PARTNER SOLUTION BRIEF

Software Defined Storage Achieves Unparalleled Performance

Leveraging AIC NVMe Cluster Servers, Mellanox Networking Technology and Seagate SSDs

The Solution

Deploying Seagate NVMe SSDs across AIC NVMe cluster servers with Mellanox networking achieves scalable, high performance application acceleration

Performance Highlights

- 750,000 4KB IOPS random read performance
- 480,000 4KB IOPS random 70% read and 30% write performance
- 13.4 GB/s 1MB sequential read and 13.4 GB/s 1MB random read throughput
- 8.2 GB/s 1MB sequential write and 2.6 GB/s 1MB random write throughput
- Scale performance and throughput as needed by adding appropriate cluster servers and network connections

Executive Summary

The datacenter is going through a drastic transition as new technologies, innovators and partnerships emerge. One such emergence is Software Defined Storage (SDS). Per IDC, SDS is “a flexible enterprise storage infrastructure that leverages the agility of virtualized resource allocation, management, and reclamation to meet extremely dynamic business requirements”. At the core, SDS decouples the storage hardware from software and enables the use of various storage deployments such as block, file and object storage. The flexibility of SDS enables administrators to deploy storage simply and dynamically.

SDS Software Provider: Dell EMC ScaleIO

Dell EMC ScaleIO is one such example of a flexible software-defined block storage with a scalable virtual SAN. ScaleIO works by installing lightweight software components on existing standard x86 servers. Also, ScaleIO allows the convergence of compute and storage on the same server node so it can run alongside other applications on the server (physical, virtual, or cloud) and can leverage any type of storage device. It provides extremely effective block IO decentralization, allowing the system to expand to thousands of storage nodes and provide massively parallel IO computing.

In order for a user to take full advantage of what ScaleIO SDS can offer, it is imperative to ensure the configuration and topology meet application storage performance requirements and enable the configuration to scale. This includes all hardware components such as servers, storage and networking they select.

Collaborative Solution

Seagate, AIC and EMC understood the need for a simple and cost effective approach to deliver the highest performance needed for demanding multiple database and multiple user environments. Together they are bringing dramatic application acceleration benefits to SDS environments with the following solution.

Seagate

Seagate flash storage is a critical component in making real world high performance and high availability SDS a reality. Seagate, for these tests, utilized the Seagate Nytro XF1440 2.5-inch NVMe SSDs. The XF1440 provides the balanced power and performance for high-density data centers which results in more computing, while utilizing less data center space and lowering costs.

Listed below are some of the key benefits of using Seagate NVMe SSDs with any software-defined-storage solution:

- Achieve higher performance and throughput per cluster server
- Achieve higher performance and throughput per cluster server
- Satisfy enterprise reliability and data integrity requirements with power loss and end-to-end data protection
- Scale storage capacity with a range of SSD models up to 1.92TB
- Optimize storage capacity and performance with read-intensive and mixed IO workload SSD models
- Reduce TCO up to 66% by using less power than other competitive NVMe SSDs

AIC SB122A-PH All-NVMe Flash Array Server Platform

The AIC SB122A-PH is an all-NVMe flash array server platform designed for high IOPs and low latency storage applications. Its PCIe connectivity comes directly from the CPUs eliminating the latency usually added by a PCIe switch. AIC SB122A-PH supports 10 NVMe SSDs and 2 PCIe 3.0 cards in a 1U form factor. This ideal building block for SDS solutions includes key features below:

- Supports two Intel® Xeon® Processors E5-2600 v3 and v4 product family
- With Intel® C612 Series Chipset to provide 5+ years product life cycle
- Flexible IO usage with Max IO™ to support up to 2 PCIe 3.0 cards
- Onboard Baseboard Management Controller for system management and IPMI control
- Front-to-back airflow and hot-swap redundant fans to provide optimal thermal conditions

Mellanox Networking

Mellanox ConnectX-4 and ConnectX-4 Lx Ethernet network adapters provide the highest performance and most flexible solution for high-performance, Web 2.0, Cloud, data analytics, database, and storage platforms. Utilizing the advantages of Open networking with Mellanox Spectrum Ethernet switches which allow a multitude of operating systems to be mounted on it and deliver the highest performance and port density with a complete chassis and fabric management solution enabling converged data centers to operate at any scale while reducing operational costs and infrastructure complexity.

