89HPES16T4G2 & 89HPES12T3G2
Hardware Design Guide

Notes

Introduction

This document provides system design guidelines for IDT 89HPES16T4G2 PCI Express® (PCIe®) 2.0 base specification compliant switch device. Information provided in this document is applicable to both 89HPES16T4G2 (PES16T4G2) and 89HPES12T3G2 (PES12T3G2) devices, even though the former is used as the primary reference within this document. The letters "G2" within the device names indicate that these devices are capable of GEN2 (5.0 GT/S) serial data speeds. The PES6T4G2 device offers 16 PCIe lanes divided into 4 ports of 4 lanes each. The PES12T3G2 device offers 12 PCIe lanes divided into 3 ports of 4 lanes each.

This document also describes the following device interfaces and provides relevant board design recommendations:

1) PCI Express Interface
2) Reference Clock (REFCLK) Circuitry
3) Reset (Fundamental Reset) Schemes
4) SMBus Interfaces
5) GPIO and JTAG pins
6) Power and Decoupling Scheme

PCI Express Interface

Port Configuration

Each of the four ports of the PES16T4G2 is statically allocated 4 lanes with ports labeled as 0, 2, 4 and 6. Port 0 is always the upstream port while the remaining ports are always downstream ports. All ports can operate at a maximum link width of x4 (i.e. 4 lanes), and support both 2.5 GT/S (Gen1) and 5.0 GT/S (Gen2) speeds.

Per the PCIe specification, each switch port is viewed as a virtual PCI-PCI bridge device. In the PES16T4G2, PCI device numbering follows the port numbering. Port 0 corresponds to Device 0 on the upstream bus. Port 2 corresponds to Device 2 on the PES16T4G2 virtual PCI bus, Port 4 to Device 4, and so on.

Note that unused PCIe TX and RX lanes are not required to have a termination and can be left open.

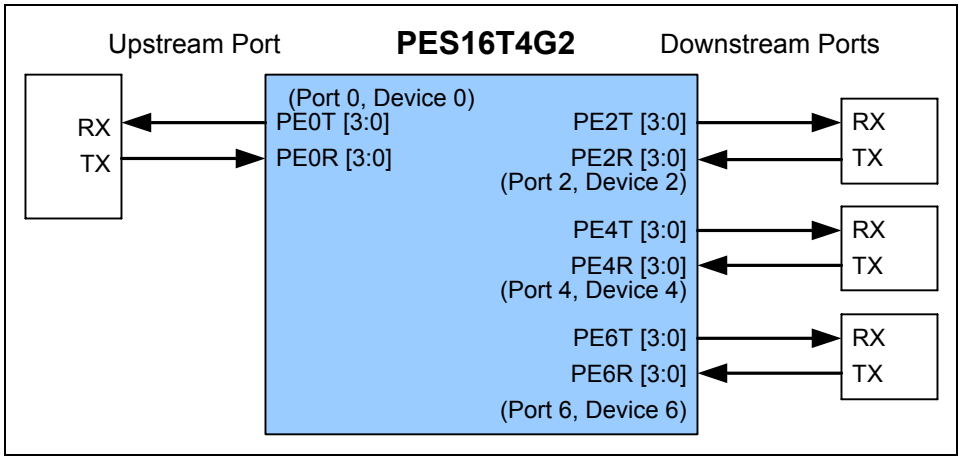


Figure 1 Port Numbering and Device Numbering

Link Width

During link training, link width is automatically negotiated. Each port is capable of independently negotiating to a x4, x2, or x1 link width. Thus, the PES16T4G2 may be used virtually in any four port switch configuration (e.g., {x4, x4, x4, x4}, {x4, x4, x1, x1}, {x4, x1, x1, x1}, etc.).

Lane Reversal

The PES16T4G2 supports automatic lane reversal outlined in the PCIe specification. This allows trace routing flexibility to avoid crossovers and potentially reduces the number of trace vias required for signal routing. Lane reversal must be done for both the transmitter and the receiver of a port.

Lane reversal mappings for the various non-trivial x4 maximum link width configurations are illustrated in Figures 3.1 and 3.2 in the IDT 89HPES16T4G2 Device User Manual.

Polarity Inversion

Each port of the PES16T4G2 supports automatic polarity inversion defined by the PCIe specification. This allows trace routing flexibility to avoid crossovers and potentially reduces the number of trace vias required for signal routing. Polarity inversion is a function of the receiver and not the transmitter. The transmitter never inverts its data. During link training, the receiver examines symbols 6 through 15 of the TS1 and TS2 ordered sets for inversion of the PExRP[n] and PExRN[n] signals. If an inversion is detected, then logic for the receiving lane automatically inverts received data. Polarity inversion is a lane function and not a link function. Therefore, it is possible for some lanes of link to be inverted and for others not to be inverted.

