

# The role of connectivity in the Data Center:

### CPU acceleration and offloading

Ramiro Alvarez – Feb 2020





Source: Crehan Research, Mellanox internal estimates



# **Mellanox Main Technology Trends**

Network Performance and Flexibility Intelligence Moving to the Network Scale-Out Architectures











### Mellanox TECHNOLOBIES

# Data Growth is Driving Mellanox Growth

Exponential Growth of Data & Real Time Analytics Requires Mellanox *High Throughput, Low Latency, Smart Interconnect* Solutions



# Leading Supplier of InfiniBand and Ethernet End-to-End Interconnect Solutions

The Smart Choice for Intelligent Compute and Storage Platforms



© 2020 Mellanox Technologies | Confidential

**Iellanox** 

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

### **Highest Performance HDR 200G InfiniBand**

Adapters	200Gb/s, 0.6us Latency 215 Million Messages per Second (10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)	
Mellanox: Quantum H Switch	40 HDR (200Gb/s) Ports 80 HDR100 (100Gb/s) Ports 16Tb/s Throughput, 15.6 Billion msg/sec	SHARP
Interconnect	Transceivers Active Optical and Copper Cables (10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)	



### **ConnectX-6 HDR InfiniBand Adapter**

#### Leading Connectivity

- 200Gb/s InfiniBand and Ethernet
  - HDR, HDR100, EDR and lower speeds
  - 200GbE, 100GbE and lower speeds
- Single and dual ports

#### Leading Performance

- 200Gb/s throughput, 0.6usec latency,
- 215 million message per second
- PCIe Gen3 / Gen4
- Integrated PCIe switch
- Multi-Host

#### 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



· Connect X. h

\*





### HDR InfiniBand Switch: QM8700, 1U Series



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

#### Superior performance

- 130ns 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



© 2020 Mellanox Technologies | Confidential

lellanox

### CS8500, Modular HDR Series

#### 800 QSFP56 ports

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

#### Superior performance

- 300ns latency
- 320Tb/s aggregate bandwidth
- Max power : 800p up to 22KW
- 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



### **Highest Performance and Scalability for Exascale Platforms**



Mellanox

### **Mellanox NIC and SmartNIC Portfolio**



#### Innova FPGA SmartNIC

- Highly programmable
- Performant Alternative
- For companies with Verilog FPGA development skills

#### **ConnectX SmartNIC**

- Best performance for price
- Built-in hardware offloads
- Extra flexibility, efficiency and performance

#### BlueField Software Programmable SmartNICs

- Highly customizable
- Leverage hardware accelerations
- Full programmability

















### **Mellanox Solutions for Efficient Cloud Deployments**



facebook. III Microsoft ビンAlibaba Group Tencent 腾讯 P PayPal 学JD.COM Bai 浴音度 ② NVIDIA。

© 2020 Mellanox Technologies | Confidential



### **Mellanox ConnectX for NFV Solutions and Telco Cloud**



- From 10Gb/s to 200Gb/s
- Best DPDK 148Mpps (100Gbs Line Rate)
- Virtual Switch Offload
  - Up to 10X message rate with ASAP<sup>2</sup>
  - VNF Acceleration
- Zero CPU % with ASAP<sup>2</sup>



### Feature Rich

- Stateless Offloads, tunneling: VXLAN, GRE
- Any Virtual Switch
- Header re-write, Hair-Pin
- SR-IOV: HW based QoS, High Availability
- RDMA



#### Variety of Solutions

- SR-IOV or VirtIO acceleration
- Support custom vSwitch
- Control plane in kernel or user-space

#### Community

- Standard software
- Upstream
- Inbox
- Key partnerships





### **DPDK with Mellanox - Industry Leading Performance**

Highest Message Rate in the Market!!!

66% lower latency compared to competition

#### Strong influence DPDK Community





### Mellanox ASAP<sup>2</sup>

### Accelerated Switching and Packet Processing

- Data plane offload
- Control Plane runs on main host CPU





Accelerated Switching & Packet Processing





### **OVS over DPDK -vs- OVS Offload - ConnectX-5**

Test	ASAP <sup>2</sup>	OVS DPDK	Benefit
1 Flow VXLAN	66M PPS	7.6M PPS (VLAN)	8.6X
60K flows VXLAN	19.8M PPS	1.9M PPS	10.4X

