

Introducing and Validating SNIA SSS Performance Test Suite

Esther Spanjer, SMART Modular

SNIA Legal Notice



- The material contained in this tutorial is copyrighted by the SNIA.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
 - Any slide or slides used must be reproduced in their entirety without modification
 - The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA Education Committee.
- Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.



SSS Performance Benchmarking Learning Objectives

- Get a good understanding of the various parameters that influence the performance characteristics of SSDs
- Get a full understanding of the proposed SNIA Performance Measurement Specification
- Provide step-by-step guidance on how to set up a test benchmark that enables comparison among the various SSS devices





SSS = Solid State Storage





Traditional hard disk drive



Solid state hard drive













Flashy fists fly as OCZ and DDRdrive row over SSD performance



	Flashy fists fly as OCZ and DDRdrive row over SSD performance • The Register - Windows Internet Explorer		
Conge	A http://www.theregister.co.uk/2011/01/14/ocz_and_ddrdrive_performance_row/	Soogle	P 🛓
And the stand is to obde and to the two own stops the second of	File Edit View Favorites Tools Help		
Control of the control of th	Google 🛛 🗸 Search * 🖗 🏢 * 👘 🖉 Share * 🔯 * 🏠 Bookmarks * 🎪 Check * 🖏 Translate	e 🔹 🍠 AutoFill 🍷 🥜	🦴 🔹 🔵 esther 🔹
<image/> Concernance	A Reshy fists fly as OCZ and DDRdrive row over 55D pe	🛅 • 🚮 • 🖶 • 🗄	🍃 Page 👻 🍈 Tools 👻
Note that the stand of the	Login Sign up	Whitepapers Reg Hardware Channel Reg	^
PC & & Chips Wree in the Control of the subcoding of the value of the control of the subcoding of the value of the	The A Register ®	(ISLON°	
Consider the entropy of the ent	Hardware Software Music & Media Networks Security Public Sector Business Science Odds &	Sods Search site	
 Print Postcomment Provided Print Print Print Provided Print Print Pri	PCs & Chips Servers HPC Storage Data Networking Virtualization Data Centre Sysadmin Bl	llog	
 Flashy fists fly as OCZ and DDRdrive row over SSD performance CTO says rival's specs are 'knowingly disingenuous' By Chris Mellor - Get more from this author Posted in Storage, 'tith January 2011 12:01 GMT Freewinklegsper - The Reg Guide to Solutions for the virtual Era Two solid state disk SSD suppliers are arguing about NAND flash performance drop-off. OCZ supplies NAND flash solid state drives (SSDs) and regularly announces high- performance products. DDRdrive hars recerity exited steath mode and makes the X1 hybrid DRAMINAND SSD. It criticises OCZ and other flash suppliers for products exhibiting a dreadful performance decline after a little use fresh out of the box. Christopher George, founder and chief technology officer of DDRdrive, talks of 'diryt tricks used to drastically inflate (OPS results', 'the egregious disparity between promised performance and reality', and 'the untoid truth about OCZ's "sustained' write (IOPS'. George gave a presentation at the Open Storage Summit 2010 on flash SSD performance degradation and we have received coupy of the silde deck. The PowerPoint pltch assets that "Flash SSDs produce one-time, unsustainable, dramatically inflated results when tested "new' or after a Secure Erase." Then George wades in on OCZ, claiming. 'Both the OCZ Vertex 2 EX and the OCZ Vertex Pro SSD use the SandForce 1500 controller, which implements compression at the drive level. Benchmarking with an lometer version that defaults to using extremely compression? El Cin tailor Takeconcetion is now connected '& Reg Wireless Network Connection is now connected '' Reg Wireless Network Connection is now connected '' Reg Wireless Network Connection is now connected '' Reg Wireless Network Connection is now connecte	Lea	arn More.	
CTO says rival's specs are 'knowingly disingenuous' By Chris Mellor - Get more from this author Posted in Storage, 14th January 2011 12:01 GMT Free wildspaper - The Reg Guide to Solutions for the Virtual Era Two solid state disk SSD suppliers are arguing about NAND flash performance drop-off. OCZ supplies NAND flash solid state drives (SSDs) and regularly announces high- performance products, DDRdrive has recently exited stealt mode and makes the X1 hydro DRAMNAND SSD. It criticises OCZ and other flash suppliers for products exhibiting a dreadful performance decline after a little use fresh out of the box. Christopher George, founder and chief technology officer of DDRdrive, talks of 'dirly tricks used to drastically inflate IOPS results', 'the egregious disparity between promised performance and reality'', and 'the untold truth about OCZ's 'sustained' write IOPS'. George gave a presentation at the Open Storage Summit 2010 on flash SSD performance 'new' or after a Secure Erase.'' Then George wades in on OCZ, claiming. 'Both the OCZ Verter 2 EX and the OCZ Verter 'new' or after a Secure Erase.'' Then George wades in on OCZ, claiming. 'Both the OCZ Verter 2 EX and the OCZ Verter 2 EX and the CoZ Verter X Pro SSD use the SandForce 1500 controller, which implements compression at the drive ivee! Benchmarking with an lometer version that defaults to using extremely compression. The Orige Verter 2 EX and the OCZ Verter X Pro SSD use the SandForce 1500 controller, which implements compression at the drive ivee! Benchmarking with an lometer version that defaults to using extremely compression. The orige Mice Mice Controller without main free more specific to controller without main defaults to using extremely compression. The orige Mice Mice Mice Mice Mice Mice Mice Mic	🗏 Print 🔍 Post comment 🕒 Retweet 📑 Facebook Alert 💟		
