SSD Performance States

Overview

SSDs have generated substantial interest in both enterprise and personal computing, whether for their low power consumption, high performance, or drop-in replacement capability (via standard hard drive interfaces like SAS and SATA).

Despite their deceptively familiar appearance, SSDs behave differently from HDDs. In this technical brief we will examine one such difference: Performance variability as a function of time.

The concept may seem unusual at first glance — SSD performance may change as the device is used? That is a good reason to examine this phenomenon and show a real-world example of it.

How SSDs Are Built

For this brief, we will limit ourselves to examining SSDs that use NAND as their basic storage media. Although design and construction details may differ from SSD to SSD, NAND-based SDDs are generally comprised of the following main components:

- Printed Circuit Board (PCB)
- Controller
- NAND
- Host Interface Connector (i.e., SATA, SAS, etc.)
- DRAM (not required for all)

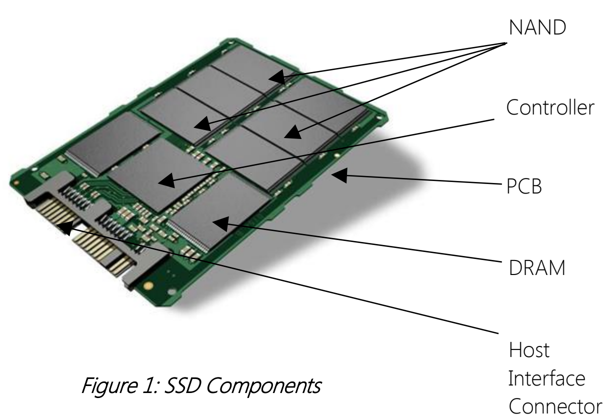

Figure 1: SSD Components

Performance

Performance is typically characterized in one of three ways: input/output operations per second (IOPS), throughput in megabytes per second (MB/s) or with a composite benchmark that gives unitless results.

This brief will look at empirical data taken from a sample SSD from a well-known manufacturer. Although the exact shape and amplitude of the resultant performance curves may differ among SSDs, most SSDs currently on the market will exhibit similar behavior: As the drive fills, performance will change nonlinearly. If the measurement period is sufficiently long, the device will exhibit one or more intermediate performance states observed during the measurement period.

The sample data in this brief was collected by Calypso Systems (www.calypsotesters.com).

Performance States: A Common Lexicon

There are consistently four different performance states for SSDs — FOB, steady, burst and transition. The time spent in each state and (to a lesser extent) the order in which they occur will vary among SSDs, but each SSD examined to date shows these performance states. Some SSDs will show some states to a greater degree, and some will enter and exit a given state more than once. The two states of primary interest are FOB and steady state, as defined below:

Fresh-Out-of-Box State (FOB)

The condition of a new/unused SSD when first received from the manufacturer is FOB state. The NAND on the SSD will have few (if any) program/erase (P/E) cycles applied when the device is in this state. (The exception would be any P/E cycling done at the factory as part of the manufacturing process.) The SSD comes with the NAND ready to have data written (that is, all storage elements are pre-erased).

Steady State

The condition of an SSD when most of the transient performance behavior has died away is steady state. An SSD may have several steady state regions in its performance curve. Steady state performance is typically indicated by a relatively small change in performance over a relatively long timeframe.

The other two performance states — burst and transition — are of less interest, but are equally well defined:

Burst State

A transient performance state, typically with very short duration and higher performance than the states immediately preceding and following, is a burst state.

Transition State

The state when an observed performance change is consistently increasing or decreasing, indicating the SSD performance is in transition, is a transition state.

