External Memory Interfaces Handbook Volume 1

Section II: Memory Standard Overviews

Subscribe

Contents

1. Selecting your Memory Component

This chapter details some of the high-speed memory selection criteria and describes some typical applications where these memories are used. It looks at the main types of high-speed memories available, memory selection based on strengths and weaknesses, and which Altera® FPGAs these devices can interface with. It concludes with some typical application examples.

This chapter highlights the memory component's capability. The Altera IP may or may not support all of the features supported by the memory.

For the maximum supported performance supported by Altera FPGAs, refer to the *External Memory Interface System Specifications* section in volume 1 of the *External Memory Interface Handbook*.

System architects must resolve a number of complex issues in high-performance system applications that range from architecture, algorithms, and features of the available components. Typically, one of the fundamental problems in these applications is memories, as the bottlenecks and challenges of system performance often reside in its memory architecture. As higher speeds become necessary for external memories, signal integrity gets more difficult. Newer devices have added several features to overcome this issue. Altera FPGAs also support these advancements with dedicated I/O circuitry, various I/O standard support, and specialized intellectual property (IP).

Memory Overview

The main considerations for choosing an external memory device are bandwidth, size, cost, latency, and power. Since no single memory type can excel in every area, system architects must determine the right balance for their design.

There are two common types of high-speed memories: DRAM and SRAM. DRAM devices are volatile memories offering a lower cost per bit than SRAM devices. A compact memory cell consisting of a capacitor and a single transistor makes this possible, as opposed to the six-transistor cell used in SRAM. However, as the capacitor discharges, the memory cell loses its state. This means that DRAM memory must be refreshed periodically, resulting in lower overall efficiency and more complex controllers. Generally, designers only choose DRAM where cost per bit is important.

DDR, DDR2, and DDR3 SDRAM

The desktop computing market has positioned double data rate (DDR) SDRAM as a mainstream commodity product, which means this memory is very low-cost. DDR SDRAM is also high-density and low-power. Relative to other high-speed memories, DDR SDRAM has higher latency-they have a multiplexed address bus, which reduces the pin count (minimizing cost) at the expense of a longer and more complex bus

cycle. DDR2 SDRAM includes additional features such as increased bandwidth due to higher clock speeds, improved signal integrity on DIMMs with on-die terminations, and lower supply voltages to reduce power. DDR3 SDRAM is the latest generation of SDRAM and further increases bandwidth, lowers power, and improves signal integrity with fly-by and dynamic on-die terminations.

RLDRAM and RLDRAM II

Reduced latency DRAM (RLDRAM) is optimized to reduce latency primarily for networking and cache applications. RLDRAM is partitioned into eight smaller banks. This partioning reduces the parasitic capacitance of the address and data lines, allowing faster accesses and reducing the probability of random access conflicts. Also, most DRAM memory types need both a row and column phase on a multiplexed address bus to support full random access, while RLDRAM supports a non-multiplexed address, saving bus cycles at the expense of more pins. RLDRAM utilizes higher operating frequencies and uses the 1.8V High-Speed Transceiver Logic (HSTL) standard with DDR data transfer to provide a very high throughput. RLDRAM II offers faster random access times, on-die termination, a delay-locked loop (DLL) for higher frequency operation, larger densities, wider data paths, and higher bus utilization compared with RLDRAM.

QDR, QDR II, and QDR II+ SRAM

SRAMs are fundamentally different from DRAMs in that a typical SRAM memory cell consists of six transistors, while a DRAM cell consists of a transistor and a capacitor used to store a charge. Inherently, SRAM is a low-density, high-power memory device, with very low latency compared to DRAM (as the capacitor in the DRAM is slow). In most cases, SRAM latency is one clock cycle.

Quad Data Rate (QDR) SRAM has independent read and write ports that run concurrently at double data rate. QDR SRAM is true dual-port (although the address bus is still shared), which gives this memory a significantly higher bandwidth. QDR SRAM is best suited for applications where the required read/write ratio is near one-to-one. QDR II SRAM includes additional features such as increased bandwidth due to higher clock speeds, lower voltages to reduce power, and on-die termination to improve signal integrity. QDR II+ SDRAM is the latest and fastest generation.

Memory Selection

One of the first considerations in choosing a high-speed memory is data bandwidth. Based on the system requirements, an approximate data rate to the external memory should be determined. Table 1–1 details the memory bandwidth for various technologies with the assumptions of a 32-bit data bus, operating at the maximum supported frequency in a Stratix® IV FPGA. The bandwidth column in this table includes a conservative DRAM bandwidth at 70 percent efficiency, which takes into consideration bus turnaround, refresh, burst length, and random access latency. The calculation assumes 85 % efficiency for QDR and QDR II SRAM.

Table 1–1. Memory Bandwidth for 32-bit Wide Data Bus in Stratix IV FPGA

Memory	Clock Frequency (MHz)	Bandwidth for 32 bits (Gbps)	Bandwidth at % Efficiency (Gbps) *(1)*
DDR3 SDRAM	533	34.1	23.9
DDR2 SDRAM	400	25.6	17.9
DDR SDRAM	200	12.8	9
RLDRAM II	400	25.6	17.9
QDR SRAM	200	25.6	21.8
QDR II SRAM	350	44.8	38.1
QDR II+ SRAM	350	44.8	38.1

Note to Table 1–1:

(1) 70% for DDR memories, 85% for QDR memories

You must also consider other memory attributes, including how much memory is required (density), how much latency can be tolerated, what is the power budget, and whether the system is cost sensitive. Table 1–2 is an overview of high-speed memories, and details some of the features and target markets of each technology.

Table 1–2. Memory Selection Overview

Parameter	DDR3 SDRAM	DDR2 SDRAM	DDR SDRAM	RLDRAM II	QDR II/+ SRAM
Performance	400–800 MHz	200–400 MHz	100–200 MHz	200–533 MHz	154–350 MHz
Altera-supported data rate	Up to 1066 Mbps	Up to 800 Mbps	Up to 400 Mbps	Up to 2132 Mbps	Up to 1400 Mbps
Density	512 Mbytes–8 Gbytes, 32 Mbytes – 8 Gbytes (DIMM)	256 Mbytes–1 Gbytes, 32 Mbytes – 4 Gbytes (DIMM)	128 Mbytes–1 Gbytes, 32 Mbytes – 2 Gbytes (DIMM)	288 Mbytes, 576 Mbytes	8–72 Mbytes
I/O standard	SSTL-15 Class I, II	SSTL-18 Class I, II	SSTL-2 Class I, II	HSTL-1.8V/1.5V	HSTL-1.8V/1.5V
Data width (bits)	4, 8, 16	4, 8, 16	4, 8, 16, 32	9, 18, 36	8, 9, 18, 36
Burst length	8	4, 8	2, 4, 8	2, 4, 8	2, 4
Number of banks	8	8 (>1 GB), 4	4	8	N/A
Row/column access	Row before column	Row before column	Row before column	Row and column together or multiplexed option	N/A
CAS latency (CL)	5, 6, 7, 8, 9, 10	3, 4, 5	2, 2.5, 3	4, 6, 8	N/A

Table 1–2. Memory Selection Overview

Parameter	DDR3 SDRAM	DDR2 SDRAM	DDR SDRAM	RLDRAM II	QDR II/+ SRAM
Posted CAS additive latency (AL)	0, CL-1, CL-2	0, 1, 2, 3, 4	N/A	N/A	N/A
Read latency (RL)	RL = CL + AL	RL = CL + AL	RL = CL	RL = CL/CL + 1	1.5 clock cycles
On-die termination	Yes	Yes	No	Yes	Yes
Data strobe	Differential bidirectional strobe only	Differential or single-ended bidirectional strobe	Single-ended bidirectional strobe	Free-running differential read and write clocks	Free-running read and write clocks
Refresh requirement	Yes	Yes	Yes	Yes	No
Relative cost comparison	Presently lower than DDR2	Less than DDR SDRAM with market acceptance	Low	Higher than DDR SDRAM, less than SRAM	Highest
Target market	Desktops, servers, storage, LCDs, displays, networking, and communication equipment	Desktops, servers, storage, LCDs, displays, networking, and communication equipment	Desktops, servers, storage, LCDs, displays, networking, and communication equipment	Main memory, cache memory, networking, packet processing, and traffic management	Cache memory, routers, ATM switches, packet memories, lookup, and classification memories

Altera supports these memory interfaces, provides various IP for the physical interface and the controller, and offers many reference designs (refer to Altera's Memory Solutions Center).

