Journal of Parallel and Distributed Computing **61**, 1209–1224 (2001)
doi:10.1006/jpdc.2001.1750, available online at http://www.idealibrary.com on IDEAL®

Optimal Routing Based on Super Topology in Optical Parallel Interconnect

Peng-Jun Wan and Liwu Liu

Department of Computer Science, Illinois Institute of Technology, Chicago, Illinois 60616

and

Yuanyuan Yang

Department of ECE, State University of New York at Stony Brook, Stony Brook, New York 11794

Received December 27, 1999; revised May 7, 2000; accepted July 20, 2000

Traditionally the routing in optical parallel interconnect is based on an embedded virtual topology. However, one important fact that has been neglected in the past is that the wavelength assignment to transceivers actually creates additional (logical) links not present in the virtual topology. Such a side-effect can be utilized to significantly reduce the number of hops between a pair of processors. This observation leads to the concept of *super topology*. This paper considers the hypercube as the embedded virtual topology. The ideas contained here are easily applicable to optical parallel interconnects employing other virtual topologies as well. We present a general framework for embedding a regular topology, the structure of the super topology, the optimal routing algorithm, the distance between any pair of processors and the diameter in the super topology.

Key Words: WDM; optical parallel interconnect; routing; virtual topology; super topology; distance; diameter.

1. INTRODUCTION

Optical passive star couplers [6, 8] provide a simple medium to connect processors in a parallel system or networking terminals in a local or metropolitan area network [11]. Figure 1 shows a typical *wavelength division multiplexing* (*WDM*) optical parallel interconnect in which each processor is connected to the star coupler via a pair of unidirectional fibers. Each processor has a set of transmitters and receivers. Each transmitter (receiver) is tuned to a specific wavelength channel from which it transmits (receives) light signals into (from) an optical fiber. The light signals entering the star coupler are evenly divided among all the output ports. A transmission from one processor to another processor is accomplished by tuning a transmitter of the sending processor and a receiver of the receiving processor to the same wavelength. Transmissions with different wavelength channels can take place

0743-7315/01 $35.00

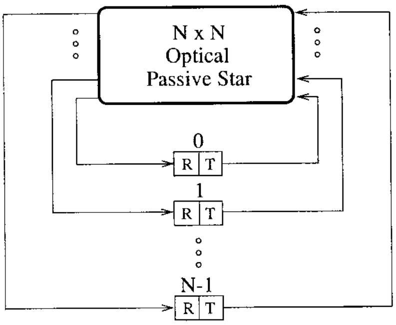

FIG. 1. An N-node optical parallel interconnect based on passive star coupler.

simultaneously. If the number of wavelength channels is less than the number of transmitters (or receivers), the wavelength channels can be shared among them in the time-division multiplexing manner, which results in *time and wavelength division multiplexing* (*TWDM*) media access protocols [5, 7].

In conventional parallel interconnect, the interconnection topology is fixed. Thus it is impossible to reconfigure it to adapt to any specific application or computation. The optical parallel interconnect based on passive star couplers, on the other hand, is reconfigurable by tuning the wavelengths channels of the optical transceivers at each processor. In general, the optical transceivers at each processor can be chosen to be either slowly tunable or fast tunable. Currently, the fast tunable transceivers cost much more than the slowly tunable transceivers. Their tuning speed is still very slow compared to the transmission speed of optical fibers and is inverse to their tunable range. In addition, they require accurate pretransmission coordination. Thus one practical and cost-effective alternate is to employ a *small* number of less expensive and readily available slowly tunable transceivers at each processor. This configuration can emulate the tunability of the fast tunable transceivers without suffering from tuning delay. In addition, a processor can take part in several communications simultaneously through different transceivers. Such concurrence cannot be achieved if a single fast tunable transceiver is used. Therefore, in this paper we consider the optical parallel interconnect based on this configuration.

In general, regular interconnection topologies are adopted for parallel interconnect topologies. In [13], an approach to realizing a regular interconnection topology has been proposed in the context of broadcast-and-select passive optical networks. The idea is to assign wavelengths to the transceivers properly such that for any link $a \rightarrow b$ in the regular topology, the node a has a transmitter and the node b has a receiver that are assigned to the same wavelength. However, the proposed approach is applicable only when both the number of transmitters and the number of receivers can divide the nodal degree of the regular topology. Such constraint leads to the limited flexibility of the system. Moreover, when neither the number of transmitters and the number of receivers at each node can divide the other, the proposed approach tends to make use of a small number of wavelengths, and therefore limits the transmission concurrence of the system. Thus the first contribution of this paper is to develop a general framework to realize a regular interconnection topology that is applicable to any number of transmitters and any number of receivers.

