Dynamic voltage frequency scaling-aware refresh management for 3D DRAM over processor architecture

J. Lim, H. Kim, H. Oh and S. Kang

Three-dimensional integrated systems that combine large-capacity dynamic random access memory (DRAM) with high-performance processors represent a promising solution to implementing high-performance computing. However, in such configurations, stacked DRAM cells will inevitably be exposed to high temperatures generated by the processor, thereby necessitating DRAM refreshes with high refresh rates driven by embedded temperature sensors. In this letter, a thermally aware refresh-control method that accounts for abrupt changes in temperature and the thermal guard band using low-power techniques such as dynamic voltage frequency scaling is proposed. Comparisons with previous systems via single- and eight-core simulations reveal that the proposed method improves efficiency with no additional overhead.

Introduction: 3D-stacked dynamic random access memory (DRAM) structures linked to a processor die with through silicon via (TSV) represent a promising high-performance computing solution. The 3D-stacked DRAM-over-processor architecture is advantageous in terms of its high bandwidth and reduced data-transmission power between the DRAM cache and the processor. In these structures, the processor consumes most of the power and generates a significantly higher amount of heat than a 3D-stacked DRAM. In the 3D-stacked architecture, the processor and DRAM are very closely located, and they increase the temperature and necessary refresh rate of the DRAM owing to a significant amount of heat generated by the processor. Thus, a refresh-control technique driven by the DRAM temperature is necessary in this architecture. However, as the thermal hotspot of the DRAM changes depending on the processor hotspot, the thermal environment of the DRAM die becomes variable and therefore difficult to monitor.

In [1, 2], refresh-control systems that employ temperature sensors in both the DRAM and the processor die were proposed. In these systems, the refresh rates determined by the respective sensors are compared and a compromise refresh rate is determined. In [1], a thermal guard band-generation method that considers both the temperature difference between sensors and the sensor data delay to calculate an efficient refresh rate considering the data reliability was also proposed.

As embodied in the DRAM refresh issue, the thermal problem is a well-known weak point in 3D-integrated systems. In [3], the reliability and performance of a 3D-integrated DRAM with processor system were evaluated. They found that although 3D integration enables high performance, high temperatures can threaten the reliability of a system. To regulate the temperature of 3D-integrated systems, the power dissipation of the processor must be regulated because the processor dissipates most of the overall power in the system. Dynamic voltage frequency scaling (DVFS) is a well-known low-power solution for processor systems in which the voltage and frequency of the circuit are dynamically adjusted. The voltage level of a DVFS operation is determined by the workload of the processor system.

Fig. 1 shows the operating features of DVFS. When the workload of the core is light, the voltage of the circuit is adjusted to a low level; conversely, when the workload is heavy, the voltage is maintained at a high level. The operating frequency is adjusted depending on the circuit voltage following a circuit frequency-to-voltage relation that can be modelled using the alpha power law [4]. Fig. 1 illustrates how the power dissipation in the circuit changes with the circuit voltage. Dynamic power dissipation is not uniform and changes rapidly within the maximum value range during operation; conversely, static power dissipation is relatively uniform and follows a gradual temperature-dependent profile. Based on these features, the power range of a circuit can be defined. The power dissipation in a normal-voltage circuit fluctuates within a range bounded by the static power and the sum of the minimum dynamic and static power. However, the power dissipation in a DVFS-applied circuit fluctuates over a wider range bounded by the static power at the minimum voltage and the sum of the maximum dynamic and static power. The resulting enlarged power-dissipation range of a DVFS-applied circuit can induce more sudden power dissipation, which can cause a significant data unreliability. To guarantee the data reliability of a 3D-integrated DRAM cell, its refresh rate must be controlled to adjust for the rapid voltage and power increment described in [1]. In this Letter, a method for efficient refresh rate control in line with specific DVFS features is proposed.

DVFS-aware temperature guard band setup: To clarify the power-consumption range of a circuit in the presence or absence of DVFS, the following power-consumption model is given:

\[ P = P_s + P_d \times \alpha, \quad P_s \propto V, \quad P_d \propto V^2 \times f \]  

where \( P_s \) is the static power, \( P_d \) is the dynamic power, \( \alpha \) is the ratio of the switching activity, \( V \) is the circuit voltage, and \( f \) is the circuit frequency. In the DVFS case, the minimum power is the static power at the minimum voltage, whereas the maximum power is the sum of the static and dynamic power at the maximum voltage.

In [1], the temperature guard band was calculated from the maximum and minimum power-consumption values of each module. This process assumes a power-consumption sequence that maximises the temperature of the DRAM cell and simultaneously minimises the temperature of the thermal sensor based on sensor data delay and location difference. If the power consumption of a module maximises the ‘temperature of the DRAM cell minus the temperature of the sensor,’ the power consumption is assumed to be the maximum; otherwise, it is assumed to be the minimum. If the gap between the minimum and maximum power is enlarged, the thermal guard band is enlarged proportionally. This enlarged guard band guarantees refresh reliability but is wasteful. By considering DVFS-specific features, the efficiency of the thermal guard band can be improved.

