Strictly Nonblocking f-cast d-ary Multi-log Networks under Fanout and Crosstalk Constraints

Yang Wang, Hung Q. Ngo
Department of Computer Science and Engineering
State University of New York at Buffalo
Amherst, New York 14260
Email: {yw43,hungngo}@cse.buffalo.edu

Xiaohong Jiang
School of Information Sciences
Tohoku University,
Sendai, Japan
Email: jiang@ecei.tohoku.ac.jp

***Abstract*—We derive conditions which are both necessary and sufficient for the d-ary multi-log switching networks to be f-cast strictly nonblocking under all combinations of fanout and crosstalk constraints. The fanout constraint tells us which stage(s) of the networks has fanout capability. The crosstalk constraint tells us whether or not two connection routes are allowed to share a link (relevant to electronic switches), or are allowed to share a switching element (crosstalk-free or not, relevant to optical switches). Thus, for any given d and f, we completely characterize the d-ary multi-log network under the f-cast strictly nonblocking constraint, the link/node-blocking constraints, and the fanout constraints.**

The most novel contribution of this paper is the analytical technique, which combines an algebraic view of the d-ary multi-log network with the max-flow min-cut theorem. Our results are more general than previously known results on several fronts: (a) d-ary networks are more general than binary networks, (b) f-cast covers both unicast ($f = 1$) and broadcast ($f = N$), (c) both link-blocking and node-blocking are considered in a unified manner, and (d) all combinations of fanout constraints are considered.

Keywords: multicast, switches, f-cast, strictly nonblocking, d-ary multi-log switching networks, crosstalk-free, fanout constraints.

I. Introduction

Many current and future Internet applications demand multicast support. To support multicast efficiently, the switching networks which serve as the switching fabric architectures at the core of electronic routers, or as the switching topology for optical cross-connects must be multicast capable.

Current multicast switch designs mostly focus on the broadcast case [1]–[3]. Although broadcast switches are certainly capable of supporting multicast with any fanout requirement, they are not scalable due to their prohibitively high hardware requirement. Almost all the multicast applications are restricted to a group of users, where broadcasting is rarely required. Hence, allocating expensive broadcast capability to each network switch is cost-inefficient for most practical purposes. Moreover, from the viewpoints of resource fairness and network security (e.g., limiting virus and worm propagation), we have other good reasons to impose a restriction on the maximum fanout of each request.

Consequently, there have been some recent research efforts on designing and analyzing the so-called *f-cast* switches, in which the maximum fanout of each request is upperbounded

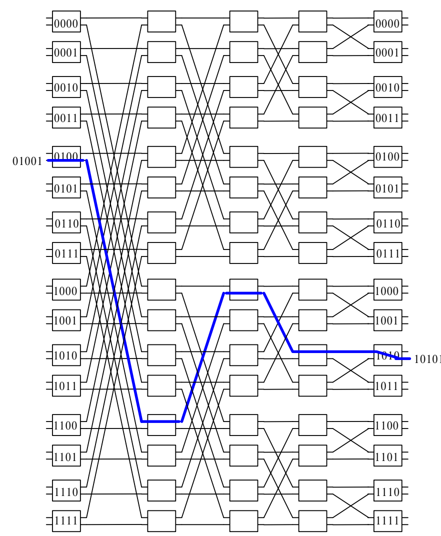

Fig. 1. The inverse Banyan network $\mathrm{BY}^{-1}(5)$

by the parameter f [4]–[8]. An f-cast switch usually requires significantly lower hardware cost than its broadcast counterpart. (This point shall be illustrated in this paper with the d-ary multi-log architecture.) Furthermore, a good design of an f-cast switch covers both the unicast design ($f = 1$) and the broadcast design ($f = N$) as special cases. Consequently, studying general f-cast switches is both mathematically pleasing and practically useful, as the results potentially can offer network designers more flexibility in selecting architectures for future multicast-intensive networks.

The design of a scalable and hardware-inexpensive switch usually employs the multistage architecture. The most popular multistage architectures are Banyan-type [9] and Clos-type [10] architectures. This paper focuses on analyzing the Banyan-type (e.g., see Figure 1). In particular, we study the general d-ary multi-log switch architecture with multiple vertically stacked inverse Banyan switches, as illustrated in Figure 2. The d-ary multi-log switches have been attractive for both electronic and photonic domains [11]–[17], because they have small depth ($O(\log N)$), absolute signal loss uniformity, and good fault tolerance. Hereafter, we use $\log_d(N, 0, m)$ to denote a d-ary multi-log switch with m vertically stacked

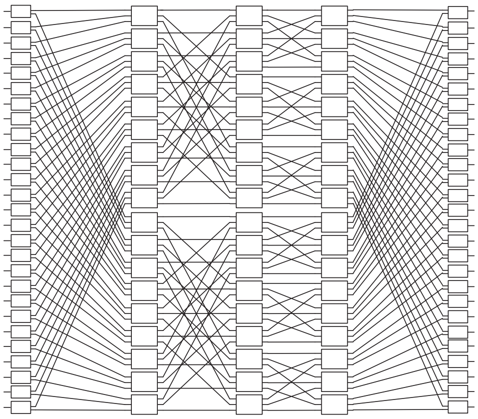

Fig. 2. Illustration of the $\log_3(27, 0, 2)$ network

inverse Banyan planes.

To support multicast (or generally f-cast) in a $\log_d(N, 0, m)$ switch, a certain degree of fanout capability must be provided in the switch. Similar to that of the fanout definitions for Clos network [5], a $1 \times m$ switch or a basic $d \times d$ switching elements (SE) in a $\log_d(N, 0, m)$ switch is said to have *fanout capability* if any one-to-many mapping between its inputs and outputs can be realized. We will say that the input stage (resp. the middle Banyan stage) of a $\log_d(N, 0, m)$ network has the fanout capability if each of its $1 \times m$ switches (resp. $d \times d$ SEs) has the fanout function. Thus, the fanout capability of a $\log_d(N, 0, m)$ switch can be provided either in its input stage, or in its central Banyan stage or in both. The additional fanout requirement of a stage usually makes its implementation more complex and costly than its unicast counterpart, so the fanout capability should be carefully allocated for the efficient design of a multicast-capable switch.

There are three levels of nonblockingness typically studied in the switching network literature: rearrangeably nonblocking (RNB), wide-sense nonblocking (WSNB), and strictly nonblocking (SNB). The reader is referred to [4] for their precise definitions. This paper focuses on analyzing the SNB f-cast $\log_d(N, 0, m)$ network.

