Interconnect Your Future

Sebastian Kalcher – Mellanox Technologies

May 2019







Mellanox Overview

\$1.09B 2018 Revenue

\$5.01 2018 Non-GAAP EPS

\$265M 2018 Cash From Operations

\$439M 2018 Ending Cash & Investments

~2,500 Employees worldwide



Ticker: MLNX

Company Headquarters: ■ Yokneam, Israel Sunnyvale, California Worldwide Offices





Mellanox Has A Long Track Record Of **Delivering Breakthrough Technology**

As Bandwidth Goes Up, Mellanox's Competitive Advantage From its Leading Offload Engine Capability Increases

Leading Edge Products Across Technology Generations

- **#1** in InfiniBand Since **2004**
- **#1** in Ethernet NIC in **2016 / 2017**
- **#2** in Ethernet Switch Silicon in **2017**



56Gb/s

100Gb/s



200/400Gbps

Mellanox Accelerates Leading HPC and AI Systems

World's Top 3 Supercomputers





Summit CORAL System World's Fastest HPC / AI System 9.2K InfiniBand Nodes







Sierra CORAL System #2 USA Supercomputer 8.6K InfiniBand Nodes







41K InfiniBand Nodes







国家超级计算无锡中则 National Supercomputing Center in Wuxi

Wuxi Supercomputing Center Fastest Supercomputer in China

Mellanox Accelerates Leading HPC and AI Systems (Examples)





Fastest HPC / AI System in Japan 1.1K InfiniBand Nodes

500



The world's Fastest Industry Supercomputer 1.6K InfiniBand Nodes



JUWELS Supercomputer 2.6K InfiniBand Nodes



NASA Ames Research Center **20K InfiniBand Nodes**





World's leading Industry Supercomputer 4.6K InfiniBand Nodes





UNIVERSITY OF TORONTO

JÜLICH

Fastest Supercomputer in Canada **Dragonfly+** Topology **1.5K InfiniBand Nodes**















5

HPC and AI Needs the Most **Intelligent Interconnect**





HDR 200G InfiniBand Accelerates Next **Generation HPC/AI Systems**





Lawrence Livermore National Laboratory



TEXAS ADVANCED COMPUTING CENTER





The Need for Intelligent and Faster Interconnect

Faster Data Speeds and In-Network Computing **Enable Higher Performance and Scale**



Must Wait for the Data **Creates Performance Bottlenecks**



Analyze Data as it Moves! **Higher Performance and Scale**



End-to-End Solutions for All Platforms

Unleashing the Power of all Compute Architectures



Highest Performance and Scalability for Intel, AMD, IBM Power, NVIDIA, Arm and FPGA-based Compute and Storage Platforms at 10, 20, 25, 40, 50, 100, 200 and 400Gb/s Speeds





Accelerating All Levels of HPC/AI Frameworks







HDR InfiniBand







Highest-Performance 200Gb/s InfiniBand Solutions

Adapters	·Connect X 6	200Gb/s Adapter, 0.6us latency 215 million messages per second (10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)	
Switch	<mark>Mellanox:</mark> Quantum ∔	40 HDR (200Gb/s) InfiniBand Ports 80 HDR100 InfiniBand Ports Throughput of 16Tb/s, <90ns Latency	S
SOC	BlueField	System on Chip and SmartNIC Programmable adapter Smart Offloads	
Interconnect	·LinkX	Transceivers Active Optical and Copper Cables (10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)	(
Software	• HPC-X [™]	MPI, SHMEM/PGAS, UPC For Commercial and Open Source Applications Leverages Hardware Accelerations	







ConnectX-6 HDR InfiniBand Adapter

Leading Connectivity

- 200Gb/s InfiniBand and Ethernet
 - HDR, HDR100, EDR (100Gb/s) and lower speeds
 - 200GbE, 100GbE and lower speeds
- Single and dual ports
- 50Gb/s PAM4 SerDes

Leading Performance

- 200Gb/s throughput, 0.6usec latency, 215 million message per second
- PCIe Gen3 / Gen4, 32 lanes
- Integrated PCIe switch
- Multi-Host up to 8 hosts, supporting 4 dual-socket servers

