WHITE PAPER

Keeping Storage from Becoming a Jurassic World

By Jon Toigo

Tintri

TABLE OF CONTENTS

Tintri

INTRODUCTION

Business-savvy IT planning has always involved the matching of compute, network and storage technologies and architectures to the performance and availability requirements of applications – and also to budgetary realities. In the real world, IT budgets impose hard boundaries on infrastructure options, forcing IT planners to choose between "safe" (but often less efficient) technologies and riskier (but often more innovative and potentially more efficient) alternatives. This dilemma is a persistent one in the history of business information technology. It tends to become more pronounced when workload requirements push *status quo* infrastructure to its limits.

Traditional storage infrastructure is now laboring under a new set of requirements, approaches, and budget realities.

Surveying the current business technology landscape, it can be argued that an inflection point is being reached in the realm of storage technology today. Simply put, traditional storage infrastructure, whose designs reflect engineering decisions that were made three or four decades ago in response to the workload processing requirements, application hosting models, and technology cost metrics of that earlier time, is now laboring under a new set of requirements, approaches, and budget realities.

Subtle changes to storage infrastructure, like the return to direct-attached topologies and the introduction of "hyper-converged" hardware/software stacks, may not prove sufficient to address the realities of virtual machine-based processing, fast silicon storage media, and hybrid cloud architectures. Planners are struggling to reassure themselves that making "safe" choices in storage infrastructure re-engineering doesn't lead to the deployment of technology that is already obsolete by the time that the project is complete.

Avoiding the construction of a *Jurassic World* in your storage infrastructure requires the consideration of how workloads have changed and what the best way is to build storage infrastructure to accommodate the new technology requirements that new workload imposes. With a vision of what the infrastructure needs to be, intelligent choices must be made to identify and contain the costs of a new infrastructure that aims to optimize application performance and improve data availability. This typically comes down to a combination of resource and service design choices and an emphasis on management capabilities that reduce labor costs and make fewer administrators more productive.

The challenges are many, but the job can be made simpler when thought leaders in the vendor community are actively revisiting storage fundamentals, deconstructing fundamental assumptions, and re-engineering products with new workload requirements in mind. This paper describes such an exercise in storage thought leadership.

HOW STORAGE REQUIREMENTS HAVE CHANGED

Like death and taxes, another of the inevitabilities of contemporary IT is the growth of digital data.

Like death and taxes, another of the inevitabilities of contemporary IT is the growth of digital data. Volumes have been written about the "digital explosion" noting a rapid acceleration in the rate at which organizations are accumulating digital data. While many of these reports are often stilted or exaggerated to serve the purposes of storage vendor sponsors (truth be told, much of the data explosion involves digitization of information once stored on analog media and has less impact on business storage requirements than analysts may suggest), companies do report substantial increases year over year in storage capacity demand.

In fact, growing storage capacity demand is a function of data growth – and of poor capacity allocation and utilization efficiency. The old saw still holds true: our storage capacity is over-subscribed and under-utilized mainly because organizations manage both their storage infrastructure and their data rather poorly. Another contributing factor that is only now coming under scrutiny is the capacity demand generated by the adoption of server virtualization technologies. The advent of hypervisor computing and the increased tendency to create workload as virtual machines has exacerbated an already problematic storage cost paradigm.

Leading analysts suggest that up to 75% of server workloads will be virtualized by 2017. This is being driven by user expectations about the benefits that server virtualization will provide to organizations including

- **Agile deployment and provisioning:** virtualizing infrastructure will enable the creation of "atomic units of compute" that will make it far easier and less time-consuming for IT to respond to fast-changing business requirements;

- **Reduced infrastructure footprint and CAPEX:** less kit will be

Evangelists have neglected to mention the impact of consolidating multiple servers into a single physical platform on application performance.

required to host more applications, reducing the hardware budget, environmental costs (power, cooling, etc.), and administration and labor costs;

- **Increased performance and availability of apps:** instantiation of virtual machines on clustered server configurations together with "template cut-and-paste" techniques for shifting workload between physical machines will reduce downtime both for planned and unplanned interruption events;

- **Simplified infrastructure and service management:** hypervisor-based system and storage management facilities will enable the consolidation of IT staff specialties and the leaning of administrative personnel requirements thereby reducing OPEX costs.

