Mellanox Spectrum®-3 Ethernet Switch IC

12.8Tb/s Ethernet Switch Optimized for Cloud, Storage and Machine Learning

†

Spectrum®3

Mellanox Spectrum®-3 is the 4th generation of Mellanox Ethernet switch ASICs and provides an unmatched combination of performance, virtualization and telemetry capabilities in a 12.8Tbps Ethernet switch silicon with 256 x 50G PAM4 SerDes. Providing maximum flexibility, the SerDes and can act independently, in pairs, quads or groups of 8, to operate at all speeds between 1GbE and 400GbE. The full range of the Mellanox Spectrum® family of products have a strong software ecosystem that can be leveraged by partners to develop and deploy solutions with a rapid time-to-market.

Consistent Performance

Consistent performance and fair bandwidth sharing are critical for multi-tenant cloud, distributed storage workloads as well as hardware-accelerated data transfers using technologies such as GPUDirect® or NVMe-oF. Mellanox Spectrum-3 features a 64MB fully-shared and monolithic packet buffer that is dynamically shared across all ports. This provides excellent burst absorption with true port-to-port cut-through latency.

The on-chip packet buffer supports high bandwidth packet reads/writes, enabling consistent and predictable cut-through performance. The monolithic packet buffer architecture simplifies buffer management and traffic scheduling, while also enabling fair resource sharing. Additionally, Mellanox Spectrum-3 supports intelligent congestion management mechanisms, including Explicit Congestion Notification (ECN) that enable robust RoCE-based, hardware-accelerated data transport and load balancing, such as adaptive routing and Priority-based Flow Control (PFC).

Advanced Network Virtualization

Mellanox Spectrum-3 is designed with a highly-flexible programmable packet pipeline that enables advanced network virtualization. The hybrid packet-forwarding pipeline comprises optimized blocks, which are streamlined in hardware to deliver the highest performing traditional functionality. The pipeline also is made up of programmable blocks for supporting new functionality. Leveraging this capability, Mellanox Spectrum-3 can deliver industry-leading packet-processing rates at scale, and concurrently with programmability and Advanced Network Virtualization.

Mellanox Spectrum-3 supports single-pass VXLAN routing and provides 10X better VXLAN VTEP and tunnel scale. In addition to VXLAN, Mellanox Spectrum-3 supports a myriad of virtualization technologies including SRV6, Network Address Translation (NAT), and more.

Highlights
- 12.8Tb/s of Ethernet switching and IP routing
- Integrated 256 x 50Gb/s PAM4 SerDes
- Flexible port configurations:
  - Up to 32 400GbE ports
  - Up to 64 200GbE ports
  - Up to 128 100/50/25/10/1 GbE ports

Consistent Performance
- Industry leading, true, cut-through latency
- Low power consumption per port
- 64MB dynamically-shared, flexible packet buffering
- Robust RoCE data path for NVMe-oF and GPUDirect®
- Adaptive Routing congestion avoidance

Advanced Network Virtualization
- Single-pass VXLAN bridging and routing
- Centralized VXLAN routing
- Comprehensive overlay & tunneling support: VXLAN-GPE, NVGRE, Geneve, NSH and MPLS/IPv6 based Segment Routing
- Network Address Translation (NAT)

Event-based Telemetry
- Reduced Mean Time to Recovery/Innocence
- Actionable telemetry with What Just Happened (WJH)
- Hardware-based buffer histograms
- Streaming telemetry
- In-band network telemetry

Cloud-Scale
- Half a million flexible entries shared across routes, MAC, MPLS, etc.
- Up to 1M on-chip IPv4 routes; extendable up to 4M
- Half a million flow tracking counters
- 100X more ACLs
- 10X better VXLAN VTEP and tunnel scale

Synchronization
- Precision Time Protocol (PTP) network synchronization

†For illustration only. Actual products may vary.

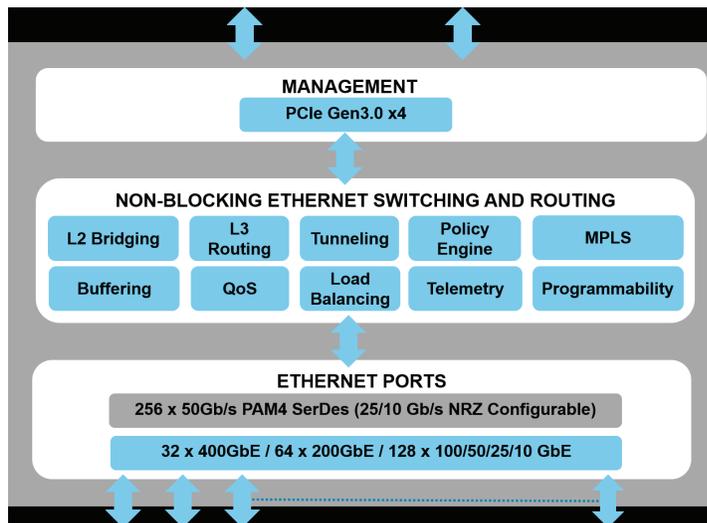# Industry-Leading Visibility with What Just Happened™ (WJH)

Mellanox Spectrum-3 provides rich, contextual and actionable network-based visibility data that seamlessly integrates with open APIs at the SDK layer. Referred to as 'What Just Happened' (WJH), this feature allows network operators to finally get instant answers to the 5 W's—When, What, Who, Where and Why, as soon as an issue arises. WJH can easily be extended and customized to support both third-party and open-source tools.

'What Just Happened' is open and extendable, making it ideal for integration into third-party network monitoring and analysis tools. For those wanting a simple turn-key solution, Mellanox NEO® is an easy way to utilize the benefits of What Just Happened.

Beyond providing hardware-based mechanisms that measure and summarize packet buffer dynamics, and over half a million purpose-built flow tracking counters, Mellanox Spectrum-3 also supports traditional hardware-based visibility features such as sFlow, in-band and streaming telemetry.

MANAGEMENT
PCIe Gen3.0 x4

NON-BLOCKING ETHERNET SWITCHING AND ROUTING
L2 Bridging | L3 Routing | Tunneling | Policy Engine | MPLS
Buffering | QoS | Load Balancing | Telemetry | Programmability

ETHERNET PORTS
256 x 50Gb/s PAM4 SerDes (25/10 Gb/s NRZ Configurable)
32 x 400GbE / 64 x 200GbE / 128 x 100/50/25/10 GbE

High Level System Block Overview

Choice of 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 Mellanox Spectrum-3 at low-cost, with short time-to-market. Users have the flexibility to choose the model best suited to their needs, including Mellanox Spectrum-3 SDK, OCP SAI (Switch Abstraction Interface) or Linux® Switch driver, which enables native Linux operating systems and applications on top of Mellanox Spectrum-3.

Table 1 - Ordering Part Number and Description

OPN	Description
MT54132A0-NCCR-W	Mellanox Spectrum®-3, 256x50Gb/s PAM4, 32 Port 400GbE Switch IC

Compatibility

Ethernet
– 32 x 400GbE ports, 64 x 200GbE ports, 128 x 100/50/25/10/1 GbE ports
– DCB (PFC, ETS, DCBX)
CPU Interoperability
– PowerPC, Intel® x86 and AMD x86

Host Interface
– PCIe Gen 3.0, 4 lanes of 8GT/s, 5GT/s or 2.5GT/s
Interoperability
– Full interoperability with standard Ethernet NICs and switches

SerDes
– 256 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
– 70mm x 70mm HFCBGA

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

60351PB
Ver 1.2