WDM Interconnection Networks Based on Uniform Graph Emulations [1]

Khaled A. Aly
Ericsson Telecom
126 25 Stockholm, Sweden
etxkaly@kk.ericsson.se

Abstract

This paper studies an interconnection network realization based on broadcast-and-select wavelength-division multiplexing that acheives scalable passive optical interconnection and reconfigurable network partitioning. A cluster-based configuration where two sets of couplers individually realize the broadcast and the select functions results in flexible scalability at low complexity by either increasing the cluster size or the number of clusters. Virtual partitioned point-to-point connectivities are supported by applying the principle of uniform graph emulation. Conflict-free wavelength assignment is modeled as a vertex coloring problem with a minimum distance constraint of 3. The paper identifies the emulation maps and derives minimal, or upper bounded, channel assignments for generalized shuffle, cube, and tree based topologies.

1 Introduction

Passive optical interconnection with broadcast-and-select (BS) wavelength-division multiplexing (WDM) has been considered for multi-access and multi-hop packet switching applications in multiprocessor systems and local and metropolitan area networks [1–9]. Relative simplicity of optical devices, low space complexity, flexible network expansion, and delay-throughput performance enhancement are among the attractive features of the BS approach. This paper applies the principle of uniform graph emulation to BS networks based on regular topologies. In this structure, referred to as Multi-domain WDM (M-WDM), scalability is achieved by trading-off WDM and space interconnection. The broadcast feature is used to achieve passively reconfigurable network partitions to subnetworks of identical topology. Applications to this approach include asynchronous multi-hop packet/cell communication in B-ISDN and computer networks and synchronous nearest-neighbor packet communication in large parallel computers. A chordal ring based structure has been considered in [10, 11] to enable arbitrary topology to be emulated through the concept of graph bandwidth reduction. Regular interconnection graphs of relevance to these applications are treated in general: (n, k) de Bruijn, k-ary n-cube, n-dimensional cube-connected cycles, and k-ary tree. The interconnection network structure is defined in Section 1. The uniform emulation priciple is described in Section 2. Emulation maps for valid partitions are derived and/or summarized in Section 3. Section 4 derives the required number of channel sets for conflict-free communication. Conclusions are summarized in Section 5.

2 Physical Interconnection Network

The network consists of m_1 clusters where each cluster is a set of m_0 nodes, for a total network size of $M = m_1 m_0$ nodes. Each node possesses a single fixed-tuned transmitter (light source). The total number of wavelength channels in the system is denoted as C. Node receivers monitor all, or a subset of, the C channels. The receiver may be realized using a multichannel acoustooptic tunable filter or a detector array and a passive grating-based wavelength-division demultiplexer, which can be shared by nodes in the same cluster. Each cluster possesses its own broadcast and select domains realized by an output and an input optical star couplers, respectively. The cluster interconnection network (CIN) refers to the fiber connection pattern between output and input couplers. Nodes sharing an output coupler transmit over an ordered set of distinct channels. At the input coupler side, several distinct channel sets can be *listened* to, depending on the CIN topology. Transmit channel sets are assigned to output couplers (i.e. channels assigned to transmitters) such that no conflicts may occur at any input coupler. The assignment guarantees that the channel sets which can be listened to through any input coupler are disjoint to support collisionless communication.

[1] Parts of this paper were presented at IEEE INFOCOM'94 and ACM SIGCOMM'94

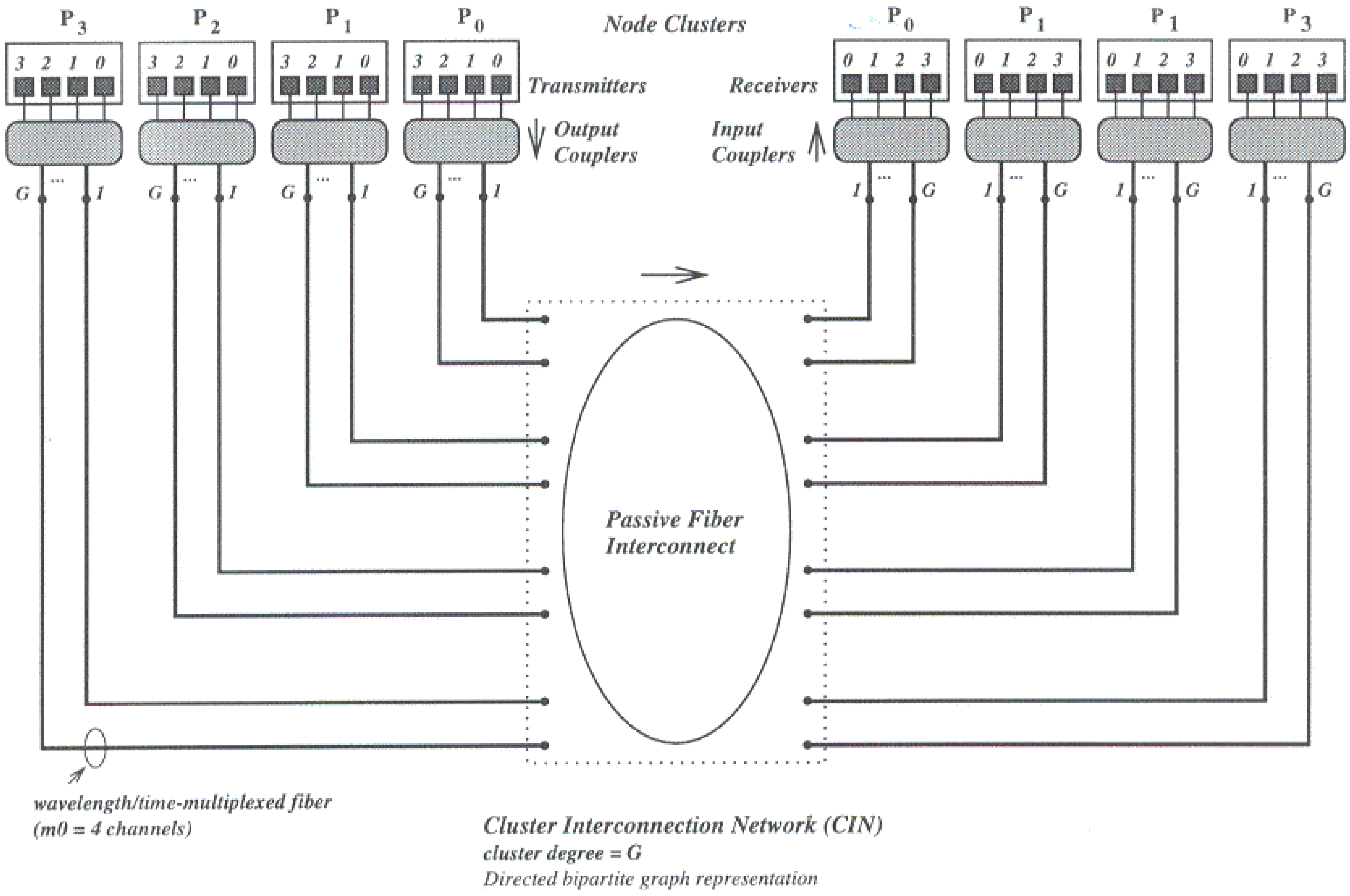

Figure 1: M-WDM processor network configuration based on the discrete broadcast-select approach ($m_1 = 4$ clusters and $m_0 = 4$ processors per cluster): transmit and receive modules are drawn *unfolded* and the CIN is represented by a directed bipartite graph. A self link (not shown) connects each cluster's output to input coupler. W ordered channel sets, generally $W \ll m_1$, are used for transmission.

