Technical Note

DDR4 Point-to-Point Design Guide

Introduction

DDR4 memory systems are quite similar to DDR3 memory systems. However, there are several noticeable and important changes required by DDR4 that directly affect the board's design:

- New V_{PP} supply
- Removed V_{REFDQ} reference input
- Changed I/O buffer interface from midpoint terminated SSTL to V_{DD} terminated pseudo open-drain (POD)
- Added ACT_n control

DDR4 added over 30 new features with a significant number of them offering improved signaling or debug capabilities: CA parity, multipurpose register, programmable write preamble, programmable read preamble, read preamble training, write CRC, read DBI, write DBI, V_{REFDQ} calibration, and per DRAM addressability. It is beyond the scope of this document to provide an in-depth explanation of these features; however, a successful DDR4 high-speed design will require the use of these new features and they should not be overlooked. The Micron DDR4 data sheet provides in-depth explanation of these features.

As the DRAM's operating clock rates have steadily increased, doubling with each DDR technology increment, DRAM training/calibration has gone from being a luxury in DDR to being an absolute necessity with DDR4. For example, if the required V_{REFDQ} calibration and data bus write training were not correctly performed, DDR4 timing specifications would have to be severely derated; but the issue is moot since the specifications require V_{REFDQ} calibration and data bus write training.

The first section of this document highlights some new DDR4 features that can help enable a successful board operation and debug. These features offer the potential for improved system performance and increased bandwidth over DDR3 devices for system designers who are able to properly design around the timing constraints introduced by this technology. The second section outlines a set of board design rules, providing a starting point for a board design. And the third section details the calculation process for determining the portion of the total timing budget allotted to the board interconnect. The intent is that board designers will use the first section to develop a set of general rules and then, through simulation, verify their designs in the intended environment.

The suggestions provided in this technical note mitigating ^tRC, ^tRRD, ^tFAW, ^tCCD, and ^tWTR can help system designers optimize DDR4 for their memory subsystems. For system designers who find the increases offered by DDR4 are not enough to provide relief in their networking subsystems, Micron offers a comprehensive line of memory products specifically designed for the networking space. Contact your Micron representative for more information on these products.

DDR4 Overview

DDR4 SDRAM is a high-speed dynamic random-access memory internally configured as an 8-bank DRAM for the x16 configuration and as a 16-bank DRAM for the x4 and x8 configurations. The device uses an $8n$-prefetch architecture to achieve high-speed operation. The $8n$-prefetch architecture is combined with an interface designed to transfer two data words per clock cycle at the I/O pins.

A single READ or WRITE operation consists of a single $8n$-bit wide, four-clock data transfer at the internal DRAM core and two corresponding n-bit wide, one-half-clock-cycle data transfers at the I/O pins.

This section describes the key features of DDR4, beginning with Table 1, which compares the clock and data rates, density, burst length, and number of banks for the five standard DRAM products offered by Micron.The maximum clock rate and minimum data rate are the operating conditions with DLL enabled or normal operation.

Table 1: Micron's DRAM Products

Product	Clock Rate (^{t}CK)		Data Rate		Density	Prefetch (Burst Length)	Number of Banks
	Max	Min	Min	Max			
SDRAM	10ns	5ns	100 Mb/s	200 Mb/s	64–512Mb	1n	4
DDR	10ns	5ns	200 Mb/s	400 Mb/s	256Mb–1Gb	2n	4
DDR2	5ns	2.5ns	400 Mb/s	800 Mb/s	512Mb–2Gb	4n	4, 8
DDR3	2.5ns	1.25ns	800 Mb/s	1600 Mb/s	1–8Gb	8n	8
DDR4	1.25ns	0.625ns	1600 Mb/s	3200 Mb/s	4–16Gb	8n	8, 16

Density

The JEDEC® standard for DDR4 SDRAM defines densities ranging from 2–16Gb; however, the industry started production for DDR4 at 4Gb density parts. These higher-density devices enable system designers to take advantage of more available memory with the same number of placements, which can help to increase the bandwidth or supported feature set of a system. It can also enable designers to maintain the same density with fewer placements, which helps to reduce costs.

Prefetch

As shown in Table 1, prefetch (burst length) doubled from one DRAM family to the next. With DDR4, however, burst length remains the same as DDR3 (8). (Doubling the burst length to 16 would result in a x16 device transferring 32 bytes of data on each access, which is good for transferring large chunks of data but inefficient for transferring smaller chunks of data.)

Like DDR3, DDR4 offers a burst chop 4 mode (BC4), which is a psuedo-burst length of four. Write-to-read or read-to-write transitions get a small timing advantage from using BC4 compared to data masking on the last four bits of a burst length of 8 (BL = 8) access; however, other access patterns do not gain any timing advantage from this mode.

Frequency

The JEDEC DDR4 standard defines clock rates up to 1600 MHz, with data rates up to 3200 Mb/s. Higher clock frequencies translate into the possibility of higher peak bandwidth. However, unless the timing constraints decrease at the same percentage as the clock rate increases, the system may not be able to take advantage of all possible bandwidths. See DRAM Timing Constraints for more information

Error Detection and Data Bus Inversion

Devices that operate at higher clock and data rates make it possible to get more work done in a given period of time. However, higher frequencies also make it more complex to send and receive information correctly. As a result, DDR4 devices offer:

- Two built-in error detection modes: cyclic redundancy cycle (CRC) for the data bus and parity checking for the command and address bits.
- Data bus inversion (DBI) to help improve signal integrity while reducing power consumption.
- Both of these features will most likely be used for development and debug purposes.

CRC Error Detection

CRC error detection provides real-time error detection on the DDR4 data bus, improving system reliability during WRITE operations. DDR4 uses an 8-bit CRC header error control: X^8+X^2+X+1 (ATM-8 HEC). High-level, CRC functions include:

- DRAM generates checksum per write burst, per DQS lane: 8 bits per write burst (CR0–CR7) and a CRC using 72 bits of data (unallocated transfer bits are 1s).
- DRAM compares against controller checksum; if two checksums do not match, DRAM flags an error, as shown in the CRC Error Detection figure
- A CRC error sets a flag using the ALERT_n signal (short low pulse; 6–10 clocks)
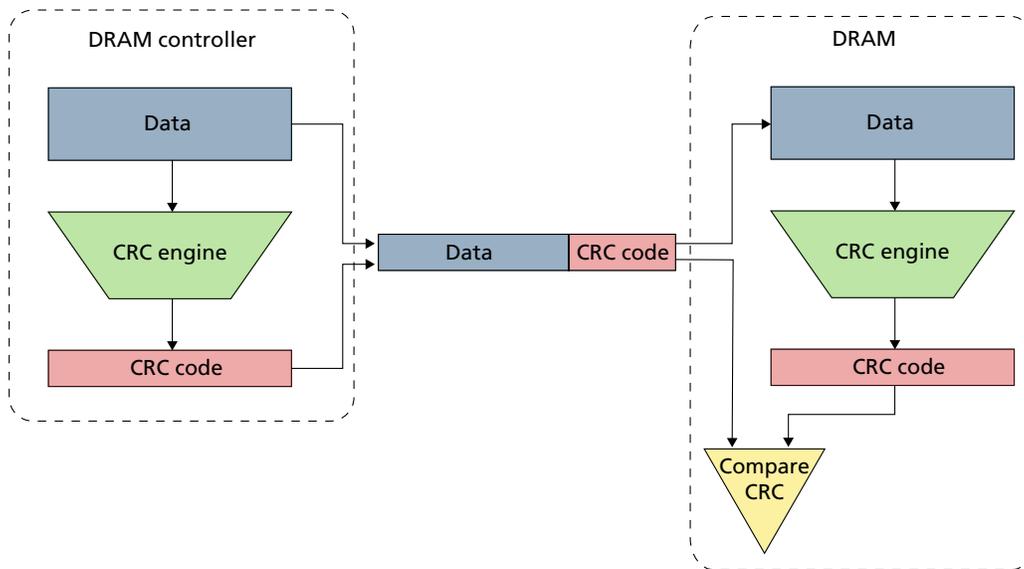
Figure 1: CRC Error Detection

**Table 2: CRC Error Detection Coverage**

Error Type	Detection Capability
Random single-bit errors	100%
Random double-bit errors	100%
Random odd count errors	100%
Random multi-bit UI error detection (excluding DBI bits)	100%

Parity Error Detection

Command/address (CA) parity takes the CA parity signal (PAR) input carrying the parity bit for the generated address and command signals and matches it to the internally generated parity from the captured address and command signals. High-level, parity error-detection functions include:

- CA parity provides parity checking of command and address buses: ACT_n, RAS_n, CAS_n, WE_n and the address bus (Control signals CKE, ODT, CS_n are not checked)
- CA parity uses even parity; the parity bit is chosen so that the total number of 1s in the transmitted signal—including the parity bit—is even
- The device generates a parity bit and compares with controller-sent parity; if parity is not correct, the device flags an error, as shown in the Command/Address Parity Operation
- A parity error sets a flag using the ALERT_n signal (long low pulse; 48–144 clocks)

Figure 2: Command/Address Parity Operation
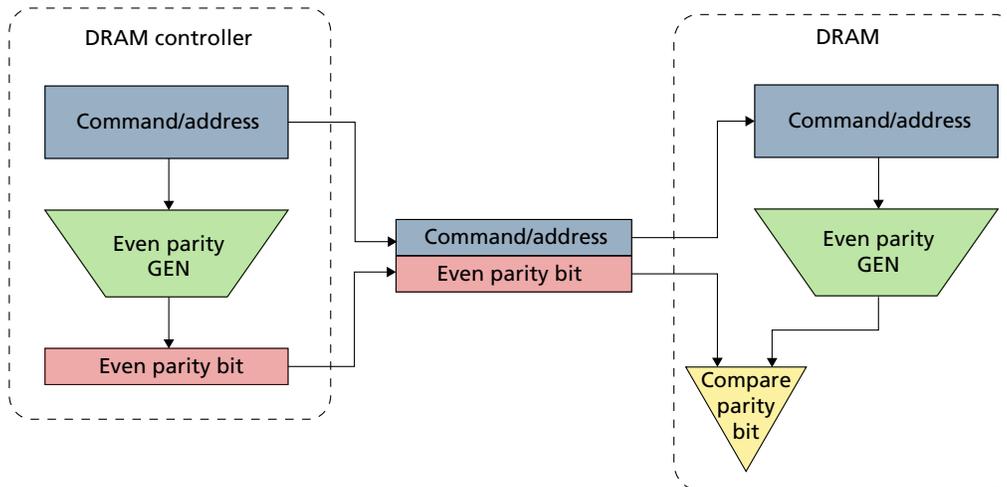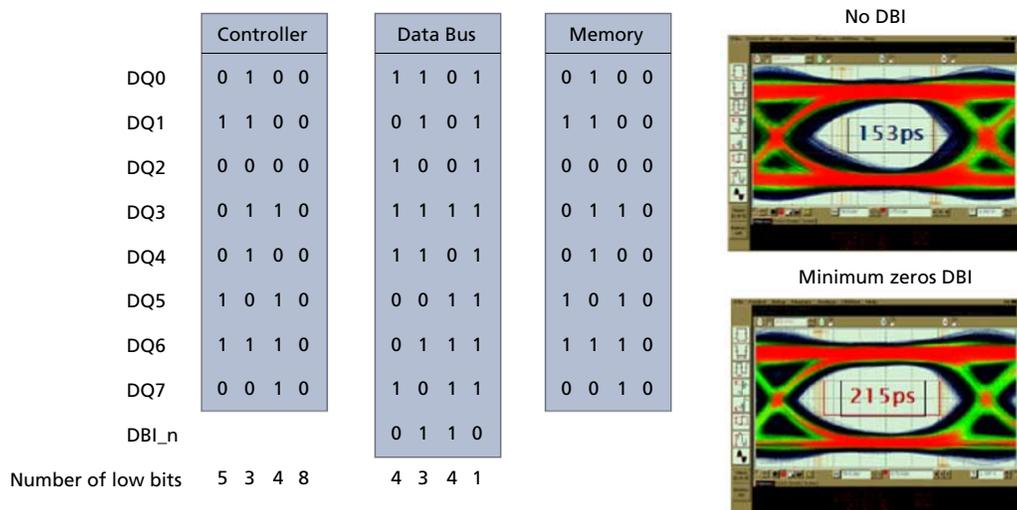