Leading Features

- In-network computing and memory for HPC collective offloads
- Security Block-level encryption to storage, key management, FIPS
- Storage NVMe Emulation, NVMe-oF target, Erasure coding, T10/DIF







HDR InfiniBand Switch: QM8700, 1U Series

40 QSFP56 ports (50G PAM4 per lane)

- 40 ports of HDR, 200G
- 80 ports of HDR100, 100G

Superior performance

- 90ns latency
- 390M packets per sec (64B)
- 16Tb/s aggregate bandwidth

Superior resiliency

- 22" depth
- 6 fans (5+1), hot swappable
- 2 power supplies (1+1), hot swappable







HDR InfiniBand Switch: CS8500, Modular Series

800 QSFP56 ports

- 800 ports of HDR, 200G
- 1600 ports of HDR100, 100G

Superior performance

- 300ns latency
- 320Tb/s aggregate bandwidth
- LCD Tablet IO panel

Water-cooled solution

- Liquid Liquid 4U CDU
- Liquid Air 42U (350mm wide) stand alone HEX
- OC 35C (air) or 40C (water) operating air range

an shug	
	UK!
	Perso
	Pro-
	10 million
1 1 (mm	







Artificial Intelligence and Deep Learning Performance







Mellanox Accelerates Leading AI Platforms (Examples)

Fastest AI Supercomputer in the World ~27000 GPUs Dual-Rail Mellanox 100G EDR InfiniBand

NVIDIA® DGX-1 and DGX-2 Platforms 2 Petaflops of AI Performance Mellanox 100G InfiniBand and Ethernet

Facebook AI infrastructure Mellanox 100G EDR InfiniBand







The set of the set of







Mellanox Accelerates Record-Breaking AI Systems



NVIDIA DGX SATURNV

- 124 DGX-1 nodes interconnected by 32 L1 TOR Switches, in 2016
- Mellanox 36 port EDR L1 and L2 switches, 4 EDR per system
- Upgraded to 660 NVIDIA DGX-1 V100 Server Nodes, in 2017
- 5280 V100 GPUs, 660 PetaFLOPS (AI)





IB Switch



Accelerating the Machine Learning Data Pipeline







Mellanox Accelerates TensorFlow



Unmatched Linear Scalability at No Additional Cost

50%

Better Performance







20

TensorFlow

Mellanox Accelerates NVIDIA NCCL 2.0



Performance Improvement

with NVIDIA[®] DGX-1 across 32 NVIDIA Tesla V100 GPUs **Using InfiniBand RDMA** and GPUDirect[™] RDMA

CNTK scaling ResNet50, images/s







© 2019 Mellanox Technologies | Confidential

Mellanox Accelerates TensorFlow 1.5



100G is a Must For Large Scale Models

7.5X

Faster Training with 100G





TensorFlow





Mellanox Ethernet Switch Systems







Open Ethernet – The Freedom to Optimize





Network-OS Choice

Open APIs



Automation

End-to-End Interconnect

© 2019 Mellanox Technologies | Confidential

Spectrum: The Ultimate 25/100GbE Switch

The only predictable 25/50/100Gb/s Ethernet switch

Full wire speed, non-blocking switch

Doesn't drop packets per RFC2544

ZPL: ZeroPacketLoss for all packets sizes







Open Ethernet SN2000 Series

SN2700 – 32x100GbE (up to 64 x 50/25/10GbE) The Ideal 100GbE ToR / Aggregation

SN2410 – 8x100GbE + 48x25GbE 25GbE → 100GbE ToR

SN2100 – 16x100GbE ports (up to 32 x50GbE , 64x25/10GbE) Ideal storage / Database 25/100GbE Switch

SN2010 – 4x100GbE + 18x10/25GbE Ideal Hyperconverged Switch 10/25GbE \rightarrow 100GbE half width ToR

- Predictable Performance
- Fair Traffic Distribution for Cloud
- Best-in-Class Throughput, Latency, Power Consumption
- Zero Packet Loss

