CN54XX/CN55XX/CN56XX/CN57XX Pass 1 Known Issues

Version: 1.4

Last modified: 24 Sep 2008

Cavium Confidential For
Geoffrey Collyer
Bell Labs, Alcatel-Lucent
05/25/2011

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 05/25/2011

1 Introduction

The following list documents the known issues with OCTEON Plus CN54XX / CN55XX / CN56XX / CN57XX documentation and silicon pass 1.0 and 1.1. The change bars in this document indicate differences from version 1.3 of this document.

All issues in this document apply to all members of the CN54XX, CN55XX, CN56XX, and CN57XX families unless stated otherwise. The term "OCTEON Plus" in this document refers to those parts.

2 Open Issues List

2.1 CN54XX/CN55XX/CN56/57XX Hardware Reference Manual

2.1.1 Coherent Memory Bus, Level-2 Cache Controller, DRAM Controllers Chapter

2.1.1.1 LMC_PLL_CTL[CLKF] Range

CN56XX/57XX Hardware Reference Manuals version 0.98 and lower specify that `LMCn_PLL_CTL[CLKF]` is 12 bits wide, and provide no guidance on the maximum value. The maximum value is 128.

2.1.1.1.1 Workaround

Always use `CLKF` values ≤ 128.

2.1.2 Signal Descriptions Chapter

2.1.2.1 SMI Interface Power

The SMI and MII interfaces are both powered from the VDD_MII supply pins. If 2.5 V is supplied on VDD_MII, the SMI*n*_MDC and SMI*n*_MDIO pins are of type CMOS25B. If 3.3 V is supplied on VDD_MII, the SMI*n*_MDC and SMI*n*_MDIO pins are of type CMOS33B.

2.1.2.1.1 Workaround

Be aware that the MII and SMI ports are powered from a common supply.

2.1.3 USB Chapter

2.1.3.1 USBN_CLK_CTL Bits <16:13>

Early versions of the CN56XX/CN57XX Hardware Reference Manual (revision 0.98 and lower) incorrectly describe bits 16 through 13 of the USBN_CLK_CTL register.

The correct description is shown below.

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 03/25/2011

Bit Pos	Field Name	Type	Reset Value	Typical Value	Field Description
<16>	—	RAZ	—	—	Reserved
<15:14>	P_RTYPE	R/W	0x0	0x0	PHY reference clock type: 0x0 12 MHz crystal between USB_XO and USB_XI pins 0x1 Reserved 0x2 12 / 24 / 48 MHz 2.5 V clock oscillator connected to USB_XO. USB_XI should be tied to GND. 0x3 Reserved
<13>	P_COM_ON	R/W	0x1	0x1	0x1 Force USB-PHY XO Bias, Bandgap, and PLL to remain powered in Suspend mode. 0x0 The USB-PHY XO Bias, Bandgap, and PLL are powered down in Suspend mode. This field must be set when POR is active.

2.1.3.1.1 Workaround

Be aware of the corrected definition.

2.1.4 System Management Interface (SMI) Chapter

2.1.4.1 SMI/MDIO Read Sampling

In revision 0.98C and later of the CN54/5/6/7XX Hardware Reference Manual, section 20.1, the figure that describes the sampling point for SMI/MDIO reads has been clarified to indicate that OCTEON Plus is sampling after the second rising edge.

2.1.4.1.1 Workaround

Be aware of the clarification, and configure software to use the appropriate settings for `SMI_CLK[SAMPLE_HI]` and `SMI_CLK[SAMPLE]`.

2.1.5 Electrical Specifications Chapter

2.1.5.1 Core Voltage for 600 MHz Parts

In revision 0.98C and later of the CN54/5/6/7XX Hardware Reference Manual, section 27.4, the nominal VDD supply for 600 MHz parts has been changed to 1.1 V.

2.1.5.1.1 Workaround

Conform to the current specification.

