Parametric Timing and Power Macromodels for High Level Simulation of Low-Swing Interconnects

Davide Bertozzi
DEIS-University of Bologna
Viale Risorgimento, 2
40136 Bologna Italy
dbertozzi@deis.unibo.it

Luca Benini
DEIS-University of Bologna
Viale Risorgimento, 2
40136 Bologna Italy
lbenini@deis.unibo.it

Bruno Ricco'
DEIS-University of Bologna
Viale Risorgimento, 2
40136 Bologna Italy
bricco@deis.unibo.it

ABSTRACT

The impact of global on-chip interconnections on power consumption and speed of integrated circuits is becoming a serious concern. Designers need therefore to quickly estimate how performance and power are affected by a given choice of the interconnection parameters (length, voltage swing, driver and receiver schematics and sizing). This work focuses on the entire communication channel (driver, interconnect, receiver), and provides high level parametric VHDL simulation models for low-swing signaling schemes. These SPICE-derived power and timing macromodels transfer electrical-level information to the RTL simulation in an event-driven fashion, as transitions occur at the input of the interconnect driver. The accuracy reached by this back-annotation technique is within 5% with respect to SPICE results, with only 4% simulation speed penalty in the worst case.

Categories and Subject Descriptors

I.6.5 [**Computing Methodologies**]: Simulation and Modeling—*Model development*

General Terms

Design

Keywords

Interconnect, Low-swing, Power, Delay, Macromodel

1. INTRODUCTION

Performance and power consumption of VLSI circuits are increasingly affected by on-chip interconnects. Scaling toward deep submicron makes signal propagation along global wires slower, thus increasing the number of clock cycles required for across-chip communication [15]. Furthermore, the increasing wire-related capacitance will be responsible for an increase in dynamic power dissipation.

Repeater insertion is a way to deal with performance degradation of on-chip interconnects, but it has a cost in terms of power consumption. In [11] it is shown that power consumption of wires and clock signals can account for up to 40% or 50% of the total on-chip power dissipation. For reconfigurable architectures the situation is even worse.

These issues are forcing design paradigm shift from device-centric to interconnect-centric [4]. Hence, many efforts are being devoted to develop interconnect performance models for design planning [7] [16], to carry out design space exploration for communication architectures [10], to develop frameworks wherein architecture design trade-offs are analyzed [12]. This extensive work confirms that designers need to estimate performance and power consumption early in the design stage. This would help investigate the most efficient communication architecture, estimate proper clock cycle time [8], and assess the effectiveness of interconnect optimization techniques [5].

The main contribution of this paper is to augment HDL simulation with accurate interconnect models that enable design space exploration of interconnect-centric architectures at a high abstraction level. The contribution is twofold:

- Considering global wires and their impact over power and performance at a high abstraction level requires information that only a low-level electrical characterization can provide.

  With respect to this issue, we pre-characterize different interconnection schemes (including driver, wire and receiver) in terms of delay and energy consumption by means of SPICE simulations. Collected data are then used to construct lookup tables, which are accessed by high level VHDL simulations in an event-driven fashion to provide power and performance estimations.

  Compared with traditional library characterization and macromodeling [9] [3], the interconnect macromodels we expose to the high level simulations are parametric. Hence, early in the design stage, the impact of wire-related parameters such as driver and receiver power supply, interconnect swing and length can be investigated. This allows investigation and selection of an optimized interconnect architecture configuration, based on the system constraints (delay or power optimization or tradeoff solution).