300ns SN2700 - 169W SN2410 - 165W SN2100 - 94W

















Performance

- 6.4Tb/s total throughput
- Support all standard speeds, 10GbE 400GbE
- 300ns industry leading true cut-through latency
- Lowest power consumption, ~4W per 100GbE port
- 42MB fully shared packet buffer

the all and a start of the start of the start of the

Advanced queue management and congestion control

Scalability

- 512K on-chip, dynamically shared forwarding table
- 2M additional IP routes w/ external memory
- SpACE Spectrum Algorithmic Classifier Engine
 - Unprecedented ACLs scale, up to 512K rules

WALLOW AND A DEWELL

Visibility

- 512K general purpose counters
- 8 SPAN/ERSPAN mirroring agents
- Multiple mirroring triggering points
- Advanced buffer monitoring
- In-band network telemetry

Flexibility

- Flexible parser, supports custom and emerging protocols
- Packet modifier, add / change packet content
- Custom tunnels over IP network





Open Ethernet SN3000 Series



Introducing speeds from 1GbE through 400GbE

SN3800 – 64x100GbE (QSFP-28) 2U

Ideal Spine/Super-spine

SN3700 – 32x200GbE (QSFP-56) 1U

High Density Leaf/200GbE Spine/Super-spine

SN3510 – 48x25/50GbE (SFP-56) + 6x400GbE (QSFP-DD) 1U

Ideal Leaf

SN3200 - 16x400GbE (QSFP-DD) 1/2U



Compact ¹/₂ U Switch









Cloud





© 2019 Mellanox Technologies | Confidential

Challenges for Cloud Networks

Provisioning Delays	Delays servicesDelays revenue
Capex & Opex	Vendor lock increases costsLabor-intensive
Complicated	Multitenant Network SegmentationGuaranteed SLAs
Obscured Operations	Lacks central controlLacks visibility



Cloud IT Requires Smarter Networks

30

The Mellanox Cloud Network Solution

Scalable	 Leaf/Spine scales easily 10K nodes in 2 tiers 100K nodes in 3 tiers
Automated	 OpenStack Integration NSX & Nutanix Integration RoH for Container-optimized deployments
Network Virtualization	 EVPN – VXLAN without Controller Best VXLAN scale and functionality
Telemetry	 Real-time network visibility Hardware accelerated streaming telemetry





Spectrum is the Best for IaaS/PaaS/SaaS

Cloud-scale Performance

- Best in class performance: PPS, latency, buffering
- Scales to 10K nodes in 2 tiers
- Cloud-scale Virtualization
 - VXLAN Routing in hardware
 - No loopbacks
 - VXLAN Scale
 - 100X more Tunnel End Points
 - 10X more VXLAN tunnels
 - VXLAN that works with ROCE
- Cloud-scale Automation
- Cloud-scale Telemetry
 - Real time visibility with hardware base histograms
 - What Just Happened (WJH)



6X more VTEPs = 6X more Server Racks











Storage & HCI





Leadership in Storage Platforms





© 2019 Mellanox Technologies | Confidential

Storage Architecture In Transition







More Ethernet Storage Traffic

Changes	Effects
Flash, Faster servers	Faster networking; 10/25/40/50/1000
Social/Mobile/Video	Huge data growth; more file and obje
Hyperconverged	Distributed "Server-SAN" on Ethernet
Cloud	Virtualization, software-defined, price
Big Data	File and distributed storage
Distributed applications	More east-west traffic

Bottom Line: More Ethernet Storage Traffic Over Scale-out Infrastructure





GbE

ect content

e pressure



Storage & Connectivity Evolution



Fibre Channel Port Shipments (in Millions)





Source: Crehan Research, Host Adapter Port Shipments, January 2018 © 2019 Mellanox Technologies | Confidential 39







© 2019 Mellanox Technologies | Confidential

Spectrum Unlocks the Maximum Flash Performance



Flash Storage is Getting a Lot Faster!







Spectrum the Best Storage Switch

Ethernet Storage Fabric needs dedicated storage switches





- 2 Switches in 1RU
- Storage/HCI port count
- **Zero Packet Loss**
- **Low Latency** \checkmark
- **RoCE optimized switches (NVMe-oF)** \checkmark
- **NEO for Network automation/visibility** \checkmark
- Native SDK on a container
- **Cost optimized**
- **NOS** alternatives









