Mellanox InfiniScale® IV

36-port 40Gb/s InfiniBand Switch Device

BENEFITS

- World-class cluster performance
- High-performance networking and storage access
- Guaranteed bandwidth and low-latency services
- Reliable transport
- I/O consolidation
- Virtualization acceleration
- Scales to tens-of-thousands of nodes

KEY FEATURES

- 2.88 Tb/s switching capacity
- IBTA 1.2 compliant
- IBTA compliant auto-negotiation
- Flexible port configurations
  - 4X, 8X, or 12X
  - 10, 20, or 40 Gb/s per 4X port
- Adaptive routing
- Congestion control
- Quality of Service enforcement
- Virtual subnets
- Port mirroring

SPECIFICATIONS

- 36 4X or 12 8X or 12 12X InfiniBand ports, or a combination of port types
- PCI Express 2.0 x4 5GT/s (1.1 compatible)
- Serial Flash interface, up to 64MB
- Dual I2C interfaces
- IEEE 1149.1 boundary-scan JTAG
- Link status LED indicators
- General purpose I/O
- 45 x 45mm FCBGA

InfiniBand provides the highest bandwidth, lowest latency, and most scalable interconnect for servers and data storage. InfiniScale IV, the fourth generation switching device from Mellanox, improves these leading attributes further, making Mellanox InfiniBand an obvious choice for the most demanding applications. Switch systems based on InfiniScale IV can be used by network architects to construct Petascale computing systems. IT managers can build large networks that carry converged traffic with the best combination of assured bandwidth and granular quality of service.

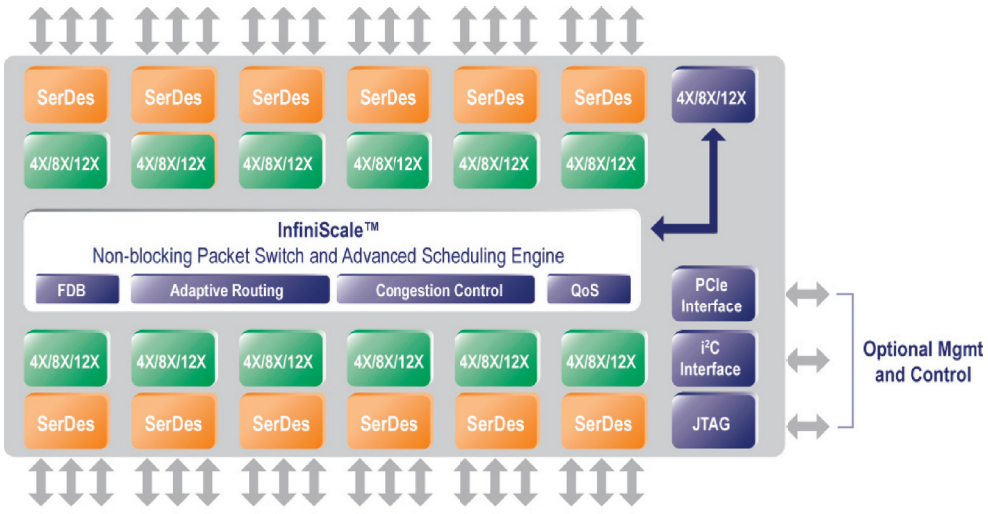
World Class Performance and Scalability

The InfiniScale IV extends the performance and capabilities of InfiniBand switching across several dimensions: more bandwidth, lower latency, more scale, and improved features. With 36 high-performance 40Gb/s 4X ports, fewer switch chips are needed to build large clusters than with customary 24 port building blocks. Besides the obvious benefit of lowering parts count, another benefit is gained – constant bisectional bandwidth (CBB) fabrics can use fewer tiers (layers) of switches, reducing the hop count and associated latency. Two tiered clusters can now be built out to 648 nodes, and three tiered clusters can be built out to greater than 10K nodes. Coupled with the lower latency inherent in InfiniScale IV, large clusters run with lower overall latency and better overall application performance.

Sustained Network Performance

Adaptive Routing has been added to the static routing capability available in previous switch families. InfiniBand supports moving traffic via multiple parallel paths. Adaptive routing dynamically and automatically re-routes traffic to alleviate congested ports. In networks 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 multiple sources are trying to



InfiniScale IV Block Diagram

InfiniScale® IV 36-port 40Gb/s InfiniBand Switch Device

reach a single destination. Congestion control, using InfiniBand 1.2 standard mechanisms, is the only proven solution to remove hotspots in large fabrics. InfiniScale IV works in conjunction with ConnectX InfiniBand adapters to restrict process traffic causing congestion, ensuring high bandwidth and low latency to all other flows.

For converged traffic, the combination of high bandwidth, adaptive routing, and congestion control provide the industry's best traffic carrying capacity. End to end Quality of Service makes sure that traffic classes can be protected, guaranteeing the delivery of critical traffic.

Utility Computing

Virtual partitioning of a cluster enables efficient use of all of its computing resources. Allocating only the compute power that each client needs enables more clients on the cluster at one time. Clusters built on InfiniScale IV can run multiple subnets, securely segregating client processes while ensuring the highest productivity of the cluster.

Switch Product Development Platforms

InfiniScale IV development platforms are available to accelerate OEMs' time to market and for running benchmark tests. These 36-port 40Gb/s rack mountable systems are available with either microGiGaCN™ or QSFP connectors compatible with passive and active copper cables and active fiber cabling solutions.

Mellanox Advantage

Mellanox is the leading supplier of industry standard InfiniBand HCAs and switch silicon. Our products have been deployed in clusters scaling to thousands of nodes and are being deployed end-to-end in data centers and Top500 systems around the world.

FEATURE SUMMARY

INFINIBAND

- IBTA Specification 1.2 compliant
- 10, 20, or 40Gb/s per 4X port
- Integrated subnet manager agent
- Integrated general service agent
- Linear forwarding table
- Hardware-based congestion control
- 256 to 4Kbyte MTU
- 9 virtual lanes: 8 data + 1 management
- 48K entry linear forwarding data base

ENHANCED INFINIBAND

- Hardware-based adaptive routing
- Hardware-based reliable transport
- Hardware-base reliable multicast
- Up to 6 virtual subnets
- Fine grained end-to-end QoS
- Port mirroring
- Supports Jumbo frame up to 10KB

COMPATIBILITY

CPU

- AMD X86
- Intel X86
- PowerPC and MIPS

PCI EXPRESS INTERFACE

- PCIe Base 2.0 compliant, 1.1 compatible
- 2.5GT/s or 5GT/s link rate x4
- Auto-negotiates to x4, x2, or x1
- Support for MSI/MSI-X mechanisms

CONNECTIVITY

- Interoperates with InfiniBand HCAs
- Drives copper cables or backplanes

MANAGEMENT AND TOOLS

- Supports Open SM or third-party subnet managers
- Diagnostic and debug tools

ADAPTER SILICON

Ordering Part Number	InfiniBand 4X Port Speed	Power (36 4X ports, typ)
MT48436A0-FCC-S	10Gb/s	TBD
MT48436A0-FCC-D	10, 20Gb/s	TBD
MT48436A0-FCC-Q	10, 20, 40Gb/s	TBD

Mellanox TECHNOLOGIES
2900 Stender Way, Santa Clara, CA 95054
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com

INFINIBAND™ TRADE ASSOCIATION