Purdue University
Purdue e-Pubs

Department of Computer Science Technical Reports | Department of Computer Science

1982

Estimating the Performance of Multi- Microcomputer Networks

Daniel A. Reed

Report Number:
81-420

Reed, Daniel A., "Estimating the Performance of Multi- Microcomputer Networks" (1982). *Department of Computer Science Technical Reports.* Paper 344.
https://docs.lib.purdue.edu/cstech/344

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries.
Please contact epubs@purdue.edu for additional information.

Estimating the Performance of Multi-microcomputer Networks

Daniel A. Reed

Department of Computer Sciences
Purdue University
West Lafayette, IN 47907

CSD-TR-420

ABSTRACT

A new paradigm for parallel computation based on large numbers of interconnected microcomputer nodes has recently emerged. The absence of any memory shared among the nodes makes the selection of an interconnection network capable of efficiently supporting message passing critical. A network performance evaluation technique based on the asymptotic analysis of a closed queueing network is used to compare six proposed interconnection networks: the simple ring, the spanning bus hypercube, the dual bus hypercube, the toroid, the cube-connected cycles, and the R-ary M-cube. Among the questions considered for each network are the theoretical upper bounds on performance, the minimum sized computation quanta efficiently supported, and the number of nodes over which each network is preferred.

November 10, 1982

Introduction

Recent, widely known advances in VLSI technology have led several groups to propose a new approach to parallel processing based on large networks of interconnected microcomputers [DESP78, FINK81, WITT81]. The one or two VLSI chips comprising each node of a network would contain a processor with some locally addressable memory, a communication controller capable of routing messages without delaying the processor, and a small number of connections to other nodes.

Suggested application areas for such networks have included parallel partial differential equations solvers [SMIT82], divide and conquer algorithms [ELDE79], and mini-max game tree searches [AKL80]. The cooperating tasks of a parallel algorithm for solving one of these problems would execute asynchronously on different nodes and communicate via message passing. Because the nodes do not share any memory, the selection of an interconnection network capable of efficiently supporting the message passing patterns of these algorithms is crucial. The selection process is complicated by the very limited interconnection fanout afforded by the VLSI implementation of the nodes and the lack of an accepted measure of network performance.

Several different characterizations of network performance have been suggested (e.g., average message delay and communication link load), but the mean internode distance has been by far the most common performance metric. Unfortunately, little attention has been given to measures that include the effects produced by the *interaction* of the nodes and the network that performs the message routing. Because the nodes are both the source and sink for messages, the rate at which the network routes messages and the rate at which the nodes perform computations are intimately related.

In this paper we examine the interdependency of the nodes and several well known interconnection networks using the asymptotic properties of queueing networks. This holistic technique provides bounds on the rate at which the system comprising both the nodes and the interconnection network is capable of processing messages. In addition, the impact of an interconnection network on the minimum size computation quanta nodes can efficiently process is investigated. Finally, we compare these results with the ranking of networks indicated by the mean message path length.

Definitions and Assumptions

Consider a message generated by some source node. This message must cross some number of communication links and pass through the communication controllers of intermediate nodes before reaching its final destination node. At the destination some computation takes place in response to the message. Each link crossing and the computation resulting from the message constitute a *visit* to that link or node. If all source-destination pairs and the probability with which they exchange messages are considered, the number of visits to each communication link and node made by an average message can be calculated.

In the absence of specific information concerning the probability with which nodes exchange messages, we assume that all nodes exchange messages with equal probability, hereafter referred to as uniform message routing. Because we are primarily interested in the traffic on the interconnection network, we exclude the degenerate case in which nodes send messages to themselves. Furthermore, we assume messages follow the shortest path from their source to destination and that if multiple shortest paths exist, they are selected with equal probability.

Given these assumptions about the message routing distribution, let i be an arbitrary system device (either a node or a link). An average message will visit device i a certain number of times. This mean number of visits is called the *visit ratio* of device i and is denoted by V_i. If S_i denotes the mean time for device i to service a message, exclusive of queueing delays, then the product $V_i S_i$ represents the total amount of service required by an average message at device i.

If the average number of messages circulating in the network is steadily increased, the utilization of at least one device must approach unity. Which device will first limit the network performance? Since $V_i S_i$ represents the average amount of service required at device i, the device with the maximum value of $V_i S_i$ will be the performance limiting factor. If X_0 denotes the rate at which a network can service messages, an absolute upper bound on X_0 is given by

$$X_0 < \frac{1}{V_b S_b} \tag{1}$$

where

$$V_b S_b = \max_i V_i S_i \ .$$

This technique, called asymptotic bound or bottleneck analysis in its most general form [DENN78], applies to any closed queueing network in a steady state (i.e., the arrival rate at each device equals the departure rate). No additional assumptions about service time or queueing distributions are necessary. This simplicity allows us to make minimal assumptions about network behavior and, consequently, leads to conclusions applicable to a wide range of intended network environments.

To simplify analysis, we further assume that all nodes require the same mean time S_{PE} to perform a computation, and all links require time S_{CL} on the average to transmit a message. It should be emphasized that this assumption is *not* required. The succeeding discussion can be applied in its entirety, albeit involving somewhat more tedious symbol manipulation, if each device has a distinct service time.

Two artifacts of our definitions result in additional simplifications. Since V_i represents the average number of visits to device i, the mean message path length is obtained directly as

$$L = \sum_{links} V_i .$$

Secondly, the uniform message routing assumption implies that all nodes generate and receive messages with equal probability so the node visit ratios for a network with K nodes are simply

$$V_{PE} = \frac{1}{K} .$$

(We exclude visits to nodes made by messages on the way to their final destination. V_{PE} represents only node visits causing computation to take place.)

The remainder of our discussion focuses on two areas:

- a brief description of several widely known interconnection networks and an example illustrating derivation of the $V_i S_i$ and
- some interconnection network comparisons in light of the performance implications of asymptotic bound analysis.

For easy reference, the notation we have defined thus far and that used in the remainder of the paper is summarized in Table I.

Interconnection Networks

The number of proposed interconnection networks is vast, and an exhaustive study would be difficult if not impossible. Thus, we have restricted our attention to the widely known networks shown in Figure I. In spite of this restriction, the complete derivation of mean path lengths and communication link visit ratios for even this small number of networks is a rather lengthy and somewhat arduous exercise in symbol manipulation that space does not permit us to indulge; detailed derivations can be found in [REED83b]. Nevertheless, a representative network is needed to illustrate the application of asymptotic bound analysis. As an example, we have selected the D-dimensional toroid. For the remaining networks, a brief geometric description of each is provided, and a summary of the mean path lengths and communication link visit ratios is given in Table II.

