Western Digital.

SSD Performance Benchmarking

Improving Workflow Efficiency with Fast Conditioning

September 2025

Table of Contents

Introduction

With the rise of super-capacity SSDs (30TB+ circa 2025), the process of testing these drives has become increasingly time-consuming and expensive. For instance, preparing a 64TB SSD for random write 4K testing takes about 144 hours, while a 122TB model can require up to 333 hours.

Nearly ten years ago, the WDC's Data Propulsion Lab—now known as the Western Digital Platform Applications Engineering—recognized these escalating testing costs. In response, they partnered with Western Digital R&D System Architecture (now Sandisk System Design Flash Business) to seek more efficient solutions. By late 2024, Sandisk System Design Flash Business had made substantial progress on this challenge. Platform Applications Engineering then resumed its role, serving as beta tester and quality assurance team. They validated the Fast Conditioning (FC) process through rigorous performance testing with the standard Spec Sheet Suite and now use FC as a routine part of their workflow.

The Need for Drive Conditioning

Conditioning SSDs before benchmarking is essential to ensure consistent, realistic, and comparable performance measurements, particularly under write-intensive workloads. Here's why conditioning is necessary:

Establishing Steady State Behavior

- SSDs often deliver their highest performance when new, in a condition known as "fresh out of the box" (FOB).
- As more NAND capacity is used and internal housekeeping mechanisms (such as garbage collection and wear leveling) activate, SSD performance typically declines and then stabilizes. This stabilized condition is referred to as the "steady state".
- While read performance is generally stable, steady state can cause a significant drop in write performance. This effect is most pronounced when all available physical and over-provisioned (OP) capacity has been used, and ongoing write activity exceeds the SSD's garbage collection rate. When the pool of free blocks drops to a low threshold, continuous garbage collection is triggered, capping write throughput at the garbage collection rate until write activity slows.
- Before benchmarking, the drive should be conditioned to reach a steady state. This ensures results are reproducible and representative of long-term, real-world usage.

Avoiding Inflated Benchmark Results — Without Conditioning

- Early test runs may capture artificially high throughput or low latency, especially for random writes.
- Drive performance may be mischaracterized under sustained, long-term workloads in enterprise environments.

Validating Long Term - Performance Trends

- Conditioning simulates the accumulation of data and wear over time, exposing how the controller, Flash Translation Layer (FTL), and NAND manages resources under pressure.
- It highlights thermal throttling, endurance limits, and QoS fluctuations, which are often invisible in short-term tests.

Conditioning transforms a test from a best-case scenario to a worst-case, yet realistic, scenario, which is vital for evaluating enterprise SSDs used in demanding workloads, such as disaggregated NVMe-oF storage supporting AI/ML pipelines.

Challenges with Increasing SSD Capacities

As SSD capacities continue to scale into the hundred-terabyte-plus range, industry standard testing methodologies face increasing pressure to adapt. A critical concern is the exponential rise in preparation time required to condition drives for meaningful performance measurement—particularly under randomized workloads. In contrast, measurement time remains largely fixed unless fundamental changes are made to the benchmarking methodology, which is uncommon due to standardization requirements.

For example, the Ultrastar® DC SN655 61.44TB NVMe™ SSD requires approximately 144 hours (6 days) to execute 4K random write preconditioning using the conventional full-drive methods that support a five-hour steady state measurement phase. This procedure is typically repeated three times to ensure statistical confidence in performance variation and repeatability.

With next-generation SSDs now exceeding 122TB, this challenge compounds significantly:

— Preparation time will double or more, potentially requiring over 12 days per test pass, depending on endurance characteristics and workload aggressiveness.

— Test labs face severe constraints on throughput, as parallel test capacity is limited by time, resource availability, and long-running conditioning.

As a result, the industry must evaluate new approaches to workload scaling, data reduction during preconditioning, and intelligent extrapolation techniques. Without such innovations, traditional benchmarking processes will become increasingly impractical for higher capacity SSDs.

SSD Benchmarking Overview

SSD performance testing requires two phases: Preparation and Measurement.

Preparation Phase

The preparation phase consists of two actions: a `nvme format -f` command and an action that puts the NAND in a state that induces garbage collection (this limits the maximum steady state performance for write workloads). Traditionally, the SSD is "conditioned" by the use of exhaustive sequential or random fills (or possibly a combination of both) to achieve steady state (induced garbage collection). The sequential or random writes match the access pattern of the workload as well as the blocksize. Notably, preparation time often significantly exceeds measurement time.

