

Retention Time Measurements and Modelling of Bit Error Rates of WIDE I/O DRAM in MPSoCs

Christian Weis*, Matthias Jung*, Peter Ehses*, Cristiano Santos[†], Pascal Vivet[†],
Sven Goossens[‡], Martijn Koedam[‡] and Norbert Wehn*

*University of Kaiserslautern, Kaiserslautern, Germany

[†]CEA LETI, Grenoble, France

[‡]Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract—DRAM cells use capacitors as volatile and leaky bit storage elements. The time spent without refreshing them is called *retention time*. It is well known that the retention time depends inverse exponentially on the temperature. In 3D stacking, the challenges of high power densities and thermal dissipation are exacerbated and have a much stronger impact on the retention time of 3D-stacked WIDE I/O DRAMs that are placed on top of an MPSoC. Consequently, it is very important to study the temperature behaviour of WIDE I/O DRAMs. To the best of our knowledge, no investigations based on real measurements were done for stacked DRAM-on-logic devices. In this paper, we first provide detailed measurements on temperature-dependent retention time and bit error rates of WIDE I/O DRAMs. To obtain the correct temperature distribution of the WIDE-I/O DRAM die we use an advanced thermal modelling tool: the DOCEA AceThermalModelerTM (ATM). The WIDE I/O DRAM retention times and bit error rates are compared to the behaviour of 2D-DRAM chips (DIMMs) with the help of an advanced FPGA-based test system. We observed data pattern dependencies and variable retention times (VRTs). Second, based on this data, we develop and validate a SystemC-TLM2.0 DRAM bit error rate model. Our proposed DRAM bit error model enables early investigations on the temperature vs. retention time trade-off in future 3D-stacked MPSoCs with WIDE I/O DRAMs in SystemC-TLM2.0 environments.

I. INTRODUCTION

Energy and thermal dissipation are limiting today's application performance on smartphones as well as high-end servers. Advanced fabrication processes based on 3D packaging enable tighter integration of systems that start to break down the memory and bandwidth walls. However, this comes at the price of increased power density and reduced heat dissipation properties of the aggressively thinned dies. In addition, poorly conductive adhesive materials used to bond dies together considerably contribute to increase the vertical thermal resistance. The thermal issues of 3D ICs cannot be solved by tweaking the technology and circuits alone. In fact, a 3D stacked SoC aggravates the thermal crisis, which can provoke errors in circuits and especially in DRAMs as they are highly sensitive to temperature changes and have to be refreshed regularly due to their charge-based bit storage property (capacitor). The *retention time* of a DRAM cell is defined as the amount of time that a DRAM cell can safely retain data without being refreshed [1]. This DRAM refresh operation must be issued periodically and causes both performance degradation and increased energy consumption (almost 50% of future DRAMs total energy), both of which are expected to worsen as the DRAM density increases [2]. Due to process variation some DRAM cells leak more than others. Many prior works assume that it is possible to keep track of the relatively few weak DRAM cells (low retention time) [3], [4] and therefore reducing the impact of refreshing by avoiding the usage of these cells. However, the effects of *data pattern dependence*

and *variable retention time* (VRT), which we also observed during our measurements, restrict heavily the usage of DRAM retention time profiling mechanisms.

In this scenario, we perform four major investigations leading to the contributions of this work:

- 1) We measure the retention times and provoked bit errors of WIDE I/O DRAM dies on top of a SoC logic die.
- 2) We derive the accurate temperature distribution map on the WIDE I/O DRAM die with the help of an advanced thermal modelling tool.
- 3) We compare the retention times and bit errors rates of WIDE I/O DRAMs and 2D-DRAMs based on DIMMs (Dual-Inline Memory Modules) with the help of an advanced FPGA-based test system.
- 4) Finally, we propose a *calibrated DRAM bit error model* based on real measured data to be integrated in our SystemC-TLM2.0 environment or in other system simulators, such as *gem5* or *DRAMsim2*. With this we can quantify the temperature vs. retention time trade-off in a early development state of future 3D stacked DRAM-on-logic-SoCs.

The remainder of this work is structured as follows: we first refer to prior work in Section II. A detailed description of the experimental measurement setup for WIDE I/O and 2D-DRAMs is presented in Section III. Then we show the derivation of the temperature map for the WIDE I/O DRAM die in Section III-B. The experimental results concerning the measurements of retention times and bit error rates are presented in Section IV. Our TLM2.0 based DRAM bit error model for fast simulations is discussed in Section V. Finally, we conclude with Section VI.

II. RELATED WORK

The academic research related to the discussed contributions in this paper are mainly grouped as follows: Measurements of DRAM device/die behaviour, systems with WIDE-I/O DRAMs, data pattern dependence and VRT, and modelling of DRAM bit errors.

A first evaluation of retention time distribution was presented by Hamamoto et al. [1] for an experimental 16 Mbit DRAM chip. More recently Kim et al. [5] measured three 1 Gbit chips. Both studies used only one type of DRAM and did not discuss data pattern dependence and VRT. A similar detailed study on data retention time in modern DDR3 DRAM devices was presented by Liu et al. [6]. However, when they analysed VRT, only a single temperature was used, while in our measurements we observed direct retention time state changes (low to high) when e.g. increasing the temperature by 10 °C.

