Optimal Network Architectures for Minimizing Average Distance in k-ary n-dimensional Mesh Networks

Matt Grange, Roshan Weerasekera, Dinesh Pamunuwa
Centre for Microsystems, Engineering Department, Faculty of Science and Technology
Lancaster University, Lancaster, United Kingdom, LA1 4YR
{m.grange,r.weerasekera,d.pamunuwa}@lancaster.ac.uk

Axel Jantsch and Awet Yemane Weldezion
Department of Electronics Systems, KTH School of Information and Communication Technologies,
Forum 120, 164 40 Kista, Sweden.
{axel,awye}@kth.se

ABSTRACT

A general expression for the average distance for meshes of any dimension and radix, including unequal radices in different dimensions, valid for any traffic pattern under zero-load condition is formulated rigorously to allow its calculation without network-level simulations. The average distance expression is solved analytically for uniform random traffic and for a set of local random traffic patterns. Hot spot traffic patterns are also considered and the formula is empirically validated by cycle true simulations for uniform random, local, and hot spot traffic. Moreover, a methodology to attain closed-form solutions for other traffic patterns is detailed. Furthermore, the model is applied to guide design decisions. Specifically, we show that the model can predict the optimal 3-D topology for uniform and local traffic patterns. It can also predict the optimal placement of hot spots in the network. The fidelity of the approach in suggesting the correct design choices even for loaded and congested networks is surprising. For those cases we studied empirically it is 100%.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: 3-D IC, Network-on-Chip, modeling techniques, design studies; C.2.1 [**Network Architecture and Design**]: Metrics—*average distance, hotspot, 3-D optimization*

1. INTRODUCTION

Two important metrics of performance for NoCs are latency and throughput, generally functions of network characteristics such as topology, interconnect characteristics, routing scheme and switch architecture, as well as application characteristics primarily defined by traffic pattern. The *average distance* is a NoC performance metric that depends on the topology and the traffic pattern only, under the assumption that the network operates well below its saturation point. We express it as a closed formula comprising a sum over all source-destination node pairs for arbitrary n-dimensional radix-k mesh networks and for arbitrary spatial traffic patterns. A traffic pattern is defined as the packet exchange probability for each source-destination pair. Average distance is an upper bound on the performance for all possible routing, switching, and flow control algorithms. For instance, a routing algorithm is optimal for a given traffic pattern if packets on average do not travel more than the average distance through the network.

We define the *distance* of communication as the minimum number of switch-points or nodes that a packet has to traverse from a source node to a destination node. It is measured in *hops*, where a *hop* is defined as the traversal of a node. The *average distance* $\bar{H}$ is the average distance of all packets in the network under a given traffic pattern, and a function of the dimension and radix in a mesh topology. It is a useful basic metric providing insight into the performance of the overall network.

The starting point in our analysis is the rigorous formulation of an expression for the average distance in $k-$ary $n-$dimensional meshes, including unequal radices in the different dimensions, that is valid for any traffic pattern. We proceed to evaluate this expression for *Uniform Random Traffic* (URT), *Local Random Traffic* (LRT), and *Hot Spot Traffic* (HST) and verify through network simulations that the resulting formula accurately predicts the average latency for unloaded[1] networks. The upshot of this is that the upper bound on performance given by the average latency of a $k-$ary $n-$mesh unsaturated network for any traffic pattern can be estimated from the general model we propose, without running network simulations, which saves both model development time and computation time.

In the case of URT, the general expression can be solved to yield a closed-form expression for the delay, which is proven

[1]We use the terms unloaded, uncongested and unsaturated interchangeably to essentially mean the same thing: the injection rate of packets into the network is low enough to ensure the network is stable.

NOCS'11 May 1-4, 2011 Pittsburgh, PA, USA

to be more exact and more general than other expressions available in the literature. We also propose empirical closed-form solutions for three specific LRT patterns based on the Response Surface Method (RSM) and illustrate the general methodology to obtain closed-form solutions for an arbitrary traffic pattern.

Due to the simplicity of the model, it is particularly useful in finding optimal network architectures to minimize average communication delay under any traffic pattern, including empirically validated traffic models for irregularly sized networks with multiple constraints. For relatively simple traffic patterns such as URT, LRT or empirically-validated models where a closed-form expression can be obtained for the delay, the optimization problem takes a few seconds of computation time. In our results we show solutions for optimal architectures under different traffic patterns and boundary conditions that can be obtained from a few iterations at most with minimal computation time using any general purpose programming language such as Matlab or C, rather than computationally expensive cycle-accurate packet-level simulations. However even for those cases where the traffic model does not allow such a formulation, the general model for average distance we propose has a computational complexity $O(N^2)$ in the worst-case where N is the number of nodes in the network, meaning that a brute-force search to find the optimal network configuration for network sizes up to say a thousand nodes is feasible on a desktop. Notably our model exhibits 100% fidelity for all simulations carried out for three types of traffic (namely URT, LRT, and HST), where the optimum architecture to minimize latency for unloaded networks remains optimal even under congestion up to the point of saturation.

The main contribution of this paper is in providing a generally valid model for the average distance and accompanying analysis that provides insight into network behavior under common traffic loads and constraints.

2. RELATED WORK

The analytic formulæ for average distance provided in the literature either cover only a special case, or the assumptions made are not fully explained leading to misunderstanding. Agarwal [1] gives

$$\bar{H} = \frac{n}{3}\left(k - \frac{1}{k}\right) \tag{1}$$

as the average distance in an n-dimensional mesh with radix k. The formula assumes that k is the same in all dimensions and no derivation or further motivation is provided.

