

# STORAGE SWITZERLAND

## ARE YOU PLANNING FOR STORAGE PERFORMANCE?

### HOW TO BE INTELLIGENT ABOUT FLASH STORAGE DEPLOYMENTS



Colm Keegan, Senior Analyst

Storage performance has become a “hot button” issue in the server virtualization era. While there is a need for high performance storage, such as flash and hybrid storage solutions, there is a risk that IT planners could just be blindly “throwing hardware” at performance issues. Moreover, the lack of insight into when key applications will hit a wall, (when storage resources cannot keep up with I/O demands being generated by applications), leaves many organizations vulnerable to not meeting service level agreements (SLAs). Before deploying any changes into a production environment, storage managers need a way to proactively identify when performance ceilings will be breached and how to evaluate the technology options for best achieving application I/O workload requirements.

#### Betting on Benchmarks

When planning for storage performance upgrades, infrastructure planners often rely on benchmark tests released by storage vendors. The challenges with this approach are two-fold: first, most vendor performance tests are produced in very tightly controlled lab settings that are designed to showcase their products in the best possible conditions. Second, these benchmarks don’t typically represent real world environments – and certainly not ones matching your applications.

In many cases, benchmarks are designed to show what the theoretical storage I/O limits are of the particular solution. Vendor benchmarks tend to represent a “best case” scenario for most IT environments, and simply are not realistic. While there are independent bodies that publish standard workload benchmarks that provide general guidelines on how certain products perform, these benchmarks cannot account for all the various permutations of a given application workload and its associated infrastructure. Instead, vendors tend to use these benchmarks as a way to demonstrate that they have “leap frogged” the competition.

#### Return on Lab Investments?

Many Fortune 1000 companies have made significant investments in lab infrastructure to independently test solutions and mitigate the risk of introducing bug-ridden code or inadequate products into their environments. Freeware testing tools like IOmeter are often used to measure how much I/O throughput can be processed by a given storage technology, however, its output has limited value since it cannot emulate the specific workload conditions of the actual production environment. It’s like buying a car based on how fast it can accelerate from 0 to 60 mph, when you commute 50 miles on the highway in heavy traffic to and from work every day – interesting, but not very relevant data.

This is especially true in virtual infrastructure where a single server host can generate multiple workloads that make demands of the shared storage infrastructure at the same time. Most benchmarks and testing tools, like IOmeter, generate single streams to represent application workloads. This is not representative of the typical workloads in today's virtualized application environments and consequently offer limited value to infrastructure planners. What's more, performance or load testing with freeware tools requires purchasing a rack of servers and configuring hundreds of VMs along with custom scripting just to approximate real-world production workload capacity. This introduces significant costs and management complexity.

### Hoping For The Best

The fact is that many organizations don't really know how a given storage solution is going to perform until after it's been implemented into the production infrastructure. Likewise, firmware upgrades to existing storage systems are often applied to production infrastructure since there is no practical way to test how the new firmware code is going to interact with every facet of the environment.

Whether it's a new storage platform, application level changes, or a firmware upgrade to an existing storage system, this constitutes a "trial-and-error" approach to storage performance validation which introduces undue risk to businesses applications. Even when new storage systems are successfully implemented into production environments, there is no way for IT planners to accurately gauge when business applications will hit the storage "performance wall".

### Planning For Success

Faulty code or undersized storage arrays can threaten application availability and/or result in service disruptions and undermine the credibility of the IT service organization. A new discipline is needed to enable IT to be more proactive with meeting and enabling consistent storage

performance service levels - "Storage Performance Planning". With the highly dynamic nature of virtualized application environments and the need for businesses to stay agile, this is not a "nice to have" capability, but rather should be a fundamental aspect of proper infrastructure design and management.

### Analyzing the Workload

In order to conduct storage performance planning, infrastructure planners need a way to understand the workload profile of production application environments and generate workload analytics that provide insight into how the workloads interact with the infrastructure. As the name implies, workload analytics is a process whereby intelligence is gathered about the unique characteristics of application workloads in a given environment.

By capturing all of the attributes of real-time production workloads, highly accurate workload models can be generated which enable application and storage infrastructure managers to stress test storage product offerings using THEIR specific workloads. The concept is to extract statistics on real-time production workloads from the storage environment to establish an I/O baseline and identify I/O growth trends.

