PXIe x8 NTB Cable Module MXP908

MXP908 - PXIe PCIe x8 Gen4 NTB Peripheral Module Users Guide
Version 1.1

Date: 9th December 2023

Table of Contents

DISCLAIMER

DOLPHIN INTERCONNECT SOLUTIONS RESERVES THE RIGHT TO MAKE CHANGES WITHOUT FURTHER NOTICE TO ANY OF ITS PRODUCTS TO IMPROVE RELIABILITY, FUNCTION, OR DESIGN.

DO NOT USE INFORMATION IN THIS GUIDE TO DESIGN YOUR OWN PCI-SIG COMPLIANT PRODUCT, ALWAYS REFERENCE THE ORIGINAL PCI-SIG SPECIFICAITON FOR DETAILS.

TO THE FULLEST EXTENT PERMITTED BY LAW, DOLPHIN WILL NOT BE LIABLE FOR ANY INDIRECT, INCIDENTAL, SPECIAL OR CONSEQUENTIAL DAMAGES (INCLUDING LOST PROFITS, LOST DATA, OR LOSS OF USE) ARISING OUT OF ANY USE OF DOLPHIN'S PRODUCTS, SOFTWARE OR SERVICE PROVIDED. DOLPHIN'S MAXIMUM LIABILITY WILL NOT EXCEED THE TOTAL AMOUNT PAID FOR THE PRODUCT BY THE PURCHASER.

PCI Express External Cabling specification 4.0 is, as of the release of MXP908 PXIe module, not completed and ratified by the PCI-SIG. The MXP908 PXIe x8 cable module is designed to the new specification, but Dolphin cannot guarantee the card will be compliant to the final 3.0 version. Dolphin firmware tools can update the CMI implementation.

LIFE SUPPORT POLICY

DOLPHIN INTERCONNECT SOLUTIONS' PRODUCTS ARE NOT AUTHORIZED FOR USE AS CRITICAL COMPONENTS IN LIFE SUPPORT DEVICES.

ENVIRONMENTAL POLICY

Dolphin is minimizing the amount of printed documentation and software CDs in its shipments; please download additional documentation and software from www.dolphinics.com.

Scope

This version of the MXP908 PXIe module Users Guide is valid for BMC firmware version 2.4 and newer.

Terms and Acronyms

Important terms and acronyms used in this manual

AOC	**A**ctive **O**ptical **C**able. PCIe fiber cable assembly available from Dolphin.
CMI	**C**able **M**anagement **I**nterface. The 2-wire management interface for communication between subsystems connected by a PCIe 4.0 cable. Details can be found in the PCI-SIG External Cabling Specification 4.0 (draft).
BMC	Board Management Controller - Microcontroller on the MXP908 PXIe module used to manage and implement the CMI communications.
eXpressWare	Dolphin's software stack for PCIe clustering and IO. Please visit www.dolphinics.com/software for more information.
Lane	One PCI Express Lane contains a differential pair for transmission and a differential pair for reception.
Link	A collection of one or more PCI Express Lanes providing the communication path between an Upstream and Downstream Port.
PCIe 4.0 cable	Cable compliant to the preliminary PCI-SIG External Cabling Specification 4.0. Support for CMI.
Port	PCIe Cable port. The MXP908 PXIe module has four x4 ports, named P1, P2, P3, P4. Ports P1 and P2 can be merged to one x8 port. Ports P3 and P4 are unused. The physical ports are identified by text on the PCIe brackets.
Wake	A mechanism used by a downstream device to request the reapplication of PC main power when in the L2 Link state / Power mode.

MXP908 PXIe Module High Level Specification

The MXP908 is a PXIe peripheral NTB PCIe x8 Gen4 cable module and supports connectivity to other Dolphin Microchip PCIe switch based NTB products like the MXH918, MXH930 and the MXS924 PCIe switch. The MXP908 supports PCIe Gen1, Gen2, Gen3 and Gen4 speeds and x1, x2, x4 and x8 link-widths. The module will operate at the highest common speed shared between the backplane slot and the card (Gen4) and the widest common cable link-width up to x8.