Liu et al. [12] provide a formula for a $k_1 \times k_2$ 2-D mesh:

$$\bar{H} = \frac{1}{3}\left(k_1 - \frac{1}{k_1}\right) + \frac{1}{3}\left(k_2 - \frac{1}{k_2}\right) \tag{2}$$

Here the case with different radices is covered but only for two dimensions. The derivation of the formula is somewhat unclear because it contains an approximation step, that replaces $k_1k_2 - 1$ by k_1k_2 without motivation or explanation. It seems that the authors intended this formula only to be an approximation. They exclude the self-traffic case (when a node is allowed to send packets to itself) which is in contradiction to Agarwal's assumptions.

Dally and Towles [3] offer

$$\bar{H} = \begin{cases} \frac{nk}{3} & k \text{ even} \\ n(\frac{k}{3} - \frac{1}{3k}) & k \text{ odd} \end{cases} \tag{3}$$

for $k+1$ nodes in each dimension. This expression is in contradiction to both Liu's and Agarwal's formulæ. However, closer inspection shows that, under the corrected assumption of radix k (rather than $k+1$), the odd case formula is identical to Agarwal's formula and the even case is an upper bound of $\bar{H}$, approaching the true value asymptotically for large k.

Holsmark [7] devotes appendix I of his Licentiate thesis to the clarification of the average distance. He concludes with

$$\bar{H} = \frac{k_1 + k_2}{3} \tag{4}$$

for a $k_1 \times k_2$ mesh assuming no self-traffic. Agarwal and Dally et al. include the self-traffic case. Taking this into account, it turns out that Holsmark's formula is consistent with Agarwal's expressions, but he covers only 1-D and 2-D meshes.

We provide a derivation for the average distance in n-dimensional meshes with the general case of unequal radices along the different dimensions. This is important as many practical on-chip networks serving a few tens or hundreds of cores in multiprocessor systems are often irregular, and the number of nodes along the x, y, and z dimensions in a 3-D system for example, are seldom equal. We show how this formula is exactly correct, and is a generalization of all the models mentioned above.

Koohi et al. [11] present abstract performance models in a spirit similar to our's. They propose power and throughput models for uniform, local, hotspot, and first matrix transpose traffic models. Their approach is more empirical since they use simulation results as starting point and analyze the effect of combinations of different traffic models to derive comprehensive throughput and power models for mixed traffic patterns. In contrast, our approach focuses on distance and latency, deriving an analytical formula which is further validated through simulation. Primarily we focus on the unloaded case but we show that design decisions based on minimum latency provided by our model also hold true in loaded networks. While Koohi et al. consider only 2-D networks, we have special interest in 3-D topologies.

In a seminal paper from 1990 Dally [4] studied the communication performance of k-ary n-dimensional tori, a similar class of topologies that we cover (meshes rather than tori). The paper analyzes average latency in networks with different dimensions (mostly between 2-20) under different cost constraints for routers and wires, ranging from "wires are free and infinitely fast" to "limited wires per router" and a linear delay wire model. The paper identifies the optimal dimension constrained by the cost and wire model assumed. It almost exclusively deals with uniform random traffic, except a short section where hot spot traffic is briefly discussed. Our work makes more specific assumptions on wire cost and wire delays based on realistic implementations for on-chip planar interconnects and in a 3-D stack. But we cover general traffic patterns and show how the model facilitates topological exploration and hot spot placement.

A significant body of research comprises the develop of network performance models for packet latency in congested networks. Even in the narrower scope of Networks-on-Chip the published literature on this topic is substantial. Most of the work (e.g. [13, 5, 14, 10, 9, 6, 16, 15]) make very specific assumptions about routing (mostly deterministic, dimension order routing) and switching (mostly wormhole switching). The majority focuses on average delay [5, 14, 10, 9, 6] while some work targets worst-case delay [16, 15]. Sometimes even more specific assumptions are made such as single flit buffers [6] or one dedicated virtual channel for each flow [8].

In contrast with all this and similar work, we do not offer a delay model under congestion or for a specific switch architecture, but we start with an analytical formula for average distance which is valid for all k-ary n-dimensional meshes and for any traffic pattern, but abstracts entirely from routing and switching techniques. We then derive average distance formulas for specific traffic patterns and we illustrate the usefulness of this abstract, ideal formulae. In particular, we use the expression for average distance to investigate optimal architectures for minimizing delay when the link delays are not necessarily equal.

3. CALCULATING AVERAGE DISTANCE

First, we derive a general average distance formula for k-ary n-dimensional networks and for arbitrary traffic patterns. Then, in section 3.1, we derive a closed form solution for Uniform Random Traffic. This is straight forward and the result is consistent with earlier published formulas. In section 3.2 we derive a closed form formula for a specific type of Local Random Traffic, which is a more complicated derivation. In section 3.3 we introduce HotSpot Traffic but do not provide a closed formula due to the intractable dependency on the precise location of the hotspots. But later on in section 4 we show how an optimal placement of hotspot nodes can be found by minimizing the average distance.

The average distance of a network is the ratio between the total distance that the packets emitted by all switches travel and the total number of packets emitted. In a network with N nodes, the total distance traveled by the packets emitted by all switches is:

$$D_t = \sum_{A \in N} \sum_{B \in N} p_{A,B} \times d_{A,B}, \tag{5}$$

where A and B are any given nodes, $d_{A,B}$ is the Manhattan distance between the node A and node B, and $p_{A,B}$ is the probability that a packet is sent from node A to node B. $p_{A,B}$ defines the traffic pattern and can be an arbitrary function. E.g. for local random traffic $p_{A,B}$ is a function of the distance $d_{A,B}$. This equation is a general formulation for any k-ary n-dimensional network.

