DOI: http://doi.org/10.25728/cs.2021.1.7

TWO-STAGE DUAL PHOTON SWITCHES IN AN EXTENDED SCHEME BASIS

E.A. Barabanova, K.A. Vytovtov, and V.S. Podlazov

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

⊠ elizavetaalexb@yandex.ru, ⊠ vytovtov_konstan@mail.ru, ⊠ podlazov@ipu.ru

Abstract. This paper proposes a new method for constructing a two-stage dual photon switch with enhanced functional characteristics in the system basis of low-channel photon switches and photon multiplexers and demultiplexers. The method yields non-blocking switches with static self-routing. Also, the method ensures a significant speed-up for the switches with the same number of channels and a significant increase in the number of channels with the same speed compared to the non-blocking dual switches known previously. The non-blocking self-routing dual photon switches presented below have the highest possible speed and a maximum possible number of channels with almost the same complexity. As is shown in the paper, the dual switches have a switching complexity comparable with full switches and, at the same time, a lower channel complexity.

Keywords: physical level, photon switch, dual switch, multistage switch, conflict-free self-routing, nonblocking switch, static self-routing, quasi-complete digraph, switching properties, direct channels, scalability and speed.

INTRODUCTION

In the papers [1–3], a technique was proposed for constructing non-blocking photon switches with static self-routing for optical supercomputer systems. A system area network is non-blocking if for any packet permutation, conflict-free paths from sources to sinks can be built in it. A system area network is selfrouting if conflict-free paths can be built locally over network nodes without their interaction based on routing information in packets only. Finally, self-routing is static if any source can independently choose conflict-free paths to its sink without interacting with other sources.

Until recently, the problem of constructing nonblocking system area networks has not been completely solved. At best, rearrangeable multistage Clos networks were proposed [4, 5]. In rearrangeable networks [4], a conflict-free implementation of any permutation of data packets is possible, but a conflictfree schedule for each permutation has to be compiled separately, and it is not self-routing.

Non-blocking Clos networks [6] are known. However, for such networks, no static or even dynamic routing procedures have been proposed so far. In addition, they have significantly greater complexity than rearrangeable Clos networks and are not applied in practice.

The *p*-ary *r*-dimensional generalized hypercube is a widely used system area network [7, 8]. However, for r > 2, it is neither non-blocking nor even rearrangeable. For making the generalized hypercube rearrangeable, the number of channels in it should be increased. For example, for p = 2, it suffices to double the number of channels of one dimension [9]. In this case, a double hypercube in which all channels are duplicated has conflict-free schedules for two permutations at once [10].

The three-dimensional hypercube with dynamic self-routing has been made non-blocking recently; see [11]. However, this required an almost threefold increase in the number of channels and the degree of constituent switches.

Nevertheless, the problem of constructing a nonblocking self-routing network has a solution in the special case of networks with the topology of a quasicomplete graph and digraph [12]. Unfortunately, the number N of users (processors) in such networks does

S

not exceed the square of the degree p of composite switches: $N = p(p-1)/\sigma + 1$ and $N = p^2$, respectively, where σ is the number of different channels between users. Moreover, they have a greater switching complexity and a smaller channel complexity compared to a complete graph. A quasi-complete graph is isomorphic to such a mathematical object as an incomplete balanced block design [13–17]; a quasi-complete digraph is isomorphic to a two-dimensional generalized hypercube (parallelogram) or a two-dimensional multiring. In particular, the two lowest dimensions of the four-dimensional hypercube Dragonfly (CRAY XC-30) [8] are implemented in the form of a 6×16 parallelogram and represent a non-blocking self-routing subnet.

Networks with the topology of quasi-complete graphs and digraphs can be extended by increasing the number of their users without changing the nonblocking and self-routing properties. This effect is achieved by the invariant extension method of system area networks [12, 18]. Unfortunately, such an extension further increases the switching complexity of extended networks, strongly restricting their scalability.

Note that most modern system area networks (Clos networks [5], generalized hypercubes [8], multidimensional tori [19], the hierarchy of complete graphs from IBM [20], and the Melanox thick tree [21, 22]) cannot implement an arbitrary permutation of packages in one session without conflict. It is often necessary to resend the packets blocked in buffers. At present, photon computers are being developed [23], whose photon networks contain no buffer memory for blocked packets in the channels.

In the papers [1–3], the concept of dual networks was introduced, and new dual networks were constructed for photon computers, implementing arbitrary permutations of packets without conflict at a slightly lower channel speed. The methodology for constructing these networks is based on four fundamental principles:

• Applying a four-channel switch of a new structure, which is *dual by the conflict resolution approach*. It combines the bus method (separation of conflicting signals to different cycles in one channel) and the switch method (separation of conflicting signals to different channels).

• Assuming the parallel transmission of signaling and control information for switches at different frequencies for each data bit. This assumption eliminates the problem of synchronizing signals from different channels. • Cascading switches so that the *I*th channel of the *J*th switch on one stage is connected to the *J*th channel of the *I*th switch on the next stage. With such exchange links, the previous and next stages must include the same number of switches, each having the same number of channels. This method allows constructing multichannel switches with a small number of stages.

