HDD Parallelism for Lower TCO
Dual Actuator Implementation

James D Borden & Timothy T Walker
Seagate Technology
24 Sept 2018
IOPs Not Growing with Capacity

HDD IOPS are governed by mechanical physics. Access times and data rates lag capacity increases, so IOPS/TiB trend down.

![Diagram showing the trend of IOPS not growing with capacity from 2016 to 2025, with two technologies: Perpendicular Magnetic Rec and Energy Assisted Magnetic Rec. The typical customer requirement range is shown as well.](image-url)

Typical customer requirement range

**2016**  **2017**  **2018**  **2019**  **2020**  **2021**  **2022**  **2023**  **2024**  **2025**
Customer Perspective: IOPs/TB 2014-2021

**Single Channel**

- Capacity (TiB)
- IOPS/TiB SLA (min HDD SLA after all caching/tiering)
- IOPS/TiB

**Multi Channel**

- Capacity (TiB)
- IOPS/TiB SLA
- IOPS/TiB

**Graph Notes:**
- Gap (min HDD SLA after all caching/tiering)
- Margin
Multi-Channel HDD Inside Look

- Separate pre-amps and connections to electronics
- Separate voice coil motors for independent seeking
- Servo algorithm cross-coupling mitigation for seek-while-track-following
- Common actuator pivot
- Each actuator maps a fixed set of heads
- No media or LBAs shared across actuators
Dual Actuator SAS Architecture

- SAS block devices do not have a separate admin endpoint, other than SES.
- One World Wide Name (WWN) is assigned to the port for device address. Each LUN has a separate unique WWN in VPD.
- Two LUNs, 0 & 1, each with independent LBA space, and each connected to an independent read channel, actuator, and set of heads.
- Each LUN maps 50% of the media – no media is shared across the actuators.
- Some commands affect both LUNs
  - TUR, Log/Mode Sense/Select,
  - FORMAT UNIT, START STOP, REMOVE I_T NEXUS
  - Reman (REMOVE ELEMENT AND TRUNCATE)
  - Sanitize
Dual Actuator PCIe-NVMe Architecture

- NVMe has clean subordinate storage support via namespaces.
- One NameSpace per actuator.
- Dual port with bifurcation; each NS shared by two ports.
- Each actuator is mapped to a unique Endurance Group/NVM set.
- All namespaces are statically created at time of manufacture, cannot be created/deleted.
RAIDing Dual Actuator HDDs

- RAID impact is from IOPS/TiB.
  - Really, both actuators in same failure domain.

Multi-LUN in single RAID:

Two single-LUN RAIDs:
Conclusion

- HDDs are very cost-competitive per TB, but not per IO.
- Current caching and tiering approaches require 8-10 IOPS/TiB; common pain-point around 14 TiB.
- Doubling HDD IOPS keeps low-TCO rotating-storage a strategic part of performant SLA storage solutions.
- Multi-actuator HDDs 90% drop in; minor engineering effort for full utilization.