DDR3/4 Combo IO Library Design in 28NM CMOS Technology

Syed Arsalan Jawed[1], Yasir Siddiqi1, Naveed[1], Sohaib Afridi[2], Muhammad Arslan[3], Khubaib Khan[1], Zain Warsi[1]

Received: 25-February-2018 / Revised: 26-April-2018 / Accepted: 17-April-2018

ABSTRACT

This paper presents a DDR3 and DDR4 dual mode combo IO library containing half duplex combo transceiver, a separate calibration cell for ZQ calibration and Power/GND cells with RC clamp for ESD protection in 28nm CMOS technology. The presented transceiver achieves 2166MT/s for DDR3 at 1.5V for a two dual rank UDIMMs/channel and 2400MT/s for DDR4 at 1.2V for a single dual rank UDIMM. The embedded break before make functionality in transmitter demonstrates a 17% reduction in crow-bar current at 2166MT/s. The slew rate control functionality demonstrates feasibility in reducing reflection on an under-damped channel as well as reducing simultaneous-switching-output noise. The ESD protection is targeted for 2KV-HBM, 8KV-MM & 5A CDM. The transmitter consumes total power of 7.5mW/Gbps@1.5V and 5.8mW/Gbps@1.2V with a break-before-make non-overlap delay of 60ps. Each cell has a footprint of 30µmx200µ and results in a continuously abutted structure when placed side by side.

KEYWORDS

SSTL, HSTL, POD, DDR3, DDR4, combo IO, multi-mode PHY, ZQ Calibration, Self-Biased Receiver, DIMM modeling, 28nm CMOS.

I. Introduction

Increasing memory bandwidth requirements dictated by applications such as cloud computing and big-data necessitate for higher speed, low-cost and low-power memory systems. With increasing speeds, communication with DIMMs, connectors and motherboard suffers from frequency attenuation, reflection and crosstalk. Therefore, though the capacity per memory module has increased the reduction in memory slots per memory channel, due to signal integrity issues, limits per channel capacity [1].Power efficiency of the link becomes a critical issue at higher data rates and heavily loaded channels [1]. The constant shift to lower voltage signaling and different schemes (such as POD) shows

[1]*Karachi Institute of Economics and Technology, Karachi.*
[2]*Tyndell Institute, Ireland.*
[3]*American University of Sharjah, UAE.*

a constant drive to achieve low energy consumption per channel. This scenario presents a multifaceted issue, which needs to be addressed by bringing in architectural changes, from the physical layer propagating through the link layer to the application layer. Some of the evolutionary changes are: (1) shift from balanced tree clock-routing topologies in DDR2 to fly-by configuration in DDR3 which necessitated for timing circuits (de-skew DLLs) in PHYs; (2) introduction of On-Chip-Termination (OCT), which necessitated for a calibration mechanism; (3) Finally evolving towards P2P interfaces where channel complexities are driving parallelism and dedicated memory controller [2].

Currently DRAM interfaces incorporate a parallel single Ended signaling due to pin-out restrictions and backward compatibility [1-3]. So far the efforts to increase data rate of single ended signaling focus either on physical design and optimization, cross-talk cancellation techniques, noise reduction through encoding and supply insensitive design as well as decaps on bandwidth enhancement through equalization [4]. GDDR5, for instance, has achieved beyond 5Gbps using Data-Bus-Inversion (DBI), data training and data equalization. This, however, is leading the system to a fundamental limit to the performance set by cross talk, which eventually results in increased system cost due to additional components such as cross-talk equalizers [5] or other bandwidth improvement techniques [8-9] or because of increased inter-wire distances and shielding.

Techniques such as parallel channels or fully buffered DIMMs are used to fulfill speed and capacity demands but again at the expense of system cost [1]. In due course, differential signaling with high speed serial IO enhancements is expected to potentially take over IO performances scaling post-DDR4. Another trend which is being exercised currently is to employ multi-mode transceivers in a PHY [7] to support multiple signaling standards using a single PHY. DDR4 and post-DDR4 future trends are moving towards a point to point configuration in which driver and receiver will have one-to-one interconnection.

This paper presents a combo DDR3/DDR4 transceiver as part of an IO library containing power, ground and Vref IO cells. Figure 1 shows the IO library composition. The half-duplex transceiver IO cell contains combo transmitter and receiver with OCT and ESD protection. The calibrator IO cell provides ZQ calibration of the OCT along with a high-sensitivity comparator for impedance matching. Power cells in VDDIO and VDDcore domain contain RC power clamps while GND IOs containing extra ground rail connections and decaps for both VDD domains. The following text is organized in four sections. Section 2 presents design description of the IOs. Section 3 presents the physical design. Section 4 presents post-layout simulations of the transceiver with a model of memory channel from controller to the DIMM. Section 5 concludes the paper.