• Balancing the speed and complexity of a multistage switch using the invariant extension method of system area networks [4], which preserves the nonblocking property and speed of the switch with an increase in the number of its channels. This method involves an extended circuit base consisting of both $p \times p$ switches for p channels and pairs of $1 \times p$ multiplexers and $p \times 1$ demultiplexers, where $p \ge 2$.

In the papers [1-3], one of the block designs of a dual 4×4 switch is a two-stage circuit of four demultiplexers and four multiplexers with feedback links through the delay lines (Fig. 1). The switch stages are interconnected by exchange links.



Fig. 1. Generalized block design of dual switch SS4: D4 – four-input demultiplexer, M4 – four-output multiplexer, DL δ – delay line of length δ signals.

If the complexity of multiplexers M4 and demultiplexers D4 is measured by the number of switching points (equal to 4), then the switching complexity of the switch is $S_1 = 32$.

The combination of two control frequencies uniquely determines the demultiplexer's mode in which the information signal can be directed to one of the four outputs. Possible combinations of control signals are presented in Table 1 below.



\$

Table 1

Control frequencies for photon 4×4 switch

Output no.	Control frequencies
1	$\lambda_1\lambda_1$
2	$\lambda_1\lambda_2$
3	$\lambda_2\lambda_1$
4	$\lambda_2\lambda_2$

The signals from the demultiplexer's outputs are supplied to the multiplexer's inputs. One of them is passed to the output, and the rest return to their delay lines DL δ . The switch implements dynamic signal delay using the feedback links through DL δ .

The dual switch SS4 provides non-blocking with static self-routing under an appropriate length δ of the delay line. The value δ depends on the number of the stage in which the switch SS4 is used.

For the first stage, $\delta = 1$. Let four signals of duration T_0 be simultaneously supplied to the inputs of the switch SS4, all received in one cycle. With dynamic signal delay at the switch outputs, there are four possible options of signal placement: one at each output, two signals in a row at two outputs, one and three signals in a row at two outputs, and four signals in a row at one output. They are shown in Fig. 2. As a result, the switch SS4 will be non-blocking for any input traffic when the period \underline{T}_1 of information signals is four cycles.





Therefore, the non-blocking self-routing switch SS4 has the following performance characteristics: the signal period $\underline{T}_1 = 4 = \underline{N}_1$ cycles, the number of channels $\underline{N}_1 = 4$, and the switching complexity $\underline{S}_1 = 32 = \underline{N}_1^{5/2}$.

In the papers [1-3], the two-stage 16×16 switch S_216 with exchange links was considered, consisting of four switches SS4 on each stage. The first stage includes DL1, and the second stage includes DL0 (no delay lines). For an arbitrary permutation of packets, S_216 turned out to be a non-blocking self-routing

switch with the following performance characteristics: the number of channels $\underline{N}_2 = 16$, the signal period $\underline{T}_2 = 4 = \underline{N}_2^{1/2}$ cycles, and the switching complexity $\underline{S}_2 = 2 \cdot 4 \cdot 32 = 256 = \underline{N}_2^2$.

In the papers [1–3], the four-stage 256×256 switch S₄256 with exchange links was considered, consisting of 16 switches S₂16 on each stage. It consists of four stages of switches SS4. The first stage includes DL1, the second stage includes DL4, the third stage includes DL15, and the fourth stage includes DL0 (no delay lines). For an arbitrary permutation of packets, S₄256 turned out to be a non-blocking self-routing switch with the following performance characteristics: the signal period $\underline{T}_4 = 49 \approx 3\underline{N}_4^{1/2}$ cycles, the number of channels $\underline{N}_4 = 256$, and the switching complexity $\underline{S}_4 = 2 \cdot 16 \cdot 256 = 8192 = \underline{N}_4^{1.625}$. Note the large digit period (low speed) and small complexity of this switch.

In the papers [1–3], the "speed–complexity" ratio was balanced using an invariant extension of smallperiod switches. In particular, the switch S₂16 was extended through external multiplexers M4 and demultiplexers D4. As a result, the non-blocking selfrouting switch S₃64 was constructed, consisting of 16 switches S₂16 and 64 demultiplexers D4 and multiplexers M4. This switch has the following performance characteristics: the number of channels $N_3 = 64$, the signal period $\underline{T}_3 = 4 = N_3^{1/3}$ cycles, and the switching complexity $\underline{S}_3 = 16.256+4.128 =$ $=4608 = \underline{N}_3^{2.028}$.

In this paper, a non-blocking switch with the quasi-complete digraph topology will be used not only to extend two-stage switches but immediately to construct them. The resulting switches have a higher speed (a smaller signal period) and more channels than those proposed in [1-3].

In Section 1, we present the structure and characteristics of a switch with the quasi-complete graph topology for any number p of ports. In Section 2, we discuss the main idea of increasing the speed and the number of channels for p = 2 on an example of constructing a non-blocking three-stage switch. In Section 3, this idea is fully realized by constructing a non-blocking two-stage switch for p = 2. In Section 4, we construct a similar two-stage switch for any p. In Section 5, we describe an extension of a two-stage switch into switches with more channels and a constant signal period.