This paper will fully investigate how the fanout constraints on the input stage and middle Banyan stage affect the cost of an f-cast SNB $\log_d(N, 0, m)$ switch. In addition to the fanout constraint, we also consider both the link-blocking and node-blocking constraints. Under the link-blocking constraint, only one request is allowed to use a link at one time, which is relevant to electronic switches [1]–[8], [13], [14], [16]. The node-blocking constraint allows only one request to use a switching element at one time, reducing the crosstalk effect in all-optical cross-connect designs [11], [12], [15], [17]–[19].

Our main contributions are as follows. We derive conditions which are both necessary and sufficient for the the $\log_d(N, 0, m)$ network to be f-cast SNB under all combinations of fanout and crosstalk constraints, where d and f are any given parameters. The most novel contribution of this paper is the analytical technique, which combines an algebraic view of the $\log_d(N, 0, m)$ network with the max-flow min-cut theorem. Our results are more general than previously known results on several fronts: (a) d-ary networks are more general than binary networks, (b) f-cast covers both unicast ($f = 1$) and broadcast ($f = N$), (c) both link-blocking and node-blocking are considered in a unified manner, and (d) all combinations of fanout constraints are considered.

The rest of the paper is organized as follows. Section II establishes basic notations and presents a simple algebraic view of $\log_d(N, 0, m)$ networks, which are used throughout the paper. Section III presents the necessary and sufficient conditions for the $\log_d(N, 0, m)$ network to be f-cast SNB under link/node-blocking constraints when only the central stage has fanout capability. Section IV presents the corresponding results when both stages have fanout capability. Section V presents the corresponding results when only the first stage has fanout capability. Section VI concludes the paper with a few remarks.

II. Preliminaries

We first establish notations which will be used throughout the paper. For any positive integers l, d, let

- $[l]$ denote the set $\{1, \ldots, l\}$;
- $\mathbb{Z}_d$ denote the set $\{0, \ldots, d-1\}$ which can be thought of as d-ary "symbols";
- $\mathbb{Z}_d^l$ denote the set of all d-ary strings of length l;
- b^l denote the string with symbol $b \in \mathbb{Z}_d$ repeated l times (e.g., $3^4 = 3333$);
- $|\mathbf{s}|$ denote the length of any d-ary string $\mathbf{s}$ (e.g., $|31| = 2$);
- $\mathbf{s}_{i..j}$ denote the substring $s_i \cdots s_j$ of a string $\mathbf{s} = s_1 \ldots s_l \in \mathbb{Z}_d^l$, when $j > i$ we agree on the convention that $\mathbf{s}_{i..j}$ is the empty string.

Let $N = d^n$. We consider the $\log_d(N, 0, m)$ network, which denotes the stacking of m copies of the d-ary inverse Banyan network $\text{BY}^{-1}(n)$ with N inputs and N outputs. We label the inputs and outputs of $\text{BY}^{-1}(n)$ with d-ary strings of length n. Specifically, each input $\mathbf{x} \in \mathbb{Z}_d^n$ and output $\mathbf{y} \in \mathbb{Z}_d^n$ have the form $\mathbf{x} = x_1 \cdots x_n$, $\mathbf{y} = y_1 \cdots y_n$, where $x_i, y_i \in \mathbb{Z}_d$, $\forall i \in [n]$.

Also, label the $d \times d$ SEs in each of the n stages of $\text{BY}^{-1}(n)$ with d-ary strings of length $n-1$. It is easy to see that an input $\mathbf{x}$ (resp. output $\mathbf{y}$) is connected to the SE labeled $\mathbf{x}_{1..n-1}$ in the first stage (resp. $\mathbf{y}_{1..n-1}$ in the last stage).

For the sake of clarity, let us first consider a small example. Consider the unicast request $(\mathbf{x}, \mathbf{y}) = (01001, 10101)$ when $d = 2, n = 5$. The input $\mathbf{x} = 01001$ is connected to the SE labeled 0100 in the first stage, which is connected to two SEs labeled 0100 and 1100 in the second stage, and so on. The unique path from $\mathbf{x}$ to $\mathbf{y}$ in $\text{BY}^{-1}(n)$ can be explicitly written out (see Figure 1):

input $\mathbf{x}$	01001
stage-1 SE	0100
stage-2 SE	**1**100
stage-3 SE	**10**10
stage-4 SE	**101**0
stage-5 SE	**1010**
output $\mathbf{y}$	10101

We can see clearly the pattern: the prefixes of $\mathbf{y}_{1..n-1}$ are "taking over" the prefixes of $\mathbf{x}_{1..n-1}$ on the path from $\mathbf{x}$ to $\mathbf{y}$. In general, the unique path from an arbitrary input $\mathbf{x}$ to an arbitrary output $\mathbf{y}$ is exactly the following:

input $\mathbf{x}$	$x_1x_2 \dots x_{n-1}x_n$
stage-1 SE	$x_1x_2 \dots x_{n-1}$
stage-2 SE	$y_1x_2 \dots x_{n-1}$
stage-3 SE	$y_1y_2 \dots x_{n-1}$
$\vdots$	$\vdots$
stage-n SE	$y_1y_2 \dots y_{n-1}$
output $\mathbf{y}$	$y_1y_2 \dots y_{n-1}y_n$

Now, consider two unicast requests $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{x}, \mathbf{y})$. In the node-blocking case, these two requests cannot be routed through the same copy of $\text{BY}^{-1}(n)$ if and only if the two corresponding paths intersect at some SE in the middle (if they were to be routed through the same copy). More precisely, $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{x}, \mathbf{y})$ are said to *node-block* each other if and only if there is some $j \in [n]$ such that $b_{1..j-1} = y_{1..j-1}$ and $a_{j..n-1} = x_{j..n-1}$. In this case, the two paths intersect at a stage-j SE. It should be noted that two requests' paths may intersect at more than one SE. In a $\log_d(N, 0, m)$ network, two requests which are node-blocking one another have to be routed through different copies of $\text{BY}^{-1}(n)$.

For any two d-ary strings $\mathbf{u}, \mathbf{v} \in \mathbb{Z}_d^l$, let $\text{PRE}(\mathbf{u}, \mathbf{v})$ denote the *longest common prefix*, and $\text{SUF}(\mathbf{u}, \mathbf{v})$ denote the *longest common suffix* of $\mathbf{u}$ and $\mathbf{v}$, respectively. For example, if $\mathbf{u} = 0100110$ and $\mathbf{v} = 0101010$, then $\text{PRE}(\mathbf{u}, \mathbf{v}) = 010$ and $\text{SUF}(\mathbf{u}, \mathbf{v}) = 10$. From the observation made in the previous paragraph, the following proposition is immediate.

