Ph.D. Dissertation

Design of Low Power Memory Controller with Adaptive Eye Detection Algorithm

적응형 눈 감지 방법을 포함한
저전력 메모리 컨트롤러의 설계

by

Mino Kim

August 2017

Department of Electrical Engineering and
Computer Science
College of Engineering
Seoul National University

ABSTRACT

Design of Low Power Memory Controller with Adaptive Eye Detection Algorithm

Mino Kim
Department of Electrical Engineering and Computer Science
College of Engineering
Seoul National University

A 4266Mb/s/pin LPDDR4 memory controller with an asynchronous feedback continuous-time linear equalizer and an adaptive 3-step eye detection algorithm is presented. The asynchronous feedback continuous-time linear equalizer removes the glitch of DQS without training by applying an offset larger than the noise, and improves read margin by operating as a decision feedback equalizer in DQ path. The adaptive 3-step eye detection algorithm reduces power consumption and black-out time in initialization sequence and retraining in comparison to the 2-dimensional full scanning. In addition, the adaptive 3-step eye detection algorithm can maintain the accuracy by sequentially searching the eye boundaries and initializing the resolution using the binary search method when the eye detection result changes. To achieve high bandwidth, a transmitter and

receiver suitable for training are proposed. The transmitter consists of a phase interpolator, a digitally-controlled delay line, a 16:1 serializer, a pre-driver and low-voltage swing terminated logic. The receiver consists of a reference voltage generator, a continuous-time linear equalizer, a phase interpolator, a digitally-controlled delay line, a 1:4 deserializer, and a 4:16 deserializer. The clocking architecture is also designed for low power consumption in idle periods, which are commonly lengthy in mobile applications. A prototype chip was implemented in a 65nm CMOS process with ball grid array package and tested with commodity LPDDR4. The write margin was 0.36UI and 148mV; and the read margin was enhanced from 0.30UI and 76mV without AF-CTLE to 0.47UI and 80mV to with AF-CTLE. The power efficiency during burst write and read were 5.68pJ/bit and 1.83pJ/bit respectively.

Keywords: LPDDR4, mobile memory, memory controller, memory interface, transceiver, adaptive eye detection algorithm, auto-DQS cleaning

Student Number: 2010-20764

CONTENTS

ABSTRACT ..... **I**

CONTENTS ..... **III**

LIST OF FIGURES ..... **V**

LIST OF TABLE ..... **VIII**

CHAPTER 1 INTRODUCTION **1**

1.1 MOTIVATION 1

1.2 THESIS ORGANIZATION 5

CHAPTER 2 LPDDR4 ..... **6**

2.1 COMPARISON BETWEEN LPDDR3 AND LPDDR4 6

2.2 SOURCE SYNCHRONOUS CLOCKING SCHEME 9

2.3 SIGNALING STANDARDS 11

2.4 MULTIPLE TRAININGS 14

2.5 RE-TRAINING AND RE-INITIALIZATION 16

CHAPTER 3 ADAPTIVE EYE DETECTION **18**

3.1 EYE DETECTION 18

3.2 1X2Y3X EYE DETECTION 20

3.3 ADAPTIVE GAIN CONTROL 22

3.4 ADAPTIVE 1X2Y3X EYE DETECTION 24

CHAPTER 4 LPDDR4 MEMORY CONTROLLER **26**

4.1 DESIGN PROCEDURE 26

4.2 ARCHITECTURE 30

4.2.1 TRANSMITTER 33

4.2.2 RECEIVER 35

4.2.3 CLOCKING ARCHITECTURE 38
4.3 CIRCUIT IMPLEMENTATION 43
4.3.1 ADPLL WITH MULTI-MODULUS DIVIDER 43
4.3.2 ADDLL WITH TRIANGULAR-MODULATED PI 45
4.3.3 CTLE WITH AUTO-DQS CLEANING 47
4.3.4 DES WITH CLOCK DOMAIN CROSSING 52
4.3.5 LVSTL WITH ZQ CALIBRATION 54
4.3.6 COARSE-FINE DCDL 56
4.4 LINK TRAINING 57
4.4.1 SIMULATION RESULTS 59
CHAPTER 5 MEASUREMENT RESULTS 72
5.1 MEASUREMENT SETUP 72
5.2 MEASUREMENT RESULTS OF SUB-BLOCK 80
5.2.1 ADPLL WITH MULTI-MODULUS DIVIDER 80
5.2.2 ADDLL WITH TRIANGULAR-MODULATED PI 82
5.2.3 COARSE-FINE DCDL 84
5.3 LPDDR4 INTERFACE MEASUREMENT RESULTS 84
CHAPTER 6 CONCLUSION 88
BIBLIOGRAPHY 90

LIST OF FIGURES

Figure 1.1.1 3-types of dynamic random-access memory. 2
Figure 1.1.2 Mobile applications requirements market research. 3
Figure 2.1.1 Comparison between LPDDR3 and LPDDR4. 7
Figure 2.1.2 Architecture of LPDDR3 and LPDDR4. 8
Figure 2.2.1 Source synchronous (a) matched scheme and (b) unmatched scheme. 10
Figure 2.3.1 (a) high speed unterminated logic and (b) pseudo open drain logic. 11
Figure 2.3.2 Low-voltage swing terminated logic 12
Figure 2.3.3 Comparison of HSUL, PODL, LVSTL 13
Figure 2.5.1 DRAM architecture with DLL 16
Figure 3.1.1 Two-dimensional eye detection 19
Figure 3.2.1 1x2y3x iteration eye detection. 21
Figure 3.3.1 Simplified block diagram of adaptive gain control scheme 22
Figure 3.3.2 Eye detection with adaptive gain control scheme 23
Figure 3.4.1 Adaptive 1x2y3x eye detection 24
Figure 3.4.2 Comparison of eye detection algorithms 25
Figure 4.1.1 Design procedure of memory controller 27
Figure 4.2.1 Architecture of proposed LPDDR4 memory controller 30
Figure 4.2.2 Block diagram of transmitter 33
Figure 4.2.3 Block diagram of receiver 35
Figure 4.2.4 Mobile DRAM state ratio 38
Figure 4.2.5 Clocking architecture of CK and CA 39
Figure 4.2.6 Clocking architecture of DQ/DQS TX 40
Figure 4.2.7 Clocking architecture of DQ/DQS RX 41
Figure 4.2.8 Block diagram of transceiver at IDLE state 42
Figure 4.3.1 Block diagram of the ADPLL 43

Figure 4.3.2 Block diagram of digitally controlled oscillator 44
Figure 4.3.3 Block diagram of ADDLL 45
Figure 4.3.4 Operation of the folded delay line.....46
Figure 4.3.5 Block diagram of triangular-modulated PI.....47
Figure 4.3.6 Timing diagram of DQS glitch issue and gate training.....48
Figure 4.3.7 Gate training fail with time variance.....49
Figure 4.3.8 DQS clean with 2x receiver 50
Figure 4.3.9 Timing diagram of automatic gate 50
Figure 4.3.10 Block diagram of CTLE with asynchronous feedback 51
Figure 4.3.11 Timing diagram of CTLE with asynchronous feedback.....52
Figure 4.3.12 Block diagram of 1:16 DES 52
Figure 4.3.13 Timing diagram of 1:16 DES 53
Figure 4.3.14 Block diagram of pull-down ZQ calibration 55
Figure 4.3.15 Block diagram of pull-up ZQ calibration.....55
Figure 4.3.16 Block diagram of coarse-fine DCDL 56
Figure 4.4.1 Initialization sequence of LPDDR4 57
Figure 4.4.2 Operation sequence of LTFSM 58
Figure 4.4.3 Boot up sequence - initialization step 1 60
Figure 4.4.4 Boot up sequence - initialization step 2 60
Figure 4.4.5 Boot up sequence - initialization step 3 61
Figure 4.4.6 Timing of command training entry 62
Figure 4.4.7 Training pattern of CS training 62
Figure 4.4.8 Setup and hold timing margin for reference voltage sweep at command training 63
Figure 4.4.9 Reference voltage sweep in CS training 64
Figure 4.4.10 1x2y3x eye detection algorithm in CA training 65
Figure 4.4.11 Training pattern of CA training: 0-A-0-B-0-C-0-D-0-E-0-A-⋯.....66
Figure 4.4.12 Result of the command training 66
Figure 4.4.13 Exit timing of the command training 67

Figure 4.4.14 Operation speed change in the end command training 67
Figure 4.4.15 Entry timing of the write leveling 68
Figure 4.4.16 Write leveling 68
Figure 4.4.17 Training code sweep of the read training 69
Figure 4.4.18 Environment of the read training 69
Figure 4.4.19 Result of the read training 70
Figure 4.4.20 Training code sweep of the write training 71
Figure 4.4.21 Write training pattern 71
Figure 5.1.1 Chip microphotograph 73
Figure 5.1.2 Package on package example 73
Figure 5.1.3. Iphone 4 PCB 74
Figure 5.1.4 Test board with placing MCU and LPDDR4 memory side by side 74
Figure 5.1.5 Top (left) and 3rd (right) layer of BGA package 75
Figure 5.1.6 2nd (left) and bottom (right) layer of BGA package 76
Figure 5.1.7 LP4_TEST board 77
Figure 5.1.8 LP4_TP board (left) and LP4_noTP board (right) 78
Figure 5.1.9 Photograph of measurement setup 79
Figure 5.1.10 Block diagram of measurement setup 79
Figure 5.2.1 Measurement results of ADPLL 81
Figure 5.2.2 Performance summary of ADPLL 81
Figure 5.2.3 Measurement results of ADDLL at 266, 533, 800 and 1066MHz 82
Figure 5.2.4 Measurement results of ADDLL at 1333, 1600, 1866 and 2133MHz 83
Figure 5.2.5 Measured DNL of ADDLL 83
Figure 5.2.6 Measurement results of DCDL 84
Figure 5.3.1 Measurement results of trainings 85
Figure 5.3.2 Shmoo plot of WRITE and READ margin test at 4266Mb/s/pin 86

LIST OF TABLE

Table 5.3.1 Performance Comparison of LPDDR4 Memory Controller.....87

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION

Dynamic random-access memory (DRAM) is divided into computing DRAM, graphic DRAM, and mobile DRAM depending on the application, as shown in Figure 1.1.1. In the early days of the DRAM market, computing DRAM and graphics DRAM led growth, but now they are reaching their growth limit. In recent years, due to the increase in demand for smartphones and tablet PCs, mobile DRAM has led the growth of DRAM market size.

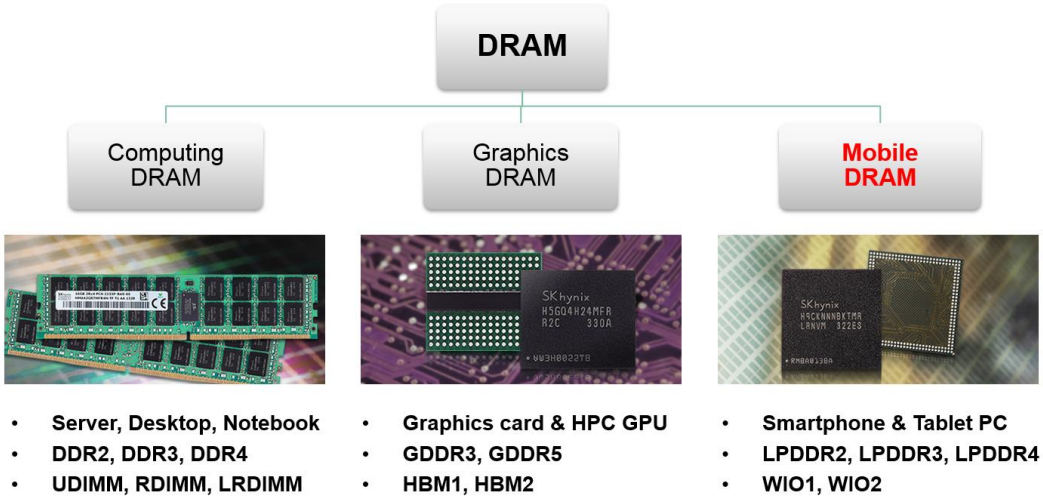


Figure 1.1.1 3-types of dynamic random-access memory.

As memory-intensive, high-quality gaming modes and video codecs become widespread in battery-powered mobile systems, there is a constant need for high-speed, low-power, mobile DRAM and controller. According to market research, as shown in Figure 1.1.2, the requirements for mobile applications are 42% for speed, 30% for battery, 18% for price, 8% for user interface and 2% for size, with high bandwidth being the highest and low power being the next [1.1.1]. To meet this demand, the bandwidth of mobile DRAM rapidly increased to 400 Mbps [1.1.2] (LPDDR) in 2007, 800 Mbps (LPDDR2) in 2012, 1600 Mbps [1.1.3] (LPDDR3) in 2013 and 4266 Mbps [1.1.4] (LPDDR4) in 2015, respectively.

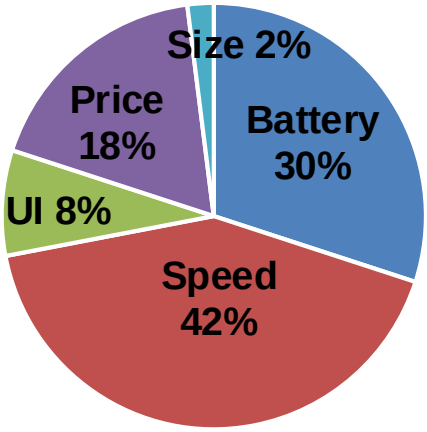


Figure 1.1.2 Mobile applications requirements market research.

Various major architectural changes that contribute to performance improvements have been implemented in LPDDR4. For example, in the case of a data (DQ) pin, the clock scheme of data strobe (DQS) was changed from the source synchronous matched scheme (SSMS) to the unmatched scheme (SSUS), the I/O interface has been changed from high-speed unterminated logic (HSUL) to low-voltage swing terminated logic (LVSTL), DQ on-die termination (ODT) become necessary, and internal reference voltage (VREF) is adopted. In the case of a command/address (CA) pin, the I/O interface has been changed from pseudo open drain logic (PODL) to low-voltage swing terminated logic (LVSTL), double data rate (DDR) becomes single data rate (SDR) and VSS termination is added.

However, the DRAM process is relatively bad compared to the logic process, and therefore it is insufficient to allow the memory to operate at high bandwidth on its own through architectural changes. To overcome this, various training operation such as ZQ calibration, CA training, DQ training, and latency training have been added to the

controller-side to improve the performance. In addition, periodic training is required to compensate for voltage and temperature variations. As a results, the bandwidth is getting higher, such trainings in controller-side become more complex, diverse, and important. Therefore, algorithms that can be trained simply and quickly effectively must be proposed.

Back to the story of the memory controller unit (MCU) and DRAM, the low-power high bandwidth is equally required for both MCU and DRAM. Therefore, not only the training algorithm but also low-power, high-bandwidth architecture and transceiver (TRX) design are very important in MCU. Since the structure and design direction of the MCU should be slightly different depending on the type of the corresponding memory, the corresponding application should be determined and designed accordingly. Considering future development potential and performance, LPDDR4, one of mobile DRAM, can be a good candidate.

Thus, in this thesis, we propose adaptive eye detection algorithm which is a simple and fast compared to the conventional eye detection algorithm and low power memory controller for LPDDR4. The proposed low power memory controller is also applicable to other DRAM applications through slight modifications to spec-sensitive parts.

1.2 THESIS ORGANIZATION

This thesis is organized as follows. First, in Chapter 2, the basics and major specification of the LPDDR4 are explained. Additionally, the discussion about the difference between LPDDR4 and LPDDR3 is presented. Discussing the LPDDR4 specifications, we understand the issues and the types of training that should be considered in the controller design. In Chapter 3, the thesis presents the adaptive 1x2y3x eye center detection algorithm with comparison of other previous eye detection algorithms. In Chapter 4, low power memory controller architecture with TRX, clocking architecture and detail sub blocks will be explained. In Chapter 5, measurement setup and experimental results are described. Finally, in Chapter 6, the thesis is summarized with the discussion of contribution.

CHAPTER 2

LPDDR4

2.1 COMPARISON BETWEEN LPDDR3 AND LPDDR4

LPDDR4 looks to lead the next generation of mobile DRAM as it operates over 4266 Mbps and achieves large power reduction per bandwidth, without such high-cost process overhead as wide I/O and through-silicon-via (TSV). The key feature of LPDDR4 is low power consumption with high bandwidth than previous LPDDR3. Figure 2.1.1 shows the comparison between LPDDR3 and LPDDR4. The difference between LPDDR3 and LPDDR4 shows which ones were key to achieving high bandwidth low power. The speed of CA is maintained by changing DDR signaling to SDR signaling. Supply voltage is reduced from 1.2V to 1.1V. There have been many changes in the architectural aspects. The 1-channel structure has been changed to a 2-channel structure, and the DQS scheme has been more simply changed from the SSMS to the SSUS. In case of I/O interface, LPDDR3 adopts HSUL for DQ and PODL for CA, LPDDR4 replace the HSUL and PODL for each DQ and CA into LVSTL. In the case of ODT, all of the optional ones have been changed to mandatory, and the VDDQ termination and the VSSQ termination have been made

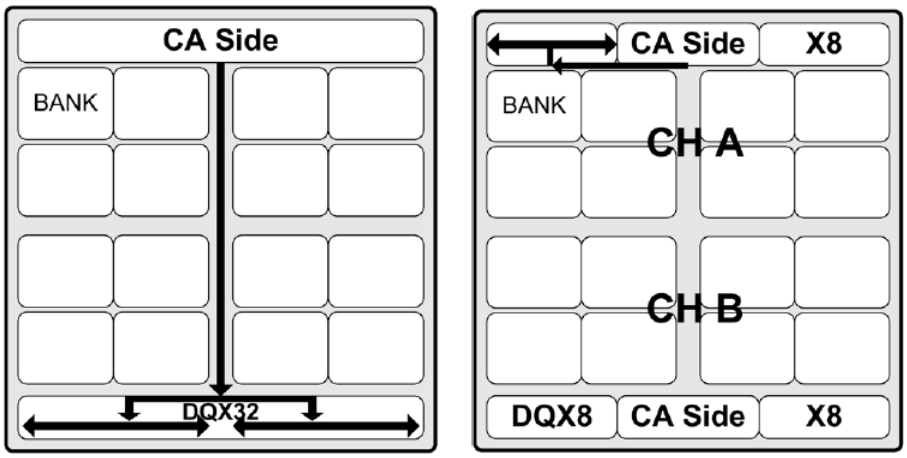
respectively for DQ CA. LPDDR4 started to provide internal VREF to improve the controller's performance. Both LPDDR3 and LPDDR4 do not use delay-locked loop (DLL) or phase-locked loop (PLL) to satisfy low power. In LPDDR4, as the speed increases and the absence of DLL increases, the proportion of training becomes relatively large.

Item		LPDDR3	LPDDR4
Speed	CLK	~1066MHz	~2133MHz
	CMD/ADDR	DDR	SDR
	DQ	DDR	DDR
Voltage	VDD2/VDDQ/VDD1	1.2/1.2/1.8	1.1/1.1/1.8
Architecture	# of Channel	1-channel	2-channel
	DQS scheme	Source Synchronous Matched Scheme	Source Synchronous Unmatched Scheme
	I/O interface	HSUL/PODL	LVSTL
	DQ ODT	No Term(VDDQ Term option)	VDDQ Term
	CA ODT	No Term	VSS Term
	VREF	External	Internal
	DLL or PLL	No DLL or PLL	No DLL or PLL
Training	CMD/ADDR	No	Yes
	DQ write	Optional	Mandatory
	DQ read	Optional	Mandatory

Figure 2.1.1 Comparison between LPDDR3 and LPDDR4.

The largest structural change in LPDDR4 is the change from a 1-channel architecture to a 2-channel architecture per die as shown in Figure 2.1.2. In 2-channel architecture, there are two independent devices on single die. The single channel of LPDDR4 consist of a uni-directional differential CK pins, 7 unit-directional single ended CA pins, 4 bi-directional

differential DQS pins, 16 bi-directional single ended DQ pins, 2 bi-directional single ended DMI pins and other control pins such as RESET, ODT_CA, and ZQ_CAL. These two devices share the ZQ and RESET signal, and use the other I/O and power pins which are significantly affected by the loading. In LPDDR3, there were CA bus and DQ bus on the top side and bottom side, respectively. Therefore, there is a problem that the signal distribution is composed across the entire chip, vulnerable to PVT variation and causing large latency. In addition, mobile DRAMs are more vulnerable to these issues because they do not use delay-locked loop (DLL) to reduce power consumption. In LPDDR4, CA bus and DQ bus are placed on each channel, reducing the length of the data path by 30% on average, and the CA signal distribution across the DRAM die disappears. By adopting such a two-channel structure, the current consumption savings of 15% can be achieved.

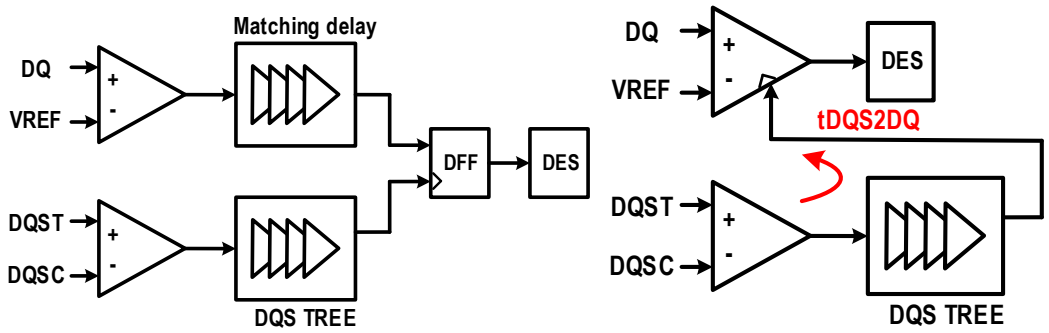


(a) (b)

Figure 2.1.2 Architecture of LPDDR3 and LPDDR4.

2.2 SOURCE SYNCHRONOUS CLOCKING SCHEME

In the communication between the DRAM and the MCU, there is a time-of-flight according to the channel environment. This time-of-flight affects not only the pin-to-pin skew but also the jitter correlation, thus reducing the data eye. A source synchronous clocking scheme is an effective scheme to compensate the unpredictable time-of-flight variation between DRAM and MCU. It is a method of solving the above-mentioned eye reduction by transmit both DQ and DQS for sampling the DQ, thereby sampling the signal affected by the same time-of-flight. Figure 2.2.1 shows the source synchronous clocking scheme of LPDDR3 and LPDDR4. In LPDDR3 IO, an extra delay element is added to the DQ signal path to match the DQS signal path delay as shown in Figure 2.2.1 (a). The matching delay of DQS TREE is called the timing parameter of the tDQS2DQ. These delay units must have full-speed bandwidth and cost large power consumption. In LPDDR4 IO, tDQS2DQ delay element has shifted from DRAM to MCU with the help of a better logic process than the DRAM process. As a result, LPDDR4 can remove replica delays and the timing of the DQS path is controlled by MCU. Therefore, a training process is added to MCU. The Time varying variations such as voltage or temperature can affect tDQS2DQ, which can be solved with a periodic training sequence. The training for compensate tDQS2DQ is WRITE training. In WRITE training in SSUS, additional matching delay logic that can compensate the parameter of tDQS2DQ should be added to DQ path in MCU.