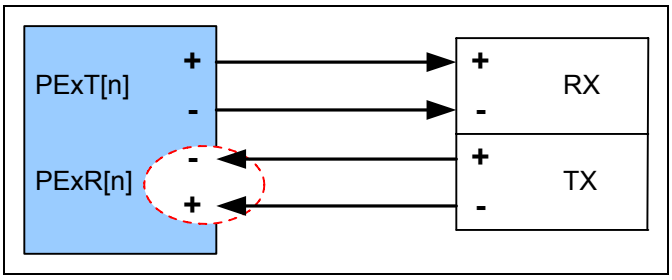


Figure 2 Polarity Inversion

AC Coupling

PCI Express signaling requires AC coupling between the transmitter and the receiver. The capacitor value must be between 75 nF and 200 nF. The 0402 package size and 100 nF value are recommended. The AC coupling capacitors should be placed near the transmitter of the device or the connector to minimize discontinuity effects.

Routing Differential Pairs

The switch includes 50 Ohm resistor on-die terminations on both the transmit and the receive pins. No external termination is required.

Individual traces within a given differential pair (positive and negative) must be matched in length to a tolerance of 5 mils. Length matching within a differential pair should occur on a segment-by-segment basis, as opposed to length matching across the total distance of the overall route. In addition, the spacing between traces of adjacent pairs must be at least 20 mils edge-to-edge to reduce crosstalk effects.

Note that trace length matching between pairs is not required because the PCIe 2.0 specifications allow up to 8ns of skew between differential pairs.

AC coupling capacitors are associated with the Tx differential pairs and should be located symmetrically on the top or bottom layer between the switch and the PCIe connectors/devices.

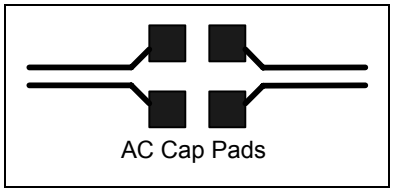


Figure 3 AC Capacitor Placement

Every effort should be made to avoid vias on the PCIe differential pairs since they can result in a signal loss of up to 0.25 dB. When a via is unavoidable, its pad size should be less than 25 mils, its hole size should be less than 14 mils and its anti-pads should be 35 mils or smaller. No extra vias should be added over and above those needed for IC pads or a connector. Vias in a differential pair should always be at the same relative location and placed in a symmetrical fashion along the differential pair as shown in Figure 4.

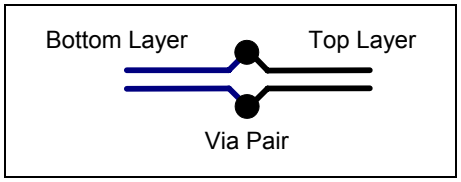


Figure 4 Via Placement

Avoid 90-degree bends or turns on traces. Wherever possible, the number of left and right bends should be matched as closely as possible. Alternating left and right turns helps to minimize length skew differences between each signal of a differential pair.

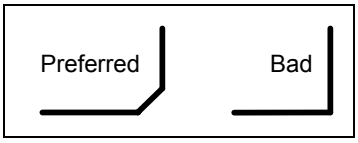


Figure 5 Acceptable Bends and Unacceptable Bends

Depending on the system topology and the maximum targeted trace length, regular FR4 material is appropriate dielectric material. In case of a backplane type of application, higher quality, lower loss material, such as Nelco 4000-13, may be used.

A HSPICE simulation kit for the switch can be requested by emailing ssdhelp@idt.com.

Serdes Reference Resistor Pins

The switch has one Serdes Reference Resistor pin per port. The 3.0K +/- 1% reference bias resistor should be located as close to these pins as possible and should be tapped immediately to the GROUND plane. This resistor must be isolated from any source of noise injection. One way to achieve this is to place the resistor on the back-side of the board, directly underneath the device. No bypass capacitors must be placed on these pins.

Reference Clock (REFCLK) Circuitry

The switch has a single differential reference clock input (PEREFCLKP[0]/PEREFCLKPN[0]) that is used to generate all of the clocks required by the internal switch logic and the SerDes.

The differential clock inputs require the signal source to drive a 0V common-mode and the REFCLK signal must meet the electrical specifications defined in the PCI Express Card Electromechanical Specification.

