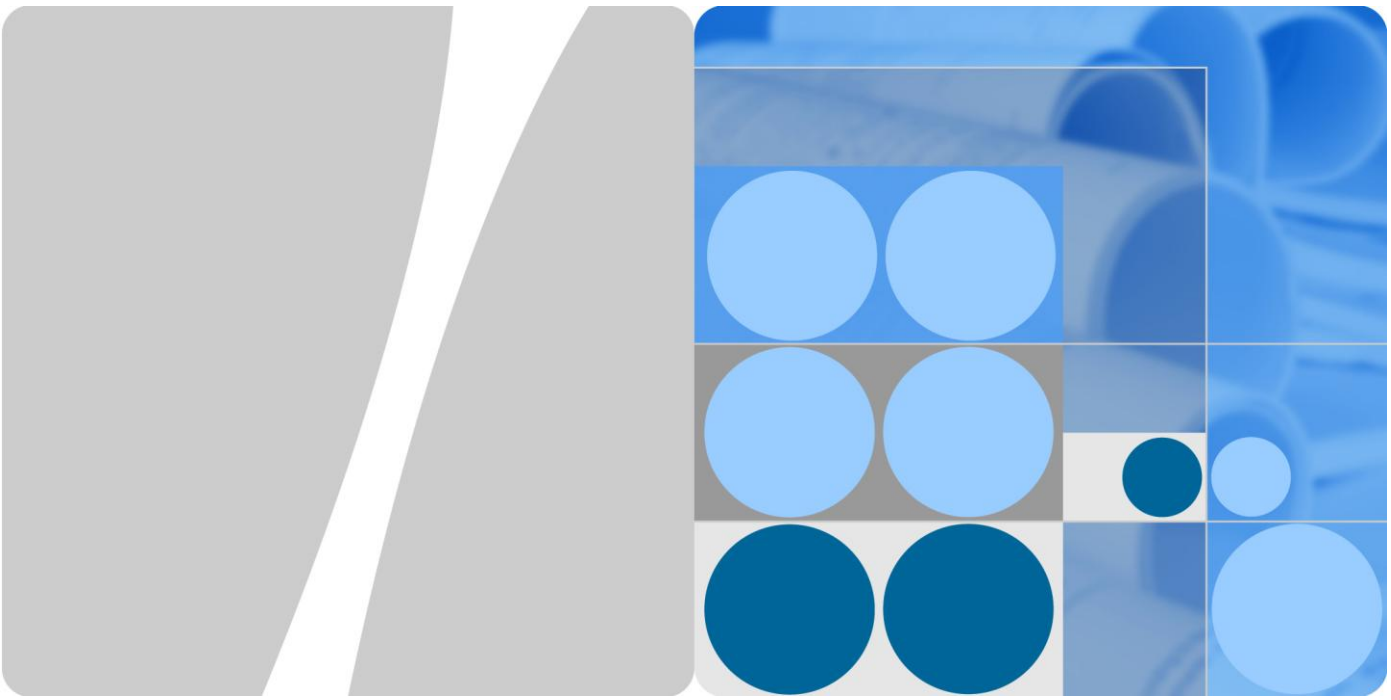


Huawei MZ520 NIC V100R001

White Paper

Issue 02

Date 2017-08-24

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://www.huawei.com

Email: support@huawei.com

About This Document

Purpose

This document describes the MZ520 in terms of its functions, appearance, features, applications, and technical specifications. You can obtain comprehensive information about the MZ520 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 02 (2017-08-24)

This issue is the second official release.

Issue 01 (2017-08-15)

This issue is the first official release.

Contents

1 Overview

1.1 Functions

The MZ520 is a 10GE converged network adapter (CNA). It is used for E9000 compute nodes and provide two ports for compute nodes to connect to switch module slots in the chassis.

The MZ520 uses the BCM57810S chip provided by Cavium, and supports NIC, Fibre Channel over Ethernet (FCoE), and Internet Small Computer System Interface (iSCSI) applications to achieve converged network solutions.

1.2 Appearance

The MZ520 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 MZ520 provides network ports for connecting to switch modules.