The circuit voltage in a DVFS-applied processor is not completely flexible owing to the following constraints in shifting the circuit voltage:

- **Cluster constraint:** in the DVFS technique, certain modules belong to the same voltage regulator and therefore have the same voltage level.
- **Timing constraint:** in general DVFS systems, voltage-level shifting occurs over a fixed time interval.

Clustering more modules at the same voltage level will induce smaller temperature variations because the potential for unbalanced power distribution is diminished in a DVFS cluster. This feature can be used to improve the efficiency of the temperature guard band. In per-chip DVFS, the DVFS cluster is at the chip level; in per-core DVFS, the cluster is at the core level, and each core can have a different voltage level. Applying the DVFS technique with clustered cores can reduce the overhead in a multicore processor [5].

If the voltage level must be maintained over a specific interval, it is possible to reduce the thermal fluctuation uncertainty by adjusting the elapsed time from temperature sensor sampling to refresh control. In our refresh-control system, a temperature-to-refresh period table is employed to determine the proper refresh rate. As the use of timing information would require additional hardware, such information is not used in the proposed refresh-control system, which supports our goal of developing an algorithm without hardware overhead.

Fig. 2 shows the proposed thermal guard band calculation method, in which \( d \) indicates the gap between voltage-shift and thermal sensor sampling, \( gb \) indicates the thermal guard band, ‘slot’ indicates the available computing time during the voltage-shift interval, and \( P_{\text{min}} \) and \( P_{\text{max}} \) indicate the minimum and maximum power values, respectively. In the computing process, the time domain is decomposed to time slots, and the temperature rise in each DVFS-cluster, and the influences of these on the thermal guard band are summed. This process is reasonable because the thermal R–C model [6] follows the superposition principle.
The use of separated computation is advantageous in terms of computational complexity. The computation complexity of the proposed method is of number of voltage levels × slot length times greater than the method of [1].

Evaluation with simulation: To evaluate the proposed method, a 3D-integrated DRAM with processor architecture was modelled as a per-core DRAM die stacked over a processor die. The temperature simulation tool HotSpot [6] and McPAT, a processor power and area modelling tool, [7] were employed. Two types of processor die were modelled: a multicore processor based on [8] and a single-core processor based on [9]. The modelled DRAM was based on the Micron DRAM described in [10]. To implement DVFS, five voltage levels, ranging from 0.5 to 0.9 V, were assumed in guard band setup and third voltage level is assumed in temperature distribution simulation. The single-core model employed a DVFS cluster with all components except for the L2 cache containing processor modules. In the multicore model, three types of DVFS clusters were simulated: a per-core cluster, a two-core cluster, and a four-core cluster. Adjacent cores were clustered to the same voltage. The most efficient sensor sampling rate, activated sensors, and refresh controller and sensor overheads were all obtained from [1].

Table 1 shows the results of the simulation. The refresh power values indicate the sum of the refresh power and the refresh controller and temperature sensor power, and the refresh interval is the average refresh interval of the 3D-stacked DRAM cells. The performances of the modelled systems were compared with the results obtained in [1] and were found to vary with processor architecture, with the single-core model demonstrating the highest efficiency. Implementation of the proposed method in the multicore model resulted in a relatively low efficiency since the DVFS-applicable area of this model is smaller than that of the single-core model and because multiple DVFS clusters induce thermal distribution uncertainty. Increasing the number of cores in the DVFS clusters improved the efficiency of the proposed method. Overall, the proposed method demonstrates high efficiency in a large DVFS cluster architecture.

Table 1: Simulation result of proposed guard band setup method

<table>
<thead>
<tr>
<th>Guard band calculation method</th>
<th>Per-core cluster</th>
<th>Two-core cluster</th>
<th>Four-core cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refresh power [1]</td>
<td>21.36 mW</td>
<td>21.36 mW</td>
<td>21.36 mW</td>
</tr>
<tr>
<td>Proposed</td>
<td>20.87 mW</td>
<td>20.81 mW</td>
<td>20.73 mW</td>
</tr>
<tr>
<td>Reduction ratio</td>
<td>2.29%</td>
<td>2.56%</td>
<td>2.94%</td>
</tr>
<tr>
<td>Proposed</td>
<td>27.08 ms</td>
<td>27.16 ms</td>
<td>27.27 ms</td>
</tr>
<tr>
<td>Refresh rate reduction</td>
<td>2.45%</td>
<td>2.74%</td>
<td>3.14%</td>
</tr>
</tbody>
</table>

Conclusion: In this Letter, a thermal guard band calculation method for 3D-stacked DRAM refresh management was proposed. The proposed method can guarantee 3D-stacked DRAM data reliability despite a DVFS-induced thermal uncertainty. In a simulation, our method improved the refresh efficiency by up to 7% without additional hardware overhead.

Acknowledgment: This work was supported by the IT R&D programme of MOTIE/KEIT [10052716, Design technology development of ultra-low-voltage operating circuit and IP for smart sensor SoC.]

© The Institution of Engineering and Technology 2017
Submitted: 5 April 2017 E-first: 13 June 2017
doi: 10.1049/el.2017.1243
J. Lim, H. Kim, H. Oh and S. Kang (Department of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea)
✉ E-mail: shkang@yonsei.ac.kr

References