SPC BENCHMARK 1™ (SPC-1™)
Data Reduction Extension

Official Specification

Revision 1.0 – Effective TBD

Storage Performance Council (SPC)

PO Box 3504
Redwood City, CA 94064-3504
Phone (650) 556-9384
www.spcresults.org

SPC Membership as of 31 March 2020

Amazon, Inc.
Austin Automation Center
– Department of Veteran Affairs
Cybernetics
Dell, Inc.
ETRI
Foundation for Research and Technology
– Institute of Computer Science
Fujitsu America, Inc.
Futurewei Technologies Co., Ltd.
The George Washington University
Hewlett-Packard Enterprise
Hitachi Vantara Corporation
Hongik University
IBM Corporation
Infortrend Technology, Inc.
Inspur Corporation
Korea Advanced Institute of Science and Technology
Lenovo
MacroSAN Technologies Co. Ltd.
NEC Corporation
NetApp, Inc.
Pennsylvania State University
Pure Storage, Inc.
Ruijie Networks Co. Ltd.
Samsung Information Systems, America
Seagate Technology LLC
Silicon Motion Technology Corporation
Sunlight.io
Telecommunication Technology Association (TTA)
Toshiba America Electronic Components, Inc.
University of California, Santa Cruz
University of Patras
Western Digital Corporation
Yunify Technologies Inc.

The SPC would also like to acknowledge the contributions of the members who helped to create the initial release of version 3 of the SPC-1 specification Austin Automation Center – Department of Veteran Affairs, DataDirect Networks, Datera, Inc., Dell, Inc., Dot Hill Systems Corp., EMC Corporation, Foundation for Research and Technology – Institute of Computer Science, Fujitsu America, Inc., The George Washington University, Hewlett-Packard Company, Hitachi Data Systems, Hongik University, Huawei Technologies Co., Ltd., IBM Corporation, Imation, Corp. Infortrend Technology, Inc., Kaminario, Inc., Korea Advanced Institute of Science and Technology, LSI Corporation, NEC Corporation. NetApp, Inc., Nimbus Data Systems, Inc., Oracle Corporation, Pennsylvania State University, QLogic Corporation, Samsung Information Systems, America. SanDisk Corporation, Seagate Technology LLC, Silicon Graphics International, Skyera, Inc.. SolidFire, Inc., Symantec Corporation, Toshiba America Information Systems, Inc., University of California, Santa Cruz, University of Patras, Violin Memory, Inc.. Western Digital Corporation. X-IO Technologies

Document History

Effective Date	Version	Description
TBD	1.0	Creation of stand-alone extension document.

Table of Contents

Clause 0 Introduction

0.1 Preamble

Benchmark extensions are optional additions to an existing benchmark ("the underlying benchmark specification") that are designed to showcase a feature or set of features, and to provide vendors with a means to differentiate their products across dimensions other than the existing performance or price-performance metrics of the benchmark

The SPC benchmark extensions are intended to be vendor and platform independent. Any vendor should be able to sponsor and publish an SPC result, with or without extensions, provided their tested configuration satisfies the performance, integrity, and availability requirements of the specification.

Rather than requiring or favoring a particular implementation, it is the goal of SPC benchmarks and extensions to provide a robust, verifiable, reproducible environment within which the relative strengths of differing design and configuration approaches can be evaluated.

0.2 General Guidelines

The purpose of SPC benchmarks is to provide objective, relevant, and verifiable data to purchasers of I/O subsystems. To that end, SPC specifications require that benchmark tests be implemented with system platforms and products that:

- Are generally available to users.
- A significant percentage of the users in the target market segment (server class systems) would implement.
- Are relevant to the market segment that the benchmark represents.

More detailed requirements can be found in the body of the SPC Benchmark-1 specification.

0.3 Measurement Guidelines

SPC benchmark results are expected to be accurate representations of subsystem performance. Therefore, stringent measurement, auditing, and reporting guidelines are mandated by this specification. In general, fidelity and candor must be maintained in reporting any anomalies in the results, even if not specified in the benchmark requirements.