ISLPED'02, August 12-14, 2002, Monterey, California, USA

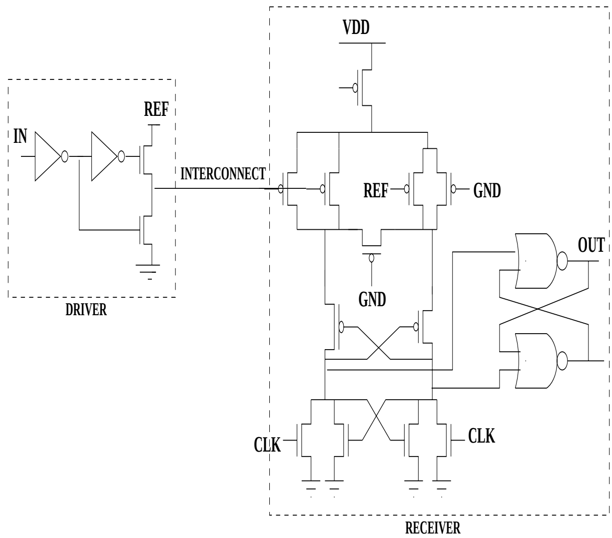

Figure 1: Pseudodifferential interconnnect

- Traditional HDLs do not provide the possibility to describe bus interfaces, which are automatically implemented by synthesis tools based on fanout knowledge, system constraints and library-derived wire capacitance information. With respect to high-level design space exploration, this methodology prevents designers from assessing the effectiveness of low-swing signaling schemes respect to traditional full-swing interconnect implementations. Instead, reducing wire voltage swing is a promising approach for on-chip power reduction [14].

  Our approach is to expose low-swing interconnects to the RTL description of the system as soft IP blocks, with pre-characterized and parameterized performance and energy consumption for design exploration.

The paper focuses on two low-swing interconnection schemes, selected among those proposed in [18] for their low energy-delay product. They are described in section II. In section III, the macromodel extraction methodology is exposed. Sections IV and V describe the details of the delay and power characterization. Finally, section VI proves the accuracy of the developed parametric VHDL models.

2. LOW-SWING SIGNALING TECHNIQUES

The low-swing signaling techniques that have been considered are pseudodifferential interconnect (PDIFF) and asymmetric source follower driver with level converter (ASDLC) [18]. Their schematics are reported in Fig. 1 and Fig. 2.

PDIFF receiver is a clocked sense amplifier followed by a static flip-flop. Although single wire per bit is used, the advantages of differential amplifiers are retained: low-input offset and good sensitivity.

As to ASDLC, the receiver is a modified voltage sense translator, and the interconnect swing is not *Vref* (as for PDIFF), but ranges from *Vref* to *Vdd-Vtn*. This scheme has a higher energy-delay product, but has been considered because of its low circuit complexity and its high operation speed.

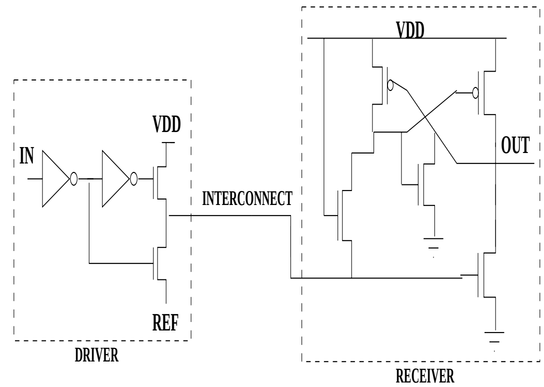

Figure 2: Asymmetric source-follower driver with level converter

3. MACROMODEL EXTRACTION METHODOLOGY

In [3] advantages and drawbacks of gate-level power simulation through library elements electrical characterization are extensively investigated.

Our approach provides more flexibility to the standard characterization technique, in that we capture the dependence of delay and power consumption of low-swing interconnects on a few key-parameters: wire length L, power supply Vdd and reference voltage $Vref$ (related to the wire voltage swing). Thus, from SPICE simulations we get 3D-matrixes, wherein each element represents the delay or energy consumption for a fixed interconnect configuration. The same experiments were run both for a metal2 wire (local interconnect) and for a metal5 global wire, considering different parameters for the different metal layers in a 0.18 um technology.

All our SPICE simulations are carried out by applying a pulse with almost vertical ramps as input signal. Although the slope of input signals should be considered in electrical level simulations, as it is closely related to the short-circuit current, our approximation does not turn out to be a major source of inaccuracy. In fact, the driver input stage consists of cascaded buffers, that can achieve large transition speed-up, thus limiting the inaccuracy of our assumption.

Further inaccuracy usually stems from neglecting the charge status of circuits internal capacitances [3]. We make the same approximation for the internal nodes of drivers and receivers, but this does not considerably impair the accuracy of our results because of the small entity of internal node capacitances.

Collected data can be used by an event-driven RTL simulator as parametric interconnect macromodels: each time a transition occurs at the driver, wire or receiver input, pre-defined lookup tables are addressed to provide timing and power information. The accessed table entries depend on the interconnect parameters set by the designer (Fig. 3).

Delay estimation is done individually on each component (driver, wire and receiver), so that their output signals can be updated in a timing accurate fashion, while power estimation is done for the whole system against transitions at the driver input. In this latter case, matrixes do exist for rising edge and falling edge transitions, as well as for quiescent periods, when little power is consumed because of the subthreshold and leakage currents.

