing transmission speed.

If a number of data units which can be continuously sent without receiving acknowledgment is increased, in other words, if window size is increased, the transmission efficiency can be increased. Fig.6 shows the transmission speed versus transmission efficiency of 3 layered system taking window size as parameter. In this case if window size of 3 is assigned to all the 3 layers, the transmission efficiency can be improved to 80.7% from 59.5% for 2400 b/s system. If window size is 7, the efficiency is 87.7%.

Window size can be increased for a part of network layers. For the speed of 2400 b/s, if window size 3 is applied to data like layer only, efficiency is 66.3%. If window size 3 is applied to transport layer only the efficiency is 70.4%. This indicate that if layer window size is used only for one layer, it should be used in higher layer.

![Graph showing transmission efficiency vs transmission rate](image1)

Similarly, Fig.7 shows the response time versus transmission speed corresponding to the cases of Fig.6. In this evaluation, it is assumed that response is returned after the window is expired. If window size is increased the sender must wait until the response is returned for a longer time. Therefore, if window size is increased, the response time tends to get worse although the efficiency is improved.

5. Evaluation for more complexed networks

Similar evaluation is possible for more complex networks as are shown in Fig.3. Table 1 shows the efficiency in %, throughput in bit/sec and response time in millisecond, for various parameters for networks shown in Fig.3 (a),(b),(c), (d),(e).

Even in cases where the transmission speed is as low as 2400 b/s, if 2 intermediate nodes are interposed, the efficiency is degraded to about 24% in the case of three layered protocol. The value is not able to be improved even when window size of one of layers is increased. If the window size of all the layers are increased to 7, the efficiency is high as 66.3% is achieved.
It should be noted that the value of efficiency is not very much improved by reducing the processing speed of protocol processing. The improvement of efficiency by protocol processing capability is at most less than 2%.

It should also be noted that response time is get worse as the efficiency is improved by enlarged window size. Especially if the window size of one layer only is increased, response time is got worse and efficiency is not improved.

6. Formulae for throughput evaluation

The value of throughput described above is obtained by a very complex procedure. However, it is possible to obtain a rough experimental formulæ to evaluate efficiency and for some of networks shown in Fig.3.

Table 2 shows the formulæ to evaluate efficiency  for protocols having two

<table>
<thead>
<tr>
<th>Table 2. Experimental formulæ for efficiency evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: efficiency of two layered networks with n intermediate nodes</td>
</tr>
<tr>
<td>for 2400~9600b/s</td>
</tr>
<tr>
<td>( (D: d_1 - a_1) ) &amp; ( \eta = \frac{w}{(w+n)}l/v + (n+2)d_1 + a_1 + a_2 + a_3 + d_2 + T + D + T' - a_3 + D - 2d_1 )</td>
</tr>
<tr>
<td>for 48kb/s</td>
</tr>
<tr>
<td>( (D: d_1 - a_1) ) &amp; ( \eta = \frac{w}{(w+n)}l/v + (n+2)(d_1 + d_2) + (n+1)(a_1 + a_2 - 2d_1 + T) + a_3 )</td>
</tr>
<tr>
<td>for 1Mb/s</td>
</tr>
<tr>
<td>( (l'/v \leq d_1 + d_2) ) &amp; ( \eta = \frac{w}{(w+n)}l/v + \frac{(n+1)(l'/v + a_1 + a_2 + 2d_1 + T) + (w+n+1)(d_1 + d_2) + a_3}{l'/v} )</td>
</tr>
</tbody>
</table>

B: efficiency of three layered networks with n intermediate nodes |
| for 2400~9600b/s |
| \( (D: d_1 - a_1) \) & \( \eta = \frac{w}{(w+n)}l/v + (n+3)d_1 + (n+2)a_1 + a_2 + a_3 + d_1 + 2d_2 + D + T + a_5 + D + N + T + a_5 + D + a_5 (2d_1 + a_5 + D + a_5 + D + N) \) |
| for 48kb/s |
| \( (D: d_1 - a_1) \) & \( \eta = \frac{w}{(w+n)}l/v + \frac{(n+3)(d_1 + d_2) + (n+2)(a_1 + a_2 + a_3) + (n+1)(2d_1 + T) + a_5}{l'/v} \) |
| for 1Mb/s |
| \( (l'/v \leq d_1 + d_2) \) & \( \eta = \frac{w}{(w+n)}l/v + \frac{(n+1)(l'/v + (n+2)(d_1 + d_2) + (n+2)(a_1 + a_2) + (n+1)(2d_1 + T)) + a_5 (2d_1 + a_5 + 2d_2 + D + N)}{l'/v} \) |
and three layers respectively of networks having n intermediate nodes. In this evaluation, window size of all the layers are assumed to be equal and denoted by $w$. Processing time for data transmission is $d_1$ and data reception is $d_2$. Processing time for supervisory data unit send is $a_1$, supervisory data unit reception is $a_2$ and supervisory data unit post processing delay is $a_3$.

Transmission times for supervisory data units in data like layer, network layer and transport layer are denoted by $D$, $N$ and $T$ respectively. The length of transport service data unit and data link protocol data unit are denoted by $l$ and $l'$ respectively. Transmission delay between two nodes is $d$ and transmission speed is $V$.

Efficiency can be evaluated by different expression depending on the transmission speed range and delay time range. The evaluation error by these formulae compared with simulation is less than a few percent.

7. Conclusion

In this paper, transmission efficiency and response time for data networks employing layered protocol architecture is evaluated for various cases considering three protocol layers. The value of efficiency are evaluated by simulation, and experimental formulae for evaluation are obtained. It is revealed that, if transmission speed is high, the efficiency some times become very low.

The networks analyzed in this paper is simple and only one transport connection is assumed. Extension of the analysis for more complex cases including multiplexed transport connection is the subject of further study.

References