

# Best Practices for SSD Performance Measurement

Doug Rollins, Senior Strategic Applications Engineer

Micron Technology, Inc.  
*Technical Marketing Brief*

SSDs have pronounced write-history sensitivity, which means that they have unique requirements for accurately measuring their performance. This technical marketing brief documents how Micron measures enterprise SSD performance, with an emphasis on the following points:

- Main goals of performance measurement (accuracy, consistency, and repeatability)
- Key assumptions about enterprise SSD requirements
- Definitions of commonly used terms
- Tools used
- Metrics of primary interest
- Test mechanics, including:
  - Preconditioning
  - Using a common measurement tool to construct a write-saturation plot

This brief cites empirical data demonstrating the main NAND-based SSD performance regions—FOB (fresh-out-of-box), transition, and steady state—and shows an example of the markedly different results that can be achieved (on the same drive) depending on the order in which the stimuli are applied. It also shows alignment to SNIA's Performance Test Specification protocol, provides explicit details on our precondition and measurement techniques (along with an example), and shows the relevance of the collected metrics.

## Main Goals of Enterprise Performance Measurement

Micron's enterprise SSD performance measurement is concerned primarily with the following standards:

- Accuracy
- Consistency
- Repeatability

It is based on some key assumptions about enterprise SSDs and the enterprise market—specifically:

- Drive fill state
- Access and I/O traffic patterns
- Decision criteria (for purchase and deployment)

## Accuracy, Consistency, Repeatability

Micron testing is designed to minimize the impact of external elements. Specifically, we minimize the influence of the operating system, host configuration, and target variables to enable reliable, relative comparisons.

To do this, we use an invariant server-grade operating system and apply only predetermined service packs and hotfixes. We disable automatic patch installation, set the target cache options manually, and ensure all host components have the same BIOS settings, firmware version, and drivers installed.

This ensures consistent run-to-run, host-to-host, and drive-to-drive measurement consistency.

We always put the drive in the same state at the beginning of each measurement, precondition each drive the same way for each measurement, and stimulate the drive to the same performance state—the test process is deterministic.

## Key Assumptions

When measuring performance for enterprise SSDs, we assume that the requirements are very different from those of client SSDs.



## Enterprise SSD Performance Measurement Assumptions

- **Drive fill state:** The drive is always 100% full.
- **Accesses:** It is being accessed 100% of the time (that is, the drive gets no interface idle time).
- **Decisions:** The enterprise market chooses enterprise SSDs based on their performance in steady state, and that steady state, full, and worst case are not the same thing.
- **Consequences of failure:** Failure is catastrophic for multiple users.

## Client SSD Performance Measurement Assumptions

- **Drive fill state:** The drive has less than 50% of its user space occupied.
- **Accesses:** It is accessed a *maximum* of 8 hours a day, 5 days a week (but typically is *written* much, much less frequently).
- **Decisions:** The client market chooses enterprise SSDs based on their performance in the FOB state.
- **Consequences of failure:** Failure is catastrophic for a single user.

Furthermore, we assume that the focus of enterprise SSD performance measurement should include the intended workload. For example, an enterprise SSD intended to accelerate a database should be measured at least using a typical database workload (typically 8K transfer size; random; two-thirds read, one-third write traffic; fully loaded queue); if the same drive is intended to be a cache for streaming files, it should be measured with that workload (larger blocks, read-intensive, fully loaded queue).

## Performance States

Because SSD performance can change as it is written, our enterprise performance measurement focuses on the steady state performance region. We use the following definitions in our measurements (see the Micron white paper "SSD Performance States" for detailed descriptions).

### Steady State

Micron uses a complex definition, from the SNIA Solid State Storage Initiative's Performance Test Specification:

- $\text{Max}(y) - \text{Min}(y)$  within the measurement window is no more than 20% of the  $\text{Ave}(y)$  within the measurement window, **AND**

- $[\text{Max}(y)$  as defined by the linear curve fit of the data with the measurement window] -  $[\text{Min}(y)$  as defined by the linear curve fit of the data with the measurement window] is within 10% of the average within the measurement window.

### Full Drive

The definition of a full drive is more straightforward:

- A drive that is full has been overwritten some multiple (could be 1X) of the user-accessible LBA space by a fixed pattern that may vary from the test stimulus (i.e., 2X user LBA space written sequentially with 128K transfers).

### Worst Case

Worst case performance is exactly what it sounds like:

- The drive has been stimulated over some fixed time with a workload *intentionally designed* to demonstrate the drive's worst possible performance.

For example, this type of stimulus may include (but is not limited to):

- Small transfers mixed with large transfers
- Intentionally misaligned writes

## Mechanics: Test Sequence

The mechanics of Micron's enterprise SSD performance measurement:

1. **Purge:** Regardless of what has been done previously to the SSD, put the drive in a known, fixed state that emulates the state in which the drive would be received from the manufacturer. This is the fresh-out-of-box (FOB) state. Micron uses the secure erase command to place sample SATA SSDs in this state; other methods may be protocol- or vendor-specific.
2. **Precondition:** Following SNIA's Performance Test Specification for workload-independent precondition, Micron next writes the drive with 128KB sequential transfers aligned to 4K boundaries.
3. **Test:** For any metric of interest, we test only one metric at a time, and we test into steady state.
4. **Record results:** At the end of the test run, results are compiled, drive performance from FOB to steady state is analyzed, and then steady state performance values are recorded.



