SNIA DSI '16
DATA STORAGE INNOVATION CONFERENCE
Innovation in Storage Products, Services, and Solutions
June 13-15, 2016 | Marriott San Mateo | San Mateo, CA

Workload Acquisition for the Enterprise Data Center

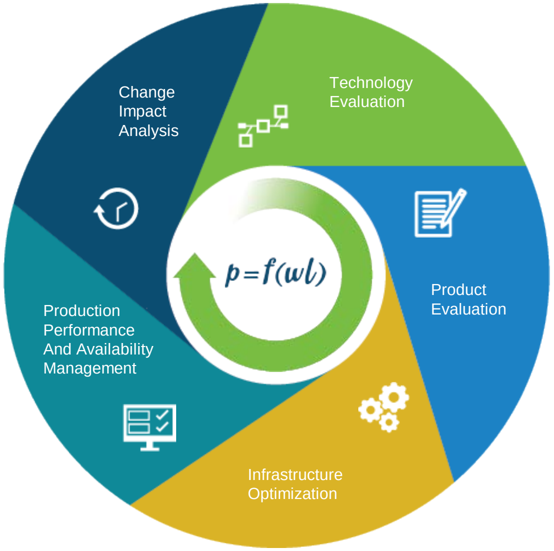
Peter Murray
Virtual Instruments

Hard Problems You're Trying to Understand

- How to most accurately test all solid state storage arrays
- Approaches for assessing storage performance
- How to select the best methodology for YOUR application(s)
- Find proven strategies to help avoid over-spending

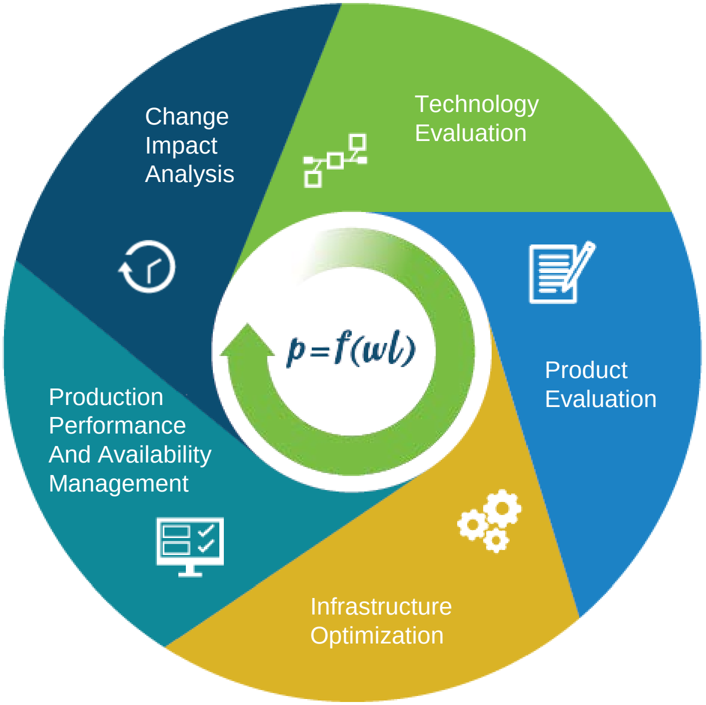
Application Emulation

- The best way to test all solid state arrays is to emulate real applications
- Important application stream characteristics
  - Temporal locality
    - When data is written/read
  - Spatial locality
    - Where data is written/read
  - Data content patterns
    - Random or compressible
    - Some patterns repeat more than others
- These characteristics are critical to understanding SS array performance



The Journey: How Did we Get Here?

