

PEX 8617

Features

■ PEX 8617 General Features

- 16-lane, 4-port PCIe Gen 2 switch
 - Integrated 5.0 GT/s SerDes
- 19 x 19mm², 324-pin PBGA package
- Typical Power: 1.93 Watts

■ PEX 8617 Key Features

- **Standards Compliant**
 - PCI Express Base Specification, r2.0 (backwards compatible w/ PCIe r1.0a/1.1)
 - PCI Power Management Spec, r1.2
 - Microsoft Vista Compliant
 - Supports Access Control Services
 - Dynamic link-width control
 - Dynamic SerDes speed control
- **High Performance**
 - Non-blocking switch fabric
 - Full line rate on all ports
 - Packet Cut-Thru with 140ns max packet latency (x4 to x4)
 - 2KB Max Payload Size
 - Read Pacing (bandwidth throttling)
 - Dual Cast
- **Flexible Configuration**
 - Ports configurable as x1, x2, x4, x8
 - Registers configurable with strapping pins, EEPROM, I²C, or host software
 - Lane and polarity reversal
 - Compatible with PCIe 1.0a PM
- **Dual-Host & Fail-Over Support**
 - Configurable Non-Transparent port
 - Moveable upstream port
 - Crosslink port capability
 - SSC Isolation
- **Quality of Service (QoS)**
 - Two Virtual Channels
 - Eight traffic classes per port
 - Weighted round-robin source port arbitration
- **Reliability, Availability, Serviceability**
 - 2 Hot-Plug Ports with native HP Signals
 - All ports Hot-Plug capable thru I²C (Hot-Plug Controller on every port)
 - ECRC and Poison bit support
 - Data Path parity
 - Memory (RAM) Error Correction
 - INTA# and FATAL_ERR# signals
 - Advanced Error Reporting
 - Port Status bits and GPIO available
 - Per port error diagnostics
 - Performance Monitoring
 - Per port payload & header counters
 - JTAG AC/DC boundary scan



PCIe Gen 2, 5.0GT/s 16-lane, 4-port Switch

The *ExpressLane*TM PEX 8617 device offers PCI Express switching capability enabling users to add scalable high bandwidth, non-blocking interconnection to a wide variety of applications including **workstations, storage systems, communications platforms, embedded systems, and intelligent I/O modules**. The PEX 8617 is well suited for **fan-out, aggregation, and peer-to-peer** applications.

High Performance & Low Packet Latency

The PEX 8617 architecture supports packet **cut-thru with a maximum latency of 140ns (x4 to x4)**. This, combined with large packet memory and non-blocking internal switch architecture, provides full line rate on all ports for performance-hungry applications such as **servers and switch fabrics**. The low latency enables applications to achieve high throughput and performance. In addition to low latency, the device supports a **max payload size of 2048 bytes**, enabling the user to achieve even higher throughput.

Data Integrity

The PEX 8617 provides **end-to-end CRC (ECRC)** protection and **Poison bit** support to enable designs that require **end-to-end data integrity**. PLX also supports data path parity and memory (RAM) error correction as packets pass through the switch.

Flexible Register & Port Configuration

The PEX 8617's 4 ports can be configured to lane widths of x1, x2, x4, or x8. Flexible buffer allocation, along with the device's **flexible packet flow control**, maximizes throughput for applications where more traffic flows in the downstream, rather than upstream, direction. Any port can be designated as the upstream port, which can be changed dynamically. The PEX 8617 also provides several ways to configure its registers.

The device can be configured through strapping pins, **I²C interface**, host software, or an optional serial EEPROM. This allows for easy debug during the development phase, performance monitoring during the operation phase, and driver or software upgrade.

Figure 1 shows some of the PEX 8617's common port configurations.

