Solid State Storage: Key to Efficient Tiered Storage Infrastructure
Agenda – Automated Tiered Storage

1. NextGen Data Center and Cloud Computing Infrastructure
2. Solid State Enabling New Systems Architecture
3. Improving Transaction Query Response Time and IOPS
4. Workload Characterization
5. Applications best utilizing Solid State Storage
6. Data Forensics and Tiered Mapping
7. Selecting Automated Storage Tiering Software
8. Key Takeaways
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©IMEX, 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 ©IMEX
Cost, Availability, Performance, Scalability, Inter-operability, Manageability & Security

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

Source: IMEX Research - Cloud Infrastructure Report ©2009-11
© 2010-11 IMEX Research, Copying prohibited. All rights reserved.
Request for data from a remote client to an enterprise data center crosses a myriad of systems and devices. Key is identifying bottlenecks & improving performance.
Application’s SLA dictates the Resources Required to meet specific requirements of Availability, Performance, Cost, Security, Manageability etc.
**Server to Storage IO Gap**

For Each Disk Operation:
Millions of CPU Operations or Hundreds of Thousands of Memory Operations can be done

<table>
<thead>
<tr>
<th>Operation</th>
<th>Time (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 cache reference</td>
<td>0.5</td>
</tr>
<tr>
<td>Branch mispredict</td>
<td>5</td>
</tr>
<tr>
<td>L2 cache reference</td>
<td>7</td>
</tr>
<tr>
<td>Mutex lock/unlock</td>
<td>25</td>
</tr>
<tr>
<td>Main memory reference</td>
<td>100</td>
</tr>
<tr>
<td>Compress 1K bytes with Zippy</td>
<td>3,000</td>
</tr>
<tr>
<td>Send 2K bytes over 1 Gbps network</td>
<td>20,000</td>
</tr>
<tr>
<td>Read 1 MB sequentially from memory</td>
<td>250,000</td>
</tr>
<tr>
<td>Round trip within same datacenter</td>
<td>500,000</td>
</tr>
<tr>
<td>Disk seek</td>
<td>10,000,000</td>
</tr>
<tr>
<td>Read 1 MB sequentially from disk</td>
<td>20,000,000</td>
</tr>
<tr>
<td>Send packet CA-&gt;Netherlands-&gt;CA</td>
<td>150,000,000</td>
</tr>
</tbody>
</table>

A 7.2K/15k rpm HDD can do 100/140 IOPS
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: as backend to DRAM &
  - SATA SSD: as front end to HDD

Source: IMEX Research SSD Industry Report ©2011
SSDs - Price Erosion & IOPS/GB

Key to Database performance are random IOPS. SSDs outshine HDD in IO price/performance – a major reason, besides better space and power, for their explosive growth.

Source: IMEX Research SSD Industry Report ©2011
# Advantage Solid State Storage (vs. HDDs)

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

** JEDEC’s Mfr’s Required Specs

.Execution:**

| #Reduced - Booting Up, - Virus Scan, - Defrag, - RAID Build, - Patching, - Data Restoration |

© 2010-11 IMEX Research, Copying prohibited. All rights reserved.
For a targeted query response time in DB & OLTP applications, many more concurrent users can be added cost-effectively when using SSDs or SSD + HDDs storage vs. adding more HDDs or short-stroking HDDs.

Source: IMEX Research SSD Industry Report ©2011
Buying SSD vs. Memory to Improve TPS

To achieve a certain TPS improvement, it's cheaper to deploy SSD vs increased buffer memory (in GB costs) needed with using HDDs alone.

