

The Effect of Network Topology and Channel Labels on the Performance of Label-based Routing Algorithms

Reza Moraveji^{1,2}, Hamid Sarbazi-Azad^{1,3}, Arash Tavakkol^{1,3}

¹IPM School of Computer Science,

²Department of Electrical and Computer Engineering, Shahid Beheshti University,

³Department of Computer Engineering, Sharif University of Technology
{Moraveji_r, Azad, Arasht}@ipm.ir

Abstract. Since network of workstations (NOWs) support for both regular and irregular topologies, designing an efficient deadlock-free routing is a point of concern. In this paper, we take a step toward the goal by developing three novel deadlock-free routing algorithms and covering three known previously reported ones [2, 3] in the body of new family called label-based routing algorithms for irregular topologies. Moreover, by simulating and comparing the newly and traditional proposed routings, it is shown that the performance of this family highly depends on the network topology and channel labeling process. Furthermore, a number of valuable results about the performance of these routing algorithms is extracted from simulations.

Keywords: Label-based routing algorithm, irregular network, network of workstation, performance evaluation.

1. Introduction

In recent years, cluster-based irregular networks such as networks of workstations (NOWs) have emerged as one of the cost-effective alternatives for traditional regular parallel computers. In such systems, an irregular high-speed network is often required in order to provide the wiring flexibility needed in network and also design of scalable systems with incremental expansion capability [4, 7].

Without a careful design for routing scheme of NOWs, deadlock may happen in these types of networks [5, 6]. Since the topology of NOWs is not predefined, designing and applying deadlock-free routing algorithms are usually done without any pre-assumption about the network topology. Therefore, the major problem of these networks is the complexity of designing a general deadlock-free routing algorithm.

The main purpose of this work, section 2, is to take a step in this direction by initially developing some deadlock-free routing schemes in the body of new proposed family of routings called label-based routing algorithms for irregular topologies.

Moreover, evaluating the performance of label-based routings in irregular networks under realistic conditions is another major concern. To this end, extensive simulation experiments have been conducted in section 3.

Section 4 concludes the paper and outlines some directions for future works in this line of research.

2. Label-Based Routing Algorithms

In order for a routing algorithm to be deadlock-free, cyclic buffer dependencies between messages and physical channels they allocate, must not occur. When the approach of labeling is used for generating deadlock-free routing algorithms, first the given topology has to be prepared for implementing the label-based routing algorithm. Let us briefly describe the way in which the topology is labeled and also the method by which the related routing schemes generated.

The main idea of label-based routing algorithms is to classify network channels by assigning predefined labels, then grouping the labeled channels in the way that there is no cyclic dependency between each group. These groups are referred to as zones in [1]. Afterwards, the generated zones are ordered in a sequence such that when a message passes through the needed channels in the zones (regarding the sequence of the zones), the sequence guaranties the message to reach its destination.

2.1. Fundamental Concepts of Graph Labeling, Deadlock-Free Zones and Routing Algorithms

The first step in generating a label-based routing algorithm is graph labeling. Since we plan to make a comparison between the previously reported routing algorithms and the newly proposed ones, in this paper we use the reported graph labeling in [1-3].

As the starting point, a spanning tree (based on BFS¹ graph traversal) is formed on the given irregular network as the base of labeling process. Nodes and channels are labeled in two stages as follows:

First stage: Nodes are labeled in ascending order regarding to spanning tree formation and according to their distances from the root of spanning tree. A channel that faces toward the lower node label is called '1' and the channel that goes away from lower node label is called '0' (figure 1).

Second stage: Subsequently, the second stage of labeling is applied to the graph in the case that an increasing number is assigned to each node in the order that nodes are visited by pre-order tree traversal. Channels are labeled using the policy of first stage (figure 1).

