

Encoding Techniques For Energy Reduction

Veeranna¹, Mallikarjun², Sai manohar³

*(Department of ECE, SBIT College/ JNTUH University, India)Email: dev.veer57@gmail.com)

** (Department of ECE, SBIT College/ JNTUH University, India Email: mallikarjun.n20@gmail.com)

***(Department of ECE, SBIT College, Khammam/JNTUH, INDIA) Email: manoharsra1312@gmail.com)

Abstract : As technology shrinks, the power dissipated by the links of a network-on-chip (NoC) starts to compete with the power dissipated by the other elements of the communication subsystem, namely, the routers and the network interfaces (NIs). In this paper, we present a set of data encoding schemes aimed at reducing the power dissipated by the links of an NoC. The proposed schemes are general and transparent with respect to the underlying NoC fabric (i.e., their application does not require any modification of the routers and link architecture). Experiments carried out on both synthetic and real traffic scenarios show the effectiveness of the proposed schemes, which allow us to save power dissipation and energy consumption without any significant performance degradation and with less area overhead in the NI.

Keywords: Coupling switching activity, data encoding, interconnection on chip, low power, network-on-chip (NoC), power analysis.

I. INTRODUCTION

Shifting from a silicon technology node to the next one results in faster and more power efficient gates but slower and more power hungry wires [1]. In fact, more than 50% of the total dynamic power is dissipated in interconnects in current processors, and this is expected to rise to 65%–80% over the next several years [2]. Global interconnect length does not scale with smaller transistors and local wires. Chip size remains relatively constant because the chip function continues to increase and RC delay increases exponentially. At 32/28 nm,

for instance, the RC delay in a 1-mm global wire at the minimum pitch is 25× higher than the intrinsic delay of a two-input NAND fanout of 5 [1].

If the raw computation horsepower seems to be unlimited, thanks to the ability of instancing more and more cores in a single silicon die, scalability issues, due to the need of making efficient and reliable communication between the increasing number of cores, become the real problem [3]. The network on chip (NoC) design paradigm [4] is recognized as the most viable way to tackle with scalability and variability issues that characterize the ultra deep submicron meter. Nowadays, the on-chip communication issues are as relevant

as, and in some cases more relevant than, the computation related issues [4]. In fact, the communication subsystem increasingly impacts the traditional design objectives, including cost (i.e. silicon area), performance, power dissipation, energy consumption, reliability, etc. As technology shrinks an ever more significant fraction of the total power budget of a complex many-core system-on-chip (SoC) is due to the communication subsystem.

In this project, we focus on techniques aimed at reducing the power dissipated by the network links. In fact, the power dissipated by the network links is as relevant as that dissipated by routers and network interfaces (NIs) and their contribution is expected to increase as technology scales. In particular, we present a set of data encoding schemes operating at flit level and on an end-to-end basis, which allows us to minimize both the switching activity and the coupling switching activity on links of the routing paths traversed by the packets. The proposed encoding schemes, which are transparent with respect to the router implementation, are presented and discussed at both the algorithmic level and the architectural level, and assessed by means of simulation on synthetic and real traffic scenarios.

The advantages of our technique over the existing techniques are as follows:

- Higher power saving
- Minimization coupling transition activities

II. DESIGN OF PROPOSED SYSTEM

I present the proposed encoding scheme whose goal is to reduce power dissipation by minimizing the coupling transition activities on the links of the interconnection network. Let us first describe the power model that contains different components of power dissipation of a link. The dynamic power dissipated by the interconnects and drivers is

$$P = [T_{0 \rightarrow 1} (C_s + C_l) + T_c C_c] V_{dd}^2 F_{ck} \dots \quad (1)$$

Where $T_{0 \rightarrow 1}$ is the number of 0 → 1 transitions in the bus in two consecutive transmissions, T_c is the number of correlated switching between physically adjacent lines, C_s is the line to substrate capacitance, C_l is the load capacitance, C_c is the coupling capacitance, V_{dd} is the supply voltage, and F_{ck} is the clock frequency. One can classify four types of coupling transitions. A Type I transition occurs when one of the lines switches when the other remains unchanged. In a Type II transition, one line switches from low to high while the other makes transition from high to low. A Type III transition corresponds to the case where both lines switch simultaneously. Finally, in a Type IV transition both lines do not change.

