

# MFLP: a low power encoding for on chip networks

Mehdi Taassori<sup>1</sup> · Meysam Taassori<sup>2</sup> · Sener Uysal<sup>1</sup>

Received: 19 March 2015 / Accepted: 22 December 2015 / Published online: 5 January 2016  
© Springer Science+Business Media New York 2016

**Abstract** Network on chip (NoC) has been proposed as an appropriate solution for today's on-chip communication challenges. Power dissipation has become a key factor in the NoCs because of their shrinking sizes. In this paper, we propose a new encoding approach aimed at power reduction by decreasing the number of switching activities on the buses. This approach assigns the symbols to data word in such a way that the more frequent words are sent by less power consumption. This algorithm dedicates the symbols with less ones to high probability data and uses transition signaling to transmit data. The proposed method, unlike the existing low power encoding, does not rely on spatial redundancy and keeps the width of the bus constant. Experimental evaluations show that our approach reduces the power dissipation up to 46 % with 2.70, 0.51, and 15.43 % power, critical path and area overhead in the NoCs, respectively.

**Keywords** Low power encoding · Network on Chip · Switching activity · Power consumption · Data compression

## 1 Introduction

The progress of VLSI technology allows researchers to design a complete system on a chip called system on chip (SoC). However, SoC has some drawbacks, such as lack of scalability and reusability. The network on chips (NoCs) have been proposed to alleviate today's communication problem of SoCs [1]. The NoCs are reusable and scalable and are able to tackle many disadvantages of SoCs [2].

---

✉ Mehdi Taassori  
mehdi.taassori@emu.edu.tr

<sup>1</sup> Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Famagusta, Mersin 10, Turkey

<sup>2</sup> Department of Computer Engineering, Sharif University of Technology, Tehran, Iran

The technological trend in portable and battery-powered devices introduces the power as a new aspect of VLSI design [3–6]. The increased power consumption causes a lot of problems such as decreasing the life time, and increasing the cost of packaging [7]. A great deal of research has been conducted to reduce the power consumption of interconnections in SoCs. Decreasing the swing voltage of power supply [8], using dual threshold voltage [9], voltage-frequency island (VFI) [10], activity postponement [11], Dynamic Voltage Scaling (DVS) [12], Dynamic Power Management (DPM) [13], statistical compression [14] and elimination of dispensable buffer slots [15] are some power reduction methods that have been presented in the literature.

One of the solutions to decrease the power consumption in chip interconnections is low power encoding [16]. This method tries to decrease the number of switching activities and consequently the dynamic power. On the other hand, the power consumption of coder and decoder are the overhead of this method which must be considered to evaluate its efficiency.

In this paper, we propose a novel low power encoding approach to decrease the number of switching activities through decreasing the number of ones included in code words and sending the code words with transition signaling. Apparently in transition signaling, the number of total switching activities is equal to the number of ones in the code words [17]. This paper introduces a new algorithm to assign code words to symbols in such a way that the more frequent symbols to be sent consume less power. To approach this goal, the proposed most frequent least power (MFLP) encoding uses a tree based infrastructure. The tree structure provides a set of symbols which assigns the less ones words to high probability data and vice versa. Based on the proposed algorithm the most frequent symbols are allocated to the least number of ones which results in the least power consumption.

Most of the low power encoding algorithms increase the width of the transmission bus to send the data [17–20], whereas the proposed method does not rely on spatial redundancy. It is also worth mentioning that even though in most of the traditional low power encoding algorithms the effect of coupling capacitors is ignored, our results show that these capacitors have an increasing contribution in power consumption in the NoCs as the VLSI technology advances and the size of the transistor shrinks. In this paper, all evaluation results consider capacitors, coupling and self, to calculate the power consumption of links. The experimental results show that by applying the proposed approach, power dissipation up to 46 % is improved and with, on an average, 14.4 % area overhead.

## 2 Literature review

Several methods have been proposed in the previous works to reduce the power consumption by encoding techniques. These include the algorithms that have been designed for data line [18], irredundant encoding [21], correlated data, like address buses [22,23], parallel and serial which are used for the parallel and serial buses, respectively [24], redundant encoding method that raise either the number of transmission bus or clock pulses to send data [25], and adaptability [26–28].

One of the most well-known low power encoding is the Bus Invert coding [18]. This coding is appropriate for the uniform distribution data and the parallel bus which have spatial redundancy. Another scheme which tries to decrease the number of transitions is limited weight coding (LWC) [17]. In this algorithm,  $W$  is defined as a weight of each code word; that is,  $W$  is equal to the number of ones included in the code words. LWC applies transition signaling after assigning the code words and can be exploited in both the parallel that have

spatial redundancy and serial buses with time redundancy. Beach coding [20] is suggested when the correlation of data pattern is computable. In this approach, the method of encoding is selected based on the pattern of data; therefore, it is strongly application dependent.

Since the power of links in NoCs is an important portion of power consumption, low power encoding also is applicable for this infrastructure. Researchers in [21] present an irredundant encoding and in [16, 29] a set of data encoding methods are proposed to decrease the link power consumption in the NoCs. In [30], a reliable data communication method to decrease the energy dissipation in NoCs is introduced. Authors in [16, 21, 29, 30] considered the effect of both self and coupling capacitances in link power dissipation. This point is worth mentioning that redundant encoding algorithm could not decrease the power consumption in NoCs because this redundancy may cause redundancy in each router which is not compensated by power reduction in links. Moreover, due to the fact that in the advanced technology the links are too close to each other the low power encoding used in the NoCs should consider the transition on coupling capacitance as well.

### 3 Proposed method

The main idea of the proposed method is to reduce the number of ones in code words. In fact, due to the transition signaling, the number of total switching activities is equal to the number of ones in code words [17]. The proposed method is a tree-based algorithm. This tree encompasses root, a number of nodes and leaves. In this tree, code words are represented according to the location of the nodes referring to the data words.