- Storage testing was black art
- Testing programs derived from disk drive utilities
  - Did not represent actual applications
  - Could not emulate spatial temporal or spatial locality
  - Did not emulate Data Content
- Difficult to emulate the varying load on many LUNs
- Difficult or impossible to configure the metadata and structure required to emulate file-based apps



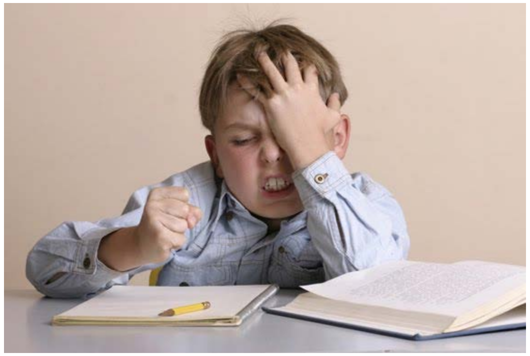
The World has Changed – Don't Miss It

- Before flash, disk drives were the storage performance bottleneck
  - Short-stroking and other techniques helped but were inadequate
  - Data reduction rarely used because it added to transaction times

- Solid state memory technologies change this model
  - Read access time is unaffected by data location
  - Any location can now be accessed as quickly as any other

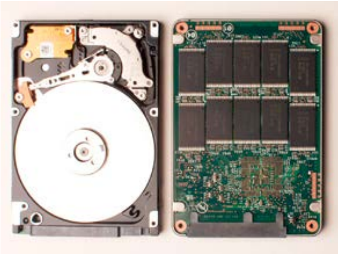
Writing to Solid State Arrays

- Solid state memory has a limited number of write cycles
- Therefore, modern solid state storage arrays avoid writing
- Write access is very different than read access
- Flash write access time is implementation dependent
  - Sequential writing may be impacted
  - Random writing can impact garbage collection
- Data reduction processing may require post-processing
  - But typically does not affect write speed



Writing is Hard

How is Flash Different?



- Addressable storage space is likely less than raw space
  - May help avoid performance issues during garbage collection
  - Other methods are available to avoid performance issues
  - Can help increase flash life
- Deduplication & compression decrease storage requirements for an app
  - More storage per nominal byte
  - But, performance may be impacted
- Advanced metadata processing & workload profiles at scale make it harder to saturate an array
  - Test at near full capacity to understand array performance
- Testing with hotspots helps model application behavior
  - Garbage collection or metadata processing may affect performance
- Software services & protocols – software runs differently on SSD than on HDD

SS Arrays Require New Storage Testing Methods

- Applications exhibit spatial and temporal locality
  - Modern solid state arrays are designed with this in mind
- Application traffic contains data content
  - Data is random or compressible
  - Data may also be de-duplicatable
  - All content types are present in most applications
- Some all solid-state storage arrays must be tested with local
  - Data reduction is a key feature - can't be turned off
  - Legacy testing apps cannot emulate the locality, content or content flocking present in applications
- New thinking and testing applications are mandatory!

Storage Performance

- Vendors have good stories, but don't confuse marketing with reality
- Vendors endorse performance testing with your workloads, derived from production environments, via synthetic workloads
- Vendors and standards organizations produce benchmarks, but they are guidelines at best
- Benchmarks don't offer configuration guidance – and don't represent <u>your</u> workloads

Why Performance Testing is Important

- Which is the best technology for my needs?
- Which is the best vendor / product for my needs?
- What is the optimal configuration for my array?
- Does performance degrade with enterprise features:
  - Deduplication
  - Compression
  - Snapshots, Clones, Replication
- What are the performance limits of a potential configuration?
- How does an array behave when it reaches its performance limit?
- Does performance degrade over time?
- Which workloads are best for an AFA? A hybrid storage array?

