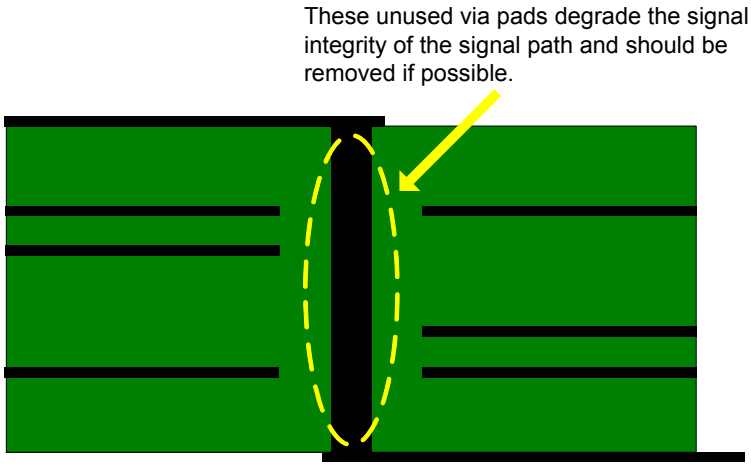


Figure 12-9 Undesirable: For Signal Vias to Have Pads on the Unused Metal Layers

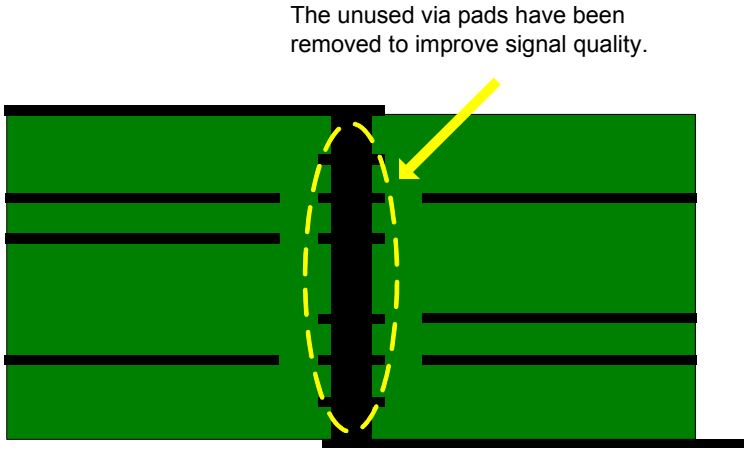


Figure 12-10 Signal Via Improved by Removing Unused Metal Layer Pads

On metal layers where signal vias need to have via pads, it is desirable to reduce capacitance between the signal vias and ground plane layers. The anti-pad diameters should be up to 20 mils larger than the via pad diameters. See Figure 12-11. Clearance between the pad and the surrounding metal should be >= 10 mils.

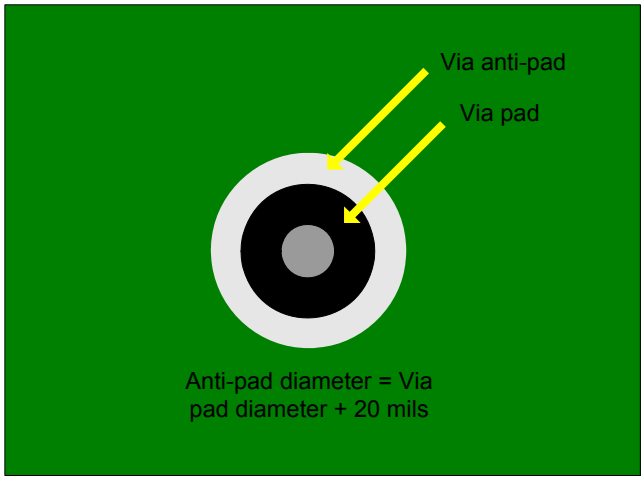


Figure 12-11 Increase Anti-Pad Diameter To Reduce Shunt Capacitance

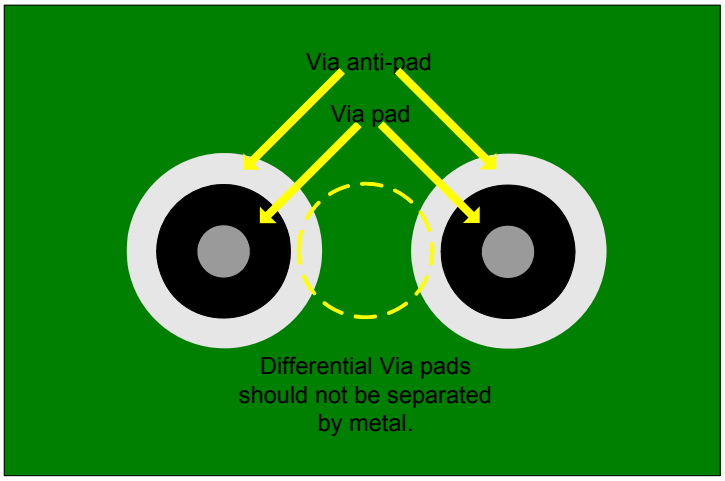


Figure 12-12 Differential Signal Via Pads Should Not Be Separated By Metal

Each time differential signal vias pass through a plane layer, within each differential pair, the anti-pads should overlap. See Figure 12-13 and Figure 12-14.