Define the *set of clusters* $\mathrm{P} = \{\mathrm{P}_0, \mathrm{P}_1, \ldots, \mathrm{P}_{m_1-1}\}$. Each node is represented by a 2-tuple (i, j), where $i \in [0, m_1 - 1]$ represents the cluster index and $j \in [0, m_0 - 1]$ represents the node index within the cluster. The *set of nodes* in cluster i, $0 \le i \le m_1 - 1$, is $\mathrm{P}_i = \{(i,0), (i,1), \ldots, (i, m_0 - 1)\}$. The channels are partitioned into W disjoint *channel sets* where each set consists of m_0 channels. The *set of all network channels* is the union of the disjoint channel sets, $\Lambda = \{\Lambda_0, \Lambda_1, \ldots, \Lambda_{W-1}\}$, where *channel set* i, $0 \le i \le W - 1$, is represented by $\Lambda_i = (\lambda_{i0}, \lambda_{i1}, \ldots, \lambda_{i(m_0-1)})$. If Λ_l is assigned to cluster P_k, then node $(k, i) \in \mathrm{P}_k$ transmits over channel λ_{li}, for all $0 \le i \le m_0 - 1$. The number of channel sets, W, required for conflict-free reception is determined by the CIN topology. Finding the lowest such number for a given topology is the subject of Section 4.

Clusters are interconnected according to a regular graph $\mathcal{H}$, whose vertices are the clusters in P and edges are the fiber links from output to input couplers. In addition to the links specified by $\mathcal{H}$, each cluster P_i possesses a self link $(\mathrm{P}_i, \mathrm{P}_i)$ between its own output and input couplers. Self cluster links enable arbitrary wavelength connectivity among nodes in the same cluster. In the considered configuration the CIN topology $\mathcal{H}$ has equal in and out degrees, denoted as G. The cluster degree, defined as the number of fiber links emanating from an output coupler or incident to an input coupler, is $G + 1$. Therefore, the dimensions of output and input couplers are $(m_0 : G + 1)$ and $(G + 1 : m_0)$, respectively.

Fig. 1 illustrates the network configuration for $m_1 = 4$ and $m_0 = 4$. The required number of channel sets for conflict-free communication depends on the CIN topology, and is studied in detail in Section 4. Since $G + 1$ fiber links are fanned-in to each input coupler, and each fiber carries a distinct channel set, the number of channel sets for any CIN topology must satisfy $W \ge G + 1$. Given a total of C channels in the system, and since each channel set consists of m_0 distinct channels, the cluster size is limited by

$$m_0 \le C/W \tag{1}$$

3 Uniform Emulation Principle

Realizing and reconfiguring network partitions is achieved by applying the principle of uniform graph emulation, described in the following. A structurally uniform emulation is defined in [12] as a structure preserving mapping of the nodes of a larger network to the nodes of a smaller network. Every node in the small network is assigned a number of nodes from the larger emulated network such that the larger network connectivity is preserved. Applying this concept, in a reverse sense, the smaller network corresponds to the physical multi-domain structure and the emulated network corresponds to the virtual point-to-point network. Nodes at the same relative position at all levels are fully connected in the wavelength domain since they belong to the same cluster. Multiple levels can be combined by any valid number, and in any order (not necessarily in the order of node labeling within each cluster). This is a particularly useful feature when sub-networks (blocks of nodes) are deallocated since no fragmentation will result. There are two distinguishing factors in M-WDM: 1) Emulation is done by clusters of nodes, and therefore every node emulates only one node of the virtual network; and 2) Cluster self-links become necessary in most cases to support communication among nodes covered by the same cluster. This corresponds to the local communication that would take place within a node among the processors which it emulates in the standard emulation of [12].

