

# Selecting the Right High-Speed Memory Technology for Your System

## Introduction

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).

This white paper 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.

## 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 and trade-offs 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. [Table 1](#) gives a general overview of the memory technologies discussed in this paper.

Table 1. Overview of Memory Interface Technologies

Memory	Bandwidth	Density	Latency	Power	Cost
DDR3 SDRAM	↑	↑	↑	↓	↓
DDR2 SDRAM	↑	↑	↑	↓	↓
DDR SDRAM	↑	↑	↑	↓	↓
RLDRAM II	↓	↑	↑	↓	↓
QDR SRAM	↓	↑	↑	↓	↓
QDRII SRAM	↓	↑	↑	↓	↓
QDRII+ SRAM	↓	↑	↑	↓	↓

*Note:*

(1) Arrows indicate approximate increasing value.

### DDR/DDR2/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/RLDRAM II*

Reduced latency DRAM (RLDRAM) is optimized to reduce latency primarily for networking and cache applications. In DDR SDRAM, the memory is partitioned into four banks, while RLDRAM is partitioned into eight smaller banks. This 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/QDRII/QDRII+ SRAM*

SRAMs are fundamentally different from DRAMs in that a typical SRAM memory cell consists of six transistors arranged to form a flipflop, 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. QDRII 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. QDRII+ SDRAM is the latest generation for this family and is faster again.

## 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 2](#) 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® III FPGA. The third column in this table includes a conservative DRAM memory bandwidth at 70 percent efficiency, which takes into consideration bus turnaround, refresh, burst length, and random access latency. For QDR and QDRII SRAM, 85 percent efficiency is used.

*Table 2. Memory Bandwidth for 32-bit Wide Data Bus in Stratix III FPGA*

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

*Note:*

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

Other memory attributes also must be considered, 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 3](#) is an overview of high-speed memories, and details some of the features and target markets of each technology.

Table 3. Memory Selection Overview

Parameter	DDR3 SDRAM	DDR2 SDRAM	DDR SDRAM	RLDRAM II	QDRII/+ 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
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 higher than DDR2	Less than DDR SDRAM with market acceptance	Lowest	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 (see Altera's [Memory Solutions Center](#)). Table 4 shows Altera's support and speeds for the various high-speed memory interfaces.

Table 4. Altera External Memory Interface Support (1)

Device	DDR3 SDRAM	DDR2 SDRAM	DDR SDRAM	RLDRAM II	QDRII/+ SRAM
Stratix IV	1,067 Mbps 533 MHz	800 Mbps 400 MHz	400 Mbps 200 MHz	1,600 Mbps 400 MHz	1,400 Mbps 350 MHz
Stratix III	1,067 Mbps 533 MHz	800 Mbps 400 MHz	400 Mbps 200 MHz	1,600 Mbps 400 MHz	1,400 Mbps 350 MHz
Stratix II/GX		667 Mbps 333 MHz	400 Mbps 200 MHz	1,200 Mbps 300 MHz	1,200 Mbps 300 MHz
HardCopy® IV	800 Mbps 400 MHz	667 Mbps 333 MHz	400 Mbps 200 MHz	1,600 Mbps 400 MHz	1,400 Mbps 350 MHz
HardCopy III	800 Mbps 400 MHz	667 Mbps 333 MHz	400 Mbps 200 MHz	1,600 Mbps 400 MHz	1,400 Mbps 350 MHz
HardCopy II		533 Mbps 267 MHz	400 Mbps 200 MHz	1,000 Mbps 250 MHz	1,000 Mbps 250 MHz
Stratix and Stratix GX			400 Mbps 200 MHz	400 Mbps 200 MHz	800 Mbps 200 MHz
Cyclone® III		333 Mbps 167 MHz	333 Mbps 167 MHz		333 Mbps 167 MHz (2)
Cyclone II		400 Mbps 200 MHz	333 Mbps 167 MHz		333 Mbps 167 MHz (2)
Arria® GX		466 Mbps 233 MHz	400 Mbps 200 MHz		

