

Power/Performance Trade-offs for Direct Networks[†]

Chirag S. Patel, Sek M. Chai, Sudhakar Yalamanchili and David E. Schimmel

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0250
e-mail: sudha.yalamanchili@ece.gatech.edu
Fax: (404) 894-9959

Abstract

High performance portable and space-borne systems continue to demand increasing computation speeds while concurrently attempting to satisfy size, weight, and power constraints. As multiprocessor systems become necessary to satisfy the computational requirements, interconnection network design will become subject to similar constraints. This paper focuses on the design of multiprocessor interconnection networks under power constraints. We analyze and study the relationship between the no-load message latency and system parameters such as network dimension, channel width, distance, and available power to determine guidelines for multiprocessor interconnection network designs. This study is an extension and application of the model developed in [8], and provides a number of interesting results: i) we have observed that under a fixed power constraint, the network dimension which achieves minimal latency is a slowly growing function of system size, ii) as we increase the available power per node for a fixed system size, the dimension at which message latency is minimized increases, iii) the overall message latency is very sensitive to the distribution of power between driving inter-router channels and switching data through the intra-router datapaths. The paper concludes with a summary of trade-offs predicted by the model and proposals for validating the model..

Keywords: multiprocessor interconnection network, power constraint modeling, low power embedded systems, spaceborne systems.

1.0 Introduction

The design of modern high performance multiprocessor interconnection networks is a process of optimizing the performance of the network in the presence of implementation constraints. Past analysis has focused on constraints such as wiring area [3], pin out [9], wire delays [6], and switching speeds [1]. This paper focuses on a similar analysis in the presence of a constraint on the available power.

The motivation for this research came from an analysis of the computational requirements of future deep space missions. Unoptimized designs have power dissipations reported in the tens of watts. This is clearly infeasible for on-board computing in deep space missions where power budgets for computations is on the order of several watts.

[†]. This research was partially supported by NASA Jet Propulsion Laboratory under contract 960586.

We need a comprehensive approach to system design that involves low power technologies, compatible system level design, and algorithms that support activity minimization. The explosive growth of telecommunications applications and portable computing has also produced a need for minimizing the consumption of limited available power. As the computational requirements of space and portable systems grow, we will begin to see architectures that employ multiple processors. The interconnection network is a major consumer of power in these systems and must come under the same scrutiny as the rest of the system if we are to be able to design networks that can be efficient consumers of power. This study focuses on the low power optimizations that can be achieved by varying the topology and channel widths of the network. We wish to study the sensitivities to application characteristics such as message injection rates, and communication/computation ratios.

The following section describes concepts for discussing the relationships between application characteristics, network topologies, physical implementation parameters, and available power. This model is used to study the trade-offs between the no-load message latency and available power in Section 4.0. The paper concludes with a summary of trade-offs and our future research directions.

2.0 Power Dissipation in Pipelined Networks

The class of networks considered in this paper are the torus connected, bidirectional, k -ary n -cubes. A k -ary n -cube is a network with n dimensions and k processors in each dimension. In torus connected k -ary n -cubes, each processor is connected to its immediate neighbors modulo k in every dimension. Routing within each dimension is orthogonal to routing in other dimensions. The mesh networks differ from tori in that there are no wrap around links in each dimension. Most contemporary parallel machines utilize such networks or variants including the Intel Paragon [12], Cray T3 models D [10] and E [11], and Ametek 2010.

Power is the rate of energy dissipation in a system. In a typical VLSI CMOS circuit, this energy is consumed in switching transistors, in leakage due to circuit device structures, and in holding the logic state of a circuit. The rate at which this energy is used by a system determines its power dissipation. We assume that a network consumes approximately the same quantity of energy to transport a message to its destination independent of the switching technique used (i.e. either packet switching or virtual cut-through switching may be employed). This analysis assumes no contention in the network. A model of contention is necessary to accurately reflect power consumption in active networks. However, as a first step we study the basic relationships between topology, power and physical constraints. We will then seek to extend this model to incorporate the effects of contention within the network.