The total number of paths N_p traversed by the emitted packets considered here is:

$$N_p = \sum_{A \in N} \sum_{B \in N} p_{A,B} \tag{6}$$

Hence the Average Distance D_{avg} is given by:

$$D_{avg} = \frac{D_t}{N_p} = \frac{\sum_{A \in N} \sum_{B \in N} p_{A,B} \times d_{A,B}}{\sum_{A \in N} \sum_{B \in N} p_{A,B}} \tag{7}$$

In order to demonstrate the formula (7), we calculate the average distance over a single dimension by considering a 1-D array shown in Fig. 1. Each node connects to every

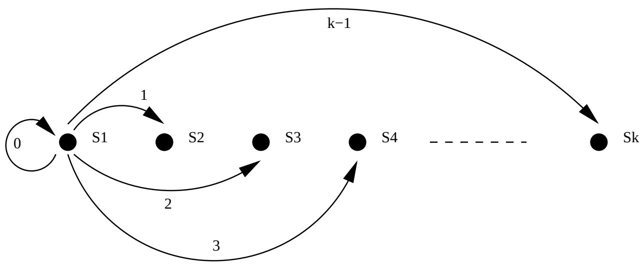


Figure 1: A $1 \times k$ Mesh Network (dimension=1 and radix=k). The distances from switch S_1 to all other switches including itself is shown.

other node, and each such path has an associated distance. Shown in the figure are the distances $h_{i,j}$ from the first node (S_1) to the other nodes in the array. The average distance for a k-ary 1-D array is:

$$\bar{H}_{1\times k} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{k} p_{i,j} \times |i-j|}{\sum_{i=1}^{k} \sum_{j=1}^{k} p_{i,j}}. \tag{8}$$

3.1 Uniform Random Traffic

For *Uniform Random Traffic*, each node generates the same number of packets uniformly distributed over time, and all destination nodes are equally likely. Hence the probabilities are the same for all source-destination pairs, leading to $p_{i,j} = p_{\text{URT}}$ where p_{URT} is a constant. Therefore, (8) can be simplified thus:

$$\bar{H}_{\text{URT}_{1\times k}} = \frac{p_{\text{URT}} \sum_{i=1}^{k} \sum_{j=1}^{k} |i-j|}{p_{\text{URT}} \sum_{i=1}^{k} \sum_{i=1}^{k} 1} = \frac{\frac{k^3}{3} - \frac{k}{3}}{k^2} \tag{9}$$

leading to

$$\bar{H}_{\text{URT}_{1\times k}} = \frac{k}{3} - \frac{1}{3k}. \tag{10}$$

In the case of URT, since the probability is not a function of distance that the packets travel, the average dimension for an n-D network can be calculated by adding the average distance for each dimension. Therefore for an n-D array, the average distance for URT can be expressed using the following closed-form equation:

$$\begin{aligned}\bar{H}_{\text{URT}} &= \sum_{i=1}^{n} \bar{H}_{\text{URT}_{1\times k_i}} \\ &= \frac{k_1}{3} - \frac{1}{3k_1} + \frac{k_2}{3} - \frac{1}{3k_2} + \cdots + \frac{k_n}{3} - \frac{1}{3k_n}\end{aligned} \tag{11}$$

When the radix along each dimension is identical, it is true that $k_1 = k_2 = k_3 \cdots = k$, and (11) reverts to (1) and the odd case of (3).

For $n = 2$ equation (11) simplifies to (2). Although Liu et al. [12] intended an approximation, they ended up providing an exact formula for the case that includes a node sending traffic to itself. Even though in their setup they excluded the self-traffic case, the approximation they included as a corrective measure resulted in exactly the right formula including self-traffic.

Finally it can be seen that taking the different assumptions into account (i.e. with and without self-traffic) equation (11) reverts to equation (4) for $n = 2$ as follows. Since the average distance is computed by dividing the sum of all distances by the number of paths, we correct for these different assumptions by multiplying with the number of paths in the self-traffic case ($(k_1k_2)^2$) and dividing by the number of paths in the case with no self-traffic ($(k_1k_2)^2 - k_1k_2$). Since the distance for self-traffic is 0, the sum of all distances does not differ in the two cases. Setting $n = 2$ and applying this correction to (11) results in:

$$\frac{1}{3}\left(k_1 - \frac{1}{k1} + k_2 - \frac{1}{k_2}\right)\left(\frac{(k_1k_2)^2}{(k_1k_2)^2 - k_1k_2}\right) = \frac{1}{3}\left(k_1 + k_2\right)$$

which turns out to be identical to Holsmark's equation (4).

Table 1: Comparison of Average Distance Calculated by the formula (7) and Simulations

Netw. Size	Prob. Model	Avg. Distance		%Error
		Formula	Simulation	
5x5x5	URT	4.8300	4.813	0.35
6x6x6	URT	5.8600	5.888	0.47
7x7x7	URT	6.8772	6.971	1.36
8x8x8	URT	7.8900	7.931	0.52
9x9x9	URT	8.9000	8.976	0.85
10x10x10	URT	9.9090	9.894	0.15
4x8x16	URT	9.9090	10.008	0.99
5x5x5	LRT:α=1.0	3.7900	3.81	0.53
6x6x6	LRT:α=1.0	4.5900	4.555	0.94
7x7x7	LRT:α=1.0	5.3900	5.418	0.52
8x8x8	LRT:α=1.0	6.1900	6.146	0.71
9x9x9	LRT:α=1.0	7.0000	6.969	0.44
10x10x10	LRT:α=1.0	7.8060	7.855	0.62
5x5x5	LRT:α=1.5	3.1800	3.163	0.53
7x7x7	LRT:α=1.5	4.4781	4.498	0.44
4x8x16	LRT:α=1.5	5.3757	5.301	1.38

3.2 Local Random Traffic

For local traffic, a variety of probabilistic models can be considered, with the probability that a given node being the destination is some inverse function of the distance from the source: $p_{A,B} = \frac{1}{\lambda_A f(d_{A,B})}$ where λ_A is the normalization constant defined by $\sum_{B \in N} p_{A,B} = 1$, and $d_{A,B}$ is the distance between nodes A and B.