#### 3.1 MFLP encoding approach

Our approach uses the tree-based structure to assign the code word with less ones to most frequent words; hence, we called it Most Frequent Least Power (MFLP) consumption coding. The objective is to minimize expectation of '1' and in turn, decrease the switching activities as well as the power consumption by using transition signaling. The tree-based structure also allows us to assign the shortest code words to more frequent symbols. Hence, this coding algorithm not only decreases the power consumption but also compresses the amount of transmitted data. In this algorithm, required statistical knowledge about frequencies of symbols are collected in previous time sliding windows; in other word, while data is passing, the frequency of data can be counted and this knowledge can be used to encode the data for next time sliding windows; evidently, the current data is being coded based on the statistical information gathered in previous time sliding window. In the other word, the encoder does not need to have any priori knowledge of data because MFLP are collecting this information while data is passing and because the sliding windows is small enough, the characteristic of data is likely be same in consecutive period of data. At first, we need to choose a parameter called division factor. According to this factor, we divide the words into two parts. This factor indicates that we are going either to decrease the power or compress data. The tree is made of nodes, where each node has a label indicating the sum of labels of its children. In the case of leaf, this label refers to the frequency of words represented by this node. This tree structure can be created reversely; after dividing the words of data in two portions according to division factor, we assign the sum of these nodes as a label of the root. The root's label represents the sum of label of its children. We continue the procedure until the leaf of the tree which refers to each word of the data. This function is implemented in hardware and inserted in coder and decoder. Figure 1 presents pseudocode of the proposed algorithm.

```

Given sorted frequencies of symbols as
{ $A_i \mid 1 \leq i \leq n \mid \forall i, j, A_i < A_j \rightarrow f_i < f_j$  // symbol  $A_i$  has frequency  $f_i$ 
; chosen division factor =  $\gamma$ 
function MFLP-tree ( $S = \{(A_j, f_j), \dots, (A_k, f_k)\}$ ) //  $j$  is first index and  $k$  is last one
 $T_1 \leftarrow \sum_{i=1}^n f_i$  // root of tree whose label is sum of all frequencies in  $S$ 
if ( $j = i$ )
  insert a node labeled  $T_1$ 
else
  {
  divide  $S$  into two subsets,  $S_1 = \{(A_j, f_j), \dots, (A_{\gamma k}, f_{\gamma k})\}$ ,  $S_2 = \{(A_{\gamma k+1}, f_{\gamma k+1}), \dots, (A_k, f_k)\}$  // two sub
  sets are generated to create children of root
  MFLP-tree ( $S_1$ ); // function is called recursively for children named  $S_1$  and  $S_2$ 
  MFLP-tree ( $S_2$ )}
end

```

**Fig. 1** Pseudocode of the proposed algorithm

In this algorithm,  $A_i$  is the word of the data whose frequency is  $f_i$  and  $S$  is a set of data words.  $T_i$  is the MFLP tree node labeled by the sum of its children's frequency.

With reference to Fig. 1, the tree construction can be further explained with the following steps:

- We sort the frequencies of symbols in descending order from higher frequencies to lower ones.
- We choose division factor ( $\gamma$ ) according to the goal, either to decrease the power consumption or to compress the amount of data.
- MFLP function constructs the tree reversely. We have to provide the frequency of data words as input of this function. It divides the data words based on  $\gamma$  in two portions as upper and lower groups. Sum of the upper and lower group frequencies is allocated to the left and right nodes, respectively. After that, it invokes itself reversely to construct interior nodes. This algorithm continues till the leaf nodes are generated.
- The labels “0” and “1” are assigned to the edge of upper and lower group, respectively.
- To figure out the code words we follow the labels of the edges. The code word is the sequence of the edge labels from root to the frequencies of the symbol.

The procedure of encoding in MFLP is composed of two steps: Counting and Coding. While data is passing from encoder the frequency of transmitted symbols can be counted in a time sliding window; this knowledge let encoder generate the tree structure and assign new code words to the symbols. These new codes are going to be used in the next sliding window. It is clear that meanwhile data is coding based on knowledge of previous window (coding), the frequencies of symbols in current window can be counted to be used in the next window (counting). It is obvious that these two steps can take place at the same time. According to the proposed algorithm, data stream should be divided into the sections with same time period namely sliding window. The frequency of data is counted in current window and will be used in the next sliding window to provide the final code words. Due to temporal locality the frequencies generated in the previous window can be used in the current window. The same procedure is applied to the decoder to figure out the frequency of received data before decoding.

In the following example, we clarify the steps of the algorithm.

- First step: the symbols should be arranged according to their frequency of occurrence in descending order. For instance, there are 13 symbols which are given to be coded. At first we organize them in alphabetical order: A,B,C,D,E,F,G,H,N,P,Q,R,S.
- Second step: This step depends on the division factor. This value should be multiplied by the number of symbols. The selection of the symbols is based on the result of the last multiplication. Top symbols should be located on the left and the others on the right. This strategy is shown in Fig. 2.

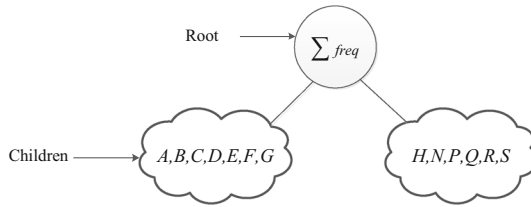


Fig. 2 Root and its children

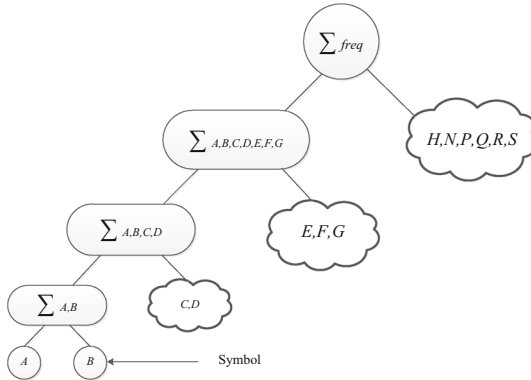


Fig. 3 Generation of symbols

It is required to repeat the second step for the symbols which are included in the left hand side. Figure 3 shows the steps to reach to the symbols.

This trend must be continued for each node either in the left hand side or in the right hand side till we get to one symbol in every set. The expected value of one is evaluated by Eq. 1.

$$E(x) = \sum_{i=0}^{\text{symbol}} F_i * N_i \tag{1}$$

where  $F_i$  is the frequency of the symbols and  $N_i$  is the number of ones for each symbol in the tree.

The tree structure assigns a code word with less ones to the more frequent data words. According to Eq. 1, the expectation of ones can be minimized by this strategy.

When time duration remains constant decreasing the power consumption can lead to decrease the energy dissipation. In the case of compression, the energy can be reduced due to decrease in the duration of time provided that either switching activity does not rise or its increment can be compensated by time reduction. Hence, there is a trade-off between the number of switching activities and compression ratio which depends on the division factor. The effect of division factor on compression and power consumption can be evaluated on these bases:

- 1- As the division factor is increased, we assign the symbols with less ones to more frequent data words resulting in less switching activities thereby reducing the power consumption.
- 2- By reducing the division factor, we can improve the compression ratio. Tree structure allocates less length symbols to more frequent data words at the expense of increasing the number of ones and consequently power dissipation.

To examine how the proposed method reduces the number of switching activities and power consumption, we evaluate the bit average by Eq. 2.

$$B_{avg} = \sum_{i=0}^{\text{symbol}} F_i * L_i \tag{2}$$

where  $F_i$  is the frequency of symbol whose length is  $L_i$ .

