PERFORMANCE EVALUATION GUIDE

Solidigm™ P41 Plus

August 2022
Revision 001

Ordering Information

Contact your local Solidigm sales representative for ordering information.

Revision History

Revision	Description	Date
001	• Initial release	August 2022

Solidigm may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Solidigm reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

Nothing herein is intended to create any express or implied warranty, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, or any warranty arising from course of performance, course of dealing, or usage in trade.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase.

Some results have been estimated or simulated using internal Solidigm analysis or architecture simulation or modeling, and provided to you for information purposes only. Any differences in your system hardware, software or configuration may affect your actual performance.

Your costs and results may vary.

Contact your Solidigm representative or your distributor to obtain the latest specifications and before placing your product order.

For copies of this document, documents that are referenced within, or other Solidigm literature, please contact your Solidigm representative.

© Solidigm. “Solidigm” and the Solidigm logo are trademarks of SK hynix NAND Product Solutions Corp (d/b/a Solidigm). “Intel” is a registered trademark of Intel Corporation. Other names and brands may be claimed as the property of others.

Contents

About This Guide

If you are evaluating platform or individual component performance in real-world or synthetic scenarios, then this guide presents considerations and recommendations when testing Solidigm solid state drives.

This guide also provides an overview of different workloads and benchmarks relevant to Solidigm solid state drives.

The target audience includes publications, OEMs, technical analysts, academia and any who plan to test or evaluate Solidigm solid state drive performance.

This guide is divided into the following sections:

- **Overview** provides an overview to establish a foundation for testing.
- **Benchmarks** explores software designed to mimic a workload on a component or system and provide an indicator of performance.
- **Setup** This section includes a testing quick start plus examines the system and component configuration settings used to measure performance of a platform or component in real-world or synthetic scenarios.
- **Analysis** explores approaches for analysis and handling variability in result sets.
- **Appendix** presents supplementary material not included in the other sections including discussions on executing a set of tests and covers considerations for specific benchmark tools.

Tool References

Product	Download Link
Solidigm Storage Tool	https://www.intel.com/content/www/us/en/download/715595?v=t
Iometer	http://www.iometer.org/
CrystalDiskMark	https://crystalmark.info/en/category/crystaldiskmark

1 Overview

1.1 Performance Testing

Solidigm recommends an approach to performance evaluation should be reproducible via controls and observation to ensure fair analysis. This guide considers general principles of performance testing and applies them to Solidigm solid state drives.

Consider the below Scenarios and Workloads sections as one set of perspectives for how to approach performance evaluation. This perspective is the basis of methodology used at Solidigm.

1.2 Scenarios

The **real-world** scenario for evaluating performance is to measure common usage where machines are configured as they are out of the box or as in general practice for most users. Simply put, the test evaluates real world usage as machines are naturally configured. Generally, real-world scenarios are associated with **system** or **platform** testing.

The contrasting scenario is to configure machines in a **synthetic** state for measuring usage under manufactured conditions that are uncommon or artificially controlled. “Synthetic” here does not necessarily mean "best" for a component or platform, nor does this use of “synthetic” describe a category of benchmarking tools. In a synthetic configuration for testing, you may be exploring a boundary condition of the platform which may put some components at an advantage and others at a disadvantage. The synthetic scenario goal is to evaluate behavior of the test target for a specific condition or set of conditions that attempt to imitate a potential state. Generally, synthetic scenarios are associated with **device** testing.

1.3 Workloads

Workloads refer to a set of activities & any sequencing. Workloads for performance evaluation can be characterized in a variety of ways:

- **Target**: component (device) or platform (system).
- **Use Case**: categories of usage from the end-user perspective.
    - *Productivity*: Usage of common office or productivity applications such as the Microsoft Office suite of applications, Web Browsers, and "light" audio/photo editing. This category of applications generally "produces" something.
    - *Content Creation*: Professional software and usage by an enthusiast or professional such as Autodesk Computer Aided Design, 3D/Audio/Video editing & rendering among other "large" workloads where something is created.
    - *Gaming*: Consumer gaming relies on multiple components and processing for a responsive experience. AAA games typically push the boundaries of audio/visual/interactive/processing capabilities of a system.
- **Tasking**: single or multi-tasking; also thought of in terms of foreground and background tasks.