Single Ring

A simple ring network consists of K nodes numbered from zero through $K - 1$ with node k connected to nodes $(k + 1) \bmod K$ and $(k - 1) \bmod K$. Messages are always routed along the shorter of the clockwise or counterclockwise paths to their destination.

Spanning Bus Hypercube (SBH)

A spanning bus hypercube [WITT81] connects each network node to D buses spanning D orthogonal dimensions; each bus is in turn connected to w nodes. This results in a network that is topologically equivalent to a D-dimensional lattice with width w in each dimension.

Expanding a spanning bus hypercube poses something of a problem because each of the w^D nodes has D connections. Increasing the number of dimensions is desirable from a performance standpoint because it reduces the mean internode distance, but it violates the constraints on node fanout imposed by integrated circuit considerations. The other alternative, expanding the width of the lattice, is limited by the number of nodes each bus can physically support. The dual bus hypercube, discussed below, was proposed to ease the expansion difficulties associated with the spanning bus hypercube while attempting to retain the advantages of buses.

Dual Bus Hypercube (DBH)

The dual bus hypercube [WITT81] is a variation of the D-dimensional spanning bus hypercube obtained by pruning $D - 2$ bus connections from each network node. In particular, one dimension, the 0-th dimension, is distinguished, and all nodes are connected to a 0-th dimension bus. In each $(D-1)$-dimensional hyperplane perpendicular to the 0-th dimension, all nodes have their second connection to buses spanning the same dimension. The second bus direction differs from hyperplane to hyperplane, but repeats if the width w of a dimension is such that $w \geq D - 1$. Note that w must be *at least* as large as $D - 1$ to insure that paths between all nodes exist.

Although the dual bus hypercube solves the expansion problems of the spanning bus hypercube, the reduction in the number of buses lengthens the average distance between nodes and results in larger visit ratios for the remaining buses. It is also important to understand that the message traffic on the two different types of buses, 0-th dimension and others, is not in general the same.

As we shall see, this difference is captured by the link visit ratios but *not* by the mean message path length.

Toroid

The toroid possesses the same topology as the spanning bus hypercube, but each bus is replaced by a ring connecting the w nodes in each dimension. This means that the number of connections to each node is increased from D to $2D$. Consequently, expanding a toroid by increasing the number of dimensions is infeasible. Fortunately, the use of dedicated links rather than buses permits expansion by increasing the lattice width.

A message is routed toward its destination by selecting a dimension in which the current node and destination node addresses differ and moving along the ring spanning that dimension in the shorter of the clockwise or counter-clockwise directions. Because the rings are bi-directional, the average distance moved in each destination is

$$L_{\text{one dimension}}^{\text{uniform}} = \frac{\sum_{k=0}^{w-1} \min\left\{k,\, w-k\right\}}{w} = \begin{cases} \dfrac{2\sum_{k=0}^{\left\lfloor \frac{w}{2} \right\rfloor} k \;-\; \dfrac{w}{2}}{w} \;=\; \dfrac{w}{4} & w \text{ even} \\[2ex] \dfrac{2\sum_{k=0}^{\left\lfloor \frac{w}{2} \right\rfloor} k}{w} \;=\; \dfrac{w^2 - 1}{4w} & w \text{ odd} \end{cases}$$

(See Table I for symbol definitions.) Since the dimensions are independent, the true mean path length is D times this distance scaled to exclude nodes routing messages to themselves:

$$L_{toroid}^{uniform} = D\left[\frac{w^D}{w^D - 1}\right] L_{one\ dimension}^{uniform}$$

$$= \begin{cases} \dfrac{Dw^{D+1}}{4(w^D - 1)} & w \text{ even} \\[2ex] \dfrac{Dw^{D-1}(w^2 - 1)}{4(w^D - 1)} & w \text{ odd} \end{cases}$$

Network symmetry and the existence of only one type of communication link allow us to immediately obtain the link visit ratios by dividing the mean message path length by the number of communication links:

$$V_{CL}^{uniform} = \frac{L_{toroid}^{uniform}}{Dw^D} = \begin{cases} \dfrac{w}{4(w^D - 1)} & w \text{ even} \\[2ex] \dfrac{w^2 - 1}{4w(w^D - 1)} & w \text{ odd} \end{cases}$$

The replacement of buses with a ring of links has two counterbalancing effects. First, the mean message path length increases because an average of $\frac{w}{2}$ link crossings are needed to achieve the same effect as a single bus crossing. However, the effect of this increase on the visit ratios is mitigated by the fact that there are w times as many links.

Cube-connected Cycles

The cube-connected cycles network [PREP81, REED82] represents an attempt to design a network whose number of dimensions can be increased without violating node fanout considerations. The network is constructed by replacing each of the 2^D nodes of a width two spanning bus hypercube with a ring of D nodes. Each of the D links incident upon each vertex of the hypercube is connected to a different node of the ring. Geometrically this means that each of the now $D2^D$ nodes is connected to the two neighboring ring nodes at its vertex and one other node in the same ring position at another vertex. Because the number of connections to each node is fixed at three, expansion by increasing the number of network dimensions becomes straightforward.

Like the dual bus hypercube, the message traffic on the cube links present in the ordinary spanning bus hypercube and the new ring links at each vertex is not the same.

R-ary M-cube

An R-ary M-cube network [BURT81] contains MR^M nodes, each of which is connected to $2R$ other nodes. The network is a generalization of the indirect binary n cube originally proposed by Pease [PEAS77] with the first and last rows of switching elements being identified, and with interchange boxes and links replaced by nodes and communication links.

Conceptually, the nodes are arranged on a horizontal cylinder in M rows, each of length R^M. Thus, each node has a row and column address of the form:

$$(i, j) \quad with \quad \begin{array}{l} 0 \le i < M \\ 0 \le j < R^M \ . \end{array}$$

A node in row i is connected to a node in row $(i + 1)$ mod M if and only if the radix R representations of their column numbers are identical except in the ith digit, with the least significant digit being considered the 0th digit.