DRAMMAND SSD. It criticises OC2 and other hash suppliers for products exhibiting a dreadful performance decline after a little use fresh out of the box. Christopher George, founder and chief technology officer of DDRdrive, taiks of "dirly tricks used to drastically inflate IOPS results", "the egregious disparity between promised performance and reality", and "the untoid truth about OC2's "sustained" write IOPS". George gave a presentation at the Open Storage Summit 2010 on flash SSD performance degradation and we have received a copy of the slide deck. The PowerPoint pitch asserts that "Flash SSDs produce one-time, unsustainable, dramatically inflate results when tested "new" or after a Secure Erase." Then George wades in on OCZ, claiming: "Both the OCZ Vertex 2 EX and the OCZ Vertex Pro SSD use the SandForce 1500 controller, which implements compression at the drive level. Benchmarking with an Iometer version that defaults to using extremely compressible	CTO says rival's specs are 'knowingly disingenuous' By Chris Mellor - Get more from this author Posted in Storage, 14th January 2011 12:01 GMT Free whitepaper – The Reg Guide to Solutions for the Virtual Era Two solid state disk SSD suppliers are arguing about NAND flash performance drop-off. OCZ supplies NAND flash solid state drives (SSDs) and regularly announces high-	USILON®	
used to drastically inflate IOPS results", "the egregious disparity between promised performance and reality", and "the untoid truth about OCZ's "sustained" write IOPS". George gave a presentation at the Open Storage Summit 2010 on flash SSD performance degradation and we have received a copy of the silde deck. The PowerPoint pitch asserts that "Flash SSDs produce one-time, unsustainable, dramatically inflated results when tested "new" or after a Secure Erase." Then George wades in on OCZ, claiming: "Both the OCZ Vertex 2 EX and the OCZ Vertex Pro SSD use the StandForce 1500 controller, which implements compressible level. Benchmarking with an Iometer version that defaults to using extremely compressible Stim un	dreadful performance decline after a little use fresh out of the box.	Learn more.	
degradation and we have received a copy of the slide deck. The PowerPoint pitch asserts that "Flash SSDs produce one-time, unsustainable, dramatically inflated results when tested 'new' or after a Secure Erase." SSD performance Then George wades in on OCZ, claiming: "Both the OCZ Vertex 2 EX and the OCZ Vertex. Pro SSD use the SandForce 1500 controller, which implements compressible level. Benchmarking with an lometer version that defaults to using extremely compressible EMC in Isilon 'takeover talks' Image: Stand Barry Comparison of the slide deck. The PowerPoint pitch asserts that "Flash SSDs produce one-time, unsustainable, dramatically inflated results when tested 'new' or after a Secure Erase." Image: Berlow Barry Comparison of the Stand Porce 1500 controller, which implements compressible Image: Stand Barry Comparison of the Stand Porce 1500 controller, which implements compressible EMC in Isilon 'takeover talks' Image: Stand Barry Comparison of the Stand Porce 1500 controller, which implements compressible Image: Stand Barry Comparison of Stand Porce 1500 controller, which implements compressible Image: Stand Barry Comparison of the Stand Porce 1500 controller, which implements compressible Image: Stand Barry Comparison Porce 1500 controller, which implements compressible Image: Stand Barry Comparison Porce 1500 controller, which implements compressible Image: Stand Barry Comparison Porce 1500 controller, which Barry Comparison Porce 1500 controleter, which Barry Comparison Porce 1500 contr	used to drastically inflate IOPS results", "the egregious disparity between promised performance and reality", and "the untoid truth about OCZ's "sustained" write IOPS".	Sainsbury's is abandoning tape	
Pro SD use the SandForce 1500 controller, which implements compression at the drive level. Benchmarking with an lometer version that defaults to using extremely compressible sinn im	degradation and we have received a copy of the slide deck. The PowerPoint pitch asserts that "Flash SSDs produce one-time, unsustainable, dramatically inflated results when tested 'new' or after a Secure Erase."	SSD performance IP cans EVA clustering CIA used 'illegal, inaccurate code to target	
	Pro SSD use the SandForce 1500 controller, which implements compression at the order vice	((9) Wireless Network Connection is now connected	
	Hartart & A A " A Messacht - A Thubu Guide B Day Training A Messacht Davids David		

"dirty tricks used to drastically inflate IOPS results"...