The applied refresh rate of a DRAM device depends on its leakiest cells. However, the number of low retention time cells is relatively small compared to the total number of cells

in a DRAM. To lower or to mitigate the impact of DRAM refresh operations, the accurate identification (profiling) of the weak DRAM cells was a prerequisite of many prior works [3], [4], [2]. Software approaches to reduce refresh power by retention time profiling and retention-aware placement of data in the DRAM are presented in [4]. A method for reducing the total refresh rate by grouping the DRAM rows into different retention time bins and applying different refresh rates on them is presented in [2]. Our measurements show that retention time profiling mechanisms should take VRT and data pattern dependence into account. The data pattern dependence on the retention time was analysed heavily by Liu et al. [6] as well. It was found that for some newer devices a simple “1” or “0” data pattern is not sufficient (only 15% coverage). In this case coverage means the ratio of the error addresses discovered in a single test versus the total number of discovered error addresses, aggregated over all tests. Consequently, our work considers different data pattern topologies. The VRT phenomenon was studied thoroughly by the electron device research groups [7], [8]. Their focus is on understanding the physical cause of VRT, such as characterizing the charge trap existing in the gate oxide of a DRAM cell’s access transistor. Recently, Shirley et al. [9] used Copulas (widely used in financial and actuarial modelling) to model the VRT probabilities. We propose a retention-aware DRAM bit error model that considers VRTs and data pattern dependence. The 3D MPSoC with WIDE I/O DRAM presented in [10] (WIOMING chip) is used for our investigations. This 3D-IC features thermal sensors and heaters, which can be used for online monitoring of temperature and the tuning of thermal models. We employ the MAGALI (SoC) evaluation board and two WIOMING devices (Chip 1 and 2) to investigate the temperature dependent retention errors (bit flips). The WIDE I/O DRAM of the WIOMING 3D-IC is based on the chip presented by Samsung in [11]. A detailed thermal characterization of 3D-ICs with WIDE I/O DRAM is presented in [12]. The effect of different alignments for DRAM dies and TSVs on overall chip temperature is evaluated. However, [12] focuses on thermal modelling only, thus it does not discuss bit error rate modelling of DRAMs. Concerning bit error rate modelling of DRAM most of research is related to soft errors caused by radiation [13], [14]. Contrary to state-of-the-art and compute-intensive prior work [9], we focus on a fast-executable, comprehensive DRAM bit error model for errors caused by retention time failures that enables early analysis of the temperature vs. retention time trade-off in future MPSoCs with 3D-stacked DRAM.

III. MEASUREMENT SETUP

In this Section, we introduce the flow of experiments finally leading to a SystemC-TLM2.0 DRAM bit error model shown in Figure 1 and we detail the setup for the conducted experiments. First, we present in Section III-A and III-B the WIDE-I/O measurement setup and the derivation of the temperature map for the WIDE-I/O DRAM die. Second, we describe the advanced FPGA-based test system in Section III-C.

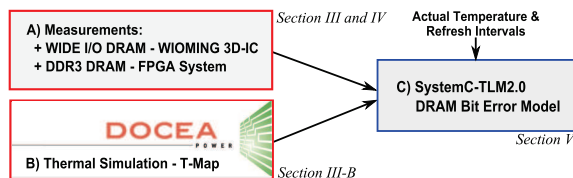


Fig. 1: Overview of Experiments and Modelling

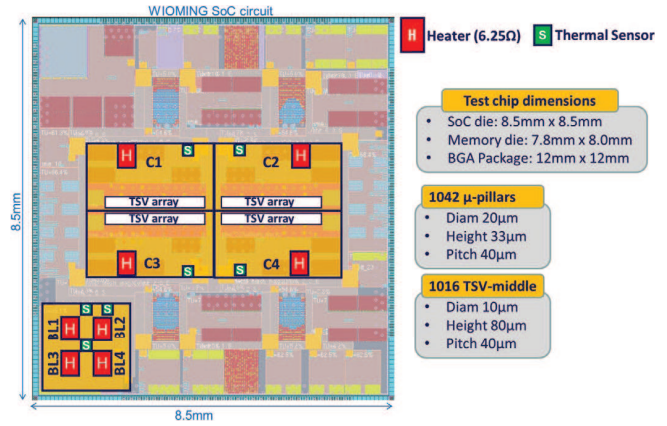


Fig. 2: Floorplan of the WIOMING 3D-IC including 4 heaters and 4 thermal sensors in the center area (C1-C4) and 4 heaters and 3 thermal sensors in the bottom-left corner (BL1-BL4)