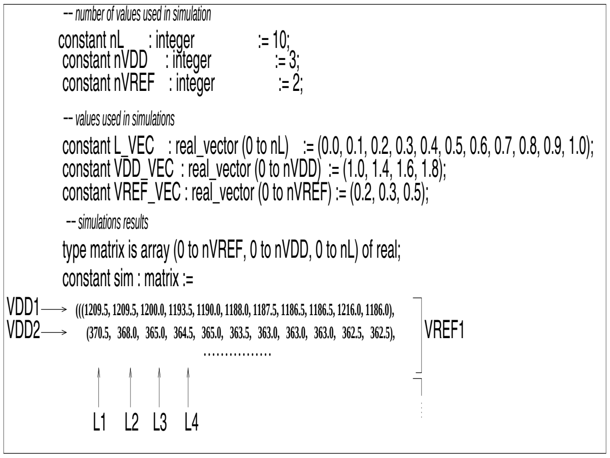

Figure 3: Example of VHDL 3D-matrix for PDIFF receiver delay estimation

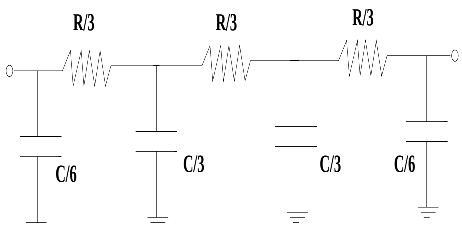

Figure 4: Interconnection model

Estimation accuracy has been derived through comparison of SPICE simulation results with VHDL estimation. Of course, as long as collected data correspond to interconnect configurations of practical interest, accuracy will be high. Unfortunately, not all configurations can be characterized at the electrical level, and hence an unavoidable discretization will ultimately impair the estimation accuracy. We deal with this problem by interpolating among data relative to annotated configurations that are closest to the selected one.

The interpolation algorithm has been chosen looking at the parabolic relations highlighted by results exposed in the following sections. So, second-order polynomial interpolation has been performed in a 3-dimensions domain. At first, SPICE simulation results have been inserted into a VHDL file and organized as packages. Each package reports data (delay or power) relative to a particular component (e.g. PDIFF driver) and for a given metal layer. This structure allows also successive integration of already collected data with new results coming from further SPICE simulations run in order to limit discretization-related inaccuracy.

When interpolation is performed, a VHDL routine uses bisection to select a 3D-submatrix from the main matrix, corresponding to the wire configurations that are closest to the configuration under test. Then the general Lagrange interpolation formula is applied many times, so to weight the simulation results with respect to *Vdd*, *L* and *Vref*.

We show that our back-annotated VHDL simulations provide estimations, both for delay and power, that differ from SPICE results by at most 5%, with negligible performance penalty on the RTL simulation speed.

	Local Interconnect	Global Interconnect
L	1mm	1.5cm
W	4λ	12λ
layer	metal2	metal5
r (Ω/mm)	296	74.1
c (fF/mm)	706.3	418.9

Table 1: Interconnection parameters

4. DELAY CHARACTERIZATION

The electrical level simulator used for schemes characterization is HSPICE 96.1. The reference technology is TSMC 0.18um, provided by MOSIS. It has 6 metal layers and λ = 0.09um.

As wires inside chips are presently dominated by RC behavior [1], rather than by transmission line behavior, we use a $\pi 3$ model for interconnects [7](Fig. 4). This is an approximation of a distributed RC model by a lumped RC network, but the error this model exhibits for delay calculation is about 3% [13].

Resistance per unit length has been derived from the MOSIS technology file, while layout parasitics extraction has been used to compute the capacitance per unit length. Both sidewall and vertical coupling capacitances have been extracted. These parameters have been computed both for a metal2 1 mm wire and for a metal5 1.5 cm wire. Collected values are shown in TABLE 1.

4.1 PDIFF scheme SPICE simulation

The original driver scheme has been enriched by cascaded inverter stages for delay optimization. Transistor sizing has been done both for metal2 and for metal5 layers (keeping *Vdd*=1.8V and *Vref*=0.5V), except for the receiver circuit, which in both cases is supposed to drive a 2fF load capacitance, corresponding to the input capacitance of a minimum area inverter.