For $f(d_{A,B}) = d^{\alpha}_{A,B}$ Table 1 compares the average distance estimated by (7) and RTL network simulations, for different traffic models and network sizes. The RTL simulator employs bufferless 5-port (2-D) and 7-port (3-D) switches, where one port serves the independent packet generating resources. The switches are bufferless and the routing decisions are based on an address minimizing "hot potato" deflection algorithm. The simulation invokes a configurable mesh to instantiate any size network in 2 or 3 dimensions. In order to comply with the RTL cycle-accurate simulations, we excluded self-traffic from formula (7) by considering $A \neq B$. To validate the model, simulations are performed at low packet injection rates to ensure an unloaded network. The results show a good agreement between the formula and the simulation results, where the discrepancy between the model and simulation is a result of stochastic deviation in the generation of packet destinations and rounding errors. Further, for URT the closed-form solution as given in (11) for 3-D and the general expression (7) show exactly the same result.

Closed-Form Formula for LRT: To provide a closed form solution for non-URT traffic, we propose an empirical equation for average distance of a 3-D NoC, using the Response Surface Method (RSM) [2], which is a collection of mathematical and statistical techniques useful for the modeling and analysis of problems in which a response of interest is influenced by several variables. This same methodology we demonstrate for our local traffic pattern can be applied to any custom traffic pattern. The second-order approximation function with three variables using RSM is defined as:

$$y = b_0 + \sum_{i=1}^{m} b_i x_i^2 + \sum_{i=1}^{m} b_{ii} x_i^2 + \sum_{i=1, i<j}^{m} \sum_{j=1}^{m} b_{ij} x_i x_j + \epsilon, \quad (12)$$

where x_i and x_j are the design variables, b_0, b_i, b_{ij} are called regression coefficients, ϵ is the error, and m is the number of variables. For various values of x_i, $i = 1, \ldots, m$ the dependent variable y is found from experiment. The relationship between a set of independent variables and the response y defined by the regression coefficients is determined using the method of least squares. In general, (12) can be written in matrix form:

$$\mathbf{Y} = \mathbf{bX} + \mathbf{E}, \quad (13)$$

where $\mathbf{Y}$ is defined to be a matrix of measured values (of size $p \times 1$), $\mathbf{X}$ to be a matrix of independent variables, $\mathbf{b}$ to be the regression coefficient matrix (of size $p \times 1$), and $\mathbf{E}$ to be the error matrix (of size $p \times 1$). Then, the solution of (13) is:

$$\mathbf{b} = (\mathbf{X^T X})^{-1} \mathbf{X^T Y}, \quad (14)$$

which are the regression coefficients for (12).

Using the RSM model, for a 3D NoC with size $k_x \times k_y \times k_z$, the average distance can be expressed as follows:

$$\begin{aligned} H_{fit} = b_0 + b_x k_x &+ b_y k_y + b_z k_z + b_{xx} k_x^2 + b_{yy} k_y^2 + b_{zz} k_z^2 \\ &+ b_{xy} k_x k_y + b_{xz} k_x k_z + b_{yz} k_y k_z \quad (15) \end{aligned}$$

For example, regression coefficients for local traffic model LRT:$\alpha = 0.5, 1.0, 1.5$ are shown in Table 2. In some cases when either of K_x, K_y or k_z is 2 the maximum error in the functional form can be as high 13%, however the error for all cases on average is shown to be lower than 1%.

3.3 Hotspot Traffic (HST)

To represent less than ideal networks, we also consider that some nodes in the network will attract more traffic than others. In this case the probability of sending packets originating from a particular node to hotspots is higher than that for non-hotspot nodes.If the fraction of packets generated at any given node that have a hotspot destination is p_{HST}, the fraction of packets addressed to non-hotspots will be $(1 - p_{\text{HST}})$. Packets for hotspots and non-hotspot can also be assigned using URT and LRT. In this analysis we assumed that 80% of packets from one node is uniformly distributed among two hotspots while the rest is uniformly distributed among all the non-hotspots. We consider different strategies for hotspot placements within the network and use our