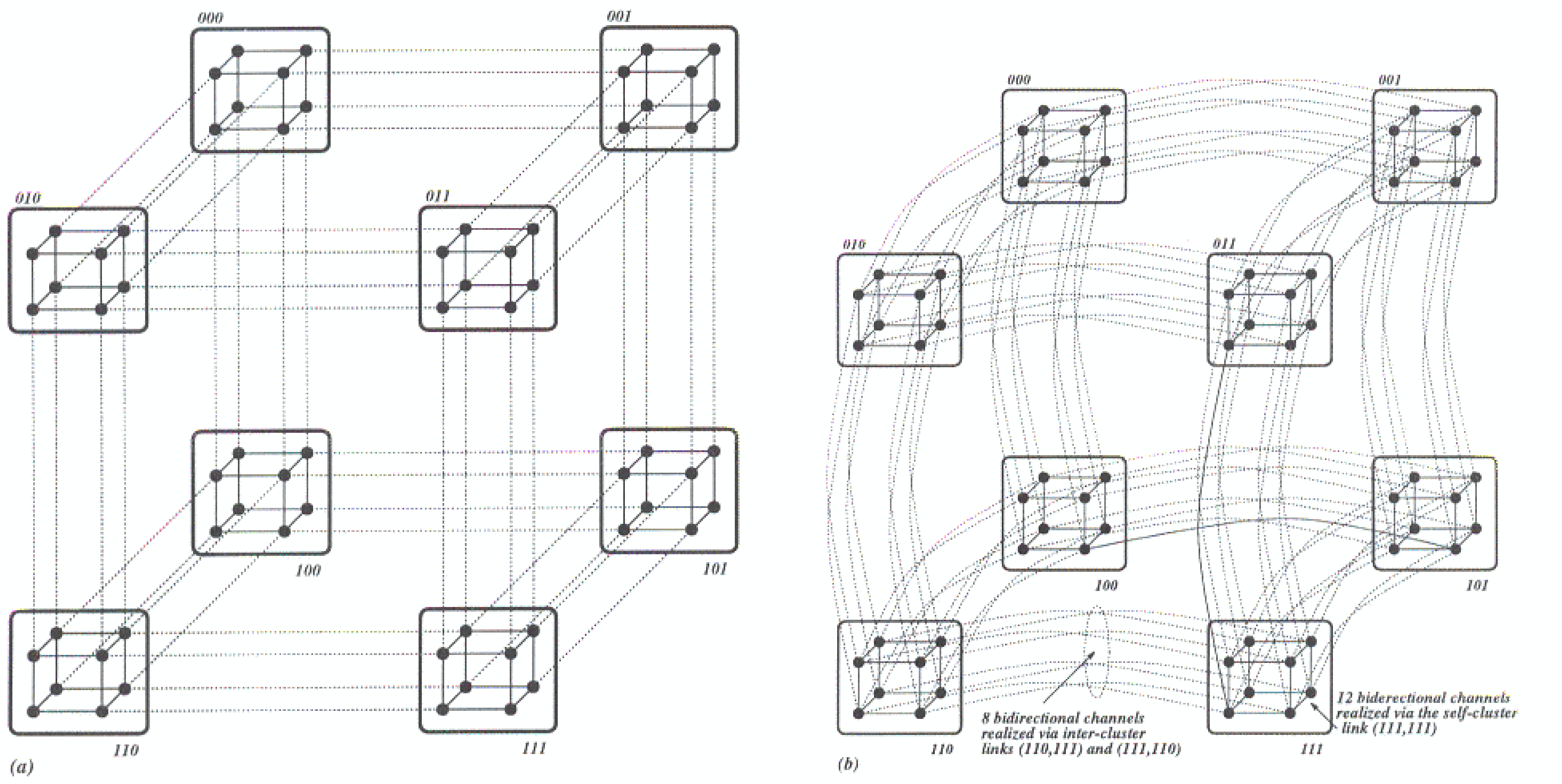

Figure 2: Virtual network emulations by an 8-cluster 3-dimensional binary cube M-WDM physical network: (a) 3-dimensional 4-ary cube (wrap-around connections not shown), (b) 6-dimensional binary cube. Virtual intra-cluster and intercluster connections are drawn in solid and dotted lines, respectively.

It is necessary to point out the distinction between uniform emulations and the special case of coverings. Let the host (physical) and the guest (virtual) networks be defined by the graphs $\mathcal{H} = (\mathrm{P}, \mathrm{E})$ and $\mathcal{H}' = (\mathrm{P}', \mathrm{E}')$, respectively. The map $f : \mathcal{H}' \rightarrow \mathcal{H}$ is an *emulation* of $\mathcal{H}'$ by $\mathcal{H}$ if and only if $(v, w) \in \mathrm{E}'$ implies that $(f(v), f(w)) \in \mathrm{E}$. The emulation is called *computationally uniform* if $|f^{-1}(x)| = |f^{-1}(y)|$ for all $x, y \in \mathrm{P}$. For simplicity, the term *uniform emulation* is used in this paper to indicate an emulation that is both structure preserving and computationally uniform. A uniform emulation is a *covering* if for all $v \in \mathrm{P}'$, the images of the neighbors of v by the map f are the distinct neighbors of $f(v)$. It can be noted from the definitions that every covering is a uniform emulation. A uniform emulation requires cluster self links while a covering does not. Example uniform emulations of a 3-dimensional 4-ary cube and a 6-dimensional binary cube by an M-WDM network whose CIN topology is a 3-dimensional binary cube are shown in Fig.s 2(a) and (b), respectively. In both figures, the virtual connections within the cluster indicates that they are emulations and not coverings.

For a partition with N_1 sub-networks, a given uniformly emulated sub-network consists of $N_0 = m_1 m_{v0}$ nodes, where m_{v0} indicates the number of participant nodes from each cluster (number of levels). The emulated sub-network is denoted as $\mathcal{H}_{n_v,k_v}$ It has the same topology of the physical network, $\mathcal{H}$, with virtual parameters $n_v \geq n$ and $k_v \geq k$. The *emulation map* $f(v) = u$ assigns all nodes v in the virtual network to cluster u.

Table 1 provides a summary of the emulation rules for the examined CIN topologies. The rest of this section defines the topologies and explains the table entries.

Table 1: Valid M-WDM network partitions

Physical network (n,k)	Virtual network (n_v,k_v)	Level size, m_1	Virtual network size, $N_0 m_1$	Number of levels, N_0/m_1	Emulation / covering map
$\mathcal{S}_{n,k}$	$\mathcal{S}_{n,k+i}$	n^k	n^i	$n^{(k+i)}$	$f(b_{k+i-1}\dots b_1 b_0)=(b_k\dots b_2 b_1)$
	shuffleNet	n^k	$(k+i)n^i$	$(k+i)n^{(k+i)}$	$f(c,b_{k+i-1}\dots b_1 b_0)=(c,b_k\dots b_2 b_1)$
$\mathcal{Q}_{n,k}$	$\mathcal{Q}_{n+i,k}$	k^n	k^{n+i}	k^i	$f(b_{n_v-1}\dots b_0)=(b_{n_v-1}\dots b_{n_v-n})$
	$\mathcal{Q}_{n,ik}$	k^n	$(ik)^n$	i^n	$f(b_{n_v-1}\dots b_0)=(\lfloor\frac{b_{n_v-1}}{i}\rfloor\dots\lfloor\frac{b_0}{i}\rfloor)$
$\mathcal{C}_n$	$\mathcal{C}_{2in}$	$n2^n$	$2in2^{2in}$	$2i2^{n(2i-1)}$	for $i=1$: $f(c,b_{n_v-1}\dots b_0)=\left(\lfloor\frac{c+1}{2}\rfloor, b_{n_v-1}b_{n_v-3}\dots b_3 b_1\right)$
$\mathcal{T}_{n,k}$	$\mathcal{T}_{\frac{n}{i},jk+1}$	$\frac{n^k-1}{n-1}$	$\frac{n_v^j-n_v}{n_v-1}$	$\frac{n_v^{k_v}-1}{n_v-1}$	$f(l,x)=(l_v,\lfloor\frac{xi^{l_v}}{n^{l-l_v}}\rfloor)$, $l_v=\lfloor\max(l-1,0)\rfloor$

3.1 (n,k) De Bruijn (Shuffle Stage)

$\mathcal{S}_{n,k}$ consists of n^k clusters, each represented by a k-digit n-ary tuple $(b_{k-1}\dots b_0)$. The interconnection is defined by an (n,k) shuffle permutation acting on individual cluster shuffle ports $0,\dots,n-1$ according to the map $\pi_{(n,k),j}(x)$, which denotes the j^{th} image of cluster x in an (n,k) shuffle stage.

