Huawei MZ620 NIC V100R001

White Paper

Issue	02
Date	2018-06-22

Huawei Technologies Co., Ltd.

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

Website: http://e.huawei.com

About This Document

Purpose

This document describes the MZ620 in terms of its functions, appearance, features, applications, and technical specifications. You can obtain comprehensive information about the MZ620 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	Indicates a hazard with a high level or medium level of risk which, if not avoided, could result in death or serious injury.
⚠ WARNING	Indicates a hazard with a low level of risk which, if not avoided, could result in minor or moderate injury.
⚠ CAUTION	Indicates a potentially hazardous situation that, if not avoided, could 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	Date	Description
02	2018-06-22	● **3 Applications** is added. ● Removed the section Compatible Compute Nodes and I/O Modules.
01	2016-11-21	This issue is the first official release.

Contents

1 Overview

1.1 Functions

The MZ620 is an InfiniBand (IB) host channel adapter (HCA). It provides two 100 Gbit/s IB enhanced data rate (EDR) ports for E9000 compute nodes to connect to switch modules in the chassis.

The MZ620 uses the Mellanox ConnectX-4 (CX4) chip and supports HCA applications. The 100 Gbit/s ports support 100/56/40 Gbit/s auto-negotiation and support 100 Gbit/s (EDR), 56 Gbit/s (FDR), and 40 Gbit/s (QDR) port applications (FDR stands for fourteen data rate; QDR stands for quad data rate). The MZ620 supports the remote direct memory access (RDMA) feature to address low-latency network applications.

1.2 Appearance

The MZ620 can be installed in slot Mezzanine1 (Mezz1 for short) or Mezzanine2 (Mezz2 for short) on a half-width compute node or in slot Mezz1, Mezz2, Mezzanine3 (Mezz3 for short), or Mezzanine4 (Mezz4 for short) on a full-width compute node. The MZ620 provides network ports for E9000 compute nodes to connect to switch modules in the chassis.

- When the MZ620 is installed in slot Mezz1 or Mezz3, its two 100 Gbit/s physical ports connect to switch module slots 2X and 3X.

- When the MZ620 is installed in slot Mezz2 or Mezz4, its two 100 Gbit/s physical ports connect to switch module slots 1E and 4E.
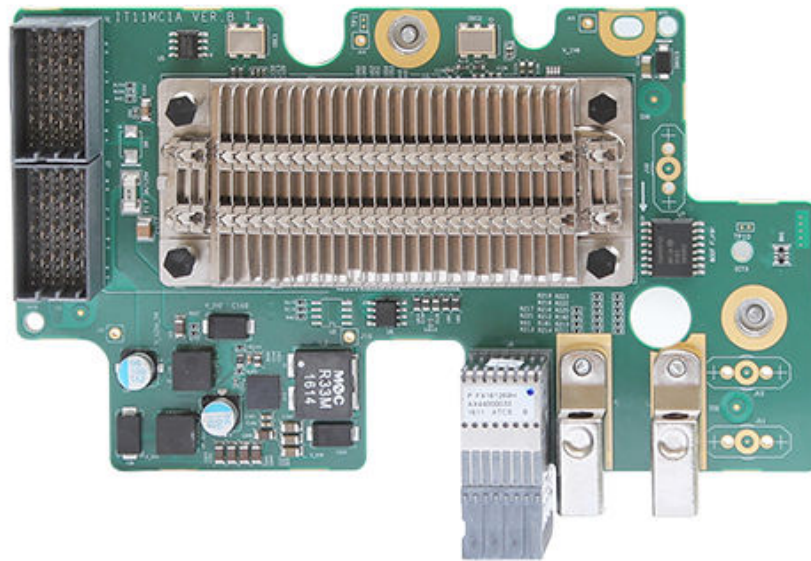
Figure 1-1 MZ620 appearance

