This is the first part of Chapter 6 from the second edition of :

Networks, Routers and Transputers:

Function, Performance and applications

Edited by: M.D. May, P.W. Thompson, and P.H. Welch

IOS Press
Van Diemenstraat 94
1013 CN Amsterdam
Netherlands

IOS Press, Inc.
P.O. Box 10558
Burke, VA 22009–0558
U.S.A.

IOS Press/Lavis Marketing
73 Lime Walk
Headington
Oxford OX3 7AD
England

Kaigai Publications, Ltd.
21 Kanda Tsukasa–Cho 2–Chome
Chiyoda–Ka
Tokyo 101
Japan

This chapter was written by C. Barnaby, V.A. Griffiths and P.W. Thompson.

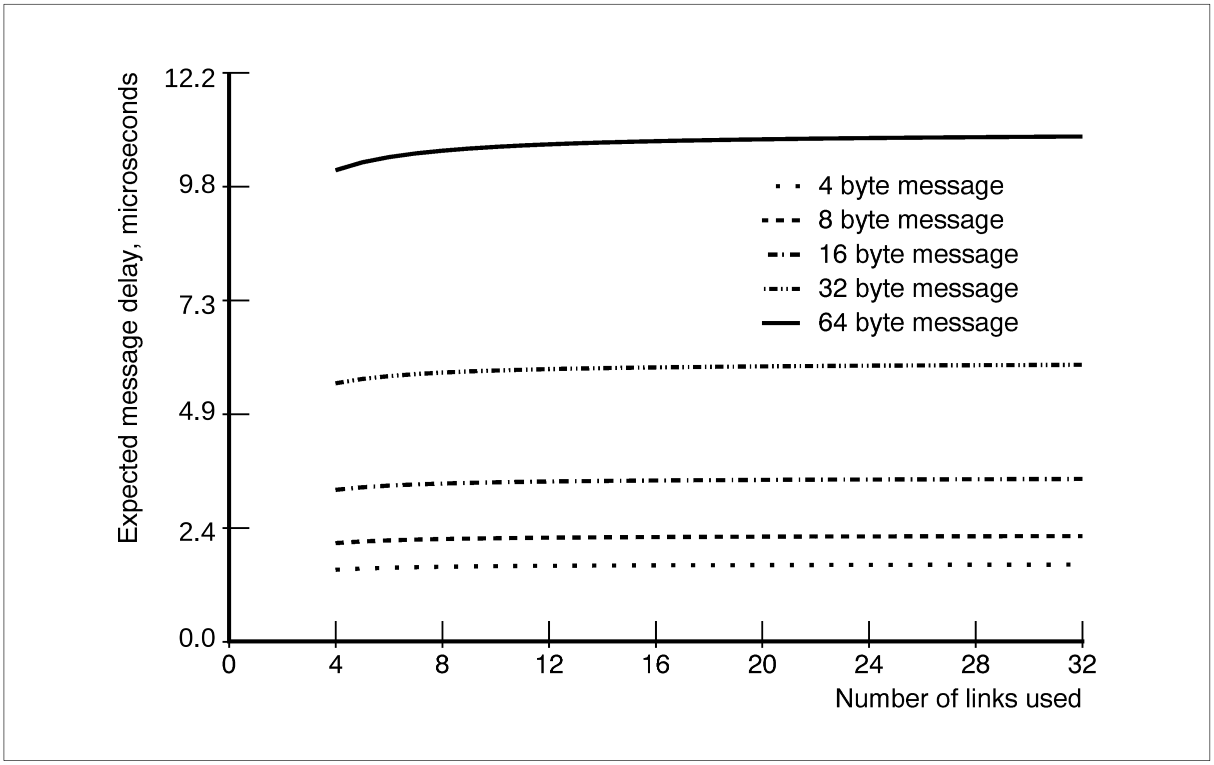

Figure 6.8 Mean packet delay vs. packet size

The model is now used to illustrate the case when $in \neq out$. The number of input links is varied, with the number of output links held constant at 32. The message size is varied. The throughput for each input link is shown in figure 6.9, and the corresponding expected packet delay is shown in figure 6.10.

