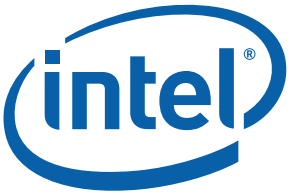

Intel® 82580 Quad/Dual GbE LAN Controller Specification Update

LAN Access Division (LAD)

322444-007EN
Revision 2.34
October 2010

Legal

Revisions

Date	Revision	Description
12/2009	1.0	Initial public release
01/14/2010	2.0	Updated errata list for PRQ.
02/03/2010	2.1	Added spec. clarification: • 1. EEPROM Required For Proper Operation of the 82580 Controller Updated errata: • 3. Possible System Hang During Enable/Disable When Connecting through nVIdia* IO55 Slot
03/01/2010	2.2	Updated sections: 1.3 Identifying Marks
06/25/2010	2.3	Updated: • 1.3 Identifying Marks. Added marking information for Engineering samples. Updated or added errata: • 6. LAN_DIS_N and AUX_PWR Pins Are Driven by 82580 When Not Used for Strapping. Added. • 7. MDIO: Com_MDIO and Destination Bits of MDICNFG Register Are Not Loaded Consistently from EEPROM. Added. • 8. Dummy Function Present When All Ports are Disabled.
7/16/2010	2.31	Added: • 9. PCIe: Link Control 2 Register Contains Incorrect Read Values. Updated: • 1. TAG Not Reset by Power-on-reset Function. Status changed to Closed. • 5. PCIe Links at x1 Instead of x4 When Attached to Certain Chipset Ports.
8/20/2010	2.32	Specification Clarification added: • 3. Use of Wake on LAN Together with Manageability Specification Change added: • 1. Update to PBA Number EEPROM Word Format.
9/10/2010	2.33	Specification Change updated: • 1. Update to PBA Number EEPROM Word Format. Text updated to address confusion about new number format.
10/19/2010	2.34	Specification Clarification added: • 1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB.

NOTE: ***This page intentionally left blank.***

1.1 Introduction and Scope

This document is an update to the *Intel® 82580 Quad/Dual Port GbE LAN Controller Datasheet*. It is intended for use by system manufacturers and software developers. All product documents are subject to frequent revision. Be sure you have the latest information before finalizing your design.

References to PCIe* in this document refer to PCIe v2.0 (5Gbps).

1.2 Product and Device Identification

Table 1. 82580 Product and Device Indentification

MM#	Stepping	Top Marking	Specification #	Status & Media	Description
905781	A1	NH82580EB	S LH5P	Production, T&R	1 Gbs, 4-port
905782	A1	NH82580EB	S LH5Q	Production, Tray	
905783	A1	NH82580EK	S LH5R	Production, T&R	1 Gbs 4-port SERDES
905784	A1	NH82580EK	S LH5S	Production, Tray	
905785	A1	NH82580DB	S LH5T	Production, T&R	1 Gbs, 2-port
905786	A1	NH82580DB	S LH5U	Production, Tray	
904983	A1	NH82580EB	Q MQR	Engineering, Tray	1 Gbs, 4 port
904984	A1	NH82580EK	Q MQS	Engineering, Tray	1 Gbs, 4-port SERDES
904985	A1	NH82580DB	Q MQT	Engineering, Tray	1 Gbs, 2-port
903315	A0	NH82580EB	Q LSY	Engineering, Tray	1 Gbs, 4-port

1.3 Identifying Marks

Engineering Samples **Production**

Figure 1. Sample Component Diagram Illustrating Identifying Marks

Refer to Figure 1:

Line 1	"intel"
Line 2	Marketing Name
Line 3	Fab Lot Number "XXXXXXXXX" (Wafer Lot no. concatenated with Assembler vendor code)
Line 4	Assembly Date Code YYWW; Engineering samples have additional Intel data; see Table 1.
Line 5	Copyright line; includes two number date code and the Pb-free mark (e1)
Line 6	Country of Origin

1.4 Nomenclature Used In This Document

This document uses specific terms, codes, and abbreviations to describe changes, errata, sightings and/or clarifications that apply to silicon/steppings. See Table 2 for a description.

Table 2. Nomenclature

Name	Description
Specification Changes	Modifications to the current published specifications. These changes will be incorporated in the next release of the specifications.
Errata	Design defects or errors. Errata may cause device behavior to deviate from published specifications. Hardware and software designed to be used with any given stepping must assume that all errata documented for that stepping are present on all devices.
Sightings	Observed issues that are believed to be errata, but have not been completely confirmed or root caused. The intention of documenting sightings is to proactively inform users of behaviors or issues that have been observed. Sightings may evolve to errata or may be removed as non-issues after investigation completes.

