Huawei MZ821 NIC
V100R001

White Paper

Issue	01
Date	2018-02-05

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 appearance, features, specifications, and applications of the MZ821.

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 (2018-02-05)

This issue is the first official release.

Contents

1 Product Introduction

1.1 Overview

The MZ821 is a single-port Intel® Omni-Path Fabric (OPF) module and can be installed on an E9000 compute node, providing a network port for connecting to a switch module.

The MZ821 adopts the Intel® Wolf River (WFR) chip and provides a 100G Omni-Path Host Fabric Interface (HFI) port to connect to the Omni-Path Architecture (OPA) switching network.

It is a key module of the Intel® Scalable System Framework and applies to high-performance computing (HPC) and other solutions that require high performance, scalability, and availability. With advanced on-load design, the fabric performance increases with the number of server nodes to support growing workloads.

1.2 Appearance

The MZ821 can be installed in slot Mezzanine 2 (Mezz 2 for short) on a half-width compute node or in slots Mezz 2 and Mezz 4 on a full-width compute node. Its 100G physical port connects only to switch module slot 1E rather than slot 4E.
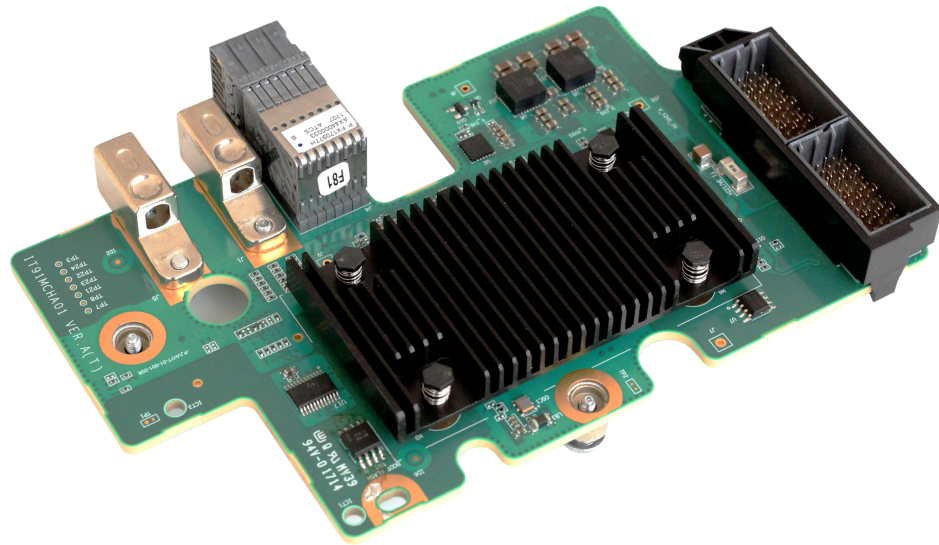
Figure 1-1 MZ821

Figure 1-2 Installation position on a half-width compute node
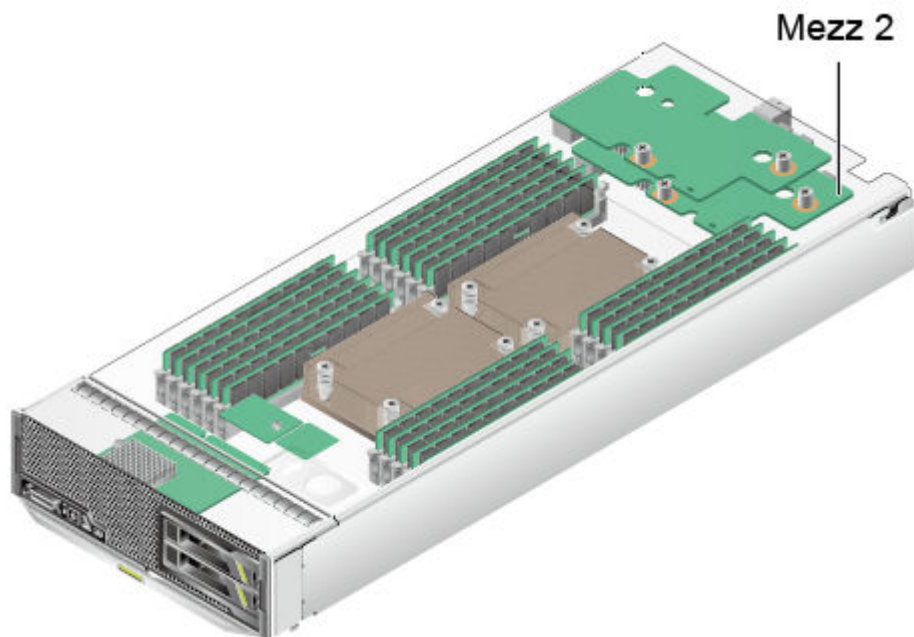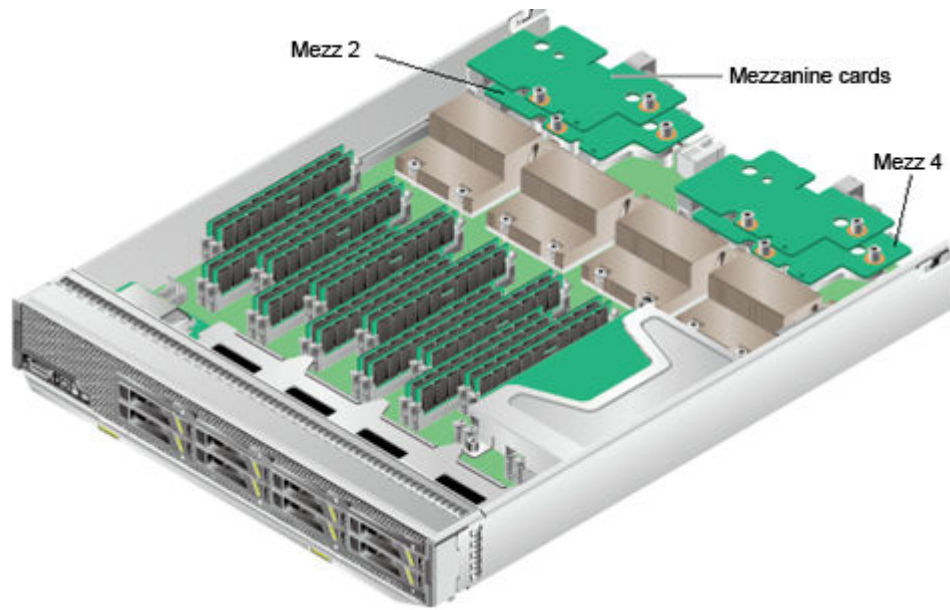
Figure 1-3 Installation positions on a full-width compute node