### 3.2 Optimality of MFLP

In this subsection, we present the mathematical proof to show that MFLP, which aims to reduce the power consumption by decreasing the number of ones in the code word, is able to reduce the expectation of ones. Therefore, the MFLP code is optimal if the expected value of ones is minimal. The frequency of symbols are ordered, so that  $F_1 \geq F_2 \geq \dots \geq F_i$ . To prove that the  $E(x)$  in MFLP code is minimal, we show that with any changes in MFLP's tree and code word the value of expected value is increased. We consider that  $C_w$  is an optimal code word which is the result of MFLP encoding. If  $F_j \geq F_k$  then  $N_k \geq N_j$ . We then swap MFLP code words. Supposing that  $C'_w$  is the code words  $j$  and  $k$  of  $C_w$  interchanged, the expected value of  $C'_w$  is shown in Eq. 3.

$$E(C'_w) = \sum_{i=0}^{\text{symbol}} F_i * N'_i \tag{3}$$

$N'_i$  is the number of ones for symbol after interchanging  $j$ th and  $k$ th code words.

$$\begin{aligned} E(C'_w) &= \sum_{i=0}^{\text{symbol}} F_i * N'_i = F_j * N_k + F_k * N_j \\ E(C'_w) - E(C_w) &= \sum_{i=0}^{\text{symbol}} F_i * N'_i - \sum_{i=0}^{\text{symbol}} F_i * N_i \\ &= F_j * N_k + F_k * N_j - (F_j * N_j + F_k * N_k) \\ &= (F_j - F_k) (N_k - N_j) \end{aligned}$$

Based on MFLP, if  $F_j \geq F_k$  then  $N_k \geq N_j$ , which means that  $E(C'_w) - E(C_w)$  should be greater than zero ( $E(C'_w) \geq E(C_w)$ ). It can be concluded that after changing the code word of MFLP, the value of expected value is increased. Hence, the minimum amount of expected value, the minimum number of ones, is related to MFLP code words and  $C_w$  is optimal.

### 4 Effective criteria in the efficiency of the proposed method

By adding coding algorithm to the system, the power consumption of coder and decoder are considered as overhead and is needed to be compensated. The power consumptions of transmission line without (5) and with (6) using encoding algorithm are calculated by:

$$P_{link} = P_{self} + P_{coupling} \tag{4}$$

$$P_{link} = \alpha_s C_{self} V_{dd}^2 f + \alpha_c C_{coupling} V_{dd}^2 f \tag{5}$$

$$P_{after} = P_{cod.} + P_{dec.} + \alpha_{as} C_{self} V_{dd}^2 f + \alpha_{ac} C_{coupling} V_{dd}^2 f \tag{6}$$

$$C_{link} = C_{self} + C_{coupling} \tag{7}$$

$P_{link}$  is power dissipation before using encoding algorithm and  $P_{after}$  is the power after inserting MFLP.  $P_{link}$  is composed of power of self capacitance ( $P_{self}$ ) and coupling capacitance ( $P_{coupling}$ ). As shown in (5),  $\alpha_s$  and  $\alpha_c$  are switching activity of the self and coupling capacitance, respectively.

$P_{after}$  is power consumption after using encoding approach.  $P_{cod.}$  and  $P_{dec.}$  are the power dissipation of coder and decoder, respectively,  $\alpha_{as}$  and  $\alpha_{ac}$  are switching activity on self and coupling capacitances after applying data coding approach.

$C_{link}$  is the total capacitance which is the summation of the self ( $C_{self}$ ) and coupling ( $C_{coupling}$ ) capacitance,  $f$  is the clock frequency and  $V_{dd}$  is the power supply of the system.

$\alpha_s$  and  $\alpha_{as}$  which are the self-switching activity before and after using encoding method are evaluated based on the number of transition (high to low and vice versa) on the link. The coupling switching activity before and after using MFLP ( $\alpha_c$  and  $\alpha_{ac}$ ) are calculated according to the direction of switching activities happening on the consecutive wires which is shown in Table 1.

The evaluation of self and coupling capacitance is based on the type of the switching activity. In Table 1 the number of self and coupling transition for different type of switching activities are depicted.

The coding algorithm can decrease the power consumption, provided that  $P_{after}$  is less than the power consumed before applying MFLP. The more the number of switching activities decreased, the more effective our method is. Efficiency factor ( $\beta$ ) is introduced in order to evaluate MFLP. Let us suppose that

$$P_{after} = P_{codec} + \alpha_{as} C_{self} V_{dd}^2 f + \alpha_{ac} C_{coupling} V_{dd}^2 f \quad (8)$$

where  $P_{codec}$  is sum of the power consumption of coder and decoder. As a result, the efficiency factor can be expressed as

$$\beta = \frac{(\alpha_s - \alpha_{as}) C_{self} V_{dd}^2 f + (\alpha_c - \alpha_{ac}) C_{coupling} V_{dd}^2 f}{P_{codec}} \quad (9)$$

MFLP can reduce the power dissipation if the value of efficiency factor ( $\beta$ ) is more than one.

Assessment of some of the parameters' effectiveness of our approach is presented below:

**Distance:** One of the most important criterion that affects the efficiency factor is the distance between the transmitter and receiver nodes. Distance has an important role on the amount of capacitance of the link and consequently on the power consumption of the NoC when the switching activity occurs. In other words, by increasing the distance between the transmitter and receiver, the value of the capacitance of links increases. This shows that reduction of the number of transitions on the link plays a more effective role in the improvement of power consumption of the NoC. It is evident that according to Eq. 8, the value of the efficiency factor ( $\beta$ ) increases due to the increased value of  $C$ . Thus, our approach is more effective in longer distances.

**Family:** With the growth of advanced VLSI technology, the transistors shrink and the length of the wire remains constant or even increases. Eventually, the capacitance of the wire gets more dominant. Therefore, based on Eq. 8, the efficiency factor increases and consequently MFLP becomes much more effective.

**Table 1** Number of self and coupling capacitances for different type of switching activities

Type	Number of Self Transition	Number of Coupling Transition
	1	1
	0	0
	2	0
	2	4

### 5 Evaluation

The power of the NoC is consumed in two parts, the routers and the links. It should be mentioned that the power of Network Interface (NI) is included in the power of router. In our experiment, the baseline network contains 16 nodes which are connected in a mesh topology whose router algorithm is XY; each router has two virtual channels. Packet length is 32 flits. We use power compiler tool from Synopsys<sup>1</sup> to calculate the power of the routers. Power compiler considers the static and dynamic power consumptions. The number of transitions is the major factor indicating dynamic power consumption in data transmission. Despite the fact that the growth of VLSI technology and shrinking the transistor size make the static power dominant part of the power consumption, the research has shown that in the NoC infrastructure

<sup>1</sup> Synopsys and Modelsim are registered trademarks.



the dynamic power still remains the prevalent portion of the power consumption due to its architecture [31–33].

The power of the links is determined by Eq. 4. We used 65nm technology for the simulations of the proposed method. According to the International Technology Roadmap for Semiconductors [34], for this technology  $V_{dd}$  is defined as 1 Volt and the clock frequency is set to 500 MHz based on the critical path of the system. The length of the metal wires is selected as 2 mm for the mesh topology. The self capacitance of the wire links and coupling capacitance are selected as 0.2 and 0.6 pF/mm, respectively. The transitions of wires are calculated by Modelsim<sup>1</sup>.