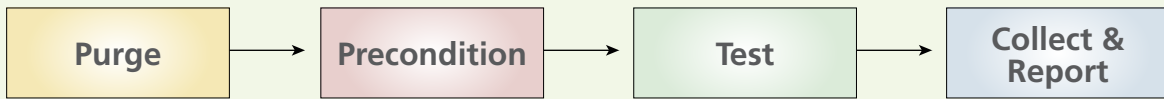


Figure 1: Test-flow sequence

## Mechanics: Setting Up the Test Platform

This section gives an example of host configuration using Microsoft Windows Server 2008; other operating systems should be treated similarly.

### Service Packs and Automatic Updates/Installation

To ensure run-to-run consistency, fix a service pack and patch level, and then disable the operating system from installing updates (automatic installation may cause an automatic system restart, without regard to a test that may be running, which can result in run-to-run variance).

**Note:** This setting may not be in accordance with your personal or corporate security policy. Micron assumes no responsibility for any consequences of this setting or its adjustment. This information is provided for reference only. Because most SSD performance measurement is done on an isolated system, this may be less of an issue, but you should know and understand your corporate security policy.

### Ensure the SSD is Recognized

1. Open the Disk Manager, locate the SSD to be tested, and then check that the drive does not contain a partition.

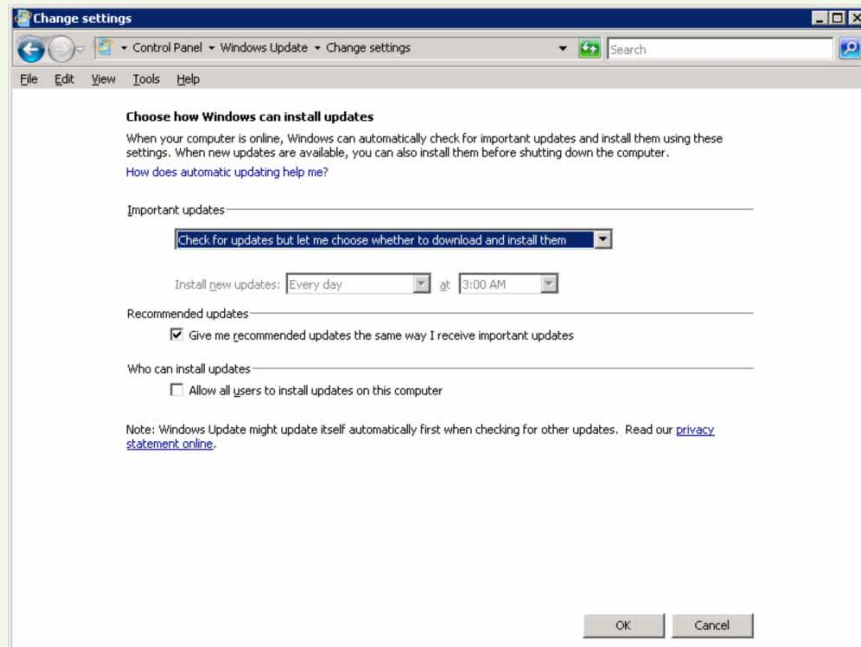


Figure 2: Change operating system settings



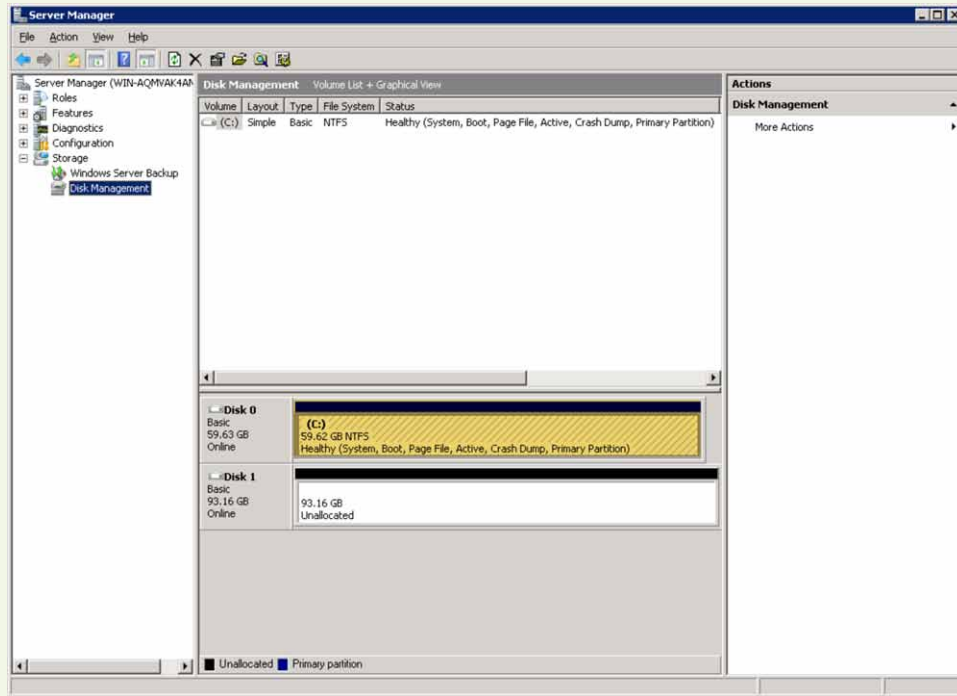


Figure 3: Locate the SSD and check partitioning

2. Although the disk number assigned to the SSD may vary, ensure that the SSD is visible, as shown in the center of Figure 3.
3. You may have to mark the SSD as online manually, and you may also have to initialize the SSD. (**Note: Initializing the SSD is not the same as formatting it. You should see a prompt for initializing the disk. Click Yes, and then select the MBR option.**)

### Set the SSD's Cache Option

If the SSD being tested does not support a complete power-backup solution (all volatile areas protected by an internal power-backup mechanism), Micron disables the write cache option, as shown in Figure 4.

