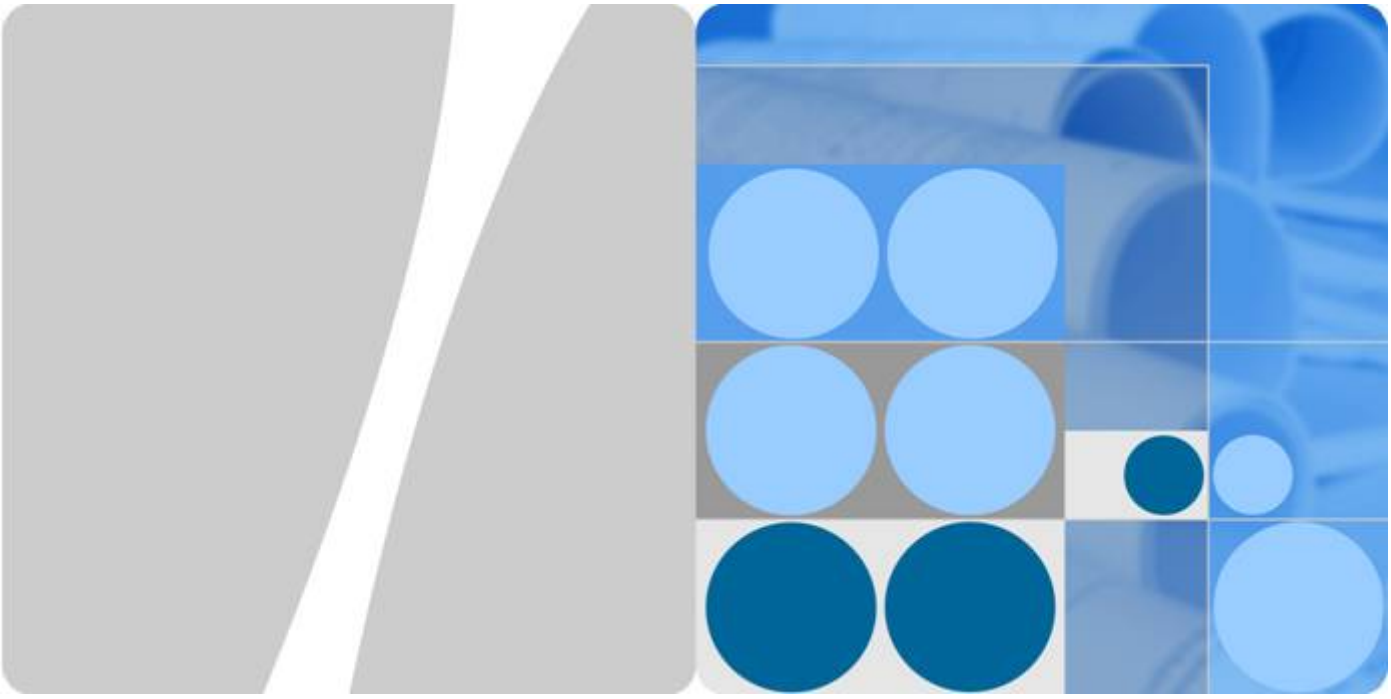


Huawei MZ312 NIC

V100R001

White Paper

Issue 07

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 MZ312 in terms of its functions, appearance, features, applications, and technical specifications. You can obtain comprehensive information about the MZ312 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 07 (2016-11-21)

This issue is the seventh official release.

Issue 06 (2016-08-17)

This issue is the sixth official release.

Issue 05 (2016-05-15)

This issue is the fifth official release.

Type	Change Description
Added	The CX318 is added .

Issue 04 (2016-03-31)

This issue is the fourth official release.

Issue 03 (2015-10-19)

This issue is the third official release.

Issue 02 (2015-05-30)

This issue is the second official release.

Issue 01 (2015-03-20)

This issue is the first official release.

Contents

Figures

Tables

1 Overview

About This Chapter

1.1 Functions

The MZ312 is a dual-chip 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 MZ312 uses two Intel 82599 chips and supports NIC applications.

1.2 Appearance

The MZ312 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 MZ312 provides network ports for connecting to switch modules:

- When the MZ312 is installed in slot Mezz1 or Mezz3, its four 10GE ports connect to switch modules in slots 2X and 3X.
- When the MZ312 is installed in slot Mezz2 or Mezz4, its four 10GE ports connect to switch modules in slots 1E and 4E.