Table 2. Nomenclature

Specification Clarifications	Greater detail or further highlights concerning a specification's impact to a complex design situation. These clarifications will be incorporated in the next release of the specifications.
Documentation Corrections	Errors, or omissions in current published specifications. These changes are incorporated in the next release of the applicable document and then dropped from the specupdate. You may also check for changes in the revision history of specific documents.
A1, B1, etc.	Stepping to which the status applies.
Doc	Document change or update that will be implemented.
Fix	This erratum is intended to be fixed in a future stepping of the component.
Fixed	This erratum has been fixed.
NoFix	There are no plans to fix this erratum.
Eval	Plans to fix this erratum are under evaluation.
Red Change Bar/ or Bold	This Item is either new or modified from the previous version of the document.

1.5 Sightings, Clarifications, Changes, Errata, Software Clarifications

See Section 1.4 above for an explanation of terms, codes, and abbreviations.

Table 3. Summary of Sightings, Clarifications, Changes, Errata, Software Clarifications

Sightings	**Status**
1. None.	N/A
Specification Clarifications	**Status**
1. EEPROM Required For Proper Operation of the 82580 Controller	N/A
2. AC JTAG Junction Temperature Limit	N/A
3. Use of Wake on LAN Together with Manageability	N/A
4. SMBus: Illegal STOP Condition	N/A
Specification Changes	**Status**
1. Update to PBA Number EEPROM Word Format	N/A
Errata	
1. TAG Not Reset by Power-on-reset Function	Closed
2. MNG Reset Clears Resource Grant With No Feedback	A1 NoFix
3. Possible System Hang During Enable/Disable When Connecting through nVIdia* IO55 Slot	A1 NoFix
4. SMBus: EEPROM Not Written On ARP If the EEPROM Semaphore Is Taken by SW	A1 NoFix
5. PCIe Links at x1 Instead of x4 When Attached to Certain Chipset Ports	A1 NoFix
6. LAN_DIS_N and AUX_PWR Pins Are Driven by 82580 When Not Used for Strapping	A1 NoFix
7. MDIO: Com_MDIO and Destination Bits of MDICNFG Register Are Not Loaded Consistently from EEPROM	A1 NoFix
8. Dummy Function Present When All Ports are Disabled	A1 NoFix
9. PCIe: Link Control 2 Register Contains Incorrect Read Values	A1 NoFix
Software Clarifications	
1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB	N/A

1.5.1 Sightings

1. None.

1.5.2 Specification Clarifications

1. EEPROM Required For Proper Operation of the 82580 Controller

Clarification: Problems identified in the PCIe logic were resolved by implementing workarounds in EEPROM. Without these workarounds, the PCIe bus may not properly configure and the possibility exists that the controller will not negotiate a PCIe connection. This leaves the device inaccessible.

Workaround: Use an EEPROM with the workarounds (v3.22 or greater) to ensure proper device configuration.

2. AC JTAG Junction Temperature Limit

Specification: IEEE 1149.6: A Boundary-Scan Standard for Advanced Digital Networks

Clarification: AC JTAG (IEEE 1149.6) is supported only up to a junction temperature of 70'C.

Workaround: None. This is a design parameter for the hardware.

3. Use of Wake on LAN Together with Manageability

Clarification: The Wakeup Filter Control Register (WUFC) contains the NoTCO bit, which affects the behavior of the wakeup functionality when manageability is in use. Note that if manageability is not enabled, the value of NoTCO has no effect.

When NoTCO contains the hardware default value of 0b, any received packet that matches the wakeup filters will wake the system. This could cause unintended wakeups in certain situations. For example, if Directed Exact Wakeup is used and the manageability shares the host's MAC address, IPMI packets that are intended for the BMC wakes the system, which might not be the intended behavior.

When NoTCO is set to 1b, any packet that passes the manageability filter, even if it also is copied to the host, is excluded from the wakeup logic. This solves the previous problem since IPMI packets do not wake the system. However, with NoTCO=1b, broadcast packets, including broadcast magic packets, do not wake the system since they pass the manageability filters and are therefore excluded.