For Altera support and the maximum performance for the various high-speed memory interfaces, refer to the *External Memory Interface System Specifications* section in volume 1 of the *External Memory Interface Handbook.*

High-Speed Memory in Embedded Processor Application Example

In embedded processor applications—any system that uses processors, excluding desktop processors—DDR SDRAM is typically used for main memory due to its very low cost, high density, and low power. Next-generation processors invest a large amount of die area to on-chip cache memory to prevent the execution pipelines from sitting idle. Unfortunately, these on-chip caches are limited in size, as a balance of performance, cost, and power must be taken into consideration. In many systems, external memories are used to add another level of cache. In high-performance systems, three levels of cache memory is common: level one (8 Kbytes is common) and level two (512 Kbytes) on chip, and level three off chip (2 Mbytes).

High-end servers, routers, and even video game systems are examples of high-performance embedded products that require memory architectures that are both high speed and low latency. Advanced memory controllers are required to manage transactions between embedded processors and their memories. Altera Arria® series and Stratix series FPGAs optimally implement advanced memory controllers by utilizing their built-in DQS (strobe) phase shift circuitry. Figure 1–1 highlights some of the features available in an Altera FPGA in an embedded application, where DDR2 SDRAM is used as the main memory and QDR II SRAM or RLDRAM II is an external cache level.

**Figure 1–1. Memory Controller Example Using FPGA**
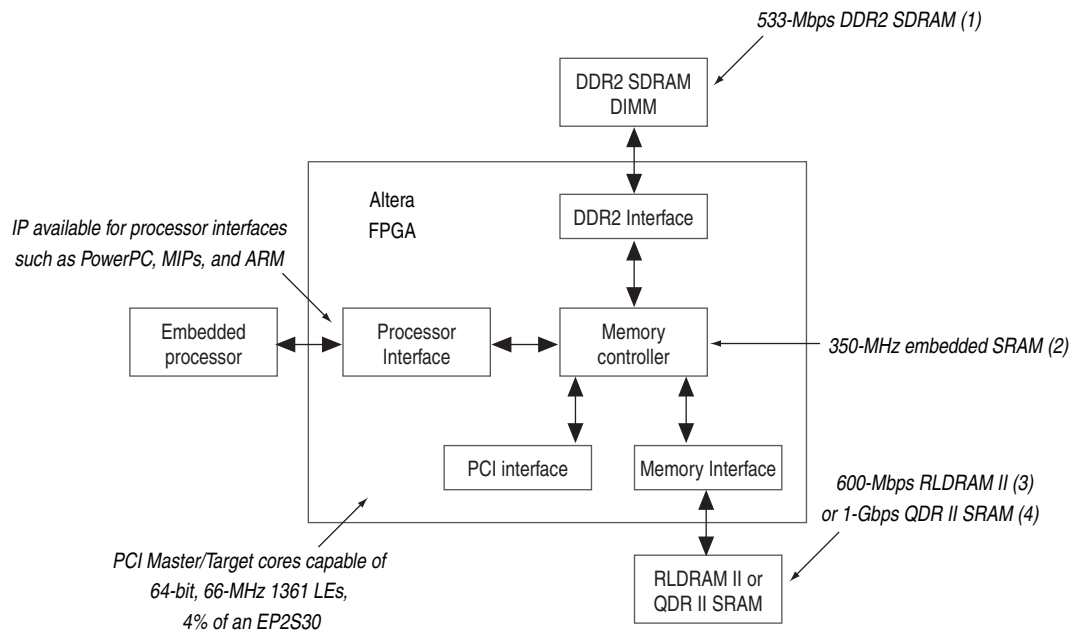
Notes to Figure 1–1:

(1) 533-Mbps DDR2 SDRAM operation using dedicated DQS circuitry, post-amble circuitry, automatic phase shifting, and six registers in the I/O element: 790 LEs, 3% of an EP2S30, and four clock buffers (for a 72-bit interface).

(2) High-speed memory interfaces such as QDR II SRAM require at least four clock buffers to handle all the different clock phases and data directions.

(3) 600-Mbps RLDRAM II operation: 740 logic elements (LEs), 3% of an EP2S30, and four clock buffers (for a 36-bit wide interface).

(4) Embedded SRAM with features such as true-dual port and 350-MHz operation allows complex "store and forward" memory controller architectures.

(5) The Quartus II software reports the number of adaptive look-up tables (ALUTs) that the design uses in the FPGA. The LE count is based on this number of ALUTs.

One of the target markets of RLDRAM II and QDR/QDR II SRAM is external cache memory. RLDRAM II has a read latency close to SSRAM, but with the density of SDRAM. A 16 times increase in external cache density is achievable with one RLDRAM II versus that of SSRAM. In contrast, consider QDR and QDR II SRAM for systems that require high bandwidth and minimal latency. Architecturally, the dual-port nature of QDR and QDR II SRAM allows cache controllers to handle read data and instruction fetches completely independent of writes.

High-Speed Memory in Telecom Application Example

Because telecommunication network architectures are becoming more complex, high-end network systems are running multiple 10-Gbps line cards that connect to multi-shelf switch fabrics scaling to Terabits per second. Figure 1–2 shows an example of a typical system line interface card. These line cards offer interfaces ranging from a single-port OC-192 to multi-port Gigabit Ethernet, and consist of a number of devices, including a PHY/framer, network processors, traffic managers, fabric interface devices, and high-speed memories.

Figure 1–2. Typical Telecom System Line Interface Card
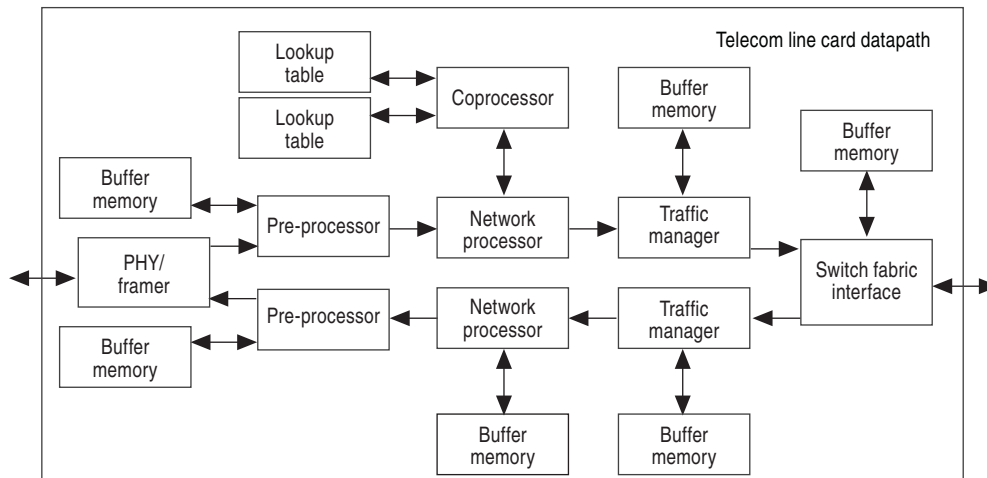
As packets traverse from the PHY/framer device to the switch fabric interface, they are buffered into memories, while the data path devices process headers (determining the destination, classifying packets, and storing statistics for billing) and control the flow of packets into the network to avoid congestion. Typically DDR/DDR2/DDR3 SDRAM and RLDRAM II are used for large buffer memories off network processors, traffic managers, and fabric interfaces, while QDR and QDR II SRAMs are used for look-up tables (LUTs) off preprocessors and coprocessors.