1. NON-BLOCKING SELF-ROUTING SWITCH WITH QUASI-COMPLETE DIGRAPH TOPOLOGY

Consider $N_1 = p^2$ dual $p \times p$ switches (SSp). For p = 4, the block design of each of them is shown in Fig. 1. In the general case, they represent a two-stage



block design with exchange links: the first stage includes p demultiplexers of the type $1 \times p$ (Dp), and the second stage includes p multiplexers of the type $p \times 1$ (Mp) with feedback links through the delay lines DL δ for each input. Each switch SSp has the switching complexity $S_1 = 2p^2$.

From $N_1 = p^2$ switches SSp, N_1 multiplexers Mp without delay lines, and N_1 demultiplexers Dp, it is possible to construct a non-blocking self-routing switch $N_1 \times N_1$ with the quasi-complete digraph topology—the switch SFN₁. The interconnections in this switch are specified by an incidence table; for p = 4, see the example in Table 2. The circuit of the switch SF16 is provided in Fig. 3.

Table 2

Interconnections in 16×16 switch SF16 with quasicomplete digraph topology

Simplex channels from users				4×4 switches SS4	Simplex channels to users				
1	2	3	4	1	1	5	9	13	
2	3	4	1	2	2	6	10	14	
3	4	1	2	3	3	7	11	15	
4	1	2	3	4	4	8	12	16	
5	6	7	8	5	5	9	13	1	
6	7	8	5	6	6	10	14	2	
7	8	5	6	7	7	11	15	3	
8	5	6	7	8	8	12	16	4	
9	10	11	12	9	9	13	1	5	
10	11	12	9	10	10	14	2	6	
11	12	9	10	11	11	15	3	7	
12	9	10	11	12	12	16	4	8	
13	14	15	16	13	13	1	5	9	
14	15	16	13	14	14	2	6	10	
15	16	13	14	15	15	3	7	11	
16	13	14	15	16	16	4	8	12	

The switch SFN_1 is non-blocking on packet permutations and for any distribution of N_1 packets among outputs when no more than p packets are sent to each output.

The switch SFN₁ has the following performance characteristics: the number of channels $N_1 = p^2$, the signal period $\tau = p$, and the switching complexity $\Sigma = S_1N_1 + 2pN_1 = 2p^2N_1 + 2pN_1 = 2p^3(p + 1)$. For any switches, we will also introduce the layout complexity by the number of simplex channels in them. For SFN₁, this complexity is given by $\Lambda = 2pN_1 = 2p^3$. All performance characteristics of the switch are combined in Table 3. We consider the switch SFN₁ to be onestage, according to the number of stages of the switches SSp. Below, the number of stages will be interpreted by the number of stages of the switches SSp. We emphasize that the switch SFN_1 is nonblocking on packet permutations and for any distribution of N_1 input packets among outputs when at most p packets are sent to each output.



Fig. 3. Switch SF16 with quasi-complete digraph topology. Boxes correspond to switches SS4, and triangles to multiplexers M4 and demultiplexers D4.

Table 3

Performance characteristics of switch SF N1

р	2	3	4	5	6	7	8
N_1	4	9	16	25	36	49	64
Σ	48	216	640	1500	3024	5488	9216
Σ	$N_1^{2.79}$	$N_1^{2.45}$	$N_1^{2.33}$	$N_1^{2.27}$	$N_1^{\ 2.24}$	$N_1^{\ 2.21}$	$N_1^{2.19}$
Λ	16	54	128	250	432	686	1024
Λ	N_1^{2}	$N_1^{1.82}$	$N_1^{1.75}$	$N_1^{1.72}$	$N_1^{1.69}$	$N_1^{1.68}$	N1 ^{1.67}

The table is extended to p = 8 since a dual photon 8×8 switch, SS8, has already been developed; see [24].

In Sections 2 and 3, the dual switches SS2 and SF4, shown in Fig. 4, will be used. Note that the one-stage switch SF4 has the same signal period and number of channels as the two-stage switch constructed from SS4 in the papers [1-3].



Fig. 4. Non-blocking self-routing switch SF4.

2. NON-BLOCKING SELF-ROUTING THREE-STAGE SWITCH

From the dual switch SF4, we construct a selfrouting two-stage network with exchange links-the network C₂16 shown in Fig. 5. It consists of two stages, and each stage includes four switches SF4. Unfortunately, this network is not a 16-channel nonblocking switch: it can have signal conflicts on the stage of the multiplexers M2, highlighted in gray. Signals in the first and second cycles may conflict. To resolve the conflicts, it suffices to provide these multiplexers with DL\delta, $\delta = 2$ (Fig. 6). Then, conditionally speaking, they will form the second stage of dual switches without demultiplexers. The result is a threestage non-blocking self-routing switch S₃16. It has the following performance characteristics: the number of channels $N_3 = 16 = 2^4$, the signal period $T_3 = 4$, the switching complexity $S_3 = 2N_1\Sigma|_{p=2} = 384 = N_3^{2.15}$, and the layout complexity $L_3 = 2N_1\Lambda + N_1|_{p=2} =$ $= 144 = N_3^{1.79}$

Let us compare the performance characteristics of this three-stage switch and the two-stage switch described in the introduction (the one composed of SS4). They have 16 channels each, the signal period ratio is $\gamma = T_3/\underline{T}_2 = 2/4 = 0.5$, and the switching complexity ratio is $\sigma = S_3/\underline{S}_2 = 384/256 = 1.5$. The product $\gamma \sigma = 0.75$ shows how many times the reduction in the signal period is less than the increase in the switching complexity.