The effective switched capacitance varies from type to type, and hence, the coupling transition activity, T_c , is a weighted sum of different types of coupling transition contributions.

Therefore

$$T_c = K_1 T_1 + K_2 T_2 + K_3 T_3 + K_4 T_4 \quad (2)$$

Where T_i is the average number of Type I transition and K_i is its corresponding weight. We use $K_1=1$, $K_2=2$, and $K_3=K_4=0$. The occurrence probability of Types I and II for a random set of data is 1/2 and 1/8, respectively. This leads to a higher Value for $K_1 T_1$ compared with $K_2 T_2$ suggesting that minimizing the number of Type I transition may lead to a considerable power reduction. Using (2), one may express (1) as

$$P = [T_{0 \rightarrow 1} (C_s + C_l) + (T_1 + 2T_2) C_c] V_{dd}^2 F_{ck} \quad (3)$$

According to, C_l Can be neglected

$$P \propto T_{0 \rightarrow 1} C_s + (T_1 + 2T_2) C_c \quad (4)$$

Here, we calculate the occurrence probability for different types of transitions. Consider that flit (t-1) and flit (t) refer to the previous flit which was transferred via the link and the flit which is about to pass through the link, respectively. We consider only two adjacent bits of the physical channel. Sixteen different combinations of these four bits could occur (Table I). Note that the first bit is the value of the generic i^{th} line of the link, whereas the second bit represents the value of its (i+1)th line. The number of transitions for Types I, II, III, and IV are 8, 2, 2, and 4, respectively. For a random set of data, each of these sixteen transitions has the same probability. Therefore, the occurrence probability for Types I, II, III, and IV are 1/2, 1/8, 1/8, and 1/4, respectively.

In the rest of this section, we present three data encoding schemes designed for reducing the dynamic power dissipation of the network links along with a possible hardware implementation of the decoder.

A. Scheme I

In scheme I, we focus on reducing the numbers of Type I transitions (by converting them to Types III and IV transitions) and Type II transitions (by converting them to Type I transition). The scheme compares the current data with the previous one to decide whether odd inversion or no inversion of the current data can lead to the link power reduction.

1) *Power Model*: If the flit is odd inverted before being transmitted, the dynamic power on the link is

$$P' \propto T'_{0 \rightarrow 1} + (K_1 T'_1 + K_2 T'_2 + K_3 T'_3 + K_4 T'_4) C_c \quad (5)$$

Where $T'_{0 \rightarrow 1}, T'_1, T'_2, T'_3,$ and T'_4 , are the self-transition activity, and the coupling transition activity of Types I, II, III, and IV, respectively. Table I reports, for each transition, the relation-ship between the coupling transition activities of the flit when transmitted as is and when its bits are odd inverted. Data are organized as follows. The first bit is the value of the generic i^{th} line of the link, whereas the second bit represents the value of its (i+1)th line. For each partition, the first (second) line represents the values at time t-1(t).

Table1: Effect of ODD Inversion on Change of Transition Types

Time	Normal			Odd Inverted		
	Type I			Types II, III, and IV		
t-1	00, 11	00, 11, 01, 10	01, 10	00, 11	00, 11, 01, 10	01, 10
t	10, 01	01, 10, 00, 11	11, 00	11, 00	00, 11, 01, 10	10, 01
	T1*	T1**	T1***	Type III	Type IV	Type II
t-1	Type II			Type I		
t	01, 10			01, 10		
	10, 01			11, 00		
t-1	Type III			Type I		
t	00, 11			00, 11		
	11, 00			10, 01		
t-1	Type IV			Type I		
t	00, 11, 01, 10			00, 11, 01, 10		
	00, 11, 01, 10			01, 10, 00, 11		

As Table I shows, if the flit is odd inverted, Types II, III, and IV transitions convert to Type I transitions. In the case of Type I transitions, the inversion leads to one of Types II, III, or Type IV transitions. In particular, the transitions indicated as

$T^*_{0 \rightarrow 1}$, $T^{**}_{0 \rightarrow 1}$, and $T^{***}_{0 \rightarrow 1}$ in the table convert to Types II, III, and IV transitions, respectively. Also, we have $T^*_{0 \rightarrow 1} = T_{0 \rightarrow 0(\text{odd})} + T_{0 \rightarrow 1(\text{even})}$ where odd/even refers to odd/even lines. Therefore, (5) can be expressed as $P \propto (T_{0 \rightarrow 0(\text{odd})} + T_{0 \rightarrow 1(\text{even})})C_s + [K_1(T_2 + T_3 + T_4) + K_2T^{***}_{0 \rightarrow 1} + K_3T^*_{0 \rightarrow 1} + K_4T^{**}_{0 \rightarrow 1}]C_c$ (6)