The following simulation results are referred to a circuit consisting of driver and wire. The receiver has been omitted because its delay is independent of the input signal, but it is measured from the falling edge of the clock, so it has been independently measured. The impact of the wire parameters *L*, *Vdd* and *Vref* on the circuit delay has been reported in Fig. 5. There are three family of curves, corresponding to three different supply voltages, and inside each family curves are further differentiated according to the wire voltage swing. All curves exhibit parabolic behavior. Given the small impact of *Vref* on delay, this voltage can be used as a degree of freedom for minimizing the energy-delay product. We find *Vref*=0.3V.

4.2 ASDLC scheme SPICE simulation

A delay optimized version of the ASDLC scheme has been simulated.

The resulting dependence of the overall scheme delay (driver, wire and receiver) on the interconnect parameters exhibits the same parabolic behavior already showed by PDIFF. So, in Fig. 6 we report a comparison between delays incurred by PDIFF, ASDLC and a standard CMOS driver and receiver, that do not make use of reduced swings (*Vdd*=1.8V, *Vref*=0.3V). PDIFF turns out to be always slower than CMOS, while for small wire lengths ASDLC

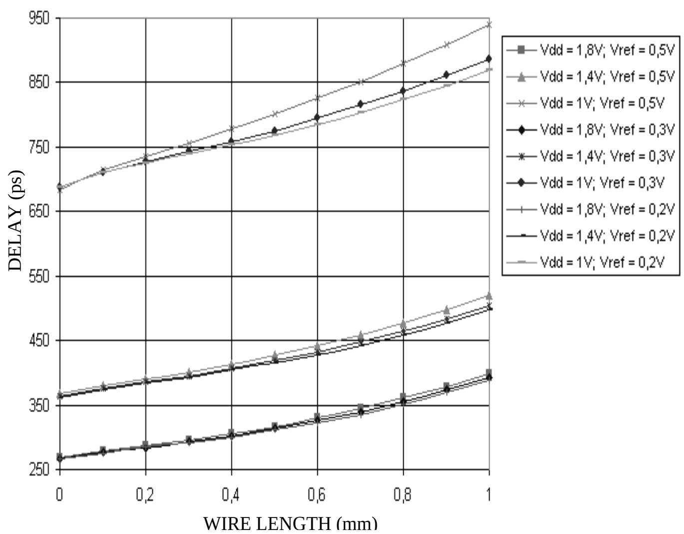

Figure 5: PDIFF scheme delay as a function of *L*, *Vdd* and *Vref*

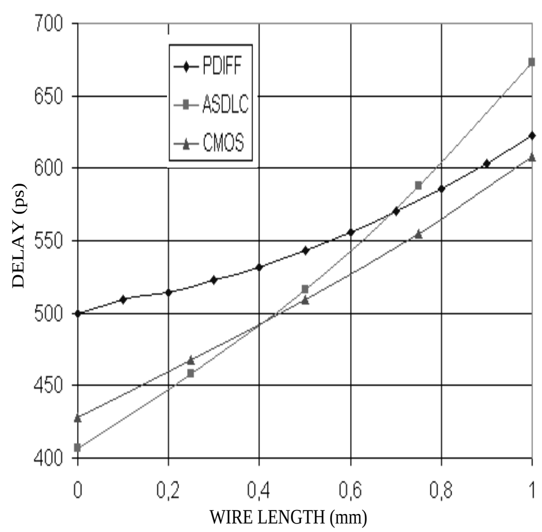

Figure 6: Comparison for metal2 wires

performs better.

From TABLE 2, derived for L=1mm, tradeoffs between the compared schemes are pointed out. The real advantage of PDIFF and ASDLC over CMOS is power consumption. PDIFF incurs less delay than ASDLC and allows the usage of very low interconnect swings, but exhibits higher receiver complexity. On the contrary, although ASDLC is a simpler scheme, it is strongly conditioned by swing constraints. In fact, *Vref* must be lower than *Vtn*, and the voltage swing can range only from *Vref* to *Vdd-Vtn*.

4.3 VHDL model for delay estimation

VHDL has already been successfully used to study the impact of VLSI scaling on microarchitectural features [2].