**Notes:**

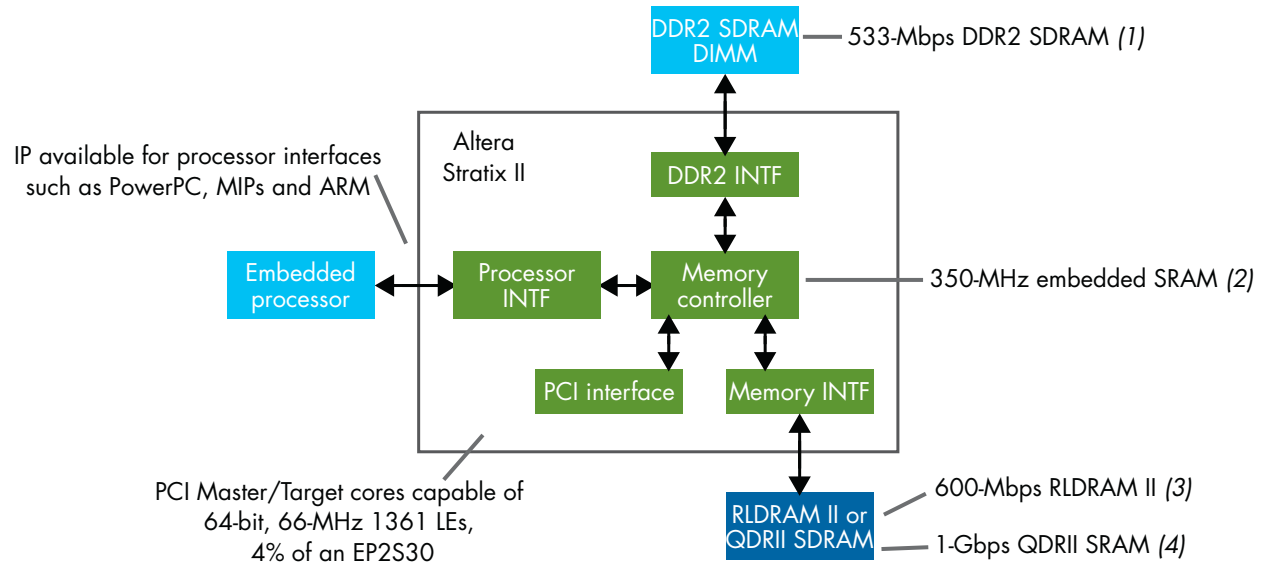
- (1) See [Altera's Memory Solutions Center](#) for the latest table.
- (2) No Altera IP support.

## 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, router boxes, 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 Stratix-series FPGAs optimally implement advanced memory controllers by utilizing their built-in DQS (strobe) phase shift circuitry. [Figure 1](#) highlights some of the features available in an Altera Stratix II FPGA in an embedded application, where DDR2 SDRAM is used as the main memory and QDRII SRAM or RLDRAM II is an external cache level.

