TOSHIBA

Lab Report

QSAN XCubeNXT "XN5124D" Unified Storage and Toshiba 22TB Nearline SAS HDD

Author: Rainer W. Kaese, Senior Manager
Business Development, Storage Products Division,
Toshiba Electronics Europe GmbH

Virtualization, containerization, surveillance, back up – for such workloads, modern enterprises and businesses need large-scale, multi-functional storage more than ever. Performance is key here and, with a view to this, the Toshiba team started a new test in its laboratory using the next-generation Toshiba Enterprise Capacity MG-Series HDDs, the 22TB SAS 12Gb/s model MG10SFA22TE, installed in QSAN's flagship XN5124D (Figure 1). This is a 4U/24-bay fully integrated unified storage that can either operate as network attached storage (NAS) providing shared folders/file system or as block storage for dedicated storage area networks (SAN) like iSCSI, Fibre Channel, or even both in parallel. QSAN promotes it as "the newest generation of XCubeNXT that harnesses industry-leading versatility for mixed workloads." Additionally, the highly available and secure multiple protocol and cross QSAN platform support is claimed as being the best Total Cost of Ownership (TCO) for capacity demanding applications.

Figure 1: The QSAN XCubeNXT XN5124D in Toshiba's HDD Lab

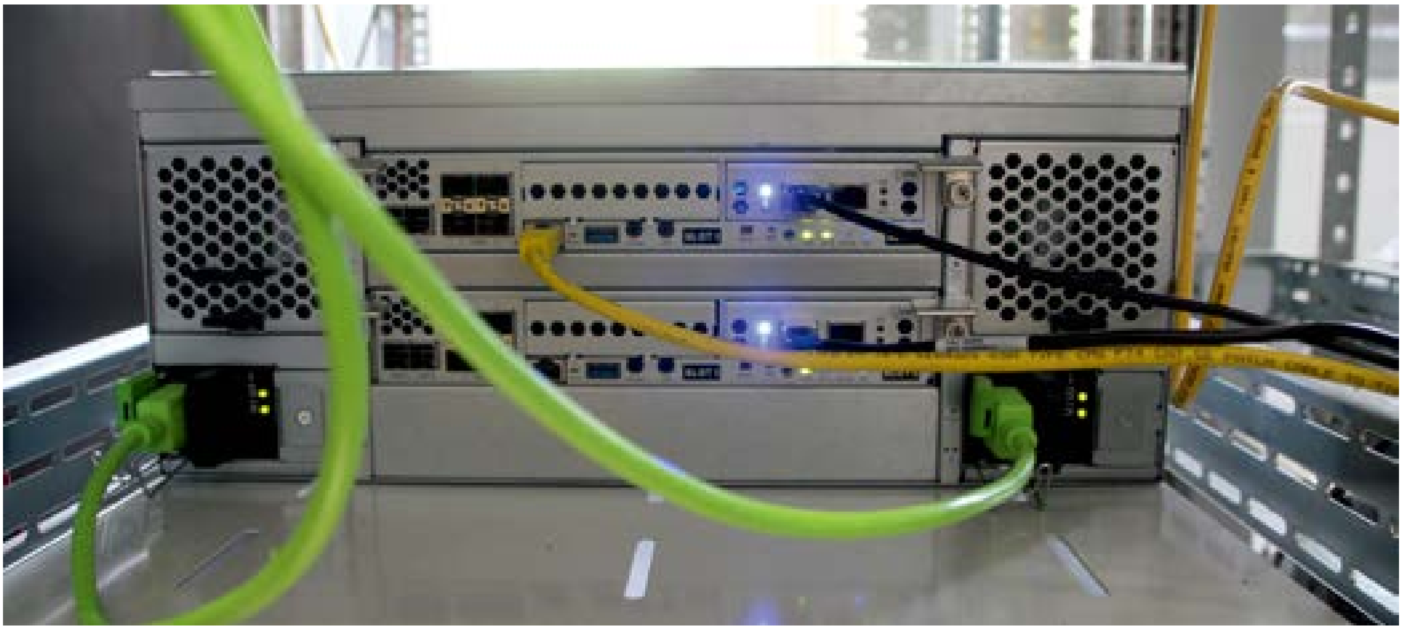

Figure 2: XN5124D Network Connections

Configuration & Performance Comparisons Maximum Performance in Focus

In previous lab reports of 24-bay storage units (like the QSAN XN8024D), the maximum performance was limited by the network interface - typically 10GbE - allowing up to 1200 MB/s. Due to this limitation, we were not able so far to evaluate the performance limits of the unit. The XN5124D model, in the lab, was the first 24-bay storage unit that came with 25GbE interfaces, so we focused to evaluate the maximum performance for different HDD configurations.