Why Performance Testing is Important

Why Performance Testing is Important

Traditional Storage Testing Approaches

- Limits finding
- Functional testing
- Error Injection
- Soak testing

Storage Performance Validation

2 core methodologies

Workload Modeling
Simulate the I/O profiles of your production environment

Performance Profiling
Fully characterize performance of arrays under wide variety of load parameters

Performance Profiling for Vendors

Performance Profiling

- Characterization under a wide range of workload conditions
- Understand sweet spots and weaknesses of an array
- Sometimes referred to as “4 corners” or “limits” testing, but you can do much more than that
- Vendors need these tests to validate portions of a storage array
- IT customers do not generally benefit from this testing
  - Applications don’t act like performance profiles
  - Some exceptions; e.g. queue depth or outstanding commands

Performance Profiling

Iteration Parameters

Access Pattern - Read %	0, 20, 40, 60, 80, 100
I/O - Constant Request Size	4KB, 8KB, 16KB, 32KB, 64KB
Port - Tx Queue Depth (FC only)	1, 2, 4, 8, 16, 32, 64, 128
Load - Throughput Value	1MB, 5MB, 10MB
Data Reduction - Uncompressed to compressed ratio	2.0, 1.5

Add Iteration Parameter

Number of configured iterations: 1440

Performance Profiling

Fibre channel performance

Finished

Started by admin

1:00:05:15 00:00

Last Log Record: 2015-01-26 11:38:35 AM | Success | Test Suite finished

Logs

Iteration Results Export to CSV

#	Status	Duration	Access Pattern - Read %	I/O - Constant Request Size	Port - Tx Queue Depth (FC only)	Load - Throughput Value	Data Reduction - Uncompressed to compressed ratio	SCSI Throughput (average)	SCSI IOs Succeeded/sec (average)	SCSI Average Response/Latency Time (average)
48	Finished	01:01	0	4KB	128	10MB	1.5	8.3 MB/sec	2115.387	6 ms
47	Finished	01:00	0	4KB	128	10MB	2	8.0 MB/sec	2044.602	.7 ms
42	Finished	01:00	0	4KB	64	10MB	1.5	7.5 MB/sec	1921.051	.5 ms
41	Finished	01:00	0	4KB	64	10MB	2	7.2 MB/sec	1837.487	.9 ms
36	Finished	01:00	0	4KB	32	10MB	1.5	6.5 MB/sec	1663.073	.3 ms
288	Finished	01:00	20	4KB	128	10MB	1.5	6.5 MB/sec	1657.239	.5 ms
35	Finished	01:00	0	4KB	32	10MB	2	6.3 MB/sec	1612.252	.5 ms
282	Finished	01:00	20	4KB	64	10MB	1.5	6.2 MB/sec	1586.806	.8 ms
281	Finished	01:00	20	4KB	64	10MB	2	6.1 MB/sec	1554.01	.1 ms
287	Finished	01:01	20	4KB	128	10MB	2	6.1 MB/sec	1545.593	.7 ms