1.4 Solidigm QLC SSD Overview

The Solidigm™ P41 Plus is the fourth generation of SSD based on Intel's advanced QLC (Quad-level-cell) NAND technology and is designed to offer the performance required for client platform usages such as web surfing, office and home productivity, media entertainment, etc. The product delivers Solidigm's industry leading quality and reliability, giving a 1.6 million hour mean time between failures (MTBF) under a 5-year limited warranty.

In everyday PC usage scenarios such as Windows search, web surfing, e-mail composition, writing project reports or delivering a presentation, the storage device on the PC is mostly idle for a significant portion of time. Solidigm™ P41 Plus, built using Intel® QLC (Quad-Level-Cell) NAND technology, is designed to thrive in these usage scenarios. Solidigm™ P41 Plus offers a large SLC span, which boosts performance of common PC workloads that are clustered in bursts between lengthy storage idles, while also providing higher capacity benefits of QLC NAND technology.

Solidigm™ P41 Plus SLC Cache Architecture

Solidigm™ P41 Plus, built using Intel® QLC NAND technology, enables 4-bits of storage per cell. The QLC NAND allows 33% more storage in the same area compared to TLC NAND. Compared to TLC NAND cell, the QLC NAND cell has a higher write and read latency. To mitigate the impact of higher latency of the QLC NAND, Solidigm has designed and implemented an innovative SLC cache, which is variable based on available unused capacity on the SSD, that delivers faster sequential and random data writes.

Data writes from the host will be directed through the variable size, high speed SLC cache and offer faster data writes through the full span of the SLC cache. When the SLC cache is reaching its full capacity, the drive firmware moves the contents of the SLC cache into the available QLC cells. The firmware also has an intelligent algorithm to opportunistically move the data from SLC cache to QLC cells during the drive idles. This feature allows the user to experience a smoother and refreshed experience from the SSD. The variable SLC adjusts in size bi-directionally based on the available capacity of drive. The drive is also designed with a dedicated static SLC cache to provide faster write performance when the drive is limited available capacity.

2 Benchmarks

The following figure shows the mechanical information for P41 Plus. All dimensions are in millimeters.

Benchmarks are designed to mimic a workload on a component or system and provide an indicator of performance. This section provides an overview of three basic types of benchmarks that can be used to measure system storage performance:

- Application-based Benchmarks
- Trace-based Benchmarks
- Synthetic Benchmarks

For specific benchmark recommendations and information please see the Benchmark Recommendations subsection in the appendix.

2.1 Application-based Benchmarks

Application-based benchmarks emulate end-user usage using scripted execution of real-world programs on a system. Application-based benchmarks measure the load and execution time of these applications and present the results as a score. Often the scores of applications that are common to a usage are combined into a subsystem score. These subsystem scores are reported along with an overall performance score that is the combination of the subsystem scores.

SYSmark 25 and UL Procyon are examples of application-based benchmarks.

Application-based benchmarks are helpful in determining the User Experience (UX) for a given system. In particular, the scores from an application-based workload are more likely to reflect the real world UX than the results from a purely synthetic workload.

One disadvantage of application-based benchmarks when trying to determine storage device speed is that they focus on CPU, memory, and graphics performance and may not properly weigh storage sub-system speed in the results. They also may not consider end-user perceivable delays such as application loads. Another disadvantage is that the scripted nature of an application-based benchmark is fixed, and the application workload may not be representative of a specific end-user usage model, especially when multiple iterations are repeatedly in sequence. Lastly, due to practical considerations, rather than constrain the total allowable runtime and total size of the benchmark (e.g., download based distribution, total amount of disk space required to run), many application-based benchmarks have a short-run duration and limited storage device usage. Therefore, these benchmarks may not be representative of storage device usage over time as a practical basis.

2.2 Trace-based Benchmarks

Trace-based benchmarks use traces, or recordings of disk I/O operations executed during a certain period of real use or script-based use. The trace is then used to "playback" the system I/O sequence on the drive to be tested. Trace-based benchmark results vary in format.

