



# **Solid State Deployments - Recommendations for POC's**

Russ Fellows – Evaluator Group  
[russ@evaluatorgroup.com](mailto:russ@evaluatorgroup.com)

# SNIA Legal Notice

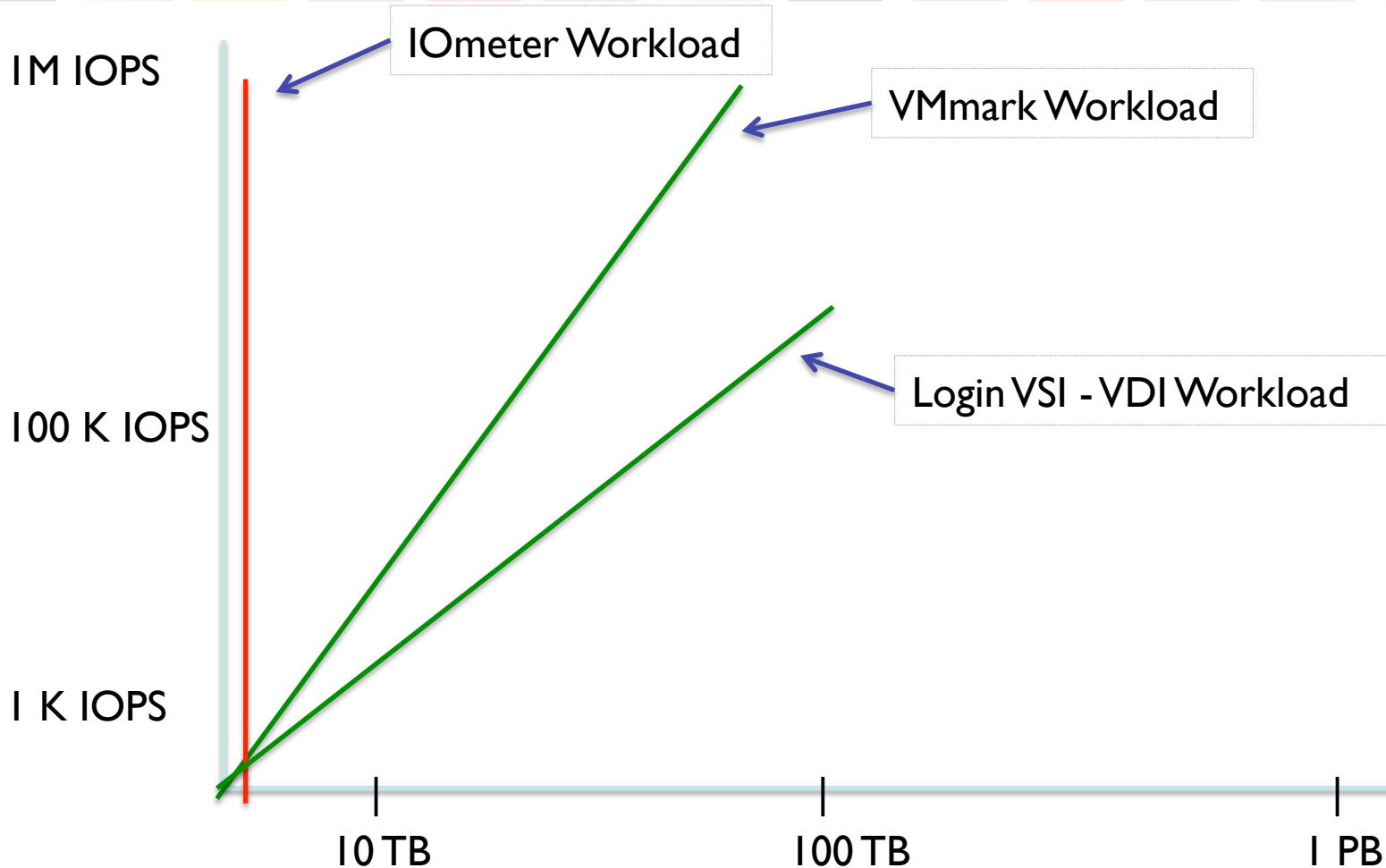
- The material contained in this tutorial is copyrighted by the SNIA unless otherwise noted.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
  - ◆ Any slide or slides used must be reproduced in their entirety without modification
  - ◆ 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.**

