

# MACH•2™

**Q: How does Exos® 2X provide up to 2x the performance of a standard single-actuator hard drive?**

**A:** Exos 2X can demonstrate up to 2x the performance of a standard single actuator hard drive because it has two independent actuators and data paths, allowing for concurrent I/O streams to and from the host.

**Q: How does an Exos® 2X SATA configuration differ from a SAS configuration?**

**A:** For the SAS configuration, each actuator is assigned to a logical unit number (LUN 0 and LUN 1). For example, one 18TB SAS drive will present itself to the operating system as two 9TB devices that the operating system can address independently, as it would with any other HDD.

The Exos® 2X SATA configuration will present itself to the operating system as one logical device since SATA does not support the concept of LUNs. The user must be aware that the first 50% of the logical block addresses (LBAs) on the device correspond to one actuator and the second 50% of the LBAs correspond to the other actuator. With both configurations, the user must send commands to both actuators concurrently to see the expected performance benefits.

**Q: How do you identify which LBAs correspond to each actuator on an Exos® 2X SATA device?**

**A:** Seagate® has worked with the T13 ATA committee to propose and implement a new log page for SATA—the Concurrent Positioning Ranges log page 47h identifies the number of LBA ranges (in this case, actuators) within a device. For each LBA range the log page specifies the lowest LBA and the number of LBAs. As a reminder, since LBA numbering starts at zero, the last LBA of either range will be the lowest LBA + the number of LBAs – 1.

**Q: How do you identify these LBA ranges in Linux?**

**A:** In Linux Kernel 5.19, the independent ranges can be found in `/sys/block/<device>/queue/independent_access_range`. There is one sub-directory per actuator, starting with the primary at “0.” The “nr\_sectors” field reports how many sectors are managed by this actuator, and “sector” is the offset for the first sector. Sectors then run contiguously to the start of the next actuator’s range.

**Q: What workloads show the best performance benefits over a single actuator?**

**A:** Exos 2X was designed for hyperscale workloads that focus on low queue-depth random read operations (low queue-depths keep command latencies low) and large transfer size sequential operations. The highest performance gains over a single actuator will be found during high transfer size sequential reads/writes (128KB transfers and larger), random reads (all transfer sizes), and random writes (128KB transfers and larger).

**Q: What performance benchmarks have shown a good benefit over a single actuator?**

**A:** Microsoft Jetstress, sequential throughput testing using fio (simulating backup operations or streaming), random read IOPS testing using fio (simulating hyperscale and CDN workloads), and more benchmarks are under investigation.

**Q: How do you create a script in fio to test a dual-actuator SATA?**

**A:** Within a fio job file, you can specify two threads—one for the first actuator (using size=50%) and one for the second actuator (using offset=50%). An example of a simple random read script to run 4KB transfers with a queue depth of 1 is provided here for Linux and Windows:

**Linux fio: Random Read 4KB Q1**

```
[global]
name=Random_Read_4KBQ1
direct=1
thread=1
ioengine=libaio
ioscheduler=noop
rw=randread
bs=4k
runtime=60
filename=/dev/sda
group_reporting

[Exos2x18_SATA_Actuator_0]
size=50%
iodepth=1

[Exos2x18_SATA_Actuator_1]
offset=50%
iodepth=1
```

**Windows fio: Random Read 4KB Q1**

```
[global]
name=Random_Read_4KBQ1
direct=1
thread=1
ioengine=windowsaio
rw=randread
bs=4k
runtime=60
filename=\\.\PhysicalDrive1
group_reporting

[Exos2x18_SATA_Actuator_0]
size=50%
iodepth=1

[Exos2x18_SATA_Actuator_1]
offset=50%
iodepth=1
```

**Q: What types of applications/solutions have shown good benefit over a single actuator?**

**A:** Content Delivery Networks (CDN), video streaming, mail servers, backup/shuttle services, Ceph, Hadoop (benefits demonstrated at a small scale with HDFS and MapReduce), and cloud applications all offer benefits over a single actuator. More solutions are under investigation, and as drive capacities grow and IOPS per TB decrease, more and more storage applications will require a dual actuator.