Setup

The XN5124D comes with a dual controller architecture, so it's best suited for high-capacity SAS Nearline HDDs. Rackmount, with dual path from network down to HDD access and dual/redundant power supply, it's definitely Enterprise suited. To achieve large-scale, we installed the current top model of the Toshiba Enterprise Capacity MG-Series, the 22TB SAS 12Gb/s model MG10SFA22TE.

By default, each controller is equipped with 4 channels of 10GbE SFP+, a 1GbE RJ45 management port, and two SFF-8644 (mini-SAS-HD) ports for capacity extension using external JBODs. Two sub-module slots enable an upgrade for different connectivity technology, such Fibre Channel, or – like in our case – 25GbE with two SFP28 ports. We connected one port of each controller unit via SFP28 direct-attached-copper cable to the applications server (Figure 2).

For reference: The model version we are testing is XN5124D-044C20, with controller firmware 4.0.2.

HDD Configurations

For the performance evaluation, we created one pool with a 24 HDD disk group in different RAID level configurations (Table 1).

RAID6 of 24 HDDs: One RAID set with double parity. This gives a net capacity of 22 HDDs, with a high data protection level due to the dual parity. If one drive fails, the data is still

RAID Level / Performance	RAID6	RAID60	RAID60	RAID60	RAID10
Number of sub-arrays	1	2	4	6	12
Pool Efficiency (%)	92	83	67	50	50
Usable Capacity (out of 528TB)	484	440	352	264	264

Table 1: RAID settings and Storage Efficiency

TOSHIBA

protected by the second parity. Pool efficiency is 92% (22 out of 24 HDDs carry user data), or 484TB usable data out of 528TB total installed capacity.

RAID60 of 24 HDDs with a different number of sub-arrays (2/4/6): RAID60 – equals RAID6 groups striped together – allows a faster, more parallel access to the drives. But, as for each sub-array, two parity units are required and the pool efficiency is reduced – the more arrays, the lower the efficiency.

RAID10 of 24 HDDs: This configuration consists of 12 sub-arrays with mirrored data on two disks. This configuration is known as rather performant in writing, especially random writing, since no parity information has to be calculated – the data is simply mirrored. As a consequence, the efficiency drops to 50%, which is typical for all mirror based configurations, and in comparison to RAID6x type of data protections, RAID10 is not strictly double protected. If one mirror fails, the data is only present on one remaining non-protected mirror disk.

On pools with the above RAID levels, we installed some volumes with one 50TB volume always connected as LUN to an iSCSI host group (Figure 3). This iSCSI LUN was connected to an application server as a Windows physical drive. We measured performance, HDD temperature, and power consumption while evaluating this physical drive with “fio”-based benchmarks.

All measurements were done with block volume status online = fully initialized, for parity RAID levels such as RAID6 and RAID60 this takes about two hours for 10TB of volume capacity.

The settings were as follows (Table 2):

Pool Settings	Disk Write Cache	Enable
	Disk Read Ahead	Enable
	Disk CMD Queuing	Enable
	Disk Standby	Disable
Volume Settings	Backgr. IO Priority	Low
	Cache Mode	Write Back
	Video Edit Mode	Disable
	Read Ahead	Enable

