STARGEN

SG2010- A4 Errata

Revision 2

Date: February 2006

Table of Contents

Bundled Link Operation

Bundled links are subject to the following limitations:

The Link Disable bit must not be set for an operational link in a bundle if traffic may be flowing over that link. Either ensure traffic is quiesced, or take the port (both links) down by clearing the Traffic Enable bit for one of the links.

A single link in an operational bundle cannot be disconnected when all four differential pairs in the same link are disconnected within 1ms of each other. If a cabling solution is desired, a custom cable that disconnects both links at the same time is required.

A fatal error may occur under very high error rate conditions, such that multiple transmission errors are received within 70ns of each other. The affected nodes must be reset. With a Bit Error Rate (BER) of 10e-12, the chances of this failure are less than 10e-20. BERs vary based on environmental conditions but are typically better than 10e-12.

Fragile Link Operation

Link going fragile: When a link is going fragile and a transmission error occurs during the first linked frame received by the SG2010, the SG2010 may clear the Traffic Enable bit (link goes down). Further, if the link partner sends a frame in the timing window between the re-synchronization after fragile and the link going down, a fatal sequence number mismatch occurs resulting in data corruption and possible hang. In the latter case the affected nodes must be reset. Given a BER of 10e-12, the odds of this happening *when a link is going fragile* is less than 10e-10. Links should go fragile very infrequently, if at all.

Single Link Operation

Credit, data corruption may occur on a link when software clears and then sets the Traffic Enable bit. In order for this to happen, the link partner must send an Update Credit frame in a 200ns window after the Traffic Enable bit is cleared. When software sets the Traffic Enable bit and traffic is directed to the link as the bit is set, data and credit corruption may occur. To avoid this, when software clears and then sets the Traffic Enable bit, it needs to ensure there is no traffic directed to that link when the Traffic enable bit is set.

The SG2010 returns a PCI target retry during port down, instead of terminating the transaction with either TRDY# or target abort, when all of the following are true:

- SG2010 is a leaf and TEN=0 and/or LP_TEN=0 on the output port
- the SG2010 is the target of a PCI write to be translated to an address-routed write frame using that port

The SG2010 returns target retry on port down, instead of terminating the transaction with TRDY# or target abort, for a PCI write to be translated to a multicast write, where the multicast write is programmed to go out both ports, but only one port has TEN=0 and/or LP_TEN=0. If both ports are disabled, or if the multicast is only programmed to go out one port, the SG2010 responds appropriately.

Link Partner and Maskable Reset Modes

The link is not always taken down during link partner reset. When the SG2010 resets and the link partner does not, the link partner does not clear its Traffic Enable bit as it should under the following conditions:

- The SG2010 is a root. For this to happen, the root's link partners must be programmed to block the propagating maskable reset.
- The SG2010 is a leaf and receives a You Are frame before it is able to send an I Am frame.

This would happen in small fabrics, for example, when the root is connected to the leaf. For this to happen, the leaf must undergo a non-propagating node reset.

If routed traffic is sent across this link, a fatal error may occur. To avoid this, before a partial fabric reset occurs, software can take the link down by clearing the Traffic Enable bit on one of the link partners to prevent traffic from flowing across the link.

Maskable Reset Mode has the following two limitations:

Leaf SG2010 not enumerated or recognized when the Maskable Reset Mode bit is set. When the SG2010's Maskable Reset Mode bit is set, the SG2010 does not reset when it receives a maskable reset comma. It does return an I Am frame if its Traffic Enable state changes. The SG2010 may send the I Am frame too quickly so that the root does not detect it. Because the root does not detect the I Am frame, it never sends a You Are frame and the leaf SG2010 is not enumerated. This is only an issue when a leaf SG2010 is connected directly to a root and the Maskable Reset Mode bit is set. Software can enumerate the leaf by sending a You Are frame using the root's software generated frame function.

Maskable Reset Mode only works fully for the first maskable reset comma received after the link comes out of reset. For subsequent maskable reset commas, the SG2010 may erroneously take the link down. The SG2010 is not reset, however. Software must bring the link up by writing the Traffic Enable bit of either the SG2010 or the link partner with a 1. Software may assign a Fabric ID to the leaf SG2010 by generating a You Are frame frame using the SG2010 link partner's SGF function.

Events

Multiple events may be dispatched during Segment Table invalidation. When the Segment Table invalidates multiple entries on reception of a path event, the SG2010 erroneously asserts the event for every entry invalidated, instead of asserting it once. This does not have an effect in polled mode. However, in list mode this may result in multiple event frames being dispatched. To avoid this, use polled mode or mask the event; otherwise be able to tolerate multiple Segment Table events.

