The material contained in this tutorial is copyrighted by the SNIA.

Member companies and individual members may use this material in presentations and literature under the following conditions:

- Any slide or slides used must be reproduced in their entirety without modification
- The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.

This presentation is a project of the SNIA Education Committee.

Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.

The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.
SSS Performance Benchmarking Learning Objectives

- Get a good understanding of the various parameters that influence the performance characteristics of SSDs
- Get a full understanding of the proposed SNIA Performance Measurement Specification
- Provide step-by-step guidance on how to set up a test benchmark that enables comparison among the various SSS devices
Definition of SSS

SSS = Solid State Storage

Traditional hard disk drive

Solid state hard drive

Apples to Apples, Pears to Pears in SSS Performance Benchmarking
© 2011 Storage Networking Industry Association. All Rights Reserved.
Flashy fists fly as OCZ and DDRdrive row over SSD performance

"dirty tricks used to drastically inflate IOPS results“…

"the egregious disparity between promised performance and reality”…

"This really isn’t an 'apples to apples' comparison”…

“you can't use an X1 dragster on the open road but you can use a Ferrari Vertex 2 EX”…

http://www.theregister.co.uk/2011/01/14/ocz_and_ddrdrive_performance_row/
The Performance Landscape
One Year Later

**MB/s or Mb/s?**

**Block Size?**

**Random or Sustained?**

**IOPS?**

**Random Precondition Sustained Speed?**

**Up to?**
### Market Segmentation

<table>
<thead>
<tr>
<th>Client SSD</th>
<th>Enterprise SSD</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="https://via.placeholder.com/150" alt="Micron" /></td>
<td><img src="https://via.placeholder.com/150" alt="SMART" /></td>
</tr>
<tr>
<td><img src="https://via.placeholder.com/150" alt="OCZ" /></td>
<td><img src="https://via.placeholder.com/150" alt="Stec" /></td>
</tr>
<tr>
<td><img src="https://via.placeholder.com/150" alt="Intel" /></td>
<td><img src="https://via.placeholder.com/150" alt="Plint" /></td>
</tr>
<tr>
<td><img src="https://via.placeholder.com/150" alt="Samsung" /></td>
<td><img src="https://via.placeholder.com/150" alt="Technology" /></td>
</tr>
</tbody>
</table>

- **Client SSD**
  - Low cost
  - C-MLC
  - 0-7% over provisioning
  - No backup power circuit
  - No Enterprise features
  - No customization
  - Warranty 1-3 yrs

- **Enterprise SSD**
  - Higher cost
  - E-MLC/SLC
  - 28-50% over provisioning
  - Backup power circuit
  - Enterprise features
  - Customization
  - Warranty 5 yrs
Performance Comparison
Enterprise vs. Client SSD

Sequential Read

- Enterprise SSD (SATA)
- Enterprise SSD (SAS)
- Client SSD 1 (SATA)
- Client SSD 2 (SATA)

Sequential Write

- Enterprise SSD (SATA)
- Enterprise SSD (SAS)
- Client SSD 1 (SATA)
- Client SSD 2 (SATA)

Random Read

- Enterprise SSD (SATA)
- Enterprise SSD (SAS)
- Client SSD 1 (SATA)
- Client SSD 2 (SATA)

Random Write

- Enterprise SSD (SATA)
- Enterprise SSD (SAS)
- Client SSD 1 (SATA)
- Client SSD 2 (SATA)
Variables influencing Performance

• **Platform**
  - Test Hardware (CPU, interface, chipset, etc)
  - Software (OS, drivers)

• **SSS Device Architecture**
  - Flash geometry, cache, flash management algorithm, etc
Variables influencing Performance

- **Platform**
  - Test Hardware (CPU, interface, chipset, etc)
  - Software (OS, drivers)
- **SSS Device Architecture**
  - Flash geometry, cache, flash management algorithm, etc
- **Workload**
  1. Write history & preconditioning: State of device before testing
The need for Preconditioning

Performance States for Various SSDs

- NM (MLC)
- NS (SLC)
- JS (SLC)
- PSM (MLC)
- JM (MLC)

Normalized IOPS (IOPS/Max(IOPS)) vs. Time (Minutes)

- FOB
- Transition
- Steady State (desirable test range)
4K Random to 128K Sequential Transition

- F.O.B. (~1hr)
- Random to Sequential Transition (~1.5hr)
- 4K Steady State
- 128K Steady State

IOPS vs. Time (Minutes)
128K Sequential to 4K Random Transition

- F.O.B. ~10 hrs
- 128K Steady State
- 4K Steady State
- Sequential to Random Transition
Variables influencing Performance