¹ Breath first search

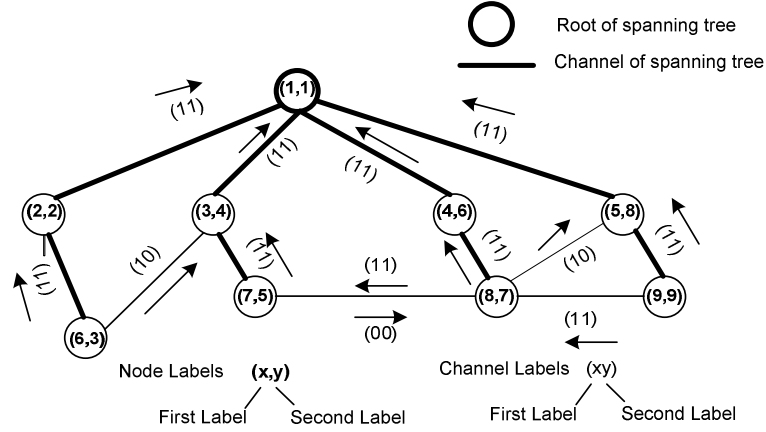


Figure 1. Node and channel labeling

Therefore, each channel is assigned two different labels and it is possible to think of a channel label as a compound label containing two distinct labels. It is obvious that there may be at most four possible channel labels for a given irregular topology. These channel labels are:

$$(11), (10), (01), (00)$$

As a result, a single '0' transition (channel) from node A to B means that the corresponding label of node A is lower than B , and a single '1' transition from node A to B represents that the node A has higher corresponding label than B . Therefore, when both labels are brought into account as a compound label, we have the following outcomes:

$$\begin{aligned} A(a_0a_1) \xrightarrow{11} B(b_0b_1) &\Rightarrow (a_0 < b_0, a_1 < b_1) \\ A(a_0a_1) \xrightarrow{10} B(b_0b_1) &\Rightarrow (a_0 < b_0, a_1 > b_1) \\ A(a_0a_1) \xrightarrow{01} B(b_0b_1) &\Rightarrow (a_0 > b_0, a_1 < b_1) \\ A(a_0a_1) \xrightarrow{00} B(b_0b_1) &\Rightarrow (a_0 > b_0, a_1 > b_1) \end{aligned}$$

Note that (a_0a_1) and (b_0b_1) are node labels.

The second step in generating a label-based routing is to group the channels such that there is no cyclic dependency between the channels of the same group. Since there are several ways to group the channels, it is possible to generate various deadlock-free groups (zones) and in turn, different deadlock-free routing algorithms.

As mentioned in label-based routing [1] there is a predefined ordering to travel between channel groups and a message cannot use channel labels belong to a previously traversed group. Consequently there is no cyclic dependency between channels of different groups. As indicated in figure 2, assume that message 1 holds channel (A, B) labeled as (10) , and it requests the use of channel (C, D) labeled as (11) and message 2 holds (C, D) and requests the use of (A, B) . Let's assume that these two channel labels $\{(10), (11)\}$ are in a same group. Now, we should consider

the situation in which message 1 and 2 just use channel labels of the mentioned group, $\{(10), (11)\}$. For message 1 we have,

$$A(a_0a_1) \xrightarrow{10} B(b_0b_1) \Rightarrow (a_0 > b_0, a_1 < b_1) \quad C(c_0c_1) \xrightarrow{11} D(d_0d_1) \Rightarrow (c_0 > d_0, c_1 < d_1)$$

$$a_0 < b_0 \dots < c_0 < d_0 \Rightarrow a_0 < d_0$$

Therefore, if message 2 wants to make a request for (A, B) while holding (C, D) , it should cross other channels such as $(00), (01)$ and it contradicts the mentioned group ordering traversal [1].

Thus, it is possible to put $(10), (11)$ in one group.

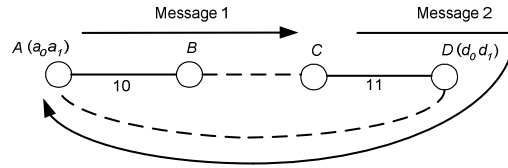


Figure 2. Cyclic dependency between (A, B) and (C, D)

The following corollary defines a general rule for creating deadlock-free zones or groups of channels without cyclic dependency.

Corollary: there is no cyclic dependency between the channels X and Y if and only if they satisfy the following equation:

$$(x_0 \odot y_0) \oplus (x_1 \odot y_1) = 1$$

where \odot is the bitwise *XNOR* operator and \oplus represents the bitwise *OR* operator. It should be noted that labels $X(x_0x_1)$ and $Y(y_0y_1)$ are channel labels not node labels and four possible channel labels were introduced formerly. Possible deadlock-free zones are listed as follows:

$$(11,10), (11,01), (10,00), (01,00)$$

The third and last step in generating a label-based routing algorithm is to order deadlock-free zones in a sequence that guaranties connectivity for the routing algorithm.