Table 2: **Coefficients for the closed-form equation for average distance for a network of size $k_x \times k_y \times k_z$, where $k_x, k_y \in 2, \ldots 10$ and $kz \in 2, \ldots 30$.**

b	LRT:$\alpha = 0.5$			LRT:$\alpha = 1.0$			LRT:$\alpha = 1.5$		
	γ=1	γ=0.5	γ=0.25	γ=1	γ=0.5	γ=0.25	γ=1	γ=0.5	γ=0.25
b_0	-0.4915	-0.4455	-0.4224	-0.4272	-0.3841	-0.3626	-0.1569	-0.1716	-0.1789
b_x	0.3556	0.3472	0.3430	0.3281	0.3087	0.2990	0.2680	0.2497	0.2406
b_y	0.3556	0.3472	0.3430	0.3281	0.3087	0.2990	0.2680	0.2497	0.2406
b_z	0.2804	0.1423	0.0732	0.2136	0.1123	0.0617	0.1493	0.0801	0.0455
b_{xx}	-0.0061	-0.0054	-0.0051	-0.0101	-0.0084	-0.0076	-0.0122	-0.0101	-0.0091
b_{yy}	-0.0061	-0.0054	-0.0051	-0.0101	-0.0084	-0.0076	-0.0122	-0.0101	-0.0091
b_{zz}	-0.0013	-0.0008	-0.0005	-0.0022	-0.0014	-0.0010	-0.0027	-0.0016	-0.0011
b_{xy}	0.0022	0.0027	0.0029	0.0055	0.0061	0.0064	0.0099	0.0096	0.0095
b_{xz}	0.0030	0.0019	0.0014	0.0055	0.0036	0.0026	0.0066	0.0044	0.0032
b_{yz}	0.0030	0.0019	0.0014	0.0055	0.0036	0.0026	0.0066	0.0044	0.0032
%Avg. Error	0.49	0.44	0.44	0.91	0.82	0.81	1.35	1.21	1.19

Table 3: Optimum Network Sizes for different traffic models under planar link clock speeds, γ, of 0.5 and 0.25 times the 3-D link clock. Uniform vertical and horizontal clock speeds through a network intuitively perform at their optimum latency in symmetrical configurations of N×N×N. δ is the ratio between the average distance for optimum network size and the average distance for cubic solution.

N	URT				LRT:$\alpha = 0.5$				LRT:$\alpha = 1.0$			
	$\gamma = 0.5$		$\gamma = 0.25$		$\gamma = 0.5$		$\gamma = 0.25$		$\gamma = 0.5$		$\gamma = 0.5$	
	N_{opt}	δ	N_{opt}	δ	N_{opt}	δ	N_{opt}	δ	N_{opt}	δ	N_{opt}	δ
27	2x2x7	0.96	2x2x7	0.78	2x2x7	0.94	2x2x7	0.78	2x2x7	0.92	2x2x7	0.78
64	2x4x8	0.98	2x3x11	0.82	2x4x8	0.97	2x2x16	0.79	2x3x11	0.95	2x2x16	0.74
125	4x4x8	0.95	3x3x14	0.82	4x4x8	0.95	3x3x14	0.80	3x3x14	0.94	3x2x21	0.77
216	4x5x11	0.96	3x4x18	0.83	4x5x11	0.95	3x4x18	0.82	4x5x11	0.94	3x3x24	0.78
343	5x5x14	0.97	4x4x22	0.84	5x5x14	0.96	4x4x22	0.82	5x5x14	0.94	3x4x29	0.79
512	5x7x15	0.97	5x5x21	0.84	5x7x15	0.96	4x5x26	0.83	5x7x15	0.95	4x5x26	0.80
729	7x7x15	0.95	5x6x25	0.84	7x7x15	0.95	5x5x30	0.83	7x7x15	0.95	5x5x30	0.81
1000	7x8x18	0.95	6x6x28	0.84	7x8x18	0.95	6x6x28	0.83	7x8x18	0.94	6x6x28	0.81

model to demonstrate optimum placements, which can be a significant factor in reducing congestion and communication bottlenecks and improving overall system performance.

4. APPLICATIONS OF THE MODEL

In design space exploration it is often of interest to find the topology that minimizes delay under various constraints imposed by technological, physical and system-level requirements. For the unsaturated case, delay is a straightforward function of the average distance, and such constrained optimization problems can be solved accurately with the proposed analytical model by treating it as the objective cost function. In this section we demonstrate how our model for average distance can be used to provide performance comparisons between networks and optimize the topological configuration of nodes in 2-D and 3-D meshes for any traffic pattern.

A 3-D mesh is an increasingly common topology with the advent of 3-D Integrated Circuits (IC). Equal radices in each dimension, translating to a cube with the same number of nodes or switches however, is an unrealistic arrangement; rather, the number of nodes in each dimension are likely to be different, especially in the vertical direction. For example, recent work [17] has shown that the lower physical delay associated with the vertical interconnects in a 3-D stacked IC can enable higher data rates in the vertical dimension by clocking die-to-die links at greater frequencies than the horizontal dimensions. This is largely due to the relatively lower parasitics of through silicon vias (TSV) used to connect vertical die layers as compared to long planar wires used on a 2-D IC.

4.1 Optimization of Network Topology

The most efficient network topology for minimizing latency under different vertical and horizontal clocking constraints can be found by solving the constrained optimization problem of minimizing

$$D_{3D} = \frac{\sum_{A \in N} \sum_{B \in N} p_{A,B} \times (|x_A - x_B| + |y_A - y_B| + \gamma |z_A - z_B|)}{\sum_{A \in N} \sum_{B \in N} p_{A,B}} \tag{16}$$

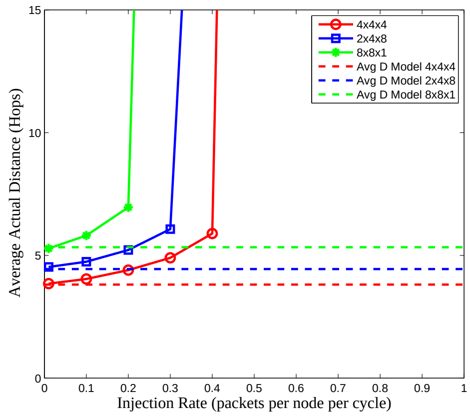
subject to $k_x \times k_y \times k_z = N$ where N is the total number of nodes. Here $d_{A,B}$ has been obtained by multiplying the distance in each dimension by the corresponding clock period, under the assumption that the periods in the x and y dimensions are equal and normalized to 1, and that the period in the z dimension is $\frac{1}{\gamma}$ shorter, where $0 < \gamma \leq 1$.