The reference clock inputs support spread spectrum clocking (SSC) for reducing EMI. If SSC is used, all clocks must come from a single source. This includes the clock for the switch itself, the clock for the devices connected to the downstream ports of the switch and the clocks for the root complex chipset or other devices (switch or bridge) connected to the upstream port of the switch. If SSC is not used, multiple clock sources are allowed for each PCI Express device in the tree.

Reference Clock Selection

The frequency of the reference clock inputs can be either 100MHz or 125Mhz, and the Reference Clock Mode Select (REFCLKM) input is used to indicate the choice of frequency. The PCIe CEM specification requires a nominal frequency of 100 MHz for the reference clock pair. Thus, in the majority of applications, the 100MHz clock input should be selected.

- For the 100MHz clock input, REFCLKM input pin must be asserted low.
- For the 125MHz clock input, REFCLKM input pin must be tied to a pull-up resistor to 3.3V power.

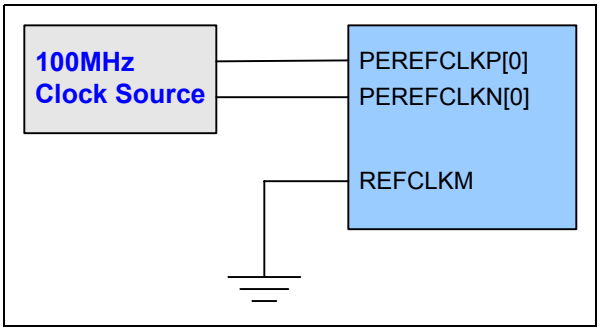


Figure 6 100MHz RefClk Selection

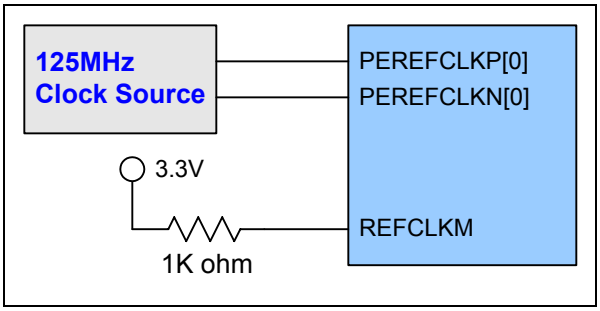


Figure 7 125MHz RefClk Selection

The differential clock for the switch and the three downstream devices can be derived from a clock buffer/generator such as ICS9DB403 or ICS9DB803. System designers can use other Gen2-compatible clock buffers/generators. The ICS9DB803 device is used on the IDT evaluation boards.

The switch provides two clock operation modes for each side of the switch - common clock and a non-common clock. System designer must configure the CCLKUS and CCLKDS pins depending on which mode is chosen for the upstream side and the downstream side of the switch. For details related to each mode, please refer to the PES16T4G2 or PES12T3G2 Device User Manual. The Spread Spectrum Clock must be disabled when the non-common clock is used on either the upstream port or downstream port.

Note that the downstream port's reference clock must be controlled by power good signal or an appropriate control signal if downstream slots support hot-plug operation.

Reset (Fundamental Reset) Schemes

The PERSTN pin is used to reset all logic inside the switch and is a Schmitt Trigger Input which can be connected to the PERST# from the system or a power-on reset circuit.

In a system, the values of Tpvperl and Tperst-clk depend on the mechanical form factor in which the switch is used. For example, the PCIe Card Electromechanical Specification, Revision 2.0, specifies minimum value of Tperst-clk=100us and Tpvperl=100ms.

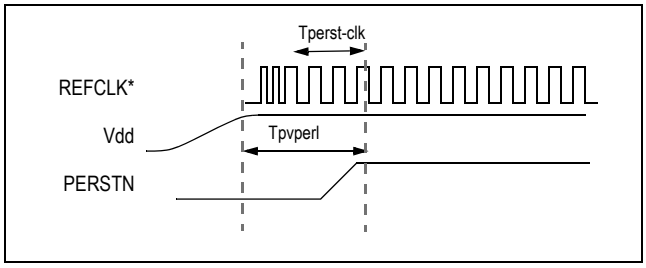


Figure 8 Fundamental Reset

For the reset signals to downstream ports, reset schemes listed below can be implemented.

- Simplified Reset Scheme
- Reset Scheme for Hot Plug Support

Simplified Reset Scheme

If Hot Plug support is not required, the simplified reset scheme can be implemented as shown in Figure 9. Add a buffer on the PERST# signal if the output from system is not able to drive a number of fan-out loads for all downstream ports.