Data Bus Inversion

New to DDR4, the data bus inversion (DBI) feature enables these advantages:

- Supported on x8 and x16 configurations (x4 is not supported)
- Configuration is set per-byte: One DBI_n pin is for x8 configuration; UDBI_n, LDBI_n pins for x16 configuration
- Shares a common pin with data mask (DM) and TDQS functions; Write DBI cannot be enabled at the same time the DM function is enabled
- Inverts data bits
- Drives fewer bits LOW (maximum of half of the bits are driven LOW, including the DBI_n pin)
- Consumes less power (power only consumed by bits that are driven LOW)
- Enables fewer bits switching, which results in less noise and a better data eye
- Applies to both READ and WRITE operations, which can be enabled separately (controlled by MR5)

Table 3: DBI Example

Read	Write
If more than four bits of a byte lane are LOW: – Invert output data – Drive DBI_n pin LOW	If DBI_n input is LOW, write data is inverted – Invert data internally before storage
If four or less bits of a byte lane are LOW: – Do not invert output data – Drive DBI_n pin HIGH	If DBI_n input is HIGH, write data is not inverted

Figure 3: DBI Example

	Controller	Data Bus	Memory
DQ0	0 1 0 0	1 1 0 1	0 1 0 0
DQ1	1 1 0 0	0 1 0 1	1 1 0 0
DQ2	0 0 0 0	1 0 0 1	0 0 0 0
DQ3	0 1 1 0	1 1 1 1	0 1 1 0
DQ4	0 1 0 0	1 1 0 1	0 1 0 0
DQ5	1 0 1 0	0 0 1 1	1 0 1 0
DQ6	1 1 1 0	0 1 1 1	1 1 1 0
DQ7	0 0 1 0	1 0 1 1	0 0 1 0
DBI_n		0 1 1 0	
Number of low bits	5 3 4 8	4 3 4 1	

Banks and Bank Grouping

DDR4 supports bank grouping:

- x4/x8 DDR4 devices: four bank groups, each comprised of four sub-banks
- x16 DDR4 devices: two bank groups, each comprised of four sub-banks

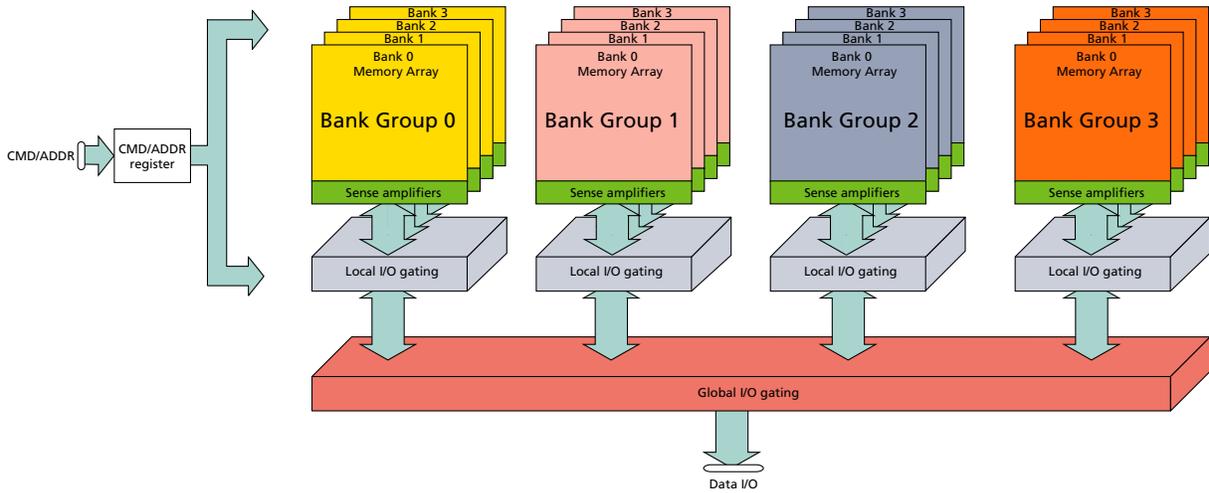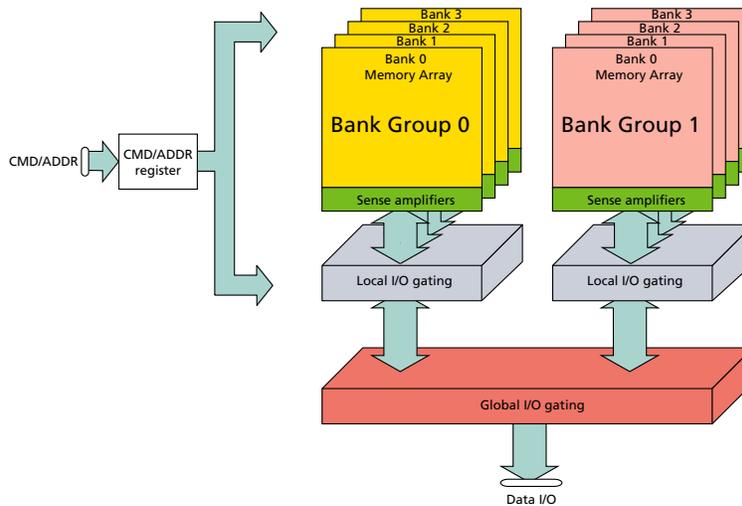**Figure 4: Bank Groupings—x4 and x8 Configurations**

**Figure 5: Bank Groupings—x16 Configuration**

Bank accesses to a different bank group require less time delay between accesses than bank accesses within the same bank group. Bank accesses to different bank group can use the short timing specification between commands, while bank accesses within the same bank group must use the long timing specifications.

Different timing requirements are supported for accesses within the same bank group and those between different bank groups:

- Long timings ($^tCCD\_L$, $^tRRD\_L$, and $^tWTR\_L$): bank accesses within the same bank group
- Short timings ($^tCCD\_S$, $^tRRD\_S$, $^tWTR\_S$): bank accesses between different bank groups

Figure 6: Bank Group: Short vs. Long Timing
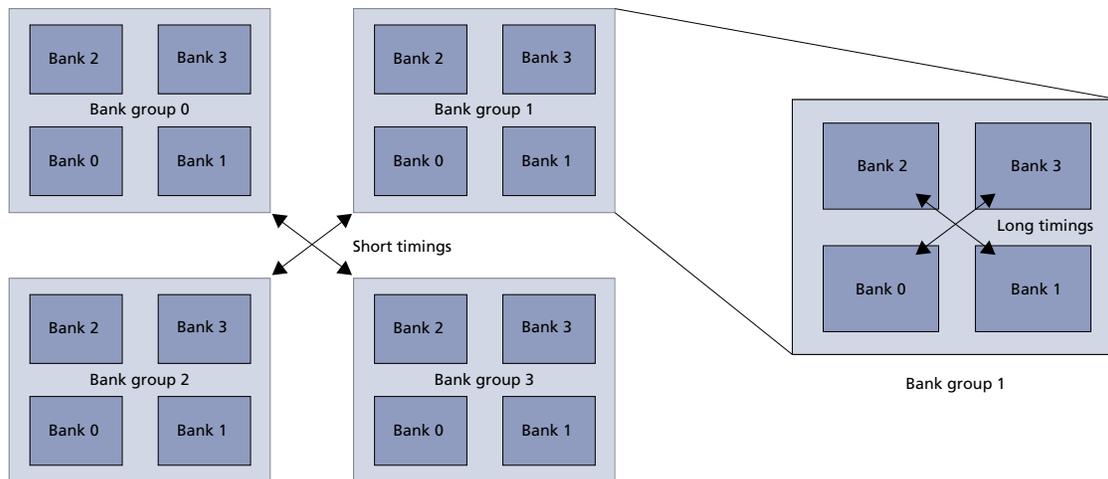
The tables below summarize the differences between DDR3 and DDR4 short and long bank-to-bank access timings ^{t}CCD, ^{t}RRD, and ^{t}WTR for DDR4-1600 through DDR4-2400. Refer to the DDR4 data sheet for timings above DDR4-2400. It is recommended the memory system utilize a ^{t}CCD_L of 5.8ns; system performance impact is likely to be negligible, if any. Accommodating a ^{t}CCD_L of 5.8ns enables the system to be backward-compatible as well as facilitate future DRAM timing adjustments.

To maximize system performance, it is important that bank-to-bank accesses are to different bank groups. If bank accessing is not controlled properly, it is possible to get less performance with a DDR4-based system versus a DDR3-based system.

Table 4: DDR3 vs. DDR4 Bank Group Timings – ^{t}CCD

Product	Parameter	1600	1866	2133	2400
DDR3	^{t}CCD	4CK	4CK	4CK	N/A
DDR4	^{t}CCD_S	4CK	4CK	4CK	4CK
DDR4	^{t}CCD_L	5CK or 6.25ns	5CK or 5.355ns	6CK or 5.355ns	6CK or 5ns

Table 5: DDR3 vs. DDR4 Bank Group Timings – ^{t}RRD

Product	Parameter	1600	1866	2133	2400
DDR3	^{t}RRD (1KB)	4CK or 5ns	4 CK or 5ns	4CK or 5ns	N/A
DDR4	^{t}RRD_S (1/2KB, 1KB)	4CK or 5ns	4 CK or 4.2ns	4CK or 3.7ns	4CK or 3.3ns
DDR4	^{t}RRD_L (1/2KB, 1KB)	4CK or 6ns	4CK or 5.3ns	4CK or 5.3ns	4CK or 4.9ns
DDR3	^{t}RRD (2KB)	4CK or 7.5ns	4CK or 6ns	4CK or 6ns	N/A
DDR4	^{t}RRD_S (2KB)	4CK or 6ns	4CK or 5.3ns	4CK or 5.3ns	4CK or 5.3ns
DDR4	^{t}RRD_L (2KB)	4CK or 7.5ns	4CK or 6.4ns	4CK or 6.4ns	4CK or 6.4ns

Table 6: DDR3 vs. DDR4 Bank Group Timings – ^{t}WTR

Product	Parameter	1600	1866	2133	2400
DDR3	^{t}WTR	4CK or 7.5ns	4CK or 7.5ns	4CK or 7.5ns	N/A
DDR4	^{t}WTR_S	2CK or 2.5ns	2CK or 2.5ns	2CK or 2.5ns	2CK or 2.5ns
DDR4	^{t}WTR_L	4CK or 7.5ns	4CK or 7.5ns	4CK or 7.5ns	4CK or 7.5ns

Manufacturing Features

DDR4 has three features that help with manufacturing: Post package repair, multiplexed address pins and connectivity test mode.