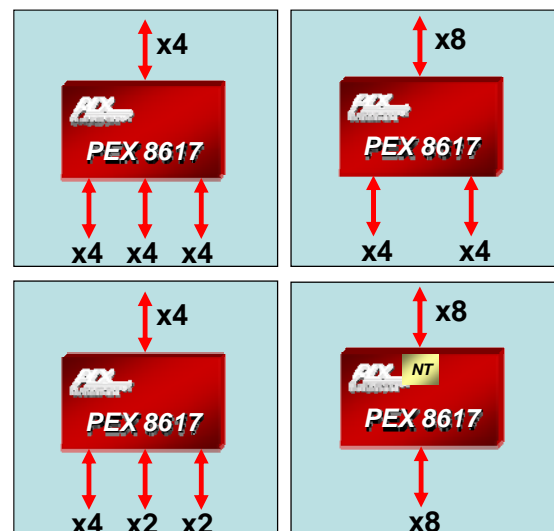


Figure 1. Common Port Configurations

Dual-Host & Failover Support

The PEX 8617 product supports a **Non-Transparent (NT) Port**, which enables the implementation of **multi-host systems** and **intelligent I/O modules** in storage, communications, and blade server applications. The NT port allows systems to isolate host memory domains by presenting the processor subsystem as an endpoint rather than another memory system. Base address registers are used to translate addresses; doorbell registers are used to send interrupts between the address domains; and scratchpad registers (accessible by both CPUs) allow inter-processor communication (see Figure 2).

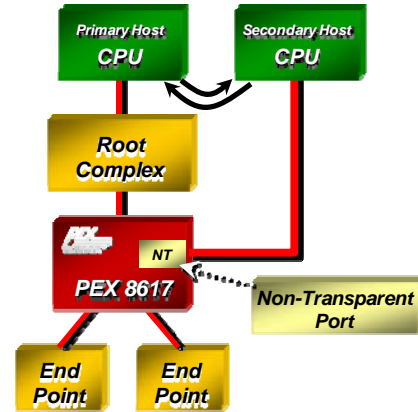


Figure 2. Non-Transparent Port

In a two-port configuration (as in Figure 1), the PEX 8617 can serve as an NT buffer, isolating two host domains via two x8 links.

Dual Cast™

The PEX 8617 supports Dual Cast, a feature which allows for the copying of data (e.g. packets) from one ingress port to two egress ports allowing for higher performance in dual-graphics, storage, security, and redundant applications.

Read Pacing™

The Read Pacing feature allows users to throttle the amount of read requests being made by downstream devices. When a downstream device requests several long reads back-to-back, the Root Complex gets tied up in serving this downstream port. If this port has a narrow link and is therefore slow in receiving these read packets from the Root Complex, then other downstream ports may become starved – thus, impacting performance. The Read Pacing feature enhances performances by allowing for the adequate servicing of all downstream devices.

Hot-Plug for High Availability

Hot-Plug capability allows users to replace hardware modules and perform maintenance without powering down the system. The PEX 8617 Hot-Plug capability feature makes it suitable for **High Availability (HA) applications**. Two downstream ports include a Standard Hot-Plug Controller. If the PEX 8617 is used in an application where one or more of its downstream ports connect to PCI Express slots, each port's Hot-Plug Controller can be used to manage the hot-plug event of

its associated slot. Every port on the PEX 8617 is equipped with a hot-plug control/status register to support hot-plug capability through external logic via the I²C interface.

SerDes Power and Signal Management

The PEX 8617 supports software control of the SerDes outputs to allow optimization of power and signal strength in a system. The PLX SerDes implementation supports four levels of power – off, low, typical, and high. The SerDes block also supports **loop-back modes** and **advanced reporting of error conditions**, which enables efficient management of the entire system.

Interoperability

The PEX 8617 is designed to be fully compliant with the PCI Express Base Specification r2.0, and is backwards compatible to PCI Express Base Specification r1.1 and r1.0a. Additionally, it supports **auto-negotiation**, **lane reversal**, and **polarity reversal**. Furthermore, the PEX 8617 is designed for Microsoft Vista compliance. All PLX switches undergo thorough interoperability testing in PLX's **Interoperability Lab** and **compliance testing at the PCI-SIG plug-fest**.

Applications & Usage Models

Suitable for **host-centric** as well as **peer-to-peer traffic patterns**, the PEX 8617 can be configured for a broad range of form factors and applications.

Host Centric Fan-out

The PEX 8617, with its symmetric or asymmetric lane configuration capability, allows user-specific tuning to a variety of host-centric applications. Figure 3 shows a typical **workstation** design where the root complex provides a PCI Express link that needs to be expanded to a larger number of smaller ports for a variety of I/O functions. In this example, the PEX 8617 has a 4-lane upstream port and three downstream ports using x4 links.