$$\pi_{(n,k),j}(b_{k-1}\dots b_0)=(b_{k-2}\dots b_2 j) \tag{2}$$

where $0\le j\le n-1$. The set of images of cluster $(b_{k-1}\dots b_0)$ by the shuffle permutation is

$$\Pi_{(n,k)}(x)=\{\pi_{(n,k),j}(x)\mid\forall j\in[0,n-1]\} \tag{3}$$

The map in Table 1 corresponding to the emulation of $\mathcal{S}_{n,k+i}$ by $\mathcal{S}_{n,k}$ is shown to be a covering by showing that: 1) every $\mathcal{S}_{n,k+i}$-image of a node x of the virtual network is mapped to a cluster that belongs to the set of $\mathcal{S}_{n,k}$-images of $f(x)$; and 2) exactly n^i distinct nodes of the virtual network are mapped to each $\mathcal{S}_{n,k}$ cluster. The first statement is true if:

$$f(\pi_{(n,k+i),a}(b_{k+i-1}\dots b_0))\in\Pi_{n,k}(f(b_{k+i-1}\dots b_0)) \tag{4}$$

for all $a\in[0,n-1]$. The left expression is reduced by definition of the shuffle permutation (Eqn. 2) to

$$f(b_{k+i-2}\dots b_0 a)=(b_{k-1}\dots b_0) \tag{5}$$

and the right expression is evaluated by the given map (in the table) to

$$\Pi_{n,k}(b_k\dots b_1)=\{(b_{k-1}\dots b_1 j)\mid\forall j\in[0,n-1]\} \tag{6}$$

Eqn.s 5 and 6 show that the relation of Eqn. 4 is true for all $a\in[0,n-1]$. This is since for all values of $b_0\in[0,n-1]$ in the RHS of Eqn. 5, there is a cluster in $\Pi_{n,k}(b_k\dots b_1)$ given an equal value of $j\in[0,n-1]$ in Eqn. 6. The map of Table 1 shows that n^i distinct nodes are mapped to each cluster, which corresponds to the difference of i digits between the two tuples $(b_{k+i-1}\dots b_1 b_0)$ and $(b_k\dots b_2 b_1)$.

An (n,k_v) shuffleNet [13] consists of k_v columns of n^{k_v} nodes, every column is connected to the next as a shuffle stage, including the last and first column as a wrap-around stage, with a total of $k_v n_v^{k_v}$ nodes. The map used for emulating an $(n,k+i)$ shuffle stage can be similarly used for each shuffleNet stage if the nodes of column $c\in[1,k+i]$ are represented by the mixed tuple $(c,b_{k+i-1}\dots b_0)$. The proof given above for a shuffle stage shows as well that the $k+i$-stage shuffleNet is covered since the map is independent of c.

A 2-D shuffle, $(n_{v_1}, k_{v_1}, n_{v_2}, k_{v_2})$, consists of $n_{v_2}^{k_{v_2}}$ (n_{v_1}, k_{v_1}) shuffle stages across the first dimension and $n_{v_1}^{k_{v_1}}$ (n_{v_2}, k_{v_2}) shuffle stages across the second dimension, with a total of $N_0 = n_{v_1}^{k_{v_1}} n_{v_2}^{k_{v_2}}$ nodes. It has been considered (as a 3-D multi-stage interconnection pattern) for 3-dimensional free space optical implementation for applications in digital optical computing and signal processing [14,15]. The realization of 2-D shuffles is significantly facilitated in the M-WDM shuffle structure due to the use of wavelength as a third dimension through the cluster self links. One of the two shuffle dimensions is emulated by the space shuffle CIN as an ordinary shuffle stage. The second orthogonal shuffle dimension is realized through the self links , which provide the depth of the interconnection pattern. This emulation is directly obtained by defining the wavelength connections since it is not restricted by the physical space connectivity. The number of nodes in each orthogonal shuffle stage is limited by the number of nodes per cluster divided by the number of nodes from each cluster used to realize the first-dimension shuffles.

3.2 k-ary n-Cube

The k-ary n-cube CIN, $\mathcal{Q}_{n,k}$ has k^n clusters, k clusters along each dimension, and the cluster index is an n-digit k-ary tuple. Cluster $(b_{n-1} \ldots b_i \ldots b_0)$ is connected to clusters $(b_{n-1} \ldots b_i' \ldots b_0)$ for all $b_i' = b_i \pm 1 \bmod k$. It was shown in [16] that there exists no coverings of a cube by a smaller one. However, emulations of a cube with larger-dimension and larger number of nodes per dimension were given in [12]. The maps are directly derived by dividing the virtual network axes or dimensions, Fig. 2.

3.3 Cube-Connected Cycles

The $\mathcal{C}_n$ CIN with $n2^n$ clusters consists of 2^n cycles, each connecting n clusters, $m_1 = n2^n$ [12]. Each cluster can be represented by a mixed tuple $(c, b_{n-1}, \ldots, b_0)$, where $c \in [0, n-1]$ is the cluster index within a cycle and the binary tuple $(b_{n-1}, \ldots, b_0)$ is the binary cube cycle index. The binary $\mathcal{C}_n$, as originally defined in [17] has 2^n clusters: 2^{n-r} cycles, each with 2^r clusters, where r is the smallest integer such that $r + 2^r \geq n$. The first definition is simpler and its emulation map which has been first derived in [12] is sketched in the following. Virtual cube-connected cycles whose dimension is an even multiple of the physical network dimension can be emulated according to the indicated map, which is given for an emulation with twice the dimension ($i = 1$). Consider an edge of the virtual network (u_1, u_2). One of the corresponding cluster indexes, c_1 and c_2, must be even and the other odd. Assume $c_1 = 2j$: if $c_2 = 2j-1$ then the cycle index in both $f(u_1)$ and $f(u_2)$ is $c = j$ and hence $f(u_1) = f(u_2)$, otherwise $c = 2j+1$ in which case the cycle index in $f(u_2)$ is $f(c_2) = j+1$ and hence $(f(u_1), f(u_2))$ is an edge in the physical host network.

