Freescale Semiconductor
Application Note

Document Number: AN3935
Rev. 1, 11/2011

MSC8156 and MSC8157 PCI Express Performance

This application note presents performance measurements of the MSC8156 and MSC8157 PCI Express controller memory accesses. It begins with an introduction to general PCI Express concepts to help the reader understand factors that affect the performance of a system. Actual test data are presented and compared to estimated performance. Finally, it provides a list of techniques for achieving optimal PCI Express performance. A CodeWarrior project is also included with this application note for the user to measure the results directly.

The MSC8156 family includes the following devices:

- MSC8151
- MSC8152
- MSC8154
- MSC8154E
- MSC8156
- MSC8156E
- MSC8251
- MSC8252
- MSC8254
- MSC8256

The members of the MSC8157 family that support the PCI Express interface include the following:

- MSC8157
- MSC8157E

Contents

freescale™ semiconductor

1 Background

Freescale multicore MSC8156 and MSC8157 DSPs include PCI Express controllers. The MSC8156 PCI Express is designed to support the *PCI Express Base Specification, Revision 1.0a* and the MSC8157 PCI Express is designed to support the *PCI Express Base Specification, Revision 2.0*. The PCI Express controllers connect the DSP devices through the High Speed Serial Interface (HSSI) to a serial interface configurable for up to a x4 interface. The MSC8156 PCI Express interface runs at a 2.5 GHz transfer rate and the MSC8157 PCI Express interface runs at a 5 GHz transfer rate.

The PCI Express controller can operate as either a root complex (RC) or an endpoint (EP) device. In RC mode, it supports configuration and I/O transactions. In EP mode, it accepts configuration transactions to the internal PCI Express configuration registers. Both RC and EP modes support message generation and acceptance. As an initiator, the controller supports memory read and write operations with a maximum transaction size of 256 bytes. As a target device, it accepts read and write operations to local memory space.

2 Encoding Overhead

Each lane of a PCI Express device transmitter implements an 8b/10b encoder that encodes 8-bit characters into 10-bit symbols. This scheme allows a clock to be embedded into the serial bit stream, eliminating the need for a high frequency clock signal (2.5 GHz for MSC8156 and 5 GHz for MSC8157) that would generate significant EMI noise. It also maintains DC balance on the bit stream because the number of 1s and 0s transmitted are as close to equal as possible. However, because each 8-bit character is encoded as a 10-bit symbol, the transmission overhead reduces the bit rate to 80% of the transfer rate per direction per lane. Equation 1 shows the bit rate for PCI Express systems with x1, x2, or x4 links.

$$\text{Bit rate} = (\text{transfer rate} * \text{lane width}) * (8/10) \qquad \text{Eqn. 1}$$

Table 1. MSC8156/MSC8157 PCI Express Bit Rate

Link Width	x1	x2	x4
MSC8156 bit rate (transfer rate = 2.5 Gbps)	2 Gbps	4 Gbps	8 Gbps
MSC8157 bit rate (transfer rate = 5.0 Gbps)	4 Gbps	8 Gbps	16 Gbps

Factors such as packet and traffic overhead and efficiency all contribute to the actual system bandwidth. These factors are discussed in the following sections.

3 Transaction Layer Packet Overhead

Communication between PCI Express devices involves the transfer of packets called Transaction Layer Packets (TLPs). TLP assembly begins at the transmitter and TLP disassembly ends at the receiver. The Transaction Layer packetizes and prepends a header to the data payload and optionally adds the end-to-end cyclic redundancy checksum (ECRC). When the TLP is forwarded to the Data Link Layer, the Sequence Number and the Link Layer CRC (LCRC) are added to the packet to guarantee successful transmission across the link. The TLP is then passed to the Physical Layer, where information is added to mark the start and end of the packet before it is sent out on the transmit side of the link. The receiver then decodes and

disassembles the inbound packet in the reverse order. Figure 1 shows a typical TLP. Please note that the max data payload size is 256 bytes. If the transfer data buffer size is more than 256 bytes, then it is broken by the PCI Express controller to several packets of 256 byte each.

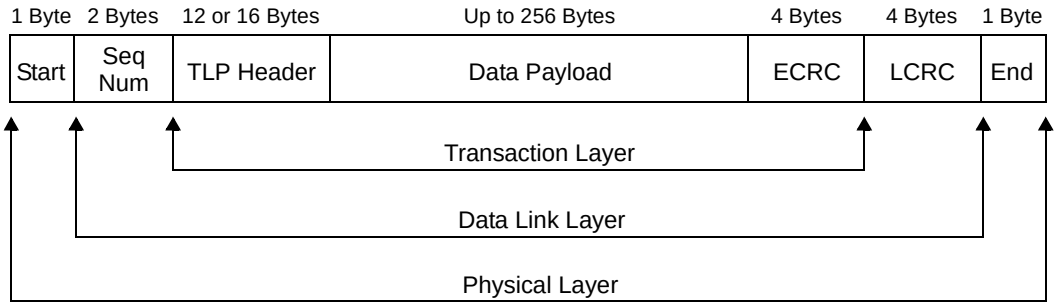

Figure 1. Transaction Layer Packet

TLP overhead varies depending on 32-bit or 64-bit addressing and the optional ECRC. The 32-bit addressable TLP header is 12 bytes, whereas the 64-bit addressable TLP header requires an additional 4 bytes of information. With 32-bit addressing and without ECRC, the minimum TLP overhead is 20 bytes.

4 ACK/NAK Overhead

The Data Link Layer ensures reliable transport of TLPs between devices by using ACK Data Link Layer Packets (DLLPs) to acknowledge reception and NAK DLLPs to indicate error reception of TLPs. This mechanism allows the receiver to notify the transmitter as to good or bad TLPs by returning an ACK or NAK DLLP. The receiver may return a single DLLP to acknowledge multiple TLPs to reduce ACK DLLP traffic on the bus. If the transmitter receives a NAK DLLP, it then resends the associated TLP(s) from its replay buffer. The transmitter holds the TLP(s) in its replay buffer until an ACK is received for the associated TLP(s) so that the packet(s) can be resent if a problem occurred in its transmission.

Figure 2 shows an example of a receiver sending a single ACK DLLP to acknowledge reception of three TLPs from the transmitter.