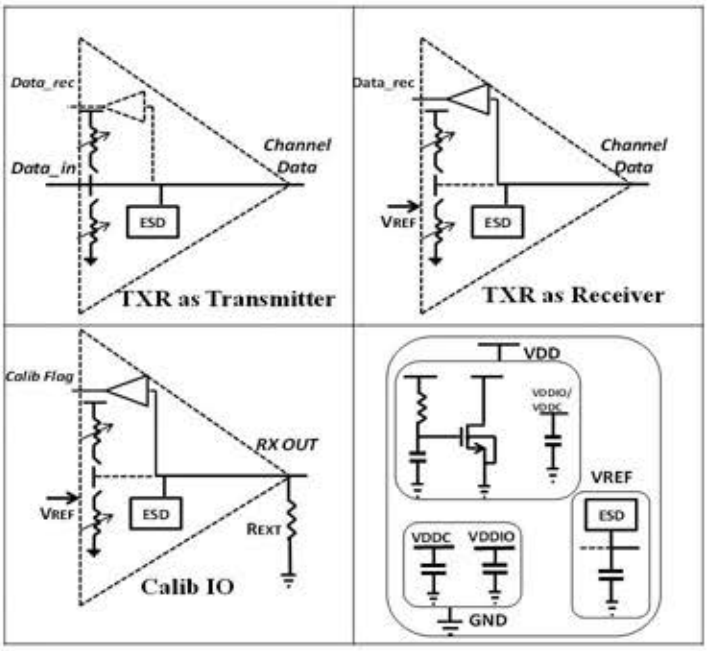

Figure 1: IO Library Composition

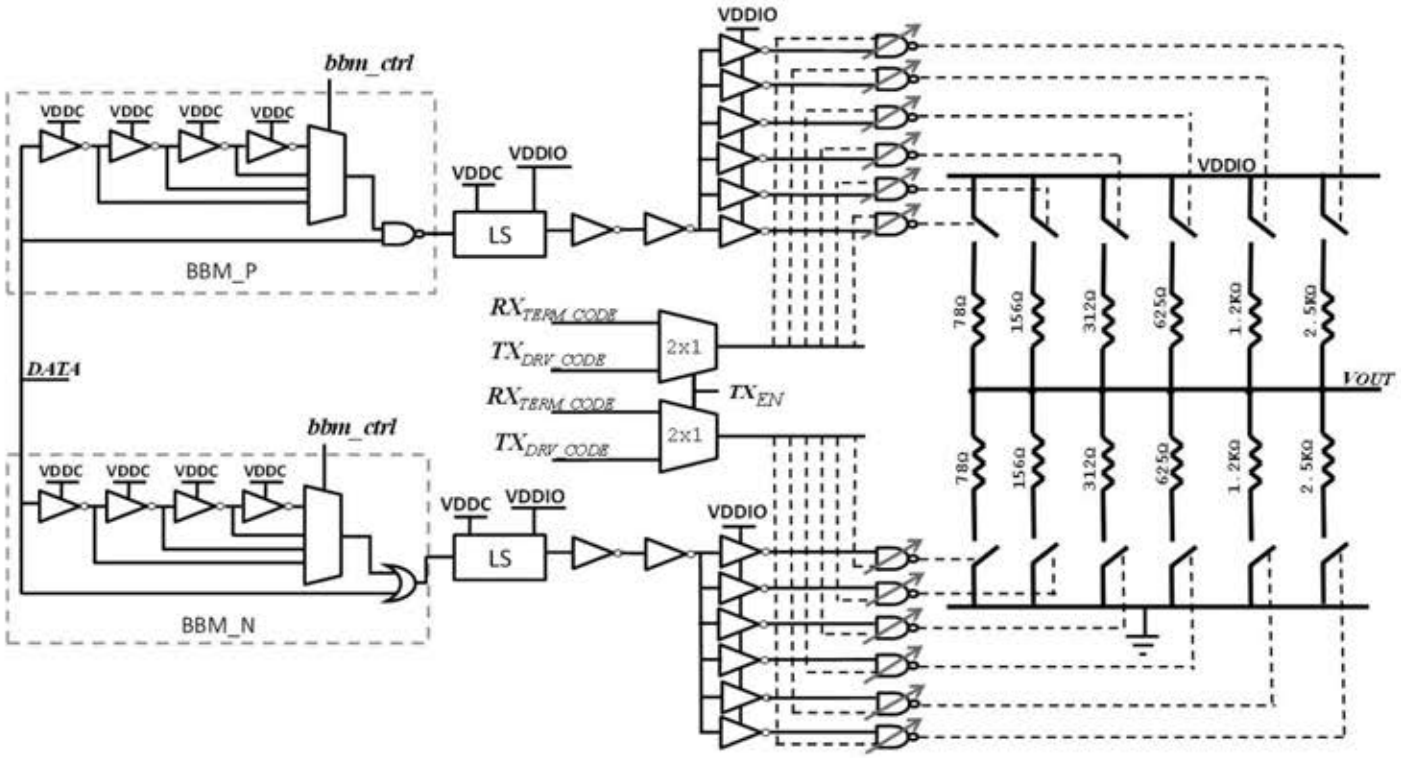

Figure 2: Transmitter Schematic

II. DESIGN DESCRIPTION

II.A. Transmitter

DDR3 and DDR4 operation requires a voltage-mode impedance matched transmitter [10-11]. The transmitter must have configurable impedance which can be calibrated to cater for PVT variations through the ZQ calibration process. The transmitter must also exhibit a linear and constant impedance during the transitions to minimize reflections.

The transmitter is designed using two identical blocks of binary weighted resistive pull-up and pull-down elements, as shown in figure 2. Each block has six binary weighted resistive legs with a conductance of 1X, 2X, 4X, 8X, 16X and 32X and are designed to provide an effective worst-case resistance down to 20 Ω, i.e. when all the legs in both blocks are operating in parallel. The unit resistance is 2.9kΩ. [2] proposes to use a single resistor with a pull-up and pull-down switch instead of two separate resistors for each leg. This, however, leads to accumulation of intrinsic capacitive load of both switches to a single node. The proposed transmitter here uses separate resistors to keep the capacitive load distributed by sacrificing a certain amount of area.

Table 1: Configurability of the Transmitter

TX_EN	Data	Remarks
0	X	TX acting as a termination depending on the provided rx_term_pu/pd code. Depending on DDR3 or DDR4 mode, termination to VDDQ/2 or VDDQ is achieved.
1	0	TX pull-down legs transmitting logic low with Rout specified by the tx_drv_pd code.
1	1	TX pull-up legs transmitting logic high with Rout specified by the tx_drv_pu code.

The resistive elements are connected to their respective power-rail through properly sized MOS switches. Each switch contributes less than 10% of the legs' resistance and therefore cannot cause more than 10% INL, resulting in a linear and controlled resistance even when the switch is transitioning between ON and OFF states. The switches are implemented using 1.5V IO devices-based transmission gates to achieve a consistent resistance for the whole voltage swing.

The transmitter has embedded break-before-make capability to reduce the crow-bar current. The BBM_dly_ctrl signal can be used to apply 20ps, 40ps. 60ps and 80ps of non-overlap duration between rise and fall transitions. Care must be taken to properly set BBM delays for both pull-down and pull-up legs to avoid distorting the output waveform. The passive multiplexer is implemented using transmission-gates to avoid excessive delay due to logic-depth in an active 4:1 multiplexer.

The pre-drivers for the pull-up/down legs are partitioned into separate inverters in VDDIO domain, as shown in figure 2, to reduce capacitive loading on the pre-drivers and also to facilitate the layout in subsequent design stages.

When the transceiver IO cell is in receive mode, the transmitter acts as a resistive termination for the receiver. The transmitter can be placed in termination-mode by de-asserting tx_en, which disconnects the break-before make switches from the core drivers. The transmitter driving impedance (the impedance codes (tx_drv_pu/tx_drv_pd) and the receiver termination (the termination codes (rx_term_pu/rx_term_pd)) can have different values. The transmitter receives both pull-up/pull-down TX impedance and RX termination codes from the core configuration registers and muxes these codes internally based on the tx_en signal, which signifies if IO is in transmit or receive mode. This mux is intentionally placed inside the IO to reduce the switching time between transmit and receive modes since only data_in_core_ddr and tx_en signals run at core speed while the others are static signals, which are set during the initial and periodic calibration procedures.

For DDR3 CTT termination, both pull-up and pull-down legs of the transmitter participate in termination, therefore, providing a termination to VDDQ/2. For DDR4, in the POD mode, only the pull-up legs are utilized as terminations and the pull-down legs are tri-stated, as shown in figure 3.

