System Level Power Estimation of System-on-Chip Interconnects in Consideration of Transition Activity and Crosstalk

Martin Gag, Tim Wegner, and Dirk Timmermann

Institute of Applied Microelectronics and Computer Engineering,
University of Rostock,
martin.gag@uni-rostock.de,
www.networks-on-chip.com

Abstract. As technology reaches nanoscale order, interconnection systems account for the largest part of power consumption in Systems-on-Chip. Hence, an early and sufficiently accurate power estimation technique is needed for making the right design decisions.
In this paper we present a method for system-level power estimation of interconnection fabrics in Systems-on-Chip. Estimations with simple average assumptions regarding the data stream are compared against estimations considering bit level statistics in order to include low level effects like activity factors and crosstalk capacitances. By examining different data patterns and traces of a video decoding system as a realistic example, we found that the data dependent effects are not negligible influences on power consumption in the interconnection system of nanoscale chips. Due to the use of statistical data there is no degradation of simulation speed in our approach.

1 Introduction

Lowering the power consumption of microsystems is one of the main topics in chip design and technology development. Not only due to the demand of energy saving and extended run times of mobile devices but also to avoid problems concerning cooling and reliability, this challenge has to be tackled.

Shrinking and further enhancements regarding technology structures are especially lowering the dynamic power consumption and the size of transistors. As logic devices are getting less and less energy dissipative and smaller, the integration density is raised. Therefore, more interconnects between these elements are needed. The power consumption of the wires mainly remains on a certain level because they cannot be made smaller and need to be at a low distance to each other raising the capacitances even under the use of ultra low-k materials. The share of energy consumed in the interconnection system increases compared to the overall energy dissipation. In the Intel 80-core e. g. the communication system is responsible for over 28 % of the overall power budget [1]. Hence, the importance of energy consumed in the interconnection system of microchips is getting bigger.

During the design process power consumption has to be estimated in every design step to be sure to meet the constraints of every part of the system as well as the whole system. The early phases of architectural, algorithmic and system design are very important parts of the whole process. Precise high level power estimation is leading to better designs, as the high level design changes are known to have more significant effects than enhancements at lower levels.

At early design stages wire-mappings and cycle-accurate behavior mostly are not known, making system level power estimations difficult. We tackle this problem with a mixture of well accepted assumptions regarding technology parameters and statistical information that represents the characteristics of the data transmitted on-chip. For this matter, different data patterns are evaluated to get significant statistics of transition probabilities and crosstalk effects. The resulting statistical data is provided to a power model. This mixture of high level information and low level assumptions will facilitate more accurate power estimation than just relying on high level design information.

In the following section this paper is related to the state of the art. Then the used power model is described. Our simulations are explained and the results are discussed before the paper is ended by a short conclusion.

2 Related Work

System level power estimation is already recognized as an important aspect in the field of chip design and system simulation. For design space exploration of Networks-on-Chip (NoCs) Kahng *et al.* give a high level power model of routers and links called Orion 2.0 [2]. This work is based on the Predictive Technology Model (PTM) [3] and calculations of capacitances by Wang *et al.* [4].

The inclusion of low level power models in system level NoC simulation is part of the work of Xi *et al.* [5]. Transition activity was included in their simulation framework, which is crucial for the correct treatment when transition encoding is utilized [6–9]. Nevertheless, no crosstalk effects were included in their simulation framework. This could be fatal as influences of coupling capacitances on on-chip buses are not negligible. Sotiriadis *et al.* derived a new low level bus model to take such deep submicron effects into account [10].

There is many work on so called crosstalk avoidance codes [11–14] and even the combination of transition and crosstalk avoidance [15] that would benefit from a system level power estimation technique respecting actual transition counts and cross coupling effects.

Using signal statistics to estimate transition activity and even crosstalk [16] is considered to claim many resources during simulation. In [17] the utilization of word level statistics was proposed to be a solution. In this paper we will show, that even bit level statistics are suitable to enhance the high level power estimations of on-chip interconnects at no simulation performance costs.

3 Modeling of Dynamic Power Dissipation on Links

The power consumed by communicating links can be divided in static and dynamic dissipation. Here we want to concentrate on the dynamic power dissipation because the static part is not influenced by the transmitted data. The well known formula

$$P_{dyn} = \frac{1}{2} \cdot a \cdot f \cdot V^2 \cdot C \tag{1}$$

where a is the transition probability, f the frequency, V the operating voltage and C the switched load capacitance, represents the dynamic power model of every logic element in CMOS systems. In the case of wires, energy consumption originates from charging ground and cross coupling capacitances. In general, capacitances to the ground and top plates are constant. The coupling capacitances are created by the left and right neighbors of a wire, which are parallel wires building a bus in most cases. The signal changes on those neighboring wires affect the effective capacitance seen by the driver through capacitive coupling. This can be considered a special case of the Miller Effect.

