Nearly-All-Optical Routing in Sparse Optical Tori

Risto T. Honkanen

Report A/2006/3

ISBN 951-781-273-6
ISSN 1795-9195

UNIVERSITY OF KUOPIO
Department of Computer Science

P.O.Box 1627, 70211 Kuopio, FINLAND

Nearly-All-Optical Routing in Sparse Optical Tori

Risto T. Honkanen
University of Kuopio, Department of Computer Science
P.O.Box 1627, 70211 Kuopio, FINLAND

May 14, 2006

Abstract

In this paper we present an all-optical network architecture and a nearly-all-optical router for it. The sparse optical torus network consists of an $n \times n$ torus, where processors are deployed diagonally. Addresses of packets are encoded and recognized by using fiber Bragg grating arrays. The optical address recognition ensures that only a few logical gates are needed to implement routing decisions at the routing nodes.

Keywords sparse optical torus, optical communication, fiber Bragg grating

1 Introduction

Most of the parallel computers on the market use an electronic communication network to transmit packets between processors. We can increase the efficiency of communication by using optics. Optical communication offers several advantages in comparison with its electronic counterpart, e.g., the possibility to use broader bandwidth and the insensitivity of optics to external interferences.

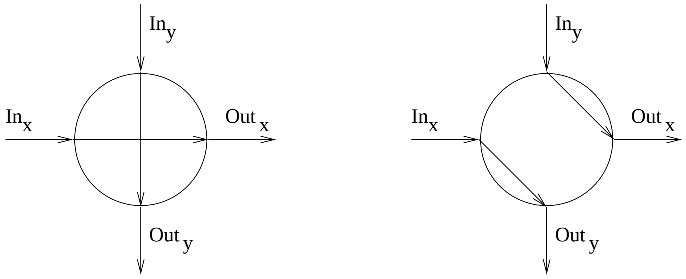
Problems that have prevented a wider use of optical communication are the lack of all-optical logic gates and the lack of optical random access memories. These disabilities reduce the functionality of optical communication in packet switched networks. For these reasons packet-routed networks usually use electronic address recognition as a routing mechanism. LaRochelle et al. have proposed a fiber Bragg grating array encoder/decoder for *optical code-division multiple access* (OCDMA) [9].

Our work is motivated by the emulation of shared memory with a network of distributed memory modules [4, 12]. If a parallel algorithm has enough parallel "slackness", i.e., large enough number of independent parallel threads in each processor, the implementation of shared memory can be reduced to efficient routing of an h-relation [12]. An h-relation is a routing problem where each processor has at most h packets to send and is the target of at most h packets [1]. An implementation of an h-relation is said to be *work-optimal* at *cost* c, if all the packets arrive at their targets in time ch. A precondition of work-optimality is that h is greater than the diameter ϕ of the network and the network can move $\Omega(n\phi)$ packets in each step, where n is the number of processors. Otherwise slackness cannot be used to "hide" the latency due to travel across the diameter [1, 7]. Some architectural solutions satisfying the above are presented, e.g, by Valiant [12] and by Sibeyn [11].

A sparse optical torus network ($\mathcal{SOT}$) consists of an $n \times n$ torus, where processors are deployed diagonally. Nodes are connected to each other by optical links [7]. The bandwidth of the system is divided in time slots, whose length t_p equals the bypass time of a packet between two consecutive routing nodes. The system operates synchronously. Every node has two incoming and two outgoing links, therefore the routing machinery can move $2n^2$ packets in each time slot.

The routing protocol used in this work is what we call a one-way lateral protocol. In the one-way lateral protocol a packet follows the horizontal path until it reaches its target column. When the packet has reached the target column, it tries to turn to its target processor. If it succeeds in turning, it has the privilege to continue to the destination. Problems arise if there is another packet sent at the same time to the same target. The packets arrive at the same router at the same time, and if they choose the same outgoing link, they collide. Collisions can be handled by using an address difference based protocol [7], or a systolic routing protocol based on the use of time-controlled routers and timed sendings of packets [6]. In this presentation we will, however, present an almost-all-optical address recognition mechanism and a fast hardware-based implementation of the one-way lateral protocol. We also sketch a theoretical analysis earlier presented in [7]. According to the simulations, an $n \times n$ $\mathcal{SOT}$ offers work-optimal routing of h-relation for $h \in \Omega(n \log n)$.

Section 2 presents the structure of routing nodes and the structure of a sparse optical torus. In Section 3 we introduce the one-sided lateral routing protocol and its implementation. Section 4 presents an analysis of our construction. Section 5 gives conclusions and discusses the future work.