### Mechanics: Test Setup and Execution – Benchmark Configuration

This section gives an example of host configuration using IOMeter® as the benchmark.

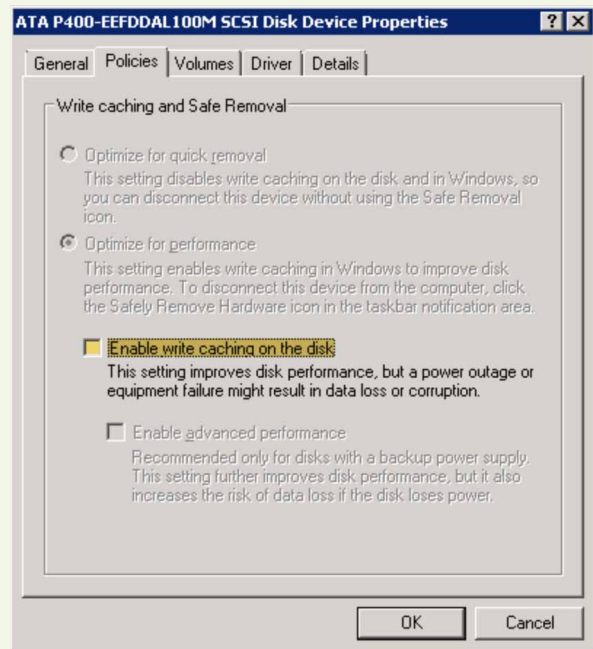


Figure 4: Disable write caching



## Launch IOMeter

IOMeter is typically launched from the Start menu, but you may have to browse to the folder where it is stored and launch it using its icon.

## Set Up the Precondition Test Specification in IOMeter

As noted before, Micron uses the SNIA SSSI TWG PTS workload-independent precondition of 128K sequential traffic aligned on 4K boundaries to write twice the user capacity. The same precondition is applied to all devices tested.

1. Set the transfer size to 128K.
2. Set the percent random/sequential distribution to 100% sequential.

3. Set the percent read/write distribution to 100% write.
4. Align the I/O on the 4K page boundaries.

## Set the Run Time, Cycling Options, and Queue Depth

As a general rule, Micron uses a run time of five minutes per iteration, giving a time-axis granularity of five minutes in the performance vs. needed time plot. In addition, the queue depth is always set to the maximum supported by the SSD to ensure that it is written at the fastest rate possible. The Cycling Options section is set to fix the queue depth (no cycling) as shown in Figure 6.

The 128K sequential write precondition access specification is now configured.

