Mellanox®
TECHNOLOGIES

ConnectX • 6 †

ConnectX®-6 EN 200Gb/s Adapter Card

World's first 200Gb/s Ethernet network adapter card, offering world-leading performance, smart offloads and In-Network Computing, leading to the highest return on investment for Cloud, Web 2.0, Big Data, Storage and Machine Learning applications

ConnectX-6 EN provides two ports of 200Gb/s for Ethernet connectivity, sub-600 ns latency and 200 million messages per second, enabling the highest performance and most flexible solution for the most demanding data center applications.

ConnectX-6 is a groundbreaking addition to the Mellanox ConnectX series of industry-leading adapter cards. In addition to all the existing innovative features of past versions, ConnectX-6 offers a number of enhancements to further improve performance and scalability. ConnectX-6 EN supports 200, 100, 50, 40, 25, and 10Gb/s Ethernet speeds.

Cloud and Web2.0 Environments

Telco, Cloud and Web2.0 customers developing their platforms on Software Defined Network (SDN) environments are leveraging the Virtual Switching capabilities of the Operating Systems on their servers to enable maximum flexibility in the management and routing protocols of their networks.

Open V-Switch (OVS) is an example of a virtual switch that allows Virtual Machines to communicate among themselves and with the outside world. Software-based virtual switches, traditionally residing in the hypervisor, are CPU intensive, affecting system performance and preventing full utilization of available CPU for compute functions.

To address this, ConnectX-6 offers Mellanox Accelerated Switching And Packet Processing (ASAP[2]) Direct technology to offload the vSwitch/vRouter by handling the data plane in the NIC hardware while maintaining the control plane unmodified. As a result, significantly higher

vSwitch/vRouter performance is achieved without the associated CPU load.

The vSwitch/vRouter offload functions supported by ConnectX-5 and ConnectX-6 include encapsulation and de-capsulation of overlay network headers, as well as stateless offloads of inner packets, packet headers re-write (enabling NAT functionality), hairpin, and more.

In addition, ConnectX-6 offers intelligent flexible pipeline capabilities, including programmable flexible parser and flexible match-action tables, which enable hardware offloads for future protocols

Storage Environments

NVMe storage devices are gaining momentum, offering very fast access to storage media. The evolving NVMe over Fabric (NVMeOF) protocol leverages RDMA connectivity to remotely access NVMe storage devices efficiently, while keeping the end-to-end NVMe model at lowest latency. With its NVMeOF target and initiator offloads, ConnectX-6 brings further optimization to NVMeoF, enhancing CPU utilization and scalability.

Security

ConnectX-6 offers a crucial innovation to network security by providing block-level encryption. Data in transit undergoes encryption and decryption as it is stored or retrieved. The encryption/decryption, based on the IEEE XTS-AES standard,
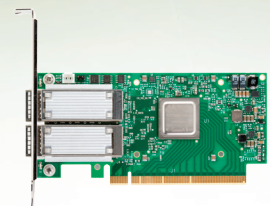
HIGHLIGHTS

NEW FEATURES
– Highest performance
  • *Up to 200Gb/s connectivity per port*
  • *Up to 200 million messages/sec*
  • *Sub 0.6usec latency*
– Block-level XTS-AES Mode Hardware Encryption
– FIPS compliant adapter card
– PCIe Gen3 and PCIe Gen4 support

BENEFITS
– Industry-leading throughput, low latency, low CPU utilization and high message rate
– Highest performance and most intelligent fabric for compute and storage infrastructures
– Advanced storage capabilities including block-level encryption and checksum offloads
– Host Chaining technology for economical rack design
– Smart interconnect for x86, Power, ARM, GPU and FPGA-based compute and storage platforms
– Intelligent network adapter supporting flexible pipeline programmability
– Cutting-edge performance in virtualized networks including Network Function Virtualization (NFV)
– Enabler for efficient service chaining capabilities
– Efficient I/O consolidation, lowering data center costs and complexity

† For illustration only. Actual products may vary.

ConnectX®-6 EN 200Gb/s Adapter Card Single/Dual-Port Adapter Card Supporting up to 200GbE with Advanced Offloading Technology

page 2

is offloaded by the ConnectX-6 hardware, saving latency and offloading CPU. ConnectX-6 block-level encryption offload enables protection between users sharing the same resources, as different encryption keys can be used.

By performing encryption in the adapter, ConnectX-6 also renders encryption unnecessary elsewhere in the network, such as in storage. Moreover, ConnectX-6 supports Federal Information Processing Standards (FIPS) compliance, alleviating the systemic need for self-encrypted disks. With this capability, customers are free to choose their preferred storage device, including byte-addressable and NVDIMMs that otherwise would be used without encryption.

Machine Learning and Big Data Environments

Data analytics has become an essential function within many enterprise data centers, clouds and Hyperscale platforms. Machine learning relies on especially high throughput and low latency to train deep neural networks and to improve recognition and classification accuracy. As the first adapter card to deliver 200Gb/s throughput, ConnectX-6 is the perfect solution to provide machine learning applications with the levels of performance and scalability that they require.

ConnectX-6 utilizes the RDMA technology to deliver low-latency and high performance. ConnectX-6 enhances RDMA network capa-

bilities even further by delivering end-to-end packet level flow control.

Socket Direct

Mellanox ConnectX-6 enables 200Gb/s transmission rate also for servers with PCIe Gen3 support. The adapter's 32-lane PCIe bus is split into two 16-lane buses, with one bus accessible through a PCIe x16 edge connector and the other bus through an x16 parallel connector to an Auxiliary PCIe Connection Card. The two cards should be installed into two adjacent PCIe x16 slots and connected using a dedicated harness.