The latency of message delivery and the dissipated power are related by the architecture of the network. For example, a network architecture with wide channel widths can transport data packets in less time compared to a network architecture with smaller channel widths. However, the wider channel network will consume energy at a higher rate because of the larger data buffers and wider buses. For example, consider the pipelined

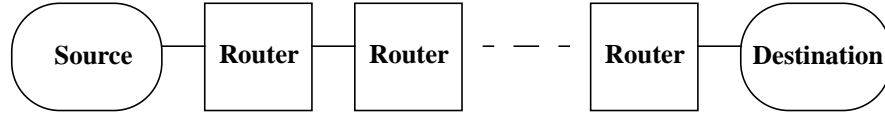


Figure 1. Interconnected Sequence of Routers in a Network

transmission of a single message across multiple routers as illustrated in Figure 1. Since our power model relies on switching transistors, we can model the network routers with a certain number of transistors, with a distinct transistor count for the routing/arbitration, switching unit, and line drivers. To find the peak power consumed during the pipelined transmission of a message, we can visualize a sequence of network nodes from the source to destination. Peak power dissipation occurs when the largest number of transistors switch, and this event might be expected to occur when the pipelined message is spread across the maximum number of routers. We use this view of power dissipation in the pipelined network to model the peak power dissipation during a message transmission. This view of power consumption is a *message-centered* view. The total available power is regarded as being distributed to each transmitted message, and each such message has a certain available power to navigate through the network towards the destination.

Another view of power dissipation involves a *router-centered* view. For each router in the system, a subset of the total available power is allocated for routing and arbitration unit (RAU), switching, and driving inter-router channels. From this perspective, each router is allocated a power budget to locally switch messages.

3.0 Power and Latency Models

From a system performance point of view messages encounter three distinct delays: routing delay, switching delay, and inter-router channel or wire delay. Routing delay is the time it takes for the routing and arbitration unit (RAU) to make a decision of routing the incoming messages. The switching delay is the time it takes for the crossbar switch to physically route the message to inter-router channels; and the wire delay is the time it takes for the message to be transferred over the channel from one processor to the next. Each of these distinct units have to do work in order to route the message which requires energy. Energy over time is defined as power dissipated in each unit. Therefore, to route a message from one processor to the next, all of the units - RAU, switch, and the wire - dissipate fixed amount of power. Therefore in our model, the critical path for a message involves the routing & arbitration unit, switching logic and inter-router wire delays. We will examine the power and time delays associated with each component. The parameters used in the model section are defined in table 1 below which is followed by the assumptions made in the derivation of the model.

| Sym bol | Description | Sym bol | Description |
|--------------------|--|--------------------|---|
| C_{in} | Input gate capacitance | P_r | Power dissipated in the RAU |
| C_o | Output capacitance of minimum transistor in the cascaded drivers | P_s | Power dissipated in the switch |
| C_{ox} | Gate oxide capacitance | P_{sc} | Short circuit power consumption |
| C_p | Next stage capacitance in the driver chain | P_s | Power dissipated in the wire |
| C_{pad} | Output pad capacitance | P_t | Total power dissipated in the system |
| C_t | Total capacitance of the node | R_o | Output resistance of minimum transistor in cascaded drivers |
| C_w | Total wiring capacitance | R_w | Wiring resistance |
| D | Average distance or number of hops | S_n | Delay constant of nFET transistor |
| f_i | Frequency of the driver | T_{clk} | Period of the clock in the driver chain |
| f_s | Frequency of the switch | T_{chain} | Time delay in the driver chain |
| f_w | Frequency of the wire | T_{line} | Time delay in the wire or the line |
| G_m | Messages injected into the network | T_{tof} | Time of flight |
| G_n | Messages handled by the network | t_r | RAU delay |
| K_1 | Duty cycle | t_s | Switch delay |
| K_m | Number of messages injected by the system | t_w | Wire delay |
| K_n | Number of messages injected by the system | V_{dd} | Supply voltage |
| K_r | Ratio of number of transistors in the RAU to switch | V_t | Threshold voltage |
| K_s | Fraction of active gates during switching | v_p | Propagation velocity of signal |
| L | Message bit length | W | Message channel width |
| l_2 | Base latency | W_1/L_1 | The size of the minimum transistor in the driver chain |

