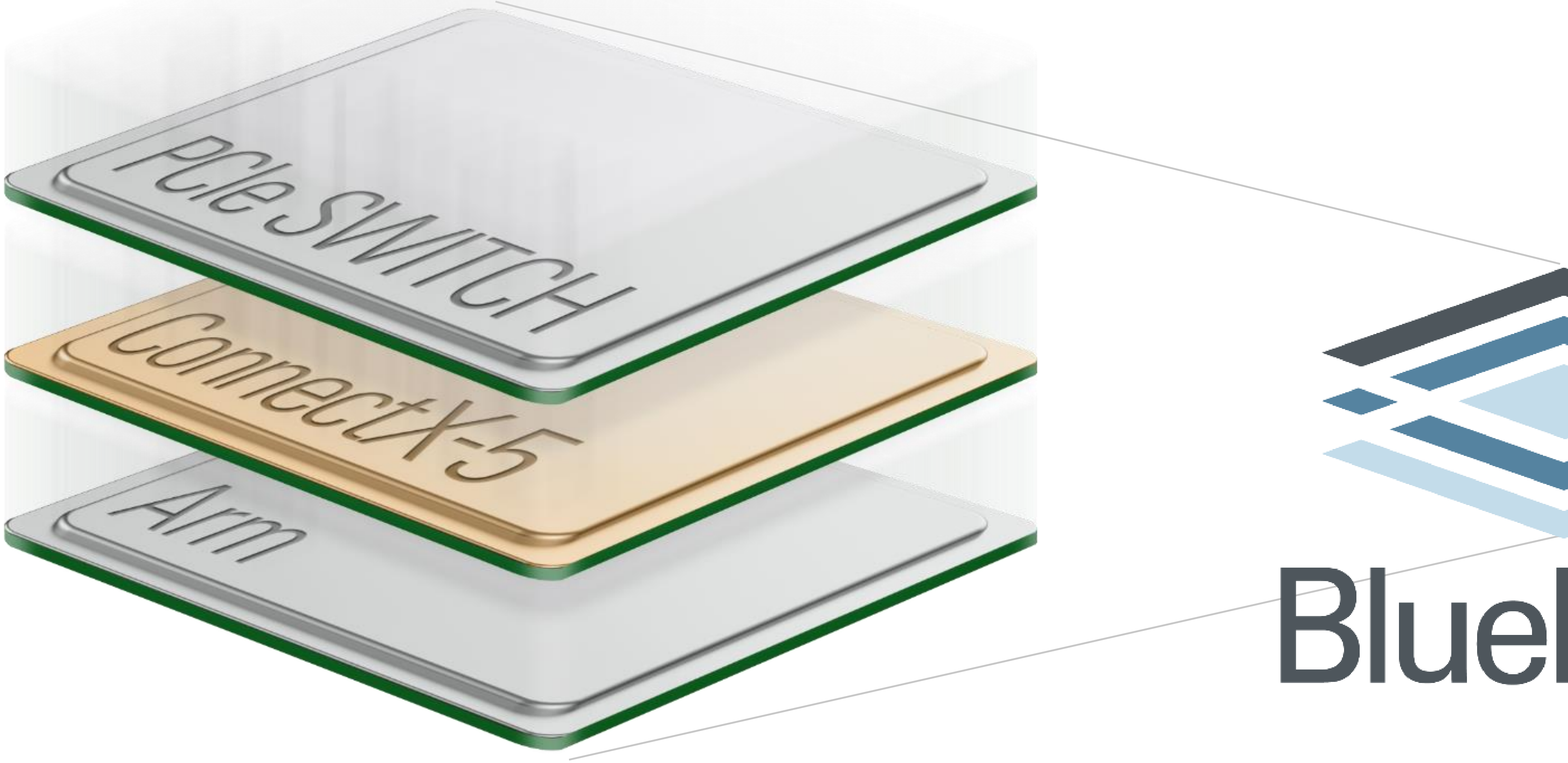
Improving Agility and Security Using Mellanox SmartNICs

Mark Taplin, Mellanox UK

November 2019

© 2019 Mellanox Technologies

BlueField – The Premier System on Chip (SoC) Solution

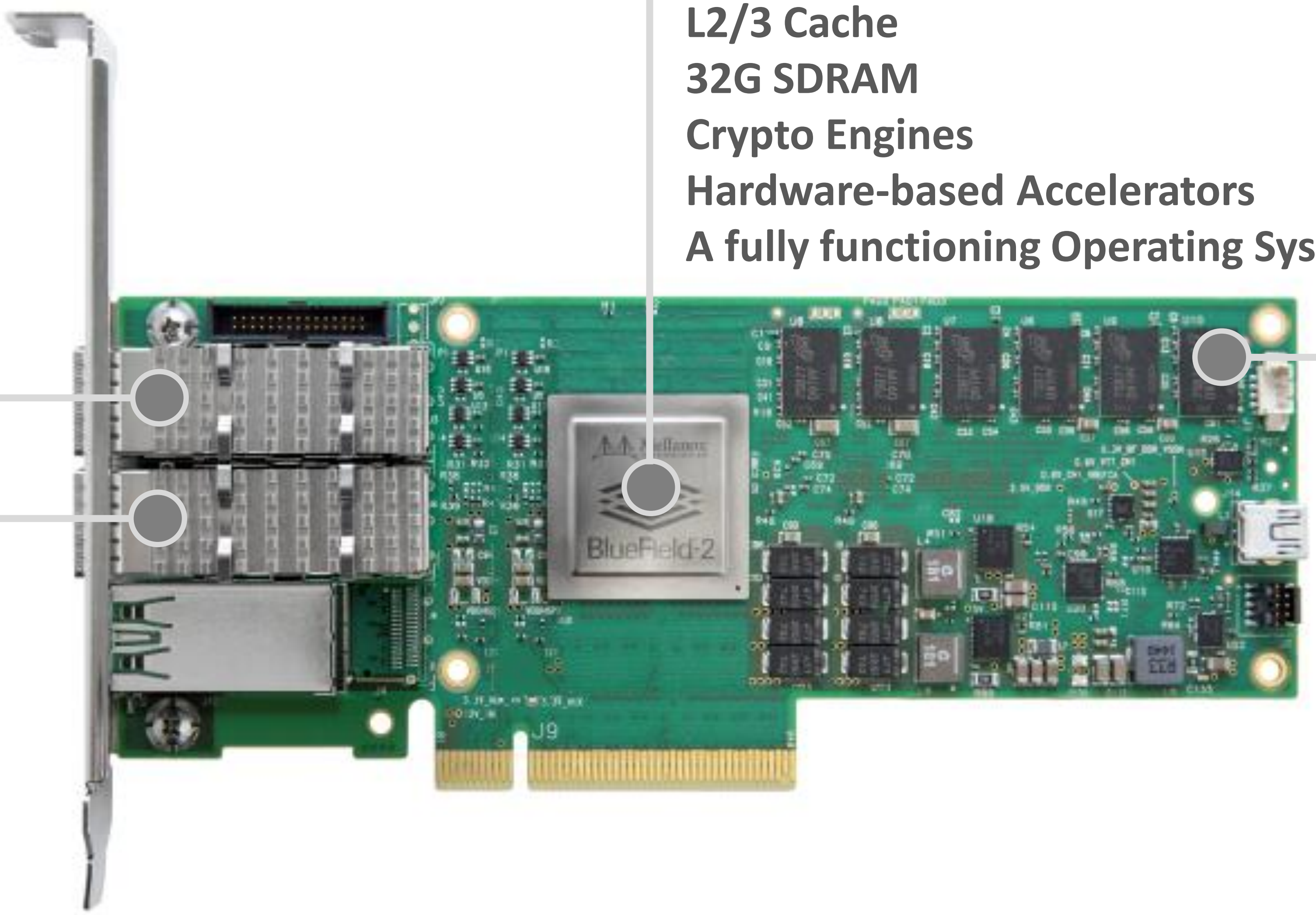


BlueField

BlueField – Field of Advantages

© 2019 Mellanox Technologies

BlueField SmartNIC is a Computer



8/16 x ARM CPU
L2/3 Cache
32G SDRAM
Crypto Engines
Hardware-based Accelerators
A fully functioning Operating System

Network Port #1

Network Port #2

Memory

© 2019 Mellanox Technologies

BlueField Software Overview

Software Packaging

- Standalone deliveries (bootloader, kernel, root file system, OFED)
- Image binaries and source code
- Host drivers (SmartNIC) and OpenBMC