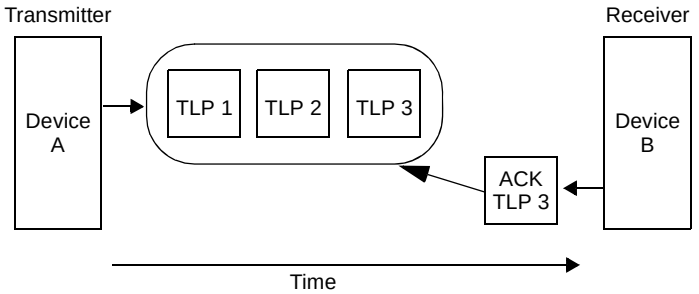

Figure 2. Example 3 TLPs and 1 ACK

All DLLPs consist of one byte to indicate the type of DLLP and three additional bytes of attributes that vary by DLLP type. A 16-bit CRC is appended to it. The physical layer then adds the framing information for a total DLLP size of 8 bytes. Figure 3 shows the format of an ACK/NAK DLLP.

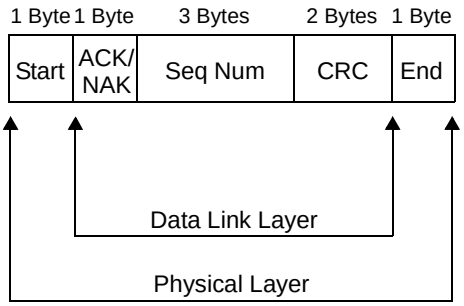

Figure 3. ACK/NAK Data Link Layer Packet

5 Flow Control Overhead

PCI Express utilizes credit-based flow control to guarantee that transmitters never send TLPs that the receiver cannot accept and which would otherwise result in buffer overruns and retries. The receiver reports the size of its receive buffers in flow control (FC) credits via FC DLLPs to the device on the opposite end of the link. The receiving port continually updates the transmitting port by sending FC DLLPs so that the transmitter only sends packets if the receiver has enough buffer space. The transmitter also keeps track of the number of credits available and decrements this counter each time a packet is sent. After the receiver processes the packet, it frees buffer space and sends an FC Update DLLP to notify the transmitter about available credits. The FC DLLP is also 8 bytes in length as shown in Figure 4.

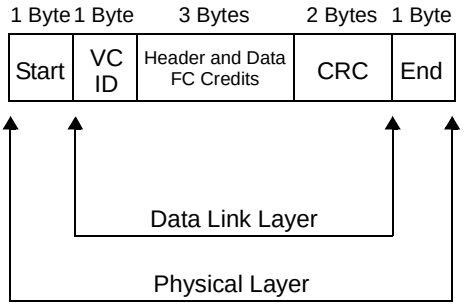

Figure 4. Flow Control Data Link Layer Packet

6 Calculating Bandwidth

Bandwidth is simply the amount of data that can be transferred in a given period of time. Taking into account the TLP overhead discussed in Section 3, “Transaction Layer Packet Overhead”, the DLLP overhead in Section 4, “ACK/NAK Overhead”, and the flow control overhead in Section 5, “Flow Control Overhead”, the full-duplex bandwidth in Gbps is calculated as shown in Equation 2.

$$Bandwidth = Bit\ rate * data\ size / (data\ size + overhead) * 2\ directions \quad \text{Eqn. 2}$$

$$overhead = TLP\ overhead + ACK\ overhead + FC\ overhead \quad \text{Eqn. 3}$$

NOTE

The amount of overhead depends on the data size. Please refer to Table 2 and Table 3 for calculation of the overheads for difference data sizes.

6.1 Write Transfer Estimates

To estimate the system bandwidth for write transfers, we l need account for the overhead discussed in the previous sections. We also need to use the following reasonable assumptions:

- 4 active lanes @ 2.5 Gbps for MSC8156 and 5 Gbps for MSC8157
- 256 bytes maximum payload size
- 20 bytes TLP packet overhead
- Assume ACK/NAK (8 bytes) and FC DLLP (8 bytes) updates for every 4 TLPs

Table 2 shows the calculated bandwidth for MSC8156 configured with x4 link width. With the given system parameters, bandwidth of 14.63 Gbps can be achieved for large packet sizes. This estimate is about 91% of the full-duplex bit rate of 16 Gbps (8Gbps * two directions). However, in a real system, actual throughput efficiency will be much less for small packets. The calculated numbers are only estimates that are shown for comparison with actual measurements, which could be better or worse depending on actual system traffic and efficiency.

Table 2. MSC8156 PCI Express Calculated Bandwidth for x4 Write Transfers

Bytes	Calculation	Bandwidth	Notes
8	8 * 8 / (8 + 20 + 8 + 8) * 2	2.91 Gbps	1 TLP (20), 1 ACK (8), 1 FC (8)
16	8 * 16/ (16 + 20 + 8 + 8) * 2	4.92 Gbps	
32	8 * 32/ (32 + 20 + 8 + 8) * 2	7.53 Gbps	
64	8 * 64 / (64 + 20 + 8 + 8) * 2	10.24 Gbps	
128	8 * 128 / (128 + 20 + 8 + 8) * 2	12.49 Gbps	
256	8 * 256 / (256 + 20 + 8 + 8) * 2	14.03 Gbps	
512	8 * 512 / (512 + 2*20 + 8 + 8) * 2	14.42 Gbps	2 TLP, 1 ACK, 1 FC
1 K	8 * 1K / (1K + 4*20 + 8 + 8) * 2	14.63 Gbps	4 TLP, 1 ACK, 1 FC
2 K	8 * 2K / (2K + 8*20 + 2*8 + 2*8) * 2	14.63 Gbps	8 TLP, 2 ACK, 2 FC
4 K	8 * 4K / (4K + 16*20 + 4*8 + 4*8) * 2	14.63 Gbps	16 TLP, 4 ACK, 4 FC
8 K	8 * 8K / (8K + 32*20 + 8*8 + 8*8) * 2	14.63 Gbps	32 TLP, 8 ACK, 8 FC
16 K	8 * 16K / (16K + 64*20 + 16*8 + 16*8) * 2	14.63 Gbps	64 TLP, 16 ACK, 16 FC
32 K	8 * 32K / (32K + 128*20 + 32*8 + 32*8) * 2	14.63 Gbps	128 TLP, 32 ACK, 32 FC

MSC8157 PCI Express bandwidth for x4 configuration can be calculated similarly, as shown in Table 3.

