



Huawei MZ610 NIC

V100R001

White Paper

Issue 01

Date 2015-05-30

Copyright © Huawei Technologies Co., Ltd. 2015. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

Trademarks and Permissions



HUAWEI and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: <http://enterprise.huawei.com>

About This Document

Purpose

This document describes the MZ610 in terms of its functions, appearance, features, applications, and technical specifications. You can obtain comprehensive information about the MZ610 by reading this document.





Intended Audience

This document is intended for:

- Huawei presales engineers
- Channel partner presales engineers
- Enterprise presales engineers

Symbol Conventions

The symbols that may be found in this document are defined as follows.

Symbol	Description
 DANGER	Alerts you to a high risk hazard that could, if not avoided, result in serious injury or death.
 WARNING	Alerts you to a medium or low risk hazard that could, if not avoided, result in moderate or minor injury.
 CAUTION	Alerts you to a potentially hazardous situation that could, if not avoided, result in equipment damage, data loss, performance deterioration, or unanticipated results.
 NOTE	Provides additional information to emphasize or supplement important points in the main text.

Change History

Issue 01 (2015-05-30)

This issue is the first official release.

Contents

About This Document.....	ii
1 Overview.....	1
1.1 Functions.....	2
1.2 Appearance.....	2
2 Features.....	4
2.1 Feature List.....	5
2.2 Feature Description.....	5
2.3 Standards Compliance.....	5
3 Applications.....	7
3.1 Compatible Compute Nodes.....	8
3.2 Connected I/O Modules.....	8
3.3 MZ610 Networking.....	10
3.4 Supported OSs.....	10
4 Technical Specifications.....	12
4.1 Technical Specifications.....	13
A Acronyms and Abbreviations.....	14

1 Overview

[1.1 Functions](#)

[1.2 Appearance](#)

1.1 Functions

The MZ610 is an InfiniBand (IB) host channel adapter (HCA) that provides two 40G IB quad data rate (QDR) ports. It is used for E9000 compute nodes, and provides network ports for connecting to switch modules in the chassis.

The MZ610 uses the Mellanox ConnectX-3 (CX3) chip and supports HCA applications. The MZ610 supports the Remote Direct Memory Access (RDMA) feature to address low-latency network applications.

1.2 Appearance

The MZ610 can be installed in slot Mezzanine1 (Mezz1 for short) or Mezzanine2 (Mezz2 for short) on a half-width E9000 compute node (such as the CH121 and CH121 V3) or in slot Mezz1, Mezz2, Mezzanine3 (Mezz3 for short), or Mezzanine4 (Mezz4 for short) on a full-width E9000 compute node (such as the CH220, CH220 V3, CH221, CH222, CH222 V3, CH240, CH242, or CH242 V3).

The MZ610 provides network ports for connecting to switch modules:

- When the MZ610 is installed in slot Mezz1 or Mezz3, its two ports connect to switch modules in slots 2X and 3X.
- When the MZ610 is installed in slot Mezz2 or Mezz4, its two ports connect to switch modules in slots 1E and 4E.

