
Anil Vasudeva
President & Chief Analyst
imex@imexresearch.com
408-268-0800
Agenda

(1) Enterprise Needs for Storage
Data Centers Infrastructure & IT Roadmap to Clouds
Storage Usage in Data Centers & in Cloud Mega-Centers
Applications mapped by Key Workload & Storage Metrics
Drivers for Flash & SSDs

(2) State of SSDs
A New SCM – Filling Price Performance Gaps
Advantage SSD: vs. HDDs
Choosing SLC vs. MLC SSDs for the Enterprise

(3) SSD Challenges & Solutions
Paranoia vs. Reality: Wearout, Life, Retention, Pricing..
Testing Stds, Eco System/ Flash, Controllers, Systems, $8.5B Mkt..

(4) SSD Market Segments by I/F
Host based vs. Stg AddOns – PCIe vs. SAS/SATA/FC

(5) New Intelligent Controllers
Meeting Enterprise Requirements - AlwaysOn24x7, Fast I/O – IOPS, BW
5-Yr Life, Failure Rates, MTBF, Power

(6) Best Practices: Best Usage of SSD in the Enterprise
Using Stg.Tiering SW: To Improve Query Response Time w SSDs
Which Apps most benefit from using SSDs (OLTP/Bl/HPC/Web2.0/Data Streaming)

(7) Key Take-Aways
IT Industry’s Journey - Roadmap

**Cloudization**
On-Premises > Private Clouds > Public Clouds
DC to Cloud-Aware Infrast. & Apps. Cascade migration to SPs/Public Clouds.

**Automation**
Automatically Maintains Application SLAs
(Self-Configuration, Self-Healing, Self-Acctg. Charges etc)

**Virtualization**
Pools Resources. Provisions, Optimizes, Monitors
Shuffles Resources to optimize Delivery of various Business Services

**Integration/Consolidation**
Integrate Physical Infrast./Blades to meet CAPSIMS
Cost, Availability, Performance, Scalability, Inter-operability, Manageability & Security

**Standardization**
Standard IT Infrastructure- Volume Economics HW/Syst SW
(Server, Storage, Networking Devices, System Software (OS, MW & Data Mgmt SW)

Source: IMEX Research - Cloud Infrastructure Report ©2009-11
Market Segments by Applications

*IOPS* for a required response time (ms)

* = (#Channels * Latency - 1)

Source: IMEX Research - Cloud Infrastructure Report ©2009-11
Data Storage Usage – In Corporate Data Centers

I/O Access Frequency vs. Percent of Corporate Data

- **Cache**
  - Logs
  - Journals
  - Temp Tables
  - Hot Tables

- **Disk Arrays**
  - Tables
  - Indices
  - Hot Data

- **Tape Libraries**
  - Back Up Data
  - Archived Data
  - Offsite DataVault

Source: IMEX Research - Cloud Infrastructure Report ©2009-11
I/O Access Frequency vs. Percent of Corporate Data

- **SSD**
  - Logs
  - Journals
  - Temp Tables
  - Hot Tables

- **FCoE/SAS Arrays**
  - Tables
  - Indices
  - Hot Data

- **Cloud Storage**
  - SATA
    - Back Up Data
    - Archived Data
    - Offsite DataVault

Source: IMEX Research - Cloud Infrastructure Report ©2009-11
NAND Flash Enabling New Markets - Consumer to Enterprise

Source: Samsung & IMEX Research

Cost Decreasing
~50% / GB / Year

Density Increasing
2X per Year

SSD
64-300 GB

NAND Fueling New Applications...

DSC Card
32MB ~ 4GB

USB Drive
64MB ~ 4GB

MP3 Player
512MB ~ 8GB

Cell Phone
512MB ~ 8GB

Component Density, Gb

NAND - $/GB

$1,000

$100

$10

$1


Component Cost, $
• **Price Erosion Trends**
  - Driven by an explosion in the use of cost-sensitive handheld mobile devices, MLC NAND has seen an explosive growth.
  - On enterprise side Clustered low cost servers used in multiple environments from DB to BI to HPC applications besides being driven by Cloud Service Providers are providing an overall growth of 107% cagr in Computing SSDs GB.
  - SSD units are forecasted to grow at 86% cagr during the 2010-14 time frame.
SSD Filling Price/Perf Gaps in Storage