The optimum network size under the given constraint is found using an extensive brute-force search algorithm, and the solutions for different N are shown in Table 3. As described in section 3.2 closed-form equations for different traffic models were obtained considering three different γ values. These equations are used for to find the optimal network sizes which considerably reduces the computational load when compared to the general expression in equation

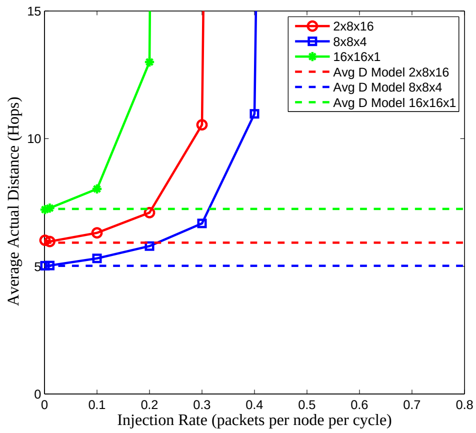
(7)

In order to assess the validity of the topological solution given by our model for the uncongested and congested cases, we use RTL cycle-accurate simulations to measure average distance for networks under varying loads. Figure 2(a) plots the average distance in hops as the packet injection rate per cycle increases. Intuitively, in a network with uniformly clocked horizontal and vertical links, the optimum configuration for minimum communication distance under uniform traffic is always a symmetrical network. The unloaded average distance for three 64-node network topologies from our model is shown as a dashed base-line. This is the absolute minimum unloaded average latency achievable in that particular architecture. In this case a 64-node 3-D mesh will be best organized as a 4×4×4 network for minimum average latency. Our model matches the baseline simulation result to within a 1% error when the injection rate is sufficiently low as to negate any contention issues.

As we increase the injection rate, contention becomes more prevalent and the average latency increases as a result of non-ideal routing conditions and link bandwidth limitations. Despite loading the networks to the point of saturation (where the network bandwidth is exceeded by the packet injection rate), Figure 2(a) demonstrates that the topology predicted as optimum by our traffic pattern-based average distance model is still valid for loaded networks as a network topology deemed as optimal in the unloaded case will be sub-optimal under congestion only if the lines intersect or cross. Figure 2(b) plots the same result for a local traffic pattern in a larger 256-node network.



(a) URT



(b) LRT

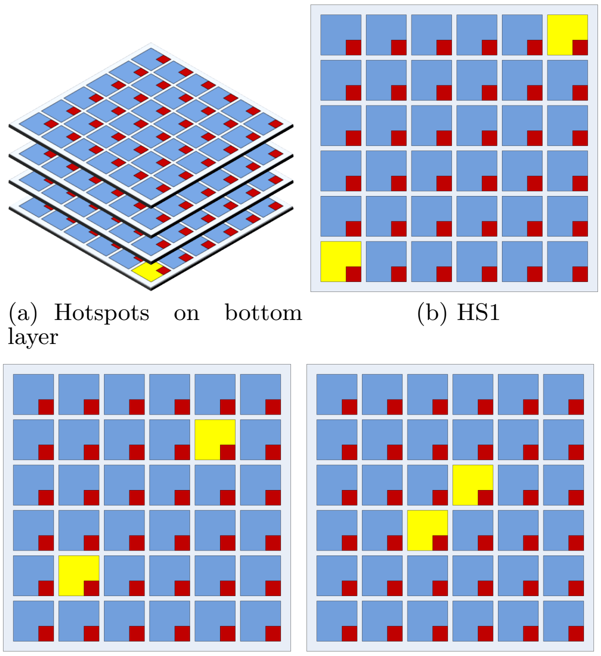
Figure 2: Average actual distance for increasing injection rates under (a) Uniform Random Traffic for a 64-node network and (b) Local traffic for a 256-node network. The model prediction for optimal network configuration (given the minimum average distance) is consistent under increasing injection rate.

4.2 Hotspot Node Placement Optimization

In a 3-D IC with TSVs providing the vertical switch-to-switch links, it is likely that the off-chip I/O will only be able to serve the outermost dies. Ball-Grid Array (BGA) or other area-array packages allow the I/O pads to be dispersed across the entire die surface, meaning that any outer node in the network could interface with off-chip devices. In a high-throughput multi-tier 3-D NoC, it is further likely that mesh nodes on the bottom layer, which have direct access to the BGA I/O will face higher traffic due to the on/off-chip communication. To model this effect we place nodes which attract high-traffic from the network (called hotspot nodes) on the bottom layer of a 3-D NoC in several configurations. The placement of the hotspot network nodes is crucial to the overall performance of the system (for example, placing hotspot nodes on the edge of a network will limit the surrounding link bandwidth due to the unused switch links).

The model we developed in section 3 can be used to estimate the average unloaded latency and the optimum placement (giving minimum average distance) of hotspots in a network. We place two hotspots on the bottom layer in three configurations as depicted in Figure 3(b): (HS1) hotspots are placed on the edge of the network in opposing corners, Figure 3(c) (HS2) hotspots are placed in opposing corners one node removed from the edge, and Figure 3(d) (HS3) the hotspots are diagonally adjacent at the center of the network. The probability that any node in the network will send a packet to either of the hotspots is 80%, where the remaining 20% of the generated packets have a uniform probability to the other non-hotspot network nodes.

