



Mellanox Windows OpenFabrics Software (MLNX_WinOF)

Mellanox InfiniBand interconnects deliver low-latency, high-performance and scalable cluster solutions with Microsoft Windows® platform

Use of clustered commodity servers, in lieu of traditional supercomputers and mainframes, offers tremendous price/performance benefits and unparalleled flexibility in deployment and long-term maintenance. To enable distributed computing transparently, high-performance computing (HPC) applications require the highest bandwidth and lowest possible latency. In enterprise data center (EDC) applications, these requirements are compounded with the need to support a large interoperable ecosystem of networking, virtualization, storage, and other applications and interfaces.

The MLNX_WinOF based on WinOF 2.0 from OpenFabrics Alliance (www.openfabrics.org) has been hardened through collaborative development and testing by Mellanox and major InfiniBand vendors to enable OEMs and System Integrators to meet the needs of HPC and EDC applications. Mellanox and Microsoft's HPC team have collaborated to provide a certified high performance driver on Mellanox InfiniBand adapters supporting Microsoft Windows HPC Server 2008.

Higher I/O Service Levels for high performance in Data Center traffic and applications

Data Center transactions and compute intensive applications require data transfers to be handled with zero-copy support reducing CPU overhead, without any kernel transitions to take advantage of higher bandwidth with very low-latency. Explicit application buffer management allows applications to transfer data directly from one application's buffer to the remote application's buffer without an intermediate operating system copy on either side using zero-copy. These capabilities enable better scalability, efficiency, and performance than is possible through a traditional socket interface. Network Direct (ND) - Microsoft's RDMA interface - takes advantage of these advanced capabilities in Mellanox InfiniBand adapters to achieve better performance for large clusters running parallel programs that utilize low-latency, high-bandwidth networks.

Running various Message Passing Interface (MPI) protocol implementations, node-to-node latency on Mellanox InfiniBand adapters is less than 2µs with a uni-directional bandwidth of 3200MB/s.

Sockets-based Applications with IPoIB, SDP and WSD

For traditional sockets-based applications, MLNX_WinOF provides a network device (NDIS) mini port driver interface supporting the field proven implementation of IP-over-IB, enabling IP-based applications to work seamlessly over InfiniBand. It also includes the IBTA (www.InfiniBandTA.org) defined Sockets Direct Protocol (SDP) enabling traditional TCP/IP sockets-based applications to capitalize on the RDMA and transport offload capabilities of InfiniBand. Also, WinSock Direct (WSD) Service provider interface enables WinSock (TCP/IP) API-based applications to capitalize on the InfiniBand RDMA and transport offload capabilities (kernel-bypass).

To enable traditional SCSI and iSCSI-based storage applications to enjoy similar RDMA performance benefits, MLNX_WinOF includes the SCSI over RDMA Protocol (SRP) initiator.

BENEFITS

- Single software stack that operates across all available InfiniBand devices and configurations such as mem-free, QDR/DDR/SDR, PCI-X and PCI Express modes
- Support for Microsoft NetworkDirect (ND) for scalability, performance and low-latency in a cluster running parallel programs
- Support for HPC applications for scientific research, oil and gas exploration, car crash tests, benchmarking
- Support for traditional IP and Sockets-based applications leveraging the benefits of RDMA
- Support for high-performance block storage applications utilizing RDMA benefits
- Support for cluster management tools

SOCKETS LAYER

- SDP and IP-over-IB component enable TCP/IP and sockets-based applications to interface seamlessly to and benefit from InfiniBand RDMA transport

ACCESS LAYER

- Supports the OpenFabrics defined Verbs API at the user and kernel levels. User level verbs allow MPI protocols and other applications to interface to Mellanox InfiniBand hardware. Kernel level verbs allow protocols like SDP, SRP and IP-over-IB to interface to Mellanox InfiniBand hardware

SCSI MID LAYER

- The SCSI Mid Layer interface enables SCSI-based block storage and management applications to interface with the SRP Initiator component and the Mellanox InfiniBand hardware

Interoperable across Multi-Vendor Solutions

As a founding member of OpenFabrics, Mellanox is driving interoperability of the OpenFabrics software across different vendor solutions. When running on servers, MLNX_WinOF interoperates with InfiniBand switches, gateways and storage targets and value-added software provided by vendors in such systems. Today, MLNX_WinOF is supported on Microsoft Windows Server Operating Systems, and supports all x86 based CPU platforms.