In this section the coder and decoder are implemented in the hardware layer and they are inserted in the local link, between the routers and process elements. In another words, this service is delivered in the transport layer of the NoC which is offered in transmitter and receiver. Hence, the data encoding is done end to end. The coding methods and the NoC infrastructure are implemented in VHDL.

## 5.1 Evaluation of the proposed algorithm

It does not matter which infrastructure the designers have chosen, either the traditional bus or the novel NoCs, this coding can be useful for all. To show the effectiveness of our algorithm, we examine its effect in decreasing the power consumption or the amount of data by using some real-life streams. We assess MFLP in the following cases: using buses as a traditional infrastructure and the NoC as a new one.

### 5.1.1 On the bus

To evaluate our approach we consider a system including a transmitter, a communication bus and a receiver. The power can be calculated in two cases: original data and coded version. The power of the link consists of power consumed in the coupling and self capacitances.

On the serial bus, length of the metal wires is assumed as 2 mm and the self capacitance of the wire links is selected as 0.2 pF/mm [34]. It is worth mentioning that on the serial bus we do not have any significant coupling capacitance. The designer needs to decide whether power reduction or decreasing the amount of data is the final goal. According to this decision we need to change the division factor. The more we increase the division factor, the more the bit average goes up. That is, we have gained more power reduction in expense of increasing the amount of data. We evaluate our approach in various division factors for the serial bus using MFLP encoding and the results are shown in Figs. 4 and 5.

In the serial system, the energy is calculated by multiplying the power consumption and time duration. It is apparent that the time duration can be estimated by:

$$T = B_{\text{avg}} * S * Clk \quad (10)$$

Where  $T$  is time duration,  $B_{\text{avg}}$  is bit average,  $S$  indicates the number of transmitted symbols, and  $Clk$  is the period of clock in the transmission system. Consequently, the bit average is able to represent the time duration because other parameters are constant with different division factors.

The energy dissipation before applying encoding algorithm and after using MFLP are evaluated based on the following formula:

$$E_{B.C.} = E_{Router} + E_{Link} \tag{11}$$

$$E_{A.C.} = E_{Router} + E_{Codec} + E_{CLink} \tag{12}$$

where  $E_{B.C.}$  is energy consumption before using coding method,  $E_{Router}$  is energy dissipation of router and NI and  $E_{Link}$  is energy which is consumed in the physical links while  $E_{A.C.}$  is energy that is consumed after coding which contains  $E_{Router}$ , energy consumed in routers,  $E_{Codec}$ , energy dissipation in coder and decoder, and  $E_{CLink}$  which is consumed in links after using coding algorithm.

The evaluation of different encoding algorithms on various media formats such as text, PDF, color image and so on is reported in [16]. We also assess MFLP and other data encoding approaches on the following data streams which belong to the several media formats like the previous works, namely: TXT: the text file in the .txt format, GIF, JPEG, BMP and PNG: the image files in .gif, .jpg, .bmp and .png format, respectively, WAV: the sound files in the .wav format, HTML: the MHTML Document file in the .mht format, PDF: a PDF format file, and DOCX: Microsoft Word Document in .docx format.

Figure 4 is normalized to the energy consumption of without coding for each benchmark. According to the results given in Figs. 4 and 5, it can be deduced that by increasing the percentage of division factor, the energy dissipation decreases, but in the extreme points of 10 and 90 %, the value of the bit average is high and is not appropriate. By increasing the percentage from 10 to 20 % and up to 50 %, the bit average decreases. This trend can be seen from 90 to 50 % as well. Hence, the optimum point for an appropriate energy consumption and bit average is 50 %, but there is flexibility in encoding to reach a tradeoff between energy dissipation and bit average. In introduced encoding algorithm, according to tree based structure used in this coding, to decrease the number of “1” in each symbol, necessarily the length of each symbol increases and as a result, the bit average goes up. In another word, in this encoding although the total number of “1” declines which means the energy consumption of link gets improved, the total length of data increases, meaning the efficiency of compressor decreases and the bit average goes down as well.

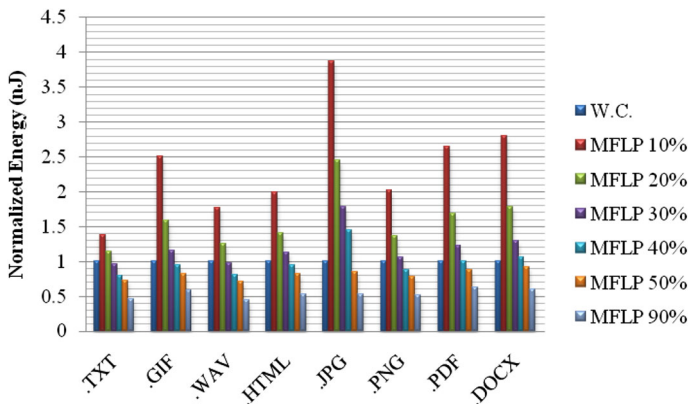


Fig. 4 Comparison of energy consumption with various division factors and without coding (W.C.) on the serial bus

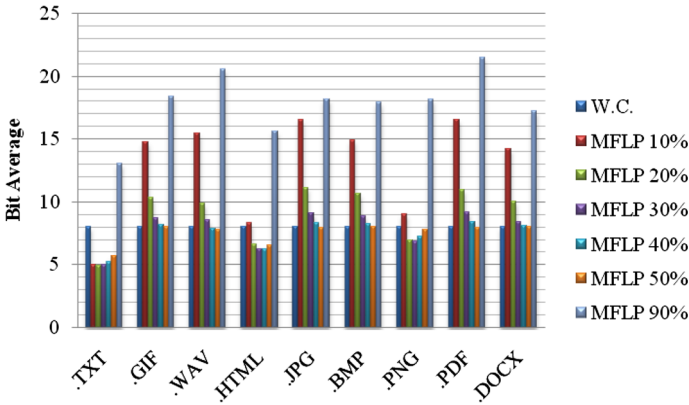


Fig. 5 Bit average with various division factors on the serial bus

Using this conclusion, we use a division factor of 50 % for the rest of the implementation. In the transition signaling approach the number of ones included in the code word is the same as the number of the transition activity [17]. Therefore, decreasing the number of ones induces the switching activity reduction. In our assessment, the switching activity reduction ratio can be defined as follows:

$$S.A. = \frac{S.N_{wc} - S.N_{MFLP}}{S.N_{wc}} * 100 \tag{13}$$

where  $S.N_{wc}$  is the number of switching activity without coding algorithm and  $S.N_{MFLP}$  is the number of switching activity with applying MFLP approach. The link power dissipation is evaluated based on the number of switching activity.