The PEX 8617 can also be used to create PCIe Gen 1 (2.5 Gbps) ports. The PEX 8617 is backwards compatible with PCIe Gen 1 devices. Therefore, the PEX 8617 enables a Gen 2 native Chip Set to fan-out to Gen 1 endpoints. In Figure 3, the PCIe slots connected to the PEX 8617's downstream ports can be populated with either PCIe Gen 1 or PCIe Gen 2 devices. Conversely, the PEX 8617 can also be used to create Gen 2 ports on a Gen 1 native Chip Set in the same fashion.

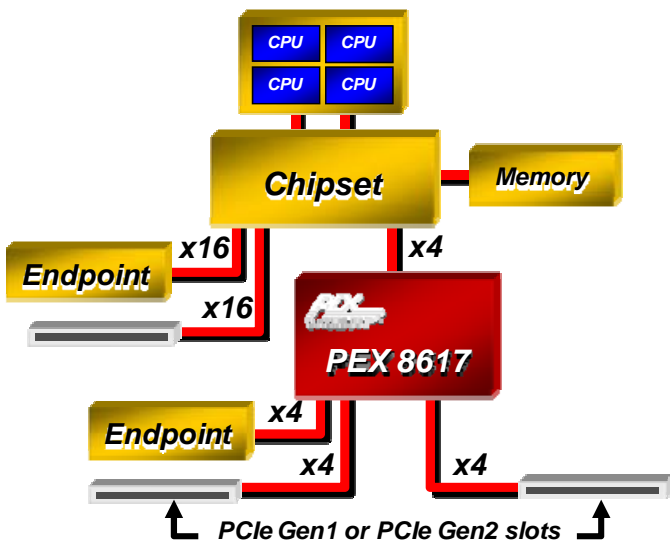


Figure 3. Fan-in/out Usage

Network Interface Cards

The PEX 8617 can also be utilized in communications applications such as Network Interface Cards (NICs). NICs, like the one shown in Figure 4, can utilize the PEX 8617 for its fan-out capabilities. In the example below, the PEX 8617 is being used on a Dual-port 10-Gigabit Ethernet (GE) NIC card. The PEX 8617 utilizes a x8 link to connect to the host and two x4 downstream links to fan-out to the 10GE ports. The peer-to-peer communication feature of the PEX 8617 allows the endpoints to communicate with each other without any intervention or management by the host.

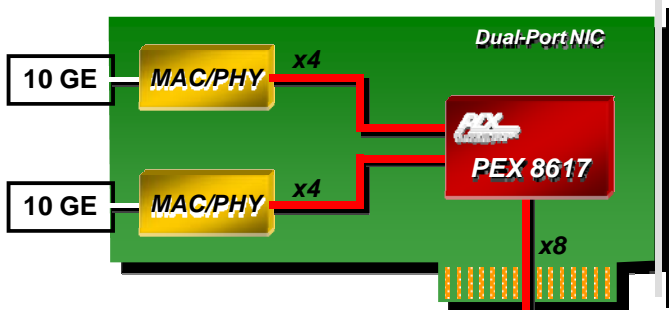


Figure 4. 10GE NIC Fan-Out

Host Bus Adapters

The PEX 8617 is well suited for Host Bus Adapter cards (HBAs) as well. Figure 5 shows a Quad-Port Fibre Channel HBAs utilizing the PEX 8617 to allow connectivity to two Fibre Channel (FC) controllers. The PEX 8617 uses a x8 link on the upstream port and two x4 downstream links to fan-out to the FC Controllers. Each FC controller supports two FC ports.

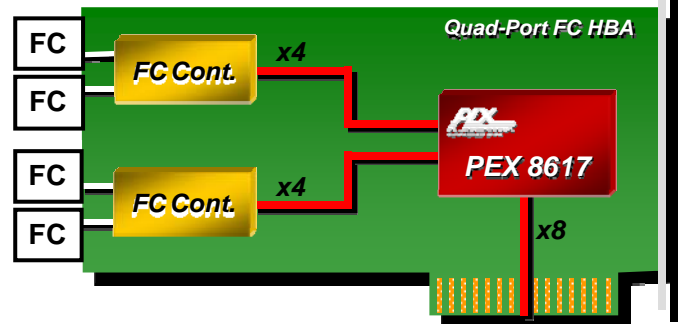


Figure 5. Quad-Port Fibre-Channel HBA

Failover Storage Systems

The PEX 8617's Dual Cast feature proves to be very useful in storage systems. In the example shown in Figure 6, the Dual Cast feature enables the PEX 8617 to copy data going to its two downstream ports to the backup system and vice versa (see yellow traffic patterns) in one transaction as opposed to having to execute two separate transactions to send data to the redundant chassis. By offloading the task of backing up data onto the secondary system, processor and system performance is enhanced. Non-Transparent (NT) ports are used to isolate the host domains of the backup system from the primary system.

