Design Con 2016

Optimal DDR4 System with Data Bus Inversion

Thomas To, Xilinx Inc.
Changyi Su, Xilinx Inc.
Juan Wang, Xilinx Inc.
Dmitry Klokotov, Xilinx Inc.
Lizhi Zhu, Xilinx Inc.
John Schmitz, Xilinx Inc.
Penglin Niu, Xilinx Inc.
Yong Wang, Xilinx Inc.

Abstract

Xilinx's Ultrascale family FPGA High Performance (HP) IO can support at least eight 72 bit DDR4 channels. The massive amount of memory IO interface makes the system tradeoffs, such as system power, interface timing and system memory speed, a difficult task. DDR4 introduces Data Bus Inversion (DBI) feature to invert transmit data bits such that fewer data bits will pull to logic LOW in PODL_12 IO standard. Therefore, the interface will consume lower power. The benefit of this feature when applies to FPGA memory interface will be investigated.

This paper will cover the relative system power improvement using DBI in DDR4 system. Because FPGA memory interface usage covers a wide range of applications, this report will vary different system activity and memory access profile. The system power saving using DBI in Write and Read operations will be compared.

Similar to system without DBI, such as DDR3, the channel interface will need initial calibration to obtain optimal channel timing. DBI in DDR4 shares a common pin with data mask (DM) and TDQS. This paper will discuss the methodology to de-skew the memory interface with DBI enabled. The concept will be illustrated to obtain optimal DQ and DQS in both Write and Read directions when DBI is enabled.

DBI feature targets system power reduction but it can have additional signal integrity benefit because it enables fewer switching bits. This results in lowering system noise and in improving channel jitter. The methodology to evaluate the impact and the results will be analyzed.

Key Terms– Memory IO interface, DDR4, System Power, Data Bus Inversion

Author(s) Biography

Hing Y (Thomas) To (SM'11) is a Technical Director in System Memory Signal Integrity & Device Power Group at Xilinx, Inc. Prior to joining Xilinx, Thomas was with NVIDIA Advanced Technology Group focused on high speed (32GTs) circuits & system channel designs and supported different test chips for different advanced process nodes such as 20nm SOC & 16nm FINFET process. Before NVIDIA, Thomas worked for Intel for more than 16 years covered and led many different types of system memory IO development such as Sandy Bridge Server DDR IO and covered many different system memory technology ranging from DDR1 to DDR4. Thomas received his PhD degree in Electrical Engineering from the Ohio State University in 1995 & he has over 37 patents in the fields of mixed signal IO circuits and system memory configurations as well as high speed clocking for high speed memory designs.

Changyi Su (S'08-M'11) is a Senior Signal and Power Integrity Engineer at Xilinx Inc. Her responsibilities include signal integrity in high-speed circuits & system channels, jitter and timing impact from power delivery network, and memory system characterization. She received a Ph.D. degree in electrical engineering from Clemson University, Clemson, SC.

Juan Wang is a Staff Signal Integrity engineer at Xilinx Inc. She has been focusing on memory interface timing analysis such as DDR4/DDR3/RLDRAM3 and corresponding lab verification. Prior to Xilinx, she worked for Juniper as signal integrity engineer for more than 5 years supporting system design 10GE/XFI/XLAUI/SFI/sGMII/rGMII/PCIE/DDR3 signal integrity modeling, simulation and measurements. Juan received her MSEE from University of Missouri-Rolla and Tsinghua University.

Dmitry Klokotov is a Senior Signal and Power Integrity Engineer at Xilinx Inc. He specializes in modeling and validation of system-level power delivery networks, and characterization of crosstalk and power supply noise induced jitter in high-speed I/O interfaces. He received his MS and PhD degrees in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign, and BS degree in Physics from Saratov State University, Saratov, Russia.

Lizhi Zhu is a staff engineer for 2 ½ years at Xilinx Inc. He has been focusing on DDR4/DDR4 validation and correlation and FPGA memory IP development. He worked at Spirent, Curtiss Wright, Extreme Networks, and Teradyne for more than 15 years as an ASIC/FPGA design engineer covering 10GE, 40GE, SRIO, PCIE, DDR3 and QDR2. He

graduated with a MSEE degree from University of Southern California and with a BS Physics degree from Peking University.

John Schmitz is a senior manager at Xilinx Inc. He manages memory IP development team covering DDR3/DDR3L/LPDDR3/RLDRAM3/QDR2+/QDR4/DDR4.

Penglin Niu is an engineer manager at Xilinx. Her team is responsible for SI/PI modeling methodology, and product SSN and PDN analysis. She was the signal integrity lead for memory interface in Xilinx before the management position. Prior to Xilinx, she worked for Intel as signal integrity lead and package design lead. She was deeply involved in high speed DDR3/DDR3L system design and high performance CPU package design. Penglin received her Ph.D. degree from University of Illinois Urbana-Champaign, and M.S. degree from University of Missouri-Rolla.

Yong Wang is currently a Director of Engineering at Xilinx leading Device Power and Signal Integrity Group since 2011. His team owns Xilinx product families' SI/PI methodology development, noise/timing/jitter analysis, interface timing such as DDR4/3, and corresponding verification/characterization. Prior to joining Xilinx, Mr. Yong Wang has been system design lead and SI/PI lead of several companies such as NVIDIA, MetaRAM, and HP/Intel. He led the world first 16GB and 32GB R-DIMM design, validation and production with patented memory buffer ASIC design when he was system lead with MetaRAM. In NVIDIA/HP/Intel, he provided technical leadership in the areas such as but not limited to, IA-64 system front-side parallel bus channel timing, serial link channel analysis, system level power modeling, and on-die power grid noise/timing analysis and timing/noise validation in the lab. Mr. Yong Wang received his M.S. degree in Electrical Engineering from Colorado State University and B.S. degree in Electrical Engineering from Peking University. Mr. Yong Wang has 21 US patents issued and several publications including best paper rewards in conferences like EPEP and ECTC.

1. Introduction

Computing demand has been growing exponentially over the past two decades [1], Figure 1 shows the performance requirement in TFLOPs (teraFlops) versus calendar years in a logarithm scale. The need is fueled by the requirement to solve many fundamental and life changing problems, such as the need to accurately model and predict weather system. This growth faces many practical limitations and challenges such as total system power limitation, memory technology and memory bandwidth, system reliability etc. System engineers have been pushing against the power and memory walls. The power efficiency and the memory bandwidth have been increased steadily over the recent years and they are expected to increase exponentially in the near future. A typical power consumption of a computing system is shown in Figure 2. [2] The system memory power dissipation could range from 19% to 48% of the total system power. Traditionally, CPU power is the dominating factor but system memory power has steadily become a major factor.