Theorem: The path between any pair of nodes is guaranteed by selecting sequence of channel labels that the first (second) bit of them is '1' followed by sequence of channel labels that the first (second) bit of them is '0' [1].

Proof: When a message chooses a channel with a compound label which contains at least one '1', it gets closer to the root node of the spanning tree. In the worst case situation the respective message reaches the root node and it is obvious that from the root node there is at least one path to each other nodes in the spanning tree (whole network). Therefore, when the channel labels are ordered in a sequence of '1's followed by '0's whether in terms of first or second bit of the channel label, there is at least one path between each pair of source and destination.

Now considering the generated deadlock-free zones and possible sequences, the following label-based routing algorithms can be defined:

1. $R_1: (11,10) \rightarrow (01,00)$ up/down routing
2. $R_2: (11,01) \rightarrow (10,00)$ left/right routing
3. $R_3: (11) \rightarrow (01,00) \rightarrow (10)$
4. $R_4: (11) \rightarrow (10,00) \rightarrow (01)$ l-turn routing
5. $R_5: (10) \rightarrow (11,01) \rightarrow (00)$
6. $R_6: (01) \rightarrow (11,10) \rightarrow (00)$

3. Empirical Performance Evaluation

The main performance metric of a NOW is the average message latency (average amount of time it takes a message to completely reach its destination). In a thorough analysis, the mentioned performance metric of the label-based routings is analyzed under different working conditions considering different irregular topologies and different spanning trees. As you will see, some interesting points are derived from the results of the analyses that were not reported or referred in the previous works on the performance evaluation of routing algorithms in irregular networks.

Analysis of this kind can be conducted through results obtained from a real implementation of the network. But a cost effective alternative is to use a simulation of the system.

3.1. Simulator

To evaluate the functionality of irregular networks under different conditions, a discrete-event simulator [10] has been developed that mimics the behavior of described label-based routing algorithms at flit level. Input data (irregular topology) to the simulator is specified in the form of adjacency matrix. Also, the spanning tree assigned to the network can be both determined by user or automatically by one of the famous BFS or DFS (with a predetermined heuristic) algorithms.

3.2. The Effect of Network Topology

When a comparison made between the performances of two or more routing algorithms, using the same working conditions such as number of virtual channels, message lengths, and traffic patterns, it is always expected that one (or more) of the compared routing algorithms shows better performance than the others. Generally, by changing the conditions for all routing algorithms the order of routing performances does not change. For example, the performance of *XY* routing [7] (ignoring the simplicity of implementation) in comparison with the west-first routing [7] (turn model) under the same working conditions is worse. By changing the topology on which the respective routing algorithms applied from $Mesh_{3 \times 3}$ to $Mesh_{5 \times 5}$, the performance of west-first routing still remains better, since the latter routing algorithm always provides more adaptivity than the former one. **As a result, in most**

cases a fair comparison provides a definite order of performances of the compared instances. As we will see in this section, this is not true for label-based routing algorithms (compared instances). The performance of the label-based routing algorithm highly depends on the topology to which the routing is applied. Therefore, it is not possible to make a general sequence of the performance of the six aforementioned label-based routing algorithms.

Another design parameter that has a strong influence on the performance of the label-based routing algorithm is the degree of irregularity of the network topology; that is, the performance of this family is involved in the variance of node degrees of the topology.

In order to show the above characteristics in irregular networks, a comparative performance evaluation is presented in the results of the figure 3 where the average message latency is plotted against the traffic generation rate. The analyzed irregular networks are as follows:

$G_1(16, 48)$, $G_2(36, 124)$, $G_3(64, 240)$, $G_4(64, 240)$, G_5 : Mesh_{8x8}, and $G_6(100, 364)$.