Using the switches SF16 in a similar way, we can construct the two-stage self-routing network N₃256 with exchange links. It consists of two stages of 16 switches SF16 on each stage. This network may have conflicts on the multiplexers M4 of the first stage. For making N₃256 a three-stage non-blocking self-routing switch, it suffices to provide these multiplexers with DL δ , $\delta = 4$ (Fig. 7). The result is a three-stage nonblocking self-routing switch S₃256. It has the following performance characteristics: the number of chan-



Fig. 5. Two-stage network N₂16.



Fig. 6. Multiplexer M2 with delay lines for first stage.



Fig. 7. Multiplexer M4 with delay lines for first stage.

nels $N_3 = 256 = 4^4$, the signal period T3 = 11, the switching complexity $S_3 = 2N_1\Sigma \Big|_{p=4} = 20480 = N_3^{1.79}$, and the layout complexity $L_3 = 2N_1\Lambda + N_1 \Big|_{p=4} =$ $=4352 = N_3^{1.51}$. The values δ and T3 are justified in Lemma 1 below.

Let us compare the performance characteristics of this three-stage switch and the four-stage switch described in the introduction in the case p = 4. They

S

have 256 channels each, the signal period ratio is $\gamma = T_3/\underline{T}_2 = 10/49 \approx 0.204$, and the switching complexity ratio is $\sigma = S_3/\underline{S}_2 = 20 \ 480/8192 = 2.5$. The product $\gamma \sigma = 0.51$ shows how many times the reduction in the signal period is less than the increase in the switching complexity.

For an arbitrary p > 2, the three-stage switch S_3p^4 is constructed by analogy from the switches SFp^2 . It has the following property.

Lemma 1. The switch S_3p^4 is non-blocking if the multiplexer Mp of the first stage has DL δ with $\delta = p$ and the signal period

$$T_3 = p^2. (1)$$

Table 4

P r o o f. At their inputs, Mp receive p signals with their possible distribution on the interval from 1 to p cycles. In a conflict situation, they are distributed maximally on the interval from 1 to p cycles. For ensuring that in any conflicts, the primary signals are not superimposed on the signals delayed in the DL δ , a sufficient condition is $\delta = p$.

The maximum number of conflicting signals occurs when all signals are distributed among p cycles. In this case, conflict resolution will require retransmitting the conflicting signals through the DL δp times. As a result, they will be distributed on the interval from 1 to p^2 cycles. And the conclusion follows. \blacklozenge

Formula (1) was confirmed by numerical simulations on a switch model with the synchronous generation of arbitrary permutations of packets.

For an arbitrary $p \ge 2$, the switches S_3p^4 have the performance characteristics shown in Table 4.

Performance characteristics of switches S₃N₃

 $N_3 = p^4$ $T_{3} = p^{2}$ δ S_3 L_3 $384 = N_3^{2.15}$ $144 = N_3^{1.7}$ 2 16 4 $3888 = N_3^{1.88}$ 9 $1053 = N_3$ 3 81 3 $20\ 480 = N_3^{1.75}$ 256 16 $4352 = N_3$ 4 4 $75\ 000 = N_3^{1.74}$ 25 5 625 $13\ 125 = N_2$ 5 1296 36 $217\ 728 = N_3^{1.71}$ 6 $32\ 400 = N_{2}$ 6 537 824 = $N_3^{1.70}$ 2401 49 7 7 $69\ 629 = N_3$ 4096 64 $1\ 179\ 648 = N_3^{1.6}$ $135\ 168 = N_3^{1.4}$

3. NON-BLOCKING SELF-ROUTING BINARY TWO-STAGE SWITCH

The three-stage switch S_316 can be turned into the non-blocking two-stage switch S_316 by *the internal parallelization method*: we should cut the M2 stage

with DL*p* from the first stage (Fig. 5) and separate the conflicting signals across two copies of the second stage (Fig. 8). The outputs of these copies are combined by an additional M2 stage. The switch S₂16 switch has the following performance characteristics: the number of channels $N_2 = 16$, the signal period $T_2 = 2$, the switching complexity $S_2 = N_1(\Sigma - pN_1) + pN_2|_{p=2} = 576 = N_2^{2.29}$, and the layout complexity $L_2 = (p+1) N_1 \Lambda + pN_2|_{p=2} = 224 = N_3^{1.95}$.



Fig. 8. Non-blocking self-routing binary two-stage switch S_216 . Channels to second stage copies and from them are indicated by dotted lines.



Note that the switch S_216 has slightly greater complexity than S_316 . In addition, the two-stage switch S_216 and the two-stage switch composed of the switches SS4 [1–3] have the same number of channels, but the former switch has half the signal period of the latter.