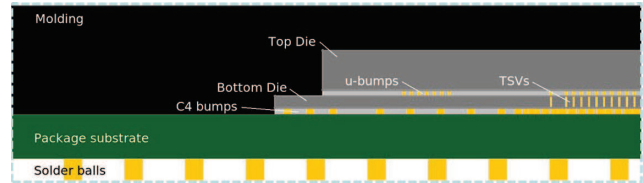


Fig. 3: WIOMING 3D-IC Package

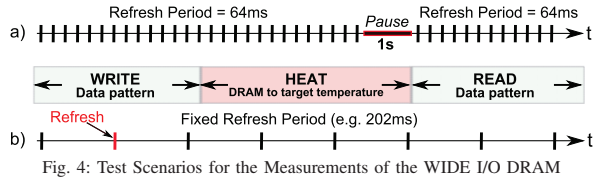
A. WIDE-I/O DRAM Chip and Board Setup

We used for all measurements of WIDE-I/O DRAM retention times the WIOMING 3D-IC chip manufactured in 65nm technology, which is shown in Figure 2. The complete system used for the retention time characterisation includes a packaged 3D test chip, a small PCB interposer and a socket mounted on a large PCB. The packaged test chips are WIDE I/O-memory-on-logic 3D-ICs, where the dies are stacked in a face-to-back configuration and are connected through TSVs and μ -bumps. The bottom die implements a WIOMING circuit [10] and is thinned down to $80\mu\text{m}$ to accommodate the integrated TSVs. Details of the package for the WIOMING 3D-IC are given in Figure 3. The WIDE I/O DRAM with 4 channels ($4 \times 128 = 512$ I/Os) on top of the WIOMING SoC die is based on the device from Samsung [11]. The WIOMING circuit (Figure 2) is instrumented with eight resistive heaters (poly resistance) to emulate hotspot power dissipation. Each heater is independently controlled via embedded software and can dissipate up to 0.659 W ($0.37\text{ KW}/\text{cm}^2$). Integrated thermal sensors (TS) are monitored in real time with 1°C of temperature resolution. Sensors accuracy is $\pm 1^\circ\text{C}$ at the calibration temperature (25°C), degrading to $\pm 4^\circ\text{C}$ at 100°C . Four TSV arrays are placed at the centre of the die, each containing 254 TSVs connected to aligned μ -bumps.

Table I shows our measurement configurations. To provoke bit errors we chose an already high temperature for the starting point of the experiments (80°C). To achieve the different temperatures we have to set the heaters in the centre (C1-C4) and in the bottom left corner (BL1-BL4) accordingly in percentage of the maximum power. Due to the maximum chip temperature limit of 130°C at the bottom left corner sensor

TABLE I: Experiment Configurations

Temp. at Sensor C3 ($^\circ\text{C}$)	80	90	95	100	103	104
Heaters Centre (C1-C4) (%)	70	80	80	95	95	95
Heaters BL1-BL4 (%)	75	85	85	85	85	85
Refresh Periods (ms)	128	143	165	183	202	1000 (wait)
Data pattern usage (hex) - all channels	FF	AA	55	Random		



(in the middle of the heaters), a maximum of 104°C could be achieved for Chip 1 and 103°C for Chip 2. The refresh periods range from 128 to 202ms and additionally we used a wait time of 1s to provoke more bit flips. A refresh period of 202ms was the highest programmable time at the memory controller of the WIOMING SoC. We developed two different tests as shown in Figure 4 on the embedded software of the system. We used test b) for the standard tests without a pause and a) for all 1s wait time (refresh pausing) tests. To improve test coverage and highlight the data pattern dependence of the retention time we run each test with several different data patterns. Figure 5 shows the heating-up process of the 3D-IC. The time spent heating the DRAM depends on the initial and target temperature (here 440s). The target temperature measured at sensor C3 (SME_21) of each test ($80\text{--}104^{\circ}\text{C}$) was kept stable for 20 seconds at the end of the heating-up process.

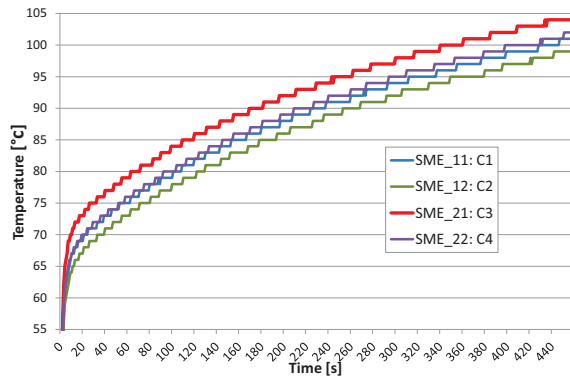


Fig. 5: Temperature Sensor read-out for each second for all SME Channels (C1-C4) during Heating-up