- When the MZ520 is installed in slot Mezz1 or Mezz3, its two 10GE ports connect to switch modules in slots 2X and 3X.
- When the MZ520 is installed in slot Mezz2 or Mezz4, its two 10GE ports connect to switch modules in slots 1E and 4E.

Figure 1-1 MZ520 appearance

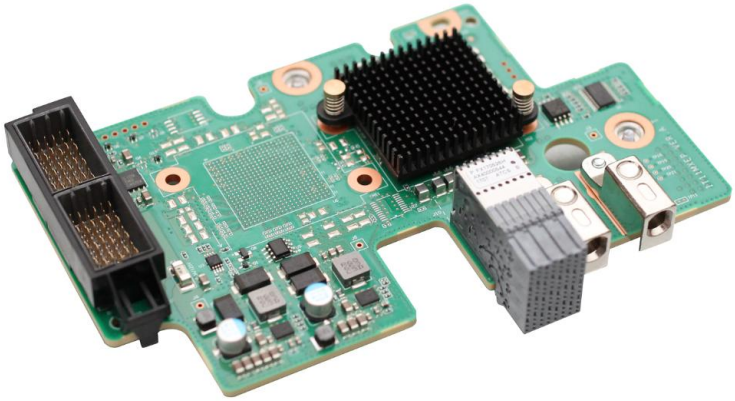


Figure 1-2 MZ520 installation positions on a half-width compute node

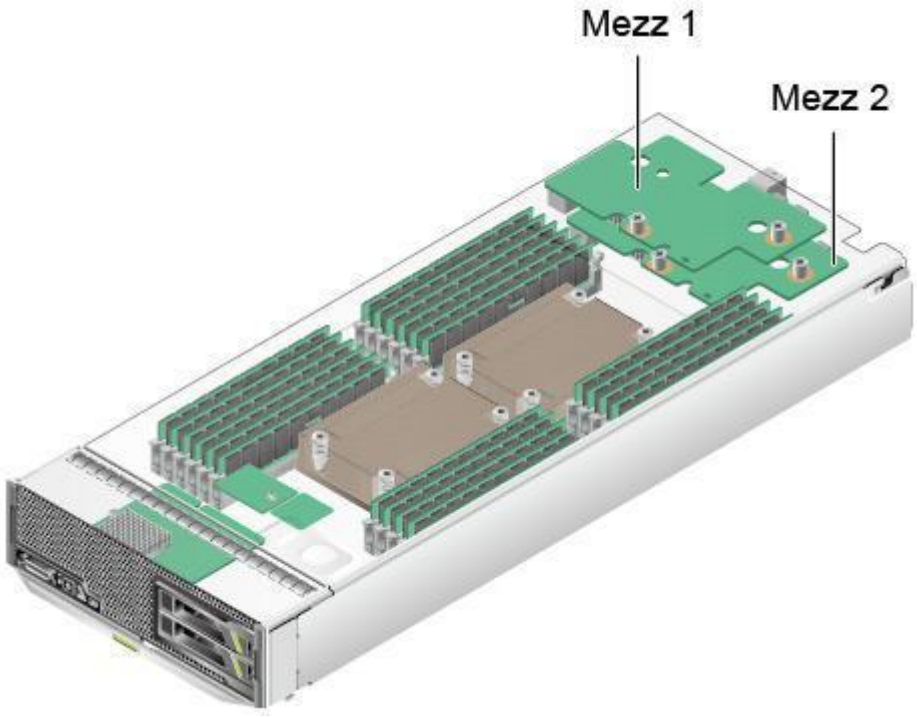
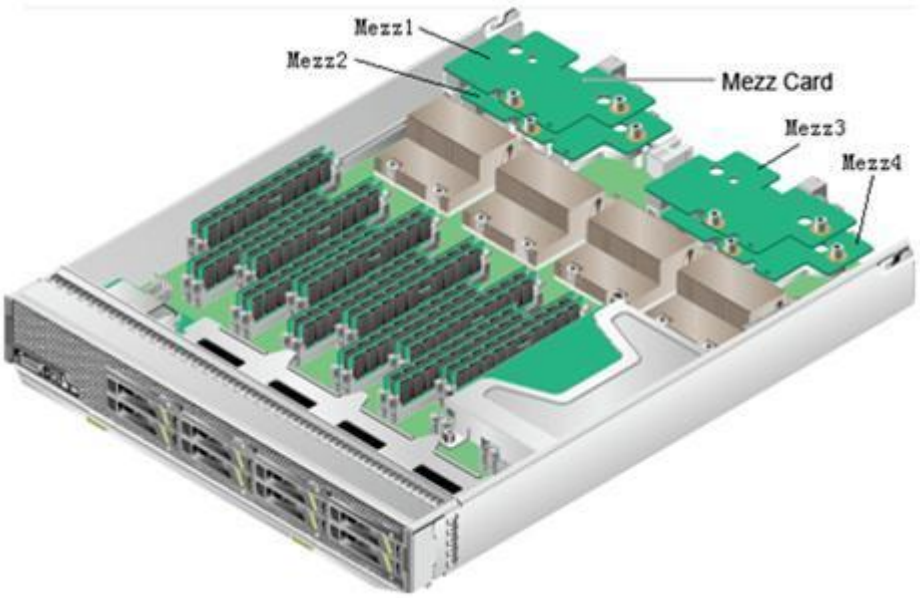


