Embedded DRAM (eDRAM) Power-Energy Estimation for System-on-a-Chip (SoC) Applications

Yong-Ha Park, Jeonghoon Kook* and Hoi-Jun Yoo
Dept. of EE, Korea Advanced Institute of Science and Technology (KAIST), Taejon, Korea
** Memory R&D Division, Hynix Semiconductor Inc., Ichon, Kyoungki-do, Korea*
yhpark@eeinfo.kaist.ac.kr jeonghoon.kook@hynix.com hjyoo@ee.kaist.ac.kr

Abstract

Embedded DRAM (eDRAM) power-energy estimation is presented for system-on-a-chip (SOC) applications. The main feature is the signal swing based analytic (SSBA) model, which improves the accuracy of the conventional SRAM power-energy models. The SSBA model combined with the high-level memory access statistics provides a fast and accurate system level power-energy estimation of eDRAM. The power-energy estimation using SSBA model shows 95% accuracy compared with the transistor level power simulation results for three fabricated eDRAMs.

1. Introduction

Although the low power eDRAM becomes important as it is widely used in SoC design, the detail study on power estimation of eDRAM has been seldom addressed. The researches for memory power-energy estimation have been focused on cache memory or SRAM, which have significant impacts on the power consumption of overall memory system [1]~[6]. In off-chip design, the power-energy consumption of DRAM can be ignored because the power-energy consumption of off-chip interconnection is larger one or two orders of magnitude than that of standalone DRAM itself [2]. In SoC design, the situation is changed. The power-energy consumption of on-chip interconnection is reduced to the same order of that of eDRAM because on-chip I/O capacitance is tremendously reduced. In order to obtain the power-energy efficient eDRAM architecture as early as possible, fast-accurate eDRAM power estimation is essential, especially for the battery driven applications. For efficient eDRAM power energy estimation, the previous SRAM power-energy models are investigated and revised for fast-accurate eDRAM power-energy estimation in SoC applications.

In section 2, the previous SRAM power-energy models are investigated and the requirements of eDRAM power-energy model are described. In section 3, the eDRAM power-energy estimation methodology, which combines the high-level memory access statistics and the eDRAM power-energy model, is presented. In section 4, the signal swing-based analytic (SSBA) eDRAM power-energy model is introduced. In section 5, the proposed model is verified and discussed with the transistor level power simulation results based on three pre-fabricated eDRAMs. Finally, the conclusion is summarized in section 6.

2. Previous Works and Requirements

There exists an analytic energy dissipation model for low power SRAM caches [1]. This model utilized memory traffic statistics, cache architectural factors and process technology factors in order to determine the energy efficient cache architecture. But this model fails to take into account of the exact signal swing so that it can't minimize the estimation error, as reported in their paper. In paper [2], the energy consumption modeling and the optimization for the general purpose SRAM were performed by the combination of the analytic model and the simulation model, which characterized the full swing of digital signal and the arbitrary swing of analog signal, respectively. The intuitive on-chip cache energy consumption model was developed only for cache hit [3]. However, energy consumption for both cache hit and cache miss must be precisely considered for hierarchical memory system, because the low cache hit rates lead to more frequent off-chip main memory access which consumes more energy than on-chip cache access [4]. This model utilized the cache-to-main memory power consumption ratio, which had the difference of one or two orders of magnitude, because power consumption by off-chip I/O interconnection usually dominated that by the DRAM itself. The performance and energy consumption according to the changes of cache size, line size, set associativity, tiling and the off-chip data organization were presented in [5]. They revealed that, for the memory exploration in a low power system, it was insufficient to consider only the cache size and the miss rate. This is because while the miss rate reduces with the increase of cache size, the energy consumption does not always decrease. Also, their analysis showed that the largest performance enhancement was obtained by the off-chip

data organization because it can reduce the number of conflict misses.

In our work, the requirements that should be satisfied for eDRAM power-energy estimation are described and previous findings are revised for the possible application to eDRAM. In SoC design, eDRAM power-energy estimation has to consider both the high-level memory access characteristics and the exact analog swing characteristics of eDRAM internal operation because of the following reasons.