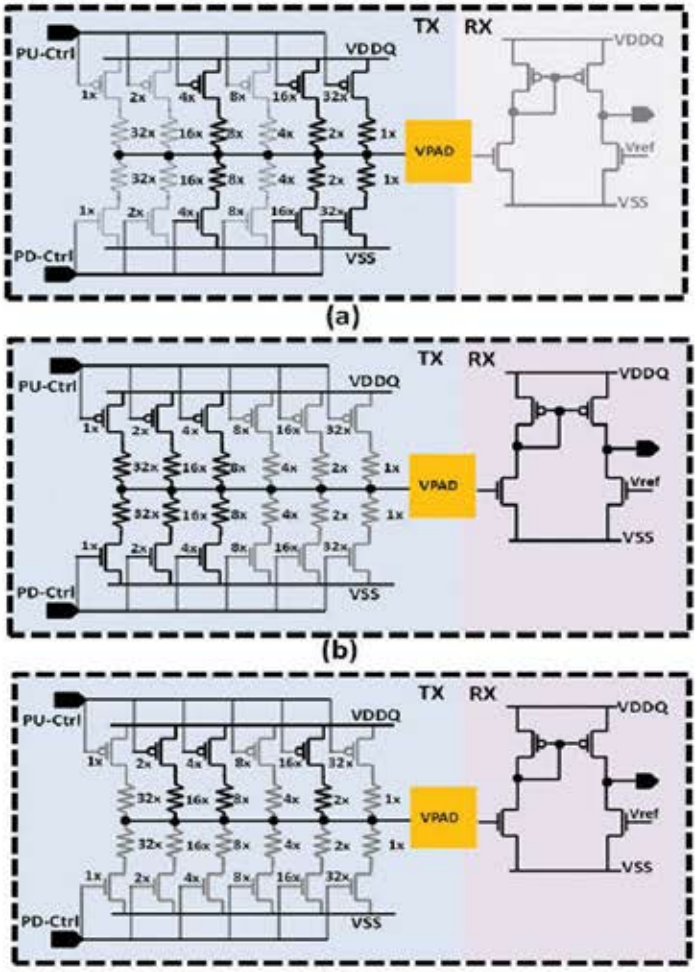

Figure 3: (a) Transmitter configured at 34Ω output impedance, (b) Transmitter configured as 120Ω center-tapped termination, (c) Transmitter configured as 80Ω pull-up termination

The transmitter has embedded slew-rate control, which is not only useful in minimizing reflections on an under-damped and under-loaded channel but is also very much needed to keep simultaneous switching noise down for higher signal:power:ground ratios at higher data rates. Multiple IOs asking for a larger impulsive current with a limited S:P:G ratio can result in jitter at the output of the transmitter. The slew-rate control is available through controllable drive strength

of the pre-driver, allowing the possibility of low-pass shaping the pre-driver's output, which in turn shapes the final output by slowing down the connection/ disconnection of the respective pull-up/down leg. A differential pad can be built by combining two IOs; one for true-phase and the other for complementary phase along with in-IO inversion to generation the complementary phase of the input.

II.B. Receiver

Due to the combo nature, the receiver must be able to accommodate a relatively large input common-mode. DDR3 signal is received at a common-mode of 0.9V due to the center-tapped-termination. DDR3L is received around 0.75V and DDR3U around 0.6V. For pseudo-open-drain reception in DDR4, the common-shifts up to 0.83V. The receiver expects a variable input swing depending upon the value of termination, DIMM/channel and the type of DIMM used. To reduce pin count, the receiver must also be self-biased.

The combo receiver is based on a self-biased differential amplifier as its first stage followed by a regenerative domain-shifter to VDDCORE and a digital buffer to drive the subsequent loads, as shown in figure 4.

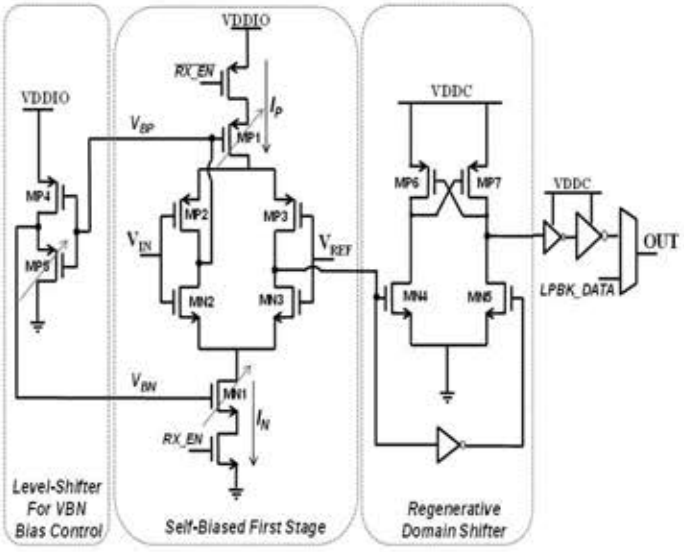

Figure 4: Self-Biased Two-stage Receiver

Table 2: Devices Functionality and Dimensions for the Self-Biased Receiver

Device(s)	Dimensions (W/L)	Device Function
MN1	90u/105n	NMOS Current Source
MP1	90u/105n	PMOS Current Source
MN2/MN3	100u/200n	NMOS Input Pair of the Stage1
MN4/MN5	10u/105n	NMOS Input pair of Stage2
MP2/MP3	100u/200n	Complementary PMOS input pair for Stage1
MP4/MP5	30u/105n	Source-follower based Level-Shifter
MP6/MP7	2u/50n	Cross-coupled positive-feedback ratioed-loads

VIN is connected to the metal stack carrying the signal while the other input to the externally provided VREF, which is the switching threshold for the receiver. Single-output of the differential amplifier is taken forward to the regenerative domain-shifter, the other complementary output is employed for self-biasing of the amplifier based on self-biased differential amplifier topology proposed in [13] with one main modification, as explained in the following text.

The self-biasing proposed in [12-13] of both PMOS and NMOS current source as shown in figure 4, requires connecting gates of MP1 and MN1 to the same voltage VBP. The 1.5V IO devices in the 28nm technology have a typical threshold voltage of around 600mV, which can go above 700mV for the slow-corner at -400C. Operating at a power supply of 1.5V, it becomes impossible to achieve an overdrive voltage of around 100mV for strong inversion for both PMOS and NMOS even if VBP is at VDDIO/2. The self-biased single-ended amplifier proposed by [12] stacks one less device as compared to [13], however, employing mirror-connected load reduces the bandwidth and at the same time reduces the input common-mode mode because it employs single input pair.

In one self-biasing scheme employed in the proposed receiver, the voltage VBP is taken slightly below VDD/2, which ensures proper biasing of the PMOS current source MP1. However, if we apply VBP directly to NMOS current source MN1, it would be biased in sub-threshold region. The amplifier employs a PMOS source-follower based level-shifter, which gives out an up-shifted version of VBP, i.e. VBN. This voltage VBN is applied to NMOS current source loads. This self-biases the receiver at a typical bias current of 300uA under typical conditions. A PMOS level-shifter also helps in keeping body-effect based non-linearity out of the equation. Consider the following circuit-analysis of the receiver:

$$I_P = I_N \quad (1)$$

$$\mu_P C_{ox}\frac{W_P}{2L_P}(V_{DD}-V_{BP}-V_{TP})^2 = \mu_N C_{ox}\frac{W_N}{2L_N}(V_{BN}-V_{TN})^2 \quad (2)$$

$$V_{BN} = V_{BP} + V_{offset} \quad (3)$$

Assuming $k_N=k_P$ and $V_{TN}=V_{TP}$ for simplicity,

$$k_N = \mu_N C_{ox}\frac{W_N}{2L_N} = \mu_P C_{ox}\frac{W_P}{2L_P} = k_P \quad (4)$$

$$V_{BP} = \frac{k_x}{2}, \text{ where } k_x = (V_{DD} - V_{TP}) \quad (5)$$

Eq.5 signifies if $K_N=K_P$ and $V_{TN}=V_{TP}$, then the remaining headroom after the PMOS threshold drop should be divided equally between NMOS and PMOS current sources. Mismatch in K_N and K_P can be employed to shift V_{BP} and therefore V_{BN} accordingly.

This self-biased circuit does not require a start-up circuit. Consider an initial condition where circuit is powered-down and each node is discharged down to zero. When power is turned on, i.e. $V_{BP}=0$, which means M_{P1} is on and M_{N1} is off. Assuming a CTT input, both input pairs are on. This pulls up the node and VBP goes high and M_{N1} turns on, eventually striking a balance between I_P and I_N.