(a) Cross state (+)   (b) Turn state (\ \)

Figure 1: Two possible states of routing nodes

2 Sparse Tori with Optical Routers

The basic component in routing nodes is the electrically controlled optical 2×2 switch, operating under a common clock. A routing node can be in one of two states. When a routing node routes signals In_y and In_x to outputs Out_y and Out_x respectively, it is said to be in the *cross* state ($+$), and when it routes incoming signals In_y and In_x to outputs Out_x and Out_y respectively, it is said to be in the *turn* state ($\backslash\backslash$). The two possible states of routing nodes are presented in Figure 1.

A 2-dimensional $n \times n$ *sparse optical torus*, $\mathcal{SOT}$, consists of n^2 optical routing nodes $\mathcal{R}_{(i,j)}$, $2n^2$ optical links, and n processors $(\mathcal{P}_0, \mathcal{P}_1, \ldots, \mathcal{P}_{(n-1)})$, where $0 \leq i, j \leq n-1$. Processors are located diagonally so that processor $\mathcal{P}_i$ resides at routing node $\mathcal{R}_{(n-i-1,i)}$ [7]. Each node has two incoming and two outgoing links In_x, In_y, Out_x, and Out_y. The two outgoing links of node $\mathcal{R}_{(i,j)}$ are connected to nodes at locations $\mathcal{R}_{((i+1) \bmod n, j)}$ and $\mathcal{R}_{(i,(j+1) \bmod n)}$. All the connections are assumed to be unidirectional and each routing node is assumed to be capable of receiving and sending one packet per link in a time unit. An example of a 6×6 sparse optical torus is represented in Figure 2.

In Figure 2 a disk indicates a routing node, a square indicates a routing node with a processor, and an arrow between two routing nodes indicates a unidirectional optical link between the nodes. The routing time of the system is divided in time slots, whose length t_p equals to the bypass time of a packet via a link between two consecutive routing nodes. We call the length of time slot t_p the packet cycle. A packet consists of so many data bits that can be passed within the packet cycle t_p between two consecutive routing nodes.

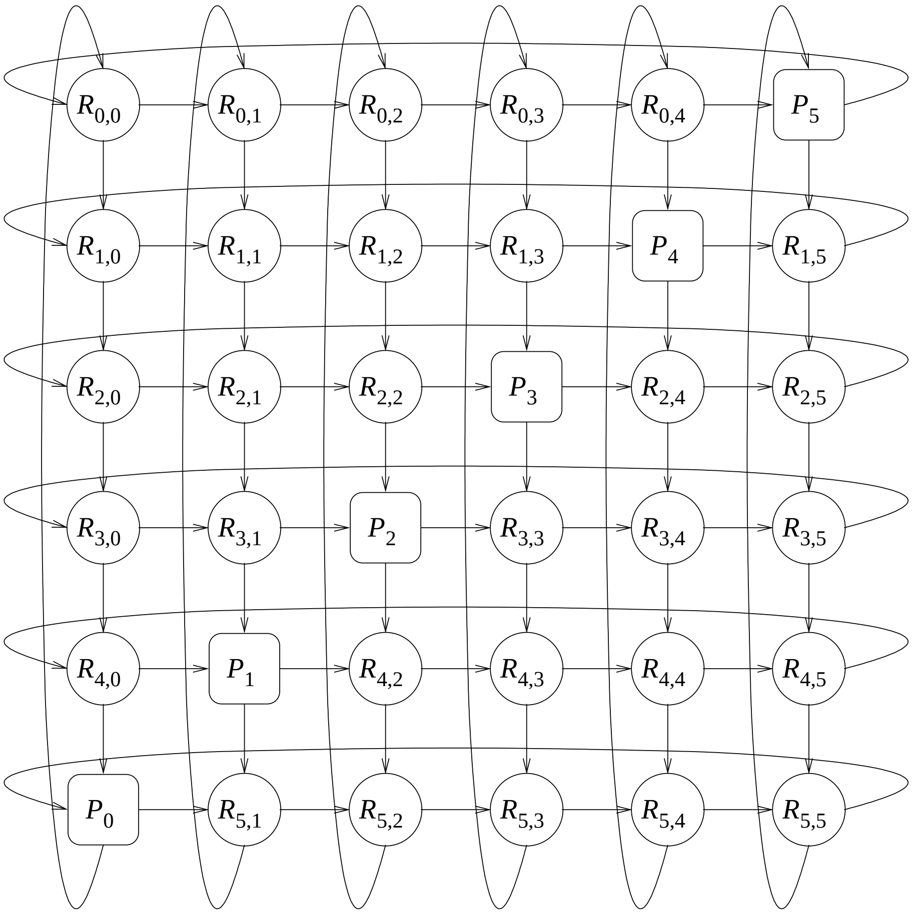


Figure 2: A 6×6 sparse optical torus.

