



# Mellanox InfiniBand FDR 56Gb/s For Server and Storage Interconnect Solutions

Accelerating Current (PCIe Gen-2 based) and Future (PCIe Gen-3 based) HPC Platforms



## InfiniBand FDR

InfiniBand FDR.....	1
InfiniBand FDR Delivers Significant Interconnect Enhancements .....	1
Summary .....	3

Choosing the right interconnect technology is essential for maximizing systems and applications performance and efficiency. Slow interconnects delay data transfers between servers, causing poor utilization of the system resources and slow execution of application. By providing low-latency, high-bandwidth and low CPU overhead, Remote Direct Memory Access (RDMA) and more, InfiniBand has become the most deployed high-speed interconnect, replacing proprietary or low-performance solutions. The InfiniBand architecture is an industry-standard fabric designed to provide high bandwidth, low-latency computing, scalability for tens of thousands of server and storage nodes and efficient utilization of compute processing resources.

InfiniBand FDR (Fourteen Data Rate, 14Gb/s data rate per lane 56Gb/s per port) is the next generation InfiniBand technology specified by the InfiniBand Trade Association. InfiniBand FDR was announced in June 2010 and is targeted for next generation high-performance computing and enterprise data centers that are looking to maximize their server and storage performance and optimize the performance, reliability, scalability and efficiency of their communications networks.

InfiniBand lane speeds continue to increase to support end user demands for improved return on investment (ROI) and performance benefits, as well as robust network capabilities to support multi-core processors and accelerators. The progression of InfiniBand product availability based on data rates is as follows:

- 2002: Single Data Rate (SDR) 2.5Gb/s per lane, 10Gb/s per port (typical port contains 4 lanes)
- 2005: Double Data Rate (DDR) 5Gb/s, 20Gb/s per port (typical port contains 4 lanes)
- 2008: Quad Data Rate (QDR) 10Gb/s, 40Gb/s per port (typical port contains 4 lanes)
- 2011: Fourteen Data Rate (FDR) 14.0625Gb/s, 56.25Gb/s per port (typical port contains 4 lanes)

## InfiniBand FDR Delivers Significant Interconnect Enhancements

InfiniBand FDR delivers greater performance, scalability and reliability when compared to products with lower data rates or compared to other available interconnect solutions.

InfiniBand FDR 56Gb/s introduces several enhancements for higher performance, scalability and reliability for performance demanding data centers.

### 1. Network Bandwidth

The InfiniBand FDR Link speed has increased to 14Gb/s per lane, or 56Gb/s per 4 lane port (typical InfiniBand implementation), a data rate increase of more than 70 percent compared to previous InfiniBand generations. The dramatic increase in network bandwidth allows the following advantages:

- For servers equipped with CPU architectures that can utilize the new bandwidth capabilities, such as PCIe Gen-3 based servers, InfiniBand FDR delivers maximum throughput from the server to the network allowing higher application performance and scalability.
- For servers equipped with CPU architectures based on PCIe Gen-2 or PCIe Gen-1, InfiniBand FDR enables building the most cost effective network infrastructures through an oversubscribed architecture. Since the server bandwidth is lower than the InfiniBand FDR throughput, a network solution can be designed with fewer switch interconnects than server interconnects without imposing any limitations on the server bandwidth. For example, a cluster based on PCIe Gen-2 servers can be connected using an oversubscribed 2:1 InfiniBand FDR switch network (which will be non-blocking in this case). Other oversubscribed (blocking or non-blocking) networks can be designed based on application needs, all utilizing fewer switch components than if built with slower fabrics.
- For any CPU architecture, the increase in network throughput reduces network congestions and hot spots, allowing applications to communicate faster and to achieve higher productivity.

## 2. Network Latency

Mellanox InfiniBand FDR interconnect solutions (based on ConnectX-3 and SwitchX) can build networks that accelerate data delivery with reduced fabric latency. The reduction in latency enables faster communication and synchronization between application processes and increases the cluster performance and the overall return on investment.