PCMark 10 Storage Benchmark test is an example of a trace-based benchmark. Trace-based benchmarking has many of the advantages of application-based benchmarking if the trace is collected from real-use or a realistic script-based activity. It has the further advantage of highlighting disk I/O behavior while avoiding the bottlenecks caused by the CPU, graphics, and memory subsystems.

One disadvantage of trace-based benchmarking is that the recorded trace may not reflect the true long-term usage of the storage device over weeks or months of time.

2.3 Synthetic Benchmarks

Synthetic benchmarks measure raw drive input/output (I/O) transfer rates. These benchmarks typically use well-defined, synthetic workloads and target only specific components such as solid state drives (SSDs), hard disk drives (HDDs) or other networking devices. These benchmarks format results as raw megabytes (MBs) in Input / Output Operations per Second (IOPS).

Storage subsystem synthetic benchmarks focus on drive performance without considering bottlenecks from other subsystems such as CPU, memory, or graphics. This makes these benchmarks useful for measuring drive performance for changing parameters such as transfer sizes or queue depths. However, because these benchmarks exercise components and systems in ways that do not reflect system usage models, the results may not reflect real-usage cases. For example, one SSD might have better synthetic benchmark scores for 512 KB random reads than other SSDs, but 512 KB random reads may not be a good indicator of overall system performance because of the rarity of that particular I/O access size in what is important to a given end user experience.

Synthetic benchmarks can be divided into two sub-groups: long and short.

- Long duration synthetic benchmarks measure performance variation over the entire run time. Iometer is a variable length benchmark that when properly configured can perform as a long duration synthetic benchmark.
- Short synthetic benchmarks are commonly used to measure component performance in the immediate present. These typically have limited configurability to see performance variations over time. CrystalDiskMark is considered a fixed, short duration synthetic.

2.4 Benchmark Considerations

To emulate real end user behavior, usage, and temporal state, there are three main challenges with typical application and trace based benchmarks: controlling the repeatability of results, accounting for the effects of “system aging”, and accurately modeling the storage footprint.

The repeatability of results can be challenging simply due to the complexity and uncertainty of the system with background processes and other runtime services. In addition, the starting state of the cache should be in a known state in order to ensure repeatability.

Accounting for the effects of “system aging” can be thought of as three sub-challenges. First, in normal system usage, the unused part of main memory will be preloaded with a complex set of applications, services, OS references, and prefetched data in anticipation of upcoming usage. The state of system memory will directly impact benchmark behavior. Second, multi-iteration benchmarks perform the same sequence of actions in repeated iterations, a behavior not often found in end users. The result of such benchmarks is an unrealistic situation where a modest amount of main memory can contain most of the necessary data and accessing storage isn’t required. Finally, accurate steady state performance measurements of a storage subsystem that implements any type of caching requires emulation of a warm-up sequence of storage I/O activity.

The storage footprint of users' systems will vary and can be quite large. Emulating a storage footprint in a workload or benchmark can be challenging because of the sheer size of the files required to achieve it.

3 Setup

This section examines the system and component configuration settings used to measure performance of a platform or component in real-world or synthetic scenarios.

3.1 Quick Start

This section recommends a set of steps to quickly get started with testing and evaluating the drive.

Plan

- Consider the audience, context, and desired learnings to plan your test.
- Identify whether your test target is the platform or a component in a real-world or synthetic scenario.

Setup

- In most cases you will either (1) run a benchmark or (2) run a workload noting the start and stop times.
- Prepare your system and storage component as dictated by your test goals, referring to the entries in the System Setup section.
- See the Benchmark Recommendations section in the appendix for a list of benchmarks relevant to storage recommended by Solidigm. Each benchmark noted in the appendix entry includes a description and tips for running.

Execute

- Run your test(s) and capture the results.

Analyze

- Look at the Analysis section of this guide for Solidigm's recommendations on interpreting and analyzing the results.

3.2 Setup Context

> When evaluating performance, Solidigm recommends configurations that reflect the way machines are configured when used in the field. For the tables below these are the settings in the "Real World Value" column.