Dell EMC ScaleIO Tested Configuration

To demonstrate this solution's benefits for SDS, the following 3-node cluster configuration was deployed and tested with ScaleIO. This includes:

- (30) Seagate NVMe SSDs
  - XF1440 800GB 2.5"
  - 10 NVMe SSDs per server
- (3) AIC 1U NVMe Cluster Servers
  - Model SB122A-PH
  - Memory 64GB 2400MHz
  - 10 NVMe 2.5" SSD bays
  - Mellanox 100Gb/s dual-port NIC
- (8) Client x86 Servers
  - Mellanox 40Gb/s dual-port NIC
  - Synthetic workload (fio version 2.15)

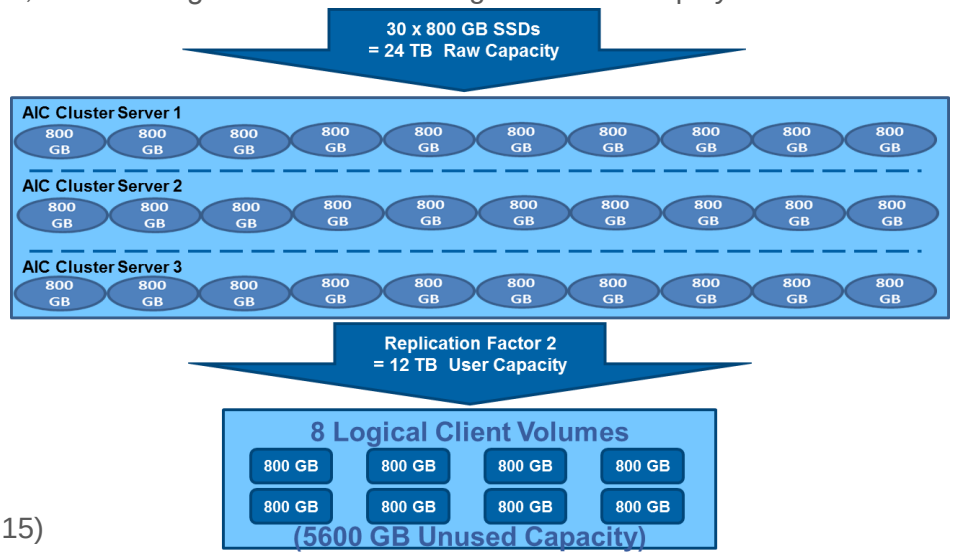

Figure 1. ScaleIO Cluster: Physical to Logical Storage Layout

Demonstrated Performance Results

Chart 1 to the right shows the 4KB small block random read, random 70% read and 30% write, and random write IO performance results.

The maximum sustained 4KB random read performance resulted in 750,000 IOPS. The maximum sustained random 70% read and 30% write performance is 480,000 IOPS, while the maximum sustained random write performance is 270,000 IOPS. It is important note the performance limit is not the Seagate SSDs in the configured solution, which could scale performance beyond this. To scale performance higher, additional cluster servers and respective network connections would be needed.