Proposition II.1. *Let* $(\mathbf{a}, \mathbf{b})$ *and* $(\mathbf{x}, \mathbf{y})$ *be two unicast requests in a* $\log_d(N, 0, m)$ *network, then the two requests node-block one another if and only if*

$$|\text{SUF}(\mathbf{a}_{1..n-1}, \mathbf{x}_{1..n-1})| + |\text{PRE}(\mathbf{b}_{1..n-1}, \mathbf{y}_{1..n-1})| \geq n - 1. \quad (1)$$

In the link blocking case, two requests *link-block* each other if and only if they share a common link in $\text{BY}^{-1}(n)$ (if they were to be routed through the same copy). More precisely, two requests $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{x}, \mathbf{y})$ are said to *link-block* each other if and only if there is some $j \in [n]$ such that $b_{1..j-1} = y_{1..j-1}$, $a_{j..n-1} = x_{j..n-1}$, $b_{1..j} = y_{1..j}$, and $a_{j+1..n-1} = x_{j+1..n-1}$. The four conditions are equivalent to just to conditions $b_{1..j} = y_{1..j}$ and $a_{j..n-1} = x_{j..n-1}$. We easily obtain the link-blocking analog of Proposition II.1 as follows.

Proposition II.2. *Let* $(\mathbf{a}, \mathbf{b})$ *and* $(\mathbf{x}, \mathbf{y})$ *be two unicast requests in a* $\log_d(N, 0, m)$ *network, then the two requests link-block one another if and only if*

$$|\text{SUF}(\mathbf{a}_{1..n-1}, \mathbf{x}_{1..n-1})| + |\text{PRE}(\mathbf{b}_{1..n-1}, \mathbf{y}_{1..n-1})| \geq n. \quad (2)$$

III. ONLY THE MIDDLE STAGE HAS FANOUT CAPABILITY

The main results of this section are summarized in the following two theorems.

Theorem III.1. *Let* $r = \lfloor \log_d f \rfloor$*, and*

$$m_{\text{nb}}(n, f, d) = f(d^{\lceil \frac{n-r-1}{2} \rceil} - 1) + d^{n - \lceil \frac{n-r-1}{2} \rceil}. \quad (3)$$

Then, the necessary and sufficient condition for $\log_d(N, 0, m)$ *to be* f*-cast strictly nonblocking in the node-blocking sense is* $m \geq m_{\text{nb}}(n, f, d)$ *when* $f \leq d^{n-1} - d^{n-2}$ *and* $m \geq m_{\text{nb}}(n, d^{n-1} - d^{n-2}, d)$ *when* $f \geq d^{n-1} - d^{n-2} + 1$.

Theorem III.2. *Let* $r = \lfloor \log_d f \rfloor$ *and*

$$m_{\text{lb}}(n, f, d) = f(d^{\lceil \frac{n-r-2}{2} \rceil} - 1) + d^{n - \lceil \frac{n-r}{2} \rceil}. \quad (4)$$

Then, the necessary and sufficient condition for $\log_d(N, 0, m)$ *to be* f*-cast strictly nonblocking in the link-blocking sense is* $m \geq m_{\text{lb}}(n, f, d)$ *when* $f \leq d^{n-2}$ *and* $m \geq m_{\text{lb}}(n, d^{n-2}, d)$ *when* $f > d^{n-2}$.

The proofs of Theorems III.1 and III.2 are similar. Due to space limitation and for the sake of presentation clarity, we will present here only the proof of Theorem III.1. The presentation should be sufficient in conveying the main ideas. There are two conditions to be shown, the sufficiency and the necessary conditions, which are proved separately in the following lemmas.

Lemma III.3 (Sufficiency)**.** *Let* $r = \lfloor \log_d f \rfloor$*. Then, a sufficient condition for* $\log_d(N, 0, m)$ *to be* f*-cast strictly nonblocking in the node-blocking sense is* $m \geq m_{\text{nb}}(n, f, d)$ *when* $f \leq d^{n-1} - d^{n-2}$ *and* $m \geq m_{\text{nb}}(n, d^{n-1} - d^{n-2}, d)$ *when* $f \geq d^{n-1} - d^{n-2} + 1$*, where the function* m_{nb} *is defined in* (3).

Proof: Suppose the network already has established routes for some set $\mathcal{R}$ of requests. Let $R = (\overline{\mathbf{x}}, \overline{Y})$ be a new multicast request compatible with the current network state, where $\overline{\mathbf{x}}$ is some input and $\overline{Y}$ is a subset of l outputs, $l \leq f$. For the network to be strictly nonblocking, we must be able to find a copy of $\text{BY}^{-1}(n)$ to route R. Let $b(\mathcal{R})$ be the number of requests in $\mathcal{R}$ which node-block R (i.e. cannot be routed through the same copy of $\text{BY}^{-1}(n)$ with R). Then, R is routable if and only if $m \geq b(\mathcal{R}) + 1$.

Firstly, consider the case when $f \leq d^{n-1} - d^{n-2}$. We will derive an upper bound B for $b(\mathcal{R})$. The upper bound B will be independent of $\mathcal{R}$ and R. Then, $m \geq B + 1$ will be a sufficient condition for $\log_d(N, 0, m)$ to be strictly nonblocking.

Since the first stage does not have fanout capacity, a request $(\mathbf{x}, Y)$ node-blocks R if and only if there is some $\mathbf{y} \in Y$ such that the unicast request $(\mathbf{x}, \mathbf{y})$ node-blocks R. Hence, without loss of generality, we can assume that $\mathcal{R}$ consists entirely of unicast requests.

To this end, write $\overline{Y} = \{\overline{\mathbf{y}}^{(1)}, \dots, \overline{\mathbf{y}}^{(l)}\}$. In order for a request $(\mathbf{x}, \mathbf{y})$ to node-block R, $(\mathbf{x}, \mathbf{y})$ must node-block the request $(\overline{\mathbf{x}}, \overline{\mathbf{y}}^{(p)})$ for some $1 \leq p \leq l$. Thus, let us elaborate on which $(\mathbf{x}, \mathbf{y})$ may node-block $(\overline{\mathbf{x}}, \overline{\mathbf{y}}^{(p)})$.