3.4 k-ary Tree

Let the cluster index in $\mathcal{T}_{n,k}$ CIN be given by (l, x) where $l \in [0, k]$ is the level indicator (with the root being at level-0 and the leaves at level k) and $x \in [1, n^l]$ is the index within level l. Uniform tree emulation by M-WDM trees of equal or higher degree is analyzed in the following. Let $n = in_v$ and $k_v = jk+1$, the relation of i and j that satisfies a uniform emulation can be found. The root cluster covers $j+1$ tree levels, and therefore $(n_v^{j+1} - 1)/(n_v - 1)$ nodes. Level j of the virtual tree emulated by the root cluster has n_v^j nodes, requiring n_v^{j+1} subordinate trees. There are n subordinate clusters, each needs to cover $n_v^{j+1}/n = n_v^j/i$ subordinate trees. Each subordinate tree has j levels, i.e. $(n_v^j - 1)/(n_v - 1)$ nodes. For uniform emulation, all clusters must contain an equal number of nodes, which is one less the number of nodes contained by the root cluster. That is, each cluster must cover $n_v(n_v^j - 1)/(n_v - 1)$ nodes. So, for uniform emulation

$$\frac{n_v(n_v^j - 1)}{n_v - 1} = \frac{n_v^j}{i} \times \frac{n_v^j - 1}{n_v - 1}$$

or

$$i = n_v^{j-1} \tag{7}$$

4 Conflict-Free Channel Assignment

To guarantee the uniqueness of channel sets used for transmission by clusters having a common adjacent cluster, distinct channel sets are assigned for transmission by any two clusters with a communication

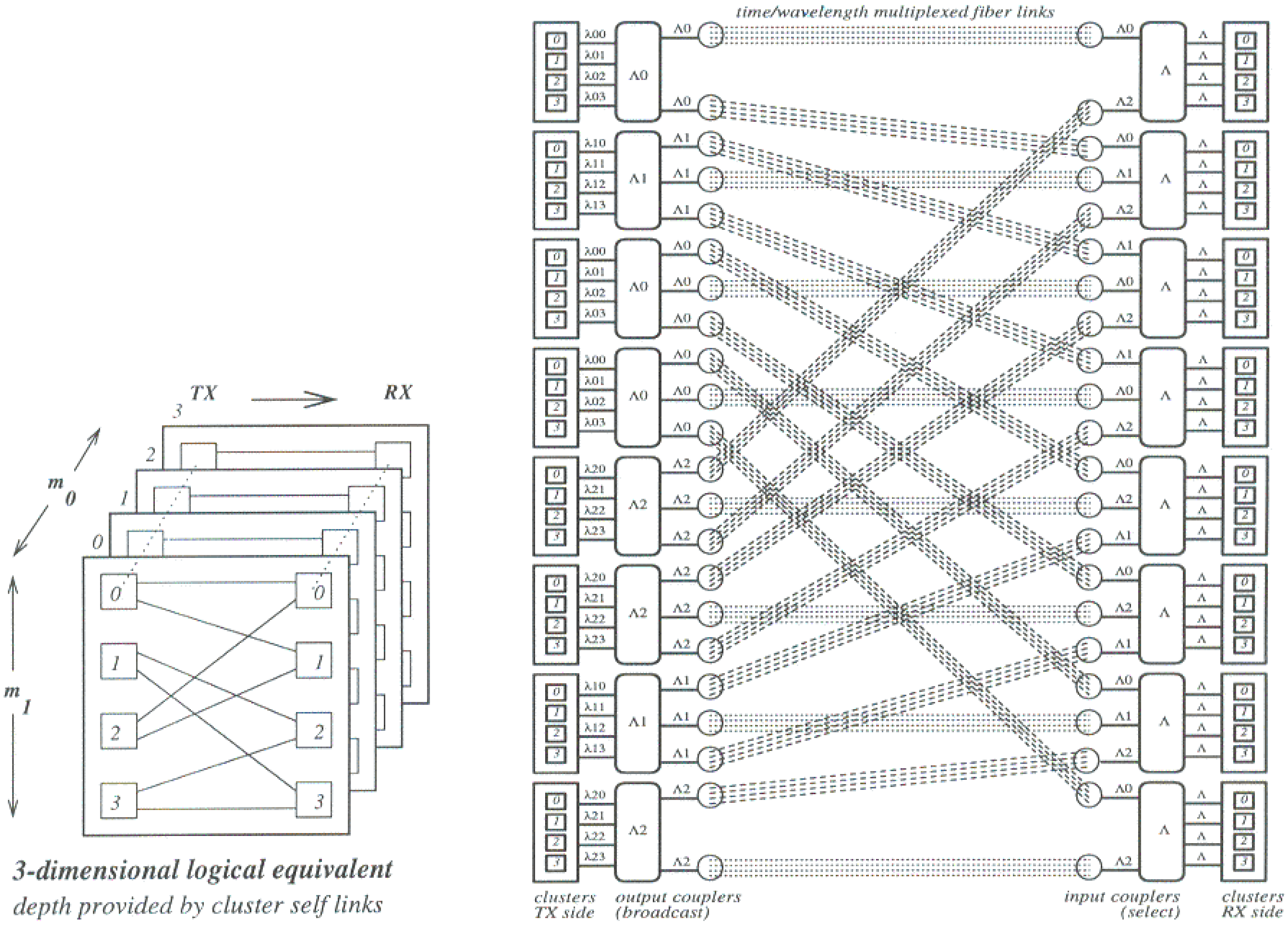

Figure 3: Channel assignment to a (2,3) M-WDM network: Clusters 0, 2 and 3 transmit over $\Lambda_0 = (\lambda_{00}, \lambda_{01}, \lambda_{02}, \lambda_{03})$, clusters 1 and 6 transmit over $\Lambda_1 = (\lambda_{10}, \lambda_{11}, \lambda_{12}, \lambda_{13})$, and clusters 4, 5 and 7 transmit over $\Lambda_2 = (\lambda_{20}, \lambda_{21}, \lambda_{22}, \lambda_{23})$. Every fiber link is shown as 4 lines, each representing a channel. Input couplers receive conflict-free the set of all channels $\Lambda = \Lambda_0 \cup \Lambda_1 \cup \Lambda_1$. Passive individual channel selection can be done at the cluster or node level.

distance not greater than 2 in the CIN graph. Two clusters separated by a distance of 1 have a common neighbor due to self cluster links and two clusters separated by a distance of 2 have a thrid common neighbor. This problem is different from the standard graph vertex coloring problem where different colors are assigned only to adjacent vertices. It is stated as follows: Find W disjoint sets of clusters Q_0, Q_1, ..., Q_{W-1}, where

$$Q_i = \{P_a, P_b, \text{ for } a, b \in [0, m_1 - 1] \mid d_{ab} > 2\} \tag{8}$$

Channel set Λ_i is then assigned to all clusters in Q_i for all $0 \leq i \leq W - 1$.