Thus, if $P > P_{-}$, it is convenient to odd invert the flit before transmission to reduce the link power dissipation. Using (4) and (6) and noting that $C_c/C_s = 4$, we obtain the following odd invert condition

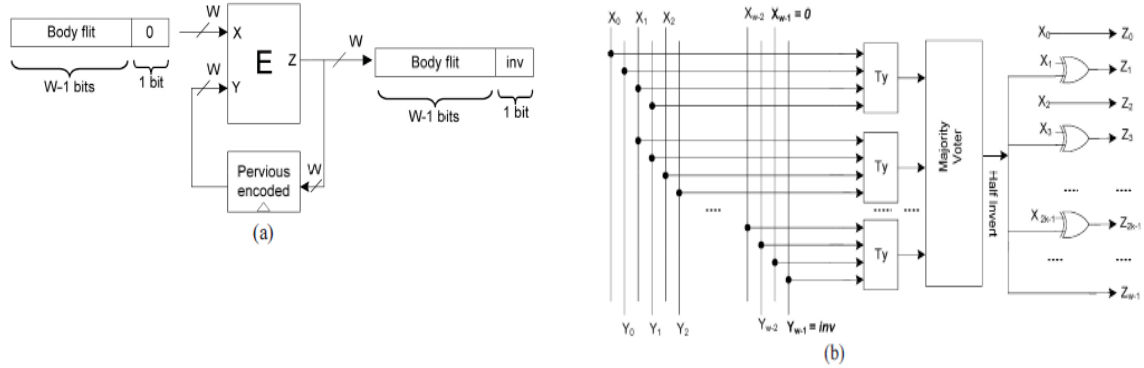


Fig. Encoder Architecture Scheme I.(a)Circuit diagram (b)Internal view of Encoder block (E)

Which is the exact condition to be used to decide whether the odd invert has to be performed. Since the terms $T_{0 \rightarrow 1(\text{odd})}$ and $T_{0 \rightarrow 0(\text{odd})}$ are weighted with a factor of 1/4, for link widths greater than 16 bits, the misprediction of the invert condition will not exceed 1.2% on average. Thus, we can approximate the exact condition as

$$T_1 + 2T_2 > T_2 + T_3 + T_4 + 2T_1^{***} \quad (8)$$

2) Proposed Encoding Architecture: The proposed encoding architecture, which is based on the odd invert condition defined by (12), is shown in Fig.. We consider a link width of w bits. If no encoding is used, the body flits are grouped in w bits by the NI and are transmitted via the link. In our approach, one bit of the link is used for the inversion bit, which indicates if the flit traversing the link has been inverted or not. More specifically, the NI packs the body flits in $w - 1$ bits [Fig. 1(a)]. The encoding logic E , which is integrated into the NI, is responsible for deciding if the inversion should take place and performing the inversion if needed. The generic block diagram shown in Fig. 1(a) is the same for all three encoding schemes proposed in this project and only the block E is different for the schemes. To make the decision, the previously encoded flit is compared with the current flit being transmitted. This latter, whose w bits are the concatenation of $w - 1$ payload bits and a “0” bit, represents the first input of the encoder, while the previous encoded flit represents the second input of the encoder [Fig. 1(b)]. The $w - 1$ bits of the incoming (previous encoded) body flit are indicated by $X_i (Y_i)$, $i = 0, 1, \dots, w - 2$. The w th bit of the previously encoded body flit is indicated by inv which shows if it was inverted ($inv = 1$) or left as it was ($inv = 0$). In the encoding logic, each T_y block takes the two adjacent bits of the input flits (e.g., $X_1X_2Y_1Y_2$, $X_2X_3Y_2Y_3$, $X_3X_4Y_3Y_4$, etc.) and sets its output to “1” if any of the transition types of T_y is detected. This means that the odd inverting for this pair of bits leads to the reduction of the link power dissipation (Table I). The T_y block may be implemented using a simple circuit. The second stage of the encoder, which is a majority voter block, determines if the condition given in (12) is satisfied (a higher number of 1s in the input of the block compared to 0s). If this condition is satisfied, in the last stage, the inversion is performed on odd bits. The decoder circuit simply inverts the received flit when the inversion bit is high.