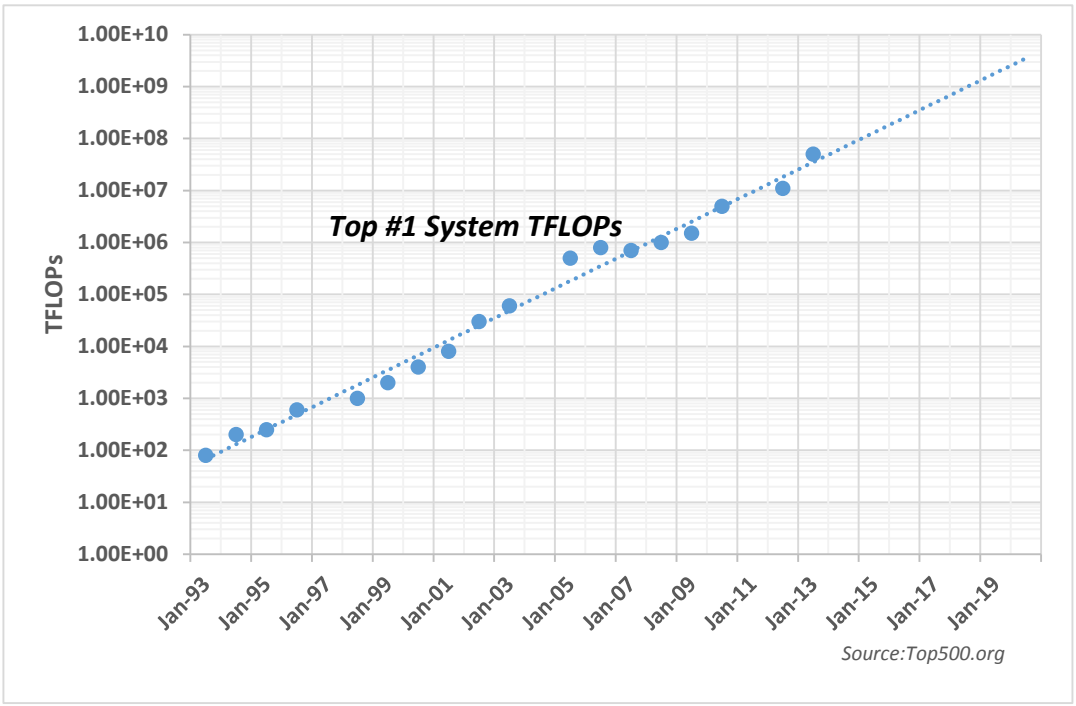


Figure 1 High Performance Computing Performance Trend

DDR4 succeeds DDR3 as the main stream system memory of choice in server and data center system. Besides many architecture improvements, DDR4 focuses on power efficiency improvement over DDR3.

Enabling the data bit inversion (DBI) in DDR4 will provide additional power savings but in order to capitalize this benefit, the DBI bit must be trained and calibrated correctly before putting into use. This paper will quantify the power saving under different work load conditions. The calibration of the interface using DBI will be presented for the optimization of the system power and performance. Section 2 will provide a background

of power improvement in DRAM technology over the past generations and it will provide an example of DDR4 memory unit power distribution and the IO power difference between DDR4 and DDR3. Section 3 illustrates the data bus inversion (DBI) concept in DDR4. System power improvement with and without DBI enabled will be compared in Section 4. Section 5 discusses the data bit training and calibration scheme together with DBI enabled. With DBI enabled, the system can also improve signal integrity and system margin. Section 6 illustrates the relative power noise improvement based on the IO power reduction when DBI is enabled. Section 7 shows the validation system set up and measurement techniques to quantify the DBI signal integrity improvement results. Section 8 will conclude the effectiveness of DBI usage and analysis key take away.

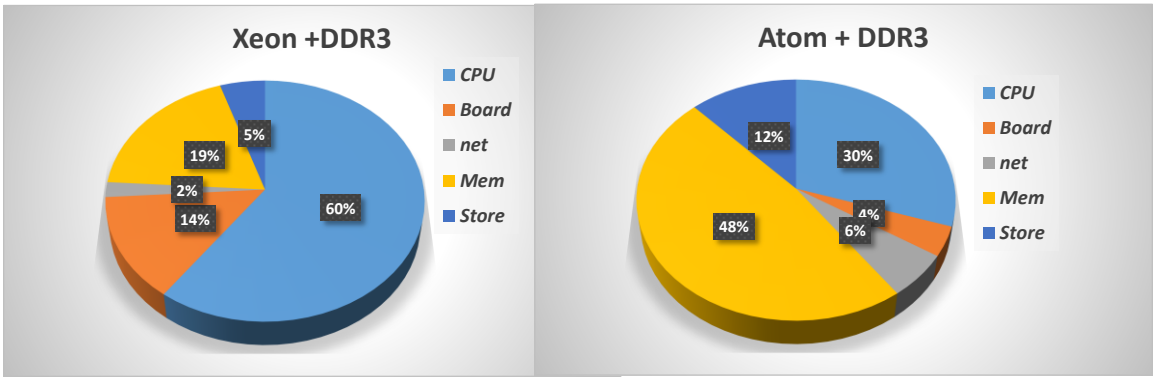


Figure 2 Typical Power Distribution in Computing System

2. System Memory Power Improvement Approach

DRAM manufacturers have been improving the total DRAM power by means of scaling down the process technology nodes almost every year, Figure 3. [3] When comparing DDR3 to DDR4 at the same technology process node, the back ground power reduction in DDR4 can be up to 35% as shown in Figure 4.[4] Furthermore, the IO voltage supply has also been scaled down from DDR, at 2.5V, to DDR4, at 1.2V, as shown in Figure 5. These are the traditional methods to improve DRAM power. But the effectiveness of the trend is slowing down. The scaling of process technology nodes and the reduction of DRAM IO voltage are expected to further slowdown for the future DRAM technology.