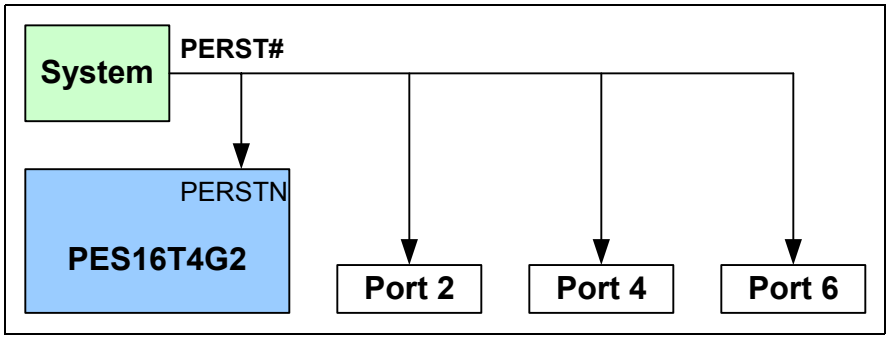


Figure 9 Simplified Reset Scheme

Reset Scheme for Hot Plug Support

Figure 10 shows an implementation where downstream endpoints have independent fundamental reset. This scheme should be used if Hot-Plug support is needed selectively on the downstream ports.

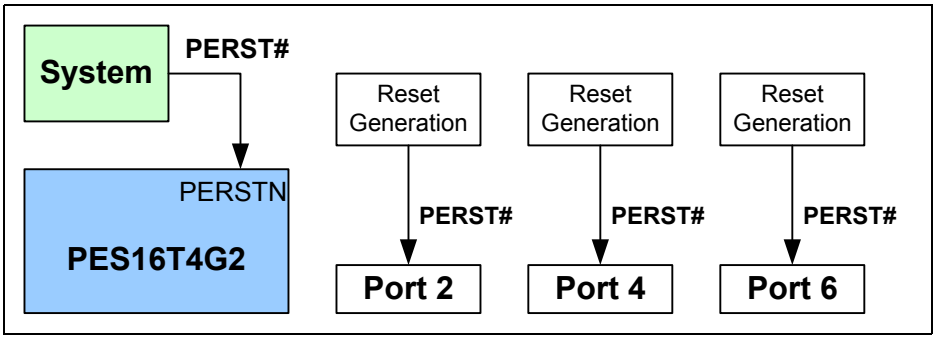


Figure 10 Reset Scheme for Hot Plug Support

Another way to generate the reset for downstream ports is by using the GPIO pins of the switch, if hot-plug support is required. The GPIO-0, GPIO-1, and GPIO-11 pins can be configured to be the reset signals for the downstream ports by programming the GPIO Function (GPIOFUNC) register via an EEPROM or slave SMBus. When a fundamental Reset occurs, all of the GPIO pins default to GPIO inputs. Therefore, the downstream port resets are tri-stated. System designers should use a pull-down on these GPIO pins to ensure that these signals are asserted low before programming GPIOFUNC register.