4. NON-BLOCKING SELF-ROUTING PARY TWO-STAGE SWITCH

Using the internal parallelization method, from the switches SFN₁ ($N_1 = p^2$, p > 2), we can construct the non-blocking self-routing *p*-ary two-stage switch





 $S_2N_1^2$ or S_2N_2 (Fig. 9) with the number of channels $N_2 = p^4$. In Fig. 9, the trapezoidal block $B_{2,1}$ indicates the switch SF N_1 without the output multiplexers Mp, which has N_1 inputs and N_1 groups of outputs with p outputs each. The square block $B_{2,2}$ indicates the complete switch SF N_1 with N_1 inputs and N_1 outputs. The triangular block $B_{2,3}$ indicates the multiplexer Mp.

A set of N_1 blocks $B_{2,1}$ is indicated by B_2^* . In this set, the outputs of blocks $B_{2,1}$ are numbered by I, J, and k, where I ($1 \le I \le N_1$) gives the number of $B_{2,1}$ in B_2^* , J ($1 \le J \le N_1$) gives the group number of p outputs of $B_{2,1}$, and k ($1 \le k \le p$) gives the output number in the group.

> A set of N_1 blocks $B_{2,2}$ is indicated by B_2 . In this set, the inputs of blocks $B_{2,2}$ are numbered by I and J, where I $(1 \le I \le N_1)$ gives the number of $B_{2,2}$ in B_2 , and J $(1 \le J \le N_1)$ gives the input number of $B_{2,2}$. There are p copies of the set B_2 .

> Between the outputs of blocks $B_{2,1}$ in the set B_2^* and the inputs of blocks $B_{2,2}$ in each of the *p* copies of the set B_2 , there are exchange links, in which the outputs no. *I*, *J*, and *k* of the set B_2^* are connected to the inputs no. *J* and *I* of the *k*th copy of the set B_2 .

> The set B_2^* and each set B_2 in the *k*th group have exchange links: the outputs no. *I*, *J*, and *k* of blocks $B_{2,1}$ are connected to inputs no. *J* and *I* of blocks $B_{2,2}$ in the *k*th copy of sets B_2 . The outputs of *k* copies of sets B_2 are combined by N_2 blocks into the outputs of the switch S_2N_2

Each block $B_{2,1}$ has the switching complexity $S_{2,1} = \sum -pN_1 = 2p^2N_1 + pN_1$ and the layout complexity $L_{2,1} = \Lambda = 2pN_1$ along with the output lines on a copy of $B_{2,2}$. Each copy of $B_{2,2}$ switching has the complexity $S_{2,2} = \sum = 2p^2 N_1 + 2p N_1$ and the layout complexity $L_{2,2} = \Lambda = 2pN_1$. Each block $B_{2,3}$ has the switching complexity $S_{2,3} = p$ and the layout complexity $L_{2,3} = p$ along with the input lines from copies of $B_{2,2}$. As a result, the switching and layout complexities of the switch $\begin{aligned} S_2 N_2 & \text{are } S_2 = N_1 S_{2,1} + p N_1 S_{2,2} + \\ &+ N_2 S_{2,3} = 2 N_2 (p^3 + 2p^2 + p) = 2 (N_2^{7/4} + \\ &+ 2 N_2^{3/2} + N_2^{5/4}) \text{ and } L_2 = N_1 L_{2,1} + \\ &+ p N_1 L_{2,2} + N_2 L_{2,3} = N_2 (2p^2 + 2p + \\ &+ p) = 2 N_2^{3/2} + 3 N_2^{5/4}, \text{ respectively. Table} \end{aligned}$ 5 shows the performance characteristics of the switches S_2N_2 compared to the switches S_3N_3 (Table 4).



Ş

According to Table 5, for the same *p*, the twostage switches S_2N_2 have a considerably greater number of channels $(N_2 = \underline{N_2}^2)$ than the two-stage switches composed of the switches SS*p* [1–3]. They both have the same signal period $(T_2 = \underline{T_2} = p)$ but different switching complexities $(S_2 \approx N_2^2 \text{ vs. } \underline{S_2} = 2\underline{N_2}^{3/2}, \text{ re$ $spectively}).$

5. EXTENSIONS OF NON-BLOCKING SELF-ROUTING #ARY SWITCHES

This section will apply the invariant extension method of system area networks [18], which is intended to enlarge the number of users by increasing the network complexity without changing transmission delays. To extend the switches S_2N_2 in this method, separate multiplexers Md and demultiplexers Ddare used, where d is a divider of N_2 , i.e., $d = p^r$ (r = 1, 2, ...). In the papers [1–3], this method was adopted in the case r = 1 for extending the switches constructed in the purely switching scheme base.

The invariant extension method consists in the following. Let us take d^2 switches S_2N_2 . Each of them is divided into $n = N_2/d$ zones with d ports in each. (A port is an "input–output" pair.) Parts of the switches S_2N_2 in any zone are also $d \times d$ switches. All together, they are the backbone element of the switch SF_1d^2 with the quasi-complete digraph topology. (For example, see Fig. 3 for d = p = 4 and Fig. 4 for d = p = 2.) To form each such switch, it suffices to connect d^2 inputs to it through multiplexers Md, and d^2 outputs through demultiplexers Dd according to the corresponding interconnection table (for example, see Table 2 for d = p = 4). These connections yield the extended switch SE₂R₂, where $R_2 = nd^2 = dN_2$. In this switch, from any input the signal comes to the unique (!) copy of the switch S₂N₂, then passing to any given output without additional delays. Therefore, SE₂R₂ is a non-blocking self-routing switch, like S₂N₂.