- PCI Express Base Specification, Rev. 4.0
- PCI Express External Cabling specification 4.0 (draft)
- PCI Express Gen4 16.0 GT/s per lane signaling – 128 GT/s total signaling.
- Compliant with PCI Express Gen1 through Gen4 IO systems, auto detection
- The MXP908 supports NTB with DMA
- Quad SFF-8644 cable connector
  - Two upper ports enabled for PCIe connectivity
  - Two lower ports reserved for future use
  - Durability max total 250 mating cycles
- Support for two PCIe 4.0 SFF-8644 cables.
- Microchip Switchtec PM40028 PFX PCI Express Gen4 chipset with DMA
- 100 nanosecond cut-through latency port to port.
- Automatic Thermal Shutdown to avoid hardware damages (can be disabled)
- Support for active optical PCIe complaint fibers (AOC)
- Single Peripheral Slot 3U PXI Express, compliant to PXI -5 PXI Express Hardware Specification rev 1.1
- SRNS clocking, Prepared for SRIS clocking
- VAUX powered board management controllers for flexible configuration and cable management
- Flash recovery option. PFX Multi configuration support
- Relative Humidity: 5% - 95% (non- condensing)
- CE and FCC markings
- Made in Norway

Packaging

The MXP908 product is delivered in a retail box includes the following components.
- MXP908 module
- Anti-static bag
- Getting started guide

PXIe Slot

The MXP908 module can be placed in any compliant PXIe slots. The MXP908 supports PCIe Gen1, Gen2, Gen3 and Gen4 speeds. The default PXIe backplane configuration is x8. The module will auto configure to the slot speed and width.

MXP908 Module Configuration

The MXP908 cable ports can either be used as a single x4 / x8 or dual x4 port configuration.
The default DIP switch setting is NTB single x4 or single x8 operations. The lower two connectors marked P3 an P4 is not in use.

Thermal Shutdown

The MXP908 module includes an overtemperature overheat protection. The BMC will automatically shut down the MXP908 module if a PCIe switch temperature beyond 105°C (221°F) is detected. Similar, if an AOC is plugged in, the MXP990 module will automatically shut down if the max operating temperature announced by the connected AOC is violated by 15°C (59°F). Please consult the AOC data sheet for details. Please consult the AOC data sheet for details.

NOTE: The MXP908 module link LEDs will start flashing yellow when a PCIe switch temperature of 95°C (203°F) is reached or when the AOC reports a temperature 5°C (41°F) lower than the max operating temperature.

NOTE: The automatic shutdown can be disabled by setting DIP switch SW2-4. Prior to disabling the automatic shutdown, please ensure you have established proper airflow.

Airflow - Operating Environment

To maximize lifetime for the product and maintain the warranty, please honor the specified operating temperature, and make sure the specified air flow is present.

TIP: After installing the Dolphin eXpressWare board management software, you can use the tool dis_diag to determine the actual board temperatures.

Cable Connections

The MXP908 module is designed to support both long and short PCIe copper cables as well as PCIe active optical cables (AOC).

The MXP924 cable ports are compliant to the SFF-8644 industry specification and supports PCIe cables compliant to the PCIe External Cabling Specification 4.0. Two x4 cables are needed for full PCIe x8 connectivity. (Ports P3 and P4 are unused)

PCI Express 4.0 Cables

When used with cables compliant to the new PCIe External Cable standard 4.0, the MXP908 module will transmit a CMI Reset message downstream. The module can be connected to a MXP924 module in Target mode or any PCIe device compliant to the new cable standard.

Active Optical PCIe Cables (AOC)

Dolphin offers active optical PCIe fiber cables up to 100 meters for the MXP908. CWAKE and CPOWERON is not supported using standard PCIe AOC cables.

Installation

Step 1 - Unpack board

The MXP908 module module is shipped in an anti-static bag to prevent static electricity damage. The module should only be removed from the bag after ensuring that anti-static precautions are taken. Static electricity from your clothes or work environment can damage your MXP908 module or other PXIe components. Always wear a grounded anti-static wrist strap while the MXP908 module is removed from the anti-static bag until it is properly installed in the PXIe chassis.