Like the dual bus hypercube and the cube-connected cycles, an R-ary M-cube contains two kinds of links bearing different amounts of message traffic, those that are connected to nodes in the same column, *column* links, and those that are connected to nodes in different columns, *cylinder* links.

Comparing Interconnection Networks

As we showed earlier, the network device at which messages must obtain the largest amount of service determines an absolute upper bound on the rate at which the network can service messages. Under our simplifying assumptions, the bound of (1) can be reduced to

$$X_0 < \frac{1}{\max\left\{ V_{PE}S_{PE},\ V_{CL}^{1}S_{CL},\ \cdots,\ V_{CL}^{T}S_{CL} \right\}} . \qquad (2)$$

if there are T distinct communication link visit ratios. Figures II-III illustrate this bound for the six networks we have discussed assuming the mean service time at both nodes and links is unity. We have omitted the simple ring from these figures because its performance bound is significantly worse than that of the other networks.

It is instructive to compare the mean message path lengths shown in Figure IV with the bounds on the network message completion rate just given. Notice that the hypercubes have small mean path lengths relative to the other networks, but the bounds on their message completion rates are considerably less than that of the other networks. The explanation for this phenomenon is quite simple: the mean path length fails to capture any notion of the message density on links. But since each link has a possibly distinct visit ratio defined as the number of times an average message crosses it, the visit ratios *do* reflect the message densities on different link types.

In addition to these bounds on the rate at which networks can process messages, asymptotic bound analysis allows us to determine several other interesting network properties:

- the minimum feasible computation quantum of a network,
- performance bounds independent of network size,
- the effects of limited bandwidth nodes, and
- performance bounds for finite workloads.

Feasible Computation Quanta

If K is the number of network nodes, the amount of service required at a node by an average message is

$$V_{PE} S_{PE} = \frac{S_{PE}}{K} .$$

A linear increase in the message completion rate with increasing network size could only be expected if the communication link VS products were no larger than this. As the figures clearly show, the message completion rate is *not*, in general, such a linear function, implying communication delays are limiting the message completion rate. Short of changing the message routing distribution, one can only adjust the ratio $\frac{S_{PE}}{S_{CL}}$ to insure that communication is not the performance limiting factor. Inspection of (2) shows that the minimum ratio of computation to communication at which communication delays are not dominant is

$$\frac{S_{PE}}{S_{CL}} = \frac{\max\{ V_{CL}^{1}, \cdots, V_{CL}^{T} \}}{V_{PE}} \tag{3}$$

$$= \; K \cdot \max\{ V_{CL}^{1}, \cdots, V_{CL}^{L} \} \; .$$

In essence, the ratio of computation time to communication time for a message must be at least K times the maximum link visit ratio if the maximum computation rate is not to be limited by communication delays.

As an example, consider the simple ring with an odd number of nodes K. For the uniform message routing distribution, we have

$$V_{CL} S_{CL} \;=\; \frac{S_{CL} \; (K \,+\, 1)}{4K} \; .$$

Applying (3) yields

$$\frac{S_{PE}}{S_{CL}} \;=\; \frac{(K \,+\, 1)}{4} \; .$$

This means that the ratio of computation time to communication time must increase at least linearly with the ring size if communication delays are not to dominate performance.

Assuming communication delays are the performance limiting factor, one can use a variation of this technique to determine the ratio of communication times needed for two different networks of the same size to have the same bound on the message completion rate. Consider now the toroid and spanning bus hypercube, each with w^D nodes and w odd. Under the uniform message routing distribution, the communication link VS products are

$$V_{CL}^{toroid} S_{CL}^{toroid} \;=\; \frac{(w^2 \,-\, 1) S_{CL}^{toroid}}{4w(w^D \,-\, 1)}$$

and

$$V_{CL}^{SBH} S_{CL}^{SBH} = \frac{(w - 1) S_{CL}^{SBH}}{w^D - 1} .$$

Equating and rearranging terms yields

$$\frac{S_{CL}^{toroid}}{S_{CL}^{SBH}} = \frac{4w}{w + 1} .$$

Simply put, the mean message transmission time for the toroid can be approximately four times larger than that for the spanning bus hypercube and the toroid still have the same bound on the message completion rate as the hypercube.

The lesson for designers of parallel algorithms is immediate and striking: the smallest feasible quantum of parallelism is dictated by the communication patterns of the algorithm *and* the network topology. Excessive parallelism leads to at best negligible performance gains and and at worst performance decreases due to increased overhead.

We note that situations do arise in which factors intrinsic to an intended application dictate the use of a specific ratio of computation to communication time for messages. In this case, the designer must be cognizant of the fact that range of optimality for a specific network does not span the entire spectrum of network sizes. This is clearly illustrated by the crossing of the bounds on message completion rate in Figures II-III. Thus, one might be justified in using one interconnection for ten nodes and a different one for one hundred nodes. One can analytically or numerically determine where the bounds for two networks cross by equating the *VS* products of the performance limiting network devices and attempting to solve for K, the network size at which they are equal.

Size Independent Performance Bounds

The *VS* products can also be used to derive performance bounds that are independent of network size. Consider the limit

$$\lim_{K \to \infty} \left[\frac{1}{\max_i V_i S_i} \right] .$$

When the limit exists, it defines an absolute upper bound on the message completion rate of a network even if it contained an *infinite* number of nodes. Using the ring with uniform message routing as an example once more,

$$X_0 < \lim_{K \to \infty} \left[\frac{1}{\max \left\{ \frac{S_{PE}}{K}, \frac{S_{CL}(K + 1)}{4K} \right\}} \right]$$

$$= \lim_{K \to \infty} \frac{4K}{S_{CL}(K + 1)}$$

$$= \frac{4}{S_{CL}} .$$

No ring based system with uniform message routing can pass messages faster than this rate.

Needless to say, any networks possessing this property are unlikely to be suitable if performance increases by incremental network expansion are important. However, these bounds apply *only* to the message routing distribution defining the visit ratios. If message routing exhibits more locality than the uniform message routing distribution (i.e., messages visit nodes closer to their source with higher probability than in the uniform routing distribution), message passing rates greater than these bounds can be achieved.