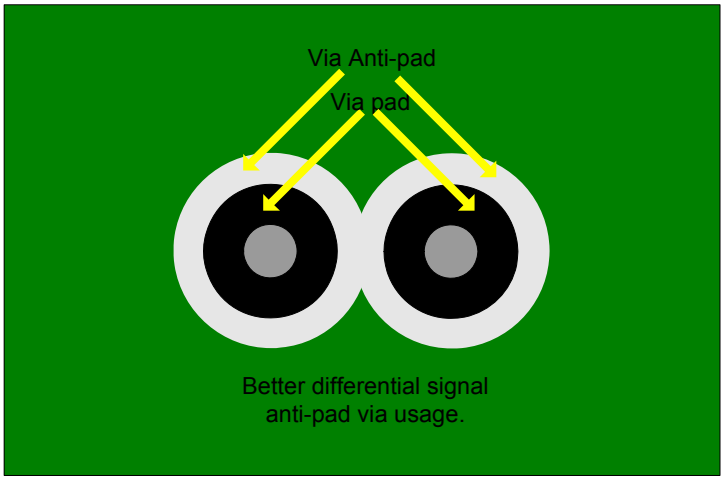


Figure 12-13 Better Differential Signals Via Anti-Pads

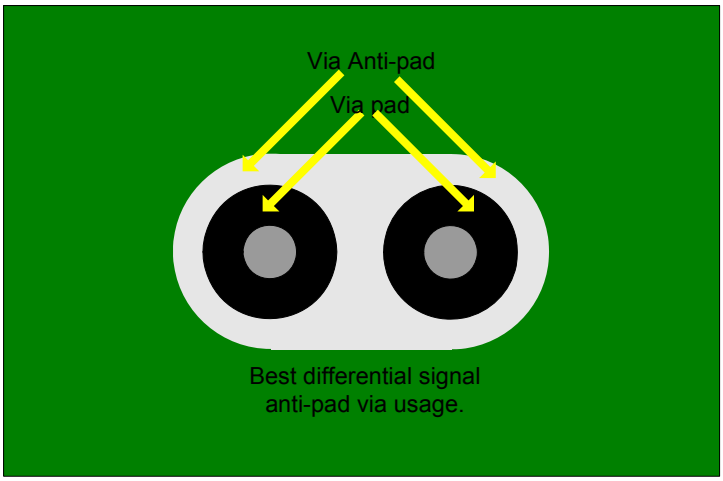


Figure 12-14 Best Differential Signals Via Anti-Pads

12.2.13 KR and SFI+ Recommended Simulations

KR and SFI+ signaling frequencies extend above 5 GHz: relatively short stubs, small discontinuities, and fairly small in-pair trace length differences can cause an undesirable increase in bit errors. Before ordering circuit boards, verify that:

- Planned KR signal trace routing on the circuit board complies with the interconnect characteristics recommended in IEEE 802.3ap sections 69.3 and 69.4.
- Planned SFI+ signal trace routing on the circuit board complies with the guidance provided in this document and complies with the interconnect characteristics recommended in the SFF-8431 specifications. Contact your Intel sales representative for more details.
- For most KR board traces:
  - If possible, export S-parameters for the planned KR signal channels, and compare them to the IEEE recommended electrical characteristics. Optimize the KR signal path until it complies with the IEEE recommendations.

    IEEE channel characteristics recommendations are for the entire length of the board channels – from the solder-pads for one KR IC device to the solder-pads for another KR device, at the far-end of the entire KR channel path. This end-to-end KR signal path typically includes two or three circuit boards, connectors, and AC coupling caps.
- For unusual routing requirements, which make it difficult to meet the IEEE channel recommendations:
  - KR board trace channels, which have been optimized by following the layout guidelines recommended within this document but which cannot be improved enough to comply with IEEE 802.3ap recommended electrical characteristics, might still work satisfactorily with the 82599 LAN silicon in KR mode (see as follows).
- With sufficient advance notice, Intel engineers can provide assistance:
  - Trace routing should be optimized prior to the next steps – request a layout review (must be willing to provide board stack-up information and the KR traces CAD artwork).
  - After KR traces have been optimized, if the IEEE recommended electrical characteristics are still not being met, then end-to-end KR board channels S-parameter models should be extracted (preferably in Touchstone* S4p format) for additional investigative simulations by Intel signal integrity engineers. Please request the required S-parameter frequency range, step size etc, before extracting Touchstone S-parameter models.
  - KR values are at 2.5 and 5 GHz and pre/post plating information must also be provided.

12.2.14 Additional Differential Trace Layout Guidelines for SFI+ Boards

As stated in the SFF-8431 specifications, SFI+ differential traces should have a nominal differential impedance that is 100 Ω with ±10 Ω tolerance and with 7% differential coupling (nominal).

Differential coupling = [(4 x Zcm) - Zdiff] / [(4 x Zcm) + Zdiff], where Zcm is the common mode impedance and Zdiff is the differential impedance. For example, when Zcm is 28.76 Ω and Zdiff is
100 Ω, the coupling is 7%.

Note: Zdiff should be 100 Ω nominal with Zdiff tolerance within +/- 10% AND when Zdiff nominal is 100 Ω, then Zcm nominal should be about 28.76 Ω, maximum.

Signal vias should be avoided on SFI+ traces. The SFI+ transmit traces must not have any vias. If there must be one or two signal vias on the SFI+ receive traces, then each via should be accompanied by one or two ground return vias, via stubs should be <= 20 mils long, AND the board's SFI+ channel insertion loss and return loss should still conform to the SFF-828431 Appendix A, SFI+ channel recommendations. To verify conformance, AC channel simulations should be performed.

Because the SFP+ module connector pads on the end of the SFI+ traces are typically wider than the SFI+ traces, in order to avoid excessive capacitance, the reference plane areas directly under each pair of SFP+ module connector pads must be voided. The voided areas under the connector pads should be rectangular-like in shape, and should be a little larger than the area occupied by each pair of SFP+ module connector pads. Simulations should be performed, to verify that 100 Ω differential impedance has been maintained. The size and shape of the reference plane voids should be adjusted to maintain the desired impedance (see Figure 12-6 for an example).

SFI+ trace lengths and trace geometry: To meet the stringent transmitter electrical requirements, some trace geometry guidance is listed in Figure 12-15 and Figure 12-16 show examples.Contact your Intel sales representative for more details.

Note: Grey is the ground plane.

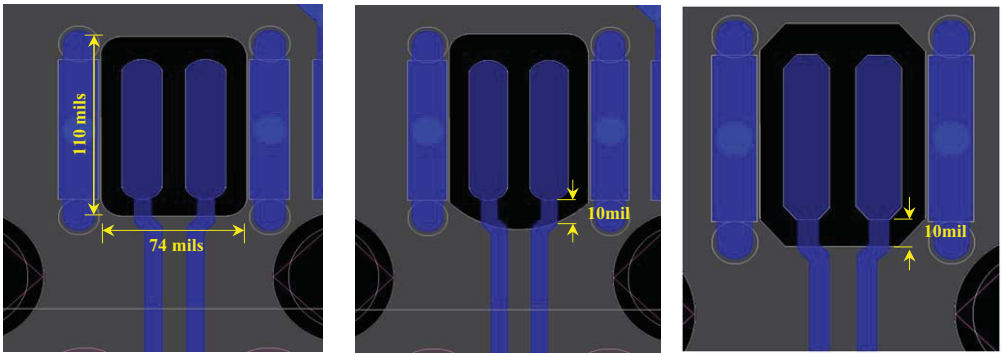


Figure 12-15 Voiding the Reference Planes Under the SFI+ Connector Pin Solder Pads