HDD becoming Cheaper, not faster

DRAM getting Faster (to feed faster CPUs) & Larger (to feed Multi-cores & Multi-VMs from Virtualization)

SSD segmenting into
- PCIe SSD Cache
- SATA SSD

Source: IMEX Research SSD Industry Report ©2011
SCM: A new Storage Class Memory

• **SCM (Storage Class Memory)**
  Solid State Memory filling the gap between DRAMs & HDDs
  Marketplace segmenting SCMs into SATA and PCIe based SSDs

• **Key Metrics Required of Storage Class Memories**
  
  **Device** - Capacity (GB), Cost ($/GB),
  
  **Performance** - Latency (Random/Block RW Access-ms); Bandwidth BW(R/W- GB/sec)
  
  **Data Integrity** - BER (Better than 1 in 10^17)
  
  **Reliability** - Write Endurance (No. of writes before death); Data Retention (Years); MTBF (millions of Hrs),
  
  **Environment** - Power Consumption (Watts); Volumetric Density (TB/cu.in.); Power On/Off Time (sec),
  
  **Resistance** - Shock/Vibration (g-force); Temp./Voltage Extremes 4-Corner (oC,V); Radiation (Rad)
Advantage: Enterprise SSDs vs. HDDs

Note: 2U storage rack, • 2.5” HDD max cap = 400GB / 24 HDDs, de-stroked to 20%, • 2.5” SSD max cap = 800GB / 36 SSDs

Source: IMEX Research SSD Industry Report ©2011
## Advantage: Enterprise SSDs vs. HDDs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SSD</th>
<th>HDD</th>
<th>Improvement SSD vs. HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concurrent Access</td>
<td></td>
<td></td>
<td>900 %</td>
</tr>
<tr>
<td>Data Access Time ms</td>
<td></td>
<td></td>
<td>&lt;1 %</td>
</tr>
<tr>
<td>IOPS</td>
<td></td>
<td></td>
<td>475 %</td>
</tr>
<tr>
<td>Read Speed</td>
<td></td>
<td></td>
<td>500%</td>
</tr>
<tr>
<td>1.0 MTBF (Million Hrs) *</td>
<td>2.1</td>
<td>1.0</td>
<td>110 %</td>
</tr>
<tr>
<td>&lt;5% Failure Rate (AFR%) **</td>
<td>&lt;=3%</td>
<td>&lt;5%</td>
<td>40 %</td>
</tr>
<tr>
<td>10^(-14) UBER **</td>
<td>10^(-16)</td>
<td></td>
<td>16 %</td>
</tr>
<tr>
<td>11.4 GB/W Power Efficiency</td>
<td>570 GB/W</td>
<td></td>
<td>5,000 %</td>
</tr>
<tr>
<td>43.1 IOPS/W Performance/Power</td>
<td>42,850 IOPS/W</td>
<td></td>
<td>100,000 %</td>
</tr>
<tr>
<td>6.8 Watts Idling Power</td>
<td>0.5 Watts</td>
<td></td>
<td>93 %</td>
</tr>
<tr>
<td>10.1 Watts Load Power</td>
<td>0.9 Watts</td>
<td></td>
<td>91 %</td>
</tr>
<tr>
<td>1.0 GB/in3 Storage Density</td>
<td>16 GB/in3</td>
<td></td>
<td>1600 %</td>
</tr>
<tr>
<td>4.2 IOPS/in3 Performance Density</td>
<td>1,250 IOPS/in3</td>
<td></td>
<td>30,000 %</td>
</tr>
<tr>
<td>Shock/Vibration/Noise Weight</td>
<td></td>
<td></td>
<td>800/1600%/30dBLess</td>
</tr>
<tr>
<td>Maintenance/Op.Time #</td>
<td></td>
<td></td>
<td>50 %</td>
</tr>
</tbody>
</table>

Source: IMEX Research SSD Industry Report ©2011

#Reduced - Booting Up, - Virus Scan, - Defrag, - RAID Build, - Patching, - Data Restoration** JEDEC's Mfr's Required Specs
SSD Challenges & Solutions: Industry Standard Testing