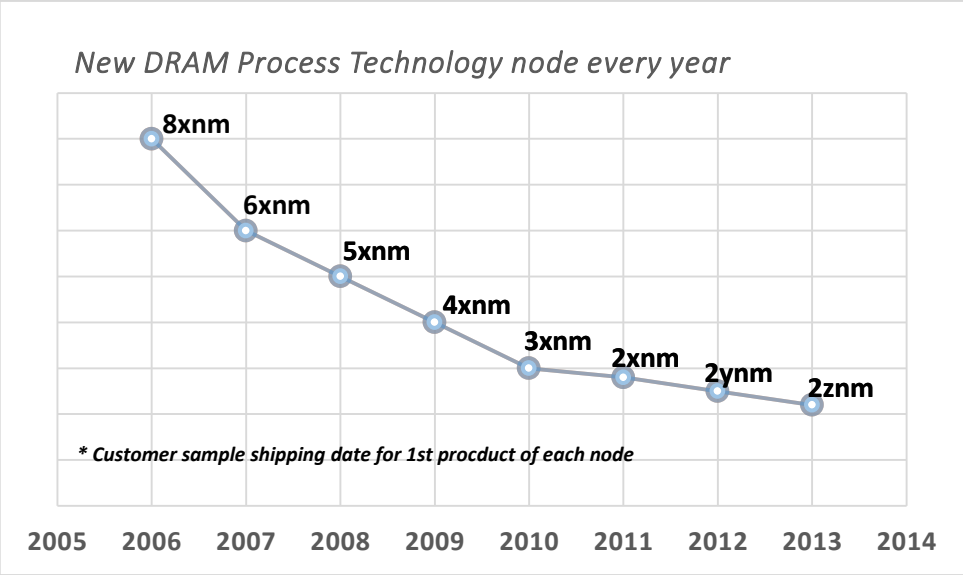


Figure 3 DRAM Technology Process Node Trend

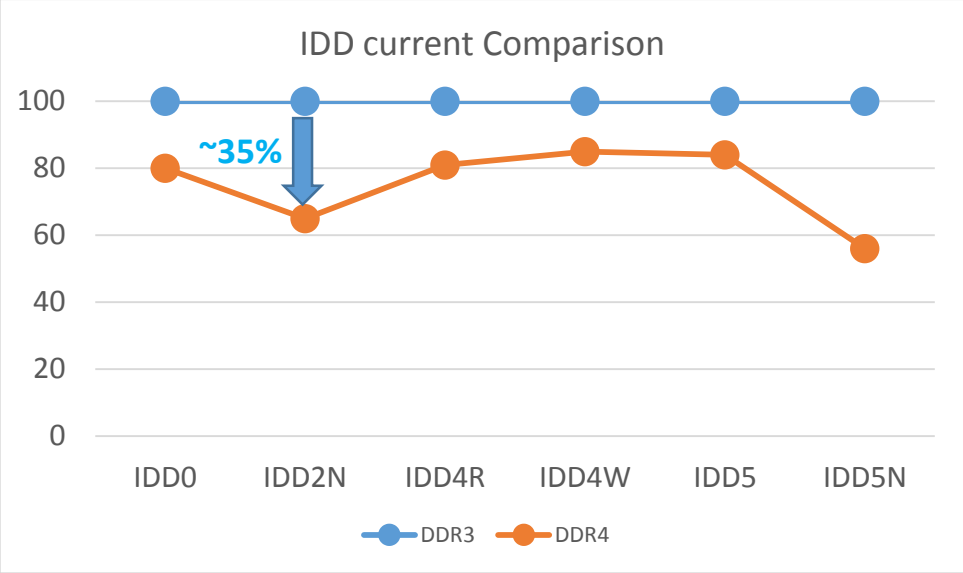


Figure 4 Power Improvement Comparison between DDR3 and DDR4

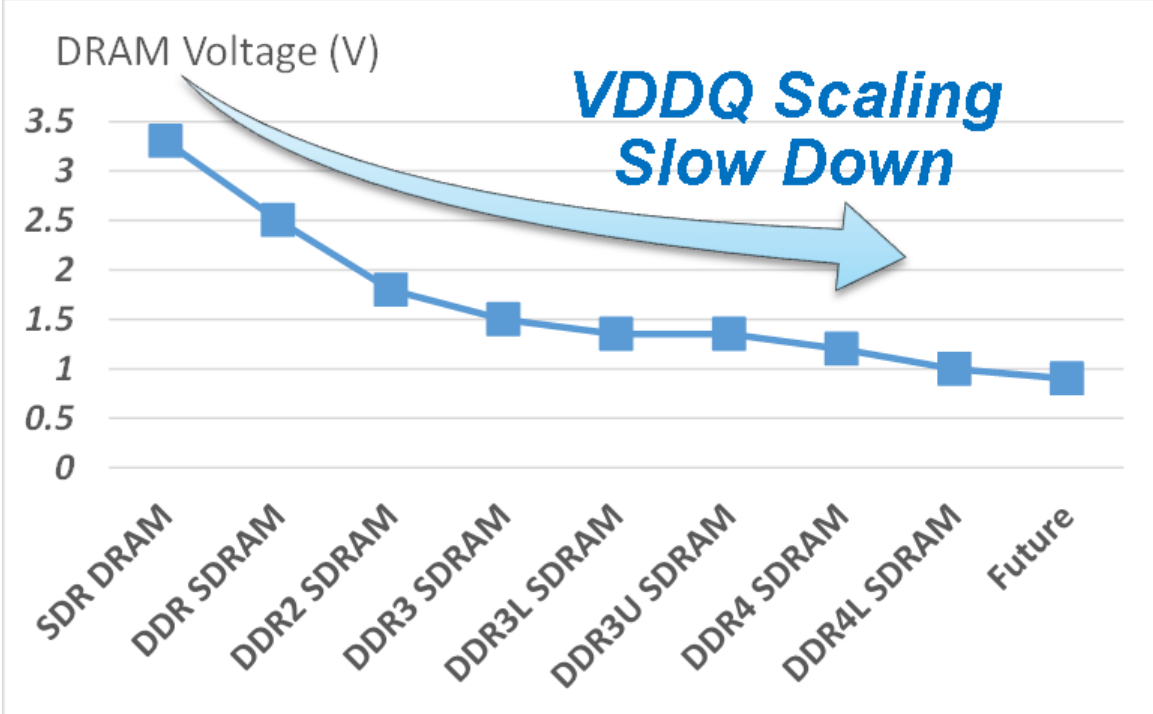


Figure 5 DDR IO Voltage Supply Scaling Trend

DDR4 memory IO changes the IO standard from DDR3 Stub Series Termination Logic (SSTL) to Pseudo Open Drain Logic (PODL_12).[5] Figure 6 below shows the IO electrical swing between SSTL and PODL_12. Because the PODL_12 dissipates DC power only when it drives a logic low, the effective power dissipation using PODL_12 is lower than DDR3 SSTL.

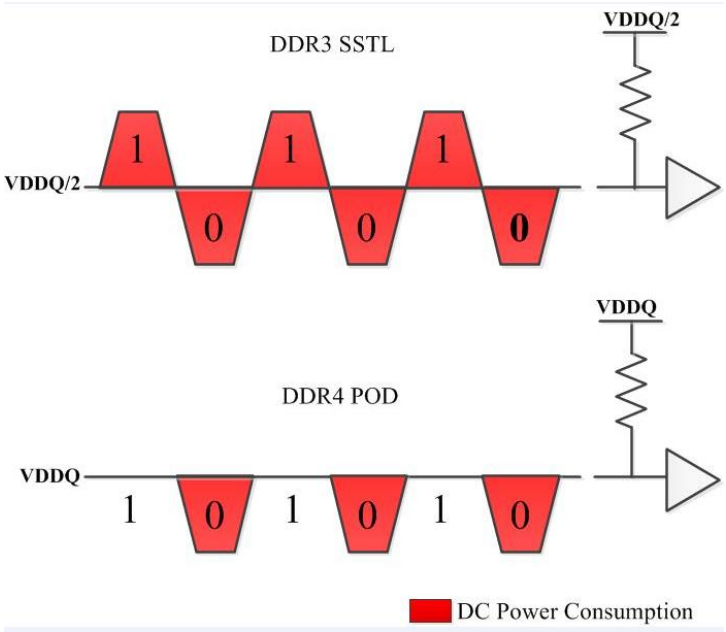


Figure 6 DDR3 SSTL and DDR4 POD Swing

An example of the relative power distribution in a DDR4 DRAM, with 70% Read and 30% Write ratio, was simulated in this study; the result is shown in Figure 7. The power analysis was based on a system that was without DBI enabled and using a similar power calculator software as listed in [6] . Comparing the power distribution, the background power is 21%, the activate power is 17% and the Read/Write/Termination power is 62% of the total power. The IO power is the dominated majority.

DBI in DDR4 will further improve the total system power and the concept is explained in the following section.