" the egregious disparity between promised performance and reality"...

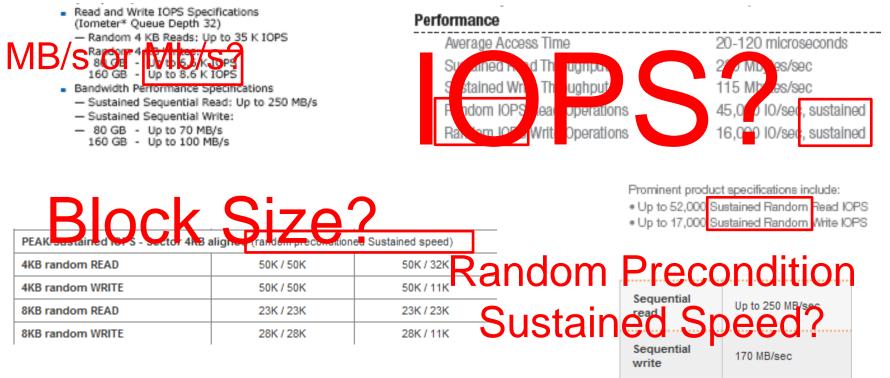
"This really isn't an 'apples to apples' comparison"...

"you can't use an X1 dragster on the open road but you can use a Ferrari Vertex 2 EX"...

http://www.theregister.co.uk/2011/01/14/ocz_and_ddrdrive_performance_row/

The Performance Landscape One Year Later





Rancomor PERFORMANCE Sustained data transfer rate O data rapefer rate Sustained data transfer rate Sustained data transfer rate Sustained data transfer rate

Market Segmentation



 Internet Internet Internet C-MLC O-7% over provisioning No backup power circuit No Enterprise features No customization Warranty 1-3 yrs 		Client SSD
 C-MLC 0-7% over provisioning No backup power circuit No Enterprise features No customization 		= Ticknology =
 0-7% over provisioning No backup power circuit No Enterprise features No customization 	•	Low cost
 No backup power circuit No Enterprise features No customization 	•	C-MLC
No Enterprise featuresNo customization	•	0-7% over provisioning
 No customization 	•	No backup power circuit
	•	No Enterprise features
• Warranty 1-3 yrs	•	No customization
	•	Warranty 1-3 yrs

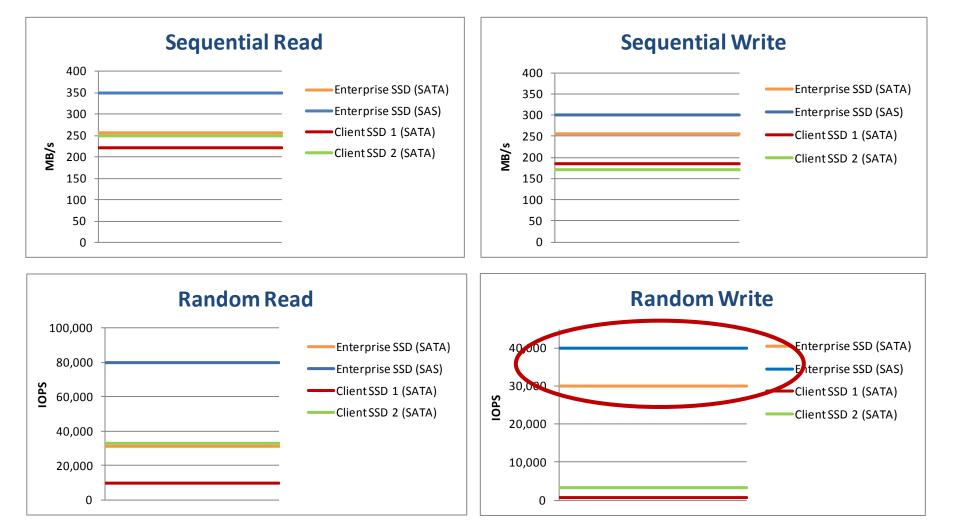




- Higher cost
- E-MLC/SLC
- 28-50% over provisioning
- Backup power circuit
- Enterprise features
- Customization
- Warranty 5 yrs

