Optimal Transmission of Data Over Multi-Hop Networks*

Ibrahim Matta, Daniel Ryan
College of Computer Science
Northeastern University
Boston, MA 02115
{matta, dtr}@ccs.neu.edu

Abstract

We study the end-to-end delay performance when transmitting data traffic over multi-hop networks. Data is normally carried in transmission units (e.g., packet, frame, cell). The performance is strongly affected by several factors, including header and payload sizes of the transmission units, and the effect of pipelining these units over the multiple links. We determine the optimal payload size as a function of the data size, the header size and the number of links. We show the performance penalty incurred because of some transmission standards.

1 Introduction

When sending a data stream of bits over a network, the total delay required for the entire data to reach its destination is strongly affected by the network characteristics, available resources, and the methods and standards used to process the data. In this paper, we study the effect of several factors on the total delay. We refer the reader to [1] for more details.

2 Minimizing End-to-End Delay

We consider a store-and-forward tandem network connecting the source and destination. We assume all links have the same propagation delay and transmission rate, and all transmission units have the same size. Thus the transmission time of all units on any link is the same, and queueing occurs only at the first link. This model may represent a connection-oriented network (e.g., ATM, frame relay) with a virtual path/circuit established between the source and destination nodes. We define the following: S, size of the data stream to be transmitted;R, the transmission rate of a link;L, number of links; P, link propagation delay; N, number of transmission units; k, the number of data bits in the transmission unit – i.e. payload size; and h, the number of header bits. Then the total end-to-end delay D is given by:

$$D = (N-1) \times \frac{(k+h)}{R} + \left(\frac{(k+h)}{R} + P \right) \times L$$

As the payload size k increases, the header overhead decreases, but because of reduced pipelining, the delay D increases. As k decreases, there is more pipelining. However, as k gets very small, the number of headers increases enough to become a significant overhead causing D to increase. The optimal value of k, denoted by k^*, that minimizes D is given by:

$$k^* = \sqrt{\frac{h \times S}{(L-1)}} \qquad (1)$$

k^* balances the effects of the two opposing factors, header overhead and pipelining, to minimize D.

Figure 1 shows the relative improvement (reduction) in total delay when an optimal size ATM cell, obtained using equation (1), is used instead of the standard 53-byte ATM cell. It shows that a variable size transmission unit would outperform a fixed size unit especially when the bandwidth allocated to the virtual data path is low. We also show in [1] that there may be a penalty in performance by fragmenting IP packets to comply with the 9180-byte MTU standard imposed when sending them over an ATM network.

References

[1] I. Matta and D. Ryan. Optimal transmission of data over multi-hop networks. Technical Report NU-CCS-96-08, April 1996. Available at http://www.ccs.neu.edu/home/matta/.

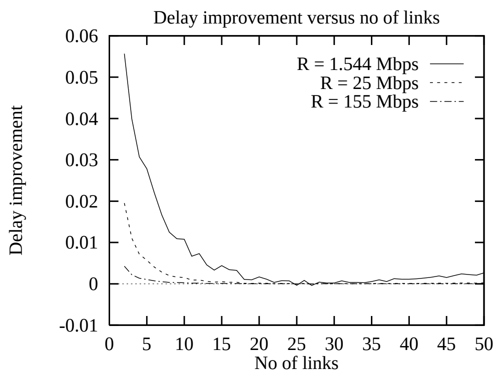

Figure 1: Delay improvement versus number of links.

*I. Matta was partially supported by NU-RSDF grant 377090. D. Ryan is also with Motorola ISG, Mansfield, MA.