Big Data Flash

Roark Hilomen presented by Jim Fitzpatrick

Eng Fellow – Data Center Solutions

9/14/2017

Agenda

1. Introduction
2. Primer to WDC (HGST/SanDisk)
   A. HDD Technology
   B. Flash Technology
3. Competitive Media Comparison
4. A System Level View
5. The Future (3+ years)
6. Summary

About Western Digital Corporation

Western Digital®

SanDisk®
a Western Digital brand

HGST
a Western Digital brand

WD
a Western Digital brand

G-Technology®

TRUST

- A data storage leader
- "Under the hood" of major OEMs
- Millions of devices shipped annually

VALUE

- Among the highest SSD/HDD MTBF rates
- Certified solutions (Msft, Oracle, VMware)
- Single source for storage devices & systems

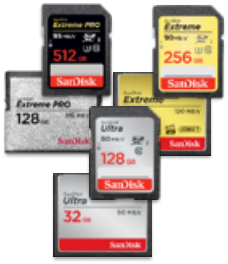
INNOVATION

- 12,000+ active patents, 500 new/yr
- $2.5 billion invested annually in R&D
- Leader in Open Community initiatives

Broad Storage Portfolio

Consumer

Imaging



WD Software

microSD™ Cards & Mobile Storage

Solid State Drives (internal and external)

Wearable Audio

USB Drives

Mobile & Industrial

Embedded & Removable

Surveillance

WD Purple

Client Computing

Laptop, Desktop & Commercial SSDs

WD Black WD Blue

Solid State Storage

FlashMAX® III PCIe

Ultrastar® SAS SSD

Ultrastar® SN100 NVMe PCIe

HDD Storage

Ultrastar® He[8] and He[10]

Ultrastar® Archive Ha[10]

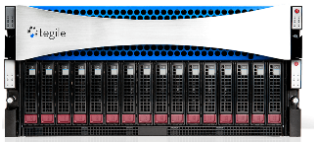
Data Center/Enterprise

WD Gold
(Data center)

WD Red

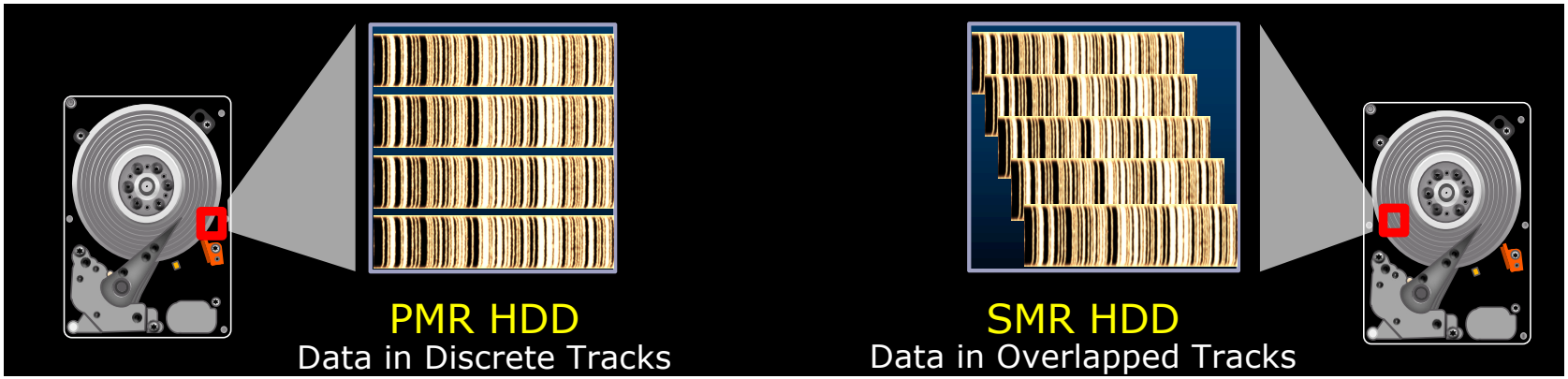
Active Archive System

Tegile



FlashSoft® Software

Primer to WDC Media Offering (HDD)

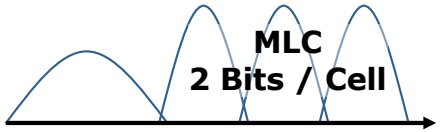


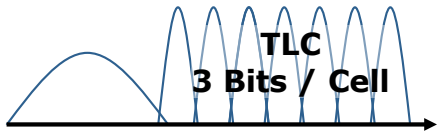
Capacity targets next 3 years

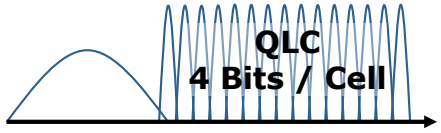
12TB → 14TB → 20TB+

Primer to WDC Media Offering (Flash)

