Bring Your SSD Testing Up to Date

Tutorial – SNIA Performance Test Specifications

PTS 2.0.1 for Solid State Storage Devices
RWSW PTS 1.0.7 for Datacenter Storage

Eden Kim
Calypso Systems, Inc.
Chair, SNIA Solid State Storage Technical Working Group (SSS TWG)
Agenda

Introduction 1:00 – 1:15

Part 1: PTS v 2.0.1 - Updates 1:15 – 2:00
   Overview
   Updated Tests & Metrics

Break 2:00 – 2:15

Part 2: RWSW PTS v 1.0.7 - Capture & Analysis 2:15 – 3:30
   Overview - Real World Workloads
   Capture – IO Captures & Reference Workloads
   Analysis – IO Stream Map & LBA Range Hit Map

Break 3:30 – 3:45

Part 3: RWSW PTS 1.0.7 - Test 3:45 – 4:45
   In-situ Target Server Self-Test
   Replay Native
   Multi-WSAT
   Individual Streams WSAT
   RWSW DIRTH

Questions 4:45 – 5:00

Flash Memory Summit 2018
Santa Clara, CA
Introduction

- Two New SNIA SSS PTS Technical Positions
  - “PTS” - Performance Test Specifications for SSDs & Datacenter Storage
    - PTS v 2.0.1 Feb. 2018 - for SSDs
    - RWSW PTS v 1.0.7 May 2018 - for Datacenter Storage

- PTS v 2.0.1 – for Solid State Storage Devices
  - Focus is on the pre-conditioning of NAND Flash to Steady State
  - Workloads are corner case stress tests and synthetic emulations of applications
  - Tests are for Comparative Performance benchmarking

- RWSW PTS 1.0.7 – for Datacenter Storage
  - Focus is on the Capture, Analysis and Test of Real World Workloads
  - Real World Workloads are Derived from IO Trace Captures of User Applications
  - Tests are for the Analysis, Qualification & Validation of Datacenter Storage
Part 1: PTS v 2.0.1 for SSDs - Overview

- Consolidates PTS Client v1.2 and Enterprise v1.1 into 1 doc
- Updates IOPS Test to 3 Steady State Variables
- Adds WSAT, DIRTH and HIR to Client Tests
- Defines 6 synthetic workloads for WSAT & DIRTH:
  a. Meta Data & Journaling: SEQ 0.5K RW
  b. Write Intensive: RND 4K W
  c. Read Intensive: RND 4K R
  d. Database OLTP: RND 8K RW65
  e. VOD: SEQ 128K RW90
  f. Composite Block Size (CBS): JEDEC 219(a) Composite workload
- Adds Optional Secondary Metrics:
  a. Power consumption; IOPS/Watt
  b. Average/Maximum RTs and 5 9s Response Time Quality of Service
  c. CPU System Usage %; CPU IO Wait
## PTS v2.0.1 – Tests & Test Flows

<table>
<thead>
<tr>
<th>Test Family</th>
<th>Tests</th>
<th>Purpose</th>
<th>Workload Type</th>
<th>Pre-condition</th>
<th>Steady State</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNIA Basic</td>
<td>IOPS</td>
<td>IO Rate Bandwidth (BW)</td>
<td>Mixed BS Loop</td>
<td>PURGE, WIPC, WDPC</td>
<td>20/10 5 Round Post Test Inspection</td>
</tr>
<tr>
<td></td>
<td>TP 128K, TP 1024K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAT</td>
<td>Single IO Response Time (RT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturation</td>
<td>WSAT</td>
<td>FOB IO, BW &amp; RT Saturation</td>
<td>Single Access Patterns</td>
<td>PURGE, WDPC</td>
<td>Time, TGBW, Drive Fills Rounds with BTW</td>
</tr>
<tr>
<td>Optimized</td>
<td>Dirth</td>
<td>IO, BW &amp; RT Saturation Quality of Service Performance over Range of Users</td>
<td>Single Access Patterns</td>
<td>PURGE, WIPC, WDPC</td>
<td>Rounds with BTW</td>
</tr>
<tr>
<td>Sustained</td>
<td>XSR</td>
<td>Garbage Collection Recovery</td>
<td>Alternating Segments</td>
<td>PURGE, WDPC</td>
<td>Time Rounds with BTW</td>
</tr>
<tr>
<td></td>
<td>HIR</td>
<td>Garbage Collection Recovery</td>
<td>Alternating Segments</td>
<td>PURGE, WDPC</td>
<td></td>
</tr>
</tbody>
</table>

- **Tests have Different Workload Types**
  - Mixed Block Size Loops
  - Single Access Pattern
  - Composite BS Mix
  - Alternating Segments

- **Tests use different PC**
  - WIPC & WDPC
  - WDPC only

- **Tests have Different SS Criteria**
  - 20/10 % Round Formula
  - 5 Round with Between Round Writes
  - 25 Rounds w/ Post Test Inspection
  - Time, TGBW, Drive Fills
<table>
<thead>
<tr>
<th>Test</th>
<th>Purpose</th>
<th>Workload Type</th>
<th>Workload Access Patterns</th>
<th>Demand Intensity</th>
</tr>
</thead>
</table>
| IOPS       | IO Rate                          | Mixed BS Loop 56 RW elements               | RW: 100R; 95:05; 65:35; 50:50; 35:65; 05:95; 100W BS: 1024K; 128K; 64K; 32K; 16K; 8K; 4K; 0.5K | Ent: T4Q32  
Client: T2Q16 |
| TP 128K; TP 1024K | IO, BW & RT Saturation | Single Access Patterns 2 RW elements | RW: 100R; 100W BS: [SEQ 128K; SEQ 1024K] | T1Q32 |
| LAT        | IO, BW & RT Quality of Service   | Single Access Patterns 9 RW elements        | RW: 100R; 65:35 RW; 100W BS: 8K; 4K; 0.5K | T1Q1 |
| WSAT       | FOB IO, BW & RT Saturation       | Single Access Patterns Composite Block Size | RND 4K R; RND 4K W; RND 8K RW65  
SEQ 128K RW90; SEQ 0.5K RW50  
CBS 14 BS Composite (JEDEC 219(a)) | Ent RND: T4Q32  
Client RND: T2Q16  
SEQ: T1Q32 |
| DIRTH      | IO, BW & RT Saturation RT QoS    | Single Access Patterns Composite Block Size | RND 4K R; RND 4K W; RND 8K RW65  
SEQ 128K RW90; SEQ 0.5K RW50  
CBS 14 BS Composite (JEDEC 219(a)) | Ent RND: T4Q32  
Client RND: T2Q16 |
| XSR        | RT & BW Recovery from Super Saturation | Alternating Segments | SEQ 1024KW; RND 8K W; SEQ 1024K W | T1Q32; T4Q32; T1Q32 |
| HIR        | RT & BW Recovery Effects of Host Idles | RND 4K W Varying Host Idles | RND 4K W; 0,5,10,15,15 & 50 sec Host Idles | Ent: T4Q32  
Client: T2Q16 |
### Basic Test Process:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>PURGE</td>
</tr>
<tr>
<td>2.</td>
<td>Workload Independent Pre-conditioning</td>
</tr>
<tr>
<td>3.</td>
<td>Workload Dependent Pre-conditioning</td>
</tr>
<tr>
<td>4.</td>
<td>Types of Workloads</td>
</tr>
<tr>
<td>5.</td>
<td>Test Settings</td>
</tr>
<tr>
<td>6.</td>
<td>Steady State Determination</td>
</tr>
<tr>
<td>7.</td>
<td>Report Data from Steady State Window</td>
</tr>
</tbody>
</table>