Table 3. MSC8157 PCI Express Calculated Bandwidth for x4 Write Transfers

Bytes	Calculation	Bandwidth	Notes
8	16 * 8 / (8 + 20 + 8 + 8) * 2	5.82 Gbps	1 TLP (20), 1 ACK (8), 1 FC (8)
16	16 * 16/ (16 + 20 + 8 + 8) * 2	9.85 Gbps	
32	16 * 32/ (32 + 20 + 8 + 8) * 2	15.06 Gbps	
64	16 * 64 / (64 + 20 + 8 + 8) * 2	20.48 Gbps	
128	16 * 128 / (128 + 20 + 8 + 8) * 2	24.98 Gbps	
256	16 * 256 / (256 + 20 + 8 + 8) * 2	28.05 Gbps	
512	16 * 512 / (512 + 2*20 + 8 + 8) * 2	28.85 Gbps	2 TLP, 1 ACK, 1 FC
1 K	16 * 1K / (1K + 4*20 + 8 + 8) * 2	29.26 Gbps	4 TLP, 1 ACK, 1 FC
2 K	16 * 2K / (2K + 8*20 + 2*8 + 2*8) * 2	29.26 Gbps	8 TLP, 2 ACK, 2 FC
4 K	16 * 4K / (4K + 16*20 + 4*8 + 4*8) * 2	29.26 Gbps	16 TLP, 4 ACK, 4 FC
8 K	16 * 8K / (8K + 32*20 + 8*8 + 8*8) * 2	29.26 Gbps	32 TLP, 8 ACK, 8 FC
16 K	16 * 16K / (16K + 64*20 + 16*8 + 16*8) * 2	29.26 Gbps	64 TLP, 16 ACK, 16 FC
32 K	16 * 32K / (32K + 128*20 + 32*8 + 32*8) * 2	29.26 Gbps	128 TLP, 32 ACK, 32 FC

6.2 Read Transfer Estimates

Memory reads use the split-transaction requests and completions. In the first phase of the transfer, the requester sends a read request MRd TLP and the completer sends an ACK DLLP to acknowledge the request. In the second phase, the completer then sends the requested data which can be split into multiple completion with data CplD TLPs. The split-transaction read transfers require more overhead than the write transfers because two separate TLPs are generated as shown in Figure 5.

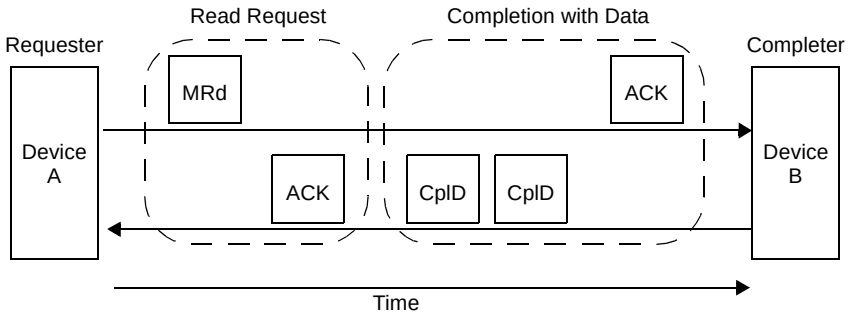

Figure 5. Memory Read Split-Transaction Example

The read request TLP contains no payload data; therefore, it results in a read request overhead of 20 bytes. MSC8156 and MSC8157 support memory read operations with a maximum transaction size of 256 bytes. This means that one read request packet is required for each packet of completion data. The full-duplex read bandwidth in Gbps is calculated as shown in Equation 4 and Equation 5.

$$Bandwidth = Bit\ rate * data\ size / (data\ size + overhead) * 2\ directions \quad \text{Eqn. 4}$$

$$overhead = read\ request\ overhead + TLP\ overhead + ACK\ overhead + FC\ overhead \quad \text{Eqn. 5}$$

The bandwidth calculations for read transfers also use the same system parameters and assumptions as the write transfers.

- 4 active lanes @ 2.5 Gbps
- 256 bytes maximum payload size
- 256 bytes maximum read request
- 20 bytes TLP packet overhead
- Assume ACK/NAK (8 bytes) and FC DLLP (8 bytes) updates for every 4 TLPs

Table 4 and Table 5 shows the PCI Express performance for MSC8156 and MSC8157. With the above settings, bandwidth of 13.65 Gbps can be achieved with MSC8156 PCI Express for large packet sizes. This number is about 85% of the full-duplex bit rate of 16 Gbps. The split-transaction mechanism results in the read performance being lower than the write performance. Again, these calculations are only used for comparison with actual measurement results.

Table 4. MSC8156 PCI Express Calculated Bandwidth for x4 Read Transfers

Bytes	Calculation	Bandwidth	Notes
8	8 * 8 / (8 + 20+ 20 + 8 + 8) * 2	2 Gbps	1 MRd, 1 CplD, 1ACK, 1 FC
16	8 * 16/ (16 + 20+ 20 + 8 + 8) * 2	3.56 Gbps	
32	8 * 32/ (32 + 20+ 20 + 8 + 8) * 2	5.82 Gbps	
64	8 * 64 / (64 + 20+ 20 + 8 + 8) * 2	8.53 Gbps	
128	8 * 128 / (128 + 20+ 20 + 8 + 8) * 2	11.13 Gbps	
256	8 * 256 / (256 + 20+ 20 + 8 + 8) * 2	13.13 Gbps	
512	8 * 512 / (512 + 4*20 + 8 + 8) * 2	13.47 Gbps	2 MRd, 2 CplD, 1ACK, 1 FC
1 K	8 * 1K / (1K + 8*20 + 8 + 8) * 2	13.65 Gbps	4 MRd, 4 CplD, 1ACK, 1 FC
2 K	8 * 2K / (2K + 16*20 + 2*8 + 2*8) * 2	13.65 Gbps	8 MRd, 8 CplD, 2 ACK, 2 FC
4 K	8 * 4K / (4K + 32*20 + 4*8 + 4*8) * 2	13.65 Gbps	16 MRd, 16 CplD, 4 ACK, 4 FC
8 K	8 * 8K / (8K + 64*20 + 8*8 + 8*8) * 2	13.65 Gbps	32 MRd, 32 CplD, 8 ACK, 8 FC
16 K	8 * 16K / (16K + 128*20 + 16*8 + 16*8) * 2	13.65 Gbps	64 MRd, 64 CplD, 16 ACK, 16 FC
32 K	8 * 32K / (32K + 256*20 + 32*8 + 32*8) * 2	13.65 Gbps	128 MRd, 128 CplD, 32 ACK, 32 FC