Performance Comparison Enterprise vs. Client SSD





Variables influencing Performance SNIA

- Platform
 - Test Hardware (CPU, interface, chipset, etc)
 - Software (OS, drivers)
- SSS Device Architecture
 - Flash geometry, cache, flash management algorithm, etc



- SSS Device Architecture
 - Flash geometry, cache, flash management algorithm, etc
- Workload
 - Write history & preconditioning: State of device before testing 1

Variables influencing Performance

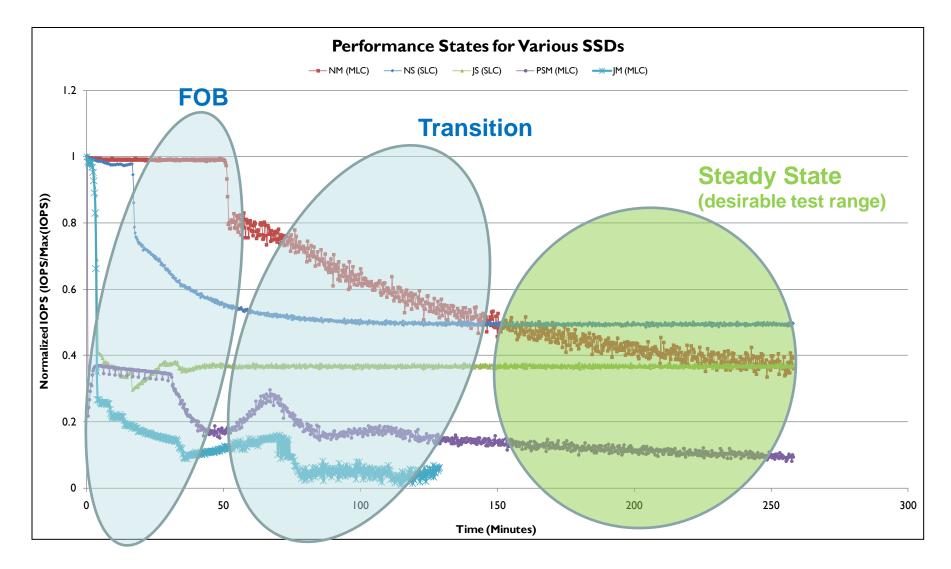
- Platform •
 - Test Hardware (CPU, interface, chipset, etc)
 - Software (OS, drivers) •





The need for Preconditioning

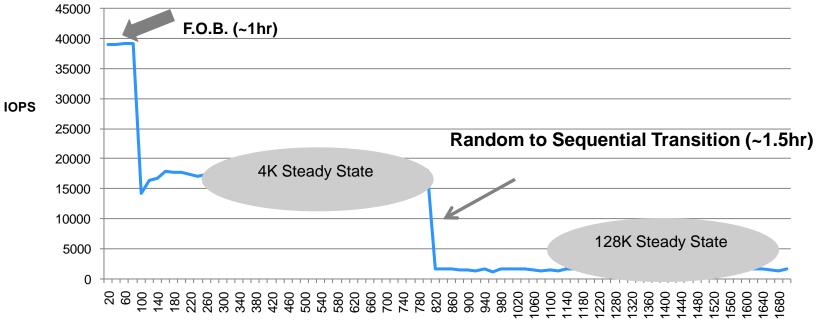








4K Random to 128K Sequential Transition

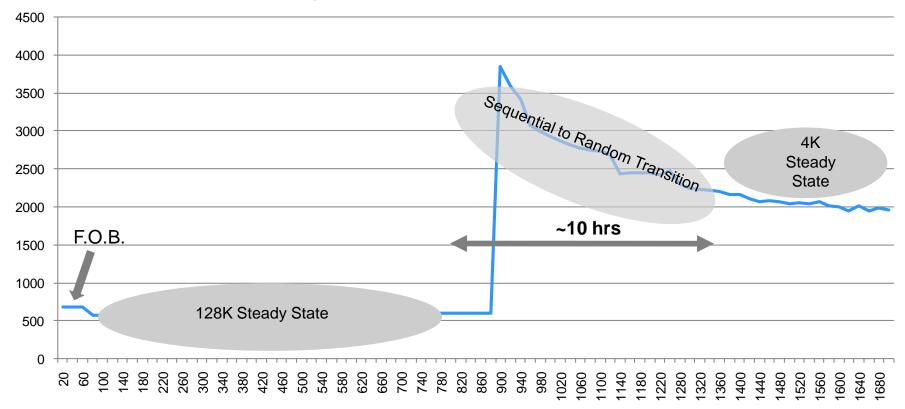


Time (Minutes)