#### 1. PURGE¹

- Return the Device to a state as if no writes have occurred by:
  - a) ATA: SECURITY ERASE
  - b) SCSI: FORMAT UNIT
  - c) NVMe: FORMAT namespace or
  - d) Vendor Specific Command / Methods

- Device PURGE creates a known and repeatable test starting point and facilitates a clear demonstration of the Steady State convergence behavior

- PURGE is a Required PTS 2.0.1 test step

¹ PTS 2.0.1 – p 21
PTS v2.0.1 – Basic Test Process: WIPC

2. Workload Independent Pre-conditioning (WIPC)

- Apply a workload to facilitate convergence to Steady State
  a) Write twice the stated User Capacity
  b) Apply SEQ 128K W at T1 Q32 or
  c) Apply SEQ 1024K W at T1 Q32
  d) WIPC is ‘independent’ of the workload of interest (‘dependent’)

- WIPC is applied to facilitate convergence to Steady State by organizing LBA tables prior to application of WDPC

- NOTE: WIPC and/or WDPC by themselves do not put an SSD into a Steady State but prepares and applies the workload of interest for calculation of the Steady State determination criteria
### Basic Test Process:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>PURGE</td>
</tr>
<tr>
<td>2.</td>
<td>Workload Independent Pre-conditioning</td>
</tr>
<tr>
<td>3.</td>
<td>Workload Dependent Pre-conditioning</td>
</tr>
<tr>
<td>4.</td>
<td>Types of Workloads</td>
</tr>
<tr>
<td>5.</td>
<td>Test Settings</td>
</tr>
<tr>
<td>6.</td>
<td>Steady State Determination</td>
</tr>
<tr>
<td>7.</td>
<td>Report Data from Steady State Window</td>
</tr>
</tbody>
</table>

#### 3. Workload Dependent Pre-conditioning (WIPC)

- Apply the Workload of Interest until convergence to Steady State
- WDPC applies the Workload of Interest for calculation of the Steady State determination criteria
- Steady State determination criteria depend on the type of workload
There are three types of PTS 2.0.1 WDPC Workloads:

a) Mixed Block Size Loop (IOPS, TP, LAT)
b) Single Access Pattern (WSAT, DIRTH, XSR, HIR)
c) Fixed Composite (CBW WSAT, CBW DIRTH)
PTS v2.0.1 – Types of Workloads: Mixed BS Loops

- **Mixed Block Size Loops – Nested RW & BS loops**
  - **IOPS Test**: 56 one minute RW/BS mix elements
    - a) (8) Read Write Mixes: 100% R, 95:05, 65:35, 50:50, 35:65, 05:95, 100% W
    - b) (7) Block Sizes: 1024K, 128K, 64K, 32K, 16K, 8K, 4K, 0.5K
  - **TP (Throughput) Test**: 2 one minute RW/BS mix elements
    - a) (2) Read Write Mixes: 100% R, 100% W
    - b) (1) Block Size: [1024K TP test, 128K TP Test]
  - **LAT (Latency) Test**: 9 one minute RW/BS mix elements
    - a) (3) Read Write Mixes: 100% R, 65:35 RW, 100% W
    - b) (3) Block Sizes: 8K, 4K, 0.5K

---

**Basic Test Process:**

1. **PURGE**
2. **Workload Independent Pre-conditioning**
3. **Workload Dependent Pre-conditioning**
4. **Types of Workloads**
5. **Test Settings**
6. **Steady State Determination**
7. **Report Data from Steady State Window**
IOPS Test: One Round = 56 one minute RW/BS

- (8) Read Write Mixes: 100% R, 95:05, 65:35, 50:50, 35:65, 05:95, 100% W
- (7) Block Sizes: 1024K, 128K, 64K, 32K, 16K, 8K, 4K, 0.5K

Basic Test Process:

1. PURGE
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window
PTS v2.0.1 – Types of Workloads: Mixed BS Loops

- TP (Throughput) Test: One Round = 2 one minute RW/BS
  a) (2) Read Write Mixes: 100% R, 100% W
  b) (1) Block Size: [1024K TP test, 128K TP Test]
PTS v2.0.1 – Types of Workloads: Mixed BS Loops

- LAT (Latency) Test: One Round = 9 one minute RW/BS\(^7\)
  - a) (3) Read Write Mixes: 100% R, 65:35 RW, 100% W
  - b) (3) Block Sizes: 8K, 4K, 0.5K

Basic Test Process:
1. PURGE
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window
PTS v2.0.1 – Types of Workloads: Single Access Pattern

- **Saturation** & **DIRTH** Single Access Pattern
  - a) Journal/Metadata: SEQ 0.5K R/W
  - b) Write Intensive: RND 4K W
  - c) Read Intensive: RND 4K R
  - d) Mixed/OLTP: RND 8K RW65
  - e) VOD: SEQ 128K RW90

---

**Basic Test Process:**

1. **PURGE**
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window
PTS v2.0.1 – Types of Workloads: Single Access Pattern


Basic Test Process:

1. PURGE
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window
PTS v2.0.1 – Types of Workloads: Alternating Segment

- XSR (Cross Stimulus Recovery): Alternating Segments

Basic Test Process:
1. PURGE
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window

SEQ 1024K W / RND 8K W / SEQ 1024K W

![Graph showing P2 TP vs Time]
Basic Test Process:

1. PURGE
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window

- Saturation\textsuperscript{12} & DIRTH\textsuperscript{13}: Composite Block Size – JEDEC 219(a)

- Composite of 12 different RND Write Block Sizes

<table>
<thead>
<tr>
<th>Block Size (KB)</th>
<th>Percent</th>
<th>Block Size (KB)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>4%</td>
<td>3.5</td>
<td>1%</td>
</tr>
<tr>
<td>1.0</td>
<td>1%</td>
<td>4.0</td>
<td>67%</td>
</tr>
<tr>
<td>1.5</td>
<td>1%</td>
<td>8.0</td>
<td>10%</td>
</tr>
<tr>
<td>2.0</td>
<td>1%</td>
<td>16.0</td>
<td>7%</td>
</tr>
<tr>
<td>2.5</td>
<td>1%</td>
<td>32.0</td>
<td>3%</td>
</tr>
<tr>
<td>3.0</td>
<td>1%</td>
<td>64.0</td>
<td>3%</td>
</tr>
</tbody>
</table>