For each $i \in \{0,\dots,n-1\}$, let X_i be the set of inputs $\mathbf{x}$ other than $\overline{\mathbf{x}}$, where $\mathbf{x}_{1..n-1}$ shares a *suffix* of length exactly i with $\overline{\mathbf{x}}_{1..n-1}$. Formally, define

$$X_i := \{\mathbf{x} \in \mathbb{Z}_d^n - \{\overline{\mathbf{x}}\} \mid \text{SUF}(\mathbf{x}_{1..n-1}, \overline{\mathbf{x}}_{1..n-1}) = i\}.$$

Similarly, for each $j \in \{0,\dots,n-1\}$, let $Y_j^{(p)}$ be the set of outputs other than $\overline{\mathbf{y}}^{(p)}$ which share a *prefix* of length exactly j with $\overline{\mathbf{y}}^{(p)}$, namely

$$Y_j^{(p)} := \left\{\mathbf{y} \in \mathbb{Z}_d^n - \{\overline{\mathbf{y}}^{(p)}\} \mid \text{PRE}(\mathbf{y}_{1..n-1}, \overline{\mathbf{y}}^{(p)}_{1..n-1}) = j\right\}.$$

By Proposition II.1, $(\mathbf{x},\mathbf{y})$ node-blocks $(\overline{\mathbf{x}},\overline{\mathbf{y}}^{(p)})$ if and only if $(\mathbf{x},\mathbf{y}) \in X_i \times Y_j^{(p)}$ for some i, j such that $i + j \geq n-1$. Consequently, $(\mathbf{x},\mathbf{y})$ node-blocks $(\overline{\mathbf{x}},\overline{Y})$ if and only if $(\mathbf{x},\mathbf{y}) \in X_i \times Y_j^{(p)}$ for some $i, j \in \{0,\dots,n-1\}$, and some $1 \leq p \leq l$, for which $i + j \geq n-1$.

Let $I = \mathbb{Z}_d^n$ be the set of inputs and $O = \mathbb{Z}_d^n$ be the set of outputs. Construct a bipartite graph $G_R = (I \cup O, E)$ which is the union of all complete bipartite graphs $X_i \times Y_j^{(p)}$ for which $i + j \geq n-1$. Then, $(\mathbf{x},\mathbf{y})$ node-blocks R if and only if $(\mathbf{x},\mathbf{y})$ is an edge of G_R. The set of requests in $\mathcal{R}$ each of which node-blocks R must be a matching of G_R, in order for $\mathcal{R}$ to be a valid request set. Consequently, $b(\mathcal{R}) \leq \nu(G_R)$, where $\nu(G_R)$ denotes the size of a maximum matching in G_R. It follows that

$$m \geq \max_{R \text{ an } l\text{-cast request, } l \leq f} \nu(G_R) + 1 \tag{5}$$

is sufficient for the $\log_d(N,0,m)$ network to be f-cast strictly nonblocking.

Let $\tau(G_R)$ denote the size of a minimum vertex cover in G_R, then $\tau(G_R) = \nu(G_R)$ by the classic König-Egevarý theorem. Set $j = \lceil \frac{n-r-1}{2} \rceil$. Note that $r \leq n-2$ in this case and thus $j \geq 1$. It is not difficult to verify that the following set is a vertex cover of G_R:

$$C := \left(\bigcup_{i=j}^{n-1} X_i\right) \cup \left(\bigcup_{p=1}^{l} \bigcup_{i=n-j}^{n-1} Y_i^{(p)}\right)$$

Since $|X_i| = |Y_i^{(p)}| = d^{n-i} - d^{n-1-i}$ for all i, we have

$$\begin{aligned} |C| &\leq (d^{n-j} - d^{n-j-1} + \cdots + d^1 - d^0) + \\ &\quad\; l\left(d^j - d^{j-1} + \cdots + d^1 - d^0\right) \\ &\leq d^{n-j} - 1 + f\left(d^j - 1\right) \\ &= m_{\text{nb}}(n,f,d) - 1 \end{aligned}$$

Thus, $\nu(G_R) = \tau(G_R) \leq |C| \leq m_{\text{nb}}(n,f,d) - 1$. Recall (5) and the lemma is proved.

Secondly, consider the case when $f \geq d^{n-1} - d^{n-2} + 1$. If $l \leq d^{n-1} - d^{n-2}$, then $b(\mathcal{R})$ can only be at most $m_{\text{nb}}(n, 2^{n-1} - 2^{n-2}, d) - 1$ by the previous analysis and the fact that $m_{\text{nb}}(n,f,d)$ is a non-decreasing function in f. If $l \geq d^{n-1} - d^{n-2} + 1$, then the number of free outputs is at most $d^n - l \leq d^n - d^{n-1} + d^{n-2} - 1$. Thus, $b(\mathcal{R}) \leq d^n - d^{n-1} + d^{n-2} - 1 = m_{\text{nb}}(n, d^{n-1} - d^{n-2}, 2) - 1$ and the lemma follows. ■

Lemma III.4 (Necessity)**.** *Let $r = \lfloor \log_d f \rfloor$. Then, a necessary condition for $\log_d(N,0,m)$ to be f-cast strictly nonblocking in the node-blocking sense is $m \geq m_{\text{nb}}(n,f,d)$ when $f \leq d^{n-1} - d^{n-2}$ and $m \geq m_{\text{nb}}(n, d^{n-1} - d^{n-2}, d)$ when $f \geq d^{n-1} - d^{n-2} + 1$. The function m_{nb} is defined in* (3).

Proof: Since a necessary condition for $f = d^{n-1} - d^{n-2}$ is obviously also a necessary condition for $f > d^{n-1} - d^{n-2}$, we only need to consider the case when $f \leq d^{n-1} - d^{n-2}$.

Suppose to the contrary that $m \leq m_{\text{nb}}(n,f,d) - 1$. Our strategy is as follows. We will specify an f-cast request $R = (\overline{\mathbf{x}}, \overline{Y})$ and a valid set $\mathcal{R}$ of requests compatible with R, where $|\mathcal{R}| = m_{\text{nb}}(n,f,d) - 1$ and every request in $\mathcal{R}$ node-blocks R. This way, we can set up a network state in which all copies of $\text{BY}^{-1}(n)$ are used for routing $\mathcal{R}$. When request R arrives, we cannot find a copy of $\text{BY}^{-1}(n)$ to route it, completing the proof of the necessary condition. Equivalently, using the language developed in the proof of the previous lemma, we will construct a request R for which the bipartite graph G_R has a matching of size exactly $m_{\text{nb}}(n,f,d) - 1$. (Each edge of the matching corresponds to a request in $\mathcal{R}$.)