B. Temperature Map on the WIDE-I/O DRAM Die

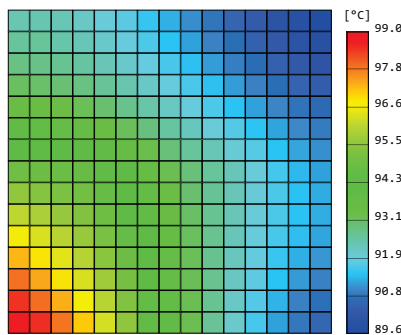
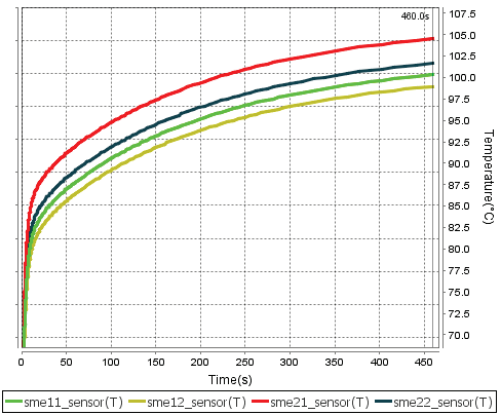


Fig. 6: Temperature Map of the WIDE-I/O DRAM die

Due to in-appropriate temperature sensors within the DRAM (not correctly enabled from SoC and with too large inaccuracy), we cannot obtain an accurate thermal map by the embedded software. For that reason we used AceExplorer in conjunction with AceThermalModeler (ATM) from DOCEA Power [15] to create an accurate temperature distribution map of the WIDE-I/O DRAM die. Such thermal model [16] has been proven to be fast enough for rapid system level exploration with a correct correlation between simulation and silicon



measurements. Thus, the simulated thermal map is sufficiently accurate, to feed the next model, which is our proposed SystemC-TLM2.0 DRAM bit error model. Our measurements of the four centre SME temperature sensors (C1-C4) during heating of 3D-IC is already shown in Figure 5. To obtain the temperature map of the DRAM die (top die) we modified the ATM-model of the WIOMING 3D-IC [16]. We set the heater power to the values we have used in our initial experiments (see also Table I) to achieve the highest temperature for Chip 1 (104°C). Figure 6 shows the simulated temperature map. We see an expected overall temperature decrease of top die (DRAM) compared to the sensors on the bottom die and Channel 3 is by far the hottest area of the DRAM die. Figure 7 verifies the correlation between simulation and measurement.

Figure 8 plots the bit errors using a fixed refresh period (202ms—test b) of all channels vs. the temperature at sensor C3 (SME_21). We see a significant amount of errors for channel 3 only. Channel 4 has a single error and the others have no bit errors. The obtained thermal map of the WIDE I/O DRAM die verifies our experiments running over all four channels of the WIDE I/O chip. Consequently, we focus on channel 3 only to show retention time depending bit flips.

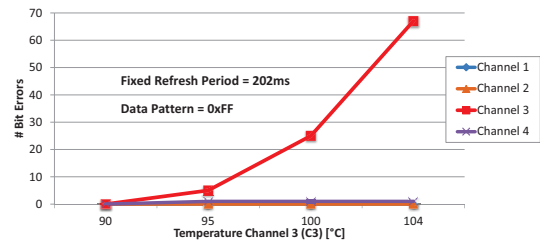


Fig. 8: All Channels at 202ms Refresh Period and $90\text{--}104^{\circ}\text{C}$ at Sensor C3 (SME_21)

C. DDR3 Test System Setup

An FPGA based setup is used for the DDR3 tests. We instantiate the memory controller from [17], which has a freely configurable refresh interval, on an ML605 FPGA board [18]. This board contains a Virtex 6 FPGA, and a Micron DDR3-1066 DIMM [19]. One of the DDR devices on the DIMM is augmented with a temperature sensor and a peltier element, which is used to measure and influence its temperature, respectively. An Arduino board converts the analog temperature reading from the sensor into a digital data stream, which is fed into a PC. A small algorithm runs on the PC. It controls the digital power supply that powers the peltier element, which completes the control loop. This setup is similar to the one used in [20], but more stable and flexible in terms of temperature set points.

A MicroBlaze processor on the FPGA is used as the test driver. In each iteration it: 1) writes the data pattern into the entire memory, 2) sets the refresh interval for the test, 3) waits for 20 seconds, 4) sets the refresh interval back to the data sheet value, 5) and then reads and checks the data, while reporting discovered errors to the PC over a UART link. Tests are repeated for varying data patterns, refresh intervals, and temperatures, similar to the 3D-IC WIDE I/O DRAM experiments.

IV. EXPERIMENTAL RESULTS

We conduct a set of experiments to measure the retention times and bit error rates of WIDE I/O DRAMs and DDR3 DRAMs. We executed the tests multiple times to manifest the results. First, the measurements based on the WIOMING 3D-IC are presented. Then, the results of the FPGA-based DDR3 DRAM test system are shown.