We have examined our approach on the parallel bus as well. In this case, we assume an eight bit bus between the transmitter and receiver whose length is 2mm; similarly, the self and coupling capacitances are considered as 2pF/mm and 6pF/mm, respectively [34]. The power of the system before coding is represented by the power of the link where the original data is passing; while the power of the encoder and decoder plus the power of links where the coded data is passing can be considered as the total power of the system after coding. The results shown in Fig. 6 illustrate that the power of the link after coding decreases so that it can compensate the power of overhead of coding. As depicted in Fig. 6, link power dissipation can be saved up to 35 %. It is obvious that the power consumption of MFLP coder and decoder is the overhead of our design.

### 5.1.2 In the network on chip (NoC)

The power consumption is one of the most important factors in the NoCs. Therefore, we assess the proposed method in this infrastructure to decrease the power consumed. The impact of MFLP is assessed on the parallel bus of the NoC. The simulation is carried out based on the specific characteristics which are explained in detail on the experimental results section. Nowadays, the link power dissipation of the NoCs is a significant portion of the total power consumption [29]. In Table 2 comparison of power consumption between proposed method and baseline is presented. In the second column of Table 2, the link power dissipation in both cases, baseline and MFLP is shown. The router’s power consumption before and after using MFLP are demonstrated in third column. As mentioned, the power of coder and decoder as

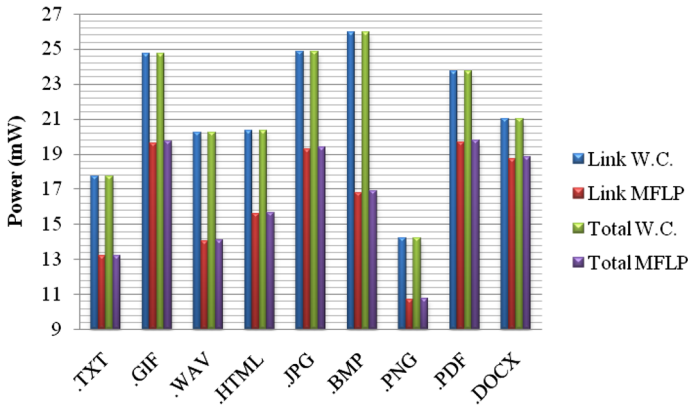


Fig. 6 Link and total power consumption on the parallel bus

Table 2 Comparison of power consumption between MFLP and without coding

Power	Link (mW)		Router (mW)		Coder & Decoder (mW)	Total (mW)	
	W.C	MFLP	W.C	MFLP		W.C	MFLP
File name					MFLP		
.TXT	27.15	19.51	53.42	53.45	0.03	80.57	72.96
.GIF	31.08	20.88	53.85	54.44	0.59	84.93	75.32
.WAV	30.15	16.22	53.87	53.88	0.01	84.02	70.1
.HTML	28.25	22.78	53.51	54.09	0.58	81.76	76.87
.JPG	33.00	20.37	54.12	54.53	0.41	87.12	74.90
.BMP	15.86	12.53	52.51	52.56	0.05	68.37	65.09
.PNG	22.19	15.29	52.12	53.53	1.41	74.31	68.82
.PDF	31.83	20.35	54.06	54.58	0.52	85.89	74.93
.DOCX	28.40	18.00	53.57	54.08	0.51	81.97	72.08

overhead of the proposed approach is presented in forth column. In the last columns the total power consumption for baseline and MFLP are depicted. As shown in Table 2, by applying MFLP, the link power dissipation can be decreased up to 46 %.

### 5.1.3 Experimental results

In this subsection, we study the effectiveness of the proposed algorithm. Figure 7, gives a comparison between the MFLP and the previous state of the art coding approaches. Regarding Fig. 7, it is obvious that LWC [17], BI [18], CDBI [19] and CABI [21] are not able to decrease the power consumption due to the one additional bit which is in the coded data. That is, these algorithms have spatial redundancy to encode data and this redundancy leads to increase the power consumption. Although, both LWC and BI cannot decrease the power consumption, the latter is better because of the simplicity of coder and decoder. That is, the overhead of BI is lower than LWC. The Beach coding [20], one of the well-known adaptive coding approaches, is application dependent and can be an appropriate solution for the application specific systems. In this case, the type of coding can be changed dynamically according to the relationship between the current data and the previous one.

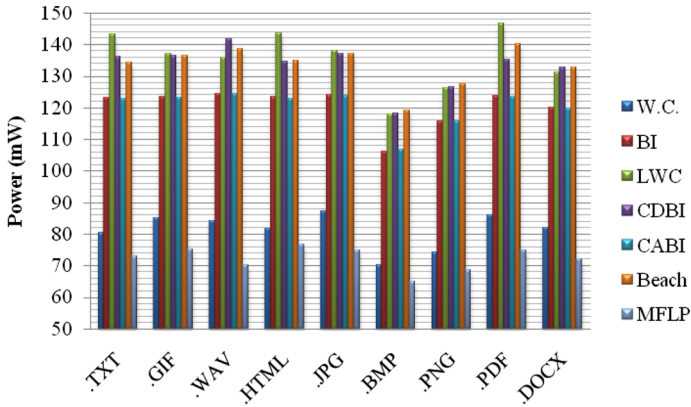


Fig. 7 Comparison of MFLP with state of the art encoding algorithms

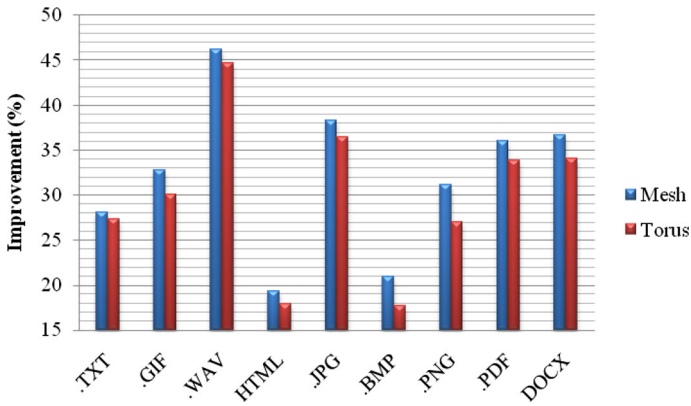


Fig. 8 Percentage of power saving compared to without coding with different topologies

### 5.2 Evaluation of sensitivity to network parameters

We assess the impact of the network parameters such as topology, routing algorithm, number of nodes, packet length and the number of virtual channels on effectiveness of our method. In this assessment, the default routing algorithm and topology is XY and mesh, respectively.