2 Features

2.1 Performance Specifications

The MZ821 supports the following features and performance specifications:

- Multi-core scaling and up to 160 contexts
- Remote Direct Memory Access (RDMA) feature to support low-latency network applications
- 16 Send DMA engines (M2IO function); Receive DMA engine arrival notifications
- Large MTU (4 KB, 8 KB, or 10 KB) for reduced per-packet processing overheads and improved resource utilization
- High-speed and low-latency MPI interface
- Each HFI can map up to 128 GB window at the granularity of 64 bytes.
- Up to nine virtual lanes for differentiated Quality of Service (QoS)
- 160M messages and 300M bidirectional messages per second
- Congestion control
- Inband management

2.2 Feature Description

QoS

The MZ821 provides powerful traffic control. 65-bit Flow Control Digits (FLITs) are added to data and assembled into much larger Link Transfer Packets (LTPs) for efficient transfer. Through traffic management at the FLIT level, OPA switch modules make extremely granular switching decisions to efficiently optimize the latency and data transfer mechanism.

On-Load Design

With on-load design, the MZ821 eliminates the need for data path firmware and external memory, while maintaining all fabric network connection state information in host memory. This reduces the possibility of data errors and makes the fabric more resilient to NIC and

fabric network faults. Additional protection against errors and downtime is provided by ECC protection on all host SRAMs and parity checking on all host buses.

3 Applications

3.1 Compatible Compute Nodes

The MZ821 can be installed in slot Mezz 2 on a half-width compute node or in slots Mezz 2 and Mezz 4 on a full-width compute node.

Table 3-1 Compute nodes that support the MZ821

Compute Node	MZ821 Installation Position
CH121 V5	Mezz 2

3.2 Compatible I/O Modules

Table 3-2 lists the I/O modules (switch modules) that can connect to the MZ821.

Table 3-2 Compatible I/O modules

I/O Module	I/O Module Slot	MZ821 (Mezz 1)	MZ821 (Mezz 2)	Typical Configuration
CX820	1E	x	√	-

3.3 Networking

The MZ821 can connect to an I/O module (switch module) to provide OPA services.

The MZ821 can work with the CX820 switch module to provide 100 Gbit/s interface bandwidth, and connect to the external OPA network through panel ports on the CX820.

Figure 3-1 Connection between the MZ821 and the CX820

3.4 Supported OSs

Table 3-3 lists the OSs supported by the MZ821.

Table 3-3 OSs supported by the MZ821

OS Type	Version
CentOS	7.2 x86_64, 7.3 x86_64
RHEL	7.2 x86_64, 7.3 x86_64
SLES	12 SP1 x86_64, 12 SP2 x86_64

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

4 Technical Specifications

4.1 Technical Specifications

Table 4-1 describes the technical specifications of the MZ821.

Table 4-1 Technical specifications of the MZ821

Item	Value
Dimensions (length x width)	148 mm x 85 mm (5.83 in. x 3.35 in.)
Power supply	12 V DC
Net weight	0.166 kg (0.366 lb)
Maximum power consumption	16 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 < 20°C (36°F)/hour
Humidity	● Operating humidity: 5% RH to 85% RH (non-condensing) ● Storage humidity: 5% RH to 95% RH (non-condensing) ● Humidity change rate < 20% RH/hour
Altitude	40°C (104°F) at 900 m (2952.76 ft) When the equipment is used at an altitude from 900 m (2952.76 ft) to 3000 m (9842.52 ft), the highest operating temperature decreases by 1°C (1.8°F) for every increase of 300 m (984.25 ft).
PCIe port bandwidth	128 Gbit/s (PCIe 3.0 x16)
Port rate	100 Gbit/s

Item	Value
Number of ports	1
Port type	OPA
Chip model/manufacturer	100HFS02BSF/Intel

4.2 Standards Compliance

The MZ821 adopts Intel's OPA architecture. The port connected to the server complies with the PCIe 3.0 standard.

The port connected to a switch module is an HFI port and complies with Intel's proprietary protocol.

A Acronyms and Abbreviations

Acronyms and Abbreviations

D	
DMA	direct memory access
E	
ECC	error checking and correction
F	
FLIT	flow control digit
H	
HFI	Host Fabric Interface
HPC	high-performance computing
I	
IO	input/output
L	
LTP	link transfer packet
M	
MPI	Message Passing Interface
O	
OS	operating system
P	
PCIe	Peripheral Component Interconnect Express
Q	
QoS	Quality of Service

R	
RDMA	Remote Direct Memory Access
S	
SRAM	static random access memory
V	
VL	virtual lane