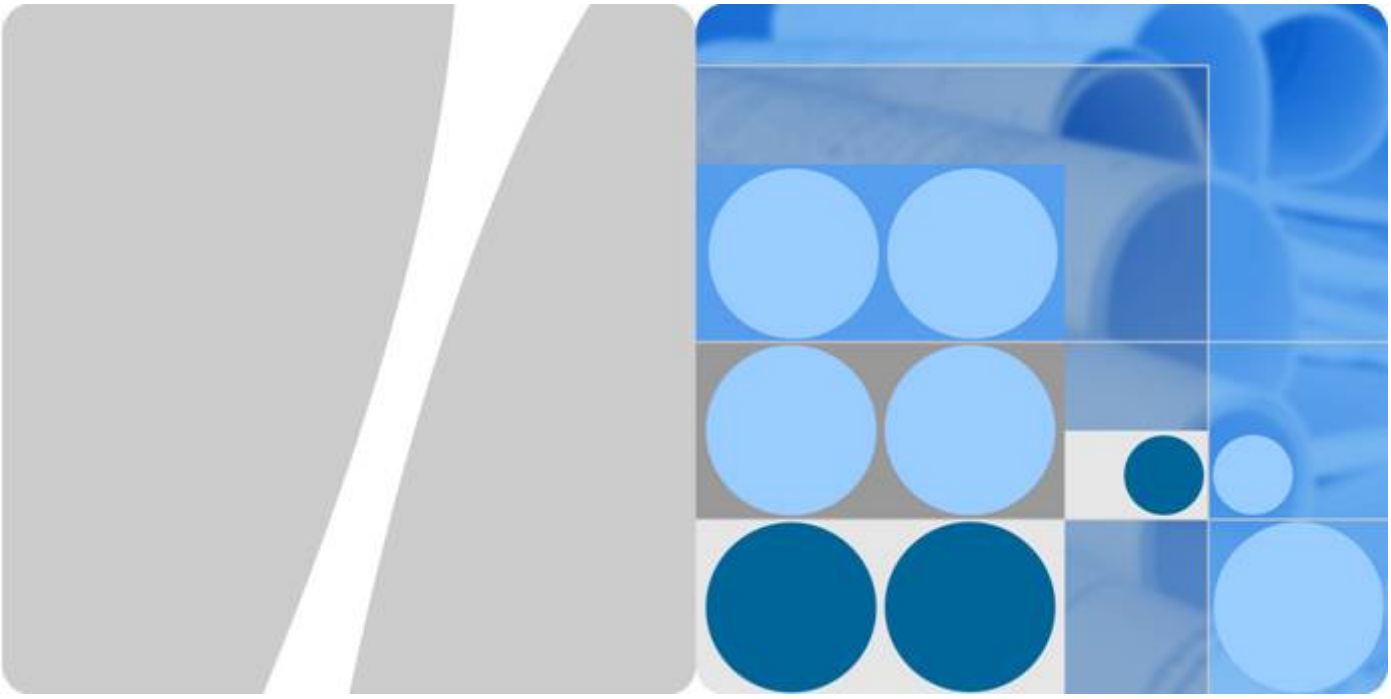


Huawei MZ512 NIC

V100R001

White Paper

Issue 09

Date 2016-11-21

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions


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://e.huawei.com

About This Document

Purpose

This document describes the MZ512 in terms of its functions, appearance, features, applications, and technical specifications. You can obtain comprehensive information about the MZ512 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 09 (2016-11-21)

This issue is the ninth official release.

Issue 08 (2016-08-17)

This issue is the eighth official release.

Issue 07 (2016-05-15)

This issue is the seventh official release.

Type	Change Description
Modify	The document content is optimized.

Issue 06 (2016-03-31)

This issue is the sixth official release.

Issue 05（2015-11-09）

This issue is the fifth official release.

Type	Change Description
Modify	Modified the note in **3.1 Compatible Compute Nodes**.

Issue 04 (2015-05-30)

This issue is the fourth official release.

Issue 03 (2014-12-02)

This issue is the third official release.

