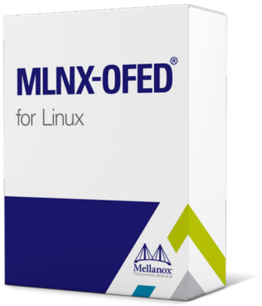
SOFTWARE

PRODUCT BRIEF



MLNX_OFED (OpenFabrics Enterprise Distribution)

High-Performance Server and Storage Connectivity Software for Field-Proven RDMA and Transport Offload Hardware Solutions

Clustering using commodity servers and storage systems is seeing widespread deployments in large and growing markets such as high-performance computing, data warehousing, online transaction processing, financial services and large scale Web 2.0 deployments.

To enable distributed computing with maximum efficiency, applications in these markets require the highest I/O bandwidth and lowest possible latency. These requirements are compounded by the need to support a large interoperable ecosystem of networking, virtualization, storage, and other applications and interfaces. The OFED Distribution from OpenFabrics Alliance (www. openfabrics.org) has been hardened through collaborative development and testing by major high performance I/O vendors. Mellanox OFED (MLNX_OFED) is a Mellanox-tested and packaged version of OFED, and supports RDMA (remote DMA) and kernel bypass APIs called OFED verbs over InfiniBand and Ethernet, allowing OEMs and System Integrators to meet the needs of end-users in their markets.

Virtual Protocol Interconnect Support

MLNX_OFED utilizes an efficient multi-layer device driver architecture to enable multiple I/O connectivity options over the same hardware I/O adapter (HCA or NIC). The same OFED stack installed on a server can deliver I/O services over both InfiniBand and Ethernet simultaneously, and ports can be repurposed to meet application and end-user needs. For example, one port on the adapter can function as a standard NIC or RoCE Ethernet NIC, and the other port can operate as InfiniBand; or, both ports can be repurposed to run as InfiniBand or Ethernet. Doing so enables IT managers to realize maximum flexibility in how they deliver converged and higher I/O service levels.

Delivering Converged and Higher I/O Service Levels

Networking in data center environments is comprised of server-to-server messaging (IPC), server-to-LAN and server-to-SAN traffic. To enable the highest performance, applications that run in the data center, especially those belonging to the target markets, expect different APIs or interfaces optimized for these different types of traffic.

BENEFITS

- Supports InfiniBand and Ethernet connectivity on the same adapter card
- Single software stack operating across all Mellanox InfiniBand and Ethernet devices
- Support for HPC applications such as scientific research, oil and gas exploration, car crash tests
- User level verbs allow protocols such as MPI and UDAPL to interface to Mellanox InfiniBand and up to 200GbE RoCE hard-ware. Kernel levels verbs allow protocols like NVMeOF, iSER, to interface to Mellanox InfiniBand and up to 200GbE RoCE hardware.
- Enhancing performance and scalability through Mellanox Message Accelerations (MXM)
- RoCEv2 enables L3 routing to provide better isolation and to enable hyperscale Web2.0 and cloud deployments with superior performance and efficiency
- Support for Data Center applications such as Oracle 11g RAC, IBM DB2, Purescale Financial services low-latency messaging applications such as IBM WebSphere LLM, Red Hat MRG Tibco
- Support for high-performance storage applications utilizing RDMA benefits
- Support I/O Virtualization technology such as SR-IOV and paraVirtualization over KVM and XenServer
- Support OVS offload with ASAP^2
- IPoIB component enable TCP/IP and sockets-based applications to benefi from InfiniBand transport
- Supports the OpenFabrics defined Verbs API at the user and kernel levels.
- SCSI Mid Layer interface enabling SCSI-based block storage and management

Efficient & High Performance HPC Clustering

In HPC applications, MPI (message passing interface) is widely used as a parallel programming communications library. In emerging large-scale HPC applications, the I/O bottleneck is no longer only in the fabric, but has extended to the communication libraries. In order to provide the most scalable solution for MPI, SHMEM and PGAS applications, Mellanox provides a new accelerations library named MXM (Mellanox Messaging) that enables MPI, SHMEM and PGAS programming languages to scale to very large clusters by improving on memory and latency related efficiencies, and to assure that the communication libraries are fully optimized over Mellanox interconnect solutions.

Accelerating Cloud Infrastructure

MLNX_OFED integrates Mellanox' ASAP^2 -Accelerated Switch and Packet Processing® technology. ASAP^2 offloads 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, minus the associated CPU load.