The calculation of the effective capacitance is a combination of ground and coupling capacitance:

$$C_{eff} = C_g + \sigma \cdot C_c \tag{2}$$

Where σ in this combination depends on switching directions of the right and left neighbor of the wire and is called the Miller Coupling Factor (MCF). There are different possible combinations which can raise but also lower the value of the effective capacitance compared to a static MCF, which is 2 on average (Tab. 1). The MCF can be calculated using the following equation, where v_i^f is one when the final value of the voltage on the i-th line is high and zero if it is low. v_i^i stands for the initial value of that line.

$$\sigma = [-1, 2, -1] \cdot \begin{bmatrix} v_{i-1}^f - v_{i-1}^i \\ v_i^f - v_i^i \\ v_{i+1}^f - v_{i+1}^i \end{bmatrix} \tag{3}$$

The resulting dynamic power consumption can be calculated with the resulting Eq. (4), where V is the initial or final voltage.

$$P_{dyn} = a \cdot f \cdot V_i^f \cdot [-\lambda, 1 + 2\lambda, -\lambda] \cdot \begin{bmatrix} V_{i-1}^f - V_{i-1}^i \\ V_i^f - V_i^i \\ V_{i+1}^f - V_{i+1}^i \end{bmatrix} \cdot C_g \tag{4}$$

Similar to the Predictive Technology Model (PTM) [3] and Orion 2.0 [2], we are using the models of Wong *et al.* [4] to calculate the technology dependent values of ground and coupling capacitances. Together with the gathered MCF these values are used for dynamic power calculation. In Addition, a component of static power is added to include leakage like it is done in Orion 2.0.

Table 1. Possible Miller Coupling Factors of a wire (i) switching from 0 to 1

i+1 \ i-1	0→0	0→1	1→0	1→1
0→0	2	1	3	2
0→1	1	0	2	1
1→0	3	2	4	3
1→1	2	1	3	2

4 Bit Level Statistics

To get the most exact values for effective coupling capacitances and transition counts, it is necessary to evaluate every bit that traverses the data bus in the system and analyze its correlation to the previous bit of this position. This is possible for all signals in gate level simulations, because all signals are known and their probable mapping to wires can be estimated. Even at system level this is possible for links connecting main modules (e. g. a bus in SoCs or the interconnection network in NoCs), if a few assumptions concerning bus mappings are made.

The evaluation of every bit transmitted through the communication system takes time during the simulation process. This may reverse speed gain achieved through high level abstractions if done during system level simulations. However, we propose to use signal statistics to account for transition activity and crosstalk effects on links. The necessary signal statistics can be obtained from a sample of data characterizing traffic on the actual link before the system simulation starts.

The time required to create offline statistics depends on the evaluated system and signal parameters but usually should be much lower than the time that is taken to process the whole real data stream. The acquisition of those signal data can be achieved by deploying cycle accurate system models or architectural models and exploiting knowledge of algorithms used in the system modules. It has to be known if the data is mostly random like compressed data or if there are inter-word correlations that are often found in uncompressed data. Of course signal traces of lower level models could be used as well, if they are available.

In our experiments we generally used two ways to gather the bit level statistics of the data. In the first method stream based evaluation software is used to examine the characteristics of general data. At first, the incoming data from a file is divided into chunks corresponding to the expected word width on the later bus structure. Then transitions between two successive words are counted and the MCF is calculated for every bit position in the data word in order to consider crosstalk. In the middle of the bus the needed energy is affected by two aggressors, while the victim lines at the fringes have only one aggressor (Fig. 1). If the stream comes to an end, the arithmetic average of transitions and MCFs of all bit positions are determined.

The second method is based on the interpretation of signal traces in Value Change Dump (VCD) format. A gate level simulation of a hardware design is used to generate the trace files. Our software extracts the interesting signals out of the signal dump. That would be the signals that will run between main modules and are possible candidates for relatively long wires i. e. claiming high capacitances in the data bus. These signals are analyzed as it is done in the stream based evaluation.

In our simulations we used the first method for general investigations of bit level statistics of common data. The Second approach was used to evaluate our estimation technique for an implemented SoC.