JEDEC Standard: Manufacturer Must Meet Requirements

<table>
<thead>
<tr>
<th>Class</th>
<th>Active Usage</th>
<th>Retention</th>
<th>Failures</th>
<th>UBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Power On</td>
<td>Power Off</td>
<td>FFR</td>
<td></td>
</tr>
<tr>
<td>Client</td>
<td>8 Hrs/day (40°C)</td>
<td>1 yr. (40°C)</td>
<td>&lt;=3%</td>
<td>&lt;10^-15</td>
</tr>
<tr>
<td>Enterprise</td>
<td>24 Hrs/day (40°C)</td>
<td>3 mo. (40°C)</td>
<td>&lt;=3%</td>
<td>&lt;10^-16</td>
</tr>
</tbody>
</table>

JEDEC Standard: Specify Endurance, Verify Spec via EVT

- Rigorous verification of Spec using EVT (Endurance Verification Test)
- JEDEC supplies the workload. Data is continuously read and verified.
- Endurance spec is max TB written to SSD over which device meets spec
- SSD must meet <=3% fail, UBER <1 in 10^16
- EVT requires high/low temp stressing
- EVT represents lifetime worth of Stress Test, so its trusted
- Accelerated Test (High Temp) & Unaccelerated Room Temp Retention Test
- Manufacturer provides ‘gauge’ informing user of % of endurance life used up

JEDEC Standards Testing & Verification for Endurance, Failures and UBER under accelerated life testing assures use of SSDs in Enterprises
## SSD Drivers & Challenges: MLC vs. SLC

### Drivers

<table>
<thead>
<tr>
<th>Raw Media Reliability</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>No moving parts</td>
<td>Higher density of MLC increases bit error rate</td>
</tr>
<tr>
<td>Predictable wear out</td>
<td>High bit error rate increases with wear</td>
</tr>
<tr>
<td>Post infant mortality</td>
<td>Program and Read Disturb Prevention, Partial Page Programming</td>
</tr>
<tr>
<td>catastrophic device failures rare</td>
<td>Data retention is poor at high temperature and wear</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Media Performance</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance is excellent (vs. HDDs)</td>
<td>NAND not really a random access device</td>
</tr>
<tr>
<td>High performance/Watt (IOPS/Watt)</td>
<td>Block oriented; Slow effective write, erase/transfer/program) latency,</td>
</tr>
<tr>
<td>Low pin count: shared command / data bus, good balance</td>
<td>Imbalanced R/W access speed</td>
</tr>
<tr>
<td>NAND Performance changes with wear, Some controllers do read/erase/modify/write, Others use inefficient garbage collection</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controller</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transparently converts NAND Flash memory into storage device</td>
<td>Interconnect</td>
</tr>
<tr>
<td>Manages high bit error rate</td>
<td>Number of NAND Flash Chips (Die); # of Buses (Real / Pipelined)</td>
</tr>
<tr>
<td>Improves endurance to sustain a 5-year life cycle</td>
<td>Data Protection (Int./Ext.RAID; DIF; ECC); Write Mitigation techniques</td>
</tr>
</tbody>
</table>

Source: IMEX Research SSD Industry Report ©2011
MLC vs. SLC SSDs - Price Erosion

Relative Price Erosion  SLC vs MLC

% Price Erosion ($/GB)

2004 2005 2006 2007 2008 2009 2010 2011e 2012e 2013e
SSD Challenges & Solutions: Endurance/Wear-out

Reason for Endurance Limitation in SSDs

Anatomy of a PE Cycle in SSDs (Roundtrip through Tunnel Oxide)

- Fundamentally – NAND Flash Memory Cell is an MOS Transistor with a Floating Gate that can permanently store charge
- Programming puts electrons in Floating Gate, Erase takes them off
- 1 Program/Erase (P/E) Cycle is a round trip by the electrons
- Electrons pass through Cell’s Tunnel Oxide. Back & Forth round trips gradually damage the Tunnel Oxide over hundred thousands of trips (Program/Erase or PE cycles) resulting in Limited Endurance (or Wear-Out by PE cycles) in SSDs

JEDEC Standards Testing & Verification for Endurance, Failures and UBER under accelerated life testing assures use of SSDs in Enterprises
SSD Challenges & Solutions: Endurance (Wear-Out)

**Challenge: Bad Block Mgmt**

- The ability to erase slows down after a number of P/E Cycles.
- If NAND Memory block fails to erase, Controller is notified and another block from spares is used instead.
- But there’s no loss of data, so a failed NAND block does not pose a problem.
- Eventually devices will run out of spares.
- The point where the % failing exceed number of spares is the most Basic Endurance Limit.