Bits per Cell

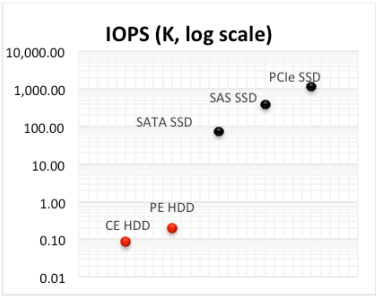


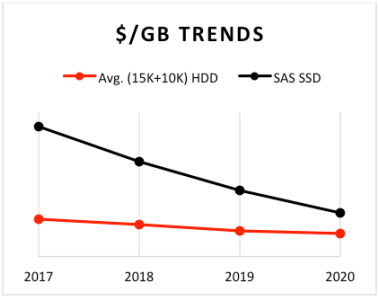




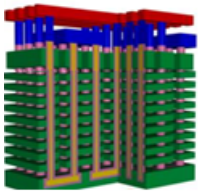
Rohit Gupta: SanDisk Blogger

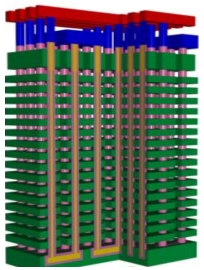
https://itblog.sandisk.com/hdds-vs-ssds-evolution-storage/





Layers





Planer (1 Layer) | 32 Layer (Samsung) | 48 Layer | 64 Layer (Today) | 96 Layer (Next Leap) | 128 Layer | ?

Primer to WDC Media Offering (NVMe 96 Layer, QLC)

Legal: The following are technology projections and not a specific product offerings

Key Characteristics

Capacity: 100TB per 2.5”/15mm
Interface: PCIE/NVMe
Read/Write IOPS: 100K/25K (Large Seq Block)
WORM (Write Once): 15+years
Write Seldom (<60): 10+years
Throughput (R/W): 4GB/s up to 1GB/s
Power (goal) : 12w(Full write), 6-9w (mixed workload), less than 2w (Idle)
** Power Management can reduce avg power use significantly

Competitive Comparison (HDD vs SSD)

Spec	HDD (14TB)	LTO8	96L QLC	Ratio SSD
Capacity	14TB	12.8TB	100TB	7x / 7.8x
Uncompressed W	180MB/s	470MB/s	1GB/s	5x, 2x
Uncompressed R	200MB/s	470MB/s	4GB/s	27x, 10x
Power Operating	7w	*	9w (mixed)	
Power Idle	5w	0	2w**	

* Based on tape drive, controller and robotics

** Based on maintenance and potential Power Management

A System Level View

Rack Level Comparison (40U Usable rack space)

- HDD Rack – 4U Server w/ 60HDD
- Tape –Quantum Scaler i6000 extrapolated using LTO8
- QLC Rack - 1 2U Server + 3x 2U PCIE Storage Chassis (96 SSD in 8U)

Spec	HDD	Tape	QLC
Capacity No Compression	8.4PB	10.5PB	48PB
Throughput/W	108GB/s	*Based on Drives	480GB/s
Throughput/R	120GB/s	*Based on Drives	2.4TB/s
Throughput/M	12GB/s	*Based on Drives	*Based on Mix 400GB/s-2TB/s
Power Mixed	6KW	*	5KW
Power Idle	4.8KW	*	1.8KW

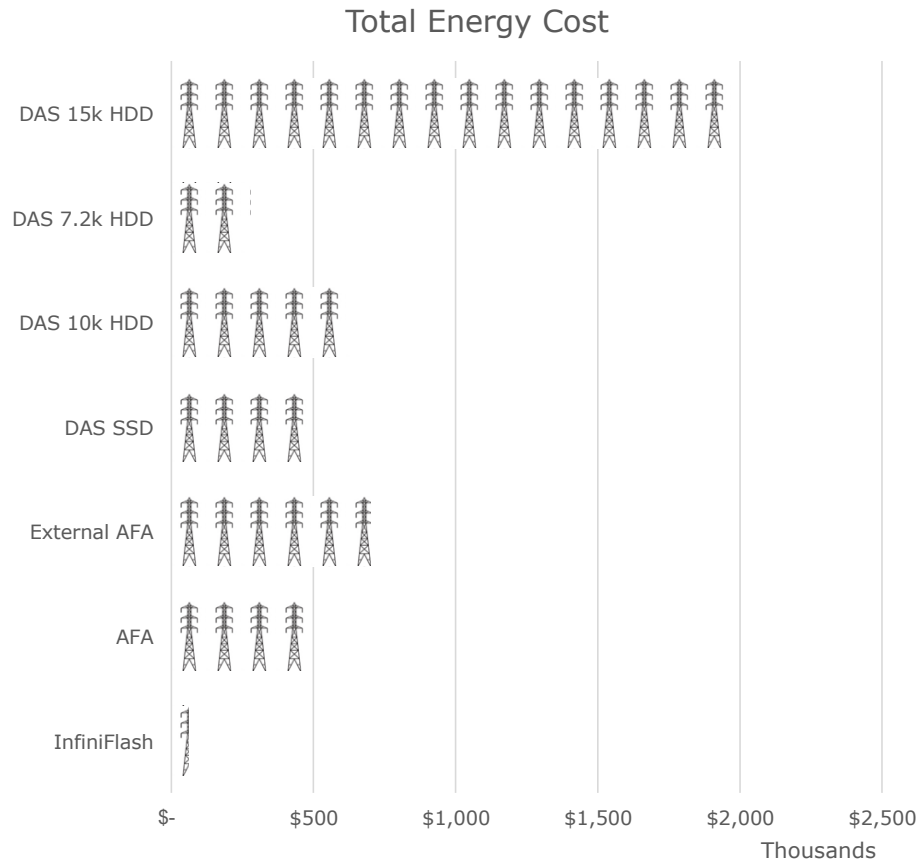
Network	Throughput (12 Uplinks)
10Gb	9.6GB/s
40Gb	36GB/s
100Gb	96GB/s
400Gb	384GB/s