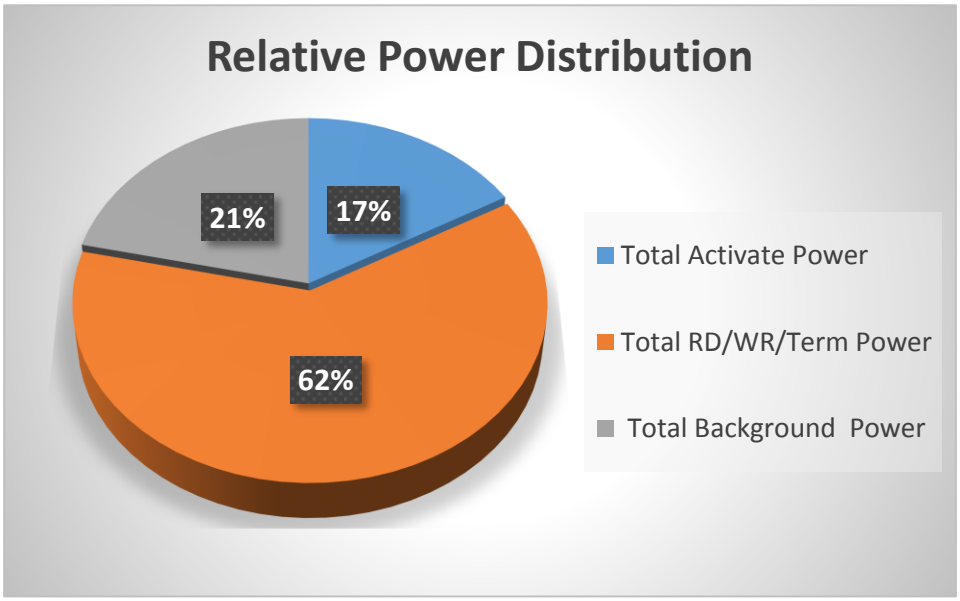


Figure 7 DDR4 DRAM Unit Relative Power Distribution (No DBI)

3. Data Bus Inversion (DBI) in DDR4 Interface

The data bus inversion scheme in DDR4 is classified as the DC DBI.[7] The DBI function can be expressed by the equation below.

$$\boldsymbol{DBI\#[k] = func(DQ(7:0)[k])} \qquad \forall \quad \boldsymbol{k} \in \mathbb{N}_{\mathbf{0}} \qquad \textbf{(1)}$$

with

$$func(\eta) = sum_{logic_{low}}(\eta) > 4$$

where $sum_{logic_{low}}(\eta)$ is the sum of number of data bits that have logic low and the data bits, DQ(7:0), are the data bits in the controller before any inversion encoding.

The DBI signal is an I/O signal sharing the same physical package ball input with data mask (DM) and one of the terminated DQS (TDQS) in the x8 and x16 DRAM device. Therefore, the DBI function is mutually exclusive from DM and TDQS mode. This bit is an additional physical bit for an 8 bit wide DRAM unit and two additional physical bits for a 16 bit wide DRAM unit. DDR4 allows DBI to be enabled in Write direction, Read direction, and in both Write & Read directions.

The content of the DBI bit is a function of other parallel DQ values within a byte. Equation (1) represents this mathematical relation where the parameter *"k"* represents the bit time location along the burst length. Figure 8 shows the data bit inversion values along the burst length graphically. When the total number of data bits that are driven to logic "LOW" are less than or equal to 4, the DBI# signal will remain logic "HIGH" (De-asserted). Otherwise, when the total number of bits that are driven to logic "LOW" is greater than 4, the DBI# at the time location *"k"* will be driven to "LOW", logically asserting the DBI # function.

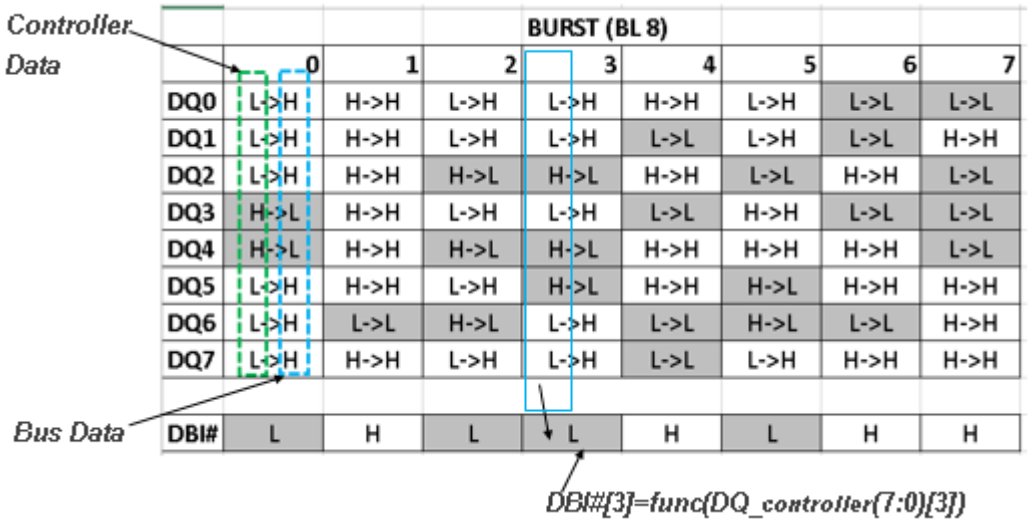


Figure 8 DBI # Functional View along Burst Length

In PODL_12 electrical standard, the IO driver only dissipates DC power when the data signal is driven logic "LOW". With data bit inversion mode is enabled, the content among all the DQ bit at location *"k"* are inverted opportunistically, resulting a fewer IO drivers that will drive LOW and hence reducing the IO power. In addition, because there are fewer bits that are switching, the interface will have lower IO supply noise.

The power saving provided by enabling DBI varies with different memory IO workload. The next section compares the total system power differences with different memory IO workload programs. In order to capitalize the benefit of using DBI feature in DDR4, the interface must be calibrated and trained together with DBI being enabled. Section 5 discusses the reason and the calibration concept with DBI. For the IO noise improvements, the data eye margin is measured and compared using the internal data eye shmoozing capability. Section 6 will show the results.

4. DDR4 DBI Power Improvement Comparison

The total system power improvement using DBI in DDR4 will vary with different Read & Write work load percentage. To compare the difference in power improvement, 11 different test cases which have different Read and Write percentage ratio were analyzed.

The Read and Write Memory IO ratio were controlled by mean of a data traffic generator implemented in the UltraScale FPGA fabric. The Read & Write percentage of these test cases are listed in Figure 9 below, the blue bars showing the memory IO Read percentage and the orange bars showing the Write percentage. These ratio cover a more typical memory usage, such as 70% of Read and 30% of Write, to the two other sides of the spectrum with higher ratio of Read and higher ratio of Write percentage.