- **Platform**
  - Test Hardware (CPU, interface, chipset, etc)
  - Software (OS, drivers)
- **SSS Device Architecture**
  - Flash geometry, cache, flash management algorithm, etc
- **Workload**
  1. Write history & preconditioning: State of device before testing
  2. Workload pattern: Read/write mix, transfer size, sequential/random
Performance depends on
- Read/Write Mix
- Block Size
- Queue Depth (not shown)

Note: Shown 3D IOPS image courtesy of Calypso Systems
Variables influencing Performance

• Platform
  • Test Hardware (CPU, interface, chipset, etc)
  • Software (OS, drivers)
• SSS Device Architecture
  • Flash geometry, cache, flash management algorithm, etc
• Workload
  1. Write history & preconditioning: State of device before testing
  2. Workload pattern: Read/write mix, transfer size, sequential/random
  3. Data Pattern: The actual bits in the data payload written to the device
Dependency on data content - 2

3D IOPS Surface Profile (IOMETER 2008)

3D IOPS Surface Profile (IOMETER 2006)
## Benchmark Suites

<table>
<thead>
<tr>
<th>Test Suite</th>
<th>Client SSD</th>
<th>Enterprise SSD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PCMark</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>HDD Score, OS and application loading timing, user simulation (surfing web, windows media player, etc)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>SysMark</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>System-level test. Measures performance based on average response time, gives score (0-250)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>IOMeter</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Sequential/Random performance, workload simulation (file server, web server workload, etc)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>HDTach/H2benchw</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Performance stability, Sequential/Burst performance, Access Time</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>HD Tune</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance stability, Sequential/Burst performance, Access Time</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>Everest</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Random Access Time (Read/Write)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>VDBench</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workload generator, performance on DAS and NAS</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>Calypso CTS</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Device (RAW) level, direct IO synthetic stimulus generator for both client and enterprise</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
The Need for Industry Standardization!

- SNIA Technical Working Group (TWG)
- Specification for testing procedures to enable comparative testing of SSS performance
- Agnostic – Does not favor any one technology
- Relevant & Repeatable – Meaningful to end users
- Practical – Complete with reasonable time and effort
- Spec 1.0 to be released soon
- Future: Split of test specification into separate Enterprise and Client categories
1. **Prepare the Device**
   - Purge/Erase → put SSD back into “original” state

2. **Workload independent preconditioning**
   - Write data 2x capacity → bring device to known state

3. **Steady State Testing (includes workload based preconditioning)**
   - Run Test Loop up until steady state is achieved
   - Performance stays within ± 10% margin

4. **Test Report**
   - Steady state convergence
   - Steady State Verification
   - Performance measurement (2D/3D)
Preconditioning

- Preconditioning is key to get repeatable results
- Preconditioning needed to get drive in Steady State, after which performance can be measured
Preconditioning

• Preconditioning is key to get repeatable results
• Preconditioning needed to get drive in Steady State, after which performance can be measured
• Two types of preconditioning
  - Workload independent – write 2x capacity with 128KB sequential writes
  - Workload dependent – run workload itself until steady state is achieved
Steady State

- Measurement window is interval for last 5 measured rounds (i.e. test loops) that show steady state results
- Steady State is achieved if BOTH conditions are met:
  - Variation of y in measurement window is within 20% of average
  - Trending of y within measurement window is within 10% of average

![Graph showing Steady State Convergence Plot](image-url)
Performance Workloads & Tests

Client Test
- Random IOPS
  - 100/0, 95/5, 65/35, 50/50, 35/65, 5/95, 0/100
  - 1024K, 128K, 64K, 32K, 16K, 8K, 4K, 0.5K
- Sequential MB/s
  - 100/0, 0/100
  - 1024K
- Latency (random access)
  - 100/0, 65/35, 0/100
  - 8K, 4K, 0.5K

Enterprise Test
- Random IOPS
  - 100/0, 95/5, 65/35, 50/50, 35/65, 5/95, 0/100
  - 1024K, 128K, 64K, 32K, 16K, 8K, 4K, 0.5K
- Sequential MB/s
  - 100/0, 0/100
  - 1024K, 64K, 8K, 4K, 0.5K
- Latency (random access)
  - 100/0, 65/35, 0/100
  - 8K, 4K, 0.5K

Version 1.0 will focus further on differentiating Client vs. Enterprise
Performance Measurements 3D

Average Latency vs. Block Size and R/W Mix %
ActiveRange = (x,y); OIO/Thread = x; Thread Count = x; Data Pattern = x

Ave Latency (ms)

R/W Mix %
0/100 65/35 100/0

Block Size
0.5KiB 4KiB 8KiB

Performance Measurements 2D

Enterprise IOPS Test - Ave IOPS vs. Block Size & R/W Mix %
ActiveRange = (x,y); OIO/Thread = x; Thread Count = x; Data Pattern = x

IOPS

1 10 100 1,000 10,000 100,000

Block Size (KiB)
1 2 4 8 16 32 64 128 256 512 1,024
# Sample Test Report included