In many designs, FPGAs connect devices together for interoperability and coprocessing, implement features that are not supported by ASIC devices, or implement a device function entirely. Altera Stratix series FPGAs implement traffic management, packet processing, switch fabric interfaces, and coprocessor functions, using features such as 1 Gbps LVDS I/O, high-speed memory interface support, multi-gigabit transceivers, and IP cores. Figure 1–3 highlights some of these features in a packet buffering application where RLDRAM II is used for packet buffer memory and QDR II SRAM is used for control memory.

Figure 1–3. FPGA Example in Packet Buffering Application
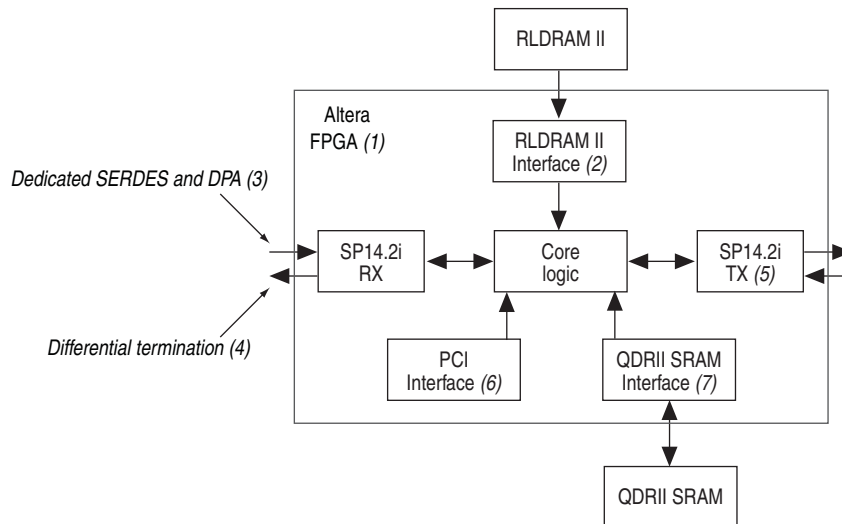
Notes to Figure 1–3:

(1)   As an example, 85% of the LEs still available in an EP2S90.

(2)   600-Mbps RLDRAM II operation: 740 LEs, 1% of an EP2S90, and four clock buffers (for a 36-bit wide interface).

(3)   Dedicated hardware SERDES and DPA circuitry allows clean and reliable implementation of 1-Gbps LVDS.

(4)   Differential termination is built in Stratix FPGAs, simplifying board layout and improving signal quality.

(5)   SPI 4.2i core capable of 1 Gbps: 5178 LEs per Rx, 6087 LEs per Tx, 12% of an ES2S90, and four clock buffers (for both directions using individual buffer mode, 32-bit data path, and 10 logical ports).

(6)   PCI cores capable of 64-bit 66-MHz 656 LEs, 1% of an EP2S90 for a 32-bit target

(7)   1-Gbps QDR II SRAM operation: 100 LEs, 0.1% of an EP2S90, and four clock buffers (for an 18-bit interface).

(8)   Note that the Quartus II software reports the number of ALUTs that the design uses in Stratix II devices. The LE count is based on this number of ALUTs.

SDRAM is usually the best choice for buffering at high data rates due to the large amounts of memory required. Some system designers take a hybrid approach to the memory architecture, using SRAM to store the packet headers and DRAM to store the payload. The depth of the memories depends on the architecture and throughput of the system.

The buffer memory for the packet buffering application of an OC-192 line card (approximately 10 Gbps) must be able to sustain a minimum of one write and one read operation, which requires a memory bandwidth of 20 Gbps to operate at full line rate (more bandwidth is required if the headers are modified). The bandwidth requirement for memory is a key factor in memory selection (see Table 1–1). As an

example, a simple first-order calculation using RLDRAM II as buffer memory requires a bus width of 48 bits to sustain 20 Gbps (300 MHz × 2 DDR × 0.70 efficiency × 48 bits = 20.1 Gbps), which needs two RLDRAM II parts (one ×18 and one ×36). RLDRAM II also inherently includes the additional memory bits used for parity or error correction code (ECC).

QDR and QDR II SRAM have bandwidth and low random access latency advantages that make them useful for control memory in queue management and traffic management applications. Another typical implementation for this memory is billing and packet statistics, where each packet requires counters to be read from memory, incremented, and then rewritten to memory. The high bandwidth, low latency, and optimal one-to-one read/write ratio make QDR SRAM ideal for this feature.

This chapter provides an overview of DDR, DDR2, and DDR3 SDRAM in Altera devices. DDR3 SDRAM is the latest generation of DDR SDRAM technology, with improvements that include lower power consumption, higher data bandwidth, enhanced signal quality with multiple on-die termination (ODT) selection and output driver impedance control. DDR3 SDRAM brings higher memory performance to a broad range of applications, such as PCs, embedded processor systems, image processing, storage, communications, and networking. DDR2 SDRAM is the second generation of DDR SDRAM technology. DDR and DDR2 SDRAMs are available as components and modules, such as DIMMs, SODIMMs, and RDIMMs

For more information about Altera DDR3 SDRAM IP, refer to the *External Memory Interface System Specifications* section in volume 1 of the *External Memory Interface Handbook*.

DDR SDRAM Overview

DDR SDRAM is a 2n prefetch architecture with two data transfers per clock cycle. It uses a single-ended strobe, DQS, which is associated with a group of data pins, DQ, for read and write operations. Both DQS and DQ ports are bidirectional. Address ports are shared for read and write operations.

Write and read operations are sent in bursts, DDR SDRAM supports burst lengths of 2, 4, and 8. The column address strobe (CAS) latency is the latency between when the read command is clocked into the memory and the requested data is presented at the memory pins. DDR SDRAM can have CAS latencies of 2, 2.5, and 3, depending on operating frequency.

DDR SDRAM devices use the SSTL-2 2.5V I/O standard and can hold between 64 MB and 1 GB of data. Each device is divided into four banks, and each bank has a fixed number of rows and columns. Only one row per bank can be accessed at a time. The `ACTIVE` command opens a row and the `PRECHARGE` command closes a row.

DDR SDRAM has a maximum frequency of 200 MHz or 400 Mbps per DQ pin.

DDR2 SDRAM Overview

DDR2 SDRAM is the second generation of the DDR SDRAM standard. It is a 4n prefetch architecture (internally the memory operates at half the interface frequency) with two data transfers per clock cycle. DDR2 SDRAM can use a single-ended or differential strobe, DQS or DQSn, which is associated with a group of data pins, DQ, for read and write operations. DQS, DQSn, and DQ ports are bidirectional. Address ports are shared for read and write operations.

Write and read operations are sent in bursts, DDR2 SDRAM supports burst lengths of 4 and 8. DDR2 SDRAM supports CAS latencies of 2, 3, 4, and 5.

DDR2 SDRAM devices use the SSTL-18 1.8-V I/O standard and can hold between 256 MB and 4 GB of data. All DDR2 SDRAM devices have at least four banks, but higher-density devices (typically 1 GB and above) have eight internal banks. With more banks available, the page-to-hit ratio is twice as great when compared to DDR SDRAM. DDR2 SDRAM also allows bank interleaving, which represents a significant advantage for applications accessing random data. Bank interleaving can be extremely effective for concurrent operations and can hide the timing overhead that are otherwise required for opening and closing individual banks.

DDR2 SDRAM also supports ODT signal options of 50, 75, or 150 Ω on all DQ, DM, and DQS and DQSn signals.

DDR2 SDRAM has a maximum frequency of 533 MHz or 1,066 Mbps per DQ pin.

