

An Energy-Efficient Ternary Interconnection Link for Asynchronous Systems

Jean-Marc Philippe
CEA-List DRT/DTSI/SARC/LCEI
F-91191 Gif-sur-Yvette
Jean-Marc.Philippe@cea.fr

Ékué Kinvi-Boh, Sébastien Pillement, Olivier Sentieys
IRISA - University of Rennes (ENSSAT)
6, rue de Kerampont 22300 Lannion, France
{kinviboh,pillemen, sentieys}@irisa.fr

Abstract— We introduce a new ternary link including a binary-to-ternary encoder and a ternary-to-binary decoder in voltage-mode multiple-valued logic (MVL). This link improves the transistor count compared to existing designs and it has no DC current path. The complete link was simulated with SPICE and a $0.13\mu\text{m}$ CMOS technology. It additionally shows interesting advantages on power consumption for global interconnects compared to full-swing signaling binary systems (up to 56.4% less energy consumption). Its low propagation delay is also an advantage in the design of high-speed on-chip links for asynchronous systems.

I. INTRODUCTION

In modern CMOS technologies, interconnects represent a significant part of the power consumption (up to 50%) [1] and of the chip area. New constraints (such as costs or speed) on systems-on-chip (SoC) with deep submicron technologies require having a low-power and high-speed interconnect [2]. Another important part of the power consumption is due to the global clock needed to synchronise the system, because of its high switching activity behaviour.

Moreover, due to the decrease of the dimensions of the devices, interconnection delay tends to be the bottleneck for the chip performances, as the wire delay becomes greater than the gate delay [3]. This makes more and more difficult to globally synchronize an SoC.

One solution to these issues is to use asynchronous systems [4]. These systems are not controlled by a clock. They are composed generally of several clock-less processing blocks and the synchronization is done by an handshake protocol that adds a signal to assert the validity of the transmitted data. Among the various protocols that can be used, we highlight the four-phase protocol as it is widely used in asynchronous communications (Fig. 1). The sender puts the data on the data bus and raises the *request* signal (phase ①). Then the receiver detects the presence of the data because of the *request* signal state, so it can process them. Once he has finished the processing, it raises the *acknowledge* signal (phase ②) that tells the sender to transmit invalid (or empty) data on the bus by asserting the *request* signal to "0" (phase ③). The final phase of the protocol is the assertion of the *acknowledge* signal to "0" once the receiver has sampled a "0" on the *request* signal (phase ④).

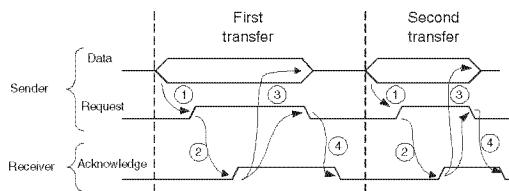


Fig. 1. Illustration of the four-phase handshake protocol.

This protocol is not transition efficient because of the return to zero, but this fact enables the control logic to be very simple [5].

The most common implementation of this protocol is the dual-rail signaling scheme with return to zero. This consists in doubling the number of wires to code the information as it is described by Fig. 2. An important disadvantage of this method is that the number of wires is very high and it makes wire routing very difficult and also time and power consuming.

Recent researches have focused on reducing the interconnect area as well as the pin requirements. One idea consists in increasing the data rate on a wire by having more than two logic states: this research field is called multiple-valued logic (MVL). This idea is also used to design high-speed inter-chip links using pulse-amplitude modulation (PAM) [6]. The voltage mode ternary equivalent of the dual-rail encoding with return to zero is also shown on Fig. 2.

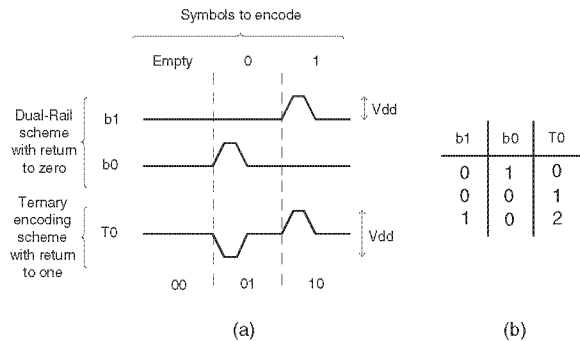


Fig. 2. Binary dual-rail and ternary encoding of the same sequence of bits.