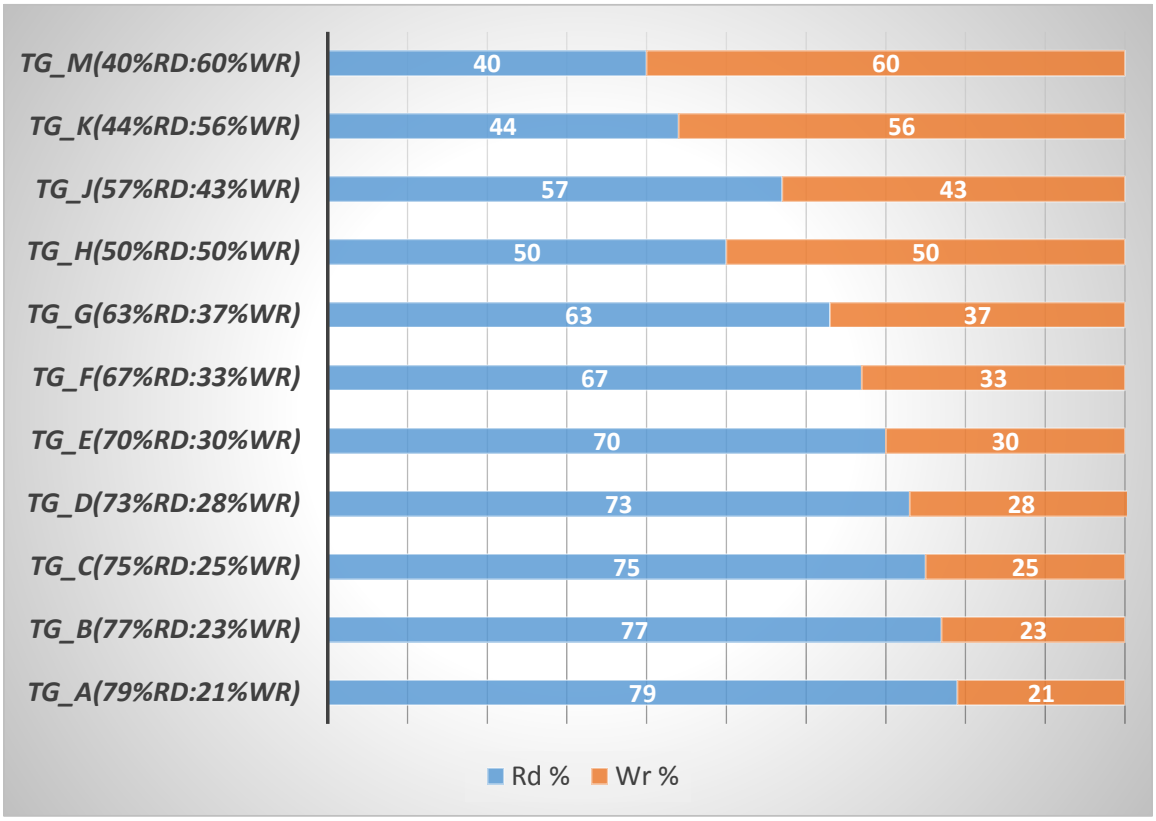


Figure 9 Memory Usage Read and Write Percentage Comparison among Different Test Programs

The system power analysis is based on a 72 bit wide DDR4 memory interface running at 2667MTs. The interface is optimized to maximize the Write and Read data eye margin as presented in previous report. [8]

To quantify the improvement using DBI among the different test cases, the system power was simulated and analyzed with each of the 11 test cases. Then the total system powers without DBI being enabled form the base of the normalization. The relative power

improvement when DBI is enabled is reference to this normalization base for each different Read and Write ratio test cases. The left side axis in Figure 10 shows the relative system power. The blue bars correspond to the system without turning on the DBI and the orange bars correspond to the one with DBI enabled. The right side axis indicates the percentage of the improvement.

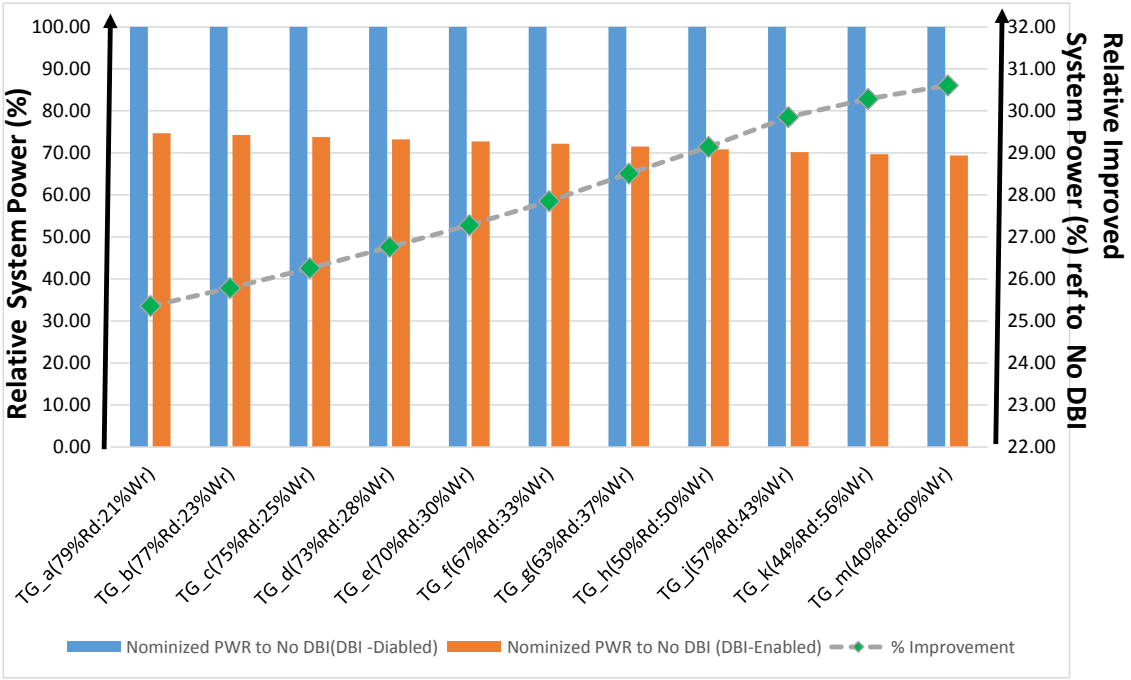


Figure 10 Relative Power Improvement Comparing across 11 different Usage Programs

The analysis comparison reveals that the power saving by enabling DBI varies from different workloads. A typical memory usage of 70% Read to 30% Write yields about 27% system power saving when DBI is enabled. The power saving increases when the memory workload leans toward a heavier Write percentage usage.

To capitalize the power saving using DBI, the DBI bit must be calibrated together with the regular DQ bits. The next section discusses the rationale behind the DQ IO bit training calibration with DBI mode enabled.

5. DQ Training and Calibration with DBI

The DBI signal ball at DRAM unit is shared among Data Mask & TDQS pin. Figure 11 is an example of a FPGA unit connecting to the DRAMs. The DRAM is an x16 unit, there is an upper DBI (U_DBI#) and a lower DBI (L_DBI#) bit for each unit. Because FPGA unit is a programmable device, the DBI at the FPGA can vary in different usage. The memory controller must ensure the DBI physical pin is optimized together with the other DQ bits with randomized calibration pattern.