Traditionally, the routing in a parallel optical interconnect or a broadcast-and-select optical network simply follows the same routing algorithm developed for the embedded regular interconnection topology [9, 10]. However, we can do better as the process to realize the regular interconnection topology actually creates some *by-products*, the additional (logical) links not present in the original regular interconnection topology. Thus the actual logical interconnection pattern contains the embedded regular interconnection topology as a subgraph and hence is referred to as *super topology*. Because of the better connectivity in the super topology, such a *side-effect* can be exploited to reduce the distance in terms of the number of hops among processors. This can be illustrated in the following simple example as illustrated in Fig. 2. Consider a parallel system of eight processors into which a 3-cube is embedded as follows. Each processor has a single transmitter and a single receiver. The transmitters at processors 000, 011, 101, 110 and the receivers at processors 001, 010, 100, 111 are assigned wavelength λ_0, while the receivers at processors 000, 011, 101, 110 and the transmitters at processors 001, 010, 100, 111 are assigned wavelength λ_1. Now we consider the routing from processor 000 to processor 111. If the routing is simply based on the routing in the 3-cube, then the shortest distance consists of three hops. However, as the transmitter at processor 000 and the receiver at the processor 111 have the same wavelength λ_0, the processor 000 can talk to the processor 111 directly, and therefore their distance is just one. A graph theoretic explanation to this improvement is the difference between the embedded 3-cube and the super topology. Figure 2 shows the super topology of the above wavelength assignment. In addition to the links in the 3-cube, four additional links are present in the super topology: the link between 000 and 111, the link between 001 and 110, the link between 010 and 101, and the link between 100 and 011. It is easy to see that the diameter of this super topology is two, while that of the 3-cube is three.

The above observation leads to the question of how much better the super topology is than the original regular interconnection topology in terms of the network properties such as routing, load balancing, and fault tolerance. The second

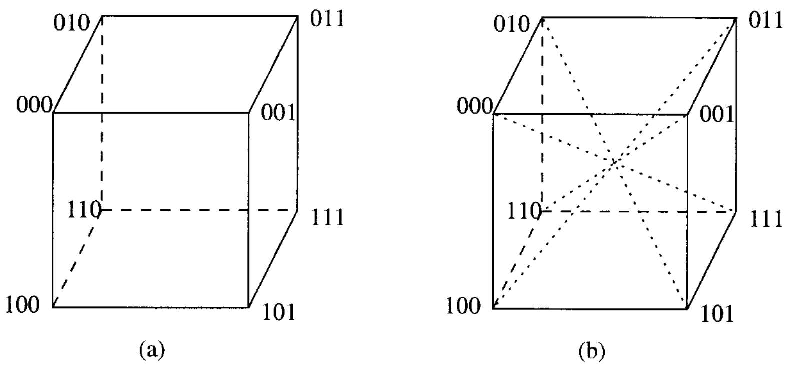

FIG. 2. The super topology of the embedded 3-cube.

contribution of our paper is to characterize various super topologies and develop their optimal routings in terms of the least number of hops.

The remainder of this paper is arranged as follows. Section 2 presents a general framework to embed a given regular topology into an optical parallel interconnect with an arbitrary number of transceivers at each processor. In addition, some key concepts such as transmission graph and connected components and their relations to the super topology are analyzed. Section 3 investigates the optimal routing in the super topology of parallel hypercube interconnect. It presents the structure of the super topology, the optimal routing algorithm, the distance between any pair of processors, and the diameter in the super topology. Finally Section 4 concludes the paper.

2. EMBEDDING OF REGULAR TOPOLOGIES

Let N be the number of processors, indexed by numbers from 0 to $N-1$. Each processor is equipped with T transmitters and R receivers. The set of transmitters at processor a is denoted by

$$\{(a, t) \mid 0 \leqslant t < T\}.$$

The set of receivers at processor a is denoted by

$$\{(a, r) \mid 0 \leqslant r < R\};$$

For each transmitter (a, t), a and t are called its *node index* and *local index*, respectively. Similarly, for each receiver (a, r), a and r are also called its *node index* and *local index* respectively. The differences between the transmitter and the receiver are specified by the context.

Let G be any regular topology on N vertices that is to be realized. If G is an undirected graph, we treat each edge in it as two unidirectional links. Let d be the nodal degree of G. In some regular interconnection topologies, d and N are independent, while in others d is a function of N. An intrinsic ordering of the outgoing links and the incoming links at each node is given a priori. In particular, the ith outgoing (incoming) link at each node is said to be along dimension i for any $0 \leqslant i < d$.

If d is a multiple of both T and R, G can be realized via the *uniform consecutive partition* scheme proposed in [13]. However, when d is not a multiple of T or R, a more delicate approach is needed. To simplify the presentation, we introduce the following definition. For any positive integer k and any set S, a k-partition of S, $\{S_0, S_1, ..., S_{k-1}\}$, is said to be *even* if

$$|S_i| = \begin{cases} \left\lceil \dfrac{|S|}{k} \right\rceil & \text{for} \quad 0 \leqslant i < |S| \bmod k, \\ \left\lfloor \dfrac{|S|}{k} \right\rfloor & \text{for} \quad |S| \bmod k \leqslant i < k. \end{cases}$$

We begin with the simplest case in which $T = R$. The d dimensions are *evenly* and consecutively partitioned into T groups

$$\{D_0, D_1, ..., D_{T-1}\}.$$

The transmitter (a, t) is responsible for the outgoing links at a along dimensions in D_t, and the receiver (a, r) is responsible for the incoming links at a along dimensions in D_r.