More detailed measurement, evaluation and disclosure requirements can be found in the body of the specification.

0.4 Related Documents

This benchmark extension relies on:

- Version 3 of SPC Benchmark-1
- Version 1 of the *SPC Pricing Guide*
- Version 1 of the *SPC Glossary (included as Appendix A)*

0.5 Document Conventions

This document follows the standard typographical conventions for SPC publications.

Generally, words and expressions will adhere to their common English usage. Where a particular term is being defined or assumed to have a benchmark-specific meaning, it appears in SMALLCAPS, and its formal definition can be found in the *SPC Glossary,* which is included here as Appendix A.

0.6 Disclaimer

While this workload models a rich multi-user environment that emulates a broad range of server applications, it neither represents the entire range of I/O requirements for server systems nor precisely mimics any particular application. In addition, the extent to which anyone is capable of achieving the results reported by a vendor is highly dependent upon how closely the customer's application maps to the SPC-1 workload. The extrapolation of SPC-1 results to other environments is therefore not recommended.

Actual system performance is highly dependent upon specific workload characteristics, platform configuration, and application-specific tuning. Relative system performance will vary as a result of these and other factors. Thus, SPC-1 should not be used as a substitute for customer application benchmarking when critical performance requirements are called for.

SPC-1 uses terminology and metrics that are similar to other benchmarks. This similarity does not imply that results from this benchmark are comparable with other benchmarks.

Clause 1 Workload Environment

1.1 Overview

All -flash arrays are increasingly popular for enterprise workloads. With high performance rotating media, customers typically buy enough STORAGE DEVICES to ensure they can achieve their required IOPS, and the resulting capacity is almost always more than sufficient. However, flash-based STORAGE DEVICES exhibit the opposite economics. Capacity is expensive, and once the array has sufficient capacity the all-flash media often has far more performance than the customers require. For this reason, customers are motivated to trade off some performance for increased capacity efficiency.

This extension allows arrays to demonstrate they are capable of enterprise-grade data reduction.

.

Clause 2 Storage Capacity and Content

In order to satisfy the requirements of this extension, the SPACE EFFECTIVENESS RATIO shall be equal to or greater than 1.25.

Clause 3 Workload and I/O Operation Profile

This extension has no impact on the workload and I/O operation profile as defined in SPC-1.

Clause 4 Benchmark Configuration and Tested Storage Configuration

This extension has no impact on the BENCHMARK CONFIGURATION or on the TESTED STORAGE CONFIGURATION as defined in SPC-1.

Clause 5 Test Methodology

This extension has no impact on the SPC test methodology as defined in SPC-1.

Clause 6 Measurement Requirements (Execution Rules)

This extension has no impact on the SPC measurement requirements as defined in SPC-1.

Clause 7 Data Persistence Requirements and Test

This extension has no impact on the Persistence test defined in SPC-1.

Clause 8 Reported Data

8.1 Comparability and Permitted Use

8.1.1 Results that execute this extension are subject to all Permitted Use requirements.

8.1.2 Public reference that includes a comparison of one or more SPC-1 RESULTS may distinguish those results that included the execution of this extension from those that did not.

8.1.3 If the testing defined in this extension is completed successfully, the result may be referred to as an "SPC-1 Data Reduction Result".

Clause 9 Pricing

9.1 This extension has no impact on the pricing requirements defined in SPC-1.

9.2 This extension is subject to the general pricing guidelines defined in version 1 of the *SPC Pricing Guidelines*.

Clause 10 Full Disclosure Report

This extension has no impact on the Full Disclosure Report requirements defined in SPC-1.

Clause 11 Measurement, Audit and Result Submission

The auditor shall confirm that the SPACE EFFECTIVENESS RATIO is equal to or greater than 1.25.

Appendix A Glossary

The SPC Glossary is used in all SPC specifications, and is available as a stand-alone document. It is included here in its entirety for ease of reference.