2.1.6 cnMIPS™ Cores Chapter

2.1.6.1 CvmMemCtl[CVMSEGENAK] Enable Requires L1 DCache Invalidate

CVMSEG is an OCTEON feature that provides core-local scratchpad memory and space to issue IOBDMA instructions. CVMSEG is located at virtual addresses `0xFFFFFFFFFFFF8000` through `0xFFFFFFFFFFFFBFFF`.

Cavium Confidential For Geoffrey Colyer Bell Labs, Alcatel-Lucent 05/25/2011

Previous revisions of the Hardware Reference Manuals did not completely describe the steps necessary when enabling CVMSEG. It is required that the L1 cache be invalidated if any CVMSEG address might appear in the cache when CVMSEG is first enabled.

Failure to invalidate the L1 cache may result in spurious Parity, Cache Error or Data Access exceptions.

Invalidate the L1 cache by issuing either an L1 DCache Virtual Tag Invalidate or an L1 DCache Invalidate, using the CACHE instruction described in section A.4 of the Hardware Reference Manual.

2.1.6.1.1 Workaround

Invalidate any L1 lines overlapping CVMSEG when enabling CVMSEG.

2.2 Silicon

2.2.1 SGMII Interfaces

2.2.1.1 CRC Generation when Preambles Disabled

When preamble generation is disabled (by setting `GMX0/1_TXn_APPEND[PREAMBLE] = 0`), and the MAC is not in HiGig mode (`GMX0/1_TX_XAUI_CTL[HG_EN] = 0`), and when CRC generation is enabled (`GMX0/1_TXn_APPEND[FCS] = 1`), the first byte of the packet will not be included in the generated CRC.

2.2.1.1.1 Workaround

No workaround is needed in the normal case, when preambles are generated by OCTEON Plus.

Alternately, disable FCS generation and generate the FCS in software, or change the receiver so it can compensate for the excluded byte.

2.2.1.2 HiGig2 Backpressure Messages Not Supported

The HiGig2 protocol defines special packets called “messages” that are used for backpressure in place of pause frames. These messages are not supported, and also can not be generated. Further, OCTEON Plus will not generate 802.3 pause frames when it is in HiGig2 mode.

2.2.1.2.1 Workaround

Use HiGig mode instead of HiGig2 mode if possible. Set the link partner to disable HiGig2 messages, if supported. If the remote can generate 802.3 pause frames, OCTEON Plus will correctly process it. Disable hardware flow control when in HiGig2 mode.

2.2.1.3 Sample Point Selection in 10 Mbps Mode

In 10 or 100 Mbps mode, the SGMII standard replicates the data 100 or 10 times (respectively). The `PCS*_MISC*_CTL_REG[SAMP_PT]` field determines which of the duplicates the receiver keeps.

The receive logic may malfunction in 10 Mbps mode unless the SAMP_PT setting is between the range of 5 to 25, inclusive.

2.2.1.3.1 Workaround

Set `PCS*_MISC*_CTL_REG[SAMP_PT]` to 25 in 10 Mbps mode.

2.2.1.4 SGMII/1000BASE-X to Core Clock Frequency Ratio

Data received from each serial network interface (SGMII or 1000Base-X) goes through a clock domain crossing FIFO to enter the core clock domain. The serial clock rate must be no greater than 5 times the core clock speed for this FIFO to operate correctly. Therefore, when operating at 2.5 Gbaud, the core clock must be at least 500 MHz.

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 05/25/201

2.2.1.4.1 Workaround

Don't use a data rate greater than 5 times the core clock when in SGMII or PICMG mode.

2.2.1.5 External Loopback Inoperative

The SGMII / 1000-BaseX interfaces have an external loopback diagnostic mode, enabled by PCSi_MISCn_CTL_REG[LOOPBCK2], which is supposed to transmit all data received. This mode may loop back corrupted data when there is clock skew between the received and transmitted data streams.

2.2.1.5.1 Workaround

Do not set PCSi_MISCn_CTL_REG[LOOPBCK2]. Instead, send the packets in via the normal OCTEON Plus receive path and loop them out with a diagnostic mode in the software.