Post Package Repair (PPR): The Micron DDR4 SDRAM has one additional row available for repair per bank (16 per x4/x8, eight per x16) even though JEDEC only requires one additional row be available for repair per bank group (four per x4/x8, two per x16). PPR enables the end user to replace one suspect row in each bank with one good spare row.

Multiplexed Command Pins: To support higher density devices without adding additional address pins, DDR4 defined a method to multiplex addresses on the command pins (RAS, CAS, and WE). The state of the newly defined command pin (ACT_n) determines how the pins are used during an ACTIVATE command. High-level multiplexed command/address pin functions include:

- ACT_n along with CS_n LOW = the input pins RAS_n/A16, CAS_n/A15, and WE_n/A14 used as address pins A16, A15, and A14, respectfully.
- ACT_n HIGH along with CS_n LOW = the input pins RAS_n/A16, CAS_n/A15, and WE_n/A14 used as command pins RAS_n, CAS_n, and WE_n, respectfully for READ, WRITE and other commands defined in the command truth table.

Connectivity Test Mode: Connectivity test (CT) mode is similar to boundary scan testing but is designed to significantly speed up testing of the electrical continuity of pin interconnections between the DDR4 device and the memory controller on a printed circuit board.

Designed to work seamlessly with any boundary scan device, CT mode is supported on all x4, x8, and x16 Micron DDR4 devices. JEDEC specifies CT mode for x4 and x8 devices and as an optional feature on 8Gb and above devices.

Contrary to other conventional shift register-based boundary scan testing, where test patterns are shifted in and out of the memory devices serially during each clock, the DDR4 CT mode allows test patterns to be entered on the test input pins in parallel and the test results to be extracted from the test output pins of the device in parallel. This significantly increases the speed of the connectivity check.

When placed in CT mode, the device appears as an asynchronous device to the external controlling agent. After the input test pattern is applied, the connectivity test results are available for extraction in parallel at the test output pins after a fixed propagation delay time

Table 7: Connectivity Test Mode Pins

Pin Type (CT Mode)	Normal Operation Pin Names
Test Enable	TEN
Chip Select	CS_n
Test Inputs	BA0-1, BG0-1, A0-A9, A10/AP, A11, A12/BC_n, A13, WE_n/A14, CAS_n/A15, RAS_n/A16, CKE, ACT_n, ODT, CLK_t, CLK_c, Parity
	DML_n, DBIL_n, DMU_n/DBIU_n, DM/DBI
	ALERT_n
	RESET_n
Test Outputs	DQ0–DQ15, UDQS_t, UDQS_c, LDQS_t, LDQS_c

Logic Equations

Test input and output pins are related to the following equations, where INV denotes a logical inversion operation and XOR a logical exclusive OR operation.

MT0 = XOR (A1, A6, PAR)
MT1 = XOR (A8, ALERT_n, A9)
MT2 = XOR (A2, A5, A13)
MT3 = XOR (A0, A7, A11)
MT4 = XOR (CK_c, ODT, CAS_n/A15)
MT5 = XOR (CKE, RAS_n,/A16, A10/AP)
MT6 = XOR (ACT_n, A4, BA1)
MT7 = x16: XOR (DMU_n / DBIU_n , DML_n / DBIL_n, CK_t)
.... = x8: XOR (BG1, DML_n / DBIL_n, CK_t)
.... = x4: XOR (BG1, CK_t)
MT8 = XOR (WE_n / A14, A12 / BC, BA0)
MT9 = XOR (BG0, A3, RESET_n, TEN)

Output Equations for a x16 DDR4 device:

|	
DQ0 = MT0	DQ10 = INV DQ2
DQ1 = MT1	DQ11 = INV DQ3
DQ2 = MT2	DQ12 = INV DQ4
DQ3 = MT3	DQ13 = INV DQ5
DQ4 = MT4	DQ14 = INV DQ6
DQ5 = MT5	DQ15 = INV DQ7
DQ6 = MT6	LDQS_t = MT8
DQ7 = MT7	LDQS_c = MT9
DQ8 = INV DQ0	UDQS_t = INV LDQS_t
DQ9 = INV DQ1	UDQS_c = INV LDQS_c

DDR4 Key Changes

As previously noted, there are at least four important changes in DDR4 that require attention when developing a DDR4 motherboard:

- New V_{PP} supply
- Removed V_{REFDQ} reference input
- Changed I/O buffer interface from midpoint terminated SSTL to V_{DD} terminated pseudo open-drain (POD)
- Added ACT_n control

V_{PP} Supply

The V_{PP} supply was added, which is a 2.5V supply that powers the internal word line. Adding the V_{PP} supply facilitated the V_{DD} transition from 1.5V to 1.2V as well as provided additional power savings, approximately 10%. Although JEDEC does not state I_{DD} and I_{PP} current limits, initial DDR4 parts have demonstrated I_{PP} current usage in the ranges of a) 2mA to 3mA when in standby mode, b) 3mA to 4mA when in the active mode, and c) 10mA to 20mA during refresh mode. It is worth keeping in mind these I_{PP} values are average currents and actual current draw will be narrow pulses in nature, in the range of 20mA to 60mA. Failure to provide sufficient power to V_{PP} will prevent the DRAM from operating correctly.

Figure 7: Ipp Current Profile
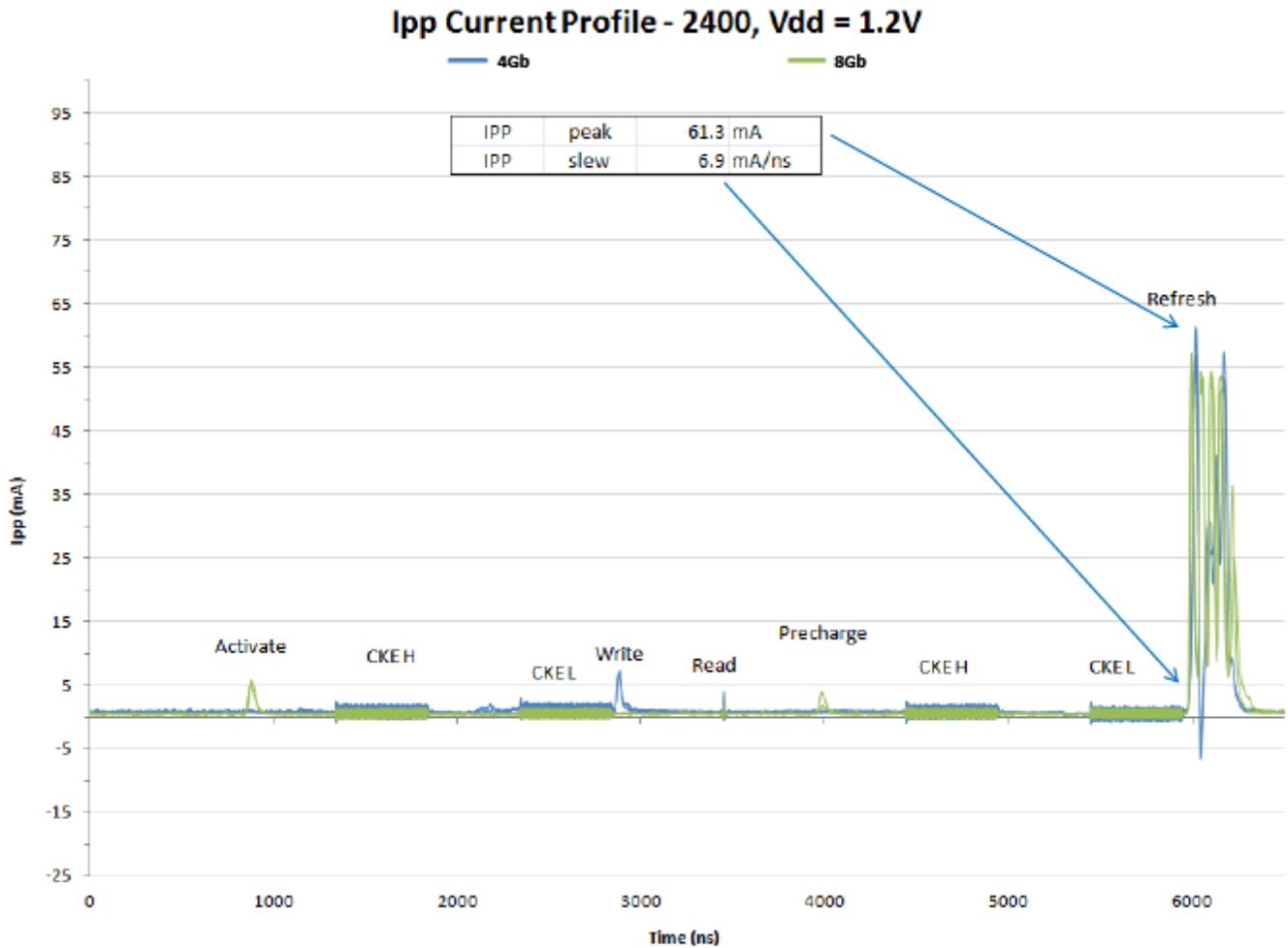
V_{REFDQ} Calibration

The V_{REFDQ} reference input supply was removed from the package interface and V_{REFDQ} is now internally generated by the DRAM. This means the V_{REFDQ} can be set to any value over a wide range; there is no specific value defined. This means the DRAM controller must set the DRAM's V_{REFDQ} settings to the proper value; thus, the need for V_{REFDQ} calibration.

JEDEC does not provide a specific routine on how to perform V_{REFDQ} calibration; however, JEDEC states allowed commands and how to enter and exit the mode. Each system will need to determine the routine to implement that provides it the best performance.

Although not to be construed as a detailed explanation of V_{REFDQ} calibration process and the most optimum methodology to employ when implementing V_{REFDQ} calibration, a general overview of how to look at the process is provided as a preview to a detailed studying of the DDR4 device specifications.

V_{REFDQ} Calibration Settings: The V_{REFDQ} can be set to either range 1 (between 60% and 92.5% of V_{DDQ}) or range 2 (between 40% and 77.5% of V_{DDQ}). Range 1 was defined with the intent of providing the choice range for module-based systems, while range 2 was defined with the intent of providing the choice range for point-to-point-based systems. Once the range is set, the internal V_{REF} can be adjusted in 0.65% V_{DDQ} ticks. Although there are specifications on tolerance of range settings, in reality these are of minimal interest when performing V_{REFDQ} calibration, as a specific value is not what is sought but rather the setting that provides the most optimum performance. Additionally, when using per DRAM addressability each DRAM may have a unique setting for its internal V_{REFDQ}.
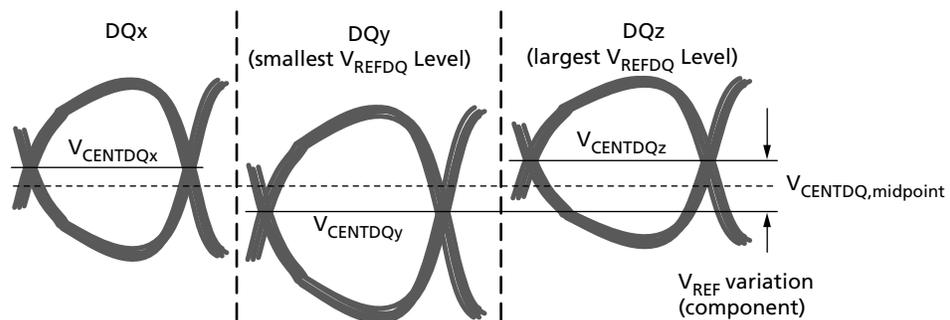
V_{REFDQ} Calibration Script: The following script is a reasonable platform to develop a V_{REFDQ} calibration routine around:

- Entering V_{REFDQ} calibration
- If range 1 then MR6 [7:6] 10* MR6 [5:0] XXXXXXX
- If range 2 then MR6 [7:6] 11* MR6 [5:0] XXXXXXX

  - Legal commands while in V_{REFDQ} calibration mode: ACT, WR, WRA, RD, RDA, PRE, DES, and MRS ** to set V_{REFDQ} values and exit V_{REFDQ} calibration mode
  - Subsequent V_{REFDQ} cal MR commands are MR6 [7:6] 10/1* MR6 [5:0] VVVVVV

- To exit V_{REFDQ} calibration, the last two V_{REFDQ} calibration MR commands are:

  - MR6 [7:6] 10/1* MR6 [5:0] VVVVVV' note VVVVV' = desired value for V_{REFDQ}
  - MR6 [7:6] 00/1* MR6 [5:0] VVVVVV' note exit V_{REFDQ} DRAM must be in idle state when exiting

*Range may only be set/changed when entering V_{REFDQ} calibration mode; changing range while in or exiting V_{REFDQ} calibration mode is illegal.

