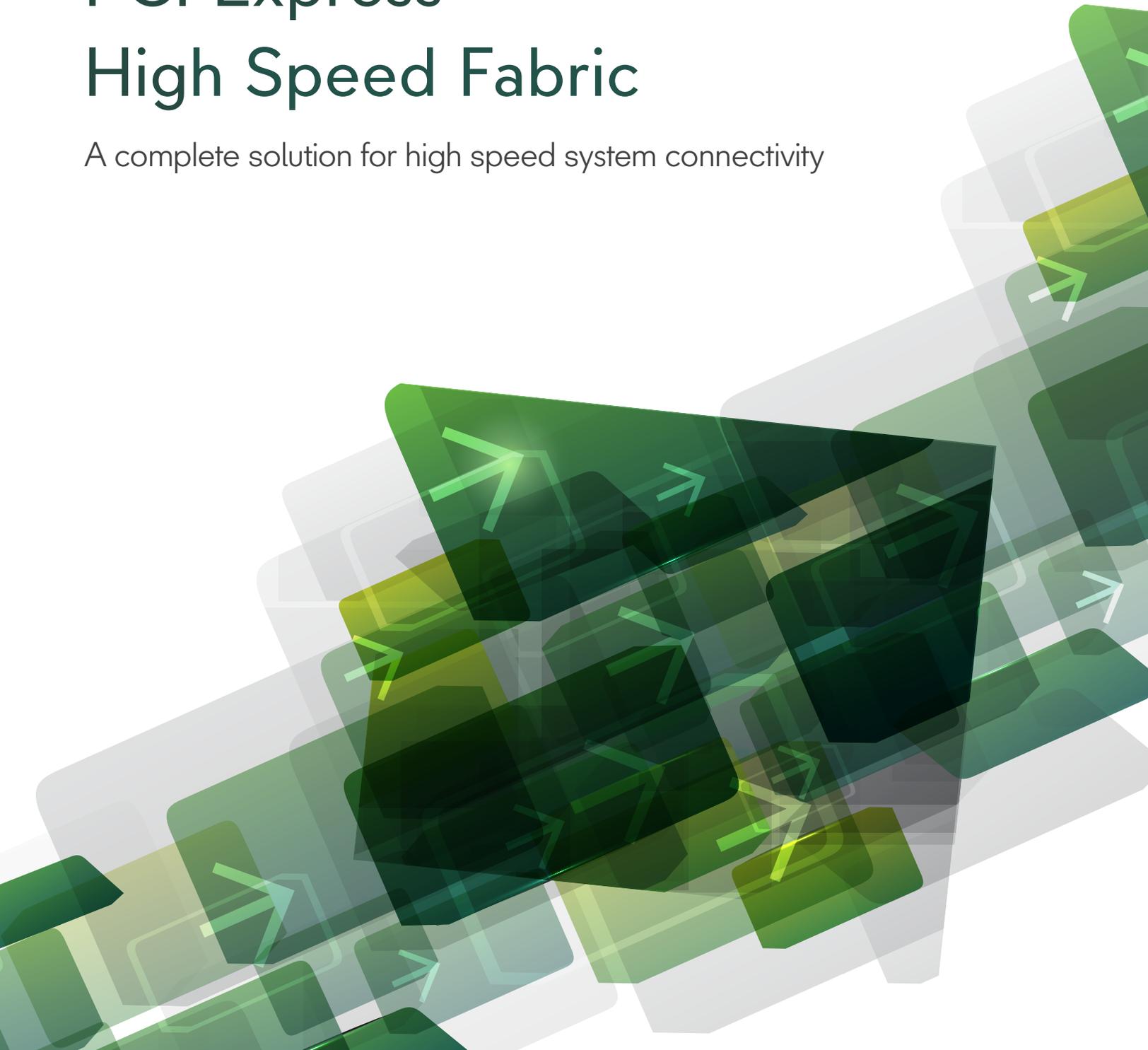




# PCI Express High Speed Fabric

A complete solution for high speed system connectivity





# Contents

- Why PCI Express? ..... 2
- PCIe Applications ..... 4
  - SmartIO Technology ..... 4
  - Reflective Memory / Multi-cast..... 5
- eXpressWare™ Software ..... 6
  - PCI Express Software Suite ..... 6
  - IPoPCIe IP over PCIe ..... 7
  - SISCI Low Level API ..... 8
  - SuperSockets™ ..... 10
- PCIe Hardware ..... 12
  - PXH830 Gen 3 PCIe NTB Adapter ..... 12
  - PXH832 Gen 3 Host/Target Adapter ..... 14
  - MXH830 Gen 3 PCIe NTB Adapter ..... 16
  - MXH832 Gen 3 Host/Target Adapter ..... 18
  - PXH810 Gen 3 PCIe NTB Adapter ..... 20
  - PXH812 Gen 3 Host/Target Adapter ..... 22
  - IXH610 Gen2 PCIe Host Adapters ..... 24
  - IXH620 Gen2 PCIe XMC Host Adapter ..... 26
  - IXS600 Gen3 PCIe Switch ..... 28



## Introduction

Maximizing application performance is a combination of processing, communication, and software. PCIe Fabrics combine all three elements. A PCIe Fabric connects Processors, I/O devices, FPGAs and GPUs into an intelligent fabric. This fabric connects devices through flexible cabling or fixed backplanes. PCIe Fabrics main goal is to eliminate system communication bottlenecks, allowing applications to reach their potential. To accomplish this, PCIe Fabrics deliver the lowest latency possible, combined with high data rates.

Dolphin's PCIe Fabric solution consists of standard computer networking hardware and eXpressWare™ PCIe software. Our standard form factor boards and switches reduce time to market, enabling customers to rapidly develop and deploy PCIe Fabric solutions for data centers and embedded systems. eXpressWare™ software enables reuse of existing applications and development of new exciting applications, both with better response times and data accessibility. eXpressWare™ SuperSocket™ and IPoPCIe software ensures quick application deployment by not requiring any application modifications. Application tuning is available with our low level SIOSI shared memory API that delivers maximum performance.

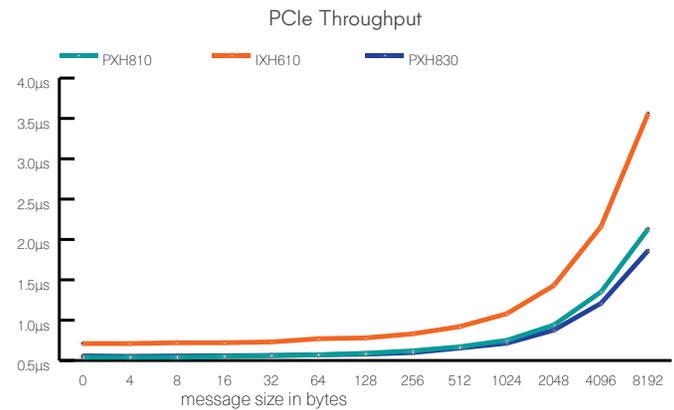
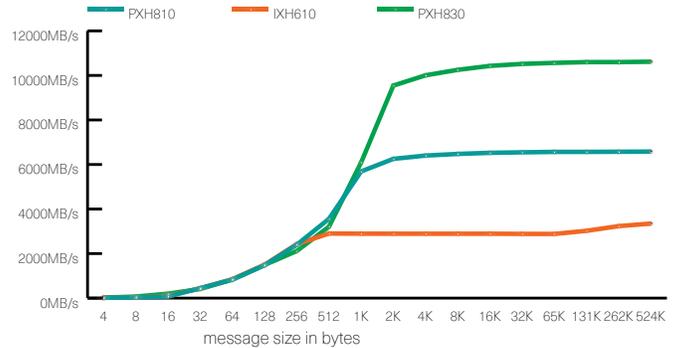
The PCI Express standard's continued performance improvements and low cost infrastructure is ideal for application and system development. The current PCI Express solutions are at 128 GT/s. The PCIe road map extends speeds to 256 GT/s and 512 GT/s, while still maintaining backward compatibility. Standard PCIe commercial components are used by Dolphin as a road map to high performance hardware. Dolphin's solution exploits the PCI Express infrastructure to deliver next generation systems with maxim application performance. Our easy to implement and deploy solution gives customers the choice of changing or not changing their existing applications, but still taking advantage of PCI Express performance.

# Why PCI Express ?

## Performance

PCI Express solutions deliver outstanding performance compared to other interconnects in latency and throughput. When compared to standard 10 Gb/s Ethernet, PCI Express latency is 1/10 the measured latency. This lower latency is achieved without special tuning or complex optimization schemes. Current solutions offer latencies starting at 540 nanoseconds memory to memory access.

In addition, Dolphin takes advantage of PCI Express high throughput. Our current Gen 3 x16 implementations achieve data rates exceeding 11 GB/s. Dolphin's eXpresWare™ software infrastructure allows customers to easily upgrade to next generation PCI Express with doubling bandwidth. No software changes are required. These products still maintain the low latency characteristic of PCI Express. An investment in low latency high performance Dolphin products yields dividends today and into the future.

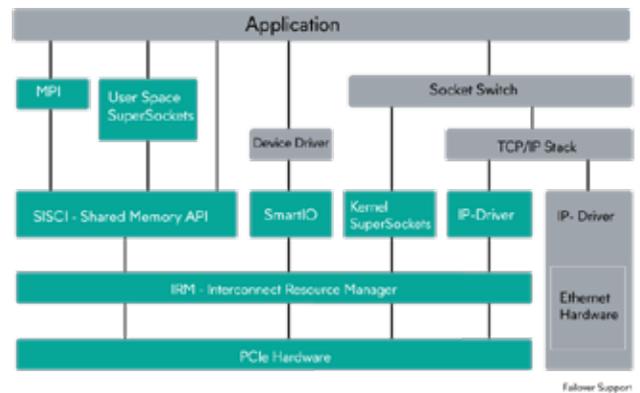


PCIe Latency

## Eliminate Software Bottlenecks

Dolphin's eXpressWare™ Software is aimed at performance critical applications. Advanced performance improving software, such as the SISCi API, removes traditional network bottlenecks. Sockets, IP, and custom shared memory applications utilize the low latency PIO and DMA operations within PCI Express to improve performance and reduce system overhead.

Software components include SuperSockets Sockets API, an optimized IPoPCiE driver for IP applications, SmartIO software for I/O optimization, and the SISCi shared memory API. Dolphin's SuperSockets software delivers latencies around 1 µs and throughput at 65 Gb/s. The SISCi API offers further application optimization by using remote memory segments and multi-cast/ reflective memory operations. Customers benefit from even lower latencies with the SISCi API in the range of 0.54 µs latency with higher throughput of over 11 GB/s. SmartIO software is used for peer to peer communication and moving devices between systems with device lending.



Dolphin eXpressWare™ Software Stack



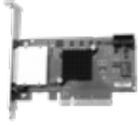
## Key Applications

- Financial Trading Applications
- High Availability Systems
- Real Time Simulators
- Databases and Clustered Databases
- Network File Systems
- High Speed Storage
- Video Information Distribution
- Virtual Reality Systems
- Range and Telemetry Systems
- Medical Equipment
- Distributed Sensor-to-Processor Systems
- High Speed Video Systems
- Distributed Shared Memory Systems

## Robust Features

- Lowest host to host latency and low jitter with 0.54μs for fast connections and data movement
- DMA capabilities to move large amounts of data between nodes with low system overhead and low latency. Application to application transfers exceeding 11 GB/s throughput.
- Management software to enable and disable connection and fail over to other connections
- Direct access to local and remote memory, hardware based uni- and multi-cast capabilities
- Set up and manage PCIe peer to peer device transfers
- High speed sockets and TCP/IP application support
- Ease installation and plug and play migration using standard network interfaces

## High Performance Hardware

	<p>Low profile PCIe Gen 2 and Gen 3 adapter cards provide high data rate transfers over standard cabling. These interface cards are used in standard servers and PCs deployed in high performance low latency applications. These cards incorporate standard iPass connectors and SFF-8644 connectors. They support both copper and fiber optic cabling, along with transparent and non-transparent bridging (NTB) operations.</p>
	<p>XMC Adapters bring PCIe data rates and advanced connection features to embedded computers supporting standard XMC slots, VPX, VME or cPCI carrier boards. PCIe adapters expand the capabilities of embedded systems by enabling very low latency, high throughput cabled expansion and clustering. Standard PCs can easily connect to embedded systems using both XMC and host adapters.</p>
	<p>PCI Express Gen 3 switch boxes scale out PCIe Fabrics. Both transparent and non-transparent devices link to a PCIe switch, increasing both I/O and processing capacity. These low latency switches scale systems while maintaining high throughput.</p>

# PCIe Applications

## SmartIO Technology

### Remote Peer-to-Peer

PCIe peer-to-peer communication (P2P) is a part of the PCI Express specification and enables regular PCIe devices to establish direct data transfers without the use of main memory as a temporary storage or the CPU for data movement as illustrated in figure 2. This significantly reduces communication latency. PCIe Fabrics expand on this capability by enabling remote systems to establish P2P communication. Intel Phi, GPUs, FPGA, specialized data grabbers can exploit remote P2P communication to reduce latency and communication overhead.