Unpack the MXP908 module from the anti-static bag using proper anti-static procedures.

Step 2 - Configure the Board for Proper Operation

Set the DIP switch settings for proper operation depends on the firmware. Please refer to the section Configuration and DIP Switches on page 10 for details.

Step 3 - Ensure proper Airflow

Please pay proper attention to ensure the selected server / slot position provides the minimum required airflow.

TIP: Immediately after software installation, please use the dis_diag tool to verify the temperature. If the chip temperature reported by dis_diag exceeds 95°C, it is strongly recommended to improve the airflow.

Step 4 - Install the MXP908 module

Before installing the module, make sure you are properly grounded to avoid static discharges that may destroy your computer or the PXIe module. Ensure you are properly grounded before opening your PXIe chassis or the anti-static bag containing the MXP908 module. Please follow your PXIe chassis manual on how to install a PXIe System switch module.

The MXP908 PXIe module supports PCI Express Gen1, Gen2, Gen3 and Gen4 backplanes. Please review the list of supported and approved PXIe chassis

Step 5 - Installing and Removing the Cable

Installing and removing cables should be done with the PXIe chassis powered off.

Connecting the Cable

Please carefully install the cable connector into the connector housing on the MXP908 module. Cable port 1 is located at the top of the face plate bracket. To install the cable, match the cable house with the connector on the MXP908 module. Use even pressure to insert the connector until it is secure. Adhere to ESD guidelines when installing the cables to ensure you don't damage the board. Computer cables should always use strain relief to protect the connected equipment from excessive force on the cable. This is especially important for cables between racks.

Cable stain relief

Strain-relief on the PCIe cables are important to protect the PCIe card from fatigue loading and damages. The cable should be mechanically connected to the rack or cabinet. This is especially important for long heavy cables and cables going between cabinets and racks.

Supported cable lengths

The MXP908 module supports both copper and active optical PCIe cables (AOC). Maximum lengths between two modules can be found in Table 1 below. The maximum distance may change when connecting to other PCIe products.

Cable	Speed	Max Distance
Copper PCIe 4.0 cables	Gen4	3 meters (4-5 meters being qualified)
PCIe Active Optical cables (AOC)	Gen4	100 meters

Table 1: Cable Specifications

Disconnecting the Cable

Please carefully pull the release tab to release the cable from the locking latches and gently pull the cable out of the connector guides.

Step 7 - Verify Installation & LEDs

The MXP908 PXIe module comes with 4 RGB LEDs which show the corresponding cable port status according to Table 2: LED below. The LEDs are visible through cut-outs in the front panel below each port.

LED color	Function
Off	No cable installed, or PXIe not powered
Yellow	Cable installed, no link
Yellow blinking	Temperature overheat warning, please immediately improve cooling to avoid thermal damages or system shutdown.
White	Cable installed, link operational Gen1 speed
Blue	Cable installed, link operational Gen2 speed
Green blinking	Cable installed, link operational Gen3 speed
Green	Cable installed, link operational Gen4 speed
Red blink	Link Reset

Table 2: LED behavior

Operation

Configuration and DIP Switches

The MXP908 module has two banks of 8 DIP switches. The default factory setting for the MXP908 module is NTB Host mode, single (up to x8) link connection.

The MXP908 module has DIP switches for setting special modes or operations, the meaning of each DIP switch depends on the loaded firmware. Please carefully read the documentation shipped with the module before modifying any DIP switch settings. Please pay close attention to ON and OFF positions written on the DIP switch.

DIP Switches

Please leave all undocumented DIP switches in the default position. Table 3: DIP SW1 settings and Table 4: DIP-SW2 settings below shows all the various DIP switch settings for the MXP908 module.

Note: DIP switch configuration options may be changed in the future versions. Please always consult the latest user guide for details. The table above is valid for version 1.5.

