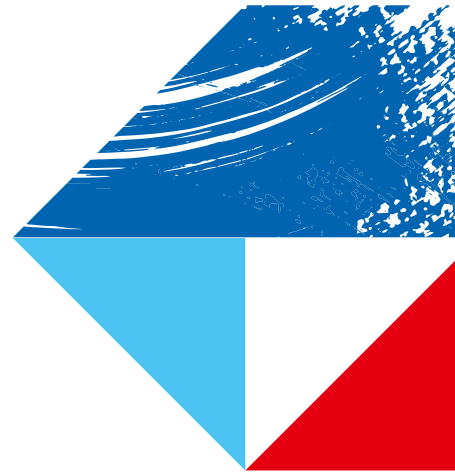# TOSHIBA

Lab Report

PROMISE VTrak J5960 JBOD

Rainer Kaese, Senior Manager Business
Development Storage Products,
Toshiba Electronics Europe GmbH

Introduction

Striving for climate neutrality and sustainability is key in
modern data centers. The aim: Keeping the power consump-
tion and the overall carbon footprint low, while processing
24/7 an enormous amount of data – data that has been and
will be stored on hard disk drives mostly.

Green initiatives affect all areas of the datacenter, from
techniques like waste heat re-usage or free cooling, down to
the use of sustainably produced inventory and state-of-the–
art Hard Disk Drive (HDD) technologies.

Here's where our lab reports starts:

The new VTrak J5960 4U 60-Bay toploader JBOD from PROMISE
is promoted as a "JBOD with green DNA", including a commit-
ment to environmental protection and sustainable production.
Toshiba Electronics Europe GmbH ("Toshiba") had the chance
to test and evaluate this JBOD in the lab, with 60 Toshiba
Enterprise HDDs with 18TB data storage capacity being used,
giving a total capacity of 1080TB.
Toshiba tested for function, performance, acoustics and power
consumption, with a focus on evaluating the green-JBOD
claimed by PROMISE.

Picture 1: PROMISE VTrak J5960 JBOD in the Toshiba lab.

TOSHIBA

Dimensions & mechanical features:

With 4U height and a chassis length of just 666mm, the J5960 is the shortest high density JBOD we have ever seen in the Toshiba HDD lab. It is just as long as a typical 66cm 2U server chassis and fits conveniently into any existing rack.
This is a big advantage over other JBODs: With many of them exceeding a 1000mm length and requiring longer racks or tending to create cabling issues.

The hot swappable IO-Modules (IOM) of J5960 are to be pulled out to the front of the JBOD – the connected cabling remains in the back of the unit. This makes field replacement of an IOM very easy, while typical rear-exchanges require access through a bunch of power and signal cables.

An interesting mechanical feature is the lid of the JBOD. It will be attached to the rack and will stay in the rack when the JBOD is pulled out for service. We have tested this, and it works really smoothly. As no lids have to be lifted, this JBOD needs to be pulled only as far as the failing drive requires. Also it can be installed at top locations of the rack.

HDDs are each fixed in their metal trays with 4 screws.

The status LEDs are normally switched off, and are only activated when the lid is removed (i.e. when the JBOD is pulled off the rack). This saves a few Watts of additional power.

Setup in the Toshiba lab

Model:	PROMISE VTrak J5960 4U-SAS-60-D BP
Firmware:	1023
Host OS:	Linux (Centos 7.9)
Host OS:	Windows (Windows Server 2019 Standard)
Host bus adapter (HBA):	Broadcom Avago HBA 9500-16e (Host IF: 8x PCIe-Gen4)
RAID adapter:	Microchip Adaptec® SmartRAID Ultra 3254-16e/e (16x PCIe-Gen4)

Tests with enterprise capacity (nearline) SAS drives:

Model name:	Toshiba MG09SCA18TE
Block size	512B emulated
Firmware	0104
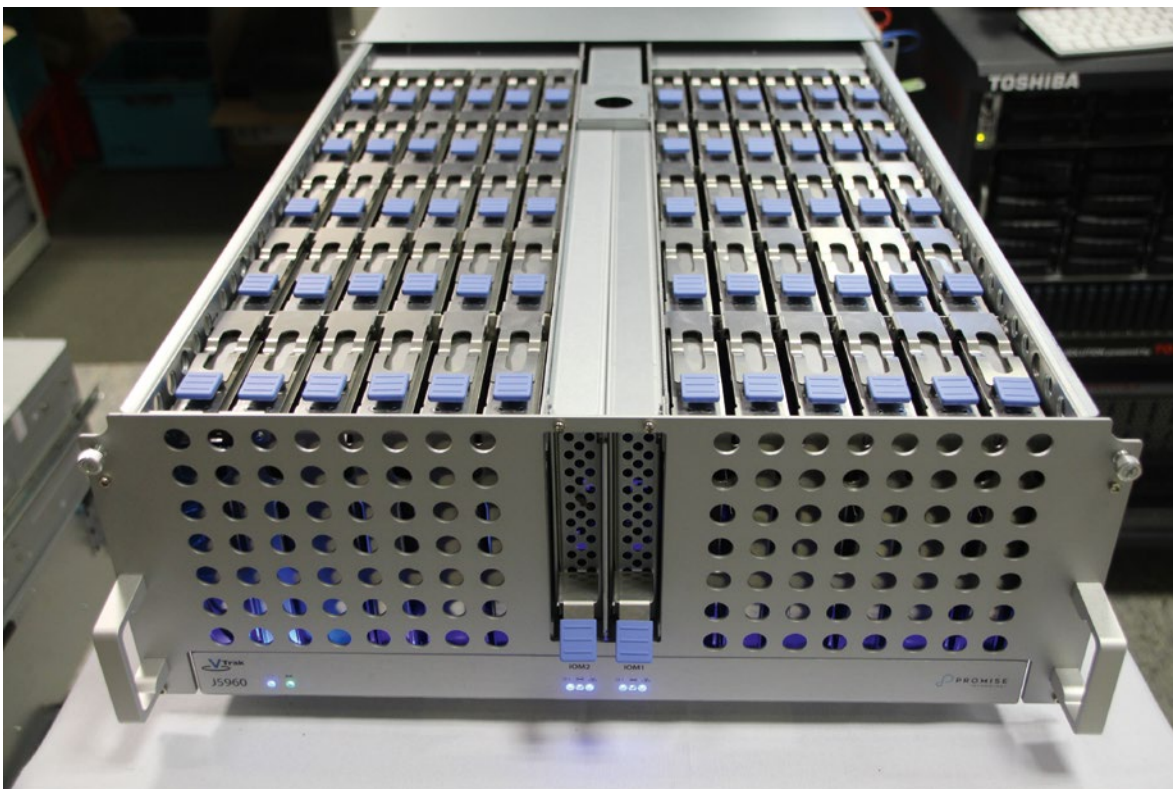