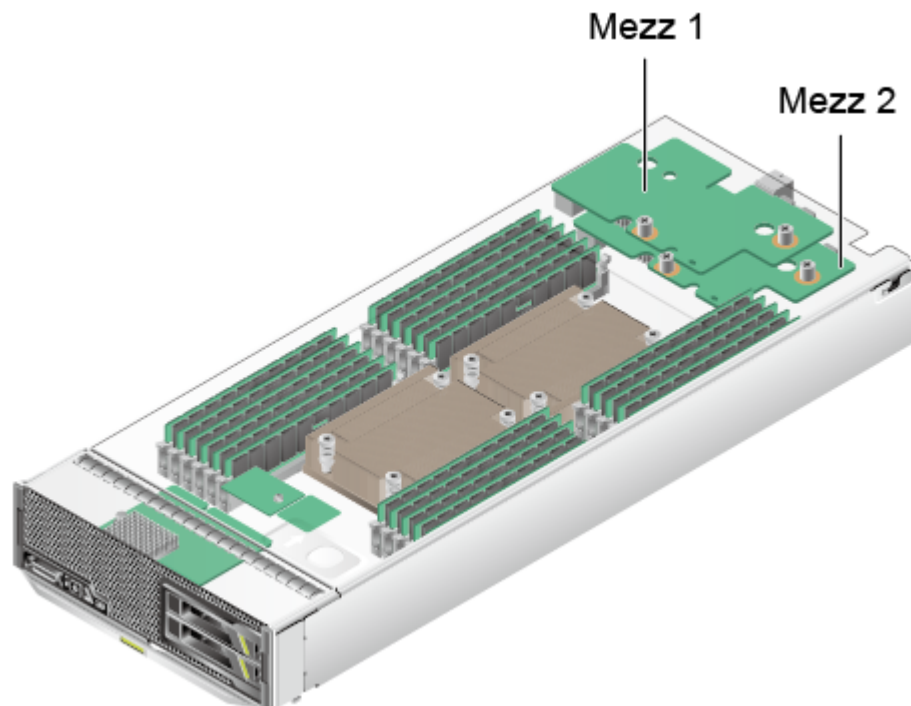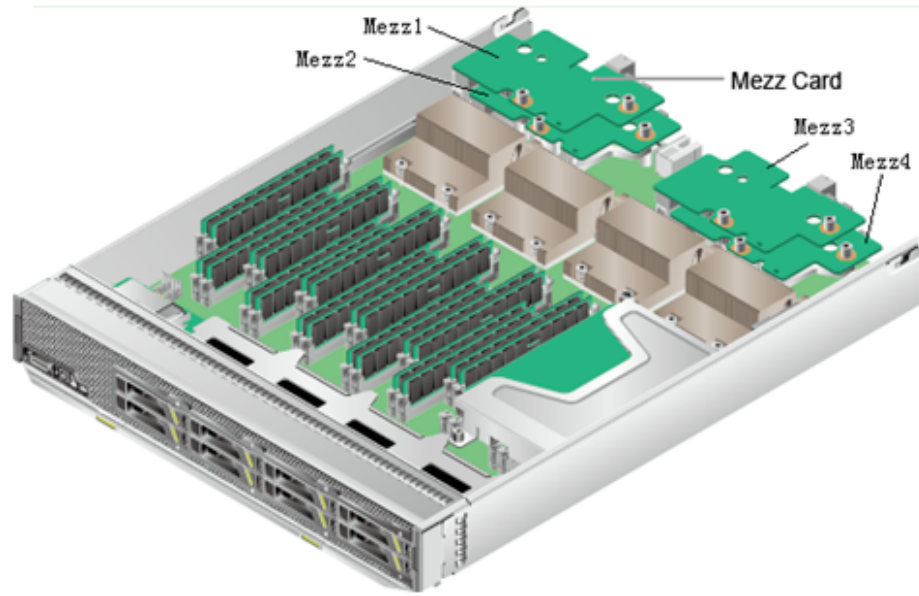**Figure 1-2** MZ620 installation positions on a half-width compute node

Figure 1-3 MZ620 installation positions on a full-width compute node

2 Features

2.1 Feature List

The MZ620 supports the following features and performance specifications:

- InfiniBand Trade Association (IBTA) 1.3 specifications
- Remote direct memory access (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
- Inband management and support for third-party subnet managers (SMs)
- PCIe 3.0 x16 system interface

📖**NOTE**

> If the flow control function is configured for the NIC, this function must also be configured for the peer device (switch module or external switch).