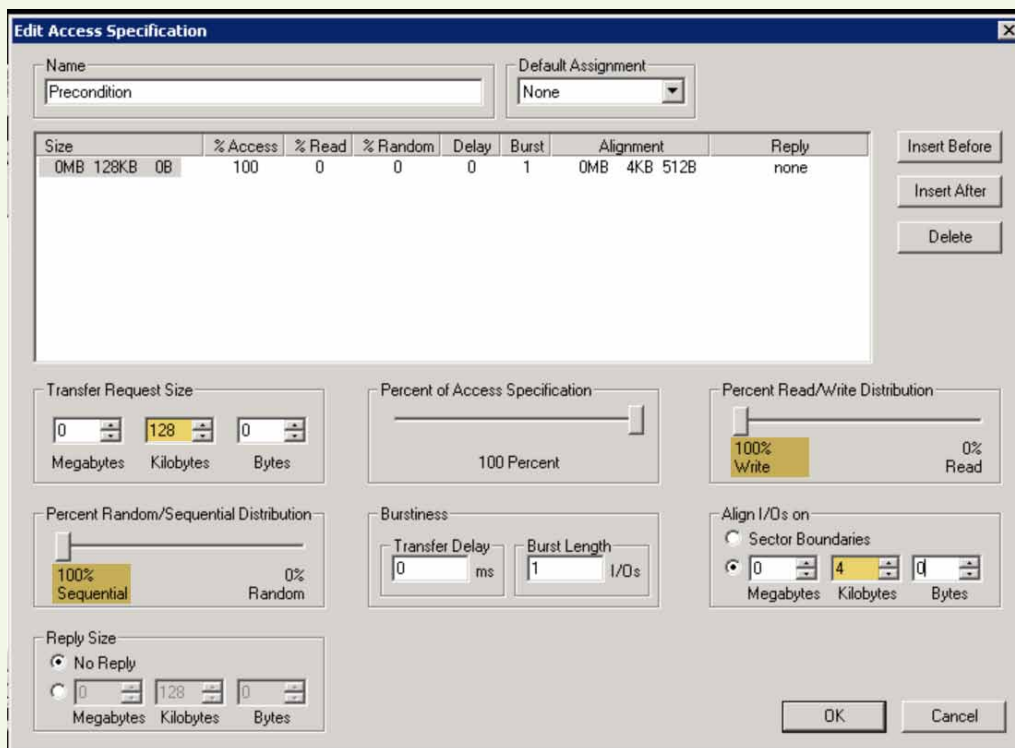


Figure 5: Precondition test specification settings



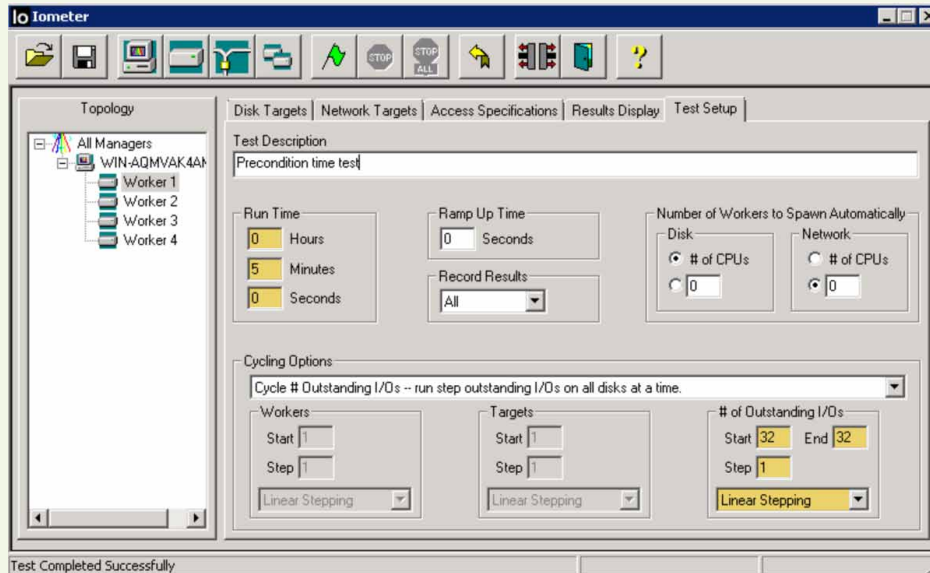


Figure 6: Set run time, cycling options, and queue depth

## Determining the Precondition Time

IOMeter tests are based on time, not the number of GB written. Because SSD performance varies substantially, Micron does not fix a time to run the precondition. Instead, we do the following:

1. Add the precondition access specification to the test run 25 (or more) times.
2. Note that the precondition access specification dwell time is set to five minutes, and the queue depth to the maximum that the device supports.
3. Execute the precondition test.
4. This test shows the SSD write performance from FOB to steady state for the precondition access specification. Examine the results to determine the steady state write rate for the precondition test.
5. Divide the drive user capacity by the steady state write rate to determine the time necessary to write the drive once (with 128K sequential traffic).
6. Double the time in step 5 to ensure that the drive is written twice.
7. Use the duration from step 6 for all future preconditioning for this drive.

**Note:** The time to precondition a drive may be unique and may vary from model to model, or from manufacturer to

manufacturer. After the duration of the precondition has been determined for a specific drive make and model, it remains constant—but only for that make and model. The precondition duration must be determined for every drive to be tested and must be applied identically for all tests on that drive.

**Note:** The number of times that the precondition access specification is added to the precondition run determines the number of values that can be shown on the time axis of a performance-vs.-time plot; the duration of the precondition access specification determines the time between those time axis values.

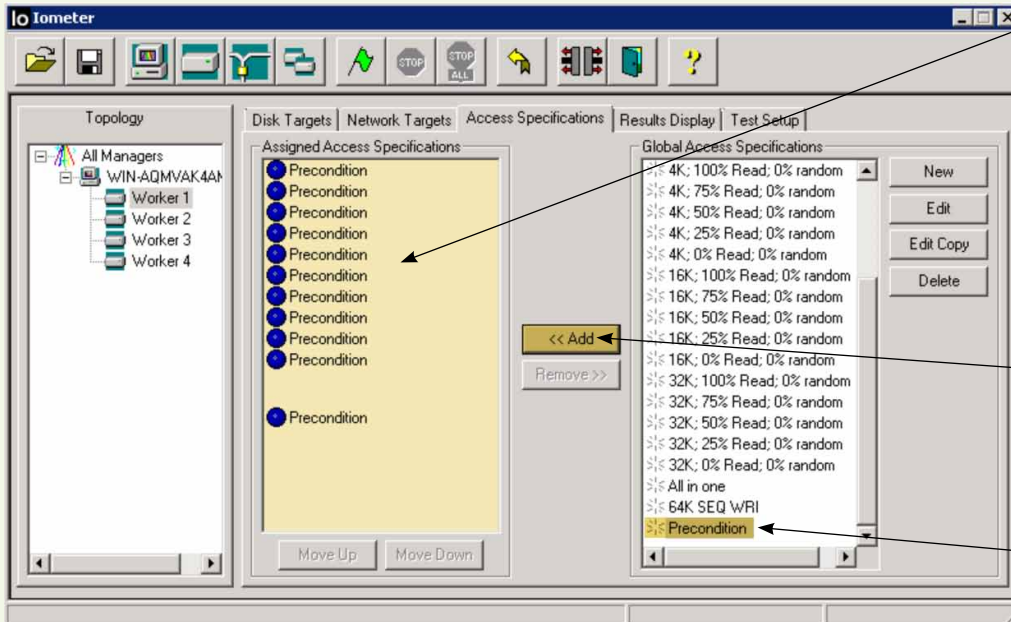
## Create the Precondition Run

Start by adding ONLY the precondition access specification to the test run (Micron's rule of thumb: add the access specification 25 times as a starting value) as shown in Figure 7.

## Execute the Precondition Run

After the precondition access specification is added to the test run 25 (or more) times, execute the run, and then save the results.





The number of times the access spec is added = the number of x-axis intervals (for a drive you've never tested, more is better); click **Add** once for each 5-minute run you want to add to the precondition

Add the **Precondition** access specification

Add only the **Precondition** access specification!

Figure 7: Add Precondition access specification

### Mechanics: Test Setup and Execution – Establishing Preconditioning

This section shows how the data collected in the benchmark configuration process is used to determine the preconditioning duration.

After parsing the resulting .csv file (via Microsoft Excel® or some other method) to remove the extraneous data, the results should look similar to Figure 8.