**Solution: Over Provisioning**

- Endurance can vary, depending on Workload.
- Endurance should match usage needs of the system to minimize costs.
  - SSD used as cache for 10 HDDs. 2 PB writes of useful life will support this.(1.1 TB writes/day for 5 years.)
- Over Provisioning by Increasing Spare blocks
  - Decreases user capacity but
  - Allows SSD to more efficiently complete random Writes
  - Improves Random Write Endurance and Performance
- Methods to Implement include:
  - Setting max LBA to limit visible drive capacity or
  - Create Smaller RAID Logical Drives or
  - Create Smaller Partitions

SSD Challenges & Solutions: Endurance (UBER)

**Challenge: Uncorrectable BER Mgmt**

- A small of written bits gets flipped (similar to HDDs)
- This is Flash Media’s **Raw Bit Error Rate (RBER)**
- ECC is used to correct/reduce this RBER
- RBER gradually increases with P/E cycles. Any bit error rate over ECC Correction capability is the **Uncorrected Bit Error Rate (UBER)**. Reaching a UBER domain user data can become corrupted.
- UBER is kept low. JEDEC Spec is 1 in $10^{16}$ errors
- The point where UBER reaches this spec, is **Another Endurance Limit**

**Solution: ECC**

Flash Media Starts with - 1 in $10^8$ (1 error/100 million bits) Read

Flash Media’s Raw Bit Errors (RBER)

Corrected by ECC

Left Uncorrected – 1 in $10^{16}$ (1 error/10,000 Trillion bits Read)

- Using modern ECC techniques based controllers, vendors are providing spec at 1 in $10^{-17}$ UBER

SSD Challenges & Solutions: Data Retention

Challenge: Data Retention

Solution: Data Retention Firmware

- Powered-On Firmware
  - To allow Higher Retention
- Balance out SSD Data Retention vs. Endurance
  - Lower Data Retention allows for higher endurance

- After PE cycles, RBER increases with time. ECC corrects bit flips but only to a certain extent.
- So the industry lives with a required UBER and required Retention Time. This, in turn, determines the Safe PE cycles that device should be exercised to, prior to reaching the UBER and Retention time. This is also another endurance limit set by retention.

### SSD Challenges & Solutions: Functional Failure Defects

**Challenge:** Electronic Component - Defects

- All ICs have defects that cause failures. In Flash early life failures are caused by such defects.
- Defects can cause functional failures not just data loss. Most of NAND defect failures are caused by PE cycles, coming in from high PE voltages causing defects to short.
- The point where % failing from defects would reach unacceptable limits is another boundary for endurance.

**Solution:** Burn-Ins, Error Avoidance Algorithm

- Vigorous SSD Burn-In & Testing
  - Remove Infant Mortality
- Compute NAND
  - $T_{\text{read}}$ to improve Read Disturbs
  - $T_{\text{PROG}}$ to reduce Program Disturbs
- SSD Error Avoidance algorithms
  - ECC ASICS
- Wear Leveling to avoid Hot Spots
- Efficient Write Amplification Factor (WAF)
  - WAF = Data written to NAND /Data Written by Host to SSD
  - WAF dependent on (a) SSD FW algorithm built into SSD (b) Over Provisioning Amount (c) App Workload

---

**Role of Defects in SSD Reliability**

- Wafer Process Defects: 61%
- Design Related & Test: 10%
- EOS/ESD: 10%
- Handling: 9%
- Process Errors: 5%
- Assembly & Test: 5%

---

**Graph:**

- **Failure Rate** vs **Time**
  - Region 1: Infant Mortality
  - Region 2: Wear Leveling
  - Region 3: Efficient WAF

---

**Source:** Intel IDF’10 & IMEX Research SSD Industry Report 2011 ©IMEX 2010-11
SSD Challenges & Solutions: Industry Standard Testing