Nodes with Limited Communication Bandwidth

Heretofore we have assumed that all communication links connected to a node operate asynchronously and in parallel. Thus, a node could be transmitting or receiving on all links to which it is connected simultaneously. This assumption quite accurately models the operation of the proposed X-tree communication controller [DESP80], but is fails totally as a model of the Micronet prototype [WITT80] in which only one link connected to a transmitter and receiver node can be active at any given time. Whereas the latter design is unlikely to be realized in a production machine, it is possible that completely parallel operation may not be adopted either.

Goodman [GOOD81] has suggested that it is much more natural to view the communication bandwidth of a VLSI implementation of a communication controller as fixed. Thus, two networks, one with C connections per node and the other with $\hat{C}$, would be perceived as having effective message transmission times of CS_{CL} and $\hat{C}S_{CL}$ respectively, if the base time to transmit a single message across a single link were S_{CL}. Intuitively, one can view the communication controller as being multiplexed among the communication links attached to the node.

This simple technique permits us to determine the possible effects of limited communication bandwidth at the nodes. By scaling each of the communication link visit ratios by the number of links connected to each node, one can obtain a set of curves similar to those in Figures II-III. Interestingly, this change of scale does not radically affect the relative rankings of the networks.

Bounds for Finite Workloads

Because systems rarely if ever operate at their theoretical capacity, it is appropriate to ask if simple characterizations of message passing rates exist for networks operating under finite workloads. The importance of this question is underlined by the observation that different functions approach their asymptotes at different rates. Thus, even though the asymptotic message passing rate of network A is greater than that of network B, the message passing rate of B may actually be larger than that of A for all workloads of interest.

If $V_i S_i$ is the total amount of service required by an average message at device i, the sum

$$R_0 = \sum_i V_i S_i$$

must be the total amount of service required by an average message at all devices. We have already established that the network message passing rate X_0 is bounded above by (2). Is there anything that can be said about the message passing rate for small numbers of messages in a network? This question has been answered affirmatively [DENN78].

Suppose only one message were present in the network. The rate at which this message is processed must be simply $\frac{1}{R_0}$. Since the message passing rate can rise at most linearly with the number of messages in the network, $X_0(N)$, the network message passing rate when N messages are present is bounded by

$$X_0(N) \leq \frac{N}{R_0} . \tag{4}$$

Combining this with (4) we have

$$X_0(N) \leq \min\left\{\frac{N}{R_0}, \frac{1}{V_b S_b}\right\} . \tag{5}$$

As Figure V illustrates, the two components of the bound in (5) must intersect, and they do so at

$$N^* = \frac{R_0}{V_b S_b} = \sum_i \frac{V_i S_i}{V_b S_b} \leq \textit{number of devices} .$$

At the critical point N^* message queueing begins to occur because at least two messages must be at the same device, one obtaining service and the other waiting for service.

The true message passing rate $X_0(N)$ is a monotonically increasing function bounded above by (5). For $1 \leq N \leq N^*$, the bound (4) provides a simple means of ranking the maximum message passing rates of networks. Most importantly, this ranking captures something of the rate at which networks approach their asymptotic message passing rate. This is illustrated for networks of 64 nodes under the uniform message routing distribution and device service times of one in Figure VI.

Summary

We have described a simple technique based on the asymptotic behavior of a queueing network that provides a substantially more robust performance metric for interconnection networks than the mean message path length. The technique applies to networks that are either lightly or heavily loaded and is asymptotically exact in both cases. Furthermore, it represents the relationship between the service rates of the nodes and the network in a manner that permits determination of the appropriate quanta of computation for a given network topology and selection of the most suitable network for a specific number of nodes. Finally, one would expect a distributed computation to exhibit more locality in its message routing distribution than that reflected by uniform routing. Asymptotic bound analysis is extensible in a straightforward manner to other message routing distributions [REED83a].

REFERENCES

AKL80 S. G. Akl, D. T. Barnard, and R. J. Doran, "Simulation and Analysis in Deriving Time and Storage Requirements for a Parallel Alpha-beta Algorithm," *Proceedings of the 1980 International Conference on Parallel Processing*, August 1980, pp. 231-234.

BURT81 F. W. Burton and M. R. Sleep, "Executing Functional Programs on a Virtual Tree of Processors," *Proceedings of the 1981 Conference on Functional Programming Languages and Computer Architecture*, October 1981, pp. 187-194.

DENN78 P. J. Denning and J. P. Buzen, "The Operational Analysis of Queueing Network Models," *Computing Surveys*, (10,3), September 1978, pp. 225-261.

DESP78 A. M. Despain and D. A. Patterson, "X-TREE: A Tree Structured Multiprocessor Computer Architecture," *Proceedings of the Fifth Annual Symposium on Computer Architecture*, ACM Sigarch Newsletter, Vol. 6, No. 7, April 1978, pp. 144-151.

ELDE79 O. I. El-Dessouki and W. H. Huen, "Distributed Enumeration on Network Computers," *Proceedings of the 1979 International Conference on Parallel Processing*, August 1979, pp. 137-146.

FINK81 R. A. Finkel and M. H. Solomon, "The Lens Interconnection Strategy," *IEEE Transactions on Computers*, Vol. C-30, No. 12, December 1981, pp. 960-965.

GOOD81 J. R. Goodman and C. H. Se´quin, "Hypertree: A multiprocessor Interconnection Topology," *IEEE Transactions on Computers*, Vol. C-30, No. 12, December 1981, pp. 960-965.

PEAS77 M. C. Pease, "The Indirect Binary n-cube Microprocessor Array," *IEEE Transactions on Computers*, Vol C-26, No. 5, May 1977, pp. 458-473.

PREP81 F. P. Preparata and J. Vuillemin, "The Cube-Connected Cycles: A Versatile Network for Parallel Computation," *Communications of the ACM*, Vol. 24, No. 5, May 1981, pp. 300-309.

REED82 D. A. Reed, "Optimal Routing in the Cube-connected Cycles Interconnection Network," *Purdue University Department of Computer Sciences Technical Report*, CSD-TR-412, August 1982, submitted to *IEEE Transactions on Computers*.

REED83a D. A. Reed and H. D. Schwetman, "Cost-Performance Bounds for Multimicrocomputer Networks," *IEEE Transactions on Computers*, Vol. C-32, January 1983, *to appear*.

REED83b D. A. Reed, "Performance Evaluation Techniques for Large Multimicrocomputer Networks," *PhD Dissertation*, Purdue University, in preparation.