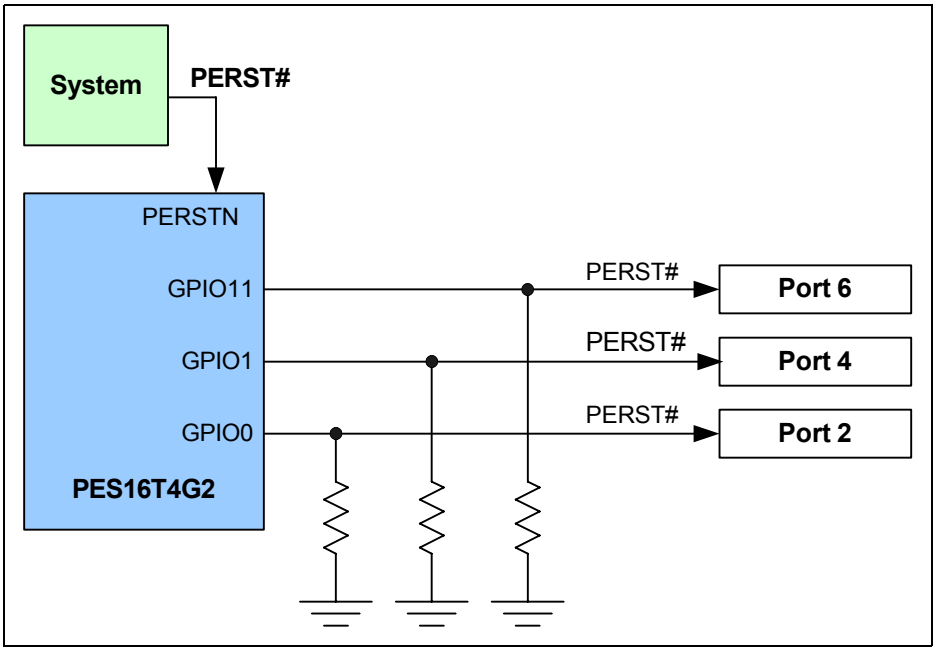


Figure 11 Reset Scheme for Hot Plug Support

RSTHALT

When this signal is asserted high during a PCI Express fundamental reset, the switch continuously returns Configuration Request Retry Completion Status (CRS) to Configuration Requests during the enumeration process. This allows system BIOS via SMBus to access internal registers before normal device operation begins. The device exits the RSTHALT state when the RSTHALT bit is cleared in the SWCTL register by a SMBus master.

This RSTHALT mode is not required in most applications. The RSTHALT pin should be pulled down externally if the application does not use a SMBus master to initialize internal registers.

SMBus Interfaces

The switch provides two SMBus interfaces.

1) The Master SMBus interface provides connection for an optional external serial EEPROM used for initialization and optional external I/O expanders.

2) The slave SMBus interface provides full access to all software visible registers in the PES16T4G2, allowing every register in the device to be read or written by an external SMBus master. The slave SMBus may also be used to preload the serial EEPROM used for initialization.

Six pins make up each of the two SMBus interfaces. These pins consist of a SMBus clock pin, a SMBus data pin, and 4 SMBus address pins. In the slave interface, these address pins allow the SMBus address to which the device responds to be configured. In the master interface, these address pins allow the SMBus address of the serial configuration EEPROM from which data is loaded to be configured. The SMBus address is set up on negation of PERSTN by sampling the corresponding address pins.

Configuring Master SMBus clock frequency

The master SMBus clock's frequency can be configured by MSMBSMODE input pin. When this pin is pulled up, the master SMBUS operates at 100KHz. When this pin is pulled down, the master SMBUS operates at 400KHz.

Note that the master SMBus clock will be in the active state during loading of the EEPROM contents or while accessing the IO Expanders. Once the operation is completed, the master SMBus clock will be in the inactive state, which is high.

Configuring a serial EEPROM address

During a fundamental reset, an optional serial EEPROM may be used to initialize any software visible register within the device. Serial EEPROM loading occurs if the Switch Mode (SWMODE [2:0]) field selects an operating mode that performs serial EEPROM initialization (e.g., Normal switch mode with Serial EEPROM initialization). The address used by the SMBus interface to access the serial EEPROM is specified by the MSMBADDR [4:1] signals, as shown in Table 1.

PES16T4G2			EEPROM
Bit	Value		Value
7	1	⟷	1
6	0	⟷	0
5	1	⟷	1
4	MSMBADDR[4]	⟷	0
3	MSMBADDR[3]	⟷	A2
2	MSMBADDR[2]	⟷	A1
1	MSMBADDR[1]	⟷	A0
0	Command R/W	⟷	Command R/W

Table 1 Serial EEPROM SMBus Address

Any serial EEPROM compatible with those listed in Table 2 can be used to store switch initialization values. EEPROM space may not be fully utilized because some of these devices are larger than the total available PCI configuration space that can be initialized in the switch.

Serial EEPROM	Size
24C32	4 KB
24C64	8 KB
24C128	16 KB
24C256	32 KB
24C512	64 KB

Table 2 PES16T4G2 Compatible Serial EEPROMs

If a serial EEPROM address is assigned to 0b1010_000, for example, all of MSMBADDR[4:1] pins should be asserted low as shown in Figure 12.