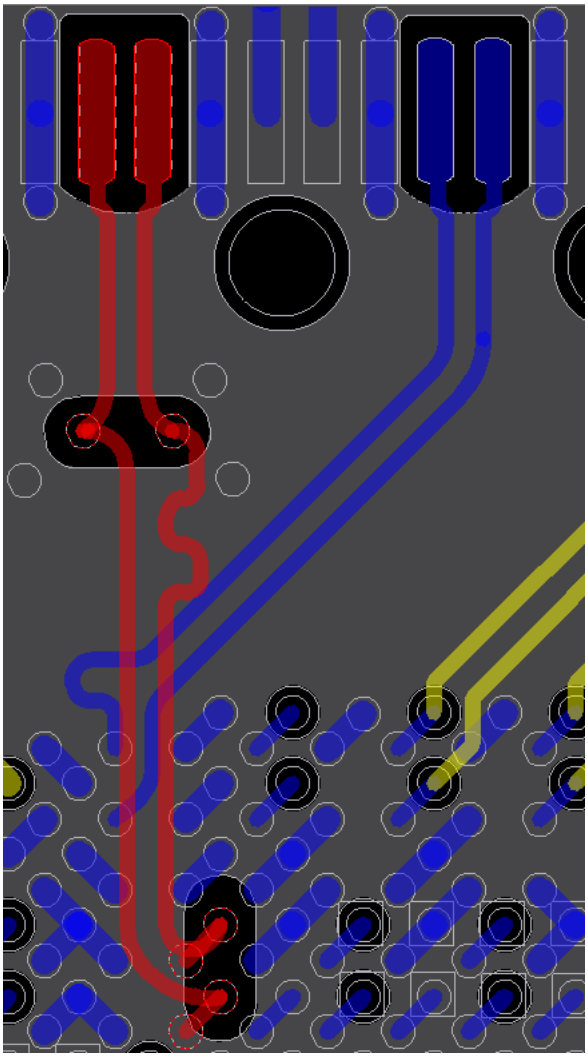
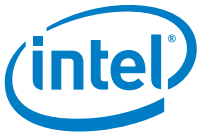


Figure 12-16 SFI+ Breakout and Routing Example From the 82599 to SFI+ Connector Pad

Note: The SFI+ transmit traces do not have any vias. Contact your Intel sales representative for more details.



12.3 Connecting the Serial EEPROM

The 82599 uses a Serial Peripheral Interface (SPI)* EEPROM. Several words of the EEPROM are accessed automatically by the 82599 after reset to provide pre-boot configuration data before it is accessed by host software. The remainder of the EEPROM space is available to software for storing the MAC address, serial numbers, and additional information. For a complete description of the content stored in the EEPROM see Section 6.0.

12.3.1 Supported EEPROM Devices

Table 12-1 lists the SPI EEPROMs that operate satisfactorily with the the 82599. SPI EEPROMs used must be rated for a clock rate of at least 2 MHz.

Table 12-1. Supported SPI EEPROM Devices

Density (Kb)	Atmel* PN	STM* PN	Catalyst* PN
128	AT25128AN-1DSI-2.7	M951286DWMN6T	CAT25C9128-TE13
256	AT25256AN-1DSI-2.7	M95256DWMN6T	

Note: Refer to Section 11.6.2.1 and Section 11.6.2.2 for minimum and recommended EEPROM sizes.

For more information on how to properly attach the EEPROM device to the the 82599, follow the example provided in the 82599 reference schematics. Contact your Intel sales representative for access.

12.4 Connecting the Flash

The 82599 provides support for an SPI Flash device that is made accessible to the system through the following:

- Flash Base Address register (PCIe Control register at offset 0x14 or 0x18).
- An address range of the IOADDR register, defined by the IO Base Address register (PCIe) Control register at offset 0x18 or 0x20).
- Expansion ROM Base Address register (PCIe Control register at offset 0x30).

12.4.1 Supported Flash Devices

The 82599 supports SPI Flash type. All supported Flashes have address size of 24 bits. Table 12-3 lists the Flash types supported.

Table 12-1 Flash Types Supported

Density	Atmel* PN	STM* PN
512 Kb	AT25F512N-10SI-2.7	M25P05-AVMN6T
1 Mb	AT25F1024N-10SI-2.7	M25P10-AVMN6T
2 Mb	AT25F2048N-10SI-2.7	M25P20-AVMN6T
4 Mb	AT25F4096N-10SI-2.7	M25P40-AVMN6T
8 Mb		M25P80-AVMN6T
16 Mb		M25P16-AVMN6T
32 Mb		M25P32-AVMN6T

For more information on how to properly attach the Flash device to the 82599, follow the example provided in the 82599 reference schematics. Contact your Intel sales representative for access.

Note: If no Flash device is used, leave FLSH_CE_N, FLSH_SCK, FLSH_SI, FLSH_SO unconnected.

12.5 SMBus and NC-SI

SMBus and NC-SI are interfaces for pass-through and configuration traffic between the Management Controller (MC) and the 82599.

Note: Intel recommends that the SMBus be connected to an MC for the EEPROM recovery solution. If the connection is to a MC, it will be able to send the EEPROM release command.

The 82599 can be connected to an external MC. It operates in one of two modes:

- SMBus mode
- NC-SI mode

The clock-out (if enabled) is provided in all power states (unless the device is disabled).