Source: IMEX Research SSD Industry Report ©2011
**Market Segments by Apps/Workloads**

- **Transaction Processing**
- **eCommerce**
- **Data Warehousing**
- **OLTP**
- **OLAP**
- **Business Intelligence**
- **Scientific Computing**
- **Imaging**
- **Web 2.0**
- **Audio**
- **Video**

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

© 2010-11 IMEX Research, Copying prohibited. All rights reserved.
I/O Access Frequency vs. Percent of Corporate Data

% of Corporate Data

% of I/O Accesses

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
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 vs. Unstructured Data**
  - Structured/SQL data access is an excellent fit for SSD
  - Exception—large, growing table spaces
  - **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
Storage performance, management and costs are big issues in running Databases

- **Data Warehousing Workloads** are I/O intensive
  - Predominantly read based with low hit ratios on buffer pools
  - High concurrent sequential and random read levels
    - Sequential Reads require high level of I/O Bandwidth (MB/sec)
    - Random Reads require high IOPS
  - Write rates driven by life cycle management and sort operations

- **OLTP Workloads** are strongly random I/O intensive
  - Random I/O is more dominant
    - Read/write ratios of 80/20 are most common but can be 50/50
    - Can be difficult to build out test systems with sufficient I/O characteristics

- **Batch Workloads** are more write intensive
  - Sequential Writes requires high level of I/O Bandwidth (MB/sec)

- **Backup & Recovery** times are critical for these workloads
  - Backup operations drive high level of sequential IO
  - Recovery operation drives high levels of random I/O

Source: IMEX Research SSD Industry Report ©2011
Applications Best Suited for SSDs:
Data Warehouse/BI

Storage Usage vs DB Capacity

DB Size (TB)
Storage Size (TB)

1-2 TB
2-5 TB
5-10 TB
>10 TB

OLTP
DW/BI

Large DB Size Growth by Market Segment

I/O Access Frequency vs. Percent of Corporate Data

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

SSD
- Logs
- Journals
- Temp Tables
- Hot Tables

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

© 2010-11 IMEX Research, Copying prohibited. All rights reserved.
Apps Best Suited for SSDs: HPC/Web 2.0

<table>
<thead>
<tr>
<th>VDI</th>
<th>Commercial Visualization</th>
<th>Bioinformatics &amp; Diagnostics</th>
<th>Decision Support Bus. Intelligence</th>
<th>Entertainment-VoD / U-Tube</th>
</tr>
</thead>
</table>

**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
Automated Storage Tiering - Next Frontier in Storage Efficiency

Data Protection
- Back Up/Archive/DR
- RAID – 0, 1, 5, 6, 10
- Virtual Tape
- Replication

Storage Efficiency
- Storage Virtualization
- Thin Provisioning
- Deduplication
- MAID

Auto Tiering
- Data Class (Tiers 0, 1, 2, 3)
- Storage Media Type (Flash/Disk/Tape)
- Policy Engines (Workload Mgmt)
- Transparent Migration (Data Placement)
- File Virtualization (Uninterrupted App. Ops. in Migration)

Source: IMEX Research SSD Industry Report ©2011

© 2010-11 IMEX Research, Copying prohibited. All rights reserved.
Automated Storage Tiering: The Killer App for Enterprise SSDs

• **Traditional Disk Mapping**
  • Volumes have different characteristics. Applications need to place them on correct tiers of storage based on usage.

• **Smart Storage Mapping**
  • All volumes appear to be “logically” homogenous to apps. But data is placed at the right tier of storage based on its usage through smart data placement and migration.

Source: IMEX Research SSD Industry Report ©2011
Automated Storage Tiering: Workload I/O Monitoring & Migration

- Automated Storage Tiering
  - Continuously monitor and analyze data access on the tiers
  - Automatically elevate hot data to "Hot Tiers" and demote cool data/volumes to "Lower Tiers"
  - Allocate and relocate volumes on each tier based on use
  - Automated Migration reduces OPEX to otherwise SANs managed manually

Source: IMEX Research SSD Industry Report ©2011
Automated Storage Tiering: Workload I/O Monitoring & Migration

- **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: IMEX Research SSD Industry Report ©2011
Automated Storage Tiering: Improving Response Time