Use of Restricted LBA Ranges

<table>
<thead>
<tr>
<th>LBA Group</th>
<th>Percentage of IOs</th>
<th>Zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>50%</td>
<td>To the first 5% LBAs</td>
</tr>
<tr>
<td>Group B</td>
<td>30%</td>
<td>To the next 15% LBAs</td>
</tr>
<tr>
<td>Group C</td>
<td>20%</td>
<td>To the remaining 20% LBAs</td>
</tr>
</tbody>
</table>
### PTS v2.0.1 – Basic Test Process: Test Settings

<table>
<thead>
<tr>
<th>Basic Test Process:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PURGE</td>
</tr>
<tr>
<td>2. Workload Independent Pre-conditioning</td>
</tr>
<tr>
<td>3. Workload Dependent Pre-conditioning</td>
</tr>
<tr>
<td>4. Types of Workloads</td>
</tr>
<tr>
<td>5. Test Settings</td>
</tr>
<tr>
<td>6. Steady State Determination</td>
</tr>
<tr>
<td>7. Report Data from Steady State Window</td>
</tr>
</tbody>
</table>

- **Test Settings are set forth for Enterprise (Ent) and Client tests**
- **Test Settings = Required or User Choice. Must be Disclosed**
- **Common Test Settings include:**
  - a) LBA Active Range: Ent = AR100; Client = AR75
  - b) Volatile Write Cache: Ent = WCD; Client = WCE
  - c) Binary Data Pattern: RND; Repeating; Binary File

- **TC/QD Setting (Demand Intensity) for Specific Tests include:**
  - a) RND access workloads (e.g. IOPS, WSAT, HIR): Ent = T4Q32, Client = T2Q16
  - b) SEQ access workloads (e.g. TP, XSR): T1Q32
  - c) Single IO Response Time Measurement (e.g. LAT): T1Q1
  - d) IO & RT Saturation (e.g. DIRTH): Outstanding IO range from T1Q1 to T32Q32
## Basic Test Process:

1. **PURGE**
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings

<table>
<thead>
<tr>
<th>Steady State Determination</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Steady State Determination</td>
</tr>
</tbody>
</table>

| 7. Report Data from Steady State Window |

<p>| 1. 20% Data Excursion/10% Slope 5 Round Formula(^{14}) |
| 2. 25 Round / Post Process Inspection(^{15}) |
| 3. 5 Round / 30 min Between Round Writes (BTW)(^{16}) |</p>
<table>
<thead>
<tr>
<th>4. Pseudo Steady State(^{17})</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Time</td>
</tr>
<tr>
<td>b) Total GB Written</td>
</tr>
<tr>
<td>c) Drive Fills</td>
</tr>
</tbody>
</table>

- All Reported Data Must from the 5 Round SS Range

---

\(^{14,15,16,17}\) PTS 2.0.1 – p 24,24,50,50
Steady State: IOPS Test – Mixed Loop BS; 5 Round 20/10 SS
**PTS v2.0.1 – “20/10” 5 Round Steady State Determination**

- **Basic Test Process:**
  1. PURGE
  2. Workload Independent Pre-conditioning
  3. Workload Dependent Pre-conditioning
  4. Types of Workloads
  5. Test Settings
  6. Steady State Determination
  7. Report Data from Steady State Window

- **“20/10” 5 Round Mixed Loop Steady State Determination**

- **(3) IOPS Steady State (SS) Determination Variables:**
  a) RND 4K W
  b) RND 64K RW65
  c) RND 1024K RW100

- **Steady State Formula**
  a) Five Consecutive Rounds
  b) 20% Data Excursion, 10% Slope Best Linear Fit Line

- **All SS Variables Must be within 5 Round Range**
- **All Reported Data Must come from the 5 Round SS Range**
PTS v2.0.1 – IOPS Mixed BS Loop SS Determination – SS Rounds 2-6

RND 4K W
SS Rounds 1-5

RND 64K RW65
SS Rounds 2-6

RND 1024K
RW100
SS Rounds 1-5

Simultaneous SS in Rounds 2-6
(3) Variable 20/10 (5) Round SS
RND 64K RW65 Round 1 Anomaly
Basic Test Process:

1. **PURGE**
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. **Steady State Determination**
7. Report Data from Steady State Window

- **Single Access Pattern “5 Round/30 Min BTW” SS Determination**
- **1 Min Rounds with 30 Min Between Round Writes (BTW):**
  a) 1 Min Measurement Rounds
  b) 30 Min Between Round Writes
  c) Run Rounds until 5 Round “20/10” SS Determination is Met

- **Single Access Pattern SS can be used for:**
  a) DIRT (Demand Intensity Response Time Histogram)
  b) HIR (Host Idle Recovery)
  c) WSAT (Write Saturation)

- **All Reported Data Must come from the 5 Round SS Range**
PTS v2.0.1 – Single Access Pattern Steady State Determination

16 (1) Min Rounds separated by 30 Min Between Round Writes (BTW)

RND 4K W
16 (1) Min Rounds with 30 Min BTW
480 Min Total Test Time
Basic Test Process:

1. PURGE
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window

- Pseudo Steady State is appropriate when there are no post SS steps
  a. Appropriate for WSAT, IOPS, TP & LAT
  b. Post SS Rounds saturate and affect Response Times (DIRTH)

- Pseudo Steady State can be by:
  a) Time (WSAT)
  b) Total GB Written (WSAT)
  c) Drive Fills (WSAT)

- 25 Round Post Process determination:
  a) Run 25 Rounds
  b) Post process inspection for Steady State (IOPS)
Basic Test Process:

1. **PURGE**
2. Workload Independent Pre-conditioning
3. Workload Dependent Pre-conditioning
4. Types of Workloads
5. Test Settings
6. Steady State Determination
7. Report Data from Steady State Window

---

**PTS v2.0.1 – Steady State Reporting Data**

- Updated v2.0.1 Reporting: Primary & Secondary Metrics
- Required - Primary Metrics include:
  - a) IOPS, TP & Response Times per individual test requirements
- Optional - Secondary Metrics include:
  - a) Average/Maximum Response Times, 5 9s Response Time Quality of Service
  - b) Power consumption; IOPS/Watt; Internal Temperature
  - c) CPU System Usage; CPU IO Wait
- Secondary Metrics - may become Required in Future PTS Revisions
- See PTS 2.0.1 SNIA Report Format
### SNIA Report Format:

1. PTS 2.0.1 Header
2. RND 4K W SS
3. RND 64K RW65 SS
4. RND 1024K R SS
5. IOPS Table
6. Secondary Metrics

---

### PTS v2.0.1 – SNIA PTS 2.0.1 Report Header

#### IOPS Test (REQUIRED) - Report Page

<table>
<thead>
<tr>
<th>Test Run Date:</th>
<th>01/19/16 08:11 PM</th>
<th>Report Run Date:</th>
<th>11/07/16 11:17 AM</th>
</tr>
</thead>
</table>