Table 2: Pool and Volume Settings

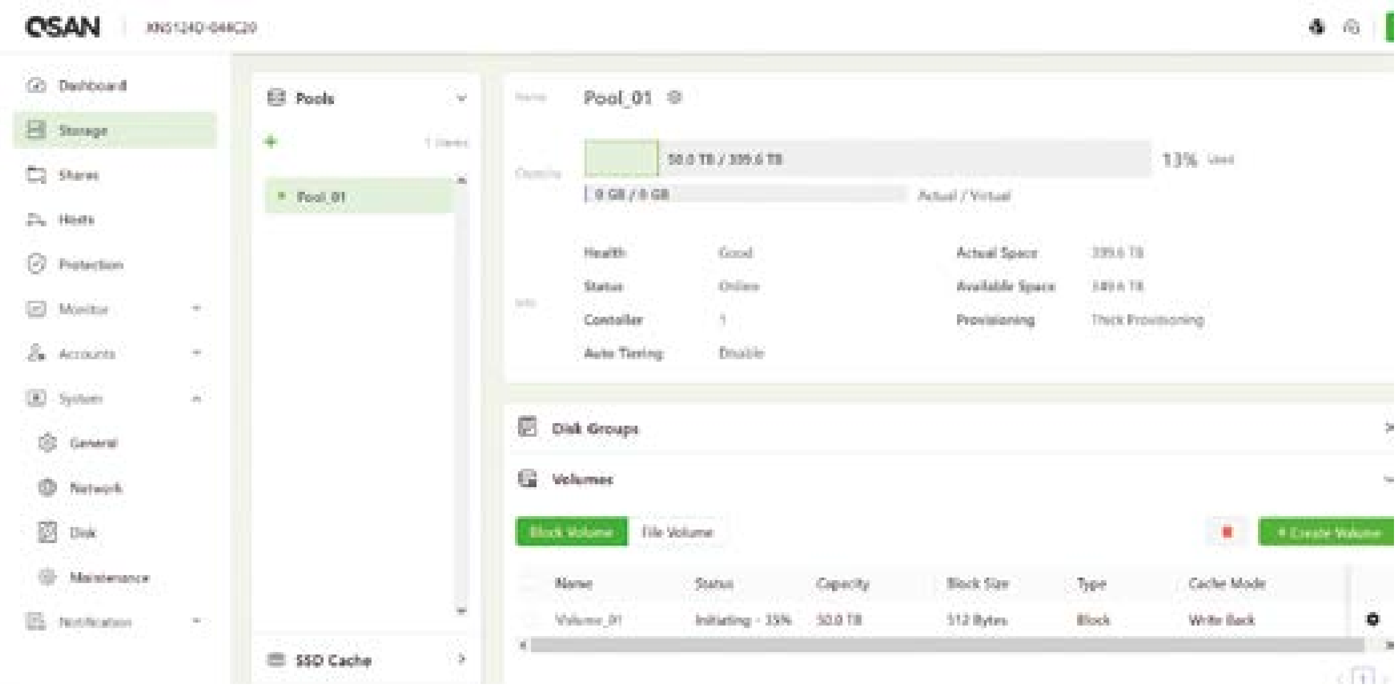

Figure 3: Pool/Volume Configuration in QSAN’s QSM GUI

Performance Measurement Results:

RAID Level / Performance	RAID6	RAID60	RAID60	RAID60	RAID10
Number of sub-arrays	1	2	4	6	12
Pool Efficiency (%)	92	83	67	50	50
Usable Capacity (out of 528TB)	484	440	352	264	264
Sequential read 1MB (MB/s)	2600	2730	2590	2750	2770
Sequential write 1MB (MB/s)	1500	1470	1500	1460	1450
Random read 4kB (IOPS)	3720	3660	3510	3340	3340
Random write 4kB (IOPS)	1600	1600	1550	2700	8160
Mix r/w 4k/64k/256k/2M (IOPS)	770	790	790	820	920

Table 3: Performance Measurement Results

A sequential reading performance of ~2700 MB/s, matches to the speed of the 25GbE iSCSI connection to the applications server, so the speed is determined by the network and not by the RAID level. Sequential writing speed is around 1500 MB/s, regardless of the RAID level. This indicates that the RAID6 level is in fact the best solution for workloads dominated by sequential access (archiving, streaming, video), due to its high storage efficiency. Any more parallel disk access arrangements, such as RAID60 or RAID10, do not bring any higher performance, they just decrease the amount of available storage capacity.

Even for random reading (important for AI/Video analytics type of applications), the RAID6 shows highest IOPS performance, since, in this configuration, the data is spread over the largest number of active disks (22 out of 24).

For workloads dominated by random writing (ie. active directory / database & email serving / virtualization), the parallel structures of RAID60 with a larger number of sub-arrays provide higher performance numbers. This is true for RAID10 as well, with no parity operation, just plain mirroring of the data even up to 8k IOPS. So, if the writing performance of the HDD mass storage is critical, RAID10 is the best configuration.

Although the storage efficiency of RAID10 is low, the protection level is not as high as in RAID6 and RAID60 – since a failure of two wrong HDDs within one single mirror would result in data loss, while in RAID6/60 any two drives can fail without loosing data.

Furthermore, for the above-mentioned random type or workloads, SSD caching may provide a significant performance benefit. But as the level of improvement depends highly on the actual workload such an option is not in the scope of this lab evaluation.

