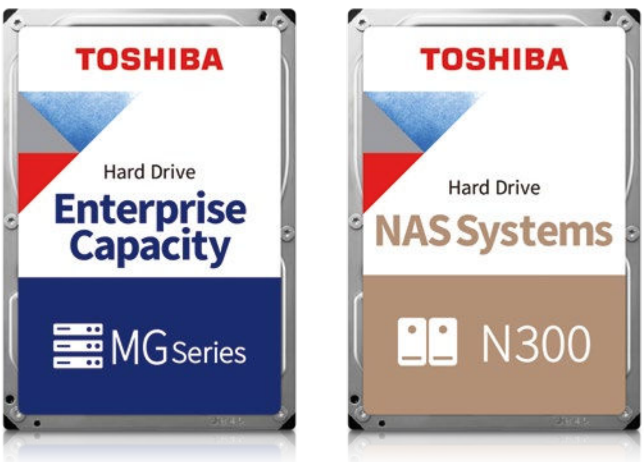


Whitepaper

Mixing Hard Disk Drives of 512 Byte native and 512 Byte emulated Block Size in RAID Systems

toshiba-storage.com

Introduction

Hard Disk Drives (HDDs) for NAS and Enterprise Storage come in two different flavors of block size. Traditionally, the internal block size of HDDs is 512 byte, with 512 byte also used at the external interface. This technology is called 512 byte native, or 512n. Newer HDD models use an internal block size of 4k byte for higher storage efficiency, but still provide 512 byte at the external interface through a transparent emulation process. This technology is called 512 byte emulated (512e). Additionally drives with the 4k byte internal block format at the interface (4k byte native) are available, but these are rare and exclusively used in specific enterprise applications.

512n is still available for older HDD models, and for backwards compatibility reasons for modern HDDs up to 4TB, while modern HDDs of more than 4TB are all 512e.

The report of diagnostics tools like smartmontool (https://www.smartmontools.org/) reports this configuration as following:

```
>smartctl -a /dev/sdx

Device Model:    TOSHIBA HDWG440
Sector Size:     512 bytes logical/physical                ( = 512 native )

Device Model:    TOSHIBA HDWG480
Sector Size:     512 bytes logical, 4096 bytes physical    ( = 512 emulated )

Device Model:    TOSHIBA MG06ACA10TA
Sector Size:     4096 bytes logical/physical               ( = 4k native )
```

From the same hardware series, typically 512e drives are slightly faster than 512n drives, due to the lower number of overhead bits in 512e, which leads to a higher data rate at the same rotational speed (512n - Toshiba MG08ADA400N: 233 MB/s vs. 512e Toshiba MG08ADA400E: 241 MB/s).

Our Evaluation

In the Toshiba HDD laboratory, we evaluated the flexibility of 512n vs. 512e replacements in RAID systems. A RAID (redundant array of independent disks) combines a bunch of physical hard disk drives into a larger array, typically with some redundancy (parity or mirror). If a drive fails, the RAID controller can rebuild the array information into a replacement drive making use of the parity or mirror information.

Following the emergence of 512e drives, we were curious to know: would an array of 512n drives accept a failing drive being replaced by a 512e drive? To confirm a free mixing of both versions, we also tested the other way round (512n replacing a failing 512e drive).

HDD Setup: The Drives (512n and 512e)

For the evaluation we used the popular Toshiba NAS HDD N300 as the 4TB version with 512n block size, and the Enterprise model of the MG-Series MG08ADA400E with 512e block size. These are the most common 4TB drives and this capacity is often used in smaller RAID sets of typically 2 or 4 bays.

We simulated a failing disk by simply unplugging it under full power and operation (hot-removal), and waiting for 10 minutes while checking the performance and function of the RAID array in this failed drive mode. Then we re-inserted the replacement drives, also under full power, and waited for RAID rebuilding without any power cycle.