The small-signal gain at Q-point is defined by $(g_{m1}+g_{m2})(r_{01}||r_{02})$ and is 20dB at 2GHz. It is important to note that input pair devices do not turn off completely in complementary large signal operation.

Receiver can be forced into power down mode by de-asserting the RX_EN signal. This is done by disconnecting the power-down switches, disabling the bias current sources. However, the receiver has a relatively long wake-up time; 20ns-30ns, due to the reason that bias generator of the receiver also gets disabled in power-down and re-biasing at the proper quiescent point takes the required time. Therefore, care must be taken and receiver must be disabled during long idle periods where the system gives enough notice to wake-up the receiver.

The first stage also embeds configurability in its bias control by using digitally configurable widths for PMOS and NMOS current bias sources. In case, the bias control loop settles to an awkward value, making it difficult for the subsequent domain-shifter to operate, these devices can used to calibrate the amplifier common-mode output.

The second stage domain-shifter is a ratioed-logic with input NMOS pair and PMOS cross-coupled regenerative load. This domain shifter also implements a bypass-able 50mV of hysteresis to minimize spurious glitches, while at the same time, the limited bandwidth of the first-stage also attenuates very high-frequency glitches already.
Receiver has internal and external loopback. The internal loopback works by returning the received data in VDDcore domain through the loopback multiplexer, as shown in figure 4. For external loopback, transmitter outputs the data on the PAD Stack and the receiver is enabled at the same time, which reads in the transmitted data.

II.C. ESD Protection of the Transceiver

The employed ESD protection is HBM-2kV, MM-8kV and CDM-5A. A two-stage diode-based ESD protection is embedded in the transceiver as shown in figure 5. The primary protection diodes (D1/D2), larger in size and capable of sinking larger currents, are also slow

to turn on especially for faster ESD events like CDM. Therefore, a secondary protection is also provided, which employs both GGMOS and and diode-strings, connected through a ballast resistor to the pad-stack. The secondary protection limits the initial current and partially clamps the voltage buildup at the receiver's gate below the snapback voltages of the MOSFETs [14-15].

The primary diodes effective on-resistance is 0.5Ω, which results in a 2.5V build-up across the diodes for a 5A-CDM. Considering routing resistance R_{ROUTE} from primary diodes to the power-clamp and then Ron of the clamp, the total affordable resistance is 1.2Ω, which results in 6V clamping leaving 2V margin away from the oxide breakdown of the IO devices.

The transmitter's transistors channel lengths and widths and layout are carried out in a way that they act as self-protecting devices. The transmitter could also serve as a limited protection device for the receiver, however, due to larger resistances in the transmitter legs, it cannot act as a CDM protection device. The total area of ESD protection is 1800μm² with primary protection consuming 80% of the area.

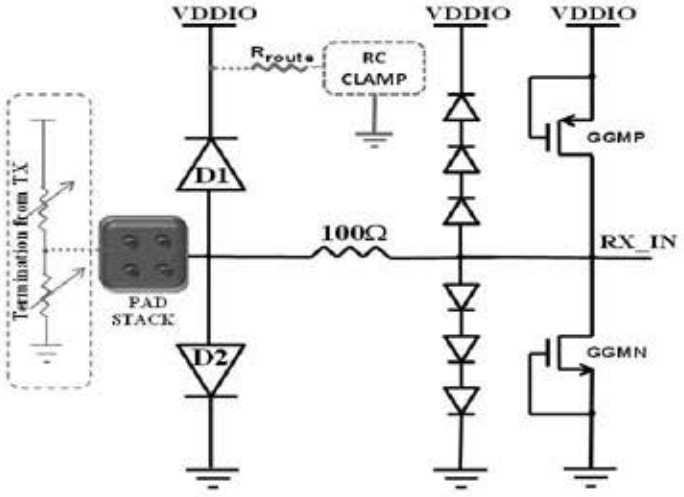

Figure 5: Two-Stage Diode Rail-Based ESD Protection for the Transceiver

II.D. The Calibrator IO

DDR3/4 IO operation requires an initial and periodic OCT impedance calibration to achieve an impedance-matched transceiver operation with the PCB traces acting as transmission lines. This calibration procedure is part of the DDR3/4 JEDEC standard PHY operation, known as ZQ-calibration. For this purpose, a separate calibrator pad, which does not take part in IO operation, is connected to a high-precision external resistance. A calibration algorithm FSM matches the calibrator IO's impedance with the external impedance. The impedance-matching code achieved through the calibration procedure for Calib IO is then copied to the actual IO pads which are residing in vicinity as part of either a DATA, ACC or CLK macro in a multi-bit PHY.

The calibrator IO is almost similar to a bidirectional transceiver IO with a slightly different receiver. The calibration requires higher gain or sensitivity, as it needs to track resistance changes in the range of 0.1Ω. However, a lower bandwidth can be afforded as this pad does not operate at the actual data-rates. Therefore, a two-stage amplifier topology is employed in the calibrator's receiver to achieve higher gain, as shown in the following figure 6.

The first stage is a self-biased differential stage based on Chappel-amplifier configuration with a level-shifter in the bias control loop [12]. The second stage is a single-ended common-source amplifying level-shifting stage for the following digital inverters. The open-loop gain variations due to PVT corners at 100MHz are between 40dB-68dB for both 1.2V and 1.5V supply voltages.

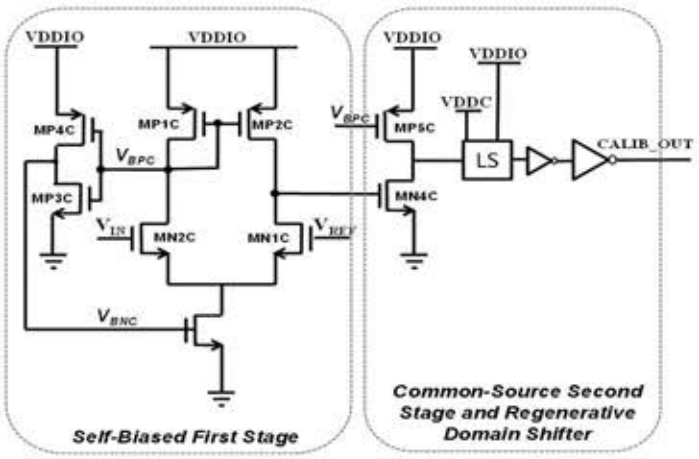

Figure 6: Self-Biased Three-Stage High-Gain Amplifier for the Calibration Operation

II.E. Power Cells and RC Clamp

The IO cells have rail-based double-diode ESD protection, which necessitates for a power clamp. The employed circuit is that of an RC power clamp, as shown in figure 7, which is placed in the power cells [14].