Figure 1-1 MZ610 appearance

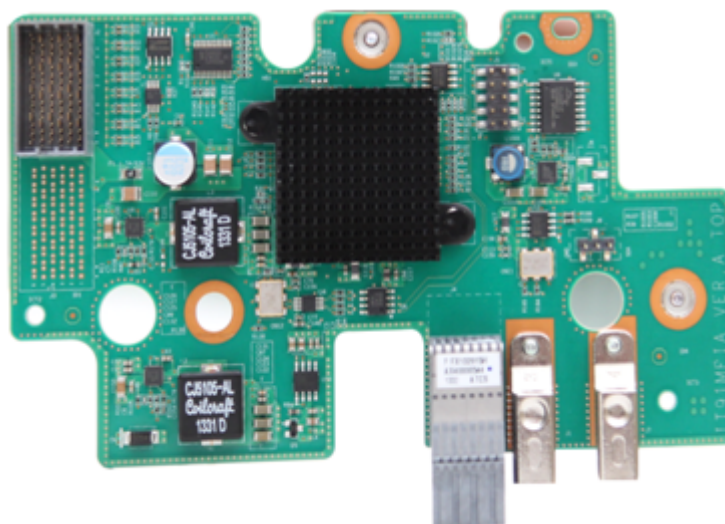


Figure 1-2 MZ610 installation positions on a half-width compute node

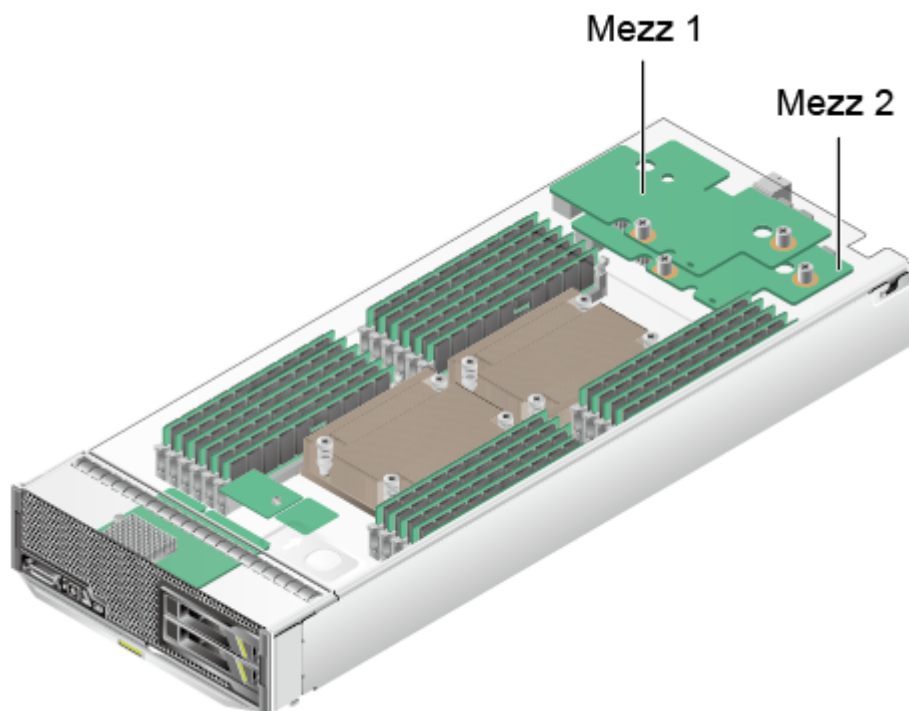
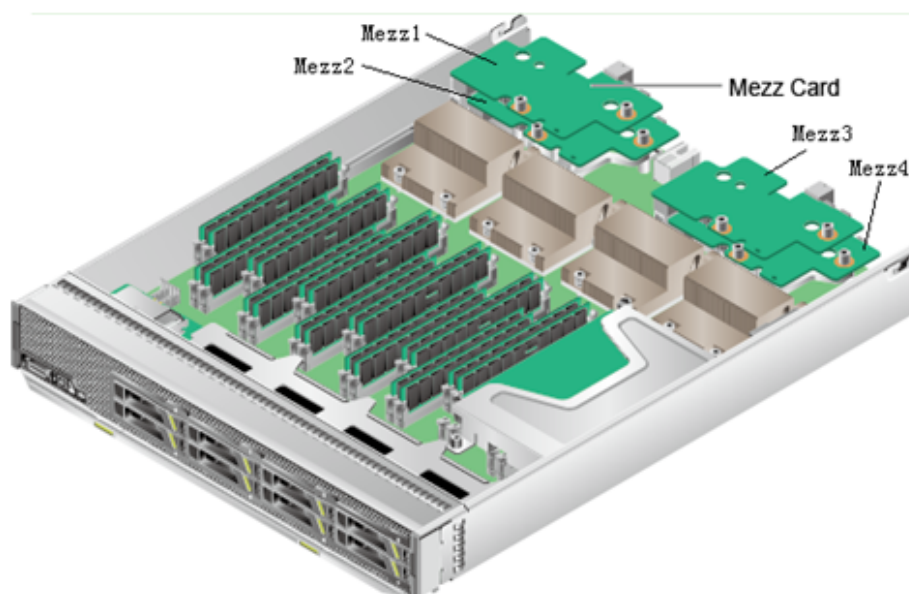


Figure 1-3 MZ610 installation positions on a full-width compute node



2 Features

- [2.1 Feature List](#)
- [2.2 Feature Description](#)
- [2.3 Standards Compliance](#)

2.1 Feature List

The MZ610 supports the following features and performance specifications:

- InfiniBand Trade Association (IBTA) 1.2.1 specifications
- RDMA
- 16 million I/O channels
- End-to-end QoS and nine virtual lanes (VLs), including eight data VLs and one control VL
- Hardware-based congestion control
- In-band management and support for third-party Subnet Managers (SMs)

2.2 Feature Description

RDMA

The MZ610 supports the RDMA feature. This feature uses the kernel bypass technology to reduce the packet processing and forwarding latency of the HCA, to reduce the CPU usage, and to implement low-latency data transmission over the data center network. With the RDMA feature, the end-to-end read and write delay of the HCA can reach 1 us when the packet length is 128 bytes. The MZ610 supports 16 million I/O channels (equivalent to IB QPs), and provides priority-based scheduling and flow control to support low-latency, high-bandwidth network transmission. The MZ610 supports OpenFabrics Enterprise Distribution for Linux (Linux OFED) and Mellanox OFED for Windows (WinOF).