ConnectX-5 provide significant performance boost Without adding CPU resources







### **Mellanox Accelerates Block, File and Object**



© 2020 Mellanox Technologies | Confidential

### **Mellanox Powers Storage Frameworks**



19

fellanox



### **Addressing Latency Demand**

#### Physical Storage Devices Cut Latency



#### Mellanox Network Solutions Cut Latency



#### **Interconnect Latency Matters!**

Lowest Latency on the Market

© 2020 Mellanox Technologies | Confidential

### Storage is Getting a Lot Faster!



### Saving Precious CPU Cycles = Free up CPU for Compute



#### NVMe-oF over RDMA Enables Running more Compute and Customers





### **10X Higher Performance with GPUDirect™ RDMA**

- Accelerates HPC and Deep Learning performance
- Lowest communication latency for GPUs





© 2020 Mellanox Technologies | Confidential



### **Mellanox SmartNICs Make Secure Cloud Possible**





### **Mellanox ConnectX for Secure Solutions**

- Inline encryption of Data-in-Motion
  - Hardware based accelerator, offloads crypto operations
  - Inline processing of IPsec (100Gbs) and TLS crypto (200Gbps)
  - Support TLS1.2 and TLS1.3. TLS ciphers: AES-GCM 128/256bit keys
- Encryption of Data-at-Rest (Storage)
  - XTS-AES 256 / 512 bit keys
    - Different keys per user or flow, FIPS Compliancy
- Key management
  - Secure storage of encryption keys (private & public)
- Stateful rules checking for Connection Tracking
  - Hardware-powered stateful Firewall solutions
- Secure Boot (Hardware Root-of-Trust)
  - Cerberus compliant, SHA2
  - HMAC based in ConnectX-6 and RSA in ConnectX-6 Dx , 6 Lx







### Mellanox BlueField-2 - The Ultimate Co-Processor

#### **High Performance IPU**

- 8 Arm<sup>®</sup> A72 CPUs @ 2GHz-2.5GHz
- Embedded ConnectX-6 Dx controller
- Dual port 10→100G, 200G
- PCIe Gen 3.0/4.0 x16

#### Advanced Hardware Accelerations

- Networking : RDMA, CC, QoS
- Storage : NVMe, Dedup (SHA), CRC64
- Storage : (De)Compression, EC, RAID
- Virtualization : ASAP2, Virtio, SR-IOV





#### Scalability and Programmability

- SmartNIC and as a controller in system
- User's applications to run over Arm cores
- Integrated control and data planes

#### Security

- SmartNICs isolated from main host
- Secured Hardware (Root-of-Trust)
- Crypto offloads : IPsec, TLS, AES-XTS
- Regular Expression (RegEx) and DPI







# Spectrum Delivers Highest Data Center ROI

0.000

Flexibility	10, 25, 40, 50, 100, 200 and 400 Gigabit Speeds, Variety of Port Configurations	
Scalability	10X Better Scalability, Optimized Data Center Utilization and End-User Experience	
Programmability	Programmable Flexible Parser and Packet Modifier, Support User-Defined Protocols	
ТСО	1.3X Lower Power Consumption Versus Competition	
Performance	1.4X Lower Latency Versus Competition, Enable Real-Time Data Analysis and Insights	
Open Ethernet	SONIC Verwork 05 Straproute Switch	









### **Flexible Ethernet**

#### Old world

\$\$\$ on legacy protocols Best performance and stability Low feature velocity

## Hybrid



#### **Fully Programable**



Write everything from scratch Implement both standard and new applications Variable feature velocity

Legacy protocols don't change Application sand box for home grown needs Extended HW longevity High feature velocity

### DC Architecture – Mellanox Leaf/Spine Deployments







Connect out via spines, L3 GW on spines or above



BGP from the Host – Kubernetes VXLAN from the host- VMware , OpenStack, Kubernetes

#### L2 over Layer 3 VXLAN



Anycast L3 GW on TORs or FW as GW located on the border leaf. EVPN Type 5 for Routes out of the fabric.

### **Open Ethernet SN-2000 Switch Portfolio**



Energy efficiency

lower power consumption!

© 2019 Mellanox Technologies

### **Mellanox SN-3000 Series Switch Family**





- Predictable Performance
- Fair Traffic Distribution for Cloud
- Best-in-Class Throughput, Latency, Power Consumption
- Advanced Telemetry with WJH™
- 2X Forwarding Scale

**SN3700 – 32x200G** (Flexible Speeds: 32x200G, 64x100G, 128x50G) 50G PAM4 based Spine Switch



© 2019 Mellanox Technologies



### **Proactive Telemetry – What Just Happened**



fellanox