**Q: How are customers testing dual-actuator drives?**

**A:** The best way to test this drive is to use it in production applications so that the true benefits are realized right from the start. Often, synthetic workloads/benchmarks can be confusing, and results can be deceiving.

**Q: How does Exos 2X compare to 10K or 15K enterprise drives in IOPS/TB?**

**A:** Exos 2X aims to solve IOPS/TB constraints in systems that are currently using high-capacity 3.5-inch nearline HDDs, whereas 10K/15K mission-critical HDDs have much smaller capacities (2.4TB/900GB top capacities, respectively) and are intended for servers. The 10K/15K solutions are in more direct competition with low-end SSDs, whereas 3.5-inch nearline HDDs are addressing the need for mass amounts of data storage with less-critical IOPS requirements—Exos 2X increases the IOPS/TB in these applications. Typically, 3.5-inch HDDs are used in tiered solutions with SSD or NVMe storage solutions to address critical performance needs.

**Q: How can I purchase Exos 2X drives to improve my IOPS/TB?**

**A:** Contact your Seagate sales representative for more information.

**Q: With new technologies like MACH.2™, what is the expected change in power consumption over a single-actuator hard drive?**

**A:** Power consumption will vary by use case, but the highest power mode on Exos 2X is sequential transfers, followed by random reads. Random reads are the most common access pattern for most of our customers' workloads. Compared to a single-actuator Exos hard drive, when testing random reads at a 4KB transfer size and a queue of 16 pending commands (4KQ16), Exos 2X power is less than 20% higher for up to a 2x gain in performance. Or to put it another way, when considering that you would need to have two single-actuator drives to achieve the random performance of a dual-actuator drive, the dual actuator is ~40% less power than two single-actuator drives.

**Q: What is the typical power capacity of slots in a data center? Is 13.5W somewhere near that number?**

**A:** Customer feedback has revealed a wide range of power capacity in data center slots, but 13.5W is a common max power target that many customers have noted. We look at the aggregate of all the drives to determine the total power you see at a system level. This power limitation is more typically on the JBOD level where many drives are populated in the same storage system. Many customers are not slot-limited by overall wattage, but rather by 5V power. Exos 2X has ~1.4A max current on 5V.

**Q: Are there any options to reduce the power consumption of a dual- actuator drive?**

**A:** The on-demand PowerBalance™ feature is available on Exos 2X, which will slightly detune seek performance and decrease sequential transfer bandwidth, lowering the overall power below 12W or 11W depending on the option chosen by the customer. Sequential read power is the highest power mode on a dual actuator but note that many systems may be bandwidth-limited by the HBA and system, meaning they are to maintain 500MB/s+ across multiple drives, naturally reducing sequential power consumption.

**Q: Why does Exos 2X SAS only use a single SAS port?**

**A:** The current electrical architecture of the drive only supports a single port as the secondary port is required for linking the SoCs.

**Q: Will dual-actuator HDDs ever support dual-port SAS for high-availability and/ or failover scenarios?**

**A:** Current electrical architecture cannot support dual-port SAS; this is a design limitation with the current generation of SoCs. Once the market and volume can support a unique SoC, this may be added.

**Q: Why is Exos 2X only 18TB when Exos X20 is already launched at 20TB?**


**A:** Exos 2X18 is a 9-disk HDD vs the 10-disk Exos X20. The areal density is very similar, but there is currently a one-disk penalty to incorporate the two actuators.

**Q: Can this drive work with software RAID or does it need compatible hardware RAID host bus adaptors (HBA)?**

**A:** Yes, this drive can work with software RAID, but considerations need to be made around setting up redundancy across actuators in the same drive so that you would not lose data in the rare event of a total drive failure. Currently there are no hardware RAID HBAs that fully support dual-actuator drives. Seagate is actively working with industry partners to ensure dual-actuator ecosystem support. Note that you may need to tune the software RAID stripe size to optimize performance based on your application.