Next, we consider the case that $0 < T < R < d$. We partition the d dimensions evenly and consecutively into T groups

$$\{D_0, D_1, ..., D_{T-1}\}.$$

Then the transmitter (a, t) will be responsible for the outgoing links at a along dimensions in D_t. For each $0 \leqslant t < R \bmod T$, we further partition D_t evenly and consecutively into $\lceil \frac{R}{T} \rceil$ subgroups

$$\{D_{t,0}, D_{t,1}, ..., D_{t,\lceil R/T \rceil - 1}\},$$

and the receiver $(a, t\lceil \frac{R}{T} \rceil + i)$ is responsible for the incoming links at a along dimensions in $D_{t,i}$ for any $0 \leqslant i < \lceil \frac{R}{T} \rceil$. For each $R \bmod T \leqslant t < T$, D_t is further evenly and consecutively partitioned into $\lfloor \frac{R}{T} \rfloor$ subgroups

$$\{D_{t,0}, D_{t,1}, ..., D_{t,\lfloor R/T \rfloor - 1}\},$$

and the receiver $(a, R \bmod T + t\lfloor \frac{R}{T} \rfloor + i)$ is responsible for the incoming links at a along dimensions in $D_{t,i}$ for any $0 \leqslant i < \lfloor \frac{R}{T} \rfloor$.

When $0 < R < T < d$, similar partitions can be performed as above. We first partition the d dimensions evenly and consecutively into R groups

$$\{D_0, D_1, ..., D_{R-1}\}.$$

Then the receiver (a, r) will be responsible for the incoming links at a along dimensions in D_r. For each $0 \leqslant r < T \bmod R$, we further partition D_r evenly and consecutively into $\lceil \frac{T}{R} \rceil$ subgroups

$$\{D_{r,0}, D_{r,1}, ..., D_{r,\lceil T/R \rceil - 1}\},$$

and the transmitter $(a, t\lceil \frac{R}{T} \rceil + i)$ is responsible for the outgoing links at a along dimensions in $D_{r,i}$ for any $0 \leqslant i < \lceil \frac{T}{R} \rceil$. For each $T \bmod R \leqslant r < R$, D_r is further evenly and consecutively partitioned into $\lfloor \frac{T}{R} \rfloor$ subgroups

$$\{D_{r,0}, D_{r,1}, ..., D_{r,\lfloor T/R \rfloor - 1}\},$$

and the receiver $(a, T \bmod R + r\lfloor \frac{T}{R} \rfloor + i)$ is responsible for the outgoing links at a along dimensions in $D_{r,i}$ for any $0 \leqslant i < \lfloor \frac{T}{R} \rfloor$.

The above partition induces a bipartite digraph in which the vertex set consists of all NT transmitters and all NR receivers, and there is a link from a transmitter to a receiver if and only if they are responsible for one common link in the interconnection topology to be realized. This graph is referred to as the *transmission graph*. It is obvious that the number of links in the transmission graph is the same as that in the interconnection topology. The partition also imposes a constraint on the wavelength assignment of the transmitters and receivers as explained below. Since the transmitters and receivers are slowly tunable, once they are tuned to some particular wavelengths, this configuration will last for a relatively long time until the next reconfiguration. Consequently, any transmitter (receiver) and its adjacent receivers (transmitters) in the transmission graph are forced to have the same wavelength channels of the transmitter (receiver). Therefore any pair of transceivers must have the same wavelength channel if there is a path between them assuming the links in the transmission graph are bidirectional. Thus all transmitters and receivers in the same connected component (ignoring the unidirectional nature of the links) of the transmission graph must have the same wavelength channel.

The connected components of the transmission graph play a very important role in the design and analysis of the optical parallel interconnect. First of all, it helps to choose the right number of transceivers in the most cost-effective way. Note that the number of connected components in the transmission graph might exceed the number of available wavelengths. As an extreme example, when $T = R = d$ each connected component consists of only one link. In this case, one or more connected components should share a wavelength channel, and the transmission concurrence is limited by the available wavelengths. Thus to minimize the system cost, we should select T and R such that the number of connected components is no more than the number of available wavelengths. So in the remainder of this paper, we assume that the number of transceivers is selected such that each connected component has a unique wavelength. Under this assumption, a wavelength assigned to a connected component is shared by all transmitters in this connected component in a time-sharing manner. Thus the number of time slots in a TDM frame of this wavelength is at least the number of transmitters in the corresponding connected component. At this point, it is hard to claim whether a higher number of transmitters would lead to higher performance. On one hand, the higher number of transmitters may result in the higher number of connected components. But on the other hand, the higher number of transmitters may also lead to the higher number of transmitters in each connected component. Thus to judge the cost–performance relation, an analytic formula for the total number of time slots required by a pair of nodes to communicate is needed.

The connected components can also help to determine the super topology of a wavelength assignment. In general, there is a link from node a to node b in the super topology if and only if a has a transmitter and b has a receiver which are in the same connected component. Thus once the structure of connected components is characterized, we are able to determine the set of neighbors of each processor and the nodal degree in the super topology. The optimal routing algorithms in the super topology are then possibly obtained.

3. OPTIMAL ROUTING IN SUPER TOPOLOGIES OF HYPERCUBES

The n-dimensional hypercube, or n-cube in short, has $N=2^n$ nodes which are labeled by n-bit binary numbers. The nodal degree of n-cube is n. For each node $0 \leqslant a \leqslant N-1$, its outgoing links are

$$a \to a \oplus 2^i, \qquad 0 \leqslant i \leqslant n-1$$

and its incoming links are

$$a \oplus 2^i \to a, \qquad 0 \leqslant i \leqslant n-1,$$

where the operator $\oplus$ is the parity operator (bit-wise exclusive or). The hypercube presents several attractive properties, such as simple self-routing, logarithmic diameter, and high fault-tolerance.