Figure 1. Memory Controller Example Using Stratix II FPGA

**Notes:**

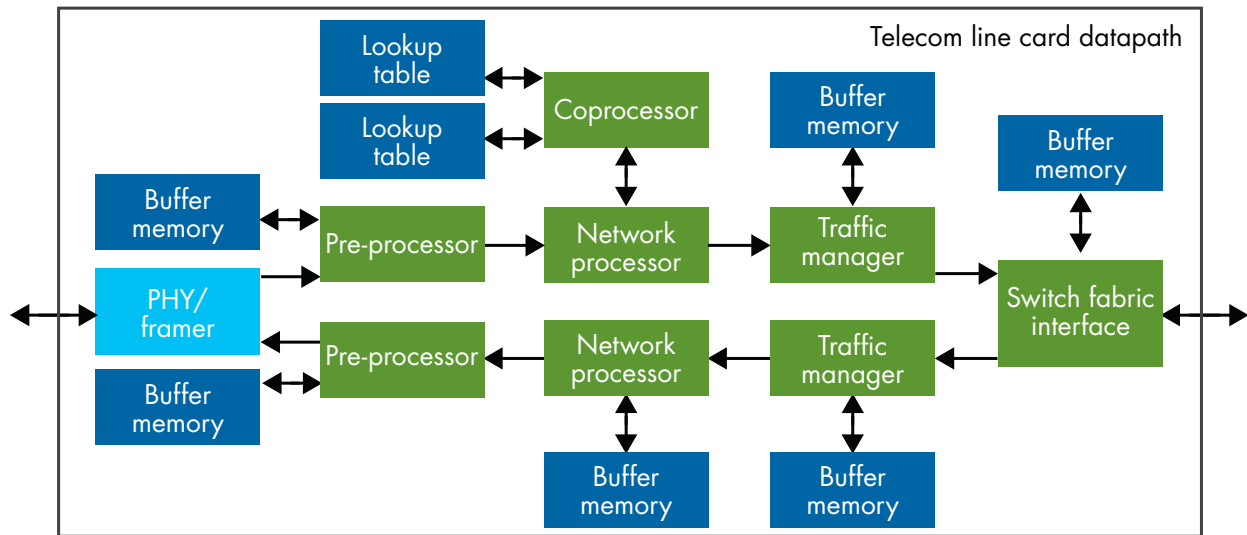
- (1) 600-Mbps RLD RAM II operation: 740 logic elements (LEs), 3% of an EP2S30, and four clock buffers (for a 36-bit wide interface).
- (2) High-speed memory interfaces such as QDR II SDRAM require at least four clock buffers to handle all the different clock phases and data directions. Stratix II FPGAs support 48 dedicated clock resources, 24 in any region of logic.
- (3) 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).
- (4) Embedded SRAM with features such as true-dual port and 350-MHz operation allows complex “store and forward” memory controller architectures.
- (5) Quartus® II software reports the number of adaptive look-up tables (ALUTs) that the design uses in Stratix II devices. The LE count is based on this number of ALUTs.

One of the target markets of RLD RAM II and QDR/QDR II SDRAM is external cache memory. RLD RAM II was developed specifically to have a read latency close to SSRAM, but with the density of SDRAM. A 16X increase in external cache density is achievable with one RLD RAM II versus that of SSRAM. In contrast, QDR and QDR II SDRAM should be considered for systems that require high bandwidth and minimal latency. Architecturally, the dual-port nature of QDR and QDR II SDRAM 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 and 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. (See Figure 2 for 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.

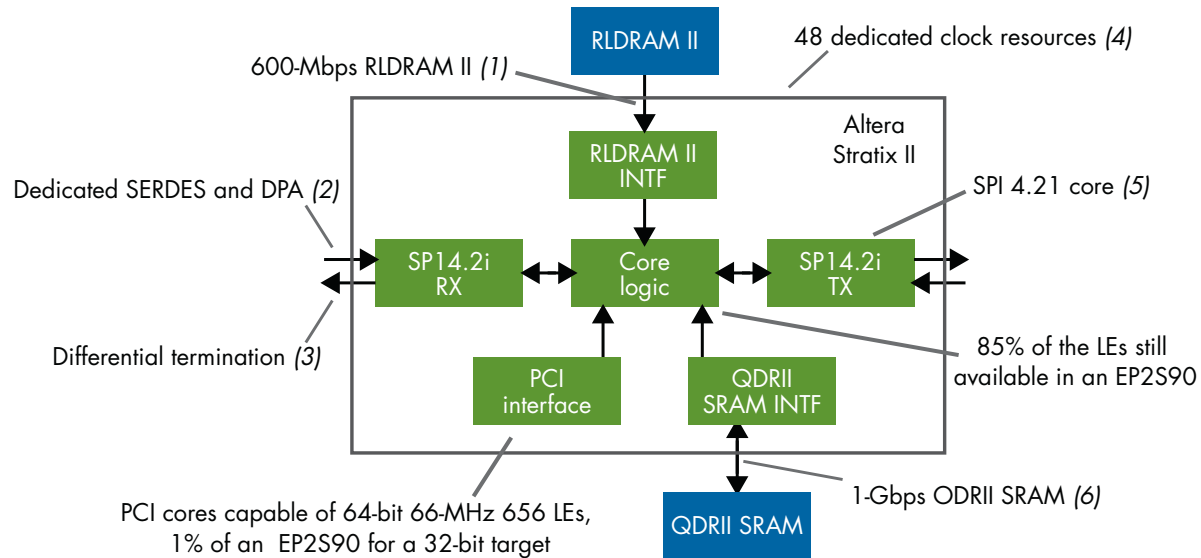
Figure 2. Typical Telecom System Line Interface Card