- **Productivity Improvements (Response Time)**
- With automated reallocation of hot spot data (~ 5-10% of total data) to SSDs, performance improvements
  - Response time reduction ~70+% IOPS increase of 200% for any I/O intensive workloads in Time-Perishable OLTP markets: Airlines Reservations, Wall Street Investment Banking Stock Transactions, Financial Institutions Hedge Funds etc.
  - **Performance boost** in Low Latency seeking Systems (High Perf. Clustered Systems)

Source: IMEX Research SSD Industry Report ©2011
Automated Storage Tiering: 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%

Source: IMEX Research SSD Industry Report ©2011
Storage Tiering - Best Practices Highlights

- **SSD-PCIe perform better** than SATA SSDs
  - Use Nehalem Class CPUs especially when using PCIe SSDs
- **Put Random Access Files on SSDs** (Index, Tables, Table Spaces)
  - Keep ample SSD Reserved Space to avoid massive SSD write deterioration
- **Put Sequentially Written Files on HDDs** since
  - HDDs better at Sequential Writes compared to SSDs
  - Removes SSD Write performance bottle necks
  - Increases SSD life
  - Archive Less Active Tables/Records to HDDs
- **Leverage Auto-Tiering Storage SW** to balance between SSDs and HDDs
  - Heat Mapping with Tier Managed Extent Pools
  - Workload Hot Spot Analysis
  - Smart Data Migration & Placement
  - Continuous Workload Monitoring
- **Use Faster Networks** (10GbE vs 1GbE) to avoid saturating DRBD
- **Target Price/Performance Economic Benefits of 150-800%**
SSS in DB Environments:
Enabling new System Architectures

SSD class memories fundamentally changing Computing Systems Architectures

Using SSDs, a leading Computer Systems company achieved:

• **Sustained 1 million IOPS** with random 4K size
• **70%RD/30%WR** with Queue depth of 16
• **System Latency 720 us** Average
• **Floor Space Less than 25%** vs. Std. Disk Storage System
• **Energy Used only 55%** vs Std. System
• **Comparable Total Cost** New system vs. Standard System
• System Test HW:
  Host - 2 Servers (26 cores, 28 GB Memory),
  Cluster - 14 Storage Controller Nodes,
  Storage - 31 Storage Arrays with 41 PCIe SSDs 160GB Each

Source: IMEX Research SSD Industry Report ©2011
### SSS in DB Environments:
**DB Improvements with Flash SSDs**

- Storage management, performance and cost - a big issue in DBs
- SSDs enable super IO performance and cost reduction in DBs

<table>
<thead>
<tr>
<th>Improve Responsiveness</th>
<th>Improve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Insert/Update/Delete Performance</td>
</tr>
<tr>
<td></td>
<td>• Random Read I/O Performance</td>
</tr>
<tr>
<td></td>
<td>• Query Response Time</td>
</tr>
<tr>
<td></td>
<td>• Sort Performance</td>
</tr>
<tr>
<td></td>
<td>• Batch Performance too.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reduce Costs</th>
<th>Reduce</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• DRAM size for Buffer Pools used to cache data on SSDs</td>
</tr>
<tr>
<td></td>
<td>• Power/Cooling Space for housing Databases</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>New Added Benefits</th>
<th>Reduce</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Database Recovery performance</td>
</tr>
<tr>
<td></td>
<td>• I/O performance impact by Flash Copy</td>
</tr>
<tr>
<td></td>
<td>• Skill levels required for DB tuning &amp; monitoring</td>
</tr>
</tbody>
</table>

Source: IMEX Research SSD Industry Report ©2011
SSS in DB Environments: Best Practices for DB/DW/BI Apps

Goals & Implementation

- Establish **Goals for SLAs** (Performance/Cost/Availability), BC/DR (RPO/RTO) & Compliance