An example of extending the switch SE₂16 to the switch SE₂32 with d = p = 2 is illustrated in Table 6 and Fig. 10.



Fig. 10. Interconnection diagram of inputs and outputs of switch SE_232 in *i*th zone of switch SE_216 .

The switch SE₂R₂ has the switching complexity $S_2^* = d^2S_2 + 2R_2d = d^22(N_2^{7/4} + 2N_2^{3/2} + N_2^{5/4} + N_2)$ and the layout complexity $L_2^* = d^2L_2 + 2R_2d = d^2(2N_2^{3/2} + 3N_2^{5/4} + N_2)$. Table 7 presents the performance characteristics of SE₂R₂ with some values of p and d. Note that the extended switch preserves its specific complexity (per channel) when increasing the number of its channels and maintaining its speed (signal period).

Table 5

р	$N_2 = p^4$	$T_2 = N_2^{1/4}$	S_2	L_2	$\gamma = S_2/S_3$	$\sigma = T_3/T_2$	γ/σ
2	16	2	$576 = N_2^{2.29}$	$224 = N_2^{1.95}$	1.50	2	0.75
3	81	3	$7776 = N_2^{2.04}$	$2187 = N_2^{1.75}$	2.00	3	0.67
4	256	4	$51\ 200 = N_2^{1.96}$	$11\ 264 = N_2^{1.68}$	2.50	4	0.63
5	625	5	$225\ 000 = N_2^{1.91}$	$40\ 625 = N_2^{1.65}$	3.00	5	0.60
6	1296	6	$762\ 048 = N_2^{1.89}$	$116\ 640 = N_2^{1.63}$	3.50	6	0.58
7	2401	7	$2\ 152\ 296 = N_2^{1.87}$	$285\ 719 = N_2^{1.61}$	4.00	7	0.57
8	4096	8	$5\ 308\ 416 = N_2^{1.86}$	$622\ 592 = N_2^{1.60}$	4.50	8	0.56

Performance characteristics of switches S₂N₂

Table 6

b /

Arranging switches SS2 in extended switch SE₂32 with d = p = 2

	Ports of switch S ₂ 16															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Copy of S ₂ 16	Composition of switches SE ₂ 32 from switches SS2															
1	1 SS2		2 5	2 SS2 3 SS2		4 S	SS2	5 SS2		6 SS2		7 SS2		8 SS2		
2	1 SS2 2 SS2		SS2	3 S	SS2 4 SS2		SS2	5 5	SS2	6 5	SS2	7 S	SS2	8 S	SS2	
3	1 SS2 2		2 5	SS2	3 SS2		4 SS2		5 5	SS2	6 5	SS2	7 S	SS2	8 S	SS2
4	1 5	SS2	2.5	SS2	3 S	SS2	4 S	SS2	5 5	SS2	6 5	SS2	7 S	SS2	8 S	SS2

р		2	3	4	5	6	7	8
	T ₂	2	3	4	5	6	7	8
	R ₂	32	243	1024	3125	7776	16 807	32 768
d = p	S_2^*	$R_2^{2.25}$	$R_2^{2.03}$	$R_2^{1.97}$	$R_2^{1.93}$	$R_2^{1.91}$	$R_2^{1.90}$	$R_2^{1.89}$
	L_2^*	$R_2^{1.89}$	$R_2^{21.74}$	$R_2^{1.69}$	$R_2^{1.66}$	$R_2^{1.65}$	$R_2^{1.64}$	$R_2^{1.63}$
$d = p^2$	R ₂	64	729	4096	15 625	46 656	117 649	262 144
	S_2^*	$R_2^{2.21}$	$R_2^{2.03}$	$R_2^{1.97}$	$R_2^{1.94}$	$R_2^{1.93}$	$R_2^{1.92}$	$R_2^{1.91}$
	L_2^*	$R_2^{1.91}$	$R_2^{1.78}$	$R_2^{1.74}$	$R_2^{1.72}$	$R_2^{1.71}$	$R_2^{1.7}$	$R_2^{1.69}$
$d = p^3$	R ₂	128	2187	16 384	78 125	279 936	823 543	2 097 152
	S_2^*	$R_2^{2.18}$	$R_2^{2.02}$	$R_2^{1.98}$	$R_2^{1.95}$	$R_2^{1.94}$	$R_2^{1.93}$	$R_2^{1.92}$
	L_2^*	$R_2^{1.92}$	$R_2^{1.81}$	$R_2^{1.78}$	$R_2^{1.76}$	$R_2^{1.75}$	$R_2^{1.74}$	$R_2^{1.74}$
$d = p^4$	R ₂	256	6561	65 536	390 625	1 679 616	5 764 801	16 777 216
	S_2^*	$R_2^{2.16}$	$R_2^{2.02}$	$R_2^{1.98}$	$R_2^{1.96}$	$R_2^{1.95}$	$R_2^{1.94}$	$R_2^{1.93}$
	L_2^*	$R_2^{1.93}$	$R_2^{1.84}$	$R_2^{1.8}$	$R_2^{1.79}$	$R_2^{1.78}$	$R_2^{1.77}$	$R_2^{1.77}$

