Structure and non-blocking properties of bidirectional unfolded two-stage switches

Structure and non-blocking properties of bidirectional unfolded two-stage switches

Labson Koloko,✉ 🆔 Hitoshi Obara, and Seiji Kumagai

Graduate School of Engineering Science, Akita University, Akita, Japan

✉Email: labinct@gmail.com

Two-stage switch networks are an emerging design option for relatively small-capacity space switches. They are classified into two categories: folded and unfolded. Although folded switches have been well studied, research on unfolded two-stage switch networks (UTSNs) remains limited. Here, non-blocking UTSNs are considered. First, a new UTSN design is presented that consists of input and output switch modules (ISMs and OSMs) using bidirectional switching techniques. The proposed UTSN is represented by $B(n, m, r)$, where n, m, and r denote the number of input ports of the ISM, number of OSMs, and number of ISMs, respectively. Second, the maximum number of rearrangements for $B(n, n, r)$ is proved to be $\lfloor (r-1)/2(n-1) \rfloor$ in general, whereas it is limited to two when $n \geq r$. The strictly non-blocking condition for $B(n, m, r)$ to be $m \geq n+1$ is also determined. Finally, it is shown that the switch hardware complexity becomes minimal at $n = \sqrt{N/2}$ and saturates at $N^2/2$ as $N \rightarrow \infty$.

Introduction: As serious attention has been recently given to space-division multiplexing technology for scaling optical network capacity [1], multistage space switch networks have become a key component for creating high-port-count optical interconnects and cross-connects [2, 3]. Although the three-stage Clos architecture is a well-established and highly practical design principle for scalable space switches [4], two-stage networks (TSNs) are emerging as a new design option for relatively small-capacity switches [5]. TSNs are classified into folded and unfolded switches. The folded TSN is equivalent to a three-stage Clos network, and its structure and non-blocking capabilities are well known [6]. In contrast, unfolded TSNs (UTSNs) are not yet completely understood; there are few types of UTSNs, all of which are rearrangeably non-blocking (RNB) [7]. To the best of our knowledge, this is the first study to consider a strictly non-blocking (SNB) UTSN. First, we present a new design principle of UTSN, which consists of input and output switch modules (ISMs and OSMs) with a bidirectional switching capability. The proposed UTSN is represented by $B(n, m, r)$, where n, m, and r denote the number of input ports to the ISM, the number of OSMs, and the number of ISMs, respectively. Second, we formulate the maximum number of rearrangements for $B(n, n, r)$ and the minimum value of m to satisfy the SNB condition for $B(n, m, r)$. Finally, we briefly estimate the complexity of the switch hardware.

Structure of $B(n, m, r)$: A design example of $B(n, m, r)$ is shown in Figure 1, where there are r ISMs and m OSMs; each is denoted by I_p, $1 \leq p \leq r$, and O_q, $1 \leq q \leq m$. Every ISM has n inputs, and the total number of inputs is given by $N = nr$, while every OSM has N outputs, of which the first and second halves are provided at the top and bottom edges. Every pair of an ISM and an OSM is interconnected with a pair of internal links, that is, one link between an outlet on the top of the ISM and an inlet on the left side of the OSM and the other link between the bottom of the ISM and the right side of the OSM. Here, we use 'inlet' and 'input' (and similarly 'outlet' and 'output') quite differently: 'inlet' and 'outlet' are internal ports, whereas 'input' and 'output' are external ports.