Linux Distribution

- Yocto Poky (BlueOS)
- CentOS Reference
- Ubuntu 18.04 Commercial Distro

Documentation

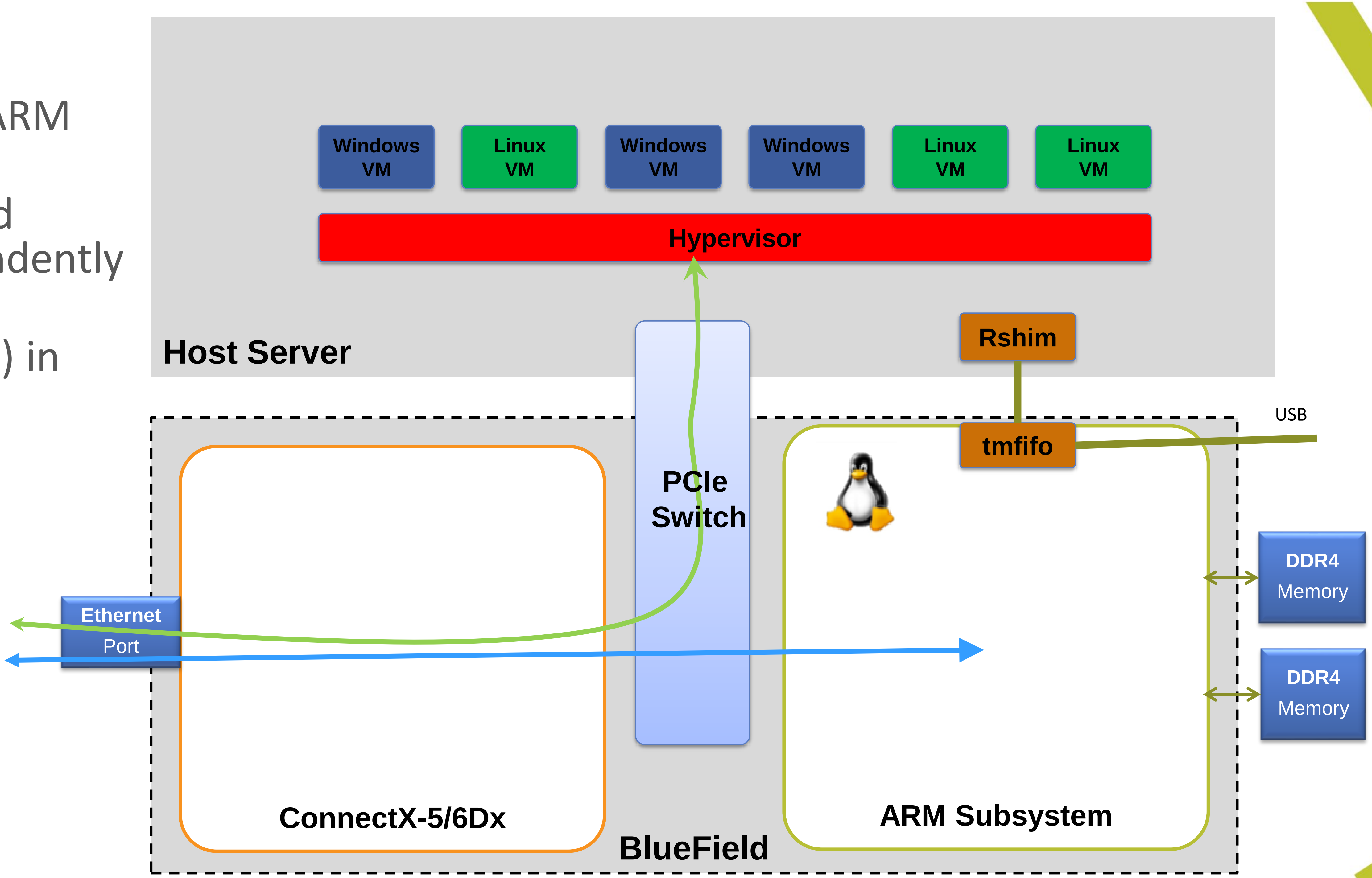
- Online Software Release Notes
- Online Software User Manual

© 2019 Mellanox Technologies

Connection Modes

Separated Hosts Mode (default configuration)

- Both the x86 and the ARM hosts are symmetric
- Each host can send and receive traffic, independently of the other host
- No OVS offload (ASAP^2) in this mode

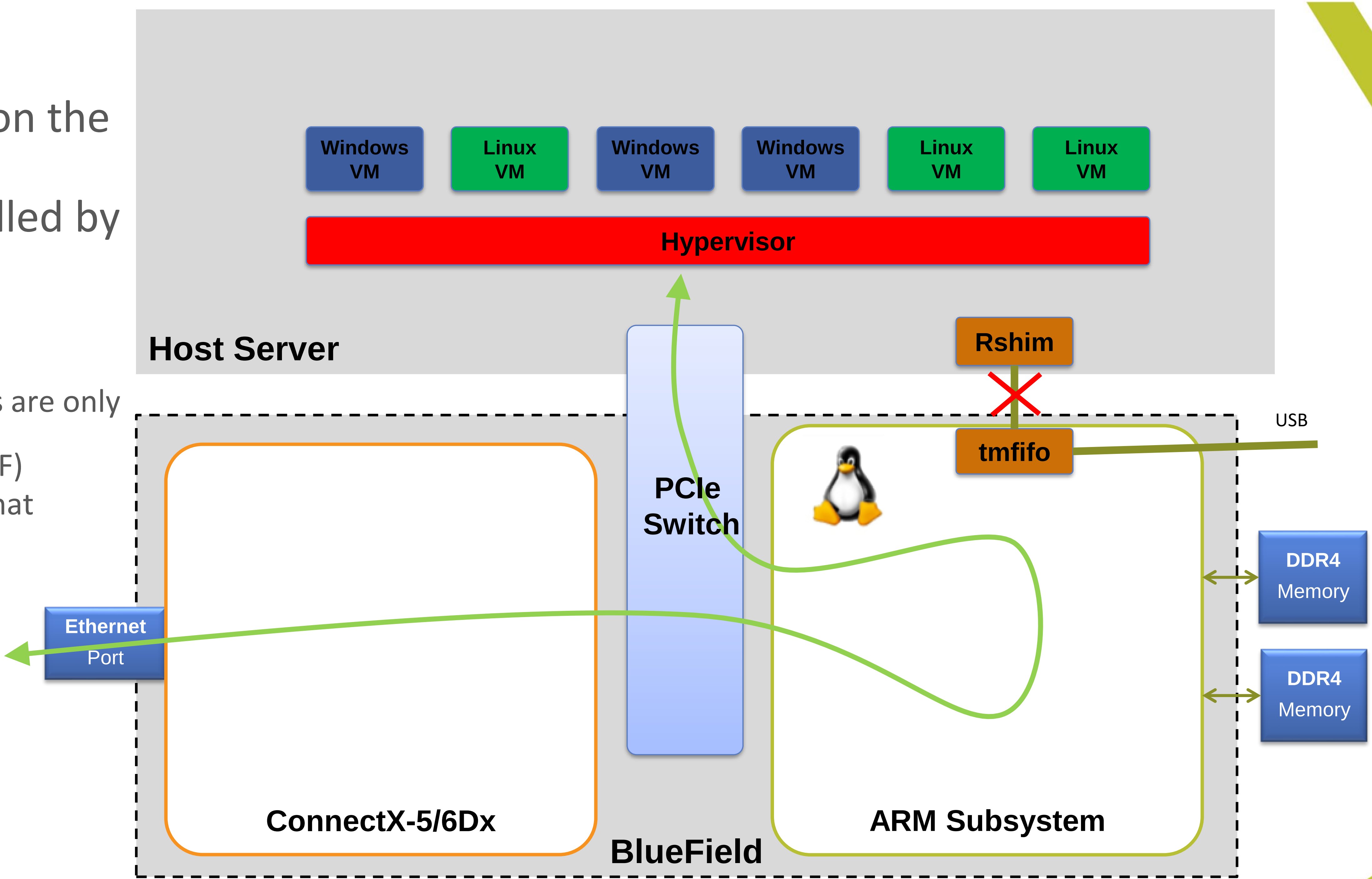


© 2019 Mellanox Technologies

SmartNIC mode (ARM Switch Ownership)