The SИСCI API supports this functionality and provides a simplified way to setup and manage remote peer-to-peer transfers. SИСCI software enables applications to use PIO or DMA operations to move data directly to and from local or remote PCIe devices.

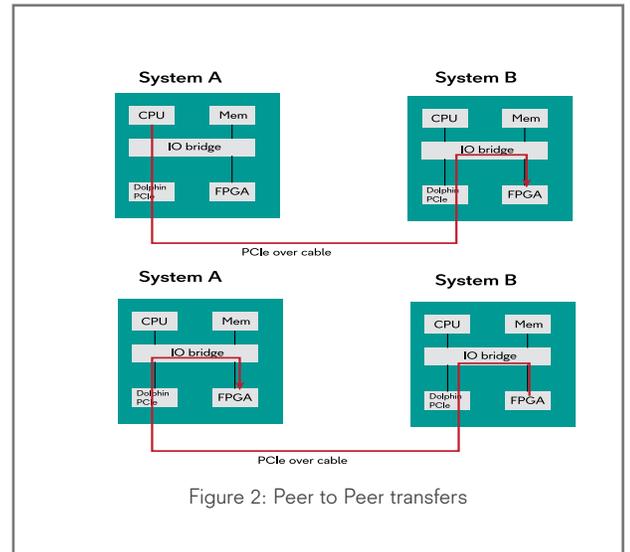


Figure 2: Peer to Peer transfers

### PCIe Device Lending Software

PCIe Device Lending offers a flexible way to enable PCIe IO devices (NVMeS, FPGAs, GPUs etc) to be accessed within a PCIe Fabric. Devices can be borrowed over the PCIe Fabric without any software overhead at PCIe speeds. Device Lending is a simple way to reconfigure systems and reallocate resources. GPUs, NVMe drives or FPGAs can be added or removed without having to be physically installed in a system on the fabric. The result is a flexible simple method of creating a pool of devices that maximizes usage. Since this solution uses standard PCIe, it doesn't add any software overhead to the communication path. Standard PCIe transactions are used between the systems. Dolphins eXpressWare software manages the connection and is responsible for setting up the PCIe Non-Transparent Bridge (NTB) mappings.

Two types of functions are implemented with device lending. These are the lending function and the borrowing function as outlined in figure 3. Lending involves making devices available on the fabric for temporary access. These PCIe devices are still located within the lending system. The borrowing function can lookup available devices. Devices can then be temporarily borrowed. When use of the device is completed, the device can be released and borrowed by other systems on the fabric or returned for local use.

Device lending also enables a SR-IOV device to be shared as a MR-IOV device. SR-IOV functions can be borrowed by any system in the PCIe Fabric. Thereby enabling the device to be shared by multiple systems. This maximizes the use of SR-IOV devices such as 100 Gbit Ethernet cards.

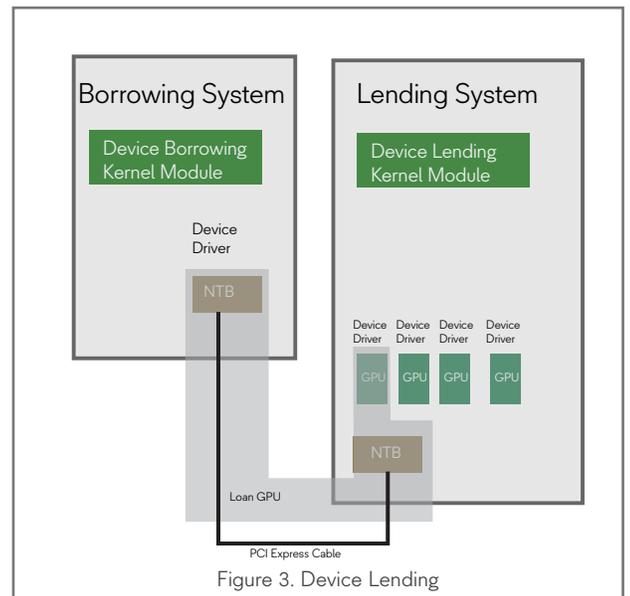


Figure 3: Device Lending



## Reflective Memory / Multi-cast

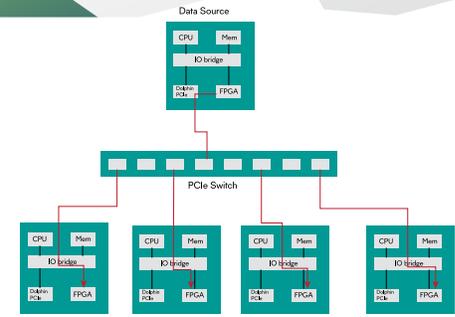
Dolphin's reflective memory or multi-cast solution reinterprets traditional reflective memory offerings. Traditional Reflective Memory solutions, which have been on the market for many years, implement a slow ring based topology. Dolphin's reflective memory solution uses a modern high speed switched architecture that delivers lower latency and higher throughput.

Dolphin's PCIe switched architecture employs multi-cast as a key element of our reflective memory solution. A single bus write transaction is sent to multiple remote targets or in PCI Express technical terms multi-cast capability enables a single Transaction Layer Packet (TLP) to be forwarded to multiple destinations. PCI Express multi-cast results in a lower latency and higher bandwidth reflective memory solution. Dolphin benchmarks show end-to-end latencies as low as 0.99µs and over 6000 MB/s dataflow at the application level. These performance levels solve many real time, distributed computing requirements.

Dolphin combines PCI Express multi-cast with the eXpressWare™ SISI (Software Infrastructure for Shared-memory Cluster Interconnect) API to allow customers to easily implement applications that directly access and utilize PCIe multi-cast. Applications can be built without the need to write device drivers or spend time studying PCIe chipset specifications.

Another main difference in Dolphin's reflective memory solution is the use of cached main system memory to store data. Cached main memory provides a significant performance and cost benefit. Remote interrupts or polling signal the arrival of data from a remote node. Polling is very fast since the memory segments are normal cached main memory and consume no memory bandwidth. The CPU polls for changes in its local cache. When new data arrives from the remote node, the I/O system automatically invalidates the cache and the new value is cached.

In addition, FPGA and GPU applications can



implement this reflective memory mechanism. The SISI API configures and enables GPUs, FPGAs, or any PCIe master device to send data directly to remote memory through the multi-cast mechanism, avoiding the need to first store the data in local memory. Data is written directly from a FPGA to multiple end points for processing or data movement. FPGAs can also receive data from multiple end points.

Reflective memory solutions are known for their simplicity, just read and write into a shared distributed memory. Our high-performance fabric increases simplicity with easy installation and setup. The SISI Developers Kit includes tools to speed development and setup of your reflective memory system. Once setup, your application simply reads and writes to remote memory.

## Features

- High-performance, ultra low-latency switched 64 GT/s and 40 GT/s data rates
- Gen 3 x8 performance up to 6000 MB/s data throughput
- Gen 2 x8 performance up to 2886 MB/s data throughput
- FPGA, GPU support
- Hardware based multi-cast
- Configurable shared memory regions
- Fiber-Optic and copper cabling support
- Scalable switched architecture
- SISI API support
- PCIe host adapters
- Expandable switch solutions

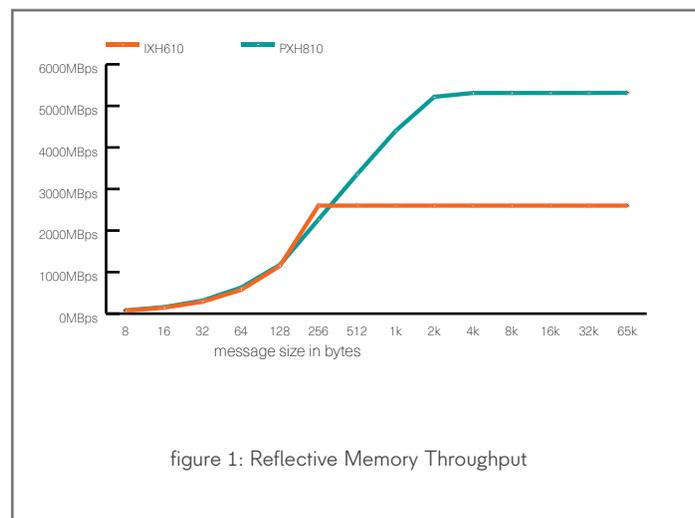
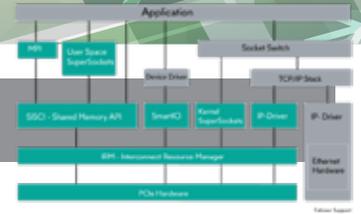


figure 1: Reflective Memory Throughput



# eXpressWare™ Software

## PCI Express Software Suite

eXpressWare™ software enables developers to easily migrate applications to PCIe Fabrics. eXpressWare's™ complete software infrastructure enables networking applications to communicate using standard PCIe over cables and backplanes. Several interfaces and APIs are supported including standard TCP/IP networking - IPoPCle driver, a low level direct remote memory access API – SISCI shared memory API and a sockets API - SuperSockets™. Each API has its benefits and can be selected based on application requirements.

The SISCI API enables customers to fully exploit the PCIe programming model without having to spend months developing device drivers. The API offers a C programming API for shared / remote memory access, including reflective memory/multi-cast functionality, peer to peer memory transfers

and RDMA capabilities. The SISCI API supports direct FPGA to FPGA, GPU to GPU, or any combination of communication with FPGA, CPUs, GPUs or memory over PCIe.

SuperSockets™ enables networked applications to benefit from a low latency, high throughput PCIe Fabric without any modifications. With SuperSockets™, a PCIe Fabric can replace local Ethernet networks. The combination of Dolphin's PCIe host adapters and switches with SuperSockets™ delivers maximum application performance without necessitating application changes. SuperSockets™ is a unique implementation of the Berkeley Sockets API that capitalizes on the PCIe transport to transparently achieve performance gains for existing socket-based network applications. Both Linux and Windows operating systems are supported,

so new and existing applications can easily be deployed on future high performance PCIe Fabrics.

Dolphin's performance optimized TCP IP driver for PCIe (IPoPCle) provides a fast and transparent way for any networked applications to dramatically improve network throughput. The software is highly optimized to reduce system load (e.g. system interrupts) and uses both PIO and RDMA operations to implement the most efficient transfer of all message sizes. The major benefits are plug and play, much higher bandwidth, and lower latency than network technologies like 10Gb/s Ethernet. The IPoPCle driver is targeted for non-sockets applications and functions that require high throughput.

- PCIe Gen 1,2,3 support
- Address based Multi-cast / reflective memory
- Point to point and switched fabric support
  - Scalable to 126 nodes
- Operating systems
  - Windows
  - Linux
  - VxWorks
  - RTX

- Low latency direct memory transfers
- Accelerated Loopback support
- Peer to peer transfers
- UDP and TCP support
- UDP multi-cast
- Cascading of switches
- FPGA and GPU direct memory transfers
- Low latency direct memory transfers

- PCIe chipset support
  - Microsemi
  - Broadcom /PLX
  - IDT
  - Intel NTB
- Cross O/S low latency data transfers
- Sockets Support
  - Berkeley Sockets
  - WinSock 2
- Fabric manager

## Specifications

Supported APIs	SISCI API Berkeley Sockets API Microsoft WinSock2/LSP support TCP/IP	Supported Platforms	x86 ARM 32 bit and 64 bit PowerPC
Application Performance	0.54 microsecond latency (application to application) Above 11 GB/s throughput	eXpressWare™ Packages	eXpressWare™ for Linux eXpressWare™ for Windows eXpressWare™ for RTX eXpressWare™ for VxWorks
Supported Components	Microsemi Broadcom/PLX IDT Intel NTB enabled servers	Dolphin Software	SuperSockets for Windows SuperSockets for Linux IPoPCle driver SISCI API IRM- Interconnect Resource Manager PCIe Fabric Manager
PCI Express	Base Specification 1.x, 2.x, 3.x Link widths 1-16 lanes		
Topologies	Switch/ point to point/ mesh		

# eXpressWare™ Software

## IPoPCle IP over PCle

Dolphin's performance optimized TCP IP driver for PCIe (IPoPCle) is targeted at non-sockets applications that require high throughput along with plug and play. This fast and transparent network driver dramatically improves network throughput. The software is highly optimized to reduce system load (e.g. system interrupts) and uses both PIO and RDMA operations to implement the most efficient transfers of all message sizes. IPoPCle offers much higher bandwidth and lower latency than standard network technologies like 40 GbE. Figure 4 illustrates the performance with Gen2 and Gen3 PCIe cards.