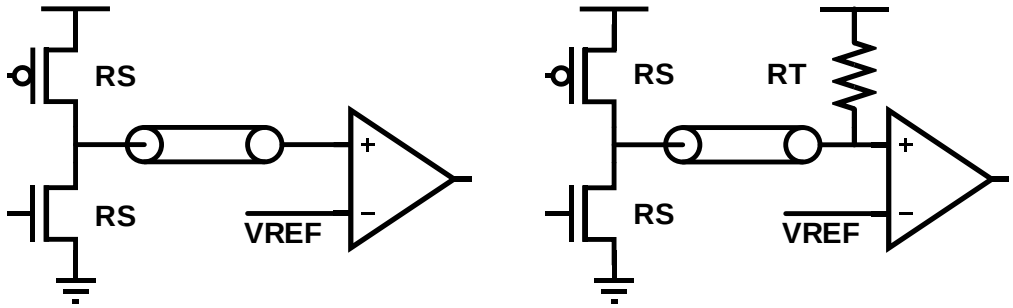
(a) (b)

Figure 2.2.1 Source synchronous (a) matched scheme and (b) unmatched scheme.

In addition, the SSUS architecture simplifies the receiver of the DRAM and allows a powerful equalizing technique called a DFE. By applying DFE, LPDDR4 can achieve higher write timing margin.

2.3 SIGNALING STANDARDS

LPDDR3 adopts the HSUL for DQ and PODL for CA as shown in Figure 2.3.1. Although zero DC power consumption in HSUL, the size of the output swing is too large as VDDQ and it is vulnerable to signal integrity (SI) because it has no termination. PODL, on the other hand, supports VDDQ termination for high speed communication, and there is power consumption at '0' but no power consumption at '1'. By using the DBI technique, power consumption can be reduced by taking advantage of these features.



(b)

Figure 2.3.1 (a) high speed unterminated logic and (b) pseudo open drain logic.

In LPDDR4, to reduce power consumption, LVSTL with VSSQ termination is adopted without using a HSUL or PODL. As shown in Figure 2.3.1, they consist of a pull-down NMOS transistor and a pull-up PMOS transistor. In contrast to HSUL or PODL, as shown in Figure 2.3.2, the LVSTL interface replaced the pull-up PMOS transistor with a NMOS transistor operating in the saturation region. The pull-up NMOS transistor reduce C_{IO} compared to PMOS and reduce power consumption in unterminated mode because it

prohibiting rail-to-rail swing due to threshold voltage drop of pull-up NMOS transistor. VSSQ termination consumes power at ‘1’ but no power at ‘0’. DBI can also be applied to LVSTL to reduce power consumption.

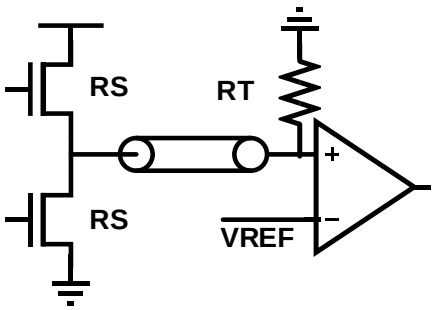
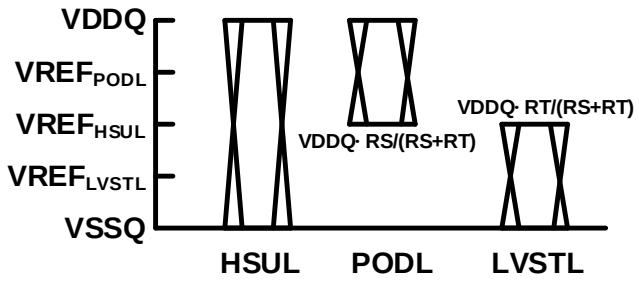


Figure 2.3.2 Low-voltage swing terminated logic

Figure 2.3.3 shows the comparison between HSUL, PODL and LVSTL. In terms of swing and power consumption, PODL and LVSTL are similar. However, the biggest difference between PODL and LVSTL is termination. Since most of the system’s ground impedance is smaller than the supply impedance, VSSQ termination instead of VDDQ or 1/2 VDDQ termination improves both noise immunity and SI characteristics. In addition, VDDQ scaling is available with VSSQ termination.



(a)

	HSUL	PODL	LVSTL
Swing	VDDQ	VDDQ×RT/(RS+RT)	
IH	0	0	VDDQ/(RT+RS)
IL	0	VDDQ/(RS+RT)	0
VREF	VDDQ/2	VDDQ/2× ((2RS+RT)/ (RS+RT))	VDDQ/2× RT/(RS+RT)
Term.	No term.	VDDQ term.	VSSQ term.

(b)

Figure 2.3.3 Comparison of HSUL, PODL, LVSTL

2.4 MULTIPLE TRAININGS

LPDDR4 requires a variety of training performed by the MCU to achieve twice bandwidth of LPDDR3. The types of training are command bus training (CBT), write leveling (WLVL), read training (RDTR), and write training (WRTR). The first training is CBT. The command bus must be trained before enabling termination for high-frequency operation. LPDDR4 provides an internal V_{REFCA} that defaults to a level suitable for un-terminated, low frequency operation. The V_{REFCA} must be trained to achieve suitable receiver voltage margin for terminated, high-frequency operation. At CBT, MCU finds the optimized timing relationship between the CA signal and its strobe clock (CK) and the optimized reference voltage V_{REFCA}. The timing relationship is related to the setup and hold times in the DRAM. The training sequence of CBT is as follows: First, MCU enables the CBT mode at low frequency. When the CBT mode is activated, values latched at the receiver on the CA bus are asynchronously output to the DQ bus. MCU finds the optimum voltage and timing based on the feedback value at high frequency. Finally, MCU exit the CBT mode and change the operation frequency.

The second training is WLVL. WLVL feature compensate CLK-to-DQS timing skew and ensure that data on a DQS domain properly crosses over into a CLK domain in DRAM. In WLVL mode, the DRAM samples the CK with the rising edge of DQS signals, and asynchronously feeds back to the MC. MCU changes the timing of DQS, finds the point where the phase between DQS and CLK crosses, and exit WLVL training.

The third training is RDTR which informs MCU when the read data arrives from

DRAM after read command is issued. The RDTR also allows MCU to recognize voltage and timing parameters relevant in read operation. LPDDR4 has five read FIFO registers to aid RDTR. The read FIFO register help to read and write the specific pattern with only the CA. After the write patterns are written to the registers by write FIFO commands, read FIFO commands can be issued to read the latched data, which is then verified against the original pattern. Since the read FIFO command behaves exactly the same as the read command, it is possible to obtain read latency, optimum reference voltage, and sampling timing at read operation by repeatedly repeating this sequence.

The last training is WRTR which is only possible if all three of previous training are well performed. As mentioned before, LPDDR4 must find the optimum sampling timing and reference voltage level as well as *tDQS2DQ* parameter. Write training proceeds in the following order: Each DQ receives specific write patterns by write command. After the write patterns are written, read commands can be issued to read the data, which is then compared against the original write pattern. This procedure is executed repeatedly for optimum sampling timing of DQ including *tDQS2DQ* and every reference voltage of DQ (V_{REFDQ}) value allowed within the predefined range, which is from 10% to 42% of VDDQ with the resolution at 0.4% of VDDQ. The *tDQS2DQ* must guarantee that it has a value in the range of 200 to 1000ps. MCU requires an absolute delay element to compensate for this timing parameter for write training.

2.5 RE-TRAINING AND RE-INITIALIZATION

In the previous sections, we introduced various architecture and interface changes of LPDDR4 to achieve low power high bandwidth. In addition, these changes are all assumptions of the multiple trainings. There are many reasons why DRAM needs training. One of them is the absence of DLL in DRAM. Figure 2.5.1 shows the conceptual block diagram of clocking with DLL in DRAM. In synchronous DRAM, the output data is aligned with an external clock after internal clock delay and timing delay from the internal clock to the output data. This delay is called *tAC*. Because *tAC* is a function of process, voltage, and temperature (PVT), READ data sampling at the MCU exhibits both static and dynamic variation. When the MCU attempts to capture incoming data bits simultaneously, the valid data window can disappear by time varying variation. The DLL in DRAM can track these variation and compensate delay parameters such as *tAC*. However, in LPDDR4, tDQS2DQ is moved from DRAM to the MCU, allowing improvement by using the application-specific integrated circuit process technology, which is better than that of DRAM.

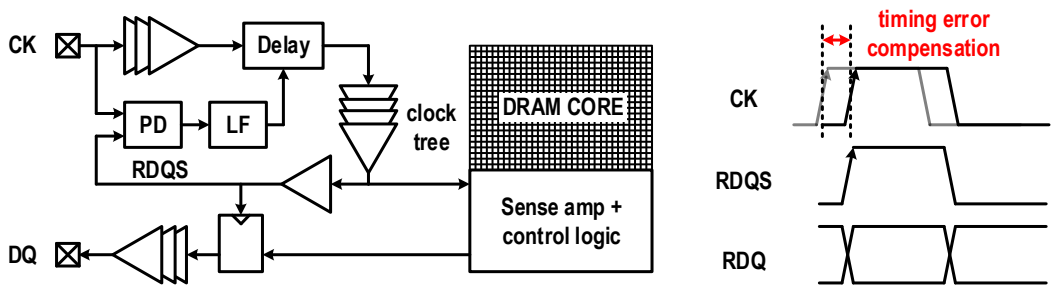


Figure 2.5.1 DRAM architecture with DLL

Since the absence of the DLL in LPDDR4, MCU needs to find other solution to compensate both static and dynamic variation. Once training or initialization can only compensate for static variation, so re-training or re-initialization is required to compensate for both static and dynamic variation. Various methods for efficient re-training have been proposed. One method is the use of built-in self-test circuit. In LPDDR4, it is recommended to use a free-running oscillator to count the oscillator cycle at regular intervals [1.1.5]. There is also a method to detect variation using an on-chip voltage or temperature sensor 2. Another method is to perform re-training or re-initialization periodically [2.5.2]. Due to the volatile nature of DRAM, it is necessary to have an auto refresh interval periodically. In this case, the training interval is defined in terms of the number of auto refresh burst cycles. Taking LPDDR4 as an example, average refresh interval (*tREFI*) is 3.9us. If re-training is performed for each 1024^{th} auto refresh burst cycle, re-training is performed every 4ms. This means that re-training and re-initialization are required fairly often even if the mobile device is not running.

In addition to the DRAM's own performance improvement, the ASIC process technology is inherently superior in performance to the DRAM process, and therefore, the DRAM functions are being handed over to the MCU. As a result, MCUs are replacing DRAM functions through various training. It should be noted that re-training or re-initialization is required to compensate for time varying variation. Various methods have been proposed for efficiently re-training the number of times, but this is not a fundamental solution. We conclude that the best solution to MCU is to apply an effective training algorithm that can be applied to both training and re-training.

CHAPTER 3

ADAPTIVE EYE DETECTION

3.1 EYE DETECTION

The multiple trainings of LPDDR4, described in Chapter 2, are a series of processes in which the MCU detects data coming from/to DRAM. Calling 'A-B' is a training target, MCU changes the timing and reference voltage of B to detect the eye. For the sake of convenience, we will refer to the 'eye center' as the optimal sampling point for the eye. CA training, WLVL, and WRTR look for the eye of the DRAM-side, looking for the eye center of CK-CA, eye left-edge of CK-DQS, and the eye center of DQS-DQ, respectively. RDTR is the process of finding the eye center of the DQ-DQS on the MC-side.

Various eye detection algorithms have been proposed in various application [3.1.1]-[3.1.8]. As shown in Figure 3.1.1, the most common way to find eye center in the eye diagram from various eye diagram is two-dimensional eye detection, which checks all point in a two-dimensional. It moves one point up, down, left or right from the starting point by checking whether the point passed or not, to find all the eye diagrams. The two-dimensional eye detection method is very accurate but has too many test points. The number of test

points required for full scanning eye detection is as follows:

$$N_{TR} = N_{time} \times N_{volt} \tag{3.1.1}$$

where N_{TR} is number of test points, N_{time} is number of time step and N_{volt} is number of voltage step. For example, in the case of LPDDR4 MC, the timing step of the x-axis is 256 points, and the reference voltage of the y-axis is 72 points, there are total of 18432 test points in total. In proportion to the number of test points, the training time, the power consumed in the training, and the register size of the link training finite-state machine (LTFSM) which performs the training increase. The optimal eye center detection algorithm based on the shape of the eye needs to know the overall shape of the eye, so it needs to use the two-dimensional eye detection method, resulting in a lot of time and complex circuitry. In LPDDR4, as mentioned before, low power consumption is the most important goal, so a simple and fast eye detection algorithm is required to replace two-dimensional eye detection even if there is some loss in accuracy.

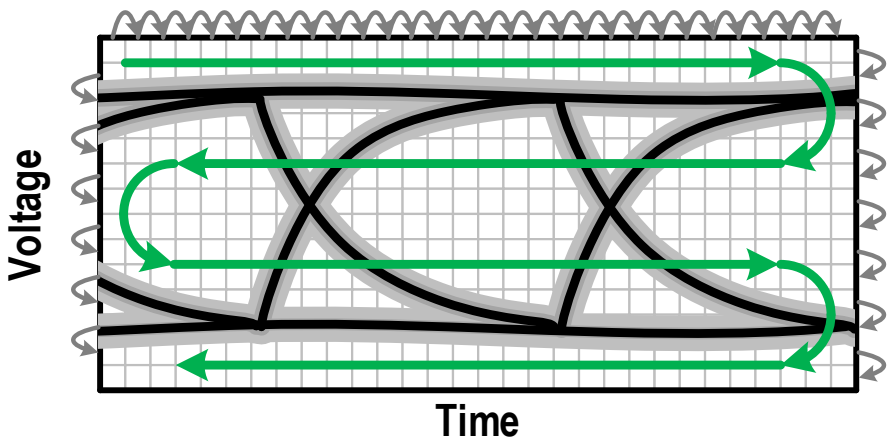


Figure 3.1.1 Two-dimensional eye detection

3.2 1X2Y3X EYE DETECTION

Instead of the two-dimensional eye center detection algorithm, the 1x2y3x eye center detection algorithm reduce the time of training. Figure 3.2.1 shows the 1x2y3x iteration eye detection algorithm. First, time direction sweep '1x' is performed. As shown in Figure, first, eye detection sweeps the x-axis direction of the sampling timing to fine the x-axis eye monitoring while the y-axis of the reference voltage were fixed. Second, the eye detection sweeps the y-axis direction of the reference voltage to find the y-axis eye monitoring while the x-axis of the timing were fixed at center point value of the first sweep called '1x'. Third, the y-axis of the reference voltage is fixed at the center point of the second sweep called '2y', and the x-axis of the timing is swept for eye detection again. The center point of the eye is fixed at center point of the second and third sweep. Thanks to the LPDDR4 specification, the value of the reference voltage of 1x sweep can be effectively started by with half value of the VOH, VDDQ/5 or VDDQ/6 [3.1.9][3.1.10]. In addition, the algorithm of saving two points, which is start and end point of eye opening, and averaging the sum of these two points is simpler than the algorithm of saving all point values of two-dimension and finding optimal eye center. The number of test points required for 1x2y3x eye detection is as follows:

$$N_{TR} = 2N_{time} + N_{volt}. \tag{3.2.1}$$

For example, if the timing step of the x-axis is 256 points, and reference voltage of the y-axis is 72 points, only 584 registers are required. And, the algorithm of finding the point of the biggest x and y margin in the two-dimensional table is needed. In the 1x2y3x

algorithm, on the other hand, only 4 registers are required to save the start and end point of the x- and y-axis, and LTFSM can find the eye center by simply averaging each.

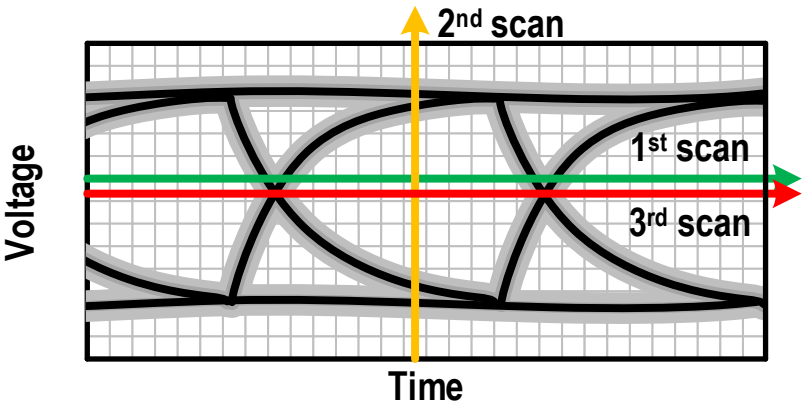


Figure 3.2.1 1x2y3x iteration eye detection.

3.3 ADAPTIVE GAIN CONTROL

The 1x2y3x eye detection method described above has succeeded in reducing the number of test points as compared with the two-dimensional eye detection method. However, there is still a need for training test points proportional to the number of time-voltage steps. Therefore, there is still a need for improvements on ways to achieve faster training while maintaining accuracy in the current state. In this paper, we propose a method to reduce training time by applying adaptive gain control (AGC) scheme to eye detection algorithm. The figure is a concept block diagram of adaptive gain control.

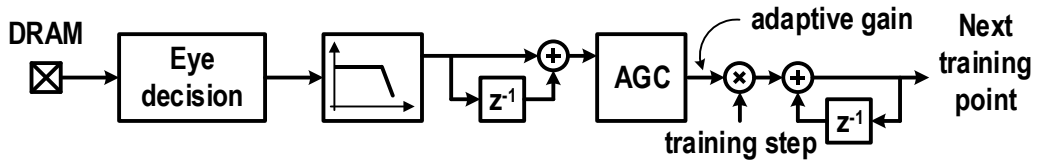


Figure 3.3.1 Simplified block diagram of adaptive gain control scheme

First, determine whether the detected training point is pass or fail. In case of continuous pass or fail, AGC block increases the gain of training step to reduce number of test points. When AGC improves gain, the resolution of training is lower in inverse proportion to gain. The reason for the high resolution is to find the eye's exact shape. And we can infer the shape of the eye by looking for the exact boundary of the eye. If we can improve adaptive gain and still have high resolution at the eye's edge, we can reduce training time without loss of accuracy. Therefore, in the proposed AGC, we reinitialize the gain by sequentially searching the training points between pass and fail in binary search

algorithm form, assuming that it is the boundary of the eye at the time of successive pass or fail.

Figure 3.3.2 shows an example of adaptive gain control scheme. The number of test points required for adaptive gain control scheme is as follows:

$$N_{TR} = N/K + K + 3\alpha \tag{3.3.1}$$

where N is number of training step, K is maximum adaptive gain and α is the number of consecutive pass or fail cycles required for the adaptive gain to increase to its maximum value. Assuming that N is larger enough to ignore K or 3α, the N_{TR} is reduced to the original test points N divided by the maximum adaptive gain value K.

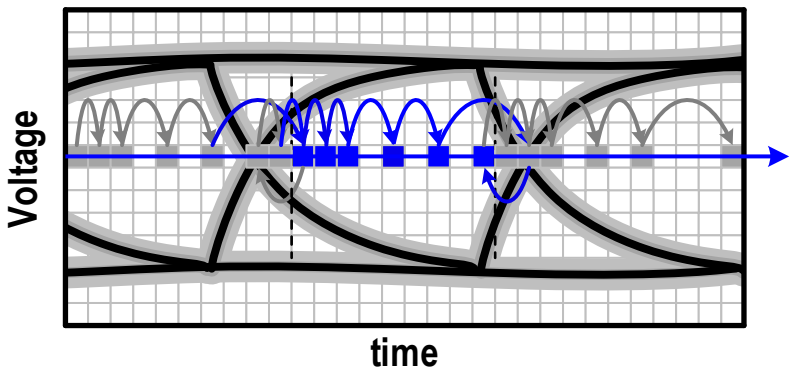


Figure 3.3.2 Eye detection with adaptive gain control scheme

3.4 ADAPTIVE 1X2Y3X EYE DETECTION

In this thesis, we implement the adaptive gain control scheme to the 1x2y3x eye detection algorithm and propose a faster and more accurate training method. Figure 3.4.1 shows an adaptive 1x2y3x eye detection algorithm and an example. Basically, eye detection is performed in the same manner as 1x2y3x.

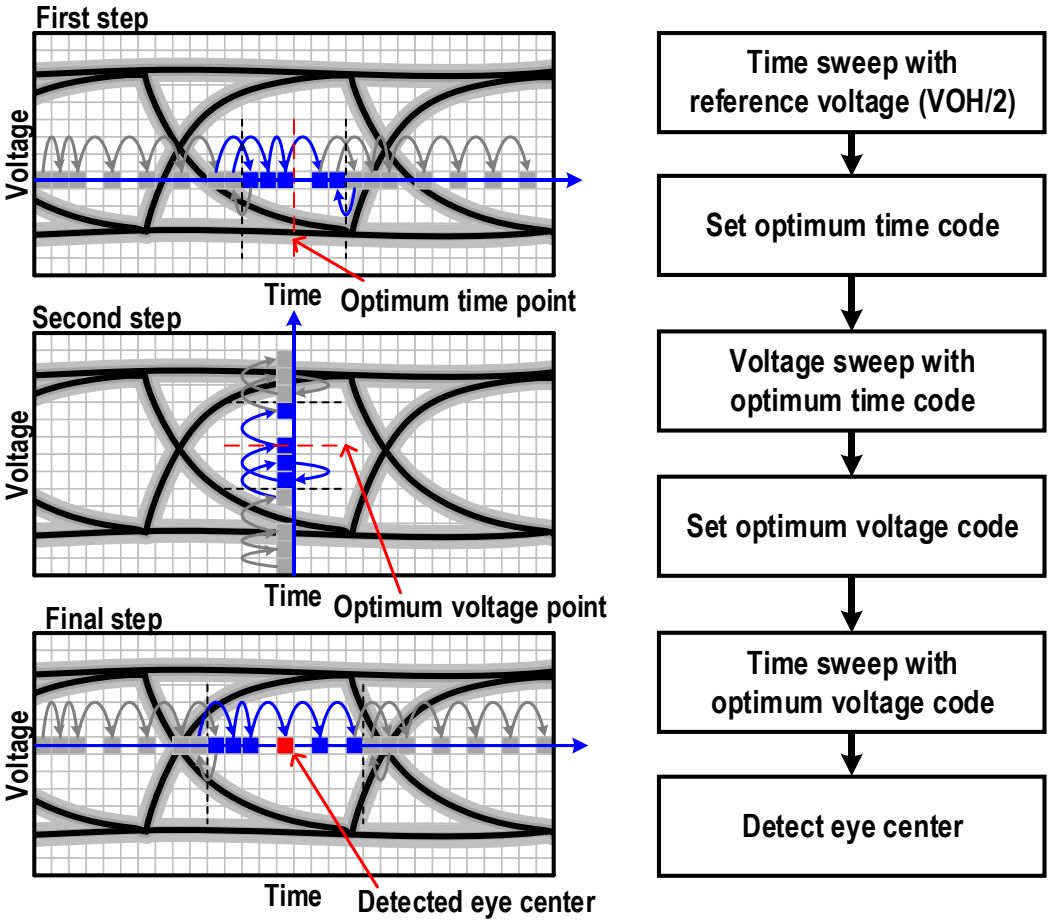


Figure 3.4.1 Adaptive 1x2y3x eye detection

Due to the adaptive gain control scheme, all three sweeps detect eye while moving test points at a wider interval in a continuous pass or fail. By using a binary search at the boundary of the eye to eliminate the loss of accuracy, the training results are the same when the adaptive gain control is enabled and disabled. Figure 3.4.2 shows the comparison between the previous eye detection algorithm and the proposed adaptive 1x2y3x eye detection algorithm. Comparing the actual design parameters, the number of time steps is 256, and the number of voltage steps is 72. The maximum adaptive gain value is designed to be 2, and the number of consecutive detection result values to reach the maximum adaptive gain is 1. Therefore, the number of test points for full scanning method, 1x2y3x iteration, and proposed adaptive 1x2y3x iteration are 18432, 584, and 307, respectively. The proposed adaptive 1x2y3x eye detection algorithm is 60 times faster than the full scanning method and detects the eye 1.9 times faster than the 1x2y3x eye detection algorithm. This indicate that the proposed adaptive eye detection algorithm can save much power than the conventional method as well as the training time.