For simplicity of discussion, we first introduce some notations. We use T and R to denote the number of transmitters and the number of receivers respectively at each processor. For any $S \subseteq \{0, 1, ..., n-1\}$ and any n-bit binary number a, we use $a|_S$ to denote the $|S|$-bit binary number consisting of the bits of a at positions in S and $a|_{\bar{S}}$ to denote the $(n-|S|)$-bit binary number consisting of the bits of a at positions not in S. For any two processors a and b, we use $H(a, b)$ to denote their distance in terms of the number of hops in the super topology.

Due to the symmetry of the hypercube, swapping the number of transmitters and receivers does not change the connectivity of the super topology. Thus we only consider the cases in which the number of transmitters is no more than the number of receivers. Section 3.1 studies the configuration in which the number of transmitters is equal to the number of receivers. Section 3.2 studies the configuration in which the number of transmitters is less than the number of receivers.

3.1. *Case* 1: $T=R$

We begin with the simplest case that $T=R$. The n dimensions are evenly and consecutively partitioned into T groups

$$\{D_0, D_1, ..., D_{T-1}\}.$$

The transmitter (a, t) is responsible for the outgoing links at a along dimensions in D_t, and the receiver (a, r) is responsible for the incoming links at a along dimensions in D_r. The next lemma characterizes the structure of any connected component in the transmission graph.

LEMMA 3.1. *The connected component containing the transmitter* (a, t) *consists of all transmitters*

$$\{(b, t) : (a \oplus b)|_{D_t} \text{ is even and } (a \oplus b)|_{\overline{D_t}} = 0\}$$

and all receivers

$$\{(b, t) : (a \oplus b)|_{D_t} \text{ is odd and } (a \oplus b)|_{\overline{D_t}} = 0\}.$$

Proof. Assume that a and b are only different at the dimensions

$$i_1, i_2, ..., i_{2k-1}, i_{2k} \in D_t.$$

Then we have the following path in the transmission graph.

$$\begin{aligned}
(a, t) &\searrow \\
&\quad (a \oplus 2^{i_1}, t) \\
(a \oplus 2^{i_1} \oplus 2^{i_2}, t) &\nearrow \\
&\searrow \\
&\quad (a \oplus 2^{i_1} \oplus 2^{i_2} \oplus 2^{i_3}, t) \\
&\quad \vdots \\
(a \oplus 2^{i_1} \oplus 2^{i_2} \oplus \cdots \oplus 2^{i_{2k-1}} \oplus 2^{i_{2k}}, t) &\nearrow \\
= (b, t).
\end{aligned}$$

Thus the two transmitters (a, t) and (b, t) are in the same subnetwork.

Now we prove the reverse direction. It is easy to prove that in each connected component all transceivers have the same local indices. Thus if two transceivers (a, t) and (b, t) are in the same subnetwork, $(a \oplus b)|_{\overline{D_t}} = 0$. If two transmitters (a, t) and (b, t) are in the same connected component, then a and b must have an even distance. Thus $(a \oplus b)|_{D_t}$ is even. ∎

Thus in the super topology, the set of neighbors of node a is

$$\bigcup_{t=0}^{T-1} \{b : (a \oplus b)|_{D_t} \text{ is odd and } (a \oplus b)|_{\overline{D_t}} = 0\}.$$

So the nodal degree of each node in the super topology is

$$\begin{aligned}
\sum_{t=0}^{T-1} 2^{|D_t|-1} &= (n \bmod T)\, 2^{\lceil n/T \rceil - 1} + (T - n \bmod T)\, 2^{\lfloor n/T \rfloor - 1} \\
&= (T + n \bmod T)\, 2^{\lfloor n/T \rfloor - 1}.
\end{aligned}$$

As

$$(T + n \bmod T)\, 2^{\lfloor n/T \rfloor - 1} = n$$

if and only if $2T \geqslant n$, each node has more neighbors in the super topology if $2T < n$ and thus may reduce the distances between some nodes. This will be confirmed later in this section.

Now consider the optimal routing in the super topology. The routing from processor a to processor b is equivalent to changing the bits of a to the bits of b

TABLE 1

Optimal Routing from a to b When $T=R$

```
Algorithm Routing1(a, b)
  if a = b, stop;
  find the smallest t with (a ⊕ b)|_{D_t} ≠ 0;
  if (a ⊕ b)|_{D_t} is odd then
    pick any a' satisfying that
      (a' ⊕ b)|_{D_t} = (a' ⊕ a)|_{\overline{D_t}} = 0 and (a' ⊕ a)|_{D_t} is odd;
    a transmits to a' via transmitter (a, t);
    Routing1(a', b);
  else
    pick any a' satisfying that
      (a' ⊕ a)|_{\overline{D_t}} = 0 and (a' ⊕ a)|_{D_t} is odd;
    pick any a'' satisfying that
      (a'' ⊕ b)|_{D_t} = (a'' ⊕ a)|_{\overline{D_t}} = 0;
    a transmits to a' via transmitter (a, t);
    a' transmits to a'' via transmitter (a', t);
    Routing1(a'', b);
End-Algorithm
```

according to certain rules. In the super topology with $T=R$, at each step any odd number of bits at positions in some D_t are allowed to be reversed simultaneously. Recall that in the original n-cube, only *one* bit can be changed at a time. Thus the distance in terms of the number of steps or hops to change a to b should be potentially smaller.