Picture 2: J5960 with open lid.
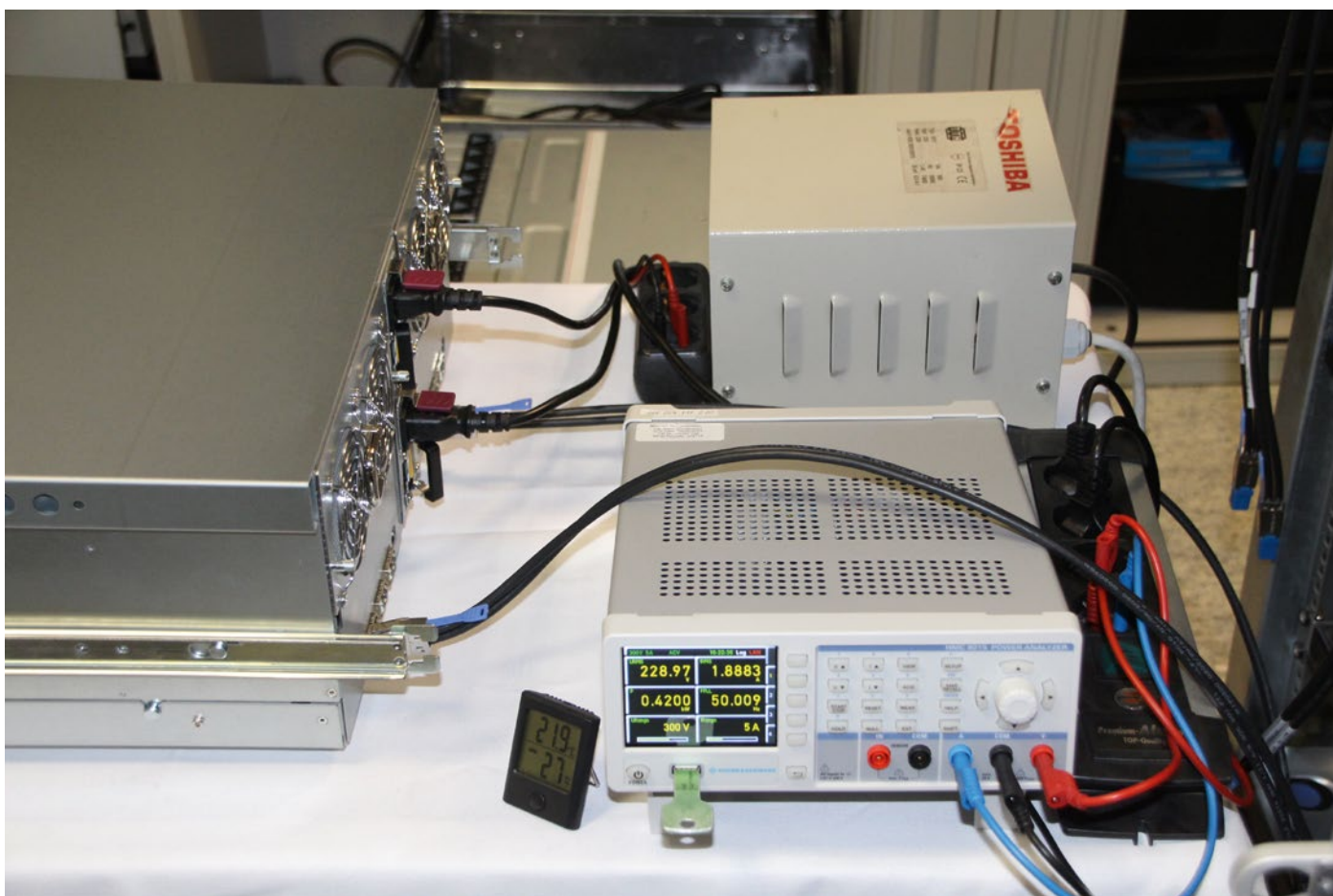
Data-rate outer diameter: 282MB/s