Enabling a Large Software Ecosystem

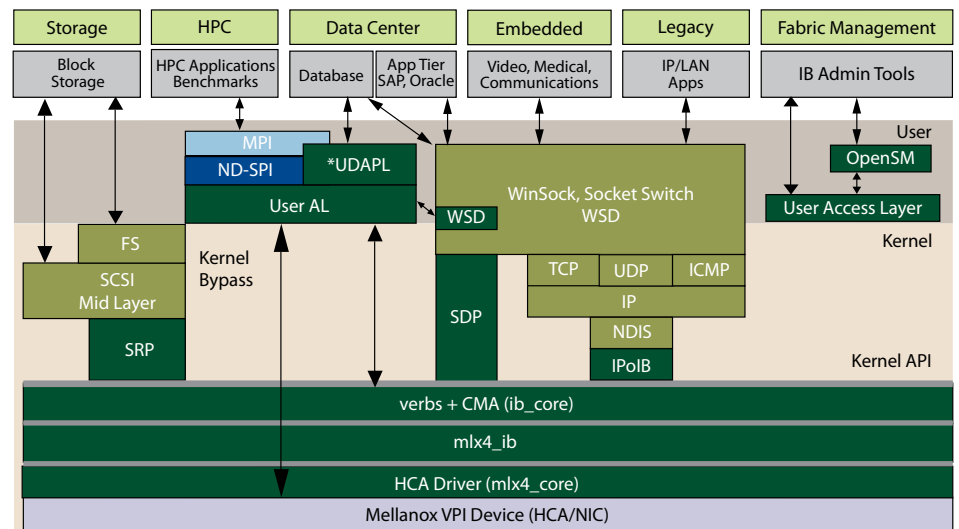
Through its collaboration with the OpenFabrics community, support for multiple industry-standard interfaces that enable off-the-shelf applications to work with InfiniBand, and support for popular OS and CPU platforms, Mellanox has created a large, interoperable software ecosystem that benefits the HPC and EDC markets.

COMPONENTS

- Driver & Access Layers
- OSU MVAPICH and Open MPI
- IP-over-IB (with Offload)
- Sockets Direct Protocol (Beta)
- SCSI over RDMA Protocol (Beta)
- Winsock Direct (WSD)
- NetworkDirect (ND)
- Subnet Manager (OpenSM)
- vstat – utility to get card status
- SdpConnect – SDP/WSD test tool
- Software Development Kit (SDK)
- Installation, Administration and Diagnostics Tools
- Performance test suites
- Related Documentation

DEVICE AND FIRMWARE SUPPORT

- InfiniHost® HCA Silicon
 - fw-23108 Rev 3.5.000 or later
- InfiniHost® III Ex HCA Silicon
 - MemFree: fw-25218 Rev 5.3.000 or later
 - Memory: fw-25208 Rev 4.8.200 or later
- InfiniHost III Lx HCA Silicon
 - fw-25204 Rev 1.2.000 or later
- ConnectX® HCA Silicon (SDR/DDR/QDR)
 - fw-25408 Rev 2.5.000 or later
- InfiniScale® Switch Silicon
- InfiniScale® III Switch Silicon
- InfiniScale® IV Switch Systems



Markets
 Windows HPC Server 2008
 MLNX_WinOF
 * Currently Not Supported by Mellanox

Platform Supplier Products and Support	
Microsoft	Windows XP, Windows Server 2003, Windows Server 2008, Windows CCS (64-bit), HPC Server 2008
Intel	x86, EM64T, EM64T x86_64, PCI-X and PCIe platforms
AMD	AMD64 and AMD64-Ex Opteron, PCI-X and PCIe platforms



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



© Copyright 2009. Mellanox Technologies. All rights reserved. Mellanox, ConnectX, InfiniBlast, InfiniBridge, InfiniHost, InfiniRISC, InfiniScale, and InfiniPCI are registered trademarks of Mellanox Technologies, Ltd. Virtual Protocol Interconnect and BridgeX are trademarks of Mellanox Technologies, Ltd. All other trademarks are property of their respective owners.