Picture 1:
512n Hard Disk Drive (N300 4TB)

Picture 2:
512e Hard Disk Drive (Enterprise MG08ADA400E)

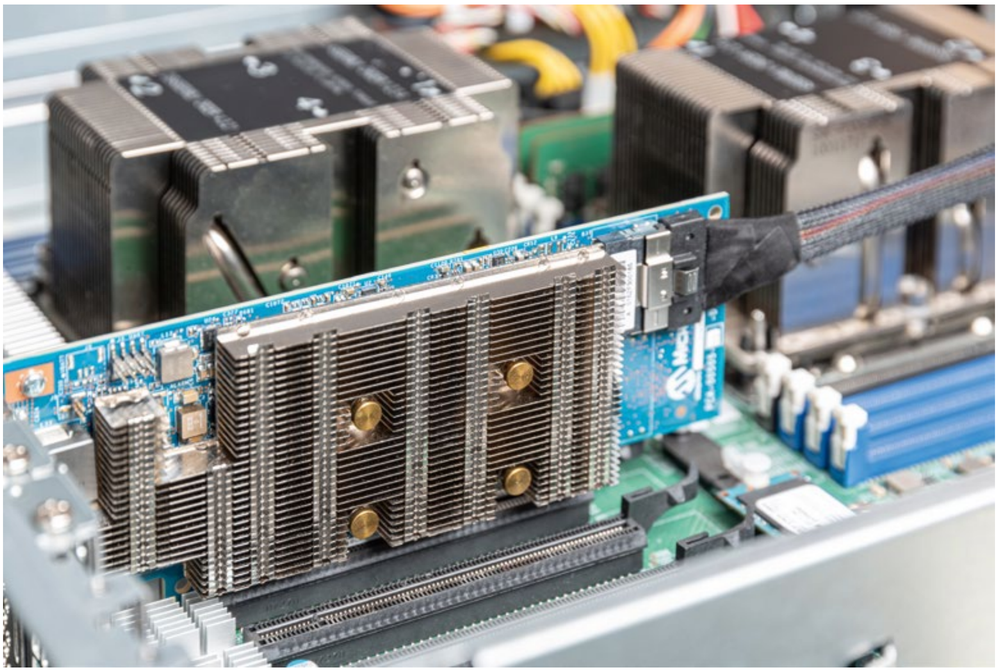
RAID System Setups: 4-Bay Desktop NAS and Internal RAID Controller

The two most common implementations of RAID configurations are integrated network attached storage (NAS) boxes, and storage sub-systems in Enterprise servers drives by RAID-Controller PCI-express add-in-cards.
Our evaluation was based on two common 4-bay desktop-NAS models – the QNAP 4-bay Desktop NAS TS-464-4G and the Synology 4-bay Desktop NAS DS-420+.

Picture 3:
QNAP 4-bay Desktop NAS TS-464-4G

Picture 4:
Synology 4-bay Desktop NAS DS-420+

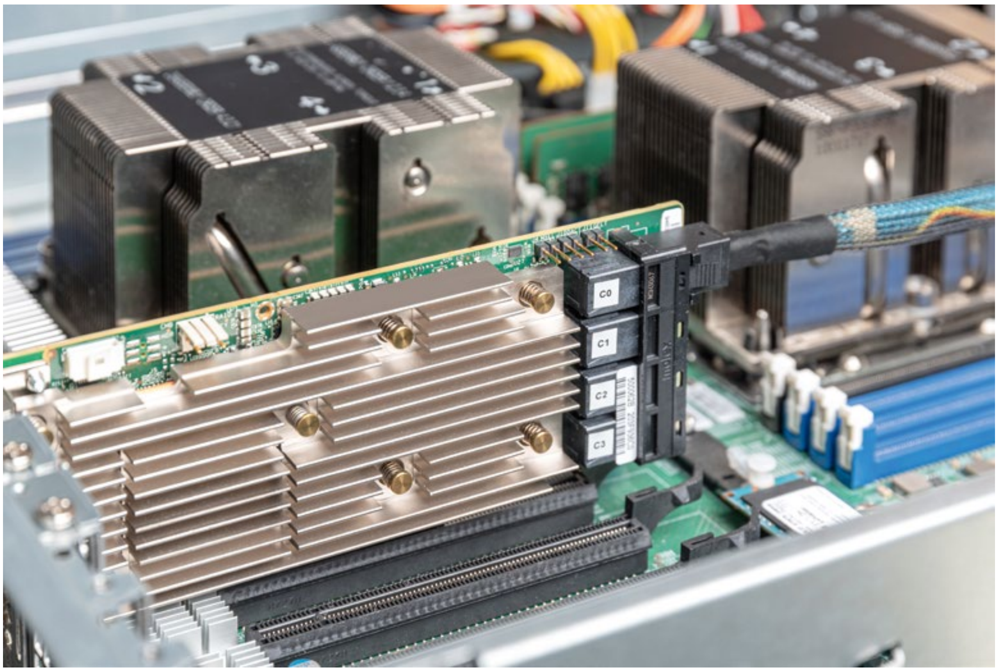
As for server RAID subsystems, we assessed several popular RAID controllers, with the HDDs installed in the front bay of the server chassis. The first was the Microchip Adaptec® SmartRAID 3204.