V_{REFDQ} Calibration Requirements: The goal is to find the best V_{REFDQ} setting that sets the internal V_{REFDQ} level to be the same as the DRAM's $V_{CENT\_DQ(pin\ mid)}$ level. Essentially, this requires the calibration process to determine what setting provides the largest optimal level for a DQ and lowest optimal level for a DQ for a given DRAM and use the setting half-way in between, as shown below.

Figure 8: V_{REFDQ} with $V_{CENT\_DQ(pin\ mid)}$

V_{REFDQ} Calibration Discussion: The following example is not to construe that there is a possible relaxation of the requirement that V_{REFDQ} calibration must be performed on each DRAM; rather, to show how much error can be induced if V_{REFDQ} calibration is not performed for each DRAM individually.

The first step is to determine the theoretical ideal V_{CENT_DQ}. This is based on the DRAM's ODT termination value used and the DRAM controller's driver impedance. Let's assume V_{DDQ} = 1.2V, the controller's R_{ON} = 34W, and the DRAM's ODT = 60W. This would make a LOW at 434mV and thereby want the internal V_{REF} set half way, which is 434mV+(1.2V - 434mV)/2 or 816mV, and is achieved setting the V_{REFDQ} setting at 0.68 V_{DDQ}, as shown below.

Figure 9: Theoretical V$_{CENT\_DQ(pin\ mid)}$
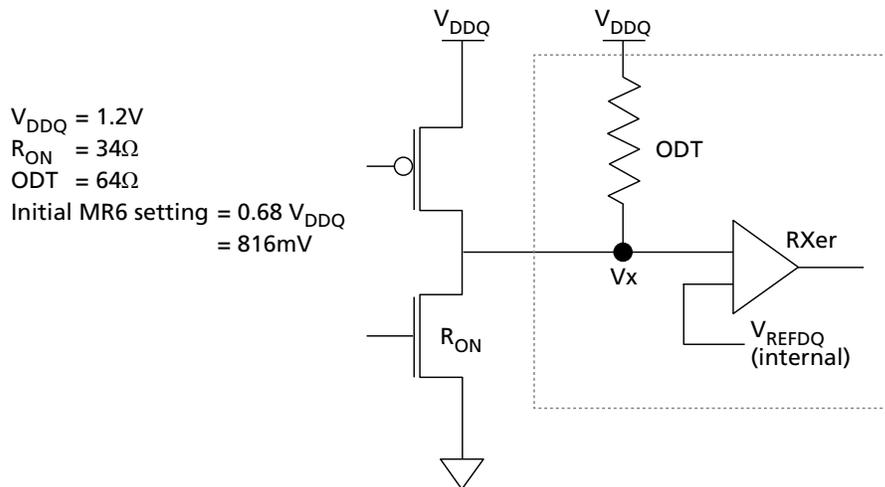
At this point, if the V_{REFDQ} register is set to $0.68 \times V_{DDQ}$, then the V_{REF} internal input is set to 816mV; however, $V_{CENT\_DQ(pin\ mid)}$ is left undefined. That is, without full calibration, $V_{CENT\_DQ(pin\ mid)}$ is not the same as the programmed value for V_{REFDQ}. Although undefined in the JEDEC specifications (since the condition is not allowed), setting the V_{REFDQ} setting at its theoretical ideal setting alone will only have the V_{REFDQ} programmed value within about ±7.5% of the correct $V_{CENT\_DQ(pin\ mid)}$ setting.

If subsequent reads and writes are performed to a rank of DRAMs at the same time when determining the largest and smallest V_{CENT_DQ} values, the final V_{REFDQ} programmed value will be within about ±4.0% of the correct $V_{CENT\_DQ(pin\ mid)}$ setting. However, if subsequent reads and writes are performed to a specific DRAM when determining the largest and smallest V_{CENT_DQ} values, the final V_{REFDQ} programmed value will then be the correct $V_{CENT\_DQ(pin\ mid)}$ setting.

Figure 10: V$_{REFDQ}$ Ranges
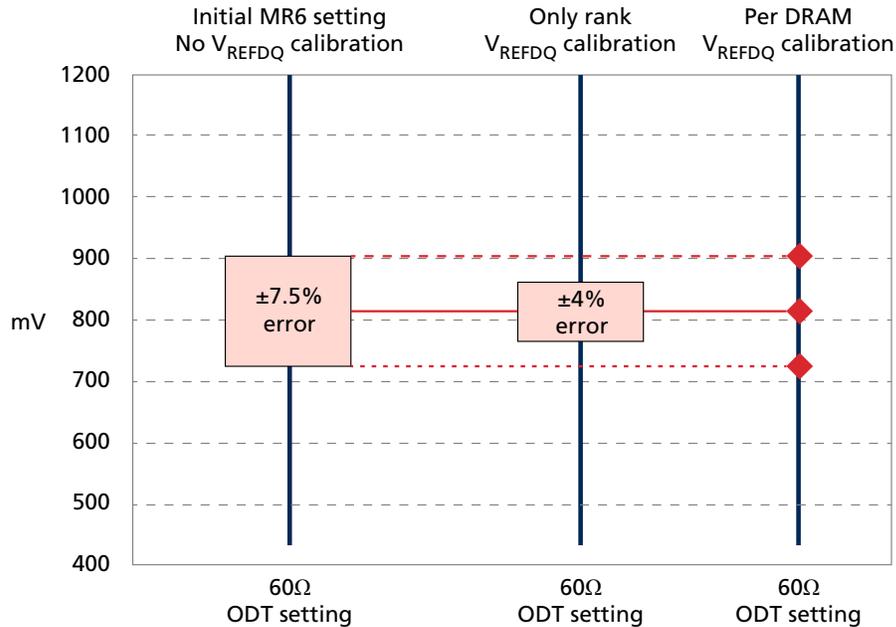
POD I/O Buffers

The I/O buffer has been converted from push-pull to pseudo open drain (POD), as seen in the figure below. By being terminated to V$_{DDQ}$ instead of 1/2 of V$_{DDQ}$, the size of and center of the signal swing can be custom-tailored to each design's need. POD enables reduced switching current when driving data since only 0s consume power, and additional switching current savings can be realized with DBI enabled. An additional benefit with DBI enabled is a reduction in crosstalk resulting in a larger data-eye.

Figure 11: DDR4 I/O Buffer vs. DDR3 I/O Buffer
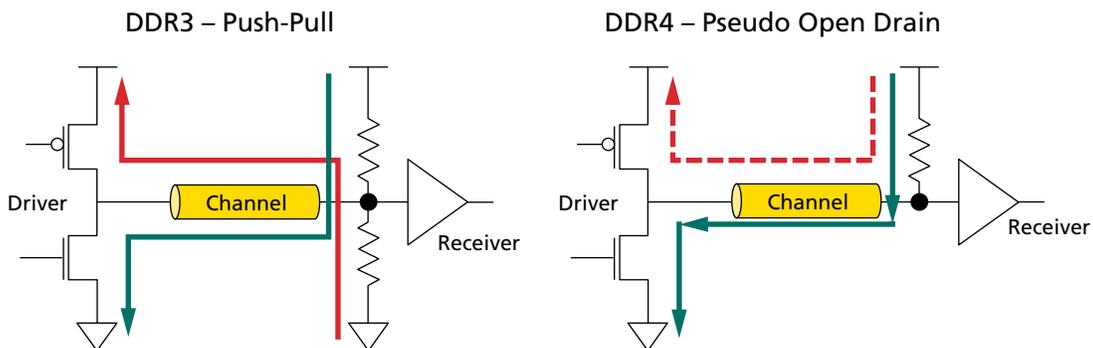
ACT_n Control

To help alleviate the demand for allocating pins after adding so many new features, DDR4 has for the first time multiplexed some of its address pins. The ACT_n determines whether RAS_n/A16, CAS_n/A15, and WE_n/A14 are to be treated as control pins or as address pins. As the nomenclature might suggest, ACT_n is an Active control when reg-

istered LOW; Activates are for latching the row address, which means when ACT_n is LOW, the three inputs RAS_n/A16, CAS_n/A15, and WE_n/A14 are treated as A16, A15, and A14, respectively. Conversely, when ACT_n is HIGH, the three inputs RAS_n/A16, CAS_n/A15, and WE_n/A14 are treated as RAS_n, CAS_n, and WE_n, respectively.

Command Bus and Address Bus Options

Two options are available for the command bus and address bus, each providing the following advantages and disadvantages:

Table 8: Bus Options

Bus Characteristics	Advantages and Disadvantages	
	Tree Bus	**Daisy Chain Bus**
Routing	Difficult	Easy
Performance	Excellent, but offers low bandwidth	Good, but offers high bandwidth
Load handling	Difficult and sensitive to large loads	Easy and unaffected by large loads
Timing skews	Minimal issues	Issues require leveling

For more details about command and address bus, see the DDR3 Point-to-Point Design Support technical note (TN 41-13) available on micron.com.

DDR4 Layout and Design Considerations

Layout is one of the key elements of a successfully designed application. The following sections provide guidance on the most important factors of layout so that if trade-offs need to be considered, they may be implemented appropriately.

Decoupling

Micron DRAM has on-die capacitance for the core as well as the I/O. It is not necessary to allocate a capacitor for every pin pair (V_{DD}:V_{SS}, V_{DDQ}:V_{SSQ}); however, basic decoupling is imperative.

Decoupling prevents the voltage supply from dropping when the DRAM core requires current, as with a refresh, read, or write. It also provides current during reads for the output drivers. The core requirements tend to be lower frequency. The output drivers tend to have higher frequency demands. This means that the DRAM core requires the decoupling to have larger values, and the output drivers want low inductance in the decoupling path but not a significant amount of capacitance.

One recommendation is to place enough capacitance around the DRAM device to supply the core and to place capacitance near the output drivers for the I/O. This is accomplished by placing four capacitors around the device on each corner of the package. Place one of the capacitors centered in each quarter of the ball grid, or as close as possible (see Decoupling Placement Recommendations Figure 12). Place these capacitors as close to the device as practical with the vias located to the device side of the capacitor. For these applications, the capacitors placed on both sides of the card in the I/O area may be optimized for specific purposes. The larger value primarily supports the DRAM core, and a smaller value with lower inductance primarily supports I/O. The smaller value should be sized to provide maximum benefit near the maximum data frequency.

