DEVELOPING HIGH-SPEED MEMORY INTERFACES: THE LatticeSCM FPGA ADVANTAGE

A Lattice Semiconductor White Paper

February 2006

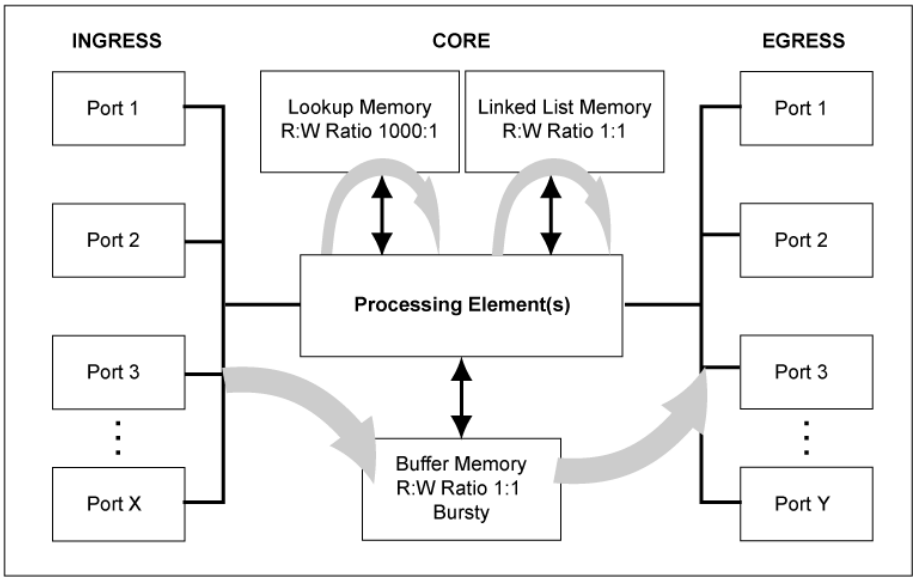
Lattice Semiconductor
5555 Northeast Moore Ct.
Hillsboro, Oregon 97124 USA
Telephone: (503) 268-8000
www.latticesemi.com

Introduction

A common problem for today's system designers is to reliably interface to their next generation high-speed memory devices. As system bandwidths continue to increase, memory technologies have been optimized for higher speeds and performance. As a result, these next generation memory interfaces are also increasingly challenging to design to. Implementing high-speed, high-efficiency memory interfaces in programmable logic devices such as FPGAs has always been a major challenge for designers.

Lattice Semiconductor offers customers a high performance FPGA platform in the LatticeSC to design high-speed memory interface solutions. The LatticeSC family implements various features on-chip that facilitate designing high-speed memory controllers to interface to the next generation high-speed, high performance DDR SDRAM, QDR SRAM, and emerging RLDRAM memory devices. The PURESPEED I/O structures on the LatticeSC, combined with the clock management resources and a high-speed FPGA fabric help customers reduce design risk and time-to-market for high-speed memory based designs. Additionally, the LatticeSCM family implements full-featured embedded high-speed memory controllers on-chip to interface smoothly to the next generation high-speed, high performance DDR I/II SDRAM, QDR II SRAM, and RLDRAM I/II memory devices. The embedded memory controllers, coupled with the I/O structures and clock management resources, are fully verified controllers for customers to use as off-the-shelf solutions.

Memories in Networking. Different functions require different approaches



Memory Applications

Memory devices are an integral part of a variety of systems. However, different applications have different memory requirements. For networking infrastructure applications, the memory devices required are typically high-density, high performance, high bandwidth memory devices with a high degree of reliability. In wireless applications, low power memory is important, especially for handset and mobile devices, while high-performance is important for basestation applications. Broadband access applications typically require memory devices where there is a fine balance between cost and performance. Computing and consumer applications require memory solutions like DRAM modules, Flash cards and others that are highly cost sensitive while delivering the performance targets for these applications. This white paper will focus primarily on memory applications in networking and communications.

Memory can be on-chip or off-chip. Next generation FPGAs like the LatticeSC have several Megabits (up to 7.8M) of RAM on-chip. These are useful for simple FIFO structures for nominal buffering requirements. Cost is the primary factor defining the amount of memory on-chip. For large memory (buffer) requirements, off-chip memory is typically used. On-chip memory is always faster, but has size restrictions due to the cost it adds to the FPGA.