Table 5. MSC8157 PCI Express Calculated Bandwidth for x4 Read Transfers

Bytes	Calculation	Bandwidth	Notes
8	16 * 8 / (8 + 20+ 20 + 8 + 8) * 2	4 Gbps	1 MRd, 1 CplD, 1ACK, 1 FC
16	16 * 16/ (16 + 20+ 20 + 8 + 8) * 2	7.11 Gbps	
32	16 * 32/ (32 + 20+ 20 + 8 + 8) * 2	11.64 Gbps	
64	16 * 64 / (64 + 20+ 20 + 8 + 8) * 2	17.07 Gbps	
128	16 * 128 / (128 + 20+ 20 + 8 + 8) * 2	22.26 Gbps	
256	16 * 256 / (256 + 20+ 20 + 8 + 8) * 2	26.26 Gbps	
512	16 * 512 / (512 + 4*20 + 8 + 8) * 2	26.95 Gbps	2 MRd, 2 CplD, 1ACK, 1 FC
1 K	16 * 1K / (1K + 8*20 + 8 + 8) * 2	27.31 Gbps	4 MRd, 4 CplD, 1ACK, 1 FC
2 K	16 * 2K / (2K + 16*20 + 2*8 + 2*8) * 2	27.31 Gbps	8 MRd, 8 CplD, 2 ACK, 2 FC
4 K	16 * 4K / (4K + 32*20 + 4*8 + 4*8) * 2	27.31 Gbps	16 MRd, 16 CplD, 4 ACK, 4 FC
8 K	16 * 8K / (8K + 64*20 + 8*8 + 8*8) * 2	27.31 Gbps	32 MRd, 32 CplD, 8 ACK, 8 FC
16 K	16 * 16K / (16K + 128*20 + 16*8 + 16*8) * 2	27.31 Gbps	64 MRd, 64 CplD, 16 ACK, 16 FC
32 K	16 * 32K / (32K + 256*20 + 32*8 + 32*8) * 2	27.31 Gbps	128 MRd, 128 CplD, 32 ACK, 32 FC

7 MSC8156 PCI Express Performance Measurements

Measurements were performed using two MSC8156 ADS boards inserted into a standard AMC backplane chassis. One device was configured as the RC and the other device as the EP. The RC device performed both outbound memory write and read transfers to/from the EP. The High-Speed Serial Interface (HSSI) DMA controllers moved data between the RC's DDR memory and the PCI Express outbound window. The EP inbound window mapped to its own DDR memory. Measurements were only made for memory transactions. Message and I/O transactions were not included in the measurements. The On-Chip Emulator (OCE) was configured to count the number of core cycles from the start of the DMA transfer to the end of the transfer. Transfer sizes range from 8 bytes to 32 KB.

Table 6 shows the MSC8156 ADS switch settings for the RC and EP devices.

Table 6. MSC8156 ADS Switch Settings

ADS Switches	ADS Settings (0=ON, 1=OFF)		Description
	RC	EP	
SW1	00000000		Default settings
SW2	00000110		Default settings
SW3	x4: 00000000 x2: 10000000 x1: 10000000		SerDes Port 2 settings
SW4	x4: 10011011 x2: 10011001 x1: 10011000		
SW5	10010000		Default settings
SW6	01010000	01000000	PCI Express ready, EP/RC mode select
SW7	00000011		Default settings

The pcie_8156 project requires CodeWarrior for StarCore v10 or higher to run the bandwidth test. The same project is used for both RC and EP devices. Use the following steps to run the test:

1. Ensure that the two MSC8156 ADS boards are connected via the AMC connector. A standard AMC backplane chassis may be used.
2. Configure the RC and EP ADS switches for x4 mode as shown in Table 6.
3. Open the project pcie_8156 in CodeWarrior for StarCore Debugger.
4. In the source file msc8156_main.c, uncomment the following definition and build the project:
   `#define EP`
5. Connect the USB TAP to the P1 JTAG connector of the EP ADS and run the project. This step opens the HSSI lanes and configures the EP device.
6. After EP configuration is complete and basic configuration information is displayed, the core enters debug mode. Terminate the project. Do not reset the ADS boards or remove power from the boards.
7. In the source file msc8156_main.c, comment the following definition and build the project:
   `//#define EP`

8. Connect the USB TAP to the P1 JTAG connector of the RC ADS and run the project. This step configures the RC device and performs various read and write tests.
9. Test results are displayed in the stdio console window.
10. Repeat the above steps for x2, and x1 modes.

7.1 Link Widths

The MSC8156 PCI Express controller supports link widths x1, x2 and x4. Wider link width means TLPs and DLLPs are transmitted faster since more bytes can be transferred in a given time. The MSC8156 PCI Express link width is determined by the reset configuration word which are recorded in the Reset Configuration Word Low Register RCWL[S2P] bits. Table 7 and Table 8 shows the results of changing the link width in the write and read tests.

Table 7. Link Width Effect on Write Performance

Bytes	Measured Bandwidth (Gbps)		
	x4	x2	x1
8	0.80	0.75	0.63
16	1.51	1.41	1.18
32	3.01	2.69	2.16
64	5.28	4.48	3.22
128	8.52	6.21	3.73
256	11.30	7.32	4.03
512	13.34	7.28	3.83
1 K	14.03	7.37	3.76
2 K	14.39	6.73	3.73
4 K	14.61	6.88	3.71
8 K	14.67	6.74	3.70
16 K	14.72	6.72	3.69
32 K	14.75	6.54	3.69

Table 8. Link Width Effect on Read Performance

Bytes	Measured Bandwidth (Gbps)		
	x4	x2	x1
8	0.58	0.55	0.43
16	1.13	1.08	0.80
32	2.17	1.99	1.34
64	3.78	3.24	1.96
128	6.03	4.41	2.55
256	8.44	5.47	3.04
512	11.02	6.17	3.29
1 K	12.40	6.23	3.42
2 K	13.19	6.64	3.49
4 K	13.68	6.64	3.52
8 K	13.91	5.62	3.54
16 K	14.03	5.81	3.54
32 K	14.09	6.19	3.55