# Understanding Solid-State Performance

- Rule 1: Vendor Reported Results are a Guide
  - ◆ At best, vendor reported results provide a guideline
- Rule 2: Never Rely Upon Vendor Results
  - ◆ Results range from under reported, to intentionally misleading
  - ◆ Typically ideal conditions designed to highlight their design
- Rule 3: Reported Results are Not Comparable
  - ◆ Due to differences in test setup, reporting, degree of their unintentional misleading results
- Rule 4: Benchmarks are Interesting
  - ◆ Will provide accurate head to head comparisons, but...
  - ◆ Only for applications that match those in the benchmark

# Why Tool Selection Matters



Total Storage Capacity Required for Testing

# Common Validation Mistakes

- Not Accounting for Solid-State's behavior
  - ◆ Writes at 0% vs. 90%, overwrite, garbage collection, etc. etc.
- Not Achieving Steady State
  - ◆ Running tests for < time than observed steady state performance
- Measuring Results in the Wrong Place
  - ◆ Do not measure on storage system (convenient but inaccurate)
- Using the wrong tools
  - ◆ Some tools can produce very misleading results

# Choose the Correct Tool

- Data Patterns are Important
  - ◆ Particularly with deduplication and compression
- The Wrong Tools will Mislead You
  - ◆ IOmeter
    - › Full Random Data = Does not compress, may be de-duplicatable
    - › Pseudo Random = Compressible, but not reproducible results
    - › No control over hot-spots, and multiple simultaneous workloads
  - ◆ SQLIO & SQLIOSIM
    - › Generates 0x0 “zero” data => Infinitely de-duplicatable and compressible
    - › Unable to generate high data rates > 2 GB/s

# Example Performance Issues

## ➤ Example 1: All-Flash System

- ◆ Initial Read / Write : 70,000 I/O's / sec. at 22 KB average
- ◆ Overwrite Read / Write : 38,000 I/O's / sec. at 22 KB average

## ➤ Example 2: Hybrid System

- ◆ SSD Performance : 6,000 IO's / sec.
- ◆ HDD Performance : < 1,000 IO's / sec.

## ➤ Example 3: All-Flash System

- ◆ Claim: 1M I/O's / sec.
- ◆ Actual: Internal, 1M I/O's @ 4KB, External 400K I/O's @ 4 KB

# Evaluation Methods



# Evaluation Types

## ➤ Performance Validation

- ◆ Evaluate a system can achieve the performance levels specified
- ◆ Verify performance goals stated in RFQ or other criteria

## ➤ Bake Off

- ◆ Compare two or more systems running agreed upon workloads
- ◆ Head to head comparison, may be run with other testing

## ➤ Proof of Concept

- ◆ Designed to evaluate a system's performance running one or more applications
- ◆ Evaluation based upon running intended production applications

# Accurately Measuring Solid-State

- **Solid-State Must be Tested Differently than HDD**
  - ◆ Performance different at 0% capacity vs. over-writes at 90%
  - ◆ Often include some type of data reduction technology
- **Hybrid Systems Have Additional Issues**
  - ◆ Vast performance differences for solid-state vs. HDD
  - ◆ Above mentioned issues effect solid-state portion
  - ◆ Must understand performance from each media independently
- **Data Reduction Technologies Impact Performance**
  - ◆ Testing must account for compression and dedupe separately

# Additional Considerations

## ➤ Hybrid Systems

- ◆ Utilize caching, tiering or both with solid-state and HDD media
- ◆ Data location determines performance, must match application
- ◆ Understand HDD, Mixed and all Solid-State profiles

## ➤ All Flash Systems

- ◆ Must test Flash with garbage collection
- ◆ Test using production environment (Data Services / Reduction)

## ➤ Data Services

- ◆ Snapshots often have performance and / or capacity impact
- ◆ Data reduction technology (thin, dedupe, compress) may each impact performance differently, depending upon workload

# Storage Efficiency Technologies \*

Capacity Technology	VM	VDI	Database
Thin Provisioning	3x	2x	3x
Compression	2x	1.5x	3x
Deduplication	1.5x	30x	0x
Average Improvement	5x	20x	5x
Max Possible	10x	100x	10x

\* Note: Guidelines based on Evaluator Group customer reports, observed industry reports and other sources