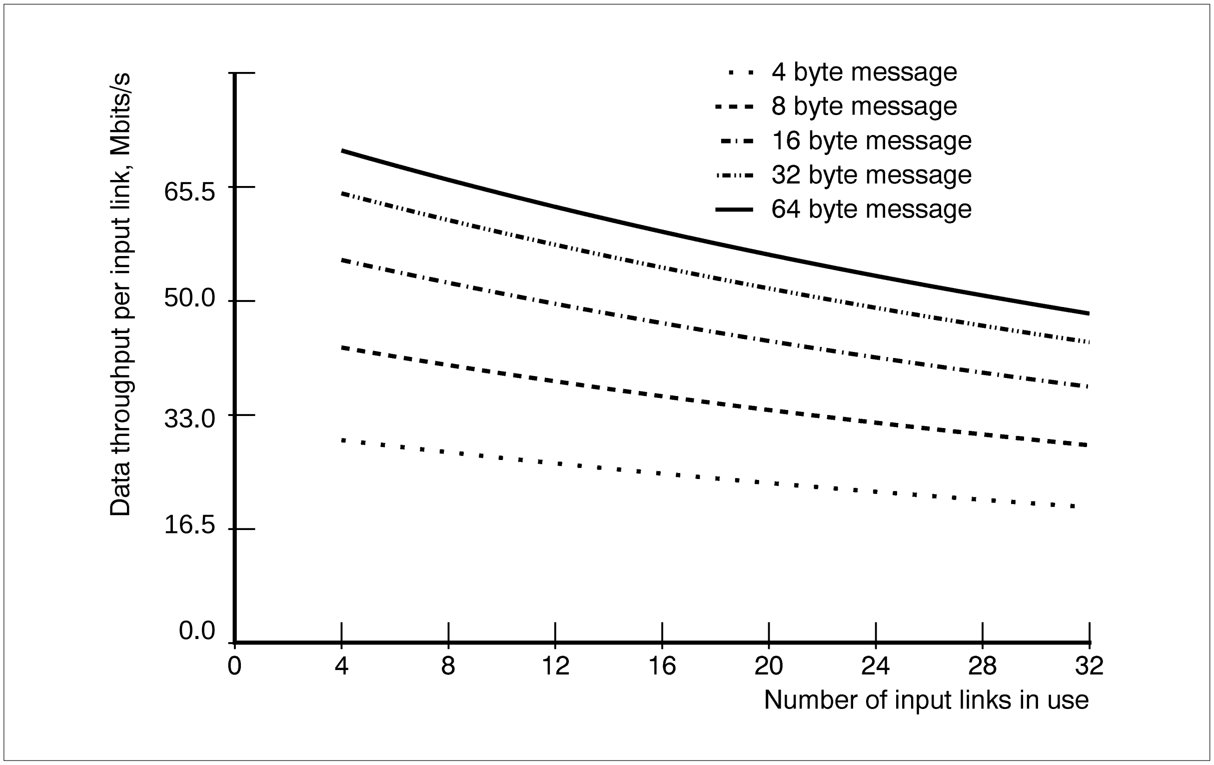

Figure 6.9 Throughput per link vs. no inputs used (32 outputs)