Workload Modeling

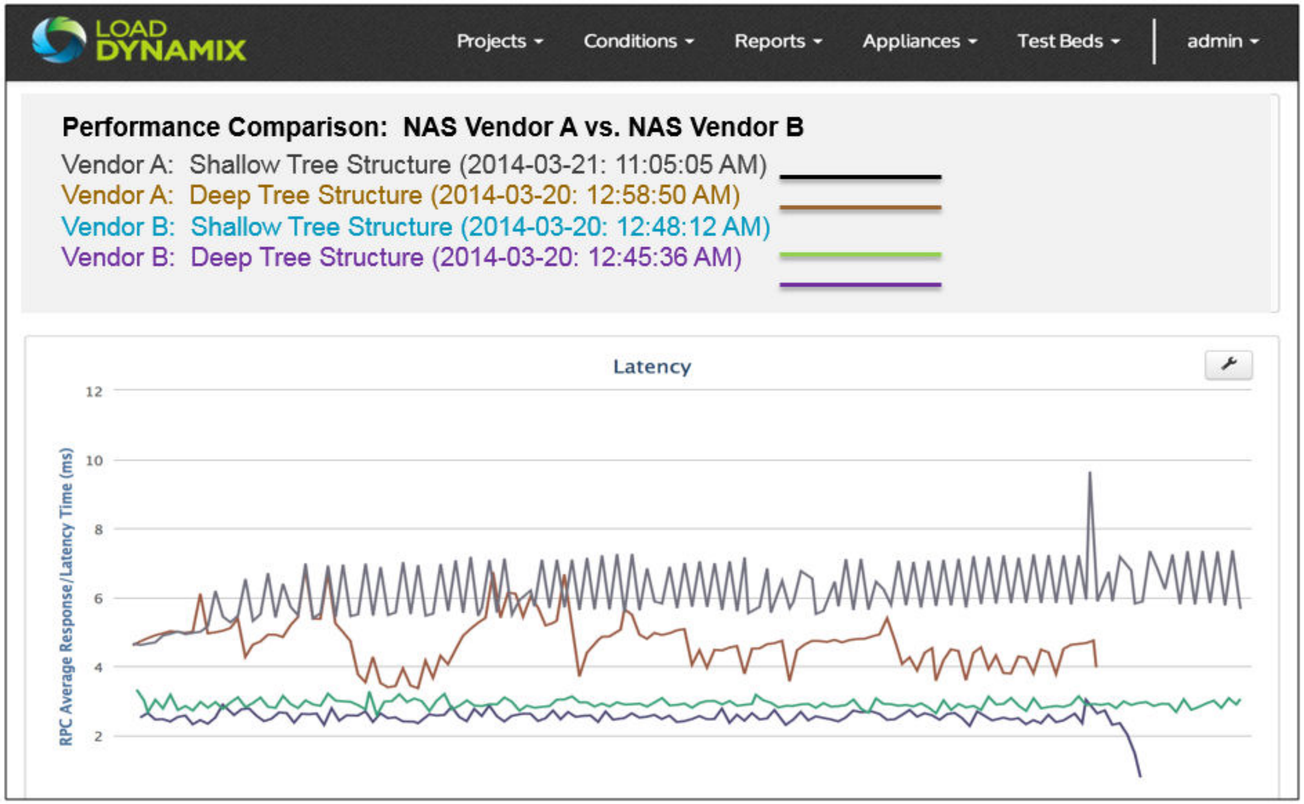
Workload Modeling
Simulate the I/O profiles of your production environment

Performance Profiling
Fully characterize performance of arrays under wide variety of load parameters

Workload Modeling

- Stresses an array using a realistic simulation of the specific production workload/s
  - For IT customers, from your current environment
  - For vendors, using customer examples or "dog food"
- Realism is paramount – realistic I/O profiles
- Packet traces offer limited utility in testing
  - Huge volume of data
  - Short Duration
  - Security concerns

Workload Modeling



Where Does Workload Modeling Come From?

- Customers ask for workload models
  - IT customers want models of their workloads
  - Vendors want “the” workload
    - Oracle, Exchange, etc.
- IT customers ask to help make better decisions about:
  - Upgrading storage hardware or software
  - Changing storage network configuration
- Vendors ask for help to:
  - Test customer examples/issues
  - Find realistic scaling limits to test app growth over time

Result: A New Modeling Method

- Cloud-based workload modeling
- Community-based workload sharing
- Workload model that be ingested into Virtual Networks load generation
- More realistic and scalable than benchmarks