The potential advantages of using MVL to design asynchronous links are the area reduction and low power consumption. In fact, we only need one wire instead of two because in dual-rail signaling, one combination of the two bits is not used. In the same way, the area reduction and the voltage reduction on the link permit to lower power consumption.

This paper introduces a low-power and high-speed ternary link using new converters architecture. It is designed with recent technologies to meet the requirements of SoC. The remainder of this paper is as follows. Section II quickly reviews some of the existing implementations of ternary links. Our approach and its feasibility are explained in section III. Section IV deals with the implementations of the encoder and the decoder. We present the experimental results in section V and finally, section VI concludes this paper.

II. RELATED WORKS ON TERNARY LOGIC

The solutions which can be found in the literature for converting two binary signals into a ternary one are composed of two parts: a binary-to-ternary encoder and a ternary-to-binary decoder. A basic description of the link is shown in Fig. 3.

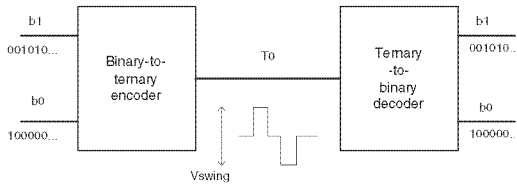


Fig. 3. Basic description of the link.

One of the most recent link is the one introduced in [7]. The authors present two encoders and one decoder that need two power supplies for the different voltage levels. The first encoder needs 9 transistors and some of them are wide and must be sized tightly. The second encoder needs 20 transistors but it is faster than the previous one. The decoder is designed as a bank of comparators and needs 12 transistors. Even if these schemes have no DC current paths in the stable state, they produce big current peaks when they make a transition because they need to break an equilibrium between transistors. Another issue is the fact that the transistors must be adequately sized to perform well. Moreover, the experiments were made with a $0.35\mu\text{m}$ with a supply voltage of 3.3V . The authors point out the limits of operability of their circuits to be about 1V . So this technique is not well suited for very deep-submicron process. In [8], the authors propose a three-level bus signaling scheme including converters from binary to ternary codes. The aim is to reduce crosstalk effects and therefore to decrease worst-case delay and power. Simulation results are given on a non-realistic 2cm bus and the coder and decoder circuits imply a large area overhead equivalent to 500 2-input AND gates for a 8-bit bus.

III. DESCRIPTION OF THE MODIFIED TRANSISTOR LIBRARY

This paper proposes a new way to design the comparators of the decoders. It consists in having a bank of inverters. The transistors thresholds are set by the fabrication process. This technology is used successfully in [9] where the authors design ternary cells using the following methodology. The optimum voltage thresholds are determined by equations 1 and 2 for PMOS and NMOS respectively.

$$V_{TH}(PMOS) = Vi - (Vo - (OP \times LSV)) \quad (1)$$

$$V_{TH}(NMOS) = Vi - (Vo + (OP \times LSV)) \quad (2)$$

Vi is the input logic level voltage limit that the transistor must respond to and Vo is the required output logic voltage level of the transistor. OP is the overlap percentage and it is set in this paper to 70%. LSV is the logic step voltage between two consecutive ternary levels.

We use two pairs of transistors with modified thresholds to design the inverters. The required thresholds for the decoder design are given for a $0.13\mu\text{m}$ technology in Table I. The power supply voltage is 1.2V in $0.13\mu\text{m}$ and LSV is therefore set to 0.6V . As an example, if Vi is set to 1.2V and Vo to 0V , the corresponding switch is designed according to equation 2 using a NMOS transistor with a modified threshold of $V_{TH} = 0.78\text{V}$ (corresponding to transistor $N+$ of Table I).

It is interesting to notice that, due to variations inherent to the fabrication process, these thresholds can vary in some limits without deteriorating the system functionality. It only impacts the noise margin by shifting the switching thresholds of the comparators. We can also notice that the switching thresholds of the inverters can be tuned finely by setting proper W/L ratios.

Transistor name	Vth (V) $0.13\mu\text{m}$
$P+$	-0.78
$N+$	0.78
$P-$	-0.18
$N-$	0.18

TABLE I
VOLTAGE THRESHOLD FOR EACH TRANSISTOR IN A $0.13\mu\text{m}$
TECHNOLOGY.

IV. DESCRIPTION OF THE CIRCUITS

This section introduces the schemes and the behavior of the encoder and of the decoder.

A. Encoder

The encoder is dedicated to the conversion of two bits into a ternary-valued signal. It uses two power supplies as it is shown in Fig. 4.