3 Routing Machinery for $\mathcal{SOT}$

In our system, processors inject packets into the network using fiber Bragg grating arrays (BGA) to encode the address of the destination. Intermediate routers make routing decisions depending on the address of the routing node, the destination addresses of the packets, and the routing protocol. Section 3.1 presents the routing protocol for $\mathcal{SOT}$ and its implementation. Section 3.2 discusses an optical code division multiple access (OCDMA). In Section 3.3 we present the optical address recognition mechanism of the system. Section 3.4 discusses the feasibility of our construction.

3.1 Routing in Sparse Optical Tori

At each time step the routers of a sparse optical torus receive at most two packets from their inputs In_x and In_y. The incoming packets should be routed to outputs Out_x and Out_y according to their destination. Routing in $\mathcal{SOT}$ can be divided in three phases. First, we have to detect and evaluate the values of the optical addresses of incoming packets. Second, we have to switch the state of the router so that it routes the packets correctly. Third, packets must be injected into the output links according to the routing decision.

Let p and q be packets at the input links In_x and In_y of router $\mathcal{R}_{(i,j)}$ addressed to processors $\mathcal{P}_a = \mathcal{R}_{(a,n-a)}$ and $\mathcal{P}_b = \mathcal{R}_{(b,n-b)}$, respectively. The routing protocol is characterized by the function $f(a,b,i,j) \in \{+, \backslash\backslash\}$, which tells the new state of the router and thus determines, how the packets are forwarded. If one of packets p and q is missing, the behavior of the router is obvious. More interesting is the case of two packets, when a collision is possible.

Figure 3 represent the routing schema for $\mathcal{SOT}$. Let $d_x(i,a) = a - i$ and $d_x(i,b) = b - i$ denote distances from the destination columns and $d_y(j,a) = n - j - a$ and $d_y(j,b) = n - j - b$ denote distances from destination row of packets P_a and P_b. In the general case the state of router can be either $+$ or $\backslash\backslash$ until at least one of the packets reaches its destination row or column. "Safe" routing areas are presented in Figure 3 as dash-dotted boxes. A number of routing protocols and mechanisms have been presented in [6, 7, 8]. Making the routing decision in optical form is, however, very difficult.

Actually, we will define f so that the value only depends on whether $a = i$ and $b = j$ or not, i.e. $f(a,b,i,j) = f'(g(a,i), g(b,j))$, where

$$g(k,l) = \begin{cases} 0 & \text{if } k = l \\ 1 & otherwise \end{cases}$$

and

$$f'(A,B) = \begin{cases} \backslash\backslash & \text{if } A = 0 \text{ and } B = 1 \\ + & otherwise \end{cases}$$

The function f' is an implementation of the one-way lateral protocol presented in [7]. The input bits are provided by optical detection and by optical address recognition. Respectively, the output bit is provided by the switching circuit of Figure 4 which presents the block diagram of protocol circuit (ProtC) design implementing the one-way lateral protocol in $\mathcal{SOT}$. Inputs A and B are connected to packet de-