**IOPS - Block Size x RW Mix Matrix**

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Mfg A</th>
<th>SSD Model</th>
<th>SSD A - NVMe SSD</th>
<th>TEST SPONSOR</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Test Platform</th>
<th>Device Under Test</th>
<th>Set Up Parameters</th>
<th>Test Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calypso RTP 3.0</td>
<td>Mfg A</td>
<td>Data Pattern</td>
<td>RND_ONCE</td>
</tr>
<tr>
<td>Motherboard 1</td>
<td>Model No.</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>CPU</td>
<td>S/N</td>
<td>AR Segments</td>
<td>Test Stimulus 1</td>
</tr>
<tr>
<td>Memory</td>
<td>Capacity</td>
<td>Pre Condition 1</td>
<td>RW Mix</td>
</tr>
<tr>
<td>Operating System</td>
<td>Firmware ver</td>
<td>SEQ 1280B W</td>
<td></td>
</tr>
<tr>
<td>Test SW</td>
<td>Interface</td>
<td>TC 1 / QD 32</td>
<td></td>
</tr>
<tr>
<td>Test SW Info</td>
<td>NAND Type</td>
<td>TC8 / QD 32</td>
<td></td>
</tr>
<tr>
<td>Test ID No.</td>
<td>PCIe NVMe</td>
<td>Steady State 3 - 7</td>
<td></td>
</tr>
<tr>
<td>HBA</td>
<td>Purge Method</td>
<td>Test Stimulus 2</td>
<td></td>
</tr>
<tr>
<td>PCIe</td>
<td>Gen 3</td>
<td>Note</td>
<td>Steady State</td>
</tr>
</tbody>
</table>
PTS v2.0.1 – RND 4K W Steady State

SNIA Report Format:
1. PTS 2.0.1 Header
2. RND 4K W SS
3. RND 64K RW65 SS
4. RND 1024K R SS
5. IOPS Table
6. Secondary Metrics

IOPS Steady State Measurement Window – RND/4KiB RW0

Steady State Determination Data

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average IOPS</td>
<td>61969.4</td>
</tr>
<tr>
<td>Allowed Maximum Data Excursion</td>
<td>12393.2</td>
</tr>
<tr>
<td>Measured Maximum Data Excursion</td>
<td></td>
</tr>
<tr>
<td>Allowed Maximum Slope Excursion</td>
<td>6196.4</td>
</tr>
<tr>
<td>Measured Maximum Slope Excursion</td>
<td></td>
</tr>
<tr>
<td>Least Squares Linear Fit Formula</td>
<td>-323.966 * R + 63389.403</td>
</tr>
</tbody>
</table>
PTS v2.0.1 – RND 64K RW65 Steady State

SNIA Report Format:
1. PTS 2.0.1 Header
2. RND 4K W SS
3. RND 64K RW65 SS
4. RND 1024K R SS
5. IOPS Table
6. Secondary Metrics

IOPS Steady State Measurement Window – RND/64KiB RW65

Steady State Determination Data
- Average IOPS: 9070.6
- Allowed Maximum Data Excursion: 9978.0
- Measured Maximum Data Excursion: 1511.3
- Allowed Maximum Slope Excursion: 6196.0
- Measured Maximum Slope Excursion: 1296.0
- Least Squares Linear Fit Formula: -323.096 * R + 63585.403
PTS v2.0.1 – RND 1024K R Steady State

### SNIA Report Format:
1. PTS 2.0.1 Header
2. RND 4K W SS
3. RND 64K RW65 SS
4. RND 1024K R SS
5. IOPS Table
6. Secondary Metrics

### IOPS Steady State Measurement Window – RND/1024KiB RW100

<table>
<thead>
<tr>
<th>Round</th>
<th>IOPS</th>
<th>Average</th>
<th>110%*Average</th>
<th>90%*Average</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Steady State Determination Data
- Average IOPS: 1571.6
- Allowed Maximum Data Excursion: 1726.7
- Measured Maximum Data Excursion: 1511.3
- Allowed Maximum Slope Excursion: 6196.0
- Measured Maximum Slope Excursion: 1256.0
- Least Squares Linear Fit Formula: -323.996 * R + 63589.403
### SNIA Report Format:

1. PTS 2.0.1 Header
2. RND 4K W SS
3. RND 64K RW65 SS
4. RND 1024K R SS
5. IOPS Table
6. Secondary Metrics

#### PTS v2.0.1 – IOPS Table

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Mfg A</th>
<th>SSD Model</th>
<th>SSD A - NVMe SSD</th>
<th>TEST SPONSOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motherboard</td>
<td>CALYPSO R2.0</td>
<td>A200</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel Xeon E5-2640 V4</td>
<td>2 / 12 / 24</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>Memory</td>
<td>32.0 GB</td>
<td>2400</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>Operating System</td>
<td>CentOS 7.3</td>
<td>2000</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>Test SW</td>
<td>CTS 5.0</td>
<td>2000</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>NVM Flash</td>
<td>MLC</td>
<td>2000</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>Test SW Info</td>
<td>120GB</td>
<td>320GB</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>Test 3D No.</td>
<td>B84-904</td>
<td>U2</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>HBA</td>
<td>SAS-12Gbps</td>
<td>MLC</td>
<td>AR</td>
<td>100%</td>
</tr>
<tr>
<td>PCIe</td>
<td>Gen 3</td>
<td>64GB</td>
<td>AR</td>
<td>100%</td>
</tr>
</tbody>
</table>

### IOPS - ALL RW Mix & BS - Tabular Data

<table>
<thead>
<tr>
<th>Block Size (KB)</th>
<th>0/100</th>
<th>5/95</th>
<th>35/65</th>
<th>50/50</th>
<th>65/35</th>
<th>95/5</th>
<th>95/5</th>
<th>100/0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>47,139.1</td>
<td>48,885.0</td>
<td>70,740.4</td>
<td>89,876.2</td>
<td>118,331.1</td>
<td>338,953.4</td>
<td>499,097.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>61,969.4</td>
<td>65,324.8</td>
<td>90,078.8</td>
<td>108,316.3</td>
<td>104,022.5</td>
<td>310,634.6</td>
<td>407,395.3</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>31,086.5</td>
<td>32,730.9</td>
<td>44,174.2</td>
<td>43,065.4</td>
<td>46,758.7</td>
<td>171,091.2</td>
<td>207,067.5</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>15,612.4</td>
<td>16,471.3</td>
<td>22,832.9</td>
<td>24,658.0</td>
<td>37,991.9</td>
<td>89,209.8</td>
<td>104,000.5</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>7,813.4</td>
<td>8,237.5</td>
<td>11,537.4</td>
<td>14,568.4</td>
<td>18,790.9</td>
<td>46,002.2</td>
<td>52,149.3</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>3,904.7</td>
<td>4,126.0</td>
<td>5,739.5</td>
<td>7,519.0</td>
<td>9,071.1</td>
<td>22,856.6</td>
<td>26,054.5</td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>1,964.8</td>
<td>2,065.9</td>
<td>2,843.5</td>
<td>3,725.5</td>
<td>4,465.9</td>
<td>10,992.1</td>
<td>12,948.0</td>
<td></td>
</tr>
<tr>
<td>1024</td>
<td>248.6</td>
<td>262.2</td>
<td>360.8</td>
<td>469.6</td>
<td>636.5</td>
<td>1,659.1</td>
<td>1,571.6</td>
<td></td>
</tr>
</tbody>
</table>