DIP-Switch SW1 settings

SW1 DIP no.	Description	ON	OFF	Default
1-8	Configuration selector, details below			OFF

Table 3: DIP SW1 settings

DIP-Switch SW2 settings

SW2 DIP no.	Description	ON	OFF	Default
1	Reserved, leave in Default OFF			OFF
2	Reserved, leave in Default OFF			OFF
3	Reserved, leave in Default OFF			OFF
4	Disable thermal shutdown	Thermal shutdown disabled	Shut down the switch if critical temperature reached.	OFF
5-8	Reserved, leave in Default OFF			OFF

Table 4: DIP-SW2 settings

DIP switch Settings

The following DIP SW1 Switch settings should be considered when configuring the **MXP908 PXIe module for NTB** operation:

Config	Use Case	MXP908 SW1 DIP settings	DIP SW1 view – 256MB BAR	32GB BAR Option *All systems may not support this.*
0	**A + C**	**Single x8 port**	**Shipping Default, Opt 4 set**	**Opt 1 and 4 set**
1	**B**	**Dual x4 ports**	**Opt 2 and Opt 4 set**	**Opt 1, 2 and 4 set**

Use Cases

The MXP908 module may be used to connect a PXIe chassis to one or more PXIe chassis having a MXP908 installed or standard PCs having a MXH914, MXH918 or MXH930 installed. The DIP switch settings are summarized in section Configuration and DIP Switches on page 10. DIP-Switch settings for the MXP924 module can be found in the MXP924 Users Guide available from www.dolphinics.com/mx

Use Case A - Connecting two PXIe chassis – PCIe Gen 4.0 x8

Both PXIe chassis have a MXP908 module configured for NTB Host operation and a direct x4 or x8 link. Alternatively, one of the systems can be a Server having a MXH914, MXH918 or MXH930 NTB adapter installed

MXP908 module Host 1 ports	MXP908 module Host 2 ports
P1	P1
P2	P2

Table 5: Required x8 cabling

Always connect a cable from Port #x to Port #x
A failure connecting any of the cables will cause the link to re-train to x4.

Use Case B – Connecting tree PXIe chassis – PCIe Gen 4.0 x4

The PXIe chassis has MXP908 modules configured for NTB Host operation.

PXIe Chassis 1 MXP908 module ports	MXP908 Switch Target 1 ports	MXP908 Switch Target 2 ports
P1	P1	
P2		P1
	P2	P2

Table 6: Required x4 cabling

Use Case C – MXS924 Switch Configurations – NTB Mode

The MXP908 module can be combined with the MXS924 24 port PCIe switch to create larger PCIe networks. Each node typically has a Dolphin MXH9xx or MXP908 NTB adapters and a x4, x8 cable connection to the MXS924 switch. All adapter cards operate in NTB mode. Please consult the MXS924 Users Guide for more details to configure the MXS924. Dip switches should be set as specified in Table 3: DIP SW1 settings on page 10. Some configuration examples can be found below, please contact Dolphin for more information how to build larger systems up to 60 nodes.

