Mellanox Adapters
Saddik EL ARGUIOUI
October, 2018
Agenda

- Mellanox Introduction

- Mellanox Solution
  - Adapters Solution
  - SmartNics Solution
Mellanox introduction
The Need for Intelligent and Faster Interconnect

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

CPU-Centric (Onload) vs. Data-Centric (Offload)

Must Wait for the Data Creates Performance Bottlenecks

Analyze Data as it Moves!

Onload Network

In-Network Computing
End-to-End Interconnect Solutions for All Platforms

Highest Performance and Scalability for
X86, Power, GPU, ARM and FPGA-based Compute and Storage Platforms
10, 20, 25, 40, 50, 56 and 100Gb/s Speeds

Smart Interconnect to Unleash The Power of All Compute Architectures
Exponential Data Growth Everywhere

Higher Data Speeds
Faster Data Processing
Better Data Security

SmartNIC
System on a Chip

Adapters
Switches
Cables & Transceivers
Main Technology Trends

Network Performance and Flexibility

Intelligence Moving to the Network

Scale-Out Architectures

HPC  Cloud  Telco  Storage  Financial  Machine Learning
## Highest-Performance 100Gb/s Interconnect Solutions

| ConnectX-5 | 100Gb/s Adapter, 0.6us latency  
175-200 million messages per second  
(10 / 25 / 40 / 50 / 56 / 100Gb/s) |
| SwitchIB-2 | 36 EDR (100Gb/s) Ports, <90ns Latency  
Throughput of 7.2Tb/s  
7.02 Billion msg/sec (195M msg/sec/port) |
| Spectrum™ | 32 100GbE Ports, 64 25/50GbE Ports  
(10 / 25 / 40 / 50 / 100GbE)  
Throughput of 3.2Tb/s |
| BlueField SoC | System on Chip and SmartNIC  
Programmable adapter  
Smart Offloads |
| LinkX™ | Transceivers  
Active Optical and Copper Cables  
(10 / 25 / 40 / 50 / 56 / 100Gb/s) |
| HPC-X™ | MPI, SHMEM/PGAS, UPC  
For Commercial and Open Source Applications  
Leverages Hardware Accelerations |
# Highest-Performance 200Gb/s Interconnect Solutions

## Connect X:6
- **Adapters**
  - 200Gb/s Adapter, 0.6us latency
  - 200 million messages per second
  - (10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)

## Mellanox Quantum
- **Switch**
  - 40 HDR (200Gb/s) InfiniBand Ports
  - 80 HDR100 InfiniBand Ports
  - Throughput of 16Tb/s, <90ns Latency

## Spectrum 2
- **Switch**
  - 16 400GbE, 32 200GbE, 128 25/50GbE Ports
  - (10 / 25 / 40 / 50 / 100 / 200 GbE)
  - Throughput of 6.4Tb/s

## BlueField SoC
- **System on Chip and SmartNIC**
  - Programmable adapter
  - Smart Offloads

## LinkX
- **Interconnect**
  - Transceivers
    - Active Optical and Copper Cables
    - (10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)

## HPC-X Software
- **MPI, SHMEM/PGAS, UPC**
  - For Commercial and Open Source Applications
  - Leverages Hardware Accelerations

---

© 2018 Mellanox Technologies | Confidential
Adapters solution
Converged Infrastructure Relies on Efficient Data Movement

- Multi-Host & eSwitch: Embedded hardware OVS switch – Advance Flow Steering Engine
- Virtual network acceleration (VXLAN, NVGRE, GENEVE)
- RDMA – Efficient Data Exchange - Low Latency, Low CPU Overhead

Efficient Data Movement (RDMA)

InfiniBand/RoCE @ 100Gb/s
10X Better Performance with GPUDirect™ RDMA

- Purpose-built for Acceleration of Deep Learning
- Lowest communication latency for acceleration devices
- No unnecessary system memory copies and CPU overhead
- Enables GPUDirect™ RDMA and ASYNC, ROCm and others
- InfiniBand and RoCE

GPUDirect™ RDMA, GPUDirect™ ASYNC
MPI Tag-Matching Offload Advantages

- 31% lower latency and 97% lower CPU utilization for MPI operations
- Performance comparisons based on ConnectX-5

Mellanox In-Network Computing Technology Deliver Highest Performance
Socket Direct - Maximizing Servers Performance

- Improving ROI
  - Higher Throughput
  - Lower Latency
  - Lower CPU Utilization
- Few Flavors (different harness length)
- Enable 100Gb/s on all Servers

Share our [White Paper](#) and [Video](#)
Socket Direct Adapter - Technology

- **Split 100Gb/s Adapters between 2 PCI slots**
  - Use two PCI x8 slots
  - Adapter and PCI extender connected by harness

- **Both CPU’s Directly Connected to Network**
  - Increases QPI bandwidth for applications
  - Enables GPU / peer direct on both slots
  - Improved bandwidth, latency and CPU utilization
Mellanox - Gold Member and Head Maintainer of DPDK project

- DPDK.org project has moved to the Linux Foundation
- Mellanox joins as **Gold** member with a seat in the Governing Board
- Mellanox is the Head Maintainer of dpdk project
  - Thomas Monjalon, Head Maintainer and DPDK Technical Board member is part of Mellanox
  - Two members in Technical Board Committee
- Mellanox assigned a full time engineer to work on DPDK CI and server administration
- The number of developers and patches contributed to DPDK.org by Mellanox is continuously growing

* As announced April 3rd 2017 - [link](http://example.com)
Mellanox Adapters Unique Network Intelligence

- Cloud
  - NVMe Emulation (for Compute)
  - ASAP² (Advance OVS Offload)
  - RoCE over VXLAN
  - Overlay Encap/Decap OVS Offload

- Storage
  - Storage Encryption Key Management
  - NVMeO/F
  - Erasure Coding T10/DIF

- Network
  - RDMA/ROCE
  - VXLAN
  - NVGRE

- HPC, AI
  - SHARP v2
  - TAG
  - Matching
  - CORE-Direct Vectors DCT
  - GPU Direct CORE-Direct
SmartNIC solution
Innova Smart and Flexible Adapters

Programmable

Crypto Offloads

A Configurable Adapter to Fit Your Needs
BlueField – Multi-Core System on a Chip (SoC)

BlueField: SoC Optimized for Storage and Networking

- **Scale Up Storage Appliance**
  - Leverage PCIe switch and NVMe over Fabric Offloads

- **Security Appliance and Smart NIC**
  - OVS Offload
  - Crypto Logic

- **Family of Products, Variety of Options**
  - 2/4/8/16 cores of A72 ARM
  - ConnectX-5 high-speed networking
  - Different number of network ports, speeds
  - 2 DDR4 channels
BlueField Intelligent Adapter

- Implementation of various types of security functions
  - Stateful L4-based firewall
  - Isolation - “compute in-front of the computer”
  - Encryption of data in motion

- Ethernet protocols accelerations
  - ASAP² OVS Offload - including custom implementation and extensions
  - ConnectX-5 Ethernet Offloads
  - RDMA over Converged Ethernet (RoCE)

Moving Features from Cloud/Switch into the Smart NIC
Thank You