Effects of NoTCO Settings WoL	NoTCO	Shared MAC Address	Unicast Packet	Broadcast Packet
Magic Packet	0b	-	OK	OK
Magic Packet	1b	Y	No wake	No wake
Magic Packet	1b	N	OK	No wake
Directed Exact	0b	Y	Wake even if MNG packet. No way to talk to BMC without waking host.	N/A

Effects of NoTCO Settings WoL	NoTCO	Shared MAC Address	Unicast Packet	Broadcast Packet
Directed Exact	0b	N	OK	N/A
Directed Exact	1b	-	OK	N/A

The Intel Windows* drivers set NoTCO by default.

Workaround: Not Applicable.

4. SMBus: Illegal STOP Condition

Problem: It is important to prevent illegal STOP conditions on the SMBus interface, even when resetting the MC.

Specifically, a STOP condition should never be generated by the MC during the high clock phase of an ACK cycle while reading packet data from the 82580 as part of a Receive TCO LAN packet transaction.

Implications: If this situation occurs, the 82580 replies with a NACK to all future commands until a power cycle. As a result, the SMBus interface becomes inoperable.

Workaround: Ensure that this illegal sequence does not occur, even during MC reset.

1.5.3 Specification Changes

1. Update to PBA Number EEPROM Word Format

Change: PBA Number Module — Word 0x8-0x9

The nine-digit Printed Board Assembly (PBA) number used for Intel manufactured Network Interface Cards (NICs) is stored in EEPROM.

Through the course of hardware ECOs, the suffix field is incremented. The purpose of this information is to enable customer support (or any user) to identify the revision level of a product.

Network driver software should not rely on this field to identify the product or its capabilities.

PBA numbers have exceeded the length that can be stored as HEX values in two words. For newer NICs, the high word in the PBA Number Module is a flag (0xFAFA) indicating that the actual PBA is stored in a separate PBA block. The low word is a pointer to the starting word of the PBA block.

The following shows the format of the PBA Number Module field for new products.

PBA Number	Word 0x8	Word 0x9
G23456-003	FAFA	Pointer to PBA Block

The following provides the format of the PBA block; pointed to by word 0x9 above:

Word Offset	Description
0x0	Length in words of the PBA Block (default is 0x6)
0x1 ... 0x5	PBA Number stored in hexadecimal ASCII values.

The new PBA block contains the complete PBA number and includes the dash and the first digit of the 3-digit suffix which were not included previously. Each digit is represented by its hexadecimal-ASCII values.

The following shows an example PBA number (in the new style):

PBA Number	Word Offset 0	Word Offset 1	Word Offset 2	Word Offset 3	Word Offset 4	Word Offset 5
G23456-003	0006	4732	3334	3536	2D30	3033
	Specifies 6 words	G2	34	56	-0	03

Older NICs have PBA numbers starting with [A,B,C,D,E] and are stored directly in words 0x8-0x9. The dash in the PBA number is not stored; nor is the first digit of the 3-digit suffix (the first digit is always 0b for older products).

The following example shows a PBA number stored in the PBA Number Module field (in the old style):

PBA Number	Byte 1	Byte 2	Byte 3	Byte 4
E23456-003	E2	34	56	03

1.5.4 Errata

1. TAG Not Reset by Power-on-reset Function

Note: Not applicable to current products.

Status: Closed

2. MNG Reset Clears Resource Grant With No Feedback

Problem: When accessing the EEPROM (via EEC register) or FLASH (via FLA register), grants may be lost due to deadlock or FW reset. Software will not be notified of the lost grant. A driver in the middle of a bit bang may renew the request and receive the grant without knowing that it is actually starting a new transaction.

Implications:

1. FLA/EEC bit banging transactions may fail.

2. Long transactions may turn into different transactions than expected.

Workaround:

1. SW should not execute bit bang sequences longer than one word at a time.

2. When SW reads the EEC/FLA, it should make sure that it still has the request and grant; if not it should renew it and re-start the transaction (this does not cover all cases but reduces the possibility of a problem).

Status: A1 NoFix

3. Possible System Hang During Enable/Disable When Connecting through nVIdia* IO55 Slot

Problem: Once the link is up, when the 82580 tries to connect with the nVidia IO55 chipset, the following may occur:

- The Link layer tries to train but not all lanes work.
- The 82580 tries retraining with fewer lanes; but during retraining may loop.
- The system hangs.

Implication: The 82580 cannot reliably connect using this configuration.

Workaround: Connecting to a different port on the system.