JEDEC Requirements: Specify Endurance, Verify Spec via EVT

<table>
<thead>
<tr>
<th>Class</th>
<th>Active Usage</th>
<th>Retention</th>
<th>Failures</th>
<th>UBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Power On</td>
<td>Power Off</td>
<td>FFR</td>
<td></td>
</tr>
<tr>
<td>Client</td>
<td>8 Hrs/day (40°C)</td>
<td>1 yr. (40°C)</td>
<td>&lt;=3%</td>
<td>&lt;10^-15</td>
</tr>
<tr>
<td>Enterprise</td>
<td>24 Hrs/day (40°C)</td>
<td>3 mo.(40°C)</td>
<td>&lt;=3%</td>
<td>&lt;10^-16</td>
</tr>
</tbody>
</table>

- Rigorous verification of Spec using EVT (Endurance Verification Test)
- Endurance spec is max TB written to SSD over which device meets spec
- JEDEC supplies the workload. Data continuously read and verified.
- Lifetime worth of Stress Test, so can be trusted,
- Accelerated High Temp and Unaccelerated Room Temp Retention Tests
- Manufacturer provides ‘gauge’ informing user of % of endurance life used up
SSD Challenges & Solutions: Goals & Best Practices

NAND Flash will have finite Endurance Limits due to limitations imposed by:

- Uncorrectable Bit Error Rates
- Functional Failures
- Data Retention Time

**Goal: Embody technologies to Improve Life (Years of Use)**

- Push TeraBytesWritten (Endurance Limit) beyond product life as required by SSD products
- Push defect rate/AFR down through Burn-Ins, Error Avoidance Algorithms and Practices, so total defects and wear-outs issues combined is <=3%
- Target data errors to be < 1 in 10^16 for Enterprise SSDs for both TBW and Retentions specs.

With advanced Errors Avoidance (ECC/Wear-Leveling etc) and capacity over-provisioning techniques, controllers successfully creating...
SSD Challenges & Solutions: Goals & Best Practices

Concerned about SSD Adoption in your Enterprise?

Be aware of Tools & Best Practices … And you should be OK!!

Best Practices

- By leveraging Error Avoidance Algorithms, and Best Practices of Verification Testing, to keep total functional failure rate <=3% (with defects and wear-outs issues combined)
- In practice, endurance ratings are likely to be significantly higher than typical use, so data errors and failures will be even less.
- Capacity Over-provisioning will provide large increases in random performance and endurance.
- Select SSD based on confirmed EVT Ratings
- Use MLC within requirements of Endurance Limits

Using Best-of-Breed Controllers to achieve <=3% AFR and JEDEC Endurance Verification Testing should allow Enterprise Capable SSDs

SCM – New Storage Class Memory

- HDD becoming Cheaper, not faster
- SSD segmenting into
  - PCIe SSD Cache
  - SATA SSD
- PCIe SSD

- CPU SDRAM
- DRAM getting Faster (to feed faster CPUs) & Larger (to feed Multi-cores & Multi-VMs from Virtualization)

Source: IMEX Research SSD Industry Report ©2011
WW Enterprise SSD 5-Yr Mkt Opportunity
Cum $B (2010-14)

Market Size
$8.6B
(5-Yr cum)

2010-14 CAGR %

60%

0%

-60%

$-

$2

$4

5-Yr Cum Market Size $B by Interface

SAS

PCle

SATA

FC

Source: IMEX Research SSD Industry Report ©2011
PCIe based SSD Storage

Target Market – Servers Storage
SSD as backend storage to DRAM as the front end
36 PCIe Lanes Availability, 3/6 GB/s Performance (PCIe Gen2/3 x8), Low Latency in micro sec, Low Cost (via eliminating HBA cost)

PCIe SSD’s attributes of high IOPS, high Bandwidth, Low Latency and lower cost are a good match for Caching
Hybrid SSD Storage

• Hybrid Storage – SAS or SATA SSD+HDD
  • Target market – External Storage Systems
  • Combines best features of SSDs - outstanding Read Performance (Latency, IOPs) and Throughput (MB/s) with extremely low cost of HDDs giving rise to a new class of storage - Hybrid Storage Devices
  • SSD as Front End to HDD
  • Controller emulates SSD as HDD
  • Use of Adaptive Memory sends High IOPS requirements to SSD while capacity requiring Apps sent to HDD
  • Simple Add on to SATA HDD Storage
  • SAS 6Gb/sec announced by multi-vendors
Hybrid SSD Storage - Perf & TCO