QoS

The MZ610 supports end-to-end QoS for IB. It supports the VL mechanism defined in IB specifications, VL arbitration, and control-domain and service-domain prioritizing. The end-to-end QoS priority is represented by the service level (SL) domain of IB packets. Each SL is mapped to a VL. SMs configure VL arbitration and the SL-to-VL mapping table using network management datagrams (MADs). The MZ610 supports nine VLs, including eight data VLs and one control VL.

2.3 Standards Compliance

Table 2-1 lists the standards and protocols that the MZ610 complies with.

Table 2-1 Standards compliance

Standard	Protocol
IBTA 1.2.1	InfiniBand Trade Association
ANSI INCITS 365-2002	SCSI RDMA Protocol (SRP)
IETF	iSCSI Extensions for RDMA

Standard	Protocol
DAPL	User Direct Access Programming Library (uDAPL)

3 Applications

[3.1 Compatible Compute Nodes](#)

[3.2 Connected I/O Modules](#)

[3.3 MZ610 Networking](#)

[3.4 Supported OSs](#)

3.1 Compatible Compute Nodes

The MZ610 can be installed in slot Mezz1 or Mezz2 on a half-width compute node or in slot Mezz1, Mezz2, Mezz3, or Mezz4 on a full-width compute node. [Table 3-1](#) lists the compute nodes that support the MZ610 and its installation positions on them.

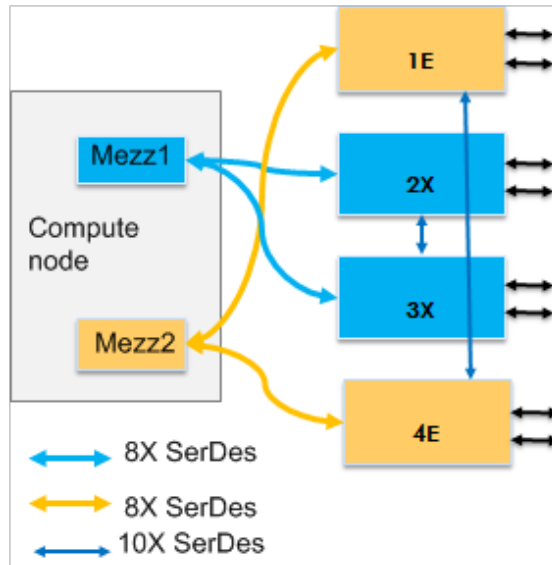
Table 3-1 Compute nodes that support the MZ610

Compute Node	Number of Mezz Module Slots	MZ610 Installation Position
CH121	2	Mezz1 and Mezz2
CH121 V3	2	Mezz1 and Mezz2
CH220	1	Mezz1
CH220 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH221	1	Mezz1
CH222	2	Mezz1 and Mezz2
CH222 V3	2	Mezz1 and Mezz2
CH240	2	Mezz1 and Mezz2
CH242	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH242 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4

3.2 Connected I/O Modules

MZ610s can connect to I/O modules (switch modules or interface boards). [Figure 3-1](#) shows the connections between the MZ610s on a half-width compute node and the ports on I/O modules.

Figure 3-1 Connections between the MZ610s on a half-width compute node and the ports on I/O modules



There are two or four groups of Serializer/Deserializer (SerDes, known as high-speed interconnect line) between each compute node and I/O module slots.

- Mezz1: 8X SerDes for connecting to I/O module slots 2X and 3X
- Mezz2: 8X SerDes for connecting to I/O module slots 1E and 4E
- Mezz3 (available only on a full-width compute node): 8X SerDes for connecting to I/O module slots 2X and 3X
- Mezz4 (available only on a full-width compute node): 8X SerDes for connecting to I/O module slots 1E and 4E

NOTE

The MZ610 provides two ports, and only 4X of each 8X SerDes is used.

Table 3-2 describes the I/O modules to which the MZ610 can connect.

Table 3-2 I/O modules to which the MZ610 can connect

I/O Module	I/O Module Slot	MZ610		Typical Configuration	Remarks
		Mezz1	Mezz2		
CX610	2X/3X	√	X	No	-
	1E/4E	X	√	Yes	You are advised to install CX610s in slots 1E and 4E.
CX611	2X/3X	√	X	No	-

	1E/4E	X	✓	No	-
--	-------	---	---	----	---

3.3 MZ610 Networking

The MZ610 can connect to I/O modules (switch modules or interface boards) to provide IB services.

