On the Performance of a Class of Multihop Shuffle Networks

Seung-Woo Seo, Paul R. Prucnal, Hisashi Kobayashi, and James B. Lim *
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544
e-mail address: seo@ee.princeton.edu

Abstract

A generalization of the well-known shuffle network is proposed for multihop lightwave communication. In the classical definition of a shuffle network, i.e., $N = kp^k$ where N is the number of nodes and k is the number of stages with base p, the realizable values of N are very discrete and many of the intermediate values of N are not realizable. In this paper, we propose a new definition of a shuffle network as $N = nk$ where n is the number of nodes per stage with base p. Based on this new definition, we divide the shuffle networks into two classes: extra-stage and reduced-stage. Study results can be used to determine an optimal network topology when given a value of N.

1 Introduction

Multihop networks with wavelength division multiplexing have been considered as a possible way for high data-rate communication (over 100 *Gbits/sec*) [1]. A bunch of nodes in multihop networks are connected by optical fiber, which has a tremendous bandwidth of tens of terabits per second. In telecommunication, each of the nodes may be a user or a station; in computer communication, each of the nodes can be a workstation or a supercomputer that is physically distributed but sharing some sources such as memories or I/O.

Regular two-connected network topologies, such as the Manhattan Street Network (MSN)[2] and the shuffle network[1], have been considered for possible multihop networks. A MSN has $N = n^2$ nodes, where n is even, that are arranged in n rows and n columns connected by unidirectional links in a grid with adjacent rows and columns having opposite directions. Efficient routing algorithms for the MSN are also developed (for example, in [4]). However, the proposed routing techniques are somewhat complicated and are not suitable for optical networks, which require information to be processed as simply as possible.

On the other hand, the classical shuffle network has a multistage topology, which has been extensively investigated in computer and communication communities. An $N(= kp^k)$ node shuffle network is characterized by two parameters p and k, and is represented as a (p, k) shuffle network, where p is the base of the network and k is the number of columns. In a (p, k) shuffle network, p^k nodes are linearly arranged in a column, and two adjacent columns are connected in a perfect shuffle by unidirectional links. The last column is wrapped around to the first column in a cylindrical manner, and packets can recirculate through the network until they arrive at their destinations. Having all of the nodes arranged in a single column is known as a single-stage shuffle exchange network. However, previous research shows that the classical shuffle network outperforms the single-stage shuffle network (for example, in [5]). Both of the MSN and the shuffle network are regular networks, which means that from every node, there exists a complete spanning tree. In a regular network, every node sees the network in an identical way, and the spanning tree rooted at each node is the same for all nodes.

In this paper, we investigate the effect of different network topologies on network performance. More specifically, we propose a new definition of a shuffle network, which differs from the conventional one. In our new definition of a shuffle network, we remove the tight relationship (i.e., $N = kp^k$ for a given N) between the number of stages (k) and the number of nodes per stage (n). The network topology becomes more flexible by allowing two independent variables whose product is N, i.e., $N = kn$. The new definition enables various distinct values of N to be realized into a shuffle network. We show that different network topologies provide us with a different network throughput; by selecting a network topology properly, performance may be improved considerably. Throughout this paper, we concentrate on shuffle networks with $p = 2$ for the comparison and simulation of various networks, while we derive general formulas for arbitrary p.

The remainder of the paper is organized as follows.

*Department of Computer Science and Engineering, The Pennsylvania State University, University Park, PA 16802.

0-7803-2486-2/95 $4.00 © 1995 IEEE

In Section 2, we define a generalized shuffle network by comparing with a classical shuffle network. The performance analysis of various network topologies is followed in Section 3. In that section, we derive an accurate model and an approximate model of the expected number of hops for different cases. Analytical models are verified through computer simulations in Section 4. Results are summarized and remarks are offered in Section 5.

2 Generalized Shuffle Networks

One important property of a classical shuffle network is its regular structure. However, this requirement always fixes a shuffle network as having k stages with p^k nodes at each stage. In this classical definition of a shuffle network, feasible numbers of nodes are so discrete many values of N cannot be realized. For example, assuming $p = 2$, if $k = 1$ then $N = 2$, if $k = 2$ then $N = 8$, if $k = 3$ then $N = 24$ and so on. As can be seen, these numbers are very discrete; thus, other given numbers, e.g., N=16, 32, $\cdots$, may not be realized into a shuffle network.