Large, fast memory devices are required in networking and communications applications, with tasks ranging from simple address lookups to traffic shaping/policing to buffer management. Each of these tasks comes with a unique set of requirements. Networking system architects have traditionally turned to Static RAM (SRAM) to solve latency issues, while incurring greater cost. For instance, low- and medium-bandwidth applications require low-latency memory, so SRAMs like ZBT (Zero Bus Turnaround) SRAM and QDR SRAM are ideal. SRAMs improve memory bandwidth by eliminating wait states or idle cycles between read and write cycles. Recently, system architects have turned to SDRAM for networking architectures, where reduced latency meets low-cost.

The figure above illustrates a typical networking architecture. At 10Gbps, address lookups with a typical read-write ratio of 1000:1 could easily be handled with Double Data Rate (DDR) SRAM. Link list management, traffic shaping and statistics gathering tasks typically have balanced a 1:1 read-to-write ratio, requiring higher-performance Quad Data Rate (QDR) SRAM. On the other hand, larger buffer memories are typically implemented in DDR SDRAMs (Synchronous DRAMs). A replacement for DRAM, SDRAM synchronizes memory access with a processor clock for faster data transfer.

Faster speeds are also achieved because SDRAM allows one block of memory to be accessed while another is being prepared for access. Unlike DRAM, SDRAM relies on static current flow rather than a dynamic stored charge, eliminating the need for continual refreshing.

Another new contender has entered the high-performance memory arena. Reduced Latency DRAM (RLDRAM) provides an SRAM-type interface with non-multiplexed addresses. RLDRAM technology provides minimized latency and reduced row cycle times that are very suitable for applications requiring critical response time and very fast random accesses, like next generation (10Gbps and beyond) networking applications.

Memory Controller Challenges

Current memory interfaces often require clock speeds in excess of 200 MHz to achieve the throughput requirements of line and switch cards. This is a major challenge in FPGA architectures. PLLs and DLLs and low-skew clock networks on-chip are essential to allow control of the clock-data relationship.

I/O Support

Next-generation memory controllers operate at HSTL (High-Speed Transceiver Logic) or SSTL (Stub-Series Transistor Logic) voltage levels. This lower voltage level swing is required to support high-speed data operation of the inputs and outputs of the memory device (and the memory controller). HSTL is the de facto I/O standard for high-speed SRAM memory devices, while SSTL is the de facto I/O standard for high-speed DDR SDRAM memories. The LatticeSC implements the industry's highest performance and lowest power programmable I/Os with on-chip termination to serve these applications.

LatticeSC PURESPEED I/Os include digitally controlled, on-chip linear input terminations and output impedance resistors. LatticeSC PURESPEED I/Os that are used as inputs (input-only or bi-directional) feature the ability to provide internal termination. Two termination configurations are available: Termination directly to VTT via programmable impedance, or Termination via a Thevenin-equivalent resistor network across VDDIO and VSS. For DDR2 memory interfaces, the terminations can be switched on/off under tri-state control to support memory bus turnaround requirements. On-chip digital terminations provide the following benefits: they alleviate board layout considerations for the termination resistor, and multiple modes and values are available, providing users with the

flexibility to select the appropriate type and value based on the memory type and I/O they are interfacing to.

When multiple output drivers switch simultaneously, the relative ground voltage within the device is raised momentarily and the power supply droops, resulting in the phenomenon known as ground bounce or Simultaneous Switching Noise (SSN). With memory data rates on the rise and package pin counts continuously increasing, this issue will continue to be more and more prevalent. Ground pin placement has been optimized in the LatticeSC packages to balance both SSO performance and board-level routing for the customer. Ground pins have been distributed throughout the package's ball array to minimize ground return loop inductances and optimize signal integrity. Within the package the signal layers have been referenced to internal ground planes in order to provide continuous return paths.

The PURESPEED I/O buffer itself is both Process/Voltage/Temperature (PVT) and slew-rate controlled. The PVT control maintains constant buffer drive strength to limit the detrimental effects of buffer overdrive under fast switching conditions. The drive strength control, combined with the buffer's linear response allows customers to match board line impedances for optimal signal integrity. This optional slew rate control can be selected to limit the di/dt switching current during signal transitions without impacting the drive strength of the I/O buffer.