SMIT82 C. U. Smith and D. D. Loendorf, "Performance Analysis of Software for an MIMD Computer," *Proceedings of the 1982 ACM Symposium on Measurement and Modeling of Computer Systems*, *Performance Evaluation Review*, Vol. 11, No. 4, pp. 151-162.

WITT80 L. D. Wittie, "Packet Switching Using Concurrent Pascal in a Network Computer," *IEEE Fall Computer Conference*, Fall 1980, pp. 358-365.

WITT81 L. D. Wittie, "Communication Structures for Large Multimicrocomputer Systems," *IEEE Transactions on Computers*, Vol. C-30, No. 4, April 1981, pp. 264-273.

Table I Notation

Symbol	*Definition*
K	number of network nodes
L	mean message path length
N	network message population
N^*	message population at which message queueing must occur
R_0	total amount of service required by an average message
S_i	mean device i service time
S_{CL}	mean communication link service time
S_{PE}	mean node service time
T	number of distinct link visit ratios
V_i	visit ratio for device i
X_0	network message completion rate

Table II Network Characteristics

Simple Ring (K nodes and K links)

$$L^{ring} = \begin{cases} \dfrac{K^2}{4(K-1)} & K \text{ even} \\[2ex] \dfrac{K+1}{4} & K \text{ odd} \end{cases}$$

$$V_{CL}^{ring} = \frac{L^{ring}}{K}$$

Spanning Bus Hypercube (w^D nodes and Dw^{D-1} buses)

$$L^{SBH} = \frac{Dw^{D-1}(w-1)}{w^D - 1}$$

$$V_{CL}^{SBH} = \frac{L^{SBH}}{Dw^{D-1}}$$

Dual Bus Hypercube (w^D nodes and $2w^{D-1}$ buses)

$$L^{DBH} = \left(\frac{2D^2 - 5D + 4}{D-1}\right)\left(\frac{(w-1)w^{D-1}}{w^D - 1}\right)$$

$$V_{CL}^{0\text{-th dimension}} = \left(\frac{D^2 - 3D + 3}{D-1}\right)\left(\frac{w-1}{w^D - 1}\right)$$

$$V_{CL}^{\text{other dimensions}} = (D-1)\left(\frac{w-1}{w^D - 1}\right)$$

Toroid (w^D nodes and Dw^D links)

$$L^{toroid} = \begin{cases} \dfrac{Dw^{D+1}}{4(w^D - 1)} & w \ even \\[2ex] \dfrac{Dw^{D-1}(w^2 - 1)}{4(w^D - 1)} & w \ odd \end{cases}$$

$$V_{CL}^{toroid} = \frac{L^{toroid}}{Dw^D}$$

Cube-connected Cycles ($D2^D$ nodes, $D2^{D-1}$ cube links, and $D2^D$ ring links)

No closed form expressions for the mean message path length and link visit ratios are known, but an optimal message routing algorithm has been derived [REED82].

R-ary M-cube (MR^M nodes, $M(R - 1)R^M$ cylinder links, and MR^M column links)

$$L^{cylinder} = M\left(\frac{R - 1}{R}\right)\left(\frac{MR^M}{MR^M - 1}\right)$$

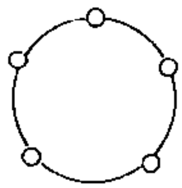
$$V_{CL}^{cylinder} = \frac{L^{cylinder}}{M(R - 1)R^M}$$

$$L^{column} = \left(\frac{MR^M}{MR^M - 1}\right)\left(\frac{(M - 1)R^M - MR^{M-1} + 1}{R^M (R - 1)}\right) +$$

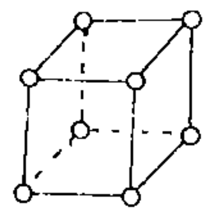
$$\left(\frac{MR^M}{MR^M - 1}\right)\begin{cases} \dfrac{M}{4} & M \ even \\[2ex] \dfrac{M}{4} - \dfrac{1}{4M} & M \ odd \end{cases}$$

$$V_{CL}^{column} = \frac{L^{column}}{MR^M}$$

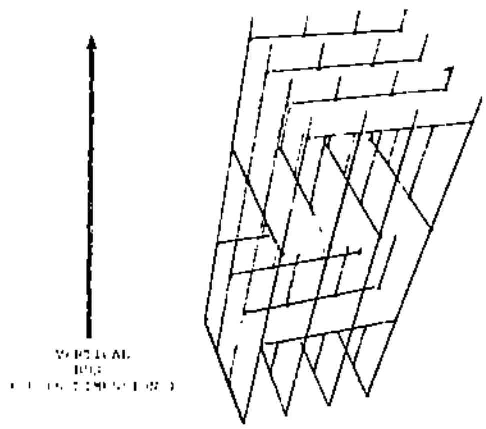
Figure 1 Interconnection networks



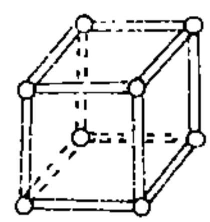
Simple Ring ($K = 5$)



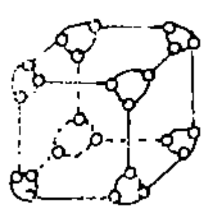
Spanning Bus Hypercube ($D = 3, \ w = 2$)



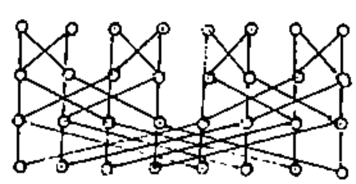
Dual Bus Hypercube ($D = 3, \ w = 4$)



Toroid ($D = 3, \ w = 2$)



Cube-connected Cycles ($D = 3$)



R-ary M-cube ($R = 2, \ M = 4$)