	Full scanning	1x2y3x iteration	
# of time step	256		
# of voltage step	72		
Accuracy	High	Medium	
Adaptive gain control	No	No	Yes
Number of test points	18432 (256×72)	584 (256+72+256)	307 (133+41+133)
Normalized test points	60	1.9	1

Figure 3.4.2 Comparison of eye detection algorithms

CHAPTER 4

LPDDR4 MEMORY CONTROLLER

4.1 DESIGN PROCEDURE

Figure 4.1.1 shows the design procedure of memory controller design. In order to design a memory controller, it is necessary to understand the specification of the target DRAM. There are various types of DRAM, and each DRAM must have a precise understanding of the specification because there are special functions that only the generation has. In this thesis, we propose the design of a memory controller suitable for LPDDR4 which is the hottest one of mobile DRAM. In order to design a system-on chip, it is effective to design it in a top-down manner. Therefore, once you understand the specification [1.1.5], we need to define the MCU architecture and core blocks. In this process, the functions required for communicating with the memory in the MCU are summarized, and the performance of the sub-blocks and the corresponding blocks to be designed is defined, and the whole top architecture is defined as the sub-blocks.

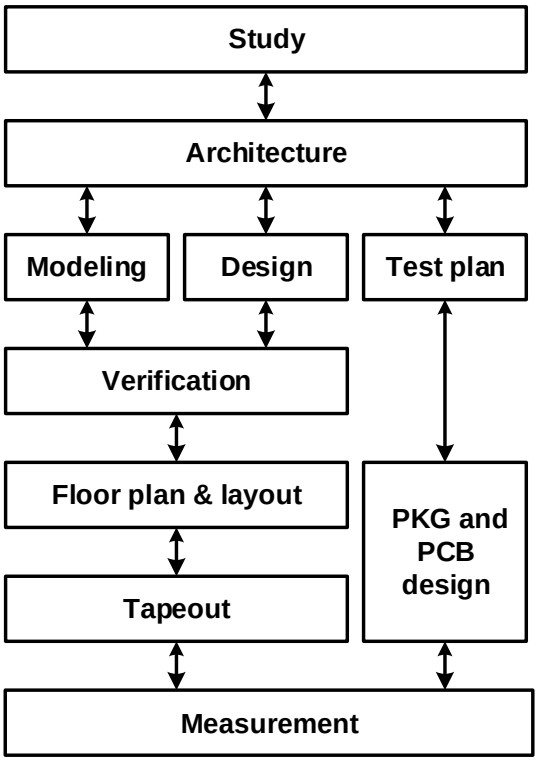


Figure 4.1.1 Design procedure of memory controller

Our proposed MCU is designed to perform point-to-point (P2P) communication with single channel of LPDDR4. Based on this condition, the architecture of MCU is summarized as follows. First, our MCU should have the same I/O with the single channel of LPDDR4. The single channel of LPDDR4 consist of an uni-directional differential CK pins, 7 unit-directional single ended CA pins, 4 bi-directional differential DQS pins, 16 bi-directional single ended DQ pins, 2 bi-directional single ended DMI pins and other control pins such as RESET, ODT_CA, and ZQ_CAL. Second, the clock speed of MCU is defined based on the operating speed of LPDDR4. The per pin speed range of the LPDDR4 is from 533Mbps to 4266Mbps with interval of 533Mbps. To meet aforementioned operation

frequency, MCU have to generate clock of 266MHz to 2133MHz with interval of 233MHz. Designing an oscillator that supports such a wide range of frequencies causes a trade-off in jitter performance or jitter performance. The frequency range was reduced from 1333MHz to 2133MHz and a 266 to 2133MHz clock was generated using a multi-modulus divider. Our MCU follows flow chart depicted in Figure to communicate with LPDDR4. The LPDDR4 memory starts at power on state, to operate normally, memory passed power on, reset, boot up state and training sequence. The training sequence consist CBT, WLVL, RDTR, and WRTR. After training, LPDDR4 goes activation state to prepare normal state. The margin tests are performed to evaluate the LPDDR4 operating performance. The *tDQSCK* (clock to DQS delay) and *tDQSDQ* (DQS to DQ delay) should be compensated by MCU to proper operation, and MCU also performs ZQ calibration, per pin de-skewing, read and write latency training, clock domain crossing, and eye center detection. To perform these functions, phase-locked loop (PLL), delay-locked loop (DLL), serializer/de-serializer (SER/DES), LVSTL driver, clock distribution circuit with skew minimization, and continuous-time linear equalizer (CTLE) are required. The proposed MCU supports power on, reset, idle, activating, bank active, read, write, command training, and MPC based training.

After architecture define, the design phase begins. Three must be done in parallel. First, it is necessary to verify the operation of the whole system through modeling. Until the detailed circuit design is completed, it is necessary to model the sub-blocks according to the defined architecture, and to check that there are no mistakes in the functional parts. The second is to design the actual circuit. We must define clearly the simulation corner

conditions that we need to verify before entering the circuit design. In the verification environment, the sub-blocks defined in the architecture configuration step are designed to meet the required performance requirements. When designing each sub-block, the design should be performed in consideration of not only the corresponding block but also other blocks to which the input output of the corresponding block is connected. Finally, the third is to establish a measurement plan. This is partly related to modeling, so we make a test plan about how to measure the designed MCU. To test proposed LPDDR4 memory controller with LPDDR4 memory, LPDDR4 is required. Generally LPDDR4 memory controller is stacked with LPDRR4 memory by package-on-package structure [4.1.1] [4.1.2]. However, it is not easy to make package-on-package structure for academic research. Thus, thin quad flat package is used for testing, and this package type is considered when layout and floor plan.

In the verification phase, both functional verification based on modeling and simulation verification based on schematic and layout are performed simultaneously. To verify the MCU, a corresponding DRAM model is required. In this thesis, the system can be configured and verified using the verilog model of LPDDR4 and channel model. The flow plan should preferably be started before layout. The layout size and pin position are defined through the flow plan and then the layout is performed. This prevents the layout size from becoming too large and optimizes the signal transmission path.

4.2 ARCHITECTURE

Figure 4.2.1 shows the architecture of proposed LPDDR4 memory controller. The input and output of MCU consist of an uni-directional differential CK, 7 unit-directional single ended CA, 4 bi-directional differential DQS, 16 bi-directional single ended DQ, 2 bi-directional single ended DMI and other control signals such as RESET, ODT_CA, and ZQ_CAL.

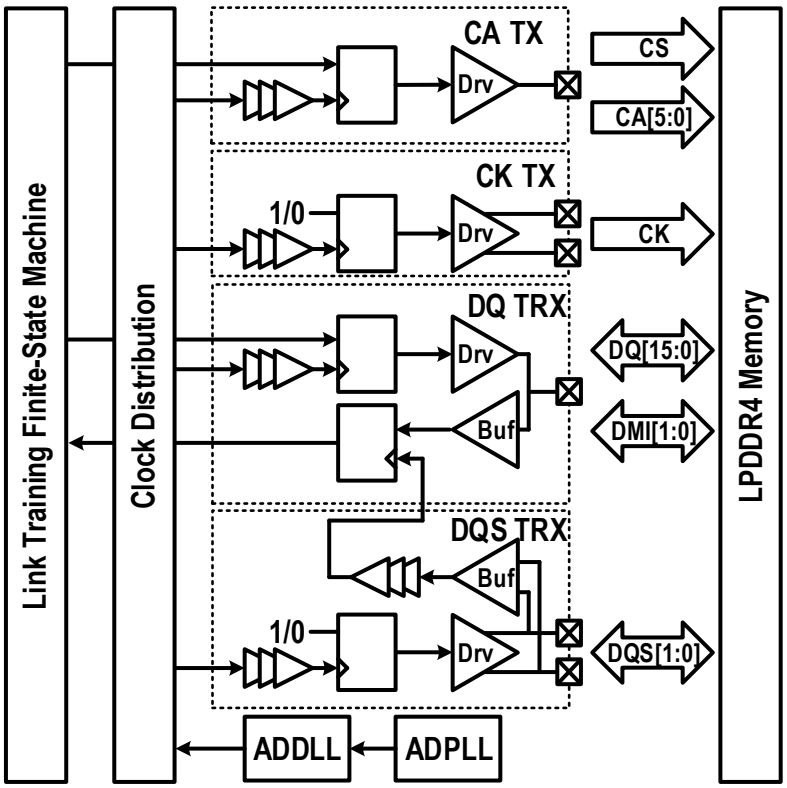


Figure 4.2.1 Architecture of proposed LPDDR4 memory controller

The MCU consists of all-digital phase-locked loop (ADPLL), all-digital delay-locked loop (ADDLL), clock distribution circuit, link training finite-state machine (LTFSM), 7 transmitter for CK and CA path, 18 transceiver for DQ and DMI, 2 transceiver for DQS. ADPLL generates the global clock PHY_CLK used in MCU and SYS_CLK used in LTFSM. PHY_CLK has an operating frequency of 266 to 2133 MHz, and SYS_CLK has an 8 divided frequency of PHY_CLK. The ADDLL consists of a global DLL and a local DLL. The global DLL helps fast locking the local DLL, the local DLL generates the multi-phase clock, and PI adjusts the clock phase to 1/64 UI steps. The LTFSM operates as SYS_CLK and generates DQ, CA signals and MCU control signals. The transmitter and the receiver are configured to satisfy the high bandwidth by supporting the training operation.

The most important part of the memory controller's architectural configuration is to maintain the source synchronous clocking scheme while keeping it in line with the overall system's training behavior. Training support should be basic and satisfy SSCS for high performance. The method incorporates a transmit clock delay line and integrating receiver yielding an increased tolerance to high frequency transmit source jitter [4.2.1]. There are two groups that must satisfy the SSCS in the memory controller. One is CA and CK, and the other is DQ and DQS. If the SSCS can be satisfied between CK and DQS, better performance can be obtained, but this can be neglected because it is difficult to achieve because the signal path inside the DRAM is quite different. In order to satisfy the SSCS, it is necessary to share a clock for generating each group signal. For example, when the CA and CK are generated with one clock, the generated CK must be designed to sample the

CA so that the SSCS can be satisfied. In order to adjust the signal timing while satisfying the SSCS, it is necessary to control the timing of the clock of the last serializer.

4.2.1 TRANSMITTER

Figure 4.2.2 shows the block diagram of transmitter. The transmitter consists of PI, digitally controlled delay line (DCDL), 16:1 SER, pre-driver (pre-DRV), and LVSTL. Transmitter should be designed with WLVL and WRTR in mind. More specifically, write DQS path should take WLVL into account, and write DQ path should be designed considering WRTR.

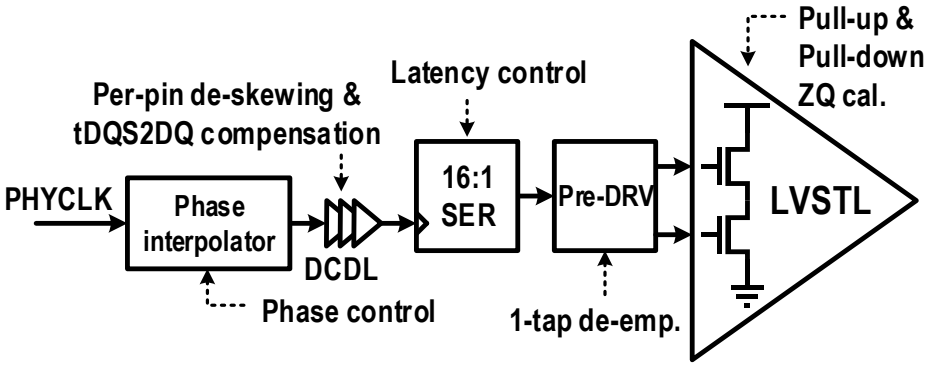


Figure 4.2.2 Block diagram of transmitter

WLVL is a training that helps the DQS-CK clock domain crossing issue of LPDDR4 by adjusting the timing of DQS generated by MC. The PI adjust the phase of DQS and CK, timing should be adjusted to 1tCK range. WRTR is the training to find the optimum sampling point between DQ and DQS. As introduced in Chapter 2, LPDDR4 uses a source synchronous unmatched scheme, so there is a skew from 200ps to 800ps called tDQS2DQ between DQ and DQS. Additionally, pin-to-pin variations between DQ and DQS must be

considered to achieve high bandwidth. Therefore, DCDL with range of 200~1000ps is placed after PI to compensate both *tDQS2DQ* and per-pin skew. To compensate *tDQS2DQ* in source synchronous unmatched scheme, PHYCLK should be delayed by over than 3 UI. For this reason, clock phase mismatch occurs between LTFSM and transmitter. In order to ensure stable data sampling considering the data transmission between LTFSM and transmitter, the SER is configured as 16:1 SER. In order to achieve high bandwidth, pre-driver supports 1-tap de-emphasis and LVSTL can have accurate pull-up and pull-down drive strength through ZQ calibration.

One thing to note about transmitter design is that all designs must be made in a limited area. This is somewhat difficult to observe at the academic level, but it must be kept in mind when creating multichannel transceiver circuits. In the chip, the number of pins is limited, and the order of the pads such as the power signal ground signal is limited for uniform power supply. In addition, the spacing between the pads and the pads is determined by the process, so it must be taken into account that both the transmitter and the receiver must be included within a limited height. The block that occupies the largest area in the transmitter is the driver. Therefore, the guidelines of the driver design should be designed first, and the blocks of the remaining transmitter should be designed to be optimized for the driver size.

Simultaneous switching output noise (SSO) in single-ended signaling is one of the major performance limiters as data-rate scales higher [4.2.2]. Research shows that improved package and motherboards can reduce the impact of SSO and improve system performance. However, academic-level research has limitations in applying these areas, so

you should implement as much of the SSO-insensitive design on-chip as possible. From an empirical point of view, the easiest and most accurate way to reduce SSO noise is to separate the output driver and the internal circuitry from the main driver of SSO noise. Unfortunately, the chip manufactured in this paper has a disadvantage that it is not susceptible to SSO noise because it does not separate the power supply. These parts should be improved in the future.

4.2.2 RECEIVER

Receiver should be designed with RDTR in mind. Figure 4.2.3 shows the block diagram of receiver. The receiver consists of VREF generator, CTLE, PI, DCDL, 1:4 DES, and 4:16 DES.

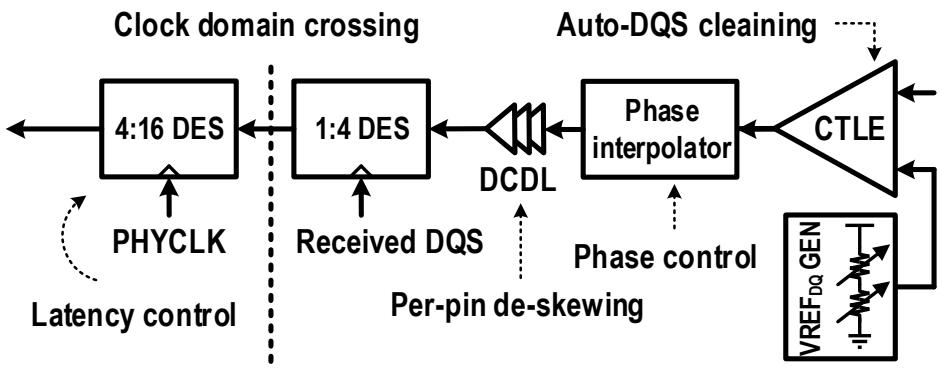


Figure 4.2.3 Block diagram of receiver

The MCU cannot use the source synchronous unmatched scheme because LPDDR4 simultaneously transmits DQ and DQS. As a result, MCU should compensate for the clock

distribution time of the DQS and therefore it is unable to use DFE which is a powerful equalizing technique. Therefore, CTLE is used to compensate the channel loss. The reference voltage of VREF generator is shared among all DQs. Since the DQS is irregularly received in the read path, the ADDLL in receiver operates in a way that locks by PHYCLK when it is not a READ operation and freezes the control code when READ operation. DCDL with range of 200ps is placed for per-pin de-skewing in each of DQ and DQS in front of DES. The clock domain crossing from DQS to PHYCLK, read latency training and byte-aligning are performed in the 1:16 DES.

As with the transmitter, the area of the receiver design must be considered first. This is because the area including both transceivers is limited by the process and pins. However, since the area of the receiver is generally smaller than that of the transmitter, there is not much difficulty in designing such a portion. Due to the nature of single-ended signaling, the reference voltage used for multiple DQs must be shared across the entire chip. This part is guarded without a special repeater, so that the top metal is laid out so that it is not affected by the noise as much as possible.

The most characteristic part of the memory controller receiver is the data strobe. Unlike a typical i/o interface, the memory controller's receiver receives data and data strokes on an irregular basis, as well as a limited number of data strobe edges to sample data. Care must be taken to ensure that all data strobe sent by the memory is taken care of and not to be missed. In addition, since the delay mismatch between the data experienced in the memory and the data strobe occurs in the memory controller, the data must be delayed in accordance with the distribution delay of the data strobe.

The memory controller is bi-directional, so the receiver must not be on at all times. If the receiver is always on, the data sent from the memory controller to the memory may loop back itself, which may cause malfunctions. Therefore, in the READ operation, the receiver is turned on. When it is not READ operation, the receiver is designed to turn off or the DQ path should be turned off. This should take into account both the timing of sending the READ command and the timing of sending the DQ in response to the READ command in memory. The activation timing of the receiver can be determined by measuring the time it takes for the memory itself to transmit the READ command and returning the DQ to the DRAM. Therefore, the memory controller is designed to use the decoder to determine the type of command currently being transmitted and, in the case of the READ command, to activate the receiver after a specific cycle, where the specific cycle is designed to be trained.

Clock domain crossing is also a difficult issue. Since the memory controller and memory are exposed to different PVT variations, timing variations occur between PHY_CLK and received DQS. To prevent this, it is necessary to generate a PHY CLK that tracks the phase of DQS, or to maximize the timing margin between the DQS and the PHY CLK. It is necessary. However, since the method of generating the PHY CLK that tracks the phase of the DQS is somewhat disadvantageous in terms of additional circuitry and complexity, we have chosen to maximize the timing margin between the DQS and the PHY CLK. To do this, the first thing to do is to deserialize the DQ as much as possible using the DQS. We deserialize DQ to DQS 1: 4 using the pre-amble edge and post-amble edge of DQS. There are three types of timing for deserializing DQ using fixed PHY CLK. Among

them, the training is configured so that the PHY CLK can sample the DQ at the timing with the widest timing margin.

4.2.3 CLOCKING ARCHITECTURE

According to [4.2.3], the state ratios of mobile DRAMs are 12%, 28%, 18%, and 42% for burst write, burst read, random write, and random read, respectively. When comparing only the burst state and the random state, it can be seen that the random state is more than burst state with ratio of 40:60. This means that the waiting time of mobile DRAM is relatively high, so that design of clocking architecture of the MCU should consider about power consumption during standby state.

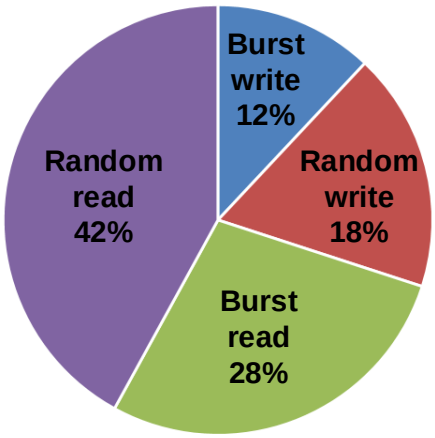


Figure 4.2.4 Mobile DRAM state ratio

As mentioned above, in order to minimize the power consumption in the idle state, we proposed a clocking structure with gated clocks. Figure 4.2.5 shows the clocking architecture of CK and CA. CK and 7 CAs share a local ADDLL and have PIs to phase

adjust independently of each other. The local ADDLL should not be shut down to track PVT variation even in idle state. Therefore, a clocking architecture is constructed by gating clocks between PI and clock tree. The control signal of the gated clock can be implemented by sending the command at the time of sending the command in LTFSM.

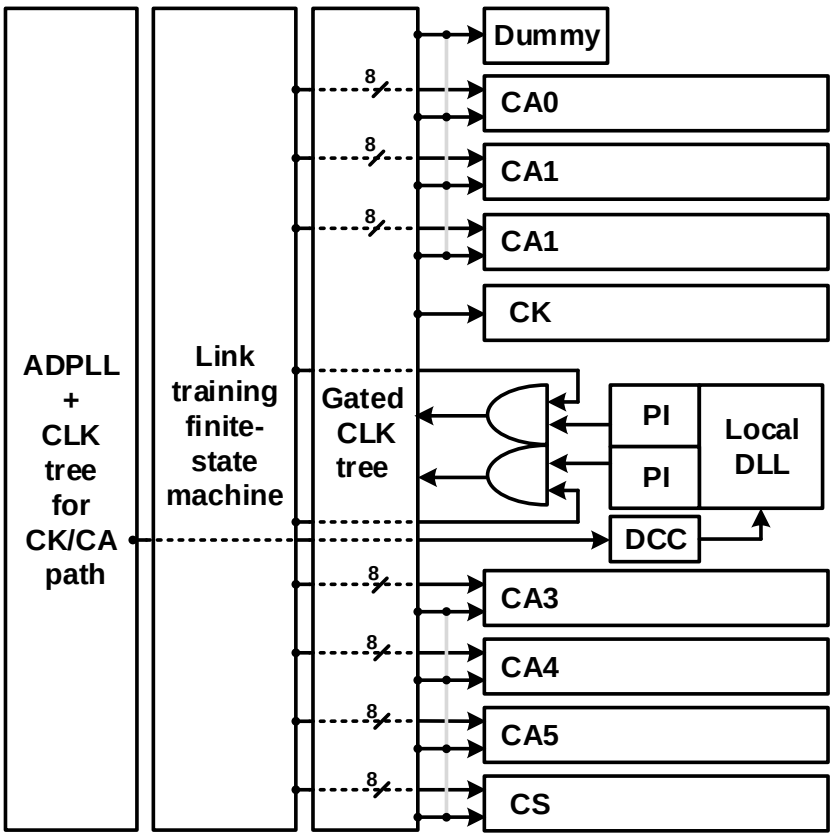


Figure 4.2.5 Clocking architecture of CK and CA

In clocking architecture, we have tried to minimize duty cycle and pin-to-pin skew to achieve high bandwidth as well as low power consumption. At the end of the gated clock tree, the metal-wire shorts the input clock to each CA to reduce the pin-to-pin skew. In the

floor plan stage, local ADDLL and PI were placed between CK CA channels and dummy load was placed to reduce load mismatch. Figure 4.2.6 shows the clocking architecture of TX of DQ, DMI and DQS. DQS, DMI and 8 DQs share a local ADDLL and have PIs to phase adjust independently of each other.