Note that at each step the reversal of bits at positions in some D_t has no impact on the bits in other positions. Thus to change the bits of a at positions in D_t to the bits of b at positions in D_t, we only have to look at $(a \oplus b)|_{D_t}$. Suppose that $(a \oplus b)|_{\overline{D_t}} = 0$ and $(a \oplus b)|_{D_t} \neq 0$. Then a single hop is needed from processor a to processor b if $(a \oplus b)|_{D_t}$ is odd, and two hops are needed if $(a \oplus b)|_{D_t}$ is even. Therefore, the routing can be performed sequentially for each $0 \leqslant t < T$. The optimal routing given in Table 1 is very similar to the well-known Z-routing in the hypercube. It is given in the recursive format for the simplicity of description.

For any binary number a and any $0 \leqslant t < T$, we define $h_t(a)$ as follows.

$$h_t(a) = \begin{cases} 0, & \text{if} \quad a|_{D_t} = 0, \\ 1, & \text{if} \quad a|_{D_t} \neq 0 \text{ and } a|_{D_t} \text{ is odd}, \\ 2, & \text{if} \quad a|_{D_t} \neq 0 \text{ and } a|_{D_t} \text{ is even}. \end{cases}$$

Then the following lemma gives the distance between any pair of processors in the super topology.

LEMMA 3.2. *When $T=R$, the distance between the processor a and the processor b in the super topology is*

$$H(a, b) = \sum_{t=0}^{T-1} h_t(a \oplus b).$$

In the next lemma we will study the diameter of the super topology.

LEMMA 3.3. *When $T=R$, the diameter of the super topology is* $\min\{n, 2T\}$.

Proof. From Lemma 3.2, the diameter is equal to the sum of the maxima of $h_t(a)$ over all $0 \leqslant t < T$. It is easy to see that for any $0 \leqslant t < T$, the maxima of $h_t(a)$ is two if $|D_t| > 1$ and is one if $|D_t| = 1$. Therefore, the diameter is at most $2T$. If $n \geqslant 2T$, then $|D_t| > 1$ for any $0 \leqslant t < T$. In particular, if $a|_{D_t} = 11$ for any $0 \leqslant t < T$, the distance between the processor a and the processor 0 is exactly $2T$. So in this case the diameter is equal to $2T$.

Now we assume that $n < 2T$. For any $0 \leqslant t < n-T$, $|D_t| = 2$ and thus the maxima of $h_t(a)$ is two. For any $n-T \leqslant t < T$, $|D_t| = 1$ and thus the maxima of $h_t(a)$ is one. Therefore the diameter is

$$2(n-T) + (2T-n) = n. \quad \blacksquare$$

Lemma 3.3 implies that the fewer the number of transmitters or receivers, the shorter the diameter. However, the fewer number of transmitters may cause a larger number of transmitters or receivers in each subnetwork and result in longer channel access delay. Indeed, the number of time slots required to complete a communication might be as large as

$$2 \sum_{t=0}^{T-1} 2^{|D_t|-1} = (T + n \bmod T)\, 2^{\lfloor n/T \rfloor},$$

which might decrease as T increases.

3.2. *Case* 2: $T < R$

Now we consider the configuration with $0 < T < R < n$. The n dimensions are evenly and consecutively partitioned into T groups

$$\{D_0, D_1, \dots, D_{T-1}\}.$$

The transmitter (a, t) is responsible for the outgoing links at a along dimensions in D_t. For each $0 \leqslant t < R \bmod T$, D_t is further evenly consecutively partitioned into $\lceil \frac{R}{T} \rceil$ subgroups

$$\{D_{t,0}, D_{t,1}, \dots, D_{t,\lceil R/T \rceil - 1}\},$$

and the receiver $(a, t\lceil \frac{R}{T} \rceil + i)$ is responsible for the incoming links at a along dimensions in $D_{t,i}$ for any $0 \leqslant i < \lceil \frac{R}{T} \rceil$. For each $R \bmod T \leqslant t < T$, D_t is further evenly and consecutively partitioned into $\lfloor \frac{R}{T} \rfloor$ subgroups

$$\{D_{t,0}, D_{t,1}, \dots, D_{t,\lfloor R/T \rfloor - 1}\}$$

and the receiver $(a, R \bmod T + t\lfloor \frac{R}{T} \rfloor + i)$ is responsible for the incoming links at a along dimensions in $D_{t,i}$ for any $0 \leqslant i < \lfloor \frac{R}{T} \rfloor$.

The next lemma presents the structure of the connected components.