12 Node, x8 Link Width using MXS924

In this use case, each NTB module/ card is connected to the switch using two cables to create a x8 Gen4 link. The aggregated link speed is 128 GT/s. The graphics in Figure 1 below shows the detailed cabling for connecting the card to the MXS924 switch. The MXS924 must be re-configured through the Ethernet interface to support this setup. Please consult the MXS924 Users Guide for details.
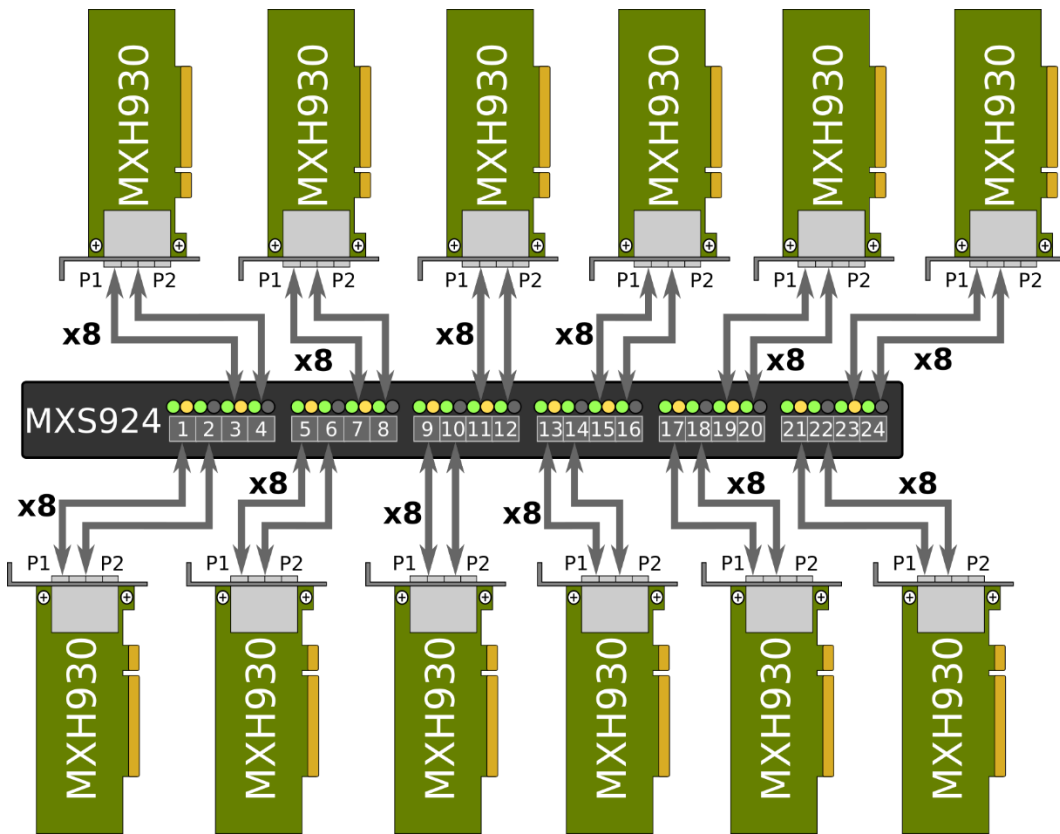

Figure 1: A 12 node PCIe x8 network using MXS924

Larger PCIe networks using the MXS924

Please consult the MXS924 Users Guide or contact Dolphin for information on larger topologies.

Firmware Upgrade

The MXP908 module design uses a Board Management Controller (BMC) to implement reset and other maintenance functions. Dolphin may from time to time publish updated firmware for the BMC, card configuration and the PCIe switch.

WARNING: Please note that standard Microsemi tools (ChipLink) cannot be used to upgrade the firmware as this will violate the warranty. Please contact Dolphin for instructions on how to upgrade the MXP908 module firmware.

Note: Please consult the MXP908 module Firmware Release Note for details on supported configurations.

Identifying the PXIe module

The module has a label-sticker with the serial number in the format 'MXP908-YY-ZZZZZZ', where YY denotes the revision (e.g. CC) and ZZZZZZ denotes the serialized production number (e.g. 012345) – this whole string makes up the serial number of the module (i.e. MXP908 module-CC-012345).

You can also get this information using lspci in Linux:

First, identify the devices for the Dolphin MXH908 card:

```
# lspci | grep "Device 4028"

01:00.0 PCI bridge: PMC-Sierra Inc. Device 4028
01:00.1 Bridge: PMC-Sierra Inc. Device 4028
02:00.0 PCI bridge: PMC-Sierra Inc. Device 4028
```

Then run lspci and identify the module. It will show up as something like

```
# lspci -s 1:0.0 -v
01:00.0 PCI bridge: PMC-Sierra Inc. Device 4028 (prog-if 00 [Normal decode])
    Flags: bus master, fast devsel, latency 0, IRQ 122
    Bus: primary=01, secondary=02, subordinate=03, sec-latency=0
    I/O behind bridge: 00002000-00002fff
    Prefetchable memory behind bridge: 00000000c8000000-00000000c81fffff
    Capabilities: [40] Express Upstream Port, MSI 00
    Capabilities: [7c] MSI: Enable+ Count=1/8 Maskable- 64bit+
    Capabilities: [8c] Power Management version 3
    Capabilities: [94] Subsystem: Dolphin Interconnect Solutions AS Device 0924
    Capabilities: [100] Advanced Error Reporting
    Capabilities: [148] Power Budgeting <?>
    Capabilities: [158] #12
    Capabilities: [188] #19
    Capabilities: [1b4] Device Serial Number 00-00-43-43-00-00-00-24
    Capabilities: [1c0] Latency Tolerance Reporting
    Capabilities: [1c8] Access Control Services
    Capabilities: [1f0] #25
    Capabilities: [1fc] #26
    Capabilities: [23c] #27
    Capabilities: [7f8] Vendor Specific Information: ID=ffff Rev=1 Len=808 <?>
    Kernel driver in use: pcieport
    Kernel modules: shpchp
```