Figure 1-1 MZ312 appearance

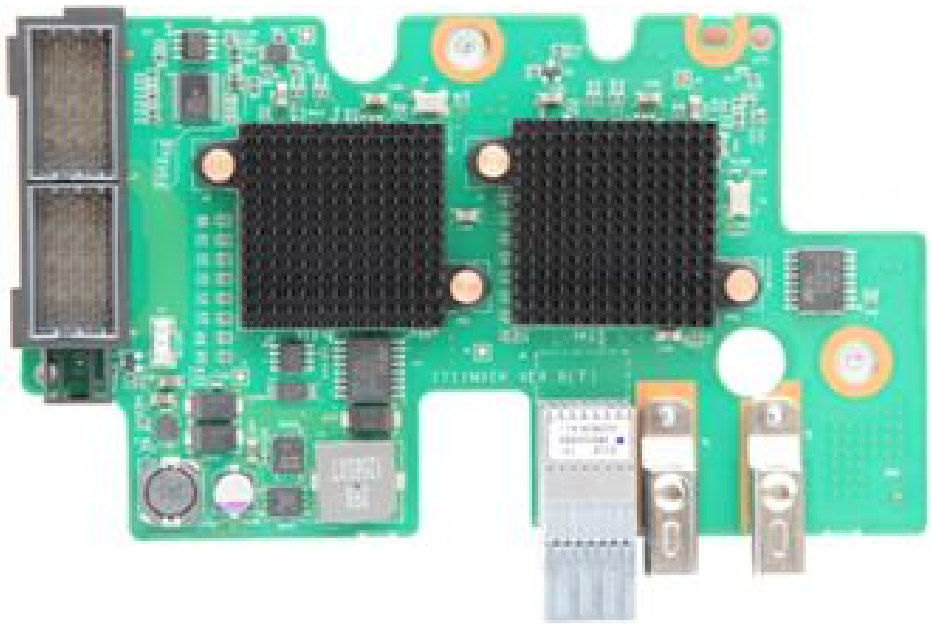


Figure 1-2 MZ312 installation positions on a half-width compute node

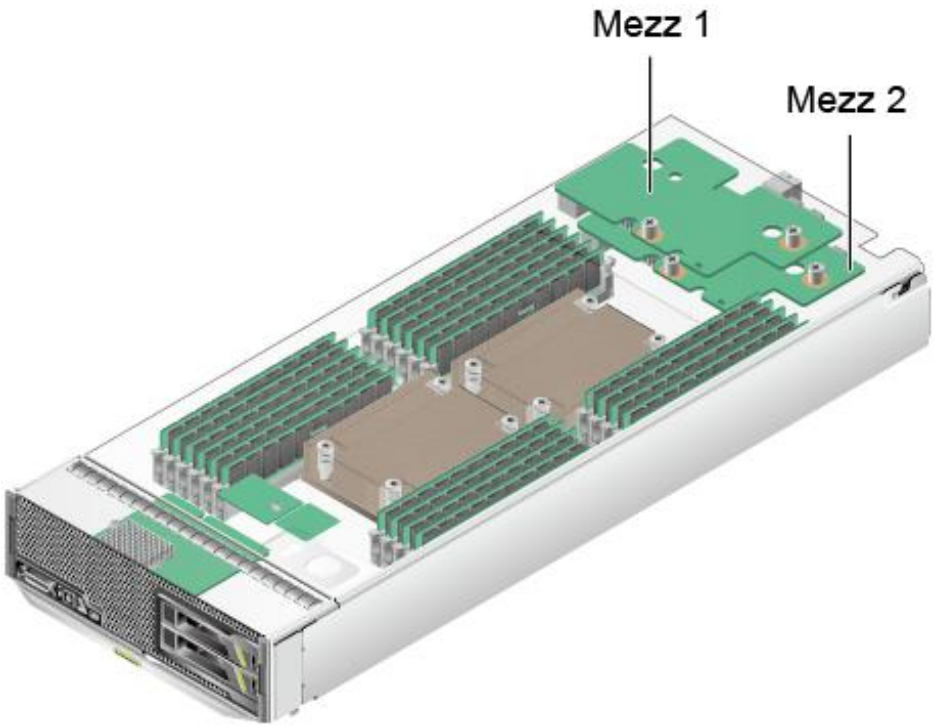
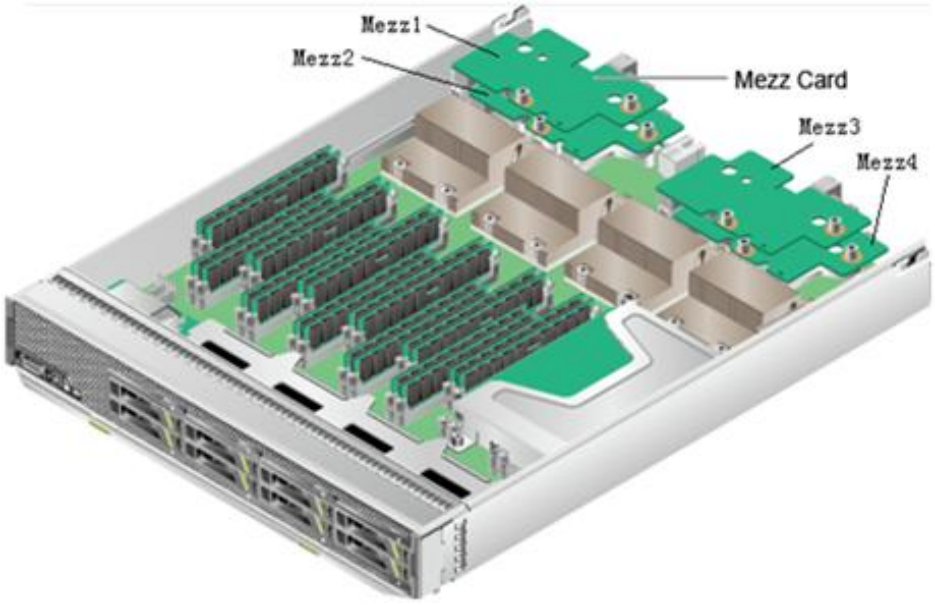