2.2 Feature Description

RDMA

The MZ620 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 MZ620 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 MZ620 supports OpenFabrics Enterprise Distribution for Linux (Linux OFED).

QoS

The MZ620 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 MZ620 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 MZ620 complies with.

Table 2-1 Standards and protocols that the MZ620 complies with

Standard	Protocol
IBTA 1.3	InfiniBand Trade Association
ANSI INCITS 365-2002	SCSI RDMA Protocol (SRP)
IETF	iSCSI Extensions for RDMA
DAPL	User Direct Access Programming Library (uDAPL)

3 Applications

3.1 Software and Hardware Compatibility

For details about the software and hardware that are compatible with the MZ620, see *Huawei Server Compatibility Checker*.

3.2 Networking

The MZ620 can connect to I/O modules (switch modules) to provide IB services.

The MZ620 can work with the CX620 switch module to provide 100 Gbit/s interface bandwidth, and connect to the external IB network through 100 Gbit/s EDR ports on the CX620. See **Figure 3-1**.

Figure 3-1 Connection between the MZ620 and the CX620
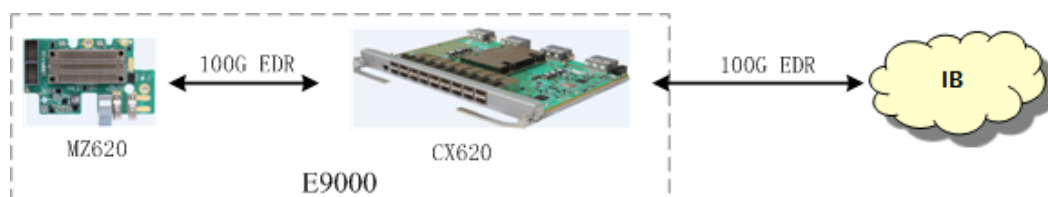
4 Technical Specifications

4.1 Technical Specifications

Table 4-1 lists the technical specifications for the MZ620.

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	0.15 kg
Maximum power consumption	18 W
Temperature	Operating temperature: 5°C to 40°C (41°F to 104°F) (ASHRAE Class A3 compliant)
	Storage temperature: –40°C to +65°C (–40°F to +149°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	40°C (104°F) at 900 m (2952 ft.)
	When the device is used at an altitude of 900 m to 5000 m (2952 ft. to 16404 ft.), the highest operating temperature decreases by 1°C (1.8°F) for every increase of 300 m (984 ft.).

Item	Specifications
Corrosive air pollutant	● Corrosion rate of the copper test piece: < 300 Å/month (in compliance with the ANSI/ISA-71.04-2013 gaseous corrosion level G1) ● Corrosion rate of the silver test piece: < 200 Å/month
Particulate pollutant	● The ISO14664-1 Class 8 requirements are met. You are advised to ask a professional organization to monitor particulate pollutants in the equipment room. ● There is no explosive, conductive, magnetic, or corrosive dust in the equipment room.
Number of ports	2
Port type	IB
Chip model/ manufacturer	ConnectX-4 (CX4)/Mellanox

A Acronyms and Abbreviations

E	
EDR	Enhanced Data Rate
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	
QoS	quality of service
QP	queue pair
R	

RDMA	Remote Direct Memory Access
S	
SL	service level
SM	Subnet Manager
V	
VL	virtual lane