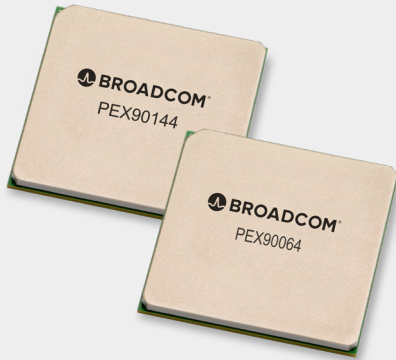


Product Brief

PEX90000 Series

PCIe Gen 6.0 Switches



Features

- PCIe 6.0 support
- Embedded ARM CPU for telemetry and management
- ExpressFabric™ PCIe switching architecture
- Host port or downstream (device) port flexibility
- Low-latency data transfers
- Standard PCIe endpoint, host, and software compatibility
- MSI-X support
- Flexible fabric topologies
- Four nontransparent bridging (NTB) ports
- Embedded MPT endpoint
- Best-in-industry SerDes technology with extended reach
- FLIT performance measurement and error injection
- Ability to connect up to four hosts to a switch
- Polarity inversion
- Lane reversal
- Policy Enforcement Agent/Viking/Ptrace
- Management Component Transport Protocol

AI Systems Using PCIe

Broadcom is introducing the PEX90000 family of PCIe Gen 6.0 switches. These switches are the internal fabric for connecting the various pieces of AI systems.

PEX90000 switches offer up to 256GB/s bidirectional raw bandwidth per x16 port. The largest PCIe Gen 6.0 switch allows the system to achieve up to 9.2Tb/s of raw bandwidth through the device.

PEX90000 PCIe switches enable designers to perform multiple tasks to create AI systems and solve the challenges of high-performing applications:

- Create the highest-performing AI systems by using the PCIe as the internal fabric
- Simplify connectivity while providing the highest PCIe switching performance available for data center servers, storage, and networks
- Reduce latency, system complexity, and power consumption in data-intensive environments
- Take advantage of industry-first features for the most demanding AI and machine learning (ML) applications

Enhanced Nontransparent Bridging 2.0 (NTB 2.0)

PEX90000 switches are equipped with the field-proven NTB technology that Broadcom has been shipping since 2004. This multihost enabling architecture has been enhanced to NTB 2.0 based on years of use and feedback by leading OEMs and ODMs. The largest PEX90000 switch (144-lane device) is equipped with four NTB 2.0-capable ports, enabling a large number of hosts and servers to be connected to the switch.

Embedded ARM CPU

Each PEX90000 series switch is equipped with an embedded ARM CPU, internal RAM, timer blocks, a watchdog timer, and vectored interrupt controllers. The embedded CPU can be used to configure desired switch functionality, I/O allocation, and I/O management.

Software-Defined PCIe Switch Fabric

The switches are designed for hybrid hardware and software platforms that offer high configurability (the number of hosts, downstream ports, and assignment of the slots or ports with those hosts). Once the configuration is complete, the data flows directly between connected devices with hardware support, enabling the fabric to offer nonblocking line speed. The solution offers an innovative approach to set up and control the PEX90000 switches, configure the routing tables, and handle errors and hot-plug events.

Key Advantages

- Standards compliant with PCIe specification 6.3
- 144-, 104-, 80-, 64-, and 32-lane configurations
- On-chip best-in-class Broadcom 64GT/s SerDes
- Low-power SerDes with PCIe sleep/power-saving modes
- Device-specific relaxed ordering
- Port reconfiguration without impacting other ports
- Independent port speeds (Gen1/2/3/4/5/6)
- Choice of link width: x2, x4, x8, or x16
- Ability to configure with serial flash, embedded CPU, and host software
- Ability to designate any port as the upstream port
- Full line rate on all ports
- Cut-through packet latency of 90 ns (x16 to x16)
- 1-KB maximum payload size
- Advanced debug and telemetry features
- Performance PAK
- DPC/eDPC support
- Read tracking for surprise removal
- All ports hot-plug capable through I2C
- SRIS/SRNS support
- ECRC and poison bit support
- Port status bits and GPIO available

Downstream Port Containment (DPC/eDPC)

Most servers have difficulty handling serious errors in I/O devices, especially when a device disappears from the system. The PEX90000 DPC/eDPC implementation allows a downstream link to be disabled after an uncorrectable error or timeout, making recovery possible in a controlled and robust manner.

Debug, Bring-Up, and Monitoring

Broadcom® PCIe switches support a rich set of features for debugging the system during initial bring-up and for monitoring the performance during runtime. These features include packet generation, eye-scope, error monitoring, port utilization, error counting, loopback, header, TLP logging, and a powerful embedded PCIe analyzer. Additional features for telemetry applications are also included.

AI and High-Speed Applications

High-performance computing (HPC) clusters are made up of high-performance processing elements that communicate through high-bandwidth, low-latency pathways to support applications such as AI, ML, medical imaging, financial trading, data analytics, and image processing. Broadcom PCIe switches are broadly used in AI and ML applications to interconnect GPUs, FPGAs, accelerators, NVMe SSDs, and NICs. The PEX90000 switch family doubles the connection bandwidth between these processing elements compared to the previous generation of PCIe 5.0 devices.

Composable systems can be built using PEX90000 switches with pools of compute, storage, and networking resources. These resources can be dynamically configured and allocated to applications or clients based on service-level agreements in a public or private cloud computing environment.

Interop Development Platform and Software Packages

The Broadcom interop development platform (IDP) allows system designers to test and evaluate PEX90000 switches in the desired configuration. This IDP can be connected to a server through a host interface board. IDPs can also be cascaded to create larger topologies. Each IDP has several standard PCIe ports.

The IDP includes drivers, software development kit, APIs, and GUI interfaces to aid in configuring and debugging in the lab and in the field. Software drivers and APIs are provided to help implement dynamic allocation of I/Os to hosts, hot add and remove, chassis management, LED management, error handling, and other key functions.

Sample Ordering Information

Part Number	Description
SS33-0B00-02	PEX90144, 144-lane PCIe 6.0 switch (Hard Secure)
SS32-0B00-02	PEX90104, 104-lane PCIe 6.0 switch (Hard Secure)
SS31-0B00-02	PEX90080, 80-lane PCIe 6.0 switch (Hard Secure)
SS36-0A00-02	PEX90064, 64-lane PCIe 6.0 switch (Hard Secure)
SS35-0A00-02	PEX90032, 32-lane PCIe 6.0 switch (Hard Secure)