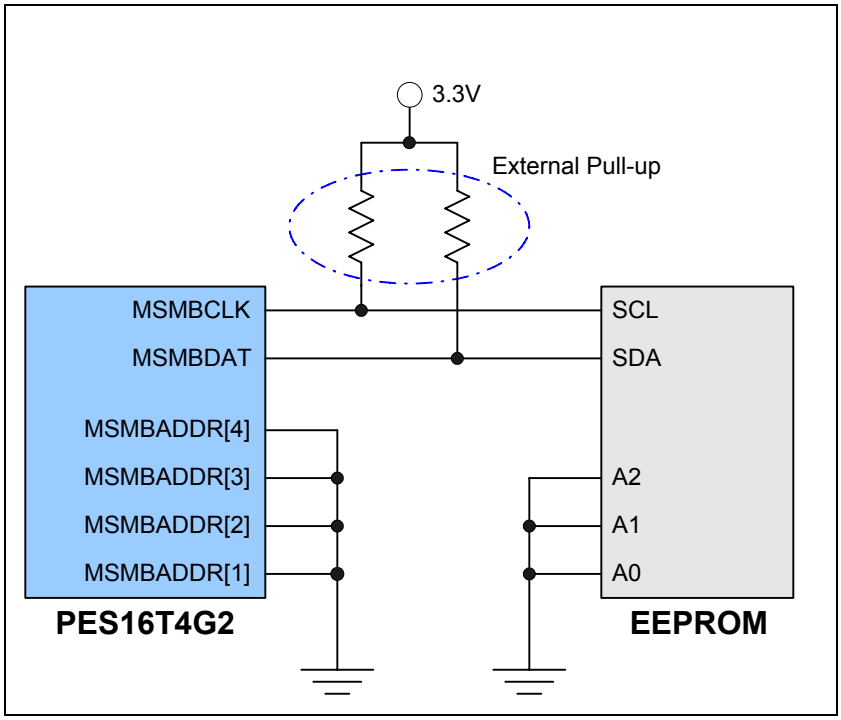


Figure 12 Example of EEPROM Address Configuration for 0b1010_000

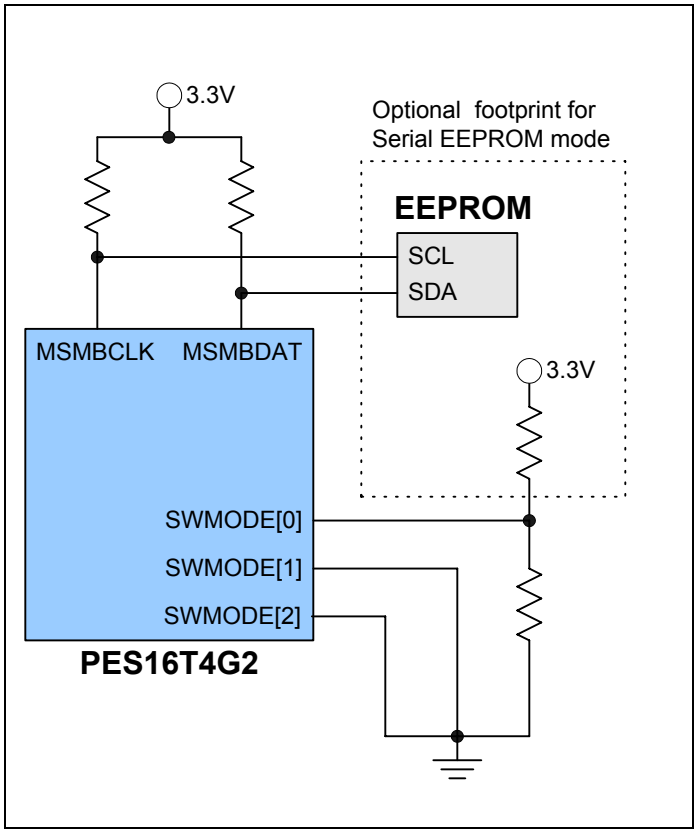


Figure 13 A Typical Implementation for SWMODE Selection and EEPROM Interface

Configuring the I/O Expander Address

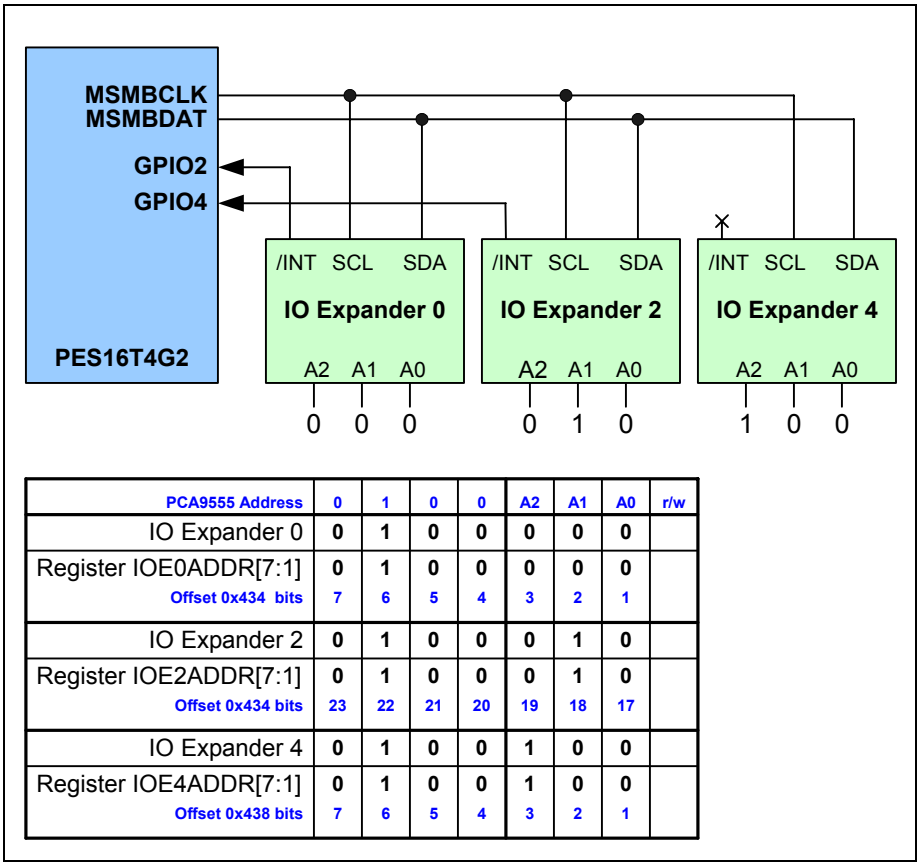
The switch utilizes external SMBus / I^2C-bus I/O expanders connected to the master SMBus interface for hot-plug and port status signals. The switch is designed to work with Phillips PCA9555 compatible I/O expanders (i.e., PCA9555, PCA9535, and PCA9539). See the Phillips PCA9555 data sheet for details on the operation of this device.

The switch supports up to three external I/O expanders numbered 0, 2, and 4.

- I/O expanders 0 and 2: hot-plug I/O signals
- I/O expander 4: link status and activity LED control

During switch initialization the SMBus/I^2C-bus address allocated to each I/O expander used in that system configuration should be written to the corresponding I/O Expander Address (IOE[0,2,4]ADDR) field. The IOE[0,2]ADDR fields are contained in the I/O Expander Address 0 (IOEXPADDR0) register while the IOE[4]ADDR fields are contained in the SMBus I/O Expander Address 1 (IOEXPADDR1) register.

Hot-plug outputs and I/O expanders may be initialized via serial EEPROM. Since the I/O expanders and serial EEPROM both utilize the master SMBus, no I/O expander transactions are initiated until serial EEPROM initialization completes.



PCA9555 Address	0	1	0	0	A2	A1	A0	r/w
IO Expander 0	0	1	0	0	0	0	0	
Register IOE0ADDR[7:1] Offset 0x434 bits	0 7	1 6	0 5	0 4	0 3	0 2	0 1	
IO Expander 2	0	1	0	0	0	1	0	