To solve this problem, we propose a generalization of the shuffle network's structure. In our new definition, we eliminate the tight relationship between k and n, instead allowing them to behave as two independent variables whose product equals N. A generalized shuffle network is defined as follows.

Definition 2.1: A generalized shuffle network (GSN), is constructed with $N = kn$ nodes where k is the number of stages and n is the number of nodes per stage. Here, n is the power of p. At each stage, n nodes are linearly arranged, and two adjacent columns are connected in a perfect shuffle by unidirectional links.

The stages in the GSN are numbered as 1, 2, $\cdots$, k from the leftmost stage to the rightmost stage. Each node at stage i, $1 \leq i \leq k-1$, has p links directed to p nodes at the next stage $i+1$. If the nodes are numbered from 0 to $n-1$ at each stage, a node j at stage i is connected to nodes, j', $j'+1$, $\cdots$, $j'+p-1$ at stage $i+1$, where $j' = (j \bmod p^{k'-1})p$, in a perfect shuffle[3]. As in the classical shuffle network, the last stage of GSN is connected to the first stage in a wrapped-around manner.

3 Performance Analysis of the GSN

In this section, we analyze the performance of the GSN. Since a GSN is characterized basically by two parameters k and n, we will focus on determining the relationship between the two parameters. Eventually, we will find out the optimal values of k and n to achieve the highest throughput of the network. It is noted that in the remainder of the paper we do not assume a specific flow control algorithm, but the results can be used with any algorithm including store-and-forward and deflection routing.

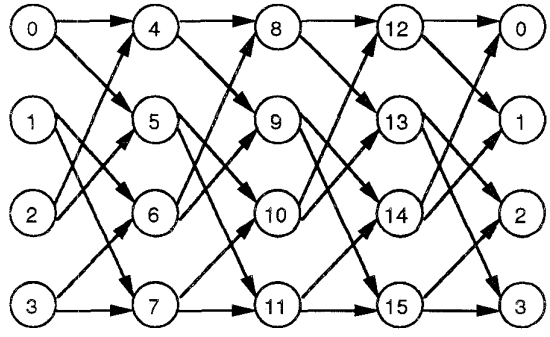

Figure 1: (16, 4, 4, 2) extra-stage shuffle network

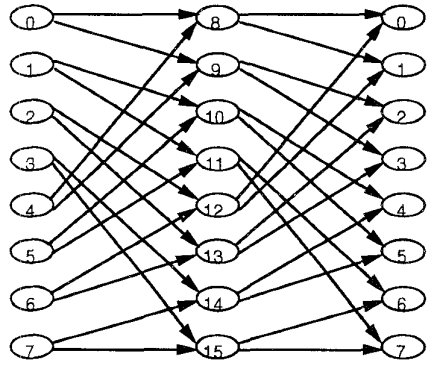

Figure 2: (16, 2, 8, 2) reduced-stage shuffle network

Let the variable k' denote the number of stages in a conventional shuffle network with n $(= p^{k'})$ nodes per stage. However, note that in a GSN, k' is a virtual number of stages, while k is the actual number of stages. In the following analysis, we consider two cases separately, i.e., $k' \leq k$ and $k' > k$. If $k \geq k'$, we call the networks as **extra-stage shuffle networks**, and if $k < k'$, we call them as **reduced-stage shuffle networks**. If $k = k'$, it is the conventional shuffle network, and is a special case of the extra-stage shuffle networks. If a node is met more than once along a path from source to destination, it is *overlapping*. In both GSN cases, a source node begins to access every node in a stage after k' hops, if overlapping is not considered. We call this point a *saturation point*. After the saturation point in extra-stage shuffle networks, the source node encounters the same number of intermediate nodes (n) at each stage at each hop, until it meets previously traversed nodes. However, in reduced-stage shuffle networks, the overlapping begins to appear before the saturation point, which makes the

exact analysis of reduced-stage shuffle networks very difficult. The extra-stage shuffle network (16, 4, 4, 2) is shown in Fig. 1. Fig. 2 shows an example of the reduced-stage shuffle networks (16, 2, 8, 2) shuffle network for $N = 16$. We now analyze the performance of a GSN in detail. In each case, we derive the expected number of hops in various network topologies.

