ConnectX®-6 VPI IC
200Gb/s InfiniBand & Ethernet Adapter IC

World’s first 200Gb/s HDR InfiniBand and Ethernet network adapter offering world-leading performance, smart offloads and In-Network Computing; leading to the highest return on investment for High-Performance Computing, Cloud, Web 2.0, Storage and Machine Learning applications.

ConnectX-6 Virtual Protocol Interconnect® is a groundbreaking addition to the Mellanox ConnectX series of industry-leading adapters. Offering unprecedented world-class performance, ConnectX-6 provides two ports of 200Gb/s for InfiniBand and Ethernet connectivity, sub-600ns latency and 215 million messages per second. With Mellanox Multi-Host® supporting up-to 8 independent hosts, integrated PCIe switch, NVMe over Fabric and security offloads, ConnectX-6 offers the highest performance and most flexible solution for the most demanding data center applications.

ConnectX-6 VPI supports HDR, HDR100, EDR, FDR, QDR, DDR and SDR InfiniBand speeds, as well as 200, 100, 50, 40, 25, and 10 Gb/s Ethernet speeds.

**HPC Environments**

Over the past decade, Mellanox has consistently driven HPC performance to new record heights. With the introduction of the ConnectX-6 adapter, Mellanox continues to pave the way with new features and unprecedented performance for the HPC market. Delivering the highest throughput and message rate in the industry, ConnectX-6 is the first adapter to deliver 200Gb/s HDR InfiniBand, 100Gb/s HDR100 InfiniBand and 200Gb/s Ethernet speeds. All this makes ConnectX-6 VPI the perfect product to lead HPC data centers toward Exascale levels of performance and scalability.

ConnectX-6 supports the evolving co-design paradigm with which the network becomes a distributed processor. With its In-Network Computing and In-Network Memory capabilities, ConnectX-6 offloads even further computation to the network, saving CPU cycles and increasing the efficiency of the network. ConnectX-6 VPI utilizes both IBTA RDMA (Remote Data Memory Access) and RoCE (RDMA over Converged Ethernet) technologies, delivering low-latency and high performance. ConnectX-6 enhances RDMA network capabilities even further by delivering end-to-end packet level flow control.

**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 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 RoCE (RDMA over Converged Ethernet) technology to deliver low-latency and high performance. ConnectX-6 enhances RDMA network capabilities even further by delivering RoCE Congestion Control, achieving end-to-end best performance.
Security

ConnectX-6 block-level encryption offers a critical innovation to network security. As data in transit is stored or retrieved, it undergoes encryption and decryption. The ConnectX-6 hardware offloads the IEEE AES-XTS encryption/decryption from the CPU, saving latency and CPU utilization. It also guarantees protection for users sharing the same resources through the use of dedicated encryption keys.

By performing block-storage encryption in the adapter, ConnectX-6 excludes the need for self-encrypted disks. This allows customers the freedom to choose their preferred storage device, including byte addressable and NVDIMM devices that traditionally do not provide encryption. Moreover, ConnectX-6 can support Federal Information Processing Standards (FIPS) compliance.

ConnectX-6 also includes a hardware Root-of-Trust (RoT), which uses HMAC relying on a device-unique key. This provides both a secure boot as well as cloning-protection. Delivering best-in-class device and firmware protection, ConnectX-6 also provides secured debugging capabilities, without the need for physical access.

Storage Environments

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

Additionally, as in previous ConnectX generations, ConnectX-6 enables Host Chaining, an innovative storage rack design by which different servers can be connected with no need for a switch.

Cloud and Web 2.0 Environments

Telco, Cloud and Web 2.0 customers developing platforms on Software Defined Network (SDN) environments are leveraging the virtual switching capabilities of their servers’ operating systems, to maximize the flexibility of managing and routing their network protocols.

Open VSwitch (OVS) allows virtual machines to communicate among themselves and with the outside world. Software-based virtual switches, traditionally residing in the hypervisor, are CPU-intensive; they affect system performance and prevent full CPU utilization for compute operations.

With Mellanox ASAP², Accelerated Switch and Packet Processing® Direct technology, significantly higher vSwitch/vRouter performance is achieved without the associated CPU load.

ConnectX-6 supports various vSwitch/vRouter offload functions including:
- Encapsulation and de-capsulation of overlay network headers
- 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 will enable hardware offloads of future protocols.

Standard & Multi-Host Management

Mellanox’s host management technology for standard and multi-host platforms optimizes board management and power, performance and memory management via NC-SI, MCTP over SMBus and MCTP over PCIe, as well as PLDM for Monitor and Control DSP0248 and PLDM for Firmware Update DSP0267.