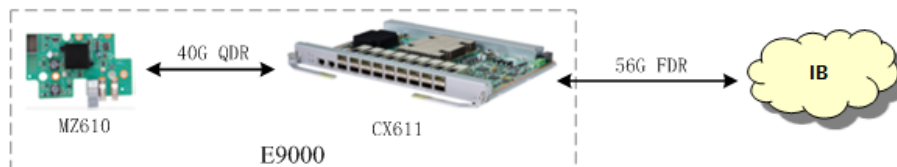
The MZ610 can work with the CX610 switch module to provide 40 Gbit/s bandwidth, and connect to the external IB network through 40G QDR ports on the CX610. See [Figure 3-2](#).

Figure 3-2 Connection between the MZ610 and the CX610



The MZ610 can work with the CX611 switch module to provide 40 Gbit/s bandwidth, and connect to the external IB network through 56G FDR ports on the CX611. See [Figure 3-3](#).

Figure 3-3 Connection between the MZ610 and the CX611



NOTE

The CX3 chip system on the MZ610 uses a PCIe 3.0 x8 port with a theoretical bandwidth of 64 Gbit/s. The two ports on the MZ610 can actually provide a total bandwidth of 56 Gbit/s instead of 80 Gbit/s because of protocol overhead and efficiency loss.

3.4 Supported OSs

[Table 3-3](#) lists the OSs supported by the MZ610.

Table 3-3 OSs supported by the MZ610

OS	Version	Remarks
Windows	Windows Server 2008 Enterprise SP2	-
	Windows Server 2008 R2 Enterprise x64	-

OS	Version	Remarks
	Windows Server 2012 Enterprise x64	-
Red Hat	Red Hat Enterprise Linux (RHEL) 6.1 x86	-
	RHEL 6.1 x86_64	-
	RHEL 6.3 x86	-
	RHEL 6.3 x86_64	-
SUSE	SUSE Linux Enterprise Server (SLES) 11 SP1 x86	-
	SLES 11 SP1 x86_64	-
	SLES 11 SP2 x86	-
	SLES 11 SP2 x86_64	-
VMware	VMware ESXi 5.0	-
	VMware ESXi 5.1	-
	VMware ESXi 5.5	-
Oracle Linux	Oracle Linux 6.2 x86	-
	Oracle Linux 6.2 x86_64	-
	Oracle Linux 6.3 x86	-
	Oracle Linux 6.3 x86_64	-

For the latest versions, see the E9000 Compatibility List.

4 Technical Specifications

4.1 Technical Specifications

4.1 Technical Specifications

Table 4-1 lists the technical specifications for the MZ610.

Table 4-1 Technical specifications

Item	Specifications
Dimensions (length x width)	148 mm x 85 mm (5.83 in. x 3.35 in.)
Power supply	12 V DC
Net weight	0.3 kg (0.66 lb)
Maximum power consumption	8 W
Temperature	Operating temperature: 5°C to 40°C (41°F to 104°F)
	Storage temperature: -40°C to +70°C (-40°F to +158°F)
Temperature change rate	15°C/h (27°F/h)
Humidity	Operating humidity: 5% RH to 85% RH (non-condensing)
	Storage humidity: 5% RH to 95% RH (non-condensing)
Altitude	<ul style="list-style-type: none">● 40°C (104°F) at 1800 m (5905.44 ft)● 30°C (86°F) at 3000 m (9842.40 ft) When the MZ610 is used at an altitude of 1800 m to 3000 m, the temperature decreases by 1°C (1.8°F) as the altitude increases by 120 m (393.70 ft).
PCIe port bandwidth	64 Gbit/s (PCIe 3.0 x8)
Port rate	40 Gbit/s
Number of ports	2
Port type	IB
Chip model/manufacturer	ConnectX-3 (CX3)/Mellanox

A Acronyms and Abbreviations

H	
HCA	host channel adapter
I	
IB	InfiniBand
IBTA	InfiniBand Trade Association
IO	input/output
M	
MAD	Management Datagram
O	
OFED	OpenFabrics Enterprise Distribution
OS	operating system
P	
PCIe	Peripheral Component Interconnect Express
Q	
QDR	quad data rate
QoS	quality of service
QP	queue pair
R	
RDMA	Remote Direct Memory Access

S	
SL	service level
SM	Subnet Manager
V	
VL	virtual lane