- The goal is to ensure reliable and repeatable testing by analyzing time series data and coded data, particularly the Coefficient of Variation (CoV) for each test.
- In SSD benchmarking (especially for enterprise-class drives), consistency is just as critical as raw performance. To evaluate this, test engineers analyze not just average performance metrics (e.g., IOPS, latency, throughput), but also the variability of those metrics over time and across repeated runs.
- CoV is a unitless statistic that normalizes variability (standard deviation) relative to the mean and provides a single number comparison between tests with different means. CoVs are directly comparable. CoV thresholds can also be set (for example, <5%) to define acceptable variation.
- A high CoV could indicate thermal throttling, garbage collection effects, or controller inconsistencies - all signs that the drive may not be ready for real-world deployment, or that it needs additional tuning.

Measurement Phase

Once Steady State is achieved, performance metrics such as IOPS, throughput, and latency are captured under stabilized conditions. Results are documented in datasheets with explicit test parameters, providing stakeholders with reliable metrics that reflect real-world SSD behavior under prolonged use.

Workload Generator and Test Harness

Flexible I/O Tester (fio) is a widely adopted open-source benchmarking tool designed to evaluate the performance characteristics of storage devices under configurable I/O workloads. Originally developed by Jens Axboe for Linux® block devices, fio supports a broad range of platforms—including Windows®, FreeBSD™, macOS®, z/OS®, and many others — and can simulate various workload types such as sequential and random reads/writes, mixed I/O, and trim operations.

fio provides granular control over parameters like block size, queue depth, I/O engine, and runtime, making it ideal for emulating real-world application behavior. Fio outputs detailed performance statistics, including IOPS, latency, bandwidth, and error counts, which are essential for assessing SSD behavior under preconditioning, steady state, or failure scenarios. Its scripting ability and extensive plugin support make fio a core utility in enterprise SSD qualification, system tuning, and performance validation workflows. fiodb is a universal test harness that can run and manage many utilities, including fio. In fact, most fiodb tests use fio as the workload generator.

Fast Conditioning: Addressing the Preconditioning Challenge

In 2016, the Western Digital Data Propulsion Lab (DPL) (now Western Digital Platforms Application Engineering) initiated discussions within Western Digital to explore solutions to the ever-increasing and expensive SSD device conditioning challenge. In 4Q2024, Western Digital R&D System Architecture (now SanDisk System Design Flash Business) developed a new process dubbed Fast Conditioning (FC) to address lengthy preparation times. It delivered their beta version of an FC solution also known as sprandom or sprandomize. In 2025, Platforms Application Engineering collaborated with the SanDisk System Design Flash Business by reviewing, conducting extensive testing and analysis, and discussing features and futures. For example, the SN655-61.44TB FC reduces Preparation Time from ~144 hours to less than 7 hours, a reduction of 95%.

- FC has been applied successfully to 4K Random Write conditioning, which traditionally takes the most time and, at this time, is the principal focus.
- For 128K Sequential Workloads, FC should reduce Preparation Time by approximately 50% by essentially eliminating one of the two sequential fills currently used.
- Fast Conditioning Impact Predictions.

The table below shows a first-order prediction of the impact of Fast Conditioning on the Ultrastar SN655 Product Family. The last three rows of this table show the Standard and FC Conditioning time, as well as the reduction in time.

Product Information for Ultrastar SN655 Product Family and First Order Estimates of FC Impact

Capacity	3.84TB	7.68TB	15.36TB	30.72TB	61.44TB
	Performance and Projections				
Read Throughput (max MB/,Seq 128KiB)	6,800 MB/s	6,800 MB/s	6,800 MB/s	6,100 MB/s	4,300 MB/s
Write Throughput (max MB/,Seq 128KiB)	2,600 MB/s	2,000 MB/s	3,700 MB/s	3,400 MB/s	3,150 MB/s
Read IOPS (max, Rnd 4KiB)	1,000 KIOPs	980 KIOPs	1,100 KIOPs	1,052 KIOPs	890 KIOPs
Write IOPS (max, Rnd 4KiB)	112 KIOPs	80 KIOPs	135 KIOPs	66 KIOPs	29 KIOPs
Read Latency (µs)	80µs	78µs	125µs	110µs	115µs
Write Latency (µs)	10µs	15µs	10µs	15µs	40µs
Standard Conditioning Time (4K RW)	2.33 H	6.51 H	7.72 H	31.57 H	143.68 H
FC.fio Preparation Time	0.41 H	1.07 H	1.15 H	2.51 H	5.41 H
Percentage Reduction in Time	82%	84%	85%	92%	96%

Impact of FC on Spec Sheet Random Suite (SSR): Actual Data

The table below shows data from an SN655-61.44 TB SSR test using the Fast-Conditioning methodology. The bottom line is that the Percentage Reduction in Time is very close to the projections shown in the table above. The projection displayed a time reduction of 96% versus 94% for the actual measurement.