	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Iops	Read Iops	Write Iops	Read MBps	Write MBps	Transacts	Connectio	Average	Average	Average	Average	Average	Maximum	Maximum	Max
2	2492	794	0	2492	794	311	5993	2492	794	0	12	83537	12	83537	0
3	2484	429	0	2484	429	308	5536	2484	429	0	12	96289	12	96289	0
4	2484	615	0	2484	615	308	6789	2484	615	0	12	90196	12	90196	0
5	2484	517	0	2484	517	308	6641	2484	517	0	12	90263	12	90263	0
6	2486	703	0	2486	703	308	3375	2486	703	0	12	97116	12	97116	0
7	2480	591	0	2480	591	307	5739	2480	591	0	13	90229	13	90229	0
8	2487	281	0	2487	281	308	4192	2487	281	0	12	96827	12	96827	0
9	2484	517	0	2484	517	308	6641	2484	517	0	12	90263	12	90263	0
10	2481	404	0	2481	404	307	6754	2481	404	0	12	99902	12	99902	0
11	2483	518	0	2483	518	307	8899	2483	518	0	12	98863	12	98863	0
12	2482	664	0	2482	664	307	833	2482	664	0	12	99244	12	99244	0
13	2483	735	0	2483	735	307	9688	2483	735	0	12	98874	12	98874	0
14	2480	876	0	2480	876	307	6095	2480	876	0	13	90187	13	90187	0
15	2459	542	0	2459	542	307	4427	2459	542	0	13	90889	13	90889	0
16	2482	123	0	2482	123	307	7854	2482	123	0	12	99525	12	99525	0
17	2481	14	0	2481	14	307	5425	2481	14	0	13	90043	13	90043	0
18	2457	778	0	2457	778	307	2222	2457	778	0	13	91827	13	91827	0
19	2459	492	0	2459	492	307	4366	2459	492	0	13	90919	13	90919	0
20	2459	744	0	2459	744	307	458	2459	744	0	13	90786	13	90786	0
21	2458	158	0	2458	158	306	8947	2458	158	0	12	99487	12	99487	0
22	2458	347	0	2458	347	307	2933	2458	347	0	13	91523	13	91523	0
23	2239	505	0	2239	505	279	5381	2239	505	0	14	28716	14	28716	0
24															

Figure 8: Preconditioning results



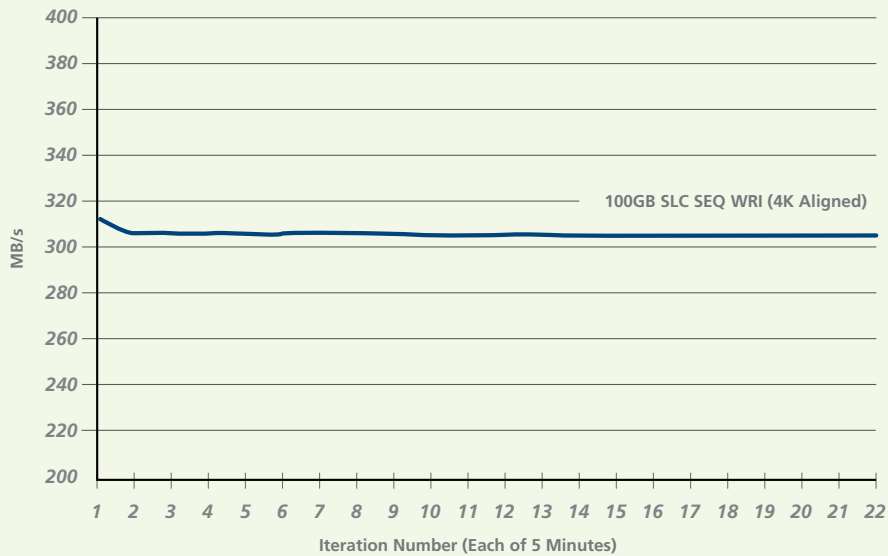


Figure 9: Write speed during precondition

### Setting Up the Test Platform: Determining the Duration of the Precondition

Because this phase of the test is concerned only with establishing the correct precondition duration, the only metric of interest in the data set in Figure 8 is the write MB/s. If this data is plotted, the result is shown in Figure 9.

From the above plot and from the definition of steady state, this example drive has a steady state write rate of about 300 MB/s.

Finally, the steady state write rate for 128K sequential, page-aligned traffic (with a queue depth of 32) sets the duration for all subsequent preconditioning.

The time to write this sample drive twice is shown in Table 1.

Drive capacity	100GB = 12,400MB
Write rate	300 MB/s
Minutes to fill (1x)	6
Minutes to fill (2x)	12
Precondition time	12 minutes

Table 1: Precondition calculations

**Note:** This scenario is the time for the drive to go from a purged state into a preconditioned state. This can be different for different drives, different media, and different firmware revisions. It is consistent for each drive under test.

### Mechanics: Test Setup and Execution – Set Up the Test Metric of Interest

With the precondition access specification and duration established, the access specification for each metric of interest needs to be set up. Note that the test flow is similar to previous steps, with only a few differences.

An example of an access specification to measure 4K random writes on an SSD with 4K pages and a maximum queue depth of 32 is shown in Figure 10.