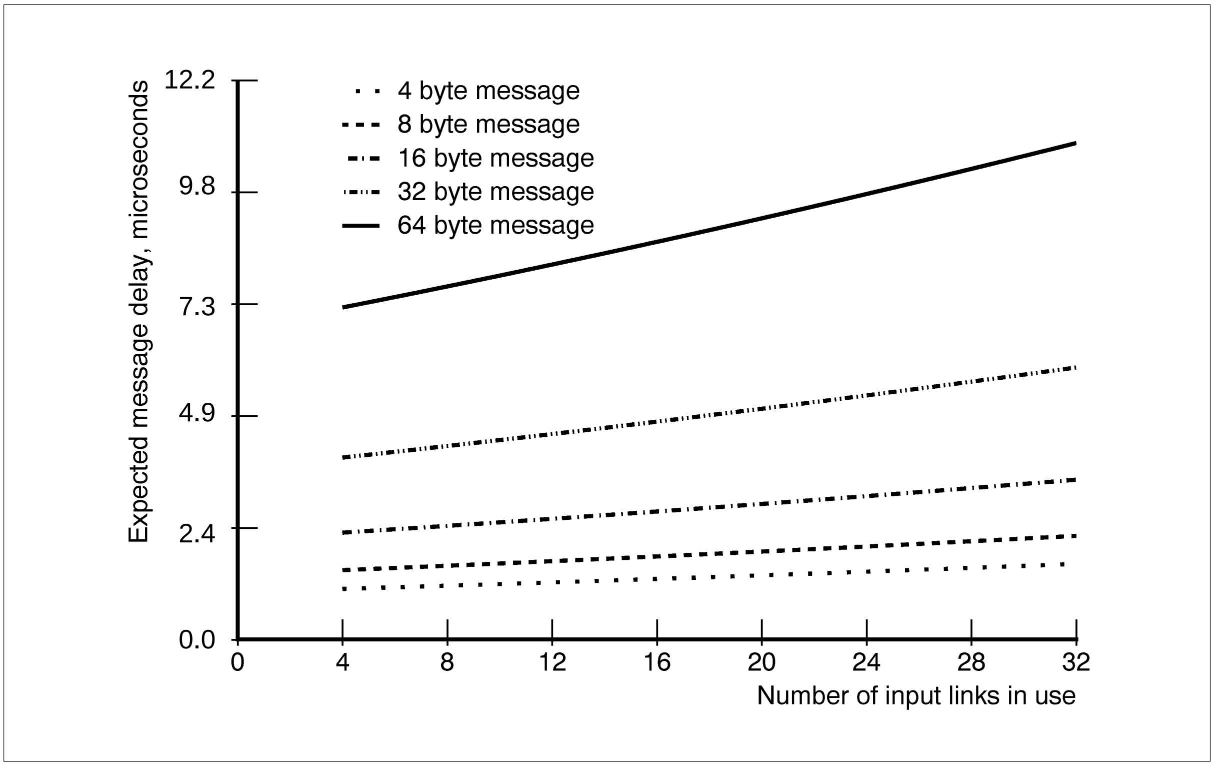

Figure 6.10 Mean delay vs. no inputs used (32 outputs)

The asymptotic results for the case $in \neq out$ describes the expected behaviour. The number of output links is held constant, first at out = 32, then at out=8. The number of input links is varied, for 32 byte messages. The throughput and delay are compared to the asymptotic curves in figures 6.11 and 6.12 with 32 output links in use. Figures 6.13 and 6.14 show 8 output links in use.