Power consumption
Idle_B:	3.36W
Sequential write:	7.62W
Sequential read:	8.71W
Random write:	6.64W
Random read:	9.47W

Basic JBOD functions:

Basic function:	ok
SAS IOM detected:	ok
Hot plug /re-insert:	ok
Smart reading:	ok
Enclosure management:	ok, tested with RJ11 serial connection

Picture 3: Toshiba HDD MG09SCA18TE in PROMISE tray.

Picture 4: Power measurement setup in the Toshiba HDD lab.

TOSHIBA

Hot Server

VTrak J5960

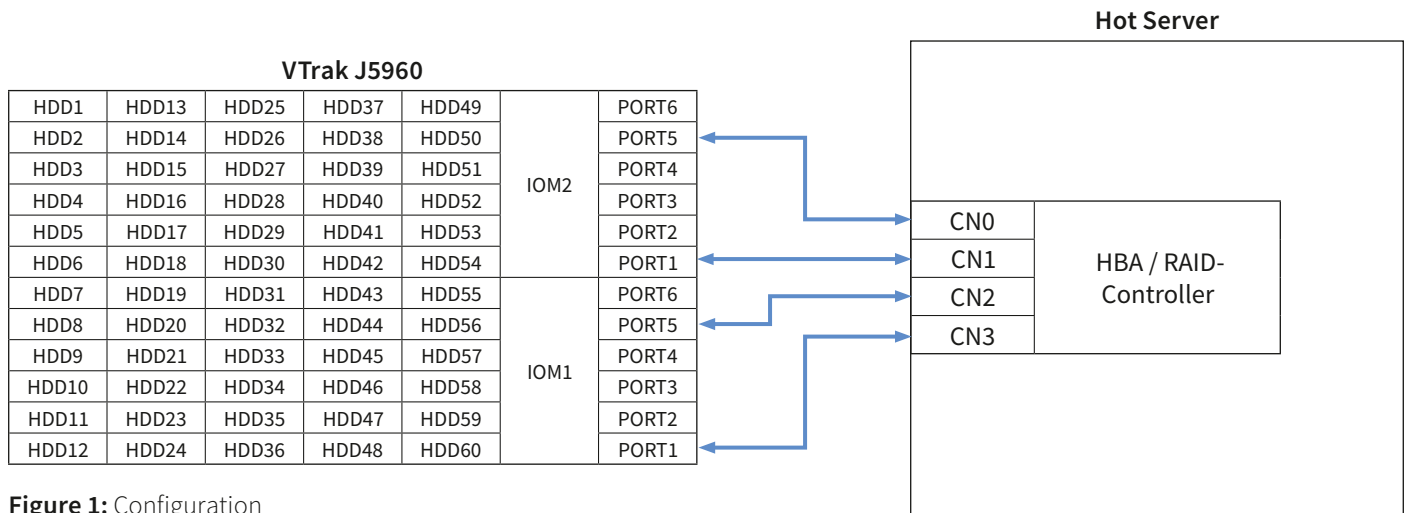|						
HDD1	HDD13	HDD25	HDD37	HDD49		PORT6
HDD2	HDD14	HDD26	HDD38	HDD50		PORT5
HDD3	HDD15	HDD27	HDD39	HDD51		PORT4
HDD4	HDD16	HDD28	HDD40	HDD52	IOM2	PORT3
HDD5	HDD17	HDD29	HDD41	HDD53		PORT2
HDD6	HDD18	HDD30	HDD42	HDD54		PORT1
HDD7	HDD19	HDD31	HDD43	HDD55		PORT6
HDD8	HDD20	HDD32	HDD44	HDD56		PORT5
HDD9	HDD21	HDD33	HDD45	HDD57		PORT4
HDD10	HDD22	HDD34	HDD46	HDD58	IOM1	PORT3
HDD11	HDD23	HDD35	HDD47	HDD59		PORT2
HDD12	HDD24	HDD36	HDD48	HDD60		PORT1

CN0
CN1 HBA / RAID-Controller
CN2
CN3

Figure 1: Configuration

For precise measurement of the power consumption, we used a high accuracy professional power analyser (R&S HMC8015).

JBOD on, no drives, SAS link to host on: 100W
JBOD with drives, maximum start-up power over 500ms: ... 850W
JBOD with raw drives at HBA Idle_B: 305W
Lambda (ratio of active and reactive power) 0.96
Noise at 1 m distance 80dB
Temperature ambient .... 23°C

VTrak J5960 has a default setting which places all (SAS) drives to Idle when not being accessed for about 2 minutes.
About 100W for a dual IOM JBOD without drives is an excellent power value. Dedicated single IOM JBODs can be as low as 80W under the same conditions, but typically dual IOM JBODs are in the 200-300W range.
About 300W as "idle with drives" is also an amazingly low value, other 60 bay JBODs start at 400W or more.
A high lambda factor of 0.96 means that the ratio of reactive power generated by the power supply units is extremely low (equating to 4% of the total power). The higher (i.e. better) the lambda factor, the smaller the reactive power.
Reactive power does not cost and does not create heat, but power supply rails have to be dimensioned for both active and reactive power – so for large data centers a high lambda is an important factor.