On-die decoupling has been included on the LatticeSC device to support the immediate sourcing of current during simultaneous switching events. These capacitors improve signal integrity and improve on-die power regulation for greater timing control of wide buses. In addition to these on-die capacitors, it is recommended that customers follow Lattice guidelines for the selection and placement of board level de-coupling capacitors.

Memory Controller Support

LatticeSC and LatticeSCM devices implement dedicated high-speed I/O circuitry to facilitate implementation of memory controllers supporting the various high-speed memory devices: DDR 1/2 SDRAM, QDR 1/2 SRAM and RLDRAM 1/2. The LatticeSCM devices also implement dedicated high-speed memory controllers on-chip in hard logic supporting these standards as well. These high-speed memory controllers are implemented using ASIC technology and are embedded on the device.

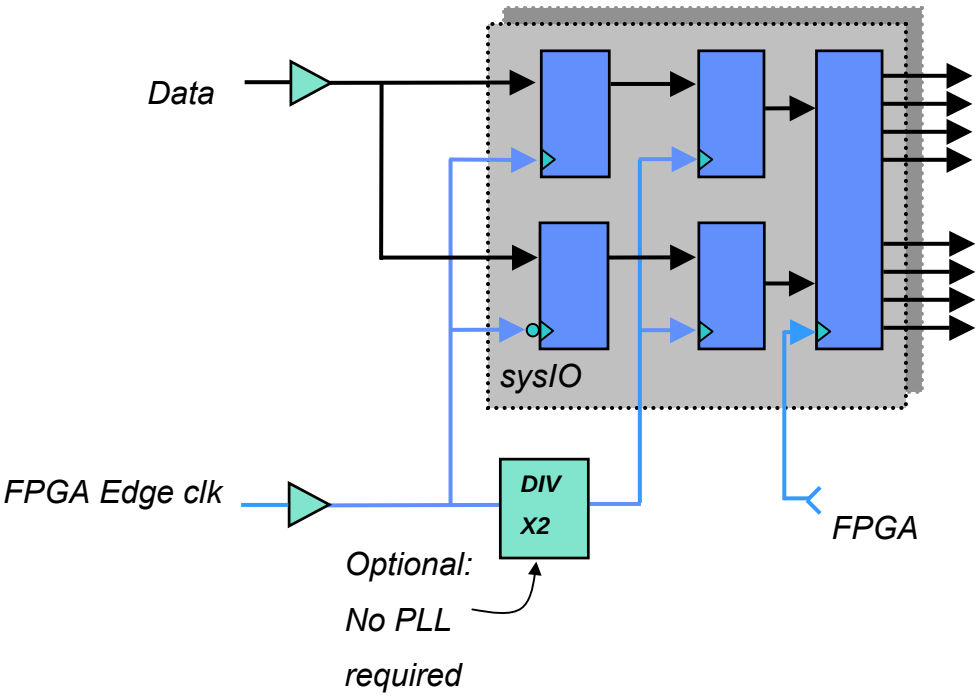
The controllers are full-featured, fully tested controllers, providing users a low-risk time-to-market solution for high-speed memory interfaces.