To ensure consistent and repeatable measurements, a stable and deterministic environment is desired. Each configuration entry listed in the tables and their associated settings have been evaluated for their impact on performance results and variability. If disabling a technology will reduce run to run variance without materially impacting performance results it will be disabled. If disabling a feature will create a meaningfully higher or lower performance score it will be left untouched from “normal” OEM default settings and Solidigm recommends additional runs to enable a statistical analysis to filter out the variability. Some settings are disabled because they have a substantial impact on performance but infrequently execute, an example being “Automatic Windows Updates”. Disabling these types of features avoids having to debug unexpected results.

> When evaluating performance with some workloads, for example low queue depth synthetic IO, it is critical to ensure that the platform features do not mask the performance of the device.

For the tables below the settings in the "Synthetic Value" column create a test environment that maximizes the performance of the storage device. C-States, for example, are critical to the ability of modern mobile platforms to save power and extend battery life. But because having C-States enabled may have a secondary effect of causing the CPU to be less responsive to IO completion, they are disabled.

3.3 All Other Things Equal

> Make sure your tests are setup the same or as intended to ensure results do not vary unexpectedly or draw conclusions from hidden causes.

Ceteris paribus is a Latin phrase that translates to "other things equal". Assume that unless otherwise mentioned "all other things being the same" is implied for test setup. This is especially important when comparing different test configurations. Care should be taken to identify independent (changed or controlled) and dependent (tested and measured) variables. Recording these variables is essential to ensuring we have a reproducible test scenario that can be objectively verified.

3.4 System Setup

The below setup items are recommended to prevent run-to-run and system-to-system variability that impact the ability to reproduce the test and results. If one of the below items is a controlled variable or the target of a test, then modify as desired.

3.4.1 Processor

Processors should be consistent with the use case of the test however different chips will have different platform level performance characteristics. Relevant specifications are base clock, turbo clock, number of cores, hyperthreading, and package power. For example, the Intel® Core i3-9350K has a base clock of 4.0 GHz, where the Core i7-9700K has a base clock of 3.6 GHz. The i3 is a lower bin chip but could potentially exhibit better storage performance on some workloads due to the faster base clock.

3.4.2 Memory

System memory should be consistent with the use case of the test, and consistent across multiple system configurations.

3.4.3 Storage

Some traditional benchmarks and tests may involve pre-filling the SSD with data, and immediately followed by further writes to evaluate performance. Although it is a good way to stress the SSD to understand its characteristics, the testing methodology may not reflect the actual use case scenarios. Solidigm recommends that sufficient idle times be given, to mimic the everyday PC usages, especially on tests that expect a heavy pre-fill condition. Because of the advanced SLC cache architecture in P41 Plus, the idle time will refresh the SLC cache by transferring the data into available QLC media and will ensure a better user experience.

Device Preparation

These steps detail the procedures necessary in order to prepare the devices so that they are updated and in a steady state before executing the measurements.

- Verify that the Solidigm™ P41 Plus SSD has the latest firmware applied.
    - To view and update the current firmware installed on the SSD, use Solidigm™ Storage Tool
- Perform Secure Erase on the drive for out-of-the box performance.

Cache Flush

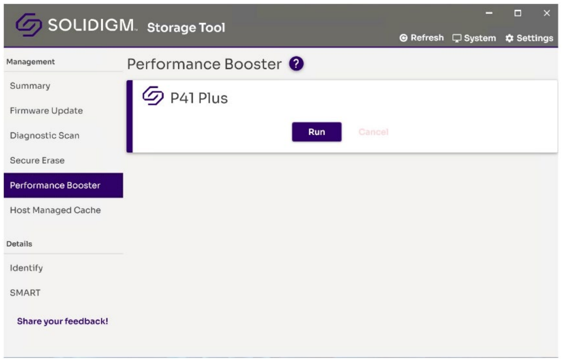
> The goal in flushing SLC is to ensure cache consistency across multiple test sets.

Solidigm™ P41 Plus has an SLC cache in addition to the QLC media.

The SLC cache should be flushed for the authenticity of test conditions when balancing consistency and reproducibility of results against the complexity of setup of in a practical time. Solidigm recommends a flush of SLC as a part of initial test preparation. This process will ensure that the entire SLC cache is available to the user for test data.

