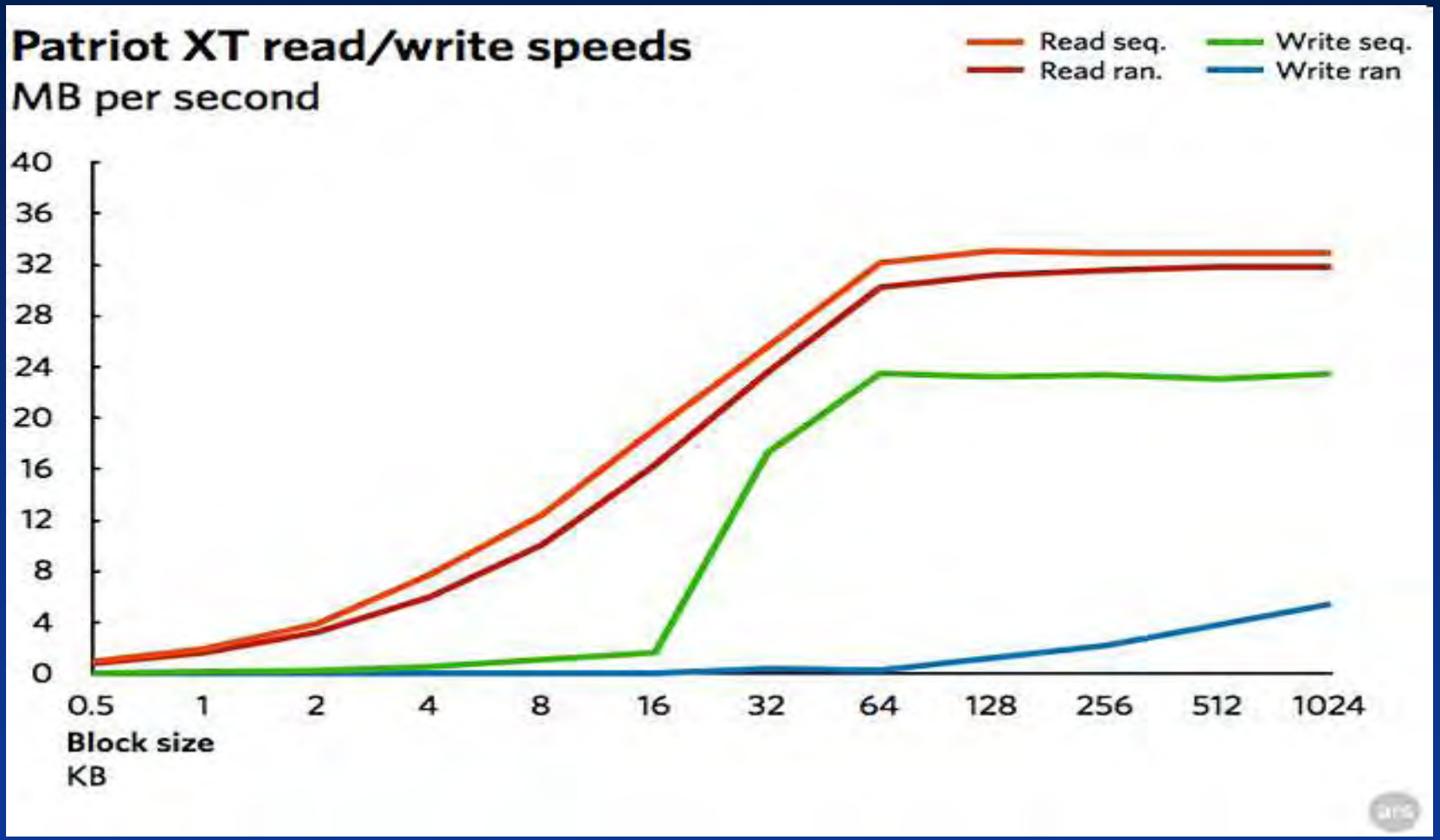
Optimizing Flash Storage With Linearization Software

