Huawei MZ311 NIC

V100R001

White Paper

Issue   08

Date   2016-11-21

HUAWEI TECHNOLOGIES CO., LTD.

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 MZ311 in terms of its functions, appearance, features, applications, and technical specifications. You can obtain comprehensive information about the MZ311 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 08 (2016-11-21)

This issue is the eighth official release.

Issue 07 (2016-08-17)

This issue is the seventh official release.

Issue 06 (2016-07-15)

This issue is the sixth official release.

Issue 05 (2016-05-15)

This issue is the fifth official release.

Type	Change Description
Modify	The document content is optimized.

Issue 04 (2016-03-31)

This issue is the fourth official release.

Issue 03 (2015-05-30)

This issue is the third official release.

Issue 02 (2014-12-02)

This issue is the second official release.

Issue 01 (2014-11-10)

This issue is the first official release.

Contents

Figures

Tables

1 Overview

About This Chapter

1.1 Functions

The MZ311 is an Ethernet NIC. It is used for E9000 compute nodes and provides four 10GE ports for compute nodes to connect to switch modules in the chassis.

The MZ311 uses the Mellanox ConnectX-3 (CX3 for short) chip and supports NIC applications. The MZ311 also supports the Remote Direct Memory Access over Converged Ethernet (RoCE) feature to provide low-latency network applications.

1.2 Appearance

The MZ311 can be installed in slot Mezzanine1 (Mezz1 for short) or Mezzanine2 (Mezz2 for short) on a half-width E9000 compute node or in slot Mezz1, Mezz2, Mezzanine3 (Mezz3 for short), or Mezzanine4 (Mezz4 for short) on a full-width E9000 compute node.

The MZ311 provides network ports for connecting to switch modules:

● When the MZ311 is installed in slot Mezz1 or Mezz3, its four 10GE ports connect to switch modules in slots 2X and 3X.

● When the MZ311 is installed in slot Mezz2 or Mezz4, its four 10GE ports connect to switch modules in slots 1E and 4E.