---

Flash Memory Summit 2018
Santa Clara, CA
**Secondary Metrics: IOPS, Ave Power mW, Temperature**

<table>
<thead>
<tr>
<th>Data</th>
<th>Block Size</th>
<th>RW 0</th>
<th>RW 5</th>
<th>RW100</th>
<th>Block Size</th>
<th>RW 0</th>
<th>RW 5</th>
<th>RW100</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IOPS</strong></td>
<td>0.5 KB</td>
<td>20,392</td>
<td>47,096</td>
<td>121,431</td>
<td>32 KB</td>
<td>3,049</td>
<td>6,560</td>
<td>14,606</td>
</tr>
<tr>
<td></td>
<td>4 KB</td>
<td>24,163</td>
<td>49,072</td>
<td>73,714</td>
<td>64 KB</td>
<td>1,525</td>
<td>3,211</td>
<td>7,620</td>
</tr>
<tr>
<td></td>
<td>8 KB</td>
<td>12,193</td>
<td>26,032</td>
<td>47,470</td>
<td>128 KB</td>
<td>760</td>
<td>1,501</td>
<td>4,057</td>
</tr>
<tr>
<td></td>
<td>16 KB</td>
<td>6,112</td>
<td>13,410</td>
<td>26,091</td>
<td>1024 KB</td>
<td>94</td>
<td>203</td>
<td>518</td>
</tr>
<tr>
<td><strong>Avg Power mW</strong></td>
<td>0.5 KB</td>
<td>3,752</td>
<td>3,568</td>
<td>2,346</td>
<td>32 KB</td>
<td>3,805</td>
<td>3,702</td>
<td>2,794</td>
</tr>
<tr>
<td></td>
<td>4 KB</td>
<td>3,767</td>
<td>3,568</td>
<td>2,336</td>
<td>64 KB</td>
<td>3,807</td>
<td>3,632</td>
<td>2,679</td>
</tr>
<tr>
<td></td>
<td>8 KB</td>
<td>3,804</td>
<td>3,734</td>
<td>2,580</td>
<td>128 KB</td>
<td>3,806</td>
<td>3,576</td>
<td>2,652</td>
</tr>
<tr>
<td></td>
<td>16 KB</td>
<td>3,809</td>
<td>3,745</td>
<td>2,691</td>
<td>1024 KB</td>
<td>3,805</td>
<td>3,573</td>
<td>3,023</td>
</tr>
<tr>
<td><strong>Temp. Celsius</strong></td>
<td>0.5 KB</td>
<td>39</td>
<td>38</td>
<td>32</td>
<td>32 KB</td>
<td>39</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>4 KB</td>
<td>39</td>
<td>39</td>
<td>32</td>
<td>64 KB</td>
<td>39</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>8 KB</td>
<td>39</td>
<td>39</td>
<td>33</td>
<td>128 KB</td>
<td>38</td>
<td>58</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>16 KB</td>
<td>39</td>
<td>39</td>
<td>34</td>
<td>1024 KB</td>
<td>38</td>
<td>38</td>
<td>37</td>
</tr>
</tbody>
</table>

- Secondary Metrics are Optional
- Summary Table: 3 RW Mixes x 8 BS
- IOPS; Average Power mW; Internal SSD reported Temp Celsius
**Secondary Metrics**: Average Response Time, 59s RT, Max RT

---

**SNIA Report Format:**

1. PTS 2.0.1 Header
2. RND 4K W SS
3. RND 64K RW65 SS
4. RND 1024K R SS
5. IOPS Table
6. Secondary Metrics

---

**Secondary Metrics are Optional**

- Summary Table: 3 RW Mixes x 8 BS
- Average Response Time, 59s RT, Maximum RT in mS
BREAK

2:00 – 2:15
Part 2: RWSW PTS v 1.0.7 - Overview

- RWSW PTS v1.0.7 describes methodologies for Real World IO Capture, Analysis & Test for Datacenter Storage
- IO Captures present IO Streams & Metrics that actually occur during real world application usage – not synthetic emulations
- RWSWs show that:
  a. Server Storage Performance depends, in large part, on the workload
  b. RWSWs are constantly changing combinations of many IO Streams and QDs
  c. IO Streams change as they traverse the HW/SW stack
  d. RWSW Replays can improve Failure Analysis & enable Storage Qualification
- RWSW PTS defines (4) RWSW Storage Tests designed to test and qualify storage to SSSI Reference and/or User captured RWSWs
Overview of RWSWs:

1. What are Real World Storage Workloads?
2. Data Center SQL Server RWSW
3. Why are RWSWs Important?

Real World Storage Workloads (RWSWs) are:

- IO Streams generated by Applications that traverse the SW stack from User space to Data Center Storage and back
- IO Streams that present to Storage at the File System, Block IO or other specified software layer
- IO Streams are modified at each layer of software abstraction by coalescing, fragmentation, appending & merging
- Data Center storage includes: SAN, NAS, DAS, JBOF, JBOD, SDS, Open, Virtualized, Object, LUN and SSD

\[19 \text{ RSW PTS v1.0.7 – page 12} \]
Data Center 24 Hour RWSW - Retail Web Portal

Real World SQL from Retail Web Portal - IO Streams, Queue Depths & IO's

Real World Workloads:
- Constantly changing combinations of IO Streams & range of QDs
- IO Streams and QDs change with Time, Events and Processes
- IO Rates are throttled by real world Applications and Users

20 SSSI Reference Capture No. 3
Why are RWSWs Important?

Understanding RWSWs Makes a Difference:

- Know what, where & when IO Streams actually occur
- Observe in-situ performance during the IO Capture process
- Evaluate RWSWs for:
  - software optimization
  - dev ops
  - load balancing
  - Interoperability
  - failure analysis replay
  - storage server & SSD qualification

![Workload Streams (by frequency)](image)

![LBA Range Hits (by frequency)](image)
An IO Capture is:  

- The tabulation of IO Stream statistics, observed at a given level in the SW Stack, during the IO Trace Capture period.
- Comprised of tables of IO statistics and metrics in binary data form – no private or personal data is captured.
- Derived from continuous IO Trace data that is parsed into Steps for visualization, analytics & playback.
- Comprised of IO Capture Steps that allow for flexibility in visualization granularity, capture length & file size.
What is an IO Stream?