Type	Change Description
Delete	The trademark Tecal is deleted from the document name and content.

Issue 02 (2014-11-10)

This issue is the second official release.

Issue 01 (2014-09-15)

This issue is the first official release.

Contents

Figures

Tables

1 Overview

About This Chapter

1.1 Functions

The MZ512 is a 10GE converged network adapter (CNA). It is used for E9000 compute nodes and provide four ports for compute nodes to connect to switch module slots in the chassis.

The MZ512 uses the BladeEngine3 (BE3) chip, and supports NIC, Fibre Channel over Ethernet (FCoE), and Internet Small Computer System Interface (iSCSI) applications to achieve converged network solutions.

NOTE

The MZ512 with firmware 4.6.423.0 or later supports the iSCSI function.

1.2 Appearance

The MZ512 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 MZ512 provides network ports for connecting to switch modules. When the MZ512 is installed in slot Mezz1 or Mezz3, its four 10GE ports connect to switch modules in slots 2X and 3X. When the MZ512 is installed in slot Mezz2 or Mezz4, its four 10GE ports connect to switch modules in slots 1E and 4E.

Figure 1-1 MZ512 appearance

Figure 1-2 MZ512 installation positions on a half-width compute node

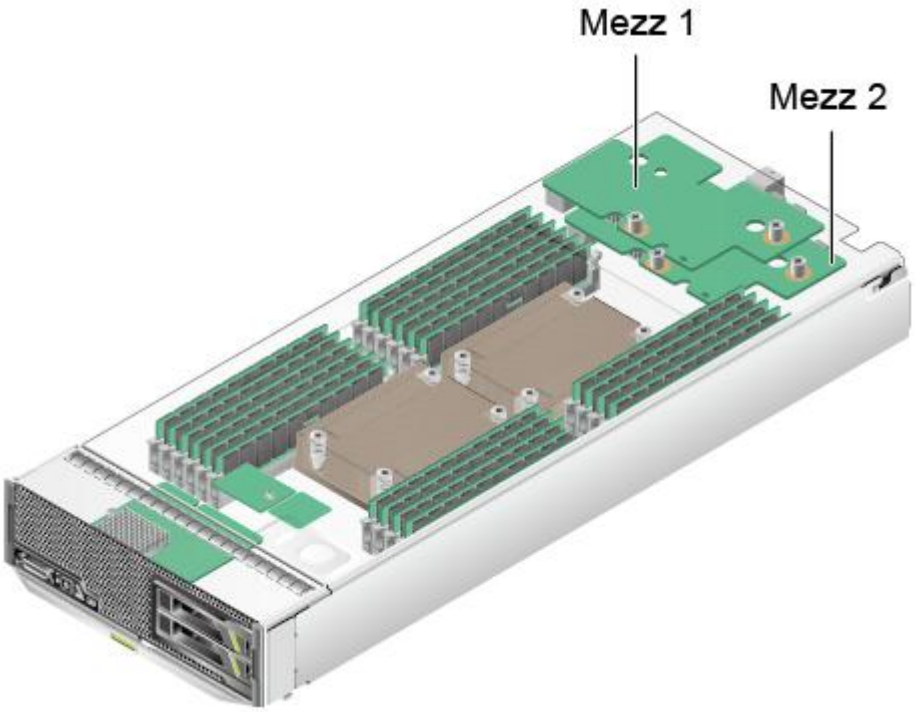
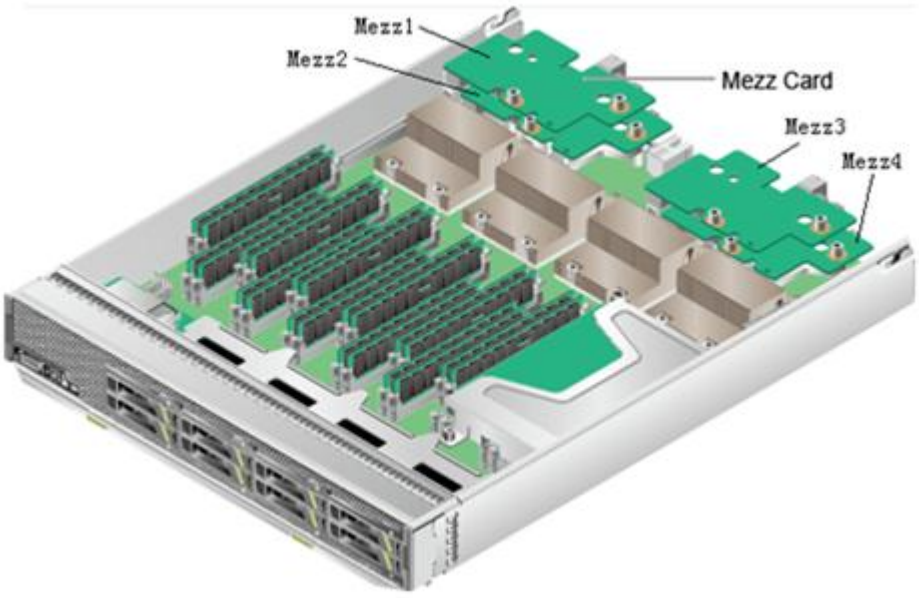


