

Data Bus Management to Reduce Dynamic Power in DRAM Devices

Ho Hyun Shin and Eui-Young Chung

School of Electrical and Electronic Engineering, Yonsei University

50 Yonsei-ro Seodaemun-gu, Seoul, Korea

E-mail : hhshin@dtl.yonsei.ac.kr, eychung@yonsei.ac.kr

Abstract

DRAM has become the most practical memory device in computing systems due to its high capacity, low latency and low cost characteristics. However, ultra-high-capacity DRAM devices are also a critical power consumption source for the computing system. Power reduction in DRAM devices is highly dependent on process technology development and sophisticated design capabilities. However, these approaches are limited in the cost sensitive DRAM industry. To reduce the power of the DRAM, we focus on the data bus, which accounts for a significant portion of the dynamic power. This paper proposes a simple and effective technique to reduce the dynamic power of the DRAM device by managing the data bus. This technique groups several adjacent data bus lines and adds control signals. When the data in the group is a specific pattern, that is, when all the data is 0 or is completely inverted compared with the previously held data, the control signal is activated instead of the data bus line. We experiment with a full system simulator GEM5 to evaluate the proposed technique. As a result, the proposed technique reduces the average dynamic power consumed in the data bus by 15.4%.

Keywords: DRAM power, data bus, DBI

1. Introduction

In the era of the Big Data, computing systems are evolving to improve data throughput with high-capacity memories. Among the various memories, DRAM is the most practical device due to its high capacity, low latency and low cost characteristics. However, as DRAM capacity increases, the power consumed by DRAM devices increases proportionally. In particular, the operating power of DRAM devices has become one of the critical power consumption sources in computing systems. To reduce the DRAM power, DRAM vendors generally rely on process technology. However, this approach faces difficulties. First, since lowering the operating voltage is related to transistor reliability issues, it is

becoming more difficult to continually improve. Second, DRAM process technology shrinks more slowly than before, hindering power savings. These facts force the DRAM designer to consider other solutions.

Dynamic power within a DRAM device consists of three major parts. The first is the active power consumed by the bit line sense amplifiers. This active power is generated whenever the DRAM controller issues an activation command. The second is the interface-related power between the DRAM device and the DRAM controller. The power consumed in the in/output buffers and DLL (delay locked loop) circuits is included here. The last is the power consumed by the peripherals to decode the command, address and data and deliver them to each array bank. Among the three power sources, the activation power is in trade-off with the characteristics of the DRAM cell, so it is not easy to reduce the power. In addition, it is difficult to reduce the interface related power because it is closely related to off-chip signal integrity.

Based on the above considerations, we are interested in the dynamic power consumed in the peripheral circuits of DRAM devices. In particular, we have studied how to reduce the load on the data bus, which consumes large amounts of power in DRAM devices. However, since it is not possible to physically reduce the capacitance of the data bus, we have examined how to minimize toggling of the data bus.

The most important consideration to reduce DRAM power is overhead. This is because the DRAM industry is very cost-sensitive and circuits designed for power savings can consume more power than expected. Based on these considerations, we examined how to reduce dynamic power with minimal overhead.

Typical DDR3 DRAM devices in PC or server system are configured by 64-byte data bus[1]. Since one rank is usually composed of 8 chips, there are 64 data bus signals in one DRAM chip. This structure may vary depending on the cache structure, but the normal cache memory operates on a 64-byte cache line basis. This data bus on the DRAM chip is toggled whenever the DRAM controller issues a read

or write command. This toggling causes dynamic power in DRAM devices. We found through SPICE simulation that the power from toggling the data bus accounts for 15% of the average DRAM operating power.

In this paper, we introduce a technique that reduce power by reducing data bus toggling in DRAM devices. To this end, we group several bus lines into one group and add control signals that characterize the data signal. The control signal indicates whether the grouped data signal is all zeros or whether the values of the signal are fully inverted for a previously held value. We also experiment with the full system simulators GEM5 and DRAMSim2 to evaluate the power reduction of the data bus[2-3].

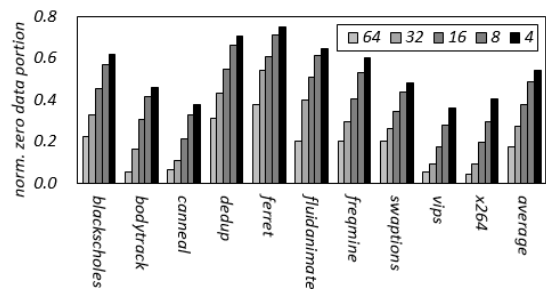
2. Motivation

The main signal sources that occupy the dynamic power of the DRAM peripheral circuits consist of three major components, which are command, address, and data signals. Among these power sources, there are many reasons why we focus on the data bus. First, since the command signals are only composed of four signals of CSB, RASB, CASB and WEB, the portion occupied by the total power is very small. Second, although the address signal is as many as 20-30, including bank, row, and column addresses, it is not easy to reduce it through special logic design techniques because it is very random in signal characteristics. On the other hand, the data signal is composed of 64 signals and occupies a very large power part. In addition, since the value of the data signal is 0 rather than 1, power can be reduced through a special signal conversion technique.