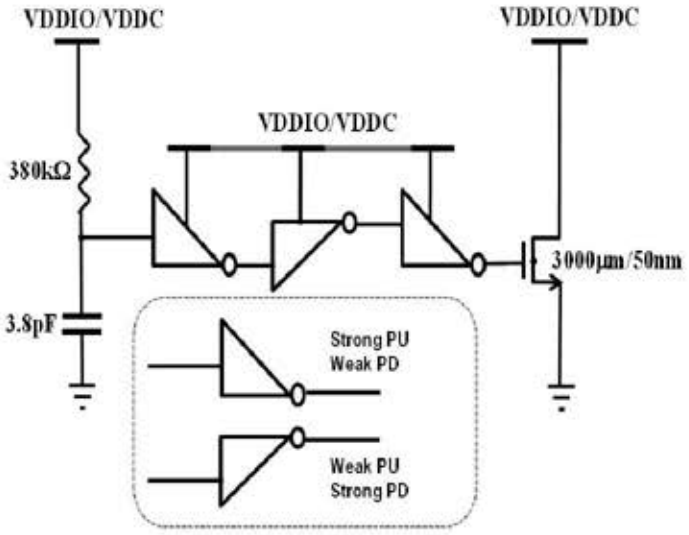

Figure 7: RC Clamp

The RC time-constant is around 1us. This is done to accommodate the fastest (CDM) and slowest (HBM) ESD event without mixing it with power-ramping, which is in milli-seconds range, therefore, keeping the ratioed-logic deactivated for the power-ramp. The drivers of the clamp are progressively sized to drive the large capacitive load. The three inverters topology also makes the design immune to spurious switching, slightly dampening the responsiveness of the clamp [14-15]. However, care must be taken in properly sizing these inverters not to deteriorate the responsiveness of the clamp especially for faster ESD events like CDM. The employed inverters are asymmetric since the detection of ESD works on a ratioed-logic.

III. LAYOUT CONSIDERATIONS

Figure 8 shows the overall floor plan of the bidirectional IO.

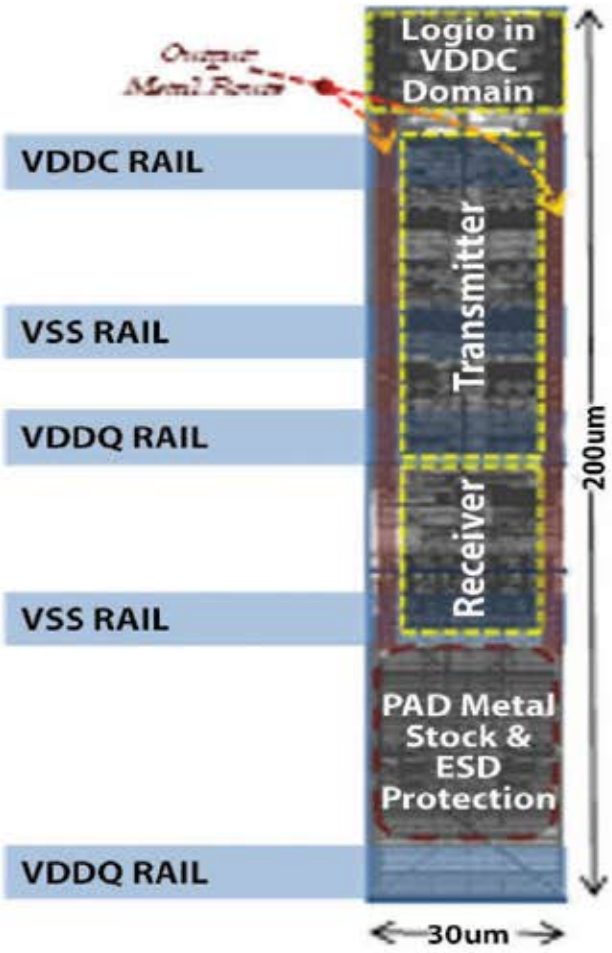

Figure 8: Floorplan of the Transceiver IO cell

The required footprint of the IO cells is to have a pitch of 30µm on x-axis to achieve a modular and scalable structure for a multibit PHY macro or slice [16]. This results in a rectangular shape which also favors signal flow in current-intensive and EM-prone conditions. The maximum dc current flow from an IO cell can be around 15mA when configured as 34Ω transmitter and 60Ω CTT at the far-end. The maximum current density before electro-migration is 1mA/µm. Therefore, to sustain 15mA, effective width of the output route should be above 15µm, as shown in figure 8.

The pads can be abutted by placing them side by side. The power and ground connections for each pad plug-in automatically to form a horizontal rail. All the pads have the same size to facilitate abutment. Passivation opening for the flip-chip

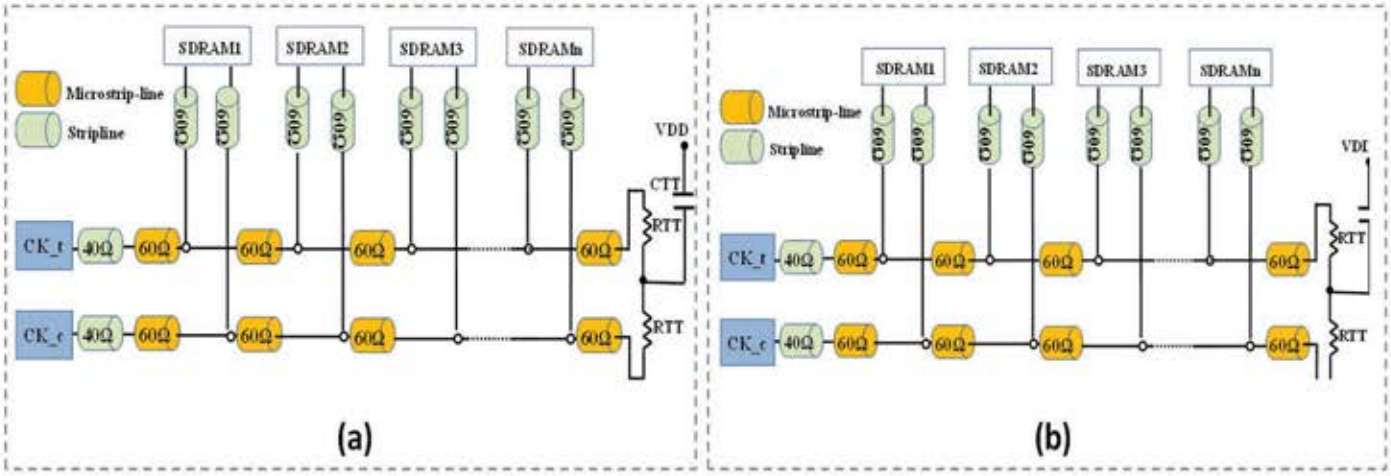

Figure 10 : Equivalent electrical model of the channel from IO to the SDRAM module, traversing through the PCB tracks

bond pad is over the ESD protection circuit, so that the ESD protection routing is minimized. Metal1-Metal6 1X metals are used for internal signal routing. The Power pads and clamps, however, utilize 2X Metal7 and Metal8 for vertical routing of the power connections. The 10X Metal9 is utilized for horizontal routing of the power lines. Horizontal power routings for IOs main circuits are kept separate from the ESD protection horizontal routing. The pad connection comes directly from the bump on M10. An abutted view of the IO cells is shown in figure 9.

