

What is the total cost of Storage testing?

by George Crump, Lead Analyst



Welcome to the new world of storage technology evaluation

IT planners are being overwhelmed with new storage solutions to help them solve old problems. Most of these solutions leverage flash memory to help them address challenges like randomized I/O, scale and overall response time. The vendor engineers and sales people that represent these products will come and go, but the IT operations professional and the supported Business Unit managers have to live with their product selection for 3 to 5 years. To ensure that your next storage system will have a bright future in the organization requires an accurate, real world assessment of how it will perform on the first day as well as in year five.

The Broken Storage Selection Process

When selecting storage technologies in the past, IT planners would start the process by interviewing several vendors. A short list of vendors would result from these interviews who would then be invited to bring in their equipment for a POC (proof of concept) to confirm it performed as promised. The IT team had time and potentially even dedicated IT staff to perform these tests. They also often had a dedicated lab. The environment they were attempting to simulate was a single, or at most, a few applications on dedicated servers with dedicated network connections. The performance potential of the systems they were testing were all limited by the underlying hard drives that populated them. They could only go so fast and hard drives were usually the gating factor. In addition, the application owners and users had modest expectations for performance; they weren't expecting hundreds of thousands of IOPS.

Because of these limitations, the grading system for storage systems was a simple "Pass - Fail". It was either good enough or it wasn't.

The Modern, Complex Storage World

In today's storage world, flash (combined with the new, higher expectations of users) breaks the old storage selection process. Users expect almost instant provisioning, and virtualization technologies have greatly enabled this, while creating a much greater strain on the network storage infrastructure. These storage systems have to generate potentially millions of IOPS with almost no latency. And users expect their data center environments to scale to thousands of virtual machines or millions of users, which now create unprecedented random I/O streams.

In addition, there are significant ways the new storage system's performance and cost is affected: (1) the performance and cost of the flash storage technology and (2) the way in which the storage system architecture is designed. Now storage systems need to be assigned a more granular grade -- an "A-F" for all test performance outcomes. The subtle grade differences are now the deciding factor between meeting user expectations or having to scrap a storage system because it can't keep up with workload demands. To make matters worse, instead of several storage vendors to choose from, there are now more than three dozen who all claim to have the required performance capabilities.

The Cost of Simulating the Real World

When creating a test environment, there are two types of costs to consider; hard costs and soft costs. The first hard cost to consider is somewhat obvious -- the cost of purchasing all the required lab testing infrastructure. While virtualization helps to keep lab costs down, it can only help so much. If the mix of VM to physical hosts is not representative of what it is in production, then the test data can be significantly skewed. Storage Switzerland recommends that test labs have sufficient equipment to stress and load test all of the core infrastructure components: servers, switches and storage both independently and as an integrated entity.

Second, since this is a storage test, the storage network has to be physically similar to the production environment. While a vendor will provide the storage system to test against, they don't typically provide the storage networking infrastructure. Certainly fewer hosts means fewer host bus adapters (HBAs) and fewer physical switches, but even this investment, particularly for a large enterprise, can be quite substantial. Third, of course, is the storage system itself. As stated above, most vendors that make the short list will provide an evaluation unit at no charge. But there are often strings attached. They will want to know how you are going to test and what will cause you to make a yes or no decision. Most will not accept "we are going to see if we like it" as the testing criteria.

Fourth, to do a real apples-to-apples comparison, the test lab environment must be kept absolutely stable and unchanging for the duration of the tests. Controllers, cables, FEPs, firmware, etc cannot change. And if you want to compare a new system to the existing production system, what are the chances that the test bed you used a year ago is still configured exactly the same? None.

The Soft Costs Will Kill You

The first soft cost and one that may catch many IT planners off-guard, is the time it takes to create and articulate a testing plan. Again, most vendors are going to want an understanding of how their unit will be tested and what the criteria for success will be. Since they are going to place the hardware in your data center they *do* have the right to ask. The time it takes to design a test that is consistent across the short list vendors can be extensive.

What makes test plan creation so expensive is deciding on how the simulation of the production application(s) I/O pattern will be simulated and at what scale? Will it be done by creating a subset of the application itself and creating simulated workers? Or will it be done by leveraging a testing utility, like Iometer, that generates an I/O pattern, but has no real tie back to the actual applications. Creating a subset of the actual environment can be expensive, but using testing utilities that don't use truly representative application I/O patterns and can't scale to full production levels will lead to poor choices.

Running The Test

The second soft cost is how long it takes to actually prepare and run the test. A big part of this issue is the human cost. How much manual intervention is required and how much of the test must be scripted? How much time do you need to write custom reports? An advantage of some I/O testing utilities is that they can be easily automated and set to run continuously, but this hardly makes up for the inability to simulate a real world workload.

The other aspect of running the test is how long can the test be run? Storage systems perform differently over time, especially as more data is stored on them. As more data is stored, NAS systems, for example, will often degrade in performance. Flash systems, on the other hand, will degrade as wear leveling technologies have to work harder and products with deduplication have to manage a growing meta-data catalog. It is important that the storage system be tested not only for initial performance but also performance over time as it reaches its capacity limits. The testing needs to be automated and hands-off so it can run for very long durations.

The Cost of Getting "IT" Wrong

The most expensive cost of a storage test is the cost of getting the test wrong. Overlooking for the moment the potential harm that can be done to one's IT career, there is the cost of having to replace a storage system or the cost of having to add significant resources to that storage system long before it is fully amortized. There's the impact to a business that results from poor performance such as lost revenues or unhappy customers. Then of course, there is the damage to the IT team's credibility for missing badly on evaluating the storage system's potential. Vendor blame can only go so far, eventually the buck stops with IT.

Solving the Problem

When it comes to storage testing, the goal is to simulate real world application workload I/O patterns, without having to create a mirror image of production infrastructure. The goal should be to capture what that I/O pattern looks like over a period of time, create a realistic workload model that reflects the production I/O patterns, and then iteratively replay that pattern based on the anticipated growth of those workloads, at exactly the scale of the production environment. At a minimum, this means capturing a mixture of random reads and writes across dozens or hundreds of parallel threads and gathering data on block and file size distributions and directory structures.

Companies like Load DynamiX provide solutions that do just that through an "appliance" approach to storage performance validation. These appliances enable highly realistic application I/O workload modeling to create workloads that can generate storage I/O for any period of time against any flash, hybrid or traditional storage platform. This process can also be used to determine the scalability of the environment to identify the true performance limits for your workloads. Essentially the IT planner turns up the volume during playback to determine if the storage system being tested could take the data center into next year or into the next decade. For the first time, IT can know, before the user, when performance will start to exceed SLAs.

Conclusion

Storage performance testing is a significant challenge in getting budget approval where IT is expected to do more with less. The cost and time it takes to simulate production operations often leads to the temptation to cut corners and either take a vendor's word for it or do only a rudimentary evaluation. Giving in to either of these impulses often leads to substantial wasted money or dissatisfied end-users. By understanding the real-time I/O patterns present in production applications and allowing those patterns to be tuned up or down for 'what-if' analysis, companies like [Load DynamiX](#) are making storage testing and validation a cost effective, repeatable process based firmly in reality.