InfiniScale® IV
4th Generation InfiniBand Switch Architecture

3 Terabits per Second Switching
InfiniBand provides the highest-bandwidth, lowest-latency, and most scalable interconnect for servers and storage. The fourth generation switching architecture from Mellanox, InfiniScale IV, improves these leading attributes further, making InfiniBand an obvious choice for the most demanding applications. InfiniScale IV silicon products will be used by system architects to construct some of the world’s most powerful computing systems, and for carrying converged traffic with the best combination of bandwidth and granular quality of service.

The InfiniScale IV architecture extends the performance and capabilities of InfiniBand switching across several dimensions: more bandwidth, lower-latency, more scale, and improved features. Extensive experience in building systems using the InfiniScale III and previous generations of InfiniBand switching products have given Mellanox a wealth of field-tested experience, leading to new insights in how to optimize performance and extend functionality.

With 36 high-performance 40Gb/s ports, fewer switch chips are needed than with customary 24-port building blocks. Besides the obvious benefit of lowering parts count, constant bisectional bandwidth (CBB) fabrics can use fewer tiers (layers) of switches, thus reducing the hop count and latency. Coupled with the lower-latency inherent in the InfiniScale IV architecture, large clusters run with lower overall latency and better overall application performance.

Adaptive Routing has been added to the static routing previously provided by the architecture. InfiniBand is optimized to move traffic via multiple parallel paths, and adaptive routing uses immediate feedback within the switch chip to move traffic from congested to less congested paths. This is done at in-chip hardware speeds, to avoid the problems created if this feedback loop is

InfiniScale IV Block Diagram
InfiniScale® IV 4th Generation InfiniBand Switch Architecture

KEY APPLICATIONS

- Virtualized data centers that require a high-bandwidth, low-latency interconnect for server and storage grids
- High-performance parallelized computing leveraging Message Passing Interface (MPI) based applications such as molecular modeling, oil and gas exploration, car crash simulations, etc.
- Clustered database applications, parallel RDBMS queries, high-throughput data warehousing
- Performance storage applications such as backup, restore, mirroring, etc.
- High bandwidth streaming content such as video-on-demand and HDTV
- Electronic Design Automation (EDA)
- Networking, telecom and industrial data acquisition

too slow. In cases where traffic patterns are more predictable, static routing has been shown to produce superior results. The InfiniScale IV architecture provides the best of both static and adaptive routing.

Some switching contention is unavoidable — for example, when two sources are trying to reach a single destination. Congestion control, using InfiniBand 1.2 standard mechanisms, is the only proven way to control hotspots in large fabrics. InfiniScale IV works in conjunction with ConnectX IB adapters to control sources when hotspots are detected, keeping traffic moving while keeping latency low and constant.

For converged traffic, the combination of high-bandwidth, adaptive routing, and congestion control provide the industry’s best traffic-carrying capacity. Adding QoS makes sure that traffic classes can be protected, guaranteeing the delivery of critical traffic

How to Build the Best Fabric

Start with the fastest fabric. Add Adaptive Routing, Congestion Control, and QoS. Reduce latency until it’s nearly non-existent. That’s the compelling benefit of the InfiniScale IV architecture.

40Gb/s and 120Gb/s Connected InfiniBand Server and Storage Diagram

© Copyright 2008. Mellanox Technologies. All rights reserved. Mellanox, ConnectX, InfiniBlade, InfiniBridge, InfiniHost, InfiniRISC, InfiniScale, and InfiniPCI are registered trademarks of Mellanox Technologies, Ltd. Virtual Protocol Interconnect and BridgeX is a trademark of Mellanox Technologies, Ltd. All other trademarks are property of their respective owners.