Note that $d^r \leq f < d^{r+1}$ and $0 \leq r \leq n-2$. For every non-negative integer $i \leq d^r - 1$, let $\mathbf{s}_i$ be the unique d-ary representation of i with exactly r symbols in $\mathbb{Z}_d$, i.e. $\mathbf{s}_i \in \mathbb{Z}_d^r$ for all $i \leq d^r - 1$. Moreover, let $u = \lceil (f - d^r)/(d-1) \rceil$, then $0 \leq u \leq d^r$.

The request $R = (\overline{\mathbf{x}}, \overline{Y})$ is defined as follows:

- $\overline{\mathbf{x}} = 0^n$ (a string of n 0s)
- $\overline{Y} = \{\overline{\mathbf{y}}^{(0)}, \dots, \overline{\mathbf{y}}^{(f-1)}\}$, where
$$\overline{\mathbf{y}}^{(p)} = \mathbf{s}_p 0 0^{n-r-1} \quad \text{for} \quad 0 \leq p \leq d^r - 1,$$

and

$$\overline{\mathbf{y}}^{(p)} = \mathbf{s}_{k_p} i_p 0^{n-r-1} \quad \text{for} \quad d^r \leq p \leq f-1,$$

where $k_p := \lfloor \frac{p - d^r}{d-1} \rfloor$ and $i_p = (p - d^r) \mod (d-1) + 1$. (The second part of the $\overline{\mathbf{y}}^{(p)}$ are defined only when $f > d^r$.)

Define the input sets X_i and output sets $Y_j^{(p)}$ as in the previous lemma. The outputs $\overline{\mathbf{y}}^{(p)}$ were specifically designed so that **all** the following sets are mutually disjoint, which the reader can straightforwardly verify: $Y_{n-1}^{(0)}, \dots, Y_{n-1}^{(f-1)}$, $Y_{n-2}^{(0)}, \dots, Y_{n-2}^{(f-1)}$, $\dots$ $Y_{r+1}^{(0)}, \dots, Y_{r+1}^{(f-1)}$, and the sets $Y_r^{(p)}$ for all p such that $u \leq p \leq d^r - 1$.

Let $j = \lceil (n-r-1)/2 \rceil$. Consider four cases as follows.

Case 1: $r \leq n-5$. Then, $n - j - 2 \geq r + 1$. Recalling the definition of G_R, we know that G_R contains all the following subgraphs, which are mutually vertex disjoint:

(a) $X_i \times \bigcup_{p=0}^{f-1} Y_{n-i-1}^{(p)}$, for $0 \leq i \leq j-1$,

(b) $X_j \times \bigcup_{p=0}^{f-1} Y_{n-j-1}^{(p)}$, and

(c) $\bigcup_{i=j+1}^{n-1} X_i \times \bigcup_{p=0}^{f-1} Y_{n-j-2}^{(p)}$.

Recall also that $|X_i| = |Y_i^{(p)}| = d^{n-i} - d^{n-i-1}$ for all i and p. In what follows, we will use extensively the trivial fact that the complete bipartite graph $K_{a,b}$ has a matching of size $\min\{a,b\}$.

For each i with $0 \le i \le j-1$ the corresponding subgraph in (a) has a matching of size

$$f(d^{i+1}-d^i) = \min\{d^{n-i}-d^{n-1-i}, f(d^{i+1}-d^i)\}.$$

The subgraph in (b) has a matching of size

$$d^{n-j}-d^{n-j-1} = \min\{d^{n-j}-d^{n-j-1}, f(d^{j+1}-d^j)\}.$$

The subgraph in (c) has a matching of size

$$d^{n-j-1}-1 = \min\{\sum_{i=j+1}^{n-1}(d^{n-i}-d^{n-1-i}), f(d^{j+2}-d^{j+1})\}.$$

In total, G_R has a matching of size

$$\begin{aligned} & \sum_{i=0}^{j-1} f(d^{i+1}-d^i) + (d^{n-j}-d^{n-j-1}) + (d^{n-j-1}-1) \\ = \; & f\left(d^j-1\right) + d^{n-j} - 1 \\ = \; & m_{\text{nb}}(n,f,d) - 1 \end{aligned}$$

as desired.

Case 2: $r = n-4$. This means $j=2$ and $n-j-1 = r+1$. We know G_R contains all the following subgraphs, which are mutually vertex disjoint:

(a) $X_i \times \bigcup_{p=0}^{f-1} Y_{n-1-i}^{(p)}$, for $0 \le i \le 1$,

(b) $\bigcup_{i=2}^{n-1} X_i \times \bigcup_{p=0}^{f-1} Y_{n-3}^{(p)}$.

Similar to the previous case, it can be seen that G_R contains a matching of size

$$\begin{aligned} & \sum_{i=0}^{1} \min\{d^{n-i}-d^{n-i-1}, f(d^{i+1}-d^i)\} \\ & + \min\left\{\sum_{i=2}^{n-1}(d^{n-i}-d^{n-i-1}), f(d^3-d^2)\right\} \\ = \; & \sum_{i=0}^{1} f(d^{i+1}-d^i) + \sum_{i=2}^{n-1}(d^{n-i}-d^{n-i-1}) \\ = \; & m_{\text{nb}}(n,f,d) - 1. \end{aligned}$$

Case 3: $r = n-3$. This means $j=1$ and $n-j-1 = r+1$. Note that the union $Y_{n-2} = \bigcup_{p=0}^{f-1} Y_{n-2}^{(p)}$ has exactly $f(d^2-d)$ vertices, and $(d^2-d)f \ge d^{n-1}-d^{n-2}$. Let Z be any subset of Y_{n-2} of size exactly $d^{n-1}-d^{n-2}$, and $W = Y_{n-2}-Z$. Then, G_R contains all the following subgraphs, which are mutually vertex disjoint:

(a) $X_0 \times \bigcup_{p=0}^{f-1} Y_{n-1}^{(p)}$,

(b) $X_1 \times Z$.

(c) $\bigcup_{i=2}^{n-1} X_i \times \left(W \cup \bigcup_{p=u}^{d^r-1} Y_{n-3}^{(p)}\right)$.