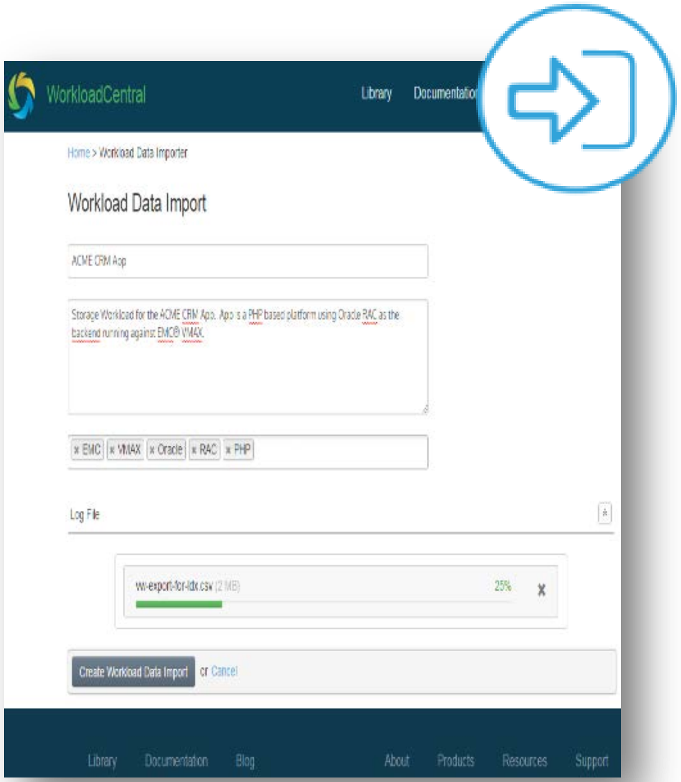
Workload Central Beta

WorkloadCentral is a free cloud-based analytics platform and community that allows you to understand analyze, create and share workloads.

Available at: www.workloadcentral.com

- Key Features:
  - Free workload analysis & creation
  - Advanced workload analytics
  - Workloads for validation, testing & benchmarking
  - Workload Library, community & discussion

Uploading Your Workload Data



The Workload Importer offers:

- Ability to upload data from any vendor or environment
- Out of the box import policies
- Analysis policies provide flexibility to define different workloads

Visualizing Your Data with the Workload Analyzer

A free downloadable, printable report and dashboard that provides:

- Workload access pattern
- Workload behavior characteristics
- Workload performance
- Workload creation

Running a Block-Based Workload Model

Workload Modeling

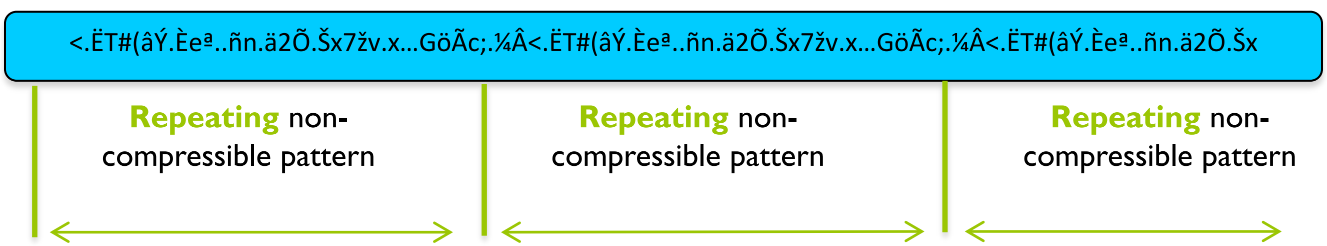
Simulate the I/O profiles of your production environment

1. Characterize Workload I/O

- Per-LUN I/O:
  - Read-Write Mix
  - Random or sequential access
  - Hot spots and hot spot drift
- Data Content
  - Randomness
  - Compressibility
  - Unique vs. duplicated blocks

2. Determine Data Content Patterns

- **Data content patterns**
  - Created during preconditioning
- **Data content streams**
  - Created during preconditioning
  - Replayed during testing
- **Consist of repeating and non-repeating patterns**
  - Random
  - Compressible
- **Consist of varying pattern lengths**



3. Build I/O Models

- Decide when to model
- Boot storm
- Everyday office load
- Backups
- End of period processing
  - Month, Quarter, year end
- Test primary models individually
- Test periodic models on top of everyday load
- Magnify load to test expected maximums