DDR3 SDRAM Overview

DDR3 SDRAM is more effective at saving system power and increasing system performance than DDR2 SDRAM. DDR3 SDRAM offers lower power by using 1.5 V for the supply and I/O voltage compared to the 1.8-V supply and I/O voltage used by DDR2 SDRAM. DDR3 SDRAM also has better maximum throughput compared to DDR2 SDRAM by increasing the data rate per pin and the number of banks (to eight banks). DDR3 SDRAM also has the following additional benefits:

■ Supports read/write levelling functionality required in the FPGA to interface with DDR3 DIMMs.

■ Supports calibrated parallel ODT via an external resistor RZQ signal termination options of RZQ/2, RZQ/4, or RZQ/6 Ω on all DQ, DM, and DQS and DQSn signals.

■ Supports controlled output driver impedance options of RZQ/6 or RZQ/7.

■ Maximum frequency of 800 MHz or 1600 Mbps per DQ pin.

■ Minimum operating frequency is 300 MHz.

☞ The DDR3 SDRAM high-performance controller only supports local interfaces running at half the rate of the memory interface.

DDR3 SDRAMs are available as components and modules, such as DIMMs, SODIMMs, and RDIMMs.

DDR3 SDRAM is internally configured as an eight-bank DRAM. DDR3 SDRAM uses an 8n prefetch architecture to achieve high-speed operation. The 8n prefetch architecture is combined with an interface that transfers two data words per clock cycle at the I/O pins. A single read or write operation for DDR3 SDRAM 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.

Read and write operations to the DDR3 SDRAM are burst oriented. Operation begins with the registration of an active command, which is then followed by a read or write command. The address bits registered coincident with the active command select the bank and row to be activated (BA0 to BA2 select the bank; A0 to A15 select the row). The address bits registered coincident with the read or write command select the

starting column location for the burst operation, determine if the auto precharge command is to be issued (via A10), and select burst chop (BC) of 4 or burst length (BL) of 8 mode at runtime (via A12), if enabled in the mode register. Before normal operation, the DDR3 SDRAM must be powered up and initialized in a predefined manner.

Differential strobes DQS and DQSn are mandated for DDR3 SDRAM and are associated with a group of data pins, DQ, for read and write operations. DQS, DQSn, and DQ ports are bidirectional. Address ports are shared for read and write operations.

Write and read operations are sent in bursts, DDR3 SDRAM supports BC of 4 and BL of 8. DDR3 SDRAM supports CAS latencies of 5 to 10.

DDR3 SDRAM devices use the SSTL-15 1.5-V I/O standard and can hold between 512 MB and 8 GB of data. The 1.5-V operational voltage reduces power consumption by 17% compared to DDR2 SDRAM.

All DDR3 SDRAM devices have eight internal banks. With more banks available, the page-to-hit ratio is twice that of DDR SDRAM. DDR3 SDRAM also allows bank interleaving, which represents a significant advantage for applications accessing random data. Bank interleaving can be extremely effective for concurrent operations and can hide the timing overhead that is otherwise required for opening and closing individual banks.

DDR, DDR2 and DDR3 SDRAM Comparison

Table 2–1 compares DDR, DDR2, and DDR3 SDRAM features.

Table 2–1. DDR, DDR2, and DDR3 SDRAM Features (Part 1 of 2)

Feature	DDR SDRAM	DDR2 SDRAM	DDR3 SDRAM	DDR3 SDRAM Advantage
Voltage	2.5 V	1.8 V	1.5 V	Reduces memory system power demand by 17%.
Density	64 MB to 1GB	256 MB to 4 GB	512 MB to 8 GB	High-density components simplify memory subsystem.
Internal banks	4	4 and 8	8	Page-to-hit ratio increased.
Prefetch	2	4	8	Lower memory core speed results in higher operating frequency and lower power operation.
Speed	100 to 200 MHz	200 to 533 MHz	300 to 800 MHz	Higher data rate.
Read latency	2, 2.5, 3 clocks	3, 4, 5 clocks	5, 6, 7, 8, 9, 10, and 11	Eliminating half clock setting allows 8n prefetch architecture.
Additive latency *(1)*	—	0, 1, 2, 3, 4	0, CL1, or CL2	Improves command efficiency.
Write latency	One clock	Read latency − 1	5, 6, 7, or 8	Improves command efficiency.
Termination	PCB, discrete to V_{TT}	Discrete to V_{TT} or ODT	Discrete to V_{TT} or ODT parallel termination. Controlled impedance output.	Improves signaling, eases PCB layout, reduces system cost.
Data strobes	Single-ended	Differential or single-ended	Differential mandated	Improves timing margin.

Table 2–1. DDR, DDR2, and DDR3 SDRAM Features (Part 2 of 2)

Feature	DDR SDRAM	DDR2 SDRAM	DDR3 SDRAM	DDR3 SDRAM Advantage
Clock, address, and command (CAC) layout	Balanced tree	Balanced tree	Series or daisy chained	The DDR3 SDRAM read and write leveling feature allows for a much simplified PCB and DIMM layout. You can still optionally use the balanced tree topology by using the DDR3 without the leveling option.

Note to Table 2–1:

(1) The Altera DDR and DDR2 SDRAM high-performance controllers do not support additive latency, but the high-performance controller II does.

DDR, DDR2, and DDR3 SDRAM Interface Pins

This section describes the DDR, DDR2, and DDR3 SDRAM interface pins.

Clock Signals

DDR, DDR2, and DDR3 SDRAM devices use CK and CK# signals to clock the address and command signals into the memory. Furthermore, the memory uses these clock signals to generate the DQS signal during a read through the DLL inside the memory. The SDRAM data sheet specifies the following timings:

- t_{DQSCK} is the skew between the CK or CK# signals and the SDRAM-generated DQS signal

- t_{DSH} is the DQS falling edge from CK rising edge hold time

- t_{DSS} is the DQS falling edge from CK rising edge setup time

- t_{DQSS} is the positive DQS latching edge to CK rising edge

These SDRAM have a write requirement (t_{DQSS}) that states the positive edge of the DQS signal on writes must be within ± 25% (± 90°) of the positive edge of the SDRAM clock input. Therefore, you should generate the CK and CK# signals using the DDR registers in the IOE to match with the DQS signal and reduce any variations across process, voltage, and temperature. The positive edge of the SDRAM clock, CK, is aligned with the DQS write to satisfy t_{DQSS}.

☞ The Altera SDRAM high-performance controllers generate the CK and CK# signals using the DDR registers in the IOE with the DQS signal and reduce any variations across process, voltage, and temperature.

DDR3 SDRAM can use a daisy-chained CAC topology, the memory clock must arrive at each chip at a different time. To compensate for this flight-time skew between devices across a typical DIMM, write leveling must be employed.

Data, Data Strobes, DM, and Optional ECC Signals

DDR SDRAM uses bidirectional single-ended data strobe (DQS); DDR3 SDRAM uses bidirectional differential data strobes. The DQSn pins in DDR2 SDRAM devices are optional but recommended for DDR2 SDRAM designs operating at more than 333 MHz. Differential DQS operation enables improved system timing due to reduced crosstalk and less simultaneous switching noise on the strobe output drivers. The DQ

pins are also bidirectional. Regardless of interface width, DDR SDRAM always operates in ×8 mode DQS groups. DQ pins in DDR2 and DDR3 SDRAM interfaces can operate in either ×4 or ×8 mode DQS groups, depending on your chosen memory device or DIMM, regardless of interface width. The ×4 and ×8 configurations use one pair of bidirectional data strobe signals, DQS and DQSn, to capture input data. However, two pairs of data strobes, UDQS and UDQS# (upper byte) and LDQS and LDQS# (lower byte), are required by the ×16 configuration devices. A group of DQ pins must remain associated with its respective DQS and DQSn pins.