Based on the collected data, we have built VHDL models able to estimate delays of the considered low-swing interconnect schemes. Interconnect parameters for design space exploration are provided to the VHDL model

Parameter	PDIFF	ASDLC	CMOS
Delay: driver+wire (ps)	393	472	544
Delay: receiver (ps)	230	201	64
Total delay (ps)	623	673	608
Power (100MHZ) (μW)	106	200	434
Swing (V)	0.3	0.96	1.8
Pτ product(μW·ns)	65.6	134.6	263.8

Table 2: Comparison of PDIFF, ASDLC and CMOS performance for a 1mm metal2 wire

through the "generic" instruction. Thus, each component of the communication channel (driver, wire and receiver) is exposed to the VHDL simulation as a soft IP block, as can be seen from the following example relative to a delay estimation model for an ASDLC driver:

```
use work.delay_driver_asdlc_metal5.all;
entity DRIVER_ASDLC is
generic (
L : real := 1.0;
VDD : real := 1.8;
VREF : real := 0.5);
port (
TX_IN : in std_logic_vector (7 downto 0); – driver inputs
TX_OUT : out std_logic_vector (7 downto 0)); – driver outputs
architecture BEHAVIORAL of DRIVER_ASDLC is
shared variable delay, : real;
shared variable delay_time: time;
begin – BEHAVIORAL
DELAY_ESTIMATION (L, VDD, VREF, delay);
conversion_to_time(delay,delay_time);
TX_OUT <= TX_IN after delay_time;
end BEHAVIORAL;
```

TABLE 3 compares the results of VHDL simulations of the interconnect schemes with SPICE results. The error made by the interpolation algorithm is within 2%, but it increases for Vdd<1.4V because of the low measurement density for supply voltages lower than 1.4V. We should however consider that the circuits under test are typically used above 1.4V, in particular ASDLC.

5. ENERGY CHARACTERIZATION

The basic idea is to annotate energy consumption of the considered interconnect architectures in response to input transitions, such as rising and falling edges of the propagating input signal. Contributions of driver, wire and receiver are summed up, giving the energy drawn by a specified interconnect configuration for each input event.

Average power consumption has also been measured when the interconnect holds a logic one and zero. Thus, the VHDL model will be able to compute the simulated time between two successive transitions on the same wire, and estimate the energy consumption inbetween.

Note that for the PDIFF receiver, energy consumed as a consequence of clock transitions has been considered as well, and used to fill separate lookup tables.

Fig. 7 and Fig. 8 show energy consumed by the low-swing schemes under test in response to a rising edge of the input

Scheme	L (mm)	V_{dd} (V)	V_{ref} (V)	Driver VHDL	Wire VHDL	Receiver VHDL	Driver SPICE	Wire SPICE	Receiver SPICE
PDIFF	0.75	1.7	0.4	307	62	231	307	62	233.5
PDIFF	0.75	1.1	0.4	611.5	69	760	611.5	69.5	645
PDIFF	0.25	1.4	0.25	382	8	349	382	7.5	349.5
PDIFF	0.25	1.3	0.3	441	8	443	428.5	8	398
ASDLC	0.8	1.5	0.35	458	74	420	452.5	80	370
ASDLC	0.4	1.8	0.4	340	19	144	339	19.5	143
ASDLC	1	1.8	0.25	339	126	202	341.5	120	202
ASDLC	0.5	1.5	0.4	450	30	338	445	30	312.5

Table 3: Comparison between interpolated data from VHDL model and SPICE results (delays in ps)

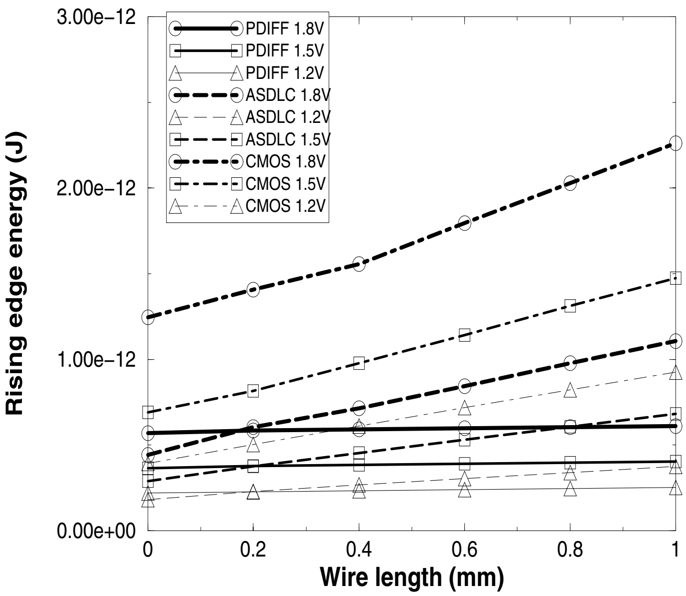

Figure 7: Rising edge energy for a metal2 wire. *Vref*=0.3V.