| Sym bol | Description | Sym bol | Description |
|--------------------|-------------------------------------|--------------------|--|
| M_r | Number of transistors in the RAU | w_{avg} | Average wire length between processors |
| M_s | Number of transistors in the switch | X | Number of inverters in cascaded driver chain |
| N | Number of nodes | α | Driver scaling factor |
| n | Network dimension | β | Minimum size transistor transconductance |
| P_{cap} | Capacitive power consumption | Γ | Total power constant coefficient |
| P_{msg} | Power dissipation per message | μ | Mobility of electrons |
| P_{node} | Power dissipation per node/router | τ | Rise time of the input clock pulse to the driver |

1. Wormhole switching for message transmission
2. The total power dissipated in the node is a function of power dissipation in the routing and arbitration unit, internal switch, and the intra-router wire. Note that in this model any power dissipation in writing and reading buffers is included on the switching delay.
3. Power dissipated in the RAU is related to the power dissipated in the switch. We use a simple model where this relationship is given by the ratio of number of transistors in the RAU to that in the switch.
4. The intra-router wire delay is given by the time of flight and the RC delay of the wire.
5. Driver chains are used to drive the off-chip physical channel and the delay through the driver chain is accounted in the total wire delay.
6. The RAU delay is equal to the switching delay. This is generally the case in most modern high speed routers.

3.1 Model Development

The time delay through the switch is referred to as the switching delay and is characterized by t_s . The delay through the Routing Arbitration Unit (RAU), time required to make a routing decision is denoted by t_r . The time to transfer a unit of the message across a physical channel on the wire is denoted by t_w . Using a wormhole switching implementation, the overall routing delay is given by,

$$t_{wormhole} = [D(t_r + t_s + t_w)] + \left[\max(t_s, t_w) \left[\frac{L}{W} \right] \right] \quad (1)$$

where D is the average number of links or distance traversed by a message through the network, L is length of the message in bits and W is the channel width of the physical

channel [4]. We proceed to determine expressions for t_s and t_w constrained by the available power. By imposing this constraint, we are constraining the values of W , t_r , t_s , and t_w , and therefore the message latency. For a fixed number of nodes, the choice of topology also affects D , W , t_s and t_w (assuming t_r is constant). We can now establish a relationship between available power and message latency as a function of topology. We first develop power dissipation models for the network components, and couple these with models of the three network delays: routing, switch, and wire.

3.2 Power Dissipation Models

The available power can be obtained from the available energy used in a finite amount of time. The total power available for the network is assumed to be equally divided among all of the nodes, and can be expressed in terms of the power per node or power per message. The power per node, P_{node} , is simply the sum of the power dissipated in the RAU, switch, and the wire connecting two adjacent nodes and can be written as:

$$P_{node} = (P_r/D) + P_s + P_w \quad (2)$$

The power per message P_{msg} , can be computed as:

$$P_{msg} = P_r + (P_s + P_w)D \quad (3)$$

where P_r , P_s and P_w are the power dissipated in the RAU, switch, and the wire between routers respectively, and D is the average number of links traversed by a message. It is important to note that (2) and (3) simply represent two different interpretations of the distribution of the total available power in the network and they are not independent equations. P_{node} and P_{msg} are related to the total power available to the network as follows:

$$P_t = NG_n P_{node} = G_m P_{msg} \quad (4)$$

where G_m and G_n represent two ways to capture the demand on network activity. G_m represents the number of messages injected into the network by all of the processing nodes and therefore captures the total demand on the network. A model for G_m can be found in [7] where the effect of processor and interconnect speed is taken into account. The model of [7] is not constrained by the total power available to the network and therefore has to be modified. The total number of messages injected into the network, G_m , under a fixed power constraint can be expressed as

$$G_m = (S_p t_w N)/(rL) \quad (5)$$

where S_p is the processor speed in (MFLOPS/sec) and r is the computation to communication ratio in (MFLOPS/bit), t_w is the time it takes to transfer a bit between routers and it is constrained by the amount of power available to do so. G_n is the number of messages handled by the processing nodes of the network. It is obtained by noting the fact that a message has to traverse D number of hops and therefore G_n is simply DG_m . It is now seen that P_{node} and P_{msg} are equivalent under the representation of $G_{m,n}$