A. WIDE-I/O DRAM Measurement Results

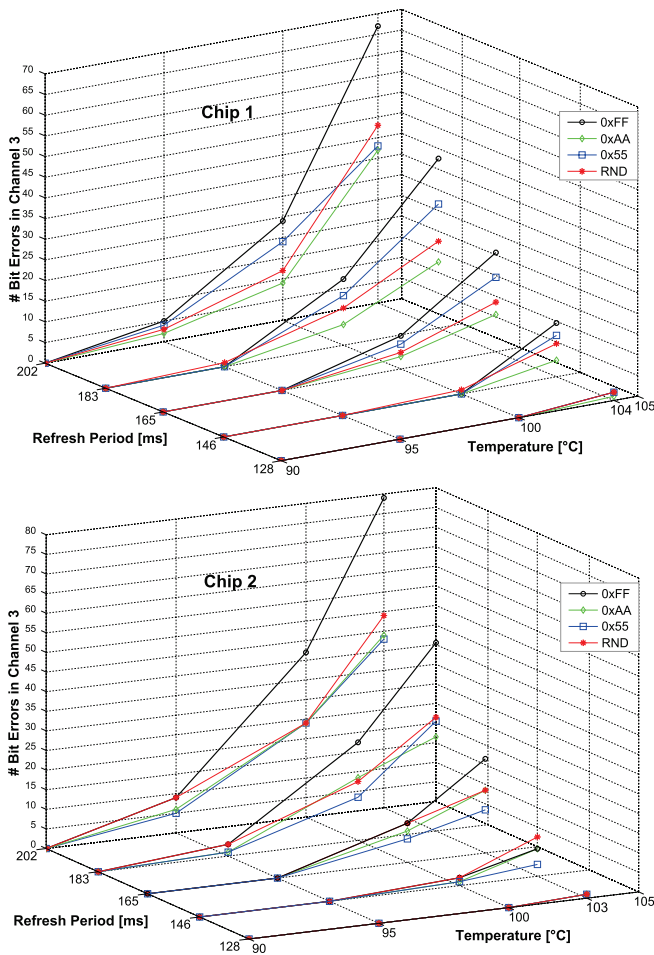


Fig. 9: Measurement Results with Fixed Refresh Periods using test b) for Chip 1 and 2

Figure 9 shows the measurement results of the two different chips. We see clearly in the two plots the data dependence of the error rates. Although the two chips behave differently for the exact error numbers, the overall trend can be seen in Table II. The 0xFF pattern detects $\approx 83\%$ of the retention errors, the others are able to find only 46-52%. The additional error coverage of the test pattern for each chip compared to the 0xFF test is shown in the rows “Add. cov. errors” in

TABLE II: Coverage at a Refresh Period of 202ms and $T=103/104^\circ\text{C}$

Data pattern	0xFF	0xAA	0x55	RND
Chip 1 (%)	82	45	46	52
Add. cov. errors (%)	0	9.8	0	11.0
Chip 2 (%)	84	47	46	52
Add. cov. errors (%)	0	6.4	4.3	7.4

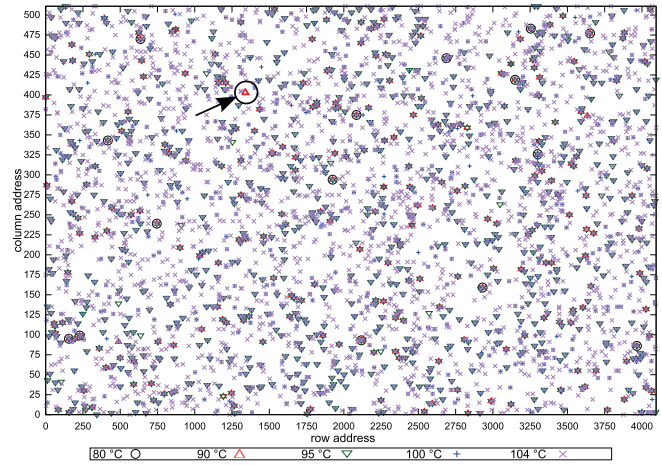


Fig. 10: Scatter-plot of accumulated Bit Errors at different Temperatures for Channel 3, Bank 1 with data pattern 0xFF and test a)

percentage. The RND and the 0xAA test contribute most to increase the coverage.

In Figure 10 we see the effect of VRTs for DRAM cells. Not all cells are permanent failing beginning from a specific temperature. For instance, the red triangle (90°C) fail shown in the circle is a typical VRT fail, as it disappears at higher temperature.

Figure 11 compares the two chips with respect to temperature dependence and variable retention times. Both chips behave in a similar way. However, the error rates of chip 2 are slightly higher. The orange part of the plot shows (also in percentage below) the fraction of error addresses not occurring at the highest temperature (e.g. 104°C and 103°C). We repeated our measurements for data pattern 0xFF and Channel 3 multiple times (10 \times). We observe an increased coverage (catching more error addresses) of $\approx 6\%$ when executing the test 10 \times . Similar results were reported in [6]; however, not for stacked WIDE I/O DRAMs.