Picture 5: Microchip Adaptec® SmartRAID 3204 in server

As this RAID controller is optimized for a smaller number of drives, we also evaluated an alternative multi-purpose controller from the previous series: the Microchip Adaptec® SmartRAID 3154.

In addition, we assessed another popular Raid Controller from Broadcom: MegaRAID SAS 9460.



Picture 6: Broadcom MegaRAID SAS 9460 in server

However, as this model is not sold anymore, we also tested the successor model MegaRAID SAS 9560 – with PCI-Express Gen4 Host Interface. The results (see Table 1) are nearly identical.

Results for 512e Replacements in 512n RAID Sets

System	Model	Initial Array	Replace Drive	Function	Rebuild Time	Post-processing
Broadcom	9460	512n	512e	YES	6h 5min	None
Broadcom	9560	512n	512e	YES	5h 45min	None
Adaptec	3154	512n	512e	YES	4h 15min	8h BPI
Adaptec	3204	512n	512e	YES	4h 5min	8h BPI
Synology	DS-420+	512n	512e	YES (1/2)	4h 0min	6h Scrub
QNAP	TS-464-4G	512n	512e	YES	6h 0min	None

(1) automatic rebuild does not start, need manual interaction
(2) System issues a warning that MG08ADA is not in the compatibility list, but rebuild still works

Table 1: 512e replacements in 512n RAID sets

Conclusion: for all configurations, failing 512n drives can be replaced by 512e models without restrictions.

Observations on Rebuild Times:

The results illustrate the different re-build strategies of the RAID engines. For Broadcom and QNAP, the entire drive is re-built: regardless of whether it's filled with data or not. After re-build has finished, no further consistency checking tasks are required.

Adaptec and Synology following a different approach. They apparently rebuild only the actual data on the drive, hence pure rebuild times can obviously be shorter in partially filled arrays. A post-processing task – background parity initialization (BPI) or scrubbing – follows the RAID rebuilding. The RAID set is not vulnerable and degraded anymore, but the post-processing task requires some performance resource, running with low priority.

For reference, Table 2 shows a list of the condition and configuration details:

Details	
512e Drive	Toshiba MG08ADA400E 4TB FW 0102
512n Drive	Toshiba N300 4TB HDWG440 FW 0601
Broadcom	MegaRAID SAS 9460-16i FW 5.010.00-0671
Broadcom	MegaRAID SAS 9560-8i FW 5.120.02-2904
Adaptec	SmartRAID 3154-16i FW 1.60
Adaptec	SmartRAID 3204-8i FW 03.01.17.056
Synology	DS-420+
QNAP	TS-464-4G
RAID Conf.	RAID5, Cache On, Fully initialized, partially filled (6TB Data)
Rebuild Time	Without Workload

Table 2: Conditions and Configurations

We used RAID5 (parity RAID) since this is a more complex configuration than alternative RAIDs based on simple mirroring (RAID1, RAID10). Rebuild times were evaluated on arrays partially filled with data, and without productive workload during rebuild. It should be noted that if workload during rebuild exists, rebuild can take significantly longer.

Results for 512n Replacements in 512e RAID Sets

We also confirmed that replacing a failing drive in a 512e RAID with a 512n drive works in all cases (Table 3). Together with the previous results, this confirms that mixing 512n and 512e is possible for all RAID configurations.