This time, G_R contains a matching of size

$$\begin{aligned} & \min\{d^n-d^{n-1}, f(d-1)\} + (d^{n-1}-d^{n-2}) \\ + \; & \min\{\sum_{i=2}^{n-1}(d^{n-i}-d^{n-i-1}), \\ & ((d^2-d)f - (d^{n-1}-d^{n-2})) + (d^r-u)(d^3-d^2)\} \\ = \; & f(d-1) + (d^{n-1}-d^{n-2}) + (d^{n-2}-1) \\ = \; & m_{\text{nb}}(n,f,d) - 1. \end{aligned}$$

Case 4: $d^{n-2} \le f \le d^{n-1}-d^{n-2}$. Similarly, G_R contains all the following subgraphs, which are mutually vertex disjoint:

(a) $X_0 \times \bigcup_{p=0}^{f-1} Y_{n-1}^{(p)}$,

(b) $\bigcup_{i=2}^{n-1} X_i \times \left(Y - (\bigcup_{p=0}^{f} Y_{n-1}^{(p)} \cup \overline{Y})\right)$.

Thus, G_R has a matching of size

$$\begin{aligned} & \min\{d^n-d^{n-1}, f(d-1)\} + \min\{d^{n-1}-1, d^n-df\} \\ & \qquad = f(d-1) + d^{n-1} - 1 = m_{\text{nb}}(n,f,d) - 1. \end{aligned}$$

■

IV. Both the input and middle stages have fanout capability

The main results of this section are summarized in the following two theorems.

Theorem IV.1. *Let $r = \lfloor \log_d f \rfloor$. Then, the necessary and sufficient condition for $\log_d(N,0,m)$ to be f-cast strictly nonblocking in the node-blocking sense is $m \ge m_{\text{nb}}(n,f,d)$, where the function $m_{\text{nb}}(n,f,d)$ was defined in* (3).

Theorem IV.2. *Let $r = \lfloor \log_d f \rfloor$. Then, the necessary and sufficient condition for $\log_d(N,0,m)$ to be f-cast strictly nonblocking in the link-blocking sense is $m \ge m_{\text{lb}}(n,f,d)$, where the function $m_{\text{lb}}(n,f,d)$ was defined in* (4)

The proofs of Theorems IV.1 and IV.2 are similar. We will present here only a proof of Theorem IV.1. Note that Theorem IV.2 is already obtained in [8]. However, the technique developed in this paper yields a fundamentally different proof of the theorem, further illustrating the strength of our technique.

Similar to the previous section, we will break the proof of Theorem IV.1 into two parts: the sufficient and the necessary conditions, which are presented in the following two lemmas.

Lemma IV.3. *Let $r = \lfloor \log_d f \rfloor$, namely $d^r \le f < d^{r+1}$. Then, a sufficient condition for $\log_d(N,0,m)$ to be f-cast strictly nonblocking in the node-blocking case is $m \ge m_{nb}(n,f,d)$.*

Proof: Suppose the network already has established routes for some set $\mathcal{R}$ of requests. Let $\overline{R} = (\overline{\mathbf{x}}, \overline{Y})$ be a new f-cast request compatible with the current network state, where $\overline{\mathbf{x}}$ is some input and $\overline{Y} = \{\overline{\mathbf{y}}^{(0)}, \dots, \overline{\mathbf{y}}^{(l)}\}$ is a subset of at most f outputs. For each $i \in [l]$, the part $(\overline{\mathbf{x}}, \overline{\mathbf{y}}^{(i)})$ can be routed independent from all other $(\overline{\mathbf{x}}, \overline{\mathbf{y}}^{(j)})$, $j \ne i, j \in [l]$, because both stages have fanout capability. Consequently, without loss of generality we can assume that $\overline{Y}$ contains only one output $\overline{\mathbf{y}}$, i.e. $\overline{R} = (\overline{\mathbf{x}}, \overline{\mathbf{y}})$.

Unlike in the proof of Lemma III.3, we cannot assume that the requests in $\mathcal{R}$ are unicast requests, because each request R in $\mathcal{R}$ can be routed through several different copies of $\text{BY}^{-1}(n)$. However, we can ignore the branches of R which do not node-block $\overline{R}$. Consequently, for every request $R = \big(\mathbf{x}, \{\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(l)}\}\big) \in \mathcal{R}$ we can assume that $(\mathbf{x}, \mathbf{y}^{(i)})$ node-blocks $\overline{R}$ for every $i \leq l$.

Similar to the proof of Lemma III.3, for each $0 \leq i \leq n-1$, define

$$X_i := \{\mathbf{x} \in \mathbb{Z}_d^n - \{\overline{\mathbf{x}}\} \mid \text{SUF}(\mathbf{x}_{1..n-1}, \overline{\mathbf{x}}_{1..n-1}) = i\}.$$

And, for each $0 \leq j \leq n-1$, let

$$Y_j := \{\mathbf{y} \in \mathbb{Z}_d^n - \{\overline{\mathbf{y}}\} \mid \text{PRE}(\mathbf{y}_{1..n-1}, \overline{\mathbf{y}}_{1..n-1}) = j\}.$$

Note that the X_i are mutually disjoint and the Y_i are mutually disjoint. Moreover, $|X_i| = |Y_i| = d^{n-i} - d^{n-1-i}, \forall i$. By Proposition II.1, for every request $R = \big(\mathbf{x}, \{\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(l)}\}\big) \in \mathcal{R}$ and for every $p \in [l]$ it must be the case that $(\mathbf{x}, \mathbf{y}^{(p)}) \in X_i \times Y_j$ for some i, j such that $i + j \geq n-1$.

Let G_R be the bipartite graph which is the union of all $X_i \times Y_j$ with $i + j \geq n-1$. Then, every request $R \in \mathcal{R}$ corresponds uniquely to a set of at most f edges incident to the same input vertex. On the output side, no output vertex appears more than once in $\mathcal{R}$. Let $b(\mathcal{R})$ be the number of edges in G_R that $\mathcal{R}$ correspond to, then $m \geq \max_{\mathcal{R}} b(\mathcal{R}) + 1$ is certainly sufficient to route R. In the worse case, each branch of each request in $\mathcal{R}$ is routed through a different copy of $\text{BY}^{-1}(n)$.

Now, let $\nu_f(G_R)$ be the maximum size of a subset S of edges of G_R satisfying the following conditions: (a) each input is incident to at most f edges in S, and (b) each output is incident to at most 1 edge in S. (Note that $\nu_1(G_R)$ is just the matching number $\nu(G)$.) Then, from the above analysis it follows that $b(\mathcal{R}) \leq \nu_f(G_R)$. Consequently, $m \geq \max_R \nu_f(G_R) + 1$ is a sufficient condition for $\log_d(N, 0, m)$ to be strictly non-blocking in the node-blocking sense.