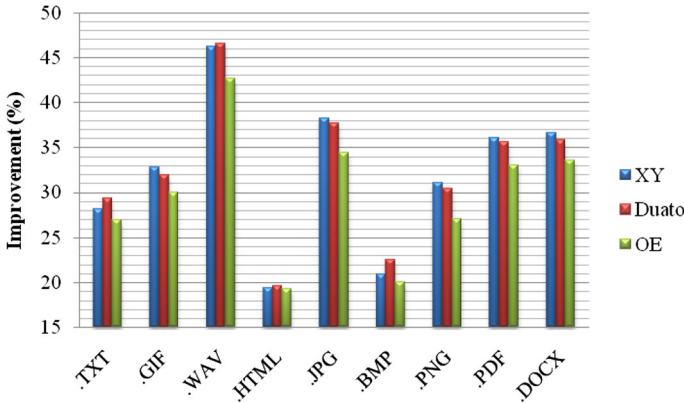
#### 5.2.1 Topology

In this subsection, we investigate the effect of different topologies on the efficiency of our method. Two of the most prevalent topologies, mesh and torus are suitable to be implemented on the NoC with 16 nodes (4 × 4) due to their two dimensional structure. Figure 8 shows the percentage of power saving acquired with the mesh and torus topology for several benchmarks as compared to the scheme with no data encoding in the NoC.

In Table 3, link, router, coder & decoder and total power consumption for Mesh and Torus topologies before applying encoding approach (W.C.) and after using MFLP algorithm are presented separately.

**Table 3** Comparison of power consumption between Mesh and Torus

Power (mW)	Mesh		Torus	
	W.C.	MFLP	W.C.	MFLP
Link	22.19	15.29	21.54	15.71
Router	52.12	53.53	53.42	53.65
Coder & Decoder	0	1.41	0	0.22
Total	74.31	68.82	74.96	69.36



**Fig. 9** Percentage of power saving compared to without coding with different routing algorithms

As shown, the mesh topology is more suitable in the case of link’s power consumption compared to the torus topology. In other words, the impact of our approach in the mesh topology is better. The reason is due to the extra link on each node in torus topology. The extra link loses the consecutiveness in the data. Hence, the power of link is increased. In terms of the total power dissipation the effect of both topologies is approximately the same.

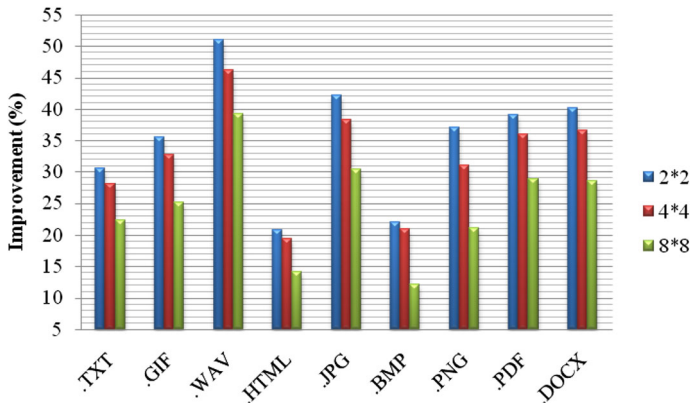
### 5.2.2 Routing algorithm

Routing algorithms can be classified into deterministic, partially adaptive and fully adaptive categories. We examine various routing algorithms, namely, XY, OE and Duato to analyze the efficacy of MFLP in power reduction. XY is a deterministic routing algorithm, OE is known as partially adaptive and Duato is fully adaptive routing algorithm. Figure 9 shows the percentage of power reduction with different routing algorithms as compared to the scheme that no data encoding algorithm is used.

To implement the Duato algorithm, we need two virtual channels to prevent deadlock. Thus, we assign two virtual channels for the other algorithms to have a fair compression. According to the results, it can be concluded, from the switching activity point of view, the Duato algorithm in average is the best and OE is the worst one. Similarly, from the link power dissipation perspective, Duato and XY can outweigh OE. With the increased consumption of the network power on the link, our approach is shown to be substantially better. On the other hand, when a routing algorithm is distributed uniformly, the power consumption of the link goes up because the more traffic is distributed smoothly, the more performance we have and, in turn, the more power is consumed. However, the results also show that OE, as a partially adaptive algorithm, cannot distribute traffic more smoothly than XY as a

**Table 4** Comparison of power consumption between XY, Duato and OE

Power (mW)	XY		Duato		OE	
	W.C.	MFLP	W.C.	MFLP	W.C.	MFLP
Link	22.19	15.29	23.73	16.50	15.09	11.01
Router	52.12	53.53	45.37	52.21	49.44	50.03
Coder & Decoder	0	1.41	0	6.84	0	0.59
Total	74.31	68.82	69.10	68.71	64.53	61.04

**Fig. 10** Percentage of power saving compared to without coding with different number of nodes

deterministic algorithm. Therefore, for the efficiency of MFLP, Duato and XY outperform the OE algorithm.

Eventually, it can be concluded that using fully adaptive routing algorithm can pass packets more smoothly which leads link power increase. The more power links consume, the more effective of coding algorithm is.

In Table 4, comparison of link, router and coder & decoder power dissipation between different routing algorithms such as XY, Duato and OE without encoding algorithm (W.C.) and after using the proposed method (MFLP) are shown.

### 5.2.3 Number of nodes

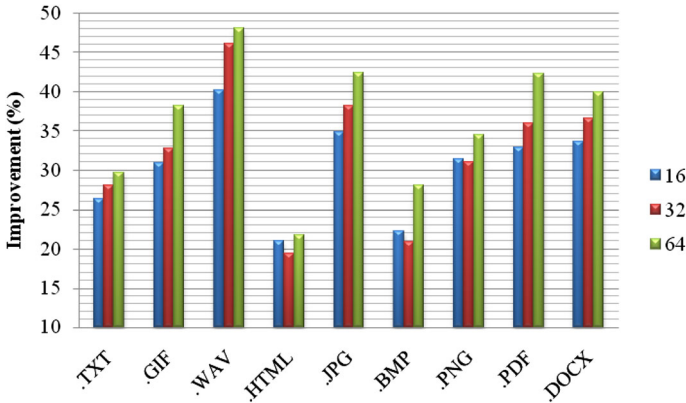
We study our method with different number of nodes. The NoCs are considered with 4, 16 and 64 nodes. Figure 10 depicts the improvement in power reduction with various numbers of nodes, compared to the case that MFLP is not used.

In Table 5, the components of the power consumption with different number of nodes in baseline and MFLP are depicted.

In this case, one criterion is effective; the consecutiveness of the data. It is evident that when the distance between the transmitter and receiver increases the chance of interference among the flits of packet goes up; therefore, the effectiveness of our approach decreases. Based on this remark, with increasing number of nodes in the NoC, the consecutiveness of the data collapses as well as the effectiveness of our approach is diminished and consequently, power dissipation increased.