Lowest Latency for Financial Services Applications

MLNX_OFED with InfiniBand, RoCE and VMA delivers the lowest latency for financial applications and the highest Packet Per Second (PPS) performance.

Web 2.0 and Other Traditional Sockets-based Applications

MLNX_OFED includes the implementation of IP-over-IB, enabling IP-based applications to work seamlessly over InfiniBand Standard UDP/TCP/IP sockets interfaces are supported over L2 NIC (Ethernet) driver implementation for applications that are not sensitive to lowest latency.

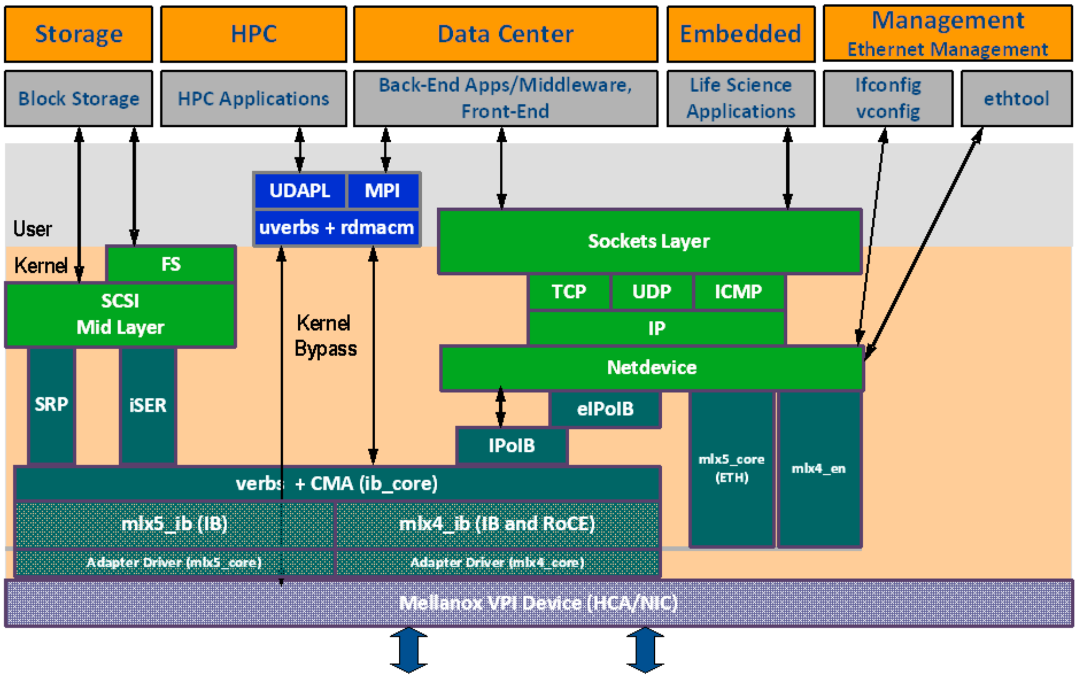
Storage Applications

To enable traditional SCSI and iSCSI-based storage applications to enjoy similar RDMA performance benefits, MLNX_OFED includes iSCSI RDMA Protocol (iSER) that interoperate with various target components available in the industry. iSER can be implemented over both InfiniBand or RoCE.

MLNX_OFED supports Mellanox storage acceleration, which is a consolidated compute and storage network that achieves significant cost-performance advantages over multi-fabric networks.

High Availability (HA)

MLNX_OFED includes high availability support for message passing, sockets and storage applications. The standard Linux channel bonding module is supported over IPoIB that enables failover across ports on the same adapter, or across adapters. Some vendor-specific fail-over/load-balancing driver models are supported as well.Database (OVSDB) or other virtual switches to create a secure solution for bare metal provisioning. The software package also includes support for DPDK, and the ability to enable IPsec and a stateful L4-based firewall.



SUPPORTED OPERATING SYSTEMS
RedHat
CentOS
Ubuntu
SLES
OEL
Citrix XenServer Host
Fedora
Debian

DEVICE SUPPORT
ConnectX-3
ConnectX-3 Pro
ConnectX-4 and ConnectX-4 Lx
ConnectX-5
ConnectX-6

COMPONENTS	
Drivers for InfiniBand, RoCE, L2 NIC	iSER Initiator
Access Layers and common verbs interface	uDAPL
VPI (Virtual Protocol Interconnect)	Subnet Manager (OpenSM)
OSU MVAPICH and Open MPI	Installation, Administration and Diagnostics
IP-over-IB	Tools
	Performance test suites

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