

ConnectX®-4 EN

单/双端口 100 Gb 以太网适配器卡

具有 100Gb/s 以太网连接的 ConnectX-4 EN 网络控制器为高性能计算、Web 2.0、云、数据分析、数据库和存储平台提供最高性能和最灵活的解决方案。

随着应用程序和社交网络所共享和存储的数据呈指数级增长，对高速和高性能计算与存储数据中心的需求也在飙升。

ConnectX®-4 EN 为要求最苛刻的数据中心、公共和私有云、Web2.0 和大数据应用以及存储系统提供卓越的高性能，使当今的企业能够满足数据爆炸产生的需求。

ConnectX-4 EN 提供单端口 100Gb/s 带宽、最低可用延迟和特定硬件卸载这一无与伦比的组合，可满足当今和下一代计算与存储数据中心的需求。

I/O 虚拟化

ConnectX-4 EN SR-IOV 技术为服务器中的虚拟机 (VM) 提供专用适配器资源以及有保证的隔离和保护。ConnectX-4 EN 的 I/O 虚拟化可为数据中心管理员提供更出色的服务器利用率，同时降低成本、功耗和线缆复杂度，从而实现在相同硬件上承载更多虚拟机和更多租户。

叠加 (Overlay) 网络

为了更好地扩展其网络，数据中心运营商经常创建叠加 (Overlay) 网络，通过采用诸如 NVGRE 和 VXLAN 之类的封装格式的逻辑隧道来传输来自各个虚拟机的流量。

虽然这解决了网络可扩展性问题，但隐藏了硬件卸载引擎的 TCP 数据包，增加了主机 CPU 的负载。ConnectX-4 通过提供先进的 NVGRE、VXLAN 和 GENEVE 硬件卸载引擎有效地解决了这一问题，这些引擎封装和解封叠加协议标头，从而支持对封装的流量执行传统卸载。利用 ConnectX-4，数据中心运营商可以在新的网络体系架构中实现本机性能。

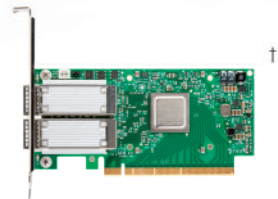
ASAP²™

Mellanox ConnectX-4 EN 提供加速交换和数据包处理 (ASAP²) 技术，用于在 Hypervisor 中执行卸载活动，包括数据路径、数据包解析、VxLAN 和 NVGRE 封装/解封等。

ASAP² 允许通过使用 SR-IOV 在网卡硬件中处理数据层面进行卸载，同时保持在当今基于软件的解决方案中使用的控制层面不被修改。因此，显著提高了性能，而不会产生相关的 CPU 负荷。ASAP² 有两种格式：ASAP² Flex™ 和 ASAP² Direct™。

OpenVSwitch (OVS) 是 ASAP² 可以卸载的虚拟交换机示例之一。

ConnectX®-4



亮点

优势

- 性能最高的硅芯片，面向需要高带宽、低延迟和高消息速率的应用程序
- 世界一流的集群、网络和存储性能
- 基于 x86、Power、ARM 和 GPU 的计算和存储平台的智能互连
- 虚拟化叠加 (Overlay) 网络 VXLAN、NVGRE、GENEVE 方面的领先性能
- 高效的 I/O 整合，降低数据中心的成本和复杂性
- 虚拟化加速
- 功率效率
- 可扩展到数万个节点

主要功能特性

- 每个端口 100Gb/s 以太网
- 1/10/20/25/40/50/56/100Gb/s 速度
- 提供单端口和双端口选项
- 纠删码卸载
- T10-DIF 签名移交
- 加速交换和数据包处理 (ASAP²)
- CPU 传输操作卸载
- 应用程序卸载
- Mellanox PeerDirect™ 通信加速
- 针对 NVGRE、VXLAN 和 GENEVE 封装流量的硬件卸载
- 端到端 QoS 和拥塞控制
- 基于硬件的 I/O 虚拟化
- RoHS-R6

基于融合以太网的 RDMA (RoCE)