Ultrastar SN655 61.44TB Data Center Time Savings for 4K Random Workload Spec Sheet Suite

Summary	Exhaustive Conditioning Method-A		Fast Conditioning Method-B	
Workload	Time Required in Hours-A	Percent of Total-A	Time Required in Hours-B	Percent of Total-B
Sequential Conditioning	15:37:09	8.94%	0:00:00	0.00%
Peak Performance Testing	0:25:05	0.24%	0:25:05	3.95%
Random Conditioning	154:45:40	88.54%	6:09:45	58.23%
Steady State Testing	4:00:12	2.29%	4:00:12	37.82%
Totals	174:48:06	100.00%	10:35:02	100.00%
Percentage Reduction in Time	—	—	93.95%	—
Net Reduction in Time (Hours)	—	—	164:13:04	—

Test Results: Coded Data

The test results are reported in two categories:

- Coded data (generally average performance) and
- Time series data (that can reveal much finer detail than coded data).

Only Phase 7: FC Results and Phase 8: RW 4K Results of the benchmark process are provided. These results are compared to the Data Sheet results. Since FC operates at full sequential bandwidth, this metric is compared to the FC Results. For RW 4K, it is another commonly reported metric.

- The FC and RW 4K results are nearly identical to the Data Sheet results.
- While FC Results and RW 4K results are the focus of this discussion, as the intersection of these tests in the time series provides the most direct demonstration of efficacy.
- All other phases were also reviewed, and they too matched almost identically the Data Sheet metrics and other performance expectations almost identically.

Additionally, the performance relationship is shown in the form of “n.nnnX” where “n.nnn” is the ratio of FC Results and Data Sheet Results while “X” signifies a ratio. Thus, the FC Results are 1.008X, and this indicates that FC performance essentially matches Data Sheet performance. This notion, “n.nnnX” can be thought of as a speedometer. Where values greater than one indicate that the new method is faster than the baseline, in this case, the Data Sheet Metric. Values larger than one indicate faster, and less than one indicates slower.

Data Sheet Results vs. Phase 7: FC Results and Phase 8: RW 4K Results

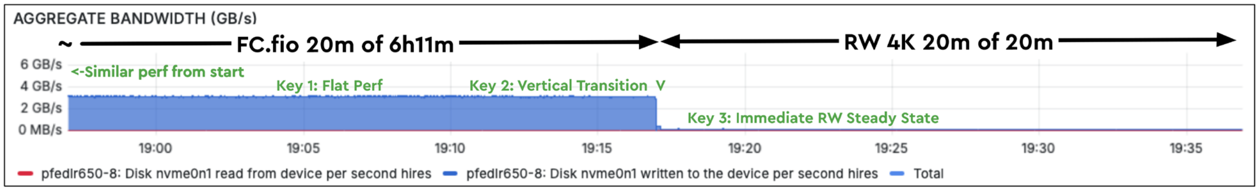
Metric	Phase 7: FC Results	Phase 8: RW 4K Results
Data Sheet Rate	3.125 GB/s	29.0 KIOPS
FC Data Rate	3.150 GB/s	28.5 KIOPS
Ratio (FC Data Rate/Data Sheet Rate)	1.008X	0.983X
Slope	-0.000179	-0.001922
Coefficient of Variation (CoV)	0.233%	1.330%

Test Results: Time Series

- The three time series charts below each cover the same time period of about 40m comprised of the last 20m of the 6h11m Spec Sheet Random (SSR) Phase 7 FC.fio and the 20m of SSR Phase 8 RW 4K.
- It is the intersection of the FC and RW 4K workloads that best demonstrates the efficacy of the solution. The FC should fill the device at the Data Sheet Sequential Write Rate, while the data rate of the RW 4K also matched the Data Sheet RW 4K rate.