LEMMA 3.4. *Suppose that $0 < T < R < n$. If $0 \leqslant t < R \bmod T$ the connected component containing the transmitter (a, t) consists of all transmitters*

$$\left\{(b, t) : (a \oplus b)|_{\overline{D_t}} = 0, (a \oplus b)|_{D_{t,i}} \text{ is even for any } 0 \leqslant i < \left\lceil \frac{R}{T} \right\rceil \right\}$$

and all receivers

$$\bigcup_{i=0}^{\lceil R/T \rceil - 1} \left\{ \left(b, t\left\lceil \frac{R}{T} \right\rceil + i\right) : (a \oplus b)|_{\overline{D_t}} = 0, (a \oplus b)|_{D_{t,i}} \text{ is odd} \right.$$
$$\left. \text{and } (a \oplus b)|_{D_{t,j}} \text{ is even for any } j \neq i \right\}.$$

If $R \bmod T \leqslant t < T$ the connected component containing the transmitter (a, t) consists of all transmitters

$$\left\{(b, t) : (a \oplus b)|_{\overline{D_t}} = 0, (a \oplus b)|_{D_{t,i}} \text{ is even for any } 0 \leqslant i < \left\lfloor \frac{R}{T} \right\rfloor \right\}$$

and all receivers

$$\bigcup_{i=0}^{\lfloor R/T \rfloor - 1} \left\{ \left(b, R \bmod T + t\left\lfloor \frac{R}{T} \right\rfloor + i\right) : (a \oplus b)|_{\overline{D_t}} = 0, (a \oplus b)|_{D_{t,i}} \text{ is odd} \right.$$
$$\left. \text{and } (a \oplus b)|_{D_{t,j}} \text{ is even for any } j \neq i \right\}.$$

Proof. We only consider the case that $0 \leqslant t < R \bmod T$; the case $R \bmod T \leqslant t < T$ can be dealt with in the same way. It is easy to verify that in the connected component containing transmitter (a, t), the local index of any transmitter must be t and the local index of any receiver must be between $t\lfloor \frac{R}{T} \rfloor$ and $(t+1)\lfloor \frac{R}{T} \rfloor - 1$. Let b be any node satisfying that $(a \oplus b)|_{\overline{D_t}} = 0$, $(a \oplus b)|_{D_{t,i}}$ is even for any $0 \leqslant i < \lceil \frac{R}{T} \rceil$. Then following the Z-routing in hypercube, we can construct a path between (a, t) and (b, t) in the transmission graph. Now we prove the reverse. Assume that (a, t) and (b, t) are in the same connected component. Then we can prove by induction on the length of the distance between (a, t) and (b, t) in the transmission graph that $(a \oplus b)|_{\overline{D_t}} = 0$, $(a \oplus b)|_{D_{t,i}}$ is even for any $0 \leqslant i < \lceil \frac{R}{T} \rceil$. Thus the set of transmitters in the connected component containing transmitter (a, t) is exactly

$$\left\{(b, t) : (a \oplus b)|_{\overline{D_t}} = 0, (a \oplus b)|_{D_{t,i}} \text{is even for any } 0 \leqslant i < \left\lceil \frac{R}{T} \right\rceil \right\}.$$

The set of receivers in this connected component is consequently

$$\bigcup_{i=0}^{\lceil R/T\rceil-1}\left\{\left(b, t\left\lceil\frac{R}{T}\right\rceil+i\right) : (a\oplus b)|_{\overline{D_t}}=0, (a\oplus b)|_{D_{t,i}} \text{ is odd and } (a\oplus b)|_{D_{t,j}} \text{ is even for any } j\neq i\right\}. \quad \blacksquare$$

From the above lemma, the set of neighbors of the node a in the super topology is

$$\bigcup_{t=0}^{R \bmod T-1}\ \bigcup_{i=0}^{\lceil R/T\rceil-1}\{b : (a\oplus b)|_{\overline{D_t}}=0, (a\oplus b)|_{D_{t,i}} \text{ is odd}, (a\oplus b)|_{D_{t,j}} \text{ is even } \forall j\neq i\}$$
$$\cup \bigcup_{t=R \bmod T}^{T-1}\ \bigcup_{i=0}^{\lceil R/T\rceil-1}\{b : (a\oplus b)|_{\overline{D_t}}=0, (a\oplus b)|_{D_{t,i}} \text{ is odd}, (a\oplus b)|_{D_{t,j}} \text{ is even } \forall j\neq i\}.$$

So the nodal degree of each processor in the super topology is

$$\sum_{t=0}^{R \bmod T-1}\left\lceil\frac{R}{T}\right\rceil 2^{|D_t|-\lceil R/T\rceil}+\sum_{t=R \bmod T}^{T-1}\left\lfloor\frac{R}{T}\right\rfloor 2^{|D_t|-\lfloor R/T\rfloor}$$

which can be simplified as follows:

- If $n \bmod T=R \bmod T$, then the degree is

$$R2^{(n-R)/T}.$$

- If $n \bmod T<R \bmod T$, then the degree is

$$R2^{\lceil (n-R)/T\rceil}+(n \bmod T-R \bmod T)\left\lceil\frac{R}{T}\right\rceil 2^{\lfloor (n-R)/T\rfloor}.$$

- If $n \bmod T>R \bmod T$, then the degree is

$$\left(R+((n-R) \bmod T)\left\lfloor\frac{R}{T}\right\rfloor\right)2^{\lfloor (n-R)/T\rfloor}.$$

Now consider the optimal routing in the super topology. We again treat the routing from a to b as the number of steps to change the bits of a to the bits of b. For any $0\leqslant t<T$, any odd number of bits at positions in some $D_{t,i}$ and any even number of bits at positions in any $D_{t,j}$ with $j\neq i$ are allowed to be reversed simultaneously within a single step. Note that at each step the reversal of bits at positions in some D_t has no impact on the bits in other positions. Thus to change the bits of a at positions in D_t to the bits of b at positions in D_t, we only have to look at $(a\oplus b)|_{D_t}$. Suppose that $(a\oplus b)|_{\overline{D_t}}=0$ and $(a\oplus b)|_{D_t}\neq 0$. Then if $(a\oplus b)|_{D_{t,i}}$ is even for all i, two hops are needed from processor a to processor b