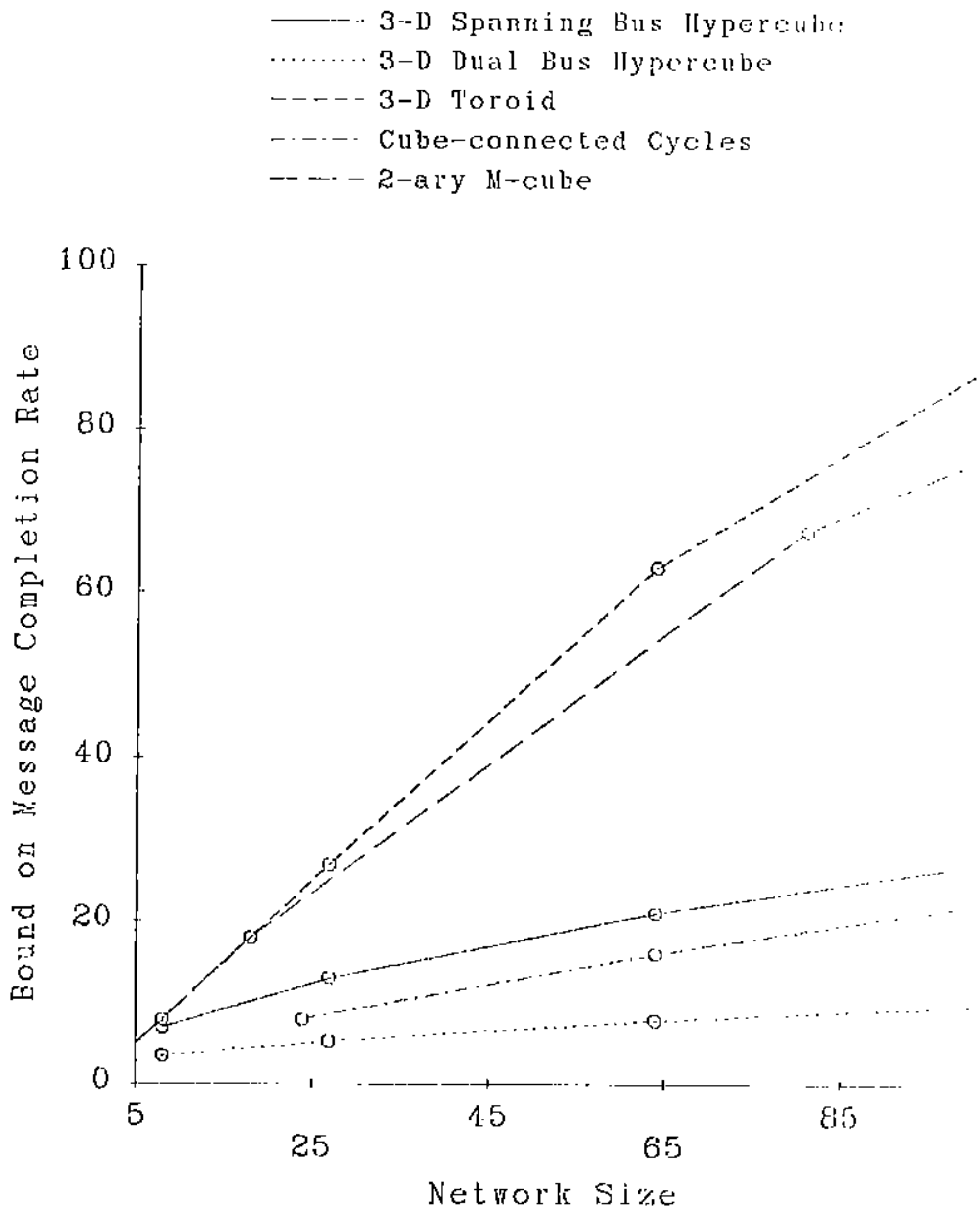


Figure II
Upper Bound on Message Completion Rate
Small Networks

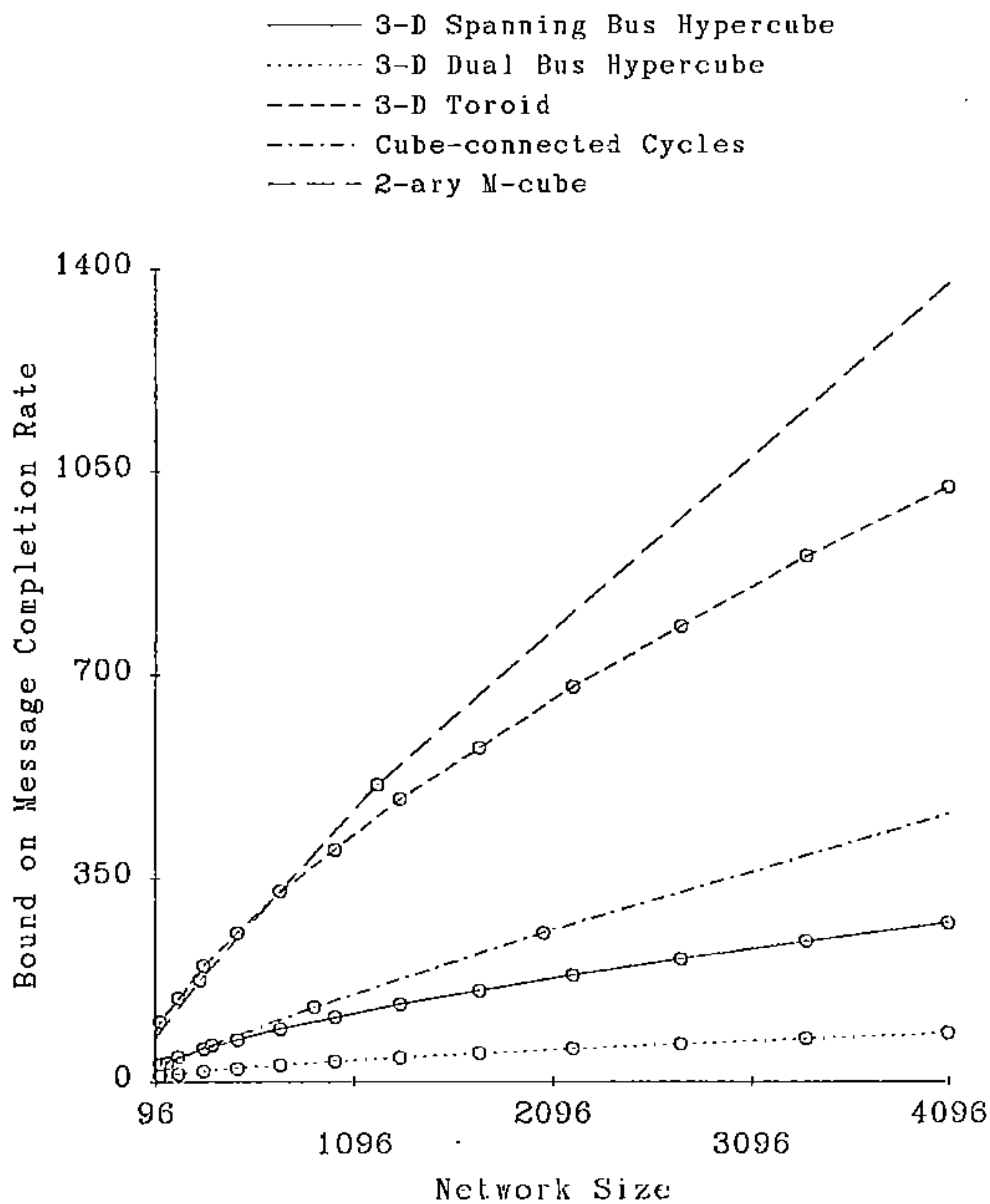


Figure III
Upper Bound on Message Completion Rate