1. The transfer size is set to 4K, and the percent random/sequential to 100% random.
2. Because the metric of interest is 100% write, the percent read/write distribution slider is set to 100% write.
3. Finally, all I/Os are aligned to 4K.
4. As before, the run time is set to five minutes, and the queue depth is set to the desired value for the test. (**Note:** Micron typically does not vary the queue



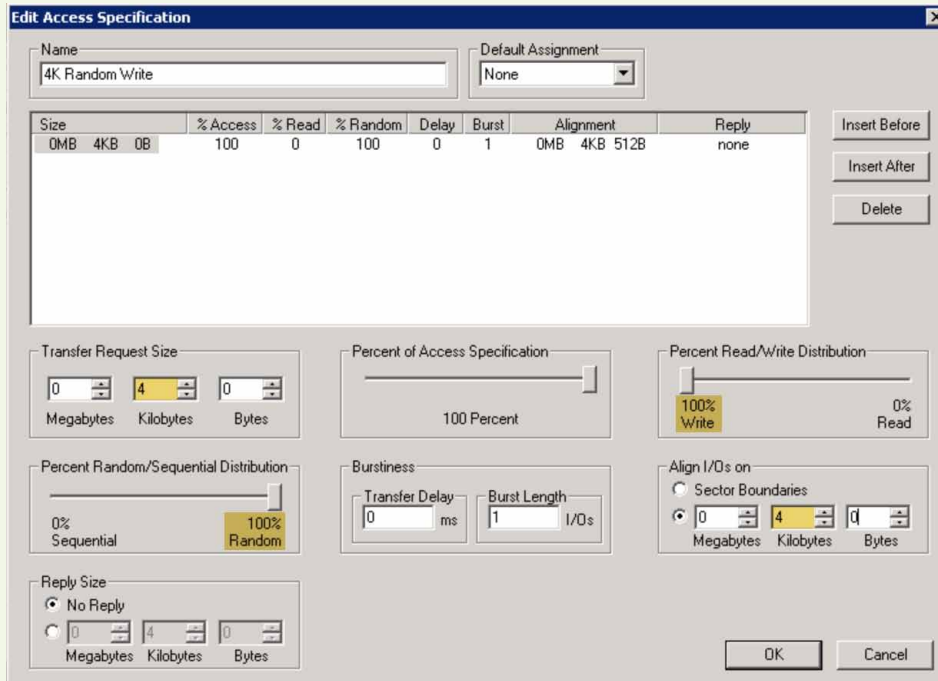
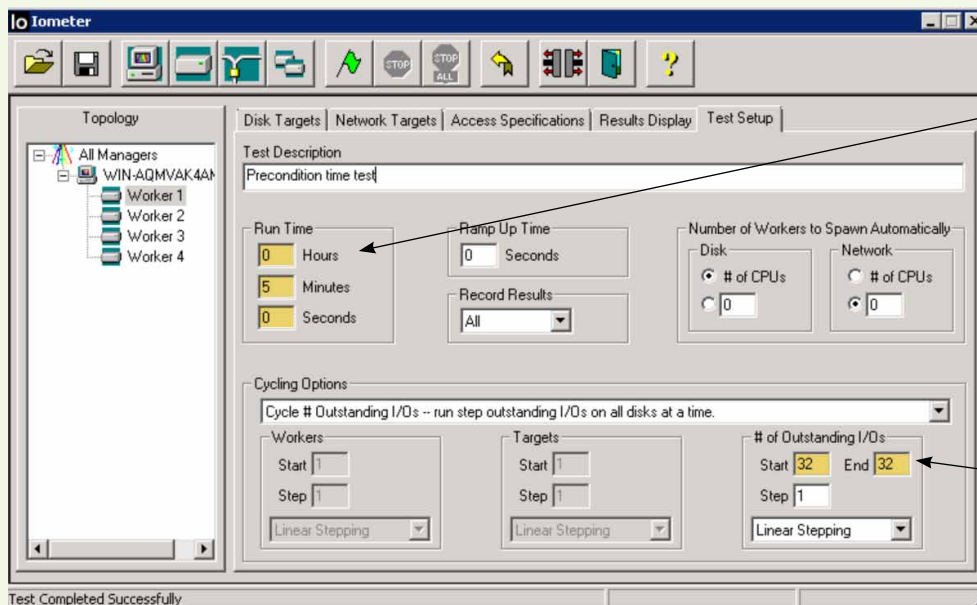


Figure 10: Set access specifications for selected metric

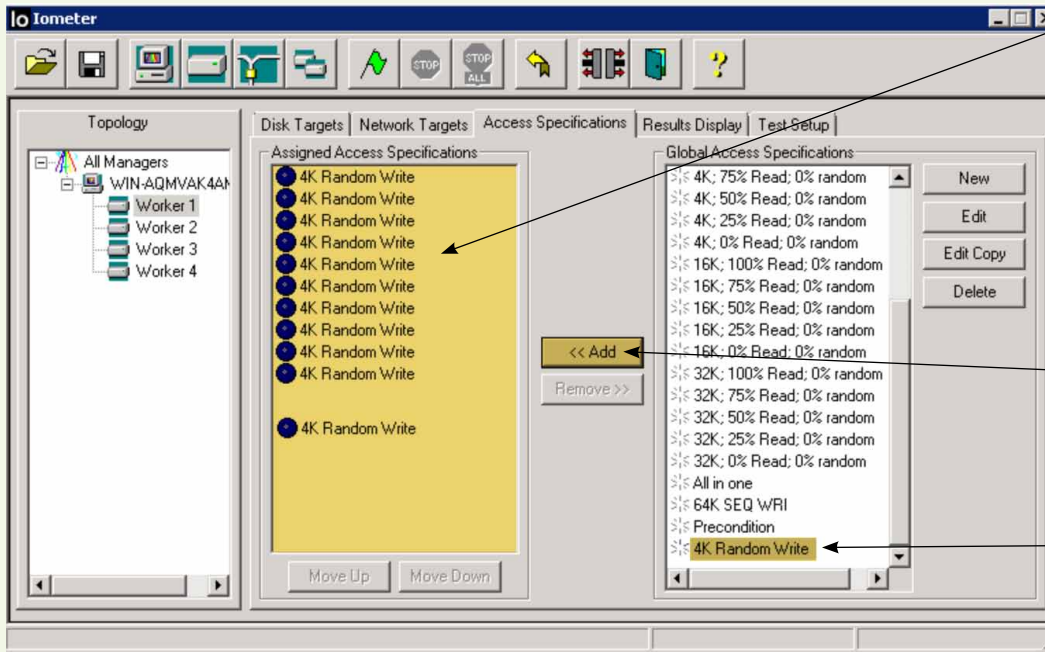


The **Run Time** will set the granularity of the results (distance between x-axis marks) in the results plot