Figure 12: Decoupling Placement Recommendations
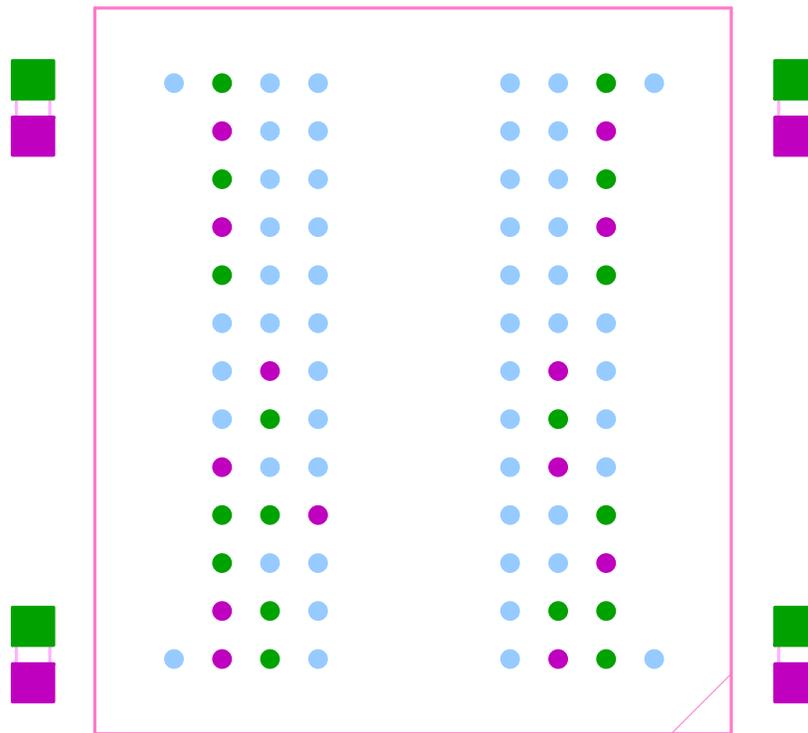
Note: 1. VDD= purple, VSS = green

Power Vias and Sharing

A DRAM device has five supply pin types: V_{DD} and V_{SS} (power the core), V_{DDQ} and V_{SSQ} (present only for the output drivers), and V_{PP}. The substrate for the device typically maintains isolation from the package balls all the way to the die where isolation is also maintained. This isolation is intended to keep I/O noise off of the core supply and core noise off of the I/O drivers. It is good practice, but not an absolute requirement, to use separate vias for V_{SS} and V_{SSQ} as well as for V_{DD} and V_{DDQ}.

There is a compromise position. Where a via connects to a V_{SS} ball on one side of the card and a V_{SSQ} ball on the other side of the card, the actual path being shared is minimized.

The path from the planes to the DRAM balls is important. Providing good, low inductance paths provides the best margin. Therefore, separate vias where possible and provide as wide of a trace from the via to the DRAM ball as the design permits.

Where there is concern and sufficient room, multiple vias are a possibility. This is generally applied at the decoupling cap to make a low impedance connection to the planes.

Return Path

If anything is overlooked, it will be the current return path. This is most important for terminated signals (parallel termination) since the current flowing through the termination and back to the source involves higher currents. No board-level (2D) simulators take this into account. They assume perfect return paths. Most simulators interpret that

an adjacent layer described as a plane is the perfect return path whether it is related to the signal or not. Some board simulators take into account plane boundaries and gaps in the plane to a degree. A 3D simulator is required to take into account the correct return path. These are generally not appropriate for most applications.

Most of the issues with the return path are discovered with visual inspection. The current return path is the path of least resistance. This may vary with frequency, so resistance alone may be a good indicator.

Trace Length Matching

Prior to designing the card, it is useful to decide how much of the timing budget to allocate to routing mismatch. This can be determined by thinking in terms of time or as a percentage of the clock period. For example, 1% (±0.5%) at 800 MHz clock is 6.25ps (1250ps/200). Typical flight times for FR4 PCB are near 6.5 ps/mm. So matching to ±1mm (±0.040 inch) allocates 1% of the clock period to route matching. Selecting 1mm is completely arbitrary. If the design is not likely to push the design limits, a larger number can be allocated.

When the design has unknowns, it is important to select a tighter matching approach. Using this approach is not difficult and allows as much margin as is conveniently available to allocate to the unknowns.

Address

For the address, the design will likely use a tree topology with branching. Making the branches uneven causes some signal integrity issues. For this reason, make all related branches match to within 1mm within each net. Different nets may have different branch lengths as long as they are matched within a branch. This is somewhat arbitrary, but there are many cases to consider, and 1mm should be adequate for all cases. There may be some exceptions.

Data Bus

For DQ, the topology is point-to-point or point-to-two-points where the two points are close together. For the data bus, the bit rate is the period of interest; that is, 625ps for an 800 MHz clock. Because 1% of this interval is 6.25ps, if the matching is held to a range of 1% (±0.5%), then ±0.5mm is the limit. Again, this is arbitrary.

Other factors to account for are vias, differences in propagation time for routing on inner layers versus outer layers, and load differences.

Propagation Delay

Propagation delay for inner layers and outer layers is different because the effective dielectric constant is different. The dielectric constant for the inner layer is defined by the glass and resin of the PCB. Outer layers have a mix of materials with different dielectric constants. Generally, the materials are the glass and resin of the PCB, the solder mask that is on the surface, and the air that is above the solder mask. This defines the effective dielectric for the outer layers and usually amounts to a 10% decrease in propagation delay for traces on the outer layers. For the design of JEDEC UDIMMs, a 10% difference accounts for the differences in propagation of the inner layers versus the outer layers. If all traces that need to match are routed with the same percentage on the outer layers versus the inner layers, this difference may be ignored for the purpose of match-

ing timing. Otherwise, this difference should be accounted for in any delay or matching calculations.

For inner layer propagation, velocity is about 6.5 ps/mm. To match all traces within 10ps, traces must be held within a range of 1.5mm, 60 mils. In most cases, this can be easily achieved. Most designs tolerate a much greater variation and still have significant margin. The engineer must decide how much of the timing budget is allocated to trace matching.

Vias

In most cases, the number of vias in matched lines should be the same. If this is not the case, the degree of mismatch should be held to a minimum. Vias represent additional length in the Z direction. The actual length of a via depends on the starting and ending layers of the current flow. Because all vias are not the same, one value of delay for all vias is not possible. Inductance and capacitance cause additional delay beyond the delay associated with the length of the via. The inductance and capacitance vary depending on the starting and ending layers. This is either complex or labor-intensive and is the reason for trying to match the number of vias across all matched lines. Vias can be ignored if they are all the same. A maximum value for delay through a via to consider is 20ps. This number includes a delay based on the Z axis and time allocated to the LC delay. Use a more refined number if available; this generally requires a 3D solver.

Timing Budget

Suggested practice is to look at the design from a timing budget standpoint to provide flexibility in the routing portion of the design, if there is suitable margin. This starts with simulation. By referencing the eye diagrams in this document, a setup and hold time can be established. From here, the parameters not included in the simulation must be added.

Typical routing for DDR4 components requires two internal signal layers, two surface signal layers, and four other layers ($2\,V_{DD}$ and $2\,V_{SS}$) as solid reference planes.

DDR4 memories have V_{DD} and V_{DDQ} pins, which are both typically tied to the PCB V_{DD} plane. Likewise, component V_{SS} and V_{SSQ} pins are tied to the PCB V_{SS} plane. Each plane provides a low-impedance path to the memory devices to deliver V_{SSQ}. Sharing a single plane for both power and ground does not provide strong signal referencing. With careful design, it is possible for a split-plane design to work adequately:

- Designs should reference data bus signals to V_{SS}.
- CA bus and clock should reference V_{DD}.
- Signals should never reference V_{PP}.

Drive Strength and Calibration

Matching the driver to the transmission line eliminates reflections that return to the driver to provide cleaner edges and a more open eye. See the DDR3 Point-to-Point Design Support technical note (TN 41-13) available on micron.com to learn the effects of mismatching a driver to the transmission line.

- DDR4 drive strengths: 48Ω and 34Ω
- Micron drive strength: 40Ω

Data Bus Topology

The improvements in the controller are reduced skew, improved setup and hold, improved package parasitic, improved calibration, and added adjustment and training features. Not all controllers have these features.

Improvements in the DRAM device are reduced skew, reduced setup and hold, improved package parasitic, improved calibration, and improved support for training.

The terminations are on-die, either in the controller or the DRAM device, with a termination resistance near the transmission line impedance.

Signal Optimization

- Avoid crossing splits in the power plane.
- Separate supplies and/or flip-chip packaging to help prevent controller SSO occurrence and the strobe/clock collapses it causes.
- Add low-pass V_{REFCA} filtering on the controller to improve noise margin.
- Minimize V_{REF} noise using spacing techniques like those recommended for signals implementing V_{REFCA}. Maintain a single reference (either ground or V_{DD}) between the decoupling capacitor and the DRAM V_{REFCA} pin.
- Minimize ISI by keeping impedances matched.
- Minimize crosstalk by isolating sensitive bits, such as strobes, and avoiding return-path discontinuities.
- Enhance signaling by matching driver impedance with trace impedance.

Simulations

For a new or revised design, Micron strongly recommends simulating I/O performance at regular intervals (pre- and post- layout for example). Optimizing an interface through simulation can help decrease noise and increase timing margins before building prototypes. Issues are often resolved more easily when found in simulation, as opposed to those found later that require expensive and time-consuming board redesigns or factory recalls.

Micron has created many types of simulation models to match the different tools in use. Component simulation models currently on micron.com include IBIS, Verilog, VHDL, Hspice, Denali, and Synopsys. Verifying all simulated conditions is impractical, but there are a few key areas to focus on: DC levels, signal slew rates, undershoot, overshoot, ringing, and waveform shape.

Also, it is extremely important to verify that the design has sufficient signal-eye openings to meet both timing and AC input voltage levels. For additional general information on the simulation process see the DDR4 SDRAM Point-to-Point Simulation Process technical note (TN 46-11) available on micron.com.