2.2.1.6 GMX Enable Sequence

The GMXi_INF_MODE[EN] bit enables the Gigabit MAC for a particular interface. The MAC must be enabled (GMXi_INF_MODE[EN] set to 1) before per-port backpressure is configured by writing to IPD_CTL_STATUS[PBP_EN].

Note that the GMXi_INF_MODE register is an RSL type CSR, where IPD_CTL_STATUS is an NCB type CSR, which means that software needs to read back the setting of GMXi_INF_MODE to ensure the write has completed before writing to IPD_CTL_STATUS.

2.2.1.6.1 Workaround

No workaround is necessary if not using per-port backpressure. Otherwise, ensure that the sequencing described above (enable the GMX before enabling IPD per-port backpressure) is observed.

2.2.2 DDR Controller

2.2.2.1 LMC*_DDR2_CTL[ODT_ENA] Non-functional

When the LMCn_DDR2_CTL[ODT_ENA] (bit 9) bit is set to 0, the DDR_ODT<7:0> lines are still driven with the values in LMCn_RODT_CTL during read cycles.

2.2.2.1.1 Workaround

Set LMCn_RODT_CTL to 0 in addition to LMCn_DDR2_CTL[ODT_ENA] to fully prevent read cycles from driving ODT lines.

The LMCn_RODT_CTL addresses are:

```
LMC0_RODT_CTL    0x0001180088000078
LMC1_RODT_CTL    0x00011800E8000078
```

2.2.2.2 Accesses to Nonexistent DDR Memory

If software sets up a memory mapping or makes an xkphys access within the DDR address range to memory that is not connected, the LMC may generate spurious ECC errors. Further, the LMC FIFOs may also get out of synchronization, resulting in all further DDR accesses being corrupted.

Prefectches are also a sufficient memory access to trigger this problem. Software needs to ensure that no prefetches are made against DDR regions that do not have attached memory.

The DDR physical address ranges, expressed in xkphys addresses, are:

0x8000 0000 0000 0000 to 0x8000 0000 0FFF FFFF

0x8000 0004 1000 0000 to 0x8000 0004 1FFF FFFF

0x8000 0000 2000 0000 to 0x8000 0003 FFFF FFFF

Once the condition has occurred, the only solution is a complete reset of the LMC (by resetting the chip).

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 05/25/2011

2.2.2.2.1 Workaround

Ensure that software does not make any accesses to non-existent DDR addresses, either through xkphys for supervisor mode, or through user-mode mappings.

2.2.3 JTAG

2.2.3.1 DC JTAG on SerDes Wires

During DC JTAG operation, some finite resistance may appear in parallel with the SERDES transmit and receive pins. This will limit the output swing in DC JTAG mode.

For pass 1.1 parts, the SERDES transmit outputs can not be driven low by DC JTAG.

2.2.3.1.1 Workaround

Be aware of the limited output swing. Do not attempt to drive 0 out on the SERDES transmit pins via DC JTAG.

2.2.4 SERDES Interfaces

2.2.4.1 Transmit Resistor

The SERDES transmit lane resistor is incorrectly implemented. This will affect device reliability as well as functionality. OCTEON Plus will not be able to detect external PCIe devices.

2.2.4.1.1 Workaround

Add an external 68 Ω resistor connected to VDD on each transmit line to improve the functionality, but not the reliability. This issue is fixed in pass 1.1.

2.2.5 PCI Express Interface

2.2.5.1 PCIE DMA Engines

The PCI Express DMA engines can malfunction when DMA instructions with a large number of scatter/gather pointers are given.

The total number of pointers (as indicated by summing the `NFST` and `NLST` fields of the DMA instruction) must be less than 16.

When running an external-only transfer, the total number of read pointers (as indicated in the `NFST` field of the DMA instruction) must be less than 12.

When running an outbound, inbound, or external-only transfer, the number of pointers in the LAST POINTERS list (as indicated in the `NLST` field) must be less than 12.

