Partial row address prefetch and auto-activate using precharge command for high density DRAM

M.-K. Lee and K.-S. Chung

Dynamic RAM (DRAM) is mainly used as the main memory. As the DRAM density increases, the address bus width required for a DRAM device is also increased. However, in a high-speed DRAM interface, as the address bus becomes wider, not only the routing area and power consumption increase but also the timing margin to maintain signal integrity also decreases. Furthermore, without increasing the address bus pins, row addressing may need multiple cycles, which will lead to significant performance degradation. In this Letter, the authors propose a novel row addressing scheme to issue single-cycle row addressing without extra address bus pins for high-density DRAM devices. The proposed scheme prefetches a part of the target row address for the next activate command using unused address bus of a precharge command. The proposed scheme also enables an auto-activate operation which activates automatically after precharge without explicitly issuing an activate command. As a result, the proposed row addressing scheme reduces the performance degradation due to multi-cycle row addressing and the power consumption on the memory bus due to explicit activate command.

Introduction: Dynamic RAM (DRAM) is mainly used as the main memory. For decades, DRAM vendors have developed memory structures and process technology to increase DRAM density. As the DRAM density increases, the address bus width is also increased. For instance, DDR3 DRAM [1] requires up to 16 bits of the address bus and DDR4 DRAM [2] requires up to 18 bits of the address bus. This increase of the address bus width has a negative effect on the high-speed DRAM interface. As the address bus becomes wider, not only the routing area and power consumption increase but also the timing margin to maintain signal integrity also decreases gradually. In addition, the wider address bus increases DRAM manufacturing cost. Therefore, it is important to prevent the address bus width from increasing in DRAM. Without adding more address bus pins, row addressing may need multiple cycles, which will lead to significant performance degradation.

To use data in a DRAM array, DRAM requires three steps in sequence: ACTIVATE, READ/WRITE, and PRECHARGE. ACTIVATE fetches the data of a specific row of the DRAM array into a row buffer. READ/WRITE is issued to access the data in the row buffer. To access a data of which row address is different from that of the fetched data in the row buffer, PRECHARGE is required to close the row buffer. In other words, PRECHARGE must precede a new activation operation. Furthermore, when issuing the precharge command, most address bus pins are not used except for one address pin [1, 2]. Using this property, it is possible to prefetch a part of the row address of the next activate operation while a precharge operation is conducted.

We propose an addressing scheme to reduce system performance degradation due to multi-cycle row addressing and power consumption in the command/address bus. In the proposed scheme, a part of the row address for the next ACTIVATE is delivered using unused address bus pins in PRECHARGE. Therefore, if the part of the row address of the activate command coincides with the prefetched row address, the target row address can be transferred in a single cycle. The proposed scheme can also activate a new row by using a prefetched row address by the precharge command without explicitly issuing an activate command.

Proposed row addressing scheme: In this Letter, we assume that a 19-bit row address is used and 16-bit address bus is available. In such a DRAM device, the memory controller needs two cycles to issue the activate command by dividing the row address into two parts. We propose a row addressing scheme to issue a single-cycle activate command through the narrow address bus than the row address bits by prefetching a part of the row address. In other words, most significant bits (MSBs) which are a set of upper bits of a row address, is prefetched along with the precharge command. Then, the activate command can be issued in one cycle by sending only the remaining row address, called as least significant bits (LSBs). However, according to our evaluation, 38.4% of precharge commands are issued when the command queue is empty. This means that there is no row address that can be prefetched. Therefore, it takes two cycles to transmit the next activate command. To avoid this two-cycle activation as much as possible, a structure called prefetch table (PT), which stores MSBs of recently accessed rows utilising access locality [3], is added in the memory controller and each bank of the DRAM device. In the conventional row address scheme, a part of the address bus of the precharge command is already used, and row addresses cannot be sent through the precharge command. However, in the proposed row address scheme, if MSBs of a pending request are found in the PT and LSBs are transmitted with the precharge command, then the corresponding row can be activated automatically without sending an activate command, which is called auto-activating. Therefore, the proposed scheme also reduces the power consumption on the address/command bus by reducing the number of activate command issues.

The memory controller selects one memory request (called target request) in the command queue and issues the activate, read/write, or precharge command depending on the target request and the DRAM state. In the proposed row addressing scheme, three types of precharge commands and two types of activate commands are defined and Table 1 shows a function of proposed commands. The precharge command is classified into PRE_Normal, PRE_Prefetch, and PRE_AutoACT depending on the prefetching and activating (auto-activating). Basically, all types of precharge commands perform a precharging operation that closes the row buffer. The activate command is classified into ACT_HIT and ACT_MISS depending on the prefetching and activating.

Table 1: Summary of proposed commands

<table>
<thead>
<tr>
<th>Command</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE_Normal</td>
<td>v</td>
<td>open</td>
<td>closed</td>
<td>MSBs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRE_Prefetch</td>
<td>v</td>
<td>v</td>
<td>open</td>
<td>closed</td>
<td>MSBs</td>
<td></td>
</tr>
<tr>
<td>PRE_AutoACT</td>
<td>v</td>
<td>v</td>
<td>open</td>
<td>open</td>
<td>LSBs</td>
<td></td>
</tr>
<tr>
<td>ACT_HIT</td>
<td>v</td>
<td>closed</td>
<td>closed</td>
<td>MSBs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(1): activating operation; (2): precharging operation; (3): prefetching operation; (4): current bank state; (5): next bank state; (6): transmitted partial row address through the address bus.