Although the ISM is an $n \times 2m$ switch, it can be implemented by an $n \times m$ bidirectional crossbar switch (BXS) and n 1×2 switches, as shown in Figure 2. Every input signal to the ISM may be switched to its outlet at either end of the column. Let (i, k) be a connection between an input i, $1 \leq i \leq n$, and an outlet k, $1 \leq k \leq 2m$, in the ISM. Similarly, let (l, j) be a connection between an inlet l, $1 \leq l \leq 2r$, and an outlet j, $1 \leq j \leq N$, in the OSM, which can be implemented by an $r \times N/2$ BXS, as shown in Figure 3. Every jth outlet in the OSM is coupled in the jth output via a passive coupler, which is shown as a dashed triangle in Figure 1. The
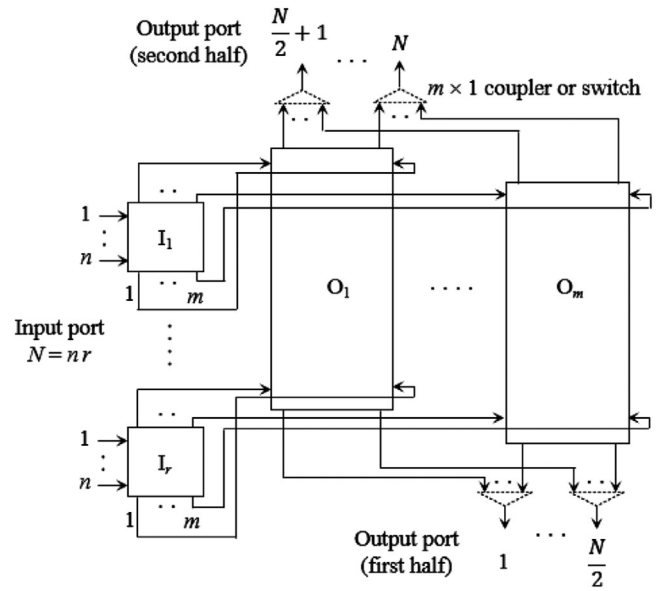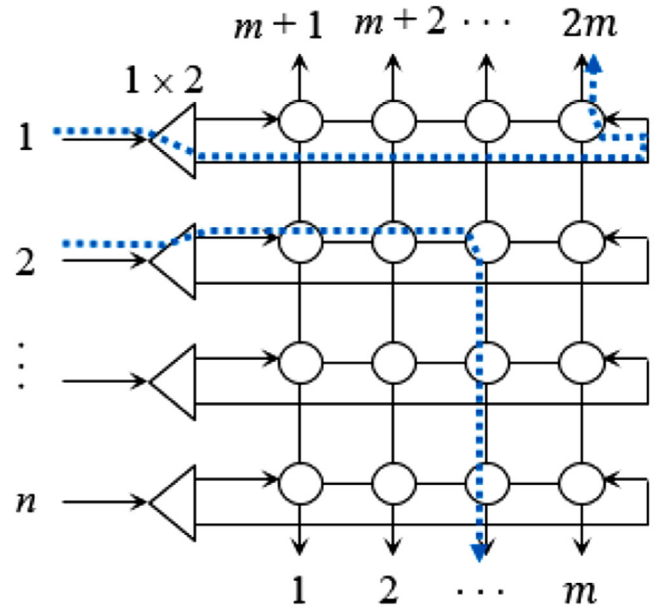

Fig. 1 *Structure of $B(n, m, r)$*

Fig. 2 *Design example of ISM using bidirectional switching*

passive coupler can be substituted by an $m \times 1$ switch at the cost of extra cross-points. It should be pointed out that only one outlet out of m outlets of the same number of OSMs may have an output signal, and the other outlets remain idle.

Because every input signal in Figure 2 is routed to either a left or right inlet through a 1×2 switch, every row of ISMs takes only a single signal at most. It can be seen that every column of ISMs and OSMs is shared by a couple of signals, that is, one signal headed for the top and the other signal down for the bottom. However, we assume that every column of the ISMs and OSMs can take a single signal at most to avoid blocking. As a result, every row of OSMs also takes a single signal, because the column of an ISM is combined with a row of an OSM. Note that these constraints will be referred to in the following discussion of the non-blocking properties. If the ISM is implemented with a conventional $n \times m$ crossbar switch: m outlets at the top of the ISM are not necessary and the 1×2 switches should be relocated to the left inlets of the OSM. In both cases, the 1×2 switches cause no exchange between any pair of connections; thus, their column cannot be counted as an independent switching stage [7].
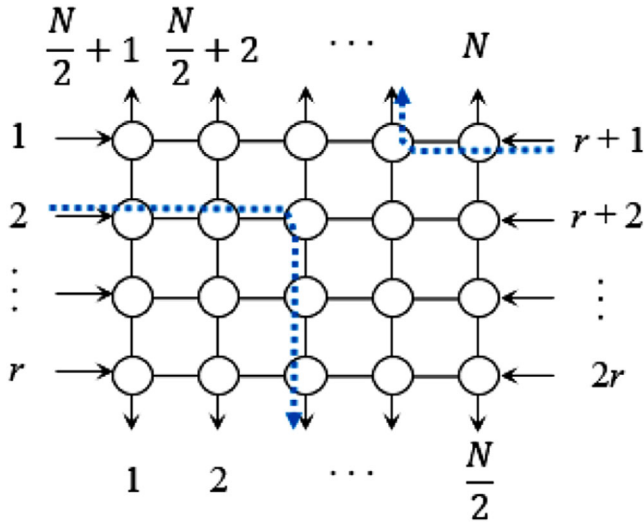
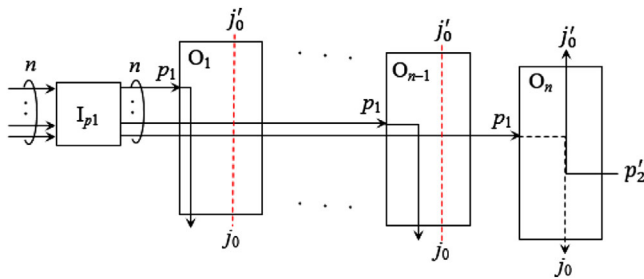**Fig. 3** *Design example of OSM using bidirectional switching*

