KumoScale Shared Accelerated Storage Software
Maximize the performance and capacity utilization of NVMe-oF™ flash storage

Organizations are adopting cloud and container architectures to help maximize the flexibility, scalability and utilization of compute and storage resources within their data centers. However, for high-performance solid state storage, the traditional shared-nothing cloud model of direct attached storage (DAS) has not only resulted in lower compute/storage capacity utilization and reduced overall performance, but also contributed to SKU proliferation.

NVMe™ over Fabrics (NVMe-oF) is designed to address the inefficiencies associated with the DAS architecture by disaggregating high performance NVMe solid state storage resources from the compute nodes and making it available across the network infrastructure.

Toshiba's KumoScale shared accelerated storage software goes one step further and abstracts the storage resources into a virtualized pool, which can then be efficiently allocated in "right-sized" capacities to compute nodes. The software was developed with a strong focus on virtualizing, managing and securing data. Its feature-rich, standards-based design provides storage management functions for abstraction, provisioning for containers and virtual machines, and integration with popular orchestration frameworks through a secure API.

TOSHIBA KUMOSCALE SOFTWARE

The prevalence of NVMe solid storage in the enterprise is steadily increasing. Servers configured with NVMe SSDs can now deliver more than enough performance required by the most demanding application workloads. The existing direct attached storage (DAS) implementation of NVMe storage presents architects with scalability and utilization limitations, which have negative CapEx and OpEx implications.
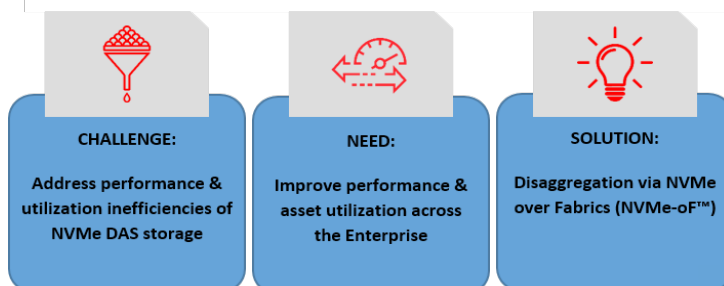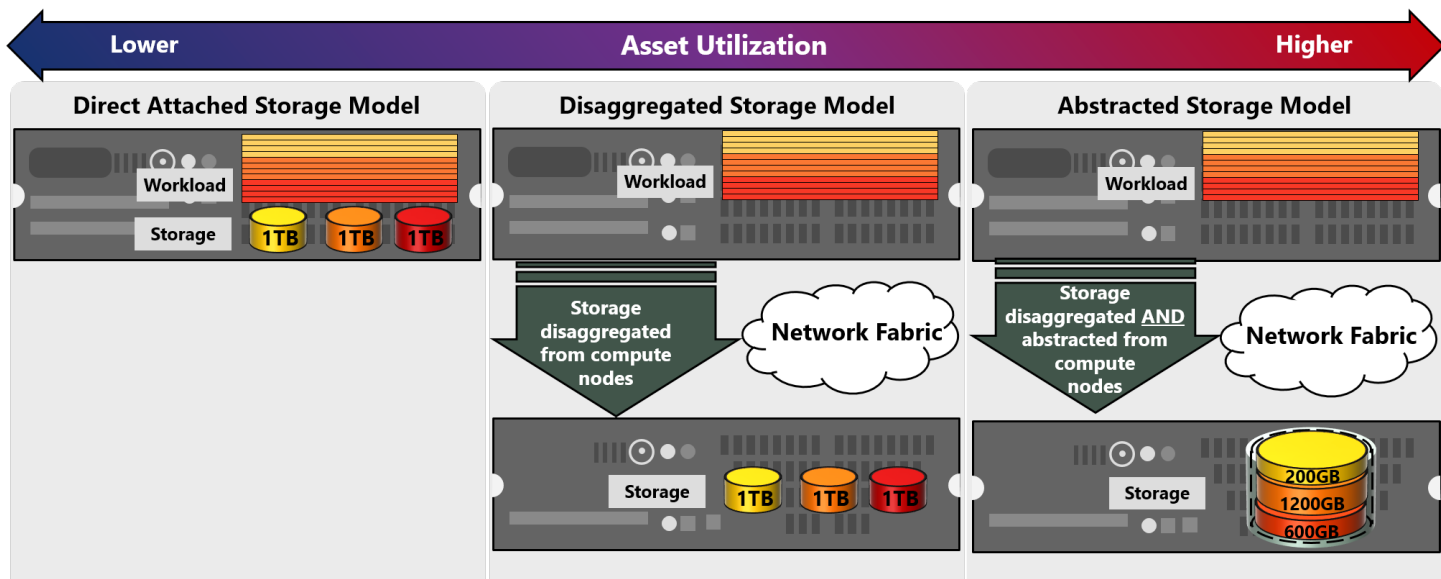