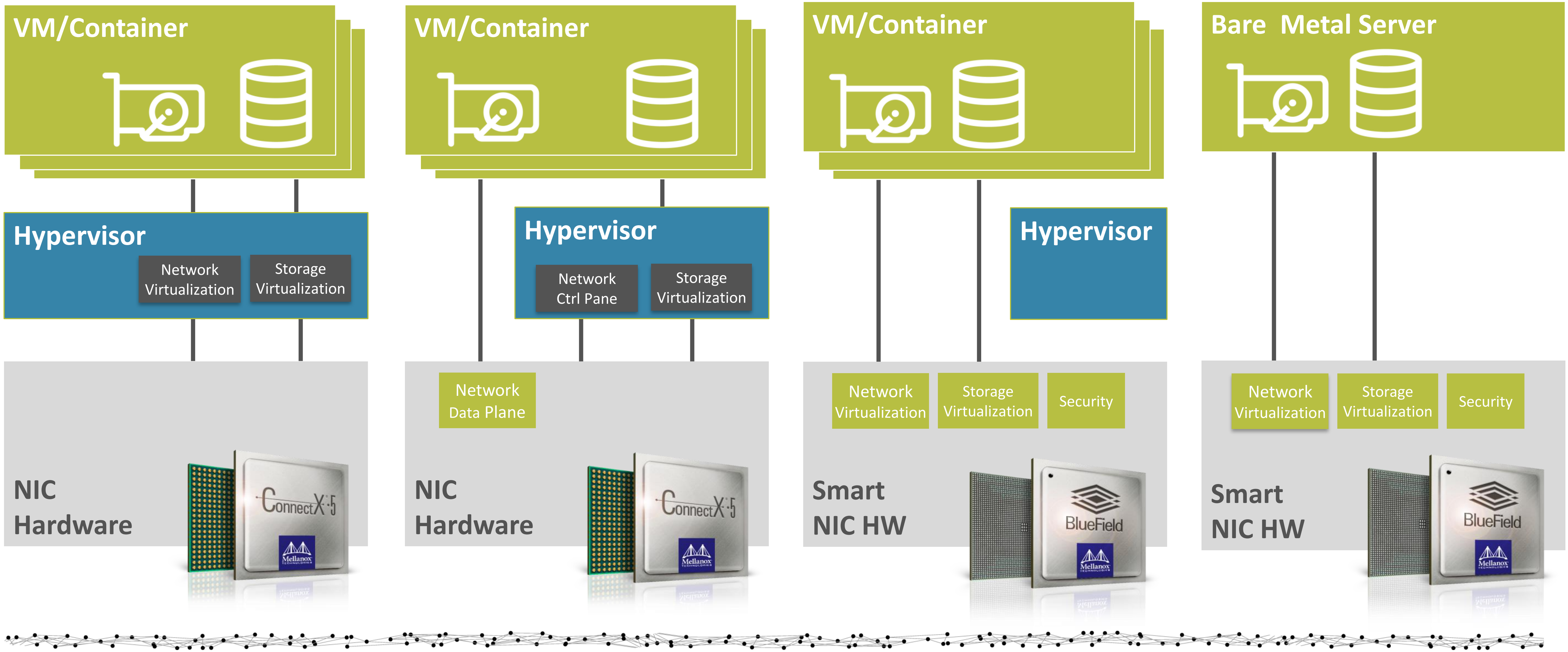
- OVS (with ASAP^2) runs on the ARM cores
- All host traffic is controlled by the PCIe switch
- Secure mode option
  - RSHIM interface is blocked
  - Port Configuration commands are only allowed from the ARM side (the x86 host is treated as a VF)
  - All device related resources that require host memory are allocated on ARM memory
  - Host PXE boot goes through

  ARM cores as well



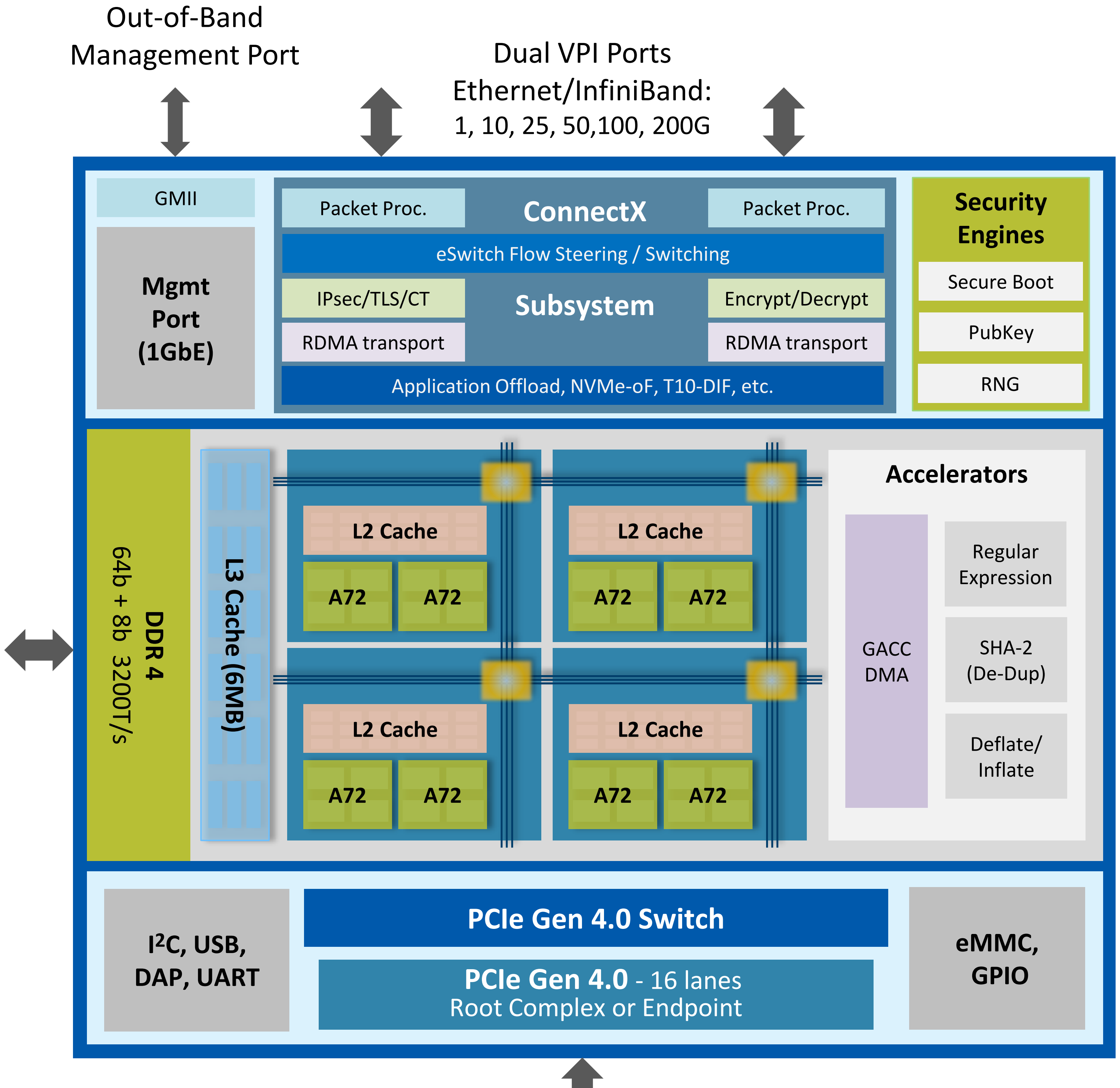
© 2019 Mellanox Technologies

Software Defined Network, Storage, Security Transition



© 2019 Mellanox Technologies

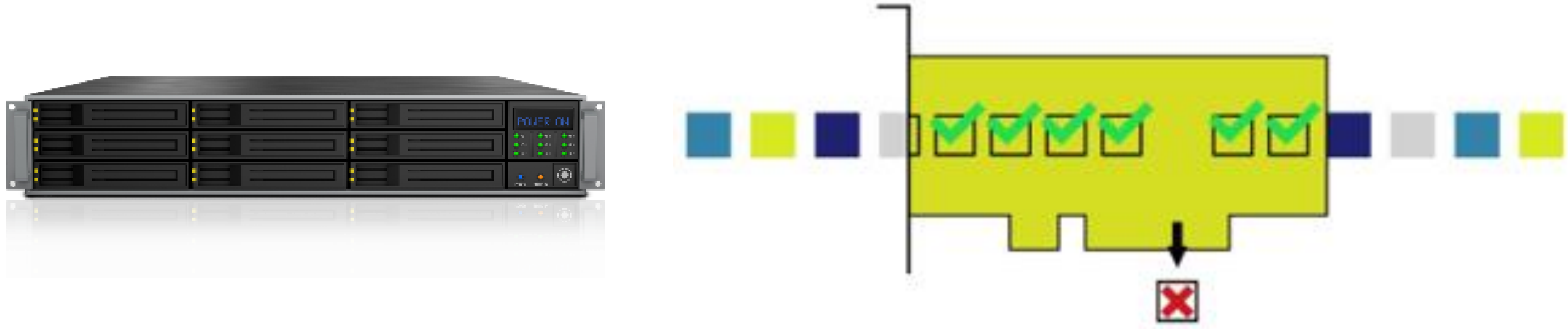
BlueField-2 Block Diagram



© 2019 Mellanox Technologies

Use case 1 - Next Generation Firewall

- Better protect your host by running next-generation Firewall on the ARM cores
- Firewall rules are programmed using OVS or Kernel TC
- Connection Tracking inspect and restrict connections to services
- Mellanox Accelerated Switching and Packet Processing (ASAP^2) seamless offload
- Firewall policy is enforced in wire speed