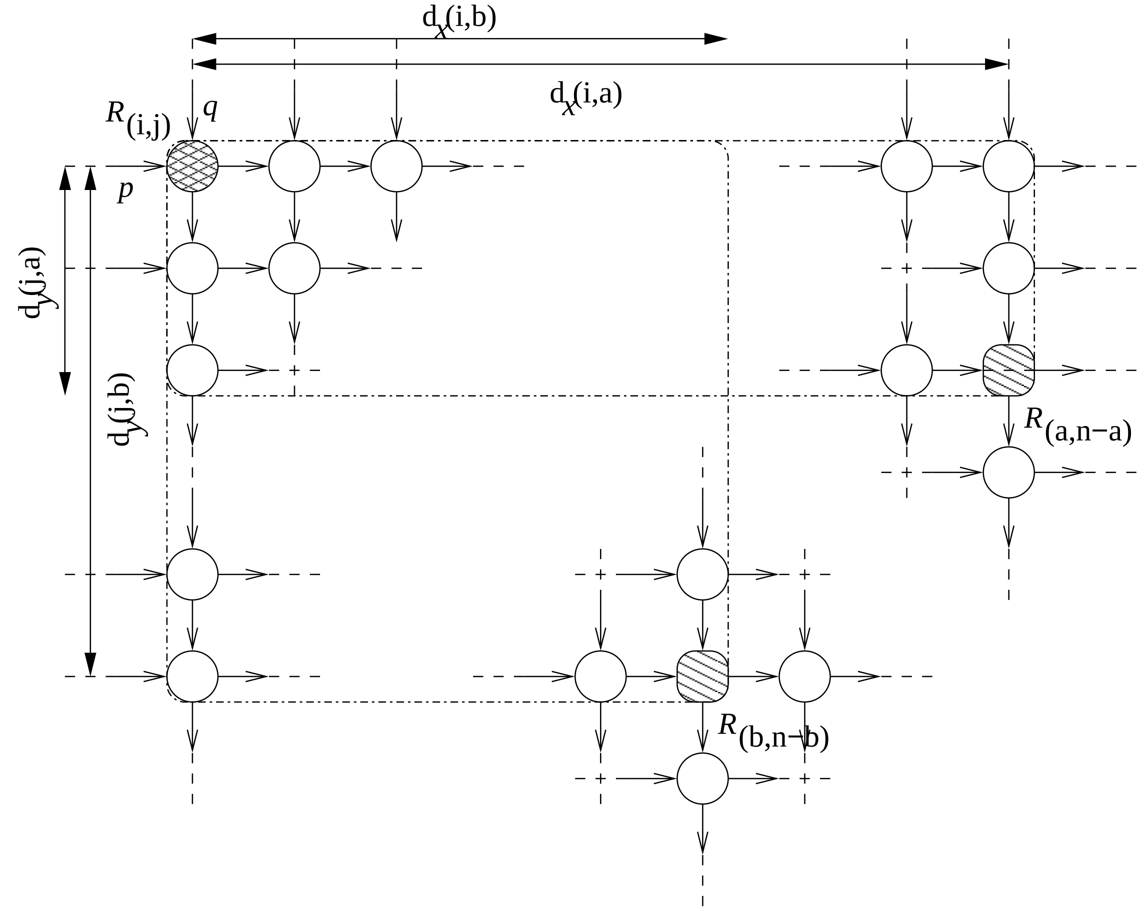


Figure 3: Example of routing in $\mathcal{SOT}$. Processors $\mathcal{P}_a$ and $\mathcal{P}_b$ resize at routers $\mathcal{R}_{(a,n-a)}$ and $\mathcal{R}_{(b,n-b)}$.

tection and address recognition circuits of incoming packets and output C controls the state of a 2×2 switch. The D flip-flop is added into the design to stabizile the state of switch during the packet cycle t_p.

3.2 Optical CDMA

LaRochelle et al. have proposed the use of fiber Bragg grating arrays (BGA) for optical code-division multiple access [9]. We consider the Bragg grating to be a periodic perturbation of the effective refractive index of an optical fiber that reflect the predetermined wavelength λ_i of light incident on the grating. The other wavelengths of the light are passed by. A fiber Bragg grating array consists of an optical fiber and a number of Bragg gratings written in the fiber [9].

Let us consider an OCDMA system which uses wavelengths $(\lambda_1, \lambda_2, \ldots \lambda_N)$ in

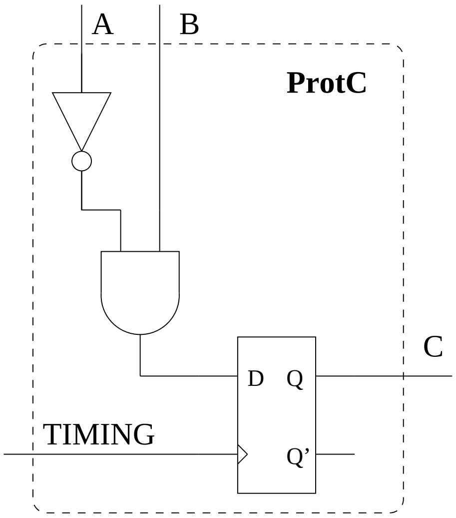


Figure 4: Circuit design implementing one-sided lateral protocol in $\mathcal{SOT}$.

its encoders and decoders. The encoding end of the transmission system consists of a broadband transmitter, a circulator, and a tunable BGA encoder [2, 3, 5, 9]. The decoding end consists of a circulator, a BGA decoder, and a detector respectively. At the encoding end the broadband pulse is multiplied to a *chip pulse sequence* having n equally separated pulses. The distance between pulses equals the bypassing time of light between Bragg gratings.

Figure 5 presents the idea of optical address recognition. At the encoding end the transmitted broadband pulse is separated into a pulse train of three pulses $\langle\lambda_2, \lambda_1, \lambda_3\rangle$ by the tunable encoder (Enc_1). Pulses are routed through the network until they reach decoder Rec_1 that has the same kind of Bragg grating array as Enc_1. To detect one large pulse at the threshold analyzer, the gratings of Rec_1 must be in inverse order compared to gratings of Enc_1.