Figure 1-3 MZ520 installation positions on a full-width compute node



2 Features

2.1 Feature List

The MZ520 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
- NIC partitioning (NPAR), 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 128 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, TSS, Receive Side Scaling (RSS), Large Receive Offload (LRO), and Large Send Offload (LSO)
- 802.1Q VLAN, supporting a maximum of 4094 VLANs
- VxLAN offload
- Data Plane Development Kit (DPDK)
- Jumbo frames of 9 KB

2.2 Feature Description

NPAR

The MZ520 supports NPAR, that is, multi-PF. Each physical 10GE port can be divided into four PFs (logical channels). Each NIC supports a total of eight PFs.

The four PFs derived from each physical 10GE port can be configured in NIC mode (Ethernet logical channels), NIC + FCoE mode (Ethernet + FCoE converged channels), or NIC + iSCSI mode (Ethernet + iSCSI converged channels). FCoE and iSCSI are both storage functions. Among the PFs of a 10GE port, at most two PFs can be used for storage functions, and only one PF can be used for FCoE. In NPAR mode, a 10GE physical port supports the following PF combinations:

NIC, NIC, NIC, NIC

NIC, NIC, NIC, NIC+FCoE

NIC, NIC, NIC, NIC+iSCSI

NIC, NIC, NIC+FCoE, NIC+iSCSI

NIC, NIC, NIC+iSCSI, NIC+iSCSI

Each PF must be configured with channel bandwidth. The four PFs on a physical 10GE port share the bandwidth of the 10GE port.

The NPAR feature leverages the PCIe function defined in PCI-SIG specifications, independent of the virtualization function, operating system (OS), or Hypervisor.

PFC

The 10GE ports on the MZ520 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 MZ520 and its connected peer device support PFC, the link between them supports PFC. The MZ520 can identify PFC frames sent by the peer device and adjusts the transmitting bandwidth. If the inbound traffic exceeds the threshold, the MZ520 sends PFC frames to instruct the peer device to control the traffic.

The MZ520 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).

ETS

The 10GE ports on the MZ520 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 MZ520 supports a maximum of 4094 VLANs. When NPAR is disabled, each physical 10GE port supports 4094 VLANs. When NPAR is enabled, each PF supports 4094 VLANs. The VLAN IDs are integers ranging from 1 to 4094.

The MZ520 does not tag or untag outbound packets, but transparently transmits them. VLAN IDs are specified by the OS or Hypervisor. The MZ520 also does not tag or untag inbound packets, but transparently transmits them to the upper layer (OS or Hypervisor).

SR-IOV

The MZ520 supports the SR-IOV feature. The NIC supports a maximum of two physical functions (10GE physical ports) and 128 VFs. Each PF supports a maximum of 64 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.

