4th Generation Server & Storage Adapter Architecture

The growing demand for better application performance is driving the increased adoption of clustered servers and server virtualization on top of multi-core processors. Connecting these servers in a way that maximizes their performance poses I/O connectivity challenges. Interconnects must support the bandwidth, reliability, and service guarantees each application demands while doing so in a cost effective manner. To meet those needs, current solutions include multiple connectivity and software options for data, storage, and management applications. These solutions are inadequate for the demands of tomorrow’s server and storage networks.

The ConnectX™ high-performance server and storage adapter architecture addresses the technical computing and enterprise data center’s demand for cost effective bandwidth, Quality of Service, and infrastructure investment protection. The architecture continues the advances in InfiniBand technology that Mellanox is known for and extends InfiniBand’s field proven price/performance value and service oriented I/O capabilities to Ethernet fabrics. It’s Ethernet capabilities conform to the IEEE standard, harness the power of convergence trends around IP, and supports all industry standard clustering, computing, storage and networking interfaces for easy application deployment.

Industry-leading Integration

The ConnectX architecture enables integration of extraordinary levels of functionality into a single adapter chip. It supports eight lane PCI Express v1.1 and v2.0 interfaces as the host interface. For network and storage connectivity, it supports two network ports. Each port can be flexibly configured as either a 4X InfiniBand port or a 10 Gigabit Ethernet XAUI/CX4 or XFI port based on OEM preferences. The InfiniBand port supports Single, Double and Quad Data Rates (SDR, DDR and QDR) delivering 10, 20 and 40Gb/s full duplex bandwidth. The network ports support either InfiniBand physical layer semantics or XAUI semantics.

Flexible Layered Architecture

ConnectX implements a layered architecture to provide maximum flexibility in the usage of software and hardware protocols it supports. It does so by maintaining industry standard application interfaces on top of an IPv4/IPv6 based network layer, and InfiniBand and 10 Gigabit Ethernet-based data link layers. The physical layer supports XAUI/CX4, XFI or InfiniBand physical link, enabling use of common fiber and copper cabling infrastructures across InfiniBand and 10 Gigabit Ethernet fabrics. The application interfaces supported transparently over either fabric includes IP, sockets, MPI, SCSI, iSCSI, Fibre Channel and NFS. The architecture supports up to 16 million flow interfaces or virtual end points.
(referred to as Channel I/O) that can be assigned to any of the above application interfaces, enabling granular levels of QoS and isolation per application or virtual machines in a virtual server environment. The key to offering such scalability and granularity is overlaying the I/O channels evenly over efficient use of multiple transport options — starting with TCP/UDP in legacy OS-compatible stateless offload modes to the use of field proven and efficient InfiniBand transport for RDMA (Remote Direct Memory Access) and full-offload modes, to use of FCP-3 (Class 3 Fibre Channel Protocol) transport for Fibre Channel over InfiniBand services.

**Optimal Transport Services**

The use of TCP/UDP transport enables operations with legacy operating systems and stacks, offloading only stateless functions into the ConnectX device. The use of InfiniBand transport enables cost-effective and low power RDMA solutions while enabling the same OpenFabrics (www.openfabrics.org) software stack and Channel I/O interfaces to work seamlessly over both InfiniBand and 10 Gigabit Ethernet. As a result, the growing ecosystem of networking, storage, database, and cluster file system applications qualified and deployed over InfiniBand can now be easily deployed over 10 Gigabit Ethernet as well. Block storage protocols such as iSER (iSCSI RDMA Extension) can be seamlessly deployed over both fabrics. FCP-3 transport services over InfiniBand enables the use of simple InfiniBand-to-Fibre Channel gateways for connectivity to Fibre Channel SANs.

**Price-Performance-Power Benefits**

ConnectX is Mellanox’s fourth generation architecture and rides on the company’s expertise in developing and deploying (in excess of 1.4 million ports) 10Gb/s and higher bandwidth server and storage connectivity products. The maturity of the design on a single chip and the level of integration enable the lowest cost per adapter in its class. Utilizing RDMA and transport services designed ground-up for hardware implementation efficiencies, ConnectX can deliver server to server connectivity bandwidth as high as 40Gb/s, end-to-end latency of as low as one microsecond, and at power consumption levels as low as 5.5 Watts per port.