## 3. Network Efficiency

The link encoding for InfiniBand FDR was modified from 8bits/10bits, used in InfiniBand SDR, DDR and QDR, to 64bits/66bits. This allows higher network efficiency for data center server and storage connectivity by reducing the ratio between control bits and data bits sent in the network. With InfiniBand FDR, the network spends more time on actual data delivery between application job processes compared to SDR, DDR, and QDR, which in turn increases the overall network productivity.

## 4. Network Reliability and Data Integrity (FEC)

InfiniBand provides a scalable and reliable high-speed interconnect for servers and storage. For data integrity and guaranteed reliable data transfer between end-nodes (servers and storage), InfiniBand uses an end-to-end hardware reliability mechanism. Each InfiniBand packet contains two Cyclic Redundancy Checks (CRCs). The Invariant CRC (ICRC) covers all fields which do not change as the packet traverses the fabric. The Variant CRC (VCRC) covers the entire packet. The combination of the two CRCs allows switches and routers to modify appropriate fields and maintain end-to-end data integrity. If a data corruption occurs due to Bit Error Rate (BER), the packet will be discarded by the switch or the adapter, and will be re-transmitted by the source to the target.

In order to accelerate the data re-transmission, a new mechanism was added to InfiniBand FDR - Forward Error Correction (FEC). FEC allows the InfiniBand devices (adapters and switches) to fix bit errors throughout the network and reduce the overhead for data re-transmission between the end-nodes. The InfiniBand FDR FEC mechanism utilizes redundancy in the 64/66-bits encoding to enable error correction with no bandwidth loss and has the ability to work over each link independently, on each of the link lanes. The new mechanism delivers superior network reliability especially for the large scale data centers, high-performance computing or web 2.0 centers, and delivers a predictable low-latency characteristic, critical for large scale applications and synchronizations.

## 5. Low Power Consumption for Green Data Centers

The InfiniBand FDR product line from Mellanox Technologies reduces the network related power consumption, resulting in dramatic savings in infrastructure and maintenance costs. Furthermore, the increased bandwidth available throughout the network makes InfiniBand FDR the most cost-effective and power-effective solution for oversubscribed network topologies, 3D-Torus topologies and many others. It also features enhanced fabric consolidation and allows multiple applications (IPC, management, storage) to share the same network with no performance degradations. The result is pure savings in capital expenditures (CAPEX) and operational expenditures (OPEX).

## 6. Scalability and Consolidation

Mellanox FDR switch product line introduces new consolidated fabric elements for higher scalability and fabric consolidation. Mellanox FDR switches include an integrated InfiniBand router and network bridges from InfiniBand to Ethernet and from InfiniBand to Fibre Channel. The integration of the InfiniBand router into the switch extends the InfiniBand available address space to  $2^{128}$  available InfiniBand address (an identifier given to end-points and switches) and therefore enables unlimited fabric size. InfiniBand FDR delivers future proofing allowing future expansion of any size to be added to the InfiniBand network. The seamless integration of the bridges yields a simplified yet high-performance connectivity from the InfiniBand network to legacy or other networks, resulting in even more savings in CAPEX and OPEX.

## 7. Performance-Driven Architecture

Mellanox InfiniBand FDR-based solutions introduce many other advantages – from MPI and SHMEM based application acceleration, GPU-based communication accelerations, distributed clock synchronization and many more. To learn more please contact your Mellanox representative, visit [www.mellanox.com](http://www.mellanox.com) or contact [HPC@mellanox.com](mailto:HPC@mellanox.com)

## Summary

The newly introduced InfiniBand FDR adapters, switches and cables deliver new, high-performance, scalable, efficient and reliable interconnect solutions for connecting servers and storage. The performance, cost-effectiveness, and scalability advantages delivered by InfiniBand FDR increase application productivity and optimize the return on investment. InfiniBand FDR is the best interconnect solution for current high-performance and data center clusters (based on PCIe Gen-1 or PCIe Gen-2) and for future planned centers based PCIe Gen-3.



350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085  
Tel: 408-970-3400 • Fax: 408-970-3403  
[www.mellanox.com](http://www.mellanox.com)