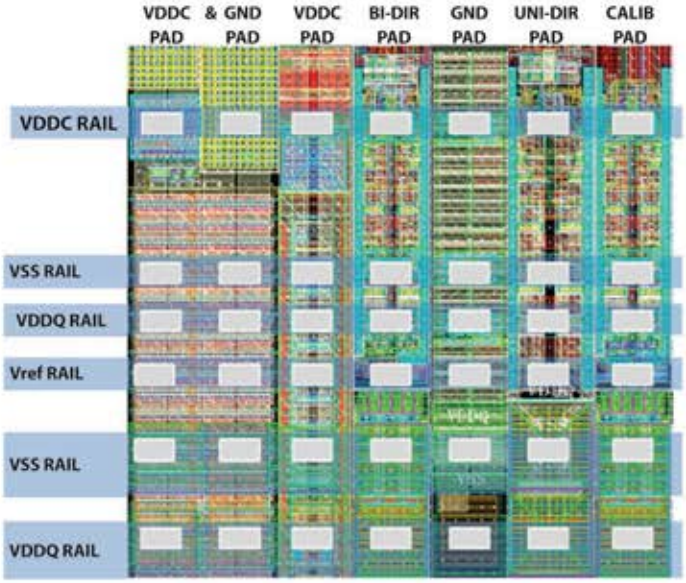

Figure 9 : Abutted view of the IO cells in a slice/ macro

The following table 3 lists down the dimensions for the IO pads:

Table 3: IO Dimensions

PAD Type	Dimensions	Remarks
Bidirectional Transceiver	200ux30u	
Unidirectional Transmitter	200ux30u	
Calibration IO	200ux30u	
VREF_IO	200ux30u	They may be implemented by separate IO cell which provides VREF routing rail and decaps or by using a disabled and ESD protected BIDIR_PAD/UNIDIR_PAD
VDDIO_PAD	200ux60u	VDDIO-VSS pair with RC clamp for IO devices
VDDC_PAD	200ux30u	VDD Core power pad with RC clamp for core devices
VSS_PAD	200ux30u	Ground pad with embedded decaps

IV. POST-LAYOUT SIMULATIONS

The extracted post-layout netlist is generated using Calibre xRC with back-annotation of transistor-level parasitic details tuned for a mixed-signal design.

To simulate IO cells, a complete model of the channel from PHY to the DIMM's memory module is needed. The PCB tracks at the targeted data rates behave like transmission lines. The complete channel, as shown in following figure X, comprises package of the PHY, break-out from PHY, track on the PCB, lead-in to the DIMM's individual memory module. This channel also

changes with the type of DIMM used, i.e. un-buffered DIMM, registered-DIMM etc., where UDIMM presents the worst case with respect to electrical loading.
The transmission lines of the channel are modeled using the field-solver Tline element in Cadence analogLib library. The length and impedances of the transmission line segments are extracted from [10-11, 17-22] and the width are accordingly set to achieve the required impedance. Similarly, the type of transmission line, i.e. either microstrip or stripline is also extracted from [10-11, 17-22]. A wideband lossy line model is employed to include skin-effect and di-electric loss as at the targeted frequencies these two effects change the channel's characteristics drastically.

DDR3 and DDR4 use fly-by topology instead of balanced Tree topology as was the case in DDR2 for clock, command and control signals. In fly by topology, the ACC signal is routed across the length of the DIMM with only termination at the far-end of the DIMM. While the data lines, are point-to-point. Therefore, it requires two complete different channels to be modelded, one for fly-by ACC signals and the other for point-to-point data signals. The channel for both cases comprises a set of transmission lines with different impedances and characteristics due to the way they are routed, either on the top-layer of the PCBs or completely shielded manner in some layer below the top layer. The channel exhibits di-electric leakage due to signal energy loss at higher frequencies through the dielectric of the PCBs. At the same time, the skin-effect also increases the resistance of the channel at higher frequencies. The length of the channel acts as an inductor, i.e. an increasing series resistor with increasing frequencies. The capacitive coupling of the conductor through the dielectric with other conductors or ground plane gives rise to a distributed capacitive load for the driver. From this perspective, the channel acts as a low-pass filter for the fast edged pulse generated by the transmitter, decreasing the edge-rate of the pulse received at the receiver. Therefore, there is a limit to the electrical load with which the transmitted signal is received with proper edge-rate for the receiver to toggle its state accordingly. Connecting more DIMMs on a particular channel increase this electrical load. On the other hand, reflections are also a major concern and it requires proper impedance matching on the driver as well as on the receiver side to minimize reflections.

The simulations are performed using Cadence-Spectre simulator. A random pattern generator was written in VerilogA. DDR3 and DDR4 use 1.5V SSTL signaling for ACC signals. DDR3 uses same 1.5V SSTL signaling for data signals as well, however, DDR4 uses 1.2V POD signaling for data signaling. The respective JEDEC standard defines the acceptable ac and dc voltage levels at the receiver for particular speed grades. The following figure shows the eye-patterns for DDR3 and DDR4 different DIMM configurations.

Table 4 shows the achieved and required ac/dc electrical parameters for the eye-patterns shown above.

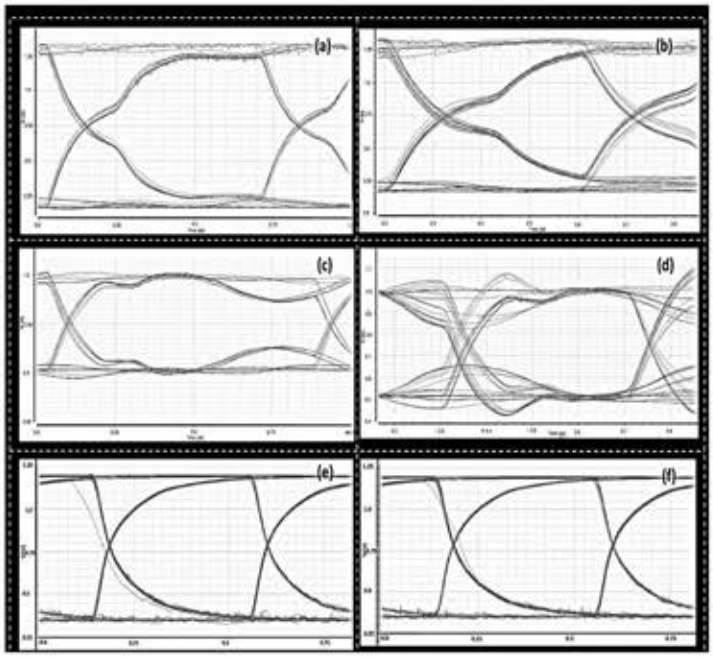

Figure 11 : (a) DDR3 Dual Rank Single DIMM at 1333MT/s, (b) DDR3 Dual Rank Single DIMM at 2166MT/s, (c) DDR3 Dual Rank Three DIMMs at 1333MT/s, (d) DDR3 Dual Rank Three DIMMs at 2166MT/s, (e) DDR4 Single Rank Single DIMM at 2400MT/s, (f) DDR4 Dual Rank

Table 4: Simulated Electrical Parameters for the above-mentioned simulations

Case	Far-End Termination	DRAM Access	Aperture (ps)	Jitter (ps)	Slew (V/ns)
DDR3 DATA 1333MT/s Dual-Rank Single-DIMM	ODT 120Ω	Write	680	7.25	1.8
DDR3 DATA 2166MT/s Dual-Rank Single-DIMM	ODT 120Ω	Write	365	13	2
DDR3 DATA 1333MT/s Dual-Rank Three-DIMMs	ODT 120Ω	Write	670	9	1.65
DDR3 DATA 2166MT/s Dual-Rank Three-DIMMs	ODT 120Ω	Write	350	28	1.5

The minimum adjustment step in the IOs impedance is around 0.4Ω over 34Ω, therefore, even in the presence of PVT variations, the impedance matching is ensured and hence the eye-patterns for SS, TT and FF corners cases from -1500C to 1250C remain almost the same.