COMPATIBILITY

PCI EXPRESS INTERFACE
– PCIe Gen 4.0, 3.0, 2.0, 1.1 compatible
– 2.5, 5.0, 8, 16GT/s link rate
– 32 lanes of PCIe Gen3 or 16 lanes of PCIe Gen4
– Support for PCIe x1, x2, x4, x8, and x16 configurations
– PCIe Atomic
– TLP (Transaction Layer Packet) Processing Hints (TPH)
– PCIe switch Downstream Port Containment (DPC) enablement for PCIe hot-plug

– Advanced Error Reporting (AER)
– Access Control Service (ACS) for peer-to-peer secure communication
– Process Address Space ID (PASID) Address Translation Services (ATS)
– IBM CAPIv2 (Coherent Accelerator Processor Interface)
– Support for MSI/MSI-X mechanisms

OPERATING SYSTEMS/DISTRIBUTIONS*
– RHEL, SLES, Ubuntu and other major Linux distributions
– Windows

– FreeBSD
– VMware
– OpenFabrics Enterprise Distribution (OFED)
– OpenFabrics Windows Distribution (WinOF-2)

CONNECTIVITY
– Interoperability with Ethernet switches (up to 200GbE, as 4 lanes of 50Gb/s data rate)
– Passive copper cable with ESD protection
– Powered connectors for optical and active cable support

ConnectX®-6 EN 200Gb/s Adapter Card Single/Dual-Port Adapter Card Supporting up to 200GbE with
Advanced Offloading Technology

page 3

FEATURES SUMMARY*

ETHERNET
– 200GbE / 100GbE / 50GbE / 40GbE / 25GbE / 10GbE / 1GbE
– IEEE 802.3bj, 802.3bm 100 Gigabit Ethernet
– IEEE 802.3by, Ethernet Consortium 25, 50 Gigabit Ethernet, supporting all FEC modes
– IEEE 802.3ba 40 Gigabit Ethernet
– IEEE 802.3ae 10 Gigabit Ethernet
– IEEE 802.3az Energy Efficient Ethernet
– IEEE 802.3ap based auto-negotiation and KR startup
– Proprietary Ethernet protocols (20/40GBASE-R2, 50/56GBASE-R4)
– IEEE 802.3ad, 802.1AX Link Aggregation
– IEEE 802.1Q, 802.1P VLAN tags and priority
– IEEE 802.1Qau (QCN) – Congestion Notification
– IEEE 802.1Qaz (ETS)
– IEEE 802.1Qbb (PFC)
– IEEE 802.1Qbg
– IEEE 1588v2
– Jumbo frame support (9.6KB)

ENHANCED FEATURES
– Hardware-based reliable transport
– Collective operations offloads
– Vector collective operations offloads
– PeerDirect™ RDMA (aka GPUDirect®) communication acceleration
– 64/66 encoding
– Enhanced Atomic operations
– Advanced memory mapping support, allowing user mode registration and remapping of memory (UMR)
– Extended Reliable Connected transport (XRC)
– Dynamically Connected transport (DCT)
– On demand paging (ODP)
– MPI Tag Matching

– Rendezvous protocol offload
– Out-of-order RDMA supporting Adaptive Routing
– Burst buffer offload
– In-Network Memory registration-free RDMA memory access

CPU OFFLOADS
– RDMA over Converged Ethernet (RoCE)
– TCP/UDP/IP stateless offload
– LSO, LRO, checksum offload
– RSS (also on encapsulated packet), TSS, HDS, VLAN and MPLS tag insertion / stripping, Receive flow steering
– Data Plane Development Kit (DPDK) for kernel bypass applications
– Open VSwitch (OVS) offload using ASAP[2]
  • *Flexible match-action flow tables*
  • *Tunneling encapsulation / de-capsulation*
– Intelligent interrupt coalescence
– Header rewrite supporting hardware offload of NAT router

HARDWARE-BASED I/O VIRTUALIZATION
– Single Root IOV
– Address translation and protection
– VMware NetQueue support
– SR-IOV: Up to 512 Virtual Functions
– SR-IOV: Up to 16 Physical Functions per host
– Virtualization hierarchies (e.g., NPAR and Multi-Host)
  • *Virtualizing Physical Functions on a physical port*
  • *SR-IOV on every Physical Function*
– Configurable and user-programmable QoS
– Guaranteed QoS for VMs

STORAGE OFFLOADS
– Block-level encryption
  • *XTS-AES 256/512 bit key*
– NVMe over Fabric offloads for target machine

– Erasure Coding offload - offloading Reed Solomon calculations
– T10 DIF - signature handover operation at wire speed, for ingress and egress traffic
– Storage Protocols:
  • *SRP, iSER, NFS RDMA, SMB Direct, NVMeOF*

OVERLAY NETWORKS
– RoCE over overlay networks
– Stateless offloads for overlay network tunneling protocols
– Hardware offload of encapsulation and decapsulation of VXLAN, NVGRE, and GENEVE overlay networks

HPC SOFTWARE LIBRARIES
– HPC-X, OpenMPI, MVAPICH, MPICH, OpenSHMEM, PGAS and varied commercial packages

MANAGEMENT AND CONTROL
– NC-SI, MCTP over SMBus and MCTP over PCIe - Baseboard Management Controller interface
– SDN management interface for managing the eSwitch
– I²C interface for device control and configuration
– General Purpose I/O pins
– SPI interface to Flash
– JTAG IEEE 1149.1 and IEEE 1149.6

REMOTE BOOT
– Remote boot over InfiniBand
– Remote boot over Ethernet
– Remote boot over iSCSI
– Unified Extensible Firmware Interface (UEFI)
– Pre-execution Environment (PXE)

* This section describes hardware features and capabilities. Please refer to the driver release notes for feature availability.

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com

MLNX-15-51603PB
Rev 1.0