At the hardware level, the TCP/IP driver provides a very low latency connection. Yet, operating system networking protocols typically introduce a significant delay for safe networking (required for non-reliable networks like Ethernet). The IPoPCle driver still implements these networking protocols increasing latency. User space applications seeking the lowest possible network latency should utilize the Dolphin SuperSockets™ technology. The IPoPCle driver will typically provide 5-6 times better throughput than 10G Ethernet.

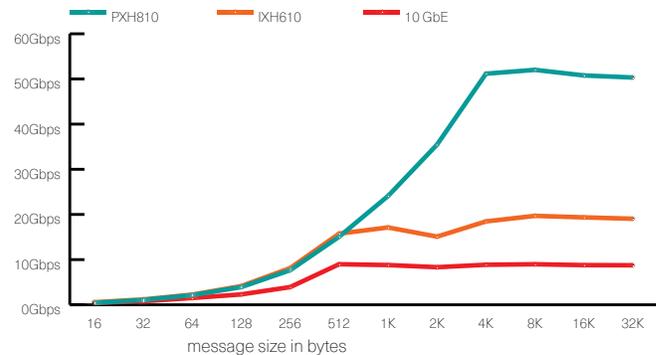


Figure 4: TCP/IP Throughput

## Features

- All networked, users space and kernel space applications are supported.
- 100% compliant with Linux Socket library, Berkeley Socket API and Windows WinSock2.
- No OS patches or application modifications required. Just install and run
- Routing between networks
- ARP support
- Both TCP and UDP supported. (UDP multi-cast/broadcast is not supported yet using Linux, but SuperSockets for Linux supports UDP multi-cast)
- Supports hot-pluggable links for high availability operation
- Easy to install

## IPoPCle Uses

The optimized TCP/IP driver is recommended for applications like

### Windows

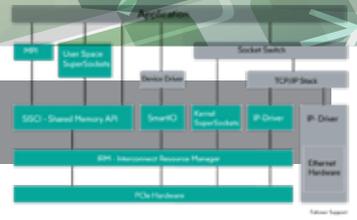
- Microsoft Hyper-V live migration
- Network file sharing (map network drive)
- Applications that require UDP (not support by SuperSockets yet).

### Linux:

- General networking
- NFS
- Cluster file systems not supported by SuperSockets
- iSCSI



# eXpressWare™ Software



## SISCI Low Level API

Dolphin’s Software Infrastructure Shared-Memory Cluster Interconnect (SISCI) API makes developing PCI Express Fabric applications faster and easier. The SISCI API is a well established API for shared memory environments. In PCI Express multiprocessing architectures, the SISCI API enables PCIe based applications to use distributed resources such as CPUs, I/O, and memory. The resulting applications feature reduced system latency and increased data throughput.

For processor to processor communication, PCI Express supports both CPU driven programmed IO (PIO) and Direct Memory Access (DMA) as transports through non-transparent bridges (NTB). Dolphin’s SISCI API utilizes these components in creating a

development and runtime environment for systems seeking maximum performance. This very deterministic environment featuring low latency and low jitter is ideal for traditional high performance applications like real time simulators, reflective memory applications, high availability servers with fast fail-over, and high speed trading applications.

The SISCI API supports data transfers between applications and processes running in an SMP environment as well as between independent servers. SISCI’s capabilities include managing and triggering of application specific local and remote interrupts, along with catching and managing events generated by the underlying PCIe system (such as a cable being unplugged). The SISCI API makes extensive

use of the “resource” concept. Resources are items such as virtual devices, memory segments, and DMA queues.

The API removes the need to understand and manage low level PCIe chip registers. At the application level, developers utilize these resources without sacrificing performance. Programming features include allocating memory segments, mapping local and remote memory segments into addressable program space, and data management and transfer with DMA. The SISCI API improves overall system performance and availability with advanced caching techniques, data checking for data transfer errors, and data error correction.

## Features

- Shared memory API
- PCI Express Peer to Peer support
- Replicated/reflective memory support
- Distributed shared memory and DMA support
- Low latency messaging API
- Interrupt management
- Direct memory reads and writes
- Windows, RTX, VxWorks, and Linux support
- Supports data transfers between all supported OS and platforms.
- Caching and error checking support
- Events and callbacks
- Example code available

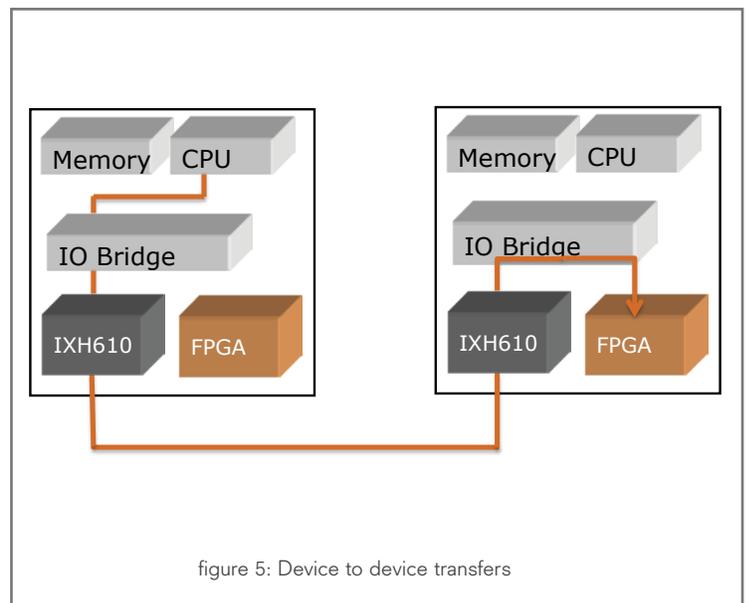


figure 5: Device to device transfers



## Why use SISCI?

The SISCI software and underlying drivers simplify the process of building shared memory based applications. For PCIe based application development, the API utilizes PCI Express non-transparent bridging to maximum application performance. The shared memory API drivers allocate memory segments on the local node and make this memory available to other nodes. The local node then connects to memory segments on remote nodes.

Once available, a memory segment is accessed in two ways, either mapped into the address space of your process and accessed as a normal memory access, e.g. via pointer operations, or use the DMA engine in the PCIe chipset to transfer data. Figure 6 illustrates both data transfer options.

Mapping the remote address space and using PIO may be appropriate for control

messages and data transfers up to e.g. 1k bytes, since the processor moves the data with very low latency. PIO optimizes small

instruction. A DMA implementations saves CPU cycles for larger transfers, enabling overlapped data transfers and computations.

DMA has a higher setup cost so latencies usually increase slightly because of the time required to lock down memory and setup the DMA engine and interrupt completion time. However, more data transfers joined and sent together to the PCIe switch in order amortizes the overhead.

The built in resource management enables multiple concurrent SISCI programs and other users of the PCIe Fabric to coexist and operate independent of each other. The SISCI API is available in user space and a similar API is available in kernel space.

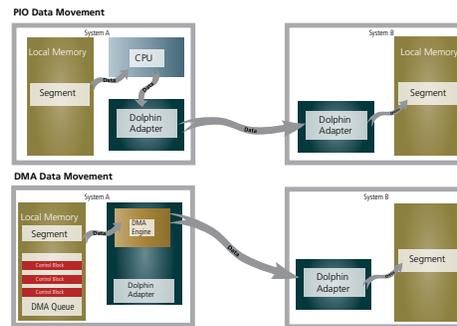
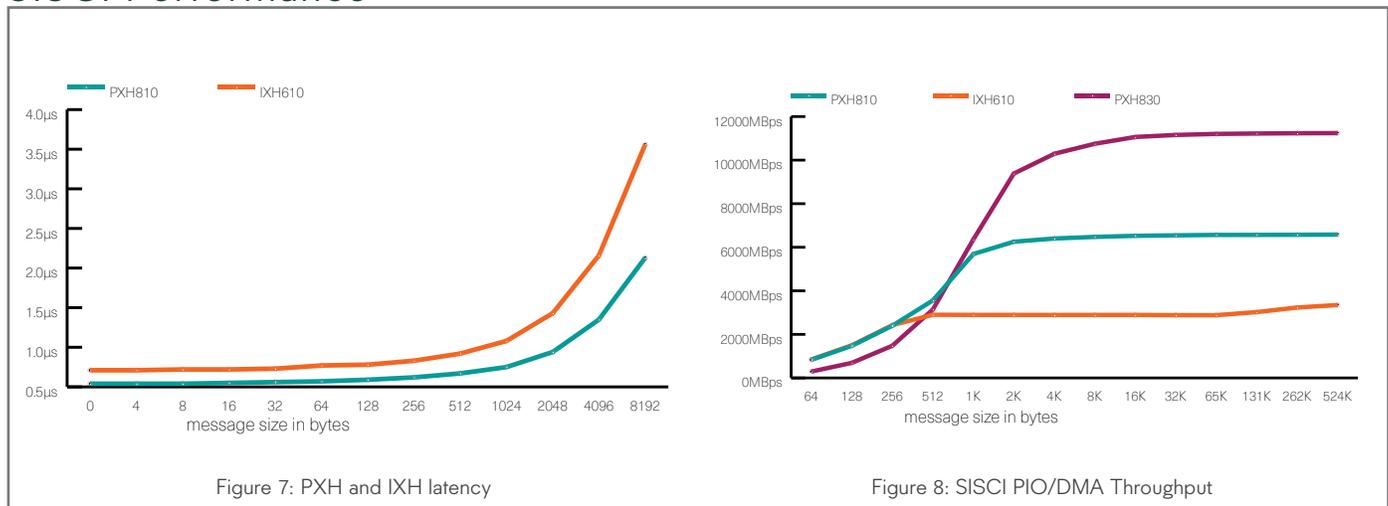


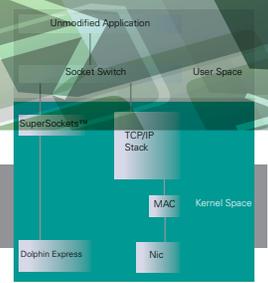
figure 6: SISCI data movement model

write transfers by requiring no memory lock down, data may already exist in the CPU cache, and the actual transfer is just a single CPU instruction – a write posted store

## SISCI Performance



The SISCI API provides applications direct access to the low latency messaging enabled by PCI Express. Dolphin SISCI benchmarks show latencies as low as 0.54μs. The chart on Figure 7 show the latency at various message sizes. The SISCI API enables high throughput applications. This high performance API takes advantage of the PCI Express hardware performance to deliver over 11 GB/s for Gen 3 and 3500 MB/s for Gen 2 of real application data throughput. Figure 8 shows the throughput at various message sizes using Dolphin IXH and PXH host adapters



# eXpressWare™ Software

## SuperSockets™

PCI Express can replace local Ethernet networks with a high speed low latency network. SuperSockets is a unique implementation of the Berkeley Sockets API. With SuperSockets , network applications transparently capitalize on the PCIe transport to achieve performance gains.

Dolphin PCIe hardware and the SuperSockets software support the most demanding sockets based applications with an ultra-low latency, high-bandwidth, low overhead, and highly available platform. New and existing Linux and Windows applications require no modification to be deployed on Dolphin’s high-performance platform.

Traditional implementations of TCP sockets require two major CPU consuming tasks: data copy between application buffers and NIC buffers along with TCP transport handling (segmentation, reassembly, check

summing, timers, acknowledgments, etc). These operations turn into performance bottlenecks as I/O interconnect speeds increase. SuperSockets eliminates the protocol stack bottlenecks, delivering superior latency performance. Our ultra-low latency remote memory access mechanism is based on a combination of PIO (Programmed IO) for short transfers and DMA (Direct Memory Access) for longer transfers, allowing both control and data messages to experience performance improvements.

SuperSockets is unique in its support for PIO. PIO has clear advantages for short messages, such as control messages for simulations systems. Transfers complete through a single CPU store operation that moves data from CPU registers into remote system memory. In most cases, SuperSockets data transfers complete before alternative technologies start their RDMA transfer.