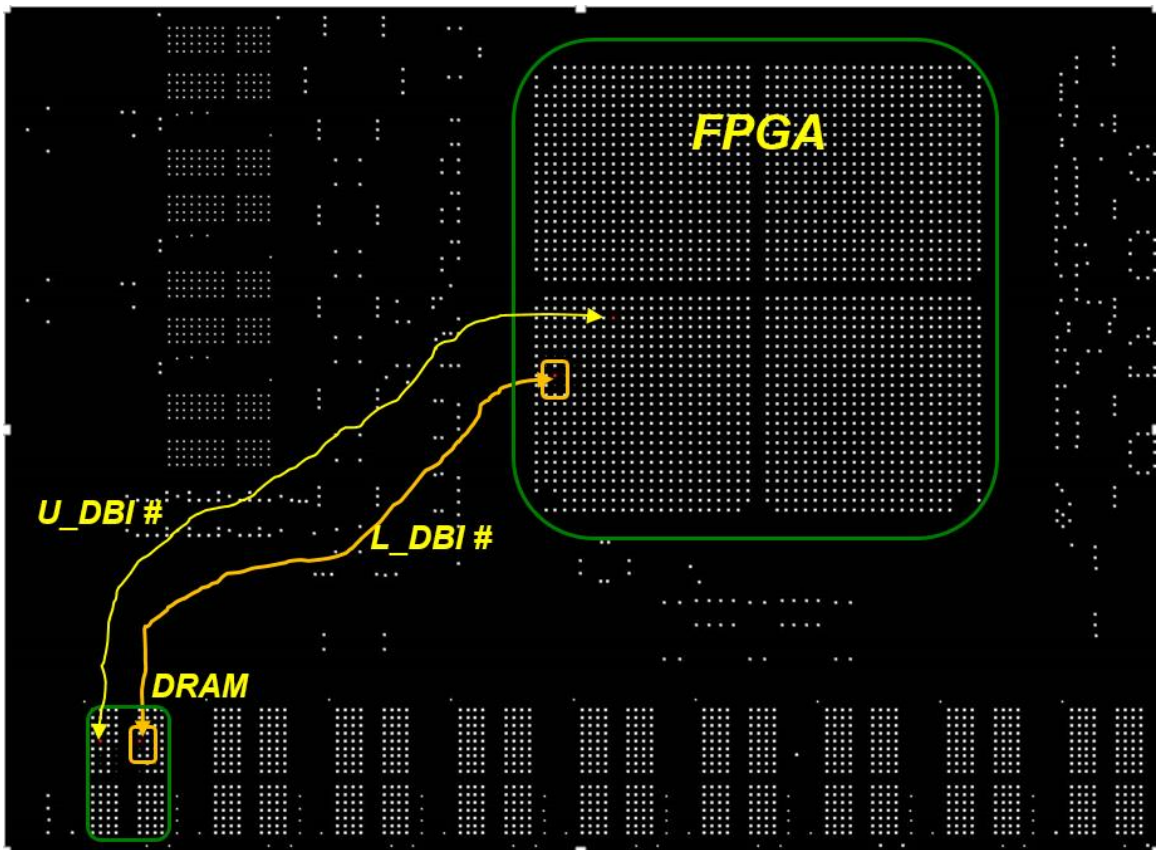


Figure 11 DBI Locations and Possible FPGA Pin Connections

Mathematically, the DBI bit stream can be expressed as:-

$$\boldsymbol{DBI\#\,[k] = func(rand\_cal(DQ(7{:}0))[k])} \qquad \forall \quad \boldsymbol{k \in \mathbb{N}_0} \qquad \textbf{(2)}$$

where "rand_cal" represents a defined randomized calibration pattern applied to all regular DQ[7:0].

When the DQ and DBI# bits toggle with randomized calibration data pattern, the data strobe (DQS) can be position to search for the optimal center position. As illustrated in Figure 12, the DQS will march towards the edge of the DQ & DBI# bits until the controller detects a data failure on the right hand side, this is the right margin. Then the relative delay between the data strobe and data will be reversed to search for the left margin. When the minimum right and left margin among the 8 bits of DQ together with the DBI bit are determined, the optimal center is the middle point of the margins.

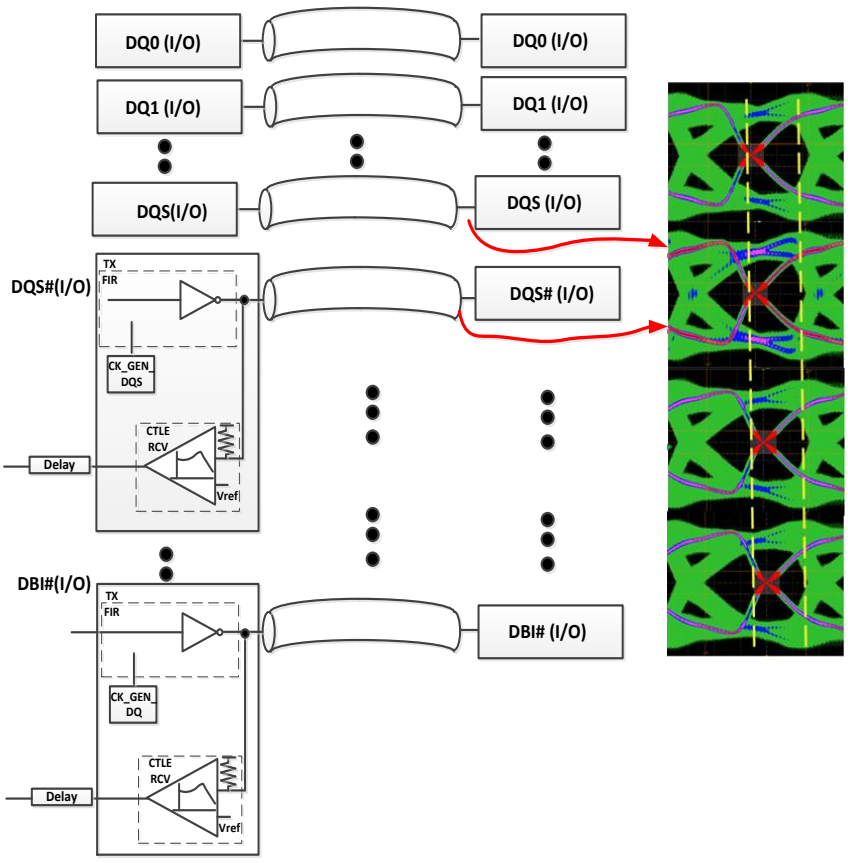


Figure 12 Data Strobe Center Positioning with Randomized Calibration Pattern

6. Power Noise Improvement with DBI Enabled

By enabling the DBI function, there are fewer IO toggle and it will provide a lower IO noise that appears on the data bus channel. The Power Deliver Noise (PDN) difference when using DBI function is quantified by using the step current method presented before in [9]. Figure 13 shows the general set up of a system-level PDN model.

When the system is without DBI enabled, the average step current of the DDR interface using a PRBS23 data pattern was captured and this step current was sourced into the PDN of the memory system. When the DBI is enabled, the average step current will be reduced. Hence, the power deliver noise will correspondingly be reduced. Figure 14 shows the voltage droop responses. The blue line shows the voltage response to the step current with no DBI enabled; the red line represents the response that has the DBI enabled. The 1st droop voltage improvement was about 38%.