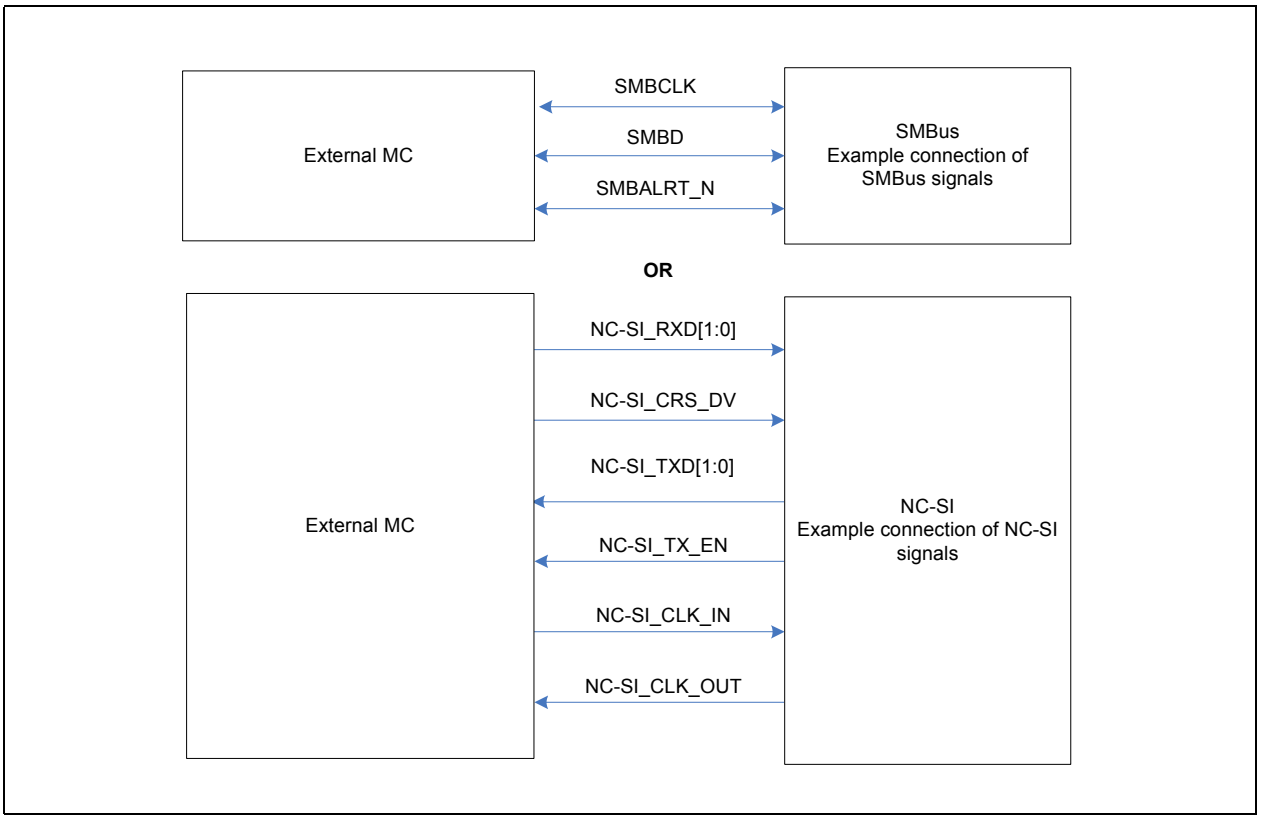


Figure 12-1 External MC Connections with NC-SI and SMBus

12.6 NC-SI

12.6.1 NC-SI Design Requirements

12.6.1.1 Network Controller

The NC-SI Interface enables network manageability implementations required by information technology personnel for remote control and alerting via the LAN. Management packets can be routed to or from a management processor.

12.6.1.2 External Management Controller (MC)

An external MC is required to meet the requirements called out in the latest NC-SI specification as it relates to this interface.

12.6.1.3 Reference Schematic

The following reference schematic (provides connectivity requirements for single and mulit-drop applications. This configuration only has a single connection to the MC. The network device also supports multi-drop NC-SI configuration architecture with software arbitration support from the MC.

Refer to the NC-SI specification for connectivity requirements for multi-drop applications.

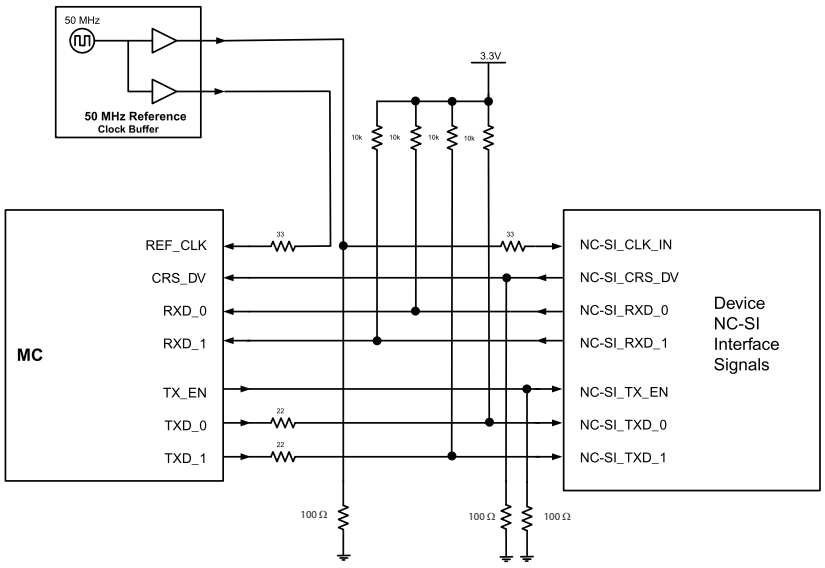


Table 12-2 NC-SI Connection Schematic: Single-Drop Configuration

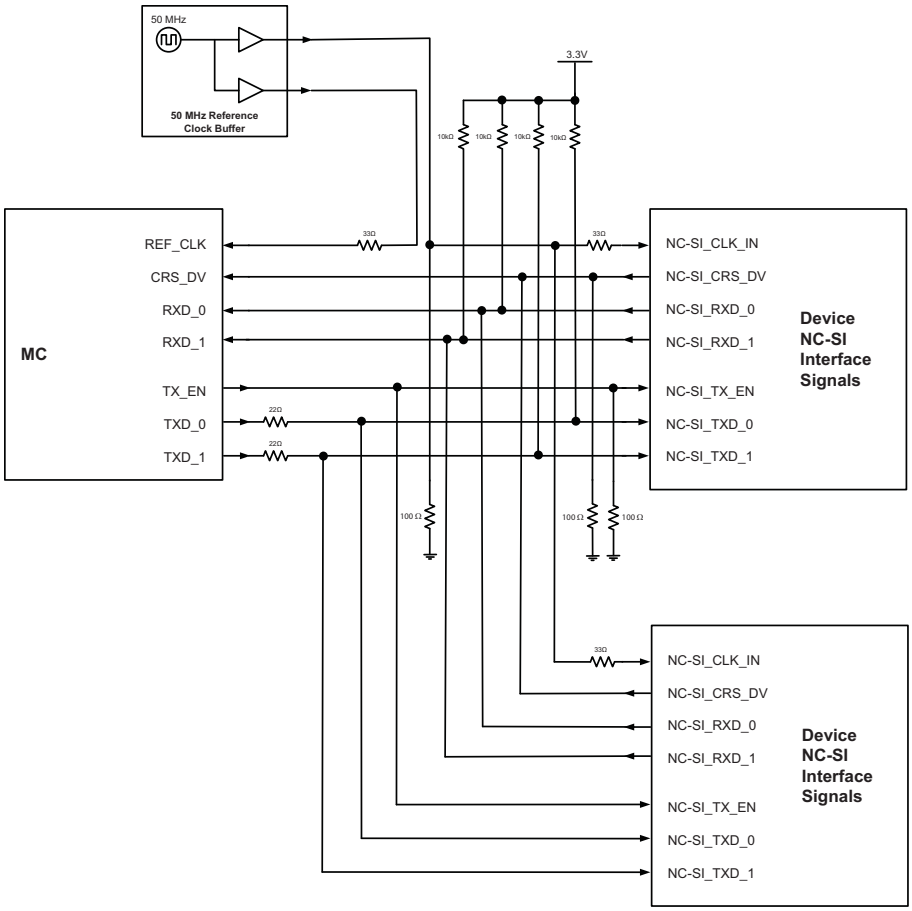


Figure 12-2 NC-SI Connection Schematic: Multi-Drop Configuration

12.6.2 NC-SI Layout Requirements

12.6.2.1 Board Impedance

The NC-SI signaling interface is a single ended signaling environment and as such Intel recommends a target board and trace impedance of 50 Ω plus 20% and minus 10%. This impedance ensures optimal signal integrity and quality.