Similar to Figure 2, we have used our analyical model to predict the optimum placement of hotspot nodes, given by the minimum average latency for several network configurations and found that in the RTL simulations even under heavily congested networks with hotspot traffic, the optimum node placement to achieve the lowest average packet latency is consistent with the model prediction at zero load. Figure 4 plots the growth of latency with injection rate for the three hotspot placement schemes in a 7×7×7 network. Despite the minimal difference in unloaded latency provided by our model between the HS2 and HS3 placement schemes, the simulation results indicate no crossover points even at heavy loads. These results demonstrate the model's usefulness in predicting network optimal topologies and node placement without computationally expensive packet-level simulations. Further to this, the model can quickly predict the optimum placement and configuration for hotspot traffic under different vertical and horizontal clocking schemes. This efficient design space exploration is enabled by the availability of a closed-form comparison metric.



(a) Hotspots on bottom layer (b) HS1

(c) HS2 (d) HS3

Figure 3: Example of hotspot node placements HS1, HS2 and HS3 for the bottom layer of a 6×6×6 network

To assess the fidelity of our model, we have conducted RTL simulation sweeps for unloaded and loaded networks for twelve different mesh configurations of sizes ranging from 8 to 1000 nodes each with the three hotspot placement schemes shown by Figure 3 and six packet injection rates of between 0.0001 and 0.8 packets-per-node-per-cycle (see Table 4 for a range of results). We consider the model to correctly predict the optimal network topology under loading if the latency curve with increasing injection rate corresponding to the zero-load optimal configuration is always below the latency curve corresponding to any other topology. Formally, fidelity can be defined as meeting the following condition: $\bar{H}_{NW_1}(IR) < \bar{H}_{NW_i}(IR)$ for any injection rate (IR) where $\bar{H}_{NW_1}(0) < \bar{H}_{NW_i}(0)$, $\bar{H}_{NW_i}$ being the average distance for the network configuration nw_i at zero load. The latter condition essentially defines $\bar{H}_{NW_1}$ as the optimal network under zero load. If the data points cross-over at any point with a stable network or $\bar{H}_{NW_1}(IR) > \bar{H}_{NW_i}(IR)$, the model prediction is deemed to have failed.

For the simulations we have conducted under local, uniform and hotspot traffic, we found that the optimum placement predicted by the model is valid throughout all the network configurations when the injection rate is below the threshold of saturation[2] for the network.

These exhaustive findings report that although the model cannot predict the exact latency of a topology under congestion (which depends on low-level architectural features such as switch design, buffering and routing algorithms among others) it can accurately and repeatedly determine the optimum network topology and placement of hotspots under

[2]We consider a network saturated when the packet injection rate exceeds the available link bandwidth.

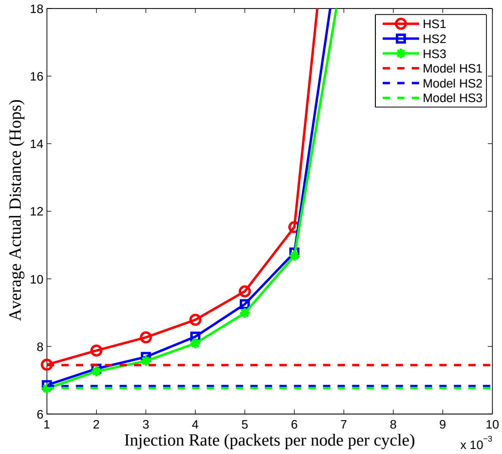


Figure 4: Hotspot simulation with increasing injection rate. The model prediction for optimal hotspot placement holds true for loaded and unloaded networks even for a relatively minor difference in minimum average latency.

any network load. This assumption holds true for the simulations we have conducted with our traffic models, however we have yet to test it under other traffic scenarios such as burst mode or worm-hole routing which may reduce the fidelity of the model.

5. CONCLUSIONS

The average traveling distance of packets in a NoC is an abstract performance metric of the network topology for a given traffic pattern and does not consider routing and switching schemes. We have rigorously formulated and presented an analytic expression for the average distance in k-ary n-dimensional meshes for any traffic pattern as a summation over the nodes in the network. It is shown to be a generalization of previously presented models and is over 99% accurate when compared against cycle-accurate RTL simulations for uniform-random, local and hotspot traffic, with any errors being traceable to stochastic deviations and rounding errors.

The formula deals with the unloaded case and does not constitute a delay model for congested networks. However it can be used as a metric for choosing between network topologies for optimal performance, with solutions that are valid even under loading. To demonstrate its usage we have applied formulæderived for URT and LRT to guide high level design decisions. In particular, we demonstrate our models ability to optimize the configuration of 2-D and 3-D symmetric and asymmetric networks given the constraints of latency, traffic pattern and different clock speeds over the on-chip and 3-D interconnect. Also, we predict the optimal placement of hot spot nodes in a network with HST based on the zero-load average distance formula and show that the placement remains optimal for loaded networks. The predictive power of our model (its fidelity) is a surprising 100% for the cases we have studied empirically.

However further studies are necessary to understand the scope and limitations of this method. In particular we have

Table 4: Simulation sweeps to assess the fidelity of the model. The optimum hotspot placement calculated by our model is shown to be valid for any traffic congestion in stable networks for over 100 data points.