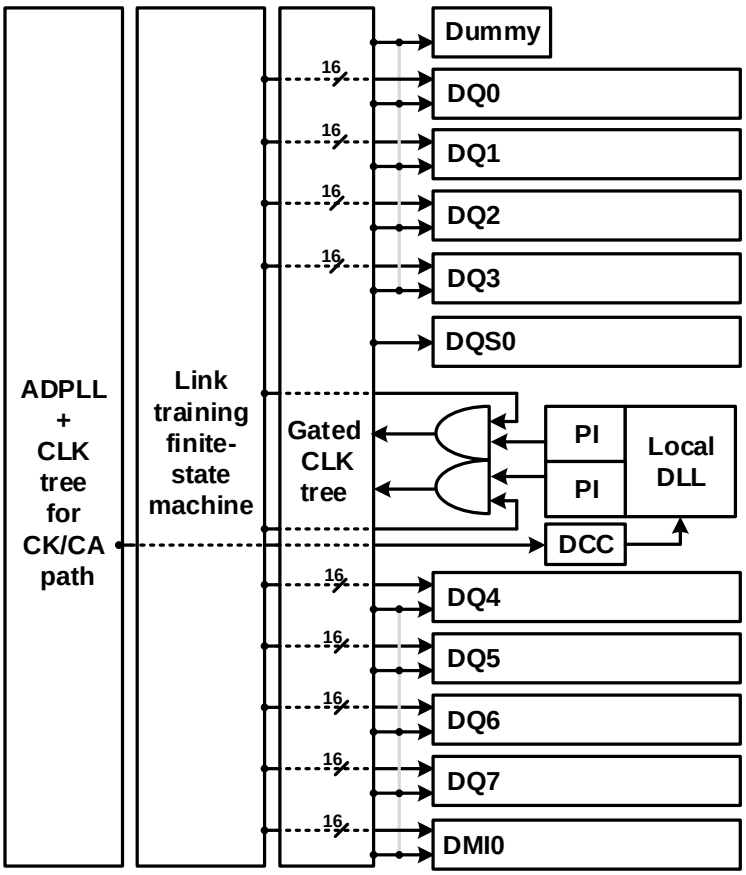


Figure 4.2.6 Clocking architecture of DQ/DQS TX

Figure 4.2.7 shows the clocking architecture of RX of DQ, DMI and DQS. Since received DQS does not transition when there is no READ operation, the same effect can be obtained without clock gating. Instead of clock gating, we turn-off the receiver circuits to

reduce static power consumption at non READ operation.

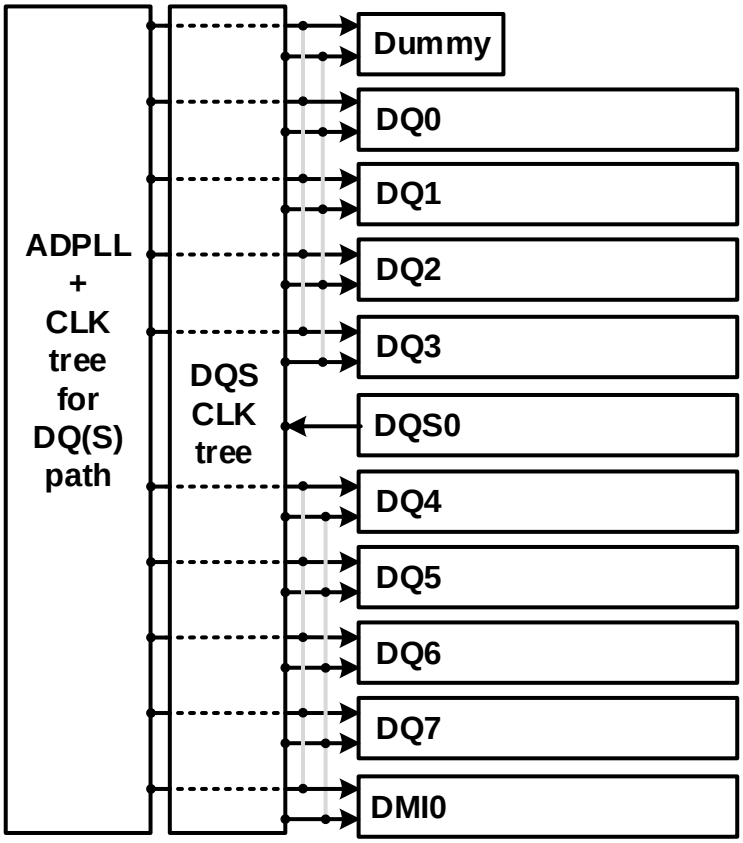


Figure 4.2.7 Clocking architecture of DQ/DQS RX

Figure 4.2.8 shows the block diagram of transceiver at IDLE state. Through the clocking architecture described above, most of the circuits are not consuming power during random operation, which improves the power efficiency of the MCU. This design is very important because it helps to determine which part meets the design expectations at the stage of verifying the system design by dividing the power consumed by each block. The

power consumption of the IDLE state in the actually designed chip is measured higher than the simulation value, but the power variation in the WRITE state and the READ state is similar to the actual design value. This allows us to determine which block was causing the design mistake. Another thing to keep in mind is to avoid SSCS breaking or glitches in the process of gating the clock. For this reason, the clock gating block is designed to have no glitch by re-sampling the control signal with the falling edge of the input signal.

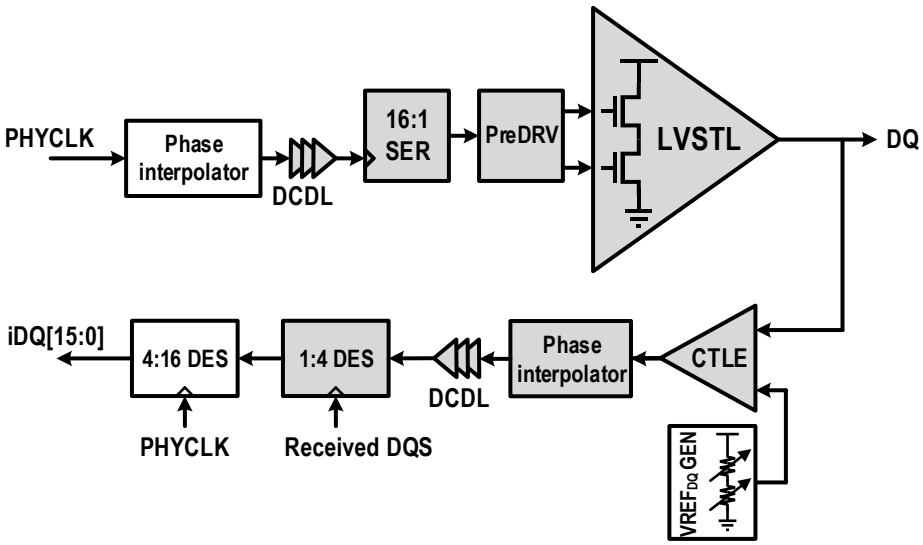


Figure 4.2.8 Block diagram of transceiver at IDLE state

4.3 CIRCUIT IMPLEMENTATION

Section 4.3 describes sub-block design of ADPLL with multi-modulus divider, ADDLL with triangular-modulated PI, CTLE with auto-DQS cleaning, 1:16 DES with clock domain crossing, LVSTL driver with ZQ calibration block, DCDL with coarse-fine delay unit and LTFSM.

4.3.1 ADPLL WITH MULTI-MODULUS DIVIDER

Figure 4.3.1 shows the block diagram of the ADPLL. The ADPLL consists of the phase-frequency detectable time-to-digital converter (PFDTDC), digital loop filter (DLF), 1st order delta-sigma modulator (DSM), digitally controlled oscillator (DCO) and 2-stage multi-modulus divider (MMDIV).

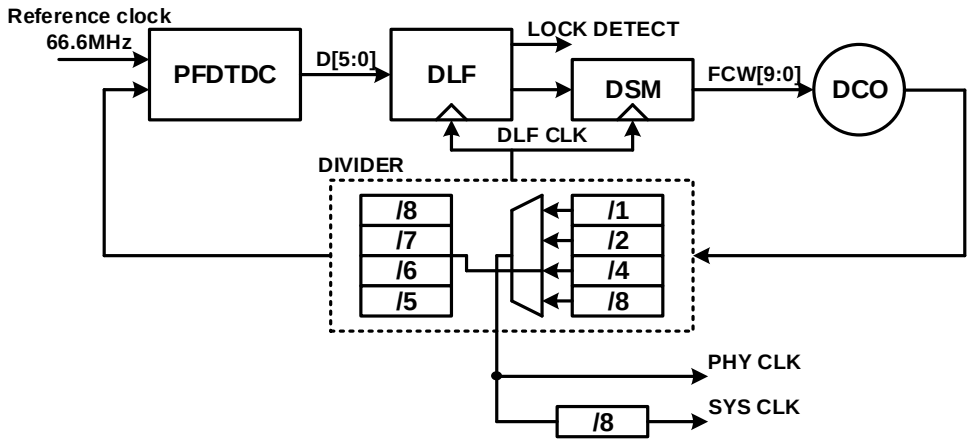


Figure 4.3.1 Block diagram of the ADPLL

A PFDTDC combines vernier TDC and the PFD to generate both high resolution and wide range TDC information with up and down information. It has dynamic range of 300ps with the resolution 10ps. The DLF provides control code of the DCO by proportional and integral path. The DSM reduces the in-band noise by shaping quantization noise and increases the effective resolution of the DCO. The 2-stage MMDIV divides the output clock of the DCO and generates both system clock of the MCU and the feedback clock of the PFDTDC. The frequency of the reference clock is 66.6MHz. The operation frequency of the DCO is from 1333MHz to 2133MHz. As shown in Figure, to provide the system clock of the LPDDR4 MCU from 266MHz to 2133MHz, the system clock is generated by selecting 1, 2, 4, or 8 dividing factors of 1st stage of MMDIV, and the feedback clock is generated by dividing factor 4 of 1st stage and dividing factors 5, 6, 7 or 8 of 2nd stage of MMDIV.

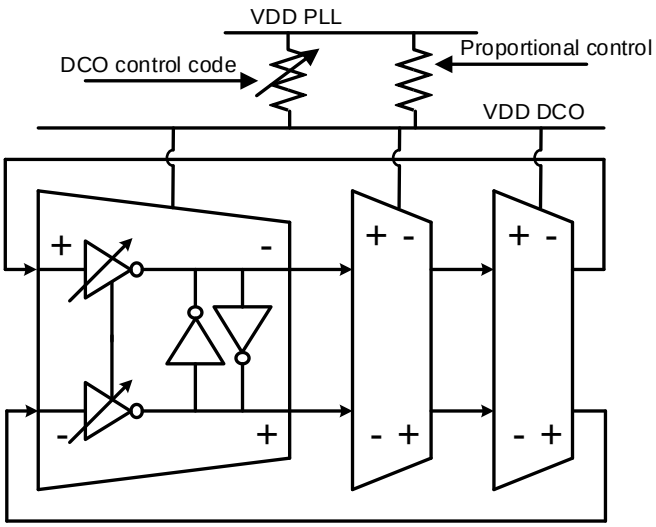


Figure 4.3.2 Block diagram of digitally controlled oscillator

The DCO adopts digitally controlled resistor (DCR) based ring oscillator [4.3.1]-[4.3.3]. The effective supply voltage of the DCO is changed to control output frequency. The jitter characteristics of the PLL is in inverse proportion to the K_{DCO}. In order to support all the frequencies from 1333MHz to 2133MHz, the DCO must be designed to have a large K_{DCO}. To achieve both wide tuning range and low K_{DCO} simultaneously, the 2-bit mode selection is implemented by adjusting the strength of the inverters of each stage of the DCO.

4.3.2 ADDLL WITH TRIANGULAR-MODULATED PI

Figure 4.3.3 shows the block diagram of the global ADDLL and the local ADDLL. The global DLL adopts an open-loop coarse TDC (Time-to-Digital Converter) architecture, and an asynchronous one-time lock completes 180° phase shift lock in several cycles. When the global DLL is locked, the lock code is sent to the local DLL and all circuits in the global DLL are powered down.

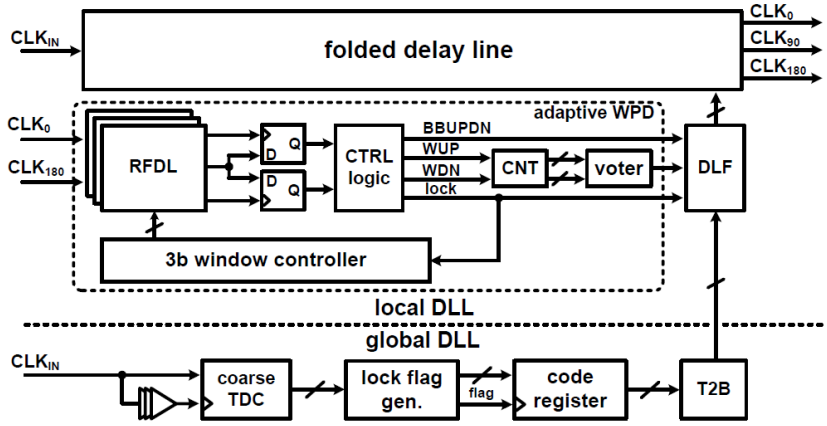


Figure 4.3.3 Block diagram of ADDLL

The local DLL immediately becomes coarse locked by delay code from global DLL. The local DLL then continually tracks the input phase to compensate for PVT variation as well as the mismatch between global and local DLLs. The local DLL generates multiphase CLK signals (CLK0, CLK90, and CLK180) and sends them to the PI. The folded delay line (DL) has coarse DL and fine DL. Conventional DL consisting of coarse DL and fine DL can cause glitch and signal distortion at the moment the coarse delay code is changed, since all of the fine delay code are changed at the same time. This increases jitter, reduces linearity and timing margins, and causes the system to malfunction. Figure 4.3.4 shows the operation of the folding DL of the invention with two delay codes with the same delay time. If the local DLL is initially locked near the change point of the CDC in the collapsed DL, digital loop filter (DLF) changes the delay code to another code away from the CDC change point, suppressing glitches and signal distortion.

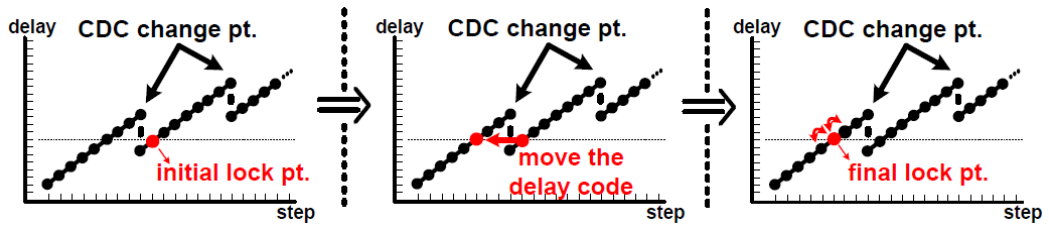


Figure 4.3.4 Operation of the folded delay line

Figure 4.3.5 shows the block diagram of triangular-modulated PI. To improve the linearity of the PI over a wide frequency range, a high slew rate between adjacent interpolating CLK phases is required but is difficult to achieve with conventional PIs. This difficulty is overcome by introducing a triangular wave generator (TWG) into the PI. TWG

use a programmable current source and capacitor array to convert the square wave of the CLK signal to a high slew rate triangle wave over a wide frequency range. Therefore, phase interpolation can be smoothly performed even if only three CLK signals are transmitted at intervals of 90 degrees from the local DLL, and complexity, power consumption, and area can be reduced.

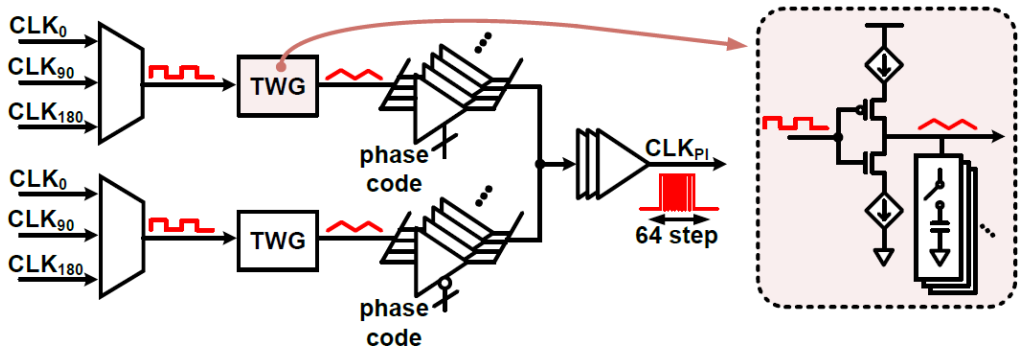


Figure 4.3.5 Block diagram of triangular-modulated PI

4.3.3 CTLE WITH AUTO-DQS CLEANING

As mentioned earlier, CTLE is basically an equalization technique that can widen the receiver's margin due to the nature of the memory interface that can't use a DFE. In this thesis, we applied asynchronous feedback to allow CTLE to remove DQS glitch apart from the equalization function.

Figure 4.3.6 shows the timing diagram of DQS glitch issue and gate training. At both the start and end times of a read operation, both DQSP and DQSN signals are grounded, as

VSSQ termination is used and the memory output driver is turned off. In the idle state, both the DQSP and DQSN signals have the same signal level, making the receiver vulnerable to noise, which can cause glitches in Y_DQS. One way to eliminate glitches is to generate gate pulses that are optimized to remove glitches through gate training in the MCU.

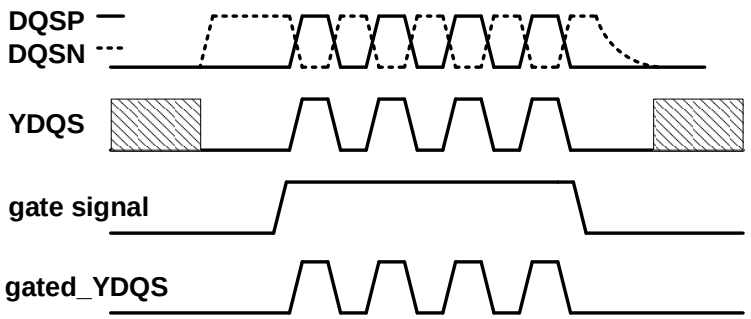


Figure 4.3.6 Timing diagram of DQS glitch issue and gate training

However, as we have already talked about, the training results are vulnerable to delay variations that change in real time. As shown in Figure 4.3.7, if the arrival times of the DQSP and DQSN signals change, the fixed gate pulse will lose the DQS. Therefore, the MCU has to do gate re-training, and during this time, the MCU will not be able to access DRAM, which will have a negative performance side effect.

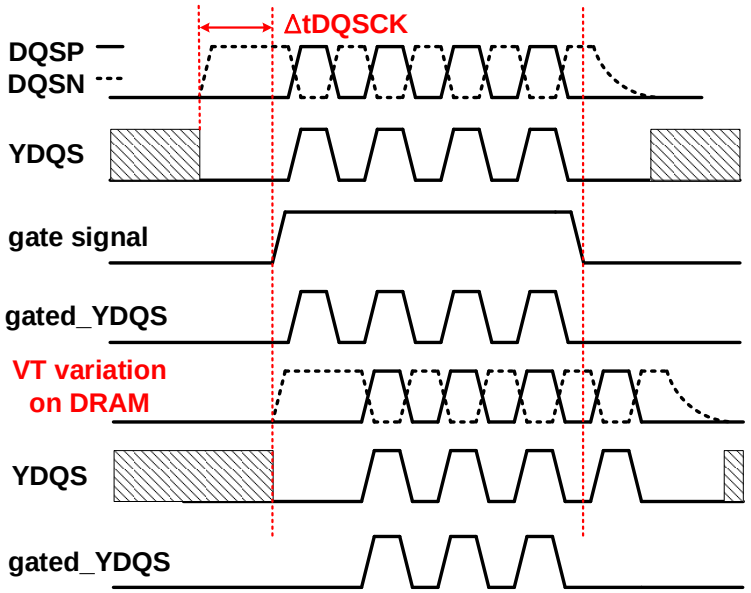


Figure 4.3.7 Gate training fail with time variance

To eliminate the DQS glitch issue without periodic retraining, as shown in Figure 4.3.8, a method has been proposed for generating a gate signal using an additional receiver that compares the DQSN to the reference voltage [4.3.5]. As shown in Figure 4.3.9, this method can produce a stable gate signal, but it has the disadvantage of using twice the power and area compared to the normal receiver because it requires using two receivers.

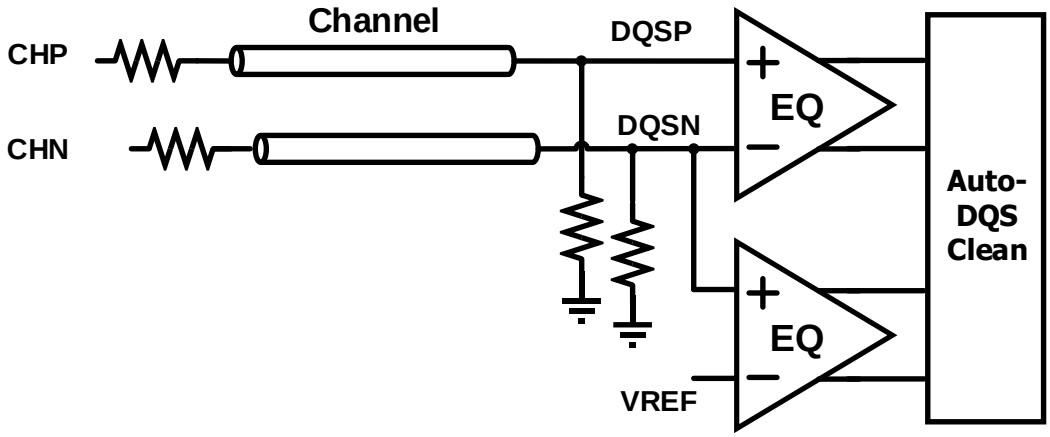


Figure 4.3.8 DQS clean with 2x receiver

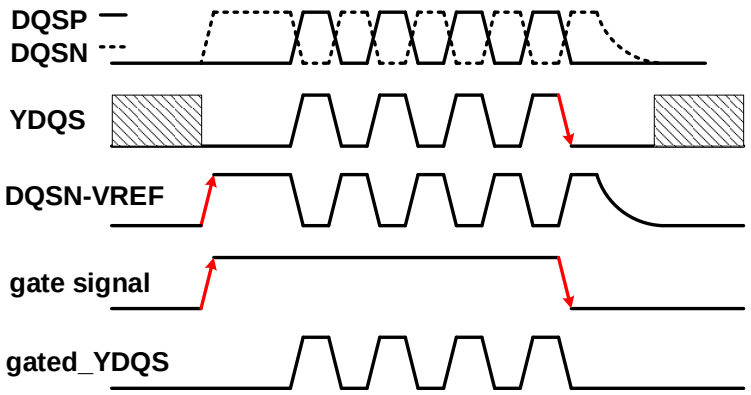


Figure 4.3.9 Timing diagram of automatic gate

In this thesis, we proposed a CTLE capable of removing DQS glitch without generating a gate signal using asynchronous feedback. Since the signal levels of DQSN and DQSP are the same in the idle state, the noise of DQSN and DQSP is easily amplified in CTLE and glitch occurs. If there is an offset between the differential inputs of CTLE, this noise is not amplified due to the difference in signal level. However, the offset in CTLE make the duty error of the output during READ operation instead of eliminating glitch

occurrence in idle state. If the sign of the offset changes according to the input in the READ operation, the CTLE can eliminate both the glitch occurrence in the IDLE state and the duty error of the output during the READ operation. We focused on the transition of the YDQS signal according to the input period when the READ operation was performed, and the YDQS signal was fed back so that the offset direction of the CTLE could be changed according to the input. Figure 4.3.10 shows the block diagram of CTLE with asynchronous feedback. In the IDLE state, the output of the SR latch is fixed to make offset of the CTLE in one direction, which interferes with the glitch occurrence. During READ operation, the CTLE offset is changed by the transition between YDQSP and YDQSN. Figure 4.3.11 shows the timing diagram of proposed receiver with and without asynchronous feedback.