The SLC cache on the P41 Plus can be manually flushed using the Solidigm Storage Tool:

1. After installing the tool from https://www.intel.com/content/www/us/en/download/715595?v=t open tool and choose the Solidigm™ P41 Plus from storage device menu.
2. Choose Performance Booster from the Features dropdown menu.



3. Click the “Run” button.
4. The status bar indicates how much cache flush is completed. Ensure 100% of process is completed.

3.4.4 BIOS

Note: **Occasionally an OEM will not expose some BIOS settings.**

See OEM BIOS documentation for support on setting details. All other settings should *generally* be left to default with any deviations documented along with an explanation of impact.

> In the following table a Real World and a Synthetic value are given for each. As you may recall, a *synthetic* value may be employed for specific test scenarios while *real-world* indicates how the system should be normally configured as a default for practical usage.

Configuration Item	Real World	Synthetic
Hyper Threading	Enabled	
EIST (Enhanced Intel Speed Step Technology)	Enabled	Disabled
Intel Turbo Mode	Enabled	
PCIe ASPM (Active State Power Management)	Enabled	Disabled
C-States	Enabled	Disabled
P-States	Enabled	Disabled

3.4.5 Operating System

While system and device tests should normally be performed across a standardized OS installation, in some test scenarios it may be more appropriate to evaluate OEM systems in their 'out of the box' configuration, to retain any OEM-specific tuning, etc. When testing multiple systems, ensure the installed KB's (or Microsoft Windows patches) are as consistent as possible. Below is a table of relevant configuration settings for Windows 11.

In the following table a Real World and a Synthetic value are given for each. As you may recall, a *synthetic* value may be employed for specific test scenarios while *real-world* indicates how the system should be normally configured as a default for practical usage.

Configuration Item	Real World	Synthetic
Settings > Accounts > Sign-in Options > Additional settings > "Automatically save my restartable apps and restart them when I sign back in"	Disabled	
Settings > Privacy & Security > Windows Security > Virus & threat protection > Manage settings > Real-time protection[1]	Off	
Settings > Privacy & Security > Windows Security > Virus & threat protection > Manage settings > Cloud delivered protection	Off	
Settings > Privacy & Security > Windows Security > Virus & threat protection > Manage settings > Automatic sample submission	Off	
Settings > Privacy & Security > Windows Security > Virus & threat protection > Manage settings > Tamper Protection	Off	
Settings > Privacy & Security > Firewall & network protection > Domain Network | Private Network | Public Network > Windows Defender Firewall	Off	
Control Panel > Indexing Options	Remove all from "Included Locations"	
Settings > Personalization > Lock Screen > Screen saver settings	None	
Settings > System > About > System Protection > Available Drives > Protection	Off	

Configuration Item	Real World	Synthetic
Settings > System > About > Advanced system settings > Advanced > Performance > Settings > Visual Effects	Let Windows choose what's best for my computer	
Settings > System > About > Advanced system settings > Advanced > Performance > Settings > Advanced >	Adjust for best performance of Programs	
Settings > System > About > Advanced system settings > Advanced > Performance > Settings > Advanced > Virtual Memory > Change	Automatically manage paging file size for all drives	No paging file
Settings > System > About > Advanced system settings > Advanced > Performance > Settings > Data Execution Protection	Turn on DEP for essential Windows programs and services only	
Control Panel > System and Security > Windows Tools > Defragment and Optimize Drives > Scheduled Optimization > Change settings	Uncheck "Run on a schedule (recommended)"	
Control Panel > System and Security > Windows Tools > Services > Windows Update	Disabled	
Control Panel > System and Security > Windows Tools > Task Scheduler > Task Scheduler Library	Disable all tasks	
Registry Editor > HKEY_LOCAL_MACHINE > SYSTEM > CurrentControlSet > Control > Session Manager > Memory Management > PrefetchParameters > EnablePrefetcher	3	0
Anti-Virus	Disabled	

Note: Real Time Protection will reset to "On" after a reboot. Take special care to disable this setting and Tamper Protection in Windows Security.