EMU Counters are not modified by write messages when PCI Master Enable is low. When the Bridge function is disabled and the SG2010 receives a write frame addressing an EMU counter register, if the Gateway function's Master Enable bit is low the EMU counter is not modified. EMU counters are still modified as expected when event frames are received. For expected usage models, it is expected that the PCI Master Enable bit would be high so that the message would be written to local memory.

Software Generated Frames

If a transmission error occurs on a Set_Credit, I_Am, or You_Are frame sent via the Software Generated Frame (SGF) mechanism, the frame may not be re-sent. The transmission error must occur within a 64ns window. Thus, with a Bit Error Rate (BER) of 10e-14, the probability of dropping a single SGF frame is 1.6E-16. BERs vary based on environmental conditions but are typically better than 10e-14. Software can send these frames twice to reduce the probability that the frames are lost. Sending these frames multiple times has no side effect. Software needs to ensure that the frames are not resent within the transmission error timeout window. A minimum delay of 5 µsec should be used for a link with four operating differential pairs, 10 µsec for a link with two operating differential pairs, and 20 µsec for a link with one operating differential pair. Software can also read the Link State Table (in the case of Set_Credit and I_Am) and the Fabric ID register (for the case of You_Are) of the SG2010's link partner to ensure that the frames were received. Note that the sending of Set_Credit and I_Am frames using SGF is outside the typical usage model and is generally not done. You_Are frames are only sent when software desires to enumerate a fabric; typically this is done by hardware.

Parallel Rom Port

The parallel ROM port is not functional when the PCI bus is operating at 66 MHz. It is also not functional when ACCTIME in the ROM Setup register is set to a value other than 0. When the ROM interface is configured to use the minimum access time (ACCTIME=00b, or 17-33MHz PCI clock cycles), and the PCI bus is operating at 33MHz (M66ENA is pulled low) the ROM interface is fully functional. This is functional through either the Expansion ROM BAR or the ROM register interface.

The following combinations using the ROM Register interface have a limited workaround:

- 33MHz PCI, ACCTIME != 00b, dword access (PRBYTE=0)

- 66MHz PCI, ACCTIME= [01b or 11b**], dword access (PRBYTE=0)

Register Interface Workaround: The address that is programmed into the ROMADDR field must be the byte location before the desired address. For example, to read four bytes starting at location 8, the ROMADDR should be programmed to be 7. If the desired address is 0, a ROMADDR[15:0] value of all 1's must be used, as the value wraps at a 16-bit boundary. ROMADDR[23:16] should match bits [23:16] of the desired address. A dummy read must be performed as the first parallel ROM access.

Note that the ROMADDR need not be dword aligned. The specification must be corrected to say that when PRBYTE=0 four bytes are accessed starting at the first address. The specification incorrectly says that ROMADDR[1:0] are ignored when PRBYTE=0.

The following combinations using the Parallel ROM Expansion BAR interface have a limited workaround:

- 33MHz PCI, ACCTIME != 00b
- 66MHz PCI, ACCTIME = [01b or 11b**]

ROM Expansion BAR Workaround: The SG2010 erroneously increments the address presented by the user by one byte when it drives it to the parallel ROM. There are two methods to compensate for this early increment. The parallel ROM may be programmed with all data shifted up by one byte or alternatively, the data may be programmed in their proper byte locations but the data returned by the SG2010 is shifted down by one byte.

The second method requires reading two dwords in order to get one dword of data. The first read carries the first three bytes in byte lanes 1 through 3, and the second read carries the fourth byte in byte lane 0.

This workaround is not effective at the 64 Kbyte boundary, as the fourth byte using either method will return the first byte location in the ROM. Thus, the last location before the 64 Kbyte boundary must not be used for valid data. This is because the offset wraps at a 16-bit boundary.

As in the register interface workaround, a dummy read is must be performed as the first parallel ROM access.

All other combinations are disallowed and must not be used:

- 33MHz PCI, ACCTIME != 00b, byte access (PRBYTE=1) using the register interface
- 66MHz PCI, byte access (PRBYTE=1) using the register interface
- 66MHz PCI, ACCTIME = [00b or 10b]

** ACCTIME=11b at 66Mhz has the same timing as ACCTIME=10b (33MHz using ACCTIME=11b has the correct specified timing)

BDSEL and LSTAT Bits

The BDSEL and LSTAT bits in the Chip Status register do not reflect the status of their respective pins, but instead will always return 0 when read.