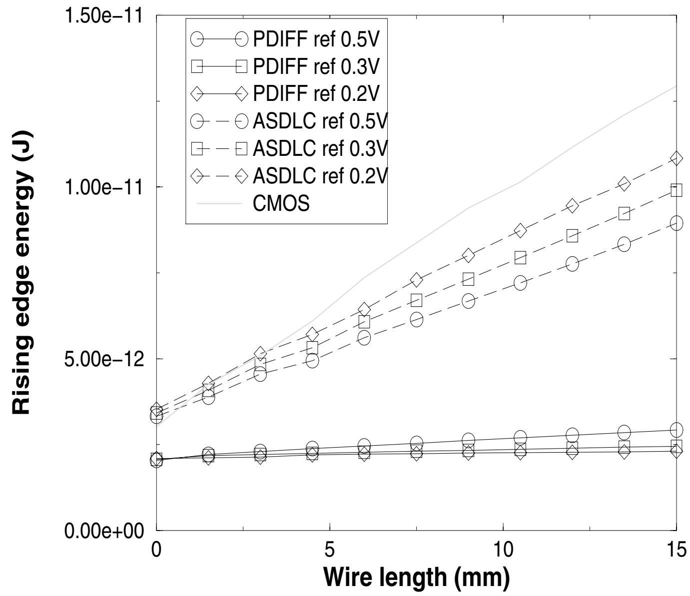

Figure 8: Rising edge energy for a metal5 wire. *Vdd*=1.8V.

signal, for metal2 and metal5 wires respectively and as a function of the interconnect length.

Metal2 experiments assume *Vref*=0.3V, and use *Vdd* as simulation parameter. CMOS traditional scheme is always the most power-hungry implementation for a given *Vdd*, while PDIFF is the less consuming one because of its lower swing. Though for very short wires the lower circuit complexity of ASDLC makes the difference, and this explains the crossing points observed on the plot.

For metal5 experiments, thinking about system buses as possible beneficiaries of this technique, we have reported rising edge related energy as a function of wire length, but using *Vref* as simulation parameter. *Vdd* is fixed at 1.8V. Note that PDIFF consumes always less energy than ASDLC, and that curves shift upwards as a consequence of a voltage swing increase.

6. VHDL POWER MODEL VALIDATION

Collected data have been embedded into a VHDL model that performs energy estimation in an event-driven fashion, as transition occurs at the interconnect driver input. The same polynomial interpolation algorithm has been used, as for delay estimation.

We have then modified the VHDL RTL description of a SPARC V8 processor called Leon [17], embedding the derived power macromodels into the source code. In practice, we assume that the read data path connecting processor caches to the external memory controller be implemented with ASDLC low-swing interconnects. Then we have run the processor bootstrap routine on top of the VHDL description, and we have traced the waveform of one bus line (about 20000 clock cycles at 50 MHz). At the end of the VHDL simulation, a routine reports the total energy consumption associated with that line, as estimated from the embedded power macromodels during application runtime.

Then we convert the traced waveform into a SPICE compliant input signal, perform electrical level simulation on the selected architecture and compare energy reports with those provided by the VHDL simulation. Results for multiple combinations of interconnect parameters are reported in Fig. 9, where the percentage error incurred by the VHDL model is shown for the ASDLC scheme.

Relative error is always within 5%, but is larger for very short as well as for very long interconnects, while for medium sized wires the accuracy is very high (below 2%). For the outer ranges of wire lengths, the main discrepancy arises

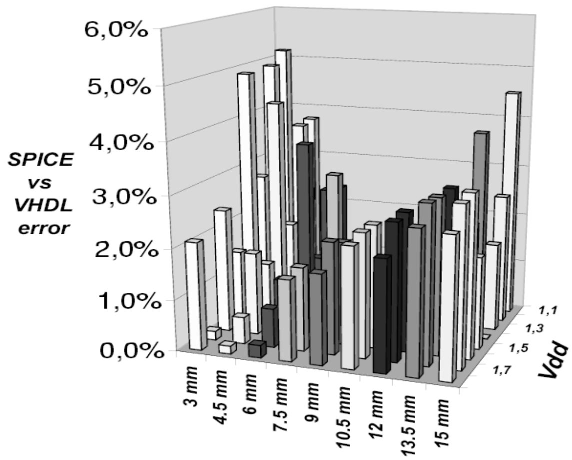

Figure 9: Percentage error between VHDL estimation and SPICE simulation for a bus line traced waveform

from energy estimation during quiescent periods, when no transitions occur. On the contrary, transition-related energy is always estimated with good accuracy.