B. Scheme II

In the proposed encoding scheme II, we make use of both odd (as discussed previously) and full inversion. The full inversion operation converts Type II transitions to Type IV transitions. The scheme compares the current data with the previous one to decide whether the odd, full, or no inversion of the current data can give rise to the link power reduction.

1) Power Model: Let us indicate with P , P' , and P'' the power dissipated by the link when the flit is transmitted with no inversion, odd inversion, and full inversion, respectively. The odd inversion leads to power reduction when $P' < P''$ and $P' < P$. The power P'' is given

$$P'' \propto T_1 + 2T_4^{**}. \quad (13)$$

Neglecting the self-switching activity, we obtain the condition $P' < P''$ as

$$T_2 + T_3 + T_4 + 2T_1^{***} < T_1 + 2T_4^{**}. \quad (14)$$

Therefore, using (9) and (11), we can write

$$2(T_2 - T_4^{**}) < 2T_y - w + 1. \quad (15)$$

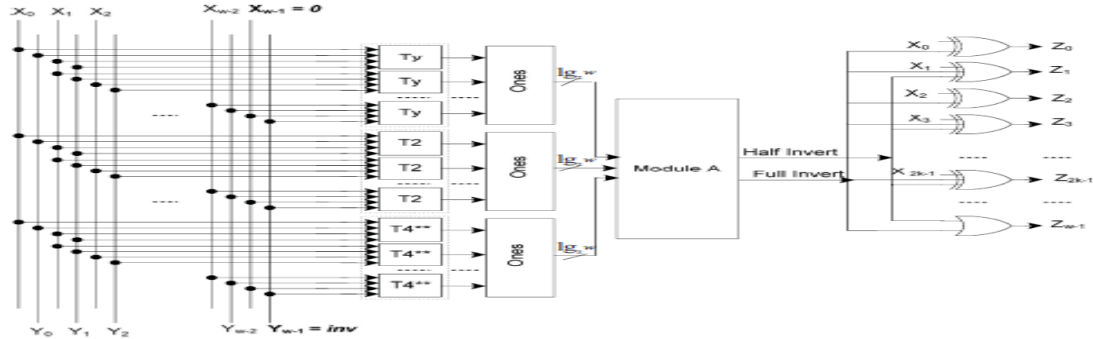


Fig. 2. Encoder architecture Scheme II.

Based on (12) and (15), the odd inversion condition is obtained as

$$2(T_2 - T_4^{**}) < 2T_y - w + 1 \quad T_y > \frac{(w-1)}{2}. \quad (16)$$

Similarly, the condition for the full inversion is obtained from $P'' < P$ and $P'' < P'$. The inequality $P < P$ is satisfied when

$$T_2 > T_4^{**}. \quad (17)$$

Therefore, using (15) and (17), the full inversion condition is obtained as

$$2(T_2 - T_4^{**}) > 2T_y - w + 1 \quad T_2 > T_4^{**}. \quad (18)$$

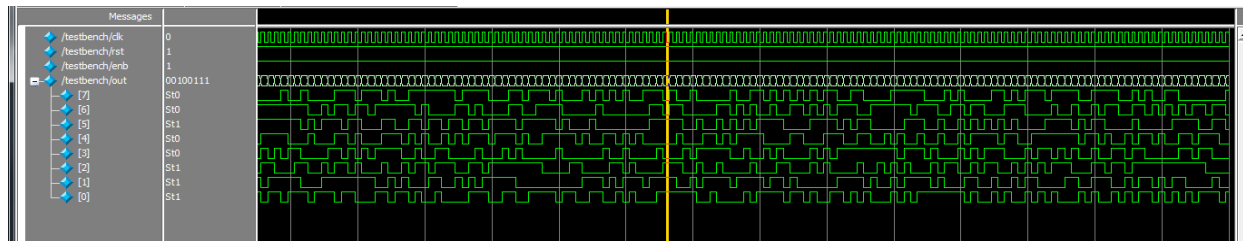
When none of (16) or (18) is satisfied, no inversion will be performed.