Write combining error case

When the SG2010 is the bus master of a memory write transaction, and it receives a target abort or a master abort in response, and at the same time is combining additional data with the data associated with that write transaction, credit corruption and other unexpected behavior may occur. Repeated occurences of this condition may eventually result in a system hang.

This error case only affects memory write transactions. This condition is not an issue if a target abort or master abort can never occur in response to a memory write, or if such a condition results in a system error.

If this condition is a normal error case in an operating system, the Write Combine Enable bit should be cleared to 0 to turn off write combining.

PCI-PCI Bridge Dual Address Cycle Errata

Under certain conditions, a Starfabric system utilizing address routing with 64-bit dual address cycles can fail to initiate the transaction on the target bus. It can also cause the SG2010 to work improperly.

The problem can occur if the lower 32-bit portion of a dual address cycle is in the BAR0 memory region of the target bridge's Gateway function. If bits 31:15 of BAR0 and a 64-bit address frame match, the BAR0 register access is performed instead of forwarding the frame onto the PCI bus. This erratum can occur in both root and leaf bridges.

When an Address Routed frame enters the SG2010 from the fabric, the address is checked against the Gateway's BAR0 and a register access is performed if they match. This is done in both root and leaf bridges. BAR0 is a 32-bit BAR and is 32KBytes large. If the Address Routed frame has a 64-bit address (DAC), the correct behavior is that the frame not match BAR0 and be forwarded to the PCI bus. With the erratum, the Gateway claims Address Routed transactions whose lower 32-bit portion of the address matches BAR0.

1. Disable the Gateway's memory address decoding by clearing the Memory Enable in the Gateway Command Configuration register. This is located at Gateway Configuration offset 4 (or offset 0x5994 in CH255 space), bit 1. This will prevent the Gateway from decoding and claiming incoming frames from the fabric.

2. For leaf bridges, hide the entire Gateway with the Gateway Configuration Mask bit in the Chip Control register. This is located at Gateway configuration offset 94 hex (i.e. 0x5994), bit 15. This effectively disables the Gateway function.

This erratum only applies to address routed frames. It does not affect 64-bit address accesses from the PCI bus. Only SG2010s with both the Bridge and Gateway functions active are affected. When the Bridge is disabled, the address routed frame is not accepted. If the Gateway address decoding is disabled, BAR0 address compares are not performed.

EMU Counter Registers

The PCI Local Bus Specification Revision 2.2 states "The target of an access where no byte enables are asserted must complete the current data phase without any state change." The EMU counter registers in the SG2010 violate this PCI protocol as these registers perform their normal operation when a fully byte disabled dword write transaction targets them.

Writing to all EMU counter registers causes register operation to occur, even when all bytes are disabled in the dword write transaction. This affects both EUMx Increment Counter registers and EMU Decrement Counter registers, which are located at offsets 0x5200h – 0x5244h in the SG2010's CSR space. Do not write to these registers unless normal register operation is intended.

Silicon Revision ID

The value of the Silicon Revision ID register was not changed between the SG2010-A3 and SG2010-A4 revision changes. Reading the Silicon Revision ID register, located at offset 0x000Ch in the SG2010-A4's StarFabric Component Header registers, returns a value of 0x00000002h. This dword value is the same as the SG2010-A3's Silicon Revision ID.

Use the SG2010's serial ROM interface to preload a unique value such as 0x00000003h in the StarFabric Component Header register called SFC Programming Interface ID, which is located at offset 0x001Ch. This register is initialized to 0x00000000h at reset, but is loadable only through the serial ROM interface. The user can utilize this register to uniquely identify the SG2010-A4 from the SG2010-A3 if it is critical in their application.

Down Stream Memory

To enable downstream address-routed memory transactions, the Root SG2010's IO Enable bit (PIO) in the port map table command registers (Port 0 - 0x4F24, bit 0, Port 1 - 0x4F2c, bit 0) needs to be set. Downstream IO and memory transaction are both enabled by the port map tables' PIO bit. The memory enable bit (PMEM), which is bit 1 in the port map tables' command register, does not effect downstream IO or memory transactions. This erratum only effects downstream address-routed memory transactions from the root SG2010 bridge. This issue is not observed

with the SG1010's port map tables or when the downstream transaction is completed on a Leaf SG2010's PCI bus.

Potential Workarounds

Potential workarounds for this erratum include:

1. Enable IO transactions throughout the system.

2. Using configuration transactions set the PIO bit in the command register (0x04, bit 0) on the P2P bridge device(s) connected to the root SG2010.

3. Using memory transactions set the PIO bit in the port map table command register (Port 0 - 0x4F24, bit 0, Port 1 - 0x4F2c, bit 0) on each active port of the root SG2010.