FUTURE HDD INTERFACES

JASON ADRIAN
STORAGE HARDWARE ARCHITECT
MICROSOFT AZURE
SAS AND SATA

- SAS and SATA are mature interfaces and have served us well for well over a decade
- SATA is the lowest cost
- SAS has mature device topology, hot plug capabilities, and signal integrity
SO WHY CHANGE?

- SAS seems like an obvious winner, but:
  - Drives often cost slightly more
  - Requires SAS controller
  - Requires SAS expander for more than 8 or 16 (sometimes 24) direct attached HDDs
- Technologies like dual actuator are easier to manage with SAS and its native LUN semantics
- Dual actuator with SATA is more challenging, and will start pushing against the 6gb PHY limitations in future generations
- Do we really need SAS controllers and expanders? Why not use PCIe like we do for NVMe SSDs?
NVME HDDS

- If HDDs move to NVMe, SW infrastructure can manage SSD and HDD platforms in a similar manner
- Can eliminate SAS controller ASICs
- Can directly connect to host CPU for low to moderate HDD counts
- Utilize PCIe switches instead of SAS expanders
  - This is a trade of protocol switches, not an improvement
- NVMe name spaces are conducive to technologies like dual actuator
NVME HDDS – HOW TO MAKE THIS SUCCESSFUL?

- NVMe HDDs need only a single PCIe lane
- Could support 2 x1 lanes for those requiring dual porting
- PCIe Gen 3 would have a reasonable trace length for system designers. Gen 4/5 likely not required for many years.
- Host CPU complex’s need to support x1 PCIe bifurcation to enable a larger direct HDD attach count
- Need to prove out surprise removal and PCIe enumeration for HDDs. This should be able to leverage much of the work from NVMe HDD design in efforts