While analysts agree that about 25% of applications (mostly transaction-oriented databases) will likely remain non-virtualized, maintaining their traditional platform requirements, virtualized environments will dominate the data center and facilitate the integration of private, on-premise compute environments with public cloud services of various types. Such architecture, according to industry leading vendors, will be required to support the "killer apps" of the future such as mobile commerce (m-commerce).

The narrative, while compelling, is incomplete. In fact, the benefits of server virtualization require much more than the replacement of server kit with hypervisor software and compatible hardware. Changes must be made to other layers of infrastructure as well – especially, to storage.

Since the inception of the hypervisor computing trend, evangelists have neglected to mention the impact of consolidating multiple servers into a single physical platform on application performance. Application/virtual machine consolidation, according to one estimate, increased the application I/O requirements of the server significantly requiring an increase in the number of physical I/O ports (LAN and storage interconnects) to between 7 and 16 per physical host. Moving data successfully between hosted applications and storage infrastructure became fraught with latencies, driving considerable attention to the mechanisms by which hosted VMs read and write data.

Over time, complaints about poor application performance produced a litany of recommended changes to storage techniques and technologies from leading hypervisor software vendors. Tighter integration with, and offloading I/O workload to, array controllers were among the first strategies advanced by hypervisor companies to address the performance problem. When this strategy proved unsuccessful, the increased use of memory (DRAM or Flash-based) caching and buffering techniques were proffered to "spoof" what was presumed to be a storage I/O path log jam. "Spoofing" is an old technique used to improve perceived application performance by buffering I/O so that speed differences between compute I/O and storage I/O are masked. This strategy proved minimally successful, but did not address the core issue, which tended to have nothing whatsoever to do with storage infrastructure but rather with the blending of I/O from different applications and the extreme randomization of write data that slowed its access speeds dramatically.

Incremental changes to storage infrastructure did not succeed in resolving the issues created by new application workload characteristics.

Bottom line: incremental changes to storage infrastructure did not succeed in resolving the issues created by new application workload characteristics. This has provoked a search for new approaches that would yield more success, even if they departed from common wisdom and practice. Of course, some approaches have been driven less by pure problem solving than by the traditional quest for market share and profit.

ENTER SOFTWARE-DEFINED STORAGE AND HYPER-CONVERGED INFRASTRUCTURE

The current strategy of the hypervisor community has not been to address the problem of the "I/O blender effect," but instead to villainize so-called "legacy storage" topologies and kits such as Storage Area Networks (SANs) and Network-Attached Storage (NAS). In their place, hypervisor vendors have proposed replacing older infrastructure with new "software-defined storage" (SDS) and "hyper-converged" infrastructures.

SDS is a confusing term that is more marketecture than architecture. All storage is software-defined, whether that software exists as part of the server operating system stack or on an array controller (the motherboard of a storage system). SDS advocates claim that their model divorces value-add services from the array controller, where the services are difficult to manage and impossible to share across storage infrastructure holistically, and places the management and control of services within the hypervisor administration console. Taking the value-add features out of the storage kit should, they claim, cause the CAPEX cost of storage to fall to the prices of the commodity hardware components of the array. Moreover, the OPEX costs for storage administration and management should decline because of management consolidation: no more management of storage on a box-by-box basis.

Combined with the consolidation of storage service management functionality at the hypervisor, SDS advocates go further to suggest that breaking up complex SAN and NAS topologies and "returning to simpler direct-attached storage models" will further reduce costs and challenges when it comes to storage resource allocation to workload. So, with SDS is a strong recommendation to "rip and replace" legacy infrastructure and return to direct-attached storage on every server. For those who prefer not to build their own kit, hypervisor vendors have been teaming with server and storage hardware vendors to create "pre-integrated appliances" that are essentially cobbles of their hypervisor and SDS software components and commodity hardware components that can be deployed readily to fulfill application hosting requirements.

Only hypervisor vendor-approved components can be used to build an SDS infrastructure.

In truth, however, most hypervisor-driven SDS and hyper-converged infrastructure solutions have proven to deliver less application performance improvement to the consumer than hardware-software lock-ins for the hypervisor vendor. Only hypervisor vendor-approved components can be used to build an SDS infrastructure. This storage must be fashioned into a minimum of three (often four) nodes with replication or mirroring between the nodes for availability, with each node requiring identical approved hardware and a separate (and typically expensive) SDS software license. The resulting infrastructure can only be used to store data from workload operating under the particular vendor's hypervisor software, thereby creating an isolated island of data storage.