The DQ signals are edge-aligned with the DQS signal during a read from the memory and are center-aligned with the DQS signal during a write to the memory. The memory controller shifts the DQ signals by –90° during a write operation to center align the DQ and DQS signals. ALTMEMPHY delays the DQS signal during a read, so that the DQ and DQS signals are center aligned at the capture register. Altera devices use a phase-locked loop (PLL) to center-align the DQS signal with respect to the DQ signals during writes and Altera devices (except Cyclone III devices) use dedicated DQS phase-shift circuitry to shift the incoming DQS signal during reads. Figure 2–1 shows an example where the DQS signal is shifted by 90° for a read from the DDR2 SDRAM.

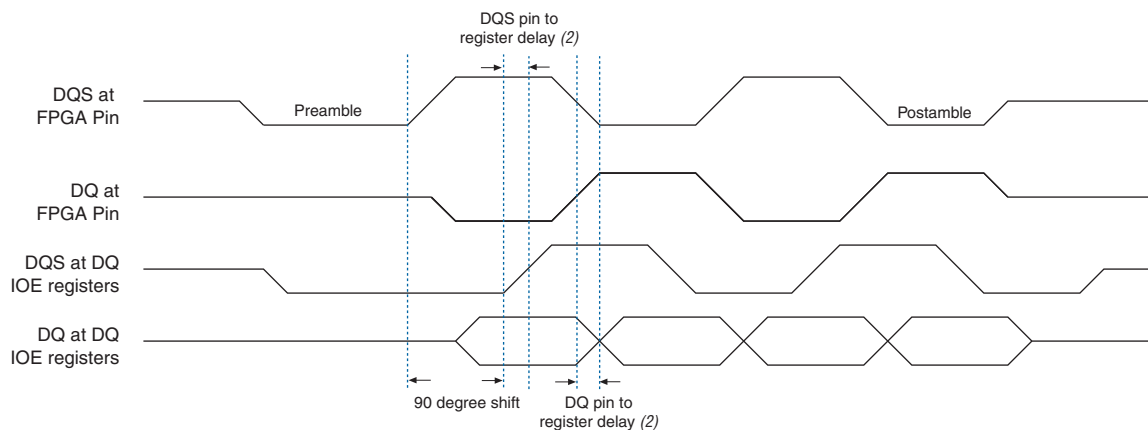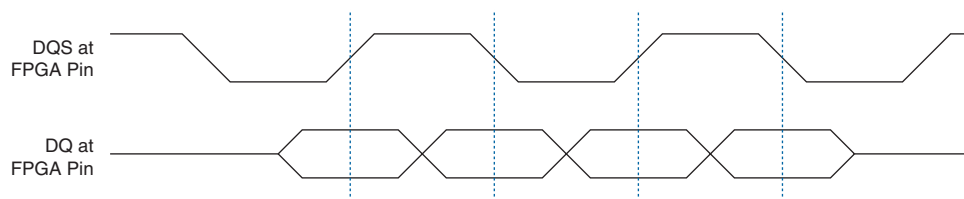**Figure 2–1. DQ and DQS Relationship During a DDR2 SDRAM Read in Burst-of-Four Mode**

Figure 2–2 shows an example of the relationship between the data and data strobe during a burst-of-four write.

Figure 2–2. DQ and DQS Relationship During a DDR2 SDRAM Write in Burst-of-Four Mode

The memory device's setup (t_{DS}) and hold times (t_{DH}) for the write DQ and DM pins are relative to the edges of DQS write signals and not the CK or CK# clock. Setup and hold requirements are not necessarily balanced in in DDR2 and DDR3 SDRAM, unlike in DDR SDRAM devices.

The DQS signal is generated on the positive edge of the system clock to meet the t_{DQSS} requirement. DQ and DM signals use a clock shifted –90° from the system clock, so that the DQS edges are centered on the DQ or DM signals when they arrive at the DDR2 SDRAM. The DQS, DQ, and DM board trace lengths need to be tightly matched (20 ps).

The SDRAM uses the DM pins during a write operation. Driving the DM pins low shows that the write is valid. The memory masks the DQ signals if the DM pins are driven high. While you can use any of the I/O pins in the same bank as the associated DQS and DQ pins, to generate the DM signal, Altera recommends that you use the spare DQ pin within the same DQS group as the respective data, to minimize skew.

The DM signal's timing requirements at the SDRAM input are identical to those for DQ data. The DDR registers, clocked by the –90° shifted clock, create the DM signals.

Some SDRAM modules support error correction coding (ECC) to allow the controller to detect and automatically correct error in data transmission. The 72-bit SDRAM modules contain eight extra data pins in addition to 64 data pins. The eight extra ECC pins should be connected to a single DQS or DQ group on the FPGA.

Address and Command Signals

Address and command signals in SDRAM devices are clocked into the memory device using the CK or CK# signal. These pins operate at single data rate (SDR) using only one clock edge. The number of address pins depends on the SDRAM device capacity. The address pins are multiplexed, so two clock cycles are required to send the row, column, and bank address. The CS#, RAS, CAS, WE, CKE, and ODT pins are SDRAM command and control pins. DDR3 SDRAM has additional pins: `RESET#`, `PAR_In` and `ERR_OUT#`. The `RESET#` pin uses 1.5-V LVCMOS I/O standard, while the rest of the DDR3 SDRAM pins use the SSTL-15 I/O standard.

The DDR2 SDRAM address and command inputs do not have a symmetrical setup and hold time requirement with respect to the SDRAM clocks, CK, and CK#.

For ALTMEMPHY or Altera SDRAM high-performance controllers in Stratix III and Stratix IV devices, the address and command clock is a dedicated PLL clock output whose phase can be adjusted to meet the setup and hold requirements of the memory clock. The address and command clock is also typically half-rate, although a full-rate implementation can also be created. The command and address pins use the DDIO output circuitry to launch commands from either the rising or falling edges of the clock. The chip select (`mem_cs_n`), clock enable (`mem_cke`), and ODT (`mem_odt`) pins are only enabled for one memory clock cycle and can be launched from either the rising or falling edge of the address and command clock signal. The address and other command pins are enabled for two memory clock cycles and can also be launched from either the rising or falling edge of the address and command clock signal.

☞ In ALTMEMPHY-based designs, the address and command clock `ac_clk_1x` is always half rate. However, because of the output enable assertion, CS#, CKE, and ODT behave like full-rate signals even in a half-rate PHY.

In Arria II GX and Cyclone III devices, the address and command clock is either shared with the `write_clk_2x` or the `mem_clk_2x` clock.

DIMM Options

Compared to the unbuffered DIMMs (UDIMM), both single-rank and double-rank registered DIMMs (RDIMM) use only one pair of clocks and two chip selects `CS#[1:0]` in DDR3. An RDIMM has extra parity signals for address, `RAS#`, `CAS#`, and `WE#`.

Dual-rank DIMMs have the following extra signals for each side of the DIMM:

■ `CS#` (RDIMM always has two chip selects, DDR3 uses a minimum of 2 chip selects, even on a single rank module)

■ `CK` (only UDIMM)

■ `ODT` signal

■ `CKE` signal

Table 2–2 compares the UDIMM and RDIMM pin options.