Measured Results

Figure 2 shows the different measured performance states of a sample SSD. The two plots represent data from the same device; the data taken is the same, only the x-axis units changed:

- Y-axis: IOPS
- Lower x-axis (blue): Total number of gigabytes written to the drive
- Upper x-axis (red): Time in minutes (increasing from left to right)

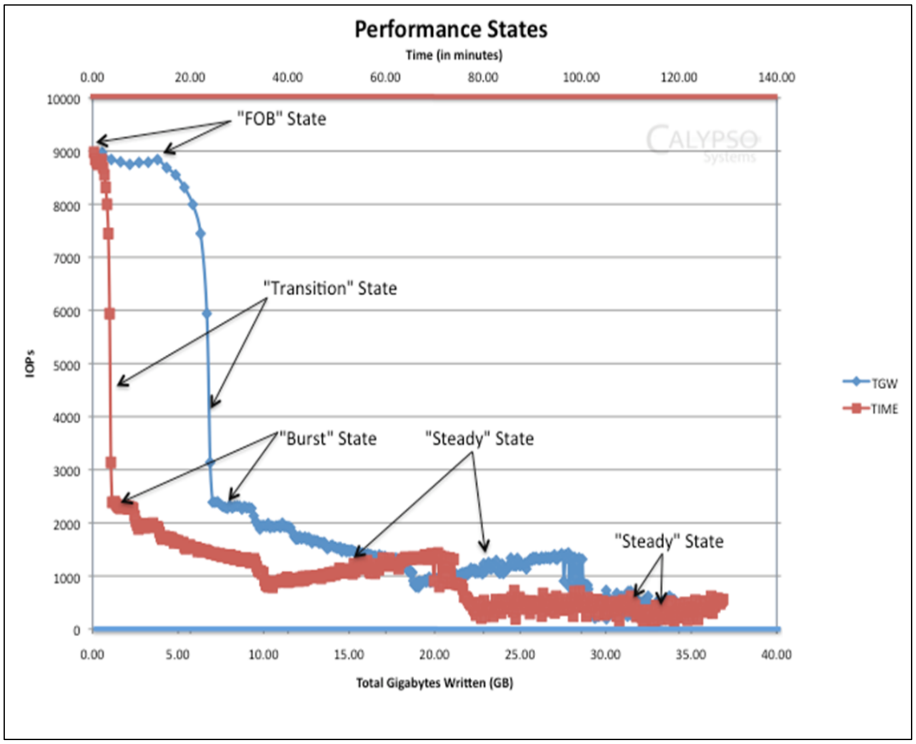

Figure 2: Performance States vs. Time and Total Gigabytes Written

The two plots do not coincide because the rate at which the drive accepts data changes with time. All measurements are expressed in IOPS, as shown on the vertical axis.

As the drive was filled, several performance states were observed:

<u>FOB State</u>: The FOB state is visible at the extreme left of the plot. This state is reached when the drive has little to no user data; all the NAND cells are erased and available to receive new data. As the SSD is written, the IOPS decrease.

<u>Transition State</u>: Immediately following the FOB state, the SSD enters a transition state marked by steadily decreasing performance (through about 7GB written).

Burst State: At the bottom of the transition state, the drive performance briefly stops decreasing and is level. This state is not long enough to be steady state, but instead reflects a short performance recovery (or burst).

Steady State: From 9GB to 28GB written, drive performance is consistent, indicating a steady state. Although performance again decreases at approximately 28GB written, the performance between 9GB and 28GB written has little variance.

Conclusion

For this test, we subjected a sample SSD to a workload specifically designed to illustrate performance variation with time.

As can clearly be seen from the performance plot, this SSD (and SSDs in general) shows substantial performance variability over time and amount of data written. When the device is in the FOB state, its NAND is pre-erased and ready to receive data — hence the performance in this state is greatest. As the SSD is written and re-written, the amount of pre-erased NAND (NAND that is ready to receive new data) and its location is governed by the SSD's internal design (controller and firmware).

Exactly when will an SSD exit the FOB state and enter the transition state, and then the steady state? Due to device level differences, this behavior is difficult to predict and, instead, is better measured empirically.

Because different drive designs offer different performance characteristics, it is also best to measure performance using a workload as close as possible to the workload the drive will experience when deployed.