**Table 5** Comparison of power consumption between 2\*2, 4\*4 and 8\*8

Power (mW)	2*2		4*4		8*8	
	W.C.	MFLP	W.C.	MFLP	W.C.	MFLP
Link	3.49	2.19	22.19	15.29	101.84	80.32
Router	13.40	14.60	52.12	53.53	216.90	237.24
Coder & Decoder	0	1.20	0	1.41	0	20.34
Total	16.90	16.80	74.31	68.82	318.74	317.56



**Fig. 11** Percentage of power saving compared to without coding with different packet lengths

**Table 6** Comparison of power consumption between different sizes of packet length

Power (mW)	16		32		64	
	W.C.	MFLP	W.C.	MFLP	W.C.	MFLP
Link	23.80	16.31	22.19	15.29	22.64	14.81
Router	53.35	54.42	52.12	53.53	51.44	52.83
Coder & Decoder	0	1.06	0	1.41	0	1.38
Total	77.15	70.73	74.31	68.82	74.08	67.64

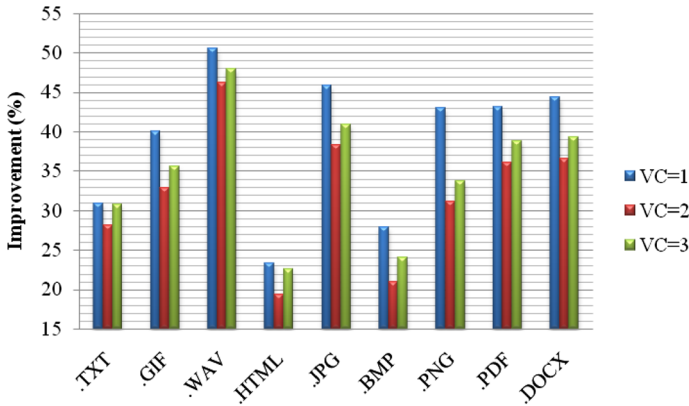
5.2.4 Size of the packet length

The proposed method is tested with different size of packet length. The topology and routing in NoC are mesh and XY, respectively. The number of nodes is 16. In Fig. 11 the amount of power reduction by applying MFLP for different number of nodes is shown.

Table 6 depicts the power dissipation of link, router and overhead of introduced algorithm with various size of packet length.

By comparing the above results, it is worth mentioning that by increasing the packet length in the NoC, the effect of MFLP goes up. We have implemented our approach in the transport layer. It means that only the data part of the flits, not header and footer, is coded only in the transmitter and receiver node. Whenever we change the size of the packet, we change the





**Fig. 12** Percentage of power saving compared to without coding with different number of virtual channels

**Table 7** Comparison of power consumption between different number of virtual channels

Power (mW)	1		2		3	
	W.C.	MFLP	W.C.	MFLP	W.C.	MFLP
Link	12.74	7.26	22.19	15.29	23.30	15.45
Router	22.29	25.29	52.12	53.53	90.25	93.30
Coder & Decoder	0	3.00	0	1.41	0	3.04
Total	35.03	32.56	74.31	68.82	113.56	108.75

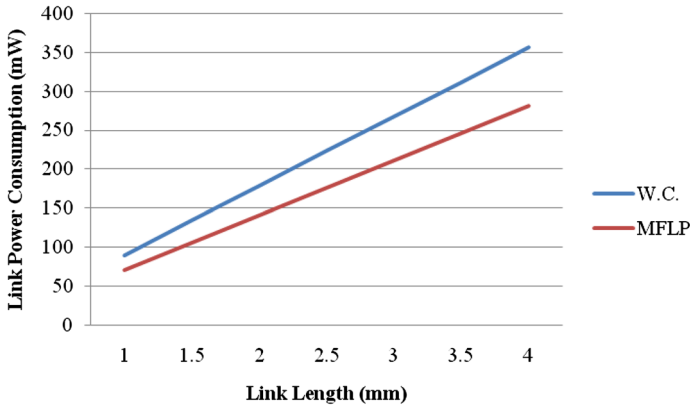
number of data. In contrast, the number of header and footer remains constant. Hence, by increasing the packet size the data increase and more data are coded. On the other hand, by decreasing the packet size, only the data section goes down and the other parts remain the same as before. In this case, the numbers of the data that are coded are less. Thus, the effect of our contribution is not much as before and the impact of our proposed method decreases.

### 5.2.5 Number of virtual channels

The number of virtual channels is effective on the throughput of the interconnection network. The significant portion of the power consumption in routers is consumed in the virtual channels. In this section we study the effect of our proposed method with different number of virtual channels. Our approach is implemented in the mesh based with XY routing algorithm. The network has 16 nodes and the packet length is 32. The result shown in Fig. 12, which is the comparison of MFLP and no coding approach, is obtained by changing the number of virtual channels.

Table 7 shows the effect of different number of virtual channels on the power consumption of link, router and coder & decoder with using the proposed algorithm.

The impact of virtual channels on the effectiveness of coding depends on two criteria. Firstly, how much order of flits in the network will remain constant while passing through network, secondly, utilization of the bus. As shown above, by increasing the number of virtual channels, sequence of data would be more subject to change and, in turn, the impact of our coding decreases. On the other hand, the growth of number of virtual channels leads to



**Fig. 13** The impact of link length on efficiency of MFLP

have less congested links and consequently, the utilization of bus goes up. The results show that power consumption of links increases and consequently, the influence of the proposed method rises.

### 5.2.6 Link length

In this subsection the impact of link length on efficiency of proposed method is studied. As depicted in Fig. 13, the link power consumption of MFLP encoding and baseline are compared in various link lengths. In Fig. 13 the vertical axis is link power dissipation and horizontal axis is link length. As shown in Fig. 13, by increasing the link length, improvement of the proposed method increases as well. The reason of this improvement is because the longer wires have the bigger capacitance and under this circumstance when the coding algorithm decreases the number of switching activities, more power improvement is possible.

## 6 Overhead

In this section the overhead of the proposed method on power consumption, critical path and area of routers is considered. The overhead is created by two extra modules, coder and the decoder of MFLP, which are inserted in routers. Entire system including encoding and decoding algorithms is implemented in VHDL and synthesized with Synopsys design compiler in 65 nm technology. According to the ITRS [34], in this technology  $V_{dd}$  is defined as 1 Volt and the clock frequency is 500MHz based on the critical path of the system. The topology is mesh with XY routing algorithm and the number of nodes is 16 while the packet length and number of virtual channels are 32 and 2, respectively. The power and area consumption of the coder and decoder are considered as the overhead of power and area which posed by our approach. It is worth mentioning that due to the fact that generating the coding and decoding trees are being done while the packets are transferring, the throughput of system remained unchanged. On the other hand, encoder and decoder can pose power, area and critical path overhead on the routers which are considered in efficiency evaluation of our method. Table 8 depicts the power, critical path and area overhead of the proposed method on routers.