First, eDRAM power-energy consumption strongly depends on the high-level memory access characteristics governed by both eDRAM architecture and other embedded components' architectures. For the example of the hierarchical memory system, the cache hit/miss ratio affecting the number of eDRAM access shows strong dependency on SRAM cache architecture or DRAM main memory data organization [5]. The other example of the non-cacheable memory system such as graphics applications, the horizontal mapping, which assigns all pixel data on a horizontal line on a screen into the same memory row address, is suitable for 2-dimemsional graphics [7]. The tile mapping, which assigns all pixel data inside a rectangular area on a screen into the same memory row address, is optimized for 3-dimemsional graphics [7]. If the horizontal mapping is used for 3-dimemsional graphics, it suffers from excessive power consumption and performance degradation because row address changes, which have different power-energy consumption characteristics from column address changes, are much more frequent than that for the tile mapping. Thus, the high-level memory access statistics should be taken into consideration in order to estimate exact system level power consumption of eDRAM.

Second, the simple but accurate analog swing model is essential for accurate eDRAM power-energy estimation. The analytic SRAM power-energy model suffers from the estimation error because of its inaccurate modeling for the analog signal [1]. Of course, the mixed SRAM power-energy model combining the analytical model for a digital signal and the simulation model for an analog-type signal can achieve more accurate results compared with the complete analytical model [2]. But the detail circuit simulation and the physical layout for the precise results need long verification time and it also has the percentage error ranged from 1% to 32% according to applied SRAM architecture [2]. The conventional power-energy models of SRAM devices trade off between verification time and accuracy, especially for analog signals. However, in DRAM operation, analog signals are more widely used than in SRAM operation; for example, various pre-charge voltage levels, small swing operation, differential signaling, and equalization scheme for bitline (BL) or data bus (DB) line. This means that the trade-off between verification accuracy and verification time causes more serious problem for early estimation than that of SRAM.

Thus, eDRAM power-energy estimation should take into consideration of both the high-level memory access characteristics and the simple-accurate power-energy model for analog signal swing.

3. eDRAM Power-Energy Consumption

For any given memory architecture, eDRAM power-energy consumption depends on five key factors such as the number of eDRAM access, the number of row activation, the number of column activation, the data transition rate and the signal swing characteristics. The number of the eDRAM access is determined by cache miss rate if cache memory is used, or it is the same as the total number of memory access if the memory system is non-cacheable. The number of row activation and the number of column activation should be separately considered because each operation shows different power-energy consumption characteristics coming from the difference between random access and burst access. The data transition rate affects the power-energy consumption of I/O interface between eDRAM and embedded logic blocks as well as internal signal interface inside an eDRAM. The signal swing characteristics especially in analog signal also should be considered for the exact power-energy model. In this section, the number of memory access, row activation and column activation will be combined for system level eDRAM power-energy estimation. The rest two factors for the SSBA model will be explained in the next section.

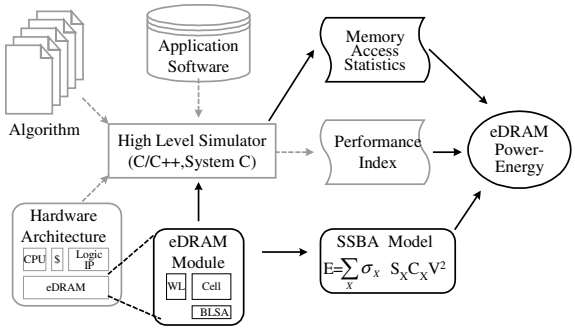

Fig.1 Hardware/software co-simulation for eDRAM power-energy estimation (gray line for conventional method and dark line for the proposed method)

A hardware-software co-simulation usually provides the performance index such as execution cycles, cache hit/miss rate, bus utilization and so on. In this study, we would like to include memory access statistics to estimate the system level eDRAM power-energy as shown in Fig. 1.