Medium and Large Networks

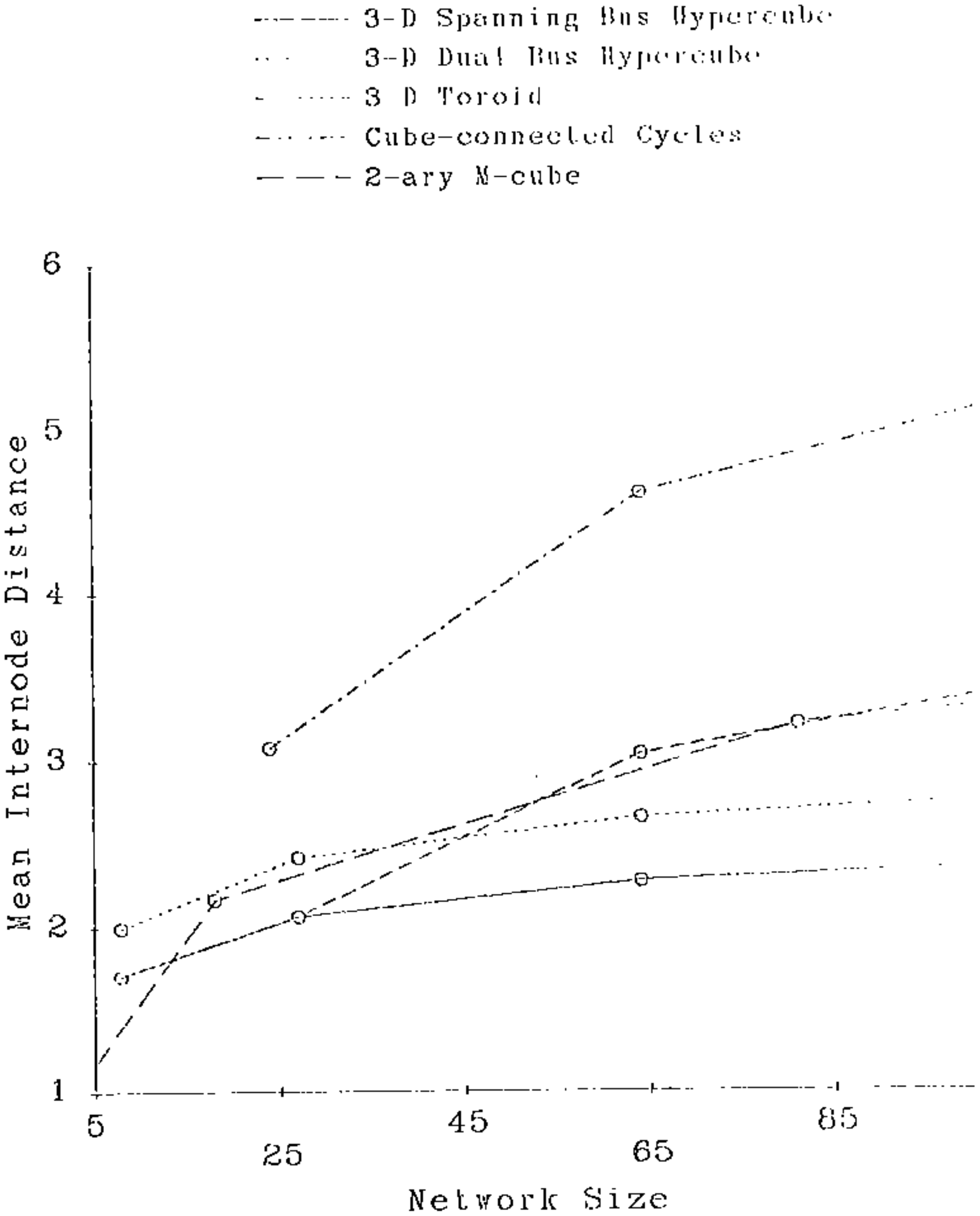


Figure IV
Mean Internode Distance - Uniform Routing
Small Networks

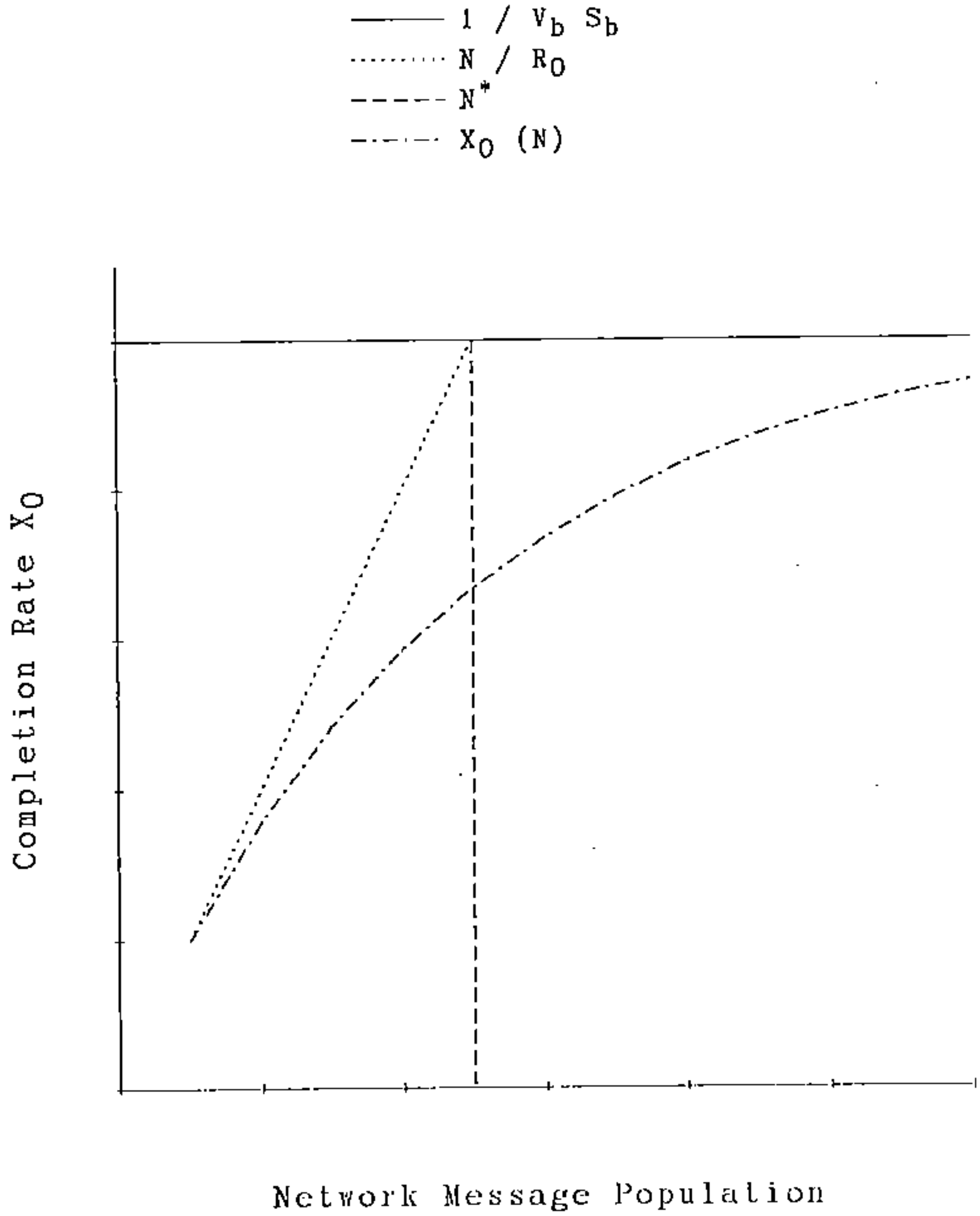


Figure V