**Table 8** Power, critical path and area overhead of MFLP

Power (mW)			Critical Path (ns)			Area ( $\mu m^2$ )		
W.C.	MFLP	Overhead (%)	W.C.	MFLP	Overhead (%)	W.C.	MFLP	Overhead (%)
52.12	53.53	2.70	1.96	1.97	0.51	36108.32	41679.82	15.43

## 7 Conclusion

This paper presents a new encoding approach with main goal of decreasing the number of switching activities and thereby improving the power consumption. Regarding the fact that sending the data by transition signaling depends on the number of switching activities, it can be concluded that reduction in the number of transitions would lead to less power dissipation. MFLP uses a tree based structure to assign a code word with the less number of ones to the symbols with high frequency. This algorithm does not rely on spatial redundancy and therefore, compared to the other encoding algorithms, it can reduce the power dissipation in the NoCs. The proposed method has demonstrated advantages in both lowering the power and providing data compression. We have evaluated our technique with some benchmarks and compared with previous state of the art encoding algorithms in the literature, MFLP is able to reduce the power consumption in the NoCs by up to 46 %. Power, critical path and area overhead of the proposed method are 2.70, 0.51 and 15.43 %, respectively.

**Acknowledgments** The authors are indebted to Dr. Dara Rahmati because of letting us use Persian tool as a NoC infrastructure as well as editors and the referees of the journal for their constructive comments that improved the quality of this paper.

## References

1. Marculescu R et al (2009) Outstanding research problems in NoC domain: system, microarchitecture, and circuit perspectives. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 28(1):3–21
2. Benini L, De Micheli G (2002) Networks on chip: a new SoC paradigm. *IEEE Comput* 35(1):70–78
3. Palma JCS, Indrusiak LS, Moraes FG, Ortiz AG, Glesner M, Reis RAL (2007) Inserting data encoding techniques into NoC-based systems. In: *Proceedings of ISVLSI*. pp 299–304
4. Pasricha S, Dutt N (2008) Trends in emerging on-chip interconnect technologies. *IPSP Trans Syst LSI Des Methodol* 1:2–17
5. Postman J, Krishna T, Edmonds C, Peh L, Chiang P (2013) SWIFT: A low-power network-on-chip implementing the token flow control router architecture with swing-reduced interconnects. *IEEE Trans VLSI* 21(8):1432–1446
6. Reehal G, Ismail M (2014) A systematic design methodology for low-power NoCs. *IEEE Trans VLSI* 22(12):2585–2595
7. Kulkarni M, Agrawal V (2011) Energy source lifetime optimization for a digital system through power management. In: *43rd Southeastern symposium on system theory*, pp 73–78
8. Svensson C (2001) Optimum voltage swing on on-chip and off-chip interconnect. *IEEE J Solid-State Circuits* 36(7):1108–1112
9. Wei L, Chen Z, Johnson M, Roy K, De V (1999) Design and optimization of dual-threshold circuits for low voltage low power applications. *IEEE Trans VLSI* 7(1):6–24
10. Shin D, Kim W, Kwon S, Han TH (2011) Communication-aware VFI partitioning for GALS-based networks-on-chip. *Des Autom Embed Syst* 15(2):89–109
11. Moyer B (2001) Low power design for embedded processors. *Proc IEEE* 89(11):1576–1587
12. Snowdon DC, Ruocco S, Heiser G (2005) Power management and dynamic voltage scaling: Myths and facts. In: *Proceedings of the workshop on power aware real-time computing*, pp 1–7

13. Benini L, Bogliolo A, De Micheli G (2000) A survey of design techniques for system level dynamic power management. *IEEE Trans VLSI* 8(3):299–316
14. Arelakis A, Stenstrom P (2014) SC2: a statistical compression cache scheme. In: Proceedings of the 41st annual international symposium on computer architecture
15. Anagnostopoulos I, Bartzas A, Filippopoulos I, Soudris D (2012) High-level customization framework for application-specific NoC architectures. *Des Autom Embed Syst* 16(4):339–361
16. Palesi M, Ascia G, Fazzino F, Catania V (2011) Data encoding schemes in network on chip. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 30(5):774–786
17. Stan MR, Burleson WR (1994) Limited-weight codes for low power I/O. In: International workshop on low power design, pp 209–214
18. Stan MR, Burleson WP (1995) Bus-Invert coding for low- power I/O. *IEEE Trans VLSI* 3:49–59
19. Kim KW, Baek KH, Shanbhag N, Liu CL, Kang S (2000) Coupling driven signal encoding scheme for low power interface design. In: Proceedings of ICCAD, pp 318–321
20. Benini L, De Micheli G, Macii E, Poncino M, Quer S (1997) System level power optimization of special purpose applications: the Beach solution. In: Proceedings of international symposium on low power electronics and design, Monterey, pp 24–29
21. Taassori M, Hessabi S (2009) Low power encoding in NOCs based on coupling transition avoidance. In: Proceedings of DSD conferences, pp 247–254
22. Mamidipaka MN, Hirschberg DS, Dutt ND (2003) Adaptive low power address encoding techniques using self-organizing lists. *IEEE Trans VLSI* 11(5):827–834
23. Cheng WC, Pedram M (2002) Power-optimal encoding for a DRAM address bus. *IEEE Trans VLSI* 10(2):109–118
24. Lee K, Lee SJ (2004) SILENT: serialized low energy transmission coding for on-chip interconnection networks. In: ICCAD, pp 448–451
25. Benini L, Macii A, Macii E, Poncino M, Scarsi R (2000) Architectures and synthesis algorithms for power-efficient bus interfaces. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 19(9):969–980
26. Lv T, Henkel J, Lekates H, Wolf W (2003) A dictionary based en/decoder scheme for low power data buses. *IEEE Trans VLSI* 11(5):943–951
27. Brahmabhatt AR, Zhang J, Wu Q, Qiu Q (2006) Low power bus encoding using adaptive hybrid algorithm. In: Design Automation Conference (DAC), pp 987–990
28. Benini L, De Micheli G (2006) Networks on chips: technology and tools. morgan kufmann publishers, Burlington
29. Jafarzadeh N, Palesi M, Khademzadeh A, Afzali-Kusha A (2014) Data encoding techniques for reducing energy consumption in network-on-chip. *IEEE Trans VLSI* 22(3):675–685
30. Jafarzadeh N, Palesi M, Eskandari S, Hessabi S, Afzali-Kusha A (2015) Low energy yet reliable data communication scheme for network on chip. *IEEE Trans Comput-Aided Des Integr Circuits Syst*. 34(12):1892–1904
31. Vitkovski R, Haukilahti A, Jantsch, Nilsson E (2004) Low-power and error coding for network-on-chip traffic. In: Proceedings of norchip, pp 20–23
32. Hale KC, Grot B, Keckler SW (2009) Segment gating for static energy reduction in networks-on-chip. In: Proceedings of network on chip architectures, pp 57–62
33. Raghunathan V, Srivastavat MB, Gupta RK (2003) A survey of techniques for energy efficient on-chip communication. In: Proceedings of design automation conference, pp 900–905
34. International Technology Roadmap for Semiconductors (ITRS), (2011) Available: <http://www.itrs.net>