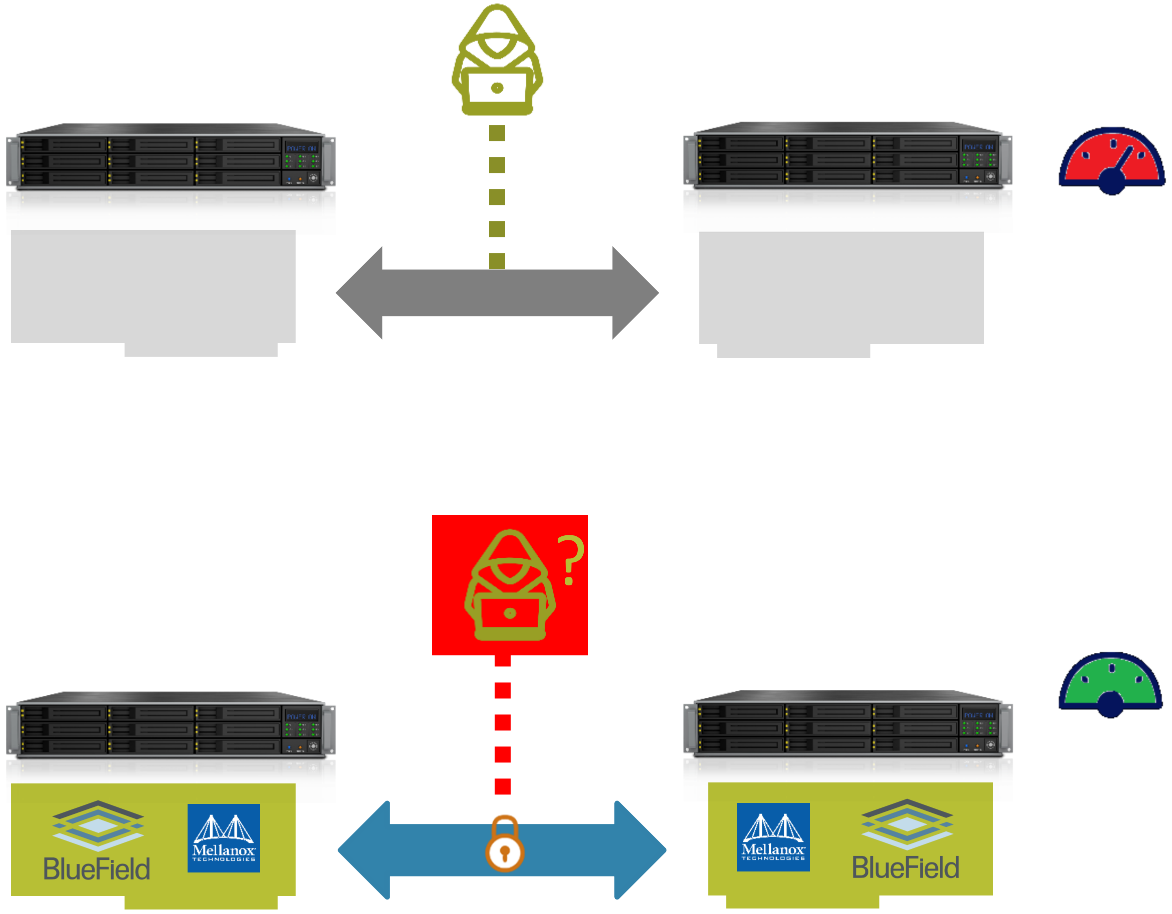
ASAP2

Accelerated Switching & Packet Processing

Use case 2a – Transparent IPSEC Encryption inside DC

Inline encryption acceleration

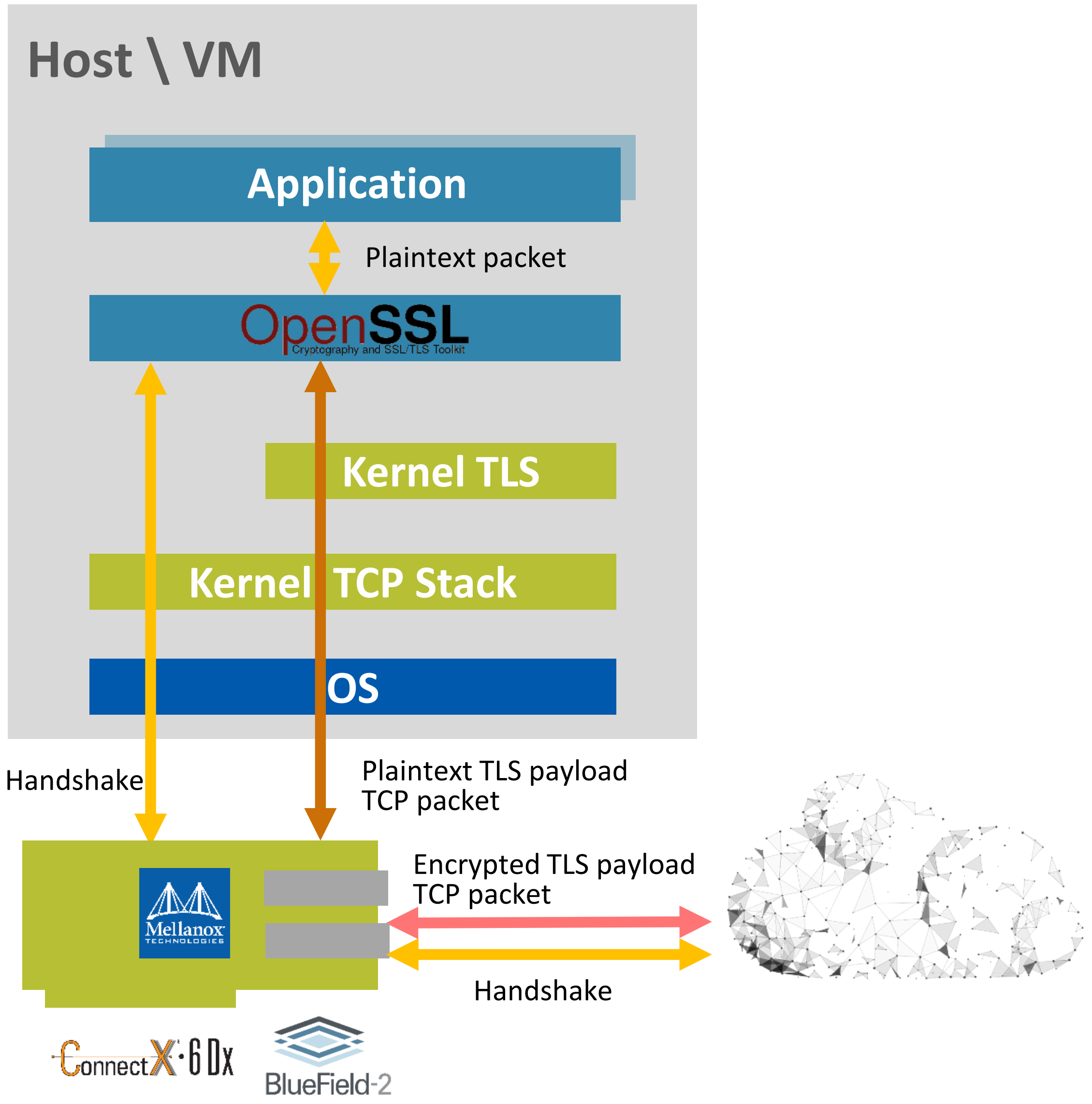
- Protection of Data-in-Motion and Data-at-Rest
- 100Gbps Encryption\decryption is done as another data-path action
- **“Zero Utilization”**- Host CPU is fully offloaded from encryption functions
- BlueField SmartNIC enable fully Isolated control plane and key management (Transparent Mode)