128K Sequential to 4K Random Transition



/rite history & preconditioning: State of device before

- Write history & preconditioning: State of device before testing
- 2. Workload pattern: Read/write mix, transfer size, sequential/random

Variables influencing Performance

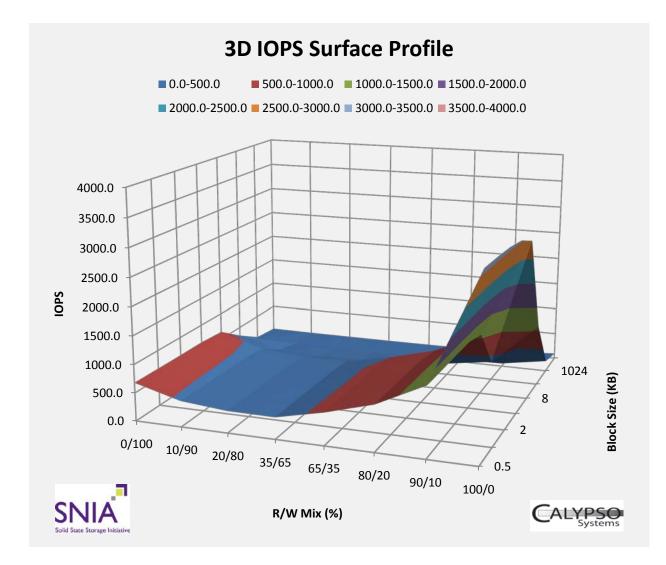
- Platform
 - Test Hardware (CPU, interface, chipset, etc)
 - Software (OS, drivers)
- SSS Device Architecture
 - Flash geometry, cache, flash management algorithm, etc
- Workload





Workload Pattern





Performance depends on

- Read/Write Mix
- Block Size
- Queue Depth (not shown)

Note: Shown 3D IOPS image courtesy of Calypso Systems

Apples to Apples, Pears to Pears in SSS Performance Benchmarking © 2011 Storage Networking Industry Association. All Rights Reserved.

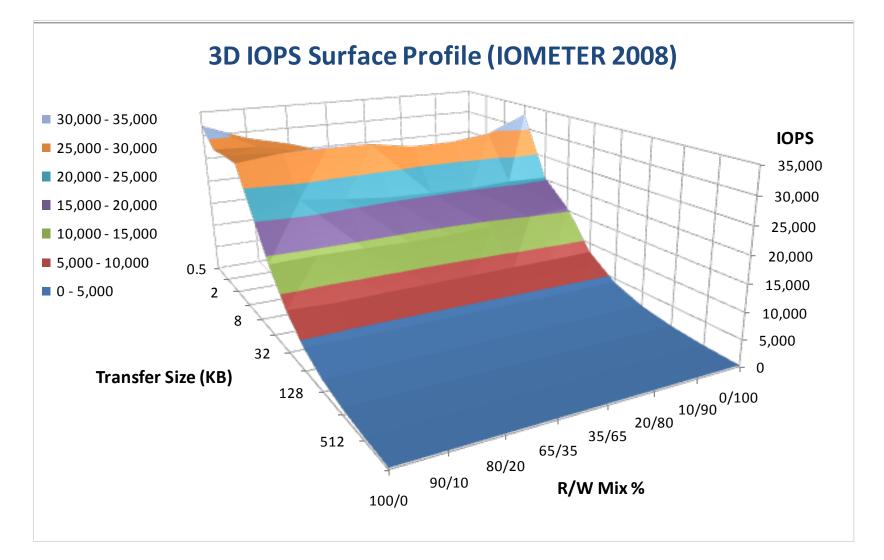
Variables influencing Performance S

- Platform
 - Test Hardware (CPU, interface, chipset, etc)
 - Software (OS, drivers)
- SSS Device Architecture
 - Flash geometry, cache, flash management algorithm, etc
- Workload
 - Write history & preconditioning: State of device before testing
 - 2. Workload pattern: Read/write mix, transfer size, sequential/random
 - 3. Data Pattern: The actual bits in the data payload written to the device