SAN TCO using HDD vs. Hybrid Storage

SAN Performance
Improvements using SSD

Source: IMEX Research SSD Industry Report ©2011
**New Intelligent Controllers: SSD Storage Architecture**

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interface Controller&lt;br&gt;&amp; Host&lt;br&gt;Signaling Mgmt, Interpret WR/RD/Status Commands for Flash Arrays, Move Data. Defect Mapping/Bad Block Mgmt, Wear Leveling, Physical&lt;&gt;Logical Translations, ECC...</td>
</tr>
<tr>
<td>2</td>
<td>Flash Controller&lt;br&gt;Signaling Mgmt, Format, Interpret WR/RD/Status Commands for Flash Arrays, Move Data. Defect Mapping/Bad Block Mgmt, Wear Leveling, Physical&lt;&gt;Logical Translations, ECC...</td>
</tr>
<tr>
<td>3</td>
<td>RAID Controller&lt;br&gt;RAID Type &amp; RD/WR/Parity Manipulation</td>
</tr>
<tr>
<td>4</td>
<td>Channels&lt;br&gt;Multiple Channel to Increase Speed between NAND Flash Arrays &amp; Flash Controller</td>
</tr>
<tr>
<td>5</td>
<td>DRAM&lt;br&gt;Increase Performance using fast DRAM Cache Buffer</td>
</tr>
<tr>
<td>6</td>
<td>Power Failure&lt;br&gt;Power Failure Protection using Big Capacitor</td>
</tr>
<tr>
<td>7</td>
<td>Power Mgmt&lt;br&gt;Power/Performance Balancing, Sleep Mode Mgmt</td>
</tr>
<tr>
<td>8</td>
<td>Encryption&lt;br&gt;Security Schemes Implementation &amp; Manipulation</td>
</tr>
</tbody>
</table>
Leveraging Long History of managing HDD’s imperfect media & high error rates

- Characterizing the quality & capabilities of media
- Allocating data based on quality of media

**HDD Media**

Adaptive Signal Processing for Media Rd/Wr/Erase
Advanced Bit Detection & Error Correction Codes
Defect Management

**Flash Media**

Adaptive Signal Conditioning for Flash Media
Auto Bit Detection & Error Correction Codes
Defect Management

Leveraging Long History of managing HDD’s imperfect media & high error rates

- Endurance for Long Life Cycle
- Reliability through RAID of Flash Elements
- Adaptive Digital Signal Processing Technology
  - Dynamically adjust Read/Write characteristics of each chip
- Tune adjustments over life of media
- ECCs - PRML Deploying Enhanced Error Correction Codes

Source: IMEX Research SSD Industry Report ©2011
New Intelligent Controllers: Meeting Enterprise Requirements

Enterprise Requirements

- **Always-On 24x7 Reliability** and performance supersede cost
- **Fast I/O Performance** required by business-critical applications and
- **5-Yr. Life Cycle Endurance** required by mission-critical applications in the enterprise.
- **Use State-of-the-Art** new sophisticated controllers and firmware technologies to run mission critical applications in the enterprise, using
  - Robust ECC, Internal RAID, Wear Leveling (To reduce hot spots), Spare Capacity, Write Amplification, Avoidance, Garbage Collection Efficiency, Wear Out Prediction Management etc.

New Gen Controllers allow SSDs to meet Enterprise Class Availability/Performance/ over 5-Year Life/Scalability/ Auto-Configuration & Auto Data-Tiering
New Intelligent Controllers: Managing Endurance in NextGen SSDs

- **Managing Endurance**
  
  To overcome NAND’s earlier endurance shortfalls due to limitation in write/erase cycles/block, intelligent controllers manage NAND SSDs using
  
  - **ECC Techniques** – Correct and guard against bit failures, same as in HDDs
  - **Wear Leveling Algorithms** – Writing data to evenly distributes it over all available cells to avoid a block of cells being overused and cause failures.
  