12.6.2.2 Trace Length Restrictions

The recommended maximum trace lengths for each circuit board application is dependent on the number drops and the total capacitive loading from all the trace segments on each NC-SI signal net. The number via's must also be considered. Circuit board material variations and trace etch process variations affect the trace impedance and trace capacitance. For each fixed design, highest trace capacitance occurs when trace impedance is lowest.For the FR4 board stack-up provided in direct connect applications, the maximum length for a 50 Ω NC-SI trace would be approximately 9 inches on a minus 10% board impedance skew. This ensures that signal integrity and quality are preserved and enables the design to comply with NC-SI electrical requirements.

For special applications which require longer NC-SI traces, the total functional NC-SI trace length can be extended with non-compliant rise time by:

- providing good clock and signal alignment
- testing with the target receiver to verify it meets setup and hold requirements.

For multi-drop applications, the total capacitance and the extra resistive loading affect the rise time. A multi-drop of two devices limits the total length to 8 inches. A multi-drop of four limits the total length to 6.5 inches. Capacitive loading of extra via's have a nominal effect on the total load.

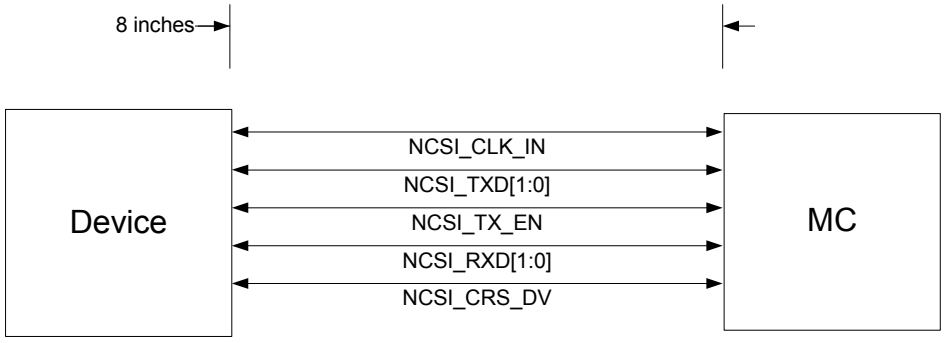


Figure 12-3 NC-SI Trace Length Requirement for Direct Connect

Table 12-3 lists how seven more vias increase the rise time by 0.5 ns. Again, longer trace lengths can be achieved.

Table 12-3 Stack Up, Seven Vias

Item	Value	Units
Trace width	4.5	mils
Trace thickness	1.9	mils
Dielectric thickness	3.0	mils
Dielectric constant	4.1	--
Loss Tangent	0.024	--
Nominal Impedance	50	W
Trace Capacitance	1.39	pf/inch

Table 12-4 lists the example trace lengths for the multi-drop topology of two and four represented in the figures that follow.

Table 12-4 Example Trace Lengths for Multi-Drop Topologies, 2 & 4

Multi-drop length parameter used in Figure 12-4 and Figure 12-5	Segment length example for multi drop configurations			
	Two drop configuration		Four drop configuration	
	Length (Inches)	Trace capacitance (Pf)	Length (Inches)	Trace capacitance (Pf)
L1	2	2.8	1.5	8.8
L2	4	5.6	2	16.1
L3	2	2.8	1	8.1
Total	8	11.1	6.5	35.6