The power dissipation in the RAU and switch is governed by the mechanism of charging and discharging the total capacitance. Therefore in our analysis of the router we assume that the power dissipated in the RAU is related to the power dissipated in the switch as follows:

$$P_r = \frac{M_r}{M_s} P_s = K_r P_s \quad (6)$$

Substitution of (4) and (6) into (2) yields the following expression for the total power dissipated in the network:

$$P_t = NG_n \left[\left(\frac{K_r}{D} + 1 \right) P_s + P_w \right] \quad (7)$$

This equation expresses the distribution of total power in the network from a system perspective and it is constrained by various system parameters such as the average distance between communicating nodes, the number of nodes, and the number of messages handled by a node. It is an equation with two unknowns P_s and P_w . We leave them as free variables to study the effect of distribution of power between the switch and inter-router channels (wire) on network performance. Now (7) can be rearranged to provide an expression for P_s

$$P_s = \frac{D(P_{node} - P_w)}{K_r + D} \quad (8)$$

This expression permits us to study the distribution of power between the wire and through the switches. Note that under a fixed constraint on total power, minimum wire power corresponds to maximum switch power dissipation and vice a versa.

The preceding equations were derived from a system architecture stand point. In order to relate the system architecture parameters to device and technology parameters, we must develop models in terms of power dissipation within the switches and physical channel interfaces, e.g., drivers. We follow the same modeling technique of viewing power dissipation as simply charging and discharging of the total capacitance. In doing so we arrive at the following equations.

$$P_s = \frac{1}{2} K_s M_s C_t V_{dd}^2 f_s \quad (9)$$

$$P_w = K_1 C_w V_{dd}^2 f_w w_{avg} W \quad (10)$$

Here, $M_s = n^2 W$ is the number of transistors in the switch assuming a full physical crossbar connecting all input ports and output ports; W is the channel width.

For the purpose of analyzing the effect of wire length we can determine this increase in wire length as [8]

$$w_{avg} = 2l_2\sqrt{N} - 1/(nk) \quad (11)$$

following the approach in [1] where N is number of nodes, n is the dimension of the network, and l_2 is the physical distance between adjacent nodes. Thus the expression is normalized to the distance between two routers in a two dimensional torus. In this manner the analysis is relative to the best that can be achieved for 2-D tori in any given technology. The above expression is based on a linear wire delay model, and is therefore conservative.

3.3 Time delay models

The delay through the switch can simply be found by rearranging (9) ($t_s = 1/f_s$) to obtain,

$$t_s = \frac{0.5K_s n^2 W C_t V_{dd}^2}{P_s} \quad (12)$$

Substitution of W from (10) into (12) yields,

$$t_s = \left(\frac{P_w}{P_s}\right) \frac{0.5K_s C_t V_{dd}^2 n^2 t_w}{K_l C_w V_{dd}^2 w_{avg}} \quad (13)$$

Normally the connection between two routers are off-chip signal paths and we have to treat this interconnect as a transmission line rather than an RC line. The geometry of these off-chip wires is “fat” to reduce the resistance of the line. However, the inductance will dominate the signal propagation and it must be taken into consideration. The scaling of the drivers in the cascaded chain is traditionally derived under the assumption that we have enough power to use the optimum number of drivers realizing minimum latency. This is unlikely to be the case under stringent power constraints, reducing the maximum number of drivers permitted. From our model captured in (7), the total power is distributed between the switch and the wire. An important question is the nature of this power distribution. What is the desired ratio? *Should we build faster drivers or should we design faster switches? What are the trade-offs on power distribution?* These are very important questions and must be addressed for future low power network designs. Therefore, in our analysis of low power network designs, we have to constrain the drivers by the maximum allowable power to the network. Under such constraint the maximum allowable drivers can be given by [8]