The detail eDRAM architectures determined by address bus type, I/O interface scheme, bank structure, row path and column path are described in the eDRAM module of Fig.1. While the application software and the correspond algorithm are executed in the high-level simulator combined with the hardware architecture and the eDRAM module as shown in Fig. 1, the memory statistics are provided for the system level eDRAM energy estimation. The assistance of the SSBA eDRAM power-energy model improves the accuracy of the system level energy estimation results. Not only the memory access statistics and the SSBA power-energy model, but also the total execution cycles from the performance index are finally combined together in order to get system level eDRAM power estimation as shown in Fig. 1.

$$E_{eDRAM} = N_{ROW} \times E_{ROW} + N_{COL} \times E_{COL} \quad \text{Eq.1 (a)}$$

$$\cong N_{ROW} \times (E_{ROW} + BL_{AVG} \times E_{COL}) \quad \text{Eq.1 (b)}$$

$$\cong N_{MEM} \times M_{C\_MISS} \times (E_{ROW} + L_{C\_LINE} \times E_{COL}) \quad \text{Eq.1 (c)}$$

$$P_{eDRAM} = E_{eDRAM} \times N_{EXE\_CYCLE} / f \quad \text{Eq.1 (d)}$$

E_{eDRAM}-Overall energy consumption of embedded DRAM
N_{ROW}-The number of row activation
N_{COL}-The number of column activation
N_{MEM}-The number of eDRAM access
N_{EXE_CYCLE}-The number of total execution cycles
E_{ROW}-Energy consumption of eDRAM row activation
E_{COL}-Energy consumption of eDRAM column activation
BL_{AVG}-Average burst length per row activation
M_{C_MISS}-Cache miss rate
L_{C_LINE}-Cache line size
f-Clock frequency

$$E_{eDRAM} = N_{ROW} \times (E_{WL} + E_{BL}) + N_{COL\_R} \times E_{DB\_R} + N_{COL\_W} \times E_{DB\_W} + N_{COL} \times E_{IO} \quad \text{Eq. 2}$$

Three kinds of methods as described in Eq. 1(a) ~ (c) are allowed to determine the eDRAM energy consumption. The general form of Eq. 1 (a) is always valid for any kind of memory system. In a synchronous memory system, Eq. 1(a) can be represented as Eq. 1(b) using the average burst length (BL_{AVG}) per row activation. If a cache memory is used, Eq. 1(c) is preferred since the performance index of cache memory is combined to determine eDRAM energy consumption. eDRAM power consumption can be derived from the power-energy relation as defined in Eq. 1(d) using the execution cycle (N_{EXE_CYCLE}) and the clock frequency (f).

System level calculation of power-energy consumption in Eq. 1 can be defined in more detail as shown in Eq. 2. E_{ROW} consists of E_{WL} and E_{BL}, and E_{COL} is composed of E_{DB_W}, E_{DB_R} and E_{IO}. Where E_{WL}, E_{BL}, E_{DB_W}, E_{DB_R} and E_{IO} are energy consumption by the eDRAM internal operations such as word line (WL) activation, bitline (BL) sensing, data bus (DB) line operation by write command (DB_W), DB line operation by read command (DB_R) and input/output (IO) line operation, respectively. From Eq. 2, we can estimate overall eDRAM power-energy consumption 2~3 orders of magnitude faster than time-consuming transistor level simulation. Furthermore, the variations of eDRAM power-energy consumption can be investigated according to the architectural or algorithmic changes of the embedded modules, not limited in eDRAM. These fast analytical results help the power-energy optimization process without the design detail. The assistance of the SSBA eDRAM power-energy model improves the accuracy of the eDRAM power-energy analysis. The detail SSBA eDRAM power-energy model of E_{WL}, E_{BL}, E_{DB_W}, E_{DB_R} and E_{IO} will be described in the following sections.

4. Signal swing based analytical (SSBA) eDRAM power-energy model

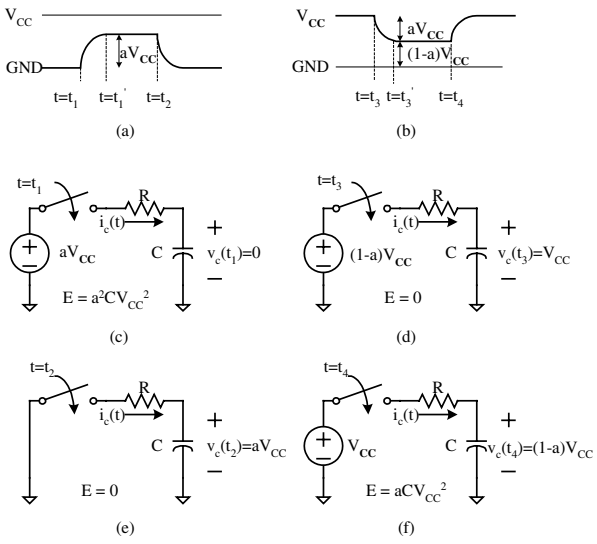

Fig.2 Arbitrary voltage swings and their equivalent models for the power-energy calculation provided from the power supply.