2) Proposed Encoding Architecture: The operating principles of this encoder are similar to those of the encoders implementing Schemes I and II. The proposed encoding architecture, which is based on the even invert condition of (28), the full invert condition of (29), and the odd invert condition of (30), is shown in Fig.. The wth bit of the previously encoded body flit is indicated by inv which shows if it was even, odd, or full inverted (inv = 1) or left as it was (inv = 0). The first stage of the encoder determines the transition types while the second stage is formed by a set of 1s blocks which count the number of ones in their inputs. In the first stage, we have added the T_e blocks which determine if any of the transition types of T_2, T_1^{**} , and T_1^{***} is detected for each pair bits of their inputs. For these transition types, the even invert action yields link power reduction. Again, we have four Ones blocks to determine the number of detected transitions for each T_y, T_e, T_2, T_4^{**} , blocks. The output of the Ones blocks are inputs for Module C. This module determines if odd, even, full, or no invert action corresponding to the outputs "10," "01," "11," or "00," respectively, should be performed. The outputs "01," "11," and "10" show that whether (28), (29), and (30), respectively, are satisfied. In this project, Module C was designed based on the conditions given in (28), (29), and (30). Similar to the procedure used to design the decoder for scheme II, the decoder for scheme III may be designed.

III. RESULTS AND DISCUSSION

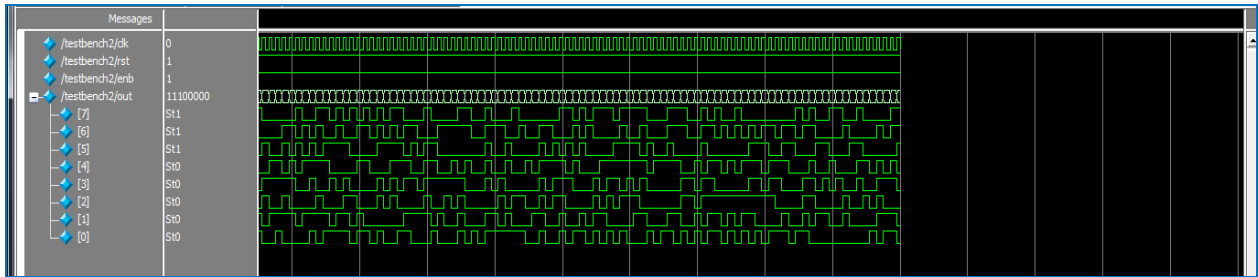
1) Scheme-1

a) Simulation waveform



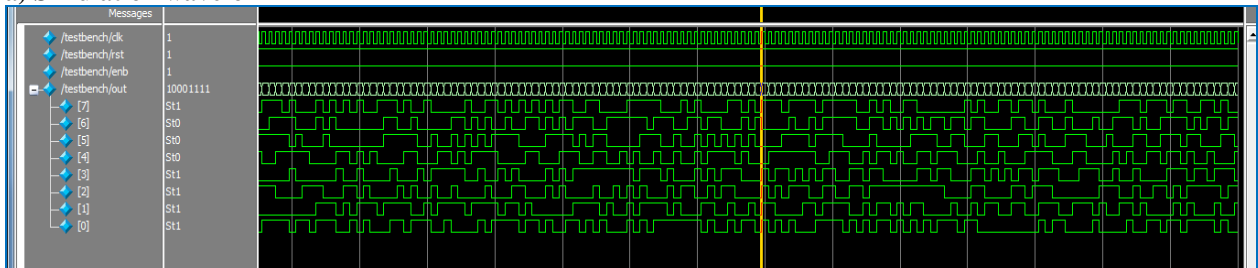
2) Scheme-2

a) Simulation waveform



3) Scheme-3

a) Simulation waveform



The proposed data encoding schemes have been assessed by means of a cycle-accurate NoC simulator based on Noxim. The power estimation models of Noxim include NIs, routers, and links. The link power dissipation was computed using where the terms $T_0 \rightarrow 1$, T_1 , and T_2 were computed based on the information obtained from the cycle accurate simulation. The following parameters were used in the simulations. The NoC was clocked at 500 MHz while the baseline NI with minimum buffering and supporting open core protocol 2 and advanced high-performance bus protocols dissipated 5mW. The average power dissipated by the worm hole-based router was 5.4 mW. Based on a 65-nm UMC technology, a total capacitance of 592 fF/mm was assumed for an inter-router wire. About 80% of this capacitance was due to the crosstalk. Using the detailed simulations, when the flits traversed the NoC links, the corresponding self and coupling switching activities were calculated and used along with the self- and coupling capacitance of 0.231 and 0.847 nf, respectively, to calculate the power ($V_{dd} = 0.5$ V and $F_{ck} = 500$ MHz).