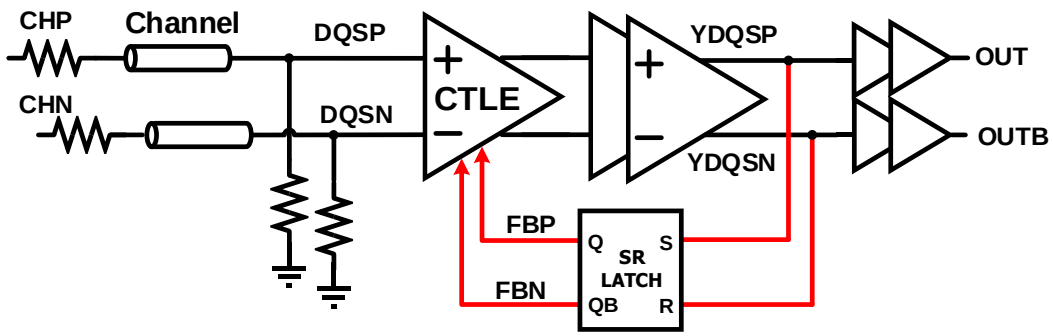


Figure 4.3.10 Block diagram of CTLE with asynchronous feedback

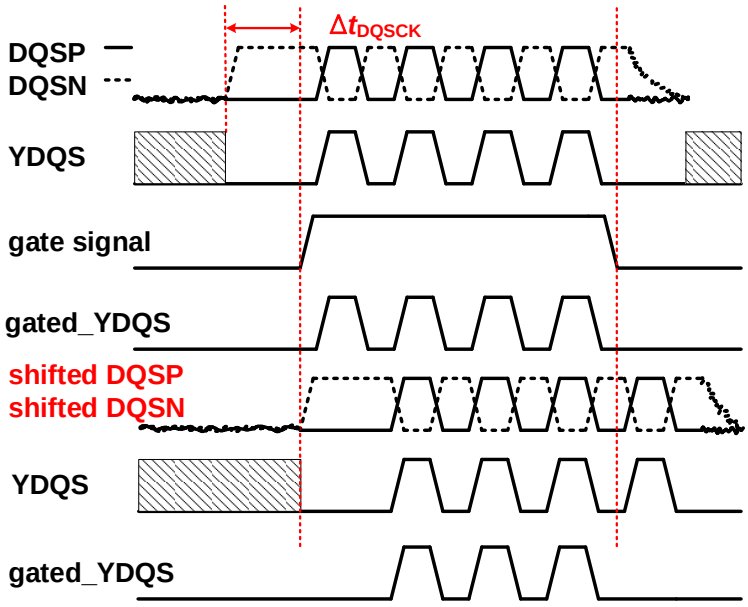


Figure 4.3.11 Timing diagram of CTLE with asynchronous feedback

4.3.4 DES WITH CLOCK DOMAIN CROSSING

Figure 4.3.12 shows the block diagram of 1:16 DES.

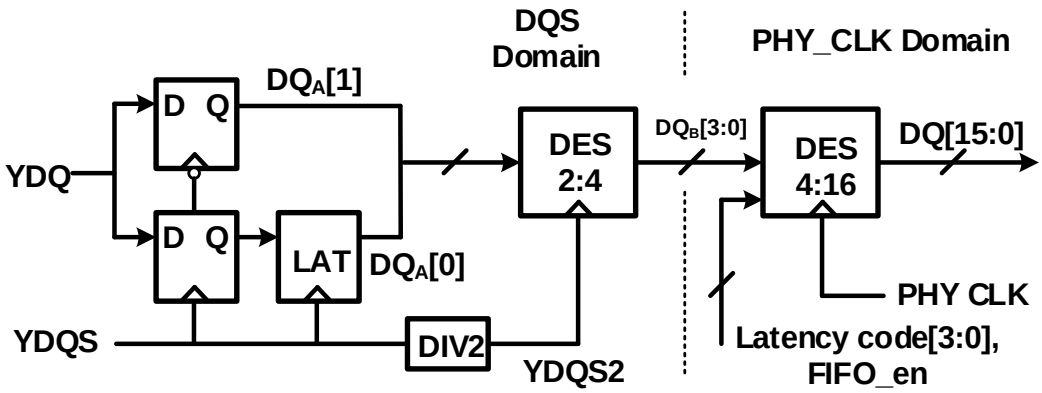


Figure 4.3.12 Block diagram of 1:16 DES

The jitter correlation is broken between the PHYCLK and the DQS as the DQS signal passes through the DRAM. In addition, the jitter of DQS becomes larger due to the noise from DRAM. As a result, the more parallel the incoming data of DQS domain DES, the more stable clock domain crossing can be achieved. However, since the number of rising edges of DQS is limited, there is a limitation on DES. We designed 1: 4 DES in DQS domain by using both pre-amble and post-amble edges. Figure 4.3.13 shows the timing diagram of 1:16 DES. In a single READ operation, YDQS has 10 rising edges, including the rising edge of pre-amble and post-amble. By using all of them, 1:4 DES can derive DQ_B by de-serializing YDQ 1: 4 with YDQS. Since DQ_B has a width of 4tCK, 1:16 DES have enough timing margin that can stably sample DQ_B with PHYCLK even if DQS has many jitter while passing through DRAM or shifted by PVT variation.

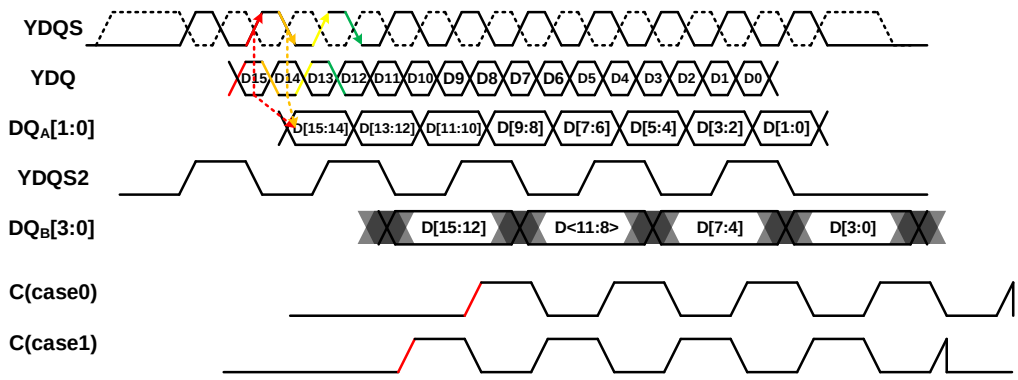


Figure 4.3.13 Timing diagram of 1:16 DES

4.3.5 LVSTL WITH ZQ CALIBRATION

To achieve high bandwidth, LVSTL must have the accurate driver strength. In addition, at READ operation, impedance matching is required to avoid reflection caused by impedance mismatching between the channel and the receiver. A pull-down of the driver acts as ground termination in the receiver mode. The values of the impedances are 240Ω, 120Ω, 80Ω, 60Ω, 48Ω, and 40Ω, which is integer divided values of 240Ω. In the proposed MCU, ZQ calibration is adopted to ensure accurate pull-down driver strength and pull-up driver strength. ZQ calibration first performs a pull-down NMOS calibration using an external resistor. Then ZQ calibration performs a pull-up NMOS calibration based on the calibrated pull-down transistor. Figure 4.3.14 shows the block diagram of pull-down ZQ calibration. Replica circuit of pull-down NMOS is connected to an external 240Ω resistor through the ZQ pad. The comparator compares the DC levels of the VOH and ZQ PAD. The counter adjusts the number of pulldown NMOS based on the output of the comparator to change the impedance so that the DC level of the ZQ pad is equal to VOH.

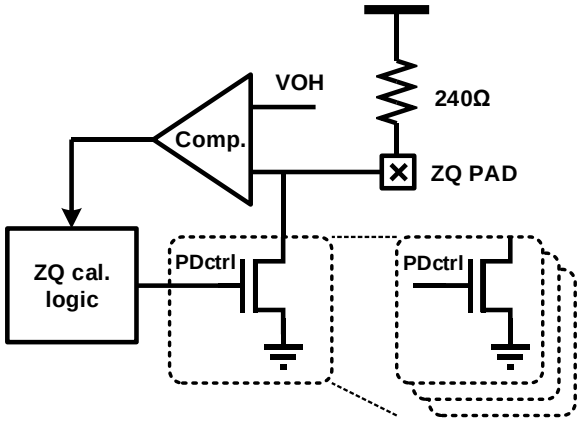


Figure 4.3.14 Block diagram of pull-down ZQ calibration

Figure 4.3.15 shows the block diagram of pull-up ZQ calibration. After the pull-down ZQ calibration is completed, the external resistor connected to the pull-down NMOS is switched to the replica circuit of the pull-up NMOS. Subsequently, the same method of pull-down ZQ calibration is repeated for the pull-up NMOS. After all ZQ calibrations are completed, the results are shared with the LVSTL in the transmitter.

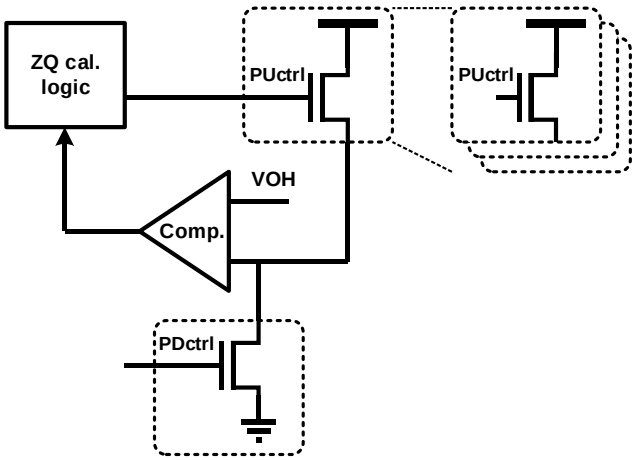


Figure 4.3.15 Block diagram of pull-up ZQ calibration

4.3.6 COARSE-FINE DCDL

When designing a DCDL, delay resolution and delay range should be defined. MCU need an asynchronous delay line to compensate tDQS2DQ and pin-to-pin skew in the transmitter. Therefore, the delay range was determined to be 1000ps by the maximum value of tDQS2DQ and pin-to-pin skew. The delay resolution should be about 4ps considering that the time margin is generally 1/64 UI and the maximum operating speed of the MCU is 4266Mbps. A delay unit with a propagation delay of 4ps requires as many as 250 delay units to achieve a delay range of 1000ps, which costs too much power and area. Coarse-fine architectures are a way to efficiently implement high resolution and wide delay lines. Figure 4.3.16 shows the block diagram of coarse-fine DCDL. We were able to achieve 4ps fine resolution and 800ps wide range simultaneously using 16 coarse-delay units and a 4-bit phase mixer type fine delay line.

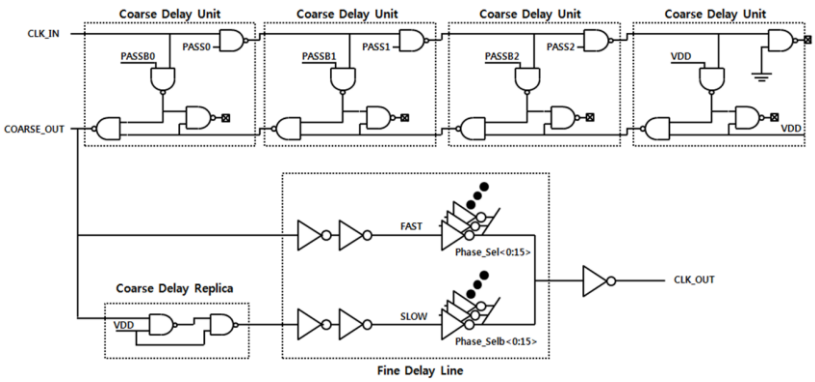


Figure 4.3.16 Block diagram of coarse-fine DCDL

4.4 LINK TRAINING

In order to design and test the MCU, the initialization and training of the DRAM must be performed by the MCU. The LTFSM generates data and commands to support memory initiation and training operations, performs calibration of the MCU, and generates control codes. As shown in Figure 4.4.1, the LPDDR4 memory can operates normally, after finish the initialization and training process from power ramp state at Ta to DQ training state at Tj.

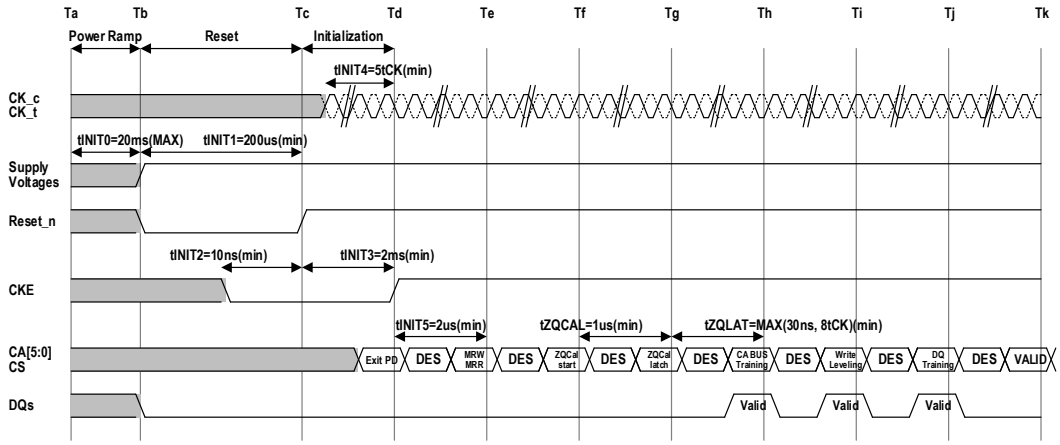


Figure 4.4.1 Initialization sequence of LPDDR4

The power ramp state means process for supplying power to the DRAM. In case of the MCU, MCU should be powered like the DRAM, besides all circuits in the MCU, such as the PLL and the DLL are prepared to train the DRAM. The meaning of prepared circuit is that entire circuits are clocked, saturated and reset. In other words, power ramp state of

the MCU means readiness for transmitting and receiving signals to and from the DRAM. Therefore, power ramp state of the MCU is the process of power supplying, reset state of the digital control circuits, and ready to operate state of the analog circuits.

The LTFSM consists of two modules. One module is a low-speed module that operates by a reference clock. The other module is a high-speed module operating by SYS_CLK. The states from time Ta to Tg in Figure 4.4.1 are low speed operation range. Prior to CBT, high-speed command transmission is not guaranteed, so the MCU must operate at low bandwidth. The low-speed module of LTFSM generates low-speed CAs before CBT and immediately after CBT. After CBT, the high-speed module replaces the low-speed module to generate CAs and DQs and control the MCU.

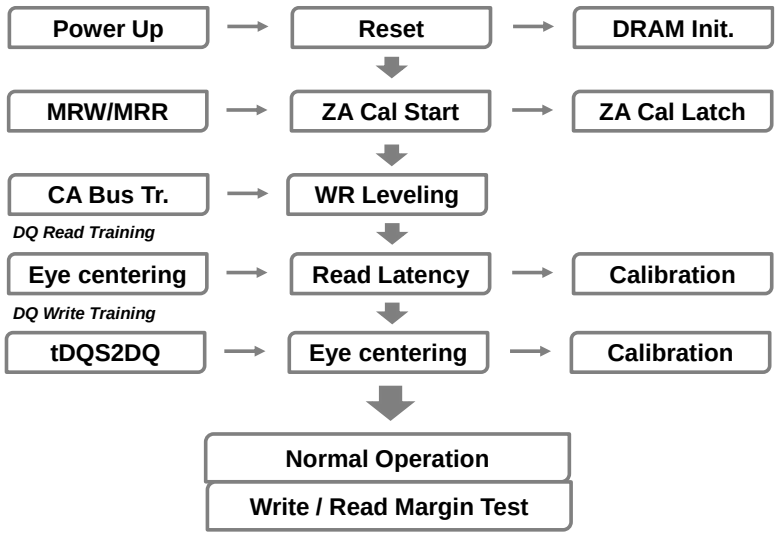


Figure 4.4.2 Operation sequence of LTFSM

4.4.1 SIMULATION RESULTS

Section 4.4.1 shows simulation results of the memory controller modeling. The modeling verification was performed by modeling the memory and channel corresponding to the whole controller PHY part as a system Verilog, and link training block using Verilog code for synthesis. Schematic is modeled to have the same structure as possible. In this environment, the memory controller and the memory are operated to communicate through the transient simulation, and the error is judged based on the pre-encoded message of the memory. From Figure 4.4.3 To Figure 4.4.5 show initialization step sequence of the boot-up operation. The boot-up operation means a step from initial power-on to the initial setting while communicating at a low speed in a state in which the memory can't guarantee high-speed communication, and preparing for high-speed command training. That is, it means initialization of the memory device. First, the MRW commands are sent sequentially to the memory. After that, ZQ calibration and ZQ latch are performed. In reality, the MRW should be written for all addresses, but the prototype chip is designed to input initial values only for the addresses that should be actually written, except for addresses that can be used as default values.

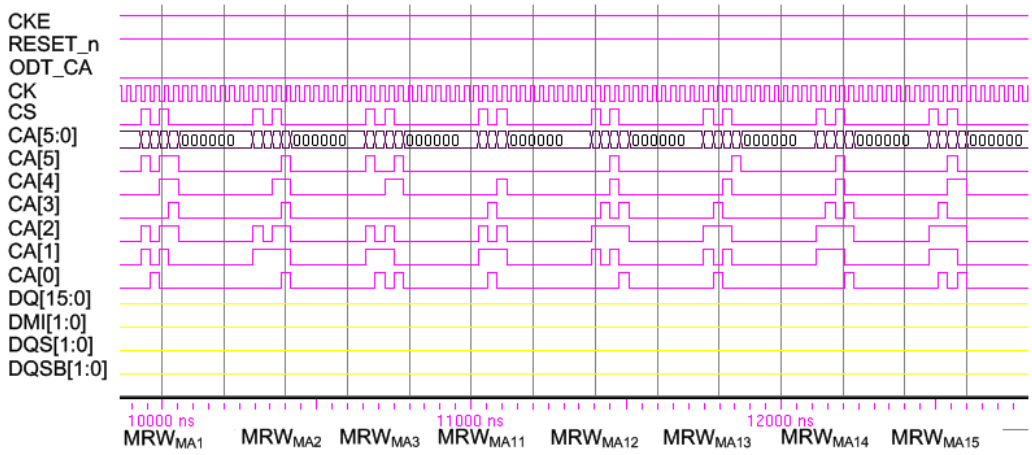


Figure 4.4.3 Boot up sequence - initialization step 1

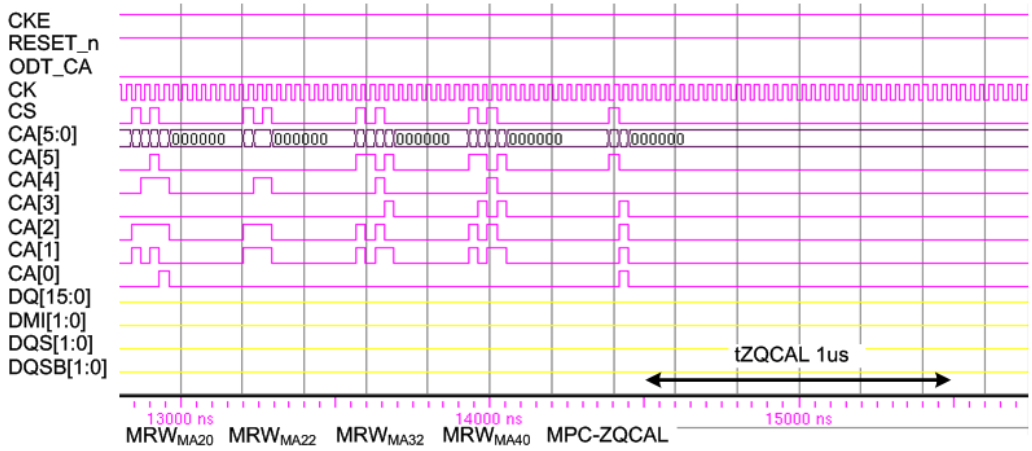


Figure 4.4.4 Boot up sequence - initialization step 2

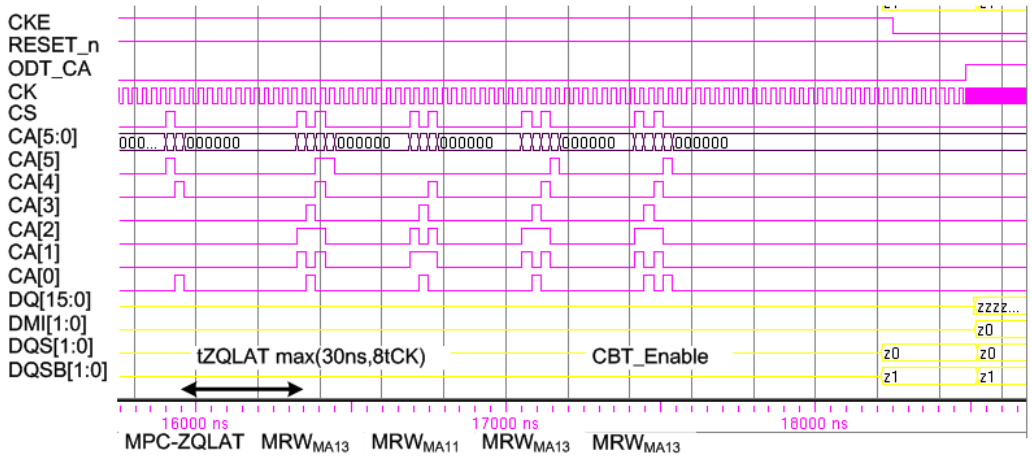


Figure 4.4.5 Boot up sequence - initialization step 3

From Figure 4.4.6 to Figure 4.4.13, modeling simulation results of the command training are shown. Figure 4.4.6 shows entry timing of the command training. When CKE signal goes low, the command training starts after the *tCAENT*. The operation speed is changed from boot frequency to the normal operation speed of the LPDDR4 memory. For example, speed is changed from 33MHz to 2133MHz. Figure 4.4.7 shows the setup and hold timing margin for the reference voltage sweep at the command training. The both margins are 2ns. During the CBT, the CS signal is sampled at the rising edge of CK to read the CA signals at the point where CS becomes H, and feed back this value to the DQ signals. In CBT, CS training is performed first and CA training is performed later. Therefore, in order to prevent the timing margin of the CA signal from affecting the CS training, the CA signal is designed to have a wide timing margin before and after the CS signal transmission.