Dependency on data content - I

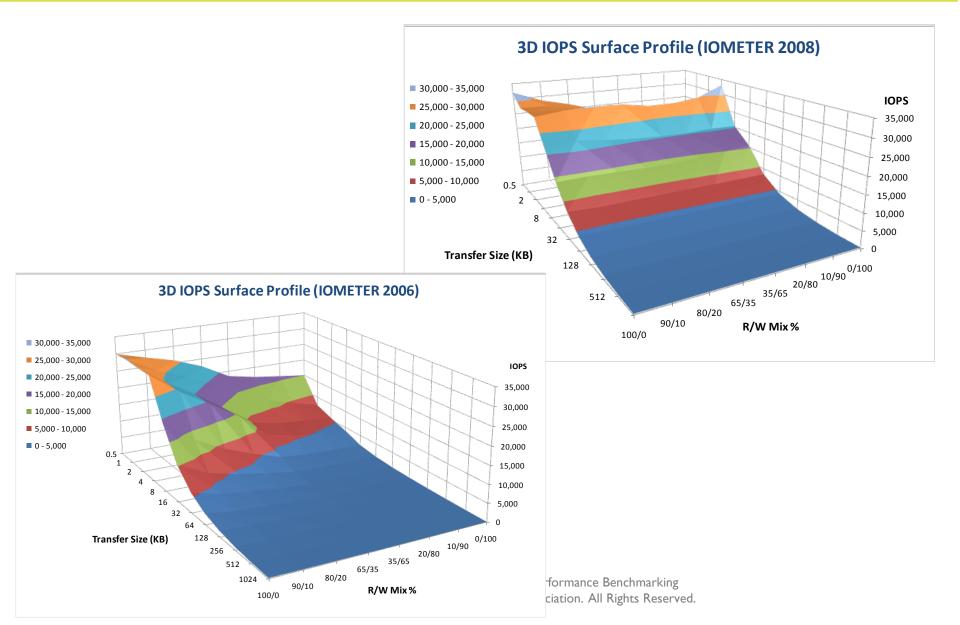


Education

SN

Dependency on data content - 2





Benchmark Suites



	Test Suite	Client SSD	Enterprise SSD
PCMark	HDD Score, OS and application loading timing, user simulation (surfing web, windows media player, etc)	\checkmark	
SysMark	System-level test. Measures performance based on average response time, gives score (0-250)	\checkmark	
IOMeter	Sequential/Random performance, workload simulation (file server, web server workload, etc)	\checkmark	\checkmark
HDTach/ H2benchw	Performance stability, Sequential/Burst performance, Access Time	\checkmark	
HD Tune	Performance stability, Sequential/Burst performance, Access Time		
Everest	Random Access Time (Read/Write)	\checkmark	
VDBench	Workload generator, performance on DAS and NAS		
Calypso CTS	Device (RAW) level, direct IO synthetic stimulus generator for both client and enterprise		





SNIA

• SNIA Technical Working Group (TWG)



ers effort

 Future: Split of test specification into separate Enterprise and Client categories



- I. Prepare the Device
 - Purge/Erase \rightarrow put SSD back into "original" state
- 2. Workload independent preconditioning
 - Write data 2x capacity \rightarrow bring device to known state
- 3. Steady State Testing (includes workload based preconditioning)
 - Run Test Loop up until steady state is achieved
 - Performance stays within ± 10% margin

4. Test Report

- Steady state convergence
- Steady State Verification
- Performance measurement (2D/3D)



- Preconditioning is key to get repeatable results
- Preconditioning needed to get drive in Steady State, after which performance can be measured



- Preconditioning is key to get repeatable results
- Preconditioning needed to get drive in Steady State, after which performance can be measured
- Two types of preconditioning
 - Workload independent write 2x capacity with I28KB sequential writes
 - Workload dependent run workload itself until steady state is achieved





- Measurement window is interval for last 5 measured rounds (i.e. test loops) that show steady sate results
- Steady State is Client IOPS Test - Steady State Convergence Plot Dependent Variable = Ave Random Write IOPS ActiveRange = (x,y); OIO/Thread = x; Thread Count = x; Data Pattern = x Variation of 12.000 measuremen 10,000 window is wi BS=512 Measurement Window 8,000 BS=4096 (Determined by 4KiB Write) 20% of avera Ave IOPS BS=8192 6,000 Slope excursion Trending of y 10% average HBS=16384 4,000 within BS=32768 BS=65536 2,000 measuremen window is w 14 16 n 12 10% of avera Round #

Performance Workloads & Tests



Client Test

- Random IOPS
 - 100/0, 95/5, 65/35, 50/50, 35/65, 5/95, 0/100
 - 1024K, 128K, 64K, 32K, 16K, 8K, 4K, 0.5K
- Sequential MB/s
 - 100/0, 0/100
 - 1024K
- Latency (random access)
 - 100/0, 65/35, 0/100
 - 8K, 4K, 0.5K