IV. CONCLUSION

In this project, we have presented a set of new data encoding schemes aimed at reducing the power dissipated by the links of an NoC. In fact, links are responsible for a significant fraction of the overall power dissipated by the communication system. In addition, their contribution is expected to increase in future technology nodes. As compared to the previous encoding schemes proposed in the literature, the rationale behind the proposed schemes is to minimize not only the switching activity, but also (and in particular) the coupling switching activity which is mainly responsible for link power dissipation in the deep sub-micrometer technology regime.

The proposed encoding schemes are agnostic with respect to the underlying NoC architecture in the sense that their application does not require any modification neither in the routers nor in the links. An extensive evaluation has been carried out to assess the impact of the encoder and decoder logic in the NI. The encoders implementing the proposed schemes have been assessed in terms of power dissipation and silicon area. The impacts on the performance, power and energy metrics have been studied using a cycle- and bit accurate NoC simulator under both synthetic and real traffic scenarios. Overall, the application of the proposed encoding schemes allows savings of power dissipation and energy consumption without any significant performance degradation and with less than area overhead in the NI.

REFERENCES

- [1] International Technology Roadmap for Semiconductors. (2011) [Online]. Available: <http://www.itrs.net>
- [2] M. S. Rahaman and M. H. Chowdhury, "Crosstalk avoidance and error correction coding for coupled RLC interconnects," in Proc. IEEE Int. Symp. Circuits Syst., May 2009, pp. 141–144.
- [3] W. Wolf, A. A. Jerraya, and G. Martin, "Multiprocessor system-on-chip MPSoC technology," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 27, no. 10, pp. 1701–1713, Oct. 2008.
- [4] L. Benini and G. De Micheli, "Networks on chips: A new SoC paradigm," Computer, vol. 35, no. 1, pp. 70–78, Jan. 2002.

- [5] S. E. Lee and N. Bagherzadeh, "A variable frequency link for a powerawarenetwork-on-chip (NoC)," *Integr. VLSI J.*, vol. 42, no. 4, pp. 479–485, Sep. 2009.
- [6] D. Yeh, L. S. Peh, S. Borkar, J. Darringer, A. Agarwal, and W. M. Hwu, "Thousand-core chips roundtable," *IEEE Design Test Comput.*, vol. 25, no. 3, pp. 272–278, May–Jun. 2008.
- [7] A. Vittal and M. Marek-Sadowska, "Crosstalk reduction for VLSI," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 16, no. 3, pp. 290–298, Mar. 1997.
- [8] M. Ghoneima, Y. I. Ismail, M. M. Khellah, J. W. Tschanz, and V. De, "Formal derivation of optimal active shielding for low-power on-chip buses," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 5, pp. 821–836, May 2006.



D.Veeranna was born on MARCH 22, 1983 , India and got B.E ECE, from CBIT Hyderabad in the Year 2005, M.Tech. (VLSI SD) from NITW University, Warangal in the Year 2008. He has got 07 Years of Teaching Experience. Presently, he is working as Asst., Professor in Department of Electronics and communication Engineering in SBIT Engineering College Khammam. His area of interest are fault tolerance, image processing



Mallikarjun Nuthanaganti was born on JULY 02, 1986 , India and got B.Tech in Electronics and Communication Engineering from JNTUH Hyderabad in the year of 2008. He completed his M.Tech in Embedded Systems branch from JNTUH university in the year of 2010. After completion of his M.Tech he joined as a Assistant Professor in ECE Department at SBIT- Khammam affiliated to JNTUH. His area of interest are Embedded system, image processing.



S.Sai manohar was born on MAY 15, 1986 , India and got B.Tech in Electronics and Communication Engineering from KU Warangal in the year of 2007. He completed his M.Tech in Embedded Systems branch from JNTUH university in the year of 2010. After completion of his M.Tech he joined as a Assistant Professor in ECE Department at SBIT- Khammam affiliated to JNTUH. His area of interest are Embedded system, VLSI.