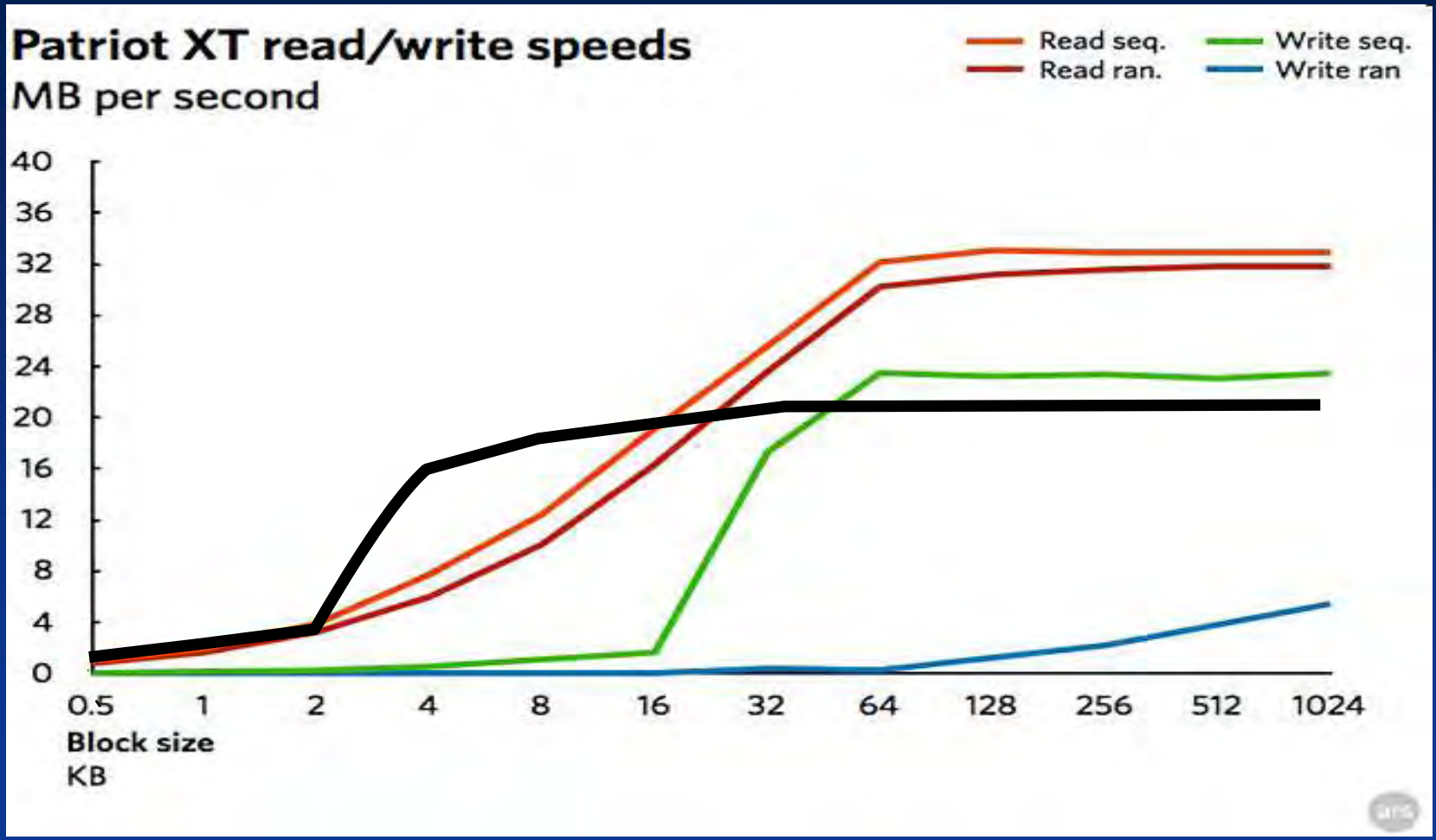
Doug Dumitru
CTO
EasyCO LLC

EasyCo

The performance curve of a thumb drive without linearization



A thumb drive's linear and random write performance with linearization



How Linearization Software Works.

- The block device is dynamically re-mapped using 4K pages.
  - 4K pages are an ideal compromise for most devices, applications, and file systems.
- Writes are always performed as long linear writes.
  - The lets the device achieve very high write data rates.
- Writes are not re-ordered
  - FIFO writes ensure file system integrity.
- RAM table keeps track of Logical Block Addresses.
- Data blocks and addresses are written together so that LBA table can be re-assembled during mount.

