An Examination of User Workloads for SSDs

Eden Kim
Calypso Systems, Inc.
An Examination of User Workloads for SSDs

- Why SSD Performance Depends on the Workload
- Why We Test with Real World Workloads
- User Workload Examples:
  - Example A: Retail Store Web Portal 24 hr Capture
  - Example B: Mac OSX You Tube 20 min Capture
  - Example C: Windows 8 Home PC 20 min Capture
  - Example D: Court Room Video Surveillance 10 hr Capture
  - Example E: Anti Virus Boot Drive Scan
SSD Performance…
It’s all about the workload

Where you measure it
- Data ingress, application space, NIC packet traffic, LUN, Array or Device
- IO Streams are affected at every level of abstraction
- Metadata, data reduction, virtualization, coalescing, fragmentation, etc.

How you define it
- Workloads are comprised of many, many IO streams
- Differing Access Patterns, Data Transfer Sizes and RW mixes
- Unique Data Content and Spatial and Temporal Locality of Reference
- Different Demand Intensity (users and jobs) and the number of outstanding IOs

How you test it
- Corner Case Benchmark, Synthetic Application or Real World workloads
- Here, we examine several applications to determine what the storage actually ‘sees’
Different ways to test SSDs…
Each one has its place

**Corner Case Benchmark Tests**
Convenient way to provide a quick comparison
SNIA Performance Test Specification (PTS-E/C) is a good way
*However, real world workloads are never a single corner case workload*

**Synthetic Application Workloads**
Synthetic approximations of commonly used applications
Allow more finely tuned test parameters and conditions
*However, application workloads differ on different systems and at different times*

**Real World IO Capture Workloads**
Creates specific test workloads based on IO captures of real world applications
IO Streams at the Data Center Storage LUN presents what the storage actually ‘sees’
*However, they change over the course of a day and no two systems are identical*
Why Test with Real World Workloads?

Because we can...

- IO trace and capture tools are available to capture & characterize real world workloads
- Test tools allow the creation of real world test workloads based on these trace captures
- Real world workloads provide another dimension to evaluate storage

We need to know what our actual workloads look like

- Everyone’s deployed application workloads are unique to their hardware / software solution
- Workload definition depends on where in the IO stack the workload is captured
- Workloads change over the course of the day depending on the use case of the storage system

We want to test storage to the actual deployed workload

- It is the best way to measure YOUR workload
- Captures can confirm what IO streams actually are presented to the storage
- Create or confirm Data Center Storage Tiering strategies - see what IO traffic goes where
Real World Workload Capture

EXAMPLE A
24 Hour 2,000 outlet retail webserver portal
How many different IO Streams were observed in 24 hours?

A. 0 – 100
B. 101 – 500
C. 501 – 1,000
D. > 1,000
How many different IO Streams were observed in 24 hours?

A. 0 – 100
B. 101 – 500
C. 501 – 1,000
D. > 1,000

Quiz Answer is....

5,038!
Example A: 24 Hour 2,000 Store Webserver Portal Capture

Sample 24hr 2016-02-15

Model: Virtual HD
Volume: 214 GB
I/Os: 4,326,159
Read: 142.9 GiB
Written: 20.8 GiB

Streams threshold: 3%

Workload Streams (by frequency)

LBA Range Hits (by frequency)

RND 64K R: 16 hits in 21.8% of drive space
Processes/hits:
- System: 16
Example A: Retail Web Portal
Workload Segment Definition / Analysis

Sample 24hr 2016-02-15

Path: \LPhysicalDrive0
Model: Virtual HD
Volume: 214 GB
IOs: 4,326,159
Read: 142.9 GiB
Written: 20.6 GiB

Workload Streams (by frequency)

 Cumulative Workload

<table>
<thead>
<tr>
<th></th>
<th>Cumulative Workload</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>SQL Server: 02:00:20</td>
</tr>
<tr>
<td></td>
<td>RND 64K R</td>
</tr>
<tr>
<td></td>
<td>SEQ 0.5K W</td>
</tr>
<tr>
<td></td>
<td>RND 8K R</td>
</tr>
<tr>
<td></td>
<td>SEQ 8K R</td>
</tr>
<tr>
<td></td>
<td>RND 4K W</td>
</tr>
<tr>
<td></td>
<td>SEQ 64K R</td>
</tr>
<tr>
<td></td>
<td>RND 8K W</td>
</tr>
<tr>
<td></td>
<td>RND 4K R</td>
</tr>
<tr>
<td></td>
<td>RND 16K R</td>
</tr>
<tr>
<td></td>
<td>SEQ 4K R</td>
</tr>
</tbody>
</table>

Total IOs of 5,038 streams: 4,326,159
Selected 6 streams: 2,784,635 (64.4%)