Compatibility

<table>
<thead>
<tr>
<th>PCI Express Interface</th>
<th>Operating Systems/Distributions*</th>
<th>Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>– PCIe Gen 4.0, 3.0, 2.0, 1.1 compatible</td>
<td>– RHEL, SLES, Ubuntu and other major Linux distributions</td>
<td>– 50G SerDes (PAM4) and 25G SerDes (NRZ) based ports</td>
</tr>
<tr>
<td>– 2.5, 5.0, 8, 16 GT/s link rate</td>
<td>– Windows</td>
<td>– Interoperability with InfiniBand switches (up to HDR, as 4 lanes of 50Gb/s data rate)</td>
</tr>
<tr>
<td>– 32 lanes as 2x 16-lanes of PCIe</td>
<td>– FreeBSD</td>
<td>– Interoperability with Ethernet switches (up to 200GbE)</td>
</tr>
<tr>
<td>– Support for PCIe x1, x2, x4, x8, and x16 configurations</td>
<td>– VMware</td>
<td>– Passive copper cable with ESD protection</td>
</tr>
<tr>
<td>– PCIe Atomic</td>
<td>– OpenFabrics Enterprise Distribution (OFED)</td>
<td>– Powered connectors for optical and active cable support</td>
</tr>
<tr>
<td>– TLP (Transaction Layer Packet) Processing Hints (TPH)</td>
<td>– Support for MSI/MSI-X mechanisms</td>
<td></td>
</tr>
</tbody>
</table>
**Features***

**InfiniBand**
- HDR / HDR100 / EDR / FDR / QDR / DDR / SDR
- IBTA Specification 1.3 compliant
- RDMA, Send/Receive semantics
- Hardware-based congestion control
- Atomic operations
- 16 million I/O channels
- 256 to 4Kbyte MTU, 26byte messages
- 8 virtual lanes + VL15

**Ethernet**
- 200GbE / 100GbE / 50GbE / 40GbE / 25GbE / 10GbE / 1GbE
- IEEE 802.3by, 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 802.3by, Ethernet Consortium 25, 50 Gigabit Ethernet, supporting all FEC modes
- 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
- Mellanox PeerDirect RDMA (aka GPUDirect®) communication acceleration
- 64/66 encoding
- Advanced memory mapping support, allowing user mode registration and remapping of memory (UMR)
- Enhanced Atomic operations
- 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 ASAP²
  - Flexible match-action flow tables
  - Tunneling encapsulation / de-capsulation
- Intelligent interrupt coalescence
- Header rewrite supporting hardware offload of NAT router

**Storage Offloads**
- Block-level encryption: XTS-AES 256/512 bit key
- NVMe over Fabric offloads for target machine
- T10 DIF - signature handover operation at wire speed, for ingress and egress traffic
- Storage Protocols: SR, iSER, NFS, RDMA, SMB Direct, NVMe-oF

**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
- RDMA memory access

**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
- Guaranteed QoS for VMs
- RDMA, SMB Direct, NVMe-oF

**Management and Control**
- NC-SI, MCTP over SMBus and MCTP over PCIe-Base Management Controller interface,
- PLDM for Monitor and Control
- SDN management interface for managing the eSwitch
- PC 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 and firmware release notes for feature availability.

---

**Table 1 - Part Numbers and Descriptions**

<table>
<thead>
<tr>
<th>InfiniBand Supported Speeds (Gb/s)</th>
<th>Ethernet Supported Speeds (GbE)</th>
<th>No. of Network Ports</th>
<th>Crypto Support</th>
<th>PCI Express Configuration</th>
<th>OPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDR100, EDR, FDR, QDR, DDR, SDR</td>
<td>100, 50, 40, 25, 10</td>
<td>2</td>
<td>No crypto</td>
<td>PCIe Gen 3.0/4.0 x32</td>
<td>MT28908A0-NCCF-EV</td>
</tr>
<tr>
<td>HDR, HDR100, EDR, FDR, QDR, DDR, SDR</td>
<td>200, 100, 50, 40, 25, 10</td>
<td>1</td>
<td>No crypto</td>
<td>PCIe Gen 3.0/4.0 x32</td>
<td>MT28904A0-NCCF-HV</td>
</tr>
<tr>
<td>HDR, HDR100, EDR, FDR, QDR, DDR, SDR</td>
<td>200, 100, 50, 40, 25, 10</td>
<td>2</td>
<td>Crypto enabled</td>
<td>PCIe Gen 3.0/4.0 x32</td>
<td>MT28908A0-CCCF-HV</td>
</tr>
<tr>
<td>HDR, HDR100, EDR, FDR, QDR, DDR, SDR</td>
<td>200, 100, 50, 40, 25, 10</td>
<td>2</td>
<td>No crypto</td>
<td>PCIe Gen 3.0/4.0 x32</td>
<td>MT28908A0-NCCF-HV</td>
</tr>
<tr>
<td>HDR, HDR100, EDR, FDR, QDR, DDR, SDR</td>
<td>200, 100, 50, 40, 25, 10</td>
<td>2</td>
<td>Crypto enabled</td>
<td>PCIe Gen 3.0/4.0 x32</td>
<td>MT28908A0-CCCF-HV</td>
</tr>
</tbody>
</table>

---

© Copyright 2020 Mellanox Technologies. All rights reserved.
Mellanox, Mellanox logo, ConnectX, Mellanox Multi-Host, Mellanox PeerDirect, GPUDirect and ASAP® are registered trademarks of Mellanox Technologies, Ltd.
All other trademarks are property of their respective owners.