Performance measurements in the Toshiba lab

For tests with SAS drives, we have connected both IOM of the JBOD with two mini-SAS-HD cables each to the 4 mini-SAS-HD ports of the16e HBA and RAID controller.

This configuration gives a theoretical JBOD/HDD access-bandwidth of 4x 4.8GB/s = 19.2GB/s, but required path aggregation by installing multipath functionality.
As for configuration with HBA, multipath has to be enabled manually in Linux/Windows. A RAID controller (such as the Microchip Adaptec® Ultra-3254 Model) will detect the configurations automatically and invoke a correct multipath setting. Manual multipath and SAS-link aggregation only works for SAS drives. Configuration with SATA drives will still benefit from the aggregation of IOPS of the 60 HDDs, but the sequential band-width is usually limited to one mini-SAS-HD link (4.8GB/s).

We`ve tested several drive configurations with "fio" – flexible IO tester software – measuring sequential, random and mixed workload performance and related power consumption.
Tests have been done for individual drives connected via HBA and in RAID configurations as physical drives, as well as for logic drives. For logic drives we also measured the performance and power of a copy (i.e. reading and writing) of a large file.

JBOD setting remarks

The VTrak J5960's default setting is "HDD to Idle after 2 min of inactivity". It takes about 1200ms for the idle HDD to transit from an Idle to an Active state.
For RAID configurations, this function is recommended to be disabled – since in a large RAID some HDDs may not be active for more than 2 mins, even if the RAID is under full data load. If some HDDs within the RAID were set to Idle mode, long latencies of 1200ms would occur if these were accessed.
Disable Idle mode switching: Via serial cable to IOM management port (115200/8/N/1), CLI command is "enclosure –m –idlep 0".

The Toshiba tests were performed with "zoning mode 0", (=default setting). This means "no zoning", so all of the drives can be accessed from both IOM.

Some RAID controllers or HBAs may reject the connection of 4 cables into two IOM. If so, "zoning mode 1" may be a workaround. Here 30 drives are connected from one SAS port to IOM1 and another 30 drives are connected from one SAS port to IOM2. Configuration appears like two JBOD of 30 drives. Change zoning mode CLI command to "enclosure –m –z 1".

The theoretical maximum bandwidth is 282MB/s (single disk) x 60 = 16.2GB/s. With 14.5GB/s and 20K+ IOPS this configuration is close to the theoretical limit.

Power figures of 640W maximum prove the green credentials of the JBOD. The temperature spread of less than 10°C difference between hottest and coldest drive with the maximum temperature of less than 14°C on top of ambient deliver efficient cooling, supporting reliability and long lifetime of the HDDs.

All drives parallel as single physical devices (multipath):

OS:	Linux (Centos 7.9)
HBA/Controller:	Broadcom HBA9500-16e
HDD:	60x Toshiba MG09SCA18TE
Configuration:	Dual IOM 2x 2 Mini-SAS HD cables (3m in length) Multipath setup for disks

Workload	Power (W)	IOPS	Bandwidth (MB/s)
Sequential write 1024K	610		13300
Sequential read 1024K	640		14500
Random write 4K	510	24100	
Random read 4K	540	33900	
Mixed 4K/64K/256K/2M	540	22600	2350
Temperature ambient	23°C		
Temperature HDD min.	27°C		
Temperature HDD max.	36°C		

All drives as RAID10, Windows physical drive:

OS:	Windows Server 2019
RAID/Adapter:	Microchip Adaptec® SmartRAID Ultra 3254-16e/e (16x PCIe-Gen4)
HDD:	60x Toshiba MG09SCA18TE
Configuration:	Dual IOM 2x 2 Mini-SAS HD cables (3m in length)

Workload	Power (W)	IOPS	Bandwidth (MB/s)
Sequential write 1024K	510		8600
Sequential read 1024K	570		15300
Random write 4K	600	12800	
Random read 4K	730	9900	
Mixed 4K/64K/256K/2M	680	6100	1800
Idle (raid backgr.)	470		
Temperature ambient	25°C		
Temperature HDD min.	29°C		
Temperature HDD max.	38°C		

Script 1 – All drives parallel as single physical devices (multipath):

```
fio --direct=1 --bs=1m --iodepth=16 --size=32g --ioengine=libaio --group_reporting --rw=write --output=se-
qwrite.log --name=/dev/mapper/mpath{a..z} -- name=/dev/mapper/mpatha{a..z} --name=/dev/mapper/mpathb{a..h}

fio --direct=1 --bs=1m --iodepth=16 --size=32g --ioengine=libaio --group_reporting --rw=read --output=se-
qread.log --name=/dev/mapper/mpath{a..z} -- name=/dev/mapper/mpatha{a..z} --name=/dev/mapper/mpathb{a..h}

fio --direct=1 --bs=4k --iodepth=16 --size=512m --ioengine=libaio --group_reporting --rw=randwrite --out-
put=randwrite.log --name=/dev/mapper/mpath{a..z} -- name=/dev/mapper/mpatha{a..z} --name=/dev/mapper/mpath-
b{a..h}

fio --direct=1 --bs=4k --iodepth=16 --size=512m --ioengine=libaio --group_reporting --rw=randread --out-
put=randread.log --name=/dev/mapper/mpath{a..z} -- name=/dev/mapper/mpatha{a..z} --name=/dev/mapper/mpath-
b{a..h}

fio --direct=1 --bssplit=4k/20:64k/50:256k/20:2M/10 --iodepth=16 --size=8g --ioengine=libaio --group_report-
ing --rw=randrw --output=mixed.log --name=/dev/mapper/mpath{a..z} -- name=/dev/mapper/mpatha{a..z} --name=/
dev/mapper/mpathb{a..h}
```

Script 2 – All drives as RAID10, Windows physical drive:

```
 fio --filename=\\.\Physicaldrive1 --direct=1 --rw=write --bs=1m --iodepth=16 --time_based --runtime=300
--group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=seqwritephysical.log

 fio --filename=\\.\Physicaldrive1 --direct=1 --rw=read --bs=1m --iodepth=16 --time_based --runtime=300
--group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=seqreadphysical.log

 fio --filename=\\.\Physicaldrive1 --direct=1 --rw=randwrite --bs=4k --iodepth=16 --time_based --runtime=300
--group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=randwritephysical.log

 fio --filename=\\.\Physicaldrive1 --direct=1 --rw=randread --bs=4k --iodepth=16 --time_based --runtime=300
--group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=randreadphysical.log

 fio --filename=\\.\Physicaldrive1 --direct=1 --rw=randrw --bssplit=4k/20:64k/50:256k/20:2M/10 --iodepth=16
--time_based --runtime=300 --group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64
--norandommap --randrepeat=0 --output=mixedphysical.log
```

As writing in RAID10 always goes to two mirrored devices in parallel, the write speed drops to half the value. Idle power in RAID configurations is almost at the level of active power since the RAID controller is always accessing the drives for some background consistency checking.