Performance characteristics of switch SE_2R_2 with some values of p and d

CONCLUSIONS

This paper has proposed a technique for constructing non-blocking self-routing photon switches of wide scalability based on new dual photon switches with a small number p of channels. In dual switches, signal conflicts are resolved by distributing them either among different channels or among different cycles. The latter method requires increasing the signal period by p times for a p-channel dual $p \times p$ switch.

Dual photon switches of wide scalability have been constructed with the minimum possible signal period, which is *p* times greater than the signal duration. Scalability is achieved using switches with the quasi-complete digraph topology and the invariant extension of any networks based on them. This method allows increasing the number of network channels while maintaining such properties as the signal period, non-blocking, and self-routing through its parallelization with increasing complexity. This method involves an extended element base consisting of $p \times p$ switches, $1 \times p$ demultiplexers, and $p \times 1$ multiplexers.

Three- and two-stage non-blocking self-routing *N*channel *N*×*N* switches with $N = p^4$ have been constructed. In the first of these, the period–complexity balance is shifted towards a lower complexity under a large period. In the two-stage switch, the balance is shifted towards the minimum period under high complexity. Its switching complexity turned out to be comparable to the complexity of a full switch, and its layout complexity turned out to be significantly less than that of a full switch. These switches have only two stages of *p*×*p* switches and two stages of 1×*p* demultiplexers and *p*×1 multiplexers.

The above switches have been extended to *R*channel $R \times R$ switches with $R = N^r$ (r = 1, 2, 3, 4) using another stage of $1 \times d$ demultiplexers and $d \times 1$ multiplexers with $d = p^r$. These *R*-channel switches are non-blocking self-routing switches with a minimum signal period. Their switching complexity is comparable to that of a full switch, and the channel complexity is significantly less.

In comparison with the papers [1-3], the main results of this work are a significant reduction in the signal period with the same number of channels and a significant increase in the number of channels with the same single period. In other words, the key novel-ty consists in developing a method for constructing non-blocking self-routing dual photon switches with the minimum possible signal period and the feasibility of achieving the maximum number of channels under almost the same complexity.

This paper, together with [1-3], has presented a technique for designing fundamentally new dual system networks with the following properties:

- They are non-blocking networks with static selfrouting of packets, i.e., networks with conflict-free self-routing on arbitrary packet permutations.

- They have the widest scalability with the maximum speed reached on them and the complexity comparable to that of a full switch.

- Their maximum speed is only 2–4 times less than the physically achievable speed.

- They allow balancing the speed-complexity ratio with decreasing their complexity to that of the non-blocking Clos network (which has no conflictfree self-routing procedures) but at a significantly lower speed of the non-blocking network.

In the future, we expect to make dual networks channel fault-tolerant by replacing the quasi-complete digraph topology with the quasi-complete graph one. In addition, when cascading dual networks, we expect to use internal parallelization (Sections 3 and 4) instead of external parallelization (invariant extension). This approach will enable the scaling of dual networks with smaller costs.

Table 7





REFERENCES

- Barabanova, E.A., Vytovtov, K.A., and Podlazov, V.S., Multistage Switches for Optical and Electronic Supercomputer Systems, *Proceedings of the 8th National Supercomputer Forum* (*NSCF-2019*), Pereslavl-Zalessky, 2019. – URL: http://2019.nscf.ru/TesisAll/02_Apparatura/037_BarabanovaE A.pdf. (In Russian.)
- Barabanova, E.A., Vytovtov, K.A., Vishnevsky, V.M., and Podlazov, V.S., The New Principle for the Construction of Optical Information Processing Devices for Information-Measuring Systems, *Sensors and Systems*, 2019, no. 9, pp. 3–9. (In Russian.)
- Barabanova, E., Vytovtov, K., Podlazov, V., and Vishnevskiy, V. Model of Optical Non-blocking Information Processing System for Next-generation Telecommunication Networks, *Proceedings of the 22nd International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications (DCCN–2019)*, Moscow, 2019. Communications in Computer and Information Science, vol. 1141, Cham: Springer, pp. 188–198. DOI: 10.1007/978-3-030-36625-4 16.
- Pipenger, N., On Rearrangeable and Non-blocking Switching Networks, J. Comput. Syst. Sci., 1978, vol. 17, pp. 307–311.
- Scott, S., Abts, D., Kim, J. and Dally, W., The Black Widow High-radix Clos Network, *Proc.* 33rd *Intern. Symp. Comp. Arch.* (*ISCA*'2006), Boston, 2006. https://www.researchgate.net/publication/4244660_The_Black Widow High-Radix Clos Network.
- 6. Benes, V.E., Mathematical Theory of Connecting Networks and Telephone Traffic, New York: Academic Press, 1965.
- Bhuyan, L.N., and Agrawal, D.P., Generalized Hypercube and Hyperbus Structures for a Computer Network, *IEEE Trans. on Computers*, 1984, vol. C-33, no. 4, pp. 323–333.
- Alverson, R., Froese, E., Kaplan, L., and Roweth, D., Cray[®] XCTM Series Network, Cray Inc., 2012. https://www.cray.com/sites/default/files/resources/CrayXCNet work.pdf.
- 9. Gu, Q.P., and Tamaki, H., Routing a Permutation in Hypercube by Two Sets of Edge-Disjoint Paths, *J. of Parallel and Distributed Comput.*, 1997, vol. 44, no. 2, pp. 147–152.
- 10.Lubiw, A. Counterexample to a Conjecture of Szymanski on Hypercube Routing, *Inform. Proc. Let.*, 1990, vol. 35(2), pp. 57–61.
- 11.Podlazov, V.S., Conflict-Free Self-Routing for a Three-Dimensional Generalized Hypercube, *Control Sciences*, 2018, no. 3, pp. 26–32. (In Russian.)
- 12.Karavai, M.F. and Podlazov, V.S., Sistemnye seti s pryamymi kanalami dlya parallel'nykh vychislitel'nykh sistem kombinatornyi podkhod. Glavy 5, 6 (System Area Networks with Direct Channels for Parallel Computing Systems. Chapters 5 and 6), URL:

https://www.ipu.ru/sites/default/files/publications/18125/15058 -18125.pdf.

- 13.Finn, A.M. and Decker, R.O., A Network Architecture for Radar Signal Processing, AIAA/IEEE 8th Digital Avionic Systems Conference, San Jose, California, 1988, pp. 614–621.
- 14.Hall, M., Combinatorial Theory, Waltham: Blaisdell Publishing Company, 1967.
- 15.Cho, Y., Chi, C., and Chung, I., An Efficient Conference Key Distribution System Based on Symmetric Balanced Incomplete Block Design, in *Lecture Notes Comput. Sci.* (*LNCS*), 2003, vol. 2657, pp. 147–154.
- 16.Lee, O., Lee, S., Kim, S., and Chung, I., An Efficient Load Balancing Algorithm Employing a Symmetric Balanced In-

complete Block Design, Lecture Notes Comput. Sci. (LNCS), 2004, vol. 3046, pp. 647–654.

- 17.Karavai, M.F., Parkhomenko, P.P., and Podlazov, V.S. Combinatorial Methods for Constructing Bipartite Uniform Minimal Quasicomplete Graphs (Symmetrical Block Designs), *Automation and Remote Control*, 2009, vol. 70, no. 2, pp. 312– 327.
- 18.Karavai, M.F. and Podlazov, V.S., An Invariant Extension Method for System Area Networks of Multicore Computational Systems. An Ideal System Network, *Automation and Remote Control*, 2010, vol. 71, no. 12, pp. 2644–2654.
- 19.Alverson, R., Roweth, D., and Kaplan, L., The Gemini System Interconnect, 18th IEEE Symposium on High Performance Interconnects, 2009, pp. 3–87.
- 20.Arimili, B., Arimili, R., Chung, V., et al., The PERCS High-Performance Interconnect, 18th IEEE Symposium on High Performance Interconnects, 2009, pp. 75–82.
- 21.URL:

https://www.mellanox.com/pdf/whitepapers/IB_vs_Ethernet_C lustering_WP_100.pdf.

- 22.URL: https://dlcdnets.asus.com/pub/ASUS/mb/accessory/PEM-FDR/Manual/Mellanox_OFED_Linux_User_Manual_v2_3-1_0_1.pdf.
- 23.Stepanenko, S.A., Photon Computer: Structure and Algorithms, Parameter Estimates, *Fotonika*, 2017, no. 7 (67), pp. 72–83. (In Russian.)
- 24.Barabanova, E., Vytovtov, K., and Podlazov, V., Model and Algorithm of Next Generation Optical Switching Systems Based on 8×8 Elements, *Proceedings of the 22nd International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications (DCCN-2019)*, Moscow, 2019, pp. 58–70. DOI: 10.1007/978-3-030-36614-8 5.

This paper was recommended for publication by V.M. Vishnevsky, a member of the Editorial Board.

Received April 30, 2020, and revised September 9, 2020. Accepted October 5, 2020.

Author information

Barabanova, Elizaveta Aleksandrovna. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, ⊠ elizavetaalexb@yandex.ru

Vytovtov, Konstantin Anatol'evich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, ⊠ vytovtov_konstan@mail.ru

Podlazov, Viktor Sergeevich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, impodlazov@ipu.ru

Cite this article

Barabanova, E.A., Vytovtov, K.A., Podlazov, V.S. Two-Stage Dual Photon Switches in an Extended Scheme Basis. *Control Sciences* **1**, 60–69 (2021). http://doi.org/10.25728/cs.2021.1.7

Original Russian Text © Barabanova, E.A., Vytovtov, K.A., Podlazov, V.S., 2021, published in *Problemy Upravleniya*, 2021, no. 1, pp. 69–81.