3.1 In Case $k' \leq k$: Extra-stage Shuffle Networks

In extra-stage shuffle networks, network performance parameters are easily calculated due to the structural regularity. The number of nodes reachable after h hops from a source is given in Table 1. It is easy to see that k' is the saturation point; after k' hops, the number of accessible nodes is fixed to n, and decreases after k hops. As a result, the expected number of hops is given by:

$$E\{hops\} = \frac{1}{kn-1}\left[\sum_{j=1}^{k'-1} jp^j + n\sum_{j=0}^{k-k'-1}(j+k')\right] + \frac{1}{kn-1}\left[\sum_{j=0}^{k'-1}(j+k)(n-p^j)\right].$$

The summations are simply calculated in a closed form by substituting k' with $log_p n$:

$$E\{hops\} = \frac{1}{kn-1}\left[\frac{n}{2}k^2 - k\frac{n-1}{p-1} - \frac{nk}{2} + nklog_p n\right]. \quad (1)$$

By substituting kn with N and by eliminating k, the optimal value of n to make the function minimal can be obtained:

$$\frac{d}{dn}E\{hops\} = \frac{1}{N-1}\left[-\frac{N^2}{2n^2} - \frac{N}{(p-1)n^2} + \frac{N}{n}log_p e\right]. \quad (2)$$

Letting $\frac{d}{dn}E\{hops\} = 0$, the minimal value of n for $k' \leq k$ can be given as:

$$n = \langle \ln p(\frac{1}{p-1} + \frac{N}{2})\rangle \quad (3)$$

where $\langle x\rangle$ means the nearest integer of power of 2 to x. It is noted that n should always be selected to satisfy $k' \leq k$; in other words, $nlog_p n \leq N$.

Furthermore, especially if $n = p^k$, the above equation can be simplified as follows:

$$E\{hops\} = \frac{N(3k-1)(p-1) - 2k(p^k-1)}{2(p-1)(N-1)} \quad (4)$$

Table 1: Number of nodes reachable at each hop for $k' \leq k$

h	# of nodes
1	p
2	p^2
$\vdots$	$\vdots$
$k'-1$	$p^{k'-1}$
k'	$p^{k'} = n$
$k'+1$	n
$\vdots$	$\vdots$
$k-1$	n
k	$n-1$
$k+1$	$n-p$
$\vdots$	$\vdots$
$k+k'-1$	$n-p^{k'-1}$

which is the same result as in [3]. In Table 2, some calculations of the expected number of hops for different n's in case $k' \leq k$ (without $\star$) are shown for $p = 2$. The results in Table 2 indicate that as the number of stages increases, the expected number of hops also increases. This phenomenon matches a simple observation of networks, because more hops are needed as the number of stages increases. This is due to the saturation effect of the network. As a guideline in selecting a network topology for $k' \leq k$, it is always preferable to keep k as small as possible.

3.2 In Case $k' > k$: Reduced-stage Shuffle Networks

In a reduced-stage network case, each of the source nodes meets a different number of intermediate nodes at each hop as it progresses to its destination. This means reduced-stage networks are not regular, and the spanning trees at each node are not identical. Since $k' > k$, some nodes (stages) may be visited more than once until a packet arrives at the saturation point (k'). Generally, a packet goes through five different parts, i.e., $1 \leq h \leq k-1$, $k \leq h \leq 2k-1$, $2k \leq h \leq \lfloor\frac{k'}{k}\rfloor k-1$, $\lfloor\frac{k'}{k}\rfloor k \leq h \leq k'$ and $k'+1 \leq h \leq k'+k-1$. The first part, $1 \leq h \leq k-1$, is trivial because there is no overlapping until $h = k$; at the h-th hop, the packet meets p^h nodes.

The second part, $k \leq h \leq 2k-1$, reflects the overlapping phenomenon between the stages, $1 \leq h \leq k-1$ and $k \leq h \leq 2k-1$. In this part, overlapping between stages begins to appear and continues until the destination stage. At the k-th stage, one is a previously visited node at the first stage as a source node, while $p^k - 1$ nodes are newly visited. Since every source node

encounters a different number of intermediate nodes along the way to the destination, one possible way is to obtain an expected number of nodes at each hop. However, even this becomes very complex if the network size increases. Hence, in order to make the analysis tractable, we derive an approximate model for the expected number of hops for $h \geq 2k$, while we derive an exact model up to $h = 2k - 1$. The expected number of nodes to visit at the k-the hop can be derived as $(p^k - 1)\frac{k}{N}p^k + p^k\frac{n-p^k}{N}k$ for all N nodes. More specifically, among n source nodes at the first stage, p^k nodes meet $(p^k - p^0)$ nodes at the $(k+1)$-th stage after k hops, and $(n-p^k)$ nodes at the first stage meet p^k nodes at the $(k+1)$-th stage in a regular manner.