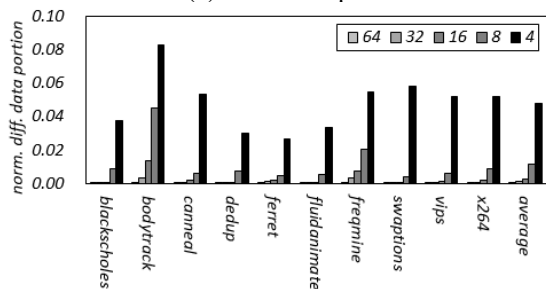
The DRAM data bus basically holds the current data on the data bus until the next read or write command is issued from the DRAM controller. This is to reduce power consumption due to pre-charge. That is, the power consumption on the data bus occurs only when the existing data is different from the currently issued data. This means that power consumption can be reduced if DRAM circuit designers exploit special techniques to suppress toggling of the data bus.

In this paper, we are interested in two data patterns. One is a pattern in which several adjacent data signals become 0 at the same time, and the other is a pattern in which data issued by the current read or write command is inverted with data stored in the existing data bus. The former case can increase the probability of power reduction because of the higher probability of occurrence, while the latter case has a lower probability of occurrence, but it can reduce power loss by preventing simultaneous toggling of multiple data buses.

To explore the data pattern in a DRAM device, we simulated a full computing system including DRAM devices. Fig. 1 shows the ratio of groups which including all zero data when the data signals of the DRAM chip are grouped. In the figure, 64, 32, 16, 8 and 4 represent the number of contiguous data signals contained in one group. As shown in Fig. 1(a), if all 64 data signals belong to one group, the average 17% group is all 0s. This ratio increases up to 54% when four data signals are organized into one group. These results show that grouping the data signals can replace zero data with another signal line, thereby reducing power consumption due to toggling of the data signals.



(a) Zero data portion



(b) Full inversion data portion

Fig. 1. Data pattern analysis with respect to various grouping topology.

Another technology we have focused on is data bus inversion(DBI). This technique is used to reduce the interface power generated between the DRAM device and the DRAM controller. However, applying a DBI to the data bus of a DRAM device is not efficient because the data inversion decision circuit is not simple and causes a lot of power. Based on these explorations, we decided to apply DBI very conservatively to DRAM devices. In other words, DBI technology applies only if the previous data is completely different from the current data. To increase the efficiency of the DBI, we group the data signals in the same way as the zero data experiments. Fig.1(b) shows the ratio of groups in which the current data is completely different from previous data. As shown in the figure, if four bus lines are grouped into one group, the case in which data is completely inverted increases sharply.

3. Proposed Scheme

(1) Bus Grouping

In the proposed scheme, it is most important to determine the grouping scale of the data bus lines. Grouping bus lines in too small a scale requires a lot of control signals, which in turn increases power consumption by themselves. In contrast, if the bus lines are grouped into too large scales, the power consumption effect is reduced. To determine the grouping scale, we experimented with a full system including the DRAM subsystem. Fig. 2 shows the results of normalizing the dynamic power occupied by the data bus, control signal and control circuit with respect to the conventional bus structure when 16, 8 and 4 bus lines are associated with one group. Fig. 2(a) and 2(b) show the experiment results of a scheme in which the control signal is activated when all signals in the group are zero and fully inverted with respect to the held signals respectively. The smaller the number of bus lines belonging to a group, the smaller the dynamic power consumed by the data bus, while the power consumed by the control signal increases. That is, since there is a trade-off point, we have determined from the experimental results the grouping level that minimizes the dynamic power. To control zero and full inversion with minimum power, eight and four bus lines must be included in one group, respectively.

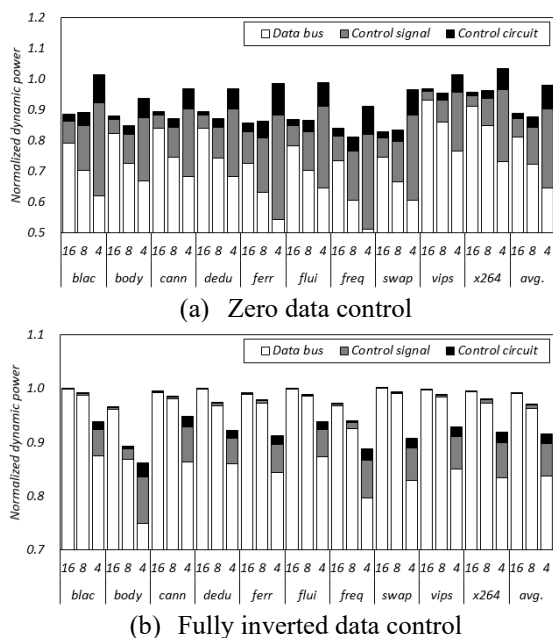


Fig. 2. Dynamic power of data bus of various grouping scales normalized to the conventional data bus structure.