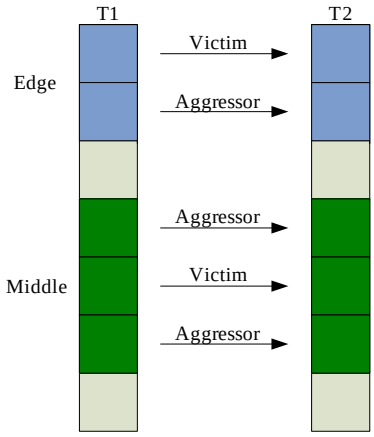

Fig. 1. Crosstalk estimation in two successive cycles at fringes and in the middle of a bus [16]

5 Simulation Results

To estimate the accuracy gain concerning power estimation with bit level statistics, different types of data were analyzed by our stream based program. As representatives for compressed data JPEG- and H.264 compressed image and video files as well as MPEG-Layer 3 encoded audio files were used. As a group of uncompressed data decoded image, audio, video and text files were used. A more practical data stream with a mixture of compressed and uncompressed data is represented by a network stream while browsing a webpage. Characteristic content of such a stream dump are uncompressed packet headers and a compressed HTML-text plus a few compressed graphics files. For comparison, we included a data pattern that maximizes crosstalk and transition probability to 100 % representing the worst case of data patterns.

To get indications for the applicability of using bit level statistics, the model of an application was investigated. The H.264 decoder [18] was simulated at register transfer level to extract signal dumps of the global connections of functional blocks like memories, entropy decoder, prediction unit etc. Those trace dumps were analyzed to extract the bit level signal statistics.

5.1 Simulation Accuracy

Traditional data independent power estimation considers a transition probability of 50 %. In Fig. 2 the results of our system level power estimation compared to a traditional one are shown. In addition, we determined the estimated power values with the actual gathered transition probability without calculating crosstalk effects to rule out the influence of the MCF.

As expected, the highly compressed data mostly consists of uncorrelated patterns. This corresponds to random data. The resulting power estimation with consideration of bit level statistics differs hardly from the traditional approach of assuming 50 % transition probability. This applies for random data as well as compressed images (JPEG), videos (H.264) and audio (MP3). The estimation error in respect to the most accurate method of using the real transition count and the crosstalk calculation shows relatively low values of up to 7.1 % (Tab. 2).

The most accurate calculation with respecting the crosstalk capacitances including the MCF shows a little bit lower power values even in the case of completely random data. That is because the fringe capacitances, which are considered to be very much lower than the coupling capacitances, were included only in this estimation mode where the deep submicron bus model was used. The other two estimation modes only assume coupling capacitances on both sides of the wire even at the fringes of the bus.

The uncompressed data shows higher autocorrelation. This results in lower power values due to fewer transitions on the wires in cases of uncompressed video as well as images (BMP), audio (WAVE) and text files. The effect is due to the most significant bits are switched more infrequently compared to the less significant ones. In these cases it is very important to choose the right word width to exploit the data characteristics. This decision is mostly implied by the application but information about this aspect can also be provided by our data analysis software. As Fig. 3 shows, transition probability of uncompressed data has a dependency on the used word width. The optimal width for uncompressed image and video data is 24 bit because typically there is 3 byte of color information per pixel in such a data structure. Our audio example consists of a 16 bit stereo wave file and shows an optimal word width of 32 bit. The text file would be optimally segmented in every multiple of 8 bit because ASCII encoding is used, which utilizes 1 byte of data per character.

The highest difference between the power estimation values was reached by uncompressed video, which consists of a scene of an animated comic in 1080p format. The method of considering realistic transition counts and calculating the crosstalk activity differs about 432.5 % from the estimation with a simple

assumption of 50 % transition activity. Just considering transitions and ignoring the MCFs of crosstalk shows a deviation of only 2.2 %.

To get more realistic data patterns a SoC was examined. This hardware design implements a H.264 decoder and is divided into functional blocks. The signals connecting those modules are considered to be intermediate wires that are long enough to produce high capacitances and make a remarkable contribution to the overall energy consumption. The extracted signal statistics lead to power estimations that are significantly lower (deviation of 84.6 %) than assuming an average transition rate of 50 %. Therefore, the average transition rates between the main modules of the SoC are more in the regions of uncompressed data than being similar to the compressed data. This leads to a better power estimation when using real signal statistics.

As simulation results show, the accuracy of the system level power estimation is raised by our approach of using signal statistics to predict transition probability. By doing so, the error of up to 432.5 % in simulations using a general assumption of 50 % transition probability is avoided. The amount of such estimation errors depends on the data itself and is higher the less compressed the data is. As our worst case data sample shows, the simple estimation could be too low by about 64.9 % in cases of practical data it is consistently too high. Crosstalk effects are not that much important to the power estimation as can be seen by the little deviations of the method using real transition counts without the application of crosstalk estimation. That is because the average MCF is mostly met by the data characteristics.