CHALLENGE:

Address performance & utilization inefficiencies of NVMe DAS storage

NEED:

Improve performance & asset utilization across the Enterprise

SOLUTION:

Disaggregation via NVMe over Fabrics (NVMe-oF™)

Figure 1. Need for Greater Performance and Asset Utilization

KEY BENEFITS

- Shared storage to increase utilization of storage assets

- Live migration capability to optimize application availability

- GUI and CLI interface options for simplified management

- Ease of integration with management APIs for popular orchestration solutions

- Automated workload allocation to deliver balanced wear leveling, across SSDs

"NVMe over Fabrics creates highly scalable and flexible storage for new cloud architectures. Toshiba KumoScale shared accelerated storage software and Mellanox ConnectX network adapters enable customers to maximize flash utilization and deliver elastic flash capacity on demand."

-- Jeremy Werner
VP SSD Marketing and Product Planning
Toshiba Memory America, Inc.

Lower	Asset Utilization	Higher

| Direct Attached Storage Model | Disaggregated Storage Model | Abstracted Storage Model |

Figure 2. Storage Models

- **Direct Attach Model**: Storage and compute node reside within a physical chassis and each compute node is given a physical drive. Unused capacity cannot be reallocated to another node within the chassis, resulting in stranded storage and low asset utilization.

- **Disaggregated without Abstraction**: Available storage is physically disaggregated from compute nodes, but the solution provides no ability to subdivide individual drives, so overall utilization is similar to the direct attached model.

- **Abstracted Storage Model**: Available storage is disaggregated and abstracted into a virtualized pool of storage which is flexibly allocated. Each compute node/application is given "right-sized" allocation of storage. Storage asset utilization and server efficiency are optimized.

NVMe-oF addresses these limitations by enabling a disaggregated pool of NVMe storage, which can be accessed by compute nodes over the existing network infrastructure.

Toshiba recognizes that storage disaggregation is only the first step towards a more efficient infrastructure, and adding storage abstraction enables full utilization of storage capacity. Toshiba's KumoScale abstracts NVMe storage resources from the physical SSDs, and enables the creation and provisioning of NVMe storage from a centralized storage pool into storage namespaces/containers, enabling application-optimized allocation of storage resources. This solution addresses the asset utilization challenges associated with the storage disaggregation model. The end result is greater capital and operational savings and full use of storage resources (Figure 2).

Ease of management was key to the design of KumoScale software. Administrative management options include a graphical user interface (GUI) and RESTful API for convenient integration into mainstream orchestration frameworks. Additionally, Kumoscale provides administrators powerful analytic and diagnostic tools to help optimize storage performance and utilization (Figure 3).
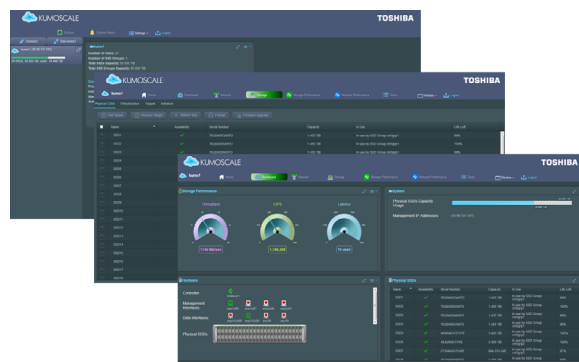
Figure 3. Toshiba Storage Node Software provides powerful diagnostic and analytic tools that simplify management and optimize storage utilization.

The table below provides an overview of some of the key attributes and capabilities of Toshiba's shared accelerated storage platform.