  - **Over-provisioning Capacity** – Extra spare raw blocks are designed-in as headroom and included to replace those blocks that get overused or go bad. Additionally provide enough room for wear-leveling algorithms to enhance reliability of the device over its life-cycle.
  
  - Typical SSD device’s specified GB device will actually contain 20-25% extra raw capacity to meet these criterions.

With advanced Errors Avoidance (ECC/Wear-Leveling etc) and capacity over-provisioning techniques, controllers successfully creating endurance for over 5-Year Product Life Cycle in Enterprise SSDs
New Intelligent Controllers
- Performance in Next Gen SSDs

- Managing Factors Impacting Performance
  - **Hardware** - CPU, Interface, Chipset ...
  - **System SW** - OS, App, Drivers, Caches, SSD specific TRIM, Purge...
  - **Device** - Flash Generation, Parallelism, Caching Strategy, Wear-Leveling, Garbage Collection, Warranty Strategy...
  - **Write History** - TBW, spares...
  - **Workload** - Random, Sequential, R/W Mix, Queues, Threads...
  - **Pre-Conditioning** - Random, Sequential, Amount ...
  - **Performance** - “Burst” First On Board (FOB), Steady State post xPE Cycles

By using New Gen Controllers, performance of MLC SSDs starting to match performance of some SLC SSDs

Additional performance gains with interleaved memory banks, caching and other techniques
• **Improving Query Response Time**
  
  • Cost effective way to improve Query response time for a given number of users or servicing an increased number of users at a given response time is best served with use of SSDs or Hybrid (SSD + HDDs) approach, particularly for Database and Online Transaction Applications.

Source: IMEX Research SSD Industry Report ©2011
AutoSmart Storage-Tiering SW: Workload I/O Monitoring/Smart Migrations

**LBA Monitoring and Tiered Placement**
- Every workload has unique I/O access signature
- Historical performance data for a LUN can identify performance skews & hot data regions by LBAs
- Using Smart Tiering identify hot LBA regions and non-disruptively migrate hot data from HDD to SSDs.
- Typically 4-8% of data becomes a candidate and when migrated to SSDs can provide response time reduction of ~65% at peak loads.

Source: IBM & IMEX Research SSD Industry Report 2011 ©IMEX 2010-11
AutoSmart Storage-Tiering SW: Enhancing Database Throughput

- **DB Throughput Optimization**
  - Every workload has unique I/O access signature and historical behavior
  - Identify hot “database objects” and smartly placed in the right tier.
  - Scalable Throughput Improvement - 300%
  - Substantial IO Bound Transaction Response time Improvement - 45%-75%

- **Productivity (Response Time) Improvement**
  - Using automated reallocation of hot spot data to SSDs (typically 5-10% of total data), significant performance improvements is achieved:
    - Response time reduction of around 70+% or
    - IOPS increase of 200% for any I/O intensive loads
  - Verticals benefitting from Online Transactions:
    - Airlines Reservations, Investment Banking, Wall St. Stock Transactions, Financial Institutions Hedge Funds etc. plus Low Latency seeking HPC Clustered Systems etc.
### Applications Best Suited for SSDs

#### Apps and impact from SSD Usage

- **Databases**
  - Databases have key elements of commit files
  - Logs, redo, undo, tempDB

- **Structured data**
  - Structured data access is an excellent fit for SSD
  - Exception—large, growing table spaces

- **Unstructured data**
  - Unstructured data access is a poor fit for SSD
  - Exception – small, non-growing, tagged files

- **OS images**
  - boot-from-flash, page-to-DRAM

#### Typical Cases - Impact on Applications

- **Financials/ATM Transactions Improvements**
  - Batch Window 22%, App Response Time 50%,
  - App I/O Rate 50%

- **Messaging Applications**
  - Cost Savings: 200+ FC HDDS into only 16 SSDs

---

Source: IMEX Research SSD Industry Report ©2011
Apps Best Suited for SSDs: DB in Memory for Data Warehouse/BI

Scale In

Scale Up

Scale Out

Storage Usage vs DB Capacity

DB Size (TB)

Storage Size (TB)