**Q: What SAS HBAs have been tested with Exos 2X drives?**

**A:**

	HBA	9600, 9500, 9400 Series Host Bus Adapters	<ul style="list-style-type: none"> <li>Officially supported in P11+</li> <li>Driver and Controller Firmware available on Broadcom.com</li> </ul>
		9300 Series Host Bus Adapters	<ul style="list-style-type: none"> <li>Tested and confirmed compatible with multiple LUNs</li> </ul>
	RAID	9600, 9500, 9400, 9300 MegaRAID Controllers	<ul style="list-style-type: none"> <li>RAID not currently supported</li> <li>RAID-in-HBA mode to be supported in 9600 generation</li> </ul>
 <b>MICROCHIP</b>	HBA	1100, 2100 Series HBAs	<ul style="list-style-type: none"> <li>SAS supported on Linux in SR2.7+</li> <li>SAS Windows support in SR2.7+</li> <li>SATA support in SR2.7.4</li> </ul>
		1200, 2200 Series HBAs	<ul style="list-style-type: none"> <li>SAS supported on Linux in SR 3.2+</li> <li>SAS Windows support in SR 3.22</li> <li>SATA support in SR3.2.4</li> </ul>
	RAID	3200, 3100 Series RAID Controllers	<ul style="list-style-type: none"> <li>RAID-in-HBA-mode support added in SR2.7+ and SR3.2+ releases</li> <li>RAID not currently supported</li> </ul>
	HBA	ExpressSAS 12Gb GT HBAs ThunderLink SH 3128	<ul style="list-style-type: none"> <li>SAS supported</li> <li>SATA supported</li> </ul>
	RAID	ExpressSAS 12Gb GT HBAs ThunderLink SH 3128	<ul style="list-style-type: none"> <li>RAID not currently supported</li> </ul>

*\*Broadcom 9600-16i support pending end-2023 – timeline subject to vendor release schedule (support enabled with special FW and Scrutiny tool – please contact Broadcom for details)*

**Q: What is Seagate doing to ensure dual-actuator ecosystem support?**

**A:** We are actively working with HBA vendors to expand the number of HBA/expander products that are compatible out of the box with Exos 2X and future dual-actuator products. Additional socialization of dual-actuator technology is happening at industry events like CES, OCP, Plugfest, and more.

**Q: Can you highlight the vibration generated and received by using 2 actuators?**

**A:** The system has been tuned so that one actuator can seek with minimal impact on the other. There are no changes to the external drive mechanical specs, such as rotational vibration, shock, acoustic disturbance, etc.

**Q: If one LUN has failed, will the drive still report Test Unit Ready when sent that command?**

**A:** A Test Unit Ready command requires both LUNs to be ready before replying with GOOD status. Note that a LUN failure is very rare, rather, head/media errors would be reported at a LUN level as happens on single-actuator drives.

**Q: Are there two copies of the defect management list?**

**A:** It is a combined defect list, and it is accessible on both LUNs for the SAS model.

**Q: Is the development of a single volume drive expected in the future (i.e., the drive itself will load balance/optimize by striping workloads evenly across both actuators)?**

**A:** Maybe. It is possible, but there are tail-latency issues with a configuration like this, yet it is the simplest way to get plug-and-play performance. There needs to be enough market to justify the firmware development/complexity.

**Q: SAS/SCSI command implementation like FORMAT UNIT, which affects both LUNs of the drive, seems like it may cause problems with RAID5, for example RAID5.**

**A:** That is true. There are additional RAID complications that could be caused by building redundancy across LUNs without assuring redundancy on additional HDDs.

**Q: How can I configure an Exos® 2X SATA drive in my Linux system?**

**A:** You can partition both actuators, stripe the actuators into a software RAID, or use as-is. Using the drive as-is would be a sufficient solution if you are migrating data to fill (or almost fill) the whole drive so that both actuators will be kept sufficiently busy. If you would like to treat each actuator as an individual device, then simple partitioning is an easy way to utilize Exos 2X SATA.