© 2019 Mellanox Technologies

Use case 2b – TLS Inline Offload

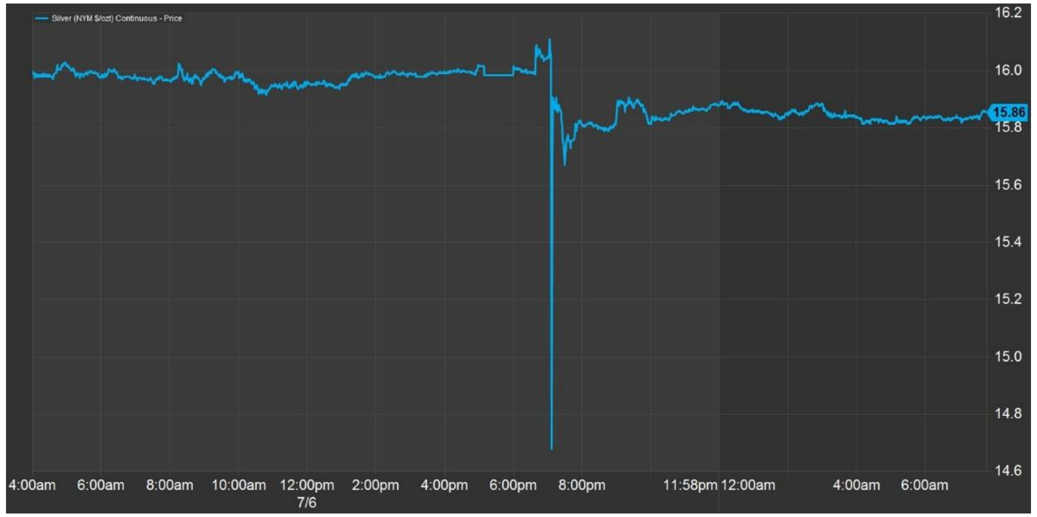
- Inline Encryption/decryption and packet authentication in NIC hardware
  - Virtualized environment and bare metal support
  - Out-of-order packets are handled by software
  - The hardware decryption is done in-order and only payload is processed
- Inline offload is enabled through Kernel TLS (kTLS)
  - Software stack unchanged (software fallback)
  - Supported in OpenSSL
  - Key management in software
- BlueField enables hardware acceleration for OpenSSL Public Key Infrastructure (PKI)

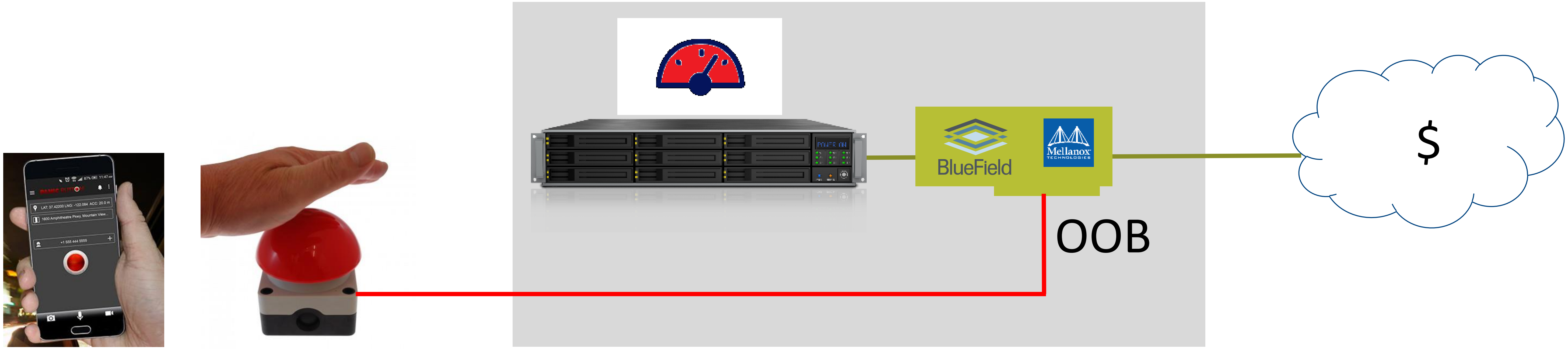


© 2019 Mellanox Technologies

Use Case 3 – Emergency Hardware Trading Stop

- ALGO trading overheats
- In house or Third Party Control to stop trading
- Allow only cancel orders through to exchange
- Rules pushed into NIC Hardware, no extra latency
- Could be via remote third party via encrypted link terminated on NIC
- Could be decision by committee

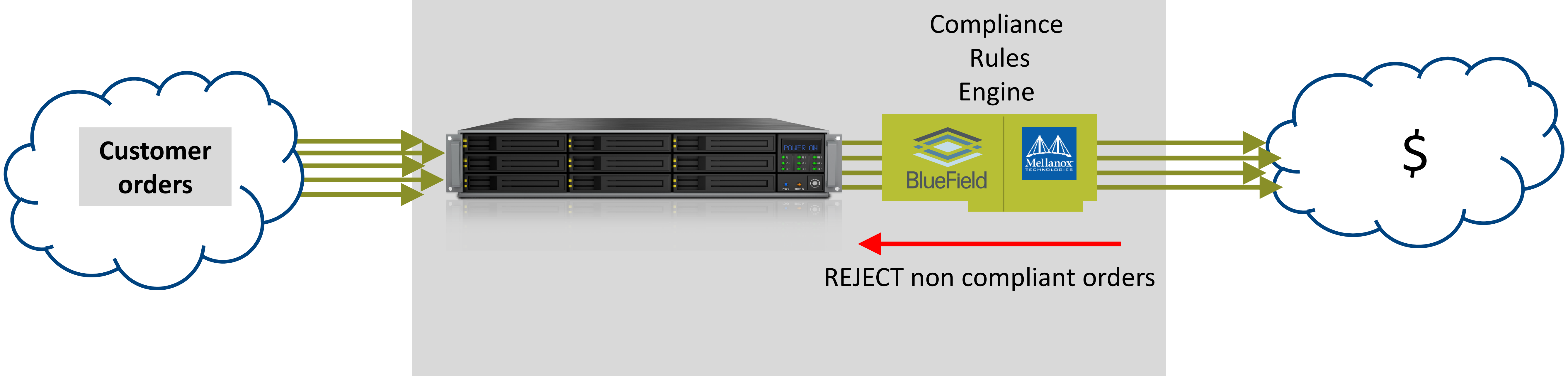




© 2019 Mellanox Technologies

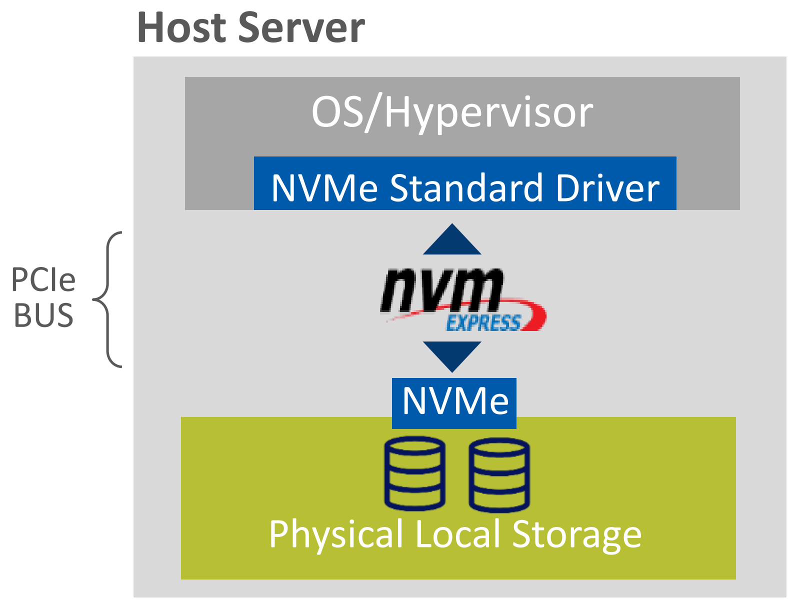
Use case 4 – Low Latency Compliance Checks