To this end, construct a flow network D_R as follows. The network has a source s, a sink t, and the set of vertices $X \cup Y$ where $X = \bigcup_{i=0}^{n-1} X_i$ and $Y = \bigcup_{i=0}^{n-1} Y_i$. There is an edge from the source s to every vertex $\mathbf{x} \in X$ with capacity f. There is an edge $(\mathbf{x}, \mathbf{y})$ with capacity 1 from X to Y if and only if $(\mathbf{x}, \mathbf{y}) \in X_i \times Y_j$ for some $i + j \geq n-1$. And finally, there is an edge from each $\mathbf{y} \in Y$ with capacity 1 to the sink t. Hence, the "middle part" of the flow network is exactly G_R with s and t adjoined on the two sides.

For every set S of edges of G_R satisfying conditions (a) and (b) above, it is easy to see that there corresponds an integral flow with value exactly the size of S. Consequently, $|S|$ is at most the maximum flow value in D_R, denoted by $\text{MAXFLOW}(D_R)$, which is equal to the capacity of a minimum s,t-cut in D_R, denoted by $\text{MINCUT}(D_R)$. Thus, $\nu_f(G_R) \leq \text{MINCUT}(D_R)$. In particular $\nu_f(G_R)$ is at most the capacity of any s,t-cut in D_R.

Let $j = \lceil (n-r-1)/2 \rceil - 1$. Consider the following s,t-cut in D_R as illustrated in Figure 3:

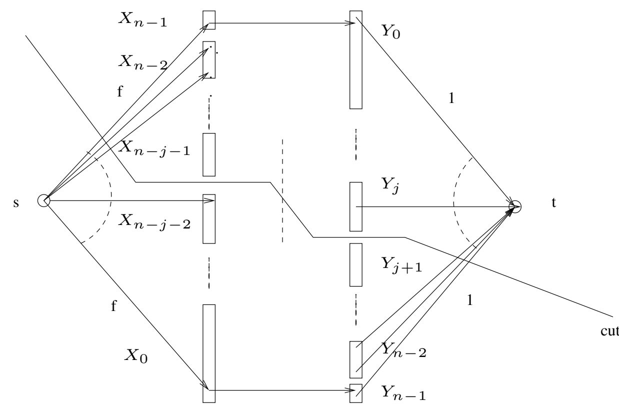

Fig. 3. The s,t-cut in the proof

$$\left(\{s\} \cup \bigcup_{i=0}^{n-j-2} X_i \cup \bigcup_{i=j+1}^{n-1} Y_i, \bigcup_{i=n-j-1}^{n-1} X_i \cup \bigcup_{i=0}^{j} Y_i \cup \{t\} \right).$$

The capacity of this cut is exactly

$$\sum_{i=n-j-1}^{n-1} f|X_i| + \sum_{i=j+1}^{n-1} |Y_i| = m_{nb}(n, f, d) - 1.$$

Consequently, $m \geq m_{nb}(n, f, d)$ is sufficient. ■

Lemma IV.4. *Let $r = \lfloor \log_d f \rfloor$, namely $d^r \leq f < d^{r+1}$. Then, a necessary condition for $\log_d(N, 0, m)$ to be f-cast strictly nonblocking in the node-blocking sense is $m \geq m_{nb}(n, f, d)$.*

Proof: Following the terminologies of the proof of the preceding lemma and the line of reasoning in the proof of Lemma III.4, it suffices to show that for every request $R = (\mathbf{x}, \mathbf{y})$, $\text{MAXFLOW}(D_R) = m_{nb}(n, f, d) - 1$.

Since we have already specified a cut with capacity $m_{nb}(n, f, d) - 1$, it is sufficient to specify a flow with value $m_{nb}(n, f, d) - 1$, which will have to be maximum due to weak duality.

Note that $j \leq n-1$, always. Let us first consider the (more general) case when $j = \lceil (n-r-1)/2 \rceil - 1 \leq n-3$, i.e. $n-j-3 \geq 0$. Consider the following flow assignment:

- For every $i = n-j-1, \dots, n-1$, and every vertex $\mathbf{x} \in X_i$, set the flow on edge $(s, \mathbf{x})$ to be f; moreover, for each such $\mathbf{x}$ choose a unique subset $Y(\mathbf{x})$ of exactly f vertices in Y_{n-1-i}. This is possible since $f|X_i| \leq |Y_{n-1-i}|$ for this range of i. Now, for every $\mathbf{y} \in Y(\mathbf{x})$, set the flow value of $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{y}, t)$ to be 1. So far, the flow-conservation constraint is satisfied at every vertex and the total flow value that t receives is exactly

$$\sum_{i=n-j-1}^{n-1} f|X_i| = \sum_{i=0}^{j} f(d^{i+1} - d^i) = f(d^{\lceil (n-r-1)/2 \rceil} - 1)$$

- Similarly, since $f|X_{n-j-2}| = f(d^{j+2} - d^{j+1}) \geq d^{n-j-1} - d^{n-j-2} = |Y_{j+1}|$, we can send flow through

X_{n-j-2} and then to Y_{j+1} and then to t so that each vertex in Y_{j+1} has a flow value of exactly 1 transits through.

- Lastly, because

$$f|X_{n-j-3}| = f(d^{j+3} - d^{j+2}) > d^{n-j-2} - 1 = \sum_{i=j+2}^{n-1} |Y_i|,$$

we can similarly send flow through X_{n-j-3} and then to $\bigcup_{i=j+2}^{n-1} Y_i$ and then to t so that each vertex in $\bigcup_{i=j+2}^{n-1} Y_i$ has a flow value of exactly 1 transits through.

Consequently, the total flow value is

$$f(d^{\lceil (n-r-1)/2 \rceil} - 1) + \sum_{i=j+1}^{n-1} |Y_i| = m_{nb}(n, f, d) - 1$$

as desired.