$$X = \frac{\ln(1 + (\alpha - 1)P_w/\Gamma)}{\ln(\alpha)} \quad (14)$$

where $\Gamma = (V_{dd}^2 f(\alpha C_{in} + C_p)) + ((\beta/12)(V_{dd} - 2V_t)^3 (\tau/T_{clk}))$ is the coefficient of the total power term. From (14), X is a function of both the scaling factor, α , the maximum available power, and the power dissipated in the wire, P_w . The motivation for

this simplification is that for a given power budget, we can distribute the power in several ways: a) completely within the switch, b) completely within the wire c) across the switch and the wire. By leaving P_w as a free variable, we can study all three cases with the number of drivers be limited by the power allocated to driving the wires.

The delay through the wire is modeled as the sum of the delay through the driver chain, the RC line, the time of flight and can be derived [8] to give

$$t_w = 3\alpha(X-1)R_oC_o + S_n \left(\frac{R_o}{\alpha^X - 1} \right) (2C_{pad} + C_w + (\alpha^X - 1)C_o) + \frac{w_{avg}}{v_p} \quad (15)$$

The delay through the RAU will be assumed to be equal to the delay through the switch ($t_r = t_s$). This is reasonable since modern routers attempt to perform routing decisions in the time it takes to drive a flit through the switch. We are now equipped with all of the necessary models to be substituted in (1) to study the overall latency through the router as a function of network topology and power distribution under a maximum allowable power constraint. Thus, the complete expression for the no load latency of a message through a k -ary n -cube network under a power constraint given as

$$t_{wormhole} = [D(t_r + t_s + t_w)] + \left[\max(t_s, t_w) \left[\frac{L}{W} \right] \right] \quad (16)$$

where t_s is defined in (13), t_w is defined in (15), W is defined in (10), and $D = (nN^{1/n})/4$ [1,4].

4.0 Performance Analysis

The goal of our analysis is to understand the relationship between the topology of the interconnect, message latency, and workload parameters under a fixed power budget. The preceding analysis has identified the relationships between power and features of the topology such as switching speeds, channel delays, and channel widths, in addition to numerous physical parameters. These relationships are complex, and their impact in determining the choice of topology is not obvious. We utilized the Maple V symbolic computational package to study these relationships [2]. Fixed power budgets are represented by total power in watts. This study focuses on the relationship between three variables: i) message latency, ii) distribution of power between the switches and wire (channels), and iii) network dimension. We are interested in the network dimension that minimizes message latency. In this section the preceding relationship is examined for small (16 nodes) and medium (256 nodes) systems. The section concludes with some observations about the behavior of the bisection bandwidth under a fixed power constraint.

4.1 Power Constrained Analysis of Small Systems

We studied the effect of a fixed power budget on the topology of small systems. The total power dissipation was initially fixed at two watts. In our model the total power budget is

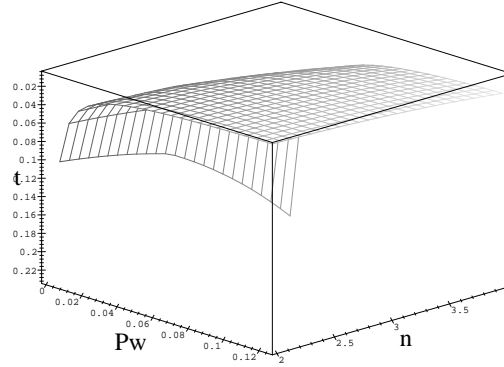


Figure 2. Message latency (in μ sec) as a function of network dimension and power distribution for 16 nodes system