3.3 Optical processing in $\mathcal{SOT}$

Disadvantages in optical packet switching are the lack of all-optical logic gates and the lack of all-optical random access memories. For these reasons packet-routed networks usually use electronic addressing as a routing mechanism. We can, however, use fiber Bragg gratings to minimize electronic components needed.

Address encoding. Address encoding in $\mathcal{SOT}$ is rather straightforward. When a processor has a packet to send destined to the processor at the ith column, it tunes its encoder to use code c_i and sends the address pulse train into horizontal outgoing

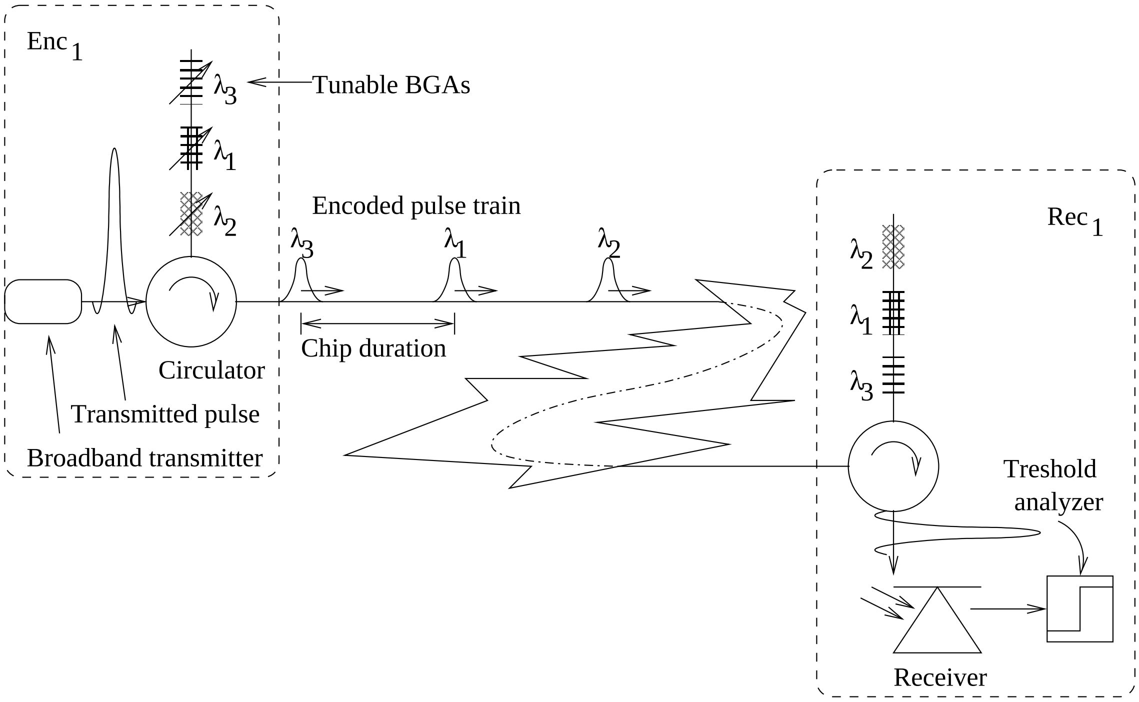


Figure 5: Block diagram presenting optical address recognition.

link before data. Data can be sent without coding after that.

Detecting turned packets. According to the one-sided lateral protocol a packet coming from the input In_y has been turned towards its target processor and must always be routed into the output Out_y. To detect this situation it needs to split the incoming signal (if there is any) in two parts and guide one part into detection. The other part is guided to a 2×2 switch after amplification. Block diagram implementing detection of the existence of turned packet (DetT) is presented in Figure 6.

Address recognition. Routing and address recognition of the packet coming from the input In_x is more complicated because this packet is routed into output Out_y only if there is no incoming packet from In_y and the packet should be turned towards its target processor.