Fig. 4 *Worst-case scenario of rearrangement in B(n, n, r)*

Non-blocking properties of B(n, m, r): Although our main objective is the SNB condition for B(n, m, r), let us begin with the rearrangement process for B(n, n, r), which provides some insights into the SNB condition. It is evident that no blocking occurs in the ISMs because the ISMs function as an incomplete $n \times 2m$ switch under the constraints. Blocking can occur in the OSM in the following worst-case scenario: Assume that $n - 1$ inputs of an ISM I$_{p1}$, $1 \leq p_1 \leq r$, are already connected with $n - 1$ outputs. Without loss of generality, we assume that the $n - 1$ input signals enter the OSMs from their left inlets. Then, the last connection request in I$_{p1}$, which is denoted by the dashed line in Figure 4, is issued. Further, assume that the request has a destination output j_0, $1 \leq j_0 \leq N/2$, while $j_0' = j_0 + N/2$ corresponds to I$_{p2}$, $1 \leq p_2 \leq r$, as shown in Figure 4, where we assume that $p_2' = p_2 + r$ and all of the n inputs to I$_{p2}$ are already in use. Note that the prime mark denotes the inlets on the right side and the outlets on the top edge. Blocking occurs in the j_0th column in O$_n$ owing to the violation of the constraints. In this case, it can be seen that every j_0th column in the OSMs, except O$_n$, is idle. If the new connection (p_1, j_0) is rerouted to the j_0th column in an OSM O$_b$, $1 \leq b \leq n - 1$, which is connected to I$_{p1}$, an existing connection denoted by (p_1, j_1) in O$_b$ should be moved to O$_n$, only where inlet p_1 is idle. However, this rearrangement can cause blocking if the j_1th column in O$_n$ is already used. Note that $j_1 \neq j_0$ holds because the blocking in the j_0th column has already been addressed. In other words, each existing connection in O$_n$ experiences a rearrangement only once. Because there are $r - 1$ existing connections in O$_n$, the rearrangement process will last $r - 1$ times at most between O$_n$ and O$_b$. Let R_1 be the number of rearrangements in this case.

Now, consider another rearrangement process that begins with p_2' in O$_n$. If the new connection (p_1, j_0) is provided in O$_n$, the existing connection (p_2', j_0') should be rerouted to the j_0th column in O$_b$. However, blocking will occur in the OSM because there is another existing connection (p_2, j_2) or (p_2', j_2'), with $j_2 \neq j_0$ and $j_2' \neq j_0'$, which should be moved to O$_n$, only where inlets p_2 and p_2' are idle. The second rearrangement process also lasted $r - 1$ times at most. Let R_2 be the number

of rearrangements in this case. Because the connections involved in R_1 and R_2 are different from each other, the following relation holds:

$$R_1 + R_2 \leq r - 1. \qquad (1)$$

Both R_1 and R_2 are integers, and the minimum number of rearrangements is expressed as

$$\min(R_1, R_2) \leq \lfloor (r - 1)/2 \rfloor, \qquad (2)$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. Note that O$_b$ was fixed in the above discussion. By examining the number of rearrangements for every O$_b$, the minimum number of rearrangements at large, denoted as R_0, can be derived as follows in a similar manner to the derivation of Equation (2):

$$R_0 \leq \left\lfloor \frac{r - 1}{2(n - 1)} \right\rfloor. \qquad (3)$$

When $n \geq r$, we have more freedom to exchange the blocked connection (p_1, j_0) with other existing connections. As shown in Figure 4, I$_{p1}$ and I$_{p2}$ have $n - 1$ and n connections under the worst-case scenario. These assumptions allow us to choose an outlet, which is not included in the $r - 1$ existing connections in O$_n$, out of the n existing connections in I$_{p2}$. Accordingly, when we move (p_2', j_0') to an appropriate OSM, of which the j_3th column is occupied by an existing connection, that is, (p_2, j_3) or (p_2', j_3'), whereas the j_3th column in O$_n$ is idle. As a result, the number of rearrangements is reduced to two at most.

