Mellanox Spectrum™-2 Ethernet Switch IC

Spectrum·2 ™ †

Highest Performance 6.4Tb/s Ethernet Switch, Enabling the Most Efficient Interconnect Solutions for Modern Data Centers

Built on the field proven and test winner Spectrum™ switch ASIC, Spectrum-2 offers increased bandwidth capacity as well as support for new 200/400GbE Ethernet standards.

Optimized for data centers and cloud applications, Spectrum-2 provides an unmatched combination of performance, power efficiency, scalability and flexibility. A comprehensive feature set and software ecosystem enable partners to develop and deploy their solutions with a short time-to-market and best ROI.

The growing demand for increased network bandwidth is one of the most critical issues facing data centers today. Applications such as machine learning and artificial intelligence, distributed storage and high-definition video streaming, create huge amounts of traffic that need to be processed and analyzed in real time. Spectrum-2 enables the use of data by supporting a wide range of Ethernet speeds – 10/25/40/50/100/200/400GbE. Supporting up to 16 ports of 400GbE, 32 ports of 200GbE, 64 ports of 100GbE, or 128 ports of 10/25/50GbE, Spectrum-2 can also be flexibly configured to any port combinations of these speeds.

Spectrum-2 is manufactured using 16nm technology with low power consumption of less than 7 watts per 200GbE port. It delivers an industry-leading cut-through latency of less than 300ns, together with 6.4Tb/s switching capacity and full wire-speed performance for VXLAN bridging and routing at all packet sizes.

BUFFERING AND QOS

Spectrum-2 uses an on-chip 42MB packet buffer. Smart architecture ensures that the entire buffer can be dynamically shared between all ports, unlike the competition in which the buffer is segregated, with each portion only accessible to some of the ports. This optimizes buffer utilization and enables Spectrum-2 to deliver superior performance in a reliable and predictable manner under varying network conditions.

Spectrum-2 also supports innovative and advanced active queue management and congestion control capabilities, such as Adaptive Flow Prioritization, a mechanism which dynamically identifies short (mice) latency-sensitive flows and prioritizes them over long (elephant) flows. Other enhancements include innovative features to detect congestion, and selective WRED/ECN which improves application performance by protecting certain packets, such as TCP control packets, from being discarded.

These capabilities enable Spectrum-2 to maximize RoCE (RDMA over Converged Ethernet) applications performance, in both lossless or lossy network configuration.

SCALABILITY

Modern data centers and hyperscale data centers require large forwarding tables. With the current shift from virtual machine-based architectures to container-based architectures these requirements are increased by an order of magnitude or more, and switches must have the forwarding capabilities

HIGHLIGHTS

– 6.4Tbps of Ethernet switching and IP routing

– Integrated 128 x 50Gb/s PAM4 SerDes

– Flexible port configurations:
  – Up to 16 400GbE ports
  – Up to 32 200GbE ports
  – Up to 64 100GbE ports
  – Up to 128 10/25/50GbE ports

– Full wire-speed performance for VXLAN bridging and routing at all packet sizes

– Sub 300ns, industry leading, true cut-through latency

– Low power consumption of less than 7 watts per 200GbE port

– 42MB dynamically-shared, flexible packet buffering

– Advanced active queue management and congestion control

– Half a million forwarding entries

– Highest RoCE performance

– 10X higher policy engine scale vs. competition

– Optimized for Software Defined Networking; support for OpenFlow version 1.0 to 1.4 and beyond

– Comprehensive overlay and tunneling support including VXLAN, NVGRE, Geneve and MPLS

– Advanced static and dynamic load balancing

– Extensive visibility capabilities at highest scale

– Programmable forwarding pipeline enables support of future protocols

– Extensive software development tools and standard open APIs

†For illustration only. Actual products may vary.

to support them. The Spectrum-2 on-chip forwarding table consists of 512K entries which can be flexibly shared between all forwarding entities such as MAC addresses, IP routes, ECMP routes and others. By using a revolutionary Algorithmic TCAM mechanism, ACL rules can be stored in the main forwarding table, in addition to a dedicated TCAM memory, to provide best-in-class ACL rules scalability. Spectrum-2 forwarding capacity may also be enlarged with up to 2M additional IP routes by using external memory to provide Internet scale router capabilities.