## Informative Annex A – Sample Test Report

### Summary Report Page

<table>
<thead>
<tr>
<th>Solid State Storage Performance Test Specification (PTS)</th>
<th>Rev. 0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device Under Test (DUT)</td>
<td>ABC Co.</td>
</tr>
<tr>
<td>Model No.:</td>
<td>ABC123</td>
</tr>
<tr>
<td>Form Factor:</td>
<td>2.5”</td>
</tr>
<tr>
<td>NAND Capacity:</td>
<td>256 GB MLC</td>
</tr>
<tr>
<td>DUT Interface:</td>
<td>SATAII, SAS HBA</td>
</tr>
</tbody>
</table>

### Test Specification:

**SNIA SSS TWG PTS v 0.9**

**SNIA**

**Apples to Apples, Pears to Pears**

© 2011 Storage Networking Industry Association. All Rights Reserved.

## Testing Summary: Tests Run

<table>
<thead>
<tr>
<th>Test</th>
<th>Preparation</th>
<th>Test Loop Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Client IOPS</td>
<td>Secure Erase 2x128K SEQ 100% QD 16 / TC 4 / OIO 64 Data Pattern RND</td>
</tr>
<tr>
<td>8.2</td>
<td>Client IOPS OPT – AR 10%</td>
<td>Secure Erase 2x128K SEQ 100% QD 16 / TC 4 / OIO 64 RND</td>
</tr>
<tr>
<td>8.3</td>
<td>Client IOPS OPT – Data</td>
<td>Secure Erase 2x128K SEQ 100% QD 16 / TC 4 / OIO 64 NON RND File as Data Pattern</td>
</tr>
</tbody>
</table>

### 9.1 Client Throughput

### 10.1 Client Latency

## General Device Description

<table>
<thead>
<tr>
<th>Device Under Test (DUT)</th>
<th>System Hardware Configuration</th>
<th>System Software Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model No.:</td>
<td>ABC123</td>
<td>Model No. RPT 2.0</td>
</tr>
<tr>
<td>Serial No.:</td>
<td>123xxx.H4F</td>
<td>Mother board Intel 3520HC</td>
</tr>
<tr>
<td>Semaphore File No.:</td>
<td>PMM/abc.123</td>
<td>Cache Intel SC3550DFP</td>
</tr>
<tr>
<td>User Capacity:</td>
<td>256 GB</td>
<td>CPU Type Intel 3 2.5GHz</td>
</tr>
<tr>
<td>Interface Speed:</td>
<td>6Gbs SATAII</td>
<td>Other SW Tool(s)</td>
</tr>
<tr>
<td>Form Factor:</td>
<td>2.5&quot;</td>
<td>DRAM Type Single</td>
</tr>
<tr>
<td>Media Type:</td>
<td>MLC</td>
<td>DDR Type 2133MHz DDR4</td>
</tr>
<tr>
<td>Major Features:</td>
<td>DUT EPP</td>
<td>ECC</td>
</tr>
<tr>
<td>NCQ:</td>
<td>YES</td>
<td>Other SW Tool(s)</td>
</tr>
<tr>
<td>Hot Plug:</td>
<td>YES</td>
<td>SAS HBA</td>
</tr>
<tr>
<td>SAS Sata Support:</td>
<td>NO</td>
<td>8Gbps 9212-4a4i</td>
</tr>
<tr>
<td>Other 1:</td>
<td>Boot HDD</td>
<td>160 GB 7200RPM</td>
</tr>
<tr>
<td>Other 2:</td>
<td>Optical Drive</td>
<td></td>
</tr>
</tbody>
</table>

## 8.1.1 Steady State Convergence Plot – All Block Sizes

- Pre-conditioning Convergence Report – All Rounds

## 8.1.2 Steady State Convergence Plot – 4K Block Sizes

- F1 Pre-conditioning Report – All Rounds

---

© 2011 Storage Networking Industry Association. All Rights Reserved.
Other Standardization Initiatives

• SSSI Group of SNIA
  • Technical Work Group (TWG) → Performance Benchmark Spec
  • Tech Dev Group → Performance Test Platform

• JEDEC 64.8
  • Specification for SSD endurance measurement

• SSDA
  • Testing of reliability (power cycling, data retention, endurance, etc) and OS compatibility (Windows 7)
Summary

• SSS Performance is dependent on many variables
• Comparing vendors is not trivial → industry standard required
• SNIA Performance Specs allows apples to apples comparison
  • Spec for review at http://www.snia.org/tech_activities/publicreview
  • Send your feedback to ssstwg@snia.org
Q&A / Feedback

Please send any questions or comments on this presentation to SNIA: tracksolidstate@snia.org

Many thanks to the following individuals for their contributions to this tutorial.
- SNIA Education Committee

David Landsman,
Easen Ho
Eden Kim
Neal Ekker,
Dan Le