Table 2–2. UDIMM and RDIMM Pin Options

Pins	UDIMM Pins (Single Rank)	UDIMM Pins (Dual Rank)	RDIMM Pins (Single Rank)	RDIMM Pins (Dual Rank)
Data	72 bit `DQ[71:0] = {CB[7:0],DQ[63:0]}`	72 bit `DQ[71:0] = {CB[7:0],DQ[63:0]}`	72 bit `DQ[71:0] = {CB[7:0],DQ[63:0]}`	72 bit `DQ[71:0]= {CB[7:0],DQ[63:0]}`
Data Mask	`DM[8:0]`	`DM[8.0]`	`DM[8.0]`	`DM[8.0]`
Data Strobe *(1)*	`DQS[8:0]` and `DQS#[8:0]`	`DQS[8:0]` and `DQS#[8:0]`	`DQS[8:0]` and `DQS#[8:0]`	`DQS[8:0]` and `DQS#[8:0]`
Address	`BA[2:0],A[15:0]−` 2 GB: `A[13:0]` 4 GB: `A[14:0]` 8 GB: `A[15:0]`	`BA[2:0],A[15:0]−` 2 GB: `A[13:0]` 4 GB: `A[14:0]` 8 GB: `A[15:0]`	`BA[2:0],A[15:0]−2` GB: `A[13:0]` 4 GB: `A[14:0]` 8 GB: `A[15:0]`	`BA[2:0],A[15:0]−` 2 GB: `A[13:0]` 4 GB: `A[14:0]` 8 GB: `A[15:0]`
Clock	`CK0/CK0#`	`CK0/CK0#, CK1/CK1#`	`CK0/CK0#`	`CK0/CK0#`
Command	`ODT, CS#, CKE, RAS#, CAS#, WE#`	`ODT[1:0], CS#[1:0], CKE[1:0], RAS#, CAS#, WE#`	`ODT, CS#[1:0], CKE, RAS#, CAS#, WE#`	`ODT[1:0], CS#[1:0], CKE[1:0], RAS#, CAS#, WE#`
Parity	—	—	`PAR_IN, ERR_OUT`	`PAR_IN, ERR_OUT`
Other Pins	`SA[2:0], SDA, SCL, EVENT#, RESET#`	`SA[2:0], SDA, SCL, EVENT#, RESET#`	`SA[2:0], SDA, SCL, EVENT#, RESET#`	`SA[2:0], SDA, SCL, EVENT#, RESET#`

Note to Table 2–2:

(1) `DQS#[8:0]` is optional in DDR2 SDRAM and is not supported in DDR SDRAM interfaces.

This chapter provides an overview of QDR II and QDR II+ SRAM in Altera devices

Synchronous static RAM (SRAM) architectures support the high throughput requirements of communications, networking, and digital signal processing (DSP) systems. The successor to quad data rate (QDR) SRAM, QDR II+ and QDR II SRAM support higher memory bandwidth and improved timing margins and offer more flexibility in system designs.

For more information about Altera QDR II and QDR II+ SRAM IP, refer to the *External Memory Interface System Specifications* section in volume 1 of the *External Memory Interface Handbook*.

QDR II+ and QDR II SRAM can perform two data writes and two data reads per clock cycle. They use one port for writing data (D) and another port for reading data (Q). These unidirectional data ports support simultaneous reads and writes and allows back-to-back transactions without the contention issues that can occur when using a single bidirectional data bus. Write and read operations share address ports.

The QDR II SRAM devices are available in ×8, ×9, ×18, and ×36 data bus width configurations. The QDR II+ SRAM devices are available in ×9, ×18, and ×36 data bus width configurations. Write and read operations are burst-oriented. All the data bus width configurations of QDR II SRAM support burst lengths of two and four. QDR II+ SRAM supports only a burst length of four. Burst-of-two and burst-of-four for QDR II and burst-of-four for QDR II+ SRAM devices provide the same overall bandwidth at a given clock speed.

Read latency is the time between the read command being clocked into memory and the time data is presented at the memory pins. For QDR II SRAM devices, the read latency is 1.5 clock cycles, while for QDR II+ SRAM devices it is 2 or 2.5 clock cycles, depending on the memory device. Write latency is the time between the write command being clocked into memory and the time data is presented at the memory pins. For QDR II+ and burst-of-four QDR II SRAM devices, the write commands and addresses are clocked on the rising edge of clock and write latency is one clock cycle. For burst-of-two QDR II SRAM devices, the write command is clocked on the rising edge of clock and the write address is clocked on the falling edge of clock. Therefore, the write latency is zero, because the write data is presented at the same time as the write command.

Altera supports both 1.5-V and 1.8-V HSTL I/O standards for QDR II+ and QDR II SRAM interfaces. QDR II+ and QDR II SRAM interfaces use a delay-locked loop (DLL) inside the device to edge-align the data with respect to the K and Kn or C and Cn pins. You can optionally turn off the DLL, but the performance of the QDR II+ and QDR II SRAM devices is degraded. All timing specifications listed in this document assume that the DLL is on.

QDR II+ and QDR II SRAM devices also offer programmable impedance output buffers. You can set the buffers by terminating the ZQ pin to V_{SS} through a resistor, RQ. The value of RQ should be five times the desired output impedance. The range for RQ should be between 175 Ω and 350 Ω with a tolerance of 10%.

QDR II+ and QDR II SRAM Interface Pin Description

This section provides a description of the clock, control, address, and the data signals on QDR II and QDR II+ SRAM devices.

Clock Signals

QDR II+ and QDR II SRAM devices have three pairs of clocks:

■ Input clocks K and K#

■ Input clocks C and C#

■ Echo clocks CQ and CQ#

The positive input clock, K, is the logical complement of the negative input clock, K#. Similarly, C and CQ are complements of C# and CQ#, respectively. With these complementary clocks, the rising edges of each clock leg latch the DDR data.

The QDR II+ and QDR II SRAM devices use the K and K# clocks for write access and the C and C# clocks for read accesses only when interfacing more than one QDR II+ or QDR II SRAM device. Because the number of loads that the K and K# clocks drive affects the switching times of these outputs when a controller drives a single QDR II+ or QDR II SRAM device, C and C# are unnecessary. This is because the propagation delays from the controller to the QDR II+ or QDR II SRAM device and back are the same. Therefore, to reduce the number of loads on the clock traces, QDR II+ and QDR II SRAM devices have a single clock mode, and the K and K# clocks are used for both reads and writes. In this mode, the C and C# clocks are tied to the supply voltage (V_{DD}).

CQ and CQ# are the source-synchronous output clocks from the QDR II or QDR II+ SRAM device that accompanies the read data.

The Altera device outputs the K and K# clocks, data, address, and command lines to the QDR II+ or QDR II SRAM device. For the controller to operate properly, the write data (D), address (A), and control signal trace lengths (and therefore the propagation times) should be equal to the K and K# clock trace lengths.

You can generate C, C#, K, and K# clocks using any of the PLL registers via the DDR registers. Because of strict skew requirements between K and K# signals, use adjacent pins to generate the clock pair. The propagation delays for K and K# from the FPGA to the QDR II+ or QDR II SRAM device are equal to the delays on the data and address (D, A) signals. Therefore, the signal skew effect on the write and read request operations is minimized by using identical DDR output circuits to generate clock and data inputs to the memory.

Command Signals

QDR II+ and QDR II SRAM devices use the write port select (WPSn) signal to control write operations and the read port select (RPSn) signal to control read operations. The byte write select signal (BWSn) is a third control signal that indicates to the QDR II+ or QDR II SRAM device which byte to write into the QDR II+ or QDR II SRAM device. You can use any of the FPGA's user I/O pins to generate control signals, preferably on the same side and the same bank.

Address Signals

QDR II+ and QDR II SRAM devices use one address bus (A) for both read and write addresses. You can use any of the FPGA's user I/O pins to generate address signals, preferably on the same side and the same banks.

Data and QVLD Signals

QDR II+ and QDR II SRAM devices use two unidirectional data buses: one for writes (D) and one for reads (Q). The read data is edge-aligned with the CQ and CQ# clocks while the write data is center-aligned with the K and K# clocks (see Figure 3–1 and Figure 3–2).