DDR and DDR2 SDRAM Controller

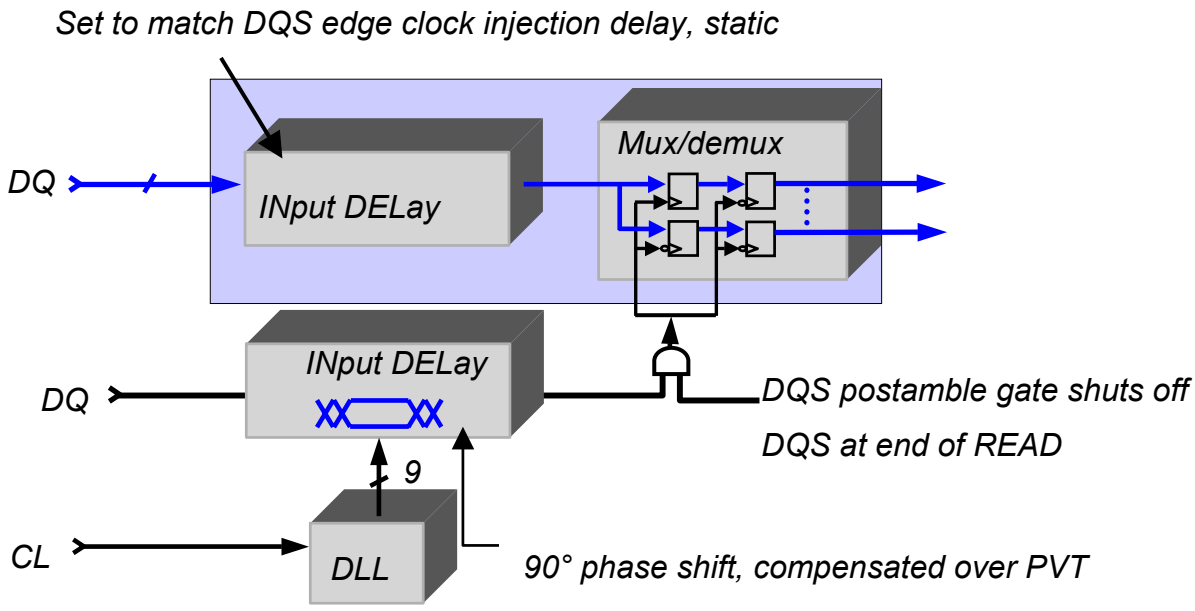
Implementing high performance DDR memory interfaces requires careful design of the read and write interface blocks of the memory controller. DDR2 memory devices pose a bigger challenge due to their higher speeds and the bi-directional DQS signal. The LatticeSCM memory controller utilizes on-chip PLLs and DLLs, along with programmable delay elements at the input buffers to align the DQS and DQ signals. These elements work together to compensate for process, voltage and temperature variations, providing reliable operation at all operating conditions and various frequencies. These devices also contain dedicated DDR register structures in the inputs (for read operations) and in the outputs (for write operations). All these blocks are critical for implementing reliable high-speed DDR and DDR2 SDRAM controllers. Typically, designers have problems implementing high-speed memory controllers in FPGAs due to the complexity of the DQS logic. LatticeSCM devices alleviate this problem by implementing full-featured, fully tested memory controllers on chip.

Dedicated Double Data Rate I/O Logic Elements

LatticeSC PURESPEED I/Os contains five blocks: input register block, output register block, tri-state register block, update block and a control logic block. These blocks contain registers for both double data rate (DDR) data transfer and the necessary clock and selection logic. The input register block also contains delay elements and registers that can be used to condition signals before they are passed to the FPGA. The delay block allows users to align signals. The delay block uses four blocks of 32 tapped delay lines to obtain coarse and fine delay resolution. These delays can be adjusted automatically via DLLs because the delay line in this block matches the delay line that is used in the 12 on-chip DLLs. The delay line can be set via configuration bits, or driven from a calibration bus that allows the setting to be controlled either from one of the on-chip DLLs or user logic. Controlling the delay from one of the on-chip DLLs allows the DQ and DQS delay to be calibrated to the memory chip reference clock and so automatically compensated for the variations in process, voltage, temperature and system speeds.



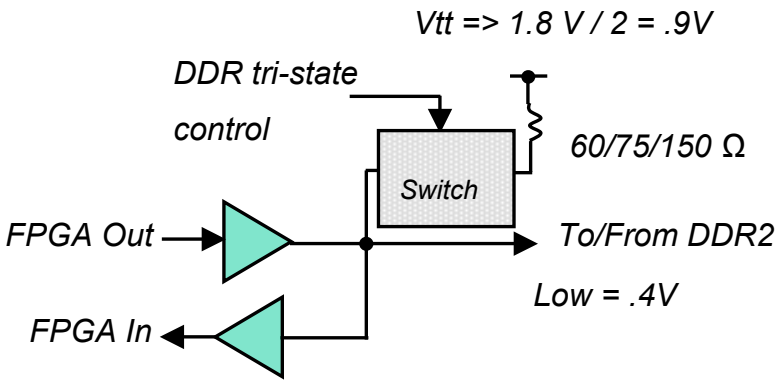
LatticeSC PURESPEED I/O DDR Input



LatticeSC DQ/DQS Implementation

Switchable I/O Termination

DDR2 memory (and RLDRAM2 in Common I/O mode) standards require that the on-chip termination to VTT be turned on when a pin is an input and off when the pin is an output. LatticeSC PURESPEED I/O also implements the DDR2 ODT control. The written signal enables the tri-state buffer while driving DQ out of the core to the memory. Termination is ON all the time and is switched OFF only during writes to the DDR2 memory device. This mode is also used for DDR Input to save power.