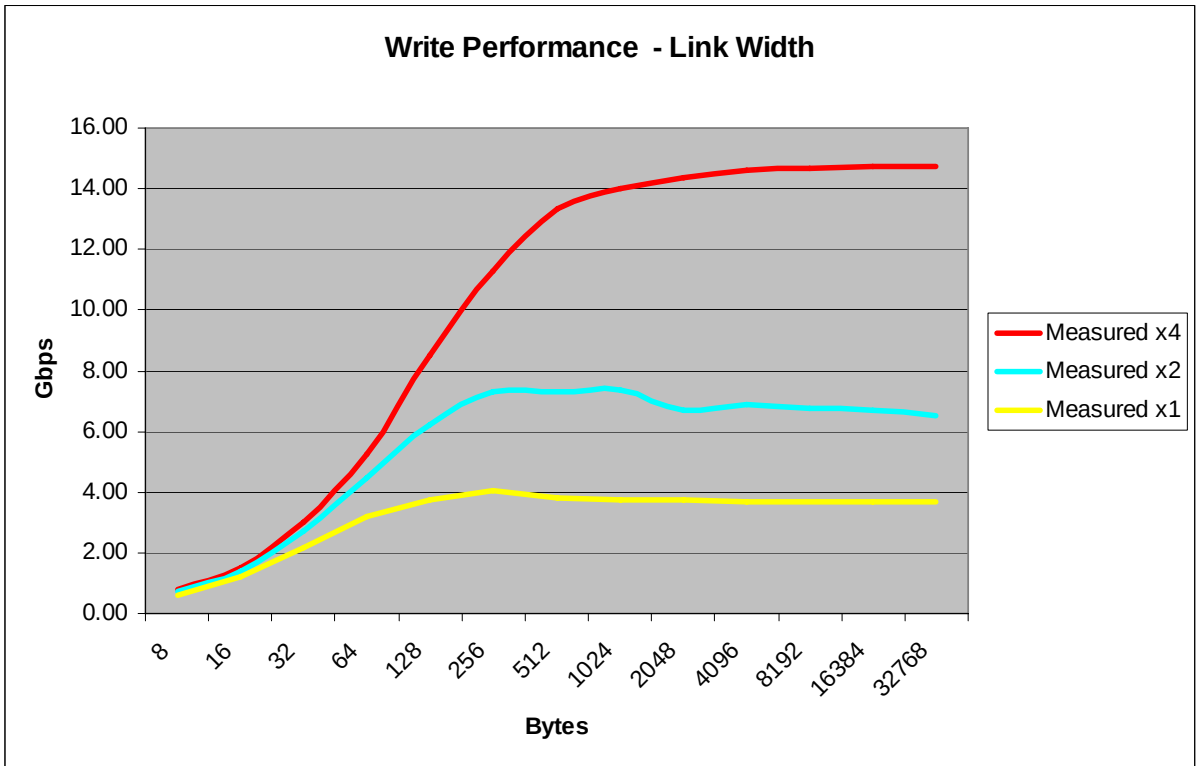

Figure 6. Write Performance x1/x2/x4

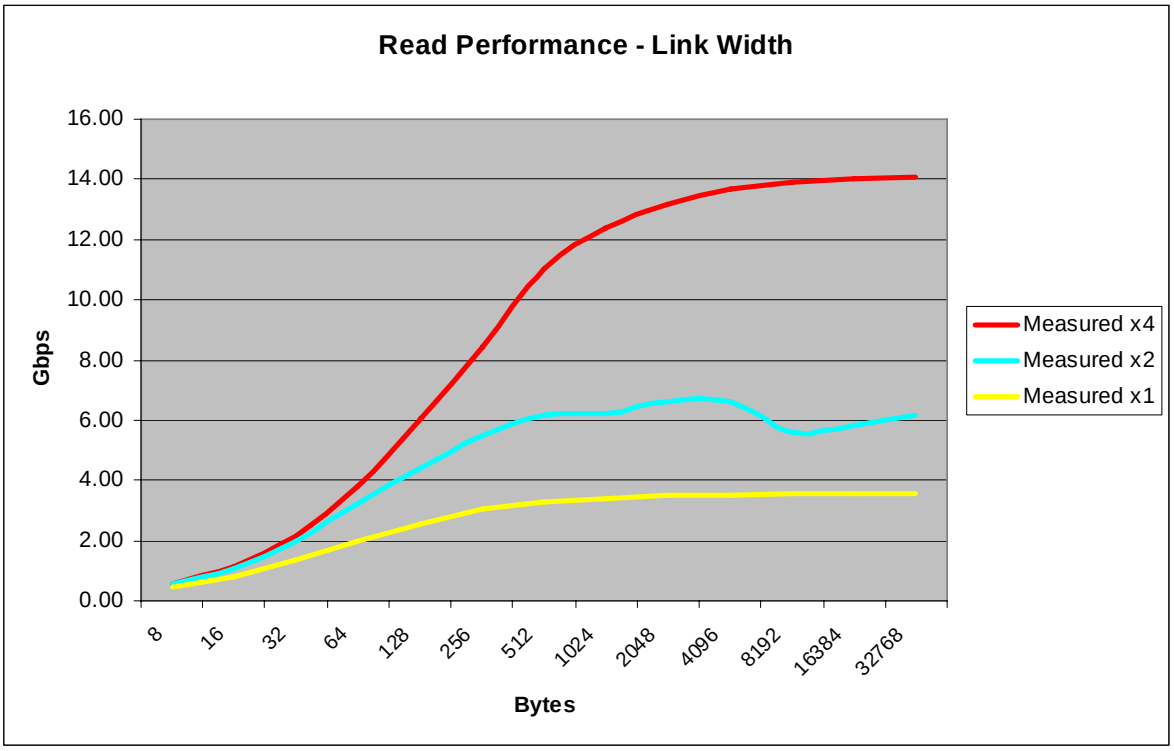

Figure 7. Read Performance x1/x2/x4

7.2 Maximum Payload Size

The size of the payload also has a direct impact on performance. The MSC8156 supports a maximum payload size (MPS) of 256 bytes which is the maximum number of bytes allowed in any TLP. The MPS can be controlled in bits 7:5 of the PCI Express Device Control Register. Transfer sizes greater than 256 bytes are split into multiple transactions. Higher MPS setting means the number of TLPs required to transmit the same amount of data is less. Because there are less TLPs, the TLP and DLLP overhead are also reduced. Bandwidth is reduced when setting the MPS to 128 bytes compared to setting it to 256 bytes.

Table 9 and Table 10 show the results of varying the MPS setting in the write and read tests.