When the voltage level is switched between $a \cdot V_{CC}$ and GND or between $(1-a) \cdot V_{CC}$ and V_{CC} shown Fig. 2 (a) and (b), the total energy amount provided from the power supply is $a^2 \cdot CV_{CC}^2$ or $a \cdot CV_{CC}^2$, respectively. Where 'a' is smaller than 1 in the case of small swing and equal to 1 in the case of full V_{CC} swing.

When the signal is discharged from the initial voltage level as shown in Fig. 2 (d) and (e), the energy amount provided from the power supply is zero. This is because

the energy stored in the capacitance is dissipated either in the distributed parasitic resistance R or in other power supply. The energy amount of $a^2 \cdot CV_{CC}^2$ and $a \cdot CV_{CC}^2$ is provided from the power supply only when the signals are charged from the initial voltage level as shown in Fig. 2 (c) and (f), respectively. Some of the energy provided from the power supply is dissipated in the distributed parasitic R. The other of them is stored in the load capacitance, C, and dissipated through R if the load capacitance is discharged. However, in the digital signal with full V_{CC} swing, the energy amount provided from the power supply is always CV_{CC}^2 since a=1.

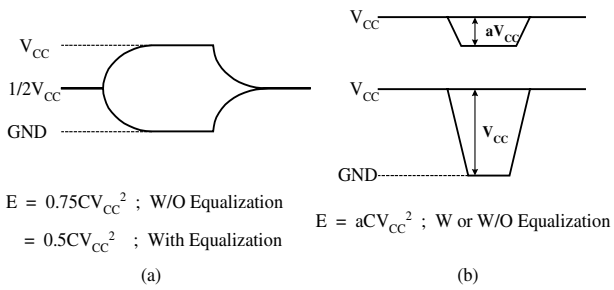

Fig.3 Energy calculation for (a) half VCC pre-charged full swing signal and (b) full VCC pre-charged arbitrary swing.

In DRAM design, differential signals with the same pre-charged voltage level are frequently used as shown in Fig. 3. Their energy amount provided from the power supply is summarized according to the pre-charged voltage level as shown in Fig. 3 (a) and (b), assumed that each signal drives the same load capacitance C. When the differential signals pre-charged with the voltage level of half V_{CC} are evaluated to GND or V_{CC} and they return to the initial voltage of half V_{CC} like Fig. 3 (a), the energy amount provided from the power supply is $0.75CV_{CC}^2$. If the charge-recycle is utilized by the equalization technique, this value is reduced to $0.5CV_{CC}^2$, because two same capacitances storing the opposite voltage level of V_{CC} and GND are naturally equalized to half V_{CC} without any energy supply from the power supply. When the only one of the differential signals, pre-charged with the voltage level of V_{CC}, is discharged to arbitrary voltage level and returns to the initial voltage level of V_{CC} (the remind signal is maintained to the voltage level of V_{CC}) like Fig. 3 (b), the energy amount from power supply is always $a \cdot CV_{CC}^2$, regardless of $0<a\leq1$, even with equalization. This is because the charge-recycle can reduce the signal swing to half but the load capacitance to be driven to the V_{CC} level increases to a double.

Overall eDRAM power-energy consumption is dominantly determined by power-energy consumption of the large capacitive load similar to the previous SRAM power-energy models [1][2][6]. This is because about 85% of DRAM power-energy is dissipated by the capacitance loads during the operation of word line (WL), high voltage supply line (RX), bit line (BL), sensing line (/S), restoring line (RTO), data bus line (DB) and I/O line (IO) as shown in Fig. 4 [10]. A timing control signal or combinational logic consumes only less than 15% of total energy consumption, which can be easily estimated by estimation methods in [6][10]. In this paper, the general eDRAM architecture as shown in Fig. 4 is considered for the SSBA power-energy model. Each bank is assumed to contain M-BL pairs, M-SAs, K-subword drivers, L-DB lines pairs, and L-IOs interface with other embedded components. Eq. 3 (a) ~(d) summarize the SSBA eDRAM power-energy models of WL, BL, DB and IO line. S_{BL}, S_{DB_R}, S_{DB_W} and S_{IO} are the signal swing coefficient for the corresponding swing characteristics as described in previous two paragraphs.