DDR4 Subsystem Attributes and Assumptions

Table 9: DDR4 Bus

Subsystem Component	Name	Description
Physical Bus		
Data	DQ/DQS/DM	3200 Mb/s (DDR)
Command/Address	CA	1600 Mb/s (SDR)
Clock	CK/CK#	1600 MHz
Bus Operations		
READ	READ	–
WRITE	WRITE	–
Data Bus Topology		
Configuration	Point-to-point	–
Trace	Mean length	5, 15, 25, 35, 45, and 60mm
	Width (min)	~0.1mm for Zo ~40Ω to 50Ω
	Spacing (min)	~0.2mm with a dialectric thickness of 0.08mm (3 mils) for Zo ~40Ω to 50Ω
PCB	Stackup, 8-layer (4 signals)	Target: Zo ~45Ω to 55Ω (FR4)

Simulation Setup and Models

- Physical bus: Data at 3200 Mb/s DDR
- Bus operation: Read
- Configuration: Point to 1 Point
- PCB model: Hspice frequency dependent W-element model with 10 coupled lines
- PCB target impedance: 50Ω ±10%
- Controller input capacitance load: 1.5pF (value from Controller IBIS model from MTK)
- Byte simulated: DQ0, DQ1,….DQ7, DQS0/DQS0# which has highest package crosstalk
- Eye measurement method: Aperture DC window (V_{REF} ±50mV) with V_{REF} centering
- Pass/Fail criteria: Aperture DC ≥70% UI, voltage margin ≥100mV, overshoot ≤200mV
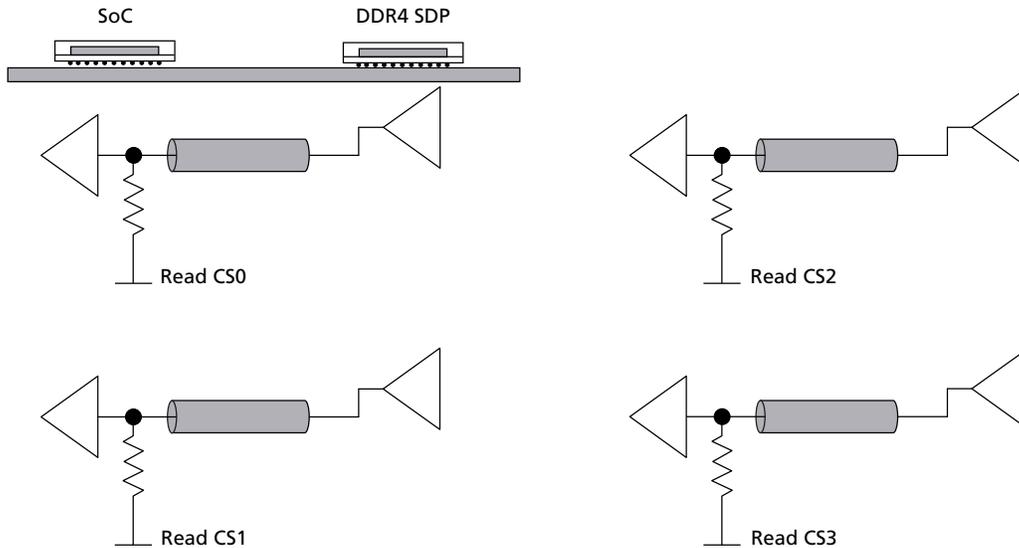
Typical Configuration

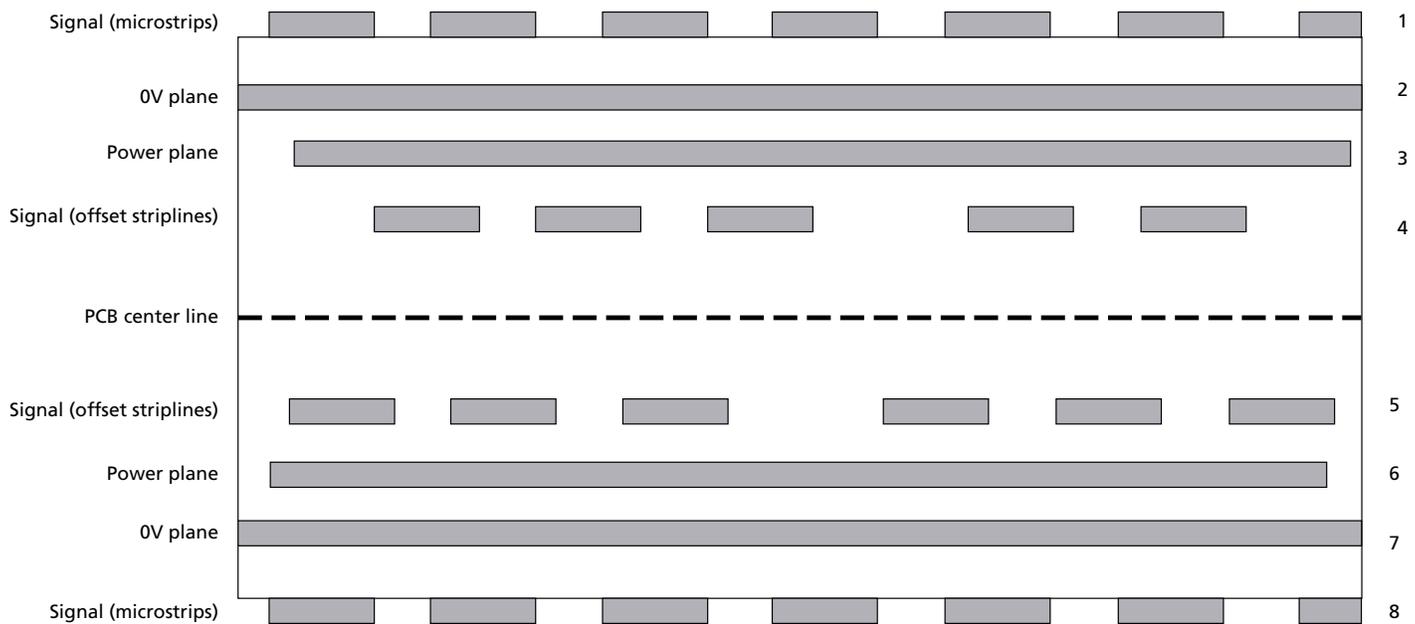**Figure 13: Typical 2GB x 4 Configuration**
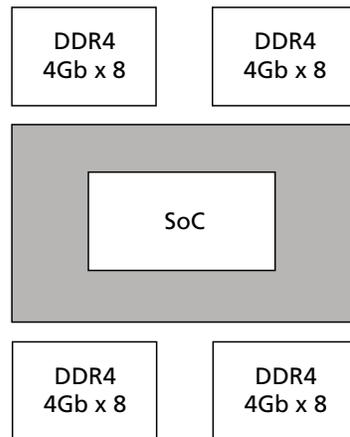
2GB DDR4 Read-Only Subsystem

Figure 14: Data Bus – ~4 DQ Channels of x8 per Component (1 Rank [CS] per Channel)

PCB Stackup

Figure 15: PCB Stackup – Example of 8 Layers (4 Signal, 4 Power Planes)

Eye Diagrams With DRAM R$_{ON}$ = 48Ω, Controller ODT = 120Ω
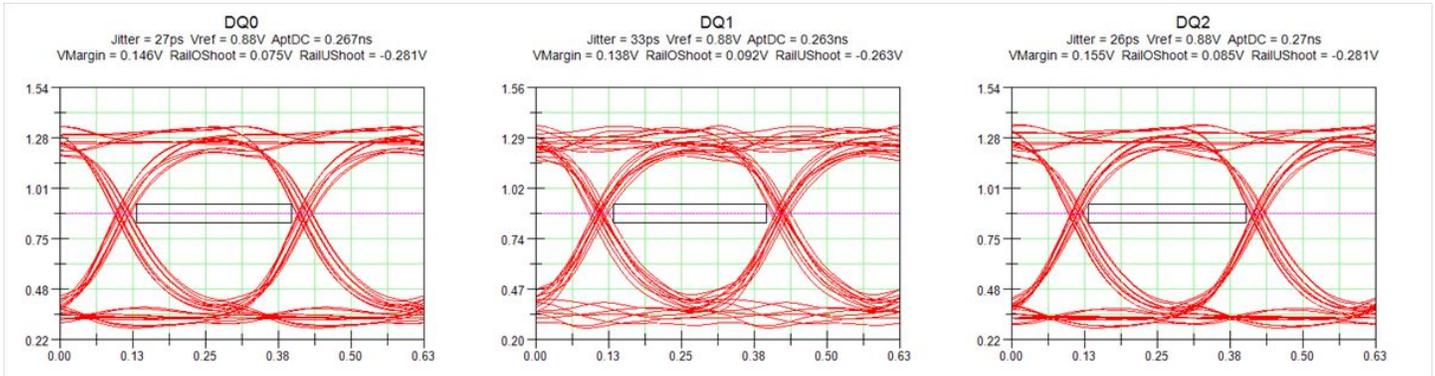
Figure 16: MB Length = 5mm
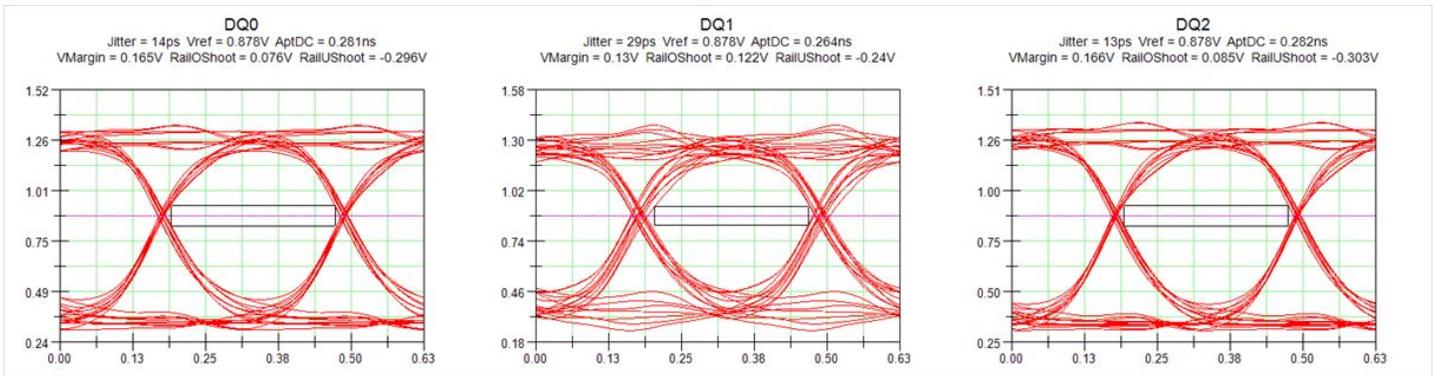