- An IO Stream is a distinct:
  - Read or Write IO (Input / Output) Operation
  - RND or SEQ Access
  - Data Transfer Size (Block Size)

- A single IO Stream can occur many times during a single IO Capture step

- Other metrics and data are associated with IO Streams (such as Response Times, Process IDs, Queue Depths)

### IO Stream Table: 2 Minute Capture Step showing IO Stream Statistics

<table>
<thead>
<tr>
<th>Access Pattern</th>
<th>RND or SEQ</th>
<th>Block Size</th>
<th>Read/Write</th>
<th>Queue Depth Ave/Max</th>
<th>% Occurrence</th>
<th>Quantity (IOs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEQ 15K W</td>
<td>SEQ</td>
<td>1536</td>
<td>W</td>
<td>1/1/11</td>
<td>1.34</td>
<td>69</td>
</tr>
<tr>
<td>SEQ 1K W</td>
<td>SEQ</td>
<td>1024</td>
<td>W</td>
<td>1/1/11</td>
<td>4.32</td>
<td>223</td>
</tr>
<tr>
<td>SEQ 0.5W</td>
<td>SEQ</td>
<td>512</td>
<td>W</td>
<td>1/1/11</td>
<td>9.24</td>
<td>477</td>
</tr>
<tr>
<td>SEQ 4K W</td>
<td>SEQ</td>
<td>4096</td>
<td>W</td>
<td>1/1/11</td>
<td>22.31</td>
<td>1152</td>
</tr>
<tr>
<td>SEQ 16K W</td>
<td>SEQ</td>
<td>16384</td>
<td>W</td>
<td>1/1/11</td>
<td>14.25</td>
<td>736</td>
</tr>
<tr>
<td>RND 4K W</td>
<td>RND</td>
<td>4096</td>
<td>W</td>
<td>1/1/11</td>
<td>9.8</td>
<td>506</td>
</tr>
<tr>
<td>RND 3.5K W</td>
<td>RND</td>
<td>3584</td>
<td>W</td>
<td>1/1/11</td>
<td>0.62</td>
<td>32</td>
</tr>
<tr>
<td>RND 3K W</td>
<td>RND</td>
<td>3072</td>
<td>W</td>
<td>1/1/11</td>
<td>0.58</td>
<td>130</td>
</tr>
<tr>
<td>RND 2.5K W</td>
<td>RND</td>
<td>2560</td>
<td>W</td>
<td>1/1/11</td>
<td>0.74</td>
<td>38</td>
</tr>
<tr>
<td>RND 8K R</td>
<td>RND</td>
<td>8192</td>
<td>R</td>
<td>1/1/11</td>
<td>0.15</td>
<td>8</td>
</tr>
<tr>
<td>RND 3K W</td>
<td>RND</td>
<td>2048</td>
<td>W</td>
<td>1/1/11</td>
<td>0.93</td>
<td>48</td>
</tr>
<tr>
<td>RND 1.5K W</td>
<td>RND</td>
<td>1536</td>
<td>W</td>
<td>1/1/11</td>
<td>1.74</td>
<td>90</td>
</tr>
<tr>
<td>RND 1K W</td>
<td>RND</td>
<td>1024</td>
<td>W</td>
<td>1/1/11</td>
<td>3.21</td>
<td>166</td>
</tr>
<tr>
<td>RND 0.5K W</td>
<td>RND</td>
<td>512</td>
<td>W</td>
<td>1/1/11</td>
<td>1.99</td>
<td>103</td>
</tr>
<tr>
<td>RND 8K W</td>
<td>RND</td>
<td>8192</td>
<td>W</td>
<td>1/1/11</td>
<td>2.73</td>
<td>141</td>
</tr>
<tr>
<td>RND 4K R</td>
<td>RND</td>
<td>4096</td>
<td>R</td>
<td>1/1/11</td>
<td>0.91</td>
<td>47</td>
</tr>
<tr>
<td>RND 12K W</td>
<td>RND</td>
<td>12288</td>
<td>W</td>
<td>1/1/11</td>
<td>1.24</td>
<td>64</td>
</tr>
<tr>
<td>RND 16K W</td>
<td>RND</td>
<td>16384</td>
<td>W</td>
<td>1/1/11</td>
<td>15.63</td>
<td>807</td>
</tr>
<tr>
<td>RND 20K W</td>
<td>RND</td>
<td>20480</td>
<td>W</td>
<td>1/1/11</td>
<td>0.58</td>
<td>30</td>
</tr>
<tr>
<td>RND 28K W</td>
<td>RND</td>
<td>28672</td>
<td>W</td>
<td>1/1/11</td>
<td>2.03</td>
<td>105</td>
</tr>
<tr>
<td>RND 36K W</td>
<td>RND</td>
<td>36864</td>
<td>W</td>
<td>1/1/11</td>
<td>0.19</td>
<td>10</td>
</tr>
</tbody>
</table>

Flash Memory Summit 2018
Santa Clara, CA

22 RWSW PTS v1.0.7 – page 30
What are IO Capture Steps?

> IO Capture Steps are:  
> a) The aggregation of I/Os and metrics into discrete time intervals  
> b) Metrics that are averaged over the interval and reported as steps

> Step Resolution:  
> a) Can be widened to optimize file size for long duration captures  
> b) Can be narrowed to observe I/O Bursts, Disk Utilization, I/O Sequentiality and Quality of Service

 Unlike continuous I/O Trace data, IO Captures are a series of discrete time interval steps

 IO Capture Steps may appear continuous or discrete depending on temporal resolution
What is an IO Stream Map?

IO Stream Maps are visual representations of IO Capture Steps where:

- Each IO Step is plotted as a time point on the x-axis showing IO Streams by Frequency (IOPS) or Amount Transferred (MB/s)
- IO Metrics are presented as data series along the y-axis
- Metrics include:
  - Process IDs (PID) for each IO
  - IO Stream Composition
  - IOPS, MB/s, RTs, QDs
- Workloads can be parsed by Time, Event, PID, IO Stream Threshold and other criteria

24 RWSW PTS v1.0.7 – page 30
What are some IO Capture Tools?

- IO Capture tools are:
  a) Public or private, fee or free
  b) Specific for OS and software layer
  c) Designed to capture IO traffic
  d) Capture various IO metrics

- Free tools include blk-trace for Linux, Perfmon for Windows and IOProfiler for cross OS platform

- IOProfiler IO Capture Applets are:
  a) Free at TestMyWorkload.com
  b) Windows, Linux or MacOS
  c) File System or Block IO layer
  d) Free upload and visualization at TestMyWorkload.com
SSSI Reference IO Capture Workloads

**IO Captures:**

1. What is an IO Capture?
2. What is an IO Stream?
3. What are IO Capture Steps?
4. What is an IO Stream Map?
5. IO Capture Tools

6. **SSSI Reference Workloads – TestMyWorkload.com**

- SSSI Reference IO Captures at TestMyWorkload.com:
  1. 24-Hr Retail Web Portal
  2. SNIA Green Storage TWG Workload
  3. 24-Hr GPS Navigation Portal