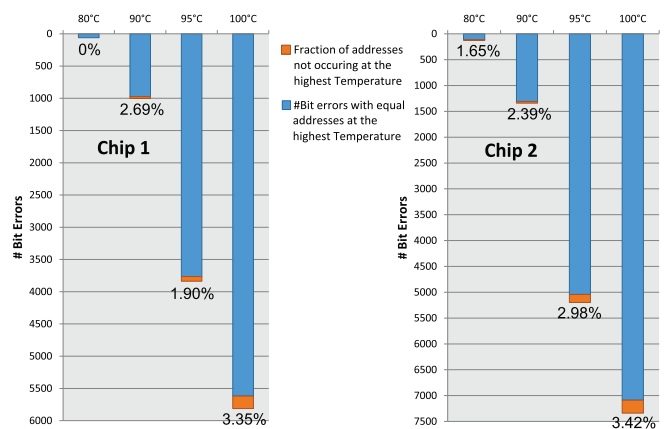


Fig. 11: Bit Errors at Channel 3 using test a) with data pattern 0xFF for both Chips, showing the VRT effect with increasing Temperature

TABLE III: Coverage of DDR3 test pattern at a Refresh Period of 1s and 2s and T=85°C

Refresh Period = 1s					
Data pattern	0xFF	0xAA	0x55	RND	0x00
Coverage (%)	43	23	28	27	28
Bit Errors (#)	47	25	31	30	31
Add. cov. errors wrt. 0xFF (#)	0	9	14	30	31
Unique errors found (#)	14	0	0	30	8
Refresh Period = 2s					
Data pattern	0xFF	0xAA	0x55	RND	0x00
Coverage (%)	39	25	27	31	28
Bit Errors (#)	471	302	328	380	341
Add. cov. errors wrt. 0xFF (#)	0	137	144	380	341
Unique errors found (#)	122	4	7	380	71

B. DDR3 DRAM Measurement Results

In Figure 12 the results of the DDR3-DRAM measurements using the FPGA test system are shown. We clearly see the data dependence as well and the trends for the different test pattern compared to the WIDE I/O chips are similar. However, we observe an interesting difference that we see bit flips from “0” to “1”, which we never have seen for the measured WIDE I/O DRAM devices. Thus, we use additionally a 0x00 data pattern to analyse these error addresses.

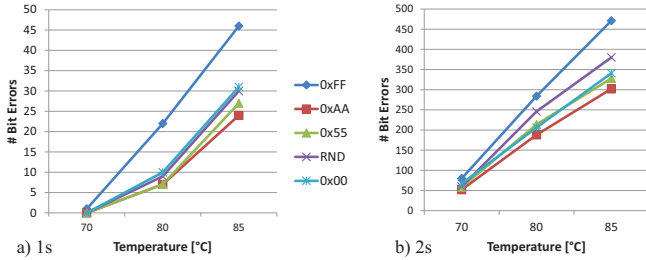


Fig. 12: Bit Errors of a DDR3-DRAM at different temperatures and refresh periods (1s and 2s)

The root cause of the 0 to 1 bit flip behaviour of the DDR3 device [19] is the different array topology. There is a high probability that Micron’s DRAM array design uses true-cells and anti-cells, where a true-cell stores the data value as it is and an anti-cell stores the inverse [6]. We see also in Table III that the number of bit errors detected by a 0x00 data pattern are nearly as high as the number of errors found by a 0xFF test pattern. Additionally, the high coverage of the 0x00 data pattern indicates the existence of such an array topology. The coverage of the test with data pattern 0xFF alone is relatively low compared to the WIDE I/O DRAM results (39-43% versus 83% for WIDE I/O). In terms of coverage gain, again the RND pattern together with the already mentioned 0x00 data pattern adds a significant amount of improvement in percentage (up to 99% in total). The causes of the phenomenon data pattern dependence, such as bitline-bitline, bitline-cell and bitline-wordline coupling, are highlighted by the unique errors found during the RND pattern test in the DDR3 device (see Table III). Although the exact DRAM array topology is unknown as it is property of each DRAM vendor, we consider the status of the neighbouring cells with an assumed array topology as essential influence to the data pattern dependence in our model, which we present in the next Section.

V. DRAM RETENTION TIME ERROR MODEL

A retention error aware DRAM model is key to analyse the impact of lower refresh rates on the executed application. Especially for error resilient applications this can be exploited, to save energy. The obtained measurement results of bit error rates based on cell retention fails, shown in Section IV, permit the creation of our DRAM retention time error model.