NPIV

The MZ520 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 MZ520 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 MZ520 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)

Standard	Protocol
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

3.1 Compatible Compute Nodes

The MZ520 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 MZ520 and its installation positions on them.

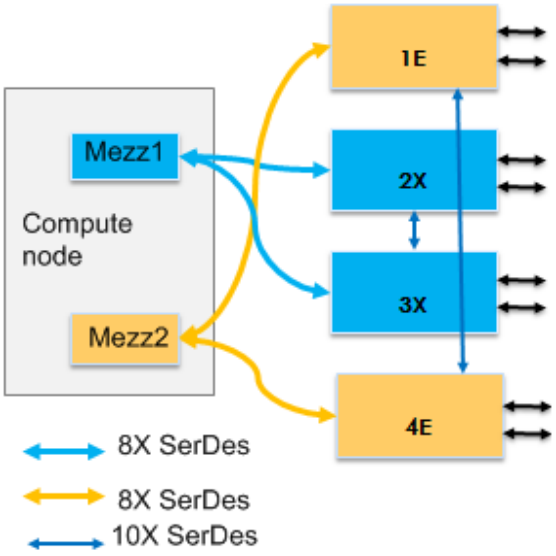
Table 3-1 Compute nodes that support the MZ520

Compute Node	Number of Mezzanine Card Slots	MZ520 Installation Position
CH121 V3	2	Mezz1 and Mezz2
CH220 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH222 V3	2	Mezz1 and Mezz2
CH242 V3 DDR4	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH225 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH226 V3	2	Mezz2 and Mezz3

3.2 Connected I/O Modules

The MZ520 can connect to I/O modules (switch modules or pass-through modules). Figure 3-1 shows the connections between the MZ520 NICs and the I/O modules.

Figure 3-1 Connections between the MZ520 NICs and the I/O modules



There are Serializer/Deserializer (SerDes) high-speed interconnection lanes between each compute node and I/O module slots.

- Mezz1: provides each 2X and 3X slot with eight SerDes lanes.
- Mezz2: provides each 1E and 4E slot with eight SerDes lanes.
- Mezz3 (available only on a full-width compute node): provides each 2X and 3X slot with eight SerDes lanes.
- Mezz4 (available only on a full-width compute node): provides each 1E and 4E slot with eight SerDes lanes.

NOTE

The MZ520 provides two ports (connect to each of the corresponding two I/O modules through one ports). Each port uses only one of the eight SerDes lanes.

Table 3-2 describes the I/O modules to which the MZ520 can connect.

Table 3-2 I/O modules to which the MZ520 can connect

I/O Module Slot	I/O Module Slot	MZ520 (Mezz1)	MZ520 (Mezz2)	Typical Configuration	Remarks
CX310	2X/3X	√	X	Yes	-
	1E/4E	X	√	Yes	-
CX311	2X/3X	√	X	Yes	-
	1E/4E	X	√	Yes	-
CX320	2X/3X	√	X	Yes	-
	1E/4E	X	√	Yes	-
CX317	2X/3X	√	X	No	It is recommended that

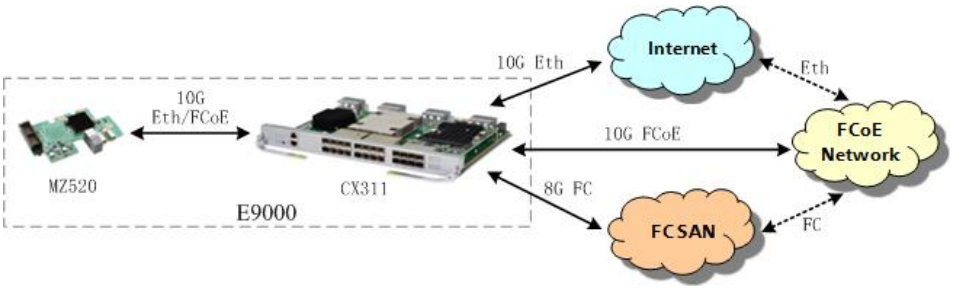
I/O Module Slot	I/O Module Slot	MZ520 (Mezz1)	MZ520 (Mezz2)	Typical Configuration	Remarks
					the CX317 not be installed in slot 2X or 3X.
	1E/4E	X	√	Yes	-
CX318	2X/3X	√	X	No	It is recommended that the CX318 not be installed in slot 2X or 3X.
	1E/4E	X	√	Yes	-

3.3 MZ520 Networking

The MZ520 can connect to I/O modules (switch modules or interface boards) to provide Ethernet and FCoE services.