3.4.6 Power

For laptops Solidigm recommends performance testing be done with the system plugged into AC power mode.

> In the table below a Real World and a Synthetic value are given for each. As you may recall, a *synthetic* value may be employed for specific test scenarios while *real-world* indicates how the system should be normally configured as a default for practical usage.

Configuration Item	Platform	Real World	Synthetic
Settings > System > Power > Screen and sleep > "turn off my screen"	Desktop Laptop	Never	
Settings > System > Power > Screen and sleep > "put my device to sleep"	Desktop Laptop	Never	
Control Panel > Hardware and Sound > Power Options > Power Plan	Desktop	Balanced	High Performance
Control Panel > Hardware and Sound > Power Options > Power Plan	Laptop	Balanced	
Control Panel > Hardware and Sound > Power Options > Change plan settings > Change advanced power settings > Turn off hard disk after	Desktop	0 minutes (never)	
[INHERIT] > Internet Explorer > JavaScript Timer Frequency	Desktop	Use defaults	Maximum Performance
[INHERIT] > Desktop background settings > Slide show	Desktop	Paused	
[INHERIT] > Wireless Adapter Settings > Power Saving Mode	Desktop	Use defaults	Maximum Performance
[INHERIT] > Sleep > Sleep after	Desktop Laptop	0 minutes (never)	
[INHERIT] > Sleep > Allow wake timers	Desktop Laptop	Use defaults	Enabled
[INHERIT] > USB settings > USB selective suspend setting	Desktop	Use defaults	Disabled
[INHERIT] > Solidigm Graphics Power Plan	Desktop	Use defaults	Maximum Performance
[INHERIT] > Power buttons and lid > Power button action	Desktop Laptop	Do nothing	
[INHERIT] > Power buttons and lid > Sleep button action	Desktop Laptop	Do nothing	
[INHERIT] > PCI Express > Link Power Management	Desktop	Use defaults	Off
[INHERIT] > Processor power management > Minimum processor state	Desktop	Use defaults	100%

Configuration Item	Platform	Real World	Synthetic
[INHERIT] > Processor power management > System cooling policy	Desktop	Use defaults	Active
[INHERIT] > Processor power management > Maximum processor state	Desktop	Use defaults	100%
[INHERIT] > Display > Turn off display after	Desktop Laptop	0 minutes (never)	
[INHERIT] > Use defaults	Desktop	Use defaults	100%
[INHERIT] > Display > Dimmed display brightness	Desktop	Use defaults	100%
[INHERIT] > Display > Enable adaptive brightness	Desktop	Use defaults	Off

4 Analysis

4.1 Median Recommendation

Solidigm uses and recommends the use of median values so an actual value from the dataset is reported as a characteristic value. Averages provide useful insight into the characteristics of the data but reported performance should always reflect a value that was measured rather than calculated. A minimum of three 'warm' runs must be collected to calculate a median value. If more than three test runs are required, data sets should always be comprised of odd valued numbers of tests from which a median value can be calculated.

The configuration recommendations in section 4 are designed to minimize variability in test outcomes without affecting the test parameters. However, if the test still produces data with a high amount of variability, the best approach to minimize that variability is to collect more data and evaluate a larger data set. In some test cases this approach is impractical and a quick measurement of the confidence in the data set is required.

4.2 Relative Standard Error

> "Relative Standard Error: A unit-free measure of the reliability of a statistic, defined as the absolute value of the ratio of the standard error to the sample estimate of the statistic, expressed as a percentage." - J. Black, N. Hashimzade & G. Myles in 'A Dictionary of Economics'

Solidigm uses and recommends relative standard error (RSE) to evaluate variability in a dataset. Thresholds vary between types of tests, but in general if relative standard error is above 5%, Solidigm views the variability as unacceptably high and more datapoints must be collected or the test set must be discarded. Depending on the context, RSE below 5% may also be unacceptable—lower is better and more datapoints carry more weight. If a test set takes more than 7 runs to reach acceptable consistency, Solidigm recommends the test set should be discarded or additional work be done to study the source of variability.