LatticeSC Switchable I/O Termination

Write Timing

During write, DQS is center aligned with the data at the memory device pins. The PLL phase-shifted clock output is used to provide this signal to the memory device.

The PURESPEED I/O in the LatticeSC does not require custom placement of the memory controller pins. Unlike other FPGAs, the LatticeSC family does not require dedicated pins for DQS (for higher speeds, pins that drive edge clocks on the devices are recommended for higher performance). This is an important benefit for designers since the device is not hampered with the stringent I/O layout requirements common with other FPGAs.

Read Timing

Source Synchronous interfaces generally require the input clock to be adjusted in order to correctly capture data at the input register. For most interfaces, a PLL is used for this adjustment; however, in DDR memories the clock (referred to as DQS) is not free running so this approach cannot be used.

In DDR memory interfaces the DQS to Master Clock relationship varies due to factors like PCB trace length and the memory device being used. The DQS signal from the DDR2 memory is generated from the K_clock sent from the LatticeSC device and then sent back to the FPGA during a read. In order to match the input buffer delay on the DQS/DQS# pins, the K_clock is looped back within the same I/O pad to the input clock routing in order to emulate the delay. This delay path thus has the same output buffer delay as K_clock (including associated extrinsic loading delay). The reference to the delay cells are fed by a control bus from the DLL, which is the same control that is used to provide a 90-degree lag on the DQS pins. On the DDR2 device, the K_clock input is used to generate the DQS strobe at t_{DQSCK}. The DQS is then received at the FPGA to capture the read data. Thus the total phase-shift that needs to be accounted for includes 2 x PCB routing delay (K_clock and DQS routing) + t_{DQSCK}. For low-speed interfaces (up to 167 MHz), and short trace lengths, the roundtrip delay can be emulated without a second PLL. For longer trace lengths and higher speeds (above 168 MHz), a second PLL is used to allow a reliable transfer of read data to the FPGA logic over process, voltage and temperature (PVT) variations. A second PLL is used to generate the phase-shifted clock that matches the round trip delay of DQS. This clock is used to clock data into the read FIFO of the memory controller, providing reliable operation over process, voltage and temperature.

The LatticeSCM family memory controller design implements DQS read circuitry that simplifies the memory interface design and ensures robust operation. The read data DQ is captured on the rising or falling edge of the data strobe DQS. Since DQ and DQS are edge-aligned, DQS needs to be delayed (ideally centered to DQ) to effectively capture the data. Methods such as using the cycle stealing delays or by pre-setting the INDEL to a given value can be used to delay the DQS with respect to the data. However, the secondary and edge clock injection delay variations over process, voltage and temperature greatly reduce the timing margin at high speeds. Using DLLs provides better margin independent of the data bus width. The core K_clock used to generate K and K# on the write interface is fed into a DLL (which operates as the master DLL) to produce a T/4 digital control output. This is a 9-bit bus that is used to control the delay cells within the PIOs used for DQS/DQS# and will provide a 90-degree time shift for the DQS/DQS# signals. This DLL can be adjusted to give additional margin on top of the 90-degree delay based on the actual system requirements. The DQ pins use the input delay in a fixed edge clock injection mode. This produces identical delays that match the clock injection delay on DQS/DQS#, allowing the DQ data to be 90 degrees away from the DQS signal at the input DDR Flip-flops.

DQS Shut-off

When neither the DDR SDRAM device nor the controller is driving DQS, the signals are in a high-impedance state. Both DQ and DQS are terminated by a pull-up resistor that pulls up the voltage to the termination voltage. According to the JEDEC specification for the SSTL-2 I/O standard and the HSTL-I standard, this is an indeterminate logic level that can be interpreted as either a logic high or logic low. During read postamble, any noise on the DQS line can cause invalid data to be registered, corrupting the read data to the user interface. Hence, it is important to gate the tristate-able DQS read strobe from the DDR2 memory. The DQS shut off logic is implemented in the LatticeSC I/O block to turn off DQS during read postamble.