Figure 1-3 MZ512 installation positions on a full-width compute node



2 Features

About This Chapter

2.1 Feature List

The MZ512 supports the following features and performance specifications:

- FCoE, supporting hardware-based FCoE Offload
- N_Port_ID Virtualization (NPIV), supporting a maximum of 255 NPIVs per port
- Internet Small Computer Systems Interface (iSCSI), supporting hardware iSCSI offload
- Universal Multi Channel (UMC), that is, multi-PF
- Priority-based Flow Control (PFC), Enhanced Transmission Selection (ETS), and Data Center Bridging eXchange (DCBX)
- Single Root-I/O Virtualization (SR-IOV), supporting a maximum of 112 virtual functions (VFs)
- NetQueue and Virtual Machine Queue (VMQ)
- Ethernet Preboot Execution Environment (PXE), FCoE SAN Boot, and FCoE iSCSI Boot
- TCP/UDP checksum offload, Receive Side Scaling (RSS), Large Receive Offload (LRO), and Large Send Offload (LSO)
- 802.1Q VLAN, supporting a maximum of 4094 VLANs
- Jumbo frames of 9 KB

2.2 Feature Description

UMC

The MZ512 supports UMC, that is, multi-PF. Each physical 10GE port can be divided into four PFs (logical channels). Each NIC supports a total of 16 PFs. The four PFs derived from each physical 10GE port can be configured as four NICs (four native Ethernet logical channels) or three NICs and one FCoE/iSCSI (three native Ethernet logical channels and one storage logical channel). Each PF must be assigned a VLAN ID and channel bandwidth. The four PFs on a physical 10GE port are independent of each other, and share the 10 Gbit/s bandwidth of the 10GE port. The UMC feature leverages the PCIe function defined in PCI-SIG specifications, independent of the operating system (OS) or Hypervisor.

NOTE

Either UMC or SR-IOV can be enabled for the MZ512.

PFC

The 10GE ports on the MZ512 support the PFC feature. This feature supports flow control and back pressure on a per class-of-service (CoS) basis. Each physical 10GE port supports a maximum of eight CoS queues. If both the MZ512 and its connected peer device support PFC, the link between them supports PFC. The MZ512 can identify PFC frames sent by the peer device and adjusts the transmitting bandwidth. If the inbound traffic exceeds the threshold, the MZ512 sends PFC frames to instruct the peer device to control the traffic.

The MZ512 can negotiate with the peer device about PFC parameters over DCBX. Typically, CoS3 is used for FCoE traffic, which is assigned by the FCoE Forwarder (FCF). When FCoE

is enabled, only two CoS queues are available: one for FCoE traffic, and the other for native Ethernet traffic.