2.2.5.1.1 Workaround

Observe the limits above when submitting commands to the DMA engines. This may require splitting commands. If using the `PCIE_DMA_CNTn` registers to signal completion, the `CA` field should be clear in all but the last of the DMA instructions submitted to the DMA engine.

2.2.5.2 PCIE Inbound DMA

An inbound DMA transfer moves data from the PCIe memory space to the L2/DRAM space. A larger inbound transfer may create the need for multiple PCIe reads and/or multiple L2/DRAM transfers, depending on the number of bytes transferred and the alignment of the components. A particular PCIe read operation may read up to 2K (or fewer depending on the setting of the Maximum_Read_Request_Size in the PCI Express capabilities device control register and the setting in `NPEI_CTL_STATUS2[MRRS]`). Also, PCIe read requests may not cross 4K alignment boundaries. Finally, the OCTEON Plus DMA engines generate 8-

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 05/25/2011

byte aligned PCIe requests in all cases, possibly masking off (internally) some of the returned bytes to make the alignment match the request.

In certain circumstances, when both of the conditions below are met, inbound DMA may lose data.

The first condition is when more than 4 PCIe transactions are required to satisfy the read components of a DMA operation (note that more PCIe read transactions may be required than specified by `NLST` for the reasons outlined above), and any except the last of the transfers is not a multiple of 8 bytes of useful data (thus, non-8B-aligned and/or not a multiple of 8B long).

The other condition is when at least 3 cache blocks writes are required by the transfer. The number of cache block writes will be at least `NFST`, but may be higher because each transfer is split to 128B natural alignment.

2.2.5.2.1 Workaround

To avoid the first condition, if all PCIe pointers provided in `NLST` are 8B aligned, and all lengths except the last in `NLST` are a multiple of 8B long.

The second condition can be avoided by limiting the length of transfers and the number of local pointers.

Splitting the DMA request into multiple requests will allow any combination of alignments and lengths to be reached, at the cost of generating extra PCIe traffic.

2.2.5.3 PCIE_DMA_CNTn Byte Counting Mode

For outbound and external-only DMA instructions, the CA bit may be set in the DMA instruction to increment one of the `PCIE_DMA_CNTn` registers (determined by the setting of the C bit in the DMA instruction) upon completion of the DMA instruction, possibly causing an interrupt on the PCIe host. `NPEI_DMA_CTL[O_ADD1]` determines whether to increment the `PCIE_DMA_CNTn` register by 1 for each transfer (`NPEI_DMA_CTL[O_ADD1] = 1`), or by the number of bytes transferred (`NPEI_DMA_CTL[O_ADD1] = 0`).

The byte-count mode does not work correctly. `NPEI_DMA_CTL[O_ADD1]` must be set to 1.

2.2.5.3.1 Workaround

Modify host software to handle the fact that only the number of transfers (rather than the number of bytes) information is available. `NPEI_DMAn_INT_LEVEL[CNT]` may need to have its value configured differently to compensate for this, and to interrupt the host in a timely fashion.

2.2.5.4 Configuration Request Retry

Early after reset, a PCIe device may respond to configuration requests with a Configuration Request Retry Status (CRS) message. This message tells the originator of the configuration request to retry the request later.

If OCTEON Plus receives a configuration retry response (CRS) to a configuration request, it will not retry the request. The request will time out and the core will receive a bus error exception.

2.2.5.4.1 Workaround

System software should wait an adequate amount of time (the PCI Express standard recommends at least 1 second) before sending a configuration request.

2.2.5.5 UR/CA Response without ECRC

If a PCIe configuration read request is issued by OCTEON Plus, and an unsupported request (UR) or completer abort (CA) response is received without ECRC, the PCIe controller will hang.

Most PCIe devices default to not send ECRC at power-on, and a naive discovery sequence is likely to generate UR or CA responses.

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 03/25/2011

2.2.5.5.1 Workaround

Build knowledge into the system software of the PCIe geography, or carefully code the discovery sequence to avoid any devices that might send UR or CA responses without ECRC.