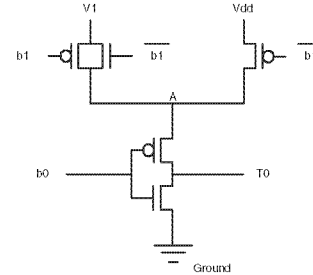


Fig. 4. Ternary encoder.

This encoder needs only 7 standard transistors (5 for the encoder itself and 2 for inverting the $b1$ input). We use the coding presented in Fig. 2b where 0, 1 and 2 are considered to be the three logic levels of the ternary link (respectively given by $V_{ss} = 0$, V_1 and V_{dd} on Fig. 4). In the framework of this paper, V_1 is set to $V_{dd}/2$.

This design enables us to have just one opened branch at a time and a very stable ternary signal. The $b0$ signal drives the central inverter to select one node between A and V_{ss} . The voltage at node A is determined by the $b1$ signal which is driving two switches (using Pass-Transistor Logic) to select the appropriate voltage.

SPICE simulation results in $0.13\mu\text{m}$ are given in Fig. 5. They are obtained by loading the output of the encoder by a 1mm wire (modeled with a $\pi 3$ model) and the decoder. We have represented all the transitions in this signal. It can be seen that the 500MHz output of the encoder is stable so it can be used in a high-speed link.

B. Decoder

The decoder that we propose is composed of 6 transistors (4 custom transistors from Table I and a standard inverter) and it has a very small area. As we can see in Fig. 6, the comparators are composed of two modified inverters whose inputs are the ternary signal.

The power supply of the decoder is the power supply of the circuit (i.e. V_{dd}). The $T0$ ternary signal drives the bank of inverters. Due to the modified thresholds, the inverters permit to isolate each of the three levels and hence to have well-formed binary signals at their outputs.

The first inverter determines the $b1$ signal from the ternary signal after an inversion and the second one determines the $b0$ signal. The truth tables of the custom inverters are given in Fig. 7.

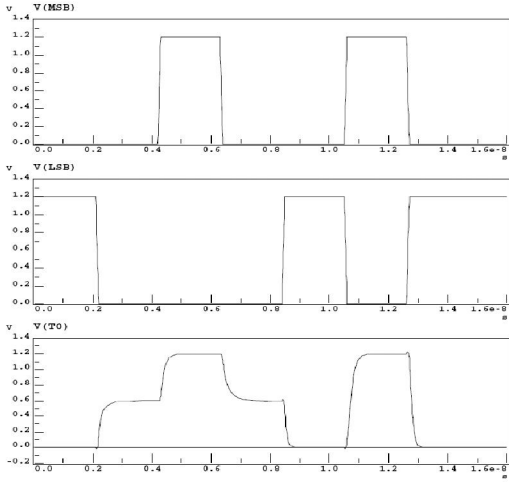


Fig. 5. SPICE simulation results for the encoder with a 500MHz input.

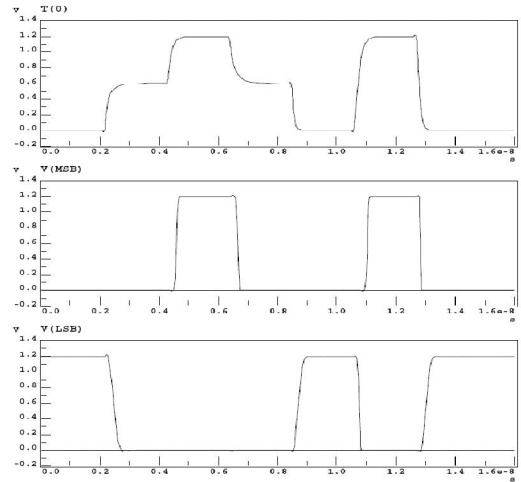


Fig. 8. SPICE simulation results for the decoder with a 500MHz input.

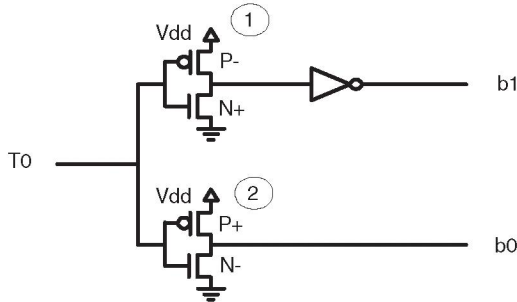


Fig. 6. Ternary decoder.