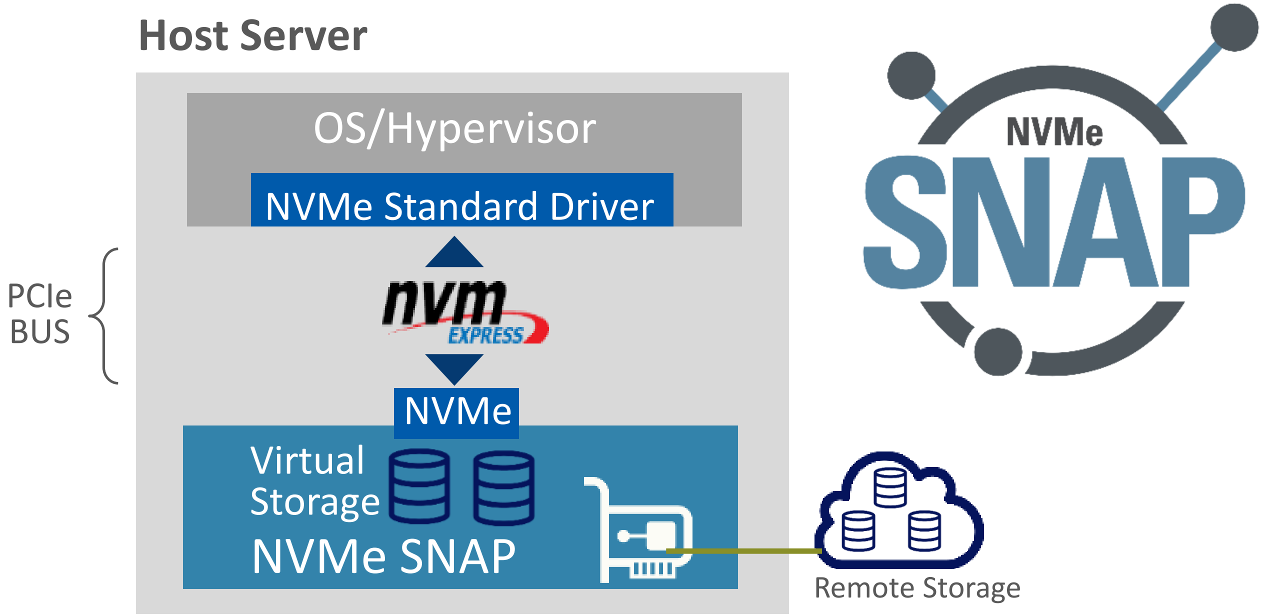
- Bank aggregates customer orders
- Compliance rules pushed into NIC Hardware, no extra latency
- Fast check for compliance before sending to exchange
- Much shorter dev cycle than FPGA/ASIC solution



© 2019 Mellanox Technologies

Use Case 5: NVMe SNAP - Hardware Emulated Storage





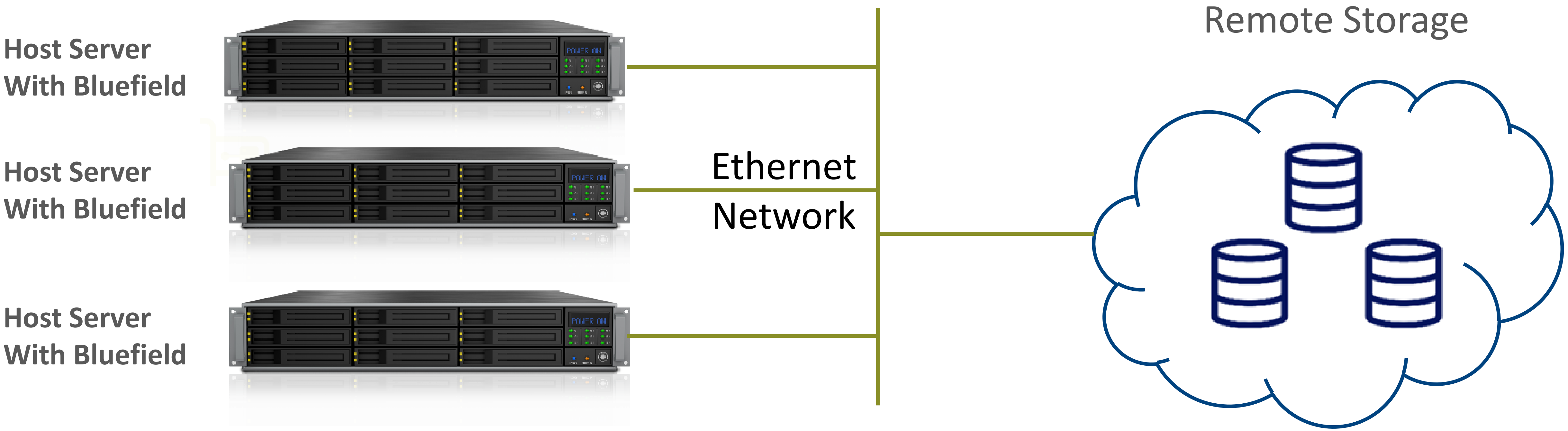
Physical Local NVMe Storage

- ✓ Serving bare-metal and hypervisor/VMs
- ✗ Bound by physical SSDs capacity
- ✗ Under-utilized storage
- ✗ Scalability
- ✗ Over-provisioning bound to compute node

NVMe SNAP Drive Emulation

- ✓ Serving bare-metal and hypervisor/VMs
- ✓ Over-provisioning, scaled to rack/cluster
- ✓ Saving OPEX and CAPEX
- ✓ OS-agnostic using inbox standard NVMe driver
- ✓ Supports all network transport types – NVMe-oF, iSCSI, iSER and even proprietary

Use Case 5: NVMe SNAP - Hardware Emulated Storage



NVMe SNAP Drive Emulation

- ✓ Serving bare-metal and hypervisor/VMs
- ✓ Over-provisioning, scaled to rack/cluster
- ✓ Saving OPEX and CAPEX
- ✓ OS-agnostic using inbox standard NVMe driver
- ✓ Supports all network transport types – NVMe-oF, iSCSI, iSER and even proprietary
- ✓ Near Local Performance

© 2019 Mellanox Technologies

How Mellanox BlueField SmartNIC Transforms Bare-Metal Cloud

© 2019 Mellanox Technologies

Cloud and NFV

Performance

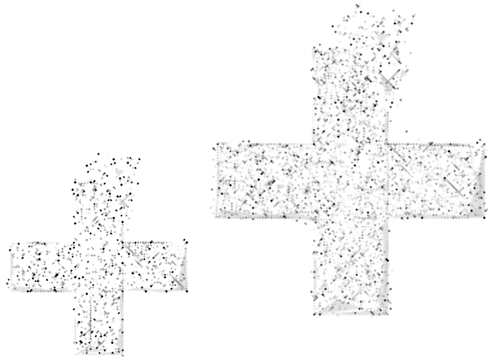
- 100Gb/s line rate @ 64B with DPDK
- Up to 10X message rate with ASAP2
- Zero CPU utilization with ASAP2

Variety of Solutions

- SR-IOV or VirtIO acceleration
- Support custom vSwitch
- Control plane in kernel or user-space
- Best in class DPDK for no-offload users
- SmartNIC for isolation and control plane offload

Feature Rich

- Any overlay tunnel: VXLAN, GRE, MPLS and more
- Any Virtual Switch
- Header re-write
- Hair-Pin