HSP	4×4×4	6×6×6			7×7×7			8×8×8			10×10×10		
	HS1	HS1	HS2	HS3	HS1	HS2	HS3	HS1	HS2	HS3	HS1	HS2	HS3
Model	4.396	7.1877	6.1285	5.606	7.4485	6.8289	6.7692	9.987	8.7897	7.598	12.79	11.509	9.5958
0.0003	4.467	7.29	6.23	5.66	7.48	6.85	6.75	9.99	8.82	7.68	12.89	11.64	9.61
0.001	4.49	7.34	6.26	5.98	7.57	7.13	7.06	10.41	9.2	8.06	13.46	12.28	10.42
0.003	4.63	7.61	6.62	6.04	8.27	7.69	7.57	11.70	10.69	9.51	115.5	117.93	120.96
0.005	4.67	8.14	7.17	6.6	9.63	9.25	8.99	38.31	36.09	38.7	301.2	309.3	308.6
0.007	4.72	8.73	8.05	7.53	25.46	20.73	19.29	194.36	202.74	209.6	388.5	391.03	385.2
0.009	4.82	10.29	10.02	9.68	177.42	180.84	178.82	288.55	290.42	299.3	414.2	417.78	409.3

not studied time variant traffic patterns and deterministic routing. First, the fidelity may be lower for bursty traffic because bursts may temporarily clog parts of the network without overloading the network as a whole in the long term. Such effects are not predicted by our model but may affect different network topologies very differently. The second limitation in our empirical evidence is the fact that we use adaptive routing in our simulations. Deterministic routing has much less capacity to balance load over the entire network. Individual links can easily become bottlenecks even though the network is only modestly loaded. Again, such local, temporary overload may not be predicted well by our zero-load model. Further experiments are required to study these effects.

In summary, although further studies are need to understand the full scope of the model, its power in predicting network performance and its usefulness as a metric in high level topology exploration is very promising given the simplicity of the basic formula, being simply the average geometric traveling distance of packets.

6. REFERENCES

[1] A. Agarwal. Limits on interconnection network performance. *IEEE Trans. on Parallel and Distributed Systems*, 2(4):398–412, 1991.

[2] G. E. P. Box and K. B. Wilson. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(1):1–45, 1951.

[3] W. Dally and B. Towles. *Principles and practices of interconnection networks*. Morgan Kaufmann, 2004.

[4] W. J. Dally. Performance analysis of k-ary n-cube interconnection networks. *IEEE Trans. on Computers*, 39(6):775–785, June 1990.

[5] S. Foroutan, Y. Thonnart, R. Hersemeule, and A. Jerraya. An analytical method for evaluating network-on-chip performance. In *Proc. Design, Automation Test in Europe Conference*, pages 1629 –1632, March 2010.

[6] Z. Guz, I. Walter, E. Bolotin, I. Cidon, R. Ginosar, and A. Kolodny. Network delays and link capacities in application-specific wormhole NoCs. *VLSI Design*, 2007.

[7] R. Holsmark. Deadlock free routing in mesh networks on chip with regions. Licentiate thesis, Department of Computer and Information Science, Linköpings Universitet, 2009.

[8] F. Jafari, Z. Lu, A. Jantsch, and M. Yaghmaee. Buffer optimization in network-on-chip through flow regulation. *IEEE Trans. on Computer Aided Design of Integrated Circuits and Systems*, 29(12):1973–1986, December 2010.

[9] A. Khonsari, M. Ould-Khaoua, and J. Ferguson. A general analytical model of adaptive wormhole routing in k-ary n-cube interconnection networks. *SIMULATION SERIES*, 35:547–554, 2003.

[10] A. Kiasari, D. Rahmati, H. Sarbazi-Azad, and S. Hessabi. A Markovian performance model for networks-on-chip. In *Proc. Euromicro Conf. on Parallel, Distributed and Network-Based Processing*, pages 157 –164, February 2008.

[11] S. Koohi, M. Mirza-Aghatabar, S. Hessabi, and M. Pedrani. High-level modeling approach for analyzing the effects of traffic models on power and throughput in mesh-based NoCs. In *IEEE/ACM Int. Conf. on VLSI Design*, pages 415 –420, January 2008.

[12] H. Liu, W. Lin, and Y. Song. An efficient processor partitioning and thread mapping strategy for mesh-connected multiprocessor systems. In *Proc. ACM symposium on Applied computing*, page 412. ACM, 1997.

[13] Z. Lu, M. Millberg, A. Jantsch, A. Bruce, P. Van der Wolf, and T. Henriksson. Flow regulation for on-chip communication. In *Proc. of Design Automation and Test Europe Conf.*, April 2009.

[14] U. Ogras, P. Bogdan, and R. Marculescu. An analytical approach for network-on-chip performance analysis. *IEEE Trans. on Comp.-Aided Design of Integrated Circuits and Systems,*, 29(12):2001 –2013, December 2010.

[15] Y. Qian, Z. Lu, and W. Dou. Worst-case flit and packet delay bounds in wormhole networks on chip. *IEICE Transactions*, 92-A(12):3211–3220, 2009.

[16] Y. Qian, Z. Lu, and W. Dou. Analysis of worst-case delay bounds for on-chip packet-switching networks. *IEEE Trans. on Comp.-Aided Design of Integrated Circuits and Systems,*, 29(5):802 –815, May 2010.

[17] A. Weldezion, M. Grange, D. Pamunuwa, Z. Lu, A. Jantsch, R. Weerasekera, and H. Tenhunen. Scalability of network-on-chip communication architecture for 3-D meshes. In *Proceedings of the 2009 3rd ACM/IEEE International Symposium on Networks-on-Chip*, pages 114–123. IEEE Computer Society, 2009.

Acknowledgments

European Union research funding under grant FP7-ICT-215030 (ELITE) is gratefully acknowledged.