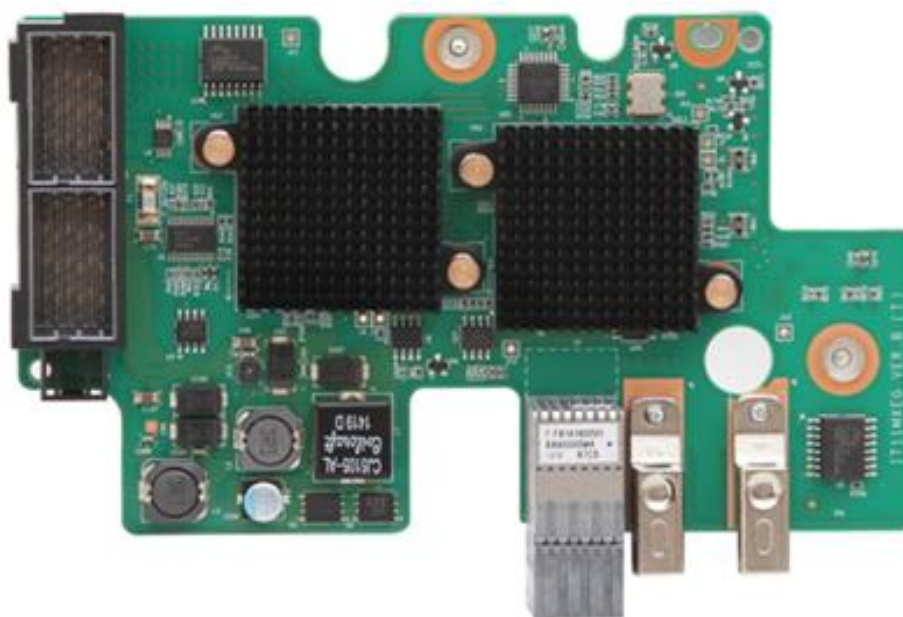**Figure 1-1** MZ311 appearance

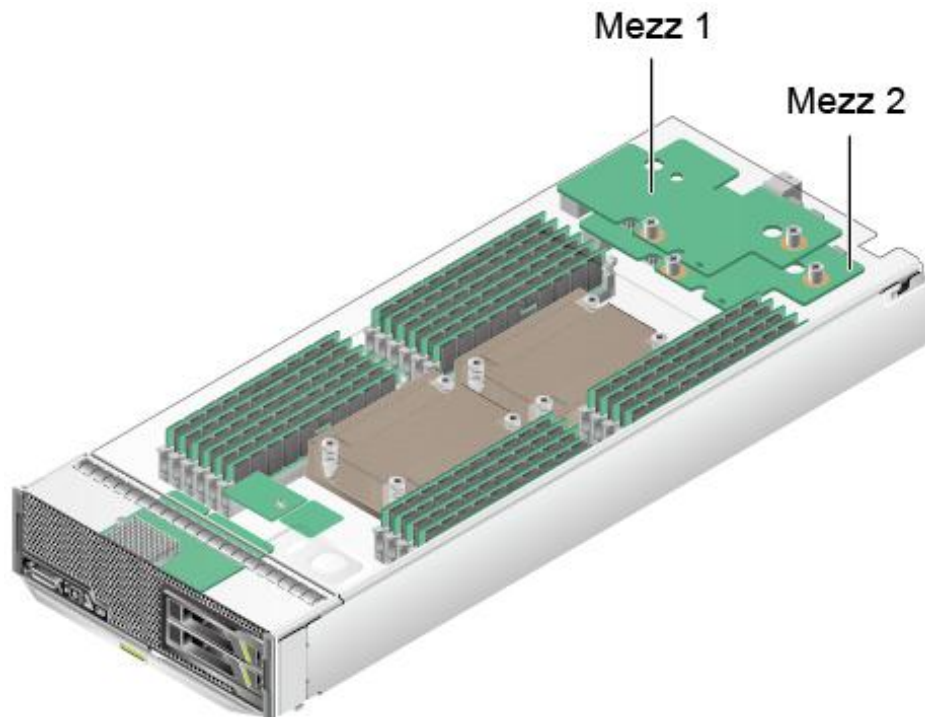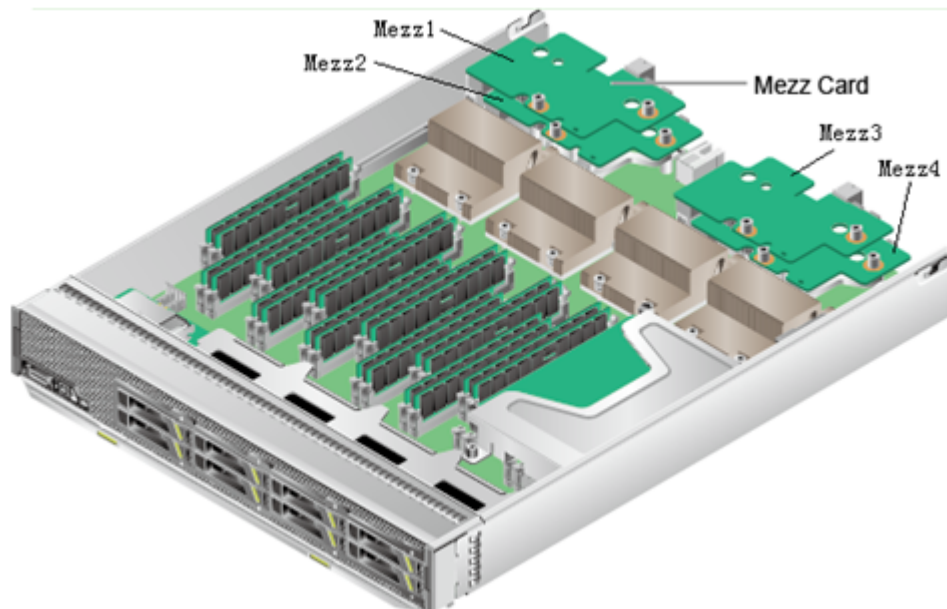**Figure 1-2** MZ311 installation positions on a half-width compute node

Figure 1-3 MZ311 installation positions on a full-width compute node

2 Features

About This Chapter

2.1 Feature List

The MZ311 supports the following features and performance specifications:

- RoCE and RoCE v2 (R-RoCE over UDP)
- Single Root I/O Virtualization (SR-IOV), supporting a maximum of 252 (2 x 126) virtual functions (VFs)
- NetQueue
- Ethernet Preboot Execution Environment (PXE)
- RoCE Offload
- 802.1Q VLAN, supporting a maximum of 4094 VLANs
- Support TCP checksum offload, UDP checksum offload, Receive Side Scaling (RSS) and Large Segmentation Offload (LSO)
- Jumbo frames of 9.6 KB
- Data Plane Development Kit (DPDK)

📖NOTE

Certain functions are related to OS and switch features. For details, contact Huawei technical support.

2.2 Feature Description

RoCE

The MZ311 supports the RoCE feature. This feature uses the Remote Direct Memory Access (RDMA) technology to reduce the packet processing and forwarding latency of the NIC, to reduce the CPU usage, and to implement low-latency data transmission over the data center network. With the RoCE feature, the end-to-end read and write delay of the NIC can reach 4 us when the packet length is 128 bytes. The MZ311 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 MZ311 supports OpenFabrics Enterprise Distribution for Linux (Linux OFED) and Mellanox OFED for Windows (WinOF).