Figure 1-3 MZ312 installation positions on a full-width compute node



2 Features

About This Chapter

2.1 Feature List

The MZ312 supports the following features and performance specifications:

- Priority-based Flow Control (PFC), Enhanced Transmission Selection (ETS), Data Center Bridging (DCB), and Data Center Bridging Exchange (DCBX)
- Single Root I/O Virtualization (SR-IOV), supporting a maximum of 256 (2 x 128) virtual functions (VFs)
- Virtual Machine Device Queue (VMDq)
- Ethernet Preboot Execution Environment (PXE)
- 802.1Q VLAN, supporting a maximum of 4094 VLANs
- Jumbo frames of 9 KB
- Data Plane Development Kit (DPDK)

2.2 Feature Description

PFC

The 10GE ports on the MZ312 support the PFC feature. This feature supports flow control and back pressure on a per class-of-service (CoS) basis. The MZ312 supports a maximum of eight CoS queues. If both the MZ312 and its connected peer device support PFC, the link between them supports PFC. The MZ312 can identify PFC frames sent by the peer device and adjusts the transmitting bandwidth. If the inbound traffic exceeds the threshold, the MZ312 sends PFC frames to instruct the peer device to control the traffic.

The MZ312 can negotiate with the peer device about PFC parameters over DCBX.

ETS

The 10GE ports on the MZ312 support the ETS feature. This feature assigns port transmitting bandwidth based on priority groups (PGs). The MZ312 supports a maximum of eight CoS queues and eight PGs (BWGs). 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 MZ312 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 MZ312 does not tag or untag packets, but transparently transmits them. VLAN IDs are specified by the OS on an E9000 compute node.

SR-IOV

The 10GE ports on the MZ312 support the SR-IOV feature. The NIC supports a maximum of 256 (2 x 128) VFs or four (2 x 2) physical functions (PFs) and 252 (2 x 126) 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.

PXE

The 10GE ports on the MZ312 support PXE.

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.

2.3 Standards Compliance

Table 2-1 lists the standards and protocols that the MZ312 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 802.1Q	VLAN Tagging
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)
DCBX	Data Center Bridging Exchange

3 Applications

About This Chapter

3.1 Compatible Compute Nodes

The MZ312 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 MZ312 and its installation positions on them.

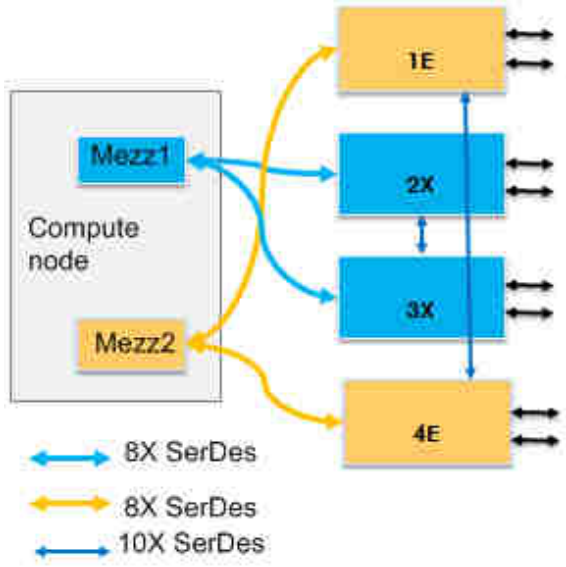
Table 3-1 Compute nodes that support the MZ312

