Flash, Fireworks, and the Cloud

Dr. Michael L. Workman
Senior Vice President
Oracle Flash Storage Systems
Oracle Corporation
OVER 6 YEARS AGO...
“There is no question that Flash is here to stay. There is also no question that architectures that were developed for HDD will have to be re-thought. New architectures will be developed by all Storage vendors to properly exploit Flash technology.”

—Mike Workman, 2009
TODAY
Network Connectivity and Compute Power *speed-up* have enabled new delivery systems...
“High-Bandwidth Network Connectivity is Crucial to Successfully Using the Cloud.”

1. San Murugesan and Irena Bojanova • August 2014, Computing Now
Continual Change: Nielsen’s Law and Moore’s Law

<table>
<thead>
<tr>
<th></th>
<th>Annual Growth Rate</th>
<th>10-Year CAGR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nielsen’s Law</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet Bandwidth</td>
<td>50%</td>
<td>57×</td>
</tr>
<tr>
<td><strong>Moore’s Law</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer Power</td>
<td>60%</td>
<td>100×</td>
</tr>
</tbody>
</table>
Everything Has Enabled the Cloud
Extremes of Security

Extreme Network Bandwidth

Extreme Performance
Has Changed Everything
HDDs IN THE ALL-FLASH WORLD
HDDs...

The Good: Linear Density and HDD Bandwidth have grown exponentially

The Bad: Latency has not

The Ugly: Shingled Magnetic Recording will make things worse
And Disk CAGR is Slowing...Cost Decline Slowing

Source: Anderson, Rosenthal, Register

Source: Preeti Gupta UCSC
Relative Differences: Flash and HDD in Storage Arrays

- Latency
- Sequential BW: RAID+Intf
- 90% Duty Cycle Power
- Cost/GB

Flash
HDD
Amdahl’s Law.

\[ S = \frac{1}{\sum_j \frac{P_j}{S_j}} \]

Speedup is limited by the fraction of the work that is not parallelizable, even using an infinite number of processors.
Where Does Sequential Data Fit for Storage Arrays?

Sequential Throughput is always limited by interconnect, SAS-2, SAS-3, or PCIe

**System Aggregate Sequential BW**

**THIS IS WHY SCALE-OUT IS MUCH MORE VIABLE THAN SCALE UP**
Flash vs HDD - Today There Is a Tangible Cost Difference

No matter who looks at it...

**eMLC vs. HDD**

- HDD
- eMLC

**Flash vs HDD: Industry Cost Trends**

MLC 8X More Expensive Than Performance HDD Through 2016

- 30X $/GB
- 75X $/IOP

- 40X
- 10X
- 40X
- 8X

Copyright © 2015 Oracle and/or its affiliates. All rights reserved.
Eventually the Cost Differential Won’t Matter

HDDs lost in the overall value equation in applications where absolute volume, fragility, power were more important than $/MB
It All Comes Down to Engineering. Eventually.
Flash Form Factor for Storage Systems

• HDDs are in their current form factor because platters are round.

• SSDs are in their current form factor because they copy HDDs.

• They copy HDDs to leverage the enclosures, including the RAS features of the enclosures.

• While expedient, this is not optimal.
…But Breaking the Form Factor of HDD is Happening
Better Form Factors Exploit Flash....But Arrays need better Reliability, Availability, Serviceability (RAS) than Servers

- Engineering packaging improvements will happen
- We need standards with RAS characteristics developed for HDD’s of the last 25 years...
- While still delivering on the performance
Flash Capacity

- Solid State Flash Storage Pool
- Optional HDD Storage Pool
- Read Cache
- Non-volatile Write Cache

Flash for non-volatile cache

Flash Cache

- Flash R/W Cache
- RAM Read Cache

IO
Sooner or later
you have to eat your peas
Storage Systems: The Trend

Today

Two Socket CU x2

Cap Disk

Perf Disk

Flash

+ 2 Years

Two Socket CU x2

Cap Disk

Perf Disk

Flash

+ 5 Years

Two Socket CU x2

Cap Disk

Perf Disk

Flash
Where do Data Reduction Technologies Fit?

Original Data

Deduplication

Deduplicated Data
Especially Oracle Database Storage:

- With Oracle databases, array-based dedupe and compression have negative impact on performance.
- In Oracle databases, each 8K database block header is unique therefore deduplication will not reduce data.
- Best practices for database security is encrypting data in the DB, but encrypted data can’t be deduped.
Oracle Database
Hybrid Columnar Compression

Oracle's Maximum Data Compression Solution

10x Up to 50x

Oracle Database
Uncompressed Data

10x to 50x Reduction

HCC Data

Save 40% on Storage
SO, WHAT DID ORACLE BUILD?
Oracle FS1 Storage Systems

Engineered for Enterprise Datacenters, Co-Engineered with Oracle Databases and Cloud

All-Flash FS1

FS1-2
All-Flash FS | Architected for All-Flash Availability

Unique Protection at Power Outage

Non-volatile DIMM +
All-Flash FS ESM

No batteries – Indefinite up time at power outage.

Retains advantages of battery-backed RAM while eliminating maintenance and persistence issues associated with batteries.
FS: Advanced High Availability Scale-Out Architecture Grows with Your Business

- Up to 16 HA Nodes
- Petabytes of Flash
- Up to 2M 50/50 R-W IOPS
- Up to 80 GB/sec Throughput
- Up to 6.6TB Cache, 128 Ports
- 512 Secure Storage Domains
- Full Suite of Data Services
The All-Flash Cloud is a Reality

- **Block IO Services with Flash Performance**
  - IOPS, Bandwidth and Latency to manage today’s transaction response time SLA’s

- **Provisioning IOPS on Flash**
  - No Read Penalties from Spinning Media

- **Scale Up and Scale Out Non-Disruptively**
  - Create, Delete, Clone, Map Operations much faster
  - Add new customers, workloads without interruption
The Future.....

- The cloud will drive an *avalanche* of flash adoption.
- The speed of adoption will be moderated by availability/price-pressure.
- Low Latency with minimal fall off will become the cornerstone “must-have” from active storage in the cloud.
- Full exploitation of flash for Storage Systems will require us to abandon the HDD form factor.
- NVMe is better for latency, but we need an NVMe approach that is RAS optimized for Storage Arrays
- Flash forces continued migration toward scale-out architectures
- Price differential will extend the life of capacity HDDs for 5 years.
Integrated Cloud
Applications & Platform Services