802.1Q VLAN

The MZ311 supports a maximum of 4094 VLANs. Each 10GE port supports a maximum of 4094 VLANs. The VLAN IDs are integers ranging from 1 to 4094.

The MZ311 does not tag or untag packets, but transparently transmits them. VLAN IDs are specified by the operating system (OS) on an E9000 compute node.

SR-IOV

The MZ311 supports the SR-IOV feature. The NIC supports a maximum of 252 (2 x 126) VFs, and each physical 10GE port on the NIC supports a maximum of 63 VFs.

The VFs derived from each physical function (PF) can be assigned to virtual machines (VMs). In this way, the mapping between VFs and VMs is set up.

2.3 Standards Compliance

Table 2-1 lists the standards and protocols that the MZ311 complies with.

Table 2-1 Standards compliance

Standard	Protocol
IEEE 802.3x	Flow Control and Back Pressure
IEEE 802.3z	1000BASE-X
IEEE 802.3ap	10GBASE-KR
IEEE 802.1ab	Station and Media Access Control Connectivity Discovery (LLDP)
IEEE 802.3ad	Link Aggregation Control Protocol (LACP)
IEEE 802.1Q	VLAN Tagging
IEEE 802.1Qbg	Edge Virtual Bridging (EVB)

3 Applications

About This Chapter

3.1 Compatible Compute Nodes

The MZ311 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 MZ311 and its installation positions on them.

Table 3-1 Compute nodes that support the MZ311

Compute Node	Number of Mezz Module Slots	MZ311 Installation Position
CH121	2	Mezz1 and Mezz2
CH121 V3	2	Mezz1 and Mezz2
CH220	1	Mezz1
CH220 V3	2	Mezz2 and Mezz3
CH221	1	Mezz1
CH222	2	Mezz1 and Mezz2
CH222 V3	2	Mezz1 and Mezz2
CH240	2	Mezz1 and Mezz2
CH242	1	Mezz1
CH242 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH140	2	Mezz1 and Mezz2
CH140 V3	2	Mezz1 and Mezz2

3.2 Connected I/O Modules

MZ311s can connect to I/O modules (switch modules or interface boards). **Figure 3-1** shows the connections between the MZ311s on a half-width compute node and the ports on I/O modules.

Figure 3-1 Connections between the MZ311s on a half-width compute node and the ports on I/O modules
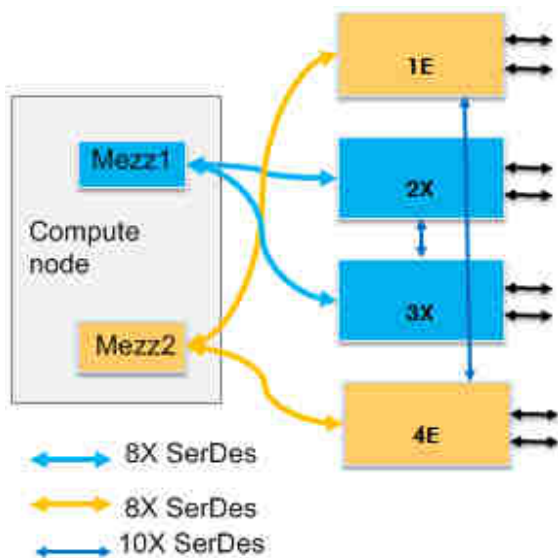
There are two or four groups of Serializer/Deserializer (SerDes, known as high-speed interconnect line) between each compute node and I/O module slots.

● 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 MZ311 provides four ports, and only 2X of each 8X SerDes is used.

Table 3-2 describes the I/O modules to which the MZ311 can connect.

Table 3-2 I/O modules to which the MZ311 can connect

I/O Module	I/O Module Slot	MZ311 (Mezz1)	MZ311 (Mezz2)	Typical Configuration	Remarks
CX910	2X/3X	√	X	Yes	-
	1E/4E	X	√	No	It is recommended that the CX910 not be installed in slot 1E or 4E.
CX310	2X/3X	√	X	Yes	-

I/O Module	I/O Module Slot	MZ311 (Mezz1)	MZ311 (Mezz2)	Typical Configuration	Remarks
	1E/4E	X	√	Yes	-
CX317	2X/3X	√	X	No	It is recommended that the CX317 not be installed in slot 2X or 3X.
	1E/4E	X	√	Yes	-
CX318	2X/3X	√	X	No	It is recommended that the CX318 not be installed in slot 2X or 3X.
	1E/4E	X	√	Yes	-

3.3 MZ311 Networking

The MZ311 can connect to I/O modules (switch modules or interface boards) to provide Ethernet services.

The MZ311 can work with the CX310 switch module to provide 40 Gbit/s bandwidth, and connect to the Internet through 10GE ports on the CX310. See **Figure 3-2**.