Table 2. Relative deviation of energy estimation techniques related to the method of considering real transition rate and crosstalk

method using	real tr. rate	50 % tr. rate
worst case	0,313	0,649
random	0,026	0,027
JPEG	0,007	0,022
H.264	0,028	0,043
MP3	0,013	0,071
web surfing	0,032	0,187
text (ASCII)	0,052	0,520
BMP	0,016	1,266
video unenc.	0,022	4,325
WAVE	0,021	0,422
H.264 decoder SoC	0,059	0,846

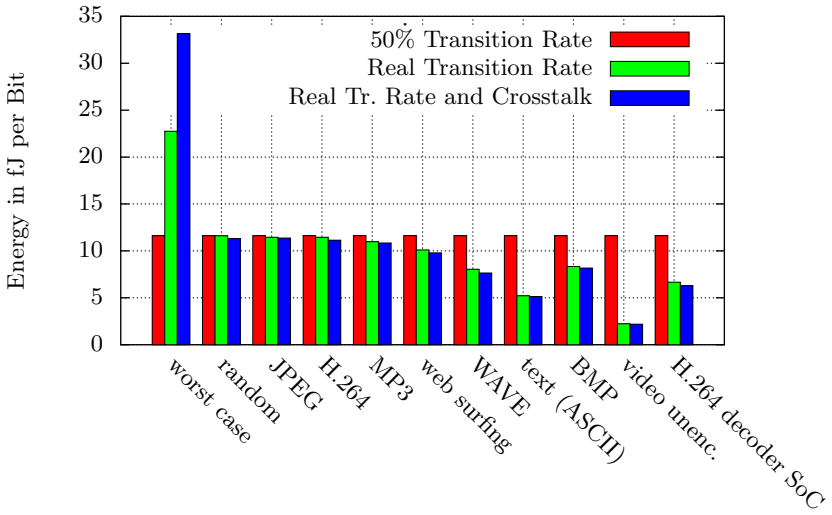

Fig. 2. Estimated average energy for transmitting one bit on an intermediate wire of 200 µm length (single spaced) in 65 nm technology for different data files evaluated by 3 different estimation techniques

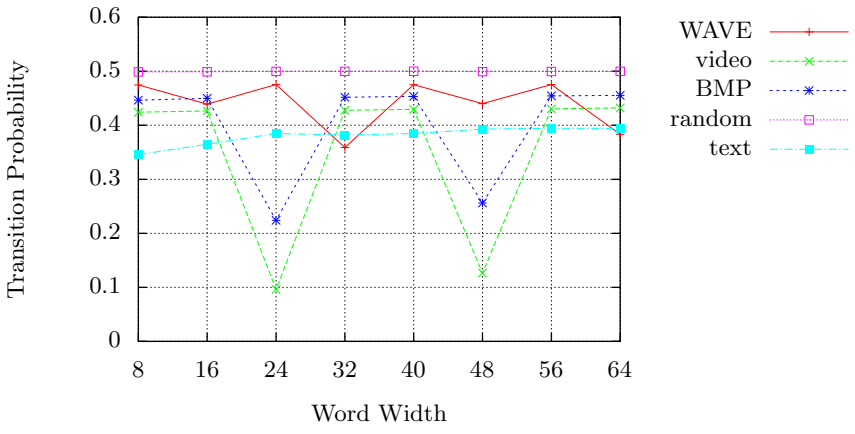

Fig. 3. Transition probability using different word width for transmission

5.2 Simulation Performance

The method of using signal statistics reduces to calculating the power equation during simulation. In this step the general time complexity of the simulation is not affected, so there is no speed penalty and system level power estimation finishes in parts of a second.

The statistical data of possible signals must be gathered prior to the simulation. This step takes time and depends on the method of statistics acquisition. In our experiment with general data files the data stream analysis lasts up to 5 seconds when processing up to 100 MB on an Intel Core2Duo workstation PC. It has to be mentioned that we did not optimize for runtime, as we assume to gather the statistics offline and then simulate high level models with few design possibilities in seconds.