Register IOE2ADDR[7:1] Offset 0x434 bits	0 23	1 22	0 21	0 20	0 19	1 18	0 17	
IO Expander 4	0	1	0	0	1	0	0	
Register IOE4ADDR[7:1] Offset 0x438 bits	0 7	1 6	0 5	0 4	1 3	0 2	0 1	

Figure 14 Example of I/O Expander Interface

Slave SMBus address interface

The slave SMBus interface provides the switch with a configuration, management, and debug interface. Using the slave SMBus interface, an external master can read or write any software visible register in the device. The address is specified by the SSMBADDR[5,3:1] signals as shown in Table 3.

Address Bit	Address Bit Value
1	SSMBADDR[1]
2	SSMBADDR[2]
3	SSMBADDR[3]
4	0
5	SSMBADDR[5]
6	1
7	1

Table 3 Slave SMBus Address

Power and Decoupling Scheme

The switch has five different types of power supply pins:

1) VDDCORE (1.0V) powers the digital core of the switch.

2) VDDPEA (1.0V) power the SERDES core and analog circuits. VDDPEA should have no more than 25mVpeak-peak AC power supply noise superimposed on the 1.0V nominal DC value.

3) VDDPEHA (2.5V) power the SERDES core and analog circuits. VDDPEHA should have no more than 50mVpeak-peak AC power supply noise superimposed on the 2.5V nominal DC value.

4) VDDPETA (1.0V) is the termination voltage used on the SERDES TX lanes. VDDPETA can be adjusted to modify the TX common mode voltage as well as the voltage swing.

5) VDDI/O (3.3V) powers the low speed IOs of the switch.

VDDCORE, VDDPEA, and VDDPETA can be derived from the same voltage source with appropriate bypass capacitors and a ferrite bead. If all voltages can not be handled by a single voltage regulator, they can be derived from separate voltage regulators.