The preciseness of modeling implies that all of the probabilities for an expected number of nodes at each hop should be calculated accurately in order to obtain the expected number of hops. However, as $\lfloor\frac{k'}{k}\rfloor$ increases, obtaining the exact values of the probabilities becomes more and more complex, because as a packet propagates further to the following parts, the previous parts affect the present part simultaneously. (Furthermore, for example, the third part, $2k \leq h \leq \lfloor\frac{k'}{k}\rfloor - 1$, may be composed of multiple stages of k.) One possible way to an approximate modeling is to assume that all of the probabilities are equal. For example, at $2k \leq h \leq 3k - 1$, the probabilities are set to be $\frac{1}{4}$. Similarly, at $3k \leq h \leq 4k - 1$, the probabilities are set to be $\frac{1}{8}$, and so on.

At $\lfloor\frac{k'}{k}\rfloor k \leq h \leq k'$, the same approximation is applied up to $h = k'$, because the network begins to saturate at $h = k' + 1$. If a network reaches a saturation point, the number nodes accessible at each hop cannot exceed n. It will be shown that the approximate model, based on equal probability assumption, predicts the expected number of nodes at each hop fairly well, and matches simulation results. Based on the assumption, the equation for the expected number of hops for reduced-stage shuffle networks, multiplied by $(kn - 1)$, is derived as follows:

$$
\begin{aligned}
&(kn-1)E\{hops\} \\
&= \sum_{j=1}^{k-1} jp^j + \sum_{j=0}^{k-1}(k+j)\{(p^{k+j}-p^j)p^k\frac{(j+1)k}{N} \\
&\quad + kp^{k+j}\frac{n-(j+1)p^k}{N}\} \\
&\quad + \sum_{i=2}^{\lfloor\frac{k'}{k}\rfloor-1}\sum_{j=0}^{k-1}(ik+j)\{p^{ik+j} - \frac{1}{2}\sum_{h=0}^{i-1}p^{(i-1)k-hk+j}\} \\
&\quad + \sum_{j=0}^{k'-\lfloor\frac{k'}{k}\rfloor k}(\lfloor\frac{k'}{k}\rfloor k+j)\{p^{\lfloor\frac{k'}{k}\rfloor k+j} \\
&\quad - \frac{1}{2}\sum_{h=0}^{\lfloor\frac{k'}{k}\rfloor-1}p^{(\lfloor\frac{k'}{k}\rfloor-1)k-kh+j}\} \\
&\quad + \sum_{j=0}^{k-2}(k'+1+j)\{n - \frac{1}{2}\sum_{h=0}^{\lfloor\frac{k'}{k}\rfloor-1}p^{\lfloor\frac{k'}{k}\rfloor k-kh+j}\}.
\end{aligned}
$$

3.2.1 Discussions

Table 2 shows some of the calculation results of the expected number of hops for different n's. The network topology for $k = 1$ corresponds to a single-stage shuffle network. Note that the expected number of hops in a single-stage network is not minimal compared to multiple stage cases, even though a single-stage network has a minimum diameter in a class of shuffle networks. This phenomenon originates from the fact that all nodes in the single-stage network are *care* nodes, which results in a low efficiency. In general, the fraction of don't care nodes in reduced-stage networks decreases as k decreases. (Note that extra-stage shuffle networks have a high fraction of don't care nodes.)

4 Simulation

Our results were verified through a computer simulation conducted for the $N = 16$ case. As expected, in the 16-node case, the (16, 2, 8, 2) GSN shows the best network throughput. We also compared the results with the 16-node MSN. The comparison shows that the MSN has an intermediate network performance when compared to various network topologies. All of these comparisons are shown in Fig. 3, where the horizontal axis indicates the interval between two generated packets at each node and the vertical axis represents the total number of packets that arrive safely at their destinations at each time slot. The results indicate that during low traffic, all types of networks show similar throughput rates, while during high traffic, various networks show significantly different throughput rates (more than 50%).