Based on the above discussions, we readily have an SNB condition as follows: When $m = n$, we need to move the blocked connection to another OSM, where blocking can occur. However, if we set $m = n + 1$, the last connection request may be provided in the $(n + 1)$-th OSM, where both p_1 and j_0 are idle. Consequently, the SNB condition for B(n, m, r) is given by

$$m \geq n + 1. \qquad (4)$$

Hardware complexity of B(n, m, r): When $m = n + 1$, the total number of cross-points becomes large and is expressed as a function of n as follows:

$$C_p(n) = \frac{N^2}{2} \left(1 + \frac{1}{n} \right) + N(n + 2), \qquad (5)$$

where the first and second terms in Equation (5) correspond to the cross-points of the OSMs and ISMs, respectively. $C_p(n)$ has a minimum value at $n_{opt} = \sqrt{N/2}$, as follows:

$$C_p(n_{opt}) = \frac{N^2}{2} + N\left(\sqrt{2N} + 2 \right). \qquad (6)$$

$C_p(n_{opt})$ converges to $N^2/2$ as $N \rightarrow \infty$. It is worth noting that some RNB UTSNs also have $N^2/2$ cross-points [7]; hence, the bidirectional UTSN has achieved SNB properties with approximately the same cross-points as RNB UTSNs.

Conclusion: We unveiled a new design principle for UTSNs using bidirectional switches. The bidirectional UTSN is represented by B(n, m, r), which consists of r ISMs, each of which has n input ports, and m OSMs. We proved that the maximum number of rearrangements for B(n, n, r) is given by $\lfloor (r - 1)/2(n - 1) \rfloor$ in general, whereas it is limited to two when $n \geq r$. We identified the SNB condition for B(n, m, r) to be $m \geq n + 1$. We also found that the UTSN hardware complexity becomes minimal at $n = \sqrt{N/2}$ and saturates at $N^2/2$. Although experimental performance analyses need to be conducted in future studies, the bidirectional UTSN could be a potential candidate for scalable space switches.

Conflict of Interest: There are no conflicts of interest to be declared.

References

1 Jinno, M.: Spatial channel network (SCN): Opportunities and challenges of introducing spatial bypass toward the massive SDM era [Invited]. *J. Opt. Commun Networking*. **11**(3), 1–14 (2019). https://doi.org/10.1364/JOCN.11.000001

2 Fiorani, M., et al.: Spatial division multiplexing for high capacity optical interconnects in modular data centers. *J. Opt. Commun. Networking*. **9**(2), A143–A153 (2017). https://doi.org/10.1364/JOCN.9.00A143

3 Jinno, M.: Architectures of spatial add/drop multiplexer and cross-connect for spatial channel networks. In: 2020 International Conference on Optical Network Design and Modelling, pp. 1–3. IEEE, Barcelona (2020). https://doi.org/10.23919/ONDM48393.2020.9132997

4 White, I. H., et al.: High port count switch architectures for data centre. Paper presented at the *Photonic Networks and Devices*, New Orleans, USA, July 2017. https://doi.org/10.1364/NETWORKS.2017.NeW1B.2

5 Lea, C.-T.: Expanding the switching capabilities of optical crossconnects. *IEEE Trans. Commun.* **53**(11), 1940–1944 (2005). https://doi.org/10.1109/TCOMM.2005.858662

6 Kabacinski, W.: Chapter 4. In: *Nonblocking Electronic and Photonic Switching Fabrics*, pp. 81–115. Springer (2005)

7 Obara, H.: Cascaded versus parallel architectures of two-stage optical crossbar switches with an extra set of inputs and outputs. *IET Optoelectron*. **12**(4), 196–201 (2018). https://doi.org/10.1049/iet-opt.2017.0096