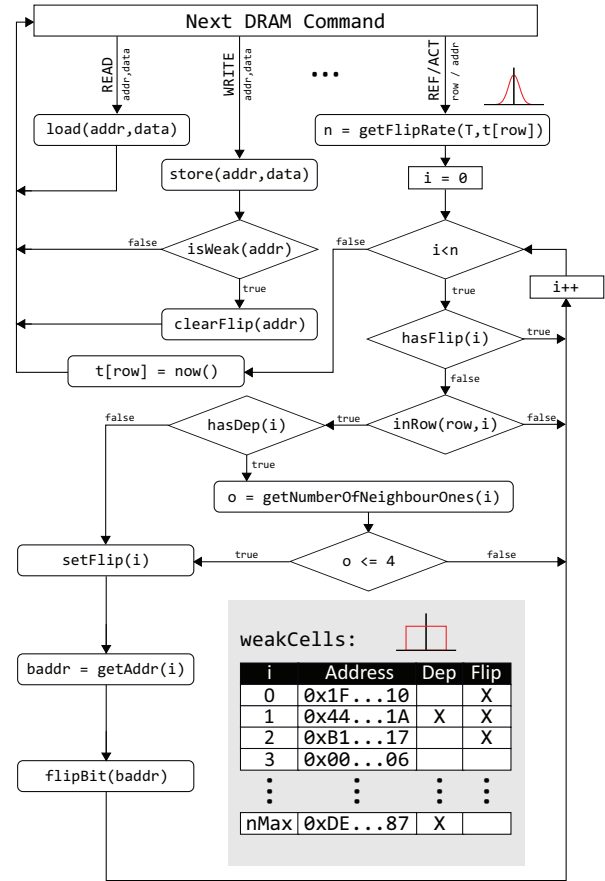


Fig. 13: Algorithm of the Retention Error Model for one DRAM Bank

Currently, the model is calibrated to the measurement results of the WIDE I/O DRAM. Thus, we use bit flips from 1 to 0, but we are not limited to this and a DDR3 like behaviour using the measurement data of the Micron device can be modelled as well.

Our proposed model is developed in C++. It is integrated in our advanced SystemC-TLM2.0 virtual-platform setup [21]. However, it can also be integrated in other simulation environments like *gem5*. Figure 13 explains the algorithm of the model.

First, we describe the basic functions of the algorithm and second, we explain the algorithm itself. The data of the memory is stored in an associative array, so that the simulator do not need to allocate the full amount of memory on the host machine. The data can be accessed with the `store(addr, data)` and `load(addr, data)` functions. With the `flipBit(baddr)` function a specific bit in a column can be flipped.

For different temperatures and refresh periods, the model gets as input a list of error numbers (mean and sigma values obtained by measurements). The mean and sigma input values are sent to a Gaussian random number generator to vary the error numbers to reproduce the effect of VRT. The results are stored in a look-up-table, which can be accessed during simulation with the function `getFlipRate`, which receives the current temperature T and the time of the last refresh or activate command on this particular row as $t[row]$ and returns the number of bit errors. The `nMax` variable is set to the maximum error value of the look-up-table, which is the actual maximum number of weak cells in the model.

At the begin of the simulation, the positions of the weak cells are selected by generating random addresses with an uniform distributed random number generator (see Figure 10). The result is stored in the table `weakCells` (see Figure 13). Each weak cell has a specific number `i`. The address of a weak cell can be obtained by calling the function `getAddr(i)`. Then, a configurable amount of weak cells are randomly selected to be a data-dependent cell. This is marked with the flag `Dep` in the table and can be tested with the function `hasDep(i)`. The flag `Flip` in the table `weakCells` indicates if this weak cell has already flipped or not. It can be tested with the function `hasFlip(i)`, can be set with `setFlip(i)` and can be cleared with `clearFlip(i)`, respectively. To identify if an arbitrary address is an element of the table, the function `isWeak(addr)` is used. The function `getNumberOfNeighbourOnes(i)` returns the number of ones in the direct neighborhood of a specific weak cell. If a weak cell lies in a specific row can be tested with the function `inRow(row, i)`.

The model is hooked in a DRAM simulator in the places of read, write, act and refresh commands for each bank:

READ: if the memory controller sends a read command to a bank the data is obtained by calling the `load` function.

WRITE: if a write command is send to a bank, it is checked if a weak cell is in the destination address. If this is the case the `Flip` flag is cleared (`clearFlip`), since new data is stored (`store`) at this address.

REFRESH and ACTIVATE: the point in time where a bit flip in the memory is manifested irreversibly, is when the memory controller issues an activate or refresh command, since the content of the cell is sensed and amplified. Because of this fact, we decided to hook the bit flip logic of the model into this place of the simulator. At the begin of these commands, it is checked how large the current failure rate (n) is for the given temperature and the last access to this row. The first n weak cells (loop) in the table `weakCells` will be flipped (`flipBit`) and marked (`setFlip`) by setting the `Flip` flag in the table. However, for data-dependent weak cells the neighbourhood of the weak cell is analysed (`getNumberOfNeighbourOnes`). Only if the number of ones in the neighbourhood is under a certain threshold the cell is flipped and marked. A flipped bit stays flipped until a new write command will overwrite the data in the cell.

Figure 14 shows the comparison of the averaged results of $30\times$ repeated model simulations and real measurements of the WIDE I/O DRAM. We see that our model implements the correct trend for the data pattern dependence and has bit error rates near to the measured values. The overhead of the retention-aware DRAM bit error model with respect to the simulation execution time is in average only 30%. Thus, our proposed model is suitable for the analysis of temperature-dependent retention errors in DRAMs for future computing systems.