# Sample Evaluation Analysis

- Three Methods for Re-Creating Production Applications
  1. Run actual Application
  2. Capture or “trace” existing production applications
  3. Find representative application workloads
- Running the Actual Application – Has Issues
  - ◆ Requires extensive planning and equipment
- Tracing – Can be Accurate with Some Issues
  - ◆ Accurate, but can be invasive
  - ◆ Tracing tools are OS / Hypervisor dependent
- Run Similar Applications – Difficult to Choose App
  - ◆ Can be a good proxy, but must be very similar

- ◆ Step 1: Analyze Expected Application Workloads
  - ◆ Single Application
    - › Take Measurements
    - › Capture I/O traces if possible
  - ◆ Mixed Workloads
    - › Measure each Application if possible
- ◆ Step 2: Validate Workloads
  - ◆ If possible, validate constructed or captured workloads against existing storage, compare to current performance
  - ◆ Other agreed upon validation
- ◆ Step 3: Run Workloads against New Systems
  - ◆ Using I/O generation, run workload against new system

# Performance Validation Process

## ➤ Choose Representative Tests

- ◆ Must use appropriate parameters for Synthetic Workloads
  - Read vs. Write ratio, Random vs. Sequential
  - Block size and size mixture (no application uses only 1 block size)
- ◆ Use Application Workloads or Traces
  - Multiple examples