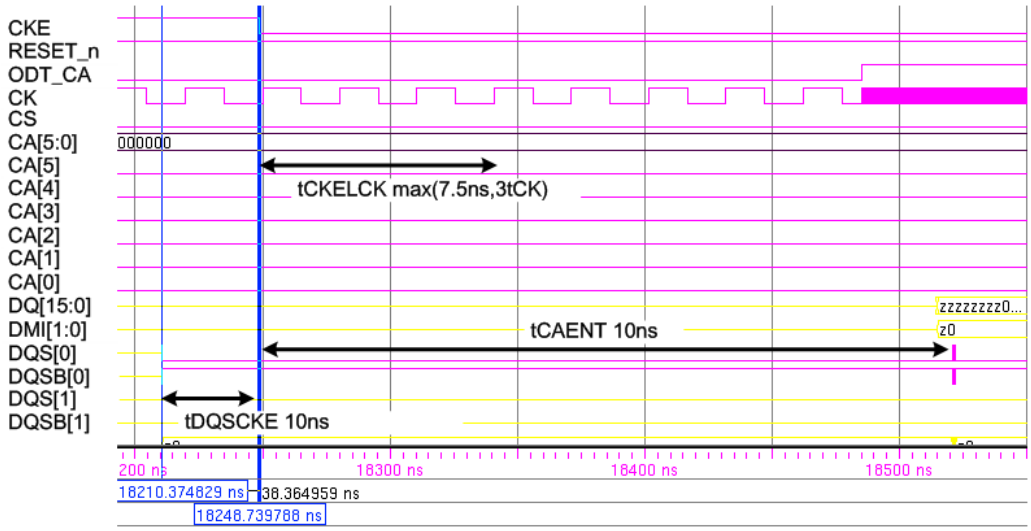


Figure 4.4.6 Timing of command training entry

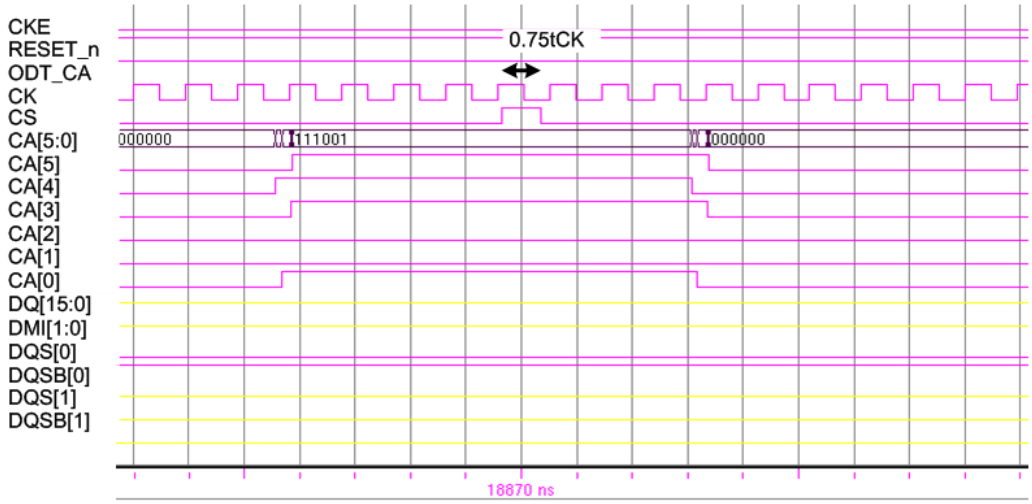


Figure 4.4.7 Training pattern of CS training

Figure 4.4.8 shows the training pattern of the CS training. The values of the CA are fixed to predefined value, and only CS signal toggle width of 0.75 tCK when the CS training. The reason why the pulse width is 0.75 tCK is to prevent duplicated sampling.

Since the CS and CA signals are transmitted at a single data rate, they are generally transmitted with a pulse width of 1 tCK. However, there is a possibility that if 1 tCK pulse width is widened due to special environmental factors, the problem of duplicated sampling may occur during training. In order to prevent this, and to obtain the center position of the training accurately, it is necessary to train the pulse using a pulse made by decreasing both ends of the existing pulse by the same amount. Therefore, a multi-phase clock is used in the CA path, and a 0.125 tCK shifted signal and a 0.475 tCK shifted signal are merged into an OR gate to generate a training signal having a pulse width of 0.75 t CK as a mask signal having a 0.75 t CK pulse width. The CS training is the x axis timing sweep training, but if the timing pass zone is not found in the predefined reference voltage, the value of the reference voltage would be changed until it finds the timing pass zone. Also the x axis value is changed, if the voltage pass zone is not found in fix timing value.

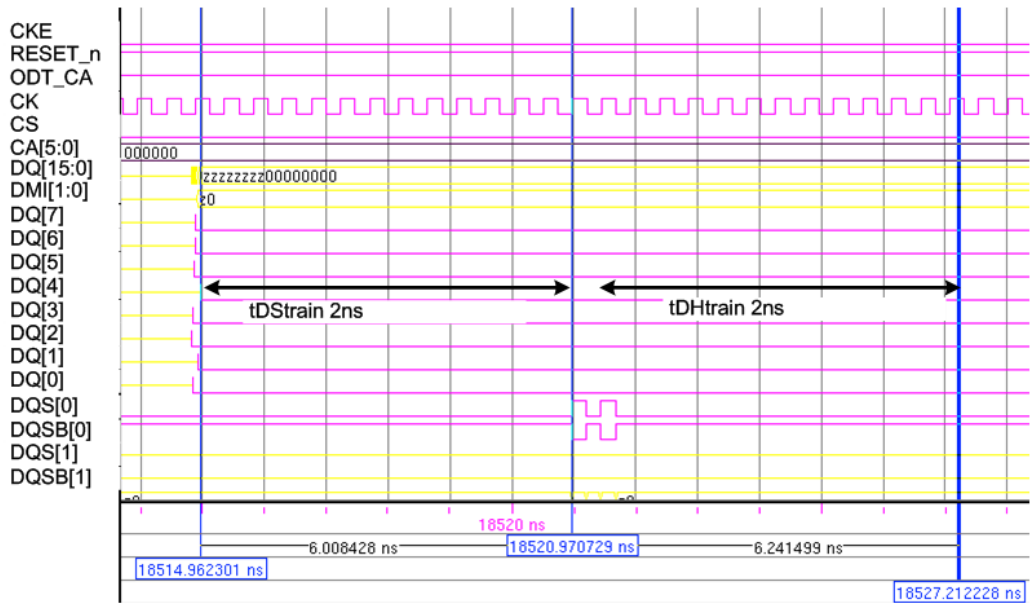


Figure 4.4.8 Setup and hold timing margin for reference voltage sweep at command training

Figure 4.4.9 shows this exceptional case of the reference voltage change. At the CS training, if there is no pass zone in particular a time code, the time code is changed and the voltage sweep is performed again to find the pass zone. This example is intended to show that it can respond to an exceptional case that rarely occurs in a real situation. In the actual environment, since the center of the eye can be moved to the default value of VREF through the ZQ calibration, there is only a slight difference, and almost no occurrence of the path region is present.

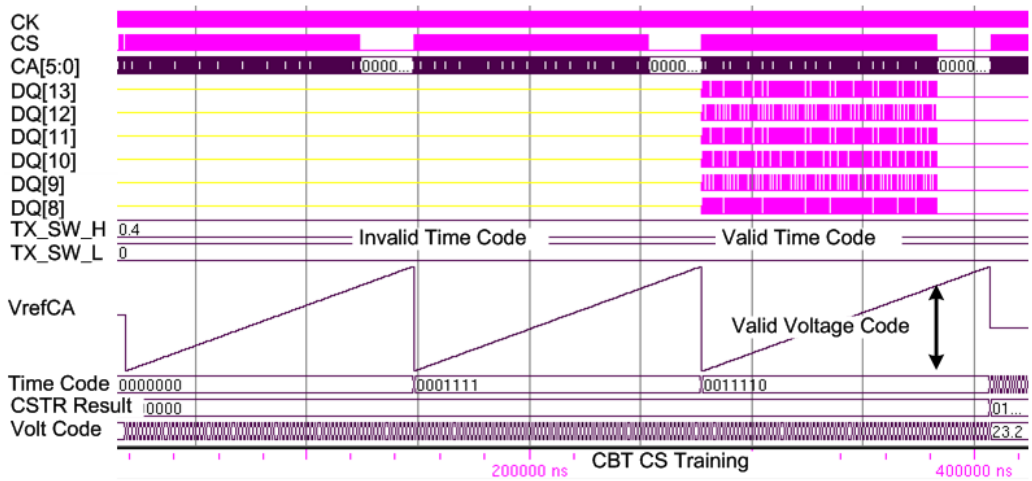


Figure 4.4.9 Reference voltage sweep in CS training

Figure 4.4.10 shows the CA training results. The 1x2y3x eye center detection algorithm is performed. First in the 1x sweep, the reference voltage code is fixed at 23.2 which means 23.2% of the VDDQ, and the timing training is performed. Second in the 2y

sweep, the timing code is fixed at 1001001 which means binary value of the phase interpolator code of 73/128, and the reference voltage training is performed. Finally in the 3x sweep, the reference voltage code is fixed at 23.2% of the VDDQ, and the timing training is performed. As the time code and voltage code change, you can see that the slope changes in the middle. This is a continuation of the process in which gain is increased and then reinitialized by adaptive training, through which the code is rapidly increased to enable the fast training claimed in this paper.

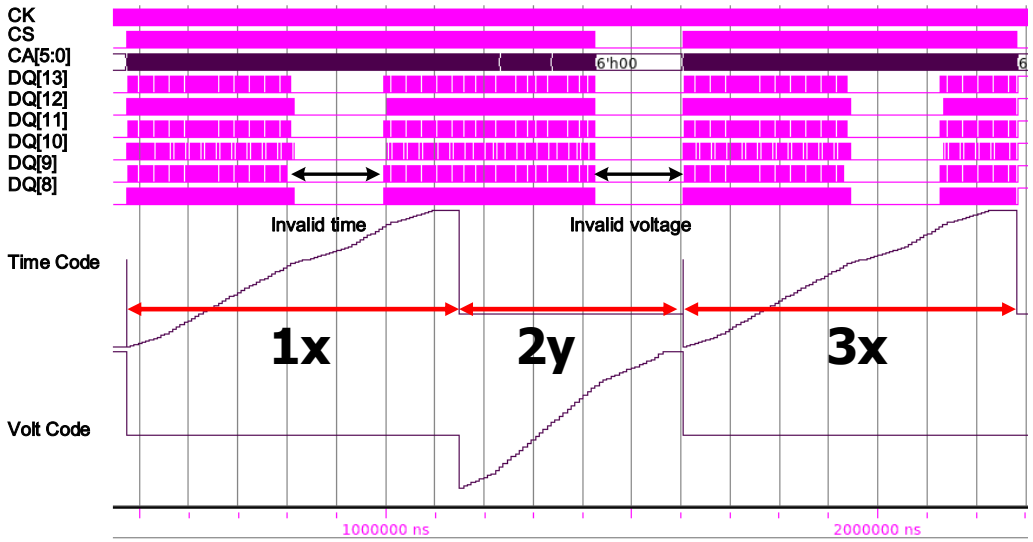


Figure 4.4.10 1x2y3x eye detection algorithm in CA training

Figure 4.4.11 shows training patterns of the CA training. The training patterns are changed in the order of A-0-B-0-C-0-D-0-E-0-A-⋯. The 5 kinds of data patterns are used to test various environment, and 0 pattern is inserted between data patterns to prevent

timing error. Figure 4.4.12 shows the result of the command training. The value of timing code is 0011110 which means binary value of the phase interpolator code of 30/128, and the value of voltage code is 14.8% of the VDDQ.

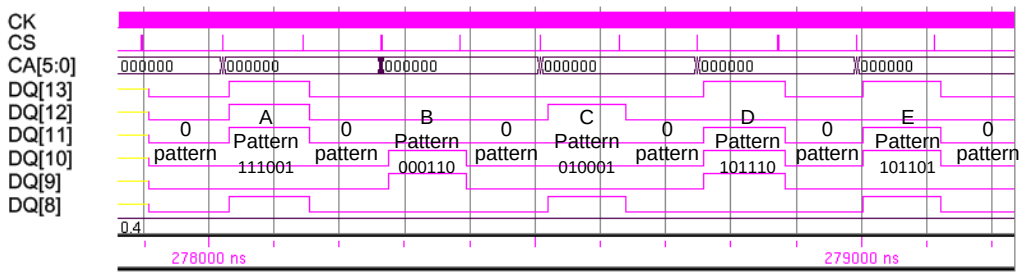


Figure 4.4.11 Training pattern of CA training: 0-A-0-B-0-C-0-D-0-E-0-A-⋯

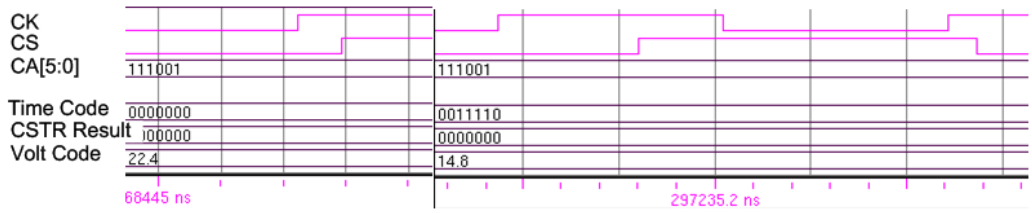


Figure 4.4.12 Result of the command training

Figure 4.4.13 and Figure 4.4.14 show the exit timing of the command training and changing of the operation speed at the end of the command training. After the command training, operation speed of the LPDDR4 memory lowers to 33MHz to write the result of the command training value at the LPDDR4 memory. As shown in Fig. 4.2.12, it will be fasted again for the write leveling.

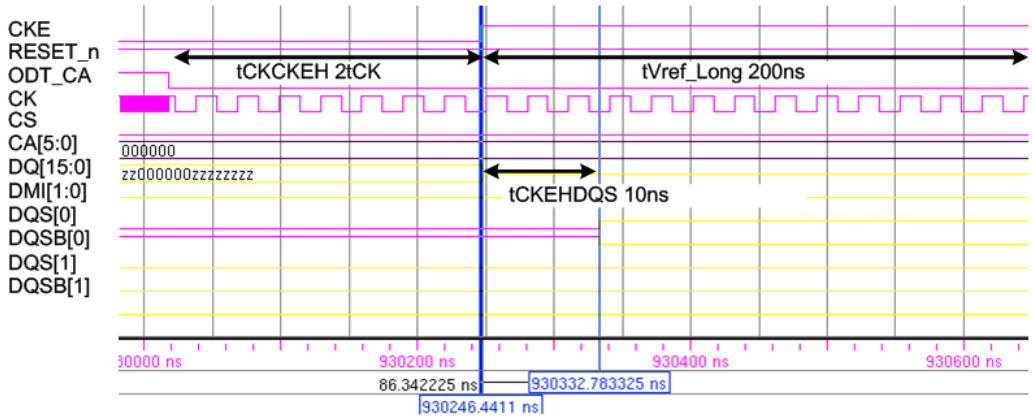


Figure 4.4.13 Exit timing of the command training

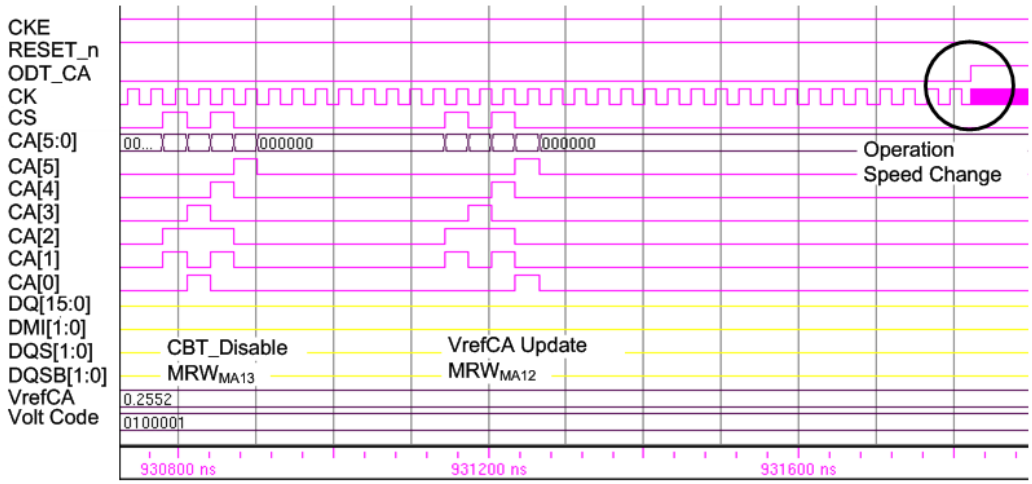


Figure 4.4.14 Operation speed change in the end command training

Figure 4.4.15 and Figure 4.4.16 show the write leveling. Figure 4.4.15 shows entry timing of the write leveling. Predefined command is sent to the memory to enter the write leveling. After *tWLDQSEN* and *tWLMRD*, the DQS and other signals are sent to the memory. As shown in Figure 4.4.16, the DQS[1] is aligned with CK at DQS1_CODE 6, and the DQS[0] is aligned with CK at DQS0_CODE 10.

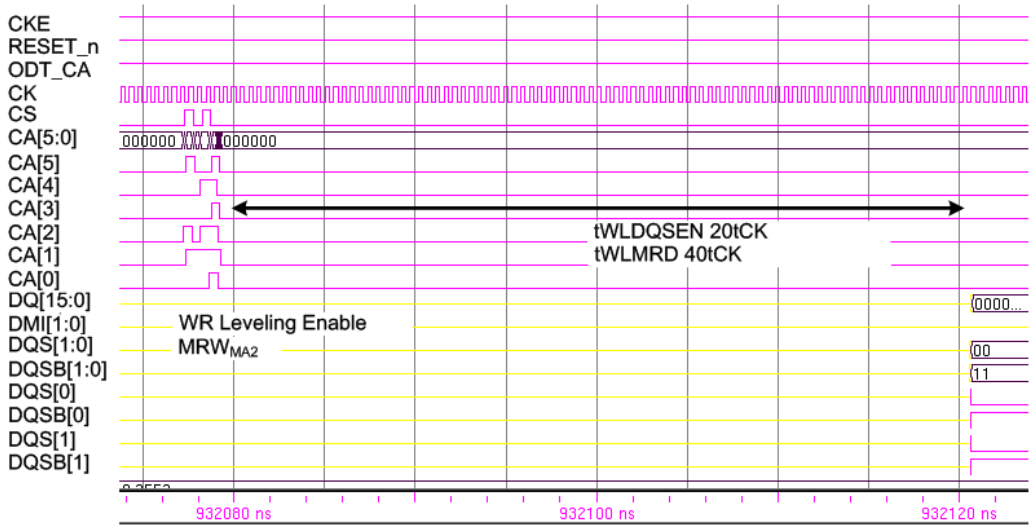


Figure 4.4.15 Entry timing of the write leveling

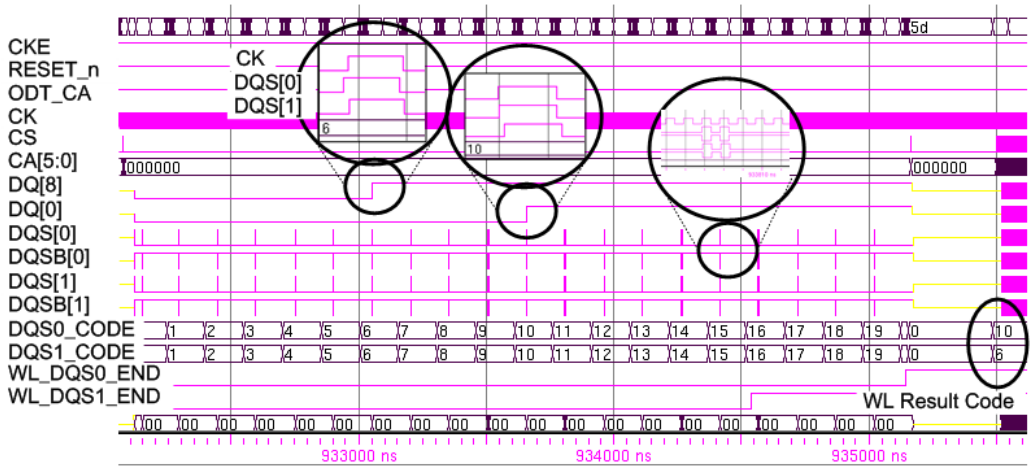


Figure 4.4.16 Write leveling

From Figure 4.4.17 to Figure 4.4.19, simulation results of the read training are shown. As shown in Figure 4.4.17, first, the phase interpolator code is swept to lock the phase interpolator code. And each delay line in DQ is swept to check the skews of the each DQ.

Figure 4.4.18 shows command transmission of the read training. The DQs receive predefined clock patterns to lock the reference voltage. Figure 4.4.19 shows the result of the read training. From DQ[0] to DQ[7], including DMI[0], the DQs have common eye open window. And DQS[0] signal leads every DQs to compensate DQS buffering delay in memory controller.

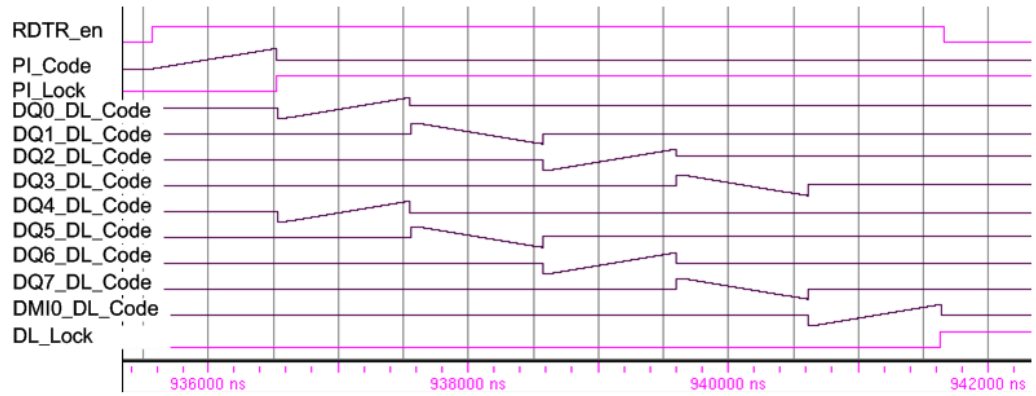


Figure 4.4.17 Training code sweep of the read training

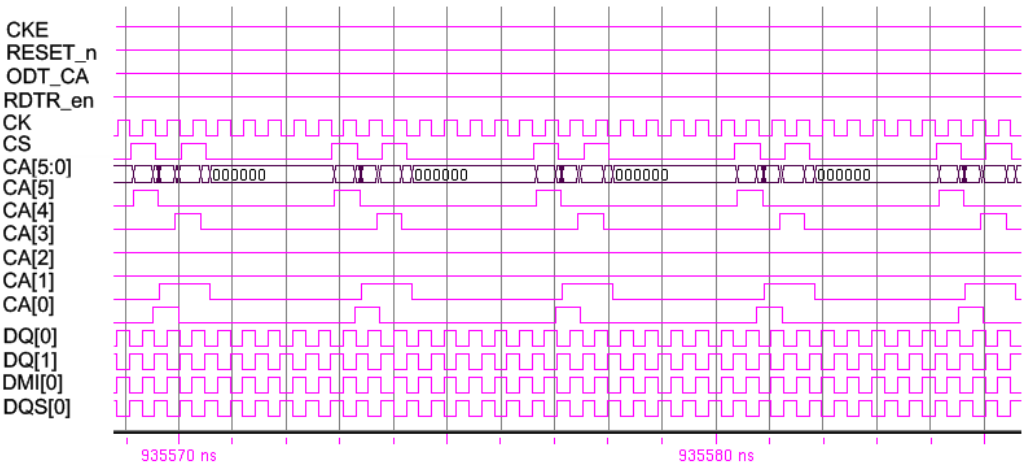


Figure 4.4.18 Environment of the read training

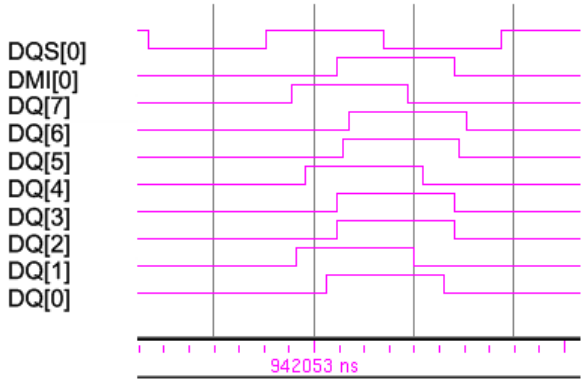


Figure 4.4.19 Result of the read training

Figure 4.4.20 and Figure 4.4.21 show the write training. In Figure 4.4.20, the min code is reflected to each DQ code to compensate the *tDQS2DQ* skew. For example, DQ[0] code is reduced from 77 to 24 by subtraction of common min code 53. Figure 4.4.21 shows the write training patterns. Consecutive 5 write commands are sent to the LPDDR4 memory. After that, 5 consecutive DQS and DQs are sent to the memory. The DQS[0] signal leads every DQs to compensate DQS buffering delay called "*tDQS2DQ* delay" in the LPDDR4 memory.