Figure 3–1. CQ and Q Relationship During QDR II+ SRAM Read
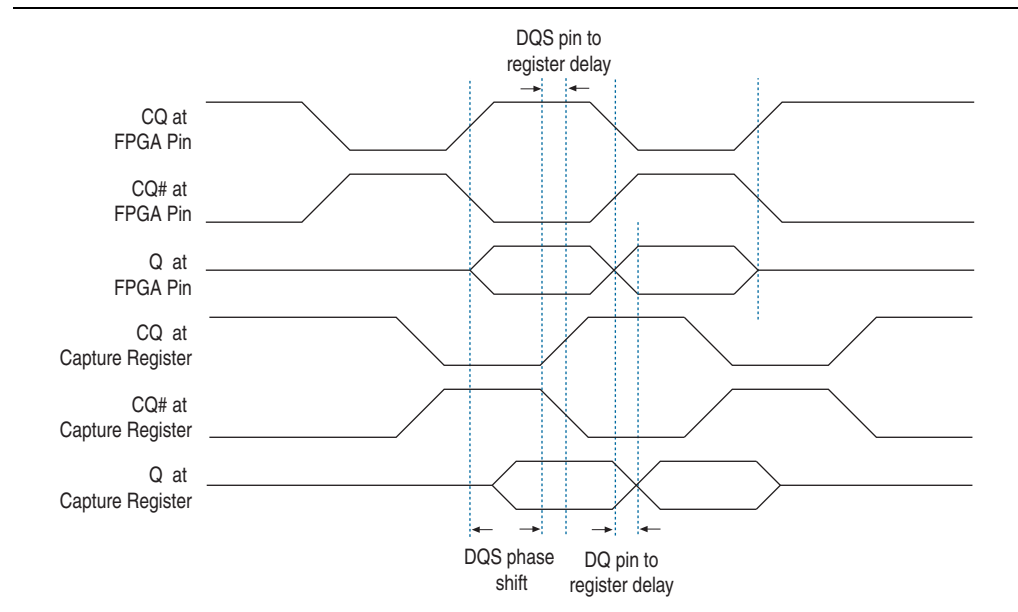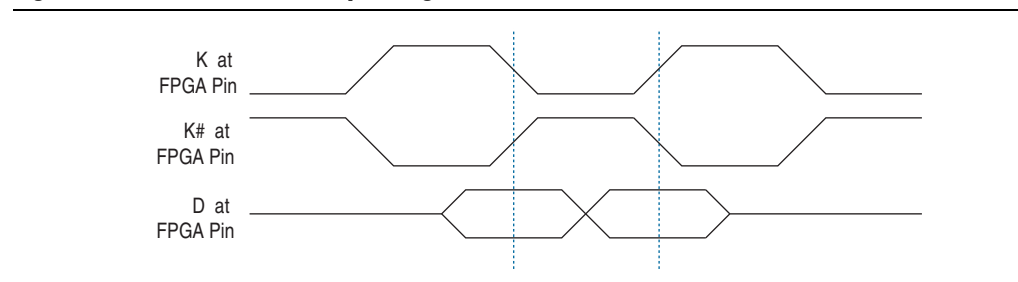

Figure 3–2. K and D Relationship During QDR II+ SRAM Write

QDR II+ SRAM devices also have a QVLD pin that indicates valid read data. The QVLD signal is edge-aligned with the echo clock and is asserted high for approximately half a clock cycle before data is output from memory.

☞ The Altera QDR II+ SRAM Controller MegaCore function does not use the QVLD signal.

This chapter provides an overview of RLDRAM II in Altera devices

Reduced latency DRAM II (RLDRAM II) is a DRAM-based point-to-point memory device designed for communications, imaging, and server systems requiring high density, high memory bandwidth, and low latency. The fast random access speeds in RLDRAM II devices make them a viable alternative to SRAM devices at a lower cost.

For more information about Altera RLDRAM II IP, refer to the *External Memory Interface System Specifications* section in volume 1 of the *External Memory Interface Handbook*.

There are two types of RLDRAM II devices: common I/O (CIO) and separate I/O (SIO). CIO devices share a single data I/O bus which is similar to the double data rate (DDR) SDRAM interface. SIO devices, with separate data read and write buses, have an interface similar to SRAM.

Compared to DDR SDRAM, RLDRAM II has simpler bank management and lower latency inside the memory. RLDRAM II devices are divided into eight banks instead of the typical four banks in most memory devices, providing a more efficient data flow within the device. RLDRAM II offers up to 2.4 Gigabytes per second (Gbps) aggregate bandwidth.

RLDRAM II uses a DDR scheme, performing two data transfers per clock cycle. RLDRAM II CIO devices use the bidirectional data pins (DQ) for both read and write data, while RLDRAM II SIO devices use D pins for write data (input to the memory) and Q pins for read data (output from the memory). Both types use two pairs of uni-directional free-running clocks. The memory uses DK and DK# pins during write operations, and generates QK and QK# pins during read operations. In addition, RLDRAM II uses the system clocks (CK and CK# pins) to sample commands and addresses and generate the QK and QK# read clocks. Address ports are shared for write and read operations.

The RLDRAM II SIO devices are available in ×9 and ×18 data bus width configurations, while the RLDRAM II CIO devices are available in ×9, ×18, and ×36 data bus width configurations. RLDRAM II CIO interfaces may require an extra cycle for bus turnaround time for switching read and write operations.

Write and read operations are burst oriented and all the data bus width configurations of RLDRAM II support burst lengths of two and four. In addition, RLDRAM II devices with data bus width configurations of ×9 and ×18 also support burst length of eight.

The read latency is the time between when the read command is clocked into the memory and the time data is presented at the memory pins. There is a similar latency for write operations called the write latency. The write latency is equal to the read latency plus one clock cycle. The RLDRAM devices have up to three programmable configuration settings that determine the row cycle times, read latency, and write latency of the interface at a given frequency of operation.

RLDRAM II devices use either the 1.5-V HSTL or 1.8-V HSTL I/O standard. You can use either I/O standard to interface with Altera FPGAs. Each RLDRAM II device is divided into eight banks, where each bank has a fixed number of rows and columns. Only one row per bank is accessed at a time. The memory (instead of the controller) controls the opening and closing of a row, which is similar to an SRAM interface.

RLDRAM II also offers programmable impedance output buffers and on-die termination. The programmable impedance output buffers are for impedance matching and are guaranteed to produce 25- to 60-ohm output impedance. The on-die termination is dynamically switched on during read operations and switched off during write operations. Perform an IBIS simulation to observe the effects of this dynamic termination on your system. IBIS simulation can also show the effects of different drive strengths, termination resistors, and capacitive loads on your system.

RLDRAM II Interface Pin Description

This section describes the RLDRAM II interface pin description.

Clock Signals

RLDRAM II devices use `CK` and `CK#` signals to clock the command and address bus in single data rate (SDR). There is one pair of `CK` and `CK#` pins per RLDRAM II device.

Instead of a strobe, RLDRAM II devices use two sets of free-running differential clocks to accompany the data. The `DK` and `DK#` clocks are the differential input data clocks used during writes while the `QK` or `QK#` clocks are the output data clocks used during reads. Even though `QK` and `QK#` signals are not differential signals according to the RLDRAM II data sheets, Micron treats these signals as such for their testing and characterization. Each pair of `DK` and `DK#`, or `QK` and `QK#` clocks are associated with either 9 or 18 data bits.

The exact clock-data relationships are as follows:

- For ×36 data bus width configuration, there are 18 data bits associated with each pair of write and read clocks. So, there are two pairs of `DK` and `DK#` pins and two pairs of `QK` or `QK#` pins.

- For ×18 data bus width configuration, there are 18 data bits per one pair of write clocks and nine data bits per one pair of read clocks. So, there is one pair of `DK` and `DK#` pins, but there are two pairs of `QK` and `QK#` pins.

- For ×9 data bus width configuration, there are nine data bits associated with each pair of write and read clocks. So, there is one pair of `DK` and `DK#` pins and one pair of `QK` and `QK#` pins each.

There are t_{CKDK} timing requirements for skew between `CK` and `DK` or `CK#` and `DK#`.

Because of the loads on these I/O pins, the maximum frequency you can achieve depends on the number of RLDRAM II devices you are connecting to the Altera device. Perform SPICE or IBIS simulations to analyze the loading effects of the pin-pair on multiple RLDRAM II devices.