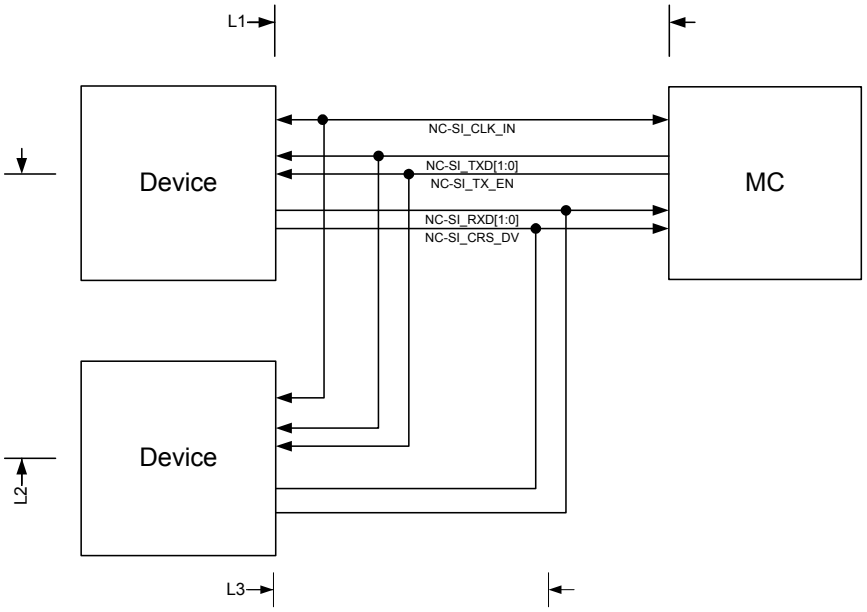


Figure 12-4 Example 2-Drop Topology

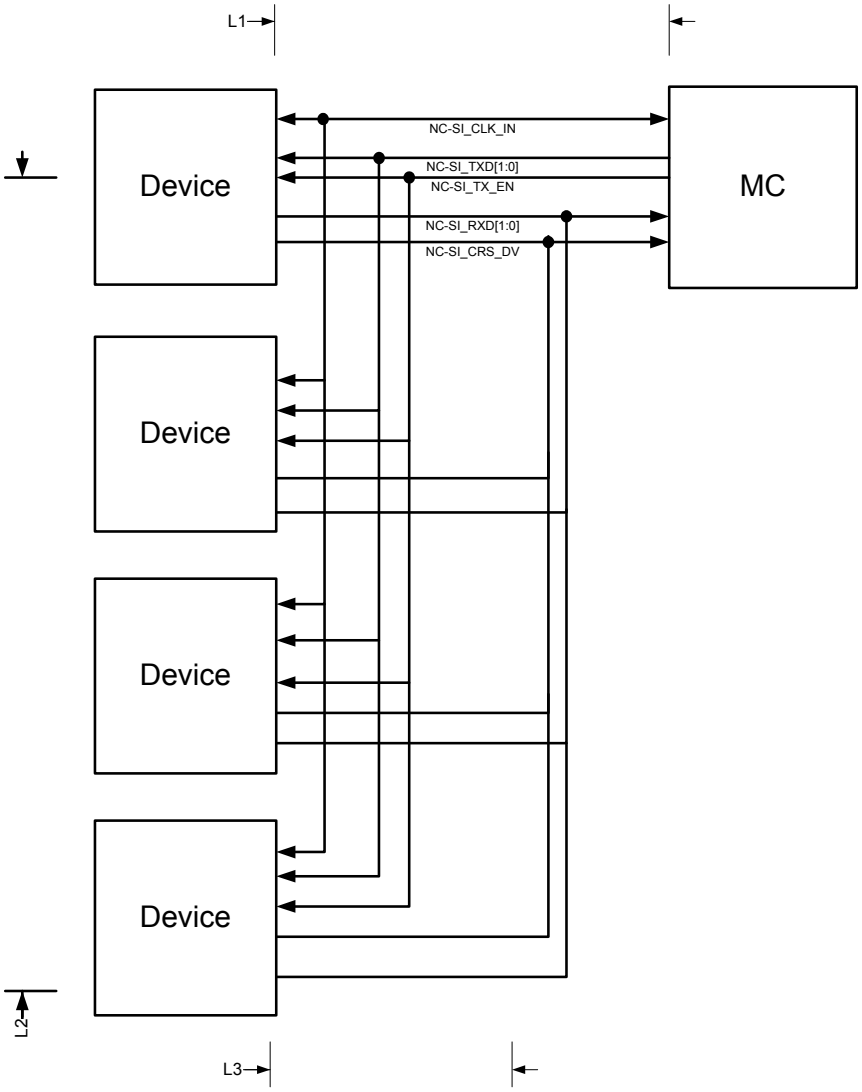


Figure 12-5 Example 4-Drop Topology

Table 12-5 Compliant NC-SI Maximum Length on a 50 Ω -10% Skew-board with Example Stack-up

Topology	Total maximum compliant linear bus size (inches)	Number of vias	Approximate Net trace capacitance minus load capacitance (pf)
4 multi-drop	6.0	1	8.3
4 multi-drop	5.5	8	8.3
2 multi-drop	8.0	1	11.1
2 multi-drop	7.5	8	11.1

Table 12-5 Compliant NC-SI Maximum Length on a 50 Ω -10% Skew-board with Example Stack-up (Continued)

Topology	Total maximum compliant linear bus size (inches)	Number of vias	Approximate Net trace capacitance minus load capacitance (pf)
Point to point	9.0	1	12.5
Point to point	8.5	8	12.5

Extending NC-SI to a maximum 11ns rise time increases the maximum trace length.

Table 12-6 Functional NC SI maximum length on a 50 Ω -10% skew board with Example Stack-up (based on actual lab-measured solution)

Topology	Total maximum functional linear bus size (inches)	Number of vias	Approximate Net trace capacitance minus load capacitance (pf)
4 multi-drop	19	1	26.4
4 multi-drop	18	8	26.4
2 multi-drop	20	1	27.8
2 multi-drop	19	8	27.8
Point to point	22	1	30.6
Point to point	21	8	30.6

12.7 Resets

After power is applied, the the 82599 must be reset. There are two ways to do this:

1. Using the internal power on reset circuit.
2. Using the external LAN_PWR_GOOD signal.

By default, the internal power on reset will reset the 82599.

If the design relies on the internal power on reset, then the power supply sequencing timing requirement between the 3.3V and 1.2V power rails has to be met. If this requirement is impossible to meet, the alternative is to bypass the internal power on reset circuit by pulling POR_BYPASS high and using an external power monitoring solution to provide a LAN_PWR_GOOD signal.

For LAN_PWR_GOOD timing requirements, see Section 4.0 and Section 5.0.

Table 12-7 Reset Context for POR_BYPASS and LAN_PWR_GOOD

POR_BYPASS	Active Reset Circuit	
If = 0b	Internal POR	
If = 1b	External Reset	

Table 12-7 Reset Context for POR_BYPASS and LAN_PWR_GOOD

	LAN_PWR_GOOD	
	If = 0b	Held in reset.
	If = 1b	Initialized, ready for normal operation.

It is important to ensure that the resets for the MC and the 82599 are generated within a specific time interval. The important requirement here is ensuring that the NC-SI link is established within two seconds of the MC receiving the power good signal from the platform. Both the 82599 and the external MC need to receive power good signals from the platform within one second of each other.

This causes an internal power on reset within the 82599 and then initialization as well as a triggering and initialization sequence for the MC. Once these power good signals are received by both the 82599 and the external MC, the NC-SI interface can be initialized. The NC-SI specification calls out a requirement of link establishment within two seconds. The MC should poll this interface and establish a link for two seconds to ensure specification compliance.

12.8 Connecting the MDIO Interfaces

The the 82599 provides one MDIO interface for each LAN port used as configuration interface for an external PHY attached to the 82599.

Connect the MDIO and MDC signals to the corresponding pins on the PHY chip. Make sure to provide a pull-up resistor to 3.3 V on the MDIO signal.

12.9 Connecting the Software-Definable Pins (SDPs)

The 82599 has eight SDPs per port that can be used for miscellaneous hardware or software-controllable purposes. These pins and their function are bound to a specific LAN device. The pins can each be individually configured to act as either input or output pins via EEPROM. The initial value in case of an output can also be configured in the same way. However, the silicon default for any of these pins is to be configured as outputs.

To avoid signal contention, all eight pins are set as input pins until after EEPROM configuration has been loaded.

Choose the right software definable pins for your applications keeping in mind that two of the eight pins (SDPx_6 and SDPx_7) are open drain. The rest are tri-state buffers. Consider that four of these pins (SDPx_0 – SDPx_3) can be used as General Purpose Interrupt (GPI) inputs. To act as GPI pins, the desired pins must be configured as inputs. A separate GPI interrupt-detection enable is then used to enable rising-edge detection of the input pin (rising-edge detection occurs by comparing values sampled at 62.5 MHz, as opposed to an edge-detection circuit). When detected, a corresponding GPI interrupt is indicated in the Interrupt Cause register.

When connecting the SDPs to different digital signals, please keep in mind that these are 3.3 V signals and use level shifting if necessary.

The use, direction, and values of SDPs are controlled and accessed using fields in the Extended SDP Control (ESDP) and Extended OD SDP Control (EODSDP) registers.

12.10 Connecting the Light Emitting Diodes (LEDs)

The 82599 provides four programmable high-current push-pull (active high) outputs per port to directly drive LEDs for link activity and speed indication. Each LAN device provides an independent set of LED outputs; these pins and their function are bound to a specific LAN device. Each of the four LED outputs can be individually configured to select the particular event, state, or activity, which is indicated on that output. In addition, each LED can be individually configured for output polarity, as well as for blinking versus non-blinking (steady-state) indication.