Power and Temperature Measurements

Maximum Power at Startup:	480 W
Power under full read/write workload:	480 W
Power when unit is idle (no read/write):	330 W
Standby Power (unit off):	10 W

RAID Level / Power Consumption	RAID6	RAID10
Sequential read 1MB (W)	410	410
Sequential write 1MB (W)	400	400
Random read 4kB (W)	400	420
Random write 4kB (W)	420	380
Mix r/w 4k/64k/256k/2M (W)	410	390

Table 4: Power Consumption Measurements

The highest power consumption measured is still below 500W, and with 400W in average for up to half a Petabyte of net storage, this model, equipped with high-capacity HDDs, such as the Toshiba MG10SFA22TE, adds a significant contribution to the important efforts of reducing power consumption for IT- and datacenter storage. Less than 1kW/PB for the complete operation is an excellent value, including high speed network interfaces.

Ambient Lab Temperature:	24 degC
Minimum HDD Temperature (smart-value):	36 degC
Maximum HDD Temperature (smart-value):	42 degC (HDDs in full read/write activity)

Toshiba MG-Series HDDs can operate up to an internal temperature of 60 degC, so a maximum temperature of 42 degC is far away from any limit. But the long-term reliability tends to decrease at average or constant temperature levels beyond

TOSHIBA

42 degC. Hence, for the highest possible long-term reliability, an ambient temperature of 24 degC is maintained, ensuring that a HDD temperature of 42 degC is not be exceeded in significant levels of time and value.

Measurement Scripts

Measurements were made with fio. The scripts are listed for reference:

```
fio --filename=\\.\Physicaldrive1 --direct=1 --rw=read --bs=1m --iodepth=128 --time_based
--runtime=300 --group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=4
--norandommap --randrepeat=0 --output=seqread.txt

fio --filename=\\.\Physicaldrive1 --direct=1 --rw=write --bs=1m --iodepth=128 --time_based --runtime=300 --group_reporting
--name=job1 --ioengine=windowsaio --thread --numjobs=4
--norandommap --randrepeat=0 --output=seqwrite.txt

fio --filename=\\.\Physicaldrive1 --direct=1 --rw=randread --bs=4k --iodepth=128 --time_based --runtime=300 --group_reporting
--name=job1 --ioengine=windowsaio --thread --numjobs=64
--norandommap --randrepeat=0 --output=randread.txt

fio --filename=\\.\Physicaldrive1 --direct=1 --rw=randwrite --bs=4k --iodepth=128 --time_based --runtime=300 --group_reporting
--name=job1 --ioengine=windowsaio --thread --numjobs=64
--norandommap --randrepeat=0 --output=randwrite.txt

fio --filename=\\.\Physicaldrive1 --direct=1 --rw=randrw --bssplit=4k/20:64k/50:256k/20:2M/10 --iodepth=128 --time_based
--runtime=300 --group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=mixed.txt
```

Conclusion

The tested QSAN XN5124D unified storage system provides large capacity, high performance, and highly available and reliable enterprise data storage and is absolutely ready to satisfy the demanding capacity and performance needs of the workloads as described at the very beginning, such as virtualization, containerization, surveillance and back up.

With 24 Toshiba 22TB Enterprise SAS HDD, the raw capacity is 528 TB, and depending on configurations, from 264 TB (RAID10) up to 484 TB (RAID6) net capacity.

Almost 3 GB/s sequential reading- and up to 1.5 GB/s sequential writing performance, even without any SSD caching is an excellent value, matching to the theoretical limits of the 25GbE network connectivity. A RAID6 pool configuration is recommended, as it combines high sequential performance with double parity protection and achieves a storage efficiency as high as 92%. Random read performance is about 3000 IOPS, and the write performance peaks up to 8000 IOPS (in RAID10 configuration).

Maximum power consumption under full load is not much more than 400W (less than 1W per TB net capacity, including expander, controller, 25GbE SFP+ and RJ network interfaces -- an excellent value for power efficiency!).

The HDD temperature is rather low with a narrow spread – thanks to appropriate and efficient cooling. This supports long lifetime and low failure rates for the spinning disks. All this is achieved with a high level of data protection level and at low costs of ownership, which is last, but not least an important argument for businesses/enterprises as well.

Note of thanks to our partner

We would like to thank our partner QSAN (www.qsan.com) for providing the XN5124D test sample, but also for their great assistance and cooperation in setup and performance testing. This cooperation is an excellent example of a value chain cooperation, providing valuable guidance for QSAN and Toshiba customers on how to design a proven storage solution.

TOSHIBA

Toshiba Electronic Components Taiwan Corporation
4F., No.168, Sec.3, Nan-jing E. Rd., Taipei, 104105 Taiwan

Toshiba Electronics Asia (Singapore) Pte Ltd
20 Pasir Panjang Rd
#12-25/26 Mapletree Business City Singapore 117439

https://toshiba.semicon-storage.com/ap-en/storage.html

Issued 07/2024