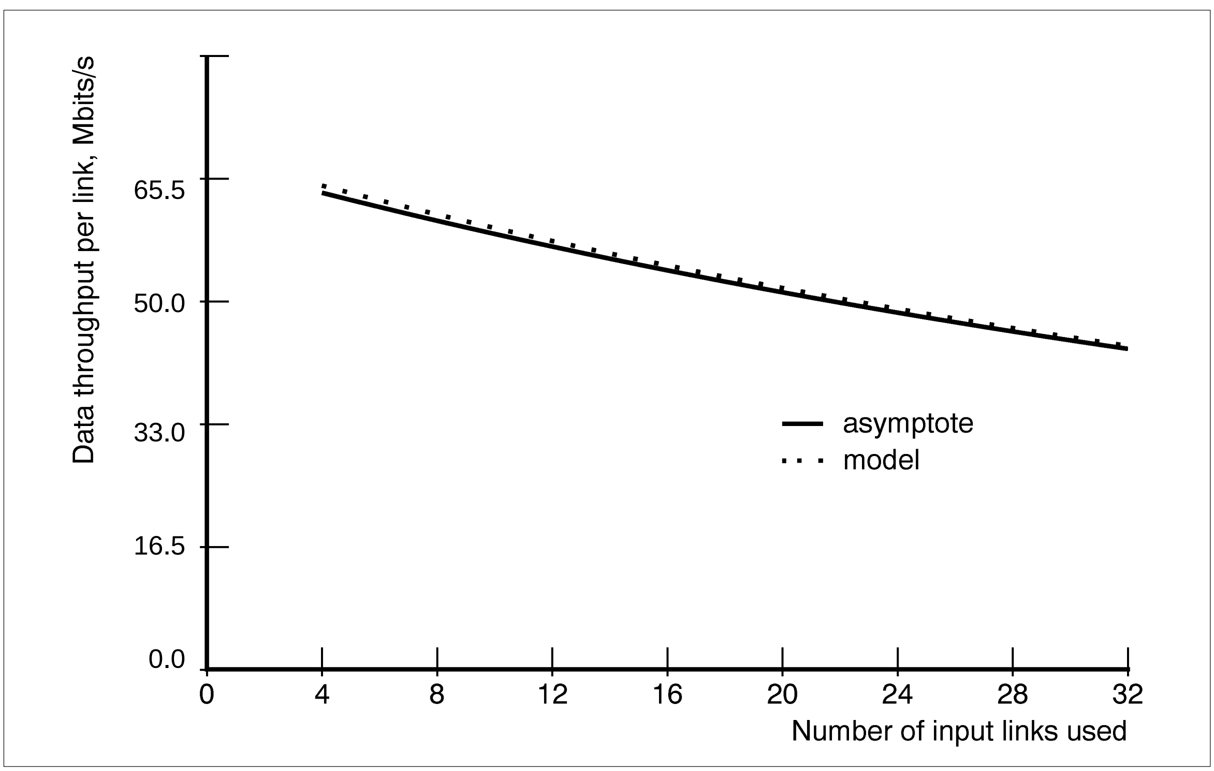

Figure 6.11 Throughput per link vs. no inputs used (32 outputs)

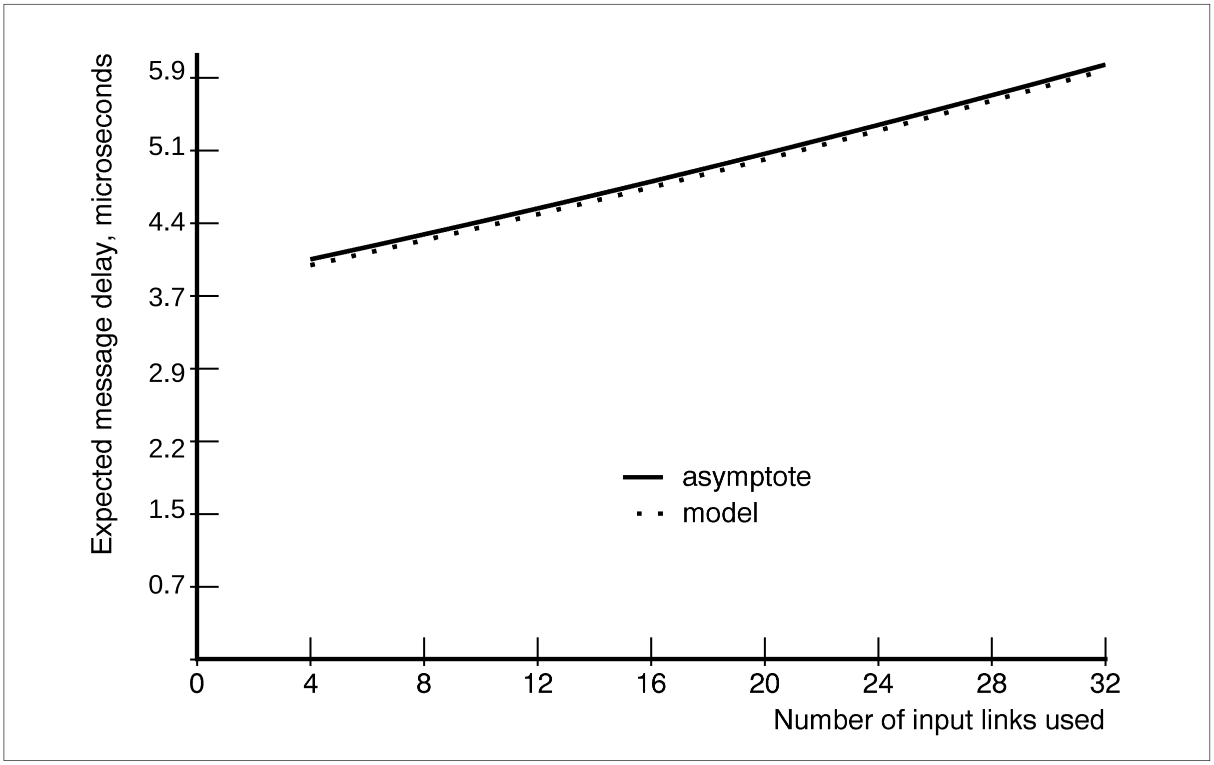

Figure 6.12 Mean delay vs. no inputs used (32 outputs)