TABLE 2

Optimal Routing from *a* to *b* When *T* < *R*

Algorithm Routing2(a, b)
  if $a = b$, *stop*;
  find the smallest t with $(a \oplus b)|_{D_t} \neq 0$;
  find the set $S = \{i : (a \oplus b)|_{D_{t,i}} \text{ is odd}\}$;
  if $S = \emptyset$ *then*
    choose any a' *satisfying that*
      $(a' \oplus a)|_{\overline{D_t}} = 0$ *abd* $(a' \oplus a)|_{D_{t,i}}$ *is odd for some* i;
    choose any a'' *satisfying that*
      $(a'' \oplus b)|_{D_t} = (a'' \oplus a)|_{\overline{D_t}} = 0$;
    a transmits to a' *via transmitter* (a, t);
    a' *transmits to* a'' *via transmitter* (a', t);
    Routing2(a'', b);
  else
    for each $i \in S$ *in the increasing order do*
      if i is the last one in S then
        choose any a' *satisfying that*
          $(a' \oplus b)|_{D_t} = (a' \oplus a)|_{\overline{D_t}} = 0$ *and* $(a' \oplus a)|_{D_{t,i}}$ *is odd*;
      else
        choose any a' *satisfying that*
          $(a' \oplus a)|_{\overline{D_t}} = 0$, $(a' \oplus a)|_{D_{t,i}}$ *is odd and* $(a' \oplus a)|_{D_{t,i}}$ *is even* $\forall j \neq i$;
      a transmits to a' *via transmitter* (a, t);
      replace a by a';
    Routing2(a, b);
End-Algorithm

if $(a \oplus b)|_{D_{t,i}}$ is even for all i; otherwise, the minimal number of hops required from processor a to processor b is equal to the number of i's with odd $a|_{D_{t,i}}$. Such procedure can be repeated sequentially for each $0 \leqslant t < T$. The recursive version of an optimal routing algorithm is given in Table 2.

For any binary number a and any $0 \leqslant t < T$, we define $h_t(a)$ as follows.

$$h_t(a) = \begin{cases} 0, & \text{if } \; a|_{D_t} = 0, \\ 1, & \text{if } \; a|_{D_t} \neq 0 \text{ and } a|_{D_{t,i}} \text{ is even for all } i, \\ 2, & \text{if } \; a|_{D_t} \neq 0 \text{ and } a|_{D_{t,i}} \text{ is odd for some } i. \end{cases}$$

The distance between any pair of processors in the super topology is given in the following lemma.

LEMMA 3.5. *When* $0 < T < R < n$, *the distance between the processor a and the processor b in the super topology is*

$$H(a, b) = \sum_{t=0}^{T-1} h_t(a \oplus b).$$

Next we will study the diameter of the super topology.

LEMMA 3.6. *When* $0 < T < R < n$, *the diameter of the super topology is* $\min\{n, \max\{R, 2T\}\}$.

Proof. The diameter is equal to the sum of the maxima of $h_t(a)$ over all $0 \leqslant t < T$. In general, if the number of receivers at each processor that are responsible for the links along dimensions in D_t is more than one, the maxima of $h_t(a)$ is equal to such number. If there is only one receiver at each processor that is responsible for the links along dimensions in D_t, then the maximum of $h_t(a)$ is equal to two if $|D_t| > 1$ and equal to one if $|D_t| > 1$.

If $R \geqslant 2T$, then $\lceil \frac{R}{T} \rceil \geqslant \lfloor \frac{R}{T} \rfloor \geqslant 2$. Thus for any $0 \leqslant t < T$, there are at least two receivers that are responsible for the links along dimensions in D_t. This implies that the maxima of $h_t(a)$ is equal to the number of receivers at each processor that are responsible for the links along dimensions in D_t. So the diameter is equal to R, the total number of receivers at each processor.

If $T < R < 2T \leqslant n$, then $\lceil \frac{n}{T} \rceil \geqslant \lfloor \frac{n}{T} \rfloor \geqslant 2$. For $0 \leqslant t < R - T$, there are exactly two receivers that are responsible for the links along dimensions in D_t; hence the maxima of $h_t(a)$ is 2. For $R - T \leqslant t < T$, $|D_t| \geqslant 2$ while there is only one receiver at each processor that is responsible for the links along dimensions in D_t. So the maxima of $h_t(a)$ is also equal to two. Therefore the diameter is $2T$.

If $T < R < n < 2T$, we show that the diameter is n. In fact, for $0 \leqslant t < R - T$, $|D_t| = \lceil \frac{n}{T} \rceil = 2$ and there are exactly two receivers that are responsible for the two links along dimensions in D_t. Hence the maxima of $h_t(a)$ is 2. For $R - T \leqslant t < n - T$, $|D_t| = \lceil \frac{n}{T} \rceil = 2$ but there is only one receiver at each processor that is responsible for the two links along dimensions in D_t. So the maxima of $h_t(a)$ is also equal to two. For $N - T \leqslant t < T - 1$, $|D_t| = \lfloor \frac{n}{T} \rfloor = 1$ and there is only one receiver at each processor that is responsible for the link along dimensions in D_t. So the maximum of $h_t(a)$ is equal to one. Therefore the diameter is

$$2(R - T) + 2(n - T) + (2T - n) = n. \quad \blacksquare$$

Combining Lemmas 3.3 and 3.6, we have the following theorem.

