SOLUTION BRIEF

In-Hardware Storage Virtualization - NVMe SNAP™ Revolutionizes Data Center Storage

Composable Storage Made Simple

Mellanox NVMe SNAP™ (Software-defined Network Accelerated Processing) technology enables hardware-accelerated virtualization of NVMe storage. NVMe SNAP makes networked storage look like a local NVMe SSD, emulating an NVMe drive on the PCIe bus. The host OS/Hypervisor makes use of its standard NVMe-driver unaware that the communication is terminated, not by a physical drive, but by the NVMe SNAP. Any logic may be applied to the data via the NVMe SNAP framework and transmitted over the network, on either Ethernet or InfiniBand protocol, to a storage target.

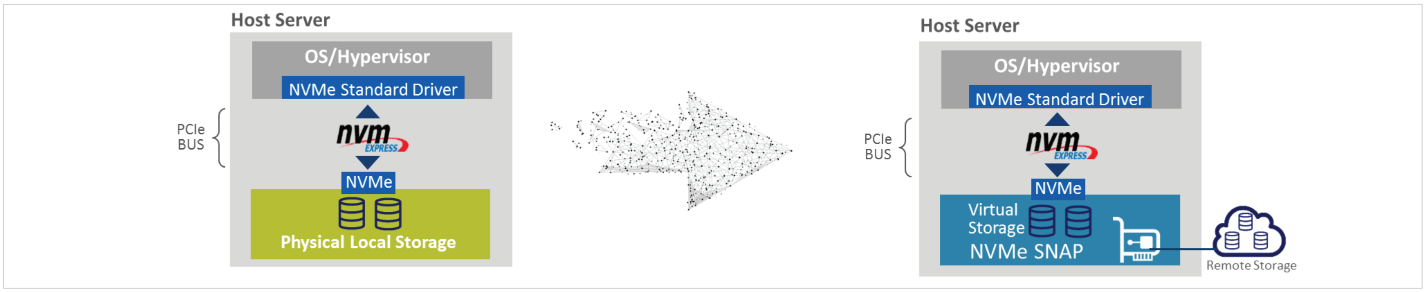

Figure 1: *Traditional host server versus a host server with virtualized storage enabled by NVMe SNAP*

NVMe SNAP is based on Mellanox BlueField® technology and combines unique hardware-accelerated storage virtualization with the advanced networking and programmability capabilities of the BlueField SmartNIC. NVMe SNAP together with the BlueField SmartNIC enable a world of applications addressing storage and networking efficiency and performance.

NVMe SNAP FRAMEWORK

NVMe SNAP empowers customers to implement their own storage solutions on top of the supplied framework. NVMe SNAP exposes an interface to the BlueField SmartNIC Arm® cores for implementing control of the storage solution. NVMe SNAP is integrated with the popular Storage Performance Development Kit (SPDK) open source project, and provides customers with the agility to program in a familiar environment, reducing integration challenges and time to market. Customers can choose between two available data paths – the first, full-offload, makes use of a hardware-offload available for NVMe SNAP which takes the data traffic from the NVMe PCIe converts it to NVMe-oF (RoCE) and transmits directly to the network, all in hardware. This option provides the best performance but lacks the ability for the software running on the Arm cores to 'touch' the data or change the storage protocol. The second option enables the SPDK running on the Arm cores to terminate the traffic coming from the NVMe PCIe, and implement any customer logic on it, and then transmit the data to the network. This latter option makes use of Arm cores in data path, and allows full flexibility to implement any type of storage solution inline. In both of these data path options, the control plane is always running in the Arm cores, orchestrating the traffic to its destination.

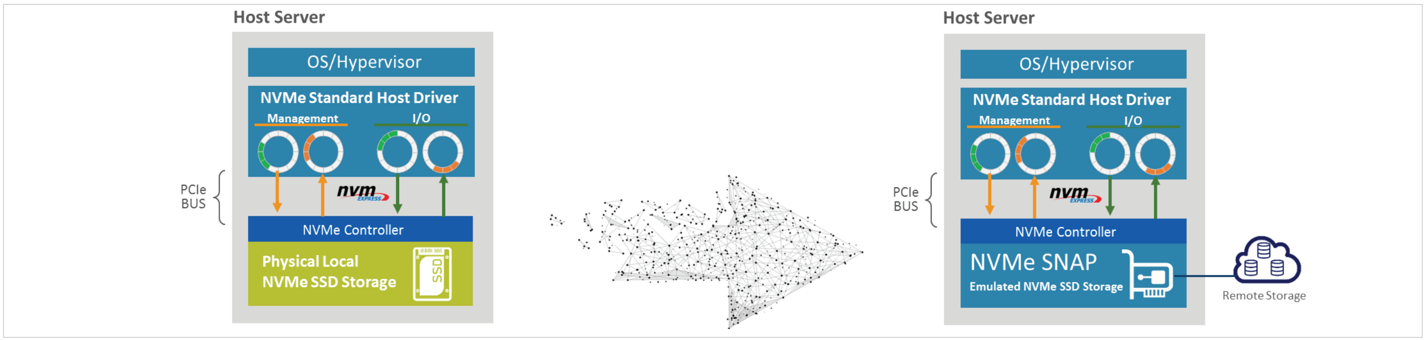

Figure 2: *NVMe SNAP empowers customers to implement their own storage solutions on top of the supplied framework*