Figure 3-2 Connection between the MZ311 and the CX310

The MZ311 can work with the CX910 switch module to provide 40 Gbit/s bandwidth, and connect to the Internet through 10GE ports on the CX910. See **Figure 3-3**.

Figure 3-3 Connection between the MZ311 and the CX910

The MZ311 can work with the CX317 pass through module to provide 40 Gbit/s, and connect to the Internet through 10GE ports on the CX317 in pass-through mode. See **Figure 3-4**.

Figure 3-4 Connection between the MZ311 and the CX317

The MZ311 can work with the CX318 pass through module to provide 40 Gbit/s, and connect to the Internet through 10GE ports on the CX318 in pass-through mode. See **Figure 3-5**.

Figure 3-5 Connection between the MZ311 and the CX318
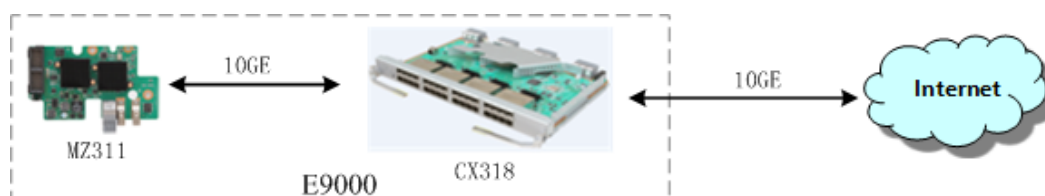
3.4 Supported OSs

Table 3-3 lists the OSs supported by the MZ311.

Table 3-3 OSs supported by the MZ311

OS	Version	Remarks
Redhat	RHEL 6.5	-
	RHEL 6.6	-
	RHEL 6.7	-
	RHEL 7.0	-
Suse	SLES 11.3	-
	SLES 12.0	-
VMware	VMware ESXi 5.5.2	-
	VMware ESXi 5.5.3	-
Windows	Windows 2008 R2 SP1	-
	Windows 2012	-
	Windows 2012 R2	-

The preceding table is for reference only. Compatible OSs for the MZ311 vary with the compute node type. For details, see the *Huawei Server Compatibility Checker*.

3.5 Connected Ethernet Switches

Table 3-4 lists the Data Center Bridging (DCB) switches to which the MZ311 can connect.

Table 3-4 DCB switches to which the MZ311 can connect

Category	Vendor	Model	Remarks
DCB switch	Huawei	CX910	The CX910 is a DCB switch module on the E9000.
		CX310	The CX310 is a DCB switch module on the E9000.
		CE6800	The MZ311 connects to CE6800 through the CX317 on the E9000.
		CE12800	The MZ311 connects to CE12800 through the CX317 on the E9000.
	Cisco	Nexus 5548	The MZ311 connects to Nexus 5548 through the CX317 on the E9000.
		Nexus 5596	The MZ311 connects to Nexus 5596 through the CX317 on the E9000.
		Nexus 2232PP	Nexus 2232PP is a Cisco Fabric Extender. The MZ311 connects to Nexus 2232PP through the CX317 on the E9000.
		Nexus 2248PQ	Nexus 2248PQ is a Cisco Fabric Extender. The MZ311 connects to Nexus 2248PQ through the CX317 on the E9000.

4 Technical Specifications

About This Chapter

4.1 Technical Specifications

Table 4-1 lists the technical specifications for the MZ311.

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 (0.35 lb)
Maximum power consumption	13.5 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 1800 m (5905.44 ft) ● 30°C (86°F) at 3000 m (9842.40 ft) When the MZ311 is used at an altitude between 1800 m and 3000 m, the highest operating temperature decreases by 1°C (1.8°F) as the altitude increases by 120 m (393.70 ft).
PCIe port bandwidth	2 x 64 Gbit/s (2 x PCIe 3.0 x8)
Port rate	10.3125 Gbit/s
Number of ports	4
Port type	Ethernet
Chip model/manufacturer	ConnectX-3/Mellanox

A Acronyms and Abbreviations

D	
DPDK	Data Plane Development Kit
L	
LACP	Link Aggregation Control Protocol
LLDP	Link Layer Discovery Protocol
LSO	Large Segmentation Offload
N	
NIC	network interface card
O	
OFED	OpenFabrics Enterprise Distribution
OS	operating system
P	
PCIe	Peripheral Component Interconnect Express
PF	physical function
PXE	Preboot Execution Environment
Q	
QP	queue pair
R	
RDMA	Remote Direct Memory Access

RoCE		RDMA over Converged Ethernet
RSS		Receive Side Scaling
S		
SR-IOV		Single Root I/O Virtualization
T		
TC		traffic class
TCP		Transmission Control Protocol
U		
UDP		User Datagram Protocol
V		
VF		virtual function
VLAN		virtual local access network
VM		virtual machine