In addition to PIO, SuperSockets implements a high-speed loopback device for accelerating local system sockets communication. This reduces local sockets latency to a minimum. For SMP systems, loopback performance is increased 10 times.

SuperSockets comes with built in high availability, providing instantaneous switching during system or network errors. If the PCI Express® Fabric fails, socket communication transfers to the regular network stack. The Linux version supports an instant fail- over and fail-forward mechanism between the PCIe and regular network.

## Features

- Windows and Linux support
- Full support for socket inheritance/duplication
- Includes local loopback socket acceleration up to 10 times faster than standard Linux and Windows
- No OS patches or application modifications required
- Easy to install with no application modifications
- Linux to Windows connectivity available soon

### Linux Specific Features

- TCP, UDP, and UDP multi-cast support
- Supports both user space and kernel space applications
- Compliant with Linux Kernel Socket library and Berkeley Sockets
- Transparent fail-over to Ethernet if high speed connection fails and falls back when problem is corrected

### Windows Specific Features

- TCP support , UDP and UDP multi-cast being implemented
- Supports user space applications
- Compliant with WinSock2 API
- Fail-over to Ethernet if high speed connection is not available at start-up



## How Does SuperSockets™ Work?

To divert socket communication without touching the application, the sockets API functions must be intercepted. This is done differently in Windows and Linux environments.

Dolphin SuperSockets on Linux differs from regular sockets only in the address family. SuperSockets implement an AF\_INET compliant socket transport called AF\_SSOCK. The Linux LD\_PRELOAD functionality is used to preload the standard sockets library with a special SuperSockets library that intercepts the socket () call and replaces the AF\_INET address family with AF\_SSOCK. All other sockets calls follow the usual code path. Target addresses within the PCI Express Fabric are accelerated by the SuperSockets module.

For Windows applications or services, a Layered Service Provider(LSP) module is

installed and automatically configured. The LSP accelerates socket transfers initiated by AF\_INET or AF\_INET6, SOCK\_STREAM endpoints. The SuperSockets stack provides a proxy application called dis\_ssocks\_run.exe that enables specific programs to use the PCI Express path. By default, the LSP is a

The network acceleration over PCI Express occurs when the interconnect topology is fully functional, the client and server programs are launched under the proxy application's control and both sides use the standard Winsock2 API calls. At runtime, a native socket is created and used for initial connection establishment. Therefore, all connections are subject to typical network administrative policies.

The supported transfer modes are blocking, non-blocking, overlapped, asynchronous window and network events. The Service Provider balances the CPU consumption based on the traffic pattern. Dedicated operating system performance counters are additionally provided.

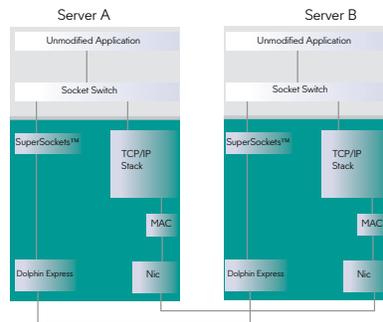


figure 9: SuperSockets™ vs. Ethernet data model

pass-through module for all applications: the network traffic passes through the NDIS stack.

## SuperSockets™ Performance

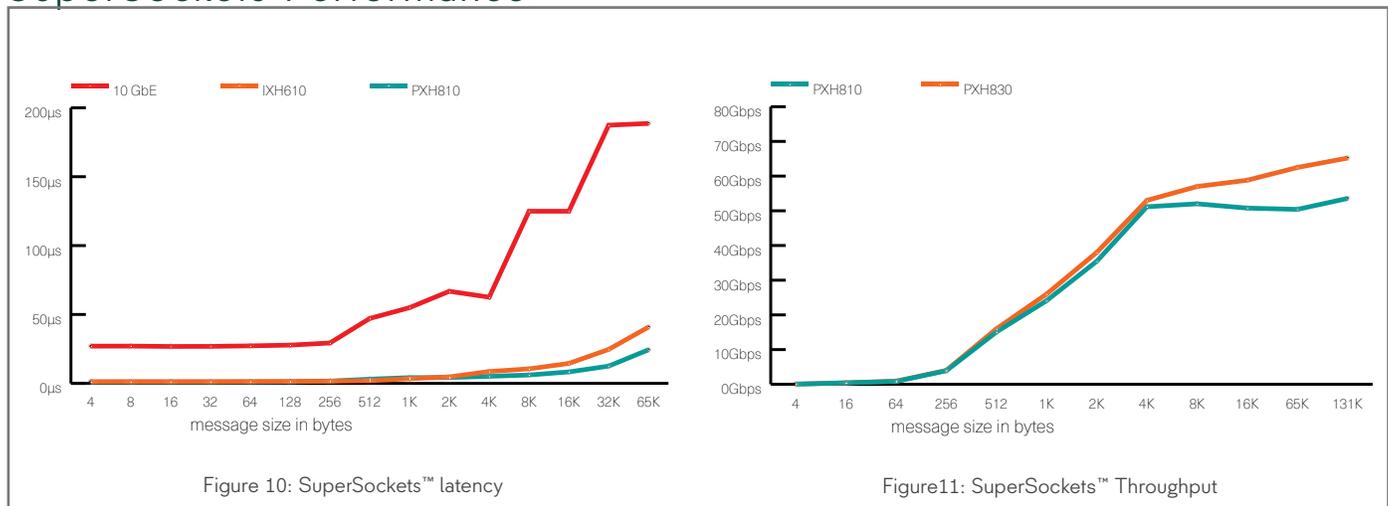


Figure 10: SuperSockets™ latency

Figure 11: SuperSockets™ Throughput

SuperSockets is optimized for high throughput, low latency communication by reducing system resource and interrupt usage in data transfers. The latency chart above shows performance results using PCI Express vs 10 Gigabit Ethernet. The socket ping-pong test shows the half RTT (Round Trip Time). The minimum latency for Dolphin SuperSockets is under 1 microseconds. SuperSockets also delivers high throughput with over 53 Gb/s of data throughput with our Gen3 PXH810 product.



# PCIe Hardware



## PXH830 Gen 3 PCIe NTB Adapter

The PXH830 Gen3 PCI Express NTB Host Adapter is a high performance cabled interface to external processor subsystems. Based on Broadcom® Gen3 PCI Express bridging architecture, the PXH830 host adapter includes advanced features for non-transparent bridging (NTB) and clock isolation. The PXH830 card have a standard Quad SFF-8644 connector and uses standard MiniSAS-HD cables.

For high performance application developers, the PXH830 host adapter combines 128 GT/s performance with an application to application latency starting at 0.54 microseconds. Inter-processor communication benefits from the high throughput and low latency. Using the latest SmartIO technology software from Dolphin, applications can now access remote PCIe devices as if they were attached to the local system.

The PXH830 performs both Remote Direct Memory Access (RDMA) and Programmed IO (PIO) transfers, effectively supporting both large and small data packets. RDMA transfers result in efficient larger packet transfers and processor off-load exceeding 11 Gigabytes per second. PIO transfers optimize small packet transfers at the lowest latency. The combination of RDMA and PIO creates a highly potent data transfer system.

The PXH830 supports our eXpressWare™ software suite which takes advantage of PCI Express' RDMA and PIO data transfer scheme. eXpressWare™ software delivers a complete deployment environment for customized and standardized applications. The suite includes a Shared-Memory Cluster Interconnect (SISCI) API as well as a TCP/IP driver and SuperSockets software. The SISCI API is a robust and powerful shared memory programming environment.

The optimized TCP/IP driver and SuperSockets™ software remove traditional networking bottlenecks, allowing standard IP and sockets applications to take advantage of the high-performance PCI Express interconnect without modification. The overall framework is designed for rapid development of inter-processor communication systems.

The PXH830 is carefully designed for maximum cable length and supports copper cables up to 9 meters at full PCI Express Gen3 speed. Fiber optics extends this distance to 100 meters.

The PXH830 card comes with a full license to the Dolphin eXpressWare software. The PXH832 Gen3 Adapter card does not include any software license and is well suited for high performance Transparent IO Expansion applications.

## Features

- PCI Express Gen3 compliant - 8.0 GT/s per lane
- Link compliant with Gen1, Gen2, and Gen3 PCIe
- Quad SFF-8644 connector
  - » PCI Express 3.0 cables
  - » MiniSAS -HD cables
- Four x4 Gen3 PCI Express cable ports that can be configured as:
  - One - x16 PCI Express port
  - Two - x8 PCI Express ports
  - Two NTB ports
- RDMA support through PIO and DMA
- Copper and fiber-optic cable connectors
- Full host clock isolation. Supports hosts running both CFC and SSC
- Non-transparent bridging to cabled PCI Express systems
- Low Profile PCIe form factor
- EEPROM for custom system configuration
- Link status LEDs through face plate



## Cluster connections

When used for multi-processor connections, the PXH830 adapter is capable of connecting up to three nodes at Gen3 x8 without a switch as shown in figure 12 or two nodes at Gen3 x16. Each port is 32 GT/s. Two ports create a 64 GT/s x8 link. Four ports create a 128 GT/s x16 link. All ports have latencies as low as 0.54 microseconds. The PXH830 supports any system with a standard x16 PCIe slot.

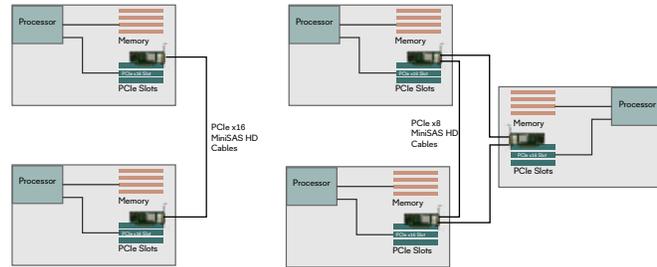


Figure 12: Switchless PXH830 Configurations

## Performance

Each connection supports 32 GT/s with a maximum of 128 GT/s. Figure 13 illustrates the latency at various packet sizes. The bottom axis are packet sizes the side axis is latency in microseconds. PXH830 latencies are as low as 0.54 microseconds.

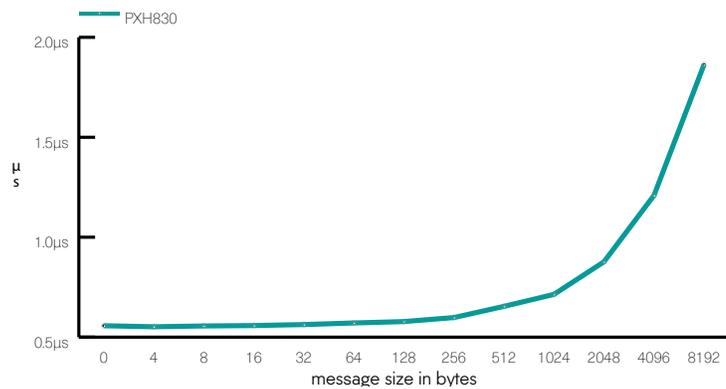


Figure 13: PXH830 Latency

## Specifications

Link Speeds	32 GT/s per port /128 GT/s	Configuration	DIP-switch NTB /Long cable /short cable /safe boot
Application Performance	0.54 microsecond latency (application to application) <130 nanoseconds cut through latency port to port	Regulatory	CE Mark FCC Class A UL94V-0 compliant RoHS Compliant
Active Components	Broadcom/PLX Gen 3 PCIe Switch with DMA	Operating Systems	Windows Linux VxWorks RTX
PCI Express	Base Specification 3	Operating Systems supported	Windows, RTX, Linux, VxWorks
Topologies	Point to point or 3 node Mesh Topology	Operating Environment	Operating Temperature: 0°C - 55°C (32°F - 131°F) Operating Temperature with AOC: 0°C - 45°C (32°F - 113°F) Air Flow: 150 LFM  Operating Temperature: 0°C - 50°C (32°F - 122°F) Air Flow: ~0 LFM  Relative Humidity: 5% - 95% (non- condensing)
Cable Connections	Quad SFF-8644 connector PCIe 3.0 and MiniSAS HD cables copper cables, 0.5 - 9 meters Fiber Optic -100 meters	Product Codes	PXH830 PCIe Fabric Adapter
Power Consumption	10 Watts (typical, 14 Watts worst case ) + 800 milliwatts (typical) pr connected x4 AOC		
Mechanical Dimensions	Low profile, Half Height - 68.90 mm (2.731 inches) x 167.65 mm (6.600 inches)		
Dolphin Software	SuperSockets Berkley Sockets API Microsoft WinSock2 /LSP support IPoPCle driver SISCI API Device Lending and SmartIO		
PCIe Bracket	Full height plate mounted Half height plate included		



# PCIe Hardware



## PXH832 Gen 3 Host/Target Adapter

The PXH832 Gen3 PCI Express Host Adapter is a high performance cabled interface to external IO subsystems. Based on Broadcom® Gen3 PCI Express bridging architecture, the PXH832 Transparent Host and Target adapter includes advanced features for transparent bridging (TB), host and clock isolation. The PXH832 card utilizes a quad SFF-8644 connector and can use standard MiniSAS-HD or the new PCIe 3.0 cable currently being standardized by the PCI-SIG.