The RSE of a sample is defined as:

$$\left|\frac{Standard\ Error\ (of\ the\ sample)}{Sample\ Estimate}\right| = \left|\frac{Standard\ Deviation\ (of\ the\ sample)}{Sample\ Estimate \cdot \sqrt{Number\ of\ Samples}}\right| = \left|\frac{s}{\bar{x} \cdot \sqrt{N}}\right|$$

In Excel, this can be expressed as a percentage with the following formula:

`=STDEV.S(range)/SQRT(COUNT(range))/AVERAGE(range)`

4.3 Performance Delta

The performance delta between two data points in a test is calculated differently depending on the test. This is because some tests are measured in seconds (lower is better), where others are measured in speed or score (higher is better).

For tests measured in time, relative performance is measured by $\mathit{\Delta} = \frac{\text{basis}}{\text{comparison}} - 1$ where if we were testing SSD$_A$ against SSD$_B$, the performance delta is $\mathit{\Delta} = \frac{\text{time}_{\text{SSDB}}}{\text{time}_{\text{SSDA}}} - 1$. This yields a value that should be treated as a percentage improvement in speed (because the values are inverted, it shows an improvement in speed, rather than time).

For tests measured in speed/score, relative performance is measured by $\mathit{\Delta} = \frac{\text{comparison}}{\text{basis}} - 1$ where if we were testing SSD$_A$ against SSD$_B$, the performance delta is $\mathit{\Delta} = \frac{\text{speed/score}_{\text{SSDA}}}{\text{speed/score}_{\text{SSDB}}} - 1$. This yields a value that should be treated as a percentage improvement in speed/score.

5 Appendix

This section presents supplementary material not included in the other sections including discussions on executing a set of tests and covers considerations for specific benchmark tools.

5.1 Benchmark Recommendations

The following table summarizes how Solidigm uses each benchmark.

Benchmark Type	Usage
Application-based benchmarks	Used as part of the functional and performance validation flows. Recommended for evaluating platform performance.
Trace-based benchmarks	Used as part of the functional and performance validation flows. Recommended for evaluating storage device performance.
Synthetic benchmarks	Measure raw input/output (I/O) performance.

5.1.1 SYSmark 25

SYSmark 25 is an example of an application-based benchmarking tool and can be used to 'score' a platform on its ability to meet user expectations.

Application-based benchmarks require more than one run to notice substantive gains. A drive's behavior will more closely match its real use behavior when tested with an application-based workload, than it will with a purely synthetic workload. Application-based benchmarks can sometimes produce a single measure for system performance, which can be used as a product ranking index. It is critical to follow the steps below to facilitate repeatability and accuracy.

Ensure the default conditioning run is performed to populate caches with frequently accessed data.

5.1.2 PCMark 10

PCMark 10 has two storage tests for device-oriented benchmarking which Solidigm recommends:

- The **Full** System Drive Benchmark uses a wide-ranging set of real-world traces from popular applications and common tasks to fully test the performance of the fastest modern drives. This benchmark includes a comprehensive set of tests executed in a very short period of time spanning >200GB of data.
- The **Quick** System Drive Benchmark is a shorter test with a smaller set of less demanding real-world traces. You can use this benchmark to test smaller system drives that are

unable to run the Full System Drive benchmark. The Quick test is a subset of tests from the Full test.

5.1.3 Iometer

Iometer is a flexible tool that enables storage device application of user-defined workloads.

Iometer test sequences should be crafted with the following considerations:

- Test file size – Solidigm recommends a test file size of 1GB-8GB (2,097,152-16,777,216 sectors). Before testing, additional files may be copied to the device to achieve desired amount of fill.
- Preconditioning for steady state – The common practice of applying a higher QD sequential or random write before running measured tests of the same type, may not attain steady state performance for some SSDs. In addition to the above conditioning pass, in order to emulate the first (potentially not cached) access of stored data, consider performing an unmeasured 'cold run' of each desired workload.
- Workload duration – Heavy workloads applied at a very high duty cycle may push smaller form factor SSDs beyond their expected thermal envelope. This can potentially trigger performance throttled conditions that would not have occurred during normal use. Thermal throttling can be alleviated by:
  - Reducing the duty cycle of the tests to better match real-world usage.
  - Active cooling may be required during testing if duty cycle cannot be reduced.