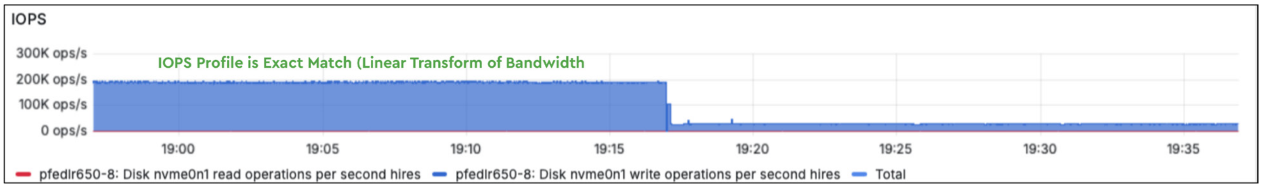
Bandwidth

This chart provides a visual summary of the test at this crucial transition point. The green text shows the points that were mentioned above. These results are prototypical.



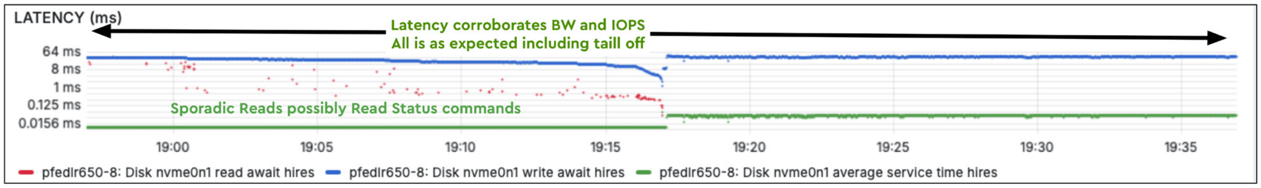
Test Results: IOPS

The IOPS chart is a linear transform of the bandwidth chart. The chart has the same general shape, and all comments and observations from the bandwidth chart also apply here. The IOPS chart is provided because the RW 4K workload is specified in IOPS, and it is appropriate to also show these results as IOPS.



Test Results: Latency

While the Bandwidth and IOPS charts are fraternal, the Latency chart offers additional information that is corroborative and provides additional information. For example, while Phases 7 and 8 are write workloads, this chart shows some sporadic reads in Phase 7. These are likely Read Status and other Command and Control IOs and do not affect overall latency. More importantly, we see that at this point in the FC.fio phase, there is a slight downward slope, indicating that the latency is decreasing while the bandwidth is constant. Since latency (aka response time) is the sum of queue time and service time, as this test is winding down, there are fewer IOs in the queue, and this would reduce the latency near the end of the test (this chart also shows Service Time that remains constant. At the very end of the test, the latency drops more aggressively as many fio processes are finishing their work, and this reduces the queue time of the IO and therefore improves the latency (response time). This is expected behavior.



Implications

The time required for the Preparation Phase is a multi-valued function including but not limited to: SSD capacity, Random and Sequential write performance, architecture, NAND (SLC, MLC, TLC, QLC, etc.), workload details, and for attached storage, the network infrastructure and storage solution, etc.

- Various SSD vendors have announced 122.88TB SSDs. One of those vendors advertises that 4K Random Write performance is 25 KIOPS, requiring ~333 hours (14 days) for traditional preparation. With FC, this time is estimated at 11 hours, a reduction of ~96%.
- Testing the DapuStor R5100 15.36TB SSDs using the FC process has shown results similar to the Ultrastar DC SN655. However, some devices may be FC-resistant due to deviations from contemporary designs.

A deeper understanding of fast conditioning and its broader applications is likely to increase. Testing confirms that performance is very similar to the traditional method (exhaustive fills) and the Fast Conditioning (one precise fill at full Sequential Write speed) that creates a similar SSD internal structure to that produced with exhaustive fills.

However, the long-term efficacy of Fast Conditioning may be influenced by several factors:

- Architectural shifts that redefine block management behavior.
- Advances in NAND management algorithms or controller intelligence.
- Integration of Fast Conditioning-like routines natively within device firmware or hardware.
- While the storage ecosystem tends to evolve incrementally, the growing diversity in SSD design and increasing capacities warrant periodic revalidation. Nonetheless, based on current product roadmaps and widespread architectural continuity, Fast Conditioning is poised to remain a reliable and efficient alternative for qualifying enterprise-class SSDs.