System	Model	Initial Array	Replace Drive	Function	Rebuild Time	Post-processing
Broadcom	9460	512e	512n	YES	6h 0min	None
Broadcom	9560	512e	512n	YES	6h 0min	None
Adaptec	3154	512e	512n	YES	4h 40min	8h BPI
Adaptec	3204	512e	512n	YES	4h 10min	8h BPI
Synology	DS-420+	512e	512n	YES (1/2)	3h 30min	6h Scrub
QNAP	TS-464-4G	512e	512n	YES	6h 30min	None

Table 3: 512n replacements in 512e RAID sets

More Flexibility: Replacing Failed Drives by Larger Capacity and Other Manufacturers

We moved on to pose some further questions.

Where no replacement candidate of the initial capacity is available or at hand, will a drive of larger capacity do the job as well? The results are shown in Table 4 below.

System	Model	Initial Array	Replace Drive	Function	Rebuild Time	Post-processing
Broadcom	9460	512n	6TB	YES	6h 10min	None
Adaptec	3154	512e	6TB	YES	4h 0min	8h BPI
Synology	DS-420+	512n	6TB	YES (1/2)	4h 0min	6h Scrub
QNAP	TS-464-4G	512e	6TB	YES	6h 30min	None

Details	
Replace Drive	N300 6TB HDWG460 FW0601

Table 4: Larger capacity replacement drive

Conclusion: a failed drive can be replaced by a model with higher capacity, but the additional capacity of the replacement drives remains unused.

What if a replacement drive or the original manufacturer is not available? Will a similar drive from another HDD manufacturer work as replacement drive? See Table 5.

System	Model	Initial Array	Replace Drive	Function	Rebuild Time	Post-processing
Broadcom	9460	512n	6TB	YES	6h 10min	None
Adaptec	3154	512e	6TB	YES	4h 0min	8h BPI
Synology	DS-420+	512n	6TB	YES (1/2)	4h 0min	6h Scrub
QNAP	TS-464-4G	512e	6TB	YES	6h 30min	None

Table 5: Other manufacturer replacement drive

Conclusion: this works fine too. HDDs of the same interface (SATA), same block size (512 byte, regardless of emulated or native) and same- or larger capacity can be mixed, regardless of the manufacturer.

However, previous evaluations by Toshiba have shown that there are limits to mixing. Drives with 512 byte and 4k byte block size cannot be mixed or used as replacement drives, and neither can drives with different interface technology (SATA and SAS). Even though an SSD is accepted by some RAID controllers as a replacement for a failed HDD and vice-versa, the actual implementation of such a use case is more of academic interest than a practical solution. And obviously replacing a failed device with one of smaller capacity won't work either.

Other RAID Systems

RAID technology is not only used for network attached storage, but also for direct attached storage (DAS) which connects a RAID based storage box to the host via dedicated USB- or Firewire connections. To confirm the replacement flexibility for DAS systems we tested a 4-bay RAID from ICY-Box, and a SafeTANK GR5640-SBA31+ from RAIDON.

Picture 7:
4-bay RAID IB-RD3640SU3 from ICY-Box

Picture 8:
4-bay RAID GR5640-SBA31+ from RAIDON

Full replacement compatibility was also confirmed for these USB-connected RAID boxes, as Table 6 shows:

System	Model	Initial Array	Replace Drive	Function	Rebuild Time	Post-processing
ICY Box	IB-RD3640SU3	512e	512n	YES	6h 50min	None
		512n	512e	YES	6h 30min	None
RAIDON	GR5640-SBA31+	512e	512n	YES	6h 45min	None
		512n	512e	YES	6h 30min	None

Table 6: Replacement flexibility for USB-connected direct attached RAID box

Toshiba is aware that end users (datacenters, large-scale enterprises, small/medium businesses and private users) may use older legacy systems for rather a long time, even beyond the original manufacturer's warranty and support timeframes. In terms of RAID controllers in particular, there is still a significant installation base of historic 3ware models. Toshiba supports testing of new HDD models on old controllers, so we tested the replacement of failing drives on older RAID sets of 512n drives with 512e replacement drives as well as in a RAID system with a more than 10-year-old 3ware 9750-4i controller (see Table 7).