Community

- Standard software
- Upstream
- Inbox
- Key partnerships

© 2019 Mellanox Technologies

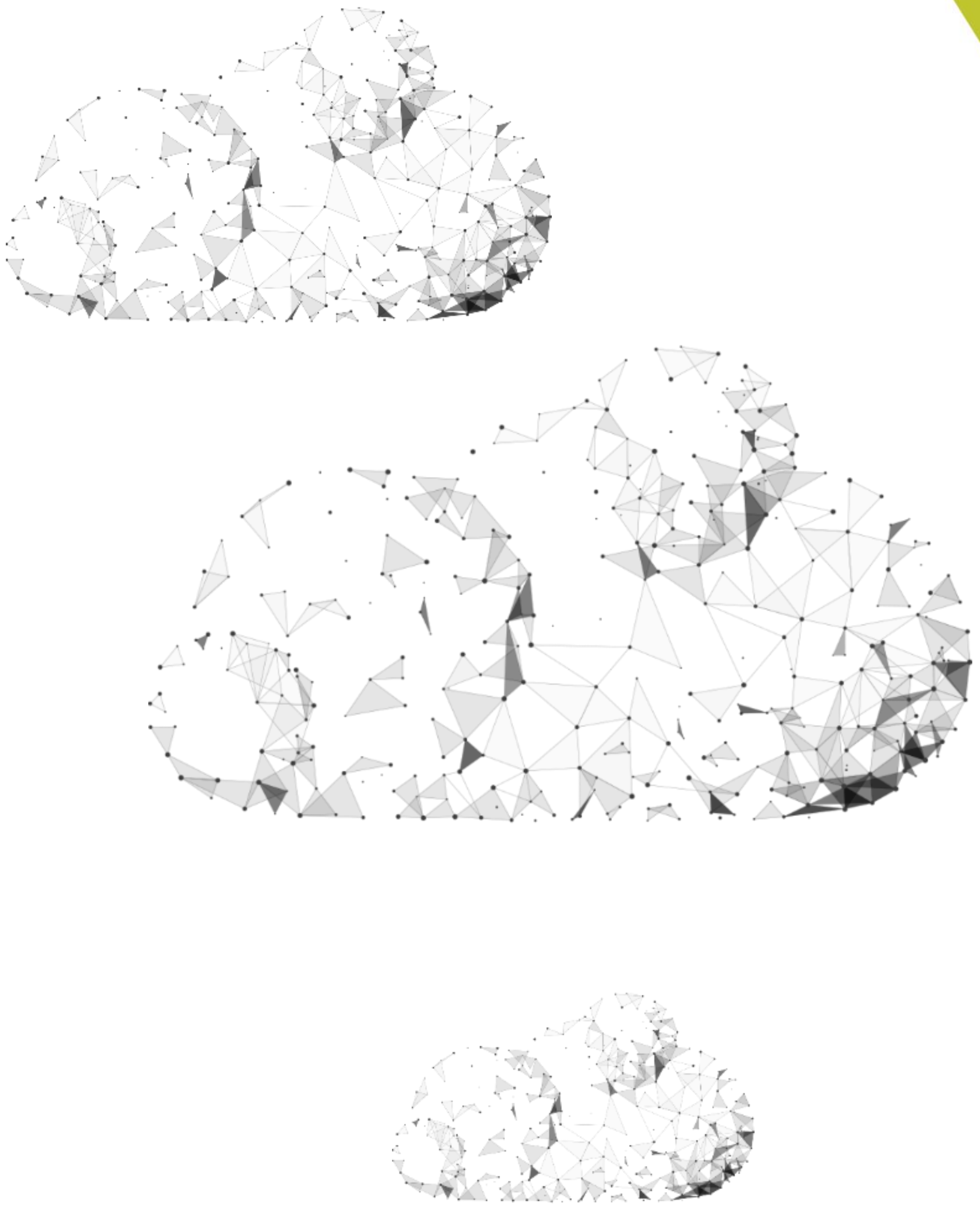
BlueField Enables SDN in Bare-Metal Clouds

- Bare-metal clouds are lacking typical SDN capabilities, leveraging ToR switch vendor solutions

- BlueField enables a full featured SDN integration and hardware acceleration for
  - Tenant networking
  - Security groups
  - Distributed virtual routers (DVR)
  - Trunk ports
  - etc.

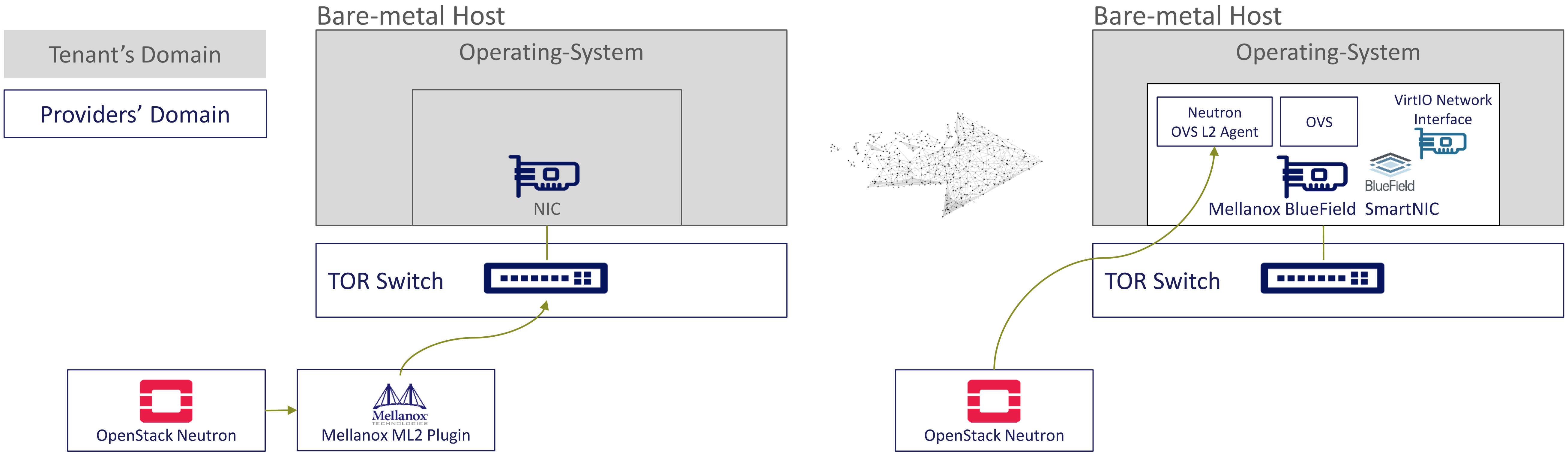
- Complete bare-metal provisioning solution powered by upstream OpenStack Ironic

- BlueField enables VirtIO network interface → No need to install network driver in bare-metal host



© 2019 Mellanox Technologies

BlueField Enables SDN in Bare-Metal Clouds



TOR Switch Networking

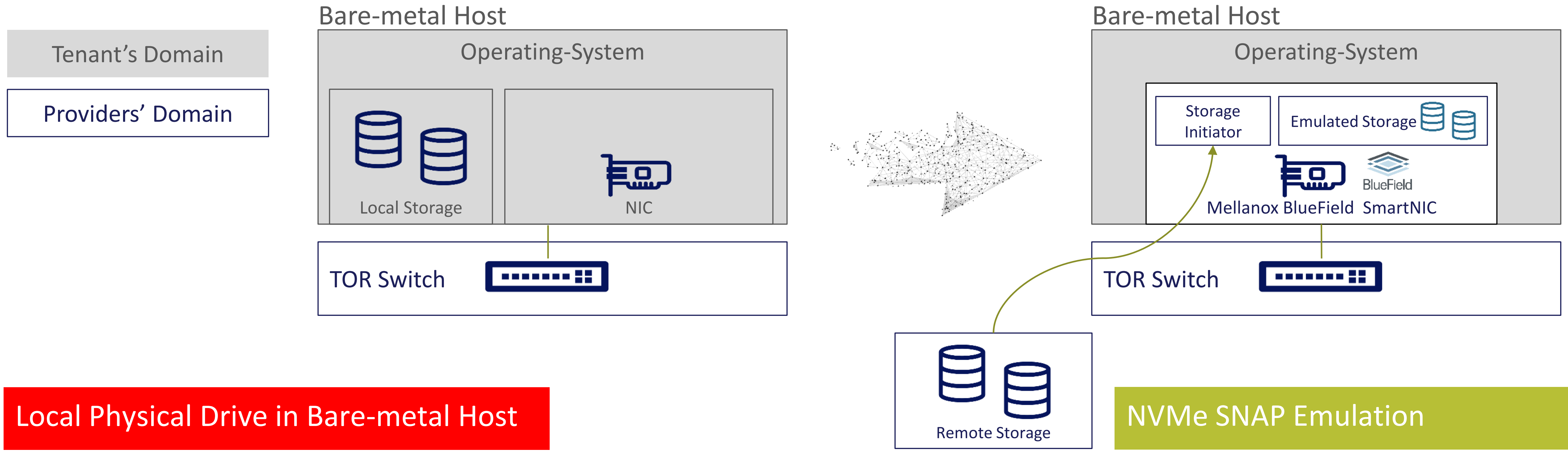
- ✗ Limited to no SDN capabilities
- ✗ Orchestration through proprietary TOR switch vendor plugins
- ✗ Mandates proprietary network driver installation in bare-metal host

SDN Integration

- ✓ Full-featured SDN capabilities
- ✓ Full orchestration through upstream OpenStack Neutron APIs
- ✓ No installation of network driver in bare-metal host
- ✓ Dynamic assignment of multiple virtual network interfaces

© 2019 Mellanox Technologies

BlueField Enables Storage Virtualization in Bare-Metal Clouds



Local Physical Drive in Bare-metal Host

- ✗ Bound by physical storage capacity
- ✗ No backup service or limited to local RAID
- ✗ No possibility to manage storage resources
- ✗ No migration of resources