Network topology is the first parameter that should be considered, while choosing the best label-based routing algorithm. Let's look at the simulation results obtained from different network sizes and network topologies. As can be seen, the sequences of routing performances are totally different from one topology to another. The following list presents the sequences of the routing performances for different topologies:

- G_1 : R₆, R₄, R₃, R₅, R₂, R₁
- G_2 : R₁, R₃, R₆, R₂, R₄, R₅
- G_3 : R₂, R₁, R₆, R₃, R₄, R₅
- G_4 : R₆, R₃, R₁, R₂, R₄, R₅
- G_5 : R₃, R₄, R₁, R₂, R₅, R₆
- G_6 : R₃, R₆, R₂, R₁, R₄, R₅

To see how different the sequences of routing performances are, consider the sequences in G_3 and G_4 . Excluding R₄ and R₅, the sequence in G_3 is the reverse of that in G_4 although these networks contain the same number of nodes and even channels. The only difference between these two networks is the way of connecting nodes or network topology. Another interesting example that exhibits the effect of network topology on the performance of label-based routing is R₁, which shows totally inconsistent behavior in G_1 and G_2 . R₁ is the best routing algorithm in G_2 while is the worst one in G_1 . We have the same scenario for R₆, in G_4 and G_5 (mesh_{8x8})

As a consequence, it is sagacious first to specify the network topology; then, choose the routing which shows the best performance on the chosen topology.

Another interesting effect is the degree of irregularity (variance of node degrees) of topology. It is evident from figure 3(e) (mesh_{8x8}) that the message latencies and generation rates for which saturations occur for the six routing algorithms are nearly the same. The reason is that although the mesh topology is not completely regular (a network is regular when all the nodes have the same degree), all of the internal nodes have the same degree of four so that the variance of node degrees goes down. The identical result can be seen for G_1 .

As a result, when the regularity of the topology decreases, or the variance of node degrees diminishes, the performance of the label-based routing algorithms are closely the same (figures 3(a) and 3(e)). It should be noted that when the network size decreases (like G_1), the probability that the variance of node degrees become smaller increases (but this may not be true in all cases).

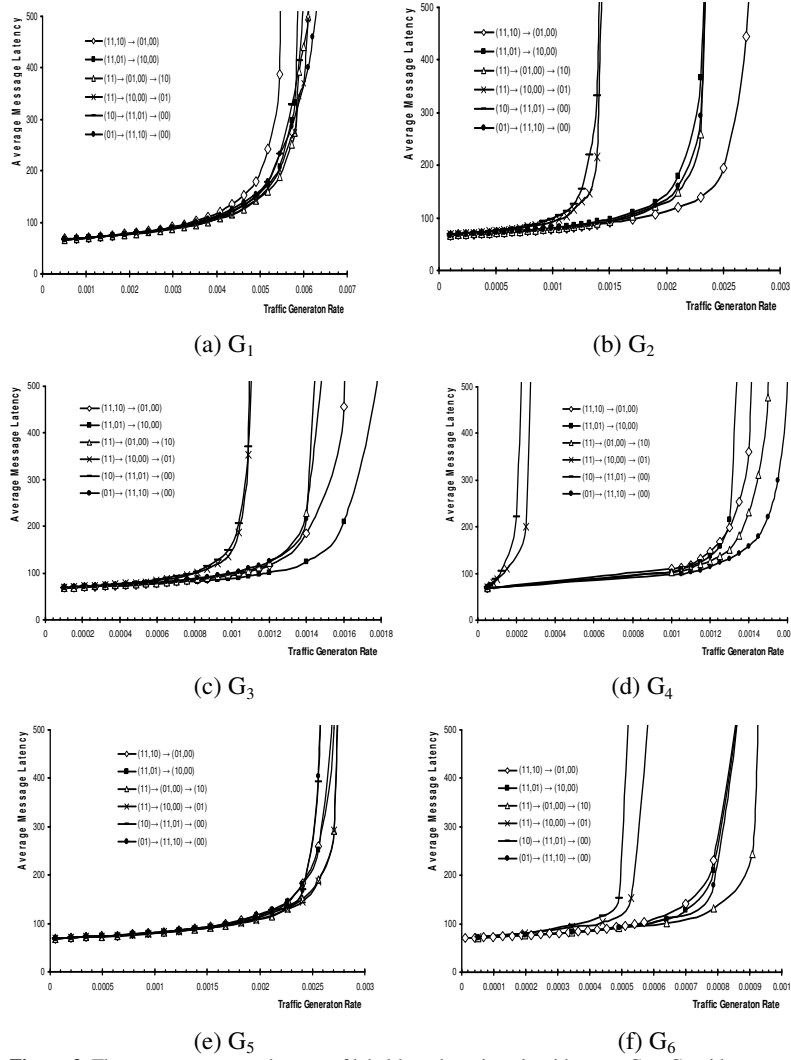


Figure 3. The average message latency of label-based routing algorithms on $G_1 - G_6$ with a message length of 64 flits