Table 1. Key attributes and capabilities of Toshiba's shared accelerated storage platform

Feature	Value	Benefits

Storage Abstraction	• Individual physical drives are presented as one or more logical pools which can then be apportioned to the specific needs of compute nodes/applications.	• "Right-sized" storage for each application/compute node
		• Delivers greater degree of asset utilization
		• Reduce:
		  • Capital acquisition
		  • Management costs
		  • Power/cooling costs
		  • Real-estate costs
			• **Improve asset utilization**
			• **Reduce OpEx & CapEx**

Powerful Management Capabilities	• Intuitive and easy to use GUI console
	• RESTful API for integration with the most popular data center orchestration, provisioning and telemetry solutions (e.g. OpenStack, Kubernetes, RSD/RedFish)
	• Active engagement with orchestration, provisioning and telemetry providers
		• Management flexibility:
		  • GUI based management console
		  • Rich RESTful API
		• Future compatibility
		  • Continued API development to support seamless integration with next generation orchestration, provisioning and telemetry solutions
			• **Simplify management**

For more information on Toshiba's KumoScale shared accelerated storage software, please visit www.KumoScale.com

COMBINED SOLUTION

The implementation of NVMe-oF architecture using Toshiba KumoScale software and the end-to-end Mellanox RDMA over Converged Ethernet (RoCE) enables high performance network Ethernet adapters and Ethernet switches.

The Mellanox ConnectX® family of network adapters provide fast, reliable, low-latency network connectivity to the KumoScale solution. They support Ethernet connection speeds of 10, 25, 40, 50 and 100 Gigabits per second (Gbps) for maximum throughput and low latencies to support performance-intensive transactional and analytic workloads. Efficient RDMA offloads via RoCE free up CPU cycles on both storage initiators and targets, allowing more CPU power to run applications, hypervisors, and storage software. ConnectX adapters also feature low power consumption and offloads for cloud, storage, overlay networks to maximize server efficiency.

Complementing the Mellanox ConnectX network adapters in compute/storage nodes, Mellanox Spectrum Ethernet switches offer transparent network fabric with deterministic high-performance, scalability and cost/operational efficiency. The Mellanox Spectrum switches always deliver line-rate throughout and consistent sub-microsecond latency at speeds of 10, 25, 40, 50, 100Gbps and any frame size. The dynamically shared buffer enables fair and predictable performance, with respect to policies and priorities defined at the control plane. Combined with faster telemetry and notifications, Mellanox Spectrum switches provide the best RoCE congestion management. All of these factors are crucial to deliver the maximum NVMe storage performance of KumoScale.

UNCOMPROMISING PERFORMANCE

In a combined test environment using Mellanox ConnectX-5 100G Ethernet adapters with Toshiba's KumoScale software running on a storage target node with 24 NVMe SSDs, performance was outstanding with minimal fabric latency impact to the benchmark workloads for both reads and writes. Overall throughput and IOPS performance measured was comparable to DAS-based environments, creating an abstracted pool of high-performance NVMe flash storage available to compute VMs, with low and predictable latency to support application service level agreements.

Table 2. Test Environment Performance

Block Size	Latency[1,2] (µs)		IOPS[3]		Throughput (GB/s)	
	Read	Write	50/50	100%W	50/50	100%W
4 KB	13	8	5.9M	4.6M	23	19
128 KB	29	34	218K	155K	27	19

Notes:
1 To provide results that are independent of SSD performance, latencies are given as a delta relative to the latency measured on the same drive in a direct-attached configuration
2 Queue depth = 1
3 Full duplex
* KumoScale performance was measured using a mid-range, single CPU platform equipped with 24 Toshiba SSDs, and is dependent on the hardware platform and installed SSDs

About Toshiba Memory America, Inc.

Toshiba Memory America, Inc. is the US-based subsidiary of Toshiba Memory Corporation, a leading worldwide supplier of flash memory and solid state drives (SSDs). From the invention of flash memory to today's breakthrough 96-layer BiCS FLASH™ 3D technology, Toshiba continues to lead innovation and move the industry forward. For more information on Toshiba Memory America, please visit www.toshiba.com/TMA and follow the company on Facebook and on Twitter: @Toshiba_Memory.

© 2017 Toshiba Memory America, Inc. All rights reserved.

About Mellanox

Mellanox Technologies (NASDAQ: MLNX) is a leading supplier of end-to-end Ethernet and InfiniBand intelligent interconnect solutions and services for servers, storage, and hyper-converged infrastructure. Mellanox intelligent interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications and unlocking system performance. Mellanox offers a choice of high performance solutions: network and multi-core processors, network adapters, switches, cables, software and silicon, that accelerate application runtime and maximize business results for a wide range of markets including high performance computing, enterprise data centers, Web 2.0, cloud, storage, network security, telecom and financial services. More information is available at www.mellanox.com.

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