NVMe SNAP USE CASES

Bare-Metal Cloud

A cloud service provider's bare metal offering will usually include local storage in each server for use by the cloud's customers. This provides easily accessible and fast performing storage under the customer's full control. But this comes at a price for the cloud service provider, limiting their ability to efficiently provision remote storage that is easier to migrate and protect. Therein lies a conflict when designing a bare metal cloud offering between what is best for the customer (local storage) and what is best and most easily composable for the cloud service provider (networked storage).

Cloud service providers can now virtualize the bare metal storage with no impact on their customers using NVMe SNAP, in effect creating a win-win situation for both. Bare-metal cloud customers continue to use their standard operating system's NVMe PCIe driver, with little to no performance degradation, while the service provider is gaining a richer offering with greater efficiency – storage is now virtualized, thin-provisioned, backed up, and can be migrated between servers, providing savings in both CAPEX and OPEX.

Disaggregate and Rack Scale

NVMe SNAP enables the disaggregation of compute and storage, optimizing data-center resource utilization (compute, storage, network, etc.).

Figure 3: *The disaggregation of compute and storage nodes increases simplicity & reduces costs*

Data center architecture has a strong influence on overall resource utilization. The traditional fixed architecture which glues together fixed ratios of compute, storage, and networking resources presents a workload resource allocation challenge that is typically addressed by virtualization's over-provisioning. But in reality, not all workloads are the same: some are compute intensive, while others are storage hungry. Overall this requires that data centers over-provision local storage, leading to under-utilized resources, resulting in an inefficient 40%-50% average storage utilization rate. This in turn translates to a higher than necessary capital and operational spending on the half (or more) of purchased storage capacity that is essentially unused.

NVMe SNAP HIGHLIGHTS

- Implements in-hardware storage virtualization
- Programmable NVMe SNAP framework enables integration into any storage solution based on any network protocol
- Enables optimization of storage resources for CAPEX and OPEX savings
- Free up compute node CPU resources by offloading both networking and storage data paths
- Two-in-one solution combines NVMe SNAP with BlueField SmartNIC:
  - Dual-port 25/100 Gb/s network adapter card
  - BlueField SoC: 16 Arm A72 cores
  - Best-in-class hardware offload accelerators
  - 16GB DDR4 memory
  - PCIe 3.0/4.0 interface
  - FHHL form factor
- Use cases:
  - Bare-metal cloud storage virtualization
  - Rack scaling designs with zero software impact
  - Introduce NVMe-oF to the Enterprise with no OS type/ version dependency

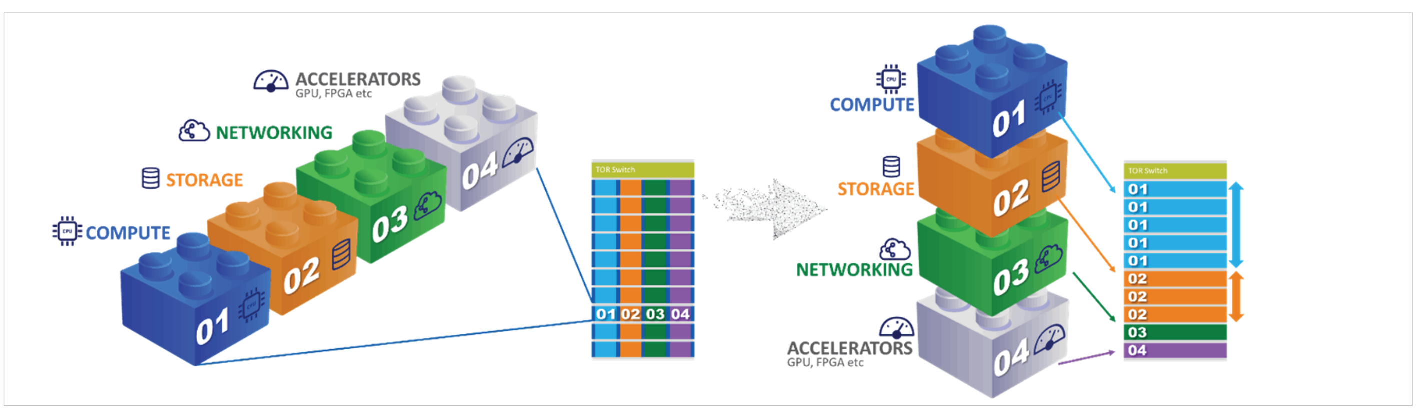

Figure 4: *NVMe SNAP enables the disaggregation of compute and storage, optimizing data-center resource utilization*

Hyper-converged storage partially addresses storage utilization, but in an ever-growing data center, adding fixed nodes with compute and storage glued together does not allow for the elastic scaling of storage independently from compute, or vice versa.