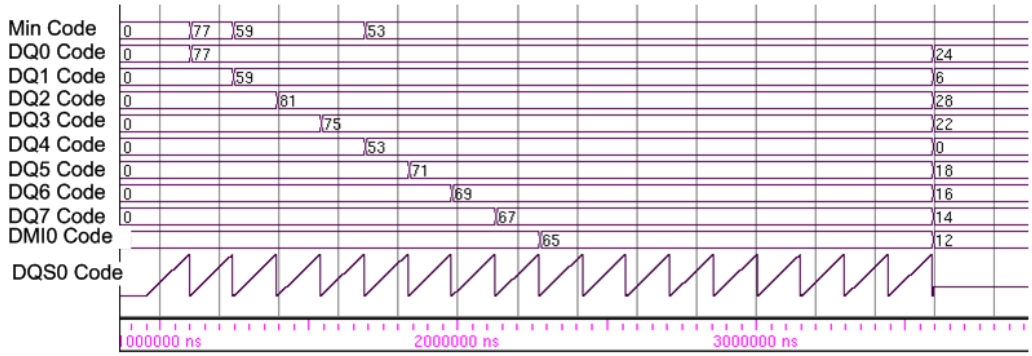


Figure 4.4.20 Training code sweep of the write training

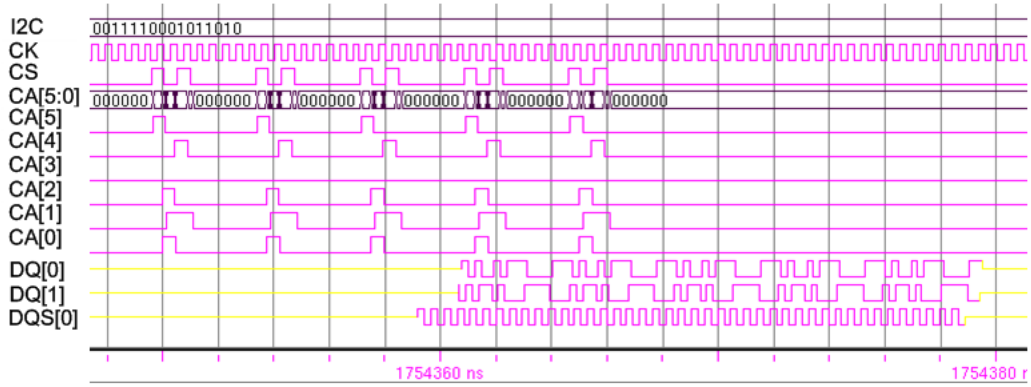


Figure 4.4.21 Write training pattern

CHAPTER 5

MEASUREMENT RESULTS

5.1 MEASUREMENT SETUP

Figure 5.1.1 shows chip microphotograph of the proposed LDDR4 memory controller. Our LPDDR4 memory controller is implemented in 65nm CMOS process and occupies an area of 12mm^2. At the right side of the chip, transceivers of DQS0, DQ0~7 and DMI0 are located. At the top side of the chip, transceivers of DQS1, DQ8~15 and DMI1 are located. The transmitters of CA0~5, CS and CK are located at upper right side of the chip. The LTFSM located at the center of the chip. The ADPLL and ADDLL are located at the bottom of the chip. The total number of pins of the chip is 176. There are four voltage domains. VDDDCO is the voltage used by the DCO of the ADPLL. VDDPLL is used by circuits other than DCO in ADPLL. VDDD is the power used by the LTFSM block. VDDQ is the power used by the transceiver and ADDLL and clock distribution circuit.

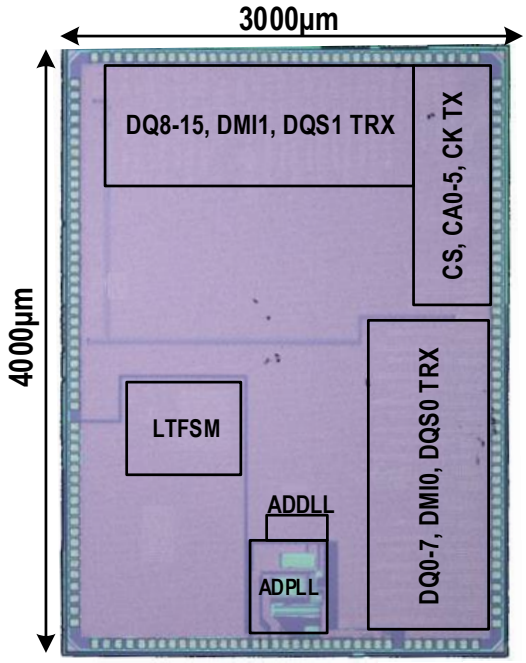


Figure 5.1.1 Chip microphotograph

As shown in Figure 5.1.2, mobile DRAM is generally configured as a package on package (PoP) type in which DRAM and MCU are overlapped. The PoP has a good signal integrity characteristics because channel length of the PoP is very short.

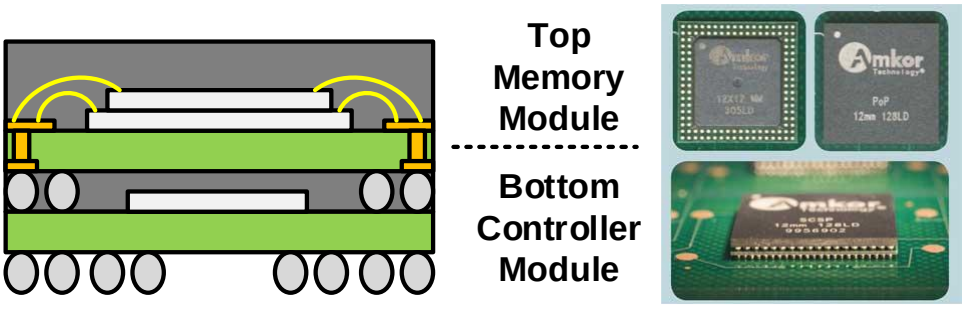


Figure 5.1.2 Package on package example

In academic research, PoP type package is hard to use, thus ball grid array (BGA) type package is used. Fortunately, the mobile memory interface does not use only the PoP type evaluation, and there are cases where two packages are arranged side by side on the PCB as shown in Figure 5.1.3.

Figure 5.1.3. Iphone 4 PCB

In this thesis, as shown in Figure 5.1.4, we design the test board by placing our MCU and LPDDR4 memory side by side.

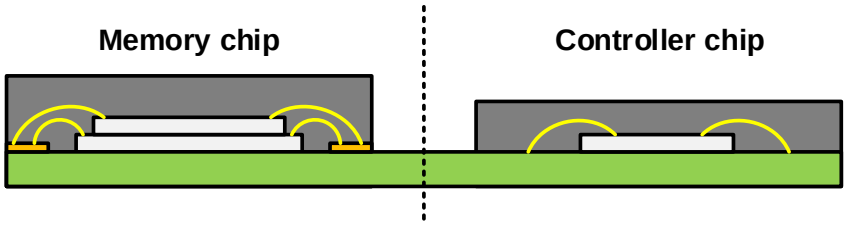


Figure 5.1.4 Test board with placing MCU and LPDDR4 memory side by side

Before we start talking about the test board, we design a BGA package that minimizes the length difference of all I/O to minimize the occurrence of skew in the package. Our BGA has 4 layers. Figure 5.1.5 and Figure 5.1.6 show the cross section from bottom layer to top layer. Top and 2nd layer is VSS domain and 3rd layer is VDD domain. In the case of VSS, since the PLL is a noise-sensitive circuit, only the VSS of PLL is provided separately from the remaining circuits.

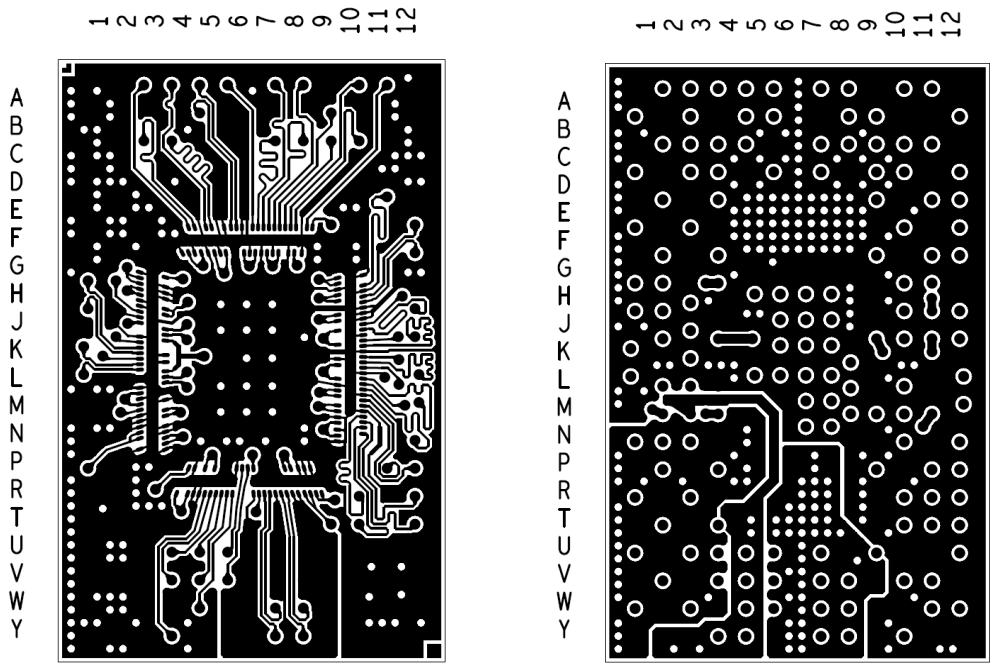


Figure 5.1.5 Top (left) and 3rd (right) layer of BGA package

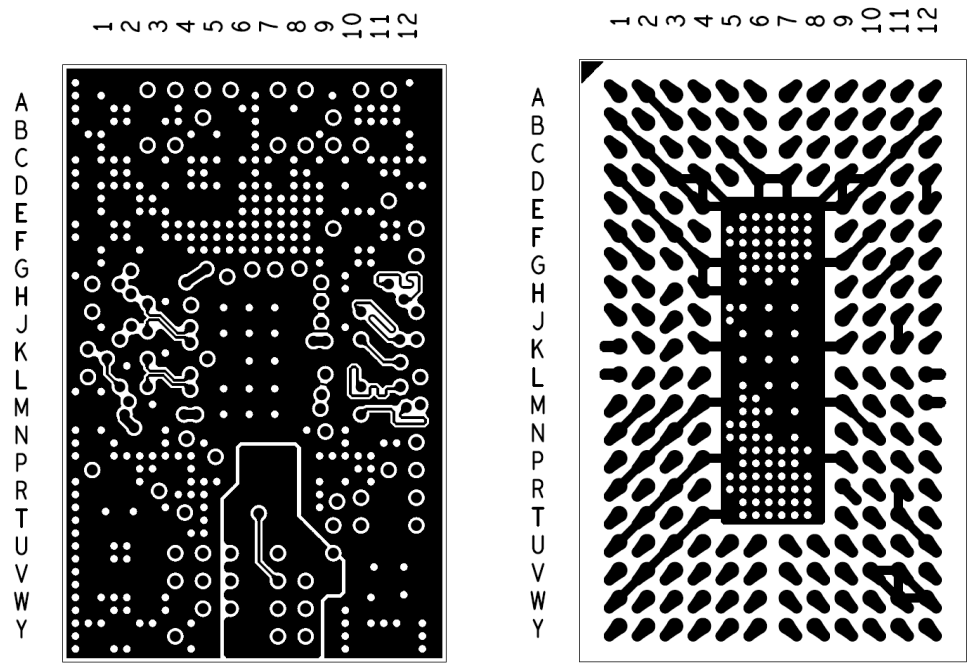


Figure 5.1.6 2nd (left) and bottom (right) layer of BGA package

In order to measure the proposed LPDDR4 memory controller, three types of test boards were fabricated. Figure 5.1.7 shows the first test board which is named LP4_TEST. LP4_TEST board is made for measurement of sub blocks of the MCU including ADPLL, ADDLL, DCDL, TX and RX. CK, DQS0, DQ0, DQ1, CS and CA0 can be measured with SMA. The rest of the I/O signals are designed to be probing through test points.

Figure 5.1.7 LP4_TEST board

Figure 5.1.8 shows the rest of test boards which is named LP4_TP and LP4_noTP. The LP4_TP board is configured so that the MCU can communicate with the LPDDR4 1-channel. The initialization sequence and the margin test result between the MCU and LPDD4 are stored in I2C block of the MCU and this value can be read to PC to check the training results. As with the LP4_TEST board for debugging, I/O signals can be measured by probing test points. LP4_noTP board is designed with test points removed by considering the effect on signal integrity.

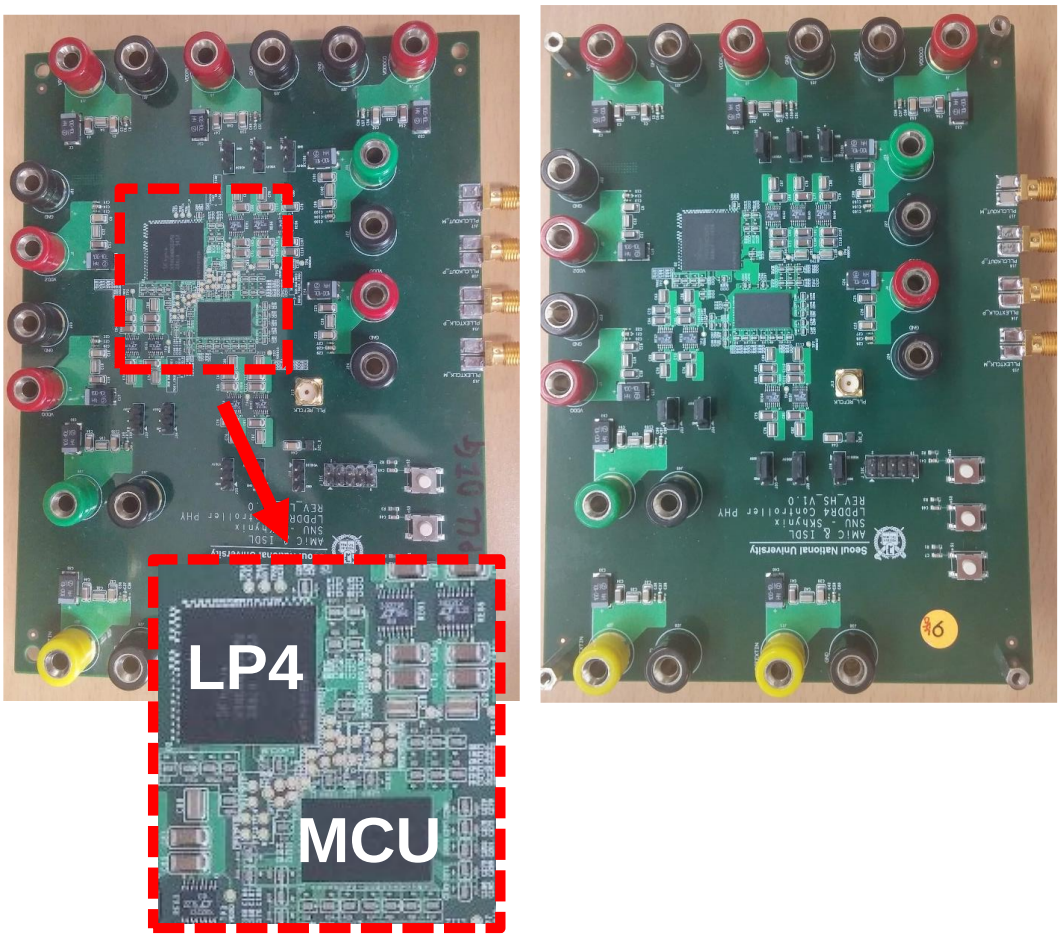


Figure 5.1.8 LP4_TP board (left) and LP4_noTP board (right)

Figure 5.1.9 and Figure 5.1.10 show the measurement setup. The measurement setup consists of a pulse function arbitrary noise generator 81160A, a DC power supply E3631A, a test board, an i2c board and a PC. 81160A generates 66MHz reference clock of the ADPLL. E3631A generates VDDD, VDDQ, VDDPLL, VDDDCO, VSS and VSSPLL. The I2C board and PC initialize the MCU and analyze the measurement results.

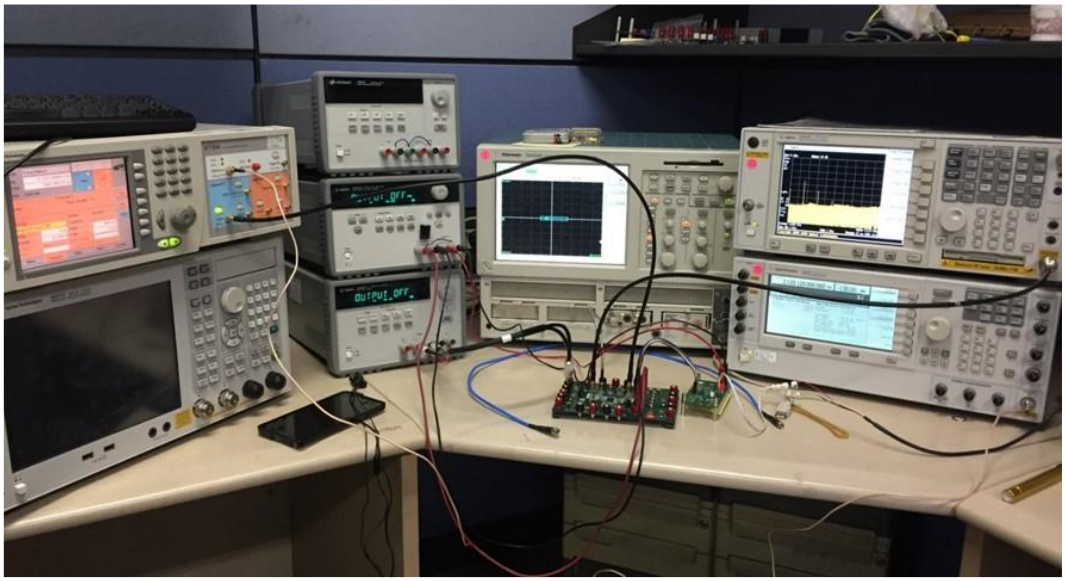


Figure 5.1.9 Photograph of measurement setup

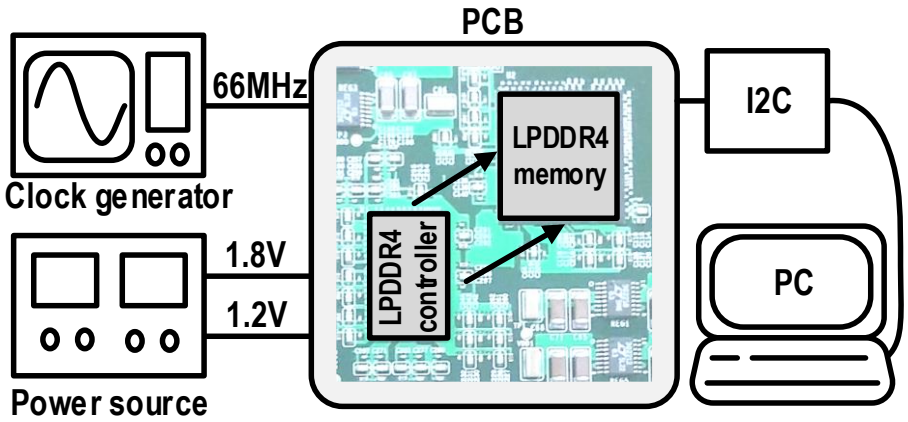


Figure 5.1.10 Block diagram of measurement setup

5.2 MEASUREMENT RESULTS OF SUB-BLOCK

5.2.1 ADPLL WITH MULTI-MODULUS DIVIDER

The ADPLL must be able to create an output clock in the range of 1333MHz to 2133MHz. In addition, low jitter of the ADPLL must be low because it is directly associated with the eye margin of signal transmission and reception. Figure 5.2.1 shows the integrated jitter measurement results of the ADPLL according to operation frequency. The integrated jitter from 1kH to 100MHz is 3.86ps, 2.85ps, 2.71ps and 2.53ps at 1333MHz, 1600MHz, 1866MHz, and 2133MHz, respectively. The circuit area of the phase-locked loop is 0.39mm2 and consumes 17.47mW at 2133MHz operation.

Figure 5.2.2 shows the performance summary of integer-N ADPLL. The reference clock is 66.66MHz and the N value is changed to 32, 28, 24, and 20 to generate clocks of 2133MHz, 1866MHz, 1600MHz and 1333MHz, respectively. The maximum power consumption is 17.47mW at 2133MHz and the worst integrated RMS jitter is 3.86ps at 1333MHz. The area of ADPLL is 0.39 mm^2.

Figure 5.2.1 Measurement results of ADPLL

Process			**TSMC 65nm LP**			
Reference	**Frequency**	**[MHz]**	**66.66**			
Divider	**Factor**	**[-]**	**32**	**28**	**24**	**20**
DCO	**Target frequency**	**[GHz]**	**2.133**	**1.866**	**1.6**	**1.333**
	Range	**[GHz]**	**2.4**	**2.2**	**1.9**	**1.4**
	Power	**[mW]**	**3.17**	**2.74**	**2.3**	**1.9**
PLL	**Integrated RMS jitter (10k-100MHz)**	**[ps]**	**2.53 (RMS)**	**2.71 (RMS)**	**2.85 (RMS)**	**3.86 (RMS)**
	Phase noise @ 1MHz	**[dBc/Hz]**	**-98.42**	**-99.53**	**-96.75**	**-96.14**
	Power	**[mW]**	**14.3**	**9.6**	**8.4**	**7.2**
Total	**Power**	**[mW]**	**17.47**	**12.34**	**10.7**	**9.1**
Area		**[mm^2]**	**0.39**			

Figure 5.2.2 Performance summary of ADPLL

5.2.2 ADDLL WITH TRIANGULAR-MODULATED PI

Figure 5.2.3 and Figure 5.2.4 shows the measurement results of the ADDLL. To verify the operation of the ADDLL, an oscilloscope was used to observe whether the waveforms of the CK0 and CK180 signals were 180° locked. As shown in Figure 5.2.3 and Figure 5.2.4, the ADDLL performs 180° of locking well in the operating frequency range of 266 to 2133MHz. The global delay-locked loop and local delay-locked loop occupy areas of 0.047mm2 and 0.027mm2 respectively. The global delay-locked loop power offed after lock and the local delay-locked loop consumes 3.71mW at 2133MHz.