## ➤ Storage Must be the Bottleneck

- ◆ Cannot be Server or Network / SAN bound

## ➤ Issues and Caveats

- ◆ Watch for issues outlined previously



# System Evaluation

## Recommendations

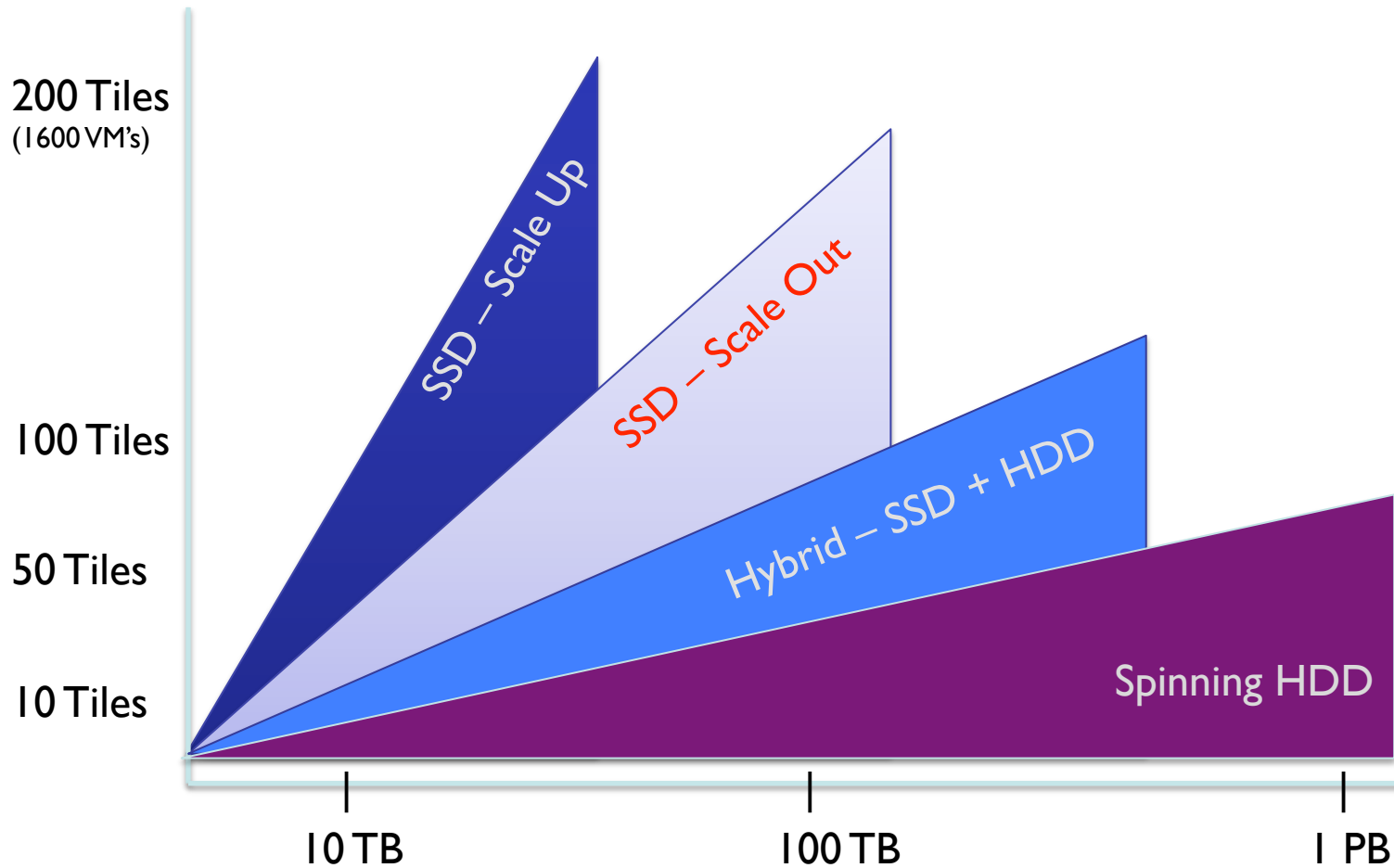
# Types of Workload Tests

- **Application**
  - ◆ An application generated workload
- **I/O Trace**
  - ◆ Similar, if not exact representation of Application workload
- **Synthetic**
  - ◆ Not generated from a Trace or an Application
  - ◆ I/O patterns, which may or may not be similar to applications
- **Workload Tool**
  - ◆ May utilize portions of application I/O, with synthetic parameters
- **Benchmark**
  - ◆ Specific workload run with explicit requirements

# Validation Testing Overview

- Pre-Conditioning is MANDATORY ( ← Read again)
  - ◆ Pre-write data to system to fill to capacity
  - ◆ Delete data, running tests at > 85% of capacity
- Run Workloads until Steady-State is Observed
  - ◆ Steady-state is on a per application basis, repeat for each test
- Run Measured Workloads at Steady-State
  - ◆ Synthetic (e.g. Storage “4 Corners” test)
    - › 100% Read, then Write (Random then Sequential)
  - ◆ Traced or Application Workloads

# Performance vs. Capacity



Total Storage Capacity Required for Testing

## ➤ Multiple Tools Available

- ◆ Windows – “WPT”, includes XPERF and other tools
- ◆ Linux – “blktrace” built-in to SCSI block device driver
- ◆ Unix – OS dependent, “filemon” on AIX, “Dtrace” on Solaris
- ◆ VMware – “vSCSI” tracing built-in