The PXH832 can be configured as four x4 PCIe connections, two x8 or one x16 PCIe connection. For GPU farms, test and

measurement equipment, medical equipment, and storage subsystem, the PXH832 delivers flexibility and 128 GT/s with extremely high data quality. The adapter is configurable as either a host adapter or target adapter using an on-board DIP-Switch to select between Host and Target operations.

The PXH832 supports the new CMI functionality when used with a PCIe 3.0 cable and can therefore connect to a CMI compliant PCIe target device. The PXH832 can also be used with a standard MiniSAS-HD cable when connecting to a PXH832 in target mode. When used with a MiniSAS-HD cable, only the reset sideband

signal will be emulated by the PXH832.

The PXH832 is carefully designed for maximum cable length and supports MiniSAS-HD copper cables up to 9 meters at full PCI Express Gen3 speed. Fiber optics extends this distance to 100 meters.

## Features

- PCI Express Gen3 compliant - 8.0 GT/s per lane
- Link compliant with Gen1, Gen2, and Gen3 PCIe
- Host and target operation
- Quad SFF-8644 connector
  - » PCI Express 3.0 cables
  - » MiniSAS -HD cables
- Four x4 Gen3 PCI Express cable ports that can be configured as:
  - » One - x16 PCI Express port
  - » Two - x8 PCI Express ports
- Copper and fiber-optic cable connectors
- Full host clock isolation. Supports hosts running both CFC and SSC
- Low Profile, half length PCIe form factor
- EEPROM for custom system configuration
- Link status LEDs through face plate



## I/O Expansion

The PXH832 adapter is capable of being configured as a host or a target adapter. Figure 14 illustrates the various configurations supported by the PXH832. It can support up to 4 PCIe expansion units without an external switch. As a host adapter, the PXH832 connect to PCIe targets via MiniSAS HD cables. It can connect to other PXH832 boards as a target or a custom target adapters. The PXH832 can be configured as a x16, x8, or x4 connection.

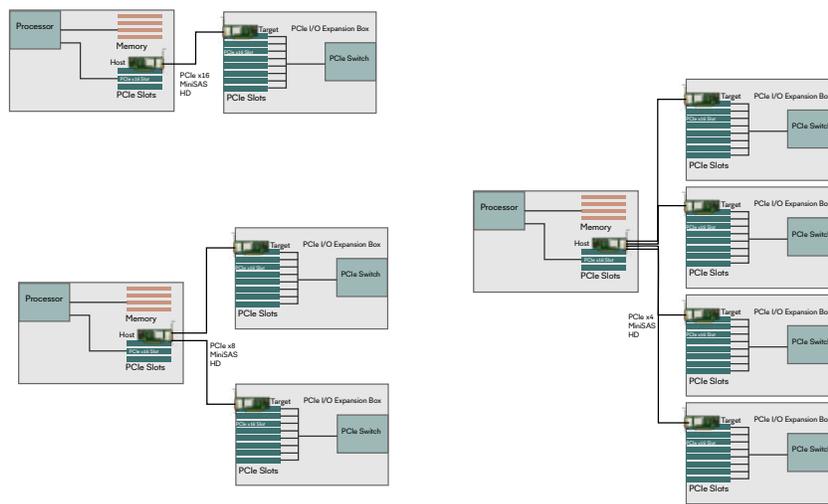


Figure 14: PXH832 I/O Expansion Configurations

## Specifications

Link Speeds	32 GT/s per port / 128 GT/s	Regulatory	CE Mark FCC Class A UL94V-0 compliant RoHS Compliant
Application Performance	<130 nanoseconds cut through latency port to port	Operating Environment	Operating Temperature: 0°C - 55°C (32°F - 131°F) Operating Temperature with AOC: 0°C - 45°C (32°F - 113°F) Air Flow: 150 LFM
Active Components	Broadcom /PLX Gen 3 PCIe Switch	Configuration	DIP-switch Host /Target /Long cable /short cable /safe boot
PCI Express	Base Specification 3	Product Codes	PXH832 Host/Target Adapter
Topologies	Transparent Host/Target up to 4 devices		
Cable Connections	Quad SFF-8644 connector PCIe 3.0 and MiniSAS HD cables copper cables, 0.5 - 9 meters Fiber Optic -100 meters		
Power Consumption	10 Watts (typical, 14 Watts worst case ) + 800 miliwatts (typical) pr connected x4 AOC		
Mechanical Dimensions	Low profile, Half Length - 68.90 mm (2.731 inches) x 167.65 mm (6.600 inches)		
PCIe Bracket	Full height plate mounted Half height plate included		



# PCIe Hardware



## MXH830 Gen 3 PCIe NTB Adapter

The MXH830 Gen3 PCI Express NTB Host Adapter is our high-performance clustering product. Based on Microsemi PXF Switchtec® Gen3.1 PCI Express bridging architecture, the MXH830 host adapter includes advanced features for non-transparent bridging (NTB) and clock isolation.

For high performance application developers, the MXH830 host adapter combines 128 GT/s performance with an application to application latency starting just above 500 nanoseconds. Inter-processor communication benefits from the high throughput and low latency. Using the latest SmartIO technology software from Dolphin, GPU / Cuda applications can now stream data to remote GPUs at the same speed as to local GPUs.

The card implements a quad SFF-8644 connector which is compliant with the latest MiniSAS-HD and PCI SIG External cable

specification 3.0 cables. Cable distance extend to 9 meters with copper cables and 100 meters with fiber cables. These four ports can be aggregated into a single x16 port that delivers a 128 GT/s data link to remote systems or partitioned to two x8 ports or four x4 ports.

The MXH830 includes an eXpressWare™ software suite license. eXpressWare provides a software infrastructure for developing PCIe applications. This software suite takes advantage of system DMA and PIO data transfer schemes to create a complete environment for customized and standardized applications. The software suite includes a Shared-Memory Cluster Interconnect (SISCI) API as well as a IPoPCle drive, SuperSockets software and SmartIO software. The SISCI API is a robust and powerful shared memory programming environment. This API enables the advanced features of PCIe such as PCIe multi-cast, remote peer-to-peer transfers, and allocation

of various size system or device memory for data transfers. The optimized IPoPCle driver and SuperSockets software remove traditional networking bottlenecks, allowing standard IP and sockets applications to take advantage of PCIe' high performance without modification. The overall framework is designed for rapid deployment of inter-processor communication systems. SmartIO software enables peer to peer transfers and device lending between systems.

The MXH830 support both switchless and switched configurations. For scaling out beyond five nodes, the MXH830 is combined with Dolphin MXS824 switch. The MXS824 is a 24 port 1U PCIe switch that can be configured a twenty-four x4 ports, twelve x8 ports, or six x16 ports. This switch can also be cascaded to create large topology that can expand up to 64 or 128 nodes.

## Features

- PCI Express Gen3 - 8.0 GT/s per lane
- Link compliant with Gen1, Gen2, and Gen3 PCIe
- Quad SFF-8644 connector
  - » PCI Express 3.0 cables
  - » MiniSAS -HD cables
- RDMA support through PIO and System DMA
- Copper and fiber-optic cable support
- Full host Clock isolation support. Automatic support for host running CFC or SSC.
- Low Profile PCIe form factor
- EEPROM for custom system configuration
- Link status LEDs through face plate
- <170ns - Cut Through latency
- MXH830 supports the following port configurations
  - » One - x16 PCIe port
  - » Two - x8 PCIe ports
  - » Four - x4 PCIe ports
- MXH830 can be configured into a 5 node cluster using 4 x4 ports to each host or a 3 node cluster using x8 connections, or a two node x16 connection
- Combined with the MXS824 can scale to up to 64 or 128 nodes

<sup>1</sup> System DMA is only available with select systems



## eXpressWare PCIe Software

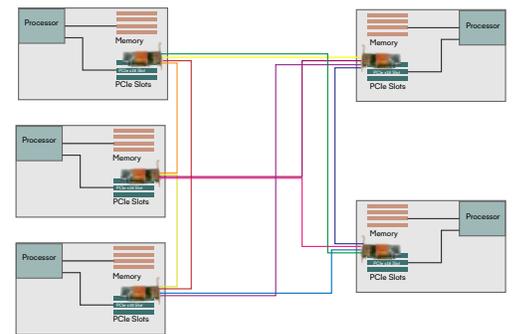
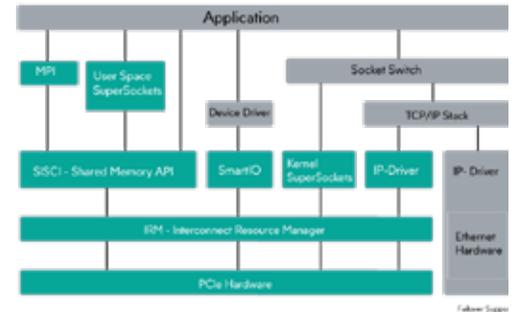
The MXH830 supports Dolphin's eXpressWare™ software suite. eXpressWare™ software enables customer to easily migrate applications to a PCIe Fabrics. A complete software infrastructure, eXpressWare enabling applications to network using standard PCI Express over cables. Several interfaces and APIs are supported including standard TCP/IP networking - IPoPCle driver, a low level direct remote memory access API - SISCI API, and a sockets API - SuperSockets. Each API has its benefits and can be selected based on application requirements.

The SISCI API enables customers to fully exploit the PCIe programming model without having to spend months developing device drivers. The API offers a C programming API for shared / remote memory access, including reflective memory/multi-cast functionality and RDMA capabilities.

SuperSockets enabled networked applications to benefit from a low latency, high throughput PCIe Fabric without any modifications. With SuperSockets, a PCIe Fabric can replace local Ethernet networks.

Dolphin's performance optimized TCP IP driver for PCIe (IPoPCle) provides a fast and transparent way for any networked applications to dramatically improve network throughput.

Dolphin SmartIO adds new features to standard PCIe IO. It enable device lending, which enables standard IO devices to be logically moved between systems. It also provides an infrastructure for peer to peer memory transfers. SmartIO supports direct FPGA to FPGA, GPU to GPU, or any combination of communication with FPGA,CPU's, GPU's or memory over PCIe.



Five node MXH830 cluster

## Specifications

Link Speeds	32GT/s per port / 128 GT/s total	Operating Environment	Operating Temperature: 0°C - 55°C (32°F - 131°F) AirFlow: 150 LFM Relative Humidity: 5% -95% non-condensing
Application Performance	~500ns latency (application to application) Above 10 Gbytes/s throughput	Dolphin Software	SuperSockets Berkley Sockets API Microsoft WinSock2/LSP support IPoPCle driver SISCI API Device Lending and SmartIO
Active Components	Microsemi Switchtec PFX Switch	Usage Modes	Non-transparent bridging
PCI Express	Base Specification 3.1 Cable Specification 3.0 ( preliminary) Card Electromechanical Specification 3.0	Regulatory	CE Mark FCC Class B UL94V-0 compliant RoHS Compliant
Topologies	Two nodes direct cable Three to five nodes mesh topology Switched Topology with MXS824	Configuration	DIP-switch x4,x8,x16 link / Safe Boot
Cut-Through Latency	<170ns	Mounting Plates	Full height plate installed Half height plate included with board
Cable Connections	SFF-8644 connector for copper / fiber cables Supports 4 - x4/ 2 - x8 or 1 - x16 connections PCIe 3.0 copper and MiniSAS-HD copper / fiber up to 9m with copper, 100m with fiber	Operating Systems supported	Windows, Linux, VxWorks
Power Consumption	12 Volt: Max 19.6 Watts 12 Volt: Max 24 Watts including AOC +3.3 Volt: Not connected +3.3 Volt AUX: Max 1 Watt	Product Codes	MXH830 Host NTB Adapter
Mechanical Dimensions	Low profile - 69.2 mm ( 2.7 inches) x 168.2 mm (6.6 inches)		



# PCIe Hardware



## MXH832 Gen 3 Host/Target Adapter

The MXH832 Gen3 Host/Target Adapter is our high performance transparent solution for connecting servers to external I/O subsystems. Based on Microsemi® Gen3.1 PCI Express bridging architecture, the MXH832 adapter enables reliable transparent PCIe cable connections and clock isolation.