2.2.5.6 PCIE Packet Input and Output

The PCI Express packet input and output ports as described in sections 9.4 and 9.5 in the CN56/57XX Hardware Reference Manual are not functional in pass 1.

2.2.5.6.1 Workaround

Move data via other means, including conventional load/store/IOBDMA requests or the PCIe DMA Engines. Cavium Networks is creating a software package to emulate this functionality. Contact your Field Application Engineer for availability.

2.2.5.7 NPEI_LAST_WIN_RDATAn Registers

The `NPEI_LAST_WIN_RDATA`*n* registers return the data from a read initiated via the `NPEI_WIN_RD_DATA` register. The data in this register is supposed to only change after a read from `NPEI_WIN_RD_DATA`, however it changes after any read in the BAR0 region.

2.2.5.7.1 Workaround

Only access `NPEI_LAST_WIN_RDATA`*n* immediately after the `NPEI_WIN_RD_DATA` read is issued, and then only once.

2.2.5.8 ECRC Destroyed in Peer-to-Peer UR/CA Completions

OCTEON Plus zeroes the ECRC field in UR/CA completions that travel peer-to-peer through both PCIe interfaces.

2.2.5.8.1 Workaround

None. Disable ECRC when doing peer-to-peer.

2.2.5.9 UR/CA to PCIe DMA Engines

Unsupported Request (UR) and Completer Abort (CA) are possible responses, generally to PCIe requests that are made to invalid addresses or are otherwise erroneous. If the OCTEON Plus DMA engines receive either of these responses, they may hang and stop processing subsequent instructions.

2.2.5.9.1 Workaround

Ensure that the DMA engines only receive valid addresses.

2.2.5.10 DMA Engines Create Only 64-bit TLPs

The PCI Express specification requires that PCIe memory space reads and writes that reference the lowest 4 GB of address space use a 32-bit TLP type. The OCTEON Plus DMA engines can only generate a 64-bit TLP type. When the PCIe DMA engine generates a read or write request to an address below 4 GB, it will erroneously generate a 64-bit TLP. The type of access (INBOUND, OUTBOUND, or EXTERNAL-ONLY) does not matter.

Note that this only affects the PCIe DMA engines, as described in the "PCIe DMA Engines" subsection in the "PCI Express" chapter of the Hardware Reference Manual. All other OCTEON Plus PCIe access mechanisms create the proper 32-bit TLPs when an address below 4 GB is accessed.

2.2.5.10.1 Workaround

No workaround is required if the target PCIe device can accept 64-bit TLPs to addresses below 4 GB. Likewise, if the DMA engines do not need to access a target device that can not handle 64-bit TLPs, no workaround is needed.

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 05/25/2011

Otherwise, the OCTEON Plus DMA engines can not be used. If the other device has DMA engines, it may make sense to use those.

Also, core-initiated loads, stores, or IOBDMA operations may be used to move the data.

2.2.5.11 DMA Engines can't be Reset

Once the PCIe DMA engines have been configured with a command buffer address by writing to the `NPEI_DMAn_IBUFF_SADDR` register and the engine started by writing a 1 to `NPEI_DMA_CONTROL[DMAn_ENB]`, it can not be reconfigured. Attempting to disable and re-enable the engine by writing to `NPEI_DMA_CONTROL[DMAn_ENB]` will not cause the new value written to `NPEI_DMAn_IBUFF_SADDR` to take effect.

2.2.5.11.1 Workaround

Only change `NPEI_DMAn_IBUFF_SADDR` once, before enabling the DMA engine.

2.2.5.12 PCIe Deskew with XAUI

Either PCIe controller on CN57XX and CN56XX can operate in x4 or x8 mode. In x8 mode, a controller uses two QLMs. PCIe controller 0 is connected to QLM 0 and optionally QLM 1. Likewise, PCIe controller 1 is connected to QLM 2 and optionally QLM 3.