Enterprise Test

• Random IOPS

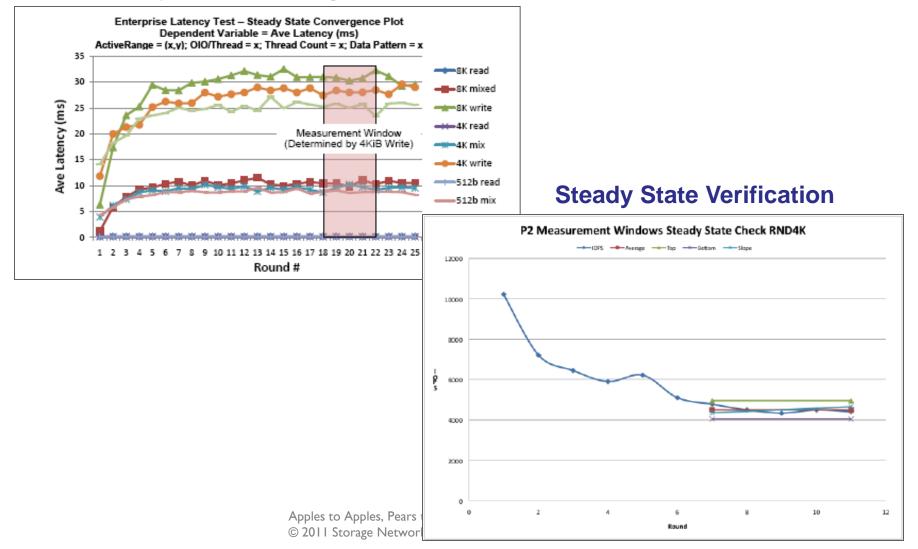
- 100/0, 95/5, 65/35, 50/50, 35/65, 5/95, 0/100
- 1024K, 128K, 64K, 32K, 16K, 8K, 4K, 0.5K
- Sequential MB/s
 - 100/0, 0/100
 - 1024K, 64K, 8K, 4K, 0.5K
- Latency (random access)
 - 100/0, 65/35, 0/100
 - 8K, 4K, 0.5K