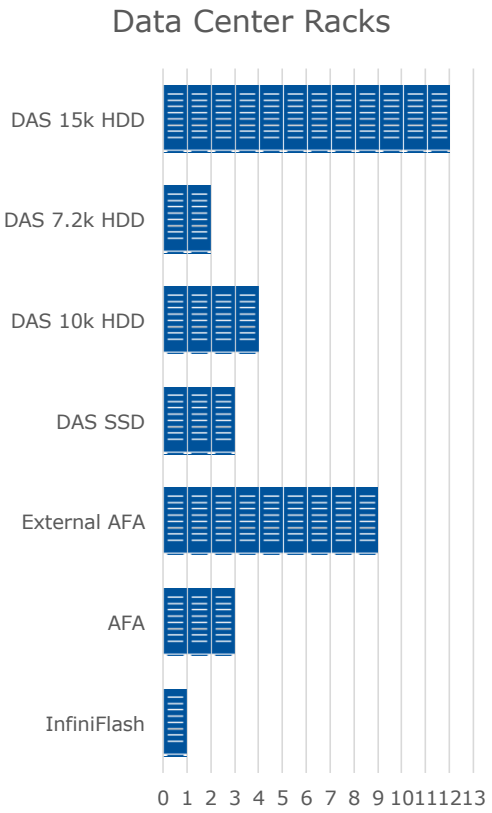
A System Level View

- **Limiting Factors**
  - Network Bandwidth
    - All media can scale on the compute side to exceed Network Bandwidth @10Gb
    - Transition to 100Gb and faster, Only SSD justify /can fully utilize
  - Software (not Optimized for Flash)
    - Object Based Erasure Coded/Compressed data to fully utilize and keep data
    - Resilient over 15+ years
- **Deployment Model**
  - As the density increases, rack level distribution becomes key to realizing throughput.
    - New Models distributing 'Archive' across active racks (Top of Rack Archive, Top of Rack Storage)
  - At 7x capacity, almost comparable power use (minus real deep archive/vaults)
    - QLC may expand where flash can be utilized
    - 10-15yr TCO on 48PB raw / Rack cost
- **Risk**
  - Limited NAND Fabrication Capacity
    - Exabyte level as we enter Zetabyte scale world
    - Choice of Fast Data or Big Data

TCO Example: 2015 Infiniflash 8TB Flash Cards

Food for thought





3.The Future (3+ years)

- Capacity Growth
  - SSD Vendors now have 3 knobs to turn for capacity
    - Layers (3d Nand)
    - Bits per Cell (1,2,3,4.. Yes more is possible, but with distinct limits)
    - **Lithography (65nm , 3xnm…)
  - Packaging
    - Infiniflash 1st in custom packaging..
    - EDSFF Ruler (U.2 capacity in 1U package)
    - 3.5” Flash drives?
    - Various ‘Flash Stick’ standards popping up (M.2,others)

- As value of data increases and cost of storage decrease – more data will be stored

3. Summary (Thoughts)

- Shifts are not made overnight. Flash + Tape, Flash + HDD + Tape will continue to exist. Ratio and length of retention/data activity will be key... Even a $0.01/GB difference is equivalent to $1M at 100 PB, moving to a Zetabyte, new architecture will be key
- As value of data increase and infrastructure cost decrease, trend is showing active data life is increasing from 1-6mo to 3-6 years and capacity of archive data growing
- Flash has limits on its fabrication capacity. Decision must be made on Fast Data Flash, Big Data Flash or Archive/Compliance Flash (QLC)

- **The dynamic is changing and it's up to the people in the room to decide how it will change**


Western
Digital®