ETS

The 10GE ports on the MZ512 support the ETS feature. This feature assigns port transmitting bandwidth based on priority groups (PGs). Each 10GE port supports a maximum of eight CoS queues and two PGs. The minimum bandwidth and maximum bandwidth are assigned to each PG. The minimum bandwidth means the committed bandwidth, which is also known as the committed information rate (CIR). The maximum bandwidth means the maximum shared bandwidth, which is also known as the peak information rate (PIR). If other PGs are underloaded, the local PG can share (borrow) their bandwidths and its maximum bandwidth can reach 10 Gbit/s.

802.1Q VLAN

The MZ512 supports a maximum of 4094 VLANs. When UMC is disabled, each physical 10GE port supports 4094 VLANs. When UMC is enabled, each PF supports 4094 VLANs. The VLAN IDs are integers ranging from 1 to 4094.

When UMC is disabled, the MZ512 does not tag or untag packets, but transparently transmits them. VLAN IDs are specified by the OS on an E9000 compute node.

When UMC is enabled for the MZ512, a PVID (native VLAN ID) must be assigned to each PF. When receiving untagged packets from the OS, the MZ512 adds PVIDs to the packets based on the PFs that the packets belong to. When receiving tagged packets from the OS, the MZ512 directly transparently transmits them, without processing the packet tags. The MZ512 checks the VLAN tag of each received packet. If a packet VLAN tag is the same as a PF PVID, the MZ512 removes the VLAN tag from the packet and then forwards the packet to the OS. If the packet VLAN tag is different from any PF PVID, the MZ512 directly forwards the packet to the OS, without processing the VLAN tag. If a packet is untagged, the MZ512 discards the packet.

NOTICE

The VLAN tags specified by the OS must be different from the PVIDs assigned to the PFs.

SR-IOV

The MZ512 supports the SR-IOV feature. SR-IOV is available only when UMC is disabled. When UMC is disabled, the NIC supports a maximum of four physical functions (10GE physical ports) and 112 VFs. Each PF supports a maximum of 28 VFs. When UMC is enabled, the NIC supports a maximum of 8 PFs and does not support VFs.

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

In VMware, each physical port or PF supports a maximum of 16 VFs.

NOTE

Either SR-IOV or UMC can be enabled for the MZ5121.

NPIV

The MZ512 supports the NPIV feature when it operates in FCoE mode. With this feature, each port supports a maximum of 255 virtual N_Port_IDs and its own physical N_Port_ID.

Virtual N_Port_IDs can be assigned to VMs. In this way, the mapping between virtual N_Port_IDs and VMs is set up.

PXE, SAN Boot, and iSCSI Boot

The MZ512 supports PXE, SAN Boot, and iSCSI Boot.

- PXE is used for remote boot over the Ethernet or IP network. It enables users to connect to the remote PXE server for loading an OS.
- SAN Boot is used for remote boot over the FC or FCoE SAN. It enables users to connect to the remote FC or FCoE disk array for loading an OS.
- iSCSI Boot is used for remote boot over the Ethernet or IP network. It enables users to connect to the remote iSCSI disk array for loading an OS.

2.3 Standards Compliance

Table 2-1 lists the standards and protocols that the MZ512 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 802.1Qbb	Priority-based Flow Control (PFC)
IEEE 802.1Qaz	Enhanced Transmission Selection (ETS)
IEEE 802.1ab	Station and Media Access Control Connectivity Discovery (LLDP)
IEEE 802.3ad	Link Aggregation Control Protocol (LACP)
IEEE 802.1Qbg	Edge Virtual Bridging (EVB)
FC-LS	FC Link Service
FC-FS-2	FC Framing and Signaling
FC-GS-5	FC Generic Service
FCP-3	Fibre Channel Protocol for SCSI
FC-BB-5	Fibre Channel-Backbone-5 (FCoE)
DCBX	Data Center Bridging eXchange

3 Application

About This Chapter

3.1 Compatible Compute Nodes

The MZ512 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 MZ512 and its installation positions on them.