Implementation

- FC has been seamlessly integrated into fiodb, our test runner and manager. The final integration will be the default fallback to the traditional methods if the exact FC.fio file is not available.
- FC.fio files must be created and validated for all permutations of the devices, including capacities, and workloads.
- This process requires some device details that, if not readily available, must be deduced through testing that probes the device to glean data like the exact physical capacity of the device (Logical Size + Over Provision (OP) Size). Without this and related information, this may devolve into somewhat of an educated guess.
- The Spec Sheet Collection (SSC) process contains two test suites: the Spec Sheet Sequential (SSS) that uses 128K Sequential Workloads and Spec Sheet Random (SSR) that uses 4K Random workloads.
- It should be noted that only write workloads have to be accounted for, as only writes have a performance drop (i.e., write cliff) when performance is reduced by a scarcity of free blocks. When this occurs, the steady state performance is defined by the rate at which free blocks are made available through garbage collection (GC). Read workloads are generally unaffected when there is a scarcity of free blocks. Reads should never wait for free blocks and therefore, should not slow.
- Tests using FC designated templates will automatically use FC if the device ID for the device under test is found in the list of devices that have a validated FC.fio file and exist for Random 4K workloads for the following devices:
  - Ultrastar SN655 Product Family is validated.
  - DapuStor R5100-15.36 TB is also validated.
- The reason for initial use in random 4K workloads is that preparation for these workloads using traditional methods can take weeks for the largest devices.
- Sequential workloads will also be addressed after implementing the random workloads. The benefit to sequential workloads is less. There are typically two sequential fills for these workloads. FC will reduce this to one fill at maximum write speed or a 50 percent reduction in conditioning time.

Futures

- Initially create FC.fio files for the largest SSDs and work through available sample sets.
- Currently, the process to do this is not well defined.
- With more experience and tools that will inevitably be created, this workflow will become further defined.

Technical Insights

SSC Overview: Traditional and Fast Conditioning

Below is a table that compares the Traditional Spec Sheet Suite benchmarking with the new Fast Conditioning method. This data shows all 30 phases (individual FIO workloads) and identifies each performance change associated with the exploitation of Fast Conditioning. Although the table indicates the removal of exhaustive fill processes, stubs of each test remain to keep all phase numbers unchanged; historical and future data access will remain consistent. In terms of extensibility, new phases (individual tests) can be added to the end of the test with no impact on other phases and the reporting structure.

The SSC is a reliable and repeatable process for measuring the performance of storage devices and solutions. The results are used in Western Digital Data Sheets. fio is the workload generator. The SSC has two tests, one for sequential (Spec Sheet Sequential - SSS @128KB) and one for random (Spec Sheet Random - SSR @4KB). The SSS has 10 phases (unique tests), and the SSR has 20 phases.

The process is simple:

- Before either suite (SSS or SSR), run a **`nvme format -f`** command on all devices under test. This action puts the device(s) in a well-known state, and it is an essential action in this process.
- Run the desired test(s) three times and to provide enough observations for valid analyses.
- Monitor the real-time results with **`iostat -mtxy 1`** so that the time series can be reviewed for anomalies.
- Capture all **`fio`** output and other relevant instrumentation.

Examining the Coefficients of Variation (CoV) is a key statistic for understanding the stability of the tests. Finally, a peer review with knowledgeable performance engineers is required. The rightmost column in this chart shows the changes for Fast Conditioning and is highlighted in yellow where there are changes.

Spec Sheet Suite Overview: Traditional and Fast Conditioning

Spec Sheet Sequential (SSS) @128KB

Suite	Phase	Phase Description	Traditional Method Time per Phase	Fast Conditioning Time per Phase
SSS	0	sw-fill	sw-fill	FC.fio-SW-128K
SSS	1	sw-fill	sw-fill	—
SSS	2	sw-js	20m	20m
SSS	3	sr-part	20m	20m
SSS	4	sr-js	20m	20m
SSS	5	sw-part	20m	20m
SSS	6	sw-js+qd64	20m	20m
SSS	7	sr-js+qd64	20m	20m
SSS	8	sw-js-32K-qd64	20m	20m
SSS	9	sr-js-32K-qd64	20m	20m
SSS Totals	10	—	(sw-fill*2@128KB) + 160m	FC.fio-SW-128K + 160m