4.1 (n, k) Shuffle

When the self link is absent, only n channel sets are required, as can be noted from the following assignment:

$$Q_i = \{n^{k-1}i, \ldots, n^{k-1}(i+1) - 1\}, \quad 0 \leq i \leq n - 1 \tag{9}$$

When self links are present, the set of clusters Q_i that may be assigned a common channel set is

$$Q_i = \{a, b \in [0, n^k - 1] \mid a \notin \Pi_n(b), b \notin \Pi_n(a), \text{and } \Pi_n(a) \cap \Pi_n(b) = \Phi\} \tag{10}$$

Expressions of the partition of the set of clusters into $n + 1$ disjoint sets, each to be assigned a common channel set, have been derived in detail in [18] and are given below for a general (n, k) M-WDM shuffle with self links. For $0 \leq i \leq n-1$:

$$Q_i = \begin{cases} \left\{ \frac{i(n^k - n^{k-(2j+1)})}{n-1}, \ldots, \frac{i(n^k - n^{k-2(j+1)})}{n-1} - 1, \right. \\ \left. \frac{i(n^k - n^{2j})}{n-1} + n^{2j}, \ldots, \frac{i(n^k - n^{2j+1})}{n-1} + (n^{2j+1} - 1) \mid j \in [0, k/2 - 1] \right\} & \text{if } k \text{ is even} \\ \left\{ \frac{i(n^k - n^{k-(2j+1)})}{n-1}, \ldots, \frac{i(n^k - n^{k-2(j+1)})}{n-1} - 1, \frac{i(n^k - 1)}{n-1}, \right. \\ \left. \frac{i(n^k - n^{2j+1})}{n-1} + n^{2j+1}, \ldots, \frac{i(n^k - n^{2(j+1)})}{n-1} + n^{2(j+1)} - 1 \mid j \in [0, (k-3)/2] \right\} & \text{if } k \text{ is odd} \end{cases} \tag{11}$$

and the last set

$$Q_n = P - \bigcup_{i=0}^{n-1} Q_i \tag{12}$$

The orders of the $n+1$ disjoint sets are obtained from the lengths of the uninterrupted sequences within each set:

$$|Q_i| = \begin{cases} \lfloor n^k/(n+1) \rfloor & k \text{ even}, 0 \leq i \leq n-1 \\ & \text{or } k \text{ odd}, i = n \\ \lceil n^k/(n+1) \rceil & k \text{ odd}, 0 \leq i \leq n-1 \\ & \text{or } k \text{ even}, i = n \end{cases} \tag{13}$$

Reference can be made to Fig. 3 as an example assignment when $n = 2$, where 3 channel sets are employed. The results for $\mathcal{S}_{n,k}$ are summarized:

$$W(\mathcal{S}_{n,k}) = \begin{cases} n+1 & \text{with self cluster links} \\ n & \text{without self cluster links} \end{cases} \tag{14}$$

4.2 k-ary n-Cube

By establishing an analogy between the conflict-free channel set assignment problem in a k-ary n-cube and single-error-correcting Hamming codes, an upper bound on the number of channel sets can be obtained for any k (Section 4.3.1) and an exact result can be obtained for $k = 2$ (Section 4.3.2). An exact result is also obtained for the special case of $k = 3$ in Section 4.3.3.

4.2.1 General Upper Bound

A k-ary (n,p) Hamming code is a linear block code with $n-p$ information digits and p parity digits in each of its k^p codewords. Provided that the Galois field $GF(k)$ exists, k^{n-p} such codes can be constructed for a given (n,k) where $p = \lceil \log_k n(k-1) + 1 \rceil$ ($GF(k)$ exists when $k = \alpha^\beta$, where α is a prime and β is a positive integer) [19]. The code rate, defined as the ratio of information-to-parity digits, is maximum when $n(k-1)+1$ is a power of k, and the code is then called a perfect code. This algebraic structure is visualized by k^p unity-radius non-interfering spheres. The centerpoints of all the spheres in a code, separated by a minimum Hamming distance of 3, are the valid codewords. The points over the surface of each sphere, are those invalid codewords with a unity distance from the centerpoint, which will be corrected to its value. In a perfect code, these non-interfering unity-radius spheres completely fill the space, i.e. no combination of parity bits represents a non-valid codeword.

The vertices of $\mathcal{Q}_{n,k}$ correspond to the n^k radix-k tuples in the vector space of $GF(k)$. The minimum communication distance between two clusters $(b_{n-1} \ldots b_0)$ and $(b'_{n-1} \ldots b'_0)$ can be expressed as $\sum_{i=0}^{n-1} (b_i - b'_i) \bmod k$. The Hamming distance is the number of different digit positions between two codeworks. Communication distance and Hamming distance are equivalent only if $k = 2$. For $k \geq 3$, a code (assignment) satisfying the Hamming distance condition also satisfies the minimum communication distance of 3. Hence, the upper bound on the number of required channel sets is the number of spheres in the corresponding Hamming code (provided $GF(k)$ exists):

$$\widehat{W}(\mathcal{Q}_{n,k}) = k^{\lceil \log_k n(k-1)+1 \rceil}, \quad n > 1 \tag{15}$$