- Free Data Analytics

- Free Export of IO Capture Steps for use with 3rd Party software tools

- Additional Workloads posted with ongoing SSSI TWG research

www.testmyworkload.com/info/demo
Analysis:  

1. Overview  
2. IO Stream Composition  
3. IO Stream PIDs  
4. Compare File System Drive C & Drive 0  
5. Compare Frequency (IOPS) v Amount Transferred (MB/s)  
6. Compare Block IO on Drive 0 & Drive 1  
7. LBA Range Hit Map  
8. Examples: Data Analytics  

- Workload Visualization with IO Stream Maps:  
  1. IO Stream Maps present the IO Streams and Metrics that occur during the IO Capture  
  2. Resolution can be adjusted to show long duration captures with reasonable file sizes  
  3. Fine grain resolution can enable micro second analysis of IO Capture events  

- Identification of IO Process IDs:  
  1. IO PIDs show IO association with individual processes (applications)  
  2. PID enable load balancing, interoperability and failure analysis  

- Observation of Specific Software layer IO Traffic:  
  1. Enables validation of software optimizations  
  2. Confirms how IO Streams change as they traverse the software stack  
  3. Allows tracking of individual IOs to storage architecture layers and logical storage
IO Stream Composition

IO Stream Map displays IO Stream Composition for each Step

- Each x-axis point is an IO Step
- Each IO Step is comprised of a unique combination of IO Streams
- Cumulative Workload:
  - Presents IO Streams by percent occurrence over the entire IO Capture duration
  - The IO Streams presented can be filtered by the ‘Threshold’ value
  - 2% Threshold shows all IO Streams that occur 2% or more of the time over the entire capture
  - At 2% Threshold, 9 IO Streams represents 78% of the total IO Streams (1,033 IO Streams)
IO Streams - Process IDs

IO Stream Process IDs (PID):

- PIDs can be shown for each IO
- Cumulative Workload shows 49 PIDs
- IO Stream Map can be filtered by:
  - IO Streams by PID – e.g. SQL IOs
  - Time point or Range
  - Event
- IO Streams at Time: 06:16:17
  - IOs can show multiple PIDs (sqlserver.exe, mysqld.exe, lsass.exe)
  - IOs can show single PIDs (sqlserver.exe)
Cumulative Workload - by Frequency v Amount Transferred

24-Hour GPS Navigation Portal
Example No. 6 TestMyWorkload.com

- IO Streams by Frequency (IOPS)
  1. Drive C File System level
  2. 24-Hour Capture

- Workload Settings
  1. IO Streams Threshold – 1%
  2. Temporal Resolution – 2 Min

- Cumulative Workload Box:
  1. IO Streams at 1% Threshold
  2. 9 IO Streams/86% of Total IOs
  3. 9.942M IOs of 11.526M IOs

- IO Stream Combinations vary for each IO Step
Cumulative Workload – Frequency v Amount Transferred

24-Hour GPS Navigation Portal

Example No. 6 TestMyWorkload.com

- IO Streams by Amount Transferred (MB/s)
  1. Drive C - File System level
  2. 24-Hour Capture

- Workload Settings
  1. IO Streams Threshold – 1%
  2. Temporal Resolution – 2 Min

- Cumulative Workload Box:
  1. IO Streams at 1% Threshold
  2. 16 IO Streams/69.8% of Total IOs
  3. 68.8 GiB of 89.8 GiB
  4. Note: 911,730b = 1,780.7KB
Comparison: 2 am sqlserver.exe Back-up - Drive 0 & Drive 1

24-Hour Retail Web Portal
Example No. 3 TestMyWorkload.com

- Cumulative Workload by IO Stream Threshold
  1. Drive 0 – Block IO
  2. Resolution – 5 min
  3. IO Streams Threshold – 3%
  4. IO Streams – 6 Streams; 64%

- 2 am Data Back-up by IO Stream Threshold
  1. IO Stream Threshold – 3%
  2. IO Streams – 1 Stream; 81.1%
  3. SEQ 64K R, 134,808 IOs

- 2 am by Process ID (sqlserver.exe)
  1. PID - sqlserver.exe
  2. SEQ 64K R, 134,808 IOs
LBA Range Hit Map: SEQ IOs

Drive C: Back-up Activity

- IO Stream Map at Time = 1:01:59
  1. Throughput: 82 MB/s
  2. SEQ 64K W: 38.5 MB
  3. SEQ 64K R: 34.4 MB
  4. Other IOs: 9.1 MB

- LBA Range Hit Map at Time 1:01:59
  1. SEQ 64K R: 0.030 GiB at LBA 92.1% +/- 0.1%
  2. SEQ 64K W: 0.034 GiB at LBA 78.5% +/- 0.1%

- LBA Map Range Hits
  1. Gray Line = 64K R
  2. Blue Line = 64K W
  3. Diagonal Lines indicate SEQ Accesses
Data Analytics: IOPS, Amount Transferred, RTs & Queue Depths

A  IO Stream Map by Frequency – IO Rate in IOPS
    IOPS, Ave/Max Response Times, 3% Threshold

B  IO Stream Map by Amount Transferred – Throughput in MB/s
    MB/s, Ave/Max Queue Depth
Data Analytics: Compression & Duplication Ratios

**A** Compression Ratio: How much more compressible is data
CR of 3.1 means data can be compressed 3.1 Times MORE

**B** Duplication Ratio – How many duplicative blocks are written
DR of 28% means that 28% of written blocks are duplicates
Data Analytics: Disk Utilization & IO Sequentiality

A Disk Utilization: IOs and Disk IO Idle Times
Disk Utilization of 84% = 16% Disk IO Idle

B IO Sequentiality – Adjacent LBA Range Hits
Diagonal LBA Range Hit lines indicate Sequential IOs
Data Analytics: Temporal Granularity & Step Resolution

A Fine Grain: 100 μS; 1 min Capture; Single IOs
IO Bursts, Disk Utilization, IO Specificity

B Coarse Grain: 1 Min; 24 hr Capture; 100’s of IOs
Long Term workload characterization
BREAK

3:30 – 3:45
Part 3: RWSW PTS v 1.0.7 - Tests

- Overview
- RWSW Tests:
  1. Replay
  2. Individual Streams WSAT
  3. Multi-WSAT
  4. RWSW Dirth
Overview - Key Test Process Concepts

- RWSW Tests are intended to:
  1. **Analyze** in-situ performance of the target server during the IO Capture
  2. **Optimize** Data Center Storage performance
  3. **Validate & Qualify** Data Center Storage

- Test Operator shall select and disclose the following:
  1. OS, IO capture tool used & the IO capture layer (File System, Block IO or other)
  2. The RWSW (which can be an SSSI Reference or User selected workload)
  3. The Applied Test Workload (which is filtered/parsed from the RWSW)
  4. The Data Center Storage to be tested (logical storage recognized by OS)