When a PCIe controller is in x4 mode, and the "other" QLM is in XAUI mode, and the PCIe lane swap is enabled (QLMn_REV_LANES is pulled high), then the PCI Express deskew logic in the QLM will malfunction, taking the link down.

The specific configurations that may have a problem are:

QLM1_MODE<1:0> = 01 and QLM0_REV_LANES = 1, or
QLM3_MODE<1:0> = 01 and QLM2_REV_LANES = 1

2.2.5.12.1 Workaround

None needed if the combinations above aren't met. Note that if lane reversal in hardware is not strapped, this will not occur.

If QLM*_REV_LANES is required due to board layout, tie it low anyways and allow the automatic reversal logic built in to PCIe to do the swap instead of strapping it in hardware. Physical lane 0 must be connected for lane reversal detection to work.

2.2.5.13 Detecting PCIe Root Complex Reset Completion

When performing a fundamental PCIe reset in Root Complex (RC) mode, software must wait a certain amount of time for the BIST inside the PCIE block to complete before continuing. The present Hardware Reference Manual implies that polling the `PESCi_CTL_STATUS2[PCIERST]` bit and waiting for it to become zero is an effective way to wait. However, repeatedly reading that register during reset initialization can cause an internal FIFO to get out of synchronization.

2.2.5.13.1 Workaround

Do not poll `PESCi_CTL_STATUS2[PCIERST]`. Delaying for 400,000 core clock cycles will allow sufficient time for the BIST to complete.

2.2.5.14 Expansion ROM Access

OCTEON Plus advertises an end-point expansion ROM BAR. The BIOS in most x86-architecture systems will attempt to map and read an expansion ROM. However, accessing the expansion ROM will generate a condition where the NPEI block is no longer able to access other resources inside OCTEON Plus. Thus, configuration space accesses will still work, but not CSR accesses or memory reads.

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 03/25/2011

2.2.5.14.1 Workaround

Certain BIOSes can be configured to bypass reading of expansion ROM. If BIOS source is available, modify the BIOS to prevent access to the expansion ROM.

Otherwise, initiating a complete PCIe-defined hot reset through the bridge upstream of OCTEON Plus (preserving the configuration space registers written by the BIOS) will recover if the BIOS has made an expansion ROM access. Example code is available from Cavium Networks that shows how to do this for x86 platforms.

2.2.5.15 PCIe Soft Reset from Remote

When in PCIe endpoint (EP) mode, soft reset to OCTEON Plus can be initiated by a store to CIU_SOFT_RST. Either a remote PCIe device or an OCTEON Plus core can initiate the store.

When a remote PCIe device initiates the store, an OCTEON-internal flow control credit is lost. After the first soft reset is issued from a remote device, performance will be degraded. After the second soft reset is issued, OCTEON Plus will no longer be able to complete PCIe memory space stores.

2.2.5.15.1 Workaround

One workaround is to have an OCTEON Plus core initiate the reset instead of the PCIe remote device. Another, probably more effective workaround is to initiate a complete a PCIe-defined hot reset, which will cause the PCIe channel to go down and reset all of OCTEON.

2.2.5.16 Unused PCIe Controllers

OCTEON Plus has two PCI Express controllers that operate independently. In some systems, both controllers may not be needed. When a controller is not needed, the primary QLM associated with the controller (QLM 0 for PCIe controller 0, QLM 2 for PCIe controller 1) should not need a reference clock.

However, if no reference clock is provided to the QLM, the PCIe controller will not be clocked, and will not come out of reset correctly. Symptoms include reporting BIST failures, and asserting PCI emulated interrupts.

Due to the way that interrupt routing works, if one of the PCIe controllers asserts an emulated interrupt continuously, the other controller will not be able to use that interrupt line to signal emulated interrupts received from external devices.

2.2.5.16.1 Workaround

If possible, provide a reference clock and power to a QLM that is connected in PCIe mode, even if it is otherwise unused.