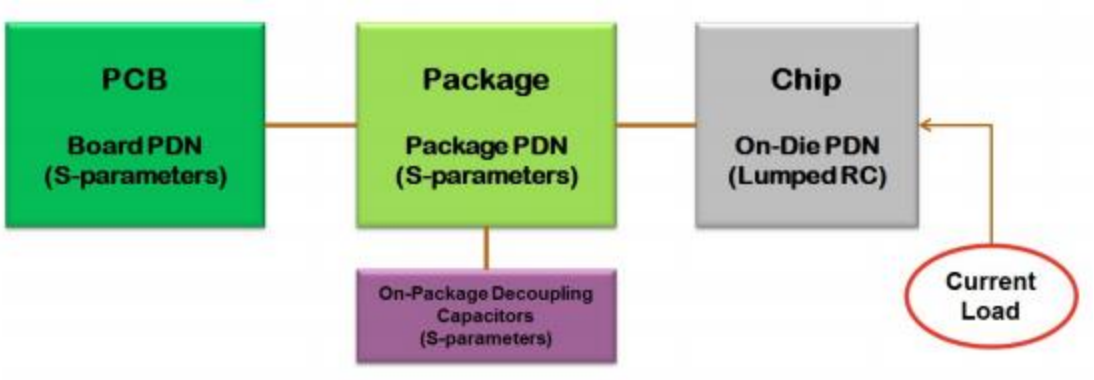


Figure 13 General Setup of a System Level PDN model

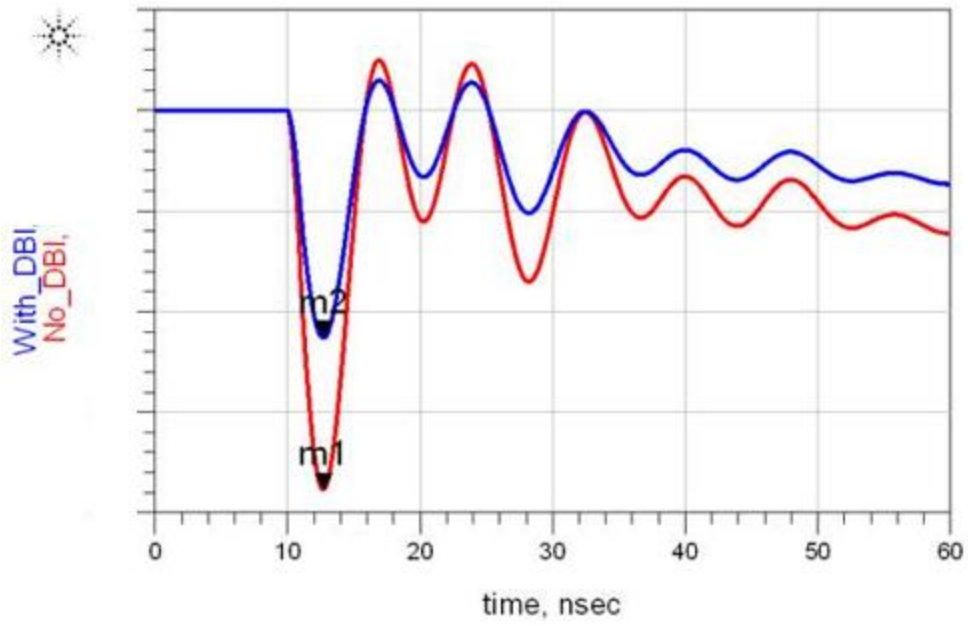


Figure 14 Voltage Droop Comparison between DBI enabled and DBI disabled

The power noise to signal jitter impact can be related by the following equation:

$$\int \left(i \times z_{PDN}(f) \right) df \times \textbf{Jitter Sensitivity} \rightarrow \textbf{Jitter} \qquad \textbf{(3)}$$

The integral sum of the product of the system switching current, *i*, and the power delivery network impedance, Z_{PDN} is the voltage noise that the system experiences. The output jitter is this voltage noise times the jitter sensitivity factor [9].

A typical power delivery network impedance plot and the phase noise plot are shown in Figure 15. The PDN impedance response varies with frequency and the circuit jitter sensitivity poses a low-pass filter characteristics. The product of these terms creates the output jitter which is a function of frequency as well.

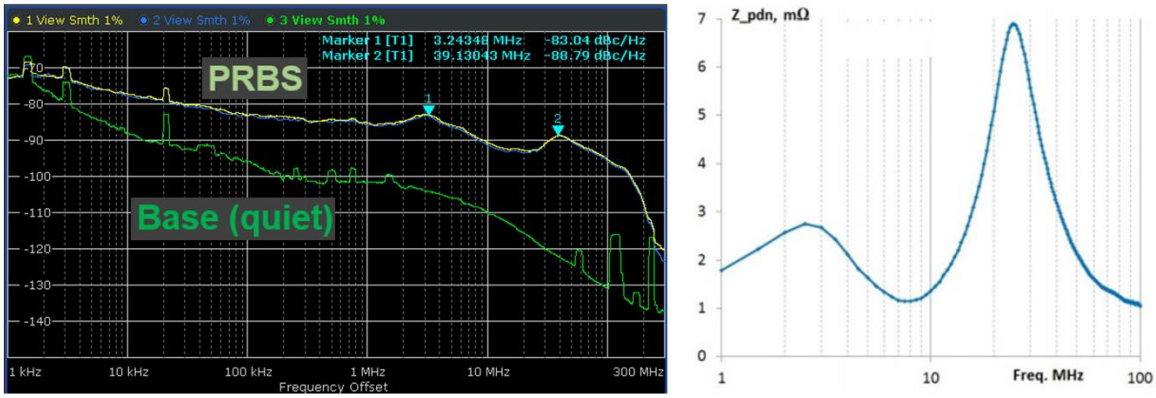


Figure 15 Phase Noise plot of a typical System and its corresponding PDN impedance profile

This noise reduction will positively affect the data eye margins which will be covered in the next section.

7. Experimental Data Validations & Results

Figure 16 is a 72 bit DQ wide interface with the capability to enable DBI. The memory interface has a work load of 50%:50% Read-to-Write ratio.

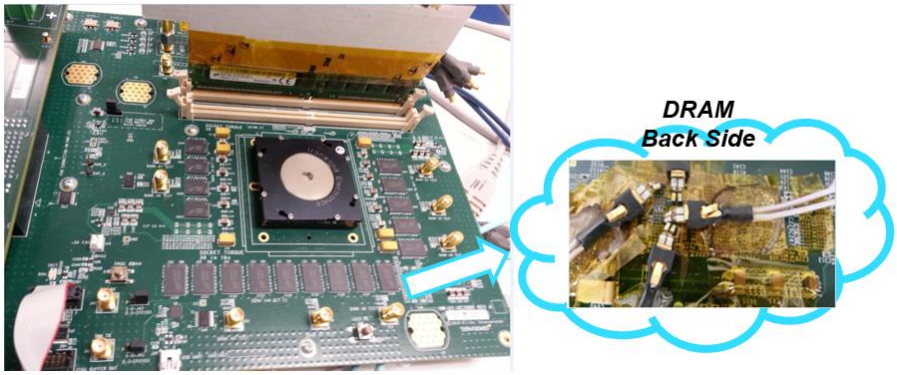
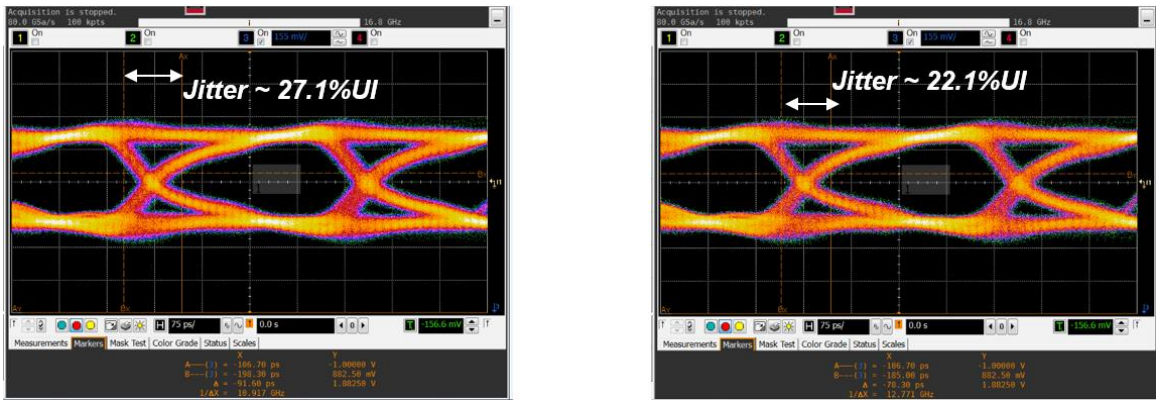


Figure 16 System Memory Validation System Board & Back Side Probing

The first validation was by direct probing. The probes were attached on the back side of one of the DRAM as shown in the insert of Figure 16. The system was running at 2930MTs. The Write direction DQ eyes, with and without enabling the DBI function, are shown in Figure 17.



(a) (b)

Figure 17 Write Data Eye (a) without DBI (b) with DBI Enabled

The data eyes were measured at the ball of the DRAM package and the DQ jitter was reduced from 27.1% of a UI to 22.1% of a UI when enabling the DBI feature.

The external measurements serves as an indication of the improvement. The actual jitter improvement was characterized by the Write and Read eye shmoo.