Address recognition circuits of routers of $\mathcal{SOT}$ are chosen so that recognition circuits of each column i recognize the same code $c_i \in \mathcal{C}$, where $\mathcal{C}$ is the set of all address codes of the system. Additionally, when using BGA's as an addressing system for $\mathcal{SOT}$, the receivers' codes are chosen to satisfy two other conditions. First, decoders of each router of the column are chosen so that the peak of the autocorrelation function

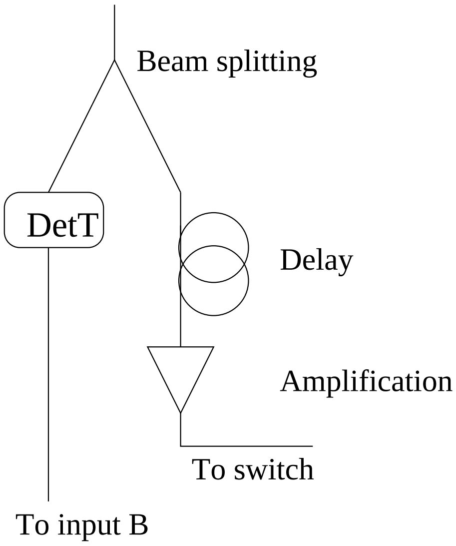


Figure 6: Block diagram presenting detection of turned packet.

$$R_k(s) = \sum_{i=0}^{N} c_k(i)c_k(i-s)$$

is maximized for each code k; parameter N is here the length of the code, and $-N+1 \leq s \leq N-1$. Second, the side-lobes of this function are minimized [3]. Demands for minimization of cross-correlation are omitted because of the lack of multiple accesses. To do that, it is needed to use the same kind of Bragg grating arrays in the encoding end, but in the inverse order.

Routing decision and routing in $\mathcal{SOT}$. Routers of $\mathcal{SOT}$ consist of a detector circuit of turned packet (DetT), an optical address recognition circuit (Rec_i), a protocol circuit (ProtC), a 2×2 optical switch, and a number of optical components to handle optical signals. First, DetT detects the existence of incoming signal from In_y and sends bit 1 to the input B of ProtC if there is an incoming packet coming from In_y, 0 otherwise. Second, Rec_i evaluates whether the packet coming from In_x (if there is any) is trying to turn and sends bit 1 to the input A of ProtC in that case, 0 otherwise. Third, ProtC makes the routing decision according its input and sets the 2×2 switch in the correct state. Attenuators and circulators [5] enlarge optical powers of signals and delay them until the routing decision has been made respectively. Last, packets are routed into the outgoing links. A block diagram of our implementation is presented in Figure 7.

3.4 Physical Feasibility of Routers for $\mathcal{SOT}$

Switches can be implemented by $LiNbO_3$ technology [10]. The switching time of $LiNbO_3$ switches lies in the range of 10–15 ps [10]. The length of packet is $l_p = \frac{N_p \times v_c}{B \times r}$, where N_p is the size of the packet in bits, $v_c \simeq 0.3$ ns/m is the speed of light in vacuum, $r \simeq 1.5$ is the refraction index of fiber [10], and B is the link bandwidth. Assuming the bandwidth to be B=100 Gb/s, the length of a bit in a fiber is $l_p \simeq 2$ mm.

In order to estimate the feasibility of a 64×64 $\mathcal{SOT}$ let us assume the link bandwidth to be $B = 100$ Gb/s, and the size of packets to be $N_p = 128$ b. The corresponding length of a packet in a fiber is $l_p \simeq 26$ cm and the length of the time slot is $t_p \simeq 1.3$ ns. Assuming the length of clock cycle of processors to be $t_{cc} = 1$ ns (corresponding to the frequency of 1 GHz), it will take 1.3 clock cycles for a packet to travel between two adjacent routing nodes. The overall amount of fibers is $L_f \simeq 2100$ m, and the routing time of packets is $t_r \simeq 82$ clock cycles for each packets. Encoders consist of a series of Bragg gratings written at the same wavelength Λ_B [3]. Each grating is tuned independently using a piezoelectric device to adjust Λ_B to the given wavelength. Tuning is applied by stretching each grating independently [3]. We consider the requested parameters to be reasonable and the architecture to be feasible to implement in the near future.

4 Analysis of Routing

At the beginning of routing each processor has h packets to send. The packets to be sent are randomly chosen by the processors from sending buffers and sent to the routing machinery using horizontal link. If more than one packet is sent to the same destination $\mathcal{P}_d$, only one of them is successful. Consider that n packets are sent to their random destinations at a given moment of time. The probability that none of them is destined to a fixed processor $\mathcal{P}_d$ is $(1 - 1/n)^n \to 1/e$. In consequence, the probability that there are packets coming to this destination is $1 - 1/e$ and exactly one of them is successful. Hence, the expected number of successful packets is $n(1 - 1/e) \simeq 0.63 \times n$.

Routing h packets in time $\theta(h)$ implies work-optimality. For large h and n and assuming that $h \in \Omega(n \log n)$ the routing time is approximately $e/(e-1) = 1.59$. In simulations we ran 10 simulation rounds for each occurence using a Java simulator. The results of simulations are presented in Figure 8. Our software simulations confirm the assumptions.