RLDRAM1 and 2 Memory Controller

Reduced Latency DRAM (RLDRAM) provides an SRAM-type interface with non-multiplexed addresses. RLDRAM II technology provides minimized latency and reduced row cycle times that are very suitable for applications requiring critical response time and very fast random accesses, like the next generation (10 Gbps and beyond) networking applications. The LatticeSCM devices also implement on-chip RLDRAM memory controllers supporting both RLDRAM1 and RLDRAM2 devices. The LatticeSC memory controller supports both types of RLDRAM2 memory: Common I/O (CIO) and Separate I/O (SIO).

The data and clocking circuitry for the RLDRAM controller is very similar to that used for the DDR2 memory controller. The same write interface as per DDR2 is used. For RLDRAM2, read data is sent edge-aligned with the read clock QK. Data is center-aligned with the write clock DK when writing to the RLDRAM device. A PLL and a DLL are used to align the data and the clock. Since the data and the clock are edge-aligned coming from the RLDRAM device, QK needs to be delayed (ideally centered to Q) to effectively capture the data. Methods such as using trace delays or pre-setting an input delay cell to a given value can be used to delay the QK with respect to the data, but this method does not scale across process, voltage and temperature changes. A DLL is used to delay the QK signals by 90°, giving the user the greatest timing margin over PVT; this is also independent of the interface speed Essentially, the same read interface as DDR2 is used without the need for DQS postamble shut-off or preamble edge detect. The main difference is the higher speed of RLDRAM, up to 400 MHz, and the optional use of the high-speed I/O gearing logic.

QDR SRAM Memory Controller

QDR SRAMs are very popular in low latency applications requiring simultaneous reads and writes. The LatticeSCM devices also implement an embedded QDR SRAM memory controller for interfacing to QDR2 SRAM memory devices. The memory controller supports both 2- and 4-word burst modes.

QDR2 SRAMs provide a separate read and write bus for read and write operations. There are three sets of clocks associated with QDR2 SRAM Devices: K and K# clocks are input clocks to the QDR2 SRAM device and are sent from the memory controller to the QDR2 SRAM device. C and C# are output clocks from the QDR2 SRAM device and are not used. CQ and CQ# are echo clocks that are better suited to capture the read data, especially at higher frequencies because they are timed exactly like the output data Q signals and can be treated as valid data indicators. Since Q data and CQ clock are edge-aligned coming from the QDR2 device, CQ needs to be delayed (ideally centered to Q) to effectively capture the data. Methods such as using the matched trace delays or preset input delay blocks can be used to delay the CQ with respect to the data, but using the DLL to delay the CQ signals by 90° gives the greatest timing margin over PVT and is independent of the interface speed.

Embedded Memory Controller Advantages

The embedded memory controllers on the LatticeSCM devices provide customers with high-performance, low-risk solutions for interfacing to their external memory chips. Customers do not need to design a memory controller using FPGA gates, saving time and FPGA real estate while designing high-speed designs requiring high-speed external memory interfaces. The embedded ASIC controllers also provide much higher performance than can be achieved by FPGA soft IP implementations. The power consumption for the ASIC implementation is approximately ½ the equivalent FPGA implementation.

Memory Controller	*FPGA Logic Resources saved*
DDR 1 / 2 Controller	1700 Look-up-Tables (LUTs) each
RLDRAM 1 / 2 Controller	2000 LUTs each
QDR SRAM Controller	400 LUTs each

FPGA Logic Resources savings

Memory Type Supported	*Maximum Operating Speeds*
DDR SDRAM	200 MHz
DDR2 SDRAM	300 MHz
QDR2 SRAM	300 MHz
RLDRAM1	300 MHz
RLDRAM 2	400 MHz

LatticeSC/SCM Supported Memory Types and Clock Speeds

Conclusion

As line rates continue to grow, high-speed SRAM and SDRAM devices are experiencing wider adoption in networking applications. These increasing system bandwidth requirements are making

memory controller designs increasingly complicated and high-risk. Facilitating robust high-performance memory interface designs was a principal goal in the design of the LatticeSC FPGAs. PURESPEED I/Os, the DQS control logic, and on-chip high performance DLLs and PLLs coupled with the high-speed FPGA fabric on the LatticeSC means that designers finally have a reliable, high-performance, low-risk solution for next-generation memory controller needs.