Finite Population Bottleneck Bounds
on the Message Completion Rate X_0

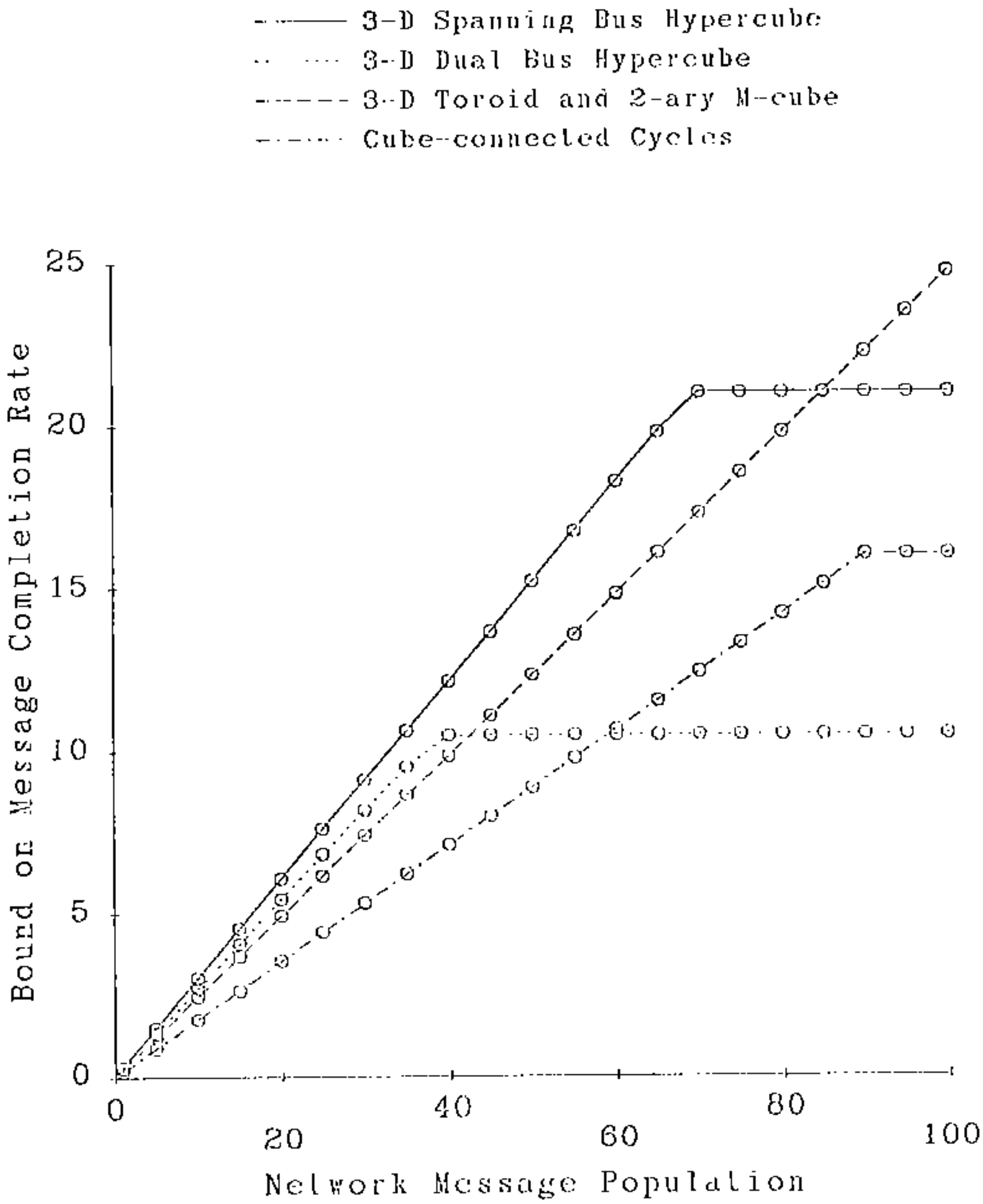


Figure VI
Bottleneck Bound on Message Completion Rate
Uniform Message Routing - 64 Node Networks