Table 3-1 Compute nodes that support the MZ512

Compute Node	Number of Mezzanine Card Slots	MZ512 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

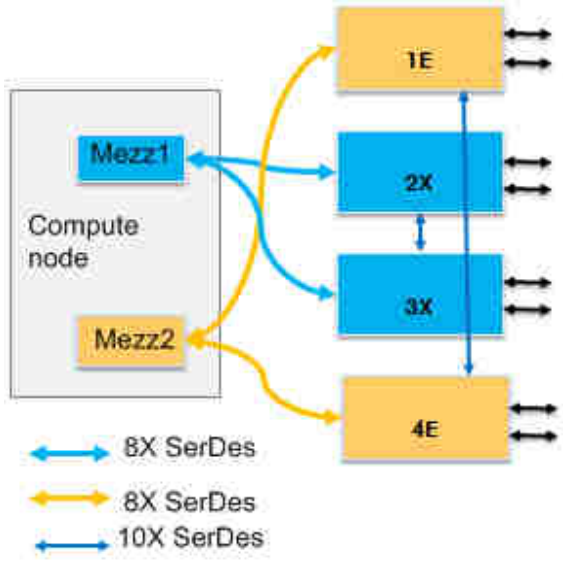
NOTE

In liquid cooling applications, the MZ512 cannot be installed on the CH140L V3 compute node.

3.2 Connected I/O Modules

The MZ512 can connect to I/O modules (switch modules or interface boards). **Figure 3-1** shows the connections between the MZ512 on a half-width compute node and the ports on I/O modules.

Figure 3-1 Connections between the MZ512s 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 MZ512 provides four ports, and only 2X of each 8X SerDes is used.

Table 3-2 describes the I/O modules to which the MZ512 can connect.

Table 3-2 I/O modules to which the MZ512 can connect

I/O Module Slot	I/O Module Slot	MZ512 (Mezz1)	MZ512 (Mezz2)	Typical Configuration	Remarks
CX310	2X/3X	√	X	Yes	-
	1E/4E	X	√	Yes	-
CX311	2X/3X	√	X	Yes	-
	1E/4E	X	√	Yes	-

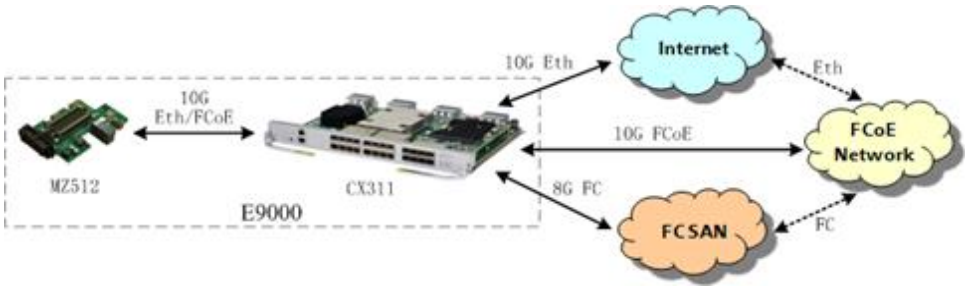
I/O Module Slot	I/O Module Slot	MZ512 (Mezz1)	MZ512 (Mezz2)	Typical Configuration	Remarks
CX317	2X/3X	√	X	No	It is recommended that the CX317 not be installed in slot 2X or 3X.
	1E/4E	X	√	Yes	-
CX910	2X/3X	√	X	Yes	-
	1E/4E	X	√	No	It is recommended that the CX910 not be installed in slot 1E or 4E.

3.3 MZ512 Networking

The MZ512 can connect to I/O modules (switch modules or interface boards) to provide Ethernet and FCoE services.