distributed equally among all of the nodes in the network, i.e., each node has a maximum dissipation of P_t/N watts. The simulations were carried out for a 16 node system and the overall latency as a function of network dimension and power dissipated in the wire is illustrated in Figure 2. *Note that the vertical axis corresponds to decreasing latency. Thus a maximum on the surface corresponds to minimum latency.* When power dissipated in the wire is a small (10-20%) fraction of the total power allocated per node, the latency is very high and is very sensitive to the network dimension. Low dimensional network are clearly preferred at any power distribution as seen in the figure. As we increase the power allocated to the wire, the latency starts to decrease until we reach a certain threshold in the value of P_w after which there is no significant change in the value of overall latency. In this regime, the overall delay is also insensitive to the network dimension. For a 16 node system, the threshold point occurs when power dissipated in the wire is approximately 40% of the total power allocated per node. When P_w is small, relatively more power is being dissipated in the switch and therefore the wire delay will be dominant; the converse is true at high values of P_w as is apparent from the figure.

From (5), it is clear that a change in r is inversely related to a change in S_p , the processor speed. Normally, improvement in the processor speed leads to lower latencies since we are able to process more messages in a given time cycle. This is not necessarily true when a system is constrained by power. Higher processor speed results in a higher message injection rate, yet the total available power is fixed. This will require the network to handle a larger number of messages with the same amount of power. If the power budget is very tight, the increased number of messages in the network will cause an increase in the overall latency. Note that as we increase processor speed, larger systems tend to favor slightly higher dimensions and smaller systems favor lower (generally two) dimensions.

Finally, the system was analyzed under various power budgets. We decreased the available power from two to one watt and observed the effect on network performance. The shape of the surface of the graph remained qualitatively the same as in Figure 2 except

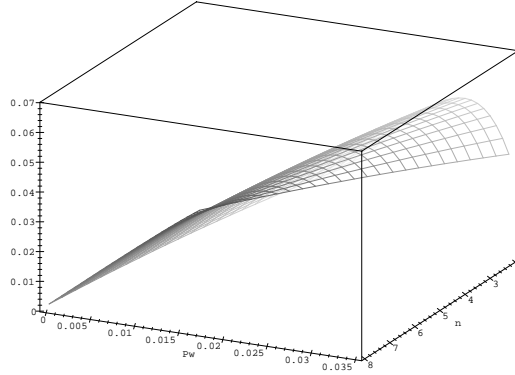


Figure 3. Power per node profile for a 256 nodes system

that the overall latency became larger. The system also became more sensitive to dimension at low values of P_w as discussed above.

4.2 Power Constrained Analysis of Medium Systems

With a power budget of two watts, we analyzed the network performance of a system of 256 nodes. In trying to simulate this system, we encountered impractical solutions for the overall latency of the network. Subsequent analysis produced the following constraint on minimum power dissipation. Consider a network of 256 nodes which admits approximately 10^3 more injected messages per unit time than a network of 16 nodes. The delivery of these messages requires a specific minimum amount of energy in order to ensure that messages are delivered at the same rate that they are injected. Maintaining this injection rate necessitates a minimum power budget for network operation to be feasible. For a given power budget per node, the number of messages switched through the node must “share” this power dissipation capacity. The profile is given in Figure 3 for a total power of 100 watts. The power, P_{node} is further allocated between driving flits through the switch and across the wire. The condition for maximum wire power dissipation is the point at which P_{node} equals P_w , which is infeasible as switching power is zero.

We increased the total power to ten watts. Even though the maximum power per node is approximately 40 mwatts, we found that wire power dissipation must be in the vicinity of 0.1 mwatt to yield minimum latency. This observation is shown in Figure 4(a). In contrast to a 16 node system, it is preferable to operate in a regime where P_w is very small. As increased power is dissipated in driving the physical channel, the system becomes more sensitive to the network dimension. As we increase the available power dissipation to 100 watts, the maximum power dissipated in the wire depends on the choice for the network dimension. This budget is approximately 0.3 watts per node. However, the optimum value clearly depends on the choice of network design. If we wanted to build low dimensional networks we have to sacrifice latency and power as seen in Figure 4(b). A three dimensional system can be built at the expense of using small amount of power in the wire (i.e. using 0.035 watts instead of 0.3 watts) and therefore penalizing the overall latency. Another interesting feature of the analysis is the response of latency to power dis-