3.3. The Effect of Spanning Tree Construction

In the previous section the effect of network topology (network size) on the performance of the six label-based routing algorithms was discussed. According to the numerous presented results, it was shown that the performance of this family of routing algorithms highly depends on the network topology. Going further through the structural details of the six label-based routing algorithms leads us to analyze the effect of forming different spanning trees created in terms of different root nodes.

The structure of a label-based routing algorithm is determined by two parameters which are number and order of zones (channel labels).

Now assume that in an arbitrary network there are two minimal paths between two nodes as follows²:

Path 1: 11 → 10 → 11 → 00 → 01 → 10 → 00

Path 2: 11 → 10 → 00 → 00 → 01 → 01 → 01

Among the six routing algorithms, only R_1 can direct the message through the two existing paths. As a result, if the sequence of channel labels in a routing algorithm, like R_1 , results in more possible minimal paths, the average distance of the network will decrease. Moreover, the sequence of channel labels in a path is determined by graph labeling. Therefore, a general consequence can be made as follows:

The performance of a label-based routing algorithm depends on the way of graph labeling.

Labeling of a network is determined by the spanning tree and the spanning tree is formed based on a root node. Thus, in an irregular topology, number of different ways that a graph can be labeled is equal to the number of different spanning trees can be formed on the graph.

The performance of a label-based routing algorithm depends on the channel labels and all the topological parameters that have direct or indirect effect on the channel labels.

Figure 4 shows the effect of spanning tree root on the average message latency of G_1 - G_6 using R_1 , R_3 , R_4 , and R_6 . It can be observed that, changing the root of the spanning tree and in turn channel labels initially causes a substantial difference in the network latencies and the generation rates for which saturation occurs. The following generated spanning trees are based on the BFS traversal of the graph.

²A sequence of channel labels marks a path in the network.