<table>
<thead>
<tr>
<th></th>
<th>Cumulative Workload</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Daily Operation</td>
</tr>
<tr>
<td></td>
<td>SEQ 0.5K W</td>
</tr>
<tr>
<td></td>
<td>RND 64K R</td>
</tr>
<tr>
<td></td>
<td>RND 8K R</td>
</tr>
<tr>
<td></td>
<td>SEQ 8K R</td>
</tr>
<tr>
<td></td>
<td>RND 4K W</td>
</tr>
<tr>
<td></td>
<td>RND 8K W</td>
</tr>
<tr>
<td></td>
<td>RND 4K R</td>
</tr>
<tr>
<td></td>
<td>RND 16K R</td>
</tr>
<tr>
<td></td>
<td>SEQ 4K R</td>
</tr>
<tr>
<td></td>
<td>RND 32K R</td>
</tr>
</tbody>
</table>

Total IOs of 4,601 streams: 1,896,820
Selected 6 streams: 1,251,852 (66.6%)

<table>
<thead>
<tr>
<th></th>
<th>Cumulative Workload</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Closing</td>
</tr>
<tr>
<td></td>
<td>RND 64K R</td>
</tr>
<tr>
<td></td>
<td>RND 8K R</td>
</tr>
<tr>
<td></td>
<td>SEQ 8K R</td>
</tr>
<tr>
<td></td>
<td>SEQ 0.5K W</td>
</tr>
<tr>
<td></td>
<td>SEQ 512K R</td>
</tr>
<tr>
<td></td>
<td>RND 16K R</td>
</tr>
<tr>
<td></td>
<td>RND 24K R</td>
</tr>
<tr>
<td></td>
<td>RND 32K R</td>
</tr>
<tr>
<td></td>
<td>RND 64K R</td>
</tr>
</tbody>
</table>

Total IOs of 1,174 streams: 888,409
Selected 4 streams: 610,837 (68.7%)
Example A: Retail Web Portal

SQL Server over 24 Hours
Real World Workload Capture

EXAMPLE B
Mac OSX You Tube 20 min Video
Example B1: Mac OSX You Tube – Block IO

Mac OSX - You Tube Video 20 min - Block IO Capture

Model: APPLE SSD SM0512F
Volume: 500 GB
IOs: 26,792
Read: 0.6 GiB
Written: 0.1 GiB

Streams threshold: 3%

Workload Streams (by frequency)

LBA Range Hits (by frequency)

- RND 4K W: 24 hits in 5 ± 1% of drive space
- Processes/hits: kernel_task: 24

2016 Data Storage Innovation Conference. © Calypso Systems, Inc. All Rights Reserved.
Example B2:
**Mac OSX You Tube – File System**

Mac OSX - You Tube Video 20 min - File System Capture

- **Volume:** 500 GB
- **Read:** 0.5 GiB
- **Written:** 0.1 GiB

Total I/Os of 1,105 streams: 60,993
Selected 6 streams: 46,994 (76.9%)
Real World Workload Capture

EXAMPLE C
Windows 8 Home Computer 20 min activity
Example C: Home PC – Block IO 10 min Capture

Sample Capture: Home Computer

Workload Streams (by frequency) for 17 selected processes

LBA Range Hits (by frequency) for 17 selected processes

Cumulative Workload

Total IOs of 502 streams: 26,269
Selected 5 streams: 18,171 (69.2%)

Total IOs of 52 streams: 778
Selected 5 streams: 613 (78.8%)

Total IOs of 24 streams: 486
Selected 3 streams: 421 (86.6%)
Real World Workload Capture

EXAMPLE D
Court Room Video Surveillance Cameras
Example D:
Court Room Video Surveillance – 10 hr Capture

Court Room Video Surveillance

Model: ST31500341AS
Volume: 1,397 GB
IOPS: 701,232
Read: 0.1 GiB
Written: 101.9 GiB

Streams threshold: 1%

Total I/Os of 2,926 streams: 701,232
Selected 5 streams: 136,574 (19.3%)
Real World Workload Capture

EXAMPLE E
PC Anti Virus Scan of HDD
Example E: Anti Virus Scan – PC HDD

Anti Virus Scan

Model: ???
Volume: 60 GB
IOs: 967,743
Read: 3.5 GB
Written: 46.7 GB

Streams threshold: 3%

Workload Streams (by frequency)

LBA Range Hits (by frequency)

SEQ 64K W: 3,775 hits in 84% of drive space
Processes/hits:
  avg.exe: 3,775
Take-Aways

- SSD Performance Depends on the SSD Workload
- IO Streams Change as they Traverse the SW Stack
- IOProfiler Captures IO Streams at the Block IO level
- See What IOs Actually get to the SSD Storage
- Be Sure you Buy the ‘Right Amount of Performance’
Understand Your SSD Workloads

To Analyze the Workloads Presented, Go to

TestMyWorkload.com

Capture & Analyze Your SSD
Real World Workloads Today!
For more information, contact Calypso Systems, Inc.

info@calypsotesters.com  www.calypsotesters.com