Enterprise Flash Storage Annual Update

Or how the data center is replacing spinning rust with solid state

Howard Marks
Chief Scientist
Your not so Humble Speaker

• 25+ years of consulting & writing for trade press
• Columnist/blogger at NetworkComputing.com
• Chief Scientist DeepStorage, LLC.
  • Independent test lab and analyst firm
• Cohost Greybeards on Storage podcast
• @DeepStorageNet on Twitter
• Email:Hmarks@DeepStorage.Net
Agenda

- Flash moves mainstream
- Server side caching falters
- 3D/TLC enters the data center
- PCIe/NVMe rising
- Advances on the horizon
Flash has gone mainstream

(Volume)

- ~400PB AFA ship 2014
  - Flash based arrays $11.3 billion
    - 1.3 AFA, 10.0
- Enterprise SDD:
  - 2012 $3 billion
  - 2013 $4.4 billion
- ~80% of VNX/FAS ship w/flash
Flash Goes Mainstream
(Function)

- Single controller rack mount SSD – DEAD
- Even upstarts have full features
  - Snapshots, two replication methods
- AFAs scale to 100s PB
- Data reduction now table stakes for price
  - Deduplication and compression
And the market matures

- **Consolidation in components**
  - HGST (Virident, Stec, Velobit)
  - Sandisk (Smart, FlashSoft, Fusion-IO)
  - Seagate (LSI)

- **Flash systems shakeout**
  - Astute networks closes
  - HGST devours Skyerea
  - Cisco shuts Whiptail
And Everyone is the Market Leader

- EMC is #1 in dollar revenue (Gartner)
- IBM is #1 in PB shipped (Gartner)
- Netapp #1 in units shipped (Gartner)
- Pure #1 in growth (700%)
Evolution of Enterprise Flash

2010
- 100K+ IOPS
- Consistent sub-millisecond latency
- Go fast for special cases

2012
- Still a point solution
- Becoming cost effective
- Limited data services
- Data reduction

2015
- Flash is mainstream
- Full data services & data reduction
- Cost effective for many applications
The All Flash Data Center?

- All flash is inevitable
- Facebook…
- Murphy’s law
- Growing our TAM

- Flash cheaper than disk, really?
  - No enterprise SSD 25X cost/GB of 8TB disk
- Kryder’s law
2012

- Market leader Violin
  - No real data services
  - Just fast, fast, fast

5

- Even mainline vendors adding data reduction
- Data services now table stakes

Dedupe increases CPU requirements
- But has minimal impact on performance
Server Side Flash - 2015

- Platforms add limited caching
  - VMware VFRC
  - Storage Spaces SSD tier & write back cache
- vSphere adds IO Filters
  - Integration points in ESXi kernel
  - “Technology preview” in 6.0
Write Through and Write Back

- Baseline Write Through
- Write Back

TPC-C IOPS

- 100 GB cache
- Dataset 330GB grows to 450GB over 3 hour test
Distributed Cache

- Duplicate cached writes across n servers
- Eliminates imprisoned data
- Allows cache for servers w/o SSD

Solutions
- PernixData
- Dell Fluid Cache
  - RDMA based
  - Integrates with Compellent
Datrium DiESL

- Host managed cache
- PCIe SSD in Host
  - Write through cache
- All flash NetShelf
  - Persistent layer
- NFS interface to vSphere
  - Per-VM data services
- Founders from Data Domain
  - Dedupe of course
Hyperconverged Infrastructure (ServerSAN)

- Use server CPU and drive slots for storage
- Software pools SSD & HDD across multiple servers
- Data protection via n-way replication
- Can be sold as hardware or software
  - Software defined/driven
- All flash versions appearing
Sample ServerSAN Products

- VMware’s VSAN
  - Scales from 4-32 nodes
  - 1 SSD, 1 HDD required per node

- Maxta Storage Platform
  - Data optimization (compress, dedupe)
  - Metadata based snapshots

- EMC ScaleIO
  - Scales to 100s of nodes
  - Hypervisor agnostic

- Atlantis Computing ILIO USX
  - Uses RAM and/or Flash for acceleration
  - Works with shared or local storage
Enterprise SSD Evolution

- **Density - Today’s largest devices**
  - SAS - 4TB
  - SATA – 2TB
  - PCIe – 4.6TB
  - PCIe vendors discontinuing 200-600GB models

- **Interfaces**
  - U.2 PCIe from several vendors
  - NVMe from all enterprise vendors
  - Server support from most vendors
U.2/SFF-8639 PCIe for 2.5” SSDs

- Adds x4 PCIe 3.0 lanes to SAS/SATA connector
  - Dual ports to x2
- Appearing on new servers
  - Making PCIe/NVMe SSDs hot swappable
- Next step for storage arrays
Diablo Puts Flash on the Memory Bus

- Memory Channel Flash (SanDisk UltraDIMM)
  - Block storage or direct memory
  - Write latency as low as 3µsec
  - Requires BIOS support

- Memory1
  - 400GB/DIMM
  - No BIOS/OS Support
  - Volatile
Flash Goes 3D

- Smaller cells are denser, cheaper, crappier
  - Today’s 1x nm cells (15-19nm) last planar node

- 3D is the future

- 3D allows larger cells
  - Makes TLC useable
    - Faster write, higher endurance

- Samsung 3D-TLC SSD
  - Others foundries sampling
The Future is PCIe

- PCIe offers:
  - Low latency, high bandwidth, RDMA

- PCIe Switch chips
  - PLX and PMC – 96 lane

- Use for:
  - Controller to controller link
  - U.2 SSDs in storage system
  - Rack scale switched system (DSSD)
  - External PCI standards exist
The Future

- All PCIe storage systems
  - As conventional storage
  - With memory interfaces
- Next-gen memory (PCM, 3d Xpoint, Etc)
  - First as write cache in SSD (2017)
  - Later as memory
- More persistent memory as memory
  - Needs application support ala SAP Hana
THANK YOU

GRACIAS
ARIGATO
SHUKURIA
BOLZIN
MERCi