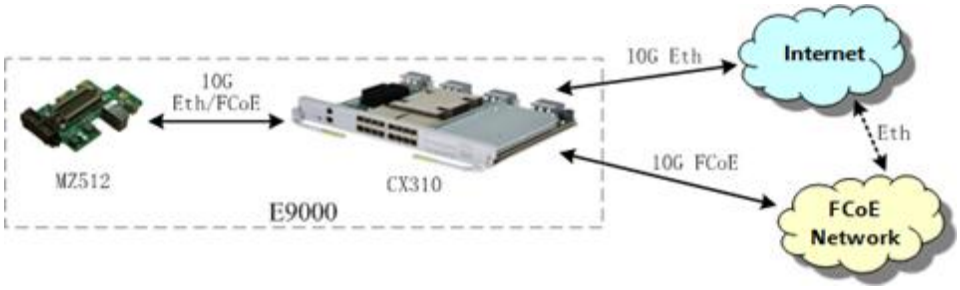
The MZ1512 can work with the CX311 switch module to provide 40 Gbit/s bandwidth, and support the Ethernet service and FCoE storage service. The MZ512 connects to the Internet, FCoE network, and FC SAN through 10GE and FC ports on the CX311 respectively to achieve converged network solutions. See **Figure 3-2**.

Figure 3-2 Connection between the MZ512 and the CX311



The MZ512 can work with the CX310 switch module to provide 40 Gbit/s bandwidth, and support the Ethernet service and FCoE storage service. The MZ512 connects to the Internet and FCoE network through 10GE and 10G FCoE ports on the CX310 respectively to achieve converged network solutions. See **Figure 3-3**.

Figure 3-3 Connection between the MZ512 and the CX310

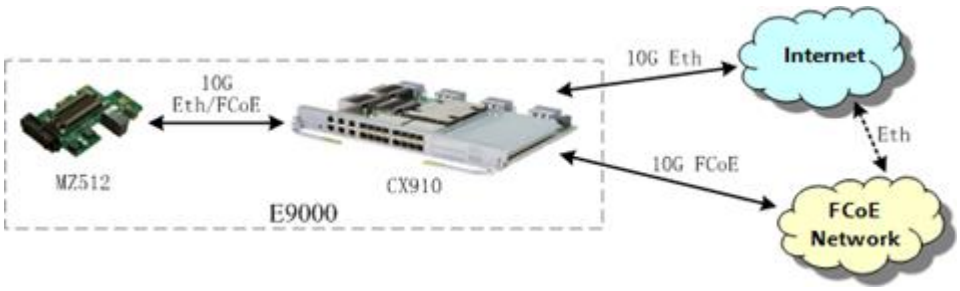


The MZ512 can work with the CX317 pass through module to provide 40 Gbit/s bandwidth, and support the Ethernet service and FCoE storage service. The MZ512 connects to the Internet and FCoE network through 10GE and 10G FCoE ports on the CX317 respectively in pass-through mode to achieve converged network solutions. See **Figure 3-4**.

Figure 3-4 Connection between the MZ512 and the CX317

The MZ512 can work with the CX910 switch module to provide 40 Gbit/s bandwidth, and support the Ethernet service and FCoE storage service. The MZ512 connects to the Internet and FCoE network through 10GE and 10G FCoE ports on the CX910 to achieve converged network solutions. See **Figure 3-5**.

Figure 3-5 Connection between the MZ512 and the CX910



3.4 Supported OSs

Table 3-3 lists the OSs supported by the MZ512.

Table 3-3 OSs supported by the MZ512

OS	Version	Remarks
Redhat	RHEL 6.5	-
	RHEL 6.6	-
	RHEL 6.7	-
	RHEL 7.0	-
	RHEL 7.1	-
	RHEL 7.2	-
Suse	SLES 11.3	-
	SLES 11.4	-
	SLES 12.0	-
	SLES 12.1	-
VMware	VMware ESXi 5.1.3	-
	VMware ESXi 5.5.2	-
	VMware ESXi 5.5.3	-
	VMware ESXi 6.0	-
	VMware ESXi 6.0.1	-
	VMware ESXi 6.0.2	-
Windows	Windows 2008 R2 SP1	-
	Windows 2012	-
	Windows 2012 R2	-

For the latest versions of OSs, see the ***Huawei Server Compatibility Checker***.

3.5 Connected Devices

Table 3-4 lists the devices to which the MZ512 can connect.