Version 1.0 will focus further on differentiating Client vs. Enterprise

Standard Reporting - 1



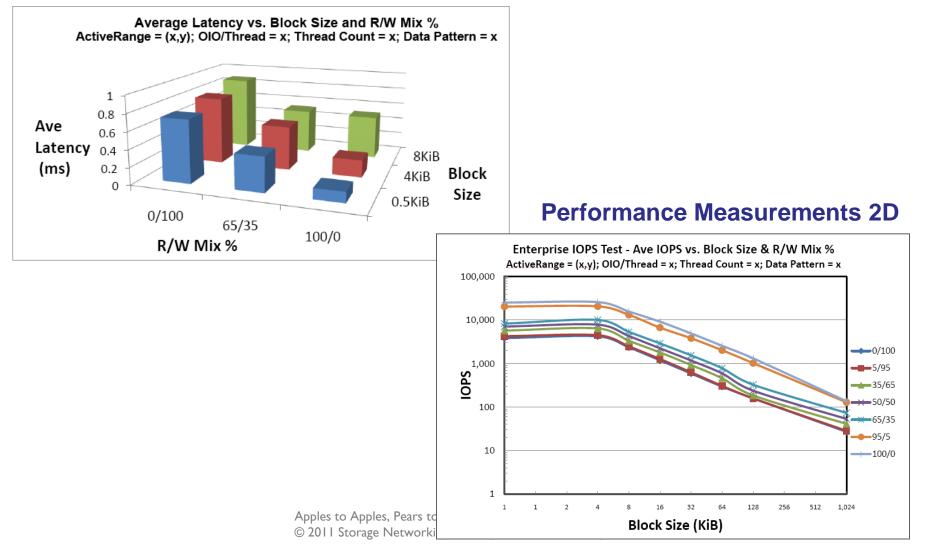
Steady State Convergence Plot



Standard Reporting - 2



Performance Measurements 3D



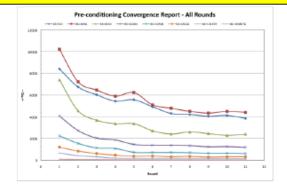
Sample Test Report included



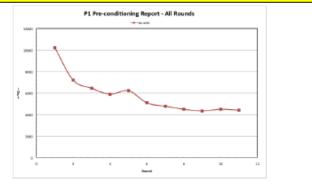
		e Annex A -		est Report		
		Summary R	eport Page			
Solid Sta	ate Storage Per	formance Test	Specification	(PTS)	Rev.	0.7
					Page I	of 5
Device Under Test (DUT)	ABC Co.	SSS TWG PTS Summary Report		SNIA SSS TWG	SNIA	
Model No.:	ABC123	Test Spec	ification:	Test Run Date:	Apr 3-2	
Form Factor:	2.5"	Test oper	meaclorn	Report Date:	June 01	
NAND Capacity: DUT Interface:	256 GB MLC SATAIL SAS HBA	SNIA SSS TW	/G PTS v 0.9	Test Sponsor: Auditor Name:	Calypso N	
bot internet.		sting Summa	ry: Tests Ru			
		Preparation	•	Test Loop	Paramet	ers
Test		Workload		1000 2000		
Test	Purge Type	Independent Preconditioning	Active Range %	QD / TC / OIO	Data Pattern	
8.1 Client IOPS	Secure Erase	2X 128K SEQ	100%	QD 16 / TC 4 / OIO 64	RN	ID
					•	
Client IOPS 8.2 OPT – AR	Secure Erase	2X 128K SEQ	10%	QD 16 / TC 4 /	RN	ID.
10%	Secure Ersse		100	OIO 64	No.	
			-			
Client IOPS 8.3 OPT- File	Secure Erase	2X 128K SEQ	100%	QD 16 / TC 4 / OIO 64	NON File as Dat	
Data				000	The as the	a ractern
9.1 Client						
7.1 Throughput						
Client				1		
10.1 Latency						
	G	eneral Device	e Descriptio	n		
Device Under	Test (DUT)	System Hardware		System Softw	are Config	uration
fanufacturer	Calvaso Systems		OS 5.4			
1odel No.	ABC123	Model No.	RTP 2.0			
ertal No. Irmware Rev No.	123.xxx.fff Ffff.hhhh.abc.123	Motherboard Chasis	Intel 5520HC Intel SC5560DP	Test SW Tool(s)	Calypso CT	5v6.5
			Intel 3.2GHz			
Jser Capacity	256 GB	СРО Туре	W5590	Other SWTool(s)		
nterface/Speed	6Gb/s SATAII No. CPUx Single					
orm Factor	2.5"	DRAM Type	1333MHz DDR3 ECC			
1edia Type	MLC	DRAM Amt	12 GB			
Major Features:		DUT I/F	SAS HBA LSI 6Gb/s 9212-			
NCQ:	YES	SAS HBA	4941			
Hot Plug	YES	SATAII PCLe	ICIOHR			
Sanitize Support: Other 1:	NO	PCI-e Boot HDD	Gen 2 (8) Iane 160 GB 7200RPM			
Other 1: Other 2:		Boot HDD Optical Drive	TEO GE 7200KPM			
Other 2:		opucar Drive			L	

Informative Annex A – Sample Test Report Client IOPS REQUIRED - Report Page										
	Solid	State Storage	e Performance	Test Specif	fication (P	PTS)		Rev.	0.7	
	20110	State Storing		rest speen		,		Page 2	of 5	
Device Under Test (DUT)		ABC Co.	8.1 Client IOPS Test SNIA SSS TWG		SNIA					
Key S Da	ta	DUT Pre	DUT Preparation		Test Loop Parameters			Steady State		
DUT	256GB MLC	Purge Done:	Yes	Required:	Data Pa	ttern	random	Convergence	YES	
DUT I/F	SAS HBA	Purge Type:	Security Erase					Rounds	7-1	
Test HW	RTP 2.0		-	Tester Choice:	OIO/Thread I6		16	Active Ra	Active Range	
os	CentOS 5.4	Workload Independent Preconditioning	2X 128K Sequential Wrt		Thread Count		4	Required:	100%	
Test SW	CTSv6.5							Optional:	N/A	

8.1.1 Steady State Convergence Plot – All Block Sizes



8.1.2 Steady State Convergence Plot – 4K Block Sizes



© 2011 Storage Networking Industry Association. All Rights Reserved.



- SSSI Group of SNIA
 - Technical Work Group (TWG) → Performance Benchmark Spec
 - Tech Dev Group \rightarrow Performance Test Platform
- JEDEC 64.8
 - Specification for SSD endurance measurement
- SSDA
 - Testing of reliability (power cycling, data retention, endurance, etc) and OS compatibility (Windows 7)





- SSS Performance is dependent on many variables
- Comparing vendors is not trivial → industry standard required
- SNIA Performance Specs allows apples to apples comparison



- Spec for review at http://www.snia.org/tech_activities/publicreview
- Send your feedback to <u>ssstwg@snia.org</u>



Please send any questions or comments on this presentation to SNIA: <u>tracksolidstate@snia.org</u>

Many thanks to the following individuals for their contributions to this tutorial. - SNIA Education Committee

David Landsman, Easen Ho Eden Kim Neal Ekker, Dan Le