Huawei MZ622 HCA

V100R001

White Paper

Issue **01**

Date **2016-11-21**

Copyright © Huawei Technologies Co., Ltd. 2016. 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

Notice

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 MZ622 in terms of its functions, appearance, features, applications, and technical specifications. You can obtain comprehensive information about the MZ622 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 (2016-11-21)

This issue is the first official release.

Contents

Figures

Tables

1 Overview

About This Chapter

1.1 Functions

The MZ622 is an InfiniBand (IB) host channel adapter (HCA) that provides two 100G IB Enhanced Data Rate (EDR) ports. It provides network ports for E9000 compute nodes to connect to switch modules in the chassis.

The MZ622 uses the Mellanox ConnectX-4 (CX4) chip and supports HCA applications. The 100G ports support 100G/56G/40G auto-negotiation and support 100G (EDR), 56G (FDR), and 40G (QDR) port applications (FDR stands for Fourteen Data Rate; QDR stands for Quad Data Rate). The MZ622 supports the Remote Direct Memory Access (RDMA) feature to address low-latency network applications.

1.2 Appearance

The MZ622 can be installed in slot Mezzanine1 (Mezz1 for short) or Mezzanine2 (Mezz2 for short) on a half-width E9000 compute node, providing network ports for connecting to switch modules.

Figure 1-1 Appearance of the MZ622
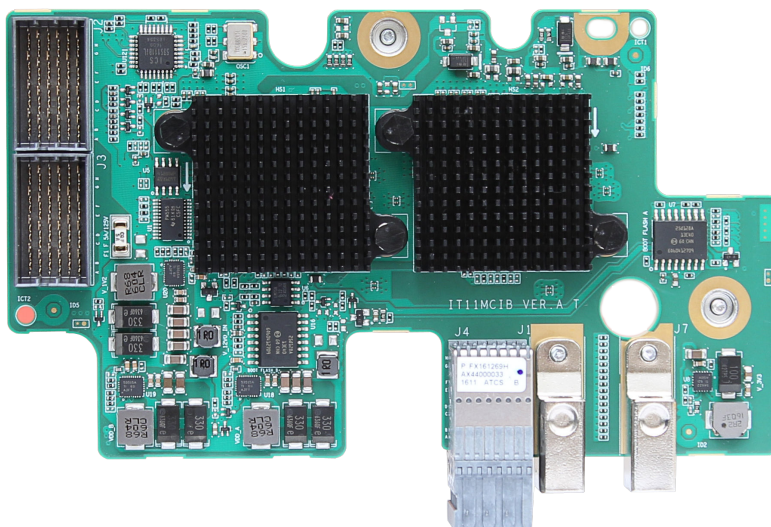
Figure 1-2 MZ622 installation positions on a half-width compute node
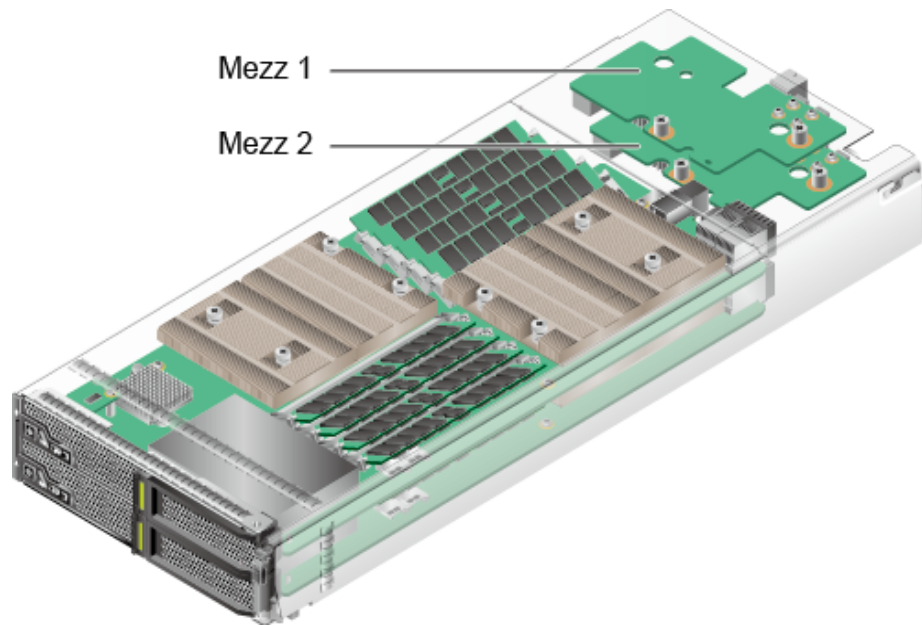
2 Features

About This Chapter

2.1 Feature List

The MZ622 supports the following features and performance specifications:

- InfiniBand Trade Association (IBTA) 1.3 specifications

- RDMA

- 32 million (2 x 16 million) I/O channels

- End-to-end QoS and 18 virtual lanes (VLs), including 2 x 8 data VLs + 2 x 1 control VLs

- Hardware-based congestion control

- In-band management and support for third-party Subnet Managers (SMs)

- Two PCIe 3.0 x8 system interfaces

2.2 Feature Description

RDMA

The MZ622 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 MZ622 supports 32 million (2 x 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 MZ622 supports OpenFabrics Enterprise Distribution for Linux (Linux OFED).