SPC Glossary, version 1.0~~.7~~, which was current as of 24 March 2020~~21 February 2020~~

A.1 A

ADDRESSABLE CAPACITY
the portion of the storage capacity of a LOGICAL VOLUME that is accessible to the WORKLOAD GENERATOR.

APPLICATION STORAGE UNIT (ASU)
the logical representation of the persistent, non-volatile storage read and or written in the course of executing a BENCHMARK.

An ASU represents is a logical interface between a BENCHMARK CONFIGURATION's data and a workload generator.

APPLICATION STORAGE UNIT CAPACITY
the total ADDRESSABLE CAPACITY of all the portions of LOGICAL VOLUMES to which an ASU is mapped.

APPLICATION STORAGE UNIT STREAM
a collection of one or more I/O STREAMs, that completely defines the I/O sent to a given ASU.

ASSOCIATED DATA
data and measurements defined by a given BENCHMARK that are used to calculate, clarify or reinforce the metrics reported as part of a RESULT.

ASU *see APPLICATION STORAGE UNIT.*

ASU CAPACITY *see APPLICATION STORAGE UNIT CAPACITY.*

ASU PRICE *the ratio of TOTAL SYSTEM PRICE to ASU CAPACITY.*

ASU STREAM *see APPLICATION STORAGE UNIT STREAM.*

AUDIT *the process that verifies that a MEASUREMENT is eligible for submission as a RESULT.*

AUDITOR *An individual who has been certified by the SPC to perform an AUDIT.*

AVAILABILITY DATE
a date by which a given product, component or configuration is released for general availability.

AVERAGE RESPONSE TIME
the sum of the RESPONSE TIMES for all MEASURED I/O REQUESTS within a given interval, divided by the total number of MEASURED I/O REQUESTS.

A.2 B

BC *see BENCHMARK CONFIGURATION.*

BENCHMARK *a collection of TESTS, TEST PHASES, documentation requirements, and comparability constraints that fully define the process for taking a MEASUREMENT and creating a RESULT.*

BENCHMARK CONFIGURATION
all hardware and software components used in the creation of a MEASUREMENT.

A.3 C

COMPLETED I/O REQUEST *an I/O REQUEST with a START TIME and a COMPLETION TIME.*

COMPLETION TIME
the time recorded by the WORKLOAD GENERATOR when an I/O REQUEST is satisfied by the TSC.

COMMITTED: *Of an IO operation, written to persistent, non-volatile storage, in such a manner that the data can be retrieved after recovery from a TSC failure.*

CRASH-CONSISTENT:
A data image (logical or physical) is considered crash consistent if there exists a point in time such that all write operations completed prior to that time are included in the image, and no write operation initiated after that time is included.

A.4 D

DATA RATE *the data volume transferred in a given interval divided by the duration of the interval, in seconds.*

A.5 E

EXTENSION *optional addition(s) to an existing BENCHMARK that showcase a feature or set of features not captured by the BENCHMARK'S existing metrics.*

EXTENSION CONFIGURATION
all hardware and software components used in the execution of an EXTENSION.

EXPECTED I/O COUNT
for any given I/O STREAM and TEST PHASE, the product of requested IO load in IOs per second, the duration of the TEST PHASE in seconds, and the INTENSITY MULTIPLIER parameter for that I/O STREAM.

EXECUTIVE SUMMARY
a high-level report summarizing a RESULT, and the configuration used to produce it.

A.6 F

FAILED I/O REQUEST
any I/O REQUEST issued by the WORKLOAD GENERATOR that could not be completed or was signaled as failed by the OS running on the HOST SYSTEM .

A FAILED I/O request has no COMPLETION TIME.

FDR *see FULL DISCLOSURE REPORT.*

FULL DISCLOSURE REPORT

a report detailing a RESULT, along with the procedures, configuration, and equipment used to produce it.

A.7 G

No terms defined.