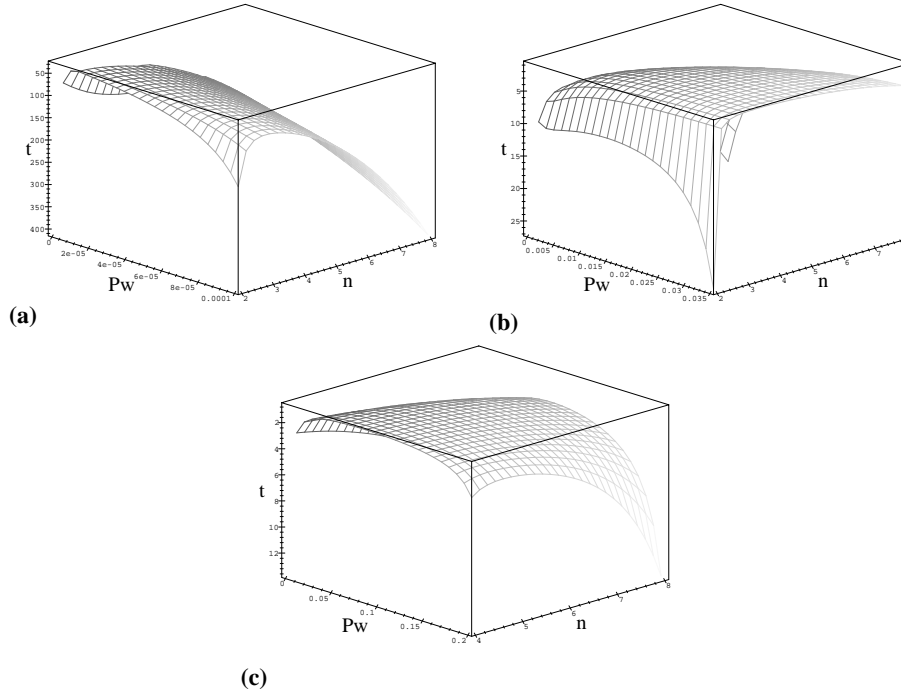


Figure 4. Message latency for various total power budget and power distribution (a) 10 watts (b) 100 watts using small fraction of maximum available power (c) 100 watts with larger fraction of maximum power

tribution. As more power becomes available (10 to 100 watts), the latency exhibits the same response to the fraction of power spent in the wire as it does in the 16 node system (i.e. above a threshold there is insignificant change in latency). This design enables the consideration of a system with three dimensions. On the other hand, if the dimension is not a crucial factor in the design, and latency is the primary concern, then we can build the network by permitting higher power dissipation in the wire and achieving smaller latencies as seen in Figure 4(c). Note that the optimal dimension in this graph is strictly a function of power dissipated in the wire. If P_w is less than 0.05 watts, the optimal dimension is four with a latency of 2.8 μ sec. The optimal P_w is approximately 0.1 watts giving latency of 0.8 μ sec at (n=4). Note that this value of the wire power correlates to the optimal value found for 16 node system (~40% of maximum power).

A preference for higher dimensional network for large systems can be explained as follows. Higher dimensional networks embedded in the plane lead to longer wire lengths for the inter-router physical channels. As available power is increased with larger system size, the negative effect (larger power consumption) of these longer wires is reduced with smaller number of switches to traverse. Moreover, the smaller number of switches that are traversed reduces the power consumption in the switches. The net effect is to favor

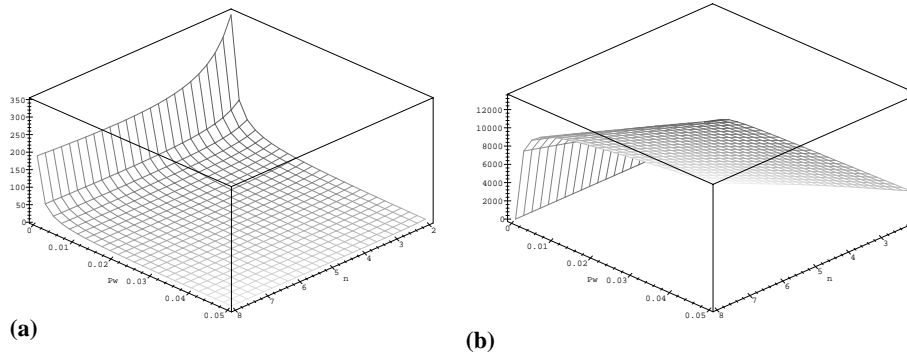


Figure 5. (a) Number of messages, and (b) bisection bandwidth of a network with 256 nodes.