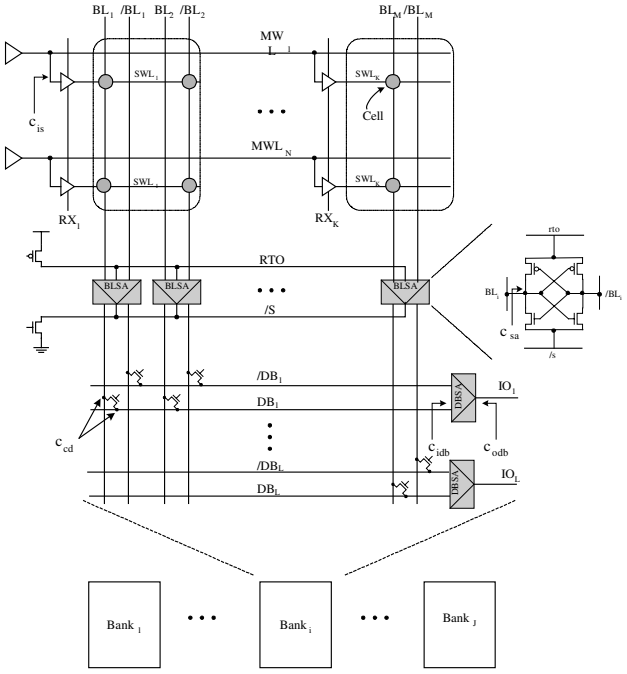

Fig. 4 eDRAM architecture for power-energy model

After one WL is activated by high voltage level of V_{PP}, half V_{CC} pre-charged full V_{CC} differential swing like Fig. 3 (a) take places on M BL pairs, signals /S and RTO. Here, /S and RTO signal are assumed to have the same capacitance as C_{SRTO} defined in Eq. 4(c) which includes the line capacitance and the junction capacitance at the corresponding segment of bitline sense amp (BLSA) connection nodes, rto and /s, as shown in Fig. 4. Power-energy consumption by the BLSA isolation logic (C_{ISO}) is also included if the shared BLSA structure is used as described in Eq. 3 (b). DB model for read and write

operation should be separated as described in Eq. 3 (c-1) and (c-2), respectively. This is because most DRAMs adopt the small DB swing for the fast read operation while they use full swing for the write operation in order to overwrite data latched at BLSA. In write operation, additional BL swing should be included as described in the last term of Eq. 3 (c-2), $\sigma_{BL}E_{bl\_w}$, because BL pairs selected by column address show full V_{CC} swing, if write data is opposite to the previous data latched at BLSA. The σ_{BL} is the bit update ratio of BL during the write operation. In conventional schemes, DB power-energy consumption is represented by the weighted sum of the power-energy model of read operation of Eq. 3 (c-1) and the power-energy model of write operation of Eq. 3 (c-2) according to read-write ratio. In read-modify-write (RMW) scheme, the weighted sum using the bit update rate of DB line (σ_{DB}) and BL (σ_{BL}) are combined as shown in Eq. 3 (c-3) where σ_{DB} is equal to σ_{BL} or 1.0 according to the DB operation scheme. The bit update ratio of IO line, σ_{IO}, is determined by the non-precharged burst I/O operation between eDRAM and other embedded modules.

$$E_{WL} = C_{WL} \cdot V_{PP}^2 \quad \text{Eq. 3 (a)}$$

$$E_{BL} = S_{BL} M (C_{BL} + C_{SRTO}) V_{CC}^{\ 2} + C_{ISO} V_{ISO}^{\ 2} \quad \text{Eq. 3 (b)}$$

$$E_{DB\_R} = L \cdot S_{DB\_R} C_{DB} V_{CC}^{\ 2} \quad \text{Eq. 3 (c-1)}$$

$$E_{DB\_W} = E_{db\_w} + \sigma_{BL} E_{bl\_w} = L \cdot S_{DB\_W} C_{DB} V_{CC}^{\ 2} + \sigma_{BL} L \cdot C_{BL} V_{CC}^2 \quad \text{Eq. 3 (c-2)}$$

$$E_{DB\_RMW} = E_{DB\_R} + \sigma_{DB} E_{db\_w} + \sigma_{BL} E_{bl\_w} \quad \text{Eq. 3 (c-3)}$$