Compute Node	Number of Mezz Module Slots	MZ310 Installation Position
CH121	2	Mezz1 and Mezz2
CH121 V3	2	Mezz1 and Mezz2
CH121H V3	2	Mezz1 and Mezz2
CH140 V3	2	Mezz1 and Mezz2
CH220 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH222 V3	2	Mezz1 and Mezz2
CH225 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH226 V3	2	Mezz1 and Mezz2
CH242 V3	4	Mezz1, Mezz2, Mezz3, and Mezz4
CH242 V3 DDR4	4	Mezz1, Mezz2, Mezz3, and Mezz4

3.2 Connected I/O Modules

MZ312s can connect to I/O modules (switch modules or interface boards). **Figure 3-1** shows the connections between the MZ312s on a half-width compute node and the ports on I/O modules.

Figure 3-1 Connections between the MZ312s 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 MZ312 provides four ports, and only 2X of each 8X SerDes is used.

Table 3-2 describes the I/O modules to which the MZ312 can connect.

Table 3-2 I/O modules to which the MZ312 can connect

I/O Module	I/O Module Slot	MZ312 (Mezz1)	MZ312 (Mezz2)	Typical Configuration	Remarks
CX310	2X/3X	√	X	Yes	-
	1E/4E	X	√	Yes	-
CX910	2X/3X	√	X	Yes	-

I/O Module	I/O Module Slot	MZ312 (Mezz1)	MZ312 (Mezz2)	Typical Configuration	Remarks
	1E/4E	X	√	No	It is recommended that the CX910 not be installed in slot 1E or 4E.
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 MZ312 Networking

The MZ312 can connect to I/O modules (switch modules or interface boards) to provide Ethernet services.

The MZ312 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 MZ312 and the CX310

The MZ312 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 MZ312 and the CX910

The MZ312 can work with the CX317 pass through module to provide 40 Gbit/s bandwidth, 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 MZ312 and the CX317

The MZ312 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 MZ312 and the CX318

3.4 Supported OSs

Table 3-3 lists the OSs supported by the MZ312.

Table 3-3 OSs supported by the MZ312

OS	Version	Remarks
Redhat	RHEL 6.5	-
	RHEL 6.6	-
	RHEL 6.7	-
	RHEL 6.8	-
	RHEL 7.0	-
	RHEL 7.1	-
	RHEL 7.2	-

OS	Version	Remarks
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	-

The preceding table is for reference only. Compatible OSs for the MZ312 vary with the compute node type. For details, see the ***Huawei Server Compatibility Checker***.

3.5 Connected DCB Switches

Table 3-4 lists the DCB switches to which the MZ312 can connect.

Table 3-4 DCB switches to which the MZ312 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 MZ312 connects to CE6800 through the CX317 or CX318 on the E9000.
		CE12800	The MZ312 connects to CE12800 through the CX317 or CX318 on the E9000.
	Cisco	Nexus 5548	The MZ312 connects to Nexus 5548 through the CX317 or CX318 on the E9000.

Category	Vendor	Model	Remarks
		Nexus 5596	The MZ312 connects to Nexus 5596 through the CX317 or CX318 on the E9000.
		Nexus 2232PP	Nexus 2232PP is a Cisco Fabric Extender. The MZ312 connects to Nexus 2232PP through the CX317 or CX318 on the E9000.
		Nexus 2248PQ	Nexus 2248PQ is a Cisco Fabric Extender. The MZ312 connects to Nexus 2248PQ through the CX317 or CX318 on the E9000.

4 Technical Specifications

About This Chapter

4.1 Technical Specifications

Table 4-1 lists the technical specifications for the MZ312.

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	0.5 kg (1.10 lb)
Maximum power consumption	17 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 MZ312 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
Chip model/manufacturer	82599/Intel

A Acronyms and Abbreviations

B	
BWG	bandwidth group
C	
CoS	class of service
D	
DCB	Data Center Bridging
DCBX	Data Center Bridging Exchange
DPDK	Data Plane Development Kit
E	
ETS	Enhanced Transmission Selection
L	
LACP	Link Aggregation Control Protocol
LLDP	Link Layer Discovery Protocol
N	
NIC	network interface card
O	
OS	operating system
P	
PCIe	PCI Express

PF	physical function
PFC	Priority-based Flow Control
PG	priority group
PXE	Preboot Execution Environment
R	
RSC	Receive Side Coalescing
RSS	Receive Side Scaling
S	
SR-IOV	Single Root I/O Virtualization
T	
TSO	TCP Segmentation Offload
V	
VF	virtual function
VLAN	virtual local access network
VM	virtual machine
VMDq	Virtual Machine Device Queue