The cases when $j = n-2$ or $j = n-1$ are done similarly. In fact, these cases only happen for small values of n (≤ 4), and are not very enlightening. ∎

V. Only the first stage has fanout capability

The proofs of the following theorems are very much similar to those in the last section, and thus omitted due to space limitation. Define

$$m^1_{\text{nb}}(n, f, d) = f d^{\lceil \frac{n-r-1}{2} \rceil} + d^{n - \lceil \frac{n-r-1}{2} \rceil}. \quad (6)$$

$$m^1_{\text{lb}}(n, f, d) = f d^{\lceil \frac{n-r-2}{2} \rceil} + d^{n - \lceil \frac{n-r}{2} \rceil}. \quad (7)$$

Theorem V.1. *Let* $r = \lfloor \log_d f \rfloor$. *Then, the necessary and sufficient condition for* $\log_d(N, 0, m)$ *to be* f*-cast strictly nonblocking in the node-blocking sense is*

$$m \geq \begin{cases} d^n & \text{when } f \geq \frac{d^n - d^{n-1} + 1}{d}, \\ df + d^{n-1} & \text{when } \frac{d^n - d^{n-1} + 1}{d} > f \geq d^{n-3}, \\ m^1_{\text{nb}}(n, f, d) & \text{when } d^{r+1} > f \geq d^r, r \leq n-4. \end{cases}$$

Theorem V.2. *Let* $r = \lfloor \log_d f \rfloor$. *Then, the necessary and sufficient condition for* $\log_d(N, 0, m)$ *to be* f*-cast strictly nonblocking in the link-blocking sense is*

$$m \geq \begin{cases} d^{n-1} - 1 & \text{when } f \geq \frac{d^n - d^{n-1} + 1}{d} \\ df + d^{n-2} - 1 & \text{when } \frac{d^n - d^{n-1} + 1}{d} > f \geq d^{n-3} \\ m^1_{\text{lb}}(n, f, d) & \text{when } d^{r+1} > f \geq d^r, r \leq n-4 \end{cases}$$

VI. Discussions

The König-Egevarý theorem used in the proofs in Section III is just a special case of the max-flow min-cut theorem, which is used in the proofs in Sections IV and V. This usage of network flows in analyzing f-cast SNB networks is the most novel contribution of this paper. We hope that the technique will be applicable in analyzing other types of SNB and RNB architectures, including the more general $\log(N, p, m)$ networks which will lead to the d-ary Benes network and likely the Clos-type networks.

Aside from that point, there are two other major observations we can make from our results. Firstly, under the same node- or link-blocking sense, the fanout constraint does not affect the complexity of SNB f-cast multi-$\log_d N$ networks for most values of f, except for very large values of f ($\geq d^{n-4}$ or so). This fact can be interpreted as a good news or a bad news. The good news is the first-stage having fanout capability is almost always as powerful as the case when all stages have fanout capability. The bad news is adding fanout capability to all stages does not help reduce the switch complexity. Secondly, the node- and link-blocking cases do not differ too much in terms of asymptotic complexity of the switch. They are roughly within a factor d of each other, with the link-blocking case requiring more hardwares. For instance, when $f = N^\epsilon$ we have both m_{nb} and m_{lb} approximately $N^{1+\epsilon/2}$.

References

[1] W. Kabacinski and G. Danilewicz, "Wide-sense and strict-sense non-blocking operation of multicast multi-$log_2 n$ switching networks," *IEEE Transactions on Communications*, vol. 50, pp. 1025–1036, Jun 2002.

[2] F.K.Hwang and B.-C. Lin, "Wide-sense nonblocking multicast $\log_2(n, m, p)$ networks," *IEEE Transactions on Communications*, vol. 51, pp. 1730–1735, Oct 2003.

[3] Y. Yang and G. M. Masson, "Nonblocking broadcast switching networks," *IEEE Trans. Comput.*, vol. 40, no. 9, pp. 1005–1015, 1981.

[4] F. K. Hwang, *The mathematical theory of nonblocking switching networks*. River Edge, NJ: World Scientific Publishing Co. Inc., 2004.

[5] H.-B. Chen and F. K. Hwang, "On multicast rearrangeable 3-stage clos networks without first-stage fan-out," *SIAM Journal on Discrete Mathematics*, vol. 20, no. 2, pp. 287–290, 2006.

[6] A. Pattavina and G. Tesei, "Non-blocking conditions of multicast three-stage interconnection networks," *IEEE Transactions on Communications*, vol. 46, pp. 163–170, Dec. 2005.

[7] X. Jiang, A. Pattavina, and S.Horiguchi, "Rearrangeable f-cast multi-$\log 2n$ networks," *IEEE Transactions on Communications*, 2007. to appear.

[8] F.K.Hwang, Y. Wang, and J. Tan, "Strictly nonblocking f-cast $\log_d(n, m, p)$ networks," *IEEE Transactions on Communications*, vol. 55, pp. 981–986, May 2007.

[9] R. R. Goke and G. J. Lipovski, "Banyan networks for partitioning multiprocessor systems," in *Proceedings of the 1st Annual Symposium on Computer Architecture (ISCA'73)*, pp. 21–28, Dec 1973.

[10] C. Clos, "A study of non-blocking switching networks," *Bell System Tech. J.*, vol. 32, pp. 406–424, Mar 1953.

[11] C.-T. Lea, "Muti-$\log_2 n$ networks and their applications in high speed electronic and photonic switching systems," *IEEE Transactions on Communications*, vol. 38, no. 10, pp. 1740–1749, 1990.

[12] M. M. Vaez and C.-T. Lea, "Strictly nonblocking directional-coupler-based switching networks under crosstalk constraint," *IEEE Transactions on Communications*, vol. 48, pp. 316–323, Feb 2000.

[13] F. Hwang, "Choosing the best $\log_2(n, m, p)$ strictly nonblocking networks," *IEEE Transactions on Communications*, vol. 46, pp. 454–455, Dec 1998.

[14] D.-J.Shyy and C.-T. Lea, "$\log_2(n, m, p)$ strictly nonblocking networks," *IEEE Transactions on Communications*, vol. 39, no. 10, pp. 1502–1510, 1991.

[15] G. Maier and A. Pattavina, "Design of photonic rearrangeable networks with zero first-order switching-element-crosstalk," *IEEE Transactions on Communications*, vol. 49, no. 7, pp. 1268–1279, 2001.

[16] C.-T. Lea and D.-J.Shyy, "Tradeoff of horizontal decomposition versus vertical stacking in rearrangeable nonblocking networks," *IEEE Transactions on Communications*, vol. 39, pp. 899–904, 1991.

[17] X.Jiang, H. Shen, M. M.R.Khandker, and S.Horiguchi, "Blocking behaviors of crosstalk-free optical banyan networks on vertical stacking," *IEEE/ACM Trans. Networking*, vol. 11, no. 6, pp. 982–993, 2003.

[18] V. R. C. et al, "Crosstalk in a lossy directional coupler switch," *J. Lightwave Technol.*, vol. 13, no. 7, pp. 1530–1535, 1995.

[19] D.Li, "Elimination of crosstalk in directional coupler switches," *Optical Quantum Electron.*,, vol. 25, no. 4, pp. 255–260, 1993.