Spec Sheet Random (SSR) @4KB				
Suite	Phase	Phase Description	Traditional Method Time per Phase	Fast Conditioning Time per Phase
SSS	0	sw-fill	sw-fill	—
SSS	1	sw-fill	sw-fill	—
SSS	2	rw-qd256	5m	5m
SSS	3	rm-qd256	5m	5m
SSS	4	rr-qd256	5m	5m
SSS	5	rw-qd1	5m	5m
SSS	6	rr-qd1	5m	5m
SSS	7	rw-fill	rw-fill	FC.fio-RW-4K
SSS	8	rw-qd256	20m	20m
SSS	9	rm-qd256	20m	20m
SSS	10	rr-qd256	20m	20m
SSS	11	rw-qd1	20m	20m
SSS	12	rr-qd1	20m	20m
SSS	13	rw-js8	20m	20m
SSS	14	rr-js8	20m	20m
SSS	15	rm-js8	20m	20m
SSS	16	rw-js4	20m	20m
SSS	17	rr-js4	20m	20m
SSS	18	rw-js2	20m	20m
SSS	19	rr-js2	20m	20m
SSS Totals	20	—	(seq-fill*2@128KB) + (rand-fill@128KB) + 265m	FC.fio-RW-4K + 265m
Grand Totals	30	—	(seq-fill*2@128KB) + 160m + (seq-fill*2@4KB) + (rand-fill@4KB) + 265m	FC.fio-SW-128K + 160m + FC.fio-RW-4K + 265m

The Fast Conditioning (FC) process is technically complex and has been developed over the last 15 years to refine SSD conditioning methods. The goal of FC is to drastically reduce the time required for SSD conditioning using a quasi-random procedure that operates at the maximum sequential write rate. It marks all blocks in a manner like explicit conditioning (exhaustive random fills) or implicit conditioning (workload-driven conditioning). Below is a streamlined breakdown of the process. While shortcuts exist, the primary objective is to maintain simplicity -- ensuring the process remains understandable, explainable, maintainable, and extensible.

Sources for required data:

- Vendor documentation
- Information accessible within the device (e.g., via nvme list all)
- Data supplied by vendors or publicly available online
- Observations from probing the device using write tests, performance analysis, random walks, or exhaustive trials

FC.fio Files

FC.fio files are fio scripts that simulate explicit or implicit workloads by writing data in specific patterns. However, these files are not readily available, and no general closed-form solution exists for systematically creating them. As FC adoption grows, its methodology will become clearer—but no defined timeline for this progression currently exists.

Initialization

Reliable, repeatable processes require well-defined starting conditions, forming the backbone of high-performance testing.

- All tests begin with a nvme format -f command, resetting the block maps to a state like Fresh Out of Box (FOB).
- Over time, NAND chips naturally degrade due to wear and errors, leading to the removal of bad blocks. However, reasonably new devices do not experience substantial performance degradation due to this wear.
- The process starts with an entirely unmapped, empty drive.

Key Parameters

- Exported Logical Capacity – The total logical space available on the drive.
- Over-Provisioning (OP) – The amount of OP allocated within the SSD (expressed as a count (OPC) or as a percentage (OPP).

Physical Capacity Calculation

Physical capacity can be estimated using the formula:

PC ≈ LC + OPC

- If two of these values are known, the third can be calculated.
- Using Logical Capacity and OP, the physical capacity of the SSD can be determined.

Segmentation

- Larger devices contain more write regions.
- Physical capacity is divided into N regions, determined by: **`Segment Size = Physical Capacity / N`**

Overlap Design

- N fio jobs execute quasi-random (not strictly random) writes across the SSD. The precise behavior is still under analysis.
- Each of the N regions overlaps with the previous one.
- Overlap size depends on the OP allocated in the prior region.

Logical Address Assignment

The writes to each of the N regions following the rules below.

- The device is written once at the full sequential write rate and is then in steady state.
- Within each region:
  — Every logical address is written exactly once.
  — The addresses are written in a pseudo-random order (similar to a shuffle).

Validate the Efficacy of the FC

Until FC methodology is fully validated, verification is required.

- Examine FC-coded data results and time-series data.
- Expected characteristics of the performance curve:
- Close to the sequential write rate of the SSD.
- Flat (no slope) during FC with a low Coefficient of Variation (CoV).
- Sharp, vertical drop-off at the end.

The closer the observed behavior aligns with this ideal model, the more reliable FC becomes.

Starting the Workloads

The first step in testing system accuracy involves initiating write workloads.

- Run a RW 4K workload—if FC is appropriately configured, it should quickly reach steady state, aligning with RW 4K datasheet specifications within minutes.
- Once the RW 4K performance is verified, proceed with additional workloads:
- Random Mixed (RM)
- Random Read (RR)

Performance Monitoring

- Use iostat -mtxy 1 to track performance in real time.
- Analyze fio-coded data results—particularly the Coefficient of Variation (CoV): **`CoV = Standard Deviation / Mean`**
- Conduct post-test regression analysis to ensure a near-zero slope, confirming steady state.

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
www.westerndigital.com

September 2025 / Revision 01