(2) Design of the Proposed Scheme

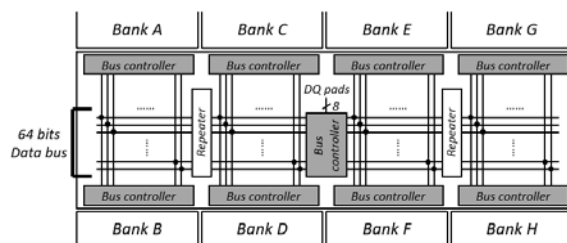


Fig. 3. Data bus architecture in a typical DRAM chip.

Fig. 3 shows the bus architecture of a typical DDR3 DRAM device with 64 data bus signals spread across the chip. To reduce power consumption by bus line toggling, we modified the bus structure as described in Section 3(1) and added the control circuitry into the bus controller as shown in Fig. 4. The circuitries of Fig. 4(a) and 4(b) operate as an encoder and a decoder respectively.

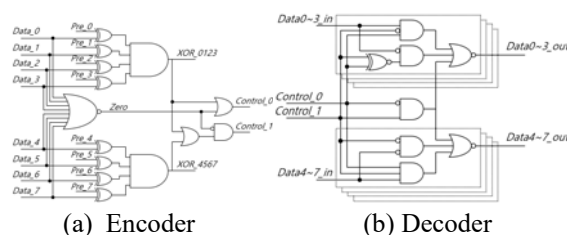


Fig. 4. Data bus control circuits.

4. Experimental Results

We define three experimental cases to verify the power reduction of the proposed scheme, named case A, B, and C. Case A divides 64 data bus lines into 8 groups and adds one control signal to each group. Here, when all eight data lines in the group are 0, a control signal is activated. Case B divides 64 data buses into 16 groups and adds one control signal to each group. Here, the control signal is activated when all four bus lines in the group are completely inverted from the previously stored data. Finally, 64 data buses are divided into 8 groups and 2 control signals are added to each group. If all the data signals in the group are 0, the control signal becomes 01. When the first and second consecutive data bus lines are completely inverted from the stored data, the control signals become 10 and 11, respectively.

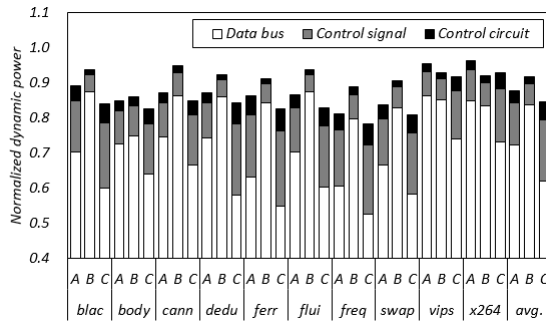


Fig. 5. Experimental results

Fig. 5 shows the experimental result. As can be seen, the schemes of A, B, and C reduce the average power by 12.2%, 8.3%, and 15.4%, respectively. In the *freq* application where the proposed scheme has the greatest effect, it shows 21.8% power reduction with case C.

5. Conclusion

Reducing power in a DRAM is one of the most difficult parts. This is because DRAM power consumption is heavily dependent on DRAM process technology, and the DRAM industry is very cost-sensitive. So, we have been thinking about how to effectively reduce power by using very simple, low-cost technology. In this paper, we introduce a technique to reduce the power consumed by bus data that takes up a lot of dynamic power of the DRAM, though not a big part, by using simple circuit techniques. The proposed scheme groups data buses into several groups and determines whether all the data in one group is 0 or completely opposite to the stored data. Based on the result of the determination, the dynamic power of the data bus is reduced by utilizing the control signal. The proposed scheme reduces the average dynamic data bus power by 15.4% and up to 21.8%.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (2016R1A2B4011799).

References

- [1] Samsung, "DDR3 SDRAM Datasheet", 2006
- [2] N. Binkert, B. Beckmann, G. Black, S. K. Reinhardt, A. Saidi, A. Basu, J. Hestness, D. R. Hower, T. Krishna, S. Sadashti, R. Sen, K. Sewell, M. Shoaib, N. Vaish, M. D. Hill, and D. A. Wood, "The Gem5 Simulator", SIGARCH Comput. Archit. News, vol. 39, no. 2, pp. 1–7, Aug. 2011.
- [3] P. Rosenfeld, E. Cooper-Balis, and B. Jacob, "DRAMSim2: A Cycle Accurate Memory System Simulator", Computer Architecture Letters, vol. 10, no. 1, pp. 16–19, Jan. 2011.