Otherwise, at boot time, observe which of CIU_INT*x*_SUM*y*[PCI_INT<3:0>] are set, and make sure the corresponding CIU_INT*x*_EN*y*[PCI_INT<3:0>] are always clear. If one or more of the CIU_INT*x*_SUM*y*[PCI_INT<3:0>] are not constantly set, they can be used.

Use PCIe MSI interrupts instead whenever possible.

Also note that NPEI_BIST_STATUS failures on a non-clocked port should be ignored.

2.2.6 Random Number Generator

2.2.6.1 RNG Output Constant

Reads from the RNG return a constant value.

2.2.6.1.1 Workaround

Use a software random number generator. Cavium Networks will be providing an alternate library, contact your field application engineer for details and availability.

Cavium Confidential For Geoffrey Collyer, Alcatel-Lucent Bell Labs, 05/25/2011

2.2.7 General

2.2.7.1 MSC_CLKOUT

OCTEON Plus provides the MSC_CLKOUT debugging output to allow observation of the core clock frequency. The MSC_CLKOUT output is not functional in pass 1.

2.2.7.1.1 Workaround

Do not depend on the MSC_CLKOUT functionality for verifying core clock reference and multiplier setting.

2.2.7.2 3.3 V Power Consumption

The 3.3 V power consumption in pass 1.x is higher than specified in the hardware reference manual. The 3.3 V I/O supply (V_{DD33}) consumes a static power of 190 mA. This is static consumption, and does not vary with the I/O frequency or activity. The static power consumption is caused by excessive current draw in the upconverter of the I/O cell.

Dynamic current consumption is in addition to the static current. For the dynamic part of the current consumption, there is no change from the HRM specification.

2.2.7.2.1 Workaround

Be aware of the increased static current. This issue will be fixed in pass 2.0 silicon.

2.2.8 Packet Input (PIP/IPD)

2.2.8.1 Multi-Buffer Packet Corruption

When large packets arrive at OCTEON Plus from one of the packet interfaces, it is possible for PIP/IPD to malfunction in some conditions. "Large" in this case means a packet that requires more than one buffer to store. The conditions required are:

(1) Packets arrive on two ports at the same time, where one port is receiving a packet larger than one buffer and the other port is a GMX port receiving a packet that will be partially dropped (with WQE Word2[RE]=1 and Word2[Opcode]=1) because of exhaustion of internal resources.

(2) The packet that is partially received but dropped has an alignment pad

(3) The data arriving from the two ports must be timed exactly so that 8B from the large packet that exactly hits the buffer boundary is followed by the packet terminated with a partial drop.

When IPD malfunctions under these conditions, data corruption will occur. A full-chip reset will be required to recover.

2.2.8.1.1 Workaround:

None needed if all received packets will fit in a single packet buffer (if using 2KB packet buffers, then only jumbo frames will exceed a single buffer).

Otherwise, ensure that internal resources will not be exhausted by using RED and/or backpressure, or run with 0 alignment pad.

2.2.9 XAUI Interface

2.2.9.1 XAUI Internal Loopback

An internal loopback mode is provided on the XAUI interface for testing purposes, and is enabled by `PCSX`*`n`*`_CONTROL1_REG[LOOPBCK1]`. However, the loopback mode does not work.

2.2.9.1.1 Workaround

Don't enable XAUI internal loopback.

Cavium Confidential For Geoffrey Colyer Bell Labs, Alcatel-Lucent 05/25/2011

2.2.9.2 XAUI Transmit Lane Swapping

The PCSXn_MISC_CTL_REG[TX_SWAP] is intended to allow swapping of the transmit lanes to ease board layout. However, this bit can not be set.

2.2.9.2.1 Workaround

If pair swapping is needed for board layout, change the configuration at the XAUI receiver. If the receiver does not support receive lane swapping, the board must have the lanes connected 1:1.

2.2.9.3 Packets Received when Interface Shut Down

When the PCSXi_MISC_CTL_REG[GMXENO] field is set to 1, no packets should be received on the corresponding XAUI interface. When packets arrive on the interface when PCSXi_MISC_CTL_REG[GMXENO] is set, the packet will be discarded, however a PCTERR interrupt will be flagged in GMXi_RX0_INT_REG for each packet. If the PCTERR interrupt is enabled in GMXi_RX0_INT_EN, many interrupts will be generated and need to be handled.