low dimensional networks in low power applications (minimize wire power dissipation) and higher dimensional networks when power is not so constrained.

This paper analyzes network performance under a fixed power budget permitting the bisection bandwidth to vary with changes in dimension and power distribution. The total number of messages in the network and the bisection bandwidth for a 256 node system is shown in Figure 5.

5.0 Concluding Remarks and Directions for Future Research.

This paper has reported our study of the relationships between a fixed power budget and the network topology from the point of view of message latency. Under a fixed total power constraint the network performance is mainly a function of system size, network dimension and the power distribution.

- For a small system size (16 nodes), the optimum latency is achieved at low dimensions (2-3) with at least 40-60% dissipation of total power in the wire and the switch.
- As we increase the size of the network, a “minimum power dissipation” is required to satisfy the traffic requirements. Power budgets that are close to the minimum value result in high latency and very high dimensional networks (7-8).
- Power budgets in excess of minimum result in lower latency, however, the networks only exhibit the optimum performance at medium dimensions (3-4).
- Power distribution between the switch and the wire in the network provide various trade-offs between the hardware and system design.

These models only begin to address the major issues in the power constrained design of multiprocessor interconnection networks. The expectation is that these models can aid the designers of embedded multiprocessor systems in which power is at a premium.

6.0 References

- [1] A. Agarwal, “Limits on interconnection network performance,” *IEEE Transactions on*

- Parallel and Distributed Systems*, October 1991.
- [2] B. W. Char, K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monagan, S. M. Watt, *Maple V*, Waterloo Maple Publishing, Virginia, 1991.
 - [3] W. J. Dally, "Performance Analysis of k-ary n-cube interconnection networks," *IEEE Transactions on Computers*, June 1990.
 - [4] J. Duato, S. Yalamanchili, and L. Ni, *Interconnection Networks, An Engineering Approach*, IEEE Press, 1997.
 - [5] S. L. Scott and G. M Thorson, "The Cray T3E networks: adaptive routing in a high performance 3D torus", Proceedings of Hot Interconnects, August 1996.
 - [6] S. L. Scott and J. R. Goodman, "The Impact of Pipelined Channels on k-ary n-cube networks", *IEEE Transactions on Parallel and Distributed Systems*, vol. 5, no. 1, pp. 2-16, January 1994.
 - [7] D. Basak and D. K. Panda, "Designing Clustered Multiprocessor Systems under Packaging and Technological Advancements," *IEEE Transactions on Parallel and Distributed Systems*, vol. 7, no. 9, pp. 962-978, Sep. 1996.
 - [8] C. S. Patel, S. M. Chai, S. Yalamanchili and D. E. Schimmel, "A Model for Power/Performance Trade-offs in Multiprocessor Network", Proceedings of International Conference on Computer Design, October 1997.
 - [9] S. Abraham and K. Padmanabhan, "Performance of Multicomputer networks under pin-out constraints", *Journal of Parallel and Distributed Computing*, vol. 12, no. 3, pp. 237-248, July 1991.
 - [10] R. E. Kessler and J. L. Schwarzmeier, "CRAY T3D: A new dimension for Cray Research", Proceedings of Compcon, pp. 176-182, Spring 1993.
 - [11] S. L. Scott, "Synchronization and communication in the T3E multiprocessor", Proceedings of the 7th International Conference on Architectural Support for Programming Languages and Operating Systems, pp. 26-36, October 1996.
 - [12] Paragon XP/S Product Overview, Beaverton, OR, Intel Corporation, Supercomputer Systems Division, 1991.