LOAD BALANCING

In order to support the ever increasing traffic demand while maintaining network efficiency, advanced load-balancing techniques are paramount. Spectrum-2 supports static, consistent and resilient load balancing and also provides users with full flexibility to control the load balancing hash functions. In addition, Spectrum-2 supports dynamic load balancing by implementing Adaptive Flowlet Routing. This capability is based on the principle that every flow actually consists of tiny bursts of traffic. These bursts (or flowlets) are separated by time gaps that are long enough to allow flowlets to take different paths without creating packet re-ordering. Hence, instead of statically assigning a flow to a single path, each flowlet can be sent over a different less-congested path, resulting in an overall more balanced network.

VISIBILITY

As networks grow in size and complexity, providing complete and in-depth visibility to the network becomes essential. Spectrum-2 supports an extensive set of telemetry capabilities that provide operators with full visibility into their network and allow them to monitor, diagnose and analyze every aspect of it. These capabilities include flexible packet mirroring trigger points, a large quantity of general purpose counters which can be assigned to any routing entity, sFlow, NetFlow and IPFIX, in-band network telemetry, and more.

PROGRAMMABILITY

Spectrum-2 implements a complete set of network protocols within the switch ASIC in the most efficient way, providing users with

all the functionality they need out-of-box. Moreover, Spectrum-2 includes a flexible parser which can be programmed to process new types of future protocols and custom tunnels, as well as a packet editor which enables users to add meta-data to packets, such as in-band network telemetry or any other information needed for new innovative instrumentation features. These programmable capabilities, along with the flexible ACL-based policy engine, which was already introduced in the first generation Spectrum™, reduce time-to-market and future proof the data center switching infrastructure.

SOFTWARE

Mellanox offers a rich set of software development tools and open APIs that enable users and partners to develop fully functional switching solutions on top of Spectrum-2 at a low-cost with a short time-to-market. Users have the flexibility to choose the model that best fits their needs, including Spectrum-2 SDK, OCP SAI (Switch Abstraction Interface) or Linux® switchdev driver which enables native Linux operating systems and applications on top of Spectrum-2.
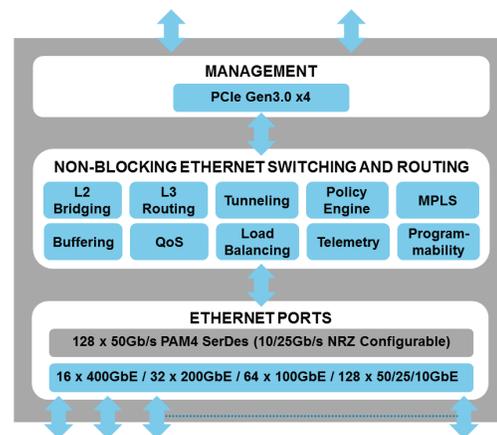
High Level Diagram

Table 1 - Part Number and Description

OPN	Description
MT53132-NCCR-V	Spectrum™-2, 32 Port Ethernet 200GbE Switch IC, No Crypto (RoHS R6)

COMPATIBILITY

Ethernet
– 16 x 400GbE ports, 32 x 200GbE ports, 64 x 100GbE ports, 128 x 1/10/25/50GbE ports
– DCB (PFC, ETS, DCBX)

CPU Interoperability
– PowerPC, Intel® x86, AMD x86 and MIPS

Host Interface
– PCIe Gen3.0, 4 lanes of 2.5GT/s, 5GT/s or 8GT/s

Interoperability
– Full interoperability with standard Ethernet NICs and switches

SerDes
– 128 x 50Gb/s PAM4 (802.3bs, 802.3cd)
– 10Gb/s (802.3ba, 802.3ae) and 25GbE (802.3bj, 802.3bm) NRZ operational modes

Connectivity
– Drives active/passive copper cables, fiber optics, PCB or backplanes

I/O Specifications
– SPI flash interface
– I²C, SMBus, MDIO
– IEEE 1149.1 boundary-scan JTAG
– IEEE 1588 1-step and 2-step

Package
– 57.5mm x 57.5mm HFCBGA

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