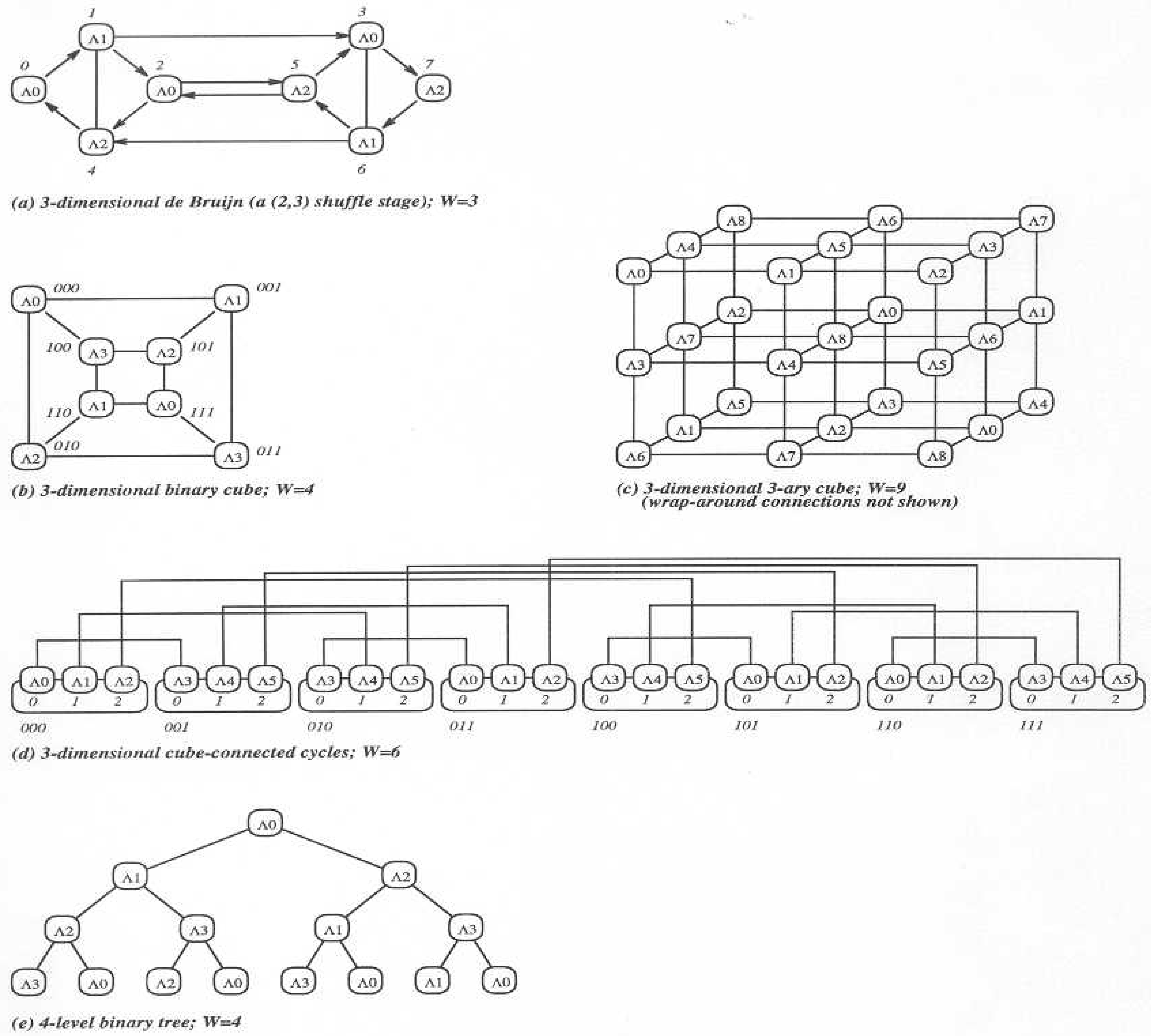

Figure 4: Conflict-free wavelength channel set assignment to clusters for considered CIN topologies.

4.2.2 Assignment when $k = 2$

When $k = 2$, Hamming and communication distance are the same, therefore

$$W(\mathcal{Q}_{n,2}) = 2^{\lceil \log_2 n+1 \rceil} \tag{16}$$

which reduces to $W = n + 1$ when 2^n is a dividend of $n + 1$. Assignment for a 3-D binary cube is shown in Fig. 4(b).

4.2.3 Assignment when k is a Multiple of 3

To show that 9 labels are sufficient for $\mathcal{Q}_{3,3}$, let each label be represented by a 3-ary 2-tuple $(l_1 l_0)$. A label is assigned to a cluster $(b_2 b_1 b_0)$ according to the map: $f(b_2 b_1 b_0) = (b_2 \oplus b_1, b_2 \oplus b_0)$, where $'\oplus'$ represents modulo 3 addition. Clusters within a distance of 1 or 2 from the one being considered are: $(a_2, a_1, a_0 \oplus 1)$, $(a_2, a_1, a_0 \oplus 2)$, $(a_2, a_1 \oplus 1, a_0)$, $(a_2, a_1 \oplus 2, a_0)$, $(a_2 \oplus 1, a_1, a_0)$, $(a_2 \oplus 2, a_1, a_0)$, $(a_2, a_1 \oplus 1, a_0 \oplus 1)$, $(a_2, a_1 \oplus 2, a_0 \oplus 2)$, $(a_2 \oplus 1, a_1, a_0 \oplus 1)$, $(a_2 \oplus 2, a_1, a_0 \oplus 2)$, $(a_2 \oplus 1, a_1 \oplus 1, a_0)$, $(a_2 \oplus 2, a_1 \oplus 2, a_0)$. By inspecting the f-images of the above tuples, it can be verified that non of them is equal to $(b_2 \oplus b_1, b_2 \oplus b_0)$ (note that they do not need be all distinct). The previous argument is extended for $k = 3i$ for all integers i and any value of n by applying the above map f. Assignment for a $\mathcal{Q}_{3,3}$ is shown in Fig. 4(c).

$$W(\mathcal{Q}_{n,3i}) = \widehat{W}(\mathcal{Q}_{2,3}), \quad i \text{ is a positive integer} \tag{17}$$

4.3 Cube-connected Cycles

The cube-connected cycles is functionally equivalent to the ShuffleNet and they can emulate each other with dilation 2 [20]. ShuffleNets can be easily emulated by an M-WDM network with a shuffle CIN, whose wavelength assignment involved a minimal number of channels. This makes the shuffle structure more diverse and efficient. An upper bound on the number of channel sets for the cube-connected cycles is obtained here by considering both the cycle and cube connections as follows.

The criterion for conflict-free assignment within a cycle (regardless of the cube connections) is that $\Lambda_i \neq \cdots \neq \Lambda_{i+2(\text{mod}n)}$. Therefore at least 3 sets are required. Let this initial assignment be given by: $(\Lambda_0, \Lambda_1, \Lambda_2) \leftrightarrow ((c,.),(c+1,.),(c+2,.))$. If $n(\text{mod}3) = 0$ the same ordered set of channel sets can be assigned to $((c+3l(\text{mod}n),.),(c+3l+1(\text{mod}n),.),(c+3l+2(\text{mod}n),.))$ for any integer l. If $n(\text{mod}3) = 1$, vertex $((n-1),.)$ must be assigned a distinct channel set Λ_3. If $n(\text{mod}3) = 2$, vertices $((n-2,.),(n-1,.))$ must both be assigned distinct channel sets (Λ_3, Λ_4). Therefore $W_c = 3 + n(\text{mod}3)$ is necessary for conflict-free assignment within the cycle.

The next step is to check the sufficiency of $2W_c$. Let $\Lambda = \{\Lambda_0, \ldots, \Lambda_{W_c-1}, \Lambda_{W_c}, \ldots, \Lambda_{2W_c-1}\}$. The cube assignment is such that if Λ_{i_c} is assigned to vertex $(c, b_{n-1}, \ldots, b_c, \ldots, b_0)$, then $\Lambda_{i_c+W_c(\text{mod}W_c)}$ is to be assigned to vertex $(c, b_{n-1}, \ldots, \bar{b}_c, \ldots, b_0)$ (the two vertices across the c^{th} dimensional axis). Since $\Lambda_{i_c} \neq \Lambda_{i_c+W_c(\text{mod}W_c)}$ and $\Lambda_{i_c} \neq \Lambda_{i_{c-1}} \neq \Lambda_{i_{c+1}}$, then vertices $(c, b_{n-1}, \ldots, b_c, \ldots, b_0)$, $(c-1, b_{n-1}, \ldots, \bar{b}_c, \ldots, b_0)$, and $(c+1, b_{n-1}, \ldots, \bar{b}_c, \ldots, b_0)$ are uniquely assigned. Therefore

$$\widehat{W}\,[\mathcal{C}_n] = 2[3 + n(\text{mod}3)] \tag{18}$$

Channel set assignment for $\mathcal{C}_3$ is illustrated in Figure 4(d).