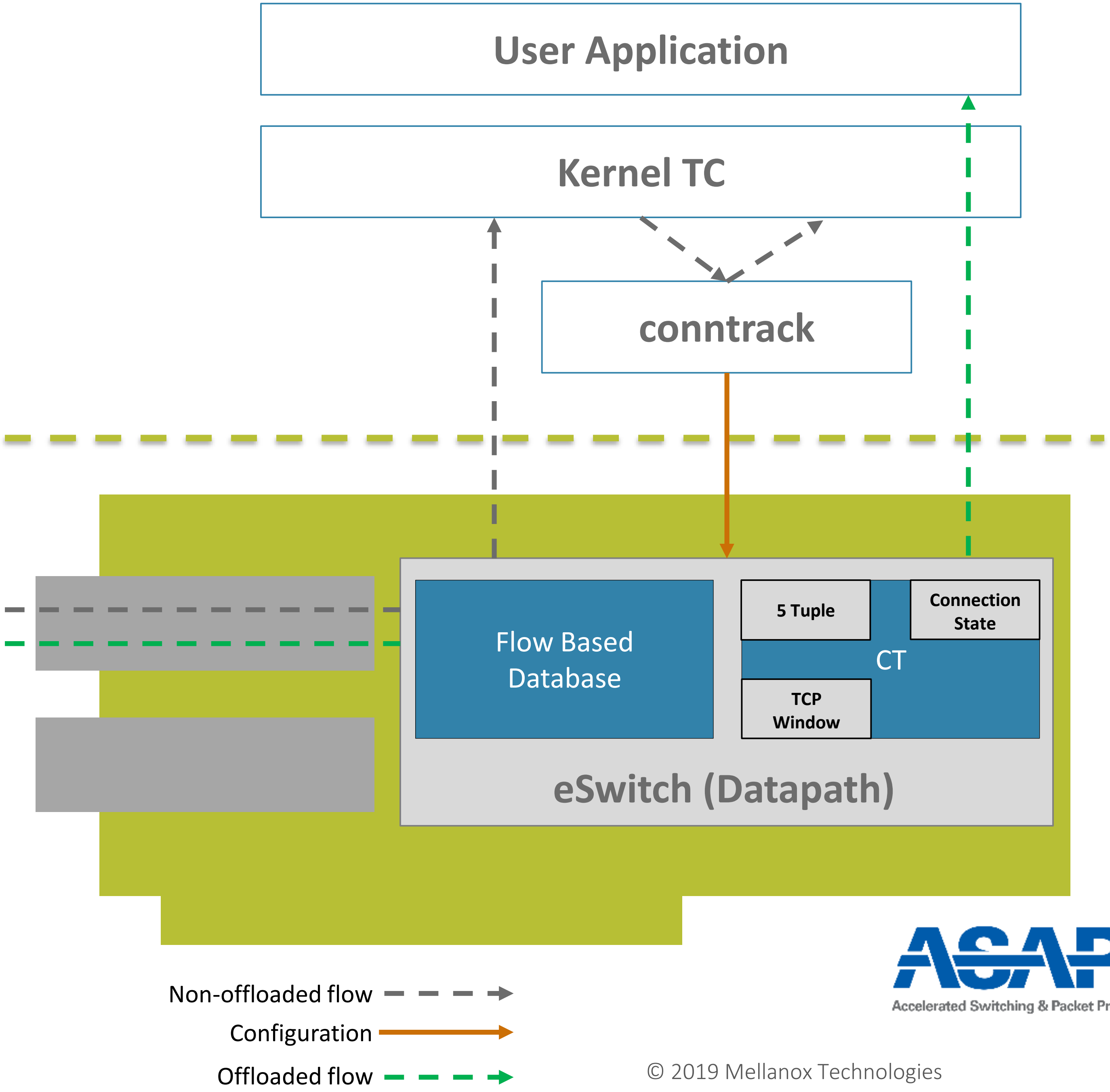
NVMe SNAP Emulation

- ✓ Same flexibility as virtualized storage
- ✓ Same performance as local storage
- ✓ OS agnostic, only NVMe driver required
- ✓ Backed-up in the storage cloud
- ✓ Dynamically allocated cloud storage
- ✓ Any wire protocol & storage management

© 2019 Mellanox Technologies

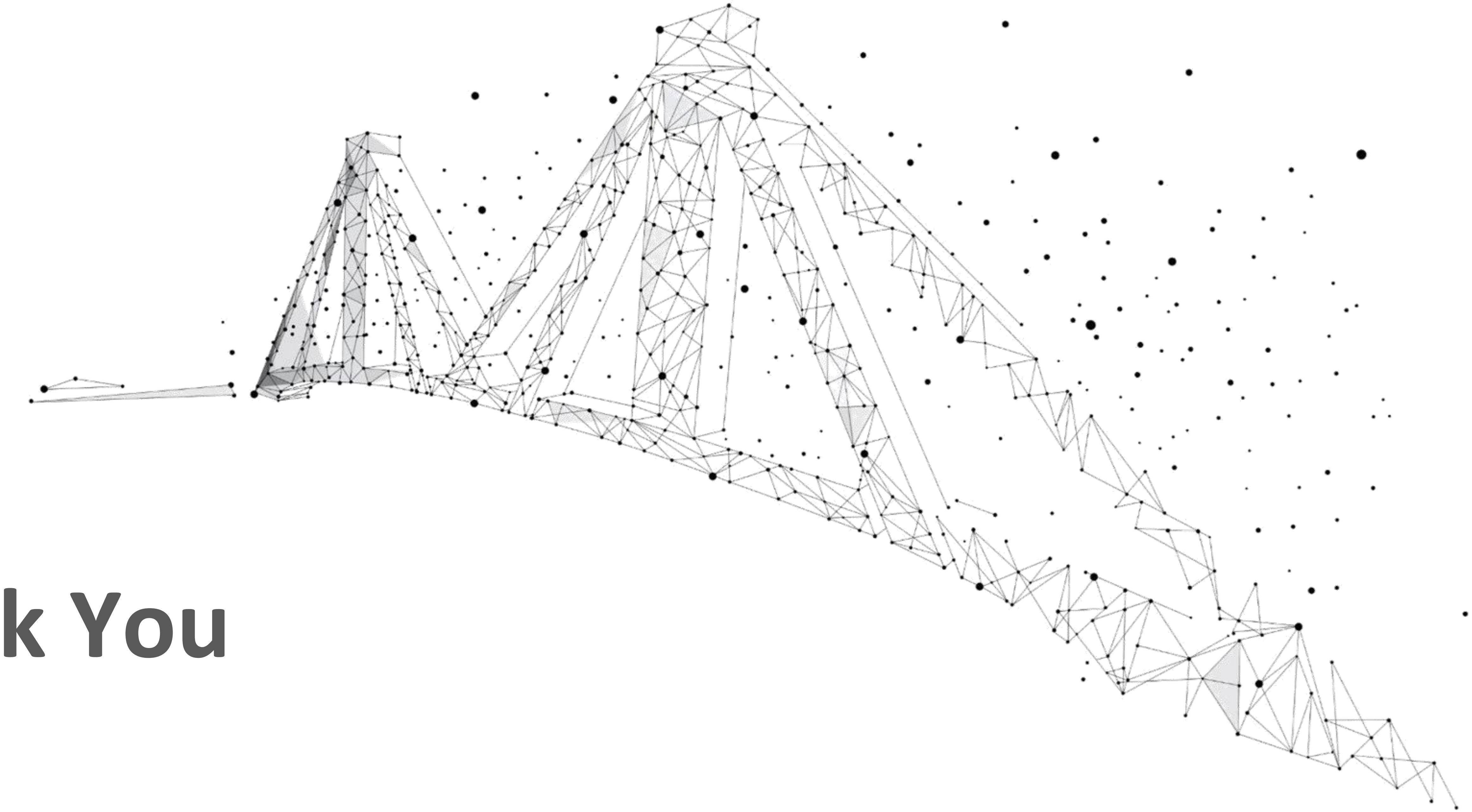
Connection Tracking Acceleration

- Hardware offload of Connection Tracking using ASAP^2
  - Allows to inspect and restrict connections to services based on their connection state
  - Check TCP connection state in hardware in real-time
  - Hardware counter per flow for activity/aging
- Offload flow
  - Packets of new connection are forwarded to the CPU (host or arm)
  - Software can then
    - Complete the TCP Hand-Shake
    - Drop the packet
  - Driver adds the connection (5-tuple) to the hardware using Kernel TC mechanism
  - Following packets of the TCP streams are forwarded directly to the VF
- Protect the host from attacks in wire speed



ASAP2
Accelerated Switching & Packet Processing

© 2019 Mellanox Technologies



Thank You

© 2019 Mellanox Technologies