$$E_{IO} = \sigma_{IO} L \cdot S_{IO} C_{IO} V_{CC}^2 \quad \text{Eq. 3 (d)}$$

$$C_{WL} = C_{mwl} + k(C_{swl} + C_{is} + C_{rx}) \quad \text{Eq. 4 (a)}$$

$$C_{BL} = C_{bl} + C_{sa} + C_{cd} \quad \text{Eq. 4 (b)}$$

$$C_{SRTO} = (C_{rto} + C_{/s}) / 2 \quad \text{Eq. 4 (c)}$$

$$C_{DB} = C_{db} + C_{cd} + C_{idb} \quad \text{Eq. 4 (d)}$$

$$C_{IO} = C_{io} + C_{odb} + C_{load} \quad \text{Eq. 4 (e)}$$

Each component of C_{WL}, C_{BL}, C_{SRTO}, C_{DB} or C_{IO} contains its line capacitance and MOS junction/gate capacitance, which is connected with corresponding line, as described in Eq. 4 (a)~(e). C_{WL} contains all of the capacitive loads including main word line capacitance (C_{mwl}), subword line capacitance (C_{swl}), subword driver input capacitance (C_{is}), and boosted voltage line capacitance (C_{rx}). C_{BL} includes all capacitive loads such as the bit line capacitance (C_{bl}), SA node capacitance (C_{sa}), and column gate capacitance (C_{cd}). C_{DB} includes all capacitive loads such as DB line capacitance (C_{db}), column gate capacitance (C_{cd}) and DB sense amp input capacitance (C_{idb}). C_{IO} is composed of the capacitive load of IO line capacitance (C_{io}), DB sense amp output capacitance (C_{odb}) and load capacitance (C_{load}).

5. Verifications and discussions

The estimation results obtained from the SSBA power-energy model is compared with the transistor level simulation results of three fabricated eDRAMs using different circuit and process technology as summarized in table I [7]~[9]. Designs A and B, which have the features of RMW scheme, are fabricated by using 0.35μm and 0.18μm embedded memory logic (EML) technology, respectively. Design C, fabricated by using 0.16μm DRAM technology, provides the conventional SDRAM interface. Design A reduces the BL power by using the partial activation, but suffers from large DB power consumption caused by large DB swing of half V_{CC} pre-charged full V_{CC} swing (HPFS) for both read and write operation. In design B, BL power consumption is much larger than DB power consumption because of the large number of BL activations and the small DB swing of full V_{CC} pre-charged small swing (FPSS) read operation. Both of the designs A and B consume more DB power than that of design C because of RMW transaction. Design C reduces BL and DB power consumption by the combination of single bitline writing (SBW) scheme and half pre-charged small swing (HPSS) read operation. Power supply of V_{CC}/V_{PP} is also scaled from 3.3V/5.0V to 2.0V/3.3V as technology scale from 0.35μm and 0.16μm.

Table I. Design characteristics used in verification

	Technology	Power (V_{CC}/V_{PP})	BL Swing	DB Swing	Etc
A	0.35μ EML	3.3/5.0	Direct + CC SA	HPFS	RMW
B	0.18μ EML	2.2/4.0	CC SA	FPSS	RMW
C	0.16μ DRAM	2.0/3.3	SBW Scheme	HPSS	R/W

The verification results of Fig. 5 show that eDRAM power-energy estimation using the SSBA model can achieve the estimation accuracy over 95% compared with the transistor level power simulation results using PowerMill. These good matches are achieved from the precise modeling of various analog signals used in each design as followings. In design A, the DB power consumption of read operation is almost equal to that of the write operation as shown in Fig. 5 because full swing takes places on both read and write operation. However, in design B, DB power consumption for read operation is smaller than that of the write operation, because of small swing DB operation for fast read. This feature is exactly

estimated by SSBA model. Furthermore, design C adopts single bitline write (SBW) scheme that allows only one of the BL pair to be activated [9]. The SSBA model exactly estimates the feature of SBW as shown in verification results. The proposed SSBA model achieves good agreements within 5% error for the various analog signals, which are difficult to be estimated by the conventional power-energy models.