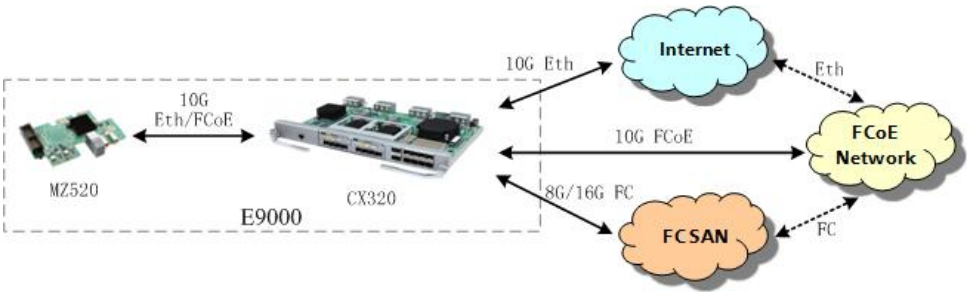
The MZ520 can work with the CX311 switch module to provide 20 Gbit/s bandwidth, and supports the Ethernet service and FCoE storage service. The MZ520 connects to the Internet, FCoE network, and FC SAN through 10GE and FC ports on the CX311 respectively to achieve converged network solutions. See Figure 3-2.

Figure 3-2 Connection between the MZ520 and the CX311



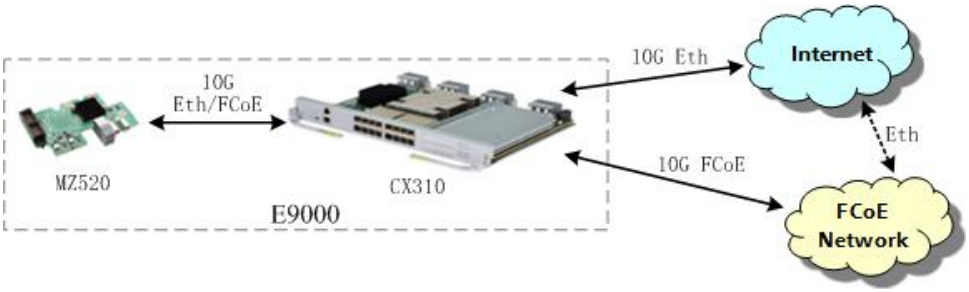
The MZ520 can work with the CX320 switch module to provide 20 Gbit/s bandwidth, and supports the Ethernet service and FCoE storage service. The MZ520 connects to the Internet, FCoE network, or FC SAN through the CX320 panel ports to achieve converged network solutions. See Figure 3-3.

Figure 3-3 Connection between the MZ520 and the CX320



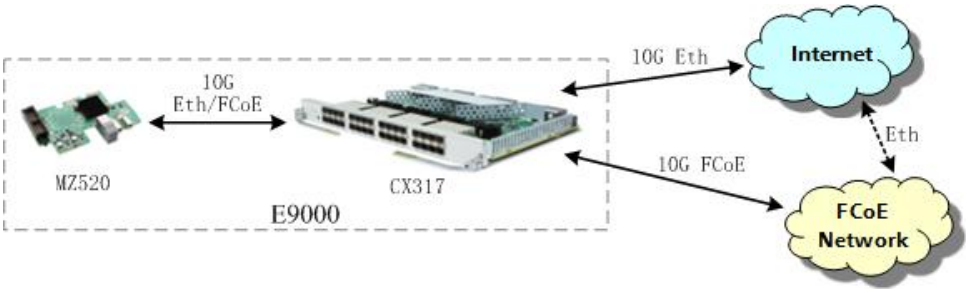
The MZ520 can work with the CX310 switch module to provide 20 Gbit/s bandwidth, and supports the Ethernet service and FCoE storage service. The MZ520 connects to the Internet and FCoE network through 10GE and 10G FCoE ports on the CX310 respectively to achieve converged network solutions. See Figure 3-4.