Managed Flash Technology

- We call this software “Managed Flash Technology”
  - MFT for short.
- The original name was “Fast Block Device”
  - But that seemed harder to market.
  - “FBD” still appears in driver names
    - fbd.sys in Windows
    - fbdusb.sys for USB sticks
    - dm-fbd.ko in Linux

Where MFT runs.

- On a Windows Host.
  - UpperFilter driver that manages a partition.
    - Works with single drives and arrays.
  - On demand driver
    - Works with portable storage like USB sticks.
  - In production release.
- On a Linux Host.
  - Device Mapper plug-in.
    - Compatible with volume manager and software raid.
    - Can load in initrd boot ramdisk.
  - In production release.

Where MFT runs.

- On a Storage Appliance.
  - Usually Linux based
  - Centralized SAN functions leverage SSD arrays
    - 2TB arrays in production use now
  - In production release
- On a Storage Controller.
  - Software available for license to raid card vendors.
- On a Drive Controller.
  - Software available for license to drive vendors.

How Performance is Improved

- Writes use 90+% of the drives available linear bandwidth.
  - Random and linear writes run at the same speed.
- Read performance is not impacted.
  - Linear regions stay linear.
- Many device / application mixes improve 10x or more.
  - Simple USB sticks do small file writes > 100x faster.
  - 1st Gen SSDs with 4K random writes > 20,000 IOPS.

How Endurance is Impacted

- Long writes actually use every Flash cell for real data, not for re-writes.
  - Simple Flash controllers can magnify wear 100x or more.
  - Linearization software reduces wear amplification to 3:1 or less.
- 2X MLC devices outlast SLC
- 3X/4X MLC become practical for SSD applications.

MFT Overhead

- DRAM Memory Overhead:
  - 1 MB per GB of array for LBA table
    - 1.25 GB per TB for large volumes (> 2TB).
  - 2x drive's write buffer size for outbound buffer
    - 128GB drive typically requires about 150MB of kernel RAM.
- Free Space Overhead:
  - Drive must have some dedicated free space.
    - Tunable for different applications.
      - As low as 5% free for workstation applications.
      - As high as 30% free for 24x7 server applications.
      - Typically 10%.

Applications

- Small to Large
  - Embedded devices
    - Smart Phones
    - Portable “computers”
  - Main Stream Computing
    - NetBooks
    - PCs
  - Commodity Servers
  - Enterprise Servers
  - SAN Servers

Real Performance Numbers

- Small Server
  - 5 x MTron 1025-32
    - 100 MB/sec read, 40 MB/sec write, MLC
  - Linux, Software raid-5
  - 59,843 IOPS read (4K random)
  - 30,882 IOPS write (4K random)

Real Performance Numbers

- Monitoring Wear
  - Software tracks writes
  - EasyCo's in-house mail server
    - Up 124 days
    - 307,123,019 reads
    - 354,977,975 writes
    - 17.715 GB/day linear writes
    - 1.58:1 wear amplification
    - 72,255 days predicted life (197 years)
  - 2 MTron 1025-64 drives mirrored to 2 SATA HDDs

Costs

- Less Expensive than 15K HDDs
  - HDD raid-10 8 drive array
    - 73 GB drives - $8.20/GB 146 GB drives - $4.93/GB
    - 2400 read IOPS - $1.00/IOPS
    - 1500 write IOPS - $1.60/IOPS
  - SSD raid-5 8 drive array
    - 64 GB drives - $3.94/GB 128/256 GB drives - $3.38/GB
    - 40,000 read IOPS - $0.039/IOPS
    - > 100,000 write IOPS - $0.016/IOPS

MFT is Available Now

- “Retail” Licenses
  - Windows
  - LinuxUSB Sticks.
- “Distribution”
  - Resellers
  - System Builders
- Software Licenses
- Technology Licenses

EasyCo LLC

220 Stanford Drive
Wallingford, PA 19086

+1 (610) 237-2000
(888) 473-7866 (US toll free)
+1 (610) 672-9549 (fax)

http://easyco.com
sales@easyco.com