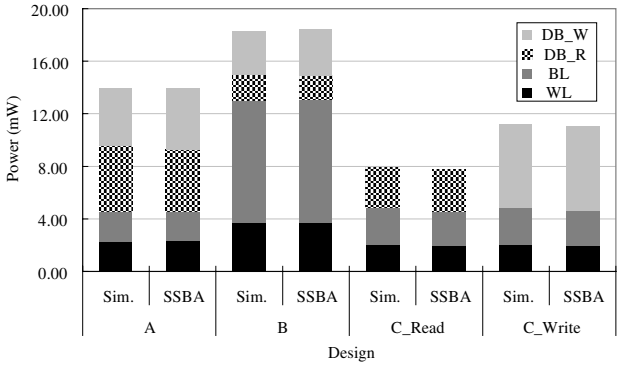

Fig. 5 The comparison of power consumption between SSBA model and transistor level simulation results

6. Conclusion

From the investigation of previous power-energy models of SRAM devices, it is shown that eDRAM power-energy estimation in SoC design have to consider both the high-level eDRAM memory access statistics and the simple-accurate analog swing characteristics of DRAM internal operation, because of following reasons. One is that eDRAM power-energy consumption shows different features with that of off-chip memory system because of the tremendous power reduction on the interconnection capacitance between eDRAM and other embedded components. Another is the eDRAM power-energy dependency on embedded memory architecture as well as other embedded components' architecture. The other is that the eDRAM power-energy estimation error is increased by various analog style signals such as multi-level pre-charge, small swing, differential signal and equalization.

In order to improve accuracy of the conventional power-energy estimation model of SRAM devices, the signal swing-based analytical (SSBA) eDRAM power-energy model is presented. The high-level eDRAM power-energy calculation methodology, which combines the SSBA model and the memory access statistics, provides fast and accurate eDRAM power-energy estimation without detail eDRAM design data. The verification results show that eDRAM power-energy estimation using the SSBA model achieves the estimation accuracy over 95% compared with the transistor level power simulation results for three fabricated designs. This is because the SSBA model exactly estimates the their special analog behaviors of three designs, which are difficult to be estimated by the conventional models.

References

[1] M.B. Kamble and K.Ghose, "Analytical Energy Dissipation Models for Low Power Caches", Proceedings of 1997 International Symposium on Low Power and Electronics Design, pp.143-148, 1997.

[2] R.J. Evans and P.D. Franzon, "Energy Consumption Modeling and Optimization for SRAM's", IEEE J. Solid State Circuits, vol.30, no.5, pp.571-579, May 1995.

[3] C.L. Su and A.M. Despain, "Cache Design Trade-offs for Power and Performance Optimization: A Case Study", Proceedings of 1995 International Symposium on Low Power and Electronics Design, pp.63-68, 1995.

[4] P. Hicks, et al., "Analysis of Power Consumption in Memory Hierarchies", Proceedings of 1997 International Symposium on Low Power and Electronics Design, pp.239-242, 1997.

[5] W.T Shiue and C.Chakrabarti, "Memory Exploration for Low Power Embedded Systems", Proceeding of the 36th ACM/IEEE Conference on Design Automation Conference, pp.140-145, 1999.

[6] D. Liu and C. Svensson, "Power Consumption Estimation in CMOS VLSI Chips", IEEE J. Solid State Circuits, vol.29, no.6, pp.663-670, June 1994.

[7] Y.H. Park, et al., "A 7.1GB/s Low Power Rendering Engine in 2D Array Embedded Memory Logic CMOS for Portable Multimedia System", IEEE J. Solid State Circuits, vol.36, no.6, pp.944-955, June 2001.

[8] C.W. Yoon, et al., "A 80/20MHz 160mW Multimedia Processor integrated with Embedded DRAM MPEG-4 Accelerator and 3D Rendering Engine for Mobile Applications", ISSCC Digest of Technical Paper, pp.142-143, Feb., 2001.

[9] Jeonghon Kook and Hoi-Jun Yoo, "A Single Bit line Writng Scheme for Low Power Reconfigurable I/O DRAM Macro", IEEE European Solid-State Circuit Conference of Digest of Technical Paper, pp. 420-423, Sept. 2000.

[10] T.Sugibayashi, et al., "A 30ns 256-Mb DRAM with a Multidivided Array Structure", IEEE J. Solid-State Circuits, vol. 28, no. 11, pp. 1092~1098, November 1993.