If you would rather gain performance benefits without partitioning, you can create a LVM striped partition which will dispatch IO evenly to both actuators. Depending on the application, the stripe size may need tuning to optimize performance. Refer to the example scripts below for partitioning and LVM striping.

### Linux: Partitioning Each Actuator of an Exos 2X SATA

```
> parted /dev/sda mklabel gpt
> parted --align optimal /dev/sda mkpart primary 0% 50%
#replace "optimal" with option "min" to save space
> parted --align optimal /dev/sda mkpart primary 50% 100%
```

### Linux: Create a LVM Striped Partition Using Both Actuators of an Exos 2X SATA

```
> parted /dev/sda mklabel gpt
> parted --align min /dev/sda mkpart primary 0% 50%
> parted --align min /dev/sda mkpart primary 50% 100%
> pvcreate /dev/sda1 /dev/sda2 #create two pvs
> vgcreate demo_test /dev/sda1 /dev/sda2 #create a vg named demo_test
> lvcreate -n lv_demo --stripes 2 demo_test -L `vgdisplay --units S
demo_test | grep "VG Size" | awk '{print $3}' 'S
#create a LVM partition lv_demolocated at /dev/demo_test/lv_demo
```

**Q: What is the recommended redundancy architecture with Exos 2X drives?**

**A:** Please review the following presentation from the 2019 Storage Network Industry Association (SNIA) conference for some recommendations: <https://www.snia.org/educational-library/multi-actuator-hdds-performance-and-integration-considerations-2019>.

Although this presentation is focused on SAS, the same considerations would apply to SATA at an actuator level.

**Q: Is there Linux IO Scheduler support for SATA multi-actuator drives?**

**A:** The prototype BFQ IO scheduler is more advantageous when workloads are imbalanced because it guarantees four out of SATA's 32 hardware queues to each actuator. This reduces the possibility of actuator starvation, even with workloads that are unbalanced across the LBA range.

Testing with FIO has demonstrated that with eight jobs on one actuator as opposed to one job on the other, performance improvements are maintained. For more on the BFQ prototype see: <https://www.snia.org/educational-library/bfq-linux-io-scheduler-optimizations-multi-actuator-sata-hard-drives-2021>.

The BFQ patch including these optimizations has been submitted for inclusion in Linux 6.0 with upstreaming targeted for Linux 6.1 or 6.2. Please download, test and report any issues via the Linux forum. To find the latest updates search for "bfq: extend bfq to support multi-actuator drives" here: <https://www.spinics.net/lists/kernel>

For more information on Budget Fair Queueing (BFQ) Linux IO scheduler optimizations for multi-actuator SATA hard drives see: <https://www.linaro.org/blog/budget-fair-queueing-bfq-linux-io-scheduler-optimizations-for-multi-actuator-sata-hard-drives/>.



[seagate.com](https://www.seagate.com)

© 2023 Seagate Technology LLC. All rights reserved. Seagate, Seagate Technology, and the Spiral logo are registered trademarks of Seagate Technology LLC in the United States and/or other countries. Exos, FastFormat, MACH.2, the MACH.2 logo, and PowerBalance are either trademarks or registered trademarks of Seagate Technology LLC or one of its affiliated companies in the United States and/or other countries. All other trademarks or registered trademarks are the property of their respective owners. When referring to drive capacity, one gigabyte, or GB, equals one billion bytes and one terabyte, or TB, equals one trillion bytes. Your computer's operating system may use a different standard of measurement and report a lower capacity. In addition, some of the listed capacity is used for formatting and other functions, and thus will not be available for data storage. Actual data rates may vary depending on operating environment and other factors, such as chosen interface and disk capacity. Seagate reserves the right to change, without notice, product offerings or specifications. SC702.2-2306US