Agenda

- Case for Multi-Channel / Multi-Actuator
- Standardization Progress
- Deploying SATA and Single-LUN SAS
- Call to action
Case for Multi-Actuator
Case for Broad Multi-Actuator Adoption

- IOPS/TB are falling
- Latency driven workloads cannot meet minimum workload QOS
- To improve read latency or IOPS without multi-actuator:
  - Short-stroke the HDD
  - Deploy lower capacity drives
  - Use caching and SSD tiering
  - Change queue depths
  - Use firmware prioritization tools
- Tradeoffs and diminishing returns to each

![Graph showing IOPS/TB vs Drive Capacity TB for Single and Dual Actuator cases.](image-url)
HDDs to Fulfill a Majority of Datasphere Storage Exabyte Shipments

HDDs support growing mass capacity demand

High Performance Storage
Compute Centric / Performance Optimized

Mass Capacity Storage
Data Centric / Capacity Optimized

Source: IDC (2018), Seagate
Flash includes non-volatile memory NAND and SSD
Multi-Actuator Today

- Dual-Actuator Dual LUN SAS
  - Delivers performance today
  - Breaks single LUN to single disk paradigm
  - Requires constructs incl. device groups, failure domains, etc.
  - Implementing RAID can be a challenge
- SATA and Single LUN SAS are required as adoption broadens
  - Move beyond select innovation deployments
  - Suits most existing storage stack paradigms
Standardized Functionality

- Available in ACS-5 (SATA)
  - Report number of actuators
    - Supports up to 29 actuators
  - Report an LBA range for each actuator
  - Plenty of room to and more actuator info

- Coming to SAS
  - Same information found in ACS-5
  - Designed to be compatible with SATA
Standardization In Process (Actuator RAID)

- Allows a single LBA space to have actuators interleaved
- Capability is host configurable
- Devices may limit the stripe size to fit device technology requirements
- May provide a drop-in performance boost for SATA and Single LUN SAS without host modification
Deploying Single LUN/SATA Multi-actuator Drives
Use Case: separate a split-address-space device into two independent devices

- Partitioning the device at the actuator boundary creates two subordinate block devices managed by the kernel
- Persistent: Kernel recreates the devices by default when it reads the partition table at device initialization
  - The backup partition table will be on the secondary actuator

```bash
#!/bin/sh
echo create two partitions on the supplied disk, dividing in half
dev=$1
device_size=`blockdev --getsize /dev/${dev}`
act1_sect0=$((device_size/2))
echo device size 512b blks: $device_size
echo act 1 first sector: $act1_sect0
echo use the following command to partition /dev/$dev
echo sudo parted /dev/$dev -s mklabel gpt mkpart act0 0% $(( $act1_sect0-1 ))s mkpart act1 $(act1_sect0)s 100%
```
Device Mapper – Separate the Actuators

- The linear device mapper target will split the underlying block device at the actuator split point
- The bare device mapper does not store any metadata anywhere
- Not persistent: The setup must be re-done by udev or a startup script – it’s not automatic
- DM does not generate any RLA (read lookahead) on its own but will pass system-generated RLA down to the devices, speculatively linking the actuators

Sample setup script:
```
dev=$1
echo 0 $((`blockdev --getsize $dev `/2)) linear $dev 0 > /tmp/actuator1
echo 0 $((`blockdev --getsize $dev `/2)) linear $dev $((`blockdev --getsize $dev `/2)) > /tmp/actuator2
dmsetup create actuator1 /tmp/actuator1
dmsetup create actuator2 /tmp/actuator2
dmsetup ls
```
LVM2 – Partitioned Disk

LVM2 is persistent across reboots.

- striped logical volume maps the data space to stripes on the actuators
- create physical volumes from partitions
- align disk partitions to split
- LVM metadata
- GPT

partition 1

0

maxlba / 2

physical vol 1

physical vol 2

Partition 1

Partition 2

/ dev/ sdb

Linear logical volumes allocated from the pv’s

volume group encloses a drive’s physical volumes.

Linear logical volumes

system metadata

single target split address space

maxlba
Call to Action
Call to Action

- Send feedback!
- Evaluate IOPS and latency needs
- Work with partners on Dual Actuator enablement
- Drive in open source and standards
Please take a moment to rate this session.

Your feedback matters to us.