Increasing DIMM-per-Channel (DPC) decreases the achieved VDC because of resistive voltage division at multiple terminations, eventually decreasing the vertical opening of the eyes. The capacitive loading from the extra signal-stubs, which not only cause impedance mismatch but also degrade settling hand reducing horizontal opening of the eyes. These effects limit DPC to a maximum of three dual-rank UDIMMs from the data-signals perspective at 2166MT/s. However, clock signals can only afford a single dual-rank UDIMM at 1033MHz. RDIMMs, with clock signal buffering, can reach three dual-rank DPC 1033MHz, however, they are still limited to maximum three DIMMs due to data signals loading.

Tables 5 to Table 7 summarize DPC with types of DIMMs and affordable speeds extracted from simulations for Data-Lines and Clock-Lines

Table 5 : Data Lines for DDR3 UDIMM

Ranks / Speed	SR1	DR1	SR1-SR2	DR1-DR2	SR1-SR2-SR3	DR1-DR2-DR3
1033MT/s	√	√	√	√	√	√
1333MT/s	√	√	√	√	√	√
1600MT/s	√	√	√	√	√	√
1866MT/s	√	√	√	√	√	√
2166MT/s	√	√	√	√	√	√

Table 6: Clock Lines for DDR3 UDIMM

Ranks / Speed	SR1	DR1	SR1-SR2	DR1-DR2	SR1-SR2-SR3	DR1-DR2-DR3
533MHz	√	√	√	√	√	X
666MHz	√	√	√	√	√	X
800MHz	√	√	√	X	√	X
933MHz	√	√	√	X	√	X
1033MHz	√	√	√	X	√	X

Table 7: Clock Lines for DDR3 RDIMM

Ranks / Speed	SR1	DR1	SR1-SR2	DR1-DR2	SR1-SR2-SR3	DR1-DR2-DR3
533MHz	√	√	√	√	√	√
666MHz	√	√	√	√	√	√
800MHz	√	√	√	√	√	√
933MHz	√	√	√	√	√	√
1033MHz	√	√	√	√	√	√

Slew-rate control can also reduce inter-symbol-interference (ISI) due to reflection by low-pass shaping and attenuating high-frequency components in the dispatched signal. These high-frequency components, related to a high edge-rate, are the primary cause of reflections. Another advantage of slew-rate control in the IO pads is to cater for SSN. Figure 12 shows the simplified model of IO cells when operated in 4:1:1 and 2:1:1 configurations.

Table 8: Simulation Results for the SSN in different configurations

Config	Case	Vert. Eye Opening (mV)	Hor. Eye Opening (ps)	Slew Rate (V/ns)	VDD SAG (mV)	VSS BOUNCE (mV)
4:1:1	With Decaps	504	914	1.1	74	84.5
4:1:1	Without Decaps	487	415	2.2	84	86.1
4:1:1	With Decaps and Slew-control	487	418	2.18	50	68

SSN causes increased jitter in the 4:1:1 configuration at 2166MT/s, however, as both slew-rate control and decaps from power-pads are employed, the jitter reduces with slight reduction in horizontal opening, as shown in the figure 12 and table 8 above.

The calibration algorithm targets and connects the the pull-up legs in the calibrator IO cell to the external precision resistor. In the first-go incrementally adds resistance in parallel to match the external resistance. In the second-go, it starts with all legs connected in parallel and then systematically removes the resistances. The final matched is the average of both runs to minimize offset.

Table 9 lists down the power consumption of the IO cells based on the post-layout simulations.

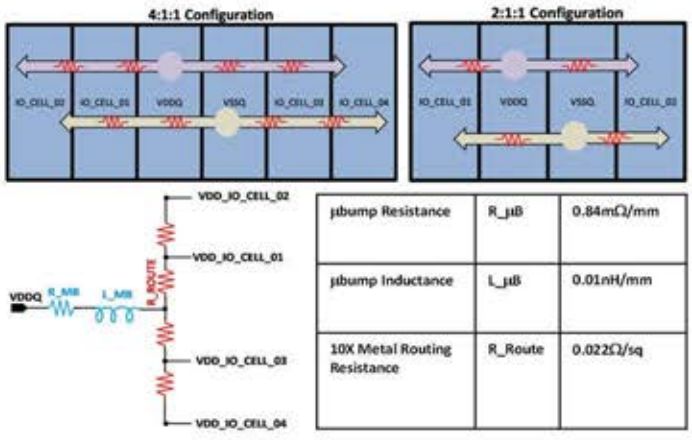

Figure 12 : Equivalent model for 4:1:1 and 2:1:1 configuration

Table 9: Power Consumption of the IOs

Component	@1.5V	@1.2V	Loading/Remarks
Receiver			
Static (mW)	0.75	0.75	
Dynamic (mW/Gbps)	16.0	7.4	
Transmitter			
Static (mW)	7.5	5.1	Single DIMM terminated at 120Ω-OCT at far-end
Dynamic (mW/Gbps) (without BBM)	10	7	Distributed RLC load of dual-rank single DIMM
Dynamic (mW/Gbps) (with BBM delay = 60ps)	8.3	5.8	Distributed RLC load of dual-rank single DIMM BBM delay results in a 17% reduction in the average dynamic power.
Leakage	0.001	0.0007	Both receiver and transmitter in power-down mode.

Table 10 compares the current work with the latest state-of-the-art available in literature. It can be seen the reported power consumption is comparable with state of the art results, however, since these are post-layout simulations and not silicon measurement results, a reasonable variation in the actual value can be expected.

V. CONCLUSIONS

This paper presented a DDR3 and DDR4 dual mode combo IO library containing half duplex combo transceiver, a separate calibration cell for ZQ calibration and Power/GND cells with RC clamp for ESD protection in 28nm CMOS technology. The post-layout extracted netlist simulations demonstrated a data-rate of 2166MT/s for DDR3 at 1.5V for a two dual rank UDIMMs/channel and 2400MT/s for DDR4 at 1.2V for a single dual rank UDIMM. The embedded break before make functionality in transmitter demonstrated a 17% reduction in crow-bar current. The slew rate control functionality demonstrated feasibility in reducing reflections on an under-damped channel as well as reducing simultaneous-switching-output noise. The transmitter consumed total power of 7.5mW/Gbps@1.5V and 5.8mW/Gbps@1.2V with break-before-make non-overlap delay of 60ps. Each cell was laid out with a footprint of 30µmx200µ, which lead to a continuously abutted structure when placed side by side.