2.2.9.3.1 Workaround

Disable the PCTERR interrupt by writing a 0 to GMXi_RX0_INT_EN[PCTERR] when setting PCSXi_MISC_CTL_REG[GMXENO].

2.2.10 Media Independent Interface

2.2.10.1 Input Packet Alignment

When the MIX unit is depositing packet data in memory, it may place the data 8 bytes further in than the pointer in the input ring specifies. This also implies that the unit may place data 8 bytes beyond the normal end of the buffer.

2.2.10.1.1 Workaround

Assume that it is not possible for the first 8 bytes of the packet to be all zero. Before writing the doorbell for an input ring entry, initialize the 8 bytes starting at the pointer to zero. Increasing the size of buffers by 8 (but do not increase the value in the len field of the input ring entry) will ensure that other buffers are not overwritten. Upon receipt, if the first 8 bytes are all zero, assume that the packet started at the offset starting location. If the assumption about 8 zero bytes at the start of a packet never occurring does not apply in the target environment, choose a suitable 64b value. Cavium Networks strongly recommends that 8B buffer start alignment be used in all cases.

2.2.11 Central Interrupt Unit

2.2.11.1 Watchdog Timeout Frequency

The watchdog timer in the CIU is intended to help systems recover from errors. The watchdog timer decrements every 192 core cycles, rather than every 256 cycles as specified. The watchdog will fire in CIU_WDOGn[LEN]×49152 cycles, rather than the expected CIU_WDOGn[LEN]×65536.

2.2.11.1.1 Workaround

Increase the value of CIU_WDOGn[LEN] by a factor of 4/3.

2.2.12 Boot Bus

2.2.12.1 Configurable DMACK Polarity

The MIO_BOOT_DMA_TIMn[DMACK_PI] register changes the polarity of the corresponding BOOT_DMACKn pin. While this functionality works, it may not be sufficient for many boards, since the polarity will revert to active-high when the chip is reset. This may cause problems on the board.

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 05/25/2011

For example, suppose OCTEON Plus is reset while performing a DMA read from a boot bus device. The BOOT_DMACK*n* signals will be driven low, which is intended to revoke DMA access. However, the signal is actually active low, so access is not revoked and the other device may continue driving the bus while OCTEON Plus is attempting to access chip select 0 flash for booting.

2.2.12.1.1 Workaround

If DMACK inversion is required, do it external to OCTEON Plus on the board.

2.2.13 Universal Serial Bus (USB)

2.2.13.1 USBN_CLK_CTL[DIVIDE] Range

The `USBN_CLK_CTL` register is used to generate the clock for the USB core by dividing the core reference clock by the value configured by the `USBN_CLK_CTL[DIVIDE]` and `USBN_CLK_CTL[DIVIDE2]` registers.

The minimum value that may be configured in `USBN_CLK_CTL[DIVIDE]` is 4. Values smaller than that may make the USB core malfunction.

2.2.13.1.1 Workaround

User values greater than or equal to 4 in `USBN_CLK_CTL[DIVIDE]`.

2.2.14 Universal Asynchronous Receiver / Transmitter (UART)

2.2.14.1 UART Line Status Interrupt

The OCTEON UART can generate multiple interrupts, and the interrupts have specific priorities as described in the Hardware Reference Manual under `MIO_UARTn_IIR`. When a Received data available interrupt occurs at the exact same time as a Receiver line status interrupt, the `MIO_UARTn_IIR` register will only show the Received data available, which is not the highest priority.

2.2.14.1.1 Workaround

Read the `MIO_UARTn_LSR` register at each received data interrupt to determine whether a line status interrupt happened at the same time.

Cavium Confidential For Geoffrey Collyer Bell Labs, Alcatel-Lucent 05/25/2011