ConnectX-4 EN 支持通过以太网网络提供低延迟和高性能的 RoCE 规范。利用数据中心桥接 (DCB) 功能以及 ConnectX-4 EN 高级拥塞控制硬件机制，RoCE 通过第 2 层和第 3 层网络提供高效的低延迟 RDMA 服务。

Mellanox PeerDirect™

PeerDirect™ 通信通过消除 PCIe 总线上各组件之间（例如从 GPU 到 CPU）不必要的内部数据复制来提供高效率的 RDMA 访问，从而大幅缩短应用程序运行时间。ConnectX-4 高级加速技术可实现更高的集群效率，并且可扩展到数万个节点。

存储加速

存储应用程序将通过 ConnectX-4 EN 提供的更高带宽获得性能提升。此外，标准块和文件访问协议可利用 RoCE 进行高性能存储访问。与多架构网络相比，将计算和存储进行整合的网络可实现显著的性能优势。

纠错码卸载

ConnectX-4 EN 提供先进的纠错码卸载功能，支持分布式独立磁盘冗余阵列 (RAID)，这是一种将多个磁盘驱动器组件组合到一个逻辑单元的数据存储技术，旨在实现数据冗余和性能改进。ConnectX-4 EN 的 Reed-Solomon 功能引入了冗余块计算，与 RDMA 一起实现高性能和可靠的存储访问。

签名移交

ConnectX-4 EN 支持 T10 数据完整性字段/保护信息 (T10-DIF/PI) 的硬件检查，可降低 CPU 开销，并加速向应用程序传递数据。由适配器对入口和/或出口数据包进行签名移交处理，可降低启动器和/或目标机器的 CPU 负载。

软件支持

所有 Mellanox 适配器卡均受 Windows、Linux 发行版、VMware、FreeBSD 和 Citrix XENServer 支持。ConnectX-4 EN 适配器支持基于 OpenFabrics 的 RDMA 协议和软件，并与 OEM 和操作系统供应商的配置和管理工具兼容。

兼容性		
PCI EXPRESS 接口	连接	操作系统/发行版*
- 符合 PCIe Gen 3.0 标准，兼容 1.1 和 2.0 2.5、5.0 或 8.0GT/s 链路速率 x16 - 自动协商为 x16、x8、x4、x2 或 x1 - 支持 MSI/MSI-X 机制	- 可与 1/10/25/40/50/100Gb 以太网交换机互操作 - 具有 ESD 保护的无源铜缆 - 用于光纤和有源线缆支持的供电连接器	- RHEL/CentOS - Windows - FreeBSD - VMware - OpenFabrics 企业发行版 (OFED) - OpenFabrics Windows 发行版 (WinOF)

订购部件号	说明	不含支架的尺寸
MCX415A-CCAT	ConnectX-4 EN 网络接口卡，100GbE 单端口 QSFP28，PCIe3.0 x16，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX416A-CCAT	ConnectX-4 EN 网络接口卡，100GbE 双端口 QSFP28，PCIe3.0 x16，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX415A-BCAT	ConnectX-4 EN 网络接口卡，40/56GbE 单端口 QSFP28，PCIe3.0 x16，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX416A-BCAT	ConnectX-4 EN 网络接口卡，40/56GbE 双端口 QSFP28，PCIe3.0 x16，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX415A-GCAT	ConnectX-4 EN 网络接口卡，50GbE 单端口 QSFP28，PCIe3.0 x16，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX416A-GCAT	ConnectX-4 EN 网络接口卡，50GbE 双端口 QSFP28，PCIe3.0 x16，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX413A-BCAT	ConnectX-4 EN 网络接口卡，40/56GbE 单端口 QSFP28，PCIe3.0 x8，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX414A-BCAT	ConnectX-4 EN 网络接口卡，40/56GbE 双端口 QSFP28，PCIe3.0 x8，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX413A-GCAT	ConnectX-4 EN 网络接口卡，50GbE 单端口 QSFP28，PCIe3.0 x8，高支架，ROHS R6	14.2cm x 6.9cm（薄型）
MCX414A-GCAT	ConnectX-4 EN 网络接口卡，50GbE 双端口 QSFP28，PCIe3.0 x8，高支架，ROHS R6	14.2cm x 6.9cm（薄型）