Table 10: Comparison with State-of-the-art

Work & Year & Tech	VDD	Power	Area	Data-rate	ESD
[2] 2012 40nm	0.9/1.0/1.5	12.7mW/Gbps for GDDR5 35.7mW/Gbps for DDR3	2700um^2	1.6Gbps for DDR3 6.4Gbps for GDDR5	None
[6] 2014 22nm	1.2V-1.5V	8mW/Gbps DDR4 14mW/Gbps GDDR5	5500um^2	3.2Gbps for DDR4 6.4Gbps for GDDR5	None
[5] 2011 40nm	1.5V	2.8mW/Gbps for TX 1.14mW/Gbps for RX 19.6mW for TX 7.98mW for RX	2628um2 (only receiver)	7Gbps GDDR5	None
[9] 2009 75nm	1.5V	-	-	7Gbps GDDR5	None
[23] 2017 28nm	1	TX Power unknown 1.76mW/Gbps for Rx	24000um2	8Gbps for post LPDDR4	Yes
[24] 2018 20nm	1	-	-	12Gbps for GDDR5	None
[25] 2017 45nm	1.1	2.45mW/Gbps for Rx	87000um2	5.8Gbps	None
This work 2014 28nm	1.2V-1.5V	7.5mW/ Gbps@1.5V 5.8mW/ Gbps@1.2V	6000um2 for the half-duplex combo transceiver	2.4Gbps for DDR4 2.166Gbps for DDR3	HBM-2kV MM-8kV CDM-5A

References

1. *Jaejun Lee, et.al., "Architecture of a Multi-slot Main Memory System for 3.2 Gbps Operation". Circuits and Systems (ISCAS) proceedings of IEEE international Symposium 2010.*
2. *A Amirkhany, et al "A 12.8-Gb/s/link Tri-Modal Single-Ended Memory Interface"., IEEE Journal of Solid-State Circuits, 2012.*
3. *Min-Sun Keel et.al., "ESD-Resilient Active Biasing Scheme for high-speed SSTL IOs". Electrical Overstress/ Electrostatic Discharge symposium (EOS/ESD), 2013.*
4. *A, Amirkhany, et. Al., "On overcoming the limitations of single-ended signaling for graphics memory interfaces", Asian Solid-State Circuits Conference, 2011.*
5. *Seung-Jun Bae, et al "A 40nm 2Gb 7Gb/s/pin GDDR5 SDRAM with a programmable DQ ordering crosstalk equalizer and adjustable clock tracking BW". IEEE Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011.*
6. *Tzu-Chien Hsueh, et al "26.4 A 25.6Gb/s differential and DDR4/GDDR5 dual mode transmitter with digital clock calibration in 22nm CMOS" IEEE Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014.*
7. *EschJr, Gerald, and Tom Chen. "Near-linear CMOS I/O driver with less sensitivity to process, voltage, and temperature variations." IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2004*
8. *Sim, Jae-Yoon. "Circuit Design of DRAM for Mobile Generation." Journal of Semiconductor Technology and Science, 2007.*
9. *Rex Kho et al. "75nm 7Gb/s/pin 1Gb GDDR5 graphics memory device with bandwidth-improvement techniques" Solid State Circuits Conference- Digest of Technical Papers, 2009.*
10. *JEDEC DDR3 SDRAM STANDARD, JESD79-3D, July 2012.*
11. *JEDEC DDR4 SDRAM STANDARD, JESD79-4. September 2012.*
12. *B. A. Chappell et. al., "Fast CMOS ECL receivers with 100-mV worst-case sensitivity" IEEE Journal of Solid-State Circuits, Feb 1988.*
13. *M. Bazes, "Two novel fully complementary self-biased CMOS differential amplifiers", IEEE Journal of Solid-State Circuits, Feb 1991.*
14. *TSMC ESD Design Guide in 28nm HPM, 2012.*
15. *Steven Voldman, "ESD : Design and Synthesis", John Wiley & Sons 2011.*
16. *DesignWare DDR4 multiPHY Datasheet from Synopsys Inc., 2013.*
17. *JEDEC Standard on DDR3 SDRAM Un-buffered DIMM Design Specification*
18. *Design Guide for Two DDR3-1033 UDIMM Systems from Micron Inc.*
19. *DDR3 SDRAM Memory Interface Termination and Layout Guidelines from Altera Inc.*
20. *TN-41-13:DDR3 Point to Point Design Support from Micron Inc.*
21. *AMP single line model data-sheet (http://www.te.com)*
22. *DDR3 DIMM Sockets Catalog from Tyco Electronics.*
23. *Huimin He, Fengman Liu et al, "Research on Optical Transmitter and Receiver Module Used for High-Speed Interconnection between CPU and Memory", Taylor and Francis Fiber and Integrated Opitcs Journal, Volume 35, pg. 212-229, 2016.*
24. *Jeongsik Yoo, Yeonho Lee, Yoonjae Choi, "A Low-Power Post-LPDDR4 Interface Using AC Termination at RX and an Active Inductor at TX", IEEE Transactions on Circuits and Systems II: Express Briefs, Issue 99, pg. 1, September, 2017.*
25. *Martin Brox, Mani Balakrishnan, Martin Broschwitz, Cristian Chetreanu et al, "An 8-Gb 12-Gb/s/pin GDDR5X DRAM for Cost-Effective High-Performance Applications", IEEE Journal of Solid State Circuits, Vol. 53, Issue 1, January 2018.*
26. *Hyun-Wook Lim, Sung Won Choi, Jeong-Keun Ahn et al, "A 5.8-Gb/s Adaptive Integrating Duobinary DFE Receiver for Multi-Drop Memory Interface", IEEE Journal of Solid State Circuits, Vol. 52, Issue 6, June 2017*

Tri-Compress: A Cascaded Data Compression Framework For Smart Electricity Distribution Systems

Syed Muhammad Atif[1], Anees Ahmed[1], Sameer Qazi[2]

Received: 30-Mrach-2018 / Revised: 26-April-2018 / Accepted: 02-May-2018

ABSTRACT

Modern smart distribution system requires storage, transmission and processing of big data generated by sensors installed in electric meters. On one hand, this data is essentially required for intelligent decision making by smart grid but on the other hand storage, transmission and processing of that huge amount of data is also a challenge. Present approaches to compress this information have only relied on the traditional matrix decomposition techniques benefitting from low number of principal components to represent the entire data. This paper proposes a cascaded data compression technique "Tri-Compress" that blends three different methods in order to achieve high compression rate for efficient storage and transmission. In the first and second stages, two lossy data compression techniques are used, namely Singular Value Decomposition (SVD) and Normalization; Third stage achieves further compression by using the technique of Sparsity Encoding (SE) which is a lossless compression technique but only having appreciable benefits for sparse data sets. Our simulation results show that the combined use of the 3 techniques achieves data compression ratio to be 15% higher than state of the art SVD for small, sparse datasets and up to 28% higher in large, non-sparse datasets with acceptable Mean Absolute Error (MAE).

KEYWORDS

Singular Value Decomposition, Sparse Matrix Representation, Smart Grid, Data Compression, Big Data

I. INTRODUCTION

Internet of Things (IoT) is a fast moving technology referring to a ubiquitous network of 'always connected' smart gadgets that can remotely perform monitoring and reporting of physical parameters of their surroundings to a remote entity through a network. The remote entity may make rapid decisions based on artificially intelligent algorithms or manual intervention by humans. Power distribution systems are also taking advantage of this concept through smart meters. Smart meters are rapidly replacing the conventional meters that allows power distribution companies to get insight of user power usage behavior and make their decision accordingly. Power grids are not only vital in distribution system but also very sensitive to fluctuations in demands. Sudden

[1] *Graduate School of Science and Engg, PAF Karachi Institute of Economics and Technology, Karachi, 75190, Pakistan, Email: syed.muhammad.atif@gmail.com*

[2] *College of Engineering, PAF Karachi Institute of Economics and Technology, Karachi, 75190, Pakistan*