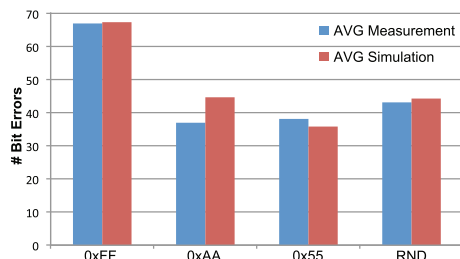


Fig. 14: Comparison of Simulation and Measurements for a Refresh Period of 202ms and a Temperature of 90 °C

VI. CONCLUSION

We measured the temperature and refresh period depending DRAM bit error rates of WIDE I/O and DDR3 DRAMs. We obtained the accurate temperature distribution map of the WIDE I/O DRAM die. Further, we observed data pattern dependencies and VRTs during the measurement of the bit error rates. Based on this measurement data and our analysis we introduced a SystemC-TLM2.0 retention-aware DRAM bit error model. Our proposed DRAM bit error model enables early investigations and explorations on the temperature vs. retention time trade-off in future 3D-stacked MPSoCs with WIDE I/O DRAMs in SystemC-TLM2.0 environments.

ACKNOWLEDGEMENTS

The authors thank the companies DOCEA Power and SYNOPSIS for their great support. This work was partially funded by the German Research Foundation (DFG) as part of the priority program Dependable Embedded Systems SPP 1500 (<http://spp1500.itec.kit.edu>), the DFG grant no. WE2442/10-1, and by projects EU FP7 288248 Flexiles, CA505 BENEFIC and ARTEMIS 621429 EMC2.

REFERENCES

- [1] T. Hamamoto, et al. *On the retention time distribution of dynamic random access memory (DRAM)*. Electron Devices, IEEE Transactions on, 45(6):1300–1309, Jun 1998.
- [2] J. Liu, et al. *RAIDR: Retention-Aware Intelligent DRAM Refresh*. In Proc. of ISCA, 2012.
- [3] J.-H. Ahn, et al. *Adaptive Self Refresh Scheme for Battery Operated High-Density Mobile DRAM Applications*. In ASSCC. IEEE Asian, Nov 2006.
- [4] R.K. Venkatesan, et al. *Retention-aware placement in DRAM (RAPID): software methods for quasi-non-volatile DRAM*. In Proc. of HPCA 2006.
- [5] K. Kim et al. *A New Investigation of Data Retention Time in Truly Nanoscaled DRAMs*. Electron Device Letters, IEEE, Aug 2009.
- [6] J. Liu, et al. *An Experimental Study of Data Retention Behavior in Modern DRAM devices: Implications for Retention Time Profiling Mechanisms*. SIGARCH Comput. Archit. News, 41, 2013.
- [7] H. Kim, et al. *Study of Trap Models Related to the Variable Retention Time Phenomenon in DRAM*. Electron Devices, IEEE Transactions on, June 2011.
- [8] H. Kim, et al. *Characterization of the Variable Retention Time in Dynamic Random Access Memory*. Electron Devices, IEEE Transactions on, Sept 2011.
- [9] C. Shirley et al. *Copula Models of Correlation: A DRAM Case Study*. Computers, IEEE Transactions on, Jun 2013.
- [10] D. Dutoit, et al. *A 0.9 pJ/bit, 12.8 GByte/s WideIO memory interface in a 3D-IC NoC-based MPSoC*. In VLSIC, Symposium on, June 2013.
- [11] J.-S. Kim et al. *A 1.2 V 12.8 GB/s 2 Gb Mobile Wide-I/O DRAM With 4*128 I/Os Using TSV Based Stacking*. IEEE Journal of Solid-State Circuits, 47, 2012.
- [12] K.-Y. Tsai et al. *Thermal characterization of a wide I/O 3DIC*. In IMPACT Conference, 6th International, pages 261–264, 2011.
- [13] B. Schroeder, et al. *DRAM Errors in the Wild: A Large-scale Field Study*. In Proceedings of, SIGMETRICS, NY, USA, 2009. ACM.
- [14] H. Shin. *Modeling of alpha-particle-induced soft error rate in DRAM*. Electron Devices, IEEE Transactions on, Sep 1999.
- [15] DOCEA Power. *AceExplorer: "http://www.doceapower.com/products-services/aceplorer.html"*, 2014.
- [16] C. Santos, et al. *System-level thermal modeling for 3D circuits: Characterization with a 65nm memory-on-logic circuit*. In 3D Systems Integration Conference (3DIC), 2013 IEEE International, Oct 2013.
- [17] S. Goossens, et al. *A Reconfigurable Real-time SDRAM Controller for Mixed Time-criticality Systems*. In Proc. CODES+ISSS, 2013.
- [18] Xilinx Inc. *ML605 Documentation, "http://www.xilinx.com/support/documentation/boards_and_kits/ug533.pdf"*, 2014.
- [19] Micron. *512MB (x64, Single Rank) 204-Pin DDR3 SDRAM SODIMM, MT4J5F6464H*, 2009.
- [20] K. Chandrasekar, et al. *Exploiting Expendable Process-margins in DRAMs for Run-time Performance Optimization*. In Proc. DATE, 2014.
- [21] M. Jung, et al. *TLM Modelling of 3D Stacked Wide I/O DRAM Subsystems: A Virtual Platform for Memory Controller Design Space Exploration*. In Proc. RAPIDO, NY, USA, 2013. ACM.