注意：所有高支架适配器在发运时都安装了高支架，并提供一个短支架作为附件。

功能特性摘要*

以太网

- 100GbE/56GbE/50GbE/40GbE/25GbE/10GbE/1GbE
- IEEE 802.3bj、802.3bm 100 Gb 以太网
- 25G 以太网 Consortium 25、50 Gb 以太网
- IEEE 802.3ba 40 Gb 以太网
- IEEE 802.3ae 10 Gb 以太网
- IEEE 802.3az 高能效以太网
- 基于 IEEE 802.3ap 的自动协商和 KR 启动
- 专有以太网协议 (20/40GBASE-R2、50/56GBASE-R4)
- IEEE 802.3ad、802.1AX 链路聚合
- IEEE 802.1Q、802.1P VLAN 标记和优先级
- IEEE 802.1Qau (QCN) – 拥塞通知
- IEEE 802.1Qaz (ETS)
- IEEE 802.1Qbb (PFC)
- IEEE 802.1Qbg
- IEEE 1588v2
- 巨型帧支持 (9.6KB)

增强的功能特性

- 基于硬件的可靠传输
- 集合操作卸载
- 矢量集合操作卸载
- Mellanox PeerDirect™ RDMA (亦称 GPUDirect®) 通信加速
- 64/66 编码
- 扩展可靠连接传输 (XRC)
- 动态连接传输 (DCT)
- 增强的原子操作
- 高级内存映射支持, 允许用户模式内存注册和重新映射 (UMR)
- 按需分页 (ODP) – 免注册 RDMA 内存访问

存储卸载

- RAID 卸载 - 纠错码 (Reed-Solomon) 卸载
- T10 DIF - 针对进口和出口流量的线速签名移交操作

叠加 (Overlay) 网络

- 针对叠加 (Overlay) 网络和隧道协议的无状态卸载
- NVGRE 和 VXLAN 叠加 (Overlay) 网络的封装和解封装的硬件卸载

基于硬件的 I/O 虚拟化

- 单根 IOV
- 每个端口多个功能
- 地址转换和保护
- 每个虚拟机多个队列
- 针对 vNIC 的增强型 QoS
- VMware NetQueue 支持

虚拟化

- SR-IOV : 最多 256 个虚拟功能
- SR-IOV : 每端口最多 16 个物理功能
- 虚拟化层次结构 (例如 NPAR)
 - » 虚拟化物理端口上的物理功能
 - » 每个物理功能上的 SR-IOV
- 1K 入口和出口 QoS 级别
- 保证虚拟机的 QoS

CPU 卸载

- 基于融合以太网的 RDMA (RoCE)
- TCP/UDP/IP 无状态卸载
- LSO、LRO 校验和卸载
- RSS (可以在封装的数据包上完成)、TSS、HDS、VLAN 插入/剥离、接收流转向

- 智能中断合并

远程启动

- 通过以太网远程启动
- 通过 iSCSI 远程启动
- PXE 和 UEFI

协议支持

- OpenMPI、IBM PE、OSU MPI (MVAPICH/2)、Intel MPI
- Platform MPI、UPC、Open SHMEM
- TCP/UDP、MPLS、VxLAN、NVGRE、GENEVE
- iSER、NFS RDMA、SMB Direct
- uDAPL

管理和控制接口

- NC-SI、MCTP over SMBus 和 MCTP over PCIe - 基板管理控制器接口
- 用于管理 eSwitch 的 SDN 管理接口
- 用于设备控制和配置的 I2C 接口
- 通用 I/O 引脚
- SPI 闪存接口
- JTAG IEEE 1149.1 和 IEEE 1149.6

* 此部分介绍硬件特性和功能。有关提供的功能特性, 请参阅驱动程序版本说明

* 此部分介绍硬件特性和功能。有关提供的功能特性, 请参阅驱动程序和固件版本说明。



北京市朝阳区望京东园七区保利国际广场 T1 15层
电话 : 010-5789 2000
www.mellanox.com