A.8 H

HOST SYSTEM *a computer system where the WORKLOAD GENERATOR executes.*

A.9 I

IN-FLIGHT I/O REQUEST

an I/O REQUEST issued by the WORKLOAD GENERATOR that does not complete within a given MEASUREMENT INTERVAL.

INTEGRATED EXECUTION

of a benchmark extension: completed during one of the test phases of a benchmark execution.

INTENSITY MULTIPLIER

the ratio of the IO load produced by a given I/O STREAM to the total IO load produced by all active I/O STREAMS.

I/O COMMAND *see I/O REQUEST.*

I/O STREAM *a single, well-defined, sequence of I/O REQUESTS.*

I/O REQUEST *a single, atomic I/O operation.*

I/O REQUEST THROUGHPUT

the total number of MEASURED I/O REQUESTS in a TEST PHASE, divided by the duration of that TEST PHASE's MEASUREMENT INTERVAL, expressed in seconds.

A.10 J

No terms defined.

A.11 K

No terms defined.

A.12 L

LOGICAL BLOCK *the smallest directly addressable unit of storage on the ASU.*

LOGICAL VOLUME *an individually addressable logical unit of storage presented to the WORKLOAD GENERATOR.*

A.13 M

MEASURED I/O REQUEST
an I/O REQUEST with a COMPLETION TIME occurring within the MEASUREMENT INTERVAL.

MEASURED INTENSITY MULTIPLIER
the percentage of all MEASURED I/O REQUESTS that were issued by a given I/O STREAM.

MEASUREMENT: *the data gathered during the execution of a BENCHMARK.*

MEASUREMENT INTERVAL
of a TEST PHASE, the time from the end of the TRANSITION to the start of the RUNOUT.

A.14 N

No terms defined.

A.15 O

ON-SITE AUDIT *an AUDIT for which the AUDITOR is physically present.*

A.16 P

PHYSICAL CAPACITY UTILIZATION
ASU CAPACITY divided by the PHYSICAL STORAGE CAPACITY.

PHYSICAL FREE SPACE
the persistent storage capacity that could be used to hold application data and the metadata required to access, maintain and protect that data, but is not in use at the time of the measurement.

PHYSICAL STORAGE CAPACITY
the total storage capacity of all of the STORAGE DEVICES in the TESTED STORAGE CONFIGURATION.

PRICED STORAGE CONFIGURATION ("PSC")*:*
the customer-orderable version of the TSC.

PRICE-PERFORMANCE
the ratio of the TOTAL SYSTEM PRICE to the primary performance metric for a BENCHMARK"):.

PRICING SPREADSHEET
a detailed computation of the total cost of ownership for a PRICED STORAGE CONFIGURATION.

PRIMARY METRIC *a metric that provides a primary basis for comparison of RESULTS.*

PROTECTED 1 *a data protection level in which the failure of any single STORAGE DEVICE in the TSC will not require user intervention to restore access to the BENCHMARK'S"): data repository.*

PROTECTED 2 *a data protection level in which the failure of any single component in the TSC will not require user intervention to restore access to the BENCHMARK'S data repository.*

PSC *see PRICED STORAGE CONFIGURATION.*

A.17 Q

No terms defined.

A.18 R

REFERENCE PRICE
the price at which component or subsystem could be ordered individually from the TEST SPONSOR or designated third-party suppler.

REMOTE AUDIT *an AUDIT for which the AUDITOR is not physically present. See ON-SITE AUDIT.*

REPLICATION *the automatic execution of all I/O operations executed against a primary storage system on a one or more, independent storage systems.*

RESPONSE TIME *for an I/O REQUEST, COMPLETION TIME minus START TIME.*

RESULT *an audited MEASUREMENT which has been submitted to the SPC for publication*

RESULTS FILES *the output of the WORKLOAD GENERATOR, created during a MEASUREMENT.*

REPORTED DATA *THE set of data, as defined by a given BENCHMARK, which fully characterizes a MEASUREMENT.*