4. Run Workload Models

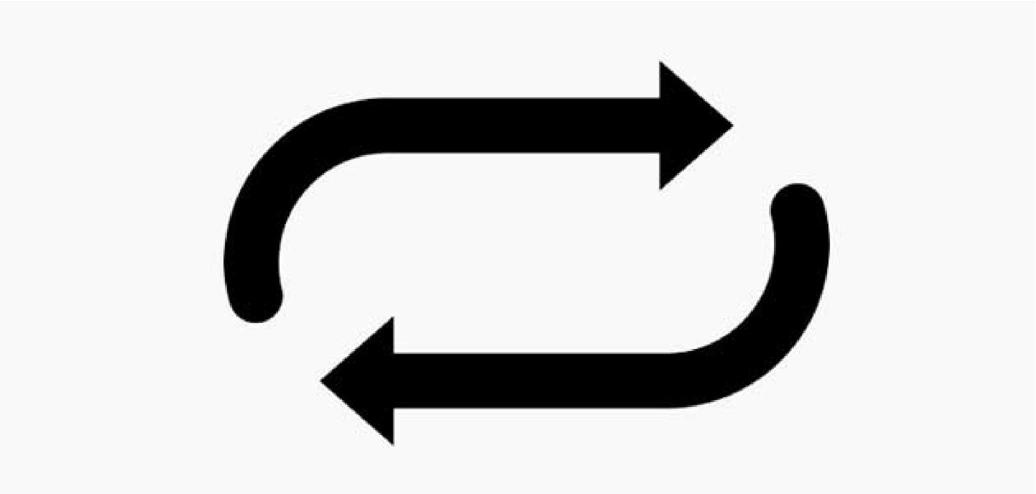
- Run most common model(s) first
  - Determine baseline performance
- Add periodic models to common model
- Combine apps if appropriate and test together

5. Test Array Features

- Test effect of MPIO
- Test effect of maintenance / other management activities
- Test at or near full capacities
- Test effect of QoS

Test in an Iterative Manner

- Run
- Analyze
- Repeat as necessary
  - Change testing to reflect business conditions

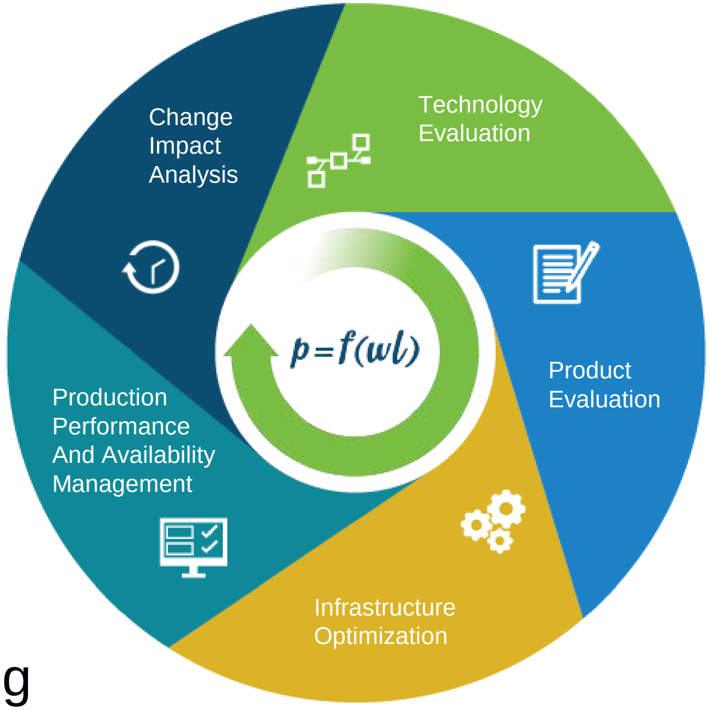


Summary

- Performance assurance
- Reduced storage costs
- Increased uptime
- Acceleration of new application deployments

Summary

- Application Testing is now mandatory
  - Black art has become repeatable
- No synthetic workload is perfect
  - But is the best approach available
  - This will only improve over time
- Customers can see:
  - How closely the model emulates apps
  - A realistic view of how an array operates
- This new model is changing storage testing



Company Overview

Global Leader in Infrastructure Performance Analytics

- ❒ Founded in 2008
- ❒ HQ in San Jose, CA
- ❒ Global 2000 Customers
- ❒ Every Major Vertical
- ❒ 44 of the Fortune 100
- ❒ Merged with Load DynamiX in April 2016