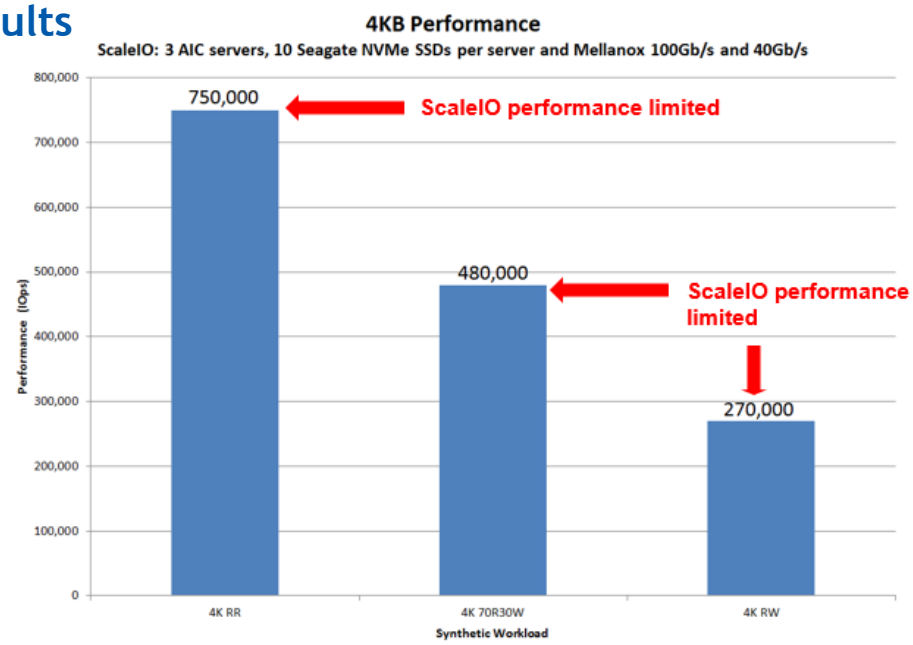

Chart 1. ScaleIO 3-Node Cluster Results

Demonstrated Throughput Results

The second chart shows the resulting 1MB large block throughput results. The first workloads indicated at the top of the chart show 1MB sequential read and 1MB random read throughput of 13.4 GB/s for both workloads. Both of these throughput numbers are limited by the three 40Gb/s client network connections to the 3 cluster servers. Similar to IOPs, additional cluster servers and respective client network connections could be added. Also additional or faster client network channels could be utilized to increase throughput.

The workload indicated below this is the 1MB sequential write throughput of 8.2 GB/s which is limited by the single 100Gb/s cluster replication or mirroring network channel. As the NVMe SSD device maximum transfer size parameter is set to 256KB, each 1MB client write is broken up into four separate 256KB write operations. It is important to note, that for or each 1MB write, a total of eight 256KB writes must occur to store user data with mirrored redundancy in the ScaleIO storage pool. The 100Gb/s data replication network is actually transferring twice the amount of data as the 8.2 GB/s (65.6Gb/s) result reported by the clients. Throughput can be increased by expanding the replication network topology.

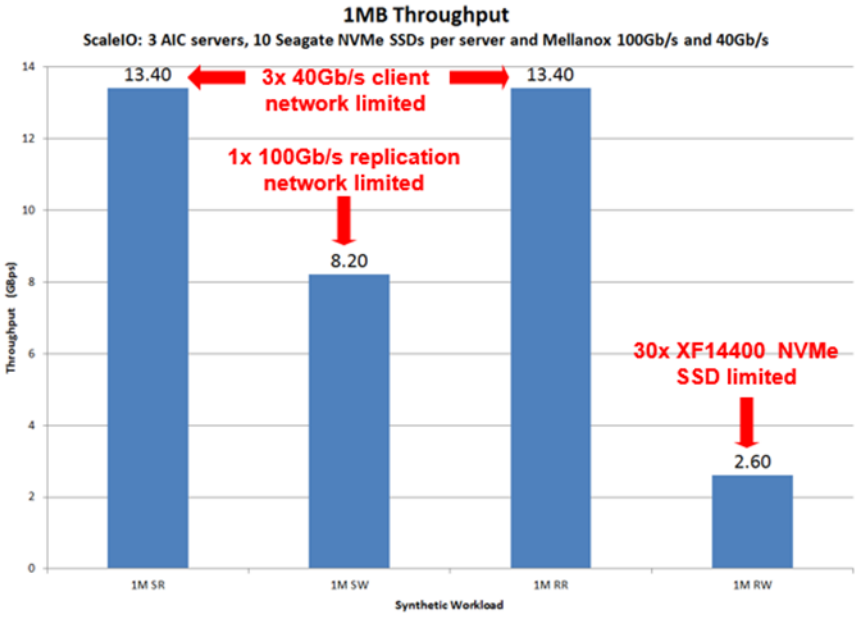

Chart 2. ScaleIO 3-Node Cluster Results – Throughput

The last workload highlighted in the right of the chart is the 1MB random write throughput of 2.6 GB/s which is limited by the 30 Seagate XF1440 NVMe SSDs. Again the cluster is actually transferring twice the amount of data as indicated by the 2.6 GB/s result reported by the clients. To scale total random write throughput higher, add additional cluster servers.

The Right Solution at the Right Time

Software Defined Storage is rewriting the datacenter and Seagate, AIC, Mellanox and Dell EMC are enabling this charge. Powering a SDS deployment using the solution described in this brief is a must in fast-tracking SDS to the highest performance, agility and operational efficiencies.

Learn More:

Representative from each company will be available to discuss the demo and answer any questions you may regarding configuration, and performance.