6 Conclusion

In this paper we showed how wrong system level power estimation could be if not aware of the data that will pass the interconnection system between the main modules. Our proposed technique takes bit level statistical data of a possible data stream in the system and makes it available to commonly accepted low level power models of interconnection links. By using this approach the actual transition activity of the interconnections and low level phenomena like cross coupling effects can be considered. It turns out that, if mainly uncompressed data is transmitted between the system components, the deviations between the power estimations are not negligible. In consequence, the consideration of bit level statistics promises to facilitate more accurate estimations. As the investigation on a realistic system showed, our technique was by 84.6 % more correct then if a general transition activity of 50 % would be assumed.

The crosstalk feature of our power estimation technique showed no mentionable effects when realistic data was used. The difference to the method considering real transition activities was 6.5 %. As we plan to integrate this work into a bigger simulation kit with different link level encoding features to exploit transition and crosstalk avoidance codes, the feature of cross coupling estimation is going to be essential for correct power estimations.

References

1. Vangal, S., Howard, J., Ruhl, G., Dighe, S., Others: An 80-tile sub-100-w teraflops processor in 65-nm cmos. IEEE Journal of Solid-State Circuits **43**(1) (2008) 29–41
2. Kahng, A., Li, B., Peh, L., Samadi, K.: Orion 2.0: A fast and accurate NoC power and area model for early-stage design space exploration. In: Design, Automation, and Test in Europe. (2009) 423–428
3. `http://ptm.asu.edu/`: Predictive Technology Model
4. Wong, S.C., Lee, G.Y., Ma, D.J.: Modeling of Interconnect Capacitance, Delay, and Crosstalk in VLSI. IEEE Transactions on Semiconductor Manufacturing **13** (Februar 2000) 108–111

5. Xi, J., Zhong, P.: A System-level Network-on-Chip Simulation Framework Integrated with Low-level Analytical Models. 2006 International Conference on Computer Design (Oktober 2006) 383–388
6. Kretzschmar, C., Siegmund, R., Müller, D.: Adaptive bus encoding technique for switching activity reduced data transfer over wide system buses. Integrated Circuit Design (2000) 66–75
7. Sotiriadis, P., Chandrakasan, A.: Bus energy minimization by transition pattern coding (TPC) in deep sub-micron technologies. In: Proceedings of the 2000 IEEE/ACM international conference on Computer-aided design, IEEE Press (2000) 322–328
8. Ramprasad, S., Shanbhag, N., Hajj, I.: A coding framework for low-power address and data busses. IEEE Transactions on Very Large Scale Integration (VLSI) Systems **7**(2) (1999) 212–221
9. Benini, L., Micheli, G., Macii, E., Sciuto, D., Silvano, C.: Address bus encoding techniques for system-level power optimization. In: Design, Automation, and Test in Europe, Springer (1998) 275–289
10. Sotiriadis, P.P., Chandrakasan, A.: A Bus Energy Model For Deep Sub-Micron Technology. IEEE Transactions on Very Large Scale Integration (VLSI) Systems **10** (2002) 341–350
11. Pande, P., Ganguly, a., Zhu, H., Grecu, C.: Energy reduction through crosstalk avoidance coding in networks on chip. Journal of Systems Architecture **54**(3-4) (2008) 441–451
12. Rahaman, M., Chowdhury, M.: Crosstalk Avoidance and Error-Correction Coding for Coupled RLC Interconnects. Crosstalk (2009) 141–144
13. Duan, C., Cordero Calle, V.H., Khatri, S.P.: Efficient On-Chip Crosstalk Avoidance CODEC Design. IEEE Transactions on Very Large Scale Integration (VLSI) Systems **17**(4) (April 2009) 551–560
14. Sankaran, H., Katkoori, S.: On-chip dynamic worst-case crosstalk pattern detection and elimination for bus-based macro-cell designs. 2009 10th International Symposium on Quality of Electronic Design (März 2009) 33–39
15. Palesi, M., Fazzino, F., Ascia, G., Catania, V.: Data Encoding for Low-Power in Wormhole-Switched Networks-on-Chip. 2009 12th Euromicro Conference on Digital System Design, Architectures, Methods and Tools (2009) 119–126
16. Gupta, S., Katkoori, S.: Intra-bus crosstalk estimation using word-level statistics. 17th International Conference on VLSI Design. Proceedings. (2004) 449–454
17. Ramprasad, S., Shanbhag, N., Hajj, I.: Analytical estimation of transition activity from word-level signal statistics. Proceedings of the 34th **16**(7) (1997) 718–733
18. Fleming, K., Dave, C., Arvind, N., Raghavan, G., Jamey, M.: H. 264 Decoder: A Case Study in Multiple Design Points. In: 6th ACM/IEEE International Conference on Formal Methods and Models for Co-Design. MEMOCODE. (2008) 165–174