## ➤ May Need Multiple Traces, or Multiple Hours

- ◆ In order to recreate multiple instances of applications

## ➤ Replay Tools

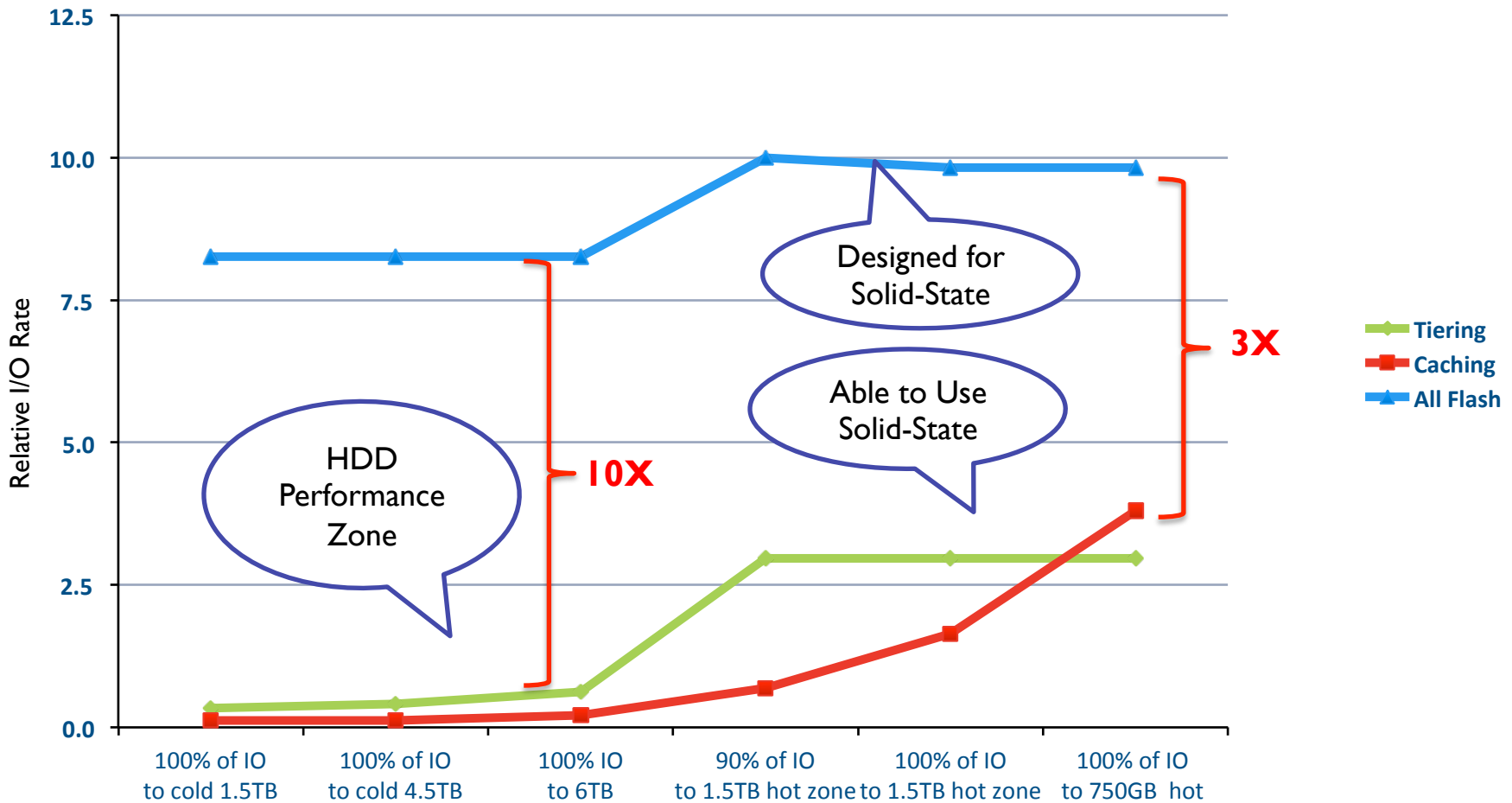
- ◆ A 1:1 requirement of OS to replay makes replay inefficient
- ◆ Look for replay capabilities that support many to 1 (M:1)
- ◆ Replay may be dependent upon trace tool

# Considerations for Hybrid Systems

- **Hybrids Behave Differently and Must be Tested Differently**
  - ◆ Must understand data promotion and demotion times
  - ◆ Conditions may apply (ran. vs. seq., read vs. write, etc.)
- **Understand Caching & Tiering Design Differences**
  - ◆ How quickly data is copied or migrated
  - ◆ Some systems may utilize both, further complicating tests
- **Two Distinct Performance Sets (bi-modal distribution)**
  - ◆ HDD performance vs. Solid-State performance
  - ◆ Conditions that may cause thrashing

# Results: Three System @ \$100K

Normalized I/O Rate @ < 20ms Avg. Resp. 6TB LUN w/ 2TB SSD



- Follow the Guidelines Provided by SNIA SSSI
  - ◆ SSSS (S4) TWG is developing guidelines for performing testing
- Determine Type of Test (POC, Validation or Bake-Off)
  - ◆ If POC, need to accurately characterize applications
- Do Not Allow Stakeholders to Perform Testing
  - ◆ Use internal, or 3<sup>rd</sup> party organization
- Do Use the Right Tools
  - ◆ Hint – IOmeter is rarely the right tool
  - ◆ Synthetic Workloads
    - › Calypso, Load DynamiX, Linux fio, vdbench
  - ◆ Application based workloads
    - › VMmark, SLOB, Login-VSI, IOmark-VM, IOmark-VDI, etc.



# Attribution & Feedback

The SNIA Education Committee thanks the following Individuals for their contributions to this Tutorial.

## Authorship History

Russ Fellows, March 2015

Updates:  
N/A

## Additional Contributors

Leah Schoeb – Intel

*Please send any questions or comments regarding this SNIA Tutorial to [tracktutorials@snia.org](mailto:tracktutorials@snia.org)*