Figure 3-4 Connection between the MZ520 and the CX310



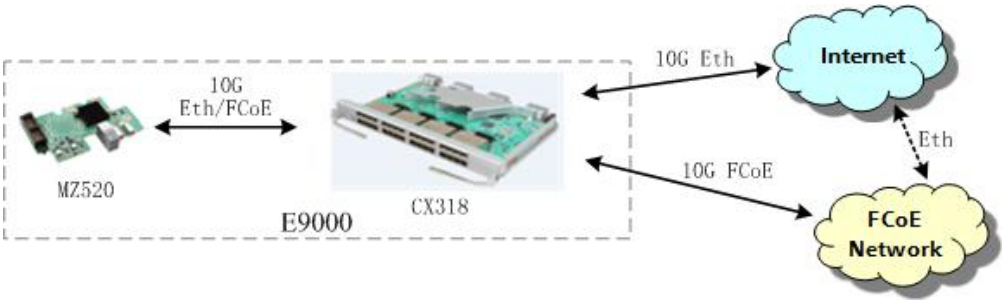
The MZ520 can work with the CX317 pass-through module to provide 20 Gbit/s bandwidth, and supports the Ethernet service and FCoE storage service. The MZ520 connects to the Internet and FCoE network through 10GE and 10G FCoE ports on the CX317 respectively in pass-through mode to achieve converged network solutions. See Figure 3-5.

Figure 3-5 Connection between the MZ520 and the CX317



The MZ520 can work with the CX318pass-through module to provide 20 Gbit/s bandwidth, and supports the Ethernet service and FCoE storage service. The MZ520 connects to the Internet and FCoE network (without intra-chassis switching) through CX318 panel ports to achieve converged network solutions. See Figure 3-6.

Figure 3-6 Connection between the MZ520 and the CX318



3.4 Supported OSs

Table 3-3 lists the OSs supported by the MZ520.

Table 3-3 OSs supported by the MZ520

OS Typ	Version
RHEL	6.6, 6.7, 6.8, 7.1, 7.2, 7.3
SLES	11.3, 11.4, 12.1, 12.2
VMware ESXi	5.5.2, 5.5.3, 6.0, 6.0.1, 6.0.2, 6.5
Windows	2008 R2, 2012, 2012 R2, 2012 Hyper-V, 2012 R2 Hyper-V, 2016

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

4 Technical Specification

4.1 Technical Specifications

Table 4-1 lists the technical specifications for the MZ520.

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	0.14 kg (0.31 lb)
Maximum power consumption	12 W
Temperature	Operating temperature: 5 ℃ to 40 ℃ (41 ℉ to 104 ℉) (ASHRAE Class A3 compliant)
	Storage temperature: –40 ℃ to +65 ℃ (–40 ℉ to +149 ℉)
Humidity	Operating humidity: 5% RH to 85% RH (non-condensing)
	Storage humidity: 5% RH to 95% RH (non-condensing)
Altitude	40 ℃ (104 ℉) at 900 m (2952.76 ft) When the MZ520 is used at an altitude between 900 m and 3000 m, the highest operating temperature decreases by 1 ℃ (1.8 ℉) as the altitude increases by 300 m (984.25 ft).
PCIe port bandwidth	40 Gbit/s (PCIe 2.0 x8)
Port rate	10.3125 Gbit/s
Number of ports	2

Item	Specifications
Port type	Ethernet and FCoE
Chip model/manufacturer	BCM57810S/Cavium

A Acronyms and Abbreviations

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	Fiber 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
NPAR	network interface card partitioning
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
TSS	Transmit Side Scaling

U	
UDP	User Datagram Protocol
V	
VEB	Virtual Ethernet Bridging
VF	virtual function
VLAN	virtual local area network
VM	virtual machine
VMQ	virtual machine queue