Fig. 1 shows the PT of the memory controller and DRAM device. The prefetching operation can be performed when the MSBs of the target request are not in the PT (called MSBs miss). When MSBs miss, the MSBs are updated on the PT of the memory controller and transferred to the DRAM device through the address bus by issuing the ACT_MISS or PRE_Prefetch. The transferred MSBs are updated in the PT of the DRAM device. The PT in the DRAM device has a circular first-in-first-out queue structure so that the oldest entry in the queue will be replaced by the new MSBs. However, if the MSBs of the target request are in the PT (called MSBs hit), the row address for the activating operation can be transferred in one cycle by issuing ACT_HIT or PRE_AutoACT. When MSBs hit, LSBs of the request and the PT index of the MSBs are transmitted through the address bus. In the DRAM device, MSBs in the PT and LSBs are combined to form the full row address for the activation as shown in Fig. 1.

Fig. 2 shows the proposed row-addressing scheme algorithm. To close the row buffer, the precharge command is issued when a new row activation is needed or the command queue is empty. If the command queue is empty, PRE_Normal is issued. When the command queue is
not empty, a target request should be selected from the command queue. Then, the MSBs of the target request are looked up in the PT as shown in Fig. 1. When MSBs miss, PRE_Prefetch is carried out. When MSBs hit, if timing parameters related to the activate operation such as \( hF\text{AW} \) and \( hRD \) are satisfied, PRE_AutoACT is carried out. PRE_AutoACT is the precharge operation where a normal precharge operation is followed by an activation operation automatically without issuing an activate command explicitly. However, the timing parameters are not satisfied, PRE_Normal is issued. In the activate command case, the ACT_Hit is issued when MSBs hit and the ACT_Miss is issued when MSBs miss. ACT_Hit is issued in the next cycle immediately after ACT_Miss is issued. In summary, if the MSBs of the request are prefetched before issuing the activate command, the memory controller can transmit the row address in one cycle.

![Fig. 2 Proposed row-addressing scheme algorithm](image)

Fig. 2 shows how the row-address field is used in each type of the precharge and activate commands. In the row-addressing scheme, the upper 8 bits of the 19-bit row address are MSBs and the lower 11 bits are LSBs. The precharge commands can be divided into PRE_AutoACT, PRE_Prefetch, and PRE_Normal according to auto-activating flag and prefetch flag. For activation, ACT_Hit is carried out when row address hit flag is high and ACT_Miss is conducted if RFH is low. In the proposed row addressing scheme, address prefetching and auto-activating are performed by transferring MSBs, LSBs, the PT index and various flags to the DRAM device through the existing address bus without additional address/command pins.

![Fig. 3 Address bus pin assignment of proposed commands](image)

Experimental results: To conduct performance and power consumption evaluation of the proposed scheme, two widely used architectural simulators, gem5 [4] and DRAMSim2 [5] are combined with some modifications to implement the proposed method. Benchmarks were selected from the SPEC CPU 2006 benchmark suite. A metric called weighted speed-up (WS) [6] was used to evaluate the performance of the proposed method. In the evaluation result, we compared the proposed scheme with a conventional multi-cycle row addressing scheme (MCR) that the activate command is sent in two cycles. For the proposed scheme and MCR, it is assumed that a row address has 19 bits and the address bus width is 16 bits.

Fig. 4 shows the normalised WS of the proposed scheme compared with the MCR. The proposed scheme has 2.04% performance improvement compared to MCR. MCR cannot send a row address in a single cycle but sends the address in two cycles. The activate operation is delayed and the command/address bus is blocked while the row address is being sent. On the other hand, the proposed scheme reduces the system performance degradation due to MCR by prefetching the partial row address using the precharge command and sending row address in a single-cycle through the activate command. Fig. 4 also shows the normalised address/command bus power consumption of the proposed scheme compared with the MCR. On average, the power consumption of the address/command bus is reduced by 29.4%. In the MCR, the power consumption in the address/command bus is increased by transmitting the activate command in two cycles. However, in the proposed row address scheme, the activate command can be transmitted in one cycle. In addition, the prefetched partial row address is used to open the new row through the precharge command without the activate command, thereby reducing the power consumption on the address/command bus.

**Fig. 4 Normalised WS and power consumption of address/command bus**

**Conclusion:** In this Letter, we presented a new row addressing scheme to issue a single-cycle activate command without adding extra address bus pins for high-density DRAM devices. The proposed scheme takes advantage of the fact that a precharge command must be preceded before the activate command and almost no address bus pins are used when issuing a precharge command. The proposed scheme prefetches a part of a row address for the next activate command. Using unused address bus pins when issuing the precharge command. The prefetched part of the row address is stored in a PT. Using the prefetched row address in a DRAM device, a single-cycle activate command with a row address wider than the address bus is possible. The proposed scheme also enables the auto-activate operation that activates a new row automatically after a precharge operation without the activate command using the proposed FT. As a result, the proposed row addressing scheme reduces the performance degradation due to multi-cycle row addressing and the power consumption on the address/command bus due to activate command issues.

**Acknowledgments:** This work was supported by Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (R719-16-1009, Development of Intelligent Semiconductor Core Technologies for IoT Devices based on Harvest Energy).

© The Institution of Engineering and Technology 2018
Submitted: 18 September 2018
doi: 10.1049/el.2018.7108

One or more of the Figures in this Letter are available in colour online.

M.-K. Lee and K.-S. Chung (Department of Electronics and Computer Engineering, Hanyang University, Seoul 133-791, Republic of Korea)
E-mail: kchung@hanyang.ac.kr

**References**

1. DDR3 SDRAM JEDEC Standard, Available at http://www.jedec.org, accessed February 2018
2. DDR4 SDRAM JEDEC Standard, Available at http://www.jedec.org, accessed February 2018