5 Conclusions and Future Work

We have proposed a novel, nearly-all-optical address recognition and routing machinery for a sparse optical torus. On one hand, the use of optical address recognition based on BGA's enables that only a few electronic components are needed to implement the one-way lateral protocol at the routing nodes. On the other hand, the optical part of the system is simple and suitable for this purpose. For these reasons we think that our construction is feasible to construct. In future work, we will study to implement so called two-way lateral protocol in the same manner and the possibility to use optical CDMA in different kinds of network structures as an addressing mechanism.

Acknowledgments

I express my gratitude to professor Martti Penttonen for his valuable discussions on the topic and his guidance on the structure of the text.

References

[1] M. Adler, J.W. Byers, and R.M. Karp. Scheduling parallel communication: The h-relation problem. In *Proceedings of Mathematical Foundations of Computer Science*, pages 1–20, Prague, Czech Republic, 1995.

[2] H. Fathallah and L.A. Rusch. Robust optical ffh-cdma communications: Coding in place of frequency and temperature controls. *Journal of Lightwave Technology*, 17(8):1284–1293, 1999.

[3] H. Fathallah, L.A. Rusch, and S. LaRochelle. Passive optical fast frequency-hop cdma communications system. *Journal of Lightwave Technology*, 17(3):397–405, 1999.

[4] S. Fortune and J. Wyllie. Parallelism in random access machines. In *Proceedings of the 10th Annual Symposium on Theory of Computing*, San Diego, California, 1978.

[5] Jeff Hecht. *Understanding Fiber Optics*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, fifth edition, 2006.

[6] R. Honkanen. Systolic routing in sparse optical torus. In P. Kilpeläinen and N. Päivinen, editors, *Proceedins of the Eight Symposium on Programming Languages and Software Tools*, pages 14–20, 2003.

[7] R. Honkanen, V. Leppänen, and M. Penttonen. Hot-potato routing algorithms for sparse optical torus. In T.M. Pinkston, editor, *Proceedings of the 2001 ICPP Workshops*, pages 302–307, Valencia, Spain, 2001.

[8] R. Honkanen, V. Leppänen, and M. Penttonen. Sparse torus as large scale routing switch. In T. Alanko and Y.A. Bogoyavlenskiy, editors, *Proceedings of Finnish Data Processing Week 2003*, volume 5, pages 175–187, Petrozavodsk, Russia, 2003.

[9] S. LaRochelle, P.Y. Cortez, H. Fatfallah, L.A. Rusch, and H. Ben Jaafar. Writing and applications fiber bragg gratings arrays. In *SPIE proceedings*, 2000.

[10] B.E.A. Saleh and M.C. Teich. *Fundamentals of Photonics*. John Wiley & Sons, Inc., 1991.

[11] J. Sibeyn. Solving fundamental problems on sparse-meshes. In *Proceedings of SWAT'98, LNCS 1432*, pages 288–299, 1998.

[12] L.G. Valiant. *Handbook of Theoretical Computer Science*, volume A: Algorithms and Complexity, chapter General Purpose Parallel Architectures, pages 943–971. Elsevier Science Publishers BV, 1990.