5 Conclusions

In this paper, we analyzed different topologies of shuffle networks. With the new definition, it is possible to realize a shuffle network in a variety of different ways with a given N. We show that different network topologies provide us with different network throughput; by selecting a network properly, significant performance improvement can be achieved. The reduced-stage shuffle networks have a shorter network diameter and offer a less number of expected hops, but experience a lot of overlapping as k becomes smaller. On the other hand, the performance of extra-stage shuffle

Table 2: Expected number of hops for different n's

	n=2	4	8	16	32	64	128	256	512
N= 8	2.286	**2.000**	2.107*	×	×	×	×	×	×
16	4.267	2.933	2.730*	2.833*	×	×	×	×	×
24	6.261	3.913	**3.261**	×	×	×	×	×	×
32	8.258	4.903	3.742	3.565*	4.417*	×	×	×	×
64	16.254	8.889	5.714	**4.635**	5.224*	5.269*	×	×	×
128	32.252	16.882	9.701	6.614	5.637*	5.753*	6.169*	×	×
160	40.252	20.881	11.698	7.610	**6.069**	×	×	×	×
256	64.251	32.878	17.694	10.604	7.561	6.253*	6.823*	7.504*	×
384	96.251	48.877	25.691	14.601	9.556	**7.536**	7.892*	×	×
512	128.250	64.877	33.691	18.599	11.554	8.532	8.063*	8.348*	8.665*

† Numbers marked with ⋆ correspond to reduced-stage shuffle networks.
‡ Numbers in bold face correspond to classical shuffle networks.
‡† Each square marked with × is not feasible.

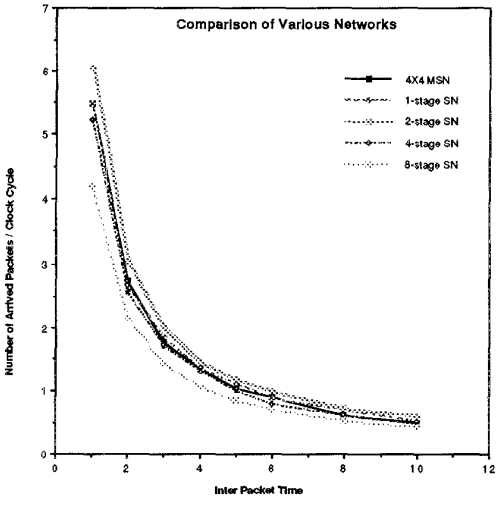

Figure 3: Comparison of various network topologies for $N = 16$

networks becomes worse as k becomes larger due to the long diameter. Study results in this paper can be used to determine an optimal network topology when given a value of N, by trading off the complexity and performance of a network.

References

[1] A.S. Acampora, M.J. Karol and M.G. Hluchyj, "Terabit lightwave networks: The multihop approach," *AT&T Tech. Journal*, Vol. 66, pp. 21 - 34, Nov./Dec., 1985.

[2] N.F. Maxemchuk, "The Manhattan street network," *Proc. of GLOBECOM '85*, pp. 255 - 261, 1985.

[3] M.G. Hluchyj and M.J. Karol, "ShuffleNet: An application of generalized perfect shuffles to multihop lightwave networks," *Proc. of IEEE INFOCOM*, pp. 4B.4.1 - 4B.4.5, 1988.

[4] N.F. Maxemchuk, "Routing in the Manhattan street network," *IEEE Trans. Commun.*, Vol. COM-35, pp. 503-512, May 1987.

[5] N.F. Maxemchuk, "Comparison of deflection and store-and-forward techniques in the Manhattan street and shuffle-exchange networks," *Proc. of IEEE INFOCOM*, pp. 800-809, 1989.

[6] S.-H. G. Chan and H. Kobayashi, "Performance analysis of shufflenet with deflection routing," *Proc. of IEEE Global Telecommunication Conference*, pp. 854-859, 1993.

[7] A.K. Choudhury and V.O.K. Li, "An approximate analysis of the performance of deflection routing in regular networks," *IEEE J. Select. Areas in Commun.*, Vol. 11, pp. 1302-1316, OCt. 1993.