For the metric of interest, set the queue depth to the desired value (this example uses 32)

Figure 11: Set run time and queue depth





From experience, this 100GB drive is always in steady state by the 25th iteration, so add the access specifications 25 times

Add the **4K Random Write** access specification

Add only the **4K Random Write** specification

Figure 12: Set remaining test specifications

Type	Iterations
SLC	25
MLC	75-100

Table 2: Micron's starting point for SLC and MLC iterations

depth per test, as it can make data compilation and analysis more difficult. Instead, we repeat the test sequence for each queue depth of interest.)

5. Configure the remainder of the test-run settings as shown in Figure 12.

**Note:** The total time to reach steady state performance for small block random write traffic varies widely. Typically, SLC-based drives do so more quickly than MLC-based drives.

This example is for an SLC drive, so the 4K random write access specification is added 25 times, and the test run is executed as a starting baseline value.

The test is executed, and the results are parsed and plotted, as shown in Figure 13.

In this example, steady state is achieved after the fourth test iteration. As each x-axis mark represents five minutes, this drive reaches 4K random write steady state in 20 minutes.



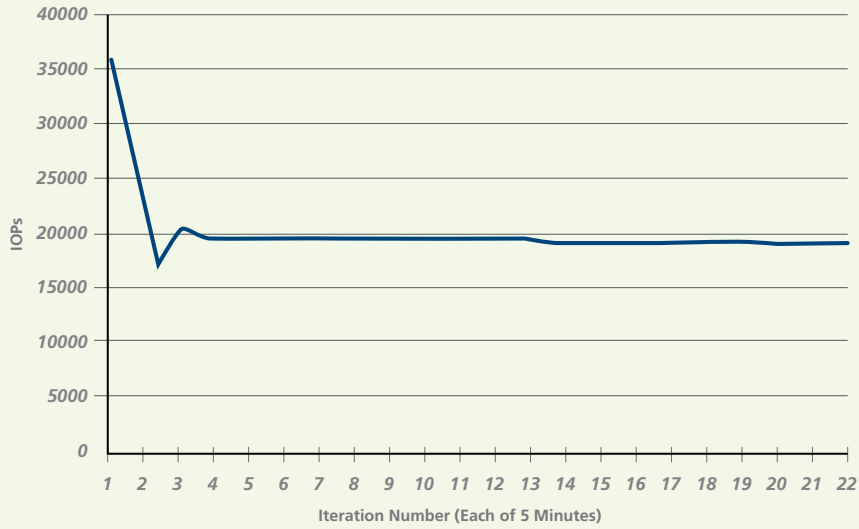


Figure 13: SLC 4K random write saturation in steady state

Element	Transfer Size	% Read	% Random	IOPs
1	4096	0	100	35460.98405
2	4096	0	100	17382.48189
3	4096	0	100	20179.53067
4	4096	0	100	19460.9479
5	4096	0	100	19568.64219
6	4096	0	100	19657.92814
7	4096	0	100	19573.7974
8	4096	0	100	19561.21846
9	4096	0	100	19572.46237
10	4096	0	100	19597.42278
11	4096	0	100	19577.95913
12	4096	0	100	19595.8534
13	4096	0	100	19442.13415
14	4096	0	100	18986.06345
15	4096	0	100	19065.45231
16	4096	0	100	18993.24672
17	4096	0	100	19016.87946
18	4096	0	100	19011.96523
19	4096	0	100	19024.61239
20	4096	0	100	19045.18973
21	4096	0	100	18990.85739
22	4096	0	100	18997.49995

Table 3: Raw IOMeter data from Figure 13 plot



## Mechanics: Test Setup and Execution – Workflow Overview

The overall workflow consists of:

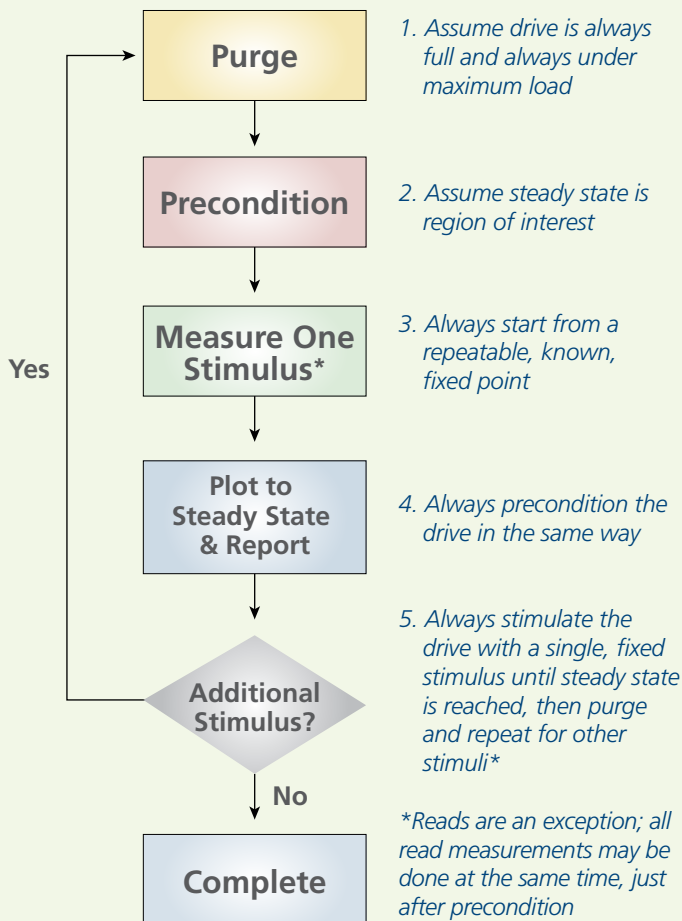


Figure 14: Test setup and execution workflow

1. Regardless of the metric being measured, each drive to be tested is put into the FOB state. This may be done via the secure erase command, a vendor-unique method, a serial port, or some other means.
2. The drive is then preconditioned with the SNIA PTS workload-independent preconditions.
3. A single, fixed stimulus is measured into the steady state region.

4. Results are then parsed and plotted; steady state performance is validated and reported.
5. If additional stimuli are of interest, the process is repeated for each additional stimuli; otherwise, the process is complete.

## Why Returning a Device to the FOB State Is Imperative – An Example of Stimulus-Sequence Variant Performance

Throughout this brief, we assert that returning the SSD to the FOB state is key to precise, consistent performance measurement. In this section, our example shows how not doing so and then simply changing the order in which the stimuli are applied can give vastly different results.

This example is an MLC intended for use in the client market, but which has been well-adopted in some enterprise applications.

The plot in Figure 15 shows steady state performance for the drive when the stimulus sequence applied is: 4K random write into steady state, then 128K sequential write into steady state, followed by 4K random write (again into steady state). The performance for each stimulus is plotted on the y-axis, with time plotted on the x-axis.

As expected, the IOPs for the 128K sequential write traffic is lower than the IOPs for the 4K random write traffic.

Note that the drive's performance for each 4K random write interval is consistent.

Note also that the drive was NOT returned to the FOB state at any point during the test—each stimulus was applied immediately following the completion of the previous stimulus.

Next, the same drive was secure erased to FOB state and the same stimulus applied, but their order was reversed. In this next plot, time is on the x-axis, performance (in IOPs) on the y-axis, but the test stimulus sequence was changed to: 128K sequential write into steady state, then 4K random write into steady state, followed by 128K sequential write (again into steady state).



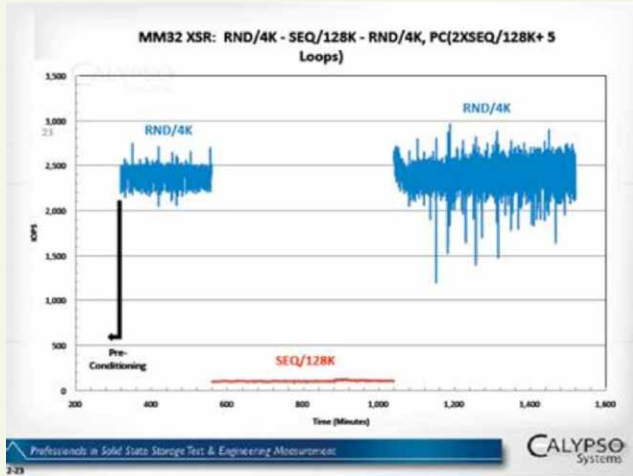


Figure 15: Stimulus sequence: Random/Sequential/Random

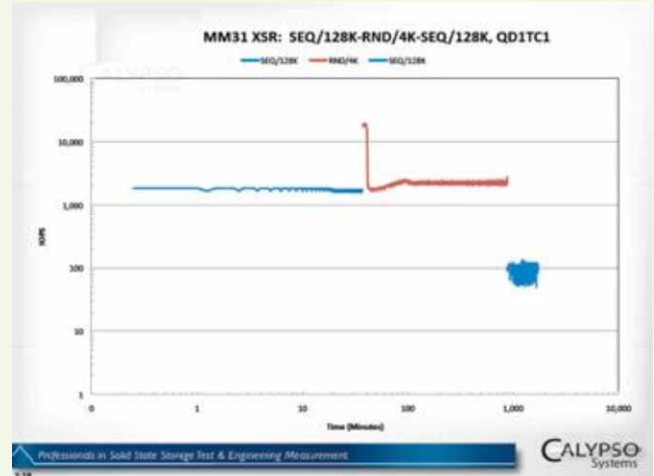


Figure 16: Sequential/Random/Sequential

*This data was collected by a third-party SSD test company.*

Based on the test data above, it would be reasonable to expect that the performance of the two 128K sequential write intervals would be the same; however, the results were quite different, as shown in Figure 16.

The performance seen in the second 128K sequential write interval, despite being written into steady state, is lower than what is seen in the first 128K write interval.

In this example, the order in which the stimuli were applied was the only variable, but the measured performance was quite different—this drive clearly demonstrates the write-history sensitivity mentioned at the beginning of this brief. If the drive had been returned to FOB in-between these tests, this variance would not have been observed, and the accuracy of the test results would have increased.

## Conclusions

Micron’s enterprise SSD performance measurement technique ensures consistent, repeatable performance measurements relevant to the enterprise customer. Through a combination of preconditioning and single-stimulus testing, we remove run-to-run and write-history-based performance fluctuation.

The method is in alignment with current industry test methods and provides a level playing field for all drives measured.

micron.com

Products are warranted only to meet Micron’s production data sheet specifications. Products and specifications are subject to change without notice.

©2011 Micron Technology, Inc. Micron and the Micron logo are trademarks of Micron Technology, Inc. All other trademarks are the property of their respective owners. All rights reserved. 10/11 EN.L P:11446