Data, DM and QVLD Signals

The read data is edge-aligned with the QK or QK# clocks while the write data is center-aligned with the DK and DK# clocks (see Figure 4–1 and Figure 4–2). The memory controller shifts the DK or DK# signal to center align the DQ and DK or DK# signal during a write and to shift the QK signal during a read, so that read data (DQ or Q signals) and QK clock is center-aligned at the capture register. Altera devices use dedicated DQS phase-shift circuitry to shift the incoming QK signal during reads and use a PLL to center-align the DK and DK# signals with respect to the DQ signals during writes.
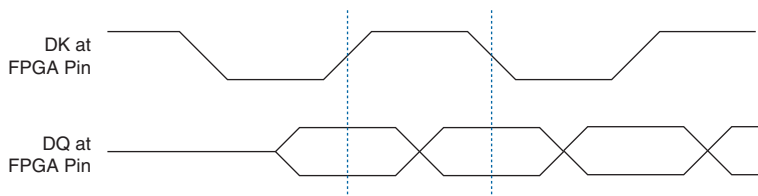
Figure 4–1. DQ and QK Relationship During RLDRAM II Read
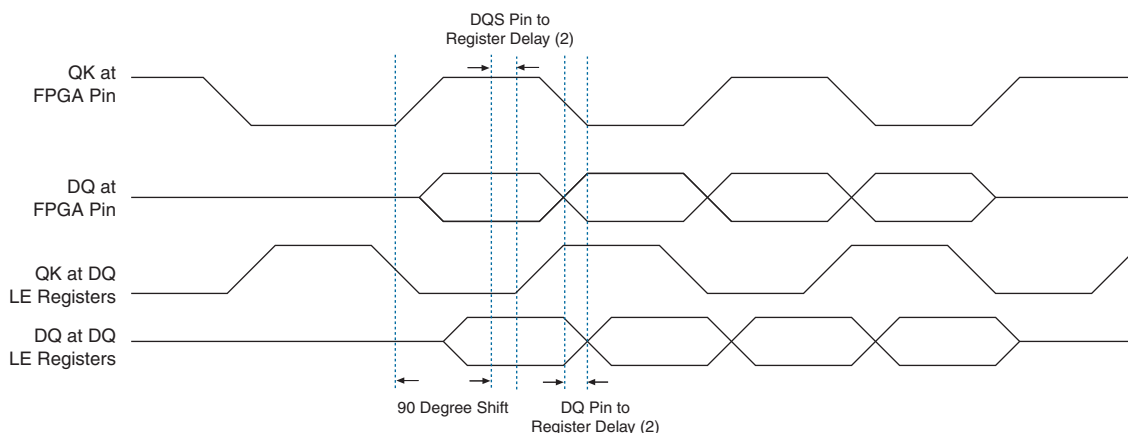
Notes to Figure 4–1:

(1) This is an example of a 90° shift. The required phase shift for your system should be based on your timing analysis and may not be 90°.

Figure 4–2. DQ and QK Relationship During RLDRAM II Write

The RLDRAM II data mask (DM) pins are only used during a write. The memory controller drives the DM signal low when the write is valid and drives it high to mask the DQ signals. There is one DM pin per RLDRAM II device.

The DM timing requirements at the input to the RLDRAM II are identical to those for DQ data. The DDR registers, clocked by the write clock, create the DM signals. This reduces any skew between the DQ and DM signals.

The RLDRAM II device's setup time (t_{DS}) and hold (t_{DH}) time for the write DQ and DM pins are relative to the edges of the DK or DK# clocks. The DK and DK# signals are generated on the positive edge of system clock, so that the positive edge of CK or CK# is aligned with the positive edge of DK or DK# respectively to meet the RLDRAM II tCKDK requirement. The DQ and DM signals are clocked using a shifted clock so that the edges of DK or DK# are center-aligned with respect to the DQ and DM signals when they arrive at the RLDRAM II device.

The clocks, data, and DM board trace lengths should be tightly matched to minimize the skew in the arrival time of these signals.

RLDRAM II devices also have a QVLD pin indicating valid read data. The QVLD signal is edge-aligned with QK or QK# and is high approximately half a clock cycle before data is output from the memory.

☞ The Altera RLDRAM II Controller IP does not use the QVLD signal.

Commands and Addresses

The CK and CK# signals clock the commands and addresses into RLDRAM II devices. These pins operate at single data rate using only one clock edge. RLDRAM II devices have 18 to 21 address pins, depending on the data bus width configuration and burst length. RLDRAM II supports both non-multiplexed and multiplexed addressing. Multiplexed addressing allows you to save a few user I/O pins while non-multiplexed addressing allows you to send the address signal within one clock cycle instead of two clock cycles. CS#, REF#, and WE# pins are input commands to the RLDRAM II device.

The commands and addresses must meet the memory address and command setup (t_{AS}, t_{CS}) and hold (t_{AH}, t_{CH}) time requirements.

This chapter provides additional information about the document and Altera.

Revision History

The following table shows the revision history for this section.

Date	Version	Changes
December 2010	2.2	Updated for 10.1 release.
July 2010	2.1	Updated for 10.0 release.
January 2010	1.1	Corrected minor typos.
November 2009	1.0	Initial release.

How to Contact Altera

To locate the most up-to-date information about Altera products, refer to the following table.

Contact (1)	Contact Method	Address
Technical support	Website	www.altera.com/support
Technical training	Website	www.altera.com/training
	Email	custrain@altera.com
Product literature	Website	www.altera.com/literature
Non-technical support (General)	Email	nacomp@altera.com
(Software Licensing)	Email	authorization@altera.com

Note to Table:

(1)   You can also contact your local Altera sales office or sales representative.

Typographic Conventions

The following table shows the typographic conventions this document uses.

Visual Cue	Meaning
Bold Type with Initial Capital Letters	Indicate command names, dialog box titles, dialog box options, and other GUI labels. For example, **Save As** dialog box. For GUI elements, capitalization matches the GUI.
bold type	Indicates directory names, project names, disk drive names, file names, file name extensions, software utility names, and GUI labels. For example, **\qdesigns** directory, **D:** drive, and **chiptrip.gdf** file.
Italic Type with Initial Capital Letters	Indicate document titles. For example, *Stratix IV Design Guidelines*.

Visual Cue	Meaning
italic type	Indicates variables. For example, *n* + 1. Variable names are enclosed in angle brackets (< >). For example, *<file name>* and *<project name>*.**pof** file.
Initial Capital Letters	Indicate keyboard keys and menu names. For example, the Delete key and the Options menu.
"Subheading Title"	Quotation marks indicate references to sections within a document and titles of Quartus II Help topics. For example, "Typographic Conventions."
`Courier type`	Indicates signal, port, register, bit, block, and primitive names. For example, `data1`, `tdi`, and `input`. The suffix `n` denotes an active-low signal. For example, `resetn`. Indicates command line commands and anything that must be typed exactly as it appears. For example, `c:\qdesigns\tutorial\chiptrip.gdf`. Also indicates sections of an actual file, such as a Report File, references to parts of files (for example, the AHDL keyword `SUBDESIGN`), and logic function names (for example, `TRI`).
↵	An angled arrow instructs you to press the Enter key.
1., 2., 3., and a., b., c., and so on	Numbered steps indicate a list of items when the sequence of the items is important, such as the steps listed in a procedure.
■ ■ ■	Bullets indicate a list of items when the sequence of the items is not important.
☞	The hand points to information that requires special attention.
?	A question mark directs you to a software help system with related information.
👣	The feet direct you to another document or website with related information.
⚠ CAUTION	A caution calls attention to a condition or possible situation that can damage or destroy the product or your work.
⚠ WARNING	A warning calls attention to a condition or possible situation that can cause you injury.
✉	The envelope links to the Email Subscription Management Center page of the Altera website, where you can sign up to receive update notifications for Altera documents.