System	Model	Initial Array	Replace Drive	Function	Rebuild Time	Post-processing
3ware	9750-8i	512n	512e	YES	6h 5min	None

Table 7: Drive replacement at legacy controller

It's interesting to see that the RAID rebuild time of this old controller is in the same range as the other RAID systems. This indicates that the RAID rebuild performance rather depends on the speed of the drive, rather than on the performance of the RAID controller.

Summary and Conclusion

Whether for RAID based storage, server thru RAID controller cards, or NAS and USB-connected RAID boxes, failing drives of 512 byte block size with SATA can be replaced by any other SATA HDD model of same or larger capacity, regardless of 512 native or 512 emulated block size. 512 native and 512 emulated drives can be mixed in any configuration.

While this is good news, there are some important points worth noting.

No back up, no mercy. It can`t be stressed enough that RAID technology is not a backup, it simply helps you to continue working in case of a storage media failure. If the RAID controller/box fails or other storage bugs occur, all your data will be lost. So please also backup the data on your RAID/NAS, ideally to a portable component like the Toshiba Canvio 2.5" external USB-connected HDD, and store it separately in an external location.

Watch your RAID. RAID technology allows you to continue working automatically in case of a HDD/HDDs failing, so you will not notice a failure by just observing function and performance. That's why RAID boxes and NAS systems have indication LEDs. Thanks to their short point-to-point connectivity, RAID boxes are usually placed close to the host computer, so they can be easily observed. But NAS boxes with network connectivity can be located anywhere within the network, often in cabinets, racks or cellars, so a red-blinking indicator LED is easily overlooked. You should therefore either login to the control dashboard of your system on a regular basis and check the status, or – even better – configure the system so that an error message is sent out to you e.g. via email.

Be prepared for disk failure. In case of a disk failing, please acquire a replacement disk in advance. Ideally it should be the same model as the original RAID set; but as this whitepaper demonstrates, there is a lot of flexibility. Just make sure it's a SATA model, it has same or higher capacity, and it's not a 4kn native block size.

Keep rebuild times short. Once the replacement drive is available, swap the failing drive and check the system to see whether automatic rebuild has started. If not, start it manually in the respective menus. You can continue working with the data, but ideally leave the RAID system idle. This shortens the rebuild time. Remember: the data is rather vulnerable during re-build as there is no redundancy anymore. Another drive failure while re-building would kill your data, so keeping rebuild times short is highly advised.

No second chances. Never ever re-insert a failed drive back into the RAID system. If it failed once, it will fail again, even if at re-insertion it may seem to work fine. And more important still, many RAID systems get confused if they recognize a drive that has been part of a RAID configuration before by finding known metadata on the drive. They may try to re-insert it into the RAID set, or assume it's part of a different RAID. And that can really mess things up.

Go configure. Check your requirements. Whether to use your four drives in RAID5 (single parity) configuration or RAID10 (striped and mirrored), or whether to buy just two drives of double the capacity and run in RAID1 (simple mirror) mainly depends on capacity and speed requirements.

More evaluation is on the way. Toshiba is working on a whitepaper with more data, background and constraints, so look out for this in the near future. Meantime, we have established that, for a typical working RAID with active data on it connected via USB3 or Gigabit Ethernet to the host, a RAID5 of 4 HDD configuration provides maximum usable data within the performance range limited by the network connectivity.

References

1. Enterprise HDD Legacy Systems: https://toshiba.semicon-storage.com/content/dam/toshiba-ss-v3/emea/en_gb/storage/technologycentre/2018-05-toshiba-whitepaper-enterprise-hdd-legacy-systems-en-screen.pdf
2. How New HDDs Interact with Old RAID Controllers: https://toshiba.semicon-storage.com/content/dam/toshiba-ss-v3/master/en/semiconductor/design-development/innovationcentre/whitepapers/How_new_hard_disk_drives_interact_with_old_RAID_controllers.pdf

Toshiba Electronics Europe GmbH

Hansaallee 181
40549 Düsseldorf
Germany

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