Status: A1 NoFix

4. SMBus: EEPROM Not Written On ARP If the EEPROM Semaphore Is Taken by SW

Problem: If an SMBus ARP address is stored at the same time the EEPROM is locked by SW, the address is not be saved.

Implication : Wrong SMBus address after FW reset.

Workaround: Set the SMBus address in the EEPROM (not through ARP).

Status: A1 NoFix

5. PCIe Links at x1 Instead of x4 When Attached to Certain Chipset Ports

Problem: When the device is connected to certain PCIe x4 ports, the PCIe link is x1 instead of x4. The affected ports on Intel devices have a Device ID of 0x2690. These ports can be found on:

- 631xESB/632xESB - I/O Controller Hub
- INTEL 3100 SCH Port B

Implications: If the PCIe link is x1, the bandwidth is limited and might not be sufficient for all device ports, resulting in a performance bottleneck.

Workaround: Do not attach the device to this chipset port. If such a connection cannot be avoided, please contact your Intel representative for assistance.

Status: A1 NoFix

6. LAN_DIS_N and AUX_PWR Pins Are Driven by 82580 When Not Used for Strapping

Problem: The LANx_DIS_N and AUX_PWR pins are inputs when PE_RST_N is asserted or when an In-Band PCIe Reset occurs. At all other times, these pins become outputs and are driven low by the 82580.

Implication If these pins are driven high from another device, there could be high current draw.

If the external pull-up resistor used on these pins is not strong enough, the wrong value could be sampled at the beginning of the PCIe reset.

Workaround When driving the LANx_DIS_N and/or AUX_PWR pins from another device, use a series resistor between the devices to reduce the current.

For both external pull-up resistors and series resistors, ensure that the resistor is strong enough to pull up the pin within 40 ns. Assuming relatively short traces between the 82580 and the resistor, a 3.3 K ohm resistor should work well. Larger resistance values should not be used.

Status: A1 NoFix

7. MDIO: Com_MDIO and Destination Bits of MDICNFG Register Are Not Loaded Consistently from EEPROM

Problem: The Com_MDIO (bit 30) and Destination (bit 31) bits of the MDICNFG register (0x0E04) are not loaded consistently from the EEPROM. In some cases, the hardware default value of 0b is used instead.

Implication Software that relies on the EEPROM-loaded value might not be able to initialize an external PHY. When not using an external PHY, there is no issue.

Workaround Software should assume that the initial values of these bits are undefined and should program them before attempting to initialize an external PHY. The EEPROM bits can be used to determine the intended settings.

Status: A1 NoFix

8. Dummy Function Present When All Ports are Disabled

Problem: When all ports are disabled, either by strapping pins or by EEPROM settings, no PCIe functions should be present. However, if the Dummy Function Enable EEPROM bit (word 0x1B, bit 14) is set to 1b, Function 0 becomes a dummy function even if all ports are disabled.

Implication Even with all ports disabled, Function 0 appears in the system PCIe enumeration with a Device ID of 0x10A6, indicating a dummy function. This should not have any effect on the system.

Workaround N/A

Status: A1 NoFix

9. PCIe: Link Control 2 Register Contains Incorrect Read Values

Description The Target Link Speed and Transmit Margin fields of the Link Control 2 Register in the PCIe configuration space are described as follows in the PCIe specification:

"For a Multi-Function device associated with an Upstream Port, the field in Function 0 is of type RWS, and only Function 0 controls the component's Link behavior. In all other Functions of that device, this field is of type RsvdP."

This means that when reading the Link Control 2 Register on functions other than Function 0, these fields should be 0b. Instead, the device returns the value written to Function 0, behavior which differs from the specification.

Implication: No functional implication since these fields are defined as reserved and ignored by software.

Workaround:NA

Status: A1 NoFix

1.5.5 Software Clarifications

Applies to Intel drivers.

1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB

Clarification: The 82580 supports 256 KB TCP packets; however, each buffer is limited to 64 KB since the data length field in the transmit descriptor is only 16 bits. This restriction increases driver implementation complexity if the operating system passes down a scatter/gather element greater than 64KB in length. This can be avoided by limiting the offload size to 64 KB.

Investigation has concluded that the increase in data transfer size does not provide any noticeable improvements in LAN performance. As a result, Intel network software drivers limit the data transfer size in all drivers to 64 KB.

Please note that Linux operating systems only support 64 KB data transfers.

For further details about how Intel network software drivers address this issue, refer to Technical Advisory TA-191.

§ §

NOTE: ***This page intentionally left blank.***