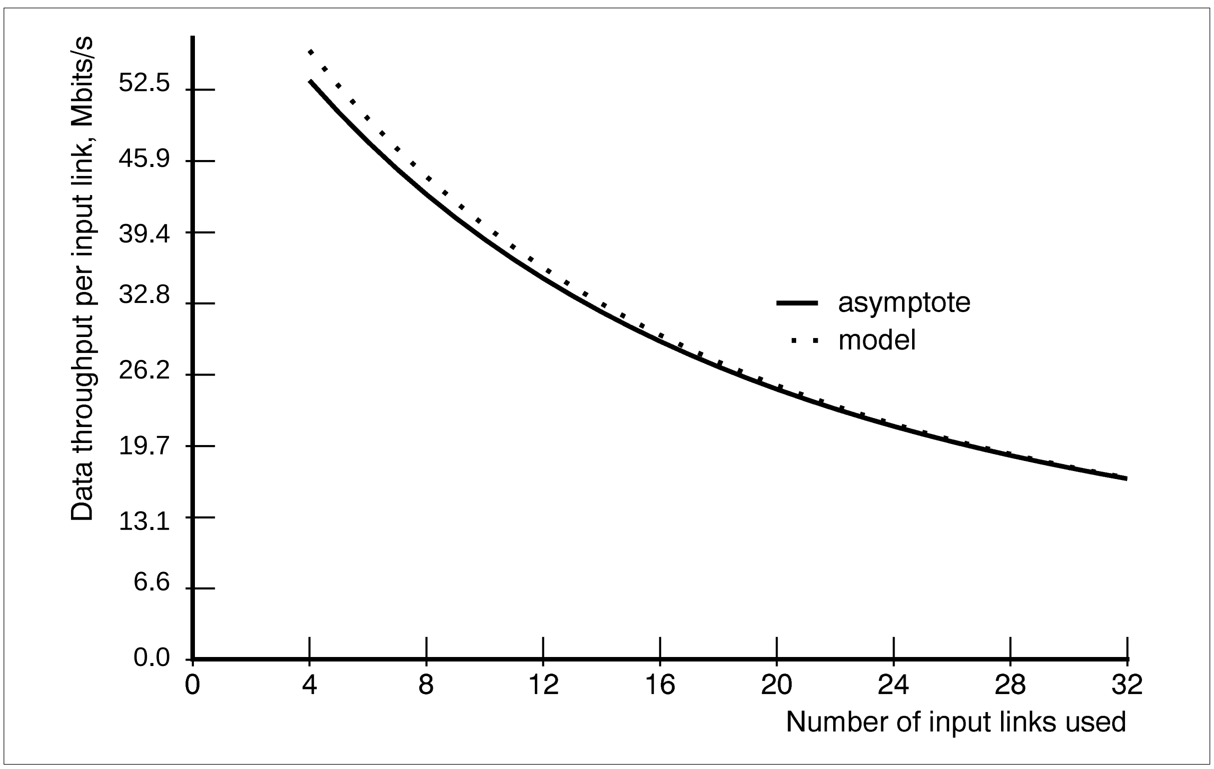

Figure 6.13 Throughput per link vs. no inputs used (8 outputs)