QoS

The MZ622 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 MZ622 supports 18 VLs, including 16 (2 x8) data VLs and two (2 x 1) control VLs.

2.3 Standards Compliance

Table 2-1 lists the standards and protocols that the MZ622 complies with.

Table 2-1 Standards compliance

Standard	Protocol
IBTA 1.3	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

About This Chapter

3.1 Compatible Compute Nodes

The MZ622 can be installed in slot Mezz1 or Mezz2 on a half-width compute node. **Table 3-1** lists the compute nodes that support the MZ622 and its installation positions on them.

Table 3-1 Compute nodes that support the MZ622

Compute Node	Number of Mezzanine Card Slots	MZ622 Installation Position
CH140 V3	2	Mezz1 and Mezz2

3.2 Compatible I/O Modules

Table 3-2 lists the I/O modules (switch modules) that can connect to the MZ622.

Table 3-2 Compatible I/O modules

I/O Module	I/O Module Slot	MZ622 (Mezz1)	MZ622 (Mezz2)	Typical Configuration	Remarks
CX611	2X/3X	√	X	No	-
	1E/4E	X	√	No	You are advised to install CX611s in slots 1E and 4E.
CX620	2X/3X	√	X	No	-
	1E/4E	X	√	Yes	You are advised to install CX620s in slots 1E and 4E.

3.3 MZ622 Networking

The MZ622 can connect to I/O modules (switch modules) to provide IB services.

The MZ622 can work with the CX611 switch module to provide 56 Gbit/s interface bandwidth, and connect to the external IB network through 56G FDR ports on the CX611. See **Figure 3-1**.

Figure 3-1 Connection between the MZ622 and the CX611
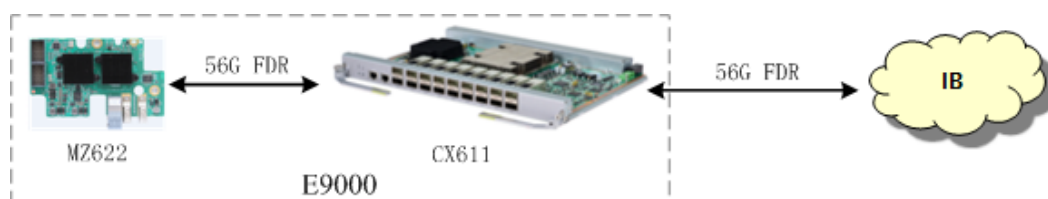
The MZ622 can work with the CX620 switch module to provide 100 Gbit/s interface bandwidth, and connect to the external IB network through 100G EDR ports on the CX620. See **Figure 3-2**.

Figure 3-2 Connection between the MZ622 and the CX620

3.4 Supported OSs

Table 3-3 lists the OSs supported by the MZ622.

Table 3-3 OSs supported by the MZ622

OS	Version	Remarks
CentOS	CentOS 6.5 x86_64	-
	CentOS 6.6 x86_64	-
	CentOS 6.7 x86_64	-
	CentOS 6.8 x86_64	-
	CentOS 7.0 x86_64	-
	CentOS 7.1 x86_64	-
	CentOS 7.2 x86_64	-
Oracle Linux	Oracle Linux 6.5 x86_64	-
	Oracle Linux 6.6 x86_64	-
	Oracle Linux 6.7 x86_64	-
	Oracle Linux 7.1 x86_64	-
Red Hat Enterprise Linux	RHEL 6.5 x86_64	-
	RHEL 6.6 x86_64	-

OS	Version	Remarks
	RHEL 6.7 x86_64	-
	RHEL 6.8 x86_64	-
	RHEL 7.0 x86_64	-
	RHEL 7.1 x86_64	-
	RHEL 7.2 x86_64	-
SUSE Linux Enterprise Server (SLES)	SLES 11 SP2 x86_64	-
	SLES 11 SP3 x86_64	-
	SLES 11 SP4 x86_64	-
	SLES 12 SP0 x86_64	-
	SLES 12 SP1 x86_64	
Ubuntu	Ubuntu 15.04 x86_64	-
	Ubuntu 15.10 x86_64	-
	Ubuntu 16.04 x86_64	-

To view the latest list of supported OSs, use **Huawei Server Compatibility Checker**.

4 Technical Specifications

About This Chapter

4.1 Technical Specifications

Table 4-1 lists the technical specifications for the MZ622.

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	0.16 kg
Maximum power consumption	21.5 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 (2952.76 ft) When the equipment is used at an altitude from 900 m to 3000 m, the highest operating temperature decreases by 1°C (1.8°F) for every increase of 300 m (393.70 ft).
PCIe port bandwidth	2 x PCIe 3.0 x8
Port rate	40 Gbit/s, 56 Gbit/s, and 100 Gbit/s
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