For high performance application, such as GPU farms, test and measurement equipment, medical equipment, and storage subsystem, the MXH832 delivers flexibility and performance with extremely high data quality. It combines 128 GT/s performance with the low latency of PCI Express, which is ideal for external I/O Expansion systems. The MXH832 card has a quad SFF-8644 cable connector and supports MiniSAS-HD cables as well as the new PCIe 3.0 cables. Using the MXH832 as both Host and Target, standard MiniSAS-HD cables enables basic

transparent Host and Target operations including propagation of host reset signal to the downstream device.

The card also supports the new PCIe 3.0 cables and the new PCI-SIG CMI (Cable Management Interface ) protocol including WAKE and as defined in the current PCI-SIG External Cable specification 3.0. When the MXH832 is used in conjunction with the new PCI-SIG external PCIe 3.0 cables, it can be connected to any compliant downstream device.

The adapter is carefully designed to ensure maximum cable length and signal quality, with a target of up to 9 meters at full PCI Express Gen3 speed. Fiber optic cables will extend this distance to 100 meters.

The MXH832 uses an on-board DIP-Switch to

select between Host and Target operations and to set other operational parameters. The adapter cable port can be configured as four x4, two x8 or one x16 PCIe connections.

The MXH832 Adapter card comes with the eXpressWare™ Board Management software for simplified monitoring and management. The adapter also comes with optional support for SMBus and configuration via a Board Management Controller.

The MXH832 card is a transparent product only. To utilize the Dolphin's PCIe eXpressWare™ clustering software, you need to purchase the MXH830 PCIe Gen3 NTB adapter card.

## Features

- PCI Express Gen3 compliant - 8.0 GT/s per lane
- Microsemi PFX 32 lane switch
- Link compliant with Gen 1, Gen 2, and Gen 3 PCIe
- Host and target operation
- Quad SFF-8644 connector
  - » PCI Express 3.0 cables
  - » MiniSAS-HD cables
- Copper and fiber-optic cables
- MXH832 supports the following port configurations
  - » One - x16 PCIe port
  - » Two - x8 PCIe ports
  - » Four - x4 PCIe ports
- Full host Clock isolation support. Automatic support for host running CFC or SSC.
- Low Profile, half length PCIe form factor
- EEPROM for custom system configuration
- Link status LEDs through face plate



## I/O Expansion

The MXH832 adapter is capable of being configured as a host or a target adapter. Figure 1 illustrates the various configurations supported by the MXH832. It can support up to 4 PCIe expansion units without an external switch. As a host adapter, the MXH832 connects to PCIe targets via an SFF-8644 connector, which support PCIe 3.0 and MiniSAS HD cables. It can connect to other MXH832 boards as a target or to custom target adapters with PCIe 3.0 cables. The MXH832 configurations support x16, x8, or x4 connections.

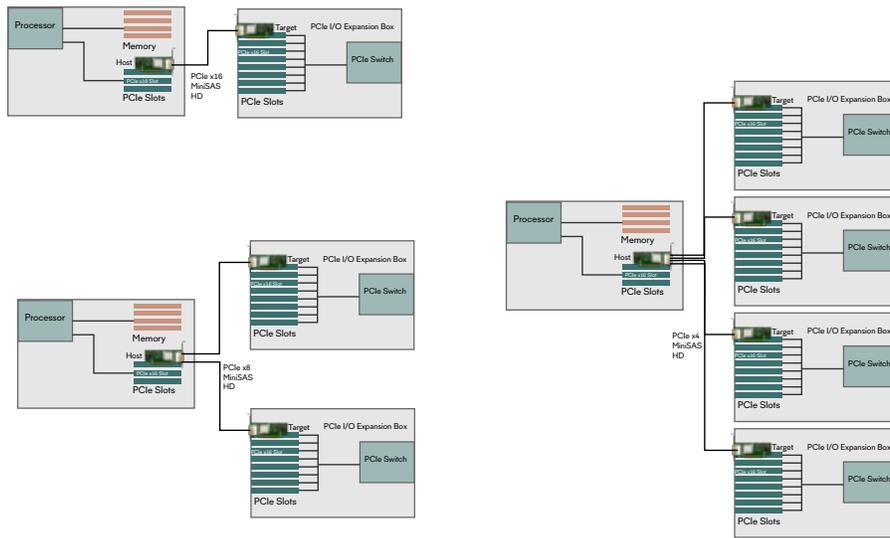


Figure 1:MXH832 I/O Expansion Configurations

## Specifications

Link Speeds	32 GT/s per port / 128 GT/s	Regulatory (pending)	CE Mark
Application Performance	170 nanoseconds cut through latency port to port		FCC Class A
Active Components	Microsemi PFX Gen 3 PCIe Switch		UL94V-0 compliant
PCI Express	Base Specification 3.1 Cable Specification 3.0 ( preliminary) Card Electromechanical Specification 3.0		RoHS Compliant
Topologies	Transparent Host/Target up to 4 devices	Operating Environment	Operating Temperature: 0°C - 55°C (32°F - 131°F) Air Flow: 150 LFM Relative Humidity: 5% - 95% (non- condensing)
Cable Connections	SFF-8644 connector for copper / fiber cables Supports 4 - x4/ 2 - x8 or 1 - x16 connections PCIe 3.0 copper and MiniSAS-HD copper / fiber Distance - up to 9m with copper, 100m with fiber	Configuration	DIP-switch Host / Target / x4,x8,x16 link / Safe Boot / Edge Clock Select
Power Consumption	12 Volt: Max 19.6 Watts 12 Volt: Max 24 Watts including AOC +3.3 Volt: Not connected +3.3 Volt AUX: Max 1 Watt	Product Codes	MXH832 Host/Target Adapter
Mechanical Dimensions	Low profile, Half Length - 69.2 mm (2.7 inches) x 168.2 mm (6.6 inches)		
PCIe Bracket	Full height plate mounted Half height plate included		



# PCIe Hardware



## PXH810 Gen 3 PCIe NTB Adapter

Based on Broadcom (PLX) Gen3 PCI Express switch architecture, the PXH810 host adapter includes advanced features for non-transparent bridging (NTB) and clock isolation.

For high performance application developers, the PXH810 host adapter combines 64 Gbit/s performance with less than one microsecond latency, significantly improving overall inter-system communication.

The PXH810 performs both Remote Direct Memory Access (RDMA) and Programmed IO (PIO) transfers, effectively supporting both large and small data packets. RDMA transfers result in efficient larger packet transfers and processor off-load. PIO transfers optimize small packet transfers at the lowest latency. The combination of RDMA and PIO creates a highly potent data transfer system.

With the implementation of clock isolation, the PXH810's signal quality is excellent. By isolating the system clock and transmitting an extremely low jitter high quality clock to downstream devices, the PXH810 offers users high signal quality and increased cable distances. The PXH810 supports PCI Express Gen3 x8 fiber optic cables up to 100 meters. The improved signal quality and fiber optic support makes the PXH810 ideal for simulation systems. The PXH810 is ideal in applications such as test and measurement equipment, medical equipment, and storage subsystem seeking high performance and data quality.

The PXH810 is carefully designed for maximum cable length and support copper cables up to 5 meters at full PCI Express Gen3 speed when connected to another

PXH810 card. Fiber optics is supported up to 100 meters at full speed.

The PXH810 can be connected to the IXS600 Switch to create PCIe Gen3 Fabrics up to 8 nodes.

The PXH810 supports the complete eXpressWare™ suite, including the SISCi API, SuperSockets™, and the IPoPCIe driver. The PXH810 card comes with a full license to the Dolphin eXpressWare™ clustering software.

## Features

- Gen 3 64 GT/s
- PCI Express Gen3 compliant - 8.0 GT/s per lane
- Link compliant with Gen1, Gen2, and Gen3 PCIe
- PCIe iPass Connectors
- One x8 PCIe port
- RDMA support through PIO and DMA
- Non-transparent bridging
- Copper connection up to 5 meters
- Fiber-optic cable connection up to 100 meters
- Clock isolation, full CFC and SSC support
- Non-transparent bridging to cabled PCIe systems
- Low Profile PCIe form factor
- EEPROM for custom system configuration
- Link status LEDs through face plate



## Cluster and I/O connections

The PXH810 adapter delivers extremely high throughput. This is demonstrated by using the SISI API. The PXH810 PCIe hardware performance delivers over 6500 MB/s of real application data throughput for high performance communication. Figure 15 shows the throughput at various message sizes using the PXH810 Gen 3 host adapters.

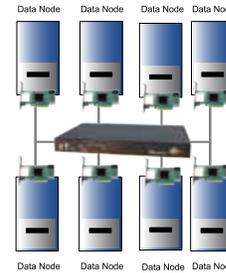


figure 15: Eight node PXH810 cluster

## Performance

When used for inter-host connections, the PXH810 adapter is capable of node to node connections or connections through a IXS600 Switch as shown in figure 16. Each connection supports 64 GT/s with latencies as low as 0.54 microseconds. Designed for x8 PCIe Systems, the PXH810 supports any system with a standard x8 or x16 PCIe slot. Gen3 operation requires a system that supports Gen3 PCI Express.

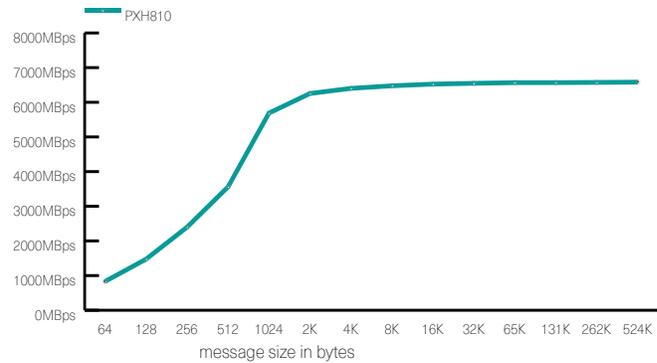


figure 16: PXH810 Throughput

## Specifications

Link Speeds	64 GT/s	Usage Modes	Host/Target Non-transparent bridging
Application Performance	0.54 microsecond latency (application to application) 6595 MB/s throughput	Regulatory	CE Mark FCC Class A UL94V-0 compliant RoHS Compliant
Active Components	PLX PCI Express Switch	Configuration	Dip-switch configurable
PCI Express	Base Specification 3	Mounting Plates	Full height plate installed Half height plate included with board
Topologies	Switch/ point to point	Operating Systems supported	Windows, RTX, Linux, VxWorks
Cable Connections	x8 iPass copper cable, fiber optic cable support Supports x8 to x4 transition cables	Product Codes	PXH810 Host NTB Adapter
Power Consumption	12 watts		
Mechanical Dimensions	Low profile - 68.90 mm ( 2.731 inches) x 120 mm (6.600 inches)		
Operating Environment	Operating Temperature: 0°C -55°C Relative Humidity: 5% -95% non-condensing		
Dolphin Software	SuperSockets Berkley Sockets API Microsoft WinSock2/LSP support IPoPCle driver SISI API		

# PCIe Hardware



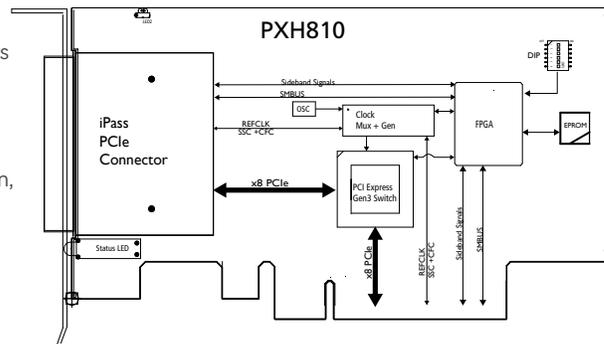
## PXH812 Gen 3 Host/Target Adapter

For customers developing I/O expansion applications with iPass® cables, the PXH812 transparent host/target adapter provides a fully compliant solution for OEMs and end users. The PXH812 is ideal in applications such as test and measurement equipment, medical equipment, and storage subsystem seeking high performance and data quality. The PXH812 adapter supports both host upstream and target downstream operations.