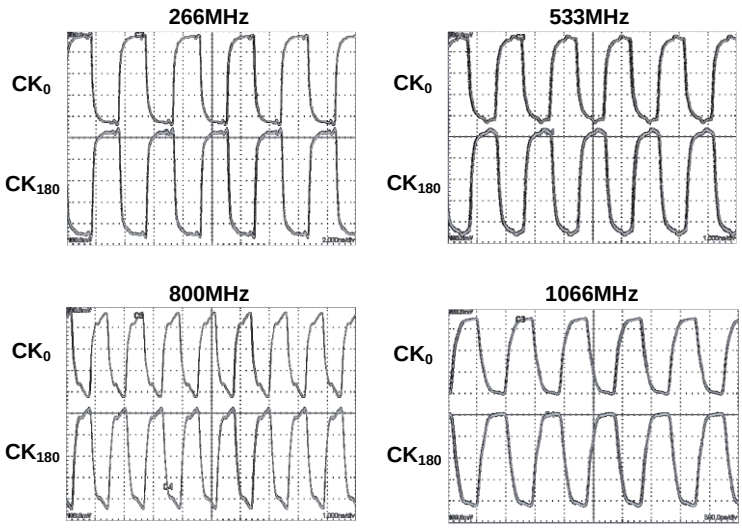


Figure 5.2.3 Measurement results of ADDLL at 266, 533, 800 and 1066MHz

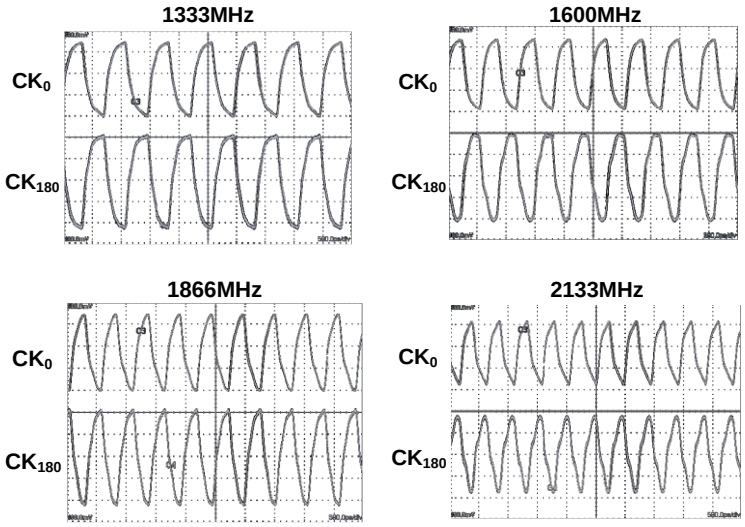


Figure 5.2.4 Measurement results of ADDLL at 1333, 1600, 1866 and 2133MHz

Figure 5.2.5 shows the measurement of differential non-linearity (DNL) at 266MHz and 2133MHz to verify the performance of the PI. The measured DNL has a value of -0.625 to 0.906 LSB and -0.727 to 0.911 LSB for 266MHz and 2133MHz, respectively.

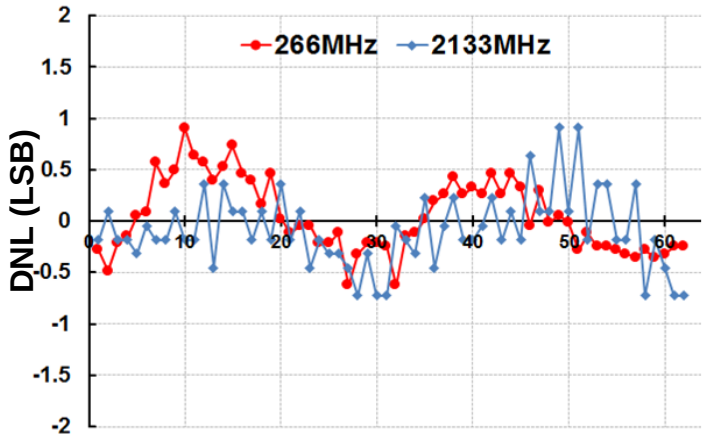


Figure 5.2.5 Measured DNL of ADDLL

5.2.3 COARSE-FINE DCDL

Figure 5.2.6 show measurement results of the coarse-fine DCDL. The coarse-fine DCDL has 256 step of delay control code, and average resolution is 4ps. The dynamic range of the coarse-fine DCDL is 1106ps.

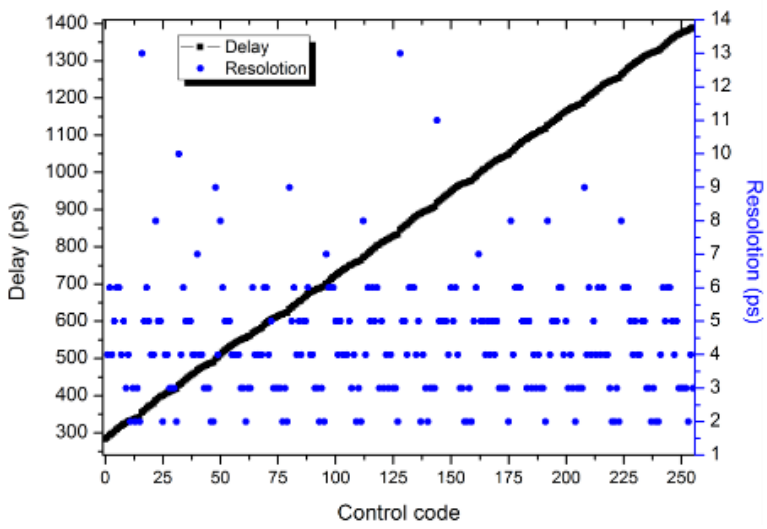


Figure 5.2.6 Measurement results of DCDL

5.3 LPDDR4 INTERFACE MEASUREMENT RESULTS

Figure 5.3.1 shows overall measurement results of the LPDDR4 memory controller operation. All operations of the LPDDR4 memory controller including the READ and WRITE operation are confirmed from 533Mbps to 4266Mbps. The resolution of time and

voltage code are 1/128tCK and 1mV, respectively. From 533Mbps to 3200Mbps, 1.25V VDDQ is used without termination. At 3733Mbps and 4266Mbps, supply voltage is risen from 1.25V to 1.3V for the internal supply voltage drop and RZQ/4 termination is enabled. The MCU stores the pass start code and the length value of the eye detection result in the register during the training process of the DRAM. During the measurement, I2C was used to retrieve the stored registers. The measured write timing margin is 23/128tCK and 278mV at 4266Mbps. The measured read timing margin is 27/128tCK and 32mV at 4266Mbps. All training sequences are verified at data-rate from 533Mb/s/pin to 4266Mb/s/pin.

Speed (Mbps)		533	1066	1600	2133	2666	3200	3733	4266
CMD TR	Time	124	122	113	120	96	87	87	81
	Voltage	403	403	403	403	403	403	403	403
Write Leveling		PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS
Read Training		PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS
Write Training		PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS
Normal Operation		PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS
Write Margin	Time	54	47	55	49	35	22	26	23
	Voltage	379.2	379.2	379.2	302.4	302.2	268.8	316.8	278.4
Read Margin	Time	47	46	41	39	46	35	28	27
	Voltage	432	436	432	352	320	304	140	32

Figure 5.3.1 Measurement results of trainings

Figure 5.3.2 shows the measured shmoo plot of write margin test and read margin test. The measured eye width and height of write operation are 0.27UI and 221mV, respectively. For a read operation, the measured eye width and height are enhanced by 94% from 0.17UI to 0.33UI and 9% from 55mV to 60mV by asynchronous feedback, respectively.

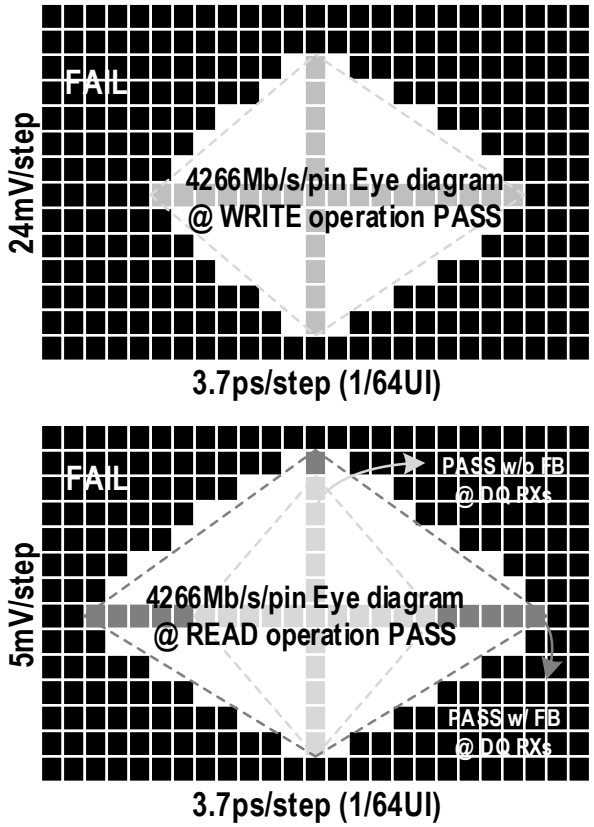


Figure 5.3.2 Shmoo plot of WRITE and READ margin test at 4266Mb/s/pin

The power efficiency of write and read operations at 4266Mbps/pin are 5.68pJ/bit and 1.83pJ/bit respectively. Table I compares this work with a LPDDR4 memory controller.

Table 5.3.1 Performance Comparison of LPDDR4 Memory Controller

Parameter		This work	[4.3.5]
Technology (nm)		65	10
Interface environment		Chip-to-chip on PCB	Package on package
Configuration		16-DQ, 2-DMI, 2-DQS, 6-CA, CK	16-DQ, 2-DMI, 2-DQS, 6-CA, CK
Data-rates (Mb/s/pin)		4266	4266
t_{DQSCK} variation compensation		Asynchronous feedback	Automatic gate by additional receiver
Training algorithm		Adaptive 3-step	Full scanning
Power efficiency (pJ/bit)	write	5.68	2.13*
	read	1.83	0.153*
Write margin		0.36UI, 148mV	0.51UI, 150mV
Read margin	without AF-CTLE	0.30UI, 76mV	0.44UI, NA
	with AF-CTLE	0.47UI, 80mV	

* Simulation result

CHAPTER 6

CONCLUSION

In this thesis, low power memory controller, which is operated with a LPDDR4, is proposed and designed with adaptive eye detection algorithm, which is used at the LPDDR4 memory training sequence and reduce the overall training time.

In order to achieve high bandwidth and low power consumption, mobile DRAM interface is increasing the weight of MCU training. MCU training is not performed on a one-shot basis but should be re-training periodically to compensate for PVT variations. Therefore, efficient training is a factor that greatly affects the power efficiency of the overall memory interface. The proposed adaptive eye detection algorithm adopts the adaptive gain control scheme for the eye detection process, which enables the 60 times faster training than the two-dimensional full scanning method and 1.9 times faster training than 1x2y3x eye center detection method. As a result, black-out time and power consumption at re-training decreased by 52.6%.

The proposed architecture of the LPDDR4 memory controller is designed based on the LPDDR4 memory specification in order to compose the memory system. We also proposed a transceiver and on-chip clock architecture and sub-blocks for MCUs with high

bandwidth and low power consumption. The proposed transmitter with write training, we achieved 0.27UI and 221mV write margin. The proposed clocking architecture reduces power consumption in IDLE states by 118% and 37%, respectively, compared to burst write or burst read states. We achieved 0.33UI and 60mV read margin through the proposed receiver, training, and CTLE with asynchronous feedback.

Fabricated in 65nm CMOS process, the proposed LPDDR4 memory controller occupies 12mm^2. The proposed LPDDR4 memory controller has verified the operation by making an evaluation board communicating with commercial LPDDR4 memory. The operation of the LPDDR4 memory system including CBT, RDTR, WLVL, WRTR, read margin test and write margin test, is verified from 533Mbps to 4266Mbps.

BIBLIOGRAPHY

[1.1.1] O.-H. Kwon, "Value-driven memory technology for the future semiconductor market," in GSA memory conference, 2010.

[1.1.2] S.-H. Kim, W.-O. Lee, J.-H. Kim, S.-S. Lee, S.-Y. Hwang, C.-I. Kim, T.-W. Kwon, B.-S. Han, S.-K. Cho, D.-H. Kim, J.-K. Hong, M.-Y. Lee, S.-W. Yin, H.-G. Kim, J.-H. Ahn, Y.-T. Kim, Y.-H. Koh, and J.-S. Kih, "A low power and highly reliable 400 Mbps mobile DDR SDRAM with on-chip distributed DCC," in IEEE Asian Solid-State Circuits Conf., pp. 34-37, 2007

[1.1.3] Y.-C. Bae, J.-Y. Park, S. J. Rhee, S. B. Ko, Y. Jeong, K.-S. Noh, Y. Son, J. Youn, Y. Chu, H. Cho, M. Kim, D. Yim, H.-C. Kim, S.-H. Jung, H.-I. Choi, S. Yim, J.-B. Lee, J. S. Choi, and K. Oh, "A 1.2V 30nm 1.6 Gb/s/pin 4Gb LPDDR3 SDRAM with input skew calibration and enhanced control scheme," in IEEE ISSCC Dig. Tech. Papers, pp. 44-46, 2012.

[1.1.4] K. Song, S. Lee, D. Kim, Y. Shim, S. Park, B. Ko, D. Hong, Y. Joo, W. Lee, Y. Cho, W. Shin, J. Yun, H. Lee, J. Lee, E. Lee, N. Jang, J. Yang, H.-k. Jung, J. Cho, H. Kim, and J. Kim, "A 1.1 V 2y-nm 4.35 Gb/s/pin 8 Gb LPDDR4 mobile device with bandwidth improvement techniques," *IEEE J. Solid-State Circuits*, vol. 50, no. 8, pp. 1945-1959, Aug. 2015.

[1.1.5] JEDEC standard LPDDR4 SDRAM specification, JESD209-4A, Nov. 2015.

[1.1.6] JEDEC standard LPDDR3 SDRAM specification, JESD209-3C, Aug. 2013.

[1.1.7] JEDEC standard LPDDR2 SDRAM specification, JESD209-2F, Apr. 2011.

[2.5.1] Y.-C. Cho, Y.-C. Bae, B.-M. Moon, Y.-J. Eom, M.-S. Ahn, W.-Y. Lee, C.-R. Cho, M.-H. Park, Y.-J. Jeon, J.-O. Ahn, B.-K. Choi, D.-K. Kang, S.-H. Yoon, Y.-S. Yang, K.-I. Park, J.-H. Choi, J.-B. Lee, and J.-S. Choi, "A sub-1.0V on-chip CMOS thermometer with a folded temperature sensor for low-power mobile DRAM." *IEEE Trans. Circuits Syst. II* vol. 64, no. 6, pp. 553-557, June 2016.

[2.5.2] Lattice, "MachXO2™ LPDDR SDRAM controller IP core user's guide", Feb.

2014.

[2.5.3] H. J. Lee and Y.-B. Kim, "A process tolerant semi-self impedance calibration method for LPDDR4 memory controller," in *IEEE Symp. Midwest Circuits and Systems*, 2015, pp.1-4.

[3.1.1] I. Fujimori and M. S. Nejad, "Eye monitoring and reconstruction using CDR and sub-sampling ADC," U.S. patent 7,460,589, Dec. 2, 2008.

[3.1.2] Y. Miki, T. Saito, H. Yamashita, F. Yuki, T. Baba, A. Koyama, and M. Sonehara, "A 50-mW/ch 2.5-Gb/s/ch data recovery circuit for the SFI-5 interface with digital eye-tracking," IEEE J. Solid-State Circuits, vol. 39, no. 4, pp. 613-621, Apr. 2004.

[3.1.3] B. Analui, A. Rylyakov, S. Rylov, M. Meghelli, and A. Hajimiri, "A 10-Gb/s two-dimensional eye-opening monitor in 0.13-μm standard CMOS," IEEE J. Solid-State Circuits, vol. 40, no. 12, pp. 2689-2699, Dec. 2005.

[3.1.4] H. Noguchi, N. Yoshida, H. Uchida, M. Ozaki, S. Kanemitsu, and S. Wada, "A 40-Gb/s CDR circuit with adaptive decision-point control based on eye-opening monitor feedback," IEEE J. Solid-State Circuits, vol. 43, no. 12, pp. 2929-2938, Dec. 2008.

[3.1.5] J. F. Bulzacchelli, T. O. Dickson, Z. T. Deniz, H. A. Ainspan, B. D. Parker, M. P. Beakes, S. V. Rylov, and D. J. Friedman, "A 78mW 11.1Gb/s 5-Tap DFE receiver with digitally calibrated current-integrating summers in 65nm CMOS," IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers, 2009, pp. 368-36.

[3.1.6] M. Loh, and A. Emami-Neyestanak, "A 3×9 Gb/s shared, all-digital CDR for high-speed, high-density I/O," IEEE J. Solid-State Circuits, vol. 47, no. 3, pp. 641-651, Mar. 2012.

[3.1.7] J. F. Bulzacchelli, C. Menolfi, T. J. Beukema, D. W. Storaska, J. Hertle, D. R. Hanson, P.-H. Hsieh, S. V. Rylov, D. Furrer, D. Gardellini, A. Prati, T. Morf, V. Sharma, R. Kelkar, H. A. Ainspan, W. R. Kelly, L. R. Chieco, G. A. Ritter, J. A. Sorice, J. D. Garlett, R. Callan, M. Brändli, P. Buchmann, M. Kossel, T. Toifl, and D. J. Friedman, "A 28-Gb/s 4-tap FFE/15-tap DFE serial link transceiver in

32-nm SOI CMOS technology," IEEE J. Solid-State Circuits, vol. 47, no. 12, pp. 3232-3248, Dec. 2012.

[3.1.8] H. Kimura, P. M. Aziz, T. Jing, A. Sinha, S. P. Kotagiri, R. Narayan, H. Gao, P. Jing, G. Hom, A. Liang, E. Zhang, A. Kadkol, R. Kothari, G. Chan, Y. Sun, B. Ge, J. Zeng, K. Ling, M. C. Wang, A. Malipatil, L. Li, C. Abel, and F. Zhong, "A 28 Gb/s 560 mW multi-standard SerDes with single-stage analog front-end and 14-tap decision feedback equalizer in 28 nm CMOS," *IEEE J. Solid-State Circuits*, vol. A 1.1V 2y-nm 4.35Gb/s/pin 8Gb LPDDR4 mobile device with bandwidth improvement techniques 49, no. 12, pp. 3091-3103, Dec. 2014.

[3.1.9] K. Song, S. Lee, D. Kim, Y. Shim, S. Park, B. Ko, D. Hong, Y. Joo, W. Lee, Y. Cho, W. Shin, J. Yun, H. Lee, J. Lee, E. Lee, J. Yang, H. Jung, N. Jang, J. Cho, H. Kim, and J. Kim, "A 1.1V 2y-nm 4.35Gb/s/pin 8Gb LPDDR4 mobile device with bandwidth improvement techniques," in *IEEE Custom Integrated Circuits Conf*. 2014.

[3.1.10] C.-K. Lee, M. Ahn, D. Moon, K. Kim, Y.-J. Eom, W.-Y. Lee, J. Kim, S. Yoon, B. Choi, S. Kwon, J.-Y. Park, S.-J. Bae, Y.-C. Bae, J.-H. Choi, S.-J. Jang, G. Jin, "A 6.4Gb/s/pin at sub-1V supply voltage TX-interleaving technique for mobile DRAM interface," in *IEEE Symp. on VLSI Circuits*, 2015. pp. 182-183.

[4.1.1] W.-J. Kim, Y.-H. Ro, and S.-W. Park, "Package on package," U.S. patent 8,446,018, May 21, 2013.

[4.1.2] A. Yoshida, J. Taniguchi, K. Murata, M. Kada, Y. Yamamoto, Y. Takagi, T. Notomi, and A. Fujita, "A study on package stacking process for package-on-package (PoP)," in *Proc. Electronic Components and Technology Conference*, 2006, pp. 825-830.

[4.2.1] J. Zerbe, B. Daly, L. Luo, W. Stonecypher, W. Dettloff, J. C. Eble, T. Stone, J. Ren, B. Leibowitz, M. Bucher, P. Satarzadeh, Q. Lin, Y. Lu, and R. Kollipara, "A 5 Gb/s link with matched source synchronous and common-mode clocking techniques," *IEEE J. Solid-State Circuits*, vol. 46, no. 4, pp. 974-985, Apr. 2011.

[4.2.2] J.-H. Kim, W. Kim, D. Oh, R. Schmitt, J. Feng, C. Yuan, L. Luo, and J. Wilson,

"Performance impact of simultaneous switching output noise on graphic memory systems," in *Electrical Performance of Electronic Packaging*, 2007, pp. 197-200.

[4.2.3] [Online]. Available: http://www.jedec.org/mobiletaiwan2012

[4.3.1] D.-H. Oh, D.-S Kim, S. Kim, D.-K. Jeong, and W. Kim, "A 2.8Gb/s all-digital CDR with a 10b monotonic DCO," in *Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2007, pp. 222-223.

[4.3.2] D.-S. Kim, H. Song, T. Kim, S. Kim, and D.-K. Jeong, "A 0.3–1.4 GHz all-digital fractional-N PLL with adaptive loop gain controller," *IEEE J. Solid-State Circuits*, vol. 45, no. 11, pp. 2300-231, Nov. 2010.

[4.3.3] H. Song, D.-S. Kim, D.-H. Oh, S. Kim, and D.-K. Jeong, "A 1.0–4.0-Gb/s all-digital CDR with 1.0-ps period resolution DCO and adaptive proportional gain control," IEEE J. Solid-State Circuits, vol. 46, no. 2, pp. 424-434, Feb. 2011.

[4.3.4] J.-H. Chae, M. Kim, H. Ko, Y. Jeong, J. Park, G.-M. Hong, D.-K. Jeong, and S. Kim, "A 266-2133MHz phase shifter using all-digital DLL and triangular-modulated phase interpolator for LPDDR4X interface." *Electron. Lett*., vol. 53, no. 12, pp. 766-768, June 2017.

[4.3.5] S.-M. Lee, J. Oh, J. Choi, S. Ko, D. Kim, K. Koo, J. Choi, Y. Nam, S. Park, H. Lee, E. Kim, S. Jung, K. Chae, S. Kim, S. Park, S. Lee, and S. Park, "A 0.6V 4.266Gb/s/pin LPDDR4X interface with auto DQS-cleaning and write-VWM training for memory controller," in *Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2017, pp. 398-399.

한글초록

본 연구에서는 비동기식 피드백 연속-시간 선형 이퀄라이저 및 적응형 3단계 눈 탐지 알고리즘을 갖춘 4266Mb/s/pin LPDDR4 메모리 컨트롤러를 제안하였다. 비동기 피드백 연속-시간 선형 이퀄라이저는 잡음보다 큰 오프셋을 적용하여 트레이닝 없이 DQS의 글리치를 제거하고 DQ 경로에서 결정 피드백 이퀄라이저로 작용하여 읽기 마진을 향상시킨다. 적응형 3단계 눈 감지 알고리즘은 2차원 전체 스캔 방식과 비교하여 초기화 동작 및 재트레이닝의 전력 소비 및 블랙 아웃 시간을 줄인다. 또한 적응형 3단계 눈 감지 알고리즘은 눈 검출 결과가 바뀌면 순차적으로 이진 방법을 사용하여 눈 경계를 탐색하고 해상도를 초기화 함으로써 정확도를 유지할 수 있다. 높은 대역폭을 달성하기 위한 트레이닝에 적합한 송신기 및 수신기를 제안하였다. 송신기는 위상 변환기, 디지털 제어 지연 라인, 16:1 시리얼라이저, 프리 드라이버 및 저전압 스윙 터미네이티드 로직으로 구성되어 있다. 수신기는 기준 전압 생성기, 연속 시간 선형 이퀄라이저, 위상 변환기, 디지털 제어 지연 라인, 1:4 디시리얼라이저 및 4:16 디시리얼라이저로 구성된다. 일반적으로 모바일 어플리케이션은 긴 유휴 기간을 가지고 있는데, 이에 적합하도록 유휴 기간에서 저전력을 소비하는 클라킹 아키텍쳐를 제안하였다. 프로토 타입 칩은 볼 그리드 어레이 패키지로 65nm CMOS 공정에서 구현되었으며 상용 LPDDR4 메모리 칩과 연동하여 측

정 하였다. 측정 결과, 쓰기 마진은 0.36UI와 148mV 였으며, 읽기 마진은 AF-CTLE를 적용하지 않은 경우 0.30UI와 76mV 였으며, AF-CTLE를 적용한 경우 0.47UI 및 80mV으로 향상되었다. 연속된 쓰기 및 읽기 동안 전력 효율은 각각 5.68pJ/bit 및 1.83pJ/bit 이었다.

주요어 : 모바일 메모리 컨트롤러; LPDDR4; 메모리 인터페이스; 송수신기; 적응형 눈 감지

학 번 : 2010-20764