The LED ports are fully programmable through the EEPROM interface (LEDCTL register). In addition, the hardware-default configuration for all LED outputs can be specified via an EEPROM field, thus supporting LED displays configurable to a particular OEM preference.

Provide separate current limiting resistors for each LED connected.

Since the LEDs are likely to be placed close to the board edge and to external interconnect, take care to route the LED traces away from potential sources of EMI noise. In some cases, it might be desirable to attach filter capacitors.

12.11 Connecting Miscellaneous Signals

12.11.1 LAN Disable

The the 82599 has two signals that can be used for disabling Ethernet functions from system BIOS. LAN0_DIS_N and LAN1_DIS_N are the separated port disable signals. Each signal can be driven from a system output port. Choose outputs from devices that retain their values during reset. For example, some ICH GPIO outputs transition high during reset. It is important not to use these signals to drive LAN0_DIS_N or LAN1_DIS_N because these inputs are latched upon the rising edge of PE_RST_N or an in-band reset end.

A LAN port can also be disabled through EEPROM settings. See Section 4.2, Reset Operation for details.

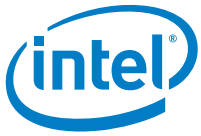


Table 12-8 PCI Functions Mapping (Legacy Mode)

PCI Function #	LAN Function Select	Function 0	Function 1
Both LAN functions are enabled	0	LAN 0	LAN 1
	1	LAN 1	LAN 0
LAN 0 is disabled	x	LAN1	Disable
LAN 1 is disabled	x	LAN 0	Disable
Both LAN functions are disabled	Both PCI functions are disabled. 82599 is in low power mode.		

Table 12-9 PCI Functions Mapping (Dummy Function Mode)

PCI Function #	LAN Function Select	Function 0	Function 1
Both LAN functions are enabled	0	LAN 0	LAN 1
	1	LAN 1	LAN 0
LAN 0 is disabled	0	Dummy	LAN1
	1	LAN 1	Disable
LAN 1 is disabled	0	LAN 0	Disable
	1	Dummy	LAN 0
Both LAN functions are disabled	Both PCI functions are disabled. 82599 is in low power mode.		

When both LAN ports are disabled following a POR / LAN_PWR_Good/ PE_RST_N/ in-band reset, the LAN_DIS_N signals should be tied statically to low. At this state the 82599 is disabled, LAN ports are powered down, all internal clocks are off and the PCIe connection is powered down (similar to L2 state).

12.11.2 BIOS Handling of Device Disable

Assume that in the following power up sequence the LANx_DIS_N signals are driven high (or is already disabled):

1. PCIe link is established following the PE_RST_N.
2. BIOS recognizes that the 82599 should be disabled.
3. BIOS drives the LANx_DIS_N signals to the low level.
4. BIOS issues PE_RST_N or an in-band PCIe reset.
5. As a result, the 82599 samples the LANx_DIS_N signals and enters the desired device-disable mode.
6. Re-enable could be done by driving high one of the LANx_DIS_N signals and then issuing a PE_RST_N to restart the 82599.

12.12 Oscillator Design Considerations

This section provides information regarding oscillators for use with the the 82599.

All designs require an external clock. There are two options for this clock source: a 25 MHz differential clock or a 25 MHz crystal. The the 82599 uses the clock source to generate clocks with frequency up to 3.125 GHz for the high speed interfaces.

The chosen oscillator or crystal vendor should be consulted early in the design cycle. Oscillator and crystal manufacturers familiar with networking equipment clock requirements can provide assistance in selecting an optimum, low-cost solution.

12.12.1 Oscillator Types

12.12.1.1 Fixed Crystal Oscillator

A packaged fixed crystal oscillator comprises an inverter, a quartz crystal, and passive components. The device renders a consistent square wave output. Oscillators used with microprocessors are supplied in many configurations and tolerances.

Crystal oscillators can be used in special situations, such as shared clocking among devices. As clock routing can be difficult to accomplish, it is preferable to provide a separate crystal for each device.

Recommended crystals are:

Table 12-10 Part Numbers for Recommended Crystals

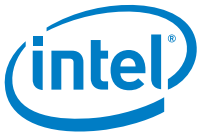
Raltron	AS-25.000-20-SMD-TR-NS7
TXC	9C25000551

12.12.1.2 Programmable Crystal Oscillators

A programmable oscillator can be configured to operate at many frequencies. The device contains a crystal frequency reference and a Phase Lock Loop (PLL) clock generator. The frequency multipliers and divisors are controlled by programmable fuses.

PLLs are prone to exhibit frequency jitter. The transmitted signal can also have considerable jitter even with the programmable oscillator working within its specified frequency tolerance. PLLs must be designed carefully to lock onto signals over a reasonable frequency range. If the transmitted signal has high jitter and the receiver's PLL loses its lock, then bit errors or link loss can occur.

PHY devices are deployed for many different communication applications. Some PHYs contain PLLs with marginal lock range and cannot tolerate the jitter inherent in data transmission clocked with a programmable oscillator. The American National Standards Institute (ANSI) X3.263-1995 standard test method for transmit jitter is not stringent enough to predict PLL-to-PLL lock failures. Therefore, use of programmable oscillators is generally not recommended.



12.12.2 Oscillator Solution

Choose a clock oscillator with a PECL or CML output. When connecting the output of the oscillator to an the 82599, use the layout information shown in Figure 11-14. Also, make sure the oscillator meets the electrical characteristics listed in Table 11-25. Note that the EuroQuartz 3HPW5761-A-25 25MHz PECL Output Crystal-Controlled Oscillator has been used successfully in 82599-based designs.

12.12.3 Oscillator Layout Recommendations

Oscillators should not be placed near I/O ports or board edges. Noise from these devices can be coupled onto the I/O ports or out of the system chassis. Oscillators should also be kept away from network interface differential pairs to prevent interference.

The reference clock should be routed differentially; use the shortest, most direct traces possible. Keep potentially noisy traces away from the clock trace. It is critical to place the termination resistors and AC coupling capacitors as close to the the 82599 as possible (less than 250 mils).

12.12.4 Reference Clock Measurement Recommendations

A low capacitance, high impedance probe (C < 1 pF, R > 500 K) should be used for testing. Probing parameters can affect the measurement of the clock amplitude and cause errors in the adjustment. A test should be done after the probe has been removed to ensure circuit operation.

12.13 Power Supplies

The 82599 requires two power rails: 3.3 V and 1.2 V. A central power supply can provide all the required voltage sources; or power can be derived from the 3.3 V supply and regulated locally using an external regulator. If the LAN wake capability is used, voltages must remain present during system power down. Local regulation of the LAN voltages from system 3.3 Vmain and 3.3 Vaux voltages is recommended.