PXH812's signal quality is excellent. By isolating the system clock and transmitting an extremely low jitter high quality clock to downstream devices, the PXH812 offers users high signal quality and increased cable distances. It supports copper cables up to 5 meters at full PCI Express Gen3 speed when connected to another PXH812 card. Fiber optics is supported up to 100 meters at full

speed. The PXH812 can be connected to the IXS600 Switch to create PCIe Gen3 scale out solutions. Combined with a single IXS600 switch, you can connect up to 7 cabled downstream devices or expansion chassis to a single host system. The PXH812 card is a transparent only product. To utilize Dolphin's PCIe clustering software, you need to purchase the PXH810 Gen3 Adapter card.

The PXH812 adapter implements an iPass connector and is carefully designed to maximize cable lengths. With the implementation of clock isolation, internal clock generation, and a well-designed PCB, the



## Features

- PCI Express Gen3 compliant - 8.0 GT/s per lane
- Link compliant with Gen1, Gen2, and Gen3 PCIe
- Host and target operation
- PCI Express iPass cables
- One x8 Gen3 PCI Express cable port that can be configured as:
  - Copper and fiber-optic cable connectors
- Full host clock isolation. Supports hosts running both CFC and SSC
- Low Profile, half length PCIe form factor
- EEPROM for custom system configuration
- Link status LEDs through face plate



## I/O Expansion

The PXH812 adapter is capable of being configured as a host or a target adapter. Figure 17 illustrates a switch configurations supported by the PXH812. It can support multiple PCIe expansion units with an external switch. As a host adapter, the PXH812 connect to PCIe targets via iPass cables. It can connect to other PXH812 boards as a target or a custom target adapters. The PXH812 supports a x8 connection.

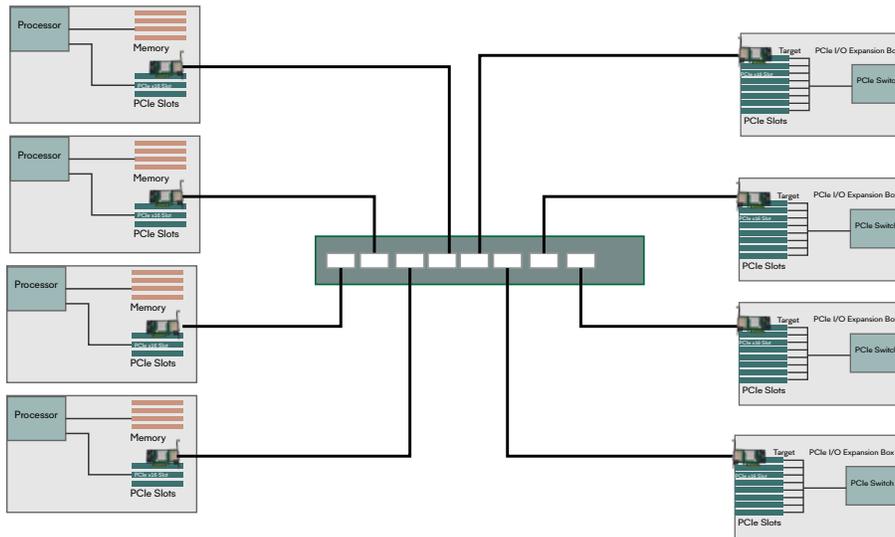


Figure 17: I/O Expansion with PXH812

## Specifications

Link Speeds	64 GT/s	Regulatory	CE Mark FCC Class A UL94V-0 compliant RoHS Compliant
Cut through latency	138 nanoseconds	EMC	EN 55022 (2010) EN 55024 (1998) + A1 (2001) + A2 (2003) EN 61000-6-2 (2005)
Active Components	Broadcom (PLX/Avago) PCIe Switch	Configuration	Dip-switch configurable host / target
PCI Express	Base Specification 3	Mounting Plates	Full height plate installed Half height plate included with board
Topologies	Switch/ point to point	Operating Systems supported	All, no driver or OS support required
Cable Connections	x8 iPass copper cable, fiber optic cable support Supports x8 to x4 transition cables	Product Codes	PXH812 Host/ Target Adapter
Power Consumption	3.3 Volt 12 watts		
Mechanical Dimensions	Low profile - 68.90 mm ( 2.731 inches) x 120 mm (6.600 inches)		
Operating Environment	Operating Temperature: 0°C -55°C Relative Humidity: 5% -95% non-condensing		
Usage Modes	Transparent Host/Target		



# PCIe Hardware



## IXH610 Gen2 PCIe Host Adapters

The IXH610 Gen2 PCIe Host Adapter is a high performance cabled interface to external processor subsystems or I/O subsystems. Based on Gen2 PCI Express bridging architecture, the IXH610 host adapter includes advanced features such as non-transparent bridging (NTB) and clock isolation.

For high performance application developers, the IXH610 host adapter combines 40 GT/s performance with less than one microsecond latency, significantly improving overall inter-system communication. Connecting remote I/O subsystems in transparent mode requires no special drivers, so deployment is fast and easy. Inter-processor communication also benefits from the high throughput and low latency.

The IXH610 performs both Direct Memory Access (DMA) and Programmed IO (PIO)

transfers, effectively supporting both large and small data packets. DMA transfers result in efficient larger packet transfers and processor off-load. PIO transfers optimize small packet transfers at the lowest latency. The combination of DMA and PIO creates a highly potent data transfer system.

Dolphin's eXpressWare™ software takes advantage of this data transfer scheme. Delivering a complete deployment environment for customized and standardized applications. The Shared-Memory Cluster Interconnect (SISCI) API is a robust and powerful shared memory programming environment. The optimized TCP/IP driver and SuperSockets™ software remove traditional networking bottlenecks. IP and sockets applications take advantage of the high performance PCI Express interconnect without modification. The overall framework

is designed to meet all the demands for rapid development of inter-processor communication systems.

With the implementation of clock isolation, the IXH610's signal quality is excellent. By isolating the system clock and transmitting an extremely low jitter high quality clock to downstream devices, the IXH610 offers users high signal quality and increased cable distances. Signal quality is essential for applications such as test and measurement equipment, medical equipment, and storage subsystem seeking high performance and data quality.

This card includes a Dolphin eXpressWare™ license.

## Features

- PCI Express 2.1 compliant - 5.0 GT/s per lane
- x8 PCI Express port - 40 GT/s
- Link compliant with Gen1 and Gen2 PCI Express
- Support Gen1, Gen2, and Gen3 PCIe Slots
- RDMA support through PIO and DMA
- PCI Express External Cabling Specification
- PCI Express x8 iPass Connectors
- Copper cables - up to 7 meters
- Clock isolation support
- Transparent bridging to cabled I/O devices
- Non-transparent bridging to cabled PCIe systems
- Low Profile PCIe form factor
- EEPROM for custom system configuration
- Link and status LEDs through face plate



## Cluster connections

When used for clustered connections, the IXH610 adapter is capable of node to node connections or connections through a IXS600 Switch as shown in figure 18. Adding industrial systems is done by connecting to the IXH620 XMC adapter. Each connection supports 40 GT/s with latencies as low as 0.74 microseconds. Designed for x8 PCIe Systems, the IXH610 supports any system with a standard x8 or x16 PCIe slot. The IXH631 uses MiniSAS HD connectors to create up to a 5 node cluster without a switch.

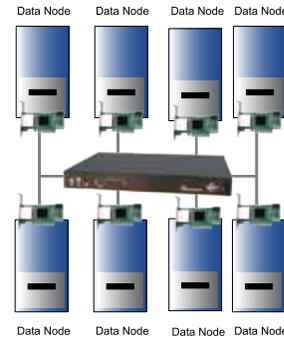


Figure 18: Eight node PCI Express Cluster

## Remote I/O Connections

The IXH610 functions as a high quality transparent connection to remote PCIe I/O subsystems. These subsystems include test equipment, I/O expansion systems, specialized equipment, and storage systems. The IXH610 is specially designed for higher signal quality and support for spread spectrum clocking. The IXH611 is used as a target adapter in I/O expansion applications.



Figure 19: I/O expansion with PCI Express

## Specifications

Link Speeds	40 GT/s	User Configuration Modes	Transparent Host Target adapter Non-transparent Host(NTB)
Application Performance	0.74 ms latency (application to application) 3,500 MB/s Throughput	Regulatory	CE Mark EN 55022, EN 55024-A1&A2, EN 61000-6-2 FCC Class A UL94V-0 compliant RoHS Compliant
PCI Express	Base Specification 2.1	Operating Systems supported	Windows Linux VxWorks RTX
Topologies	Point to point, Switched	Product Codes	IXH610 - Host / NTB Adapter IXH611 - Host/Target Adapter IXH614 - 104Mhz Overclocked Host Adapter IXH631 MiniSAS HD Host Adapter
Cable Connections	x8 iPass copper cable support Supports x8 to x4 transition cables		
Power Consumption	7 watts		
Mechanical Dimensions	PCI Express Card Electromechanical Specification 2.0		
Operating Environment	Operating Temperature: -10°C -60°C Relative Humidity: 5% -95% non-condensing		
Dolphin Software	SuperSockets Berkeley Sockets API Microsoft WinSock2/LSP support SISCI API IPoPCle		



# PCIe Hardware



## IXH620 Gen2 PCIe XMC Host Adapter

As performance needs increase in embedded systems, high performance embedded system designers are implementing PCI Express for inter-system communication. The IXH620 XMC cable adapter enables high speed PCI Express inter-system connections to external systems including servers, single board computers, and I/O subsystems.

The IXH620's 40 GT/s throughput and sub one microsecond latency deliver superior performance and low latency to systems supporting standard XMC slots or XMC VPX, VME or cPCI carrier boards. The IXH620's x8 cable connection uses an iPass™ connector as a standardized connection method. The XMC adapter supports either upstream or downstream target configurations. To connect remote I/O or processing, the IXH620 implemented

IDT's transparent or non-transparent bridging (NTB) functions. Used in transparent mode, standard PCI Express devices and drivers require no modifications. In NTB mode, the adapter facilitates inter-processor communication through Programmed IO (PIO) or Remote Direct Memory Access (RDMA).

The IXH620 implements a system clock isolation system for excellent signal quality. By isolating the system clock and transmitting an extremely low jitter high quality clock to downstream devices, IXH620 users benefit from improved signal quality, reliability, and cable distances.

The IXH620 comes with Dolphin's eXpressWare™ software suite that reduces time to market for customer applications. The Shared-Memory Cluster Interconnect (SISCI)

API is a robust and powerful programming environment for easy development of shared memory applications. Shared memory applications benefit from the 0.74 microsecond inter-system latency and more than 3500 Megabytes/s throughput. The optimized TCP/IP driver and SuperSockets software remove traditional networking bottlenecks. IP and Sockets applications can take advantage of the high performance PCI Express interconnect. Sockets applications experience 1.25 microsecond latency and 23 Gigabit/second user payload throughput.

These powerful features make XMC adapter an ideal interconnect for applications such as military and industrial systems that seek high performance and flexibility.

This card includes a Dolphin eXpressWare™ license.

## Features

- PCI Express 2.1 compliant - 5.0 GT/s per lane
- x8 PCIe port - 40 GT/s
- Link compliant with Gen1 PCI Express
- VITA 42.0-2005, ANSI/VITA 42.3-2006 compliant
- RDMA support through PIO and DMA
- PCI Express External Cabling Specification
- PCIe x8 iPass Connectors
- Copper cable connections up to 7 meters copper connections
- Clock isolation support
- Transparent bridging to cabled I/O devices
- Non-transparent bridging to cabled PCI Express systems
- Short XMC form factor
- EEPROM for custom system configuration
- XMC P15 connector
- Link and status LEDs through front panel



## Host to host connections

The IXH620 connects single board computers or systems running a distributed processing environment. Industrial or military customers requiring redundancy or increased compute resources use the IXH620 by itself or in conjunction with the IXS600 switch. Figure 20 illustrates connecting two single board computers with the IXH620. Fast data transfers are enabled through Dolphin's shared memory mechanism.

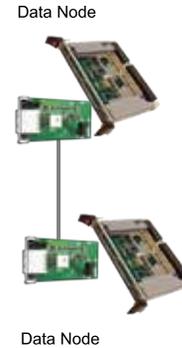


Figure 20: two node connection with XMC adapter

## Remote I/O subsystems

To extend the I/O capabilities of an XMC enabled system, the IXH620 supports host adapter or target adapter configurations. Industrial and Military customers requiring increased I/O bandwidth for graphics, processing, or data collection can add IXH620 enabled XMC carrier cards and chassis. Figure 21 illustrates connecting a standard server to an XMC enabled chassis to attach additional XMC cards for test system or increased functionality.