- Test Storage Preparation:
  1. It is recommended to apply the RWSW to Steady State (SS) when possible
  2. It may be impractical to PURGE, Pre-condition (PC) or bring RWSW to SS
  3. Test operators shall select an appropriate PC regime and disclose it in the test results
Define the Test Workload

- Run IO Capture to obtain RWSW at desired software level:
  1. File System
  2. Block IO
  3. User selected (virtualized, cluster, LUN, hypervisor, custom, other)

- Create / Select Real World Storage Workload:
  1. SSSI Reference Workload
  2. User selected/defined RWSW

- Create Applied Test Workload:
  1. Filter/parsing the selected IO Capture as desired (by Time, Event, PID, etc.)
  2. Set the IO Stream Threshold
  3. Define the Applied Test Workload & IO Stream Distribution

Overview - RWSW Tests:

1. Key Test Process Concepts
2. Define the Test Workload
3. Applied Test Workloads
4. Basic Test Flow
5. Test Settings
Create the Applied Test Workload:

- First, the IO Capture workload is filtered, or parsed, as the user desires – often the Cumulative Workload is selected.
- The filtered IO Streams and metrics are used to create an Applied Test Workload.
- IO Stream percentages of the desired workload are normalized to equal 100%.
- The Applied Test Workload is used in the RWSW tests as specified, but can be:
  1. A replay of the IO Capture sequence.
  2. A fixed composite of the IO Streams.
  3. Individual IO Streams tested to SS.
  4. A fixed composite DIRTH test.
Overview – RWSW Tests:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Key Test Process Concepts</td>
</tr>
<tr>
<td>2.</td>
<td>Define the Test Workload</td>
</tr>
<tr>
<td>3.</td>
<td>Applied Test Workloads</td>
</tr>
<tr>
<td>4.</td>
<td>Basic Test Flow</td>
</tr>
<tr>
<td>5.</td>
<td>Test Settings</td>
</tr>
</tbody>
</table>

- **RWSW PTS Basic Test Flow:**
  - a) PURGE
  - b) Apply Pre-Conditioning
  - c) Run to Steady State
  - d) Apply RWSW
  - e) Report Measurements
Test Settings

- **PURGE & Pre-Conditioning:**
  - a) It may be impractical to PURGE and run Pre-conditioning
  - b) Target Data Center Storage may be LUNs or other logical storage

- **Steady State**
  - a) Apply RWSW until performance meets the Steady State Criteria
  - b) Measurement values should be relatively time invariant

- **Active Range:** AR=100 for Enterprise; AR=75 for Client

- **Data Pattern:** DP= Random

- **Test Operator may select other test settings so long as settings are Disclosed**
RWSW Tests

- **Purpose of Tests:**
  1. To analyze in-situ target server storage performance using IO Captures
  2. To characterize and test Data Center storage performance using RWSWs

- **RWSW PTS sets forth five tests:**
  1. Target Server Self Test – Pseudo test that reports metrics during IO Capture process
  2. Replay Native – reproduces the sequence and combinations of IO Streams for testing
  3. Individual Streams WSAT – tests each individual IO Stream to Steady State
  4. Multi-Stream WSAT – applies the fixed composite of IO Streams for every test step
  5. RWSW DIRT – applies the fixed composite IO Stream steps across a range of OIO

- **Results Reporting:**
  1. All Test Settings, Workload Composition and Test Variables shall be disclosed
  2. Additional settings, metrics & reports are Optional
  3. Test Operator may select SSSI Reference Workloads or apply User selected RWSWs

---

Sample test data in this section is based on SSSI Reference IO Capture Workload No. 3 – 24-Hour SQL Server Retail Web Portal. Demo and example located at [www.testmyworkload.com/info/demo](http://www.testmyworkload.com/info/demo)
Target Server Self-Test: In Situ Performance

Target Server Self-Test presents performance of the Target Server during the IO Capture Process:

- **Target Server Self-Test:**
  1. Is a pseudo-test, not an actual test
  2. Is a compilation of measurements based on metrics taken during the IO Capture
  3. Server Performance is throttled by the Users, HW/SW & applications (and thus can be lower than lab tests that use the RWSW as the test workload)

- **Target Server Self-Test Reports** include:
  1. IO Streams Distribution vs Segments
  2. IOPS & TP vs Time
  3. Response Time Latency vs Time
  4. Outstanding IO (OIO) vs Time

Thread Count x Queue Depth

---

Flash Memory Summit 2018
Santa Clara, CA
Replay Native Test

Replay Native reproduces the sequence and combinations of IO Streams & QDs of the IO Capture

- **Replay Native Workload:**
  1. Sequence of IO Streams & QDs observed during the capture are applied to each test step
  2. RWSW DIRTH test or other RWSW Test recommended to be used as PC as PC & SS are difficult w/ Replay (due to changing step combinations)

- **Replay Native Reports include:**
  1. IO Streams Distribution v Segments (Cumulative Workload listed as listing every step composition is impractical)
  2. IOPS & TP v Time
  3. Response Time Latency v Time
  4. Outstanding IO (OIO) v Time
     - Thread Count x Queue Depth
Individual Streams WSAT Test

Individual Streams WSAT applies each individual IO Stream to Steady State

- **Individual Streams WSAT Workload:**
  1. Each IO Stream is applied to SS
  2. Ind. Stream SS can be compared to mfgr benchmark tests

- **Individual Streams WSAT Reports include:**
  1. IO Streams Distribution v Segments
  2. Segment IOPS v Time (run to SS)
  3. IOPS & Response Time by Segment
  4. Throughput & Power by Segment
Multi-Streams WSAT Test

Multi-Streams WSAT applies the fixed composite of IO Streams to Steady State

- Multi-Streams WSAT Workload:
  1. The fixed composite of the IO Streams is applied to SS
  2. PURGE & Pre-conditioning are optional, SS is required

- Multi-Streams WSAT Reports include:
  1. IO Streams Distribution v Segments
  2. Segment IOPS v Time (run to SS)
  3. IOPS & Response Time by Segment
  4. Throughput & Power by Segment
RWSW DIRTH Test

RWSW DIRTH applies the fixed composite of IO Streams to Steady State then varies OIO:

- **RWSW DIRTH Workload:**
  1. Applies the fixed composite of IO Streams to SS followed by OIO loop
  2. Shows IO & Bandwidth saturation across a range of OIO from 1 to 1,024

- **RWSW DIRTH Reports include:**
  1. IO Streams Distribution v Segments
  2. RT Histogram for Max IOPS 16,575
  3. Quality of Service: OIO 1, 8, 32
  4. ART, 5 9s & TP v Total OIO Range
Thank You

감사합니다

Thank You

Danke

Ευχαριστίες

Dalu

Köszönöm

Tack

Cпасибо

Dank

Gracias

Merci

谢谢

Seé

ありがとう

www.TestMyWorkload.com

edenkim@calypsotesters.com