Make sure that all the external voltage regulators generate the proper voltage, meet the output current requirements (with adequate margin), and provide the proper power sequencing. See Section 11.0.

12.13.1 Power Supply Sequencing

Due to the current demand, a Switching Voltage Regulator (SVR) is highly recommended for the 1.2 V power rail. Regardless of the type of regulator used, all regulators need to adhere to the sequencing shown in Section 11.0 of this document to avoid latch-up and forward-biased internal diodes (1.2 V must not exceed 3.3 V).

The power supplies are all expected to ramp during a short power-up interval (recommended interval 20 ms or faster). Do not leave the 82599 in a prolonged state where some, but not all, voltages are applied.

12.13.1.1 Using Regulators With Enable Pins

The use of regulators with enable pins is very helpful in controlling sequencing. Connecting the enable of the 1.2 V regulator to 3.3 V ensures that the 1.2 V rail ramps after the 3.3V rail. This provides a quick solution to power sequencing. Make sure to check design parameters for inputs with this configuration. Alternatively, power monitoring chips can be used to provide the proper sequencing by keeping the voltage regulators with lower output in shutdown until the one immediately above doesn't reach a certain output voltage level.

12.13.2 Power Supply Filtering

Provide several high-frequency bypass capacitors for each power rail (Table 12-11), selecting values in the range of 0.001 µF to 0.1 µF. If possible, orient the capacitors close to the 82599 and adjacent to power pads.

Traces between decoupling and I/O filter capacitors should be as short and wide as practical. Long and thin traces are more inductive and would reduce the intended effect of decoupling capacitors. Also for similar reasons, traces to I/O signals and signal terminations should be as short as possible. Vias to the decoupling capacitors should be sufficiently large in diameter to decrease series inductance.

Table 12-11 Minimum Number of Bypass Capacitors per Power Rail

Power Rail	Total Bulk Capacitance	1.0 µF	0.1 µF	0.001 µF
3.3 V	44 µF	0	8	0
1.2 V	132 µF	6	36	12

12.13.3 Support for Power Management and Wake Up

A designer must connect the MAIN_PWR_OK and the AUX_PWR signals on the board. These are digital inputs to the the 82599 and serve the following purpose:

MAIN_PWR_OK signals the the 82599 controller that the main power from the system is up and stable. For example, it could be pulled up to the 3.3V main rail or connected to a power well signal available in the system.

When sampled high, AUX_PWR indicates that auxiliary power is available to the 82599, and therefore it advertises D3cold wake up support. The amount of power required for the function, which includes the entire network interface card, is advertised in the Power Management Data register, which is loaded from the EEPROM.

If wake-up support is desired, AUX_PWR needs to be pulled high and the appropriate wake-up LAN address filters must also be set. The initial power management settings are specified by EEPROM bits. When a wake-up event occurs, the 82599 asserts the PE_WAKEn signal to wake the system up. PE_WAKEn remains asserted until PME status is cleared in the the 82599 Power Management Control/Status Register.

12.14 Connecting the JTAG Port

The the 82599 contains a test access port (3.3 V only) conforming to the IEEE 1149.1-2001 Edition (JTAG) specification. To use the test access port, connect these balls to pads accessible by your test equipment.

For proper operation, a pull-down resistor should be connected to the JTCK and JRST_N signals and pull-up resistors to the JTDO, JTMS and JTDI signals.

A Boundary Scan Definition Language (BSDL) file describing the the 82599 10 Gigabit Ethernet Controller device is available for use in your test environment.

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

13.0 Thermal Design Recommendations

This section provides a method for determining the operating temperature of the 82599 in a specific system based on case temperature. Case temperature is a function of the local ambient and internal temperatures of the component. This document specifies a maximum allowable Tcase for the 82599.

13.1 Thermal Considerations

In a system environment, the temperature of a component is a function of both the system and component thermal characteristics. System-level thermal constraints consist of the local ambient temperature at the component, the airflow over the component and surrounding board, and the physical constraints at, above, and surrounding the component that may limit the size of a thermal enhancement (heat sink).

The component's case/die temperature depends on:

- Component power dissipation
- Size
- Packaging materials (effective thermal conductivity)
- Type of interconnection to the substrate and motherboard
- Presence of a thermal cooling solution
- Power density of the substrate, nearby components, and motherboard

These parameters are pushed by increased performance levels (higher operating speeds, MHz) and power density (more transistors). As operating frequencies increase and packaging size decreases, the power density increases and the thermal cooling solution space and airflow become more constrained. The result is an increased emphasis on system design to ensure that thermal design requirements are met for each component in the system.

13.2 Importance of Thermal Management

The thermal management objective is to ensure that all system component temperatures are maintained within functional limits. The functional temperature limit is the range in which the electrical circuits are expected to meet specified performance. Operation outside the functional limit can degrade system performance, cause logic errors, or cause device and/or system damage. Temperatures exceeding the maximum operating limits may result in irreversible changes in the device operating characteristics. Note that sustained operation at component maximum temperature limit may affect long-term device reliability.

13.3 Packaging Terminology

The following terminology is used in this chapter:

- FCBGA Flip Chip Ball Grid Array: A surface-mount package using a combination of flip chip and BGA structure whose PCB-interconnect method consists of solder ball array on the interconnect side of the package. The die is flipped and connected to an organic build-up substrate with C4 bumps. An integrated heat spreader (IHS) may be present for larger FCBGA packages for enhanced thermal performance (but IHS is not present for the 82599).
- Junction: Refers to a P-N junction on the silicon. In this document, it is used as a temperature reference point (for example, JA refers to the “junction” to ambient thermal resistance).
- Ambient: Refers to local ambient temperature of the bulk air approaching the component. It can be measured by placing a thermocouple approximately 1”inch upstream from the component edge.
- Lands: The pads on the PCB to which BGA balls are soldered.
- PCB: Printed circuit board.
- Printed Circuit Assembly (PCA): An assembled PCB.
- Thermal Design Power (TDP): The estimated maximum possible/expected power generated in a component by a realistic application. Use Maximum power requirements listed in Table 13-2.
- LFM: Linear feet per minute (airflow).
- JA (Theta JA): Thermal resistance junction-to-ambient, °C/W.
- Ψ_{JT} (Psi JT): Junction-to-top (of package) thermal characterization parameter, °C/W. ΨJT does not represent thermal resistance, but instead is a characteristic parameter that can be used to convert between Tj and Tcase when knowing the total TDP. Ψ_{JT} is easy to characterize in simulations or measurements, and is equal to Tj minus Tcase divided by the total TDP. This parameter can vary by environment conditions like heat sink and airflow.