**Service Oriented I/O Benefits**

ConnectX excels over other connectivity solutions in its ability to provide stringent levels of services to applications and virtual machines in virtual server environments. At the core of such services is the Channel I/O based flow interfaces applied coherently over a converged set of clustering, communications, storage and management functions. A server farm equipped with ConnectX is therefore the perfect complement to service oriented architecture-based software deployments in the data center, where I/O services are provided to consumers as utility services.

An application plugging into such services over a suitably enabled network can support end-to-end QoS and congestion management, leading to guaranteed levels of services. Multi-level QoS including marking, shaping and queuing are supported. End-to-end congestion avoidance and management is achieved through use of hardware based forward and backward congestion notification mechanisms and throttling of traffic in sending nodes on a per flow interface (or Channel I/O) level granularity.

Virtual machines plugging into Channel I/O-based services can enjoy direct adapter hardware access, partitioning, isolation and intra-VM...
Design Flexibility for OEMs

With ConnectX-based adapter solutions server OEMs can not only reduce their design cycles and costs significantly, but can also expand their target markets. ConnectX adapter network ports are configurable to support either InfiniBand or Ethernet. Both conservative stateless offload and high performance RDMA can be enabled over either option. A legacy TCP/IP stack based implementation or a common OpenFabrics RDMA stack can be applied over either option, cutting down software qualification cycles drastically, for all application types – clustering, communications, storage and management. Physical layer and connector requirements are common as well resulting in one backplane design and layout cycle. In blade servers, which is the fastest growing server market segment, a single ConnectX-based I/O mezzanine card can enable the addressing of multiple data center applications while cutting down on blade form factors (such as number of PCI slots) and I/O power requirements.

Future Proofing for End Users

Increasing deployments of multi-core CPU-based servers, virtualization, unified I/O combined with explosive growths in data volumes that need to be processed by data center applications are placing tremendous burdens on traditional I/O connectivity options such as Fast/Gigabit Ethernet and Fibre Channel. With ConnectX, end users with server farms using Fast or Gigabit Ethernet for connectivity and Fibre Channel-based SANs can avoid fork lift upgrades as they move to higher performance and unified (storage and network) connectivity options. For example, the data center manager may opt to keep his applications, OS and TCP/IP stacks unchanged, upgrading only the connectivity infrastructure to 10 Gigabit Ethernet or InfiniBand. In another situation, he may chose to only move a few applications to utilize RDMA benefits, and do so first over 10 Gigabit Ethernet and later move some of those applications to InfiniBand. In another situation, he may choose to keep his legacy Fibre Channel SAN, and Gigabit Ethernet LAN intact. In another situation, he may chose to upgrade some of his storage applications to use iSCSI and when he is comfortable, move a subset to iSCSI RDMA (using iSER) over 10 Gigabit Ethernet or InfiniBand. And the choices go on and on. ConnectX adapter solutions offer ultimate flexibility to IT managers in how they want to grow and scale their network and storage capabilities.

Deployment Scenarios

A typical deployment scenario (see figure below) shows use of ConnectX-based server blade mezzanine cards in server blades with one InfiniBand port and one 10 Gigabit Ethernet port. The reliable InfiniBand backplane is used for server to server and server to storage connectivity for bandwidth hungry and low latency mid-tier and back-end tier storage applications. Native InfiniBand storage targets are used for accelerating storage performance. The 10 Gigabit Ethernet port is used for front-end tier applications where bandwidth, reliability and latency requirements are not as stringent, and LAN and WAN connectivity is important. Routing between InfiniBand and 10 Gigabit Ethernet can be achieved through simple software based routing on the servers or through use of a low cost InfiniBand-Ethernet gateway that uses the...
ConnectX solution. The InfiniBand fabric may connect to an existing Fibre Channel SAN using a low cost InfiniBand-Fibre Channel gateway that uses the ConnectX solution, specifically the Fibre Channel over InfiniBand implementation using the FCP-3 transport services.

Meeting The Needs Of Server And Storage Networks

The ConnectX architecture offers the necessary building blocks for designing leading price/performance server and storage connectivity solutions for today and tomorrow. Service oriented I/O ensures scalable and granular quality of service for a converged set of data center applications, in both native operating system and virtualized server environments. Multiple network and storage connectivity options on a single ConnectX adapter enhances the server OEMs target markets and time to market. Finally, end users can use ConnectX-based solutions for incremental and demand-based I/O growth and scaling.