Table 9. Max Payload Size Effect on Write Performance

Bytes	Calculated Bandwidth (Gbps)		Measured Bandwidth (Gbps)	
	MPS = 256 bytes	MPS = 128 bytes	MPS = 256 bytes	MPS = 128 bytes
8	2.91	2.91	0.80	0.77
16	4.92	4.92	1.51	1.46
32	7.53	7.53	3.01	2.92
64	10.24	10.24	5.28	5.12
128	12.49	12.49	8.52	8.36
256	14.03	13.13	11.30	10.24
512	14.42	13.47	13.34	11.96
1 K	14.63	13.47	14.03	12.66
2 K	14.63	13.47	14.39	12.99
4 K	14.63	13.47	14.61	13.22
8 K	14.63	13.47	14.67	13.41
16 K	14.63	13.47	14.72	13.39
32 K	14.63	13.47	14.75	13.40

Table 10. Max Payload Size Effect on Read Performance

Bytes	Calculated Bandwidth (Gbps)		Measured Bandwidth (Gbps)	
	MPS = 256 bytes	MPS = 128 bytes	MPS = 256 bytes	MPS = 128 bytes
8	2	2.00	0.58	0.58
16	3.56	3.56	1.13	1.14
32	5.82	5.82	2.17	2.17
64	8.53	8.53	3.78	3.78
128	11.13	11.13	6.03	6.03
256	13.13	12.34	8.44	8.14
512	13.47	12.34	11.02	10.50
1 K	13.65	12.49	12.40	11.72
2 K	13.65	12.49	13.19	12.43
4 K	13.65	12.49	13.68	12.81
8 K	13.65	12.49	13.91	13.02
16 K	13.65	12.49	14.03	13.13
32 K	13.65	12.49	14.09	13.19