Figure 18 is the 2D Data eye shmoo set up. The DQ and DQS relative delay position was scanned from the center optimal point to the edges. In doing so, the pass & fail region could be identified for each selected reference voltage for the receiver. For all the adjustable reference voltages, the pass & fail regions were then created. Combining these measurements, a 2 dimensional eye shmoo was formed. The Read and Write directions eye shmoo were created using a similar method [8].

The eye shmoo characterization method provides a more actual measure of the quality of the Read and Write margins because they are a true measure of the system pass and fail condition.

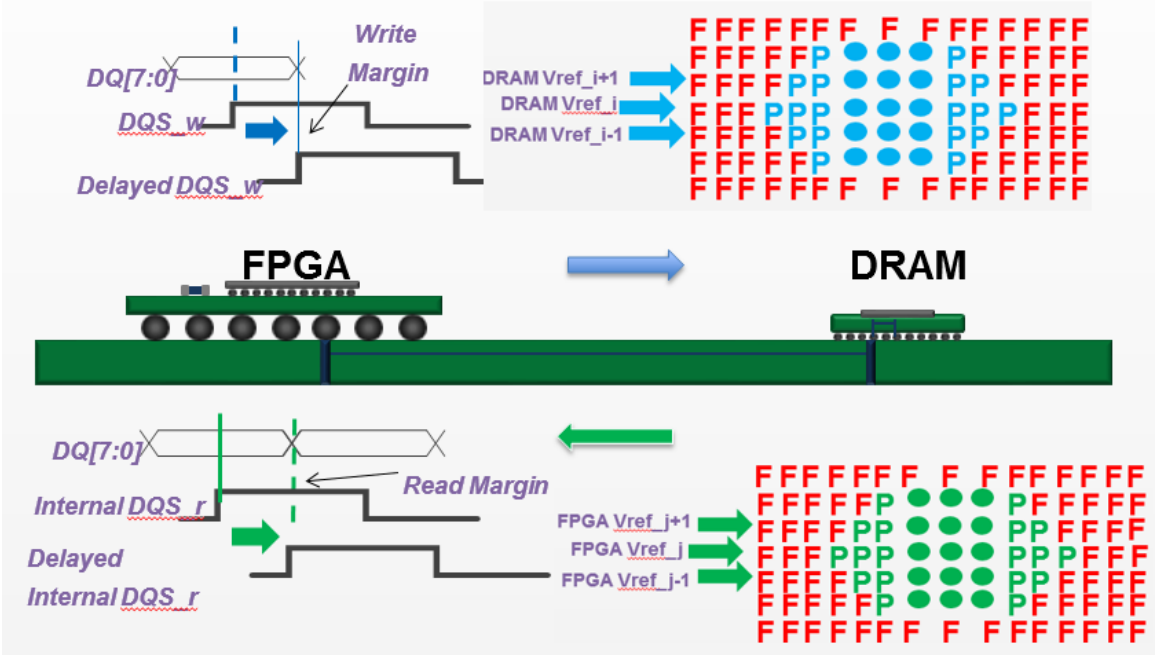


Figure 18 Write and Read 2 Dimension Data Eye Shmoo

Figure 19 shows the eye shmoo with and without DBI comparison measurement results in both Write and Read directions. The comparisons were done at the optimal input reference voltage settings in the corresponding Read and Write directions. The grey regions in the figures indicate the eye shmoo when the system does not enable DBI, the outer blue region in the horizontal bars show the improved eye widths. In the Write direction, the eye margin improved by about 7% with respect to no DBI being enabled. In the Read direction, the eye margin improvement is about 11% compared to the case without DBI. The difference in eye width improvement in Read and Write direction indicates that the relative power noise improvement in FPGA (Write) is different from the DRAM (Read).

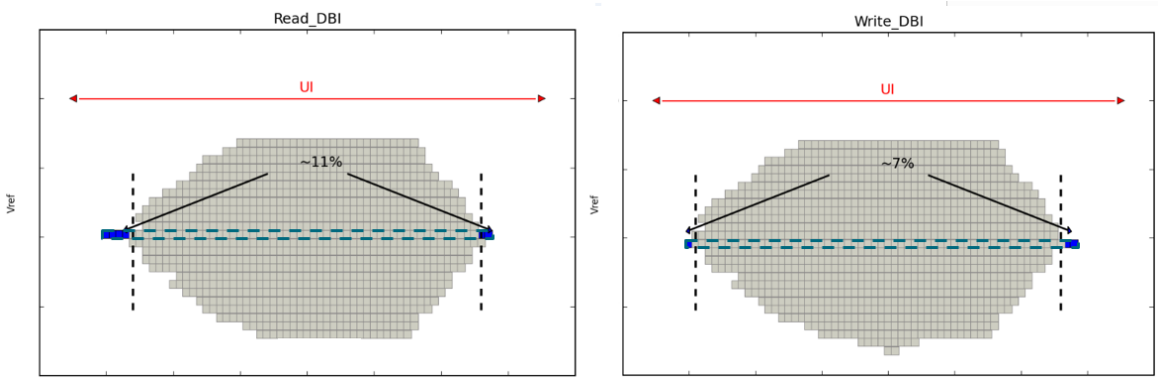


Figure 19 DDR4 System Channel Read & Write Eye Shmoo

8. Summary and Conclusions

The data bus inversion function in DDR4 provides system power improvement. This report shows the amount of power improvement depends on the work load conditions. The relative power improvement in the Write direction is more than the relative power improvement in the Read direction. Even though the exact power improvement depends on the implementation and it can vary, most of the DDR controller PHY carries the heavy lifting of signal quality enhancement, such as utilizing transmitter equalization [8], reducing the number of IO toggling in the Write direction in such a controller will provide more power saving relative to Reading from DDR4 DRAM. Therefore, it creates a larger system power improvement.

In the Power Delivery Noise with and without DBI being enabled comparison section, the corresponding step current reduction reduces the voltage droop response. The power noise improvement translates to lower data eye jitter. This report also discusses the importance of training the IO interface together with the DBI bit being enabled in order to capture the benefits.

An example of the DBI system jitter improvement was validated by measuring the Write data jitter and also by means of 2D data eye shmoo in both Write and Read directions. The validation example also indicates a difference data jitter improvement amount in Write and Read directions.

References

[1] "Top 500 Supercomputing Sites," 2015. [Online]. Available: www.top500.org.

[2] K. Malladi, F. Nothaft, K. Periyathambi, B. Lee, C. Kozyrakis and M. Horowitz, "Towards energy-proportional datacenter memory with mobile DRAM," *Computer Architecture (ISCA), 2012 39th Annual International Symposium on*, pp. 37-48, 2012.

[3] S. Patel, "Server Memory Power Trends„" in *JEDEC Server Memory Forum*, Shenzhen, 2012.

[4] E. Kim, "Application Specific Memory for a Server Centric World," in *Server Memory Forum*, Shenzhen, 2012.

[5] JEDEC, "POD12-1.2V Pseudo Open Drain Interface, JESD8-24," 2011.

[6] Micron, "DDR4 SDRAM System Power Calculator," Micron Technology, Inc., 2015. [Online]. Available: http://www.micron.com/support/power-calc.

[7] JEDEC, "DDR4 SDRAM, JESD79-4," 2012.

[8] Hing Y (Thomas) To et.al. , "UltraScale FPGA DDR4 2400Mbps System Level Design Optimization and Validation, DesignCon 2015," Santa Clara, 2015.

[9] Dmitry Klokotov, Jin Shin, Yong Wang, "Distributed Modeling and Characterization of On-Chip/System Level PDN and Jitter Impact , DesignCon 2104," Santa Clara, 2014.