Figure 17: MB Length = 10mm
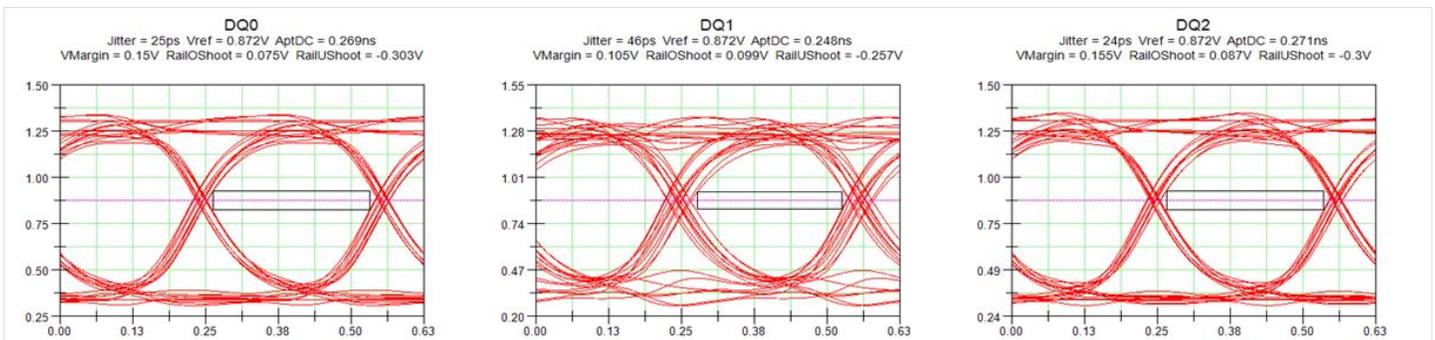

Figure 18: MB Length = 20mm

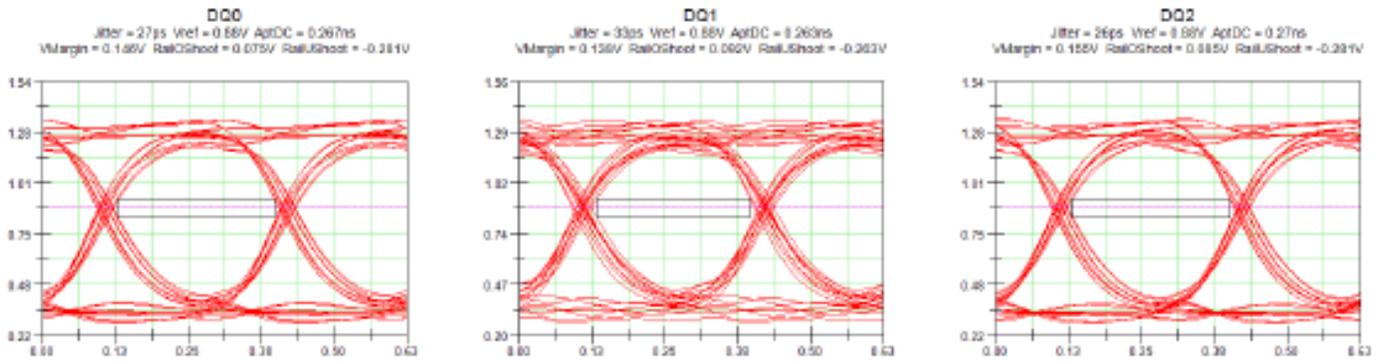**Figure 19: MB Length = 30mm**

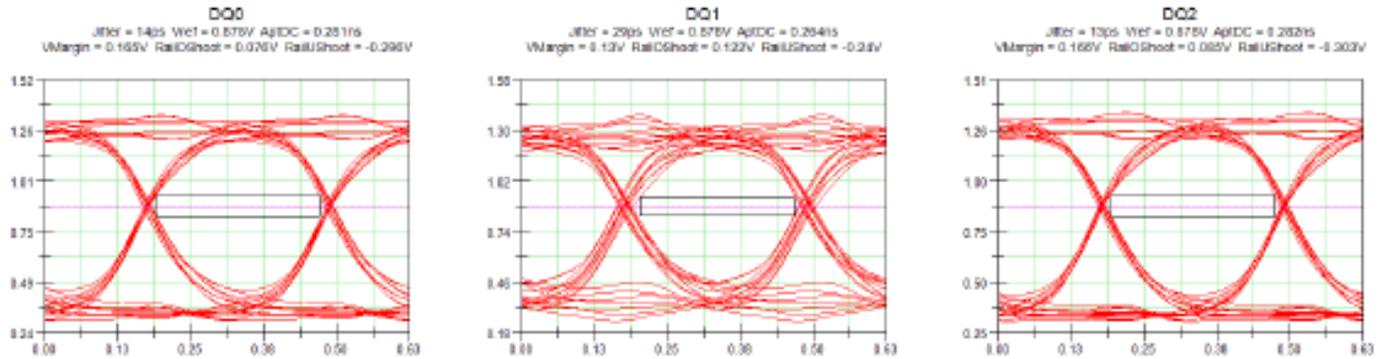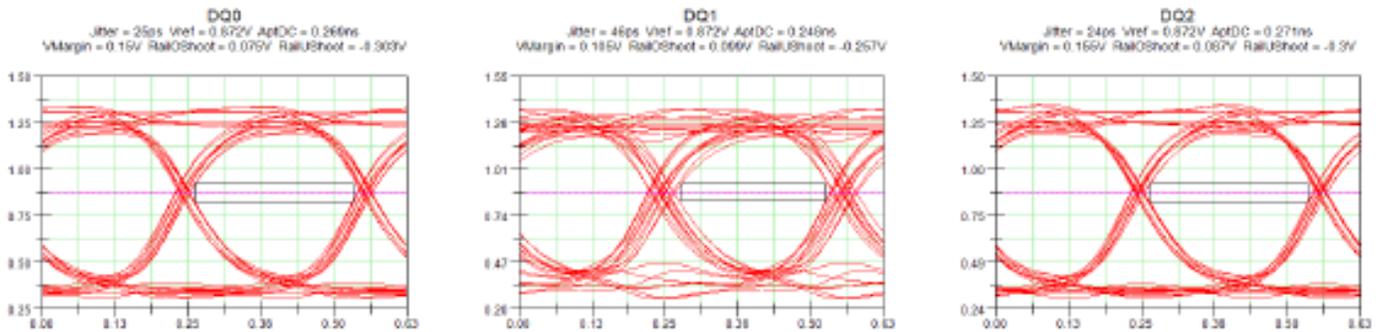**Figure 20: MB Length = 40mm**

Figure 21: MB Length = 50mm

25

Data Read R_{ON} and ODT Recommendations

Table 10: Recommendations- DRAM Driver Impedance and Controller ODT Settings

DRAM (Driver) R_{ON}	Controller ODT	V_{DDQ}	Package-Z		Mother Board Z	Memory PVT
			Memory	Controller		
Mother Board Length: 5 to 10mm						
34	34, 40, 48, 60, 80, 120	1.26	54	54	55	Slow
40	34, 40, 48, 60, 80, 120, 240	1.26	54	54	55	Slow
48	34, 40, 48, 60, 80, 120, 240, Off	1.26	54	54	55	Slow
Mother Board Length: 11 to 20mm						
34	34, 40, 48, 60, 80, 120	1.26	54	54	55	Slow
40	34, 40, 48, 60, 80, 120, 240	1.26	54	54	55	Slow
48	34, 40, 48, 60, 80, 120, 240, Off	1.26	54	54	55	Slow
Mother Board Length: 21 to 30mm						
34	34, 40, 48, 60, 80, 120	1.26	54	54	55	Slow
40	34, 40, 48, 60, 80, 120, 240	1.26	54	54	55	Slow
48	34, 40, 48, 60, 80, 120, 240, Off	1.26	54	54	55	Slow
Mother Board Length: 31 to 40mm						
34	34, 40, 48, 60, 80, 120	1.26	54	54	55	Slow
40	34, 40, 48, 60, 80, 120, 240	1.26	54	54	55	Slow
48	34, 40, 48, 60, 80, 120, 240, Off	1.26	54	54	55	Slow
Mother Board Length: 41 to 50mm						
34	34, 40, 48, 60, 80, 120	1.26	54	54	55	Slow
40	34, 40, 48, 60, 80, 120, 240	1.26	54	54	55	Slow
48	34, 40, 48, 60, 80, 120, 240, Off	1.26	54	54	55	Slow
Mother Board Length: 51 to 60mm						
34	34, 40, 48, 60, 80, 120	1.26	54	54	55	Slow
40	34, 40, 48, 60, 80, 120, 240	1.26	54	54	55	Slow
48	34, 40, 48, 60, 80, 120, 240, Off	1.26	54	54	55	Slow

Notes: 1. Passing criteria: Aperture DC >= 70%; Voltage margin >= 100mV; Overshoot <= 200mV.
2. Based on simulation optimum signal integrity is achieved with controller ODT of 34Ω, 40Ω, 48Ω, 60Ω, or 80Ω.
3. Controller ODT of 120Ω, 240Ω, or Off yields acceptable signal integrity with recommended drive strength; therefore, these controllers are recommended in case the weaker ODT is beneficial, such as in the need to minimize power consumption.

4-Layer Design Recommendations

Figure 22: PCB Stackup—Example of 4 Layers

Signal (microstrips)	1
0V plane	2
Power plane	3
Signal (microstrips)	4
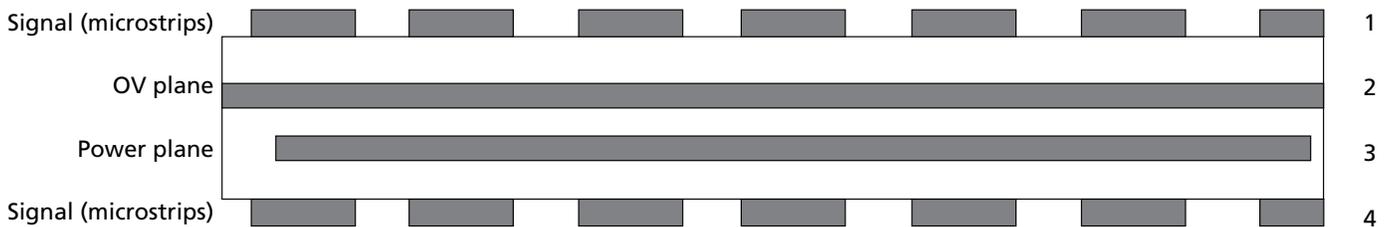
- All high-speed nets (DQ/DM/DQS, Address/Command, Control, Clock) should remain on the same reference plane (either power or ground), all the way from the DRAM pin to the controller pin.
- To help eliminate crosstalk due to vias, place the reference via (power or ground) next to each high-speed via that transitions to another layer.
- The clock pair should keep the same reference plane, all the way from the controller pins to the DRAM pins.
- Place decoupling capacitors as close as possible to the device.
- Perform signal integrity simulation to optimize Address/Command, Clock, DQ/DM/DQS termination and drive strength.
- Perform simulation to optimize on-board decoupling capacitor placement and values.
- To reduce power impedance at lower frequency, add more capacitors (two capacitors for each 4 signals is recommended).
- To reduce power impedance at higher frequency, make the V_{TT} plane tightly coupled to the ground plane and as large as possible.
- To reduce current and V_{TT} noise, reduce the controller's drive strength and increase termination resistor values while adhering to Address/Command bus timing specifications.

Unused Pin Guidance

Table 11: Unused Pin Details

Pin	Type	Recommendation
DM_n/DBI_n/TDQS or LDM_n/LDBI_n	I/O	**Data Mask/Data Bus Inversion/Termination Data Strobe:** Float if DBI, DM and TDQS features are disabled. Alternatively, connect to V_{DDQ} via a large (10,000Ω) pull-up resistor.
TDQS_c	Output	**Termination Data Strobe (x8 Devices Only):** Float if TDQS feature is disabled.
PAR	Input	**CA Parity:** Float or connect to ground if the command/address (CA) parity feature is disabled.
TEN	Input	**Connectivity Test Mode:** Connect to ground if the TEN feature is not used. If the TEN feature is used, it is recommended that the pin be connected to ground via a 1000Ω pull-down resistor.
ALERT_n	Output	**Alert:** If CA parity or write CRC modes are not used, ALERT_n may float unless connectivity test is used, ALERT_n should then be bound to V_{DDQ}. ALERT_n is an open drain output. Multiple devices can be connected together with a pull up at the end.
ODT	Input	**On-Die Termination:** Float if ODT pin is not used. Some cases do not require the use of the ODT pin (ODT does not need to be routed to provide sufficient termination). Single-rank (single-slot or memory-down) memory arrays only require R_{TT_WR} to be enabled; R_{TT_NOM} and R_{TT_Park} can be disabled. Dual-rank (single-slot or memory-down) memory arrays only require R_{TT_WR} and R_{TT_Park} to be enabled; R_{TT_NOM} can be disabled. Refer to the ODT section of this technical note or the data sheet for additional information.
DQ	I/O	**Data Input/Output:** Connect any unused DQ individually to V_{DDQ} via a large (10,000Ω) pull-up resistor or float. If only one of two bytes of a x16 device are used, assign the lower byte for data transfers and individually connect the upper byte (DQ[15:8]) to V_{DDQ} via a large (10,000Ω) pull-up resistor or float.
UDQS_t UDQS_c	I/O	**Data Strobe (Upper):** The only time a DQS strobe (true and compliment) should not be utilized is when the upper byte of a x16 device is not used. When the upper DQS strobe is not used, the UDQS_t should be connected to either V_{DDQ} or V_{SS}/ V_{SSQ} via a resistor in the 200Ω range. The UDQS_c should be connected to the opposite rail via a resistor in the same 200Ω range.
UDM_n/UDBI_n	I/O	**Upper Data Mask/Upper Data Bus Inversion:** Float if DBI and DM features are disabled. Alternatively, connect to V_{DDQ} via a large (10,000Ω) pull-up resistor.
C0/CKE1 C1/CS1_n C2/ODT1	Input	**Stacked-Die Address Inputs:** Float for single-die package (SDP) devices. For dual-die package (DDP) devices, use CKE1,CS1_n, ODT1 as directed by the data sheet. If 3DS-2H, use C0 and float C1 and C2; if 3DS-4H, use C0 and C1 and float C2; if 3DS-8H, use C0, C1 and C2.

