Tiered Storage and Caching Decisions

Bob “Grizzly” Griswold
Director of Industry Software Architecture, WD

September 17, 2012
How did I get this speaking slot?

- Western Digital is a Gold Sponsor of this Event
- The last guy, Mr. Mitchell, couldn’t make it
- I’m friends with some of these people
- I’m good at explaining things

You will hear about how I think about:

- Tiered Storage, SW and IO caching, Cost of your storage
- Why you cannot ignore the system, IO impacts it all
- Why better decisions on current and future needs pay-off

Sources available
What is Tiered Storage and Caching?

- Definitions and references found using your FSE
  - “… the assignment of different categories of data to different types of storage media in order to reduce total storage cost.”
  - “… is a data storage environment consisting of two or more kinds of storage delineated by differences in at least one of these four attributes: Price, Performance, Capacity and Function.”
  - “… is the leveraging of multiple storage media types across an enterprise that constitutes "tiered storage" and that comprises the 'hardware component' of an intelligent Information Lifecycle Management (ILM) strategy.”
  - “… is an underlying principle of ILM (information lifecycle management). It is a storage networking method where data is stored on various types of media based on performance, availability and recovery requirements.”
  - “… Intelligent caching lets you create a large pool of shared storage divided into virtual classes of service rather than create many hard-to-manage tiers.”
  - “… Tiered Storage combines flash-based storage with HDD storage to create a high-performance storage volume!”
  - “Automated tiering and caching often get confused.”

Sources available
Isn’t Tiered Storage and Caching the Same?

Source: John Squires – Founder and VP Engineering of Conner Peripherals

Source: http://archive.computerhistory.org/resources/access/text/Oral_History/102657961.05.01.acc.pdf
Source: http://en.wikipedia.org/wiki/Hard_disk_drive
Tiered Storage impacts Caching

- **Zones as defined by the “Industry”**
  - Tier 0 or Zone 0: SSDs, Flash, Flash-based NVMe
  - Tier 1 or Zone 1: SAS, SAN-based Fibre Channel, NAS
  - Tier 2 or Zone 2: SAN-based SATA, Cloud Storage
  - Tier 3 (Zombie storage?): Tape, Optical, rock-n-chisel

- **Zone 1.5 (Not Listed…)**
  - If Caching decisions and Tiered Zones go together than…
    - Application Client accelerations (SQL, Exchange, Oracle, DBs)
    - Host boot or paging and boot HDDs accelerators
    - Hybrid HDDs and OS / Applications that accelerate *all* host-based commands

- **Zones 0.0: Don’t forget the system**
  - You are building tiers and caching for the Applications, not the storage
Inherent Cost/Performance Differences in Storage

**HDD**
- Head Gimbal Assembly (HGA)
- Replicated over surface of disk by mechanical motion
- 200mm Slider Wafer
- Diameter 2.5” or 3.5”
- = 80,000 Sliders
- = 40,000 Terabytes
- Price per GB in HDD ~ $0.08

**SSD**
- 300mm NAND Wafer
- = 250 NAND Die
- = 2 Terabytes
- Price per GB in SSD ~ $0.80
- No mechanical motion required and parallelized for higher throughput

- HDD: Low $/GB through “mechanical advantage”
- NAND Flash: High performance through parallelism
Tiered and Caching ecosystems – New SSHDs

Overall value proposition is favorable with Hybrid when all attributes are viewed collectively
Tiered and Caching Storage Solutions

- **Everything in Software**
  - Easy deploy, *usually* cost lower, easy education, system focus
  - Updates on every host, may overload FS-based caches

- **Everything in hardware**
  - Little education required, Set-it-and-forget it operation, robust operation
  - Expensive, proprietary, harder to update in-place

- **Striking a balance**
  - Only hosts applications and systems that need solution have ROI
  - Host-based applications and software update easily
  - Cooperative methods require interoperability – many vendors support
  - Keeps simple just that, allows new solutions to migrate
Influencing Tiers and Acceleration in Storage

- Application
- OS UX
- File system / Partitions/Volumes
- Port / Protocol driver
- Miniport / HW driver
- Host controller (HC)
- Upper Disk Filter driver
- Disk class driver
- Firmware (RTOS / CPU)
- DRAM / FLASH
- Media controller
- Interface controller
- ASIC
- Logical Block Addresses
- Device protocol
- Transport
- Clusters
- Files
- Link level (Serial or Parallel)

Complexity vs. Performance
Effecting and Influencing Windows Storage

** Windows XP
- File Systems
- Applications
- IO Managers & Filters
- Volumes & Partitions
- Disk Class & Filters
- HW Drivers
- Disk ASIC
- Hard Disk
- Many third-party drivers, more open on device variety

** Windows Vista / 7
- File Systems
- Applications
- ** IO Managers & Filters
- Volumes & Partitions
- Disk Class & Filters
- ** HW Drivers
- Disk ASIC
- ** Hard Disk
- Chipsets influence, storage devices strictly defined, filters expand for stability and device features

** Windows 8
- ** File Systems
- Applications
- ** IO Managers & Filters
- Volumes & Partitions
- ** Disk Class & Filters
- ** HW Drivers
- ** Disk ASIC
- ** Hard Disk
- Chipsets variety shrinks – storage devices locked to class, filters restricted, applications become focus for storage feature and acceleration enablement

** Places where third-parties influence
Conclusions

- Understand your workload – Never assume: Measure
  - Working with more than one partner expands understanding
  - Make use of *free* educational resources – you pay for them
  - Watch what others do first

- Know that you are impacting the system
  - Application client performance desires acceleration closer to the CPU
  - Spread your costs to best accelerate data placement

- Striking a balance; again
  - Most metadata is dynamic in application need, once used, it’s flushed
  - File systems best cache file systems, avoid redundancy here
  - SAN- and NAS-based accelerators still need cooperation at host
  - OS-adopted standards for accelerations, go here