THEOREM 3.1. *When* $T \leqslant R$, *the diameter of the super topology is*

$$\min\{n, \max\{R, 2T\}\}.$$

When $T \geqslant R$, *the diameter of the super topology is*

$$\min\{n, \max\{T, 2R\}\}.$$

4. CONCLUSION

In this paper, we first presented a general framework to embed any regular topology in the parallel interconnect with an arbitrary number of transceivers at each processor. We then characterized the structure of the super topology of the parallel hypercube interconnect by using the structure of the connected components in the transmission graph. The super topology has richer connectivity than the hypercube itself. The optimal routing between any pair of processors in the super topology was then developed. In addition, the analytic formula for the distance between any pair of processors was also obtained. Finally, the diameter of the super topology was calculated.

One possible future work is to develop the optimal routing algorithms in the super topology for common communication patterns such as all-to-all personalized communication.

The ideas and approaches contained in this paper are easily applicable to optical parallel interconnects realizing other topologies such as the de Bruijn graph [12], the star graph [1], and the rotator graph [4].

REFERENCES

1. S. B. Akers, D. Harel, and B. Krishnamurthy, The Star Graph: An attractive alternative to the *n*-cube, *in* "Proc. Int. Conf. Parallel Processing," pp. 393–400, 1987.
2. C. A. Brackett, On the capacity of multiwavelength optical-star packet switches, *IEEE Lightwave Magazine* (May 1991), 33–37.
3. M. S. Chen, N. R. Dono, and R. Ramaswami, A media access-protocol for packet-switched wavelength division multiaccess metropolitan networks, *IEEE J. Selected Areas Commun.* **8**, 6 (Aug. 1990), 1048–1057.
4. P. F. Corbett, Rotator graphs: An efficient topology for point-to-point multiprocessor networks, *IEEE Trans. Parallel Distrib. Systems* **3**, 5 (Sept. 1992), 622–626.
5. T. Q. Dam, K. A. Williams, and D. H. C. Du, "A Media-Access Protocol for Time and Wavelength Division Multiplexed Passive Star Networks," Technical Report 91-63, Computer Science Dept., University of Minnesota.
6. C. Dragone, Efficient $N \times N$ star coupler based on Fourier optics, *Electron. Lett.* **24**, 15 (July 1988), 942–944.
7. M. G. Hluchyj and M. J. Karol, ShuffleNet: An application of generalized perfect shuffles to multihop lightwave networks, *J. Lightwave Technol.* **9**, 10 (Oct. 91), 1386–1396.
8. R. A. Linke, Frequency division multiplexed optical networks using heterodyne detection, *IEEE Network Magazine* **3**, 2 (Mar. 1989), 13–20.
9. B. Mukherjee, WDM-based local lightwave networks-Part I: Single-hop systems, *IEEE Network* **6** (May 1992), 12–27.
10. B. Mukherjee, WDM-based local lightwave networks-Part II: Multihop systems, *IEEE Network* **6** (July 1992), 20–32.
11. B. Mukherjee, "Optical Communication Networks," McGraw–Hill, New York, 1997.
12. K. Sivarajan and R. Ramaswami, Multihop lightwave networks based on De Bruijn graphs, *in* "INFOCOM 90," Vol. 2, pp. 1001–1011, 1990.
13. P.-J. Wan, TWDM multichannel lightwave hypercube networks, *Theoret. Comput. Sci.* **194**, 1–2 (1998), 123–136.

PENG-JUN WAN has been an assistant professor in computer science at Illinois Institute of Technology since 1997. He received his Ph.D. in computer science from the University of Minnesota in 1997, his M.S. in operations research and control theory from the Chinese Academy of Science in 1993, and his B.S. in applied mathematics from Qsinghua University in 1990. His research interests include algorithm design and optimizations in optical networks and wireless networks.

LIWU LIU has been a Ph.D. student at Illinois Institute of Technology since 1998. He received his M.S. in computer science in 1997 and his B.S. in computer science in 1995, both from Qsinghua University. He is currently working on optical networking design and optimizations.

YUANYUAN YANG received the B.Eng. and M.S. in computer science and engineering from Tsinghua University, Beijing, China, in 1982 and 1984, respectively, and the M.S.E. and Ph.D. in computer science from Johns Hopkins University, Baltimore, Maryland, in 1989 and 1992, respectively. Dr. Yang is an associate professor of computer engineering with a joint appointment in computer science at the State University of New York at Stony Brook. Dr. Yang's research interests include parallel and distributed computing and systems, high speed networks, optical networks, high performance computer architecture, and fault-tolerant computing. Dr. Yang has published extensively in major journals, book chapters, and refereed conference proceedings related to these research areas. Dr. Yang holds two U.S. patents in the area of multicast communication networks. She is an associate editor for the IEEE Transactions on Parallel and Distributed Systems. Dr. Yang has served as a chair or on the program/organizing committees of a number of international conferences and workshops in her areas of research. She is a distinguished visitor of IEEE Computer Society, a senior member of the IEEE, and a member of the ACM, IEEE Computer Society and IEEE Communication Society.