Notes: 1. For the connectivity test (CT) to provide consistent results, all inputs must be in a known, defined state; whether driven from the controller or tied to V_{DD}/V_{DDQ} or ground. If a CT input is floating, pins associated with the floating CT input state may be indeterminate.
2. If upper byte of a x16 data bus is not used, use of Write CRC may be limited.

JEDEC DDP – Single Rank x16

A DDP composed of two x8s in a single rank improves the internal timing performance of the x16 configuration device. Some pins differ between the SDP and the DDP packages, but board design recommendations below support both SDP and DDP devices.

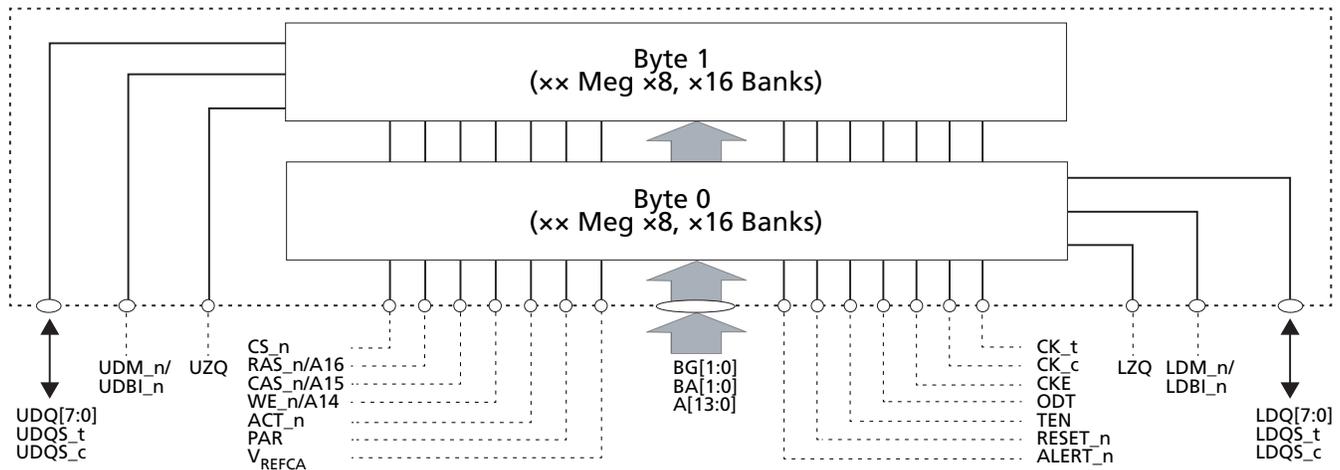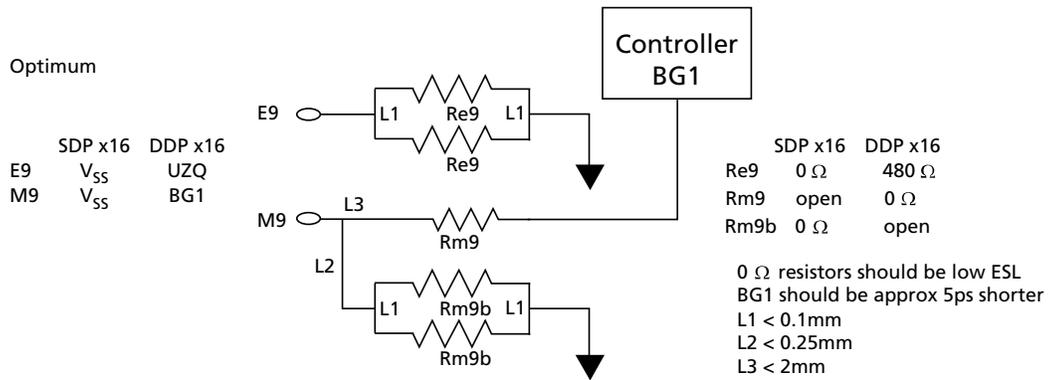**Figure 23: Device Performance – Two ×8s in a Board Space of One ×16**
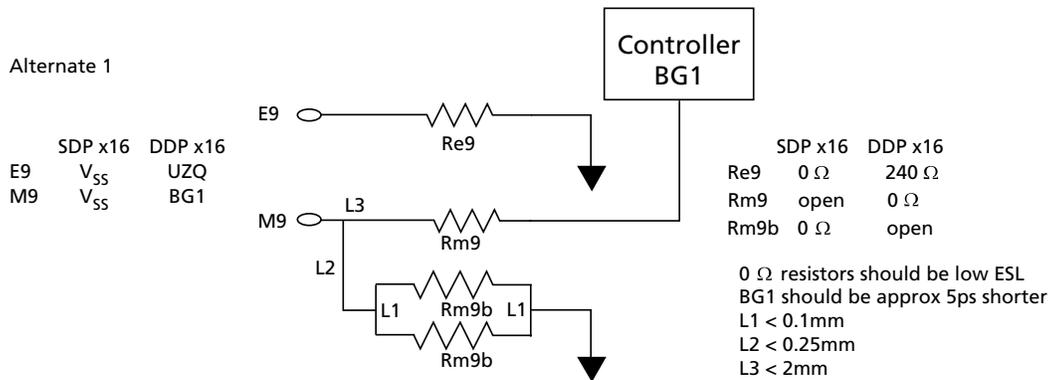

Table 12: JEDEC ×16 DDP Pin-Out

	DDP and SDP Symbols			DDP and SDP Symbols				DDP Symbols (x16)			SDP Symbols (x16)		
Pin	**1**	**2**	**3**	**4**	**5**	**6**	**Pin**	**7**	**8**	**9**	**7**	**8**	**9**
A	V_{DDQ}	V_{SSQ}	UDQ0	–	–	–	A	UDQS_c	V_{SSQ}	V_{DDQ}	UDQS_c	V_{SSQ}	V_{DDQ}
B	V_{PP}	V_{SS}	V_{DD}	–	–	–	B	UDQS_t	UDQ1	V_{DD}	UDQS_t	DQ9	V_{DD}
C	V_{DDQ}	UDQ4	UDQ2	–	–	–	C	UDQ3	UDQ5	V_{SSQ}	DQ11	DQ13	V_{SSQ}
D	V_{DD}	V_{SSQ}	UDQ6	–	–	–	D	UDQ7	V_{SSQ}	V_{DDQ}	DQ15	V_{SSQ}	V_{DDQ}
E	V_{SS}	UDM_n/ UDBI_n	V_{SSQ}	–	–	–	E	LDM_n/ LDBI_n	V_{SSQ}	UZQ	LDM_n/ LDBI_n	V_{SSQ}	V_{SS}
F	V_{SSQ}	V_{DDQ}	LDQS_c	–	–	–	F	LDQ1	V_{DDQ}	LZQ	DQ1	V_{DDQ}	ZQ
G	V_{DDQ}	LDQ0	LDQS_t	–	–	–	G	V_{DD}	V_{SS}	V_{DDQ}	V_{DD}	V_{SS}	V_{DDQ}
H	V_{SSQ}	LDQ4	LDQ2	–	–	–	H	LDQ3	LDQ5	V_{SSQ}	DQ3	DQ5	V_{SSQ}
J	V_{DDDLL}	V_{DDQ}	LDQ6	–	–	–	J	LDQ7	V_{DDQ}	V_{DD}	DQ7	V_{DDQ}	V_{DD}
K	V_{SS}	CKE	ODT	–	–	–	K	CK_t	CK_c	V_{SS}	CK_t	CK_c	V_{SS}
L	V_{DD}	WE_n/ A14	ACT_n	–	–	–	L	CS_n	RAS_n/ A16	V_{DD}	CS_n	RAS_n/ A16	V_{DD}
M	V_{REFCA}	BG0	A10/AP	–	–	–	M	A12/ BC_n	CAS_n/ A15	BG1	A12/ BC_n	CAS_n/ A15	V_{SS}
N	V_{SS}	BA0	A4	–	–	–	N	A3	BA1	TEN	A3	BA1	TEN
P	RESET_n	A6	A0	–	–	–	P	A1	A5	ALERT_n	A1	A5	ALERT_n
R	V_{DD}	A8	A2	–	–	–	R	A9	A7	V_{PP}	A9	A7	V_{PP}
T	V_{SS}	A11	PAR	–	–	–	T	V_{SS}	A13	V_{DD}	NC	A13	V_{DD}

Figure 24: Optimum Layout – DDP ×16 and SDP ×16 Compatibility

	SDP x16	DDP x16
E9	V$_{SS}$	UZQ
M9	V$_{SS}$	BG1

	SDP x16	DDP x16
Re9	0 Ω	480 Ω
Rm9	open	0 Ω
Rm9b	0 Ω	open

0 Ω resistors should be low ESL
BG1 should be approx 5ps shorter
L1 < 0.1mm
L2 < 0.25mm
L3 < 2mm

Note: 1.   Mitigates V$_{SS}$ offset; Parallel resistors when connecting to V$_{SS}$ reduces inductance.

Figure 25: Alternate One Layout – DDP ×16 and SDP ×16 Compatibility

	SDP x16	DDP x16
E9	V$_{SS}$	UZQ
M9	V$_{SS}$	BG1

	SDP x16	DDP x16
Re9	0 Ω	240 Ω
Rm9	open	0 Ω
Rm9b	0 Ω	open

0 Ω resistors should be low ESL
BG1 should be approx 5ps shorter
L1 < 0.1mm
L2 < 0.25mm
L3 < 2mm

Note: 1.   Mitigates V$_{SS}$ offset on M9 ball.

Revision History

Rev. F – 8/18

- Updated Layout and Design Considerations section

Rev. E – 2/18

- Corrected DQSU and DQSL to UDQS and LDQS the Overview and Unused Pin Guidance sections
- Updated the Alert_n pin recommendation in the Unused Pin Guidance section
- Updated Note 1 in the Unused Pin Guidance section

Rev. D – 11/17

- Replaced incorrect VDD1 reference with correct VPP reference

Rev. C – 7/17

- Correcting and adding information to the unused pin table.

Rev. B – 10/16

- Added Layout and Design Considerations section

Rev. A – 12/15

- Initial release
- 12/10 added new section on x16 DDP