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/2 SDRAM and RLDRAM II are used for large buffer memories off network processors, traffic managers, and fabric interfaces, while QDR and QDRII 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 are designed to 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 (using Stratix GX), and IP. Figure 3 highlights some of these features in a packet buffering application where RLDRAM II is used for packet buffer memory and QDRII SRAM is used for control memory.

Figure 3. Stratix II FPGA Example in Packet Buffering Application

**Notes:**

- (1) 600-Mbps RLDram II operation: 740 LEs, 1% of an EP2S90, and four clock buffers (for a 36-bit wide interface).
- (2) Dedicated hardware SERDES and DPA circuitry allows clean and reliable implementation of 1-Gbps LVDS.
- (3) Differential termination is built in Stratix FPGAs, simplifying board layout and improving signal quality.
- (4) Stratix II FPGAs support 48 dedicated clock resources, 24 in any region of logic.
- (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) 1-Gbps QDRII SRAM operation: 100 LEs, 0.1% of an EP2S90, and four clock buffers (for an 18-bit interface).
- (7) 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 systems 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 very much 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 2). 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 ( $300 \text{ MHz} * 2 \text{ DDR} * 0.70 \text{ efficiency} * 48 \text{ bits} = 20.1 \text{ Gbps}$ ), which needs two RLDram II parts (one x18 and one x36). RLDram II also inherently includes the additional memory bits used for parity or error correction code (ECC).

QDR and QDRII 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.

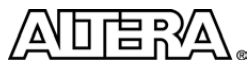
**Conclusion**

As memories advance in complexity, system designers face the challenge of selecting the proper memory for their applications. Because Altera is aware of the importance of memories in high-performance systems, the company has developed the Stratix and Cyclone series FPGAs to be part of the system solution for applications that use external memories, such as embedded processing and telecommunications. Altera is dedicated to identifying features required

in each generation of FPGA to work with emerging generations of memory, such as dedicated DQS phase shifting circuitry, DDR I/O structures, various SSTL/HSTL I/O standards, on-chip termination, and abundant clock resources. Altera's Memory Solutions Center provides reference designs and IP that includes RTL, schematics, demonstration boards, simulation models, characterization reports, board layout guidelines, SSN guidelines, software support, reference designs, and application notes.

### Further Information

- Altera's Memory Solutions Center:  
[www.altera.com/memory](http://www.altera.com/memory)



101 Innovation Drive  
San Jose, CA 95134  
[www.altera.com](http://www.altera.com)

Copyright © 2008 Altera Corporation. All rights reserved. Altera, The Programmable Solutions Company, the stylized Altera logo, specific device designations, and all other words and logos that are identified as trademarks and/or service marks are, unless noted otherwise, the trademarks and service marks of Altera Corporation in the U.S. and other countries. All other product or service names are the property of their respective holders. Altera products are protected under numerous U.S. and foreign patents and pending applications, maskwork rights, and copyrights. Altera warrants performance of its semiconductor products to current specifications in accordance with Altera's standard warranty, but reserves the right to make changes to any products and services at any time without notice. Altera assumes no responsibility or liability arising out of the application or use of any information, product, or service described herein except as expressly agreed to in writing by Altera Corporation. Altera customers are advised to obtain the latest version of device specifications before relying on any published information and before placing orders for products or services.