T0	Output of inverter ①	T0	Output of inverter ②
0	1	0	1
1	1	1	0
2	0	2	0

(a) (b)

Fig. 7. Truth tables of the custom inverters of the ternary decoder.

SPICE simulation results in $0.13\mu\text{m}$ are given in Fig. 8. Each of the decoder outputs is loaded by a standard inverter. The input is generated using SPICE with a fall time and a rise time of 0.1ns for all the transitions.

It can be seen that the decoder can operate with a high-frequency input and hence is adapted to the design of high-speed on-chip links.

V. PERFORMANCES

We have simulated the entire link with SPICE using UMC $0.13\mu\text{m}$ CMOS technologies. The encoder and the decoder are linked by a wire modeled using the $\pi 3$ model. All transistors are designed with common W of 12λ for the PMOS and 6λ for the NMOS. We can expect an improvement of our link by optimizing these dimensions. The link was modeled using UMC rules for a metal-2 layer with a power supply voltage of 1.2V . The extreme voltage levels were set to the ground and the power supply V_{dd} . The intermediate voltage

level was set to $V_{dd}/2$. In order to optimize the delay, we have also simulated the same link with double-sized PMOS for the central inverter of the encoder. We can compare the two implementations from both power consumption and delay points of view.

A. Energy Consumption

This section compares the proposed ternary interconnect link structures with two binary wires modeling a dual-rail signaling scheme. The drivers and receivers of the binary model are simple inverters. The $b1$ and $b0$ inputs are two random signals encoded by a return to zero protocol with a level duration of 10ns to measure the energy consumption of the system with a very long wire up to 10mm . The input rise and fall times are 0.1ns for all signals. The outputs of the decoder and of the two last inverters in the binary links have a load of one common inverter ($W = 12\lambda$ for the PMOS and 6λ for the NMOS). The energy consumption is detailed in Fig. 9 for wire lengths from 1mm to 10mm .

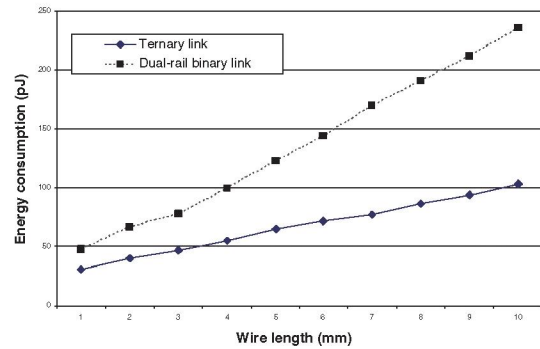


Fig. 9. Energy Consumption (pJ) as a function of the interconnect wire length (mm) for the $0.13\mu\text{m}$ technology.

The energy consumption of the modified ternary link is not shown on that figure because it is very closed to the one of the original link. This figure shows that the proposed ternary link consumes less energy than the equivalent dual-rail binary one. The gain is up to 56.4% in a $0.13\mu\text{m}$ technology.

B. Delay

We measured the propagation delay and the rise and fall times for both the decoder and the encoders and the total propagation delay of the interconnection link for all the possible ternary transitions. Our benchmark is composed of signals with rise and fall times of $0.1ns$. Hence, the slope is not the same for all transitions. In the different tables, each transition is defined by a number given in Fig. 10. An 'X' in the tables means that there is no transition at the output for this transition at the input.

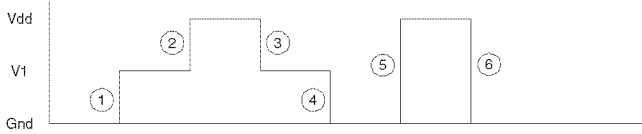


Fig. 10. Ternary input signal transitions used in the testbench.

We define the propagation delay as the delay between the time the input reaches 50% of its transition and the time the output reaches 50% of its transition, even for a transition in the ternary case. The rise (or respectively fall) time is defined as the time needed for a signal to increase from 10% to 90% (or decrease from 90% to 10%) of its maximal value.

1) *The decoder*: The propagation delay and the rise and fall times of the decoder are given in Table II. We can see that the worst case delays are 118ps for the propagation delay and 124ps for the rise/fall time. The very simple inverter-based structure of the decoder enables it to be very fast.