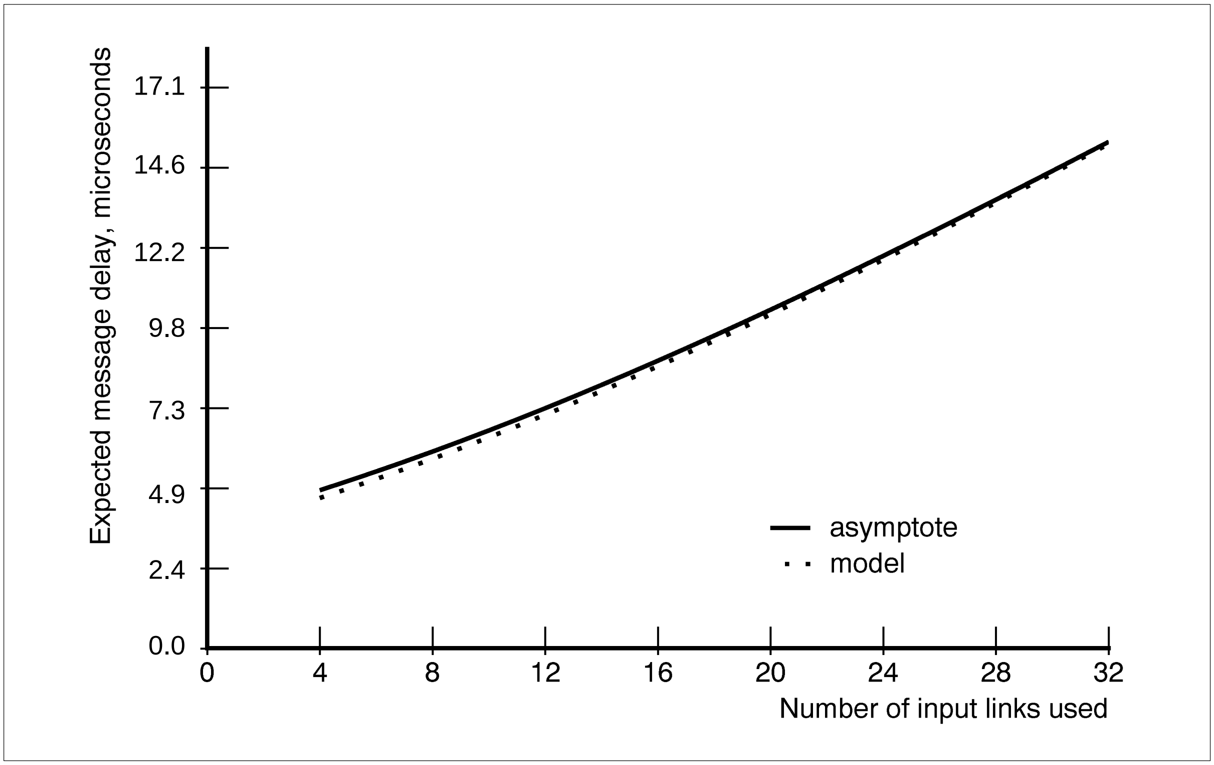

Figure 6.14 Mean delay vs. no inputs used (8 outputs)

These graphs show that the limits of the expressions are actually a very good approximation to the exact model as long as there are more than a few links in use for both input and output. The factor common to the expression for link throughput, and delay, is $r\left(1 - e^{-1/r}\right)$. This is plotted in figure 6.15.