RUNOUT *of a TEST PHASE, the time period immediately following the MEASUREMENT INTERVAL during which the IO load presented by the WORKLOAD GENERATOR to the TSC remains constant long enough for any IO issued during the MEASUREMENT INTERVAL to complete.*

A.19 S

SER *see SPACE EFFECTIVENESS RATIO.*

SOR *see SPACE OPTIMIZATION RATIO.*

SNAPSHOT *a logical, point-in-time, CRASH-CONSISTENT image of one or more LOGICAL VOLUMES.*

SNAPSHOT SET *a crash-consistent collection of SNAPSHOTS, taken and managed as a unit.*

SPACE EFFECTIVENESS RATIO ("SER")
the ratio of the total amount of data that the TSC can hold to its PHYSICAL CAPACITY.

SPACE OPTIMIZATION RATIO ("SOR")
the size of a data set as generated by the WORKLOAD GENERATOR divided by the amount of incremental space consumed by that data set.

SPC RESULT *see RESULT.*

SSU *see STIMULUS SCALING UNIT.*

START TIME *for an I/O REQUEST, the time recorded by the WORKLOAD GENERATOR when the request is submitted for execution on the TSC.*

STEADY STATE *a state in which the behavior of the TSC is stable and sustainable while the load presented to the TSC by the WORKLOAD GENERATOR is constant.*

STIMULUS SCALING UNIT
a logical abstraction that captures the key elements in the IO demands of an application's user population.

STORAGE DEVICE *a discrete, physical hardware component, such as an HDD or an SSD, that provides permanent data storage.*

A STORAGE DEVICE must be capable of storing data indefinitely without external power. The requirement excludes components that provide volatile data storage, such as a read and/or write cache.

SYNCHRONOUS REPLICATION
REPLICATION IN WHICH THE INITIAL I/O OPERATION IS NOT MARKED AS COMPLETE UNTIL THE RELATED OPERATION HAS COMPLETED ON THE OTHER, INDEPENDENT STORAGE SYSTEM(S).

SUBMISSION IDENTIFIER
a unique identifier, assigned by the SPC, for each new RESULT.

SUPPORTING FILES
a collection of data, documentation, and illustrations used to demonstrate the validity of a RESULT.

A.20 T

TARGET COUNTRY
the country in which the PRICED STORAGE CONFIGURATION is available for sale no later than the AVAILABILITY DATE, and in which the required hardware maintenance and software support is provided either directly from the TEST SPONSOR or indirectly via a third-party supplier

TEST *a collection of one or more TEST PHASES sharing a common objective.*

TEST PHASE *the smallest logical component of a TEST, during which a data is collected to satisfy the requirements of a BENCHMARK.*

TEST SPONSOR *a distinctly identifiable entity that acts as the sponsor of an RESULT.*

TESTED STORAGE CONFIGURATION
all software and hardware necessary to implement and support the storage configuration defined for a MEASUREMENT.

TESTED STORAGE PRODUCT
a distinct, customer orderable product, which is the focal point of a RESULT.

TOTAL SYSTEM PRICE
the total cost of ownership for the PRICED STORAGE CONFIGURATION.

TRANSITION *of a TEST PHASE, a time period during which the IO load presented by the WORKLOAD GENERATOR to the TSC is changing, either increasing or decreasing.*

TSC *see TESTED STORAGE CONFIGURATION.*

TSC BOUNDARY *the boundary between the HOST SYSTEM and TSC.*

TSC EXECUTIVE the software component of the *TSC*.

TSP *see TESTED STORAGE PRODUCT.*

A.21 U

No terms defined.

A.22 V

No terms defined.

A.23 W

WORKLOAD *a collection of ASU STREAMS.*

WORKLOAD GENERATOR
a user-space application, provided by the SPC, that produces benchmark-specific IO STREAMS.

A.24 X

No terms defined.

A.25 Y

No terms defined.

A.26 Z

No terms defined

Index

M

O

P

R

S

T

W