All drives as RAID10, Windows logic volume:

OS:	Windows Server 2019
RAID/Adapter:	Microchip Adaptec® SmartRAID Ultra 3254-16e/e (16x PCIe-Gen4)
HDD:	60x Toshiba MG09SCA18TE
Configuration:	Dual IOM 2x 2 Mini-SAS HD cables (3m in length)

Workload	Power (W)	IOPS	Bandwidth (MB/s)
Sequential write 1024K	520		6900
Sequential read 1024K	550		15000
Random write 4K	520	11100	
Random read 4K	540	29500	
Mixed 4K/64K/256K/1M	550	8100	2400
Windows copy	500		550
Idle (raid backgr.)	470		
Temperature ambient	25°C		
Temperature HDD min.	29°C		
Temperature HDD max.	39°C		

Script 3 – All drives at RAID10, Windows logic volume:

```
 fio --filename=test --size=1T --direct=1 --rw=write --bs=1m --iodepth=16 --time_based --runtime=300
--group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=seqwritelogical.log

 fio --filename=test --size=1T --direct=1 --rw=read --bs=1m --iodepth=16 --time_based --runtime=300 --group_
reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0 --out-
put=seqreadlogical.log

 fio --filename=test --size=1T --direct=1 --rw=randwrite --bs=4k --iodepth=16 --time_based --runtime=300
--group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=randwritelogical.log

 fio --filename=test --size=1T --direct=1 --rw=randread --bs=4k --iodepth=16 --time_based --runtime=300
--group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64 --norandommap --randrepeat=0
--output=randreadlogical.log

 fio --filename=test --size=1T --direct=1 --rw=randrw --bssplit=4k/20:64k/50:256k/20:2M/10 --iodepth=16
--time_based --runtime=300 --group_reporting --name=job1 --ioengine=windowsaio --thread --numjobs=64
--norandommap --randrepeat=0 --output=mixedlogical.log
```

This benchmark for a logic volume in Windows does not exercise the entire drive size, but operations on a test file size of 1TB. Hence it is more realistic to actual use cases. The more than 32k IOPS is due to the fact that the seek operation does not cover the entire capacity range. This also results in significantly lower power consumption for write operations, compared to full 500TB seeks for physical drives from above.

Configuration with SATA drives

We also tested the VTrak J5960 JBOD with SATA drives (Model MG09ACA18TE). As the SATA interface has only one signal path, we used a single IOM configuration (IOM2 was plugged out, and all four cables were attached to IOM1).

The sequential performance is typically limited to the level of one mini-SAS HD cable (about 4.3GB/s for sequential reading and writing). IOPS values match well with the SAS based results, as these are not bandwidth limited (10k IOPS for full capacity, 30k IOPS for logic drives with 1TB data range).

The power consumption with SATA HDDs and single IOM operation is about 70~80W lower than for the equivalent setup with SAS drives and a dual IOM. This is due to the single SATA drive itself having about 0.4~0.8W lower power consumption (depending on load) than the same drive with an SAS interface and the missing second IOM.

System considerations

A network bandwidth of 100Gbit/s and a storage bandwidth of 12.5GB/s do match quite well. For systems requiring such highest sequential performance we recommend using the nearline SAS enterprise capacity HDDs of the Toshiba MG-Series in dual IOM setups.

If network bandwidth is 25Gbit/s or lower and highest capacity is the main target, nearline SATA drives in single IOM configurations may be used as well, as the storage bandwidth is typically limited to 4GB/s, which again matches to the 25Gbit/s network link speed.

Summary

The PROMISE VTrak J5960 is an energy-efficient and easy-to-maintain 60 bay top-loader JBOD. In its class it's the most compact model with a total length of just 666mm. Fully populated with Toshiba 18TB drives, it offers a total capacity of more than one PB at just around 500W of power consumption.

Storage configurations based in this JBOD have been evaluated and Toshiba was able to show aggregates performances of 60 HDDs up to 15GB/s sequential throughput and over 30k of random IOPS, while the efficient cooling and airflow management of the JBOD supports long lifetime and high reliability of the HDDs by keeping their temperature in full operation always lower than 14°C above ambient temperature.

TOSHIBA

Toshiba Electronics Europe GmbH

Hansaallee 181
40549 Düsseldorf
Germany

info@toshiba-storage.com
toshiba-storage.com