We have also measured performance penalty incurred by the VHDL simulation when power estimation is carried out. For the Leon bootstrap routine, we noted a 2% penalty in terms of execution time, which achieves 4% when both power and delay estimations are active.

7. CONCLUSION

This work provides high level parametric VHDL simulation models for power and delay estimation of low-swing interconnect schemes. They are implemented as soft IP blocks, thus enabling design space exploration of interconnect-centric architectures at a high abstraction level. The accuracy of VHDL estimations is within 5% with respect to SPICE results, with negligible execution time overhead.

8. REFERENCES

[1] H. Bakoglu. *Circuits, Interconnections and Packaging for VLSI*. Addison-Wesley, 1990.

[2] T. Bautista and A. Nunez. “Quantitative Study of the Impact of Design and Synthesis Options on Processor Core Performance”. *IEEE Trans. on VLSI Systems*, 9(3):461–473, June 2001.

[3] A. Bogliolo, L. Benini, G. De Micheli, and B. Ricco’. “Gate-Level Power and Current Simulation of CMOS Integrated Circuits”. *IEEE Trans. on VLSI Systems*, pages 473–488, December 1997.

[4] J. Cong. “An Interconnect-Centric Design Flow for Nanometer Technologies”. *Symposium VLSI Technology Systems Applications*, pages 54–57, June 1999.

[5] J. Cong, L. He, K. Khoo, and D. Pan. “Interconnect Design for Deep Submicron ICs”. *Int. Conf. on Computer Aided Design*, pages 478–485, November 1997.

[6] J. Cong, L. He, C. Koh, and P. Madden. “Performance Optimization of VLSI Interconnect Layout”. *Integr. VLSI J.*, 21:1–94, 1996.

[7] J. Cong and Z. D. Pan. “Interconnect Performance Estimation Models for Design Planning”. *IEEE Trans. on CAD of ICs and Systems*, 20(6):739–752, June 2001.

[8] P. D. Fisher and R. Nesbitt. “The Test of Time”. *Circuits and Devices*, pages 37–44, March 1998.

[9] B. George et al. “Power Analysi and Characterization for Semiconductor Design”. *Int. Workshop Low Power Design*, pages 215–218, 1994.

[10] K. Lahiri, A. Raghunathan, and S. Dey. “Efficient Exploration of the SoC Communication Architecture Design Space”. *ICCAD-2000*, pages 424–430, November 2000.

[11] D. Liu et al. “Power Consumption Estimation in CMOS VLSI Chips”. *IEEE JSSC*, 29:663–670, June 1994.

[12] C. A. Moritz, D. Yeung, and A. Agarwal. “SimpleFit: A Framework for Analyzing Design Trade-Offs in Raw Architectures”. *IEEE Trans. on Parallel and Distributed Systems*, 12(7):730–742, July 2001.

[13] J. M. Rabaey. *Digital Integrated Circuits: A Design Perspective*. Prentice-Hall, 1996.

[14] C. Svensson. “Optimum Voltage Swing on On-Chip and Off-Chip Interconnect”. *IEEE Journal of Solid-State Circuits*, 36(7):1108–1112, July 2001.

[15] D. Sylvester and C. Hu. “Analytical Modeling and Characterization of Deep-Submicrometer Interconnect”. *Proceedings of the IEEE*, pages 634–664, May 2001.

[16] S. Takahashi, M. Edahiro, and Y. Hayashi. “Interconnect Design Strategy: Structures, Repeaters and Materials with Strategi System Performance Analysis (S2PAL) Model”. *IEEE Trans. on Electron Devices*, 48(2):239–251, February 2001.

[17] http://www.gaisler.com. Gaisler Research Website.

[18] H. Zhang, V. George, and J. M. Rabaey. “Low Swing On-Chip Signaling Techniques: Effectiveness and Robustness”. *IEEE Trans. on VLSI Systems*, 8(3):264–272, June 2000.