Networked storage technology has been available for quite some time, but often at the price of high latency or low throughput, or both. With new technologies such as NVMe-over-Fabrics (NVMe-oF) using RoCE (RDMA over Converged Ethernet) this is no longer an issue.

Networked storage together with NVMe SNAP hardware-accelerated storage virtualization technology enables data centers to remove their physical disks entirely from their compute nodes while connecting the compute nodes to a storage cluster transparently and seamlessly. This disaggregation of compute and storage enables storage to be part of a composable infrastructure and offers huge savings in acquisition cost, maintenance, and operating overhead, while leveraging the simplicity of this design with no software alterations or performance impact on the infrastructure at all. Compute nodes can now be added independently of storage and vice versa, allocating exactly the right amount of storage for each compute node and optimizing for any workload. The data center can now manage logical pools of resources that maximize utilization of the entire rack or cluster. (Figure 4.)

NVMe-oF for the Enterprise

The adoption of NVMe and NVMe-oF technology is on the rise, but there are still barriers to entry in software support from major OS suppliers such as Microsoft® Windows/HyperV and VMware® ESXi, which in effect are causing a "chicken and egg" problem blocking the adoption of NVMe-oF in the enterprise. Storage solution vendors adopted the NVMe-oF protocol to offer higher throughput with lower latency, but struggle to deploy into diversified enterprises with multiple OS types and versions with little or no NVMe-oF driver support. With NVMe SNAP on the compute (initiator) side, the lack of NVMe-oF software driver support is bridged, seamlessly to the running OS, implementing NVMe-oF protocol in hardware with absolutely no need for support from the running OS.

ONE AND ONE MAKE THREE

NVMe SNAP on Mellanox BlueField enables in-hardware storage virtualization while leveraging the smart adapter's Arm programmability offering a highly flexible solution. Customers can also make use, in parallel, of the BlueField infrastructure to implement network virtualization offloads, such as running the vSwitch control on Arm cores while offloading the data path to the ConnectX technology in the SoC, maximizing virtualization scalability and efficiency. (Figure 5.)

Data centers can benefit from better server utilization, allowing more virtual machines and more tenants on the same hardware, while reducing the TCO, power consumption and cabling complexity.

Customers can now develop tailor-made virtualization solutions on top of BlueField SmartNICs with the NVMe SNAP framework utilizing BlueField SoC Arm cores to leverage its built-in hardware acceleration engines.

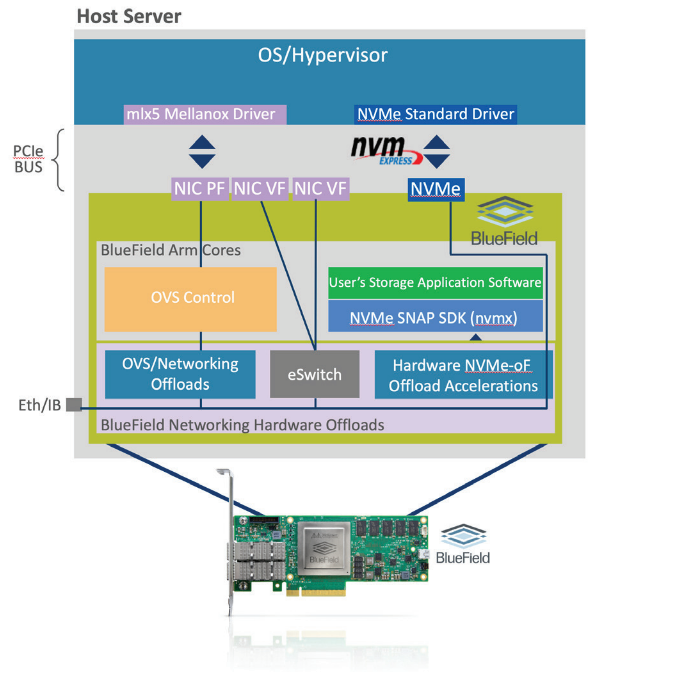

Figure 5: *BlueField SmartNIC with NVMe SNAP serves as a smart network adapter for both storage and network virtualization simultaneously*

FIND OUT MORE

Take advantage of Mellanox BlueField 25/100 Gb/s SmartNICs and NVMe SNAP technology with in-hardware storage virtualization to improve your storage and networking infrastructure.
More information can be found on www.mellanox.com:

- Mellanox NVMe SNAP page
  https://www.mellanox.com/page/products_dyn?product_family=300
- Mellanox Storage Solutions page
  https://www.mellanox.com/solutions/storage/
- BlueField Ethernet SmartNIC Product Brief
  http://www.mellanox.com/related-docs/prod_adapter_cards/PB_BlueField_Smart_NIC_ETH.pdf
- BlueField SoC Product Brief
  http://www.mellanox.com/related-docs/npu-multicore-processors/PB_Bluefield_SoC.pdf
- BlueField software suite documentation
  https://docs.mellanox.com/category/bluefieldsw

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