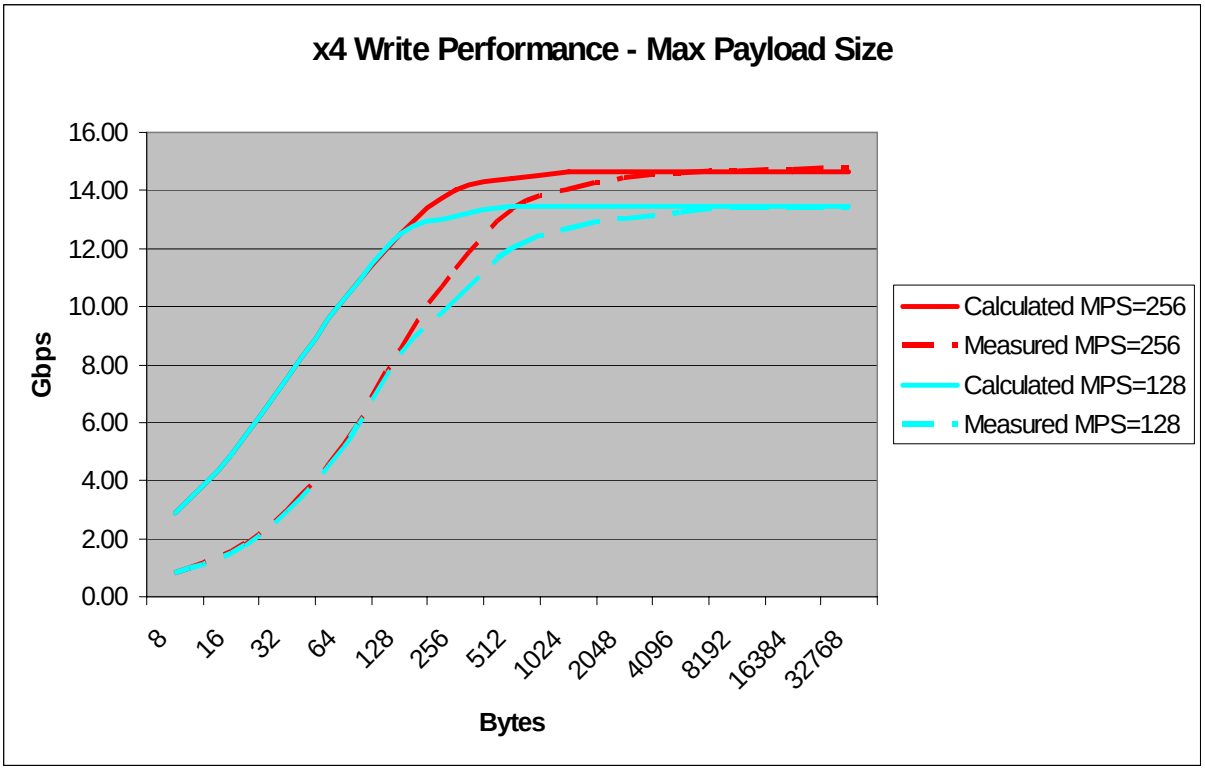

Figure 8. Write Performance - Maximum Payload Size

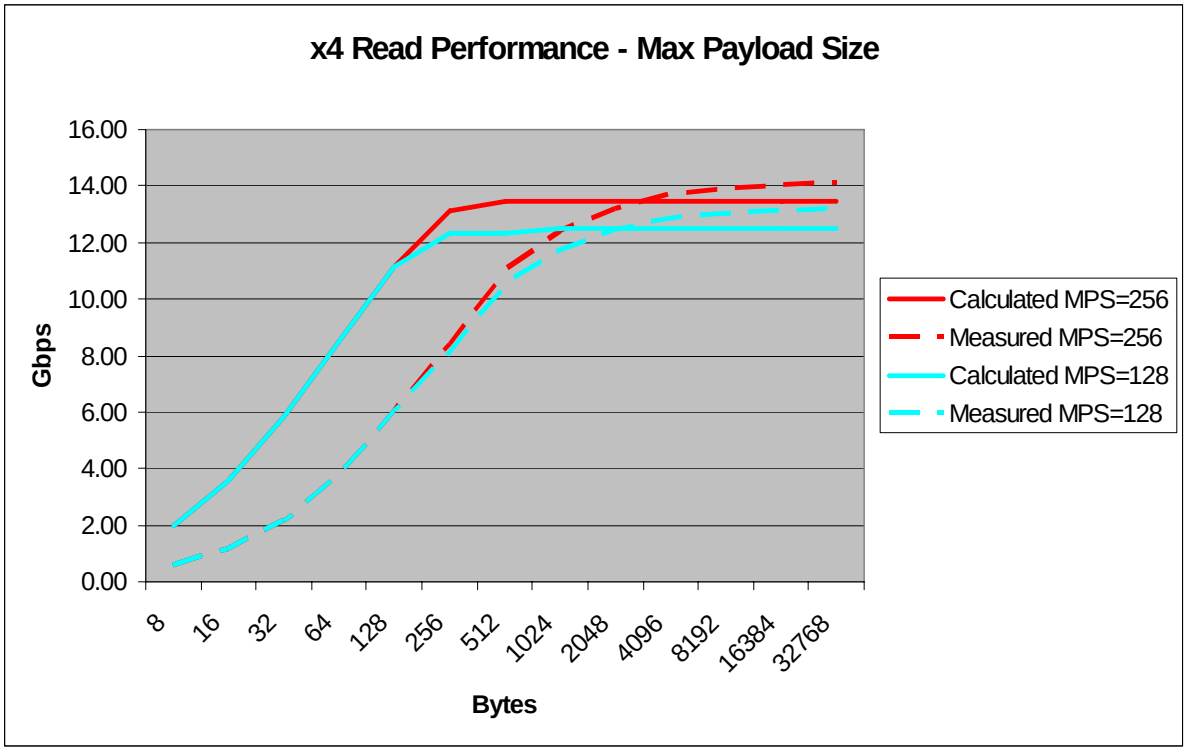

Figure 9. Read Performance - Maximum Payload Size

7.3 Back-to-Back Transfers

Instead of waiting for the receipt of an ACK DLLP after sending the next request, more requests can be inserted during this idle time to improve performance. Higher bandwidth was achieved when memory writes were pipelined using multiple DMA channels. The effects were especially noticeable in smaller transfer sizes. Similarly, pipelining read requests increased throughput because it eliminated the delay for the completions with data except for the first read request. The requester does not have to wait for a CplD before making subsequent read requests. Since the MSC8156 has an initial non-posted header credit of 8, the requester knows it can issue a maximum of 8 non-posted requests before it reaches the credit limit. All tests presented in previous sections activated 8 DMA channels consecutively to perform the memory transfers. This test demonstrates how better bandwidth utilization can be achieved with 8 back-to-back requests instead of a single request, especially for small transfer sizes. Table 11 and Table 12 show the results of activating four DMA channels to increase performance.

Table 11. Back-to-Back Write Performance

Bytes	Calculated Bandwidth (Gbps)	Measured Bandwidth (Gbps)		
		Single Request	4 back-to-back	8 back-to-back
8	2.91	0.31	0.67	0.80
16	4.92	0.56	1.32	1.51
32	7.53	1.12	2.68	3.01
64	10.24	1.88	4.30	5.28
128	12.49	3.24	6.27	8.52
256	14.03	5.08	8.61	11.30
512	14.42	7.68	11.46	13.34
1 K	14.63	10.32	13.24	14.03
2 K	14.63	12.05	13.96	14.39
4 K	14.63	13.39	14.36	14.61
8 K	14.63	14.05	14.56	14.67
16 K	14.63	14.40	14.66	14.72
32 K	14.63	14.58	14.72	14.75

Table 12. Back-to-Back Read Performance

Bytes	Calculated Bandwidth (Gbps)	Measured Bandwidth (Gbps)		
		Single Request	4 back-to-back	8 back-to-back
8	2.00	0.14	0.41	0.58
16	3.56	0.29	0.76	1.13
32	5.82	0.53	1.36	2.17
64	8.53	0.96	2.44	3.78
128	11.13	1.55	4.01	6.03
256	13.13	2.45	6.00	8.44
512	13.47	4.24	9.10	11.02
1 K	13.47	6.45	11.01	12.40
2 K	13.47	8.92	12.41	13.19
4 K	13.47	11.04	13.19	13.68
8 K	13.47	12.33	13.65	13.91
16 K	13.47	13.20	13.91	14.03
32 K	13.47	13.66	14.04	14.09

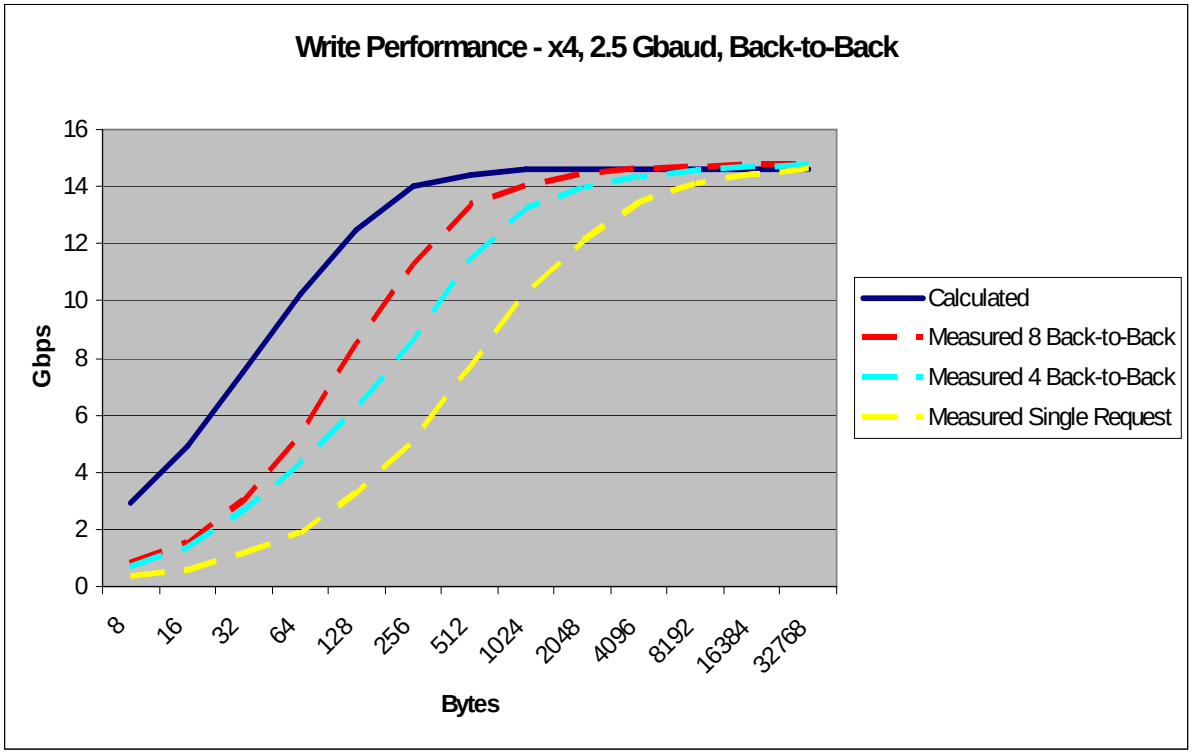

Figure 10. Write Performance - Back-to-Back Requests

Table 13. Read Performance - Back-to-Back Requests

8 MSC8157 PCI Express Performance Measurements

Measurements were performed using two MSC8157 ADS boards connected via an AMC connector. The same as MSC8156 PCI Express performance measurement, one MSC8157 was configured as the RC and the other as the EP. The RC device performed both outbound memory write and read transfers to/from the EP. Transfer sizes range from 8 bytes to 32 KB.

Table 14 shows the MSC8157 ADS switch settings for the RC and EP devices.