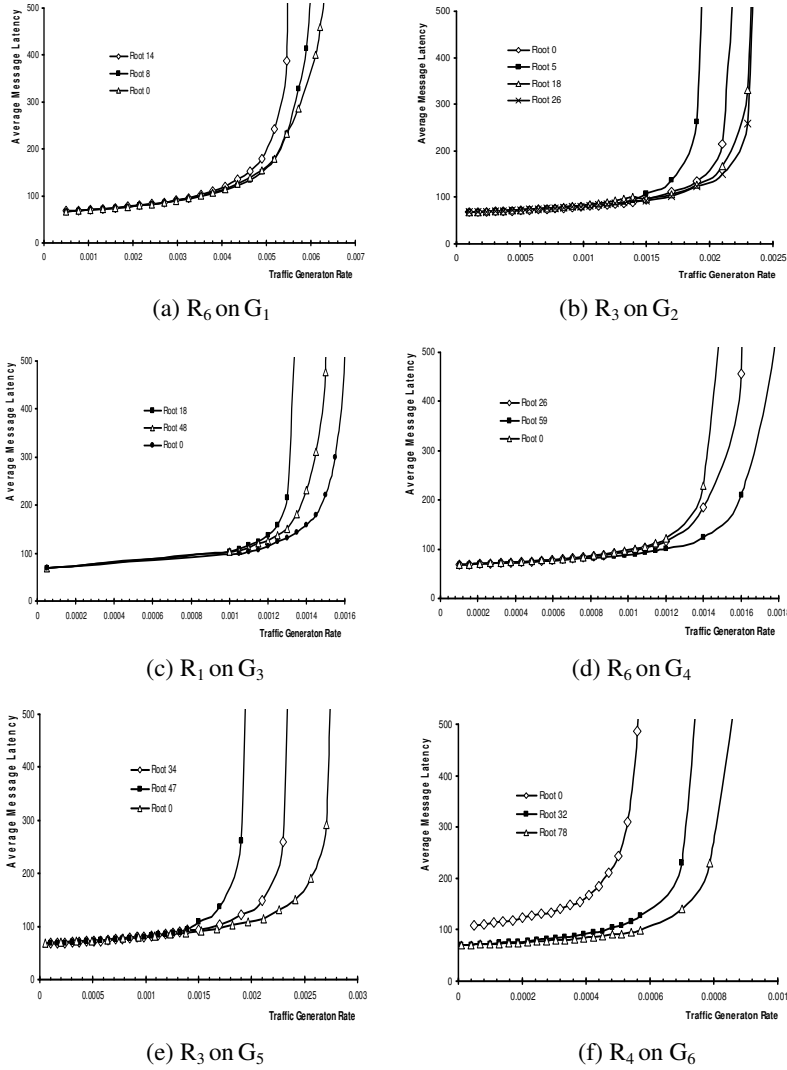


Figure 4. The average message latency of label-based routing algorithms on $G_1 - G_6$ with different spanning tree roots and message length of 64 flits

4. Conclusion

First, in addition to cover three previously reported routings, we proposed three novel deadlock-free routing algorithms in the body of new family called label-based routing algorithms. Second, the work has confronted the task of evaluating the performance of the mentioned family in irregular networks under realistic conditions.

Third, by analyzing the experimental results, we revealed that the network topology, channel labels, and other topological parameters related to channel labels have great influence on the performance of label-based routing algorithms. Therefore, it is not possible to make a general sequence of the performance of the six aforementioned label-based routing algorithms.

Regarding to previous work which exhibits the reaction of routing algorithms in regular networks in case of analytical models [8, 9], further research in this line may consider developing such models for irregular networks. Moreover, investigating a general routing methodology for irregular networks and proposing some heuristics to compute the best spanning tree on the given topology can be considered for future work.

References

- [1] Moraveji, R., Sarbazi-azad, H.: A General Methodology of Routing in Irregular Networks. Technical Report, IPM School of Computer Science, Tehran, Iran (2007)
- [2] Schroeder M.D., et al.: Autonet: a High-speed, Self configuring Local Area Network Using Point-to-point Links. *J. Selected Areas in Communication*. 9, 1318--1335 (1991)
- [3] Koibuchi, M., Funahashi, A., Jouraku, A., Amano, H.: L-Turn Routing: An Adaptive Routing in Irregular Networks. *International Parallel Processing Conference*. pp. 383-392 (2001)
- [4] Sancho, J.C., Robles, A., Duato, J.: An Effective Methodology to Improve the Performance of the Up*/Down* Routing Algorithm. *IEEE Transaction on Parallel and Distributed Systems*. 15, 740--745 (2004)
- [5] Lysne, O. Skeie, T. Reinemo, S., Theiss, I.: Layered Routing in Irregular Networks. *IEEE Transaction on Parallel and Distributed Systems*. 17, 51--65 (2006)
- [6] Puente, J.A., Gregorio, F., Vallejo, R., Beivide, C. Izu.: High-performance Adaptive Routing for Networks with Arbitrary Topology. *J. System Architecture*. 52, 345--358 (2006)
- [7] Duato, S. J., Yalamanchili, L. Ni.: *Interconnection Networks: An Engineering Approach*. IEEE Computer Society Press (2003)
- [8] Moraveji, R., Sarbazi-Azad, H., Nayebi, A., Navi, K.: Performance Modeling of Wormhole Hypermeshes under Hot-spot Traffic. 2nd International Computer Science Symposium in Russia (LNCS), pp. 290-320 (2007)
- [9] Moraveji, R., Sarbazi-Azad, H., Abbaspour, M.: Optimal Placement of Frequently Accessed IPs in Mesh NoCs. 12th Asia-Pacific Computer Systems Architecture Conference (LNCS). pp. 126-138 (2007)
- [10] Nayebi, A. Meraji, S. Shamaei, A. Sarbazi-Azad, H.: Xmulator: A Listener-Based Integrated Simulation Platform for Interconnection Networks. First Asia International Conference on Modeling and Simulation. pp. 128-132 (2007)