Table 3-4 Devices to which the MZ512 can connect

Category	Vendor	Model	Remarks
FC storage array	EMC	Symmetrix DMX-4	-
		Symmetrix VMAX	-

Category	Vendor	Model	Remarks
		Symmetrix VMAXe	-
		VNX	-
	Huawei	HVS85T	-
		S2200T	-
		S2600T	-
		S5500T	-
		S3900	-
		S5600T	-
		S5800T	-
		S5900	-
		S6800T	-
		S6900	-
	Futijsu	DX80	-
	Synology	DS3611xs	-
		DS3612xs	-
		RS3411RPxs	-
		RS3411xs	-
		RS3412RPxs	-
		RS3412xs	-
FCoE switch	Cisco	Nexus 2232PP Fabric Extender	-
		Nexus 5010	-
		Nexus 5020	-
		Nexus 5548P	-
		Nexus 5548UP	-
		Nexus 5596UP	-
	Juniper	QFX3000 QFabric	-
		QFX3500	-
	Huawei	CX311	The CX311 is a converged switch module on the E9000 and supports the FCF function.

Category	Vendor	Model	Remarks
	Brocade	Brocade 8000	-
		VDX6730	-
FC switch	Brocade	Brocade 300	The MZ512 connects to Brocade 300 in NPV mode through the CX311 on the E9000.
		Brocade 5100	The MZ512 connects to Brocade 5100 in NPV mode through the CX311 on the E9000.
		Brocade 5300	The MZ512 connects to Brocade 5300 in NPV mode through the CX311 on the E9000.
		Brocade 8510	The MZ512 connects to Brocade 8510 in NPV mode through the CX311 on the E9000.
	Huawei	SNS2124	OEM Brocade 300 The MZ512 connects to SNS2124 in NPV mode through the CX311 on the E9000.
		SNS2248	OEM Brocade 6510 The MZ512 connects to SNS2248 in NPV mode through the CX311 on the E9000.
	Cisco	MDS 9148	The MZ512 connects to MDS 9148 in NPV mode through the CX311 on the E9000.
		MDS 9513	The MZ512 connects to MDS 9513 in NPV mode through the CX311 on the E9000.
		MDS 9505	The MZ512 connects to MDS 9505 in NPV mode through the CX311 on the E9000.
		MDS 9222i	The MZ512 connects to MDS 9222i in NPV mode through the CX311 on the E9000.

4 Technical Specification

About This Chapter

4.1 Technical Specifications

Table 4-1 lists the technical specifications for the MZ512.

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	0.26 kg (0.57 lb)
Maximum power consumption	24 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 MZ512 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 40 Gbit/s (2 x PCIe 2.0 x8)
Port rate	10.3125 Gbit/s
Number of ports	4
Port type	Ethernet and FCoE
Chip model/ manufacturer	BladeEngine3 (BE3)/Emulex

A Acronyms and Abbreviations

C	
CNA	converged network adapter
CoS	class of service
D	
DCB	Data Center Bridging
DCBX	Data Center Bridging Exchange
E	
ETS	Enhanced Transmission Selection
EVB	Edge Virtual Bridging
F	
FC	Fibre Channel
FCF	FCoE Forwarder
FCoE	Fibre Channel over Ethernet
I	
iSCSI	Internet Small Computer System Interface
L	
LACP	Link Aggregation Control Protocol
LLDP	Link Layer Discovery Protocol
LRO	Large Receive Offload
LSO	Large Segmentation Offload

N	
NIC	network interface card
NPIV	N_Port_ID Virtualization
O	
OS	operating system
P	
PCI-SIG	Peripheral Component Interconnect Special Interest Group
PCIe	Peripheral Component Interconnect Express
PF	physical function
PFC	Priority-based Flow Control
PG	priority group
PVID	port default VLAN ID
PXE	Preboot Execution Environment
R	
RSS	Receive Side Scaling
S	
SAN	storage area network
SR-IOV	Single Root I/O Virtualization
T	
TCP	Transmission Control Protocol
U	
UDP	User Datagram Protocol
UMC	Universal Multi Channel
V	
VEB	Virtual Ethernet Bridging
VF	virtual function
VLAN	virtual local area network

VM	virtual machine
VMQ	virtual machine queue