Large DB Size Growth by Market Segment

OLTP

DW/BI

Corporate Data Usage

I/O Access Frequency vs. Percent of Data

Cache

Logs

Journals

Temp Tables

Disk Arrays

Hot Tables

Indices

Data

Cache

SSD

VTL & Tape Libraries

Back Up Data

Archived Data

Data Source: IMEX Research
Cloud Infrastructure Report ©2009-11

© 2010-11 IMEX Research, Copying prohibited. All rights reserved.
## Mitigating Boot Storms

- Boot Storm created by simultaneous Logins by users at start of office day
- Over provisioning SAN capacity just for short morning burst expensive, while sitting almost idle rest of the day
- Three potential solutions with pros and cons include:
  - (1) DAS Storage, (2) Virtual Desktop Images on SSD (3) SS Cache to accelerate SAN,

<table>
<thead>
<tr>
<th>Solution</th>
<th>Pros</th>
<th>Con</th>
</tr>
</thead>
</table>
| New DAS Storage                 | - Popular with Desktop SW Vendors  
- Lowered Cost                 | - Additional Cost for Dedicated Storage  
- Wasted existing SAN Storage  
- Data Protection & Mgmt Challenges |
| SSD                             | - SSD Ideal for read intensive app  
- Instant-On/Immediate Boot Response  
- Images Stored with High Efficiency | - Using most expensive storage  
- High Speed Needed just for an hour  
- Not simple shoe-in w existing storage |
| SAN w SSD Accelerator           | - Possibly best way to solve problem  
- Small SSD optimized for image store  
- No change to existing Data Protection | - Not feasible without existing SAN  
- SSD in SAN Integration still a challenge |

A perfect balance of access and storage is achieved through Integrating SATA HDDs with SSDs and using Automatic Tiering Solutions

Source: IMEX Research SSD Industry Report ©2011
<table>
<thead>
<tr>
<th>Apps Best Suited for SSDs: HPC/Web 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Smart Mobile Devices</strong></td>
</tr>
<tr>
<td><strong>Commercial Visualization</strong></td>
</tr>
<tr>
<td><strong>Bioinformatics &amp; Diagnostics</strong></td>
</tr>
<tr>
<td><strong>Decision Support Bus. Intelligence</strong></td>
</tr>
<tr>
<td><strong>Entertainment-VoD / U-Tube</strong></td>
</tr>
</tbody>
</table>

**Data:** IMEX Research & Panasas

- **Instant On Boot Ups**
  - Rugged, Low Power
  - 1GB/s, __ms

- **Rendering (Texture & Polygons)**
  - Very Read Intensive, Small Block I/O
  - 10 GB/s, __ms

- **Data Warehousing**
  - Random IO, High OLTPM
  - 1GB/s, __ms

- **Most Accessed Videos**
  - Very Read Intensive
  - 4 GB/s, __ms
Key Takeaways

• Optimize Infrastructure to meet needs of Applications/SLA
• Solid State Storage creating a paradigm shift in Storage Industry
  • Leverage the opportunity to optimize your computing infrastructure with SSD adoption after making a due diligence in selection of vendors/products, industry testing and interoperability
• Enterprise SSD Market Segments: PCIe vs. SAS/SATA
  • 5-Year cum Market $8.6B Segments by Revenues: 36% PCIe, 33% SAS, 24% SATA, 6% FC based SSDs
• Understand Drivers and Challenges of SSDs for Enterprise Use
• Intelligent Controllers key to adoption & success of SSDs
  • Mitigate Endurance, Wear-Out, Life issues
• Optimize Transactions for Query Response Time vs. # of Users
  • Improving Query Response time for a given number of users (IOPs) or Serving more users (IOPS) for a given query response time
• Select Automated Storage Tiering Software
Acknowledgements

Many thanks to the following individuals for their industry vision* and leadership in advancing the Solid State Storage industry
(*in preparation of some slides in this presentation)

• N. Mielke, Fellow & Director of Reliability Methods, Intel
• Jim Elliott, VP Marketing, Samsung
• John Scaramuzzo, SVP & GM, Smart Modular Technologies
• Michael Raam, President & CEO, SandForce
• Manouch Moshayedi, CEO, STEC
• Dean Klein, VP Memory Systems Development
• Joseph White, Chairman SSD Initiatives, SNIA

Author: Anil Vasudeva, President & Chief Analyst, IMEX Research