4.4 k-ary Tree

A two-level tree ($k = 1$) can be assigned $n+1$ channel sets without conflicts. The following is a conflict-free assignment for $\mathcal{T}_{n,k}$, $k \geq 2$, that uses exactly one more channel set in addition to the original $n+1$ sets assigned to the 2-level tree emanating at the root. $\mathcal{T}_{n,k}$, $k \geq 2$, can be recursively constructed from the $n+1$ 2-level sub-trees $\text{R}, \text{S}_1, \ldots, \text{S}_n$, representing a root sub-tree and its n children sub-trees, respectively. A 2-level n-ary sub-tree can be represented in general by the ordered set $((l,x),(l+1,nx),\ldots,(l+1,(n+1)x-1))$. Assuming that the sub-tree whose root cluster index is the tree root has been assigned $(\Lambda_i, \Lambda_{(i+1)\text{mod}(n+1)}, \ldots, \Lambda_{(i+n)\text{mod}(n+1)})$, the assignment is then recursively made to the sibling clusters of each sub-tree starting from the sub-trees whose roots are the level-1 clusters as follows:

$$\text{S}_j \leftarrow (\Lambda_{(i+j)\text{mod}(n+1)}, \Lambda_{(i+j+1)\text{mod}(n+1)}, \ldots, \Lambda_{(i-1)\text{mod}(n+1)}, \Lambda_{(i+1)\text{mod}(n+1)}, \ldots, \Lambda_{(i+j-1)\text{mod}(n+1)}) \tag{19}$$

It follows from the given assignment that $n+2$ channel sets are required and sufficient when k is 2 (i.e. 3 levels) or more. To summarize:

$$W(\mathcal{T}_{n,k}) = \begin{cases} n+1 & k=1 \\ n+2 & k \geq 2 \end{cases} \tag{20}$$

Channel set assignment for $\mathcal{T}_{2,3}$ is illustrated in Figure 4(e).

5 Conclusion

This paper proposed a structure for realizing scalable WDM interconnection networks with passive reconfigurable partitioning. Applications are found in massively parallel computers as well as multicomputer and local/metropolitan area networks. The concept is based on uniform graph emulations and has been applied to a number of generalized regular topologies: shuffle, cube, and tree based. Emulation maps and partitioning rules were given. Conflict-free channel assignment was considered for the mentioned topologies. It can be concluded that the shuffle realization is most efficient in terms of its fiber utilization, since the number of required channels is proportional to the network degree. Shuffle-based interconnection has very desirable characteristics for multi-hop packet communication. Partitioning the interconnection fabric in a passive way can be used to allocate proper switching resources to different traffic types in a B-ISDN environment. Partitioning in a massively data parallel computer has the clear advantage of mapping the application to the most efficient machine size. Scalability and complexity advantages of the proposed

M-WDM implementation can be specially seen in terms of the wiring reduction for large hypercube networks. The paper has focused on the realization aspects of this concept, considering that the advantages of scalability and partitioning are generally well known.

References

[1] M. Karol and S. Shaikh, "A simple adaptive routing scheme for congestion control in shufflenet multihop lightwave networks," *IEEE Journal on Selected Areas of Communications*, vol. 9, pp. 1040–1051, Sept. 1991.

[2] B. Mukherjee, "Architectures and protocols for WDM-based local lightwave networks," *IEEE Network*, July 1992.

[3] D. Guo and A. Acampora, "Scalable multihop WDM passive ring with optimal wavelength assignment and adaptive wavelength routing," *IEEE Journal on Lightwave Technology*, vol. 14, pp. 1264–1277, June 1996.

[4] N. Dono, P. Green, K. Liu, R. Ramaswami, and F. Tong, "A wavelength division multiple access network for computer communication," *IEEE Journal on Selected Areas of Communications*, vol. 8, pp. 983–994, Aug. 1990.

[5] C. Qiao and R. Melhem, "Time-division optical communications in multiprocessor arrays," *IEEE Transactions on Computers*, vol. 42, pp. 577–590, May 1993.

[6] P. Lalwaney, L. Zenou, A. Ganz, and I. Koren, "Optical interconnects for multiprocessors: Cost-performance tradeoffs," in *Proc. IEEE Supercomputing'92*, pp. 278–285, Nov. 1992.

[7] P. W. Dowd, K. Bogineni, K. A. Aly, and J. Perreault, "Hierarchical scalable photonic architectures for high-performance processor interconnection," *IEEE Transactions on Computers*, vol. 42, pp. 1105–1120, Sept. 1993.

[8] T. Szymanski, "Hypermeshes: Optical interconnection networks for parallel computation," *Journal of Parallel and Distributed Computing*, 1993.

[9] P. Dowd *et al.*, "Lightning network and systems architecture," *IEEE Journal on Lightwave Technology*, vol. 14, pp. 1371–1387, June 1996.

[10] K. A. Aly and P. W. Dowd, "Parallel computer reconfigurability through optical interconnects," in *Proc. 22nd International Conference on Parallel Processing*, pp. I105 – I108, August 1992.

[11] K. A. Aly and P. W. Dowd, "WDM cluster ring: A low-complexity partitionable reconfigurable processor interconnection structure," in *Proc. 23rd International Conference on Parallel Processing*, pp. I150 – I153, August 1993.

[12] J. P. Fishburn and R. A. Finkel, "Quotient networks," *IEEE Transactions on Computers*, vol. C-31, pp. 288–295, Apr. 1982.

[13] M. G. Hluchyj and M. J. Karol, "ShuffleNet: An application of generalized perfect shuffles to multihop lightwave networks," *IEEE Journal on Lightwave Technology*, vol. 9, pp. 1386–1397, Oct. 1991.

[14] J. Giglmayr, "Classification scheme for 3-D shuffle interconnection patterns," *Applied Optics*, vol. 28, pp. 3120–3128, 1989.

[15] J. Giglmayr, "d-Dimensional ($d \geq 3$) Shuffle Interconnections," *Applied Optics*, vol. 31, pp. 1695–1708, Apr. 1992.

[16] H. Bodlaender, "The Classification of Coverings of Processor Networks," *Journal of Parallel and Distributed Computing*, vol. 6, pp. 166–182, 1989.

[17] F. P. Preparata and J. Vuillemin, "The cube-connected cycles: A versatile network for parallel computation," *Communications ACM*, pp. 300–309, May 1981.

[18] K. A. Aly, "Conflict-free channel assignment for an optical cluster-based shuffle network configuration," *Computer Communication Review*, vol. 24, no. 4, pp. 201–210, 1994.

[19] R. E. Blahut, *Theory and Practice of Error Control Codes*. Addison Wesley, 1984.

[20] F. T. Leighton, *Introduction to Parallel Algorithms and Architectures*. San Mateo, California: Morgan Kaufmann Publishers, 1992.