A ferrite bead can be used to attenuate the power noise and improve the analog circuit performance in a noisy environment. The following three parameters should be considered when you select a ferrite bead for power rails:

1) very low DC resistance

2) impedance of 50 ~ 120 ohms at 100MHz

3) provides enough DC current

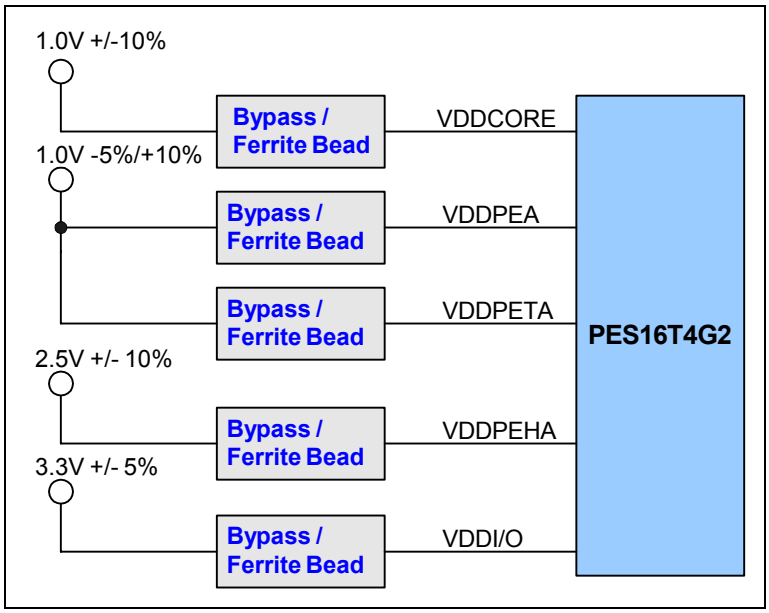


Figure 15 Board Power Supply Arrangement

Power Consumption

The typical and maximum power consumption can be found in the appropriate switch data sheet (see Reference Documents at the end of this guide).

Power-Up/Power-Down Sequence

During power supply ramp-up, VDDCORE must remain at least 1.0V below VDDI/O at all times. There are no other power-up sequence requirements for the various operating supply voltages. The power-down sequence can occur in any order.

Decoupling Scheme

1) One bypass capacitor per power pin is recommended if board layout allows. 0402 package ceramic capacitors are recommended for 0.1µF and 0.01µF capacitors.

2) Bypass Capacitors must be placed as close to the device pins as possible based on space availability. Note that some of the vias need to be shared in order to create space for placing a capacitor next to a pin.

3) A bigger capacitor should be used to filter out low frequency noise. Larger 1µF and 47µF capacitors should be added around the part. Two bigger capacitors per voltage supply are appropriate. One option is to spread out the big capacitors at four corners, top and bottom layers of the chips.

4) Short and wide traces should be used to minimize resistance and inductance.

5) Prioritize the bypass capacitors in the following order for each power supply:

1. VDDCORE
2. VDDPEA / VDDPETA
3. VDDPEHA
4. VDDI/O

GPIO and JTAG pins

GPIO pins

The switch has a number of General Purpose I/O (GPIO) pins that may be individually configured as general purpose inputs, general purpose outputs, or alternate functions. GPIO pins are controlled by the General Purpose I/O Function (GPIOFUNC), General Purpose I/O Configuration (GPIOCFG), and General Purpose I/O Data (GPIOD) registers in the upstream port's PCI configuration space. Please refer to the device datasheet for additional details.

The internal pull-up resistors value for the GPIO pins under typical condition is about 92K ohm.

JTAG pins

The switch provides the JTAG Boundary Scan interface to test the interconnections between integrated circuit pins after they have been assembled onto a circuit board. For details of the interface, please refer to to the appropriate switch user manual (see Reference Documents below).

The JTAG_TRST_N pin must be asserted low when the switch is in normal operation mode (i.e. drive this signal low with an external pull-down or control logic on the board if the JTAG interface is not used).

Reference Documents

PES12T3G2 Data Sheet

PES12T3G2 User Manual

PES16T4G2 Data Sheet

PES16T4G2 User Manual

PCI Express Base Specification, Revision 1.1, PCI-SIG

PCI Express Base Specification, Revision 2.0, PCI-SIG

PCI Express Card Electromechanical Specification Revision 1.1, PCI-SIG

PCI Express Card Electromechanical Specification Revision 2.0, PCI-SIG

PCI to PCI Bridge Architecture Specification, Revision 1.2, PCI-SIG

SMBus Specification, Revision 2.0

Intel PCI Express Electrical Interconnect Design

Revision History

September 14, 2007: Initial publication.

October 03, 2007: Updated RefClk section.