5.1.4 CrystalDiskMark

CrystalDiskMark is a simple disk benchmark software for measuring sequential and random performance (Read/Write/Mix) and contains profiles for real world and peak performance.

Solidigm recommends using the “Default” and “Default [+ Mix]” profiles for evaluating storage device performance. When using the “Default [+ Mix]” profile, the default setting of “R70%/W30%” is appropriate for approximating end-user mixed workloads.

Additionally in Crystal Disk Mark 8, toggling the application Settings menu from “Default” to “NVMe SSD” introduces two new workloads that also represent valuable data points for quantifying storage device performance.

5.2 Example Test Results

5.2.1 Iometer Synthetic

The following test results reflect synthetic numbers that the Solidigm™ P41 Plus is able to achieve. The results are based upon measurements from Iometer test tool.

Performance Parameters	512GB	1TB	2TB
Sequential Read (MB/s) (T=1 QD=64)	3500	4125	4125
Sequential Write (MB/s) (T=1 QD=64)	1625	2950	3325
Random Read (KIOPS) (T=4 QD=64 each worker)	115	225	390
Random Write (KIOPS) (T=4 QD=64 each worker)	390	520	540

System Configuration: Intel® Core™ i9-11900K @ 3.50GHz, Asus PRIME Z590-A motherboard, BIOS: American Megatrends Inc. 0605 (3/11/2021), Memory: 32GB (2x16GB) DDR4-4000, Microsoft Windows 11 Enterprise 64-bit using native NVMe storage driver.

5.2.2 Iometer Low Queue Depth

The following test results are synthetic test results for the Solidigm™ P41 Plus as measured by Iometer but measured at low queue depth. Low queue depth is an important performance metric as many real world workloads execute at low queue depth.

Performance Parameters	512GB	1TB	2TB
Sequential Read (MB/s) (T=1 QD=1)	1300	1300	1300
Sequential Write (MB/s) (T=1 QD=1)	1600	2700	2750
Random Read (KIOPS) (T=1 QD=1)	20	20	20
Random Write (KIOPS) (T=1 QD=1)	90	90	90

System Configuration: Intel® Core™ i9-11900K @ 3.50GHz, Asus PRIME Z590-A motherboard, BIOS: American Megatrends Inc. 0605 (3/11/2021), Memory: 32GB (2x16GB) DDR4-4000, Microsoft Windows 11 Enterprise 64-bit using native NVMe storage driver.

5.2.3 Mixed Workload

Real world usage is typically a mixture of read and write activity. The following results show how the Solidigm P41 Plus performs under a mixed workload of reads and writes as measured by Iometer.

Performance Parameters	512GB	1TB	2TB
Mixed Random 4k IO 80% Reads/20% Writes (KIOPS) (T1 QD1)	22	22	22
Mixed Random 4k IO 70% Reads/30% Writes (KIOPS) (T1 QD1)	24	24	24

System Configuration: Intel® Core™ i7-11700K @3.6GHz, Graphics: Nvidia GeForce GTX 1080, Motherboard: ASUSTEK PRIME Z590-A, BIOS: Rev 0402; DRAM: 16GB (2x8GB) DDR4-3200 Corsair CMD32GX4M4C3200C16; OS: Microsoft Windows 11 (21H2 Build 22000.93) using native NVMe storage driver.

5.2.4 Trace-based Benchmark

The following are results from the popular PCMark 10 Storage Benchmark test.

Performance Parameters	512GB	1TB	2TB
PCMark 10 Storage Full Score	2440	2970	3085
PCMark 10 Storage Full Bandwidth (MB/s)	395	470	490

System Configuration: Intel® Core™ i7-11700K @3.6GHz, Graphics: Nvidia GeForce GTX 1080, Motherboard: ASUSTEK PRIME Z590-A, BIOS: Rev 0402; DRAM: 16GB (2x8GB) DDR4-3200 Corsair CMD32GX4M4C3200C16; OS: Microsoft Windows 11 (21H2 Build 22000.93) using native NVMe storage driver.