Ternary decoder		b_1	b_0	Ternary decoder		b_1	b_0
Rise/fall times (ps)	①	X	96	Propagation delay (ps)	①	X	68
	②	60	X		②	105	X
	③	43	X		③	82	X
	④	X	124		④	X	72
	⑤	55	41		⑤	118	17
	⑥	27	113		⑥	35	86

TABLE II

RISE/FALL TIMES AND PROPAGATION DELAYS OF THE TERNARY DECODER IN THE $0.13\mu m$ TECHNOLOGY ACCORDING TO TRANSITION NUMBER OF FIG. 10.

2) *The encoder*: We measured the propagation delay and the rise and fall times of the encoders for all the transitions. The worst case results are presented in Table III. The encoder is loaded by a $1 - mm$ wire and the decoder. We can see that the worst case rise/fall time is about 881ps for the first considered transition and the corresponding propagation delay is 449ps.

Doubling the size of the PMOS of the central inverter of the encoder is very interesting for improving the delays. This simple modification enables the encoder to be 33% faster than the original one for the rise and fall times and 28% faster for the propagation delay. We can notice also that doubling the width of this transistor does not increase the power consumption as it is shown by former measures.

3) *Propagation delay of the global interconnection link*: We measured the propagation delay of the global interconnection link (the encoder, a wire and the decoder) for all the possible transitions of the binary inputs and for two wire lengths ($1mm$ and $5mm$).

Ternary encoder		Original	Modified
Rise/fall times (ps)	①	881	590
	②	232	194
	③	411	326
	④	96	93
	⑤	270	221
	⑥	114	115
Propagation delay (ps)	①	449	322
	②	114	99
	③	125	109
	④	53	57
	⑤	154	133
	⑥	71	75

TABLE III

RISE AND FALL TIMES AND PROPAGATION DELAYS OF THE ORIGINAL AND MODIFIED TERNARY ENCODERS IN THE $0.13\mu m$ TECHNOLOGY ACCORDING TO TRANSITION NUMBER OF FIG. 10.

The worst case propagation delay for the original encoder is about 724ps for a $1 - mm$ wire. Using the modified encoder improves the propagation delay to 539ps (which represents a decrease of 25%). With a $5 - mm$ wire, the propagation delays increase to 2.27ns and 1.6ns respectively. The modified encoder permits a delay reduction of 29.5%.

VI. CONCLUSION

A new ternary link designed for asynchronous systems is presented in this paper. This approach can decrease the number of wires compared to traditional techniques and can permit to save silicon area because it divides by two the number of required wires. It can also be used to increase the inter-wire distance, and thus to reduce cross-talk noise. This link was simulated with SPICE models on a recent UMC technology. It has up to 56.4% less power consumption than a dual-rail full-swing signaling system for long global interconnects. This link is also adapted to design high-speed asynchronous interconnects due to its low propagation delay and to the possibility of adapting the width of the central PMOS.

REFERENCES

- [1] H. Zhang, V. George, and J. M. Rabaey, "Low-Swing On-Chip Signaling Techniques: Effectiveness and Robustness," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 8, no. 3, pp. 264–272, June 2000.
- [2] ITRS, "http://public.itrs.net/files/2003itrs/home2003.htm," International Technology Roadmap for Semiconductors, Tech. Rep., 2003.
- [3] R. Ho, K. Mai, and M. Horowitz, "The future of wires," *Proc. IEEE*, vol. 89, no. 4, pp. 490–504, April 2001.
- [4] S. Hauck, "Asynchronous design methodologies: an overview," *Proceedings of the IEEE*, vol. 83, no. 1, pp. 69–93, 1995.
- [5] W. J. Dally and J. W. Poulton, *Digital Systems Engineering*. Cambridge University Press, 1998.
- [6] M. Pedram and J. M. Rabaey, *Power Aware Design Methodologies*. Kluwer Academic Publishers, June 2002, ch. 8, pp. 201–239.
- [7] T. Felicijan and S. B. Furber, "An Asynchronous Ternary Signaling System," *IEEE Transactions On Very Large Scale Integration (VLSI) Systems*, vol. 11, no. 6, pp. 1114–1119, December 2003.
- [8] Y. Zhang, T. Blalock, and M. Stan, "A three-level toggle-avoid bus signaling scheme," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, 2005, pp. 1843–1846.
- [9] E. Kinvih-Boh, M. Aline, O. Sentieys, and E. D. Olson, "MVL circuit design and characterization at the transistor level using SUS-LOC," in *33th International Symposium on Multiple-Valued Logic (ISMVL'03)*, Tokyo (Japan), May 2003, pp. 105–110.