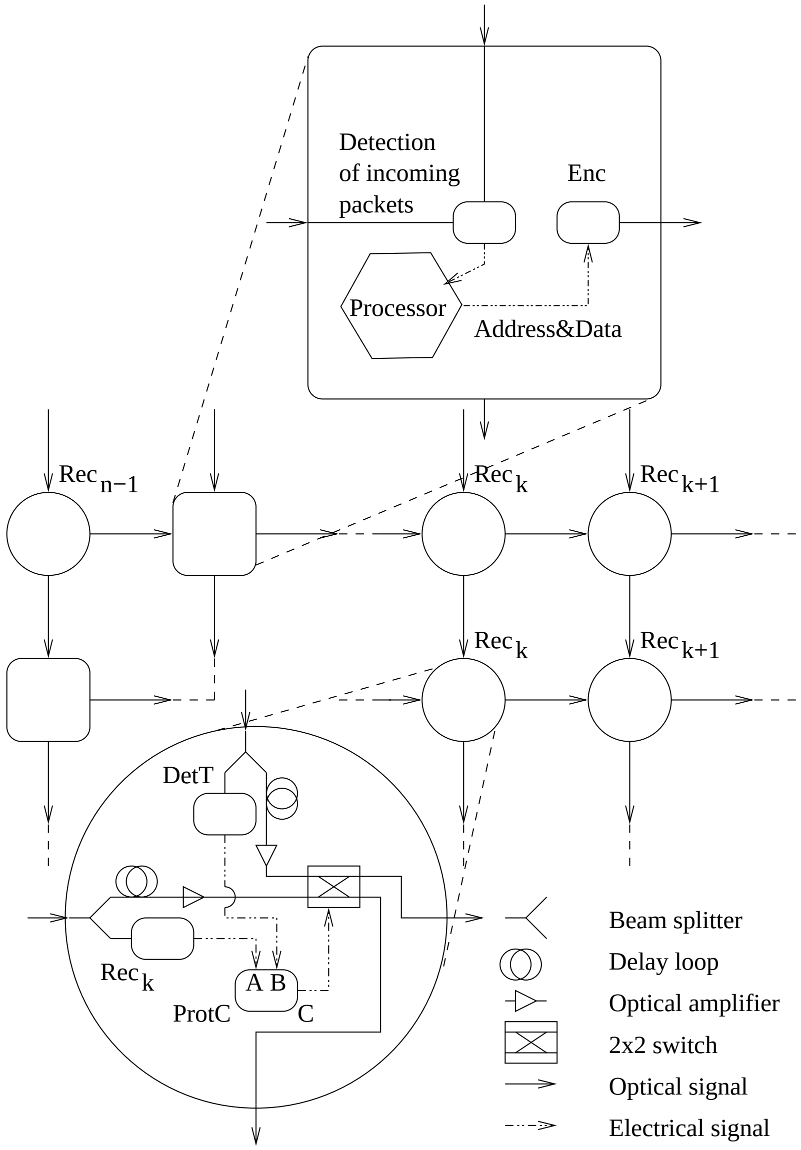


Figure 7: Block diagram presenting the routing machinery for $\mathcal{SOT}$.

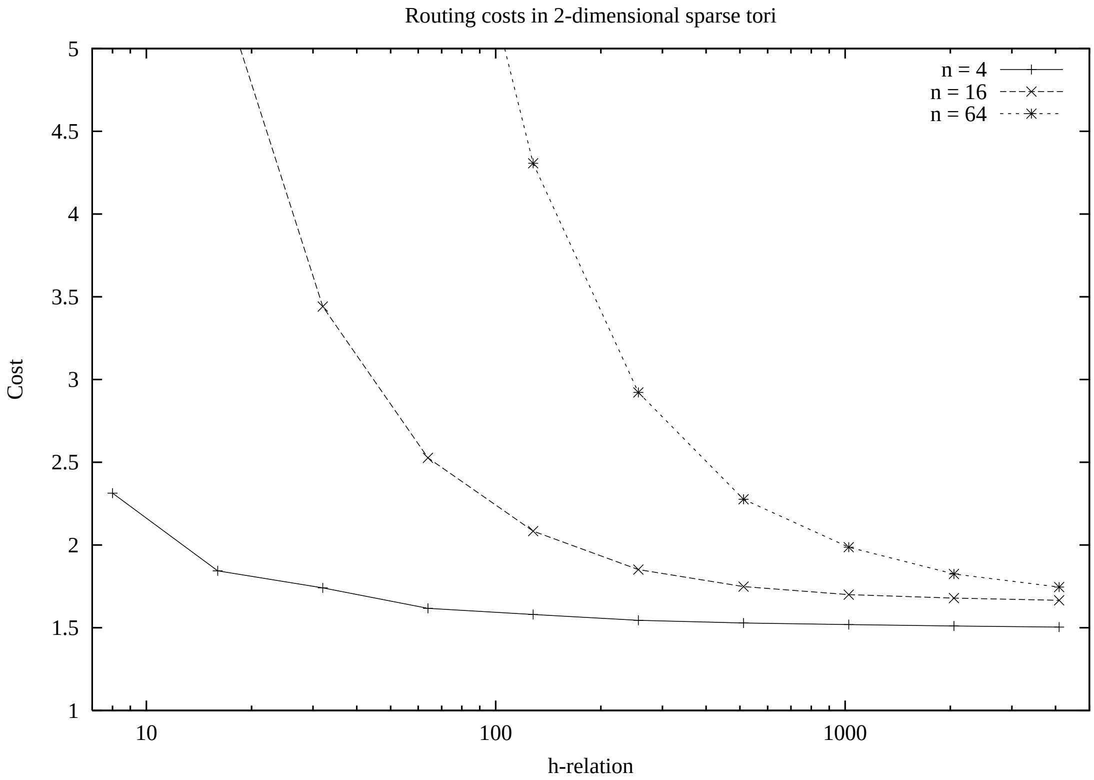


Figure 8: Routing costs when the size of h-relation varies.