- **Increase Performance for DB, Data Warehousing, OLTP Apps:**
  - Random I/O > 20x, Sequential I/O Bandwidth > 5x
  - Remove Stale data from Production Resources to improve performance

- **Use Partitioning Software to Classify Data**
  - By Frequency of Access (Recent Usage) and
  - Capacity (by percent of total Data) using general guidelines as:
    - Hyperactive (1%), Active (5%), Less Active (20%), Historical (74%)

Implementation

- **Optimize Tiering** by Classifying Hot & Cold Data
  - Improve Query Performance by reducing number of I/Os
  - Reduce number of Disks Needed by 25-50% using advance compression software achieving 2-4x compression

- **Match Data Classification vs. Tiered Devices** accordingly
  - Flash, High Performance Disk, Low Cost Capacity Disk, Online Lowest Cost Archival Disk/Tape

- **Balance Cost vs. Performance** of Flash
  - More Data in Flash > Higher Cache Hit Ratio > Improved Data Performance

- **Create and Auto-Manage Tiering** (Monitoring, Migrations, Placements) without manual intervention

Source: IMEX Research SSD Industry Report ©2011
SSS in DB Environments: Enabling new Economics

SAN Performance
Improvements using SSD

SAN TCO using HDD vs. Hybrid Storage

Source: IMEX Research SSD Industry Report ©2011
Industry Status

- Integrated Storage Tiering Products offered by over top 10 Storage Vendors
- **Major Storage Vendors**
  - Automated Volume Level Tiering (SSD & HDD)
- **New Storage Start Ups**
  - Integrated Flash Caching & Block Level Tiering
- **Cloud Vendors**
  - Adding Shared Cloud for Lowest Cost Backup & Restore Storage

Source: IMEX Research SSD Industry Report ©2011
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

© 2010-11 IMEX Research, Copying prohibited. All rights reserved.
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.

<table>
<thead>
<tr>
<th>SATA3 I/F</th>
<th>New Intelligent Controller (2nd Generation)</th>
<th>RS232,GPIO,I2C, JTAG I/F</th>
</tr>
</thead>
<tbody>
<tr>
<td>6Gb/s,32 NCQ</td>
<td>Optimized Write</td>
<td>NAND Flash I/F</td>
</tr>
<tr>
<td>PHY</td>
<td>Block Mgmt/ Wear Leveling</td>
<td>• Toggle, ONFI-2</td>
</tr>
<tr>
<td>Link</td>
<td>Garbage Collection</td>
<td>• SLC/ MLC/ eMLC</td>
</tr>
<tr>
<td>Transport</td>
<td>Read/Disturb Management</td>
<td>• 8ch/16 Byte Lanes</td>
</tr>
<tr>
<td>Command</td>
<td>RAID w/o Std. Parity OH</td>
<td>• 3x,2x nm Supp</td>
</tr>
<tr>
<td>AES 256/128 Encryption</td>
<td>TCG Compliance</td>
<td>• 512 GB Capable</td>
</tr>
<tr>
<td>AES 256/128 Encryption</td>
<td>TCG Compliance</td>
<td></td>
</tr>
<tr>
<td>Encryption</td>
<td>55b/512B BCH ECC</td>
<td></td>
</tr>
</tbody>
</table>

New Gen Controllers allow SSDs to meet Enterprise Class Availability/Performance/ over 5-Year Life/Scalability/ Auto-Configuration & Auto Data-Tiering

Source: SandForce

Source: IMEX Research SSD Industry Report ©2011
Key Takeaways

• **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

• **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**
  • **Data Forensics 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. Non-disruptively migrate hot data from HDD to SSDs.

• **Optimize Infrastructure to meet needs of Applications/SLA**
  • **Performance Economics/Benefits**
  • Typically 4-8% of data becomes a candidate and when migrated to SSDs can provide response time reduction of ~65% at peak loads

Source: IMEX Research SSD Industry Report ©2011