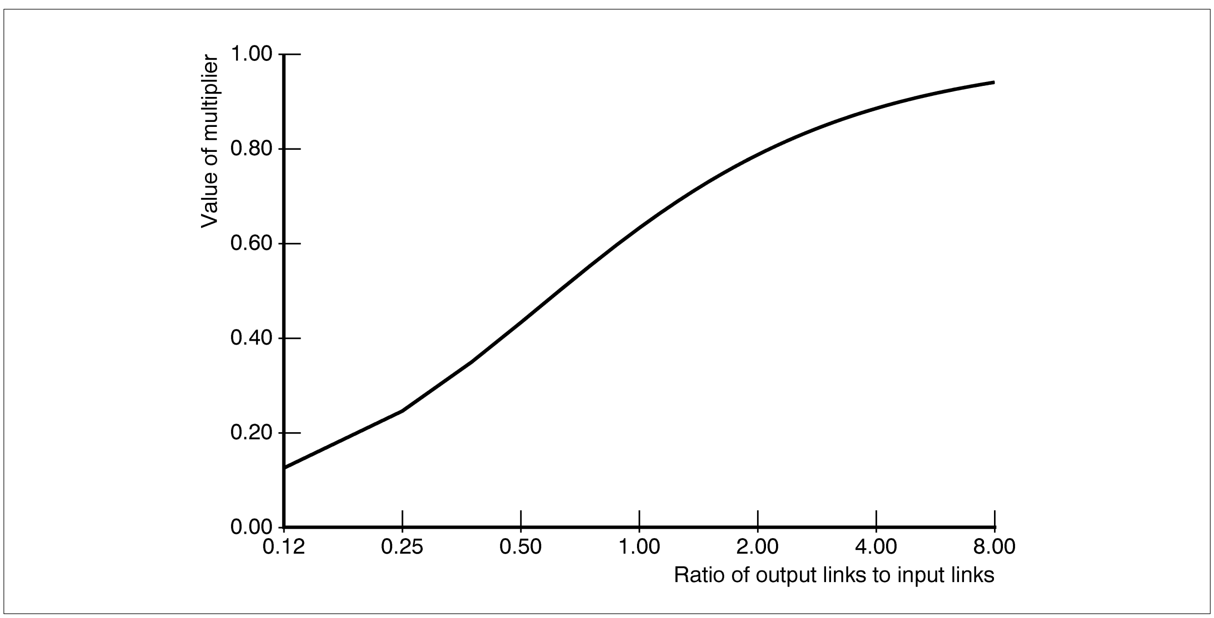

Figure 6.15 Variation of $r\left(1 - e^{-1/r}\right)$ with r

The approximation is dependent upon the ratio of output links to input links, and not the absolute number of input links and output links in use. This suggests that the throughput for each of 8 input links, choosing among 16 output links, will be about the same as the throughput of each of 16 input links, choosing between 32 output links.

In the expression for throughput, the value k/S describes the amount of data in time S: the data throughput rate. The factor $r\left(1 - e^{-1/r}\right)$ describes the proportion of output links which are used.

The expression for delay depends on S, with the factor $\frac{1}{r\left(1 - e^{-1/r}\right)}$ determining the number of slot times which the packet takes to get across the switch.

6.3.7 Maximum Routing Delay

Finally we consider the effect of the very worst case contention on the transit time of a packet through a C104. This means the time between the header arriving on an input link of the C104 to the time that the header is transmitted from the chosen output link.

In the worst case, 32 inputs contend for the same output[20], so the unlucky packet must wait for 31 others to be routed before it can proceed. The very worst case is when all 32 packets arrive simultaneously; in all other cases some of the routing of the first packet will have been done by the time the unlucky packet arrives.

Although every packet header must be received and the corresponding routing decisions taken, this occurs concurrently for all 32 packets. So the unlucky packet is delayed only by the time taken to receive its own header and perform the corresponding routing decision. The worst case is with two-byte headers, which take $2 \times$ 100ns + *(link input latency)* to receive. Making the routing decision and performing the arbitration takes about 60ns; the first packet can then start to be transferred across the crossbar. Each successive packet will start to be transferred immediately after the previous one finishes; the whole process will be limited by the speed of the output link[21].

Thus a total packet transfer time of $31 \times L_{packet} \times$ 100ns is required before the unlucky packet gets across the crossbar; it then has to reach the outside world through a FIFO and the link output circuitry. The delay through the FIFO is minimal, but the link output latency should be considered.

Thus the total is: $(L_{header} + 31 \times L_{packet}) \times$ 100ns + $4 \times$ 20ns + (*total link latency*). In a T9000 system packets (in the worst case) are 32 bytes plus a terminator plus a routing header of length L_{header} plus (usually) a virtual channel header (typically another two bytes), so L_{packet} is typically at most 37. The link latency is small compared to the other terms, so this gives a total of about 115µs.

Note that this analysis assumes that the congested output link transfers data at full speed the whole time. If this is not the case (for example if it is connected to another C104, where there is contention for an output link...) then the time must be increased. Note however that the effect of this multiplication is minimized by using large fan-out routers such as the C104.

6.4 Summary

A variety of models have been developed to describe data throughput on the DS-Link and through the C104 router. The first takes into account the overhead of acknowledge packets and flow control, use of one or two byte headers, and unidirectional or bidirectional link use. This model has been used to give the asymptotic throughput of the DS-Link, as the message size gets very large. For large messages and a maximum packet size of 32 bytes, the lowest throughput value of the link is 8.27 Mbytes per second. This occurs when two-byte headers are used along with bidirec-

20. Note that this includes the unusual - but not impossible - case that one packet is being routed directly back out of the link on which it arrived.

21. All the links are assumed to run at the same speed.

tional link use. Further models consider the effect of latency on bandwidth given the particular protocols used, both at the token and the message levels of the protocol. The final models show the effect of output contention in a single C104, both in an average and a worst case.