Figure 21: connect embedded equipment to hosts

## Specifications

Link Speeds	40 GT/s	User Configuration Modes	Transparent/non-transparent(NTB)
Application Performance	0.74 microsecond latency (application to application)	Regulatory	CE Mark EN 55022,EN 55024-A1&A2, EN 61000-6-2 FCC Class A UL94V-0 compliant RoHS Compliant
Specifications	PCI Express Base Specification 2.1 VITA 42.0-2005, ANSI/VITA 42.3-2006	Operating Systems supported	Windows Linux VxWorks RTX
Topologies	Point to point, Switched	Product Codes	IXH620 Host /Target / NTB adapter
Cable Connections	x8 iPass copper cable support Supports x8 to x4 transition cables		
Power Consumption	7 watts		
Mechanical Dimensions	XMC Short form factor		
Operating Environment	Operating Temperature: 0°C -55°C Relative Humidity: 5% -95% non-condensing		
Dolphin Software	SuperSockets Berkeley Sockets API Microsoft WinSock2/LSP support SISCI API IPoPCle		



# PCIe Hardware



## IXS600 Gen3 PCIe Switch

PCI Express provides low latency, highly efficient switching for high performance applications. The IXS600 Gen3 PCI Express switch delivers a powerful, flexible, Gen3 switching solution. This powerful switch enables I/O scaling and inter-processor communication by combining transparent and non-transparent bridging capabilities with Dolphin's software and clustering technology. IXS600 users can connect multiple PCI Express devices or create a highly efficient compute cluster with PCs, servers, or SBCs with XMC sites.

The IXS600 is the switching element of Dolphin's product line. This eight port, 1U cluster switch delivers 64 GT/s of non-blocking bandwidth per port at ultra-low latencies. Each x8 PCI Express port delivers maximum bandwidth to each device while maintaining backwards compatibility with Gen1 and Gen2 components. As

with other Dolphin products, the IXS600 utilizes standard iPassconnectors to string components via copper or fiber-optic cabling. IXS600 customer can link multiple standardized PCI Express products such as PXI chassis, storage, and I/O expansion units.

For Non Transparent Bridging (NTB) or clustering applications, the IXS600 integrates with Dolphin's PCI Express Host Adapters or XMC Adapter. The total NTB solution solves many of the problems related to PCI Express operations, such as power sequencing and standard software. Dolphin customers avoid the severe power-on requirements normally associated with PCI Express. Hosts, cables, and the switch can be hot-swapped and power cycled in any sequence for real plug and play.

The IXS600 switch can also be partitioned into several virtually independent partitions,

e.g. mixing NTB and Transparent functionality on separate ports.

The IXS600 switch supports Dolphin's comprehensive software suite. The Shared-Memory Cluster Interconnect (SISCI) API is a robust and powerful programming environment for easy development of shared memory applications. The optimized TCP/IP driver and SuperSockets software remove traditional networking bottlenecks. IP and Sockets applications can take advantage of the high performance PCI Express interconnect. This comprehensive solution is ideal for real-time, technical and high performance computing, cloud computing, and enterprise business applications.

The IXS600 switch is compatible with all Dolphin Gen2 and Gen3 iPass based PCI Express products.

## Features

- PCI Express 3.0 compliant -8.0 GT/s per lane
- Eight PCI Express Gen3 x8 ports
- PCI Express x8 iPass Connectors
- Auto-training to lower lane widths
- Supports x4 lanes with a transition cable
- Link compliant with Gen1 and Gen2 PCI Express
- Transparent and Non Transparent support
- PCI Express External Cabling Specification
- Fiber-optic and copper cable support
- Hot Plug PCI Express cabling support in NTB mode
- Built in management processor
- Boot configuration and user data support for cluster and system configuration
- 19 Inch 1U rack mountable chassis
- Front and rear operational status and alert LEDs
- Redundant Fans



## Cluster connections

The IXS600 is a key element in Dolphin's inter-processor connectivity strategy. Industrial, military, or enterprise customer can create diverse multiprocessing configurations. The hybrid configuration illustrated shows single board computers and data nodes connected through the switch. The IXH610 host adapter and IXH620 XMC adapter are used to connect to different compute nodes.

Increasing the number of I/O components in a system is accomplished by using the IXS600 with PCIe I/O expansion boxes. Figure 22 illustrates the IXS600 connecting 7 additional I/O expansion boxes. These boxes can accommodate components such as sensors, graphics, coprocessors, video capture cards, and other I/O devices.

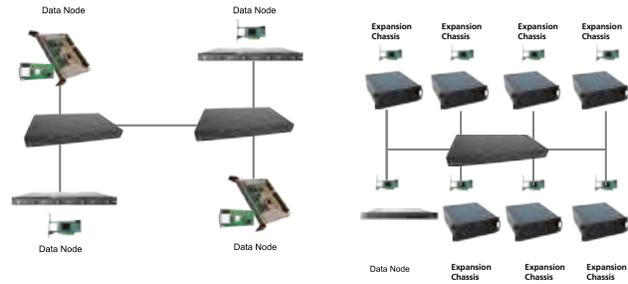


Figure 22: I/O and hybrid configuration

## Scalability

Scalability is achieved by connecting multiple switches. Multiple IXS600 switches are used for larger reflective memory applications or creating larger processing clusters.

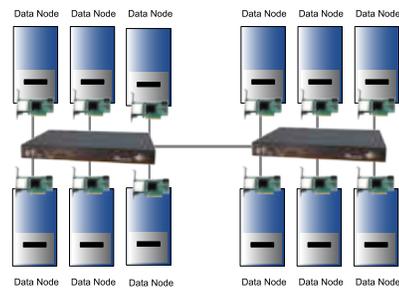


Figure 23: Scalability of switches

## Specifications

Link Speeds	64 GT/s per port	User Configuration Modes	Transparent/non-transparent(NTB)
Application Performance	6595 MB/s application data rate 200ns port latency	Operating Environment	Operating Temperature: 0°C -55°C Relative Humidity: 5% -95% non-condensing
Ports	8 - x8 non-blocking 64 GT/s ports	Regulatory	CE Mark EN 55022,EN 55024-A1&A2, EN 61000-6-2 FCC Class A UL94V-0 compliant RoHS
Cable connections	x8 iPass copper cables	Product Codes	IXS600 Gen3 100 Mhz
Management	Ethernet management port Operational status and alert LEDs		
Power	Auto-sensing power supply 110 - 240 V AC 50-60Hz Power consumption : Max 80 Watts		
Mechanical	1U, 19 inch rackmountable chassis 440mm (W) x 300mm (D) x 45mm( H) Redundant Fans		



## Hardware Comparison - PCIe Gen 3 with SFF-8644 Connector

Feature	PXH830	PXH832	MXH830	MXH832
Use	NTB	Transparent Host/ Target	NTB	Transparent Host/ Target
Form factor	Short/ Low profile PCIe	Short/ Low profile PCIe	Short/ Low profile PCIe	Short/ Low profile PCIe
PCIe max connection speed	x16 Gen3	x16 Gen3	x16 Gen3	x16 Gen3
NTB	Yes	No	Yes	No
Transparent Target	No	Yes	No	Yes
Transparent Host	No	Yes	No	Yes
Clocking	100Mhz	100MHz	100Mhz	100MHz
Max copper cable length	9 Meter	9 Meter	9 Meter	9 Meter
Fiber support	100 Meter	100 Meter	100 Meter	100 Meter
Switch support	MXS824	MXS824	MXS824	MXS824
Scalability	3 nodes without switch 64/128 nodes with switch	1-4 expansion without switch	5 nodes without switch 64/128 nodes with switch	1-4 expansion without switch
Number of reflective memory segments	4(16)	N/A	4 (16)	N/A
Max PIO Performance	10.2 GB/s	N/A		N/A
Max DMA Performance	11 GB/s	N/A		N/A
Scipp latency back to back	0.54 $\mu$ s	N/A		N/A
Scipp latency switch	0.71 $\mu$ s	N/A		N/A



## Hardware Comparison - PCIe Gen 2/3 with iPass Connector

Feature	PXH810	PXH812	IXH610/IXH611	IXH620
Use	NTB	Transparent Host/ Target	NTB/ Transparent Host/ Target	NTB/ Transparent Host/Target
Form factor	Short/ Low profile PCIe	Short/ Low profile PCIe	Short/ Low profile PCIe	Short XMC
PCIe max connection speed	x8 Gen3	x8 Gen3	x8 Gen2	x8 Gen2
NTB	Yes	No	Yes	Yes
Transparent Target	No	Yes	IXH610-No IXH611-Yes	Yes
Transparent Host	No	Yes	Yes	Yes
Clocking	100Mhz	100MHz	100MHz	100MHz
Max copper cable length	5 Meter	5 Meter	7 Meter	7 Meter
Fiber support	100 Meter	100 Meter	No	No
Switch support	IXS600	IXS600	IXS600	IXS600
Scalability NTB, general use	14 nodes	N/A	20 nodes	20 nodes
Scalability Reflective memory functionality only	14 nodes	N/A	56 nodes	56 nodes
Transparent scalability	256 bus numbers	256 bus numbers	256 bus numbers	256 bus numbers
Maximum reflective memory segment size	2 GB	N/A	2 GB	2 GB
Number of reflective memory segments	4 (16)	N/A	4	4
Total reflective memory size	4 x 2 GB	N/A	4 x 2 GB	4 x 2 GB
Max PIO Performance	5312 MB/s	N/A	2950 MB/s	2950 MB/s
Max DMA Performance	6695 MB/s	N/A	3500 MB/s	3500 MB/s
Scipp latency back to back	0.54 $\mu$ s	N/A	0.74 $\mu$ s	0.74 $\mu$ s
Scipp latency switch	0.71 $\mu$ s	N/A	0.9 $\mu$ s	0.9 $\mu$ s



# Product Comparison Charts

## Software Comparison

Feature	SISCI	SuperSockets	IPoPCle Driver
Linux Platform Support	x86, x86_64 ARM 32/64. PPC NVidia Tegra	x86, x86_64 ARM 64 PPC	x86, x86_64 ARM 64 PPC
Kernel 2.6 - 4.x			
Windows Vista, 7, 10, Server 2008-2016 Platform Support	x86, x86_64	x86, x86_64	x86, x86_64
RTX Platform Support	x86, x86_64	No	No
VxWorks 6.9 , 7.0	x86, x86_64 PPC	No	No
Latency	0.54 $\mu$ s Linux	0.98 $\mu$ s Linux	5.6 $\mu$ s Linux
Max Bandwidth	11 GB/s Linux	65 Gb/s	65 Gb/s
Reflective Memory Support	Yes	N/A	N/A
Address based Multi-cast support	Yes	N/A	N/A
TCP Support	N/A	Yes	Yes
UDP Support	N/A	Linux -Yes Windows - No	Yes
UDP Multi-cast Support	N/A	Linux -Yes Windows - under development	Linux -under development Windows - Yes
Application modifications required	Yes	No	No
Cross O/S and Platform data transfer support	All	User space Linux / Windows	No
Peer to Peer Transfers support	Yes	N/A	N/A
Accelerated Loopback Support	Yes	Yes	No
Open Source code available	Yes	No	No
Licensing of software available	Yes	Yes	Yes
Chipset support	Microsemi Broadcom/PLX IDT Intel NTB	Microsemi Broadcom/PLX IDT Intel NTB	Microsemi Broadcom/PLX IDT Intel NTB



Dolphin Interconnect Solutions has been a global provider of ultra-low latency, high-bandwidth computer interconnect solutions for high speed real-time systems, clustered databases, general networking, web services and industrial applications for more than 25 years. For more information, please visit [www.dolphinics.com](http://www.dolphinics.com).

Visit us on LinkedIn and Facebook:

- <https://www.linkedin.com/company/dolphin-interconnect-solutions>
- <https://www.facebook.com/dolphinInterconnect>



©2017 Dolphin Interconnect Solutions AS. All rights reserved. Dolphin ICS and the Dolphin Logo are registered trademarks. PCI, PCI Express, PCIe, Express Module, and PCI-SIG are trademarks or registered trademarks of PCI-SIG. All other trademarks are the property of their respective owners. All statements herein are based on normal operating conditions, are provided for information purposes only and are not intended to create any implied warranty of merchantability or fitness for a particular purpose. Dolphin Interconnect AS reserves the right to modify at any time without prior notice these statements, our products, their performance specifications, availability, price, and warranty and post-warranty programs.