Table 14. MSC8157 ADS Switch Settings

MSC8157 switch	RCW BIT	Value	Comments
SW5.7	SCLK1	1	
SW5.8	SCLK0	1	
SW6.1-SW6.7	SP	0000001	link width x4
		0000101	link width x2
		0000110	link width x1
SW10.4	RC	0/1	One ADS in RC mode while the other in EP mode

The MSC8157 PCI Express performance are measured at different link widths x1, x2, and x4. The link width is determined by the reset configuration word RCW[SP] bits. Wider link width provider higher performance.

The MSC8157 pex_performance project requires CodeWarrior for StarCore v10.2.x or higher to run the bandwidth test. The same project is used for both RC and EP devices. Use the following steps to run the test:

1. Ensure that the two MSC8157 ADS boards are connected via the AMC connector. The AMC X-over card was used in this application note.
2. J16 is open on both ADS boards.
3. Configure the RC and EP ADS switches as shown in Table 14.
4. Open two CCS connections for the two ADS boards. Configure the Port and uTAP for each CCS. The uTAP number can be found on the ADS board and the Port number is the last 4 digits of the uTAP number.
5. Open two CW IDEs, and open the pex_performance project in each of them.
6. Set CCS server to Manual launch with corresponding server IP and port number in each CW project.
7. Start running both CW projects at the same time.

Table 15 and Table 16 shows the MSC8157 PCI Express on write and read performance with different link width settings.

Table 15. MSC8157 PCI Express Write Performance with Different Link Width and Packet Size

Data Size (Bytes)	Measured: Two MSC8157 Connected			Calculated: PCI Express Ideal		
	x1	x2	x4	x1	x2	x4
32K	7.12	14.2	27.32	7.31	14.63	29.26
16K	7.13	14.2	27.32	7.31	14.63	29.26
8K	7.11	14.2	27.32	7.31	14.63	29.26
4K	7.13	14.2	27.32	7.31	14.63	29.26
2K	7.13	14.2	27.32	7.31	14.63	29.26
1K	7.13	14.23	27.3	7.31	14.63	29.26
512	7.13	14.21	27.31	7.21	14.42	28.85
256	7.13	14.2	27.28	7.01	14.03	28.05
128	6.43	12.12	18.63	6.24	12.49	24.98
64	5.34	8.48	12.02	5.12	10.24	20.48
32	3.97	5.19	6.82	3.76	7.53	15.06
16	2.28	3.27	3.89	2.46	4.92	9.85
8	1.32	1.72	2.06	1.45	2.91	5.82

Table 16. MSC8157 PCI Express Read Performance with Different Link Width and Packet Size

Data Size (Bytes)	Measured: Two MSC8157 Connected			Calculated: PCI Express Ideal		
	x1	x2	x4	x1	x2	x4
32K	6.6	12.9	24.78	6.83	13.65	27.31
16K	6.6	13.06	24.76	6.83	13.65	27.31
8K	6.6	12.91	24.75	6.83	13.65	27.31
4K	6.6	12.9	24.71	6.83	13.65	27.31
2K	6.6	12.94	24.46	6.83	13.65	27.31
1K	6.6	13.02	24.66	6.83	13.65	27.31
512	6.58	13	24.64	6.74	13.47	26.95
256	6.43	11.72	19.53	6.56	13.13	26.26
128	5.42	10.25	15.08	5.57	11.13	22.26
64	4.49	7.29	9.48	4.27	8.53	17.07
32	2.97	3.92	5.1	2.91	5.82	11.64
16	1.65	2.05	2.45	1.78	3.56	7.11
8	0.86	1.09	1.26	1.00	2.00	4.00

9 Summary

This following list summarizes ways to improve PCI Express performance.

- Because read transfers generate more overhead than write transfers, it is preferable to generate write transactions instead. For example, suppose two MSC8156 or MSC8157 devices are connected via PCI Express and Device A needs to read data from Device B. Instead of the Device A sending read requests to the Device B, Device A should configure the HSSI DMA controller of Device B to move data out (generate writes). Higher performance can be achieved by pushing data out.
- Set the MSC8156/MSC8157 PCI controller maximum payload (MPS) size to 256 bytes to increase packet efficiency. The MSC8156 and MSC8157 support a maximum payload size of 256 bytes. As a receiver, configure the Device Control Register[7:5] = 001 for 256 bytes. As the MPS increases, fewer TLPs are required to transfer a given amount of data.
- Set the link width to x4 for faster packet transmission. For MSC8156, this mode is selected in the reset configuration word and reflected in the Reset Configuration Word Low Register[28:24] = 01101. For MSC8157, the PCI Express link width is set in the Reset Configuration Word Low Register[28:22].
- Issue back-to-back requests to for better bus utilization. Streaming requests allows the completer to submit ACK and FC Update packets after a few TLPs have been transmitted. This method minimizes the overhead of ACK and FC Update packets and maximizes bandwidth utilization.

10 Revision History

Table 17 provides a revision history for this application note. Note that this revision history table reflects the changes to this application note template, but can also be used for the application note revision history.

Table 17. Document Revision History

Rev. Number	Date	Substantive Change(s)
0	4/2010	First public version.
1	11/2011	The MSC8157 family was added throughout the document.

How to Reach Us:

Home Page:
www.freescale.com

Web Support:
http://www.freescale.com/support

USA/Europe or Locations Not Listed:
Freescale Semiconductor, Inc.
Technical Information Center, EL516
2100 East Elliot Road
Tempe, Arizona 85284
+1-800-521-6274 or
+1-480-768-2130
www.freescale.com/support

Europe, Middle East, and Africa:
Freescale Halbleiter Deutschland GmbH
Technical Information Center
Schatzbogen 7
81829 Muenchen, Germany
+44 1296 380 456 (English)
+46 8 52200080 (English)
+49 89 92103 559 (German)
+33 1 69 35 48 48 (French)
www.freescale.com/support

Japan:
Freescale Semiconductor Japan Ltd.
Headquarters
ARCO Tower 15F
1-8-1, Shimo-Meguro, Meguro-ku
Tokyo 153-0064
Japan
0120 191014 or
+81 3 5437 9125
support.japan@freescale.com

Asia/Pacific:
Freescale Semiconductor China Ltd.
Exchange Building 23F
No. 118 Jianguo Road
Chaoyang District
Beijing 100022
China
+86 010 5879 8000
support.asia@freescale.com

For Literature Requests Only:
Freescale Semiconductor
Literature Distribution Center
+1-800 441-2447 or
+1-303-675-2140
Fax: +1-303-675-2150
LDCForFreescaleSemiconductor
@hibbertgroup.com

Document Number: AN3935
Rev. 1
11/2011

freescale™
semiconductor