### Better Predictability

This I/O baseline and trend data can then be used to run performance validation scenarios to demonstrate whether the prospective solution will be capable of meeting current and future workload requirements. In short, the idea of storage performance planning is to determine up front what the I/O limitations are of any given storage platform using near real-time production workload data extracted from the actual user environment. This allows storage planners to predict when more resources, like network bandwidth, storage IOPs, etc., will be needed to maintain application service levels and whether technologies like flash storage or SSDs will be appropriate.

## Trust But Verify

The key is to simulate application workloads in a lab setting first, prior to introducing any changes into the production storage environment. By generating highly realistic application workload emulations to test new storage components, like flash or SSD modules for existing arrays, it is possible for IT planners to ascertain whether there will be any investment benefit.

For example, one of the major challenges of evaluating performance in a hybrid storage array (storage that mixes SSD with hard disk drives) is determining how a cache miss will impact overall performance. By applying actual business application workloads against a proposed hybrid solution and running those workloads over a sufficient period of time, it should be possible for storage architects to analyze how application response time can be impacted during a cache miss event.

As previously discussed, another practical use of effective storage performance planning is for validating new infrastructure firmware code and application level changes. Storage suppliers and fibre channel switch manufacturers may extoll the enhanced performance features available through new firmware code. Application and database admins may implement a new procedure. However, unless there is a way to put these “improvements” through a change validation process that uses true production representative workloads, there is no way to determine if there is a performance benefit or worse, if it will negatively impact application availability.

## Time and Resource Crunch

One of the primary challenges in evaluating the performance of new storage technologies is dedicating the people time and lab resources to collect production workload data. That data is then used to generate a realistic test scenario, at true production scale, which is required to produce a meaningful outcome. While many

companies have invested in lab infrastructure, these environments frequently are used for testing and developing internal and/or 3rd party business software applications. Rarely, if ever, are they fully dedicated for validating the performance infrastructure systems.

As Storage Switzerland covered [in a recent article](#), IT staffs are not increasing to meet the ever growing demands of today’s data center environments. In fact IT staffs are increasingly being asked to “do more with less”.

Automated solutions are needed for infrastructure performance validation, including the ability to understand and model storage I/O workload metrics and eliminate testing as a “guesswork” exercise.

Consequently, most organizations don’t truly know if storage upgrades or technology (e.g. solid solid state) refreshes are going to do the job until they’ve been running in their production environment for a certain period of time. This leaves critical business application environments exposed to service disruptions or even total outages that can adversely affect company brand and top line revenues.

## Automating Storage Performance Planning

This calls for an automated workload modeling and analytics solution which can characterize actual production workloads so that they can then be modeled, reproduced and used for evaluating various storage vendor technology offerings. Solutions like the [Load DynamiX](#) Performance Validation Appliance enable infrastructure planners to bring more automation into storage performance planning and validation process.

Their purpose-built storage performance validation appliance can be used to help organizations pre-determine if new components for storage platforms, like SSD modules for existing arrays, new storage platforms or even if new firmware versions will deliver value and are safe for deploying into the production environment.

The benefits realized include a reduced risk of application downtime and degraded performance occurring as the result of an undersized solution or faulty firmware/ application code. In addition, since the appliance generates load based on emulations of existing application workloads, it can actually help businesses pinpoint the exact storage configuration to avoid over spending on storage infrastructure upgrades. You will know the performance trade-offs between SSDs and HDDs and how much flash to deploy. What's more, the appliance can also do predictive analysis to determine ahead of time when the current storage environment will reach a performance wall; enabling organizations to avoid storage latency issues and consistently maintain SLA's.

## Conclusion

The growth of business data coupled with the need to remain agile in a highly competitive global economy are placing significant burdens on IT infrastructure environments. Application workloads are constantly changing or 'drifting'. The need to consistently deliver a reliable business service -- regardless of the continuous changes taking place in the data center -- calls for automated performance validation solutions like Load DynamiX. By using Load DynamiX to generate tests based on your actual production workloads, organizations can be more confident about their technology purchasing decisions and avoid the "wall of worry" that often accompanies infrastructure changes.

### **About Storage Switzerland**

Storage Switzerland is an analyst firm focused on the virtualization and storage marketplaces. For more information please visit our web site: <http://www.storage-switzerland.com>

*Copyright © 2013 Storage Switzerland, Inc. - All rights reserved*