Second, do

```
# lspci -s 1:0.0 -v | grep -E "Subsystem|Serial"
        Capabilities: [a4] Subsystem: Dolphin Interconnect Solutions AS Device 0908
        Capabilities: [100] Device Serial Number 00-00-43-43-00-00-00-ff
```

This shows the module as revision 0x4242 (hexadecimal values of the 'CC' letters in the ASCII table), with the production number 0x000000ff (00000255 in decimal).

Support

More information about the product, support and software downloads are available at http://www.dolphinics.com/mx.

Please file a support ticket at www.dolphinics.com/csp if you have any questions.

Technical Information

Board revision history

The following table gives a general overview of the hardware revision history.

MXP908 module revision	Capabilities
MXP908-BB	• Initial version for testing. Not available
MXP908-CD	• Production version

PCIe Cable Port Mapping

The MXP908 module has a quad SFF-8644 connector. Only P1 and P2 are connected.

Compliance and Regulatory Testing

EMC Compliance

The MXP908 module is tested to comply with the following standards:

- EN 61326-1:2013
- EN 61000-6-1 :2007
- KS C IEC 61326-1:2008
- KS C IEC 61000-6-1:2002
- 47 CFR Part 15. Subpart B (Clause 15.107 and 15.109) in conjunction with ANSI C63.4:2014

This does not ensure that it will comply with these standards in any random PXIe chassis. It is the responsibility of the integrator to ensure that their products are compliant with all regulations where their product will be used.

FCC Class A

This equipment is tested to comply with the limits for a Class A digital device, pursuant to part 15 of the FCC Rules.

RoHS Compliance

The MXP908 module is RoHS compliant. A Compliance certificate issued by the manufacturer is available upon request.

WEEE Notice

The MXP908 module is labelled in accordance with European Directive 2002/96/EC concerning waste electrical and electronic equipment (WEEE). The Directive determines the framework for the return and recycling of used appliances as applicable throughout the European Union. This label is applied to products to indicate that the product is not to be thrown away but returned to your local approved WEEE waste collector.

Limited Warranty

Dolphin Interconnect Solutions warrants this product to be free from manufacturing defects under the following terms:

Warranty Period

Dolphin warrants the product for one (1) year from the date of purchase. Extended warranties are available for purchase.

Coverage

To the extent permitted by applicable law, this warranty does not apply to:

- Damages caused by operator error or non-compliance with instructions available for the product.
- Use or attempt to use or program firmware not approved by Dolphin.
- Damage which results from accident, abuse, misuse, neglected improper handling or improper installation; moisture, corrosive environments, missing cable strain-relive, high voltage surges, shipping or abnormal working conditions.
- Damages which result from violating the specified operating or storage temperatures and airflow.
- Damages caused by acts of nature, e.g. floods, storms, fire, or earthquakes.
- Damage caused by any power source out of range or not provided with the product.
- Normal wear and tear.
- Attempts to repair, modify, open or upgrade the product by personnel or agents not authorized by Dolphin.
- Products that have had the product serial number tampered with or removed.
- Damage to the product caused by products not supplied by Dolphin.

Service Procedure

If the product proves defective during the Warranty Period, you should contact the seller that supplied you with the product, or if you purchased it directly from Dolphin, email returnrequests@dolphinics.com to obtain a valid RMA number and instructions. Products returned to Dolphin without a proper RMA number will not be serviced under this warranty.