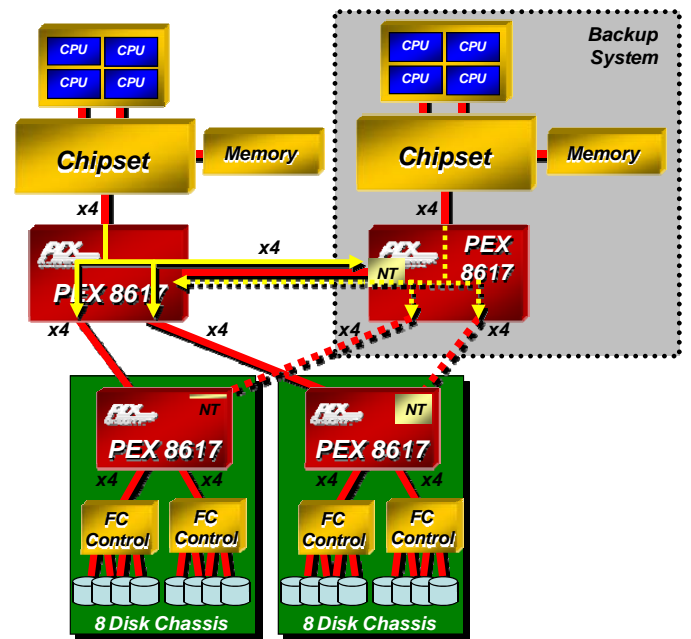


Figure 6. Dual Cast in Storage Systems

Software Usage Model

From a system model viewpoint, each PCI Express port is a virtual PCI to PCI bridge device and has its own set of PCI Express configuration registers. It is through the upstream port that the BIOS or host can configure the other ports using standard PCI enumeration. The virtual PCI to PCI bridges within the PEX 8617 are compliant to the PCI and PCI Express system models. The Configuration Space Registers (CSRs) in a virtual primary/secondary PCI to PCI bridge are accessible by type 0 configuration cycles through the virtual primary bus interface (matching bus number, device number, and function number).

Interrupt Sources/Events

The PEX 8617 switch supports the INTx interrupt message type (compatible with PCI 2.3 Interrupt signals) or Message Signaled Interrupts (MSI) when enabled. Interrupts/messages are generated by PEX 8617 for Hot-Plug events, doorbell interrupts, baseline error reporting, and advanced error reporting.

Development Tools

PLX offers hardware and software tools to enable rapid customer design activity. These tools consist of a hardware module (PEX 8617RDK), hardware documentation (available at www.plxtech.com), and a Software Development Kit (also available at www.plxtech.com).

ExpressLane PEX 8617RDK

The PEX 8617RDK is a hardware module containing the PEX 8617 which plugs right into your system. The PEX 8617RDK can be used to test and validate customer software, or used as an evaluation vehicle for PEX 8617 features and benefits. The PEX 8617RDK provides everything that a user needs to get their hardware and software development started. For more information, please refer to the PEX 8617RDK Product Brief.

Software Development Kit (SDK)

PLX's Software Development Kit is available for download at www.plxtech.com/sdk. The software development kit includes drivers, source code, and GUI interfaces to aid in configuring and debugging the PEX 8617. For more information, please refer to the PEX 8617RDK Product Brief.



PLX Technology, Inc.
870 Maude Ave.
Sunnyvale, CA 94085 USA
info@plxtech.com
www.plxtech.com

Product Ordering Information

Part Number	Description
PEX8617-BA50BC	16-Lane, 4-Port PCI Express Switch (19x19mm ²)
PEX8617-BA50BC G	16-Lane, 4-Port PCI Express Switch, Pb-Free (19x19mm ²)
PEX8617BA-AIC4U4D RDK	PEX 8617 Rapid Development Kit

Please visit the PLX Web site at <http://www.plxtech.com> for sampling.