Such a strategy quickly becomes implausible for smaller firms that cannot afford to "rip and replace" current infrastructure or who lack the deep pockets required to fund an expensive redundant multi-nodal storage architecture. For larger firms, the cost problems are probably

Current hypervisor-vendor driven models for storage are less revolutionary than they are self-serving.

dwarfed by the problems created by isolating storage infrastructure and data in hypervisor-centric stovepipes. According to recent surveys, larger firms are diversifying their hypervisor software choices (using more than one hypervisor), and maintaining some applications in traditional physical infrastructure and operating system configurations for the sake of performance. This has the effect of multiplying the number of storage targets that need to be independently managed, administered, protected, allocated, and shared. Sharing data across hypervisor-imposed storage boundaries is already becoming a hassle.

These outcomes of SDS/H-CI do not, in any case, represent a solution to the problem of application performance that supposedly drove the development of the technology in the first place. Changing physical infrastructure topology to direct-attached storage may have reduced the number of SAN or LAN switches, the amount of cabling, and the distance that data needs to travel from the server processor to the storage device, but none of these innovations are new: all storage is direct-attached, after all. Nor is the re-centralization of storage services in a software layer on a server a new, or even particularly successful, strategy for improving management and control and reducing storage costs – especially when each hypervisor vendor's SDS model is exclusive and creates a storage stovepipe that cannot be managed in common with other vendor's stovepipes. And the addition of flash or DRAM buffering and spoofing to improve the perceived performance of application I/O is really only a band-aide with a long pedigree in storage array engineering, rather than a new or effective hedge against randomized I/O and the latencies it ultimately creates.

In short, current hypervisor-vendor driven models for storage are less revolutionary than they are self-serving. What is really needed are new storage techniques and technologies that map to contemporary workload requirements simply. It is unlikely that this requires a "revolution," but storage probably needs to do more than "evolve" incrementally – a strategy that didn't work for the dinosaurs in the long run. What is needed is a questioning of fundamentals that underpin current storage to locate and correct assumptions that no longer apply.

STORAGE FUNDAMENTALS NEED TO BE REVISITED

From the preceding analysis of the current challenges confronting application performance and storage I/O handling, the outlines of a critique of traditional storage emerge. These are just a starting point, of course, for on-going development by thought leaders in the storage industry.

File System Changes are Necessary and Overdue

Contemporary file systems use the metaphor of a filing cabinet or library card system.

If the computing model has truly changed to one of virtual machines flowing from physical server to physical server, changes are needed in the way that workload data is stored and managed. Contemporary file systems use the metaphor of a filing cabinet or library card system, placing data onto physical devices using structures such as directories, sub-directories, volumes, and logical unit numbers (LUNs). These constructs support hierarchical stores of large amounts of data, but they are not necessarily optimized for the storage of complex entities such as virtual machine data files that contain their own block structures and reference components.

Needed is a re-think of the metaphor guiding the file system to better align with the concept of an object that participates in a much flatter (as opposed to hierarchical) schema and a much more efficient use of storage capacity. Some refer to this as object storage, but it may be simply a matter of a structure that can be managed more efficiently and conveniently using simple RESTful primitives.

Queuing and buffering needs to align to workload

To the extent that SDS has yielded any improvements in application performance, they have likely resulted from the strategy of buffering writes with DRAM or flash caches. This "spoofing" enables the application to proceed to its next operation immediately while data writes to disk storage occur in the background at whatever speed the storage infrastructure can support.

This does not address the root cause of application performance slowdowns however, which often have little or nothing to do with storage I/O speeds and feeds. In fact, investigations of laggard application performance often reveal a very shallow queue depth for pending writes to storage; instead, high processor cycle rates suggest that the log jam is probably in the application code or VM itself.

When VMs are, in fact, found to be I/O-bound, this is often the fault of slow read performance in the backend storage infrastructure created not by topology but by randomized I/O and scattered data layout. There are numerous approaches to rectifying the problem that focus on reorganizing write data prior to writing it physically to disk. Most, however, entail the use of precious processing resources or require additional storage capacity.

Needed is a workload-oriented approach to write queuing that segregates I/O with reference to the VM that generates it.

Needed is a workload-oriented approach to write queuing that segregates I/O with reference to the VM that generates it. Such a strategy must enable parallel I/O transfers in a non-randomized fashion to prevent the I/O blender effect.

Leverage hybrid memory and disk for optimal performance and cost efficiency

While evangelists like to talk about the rise of the all-silicon data center, the reality is that most data centers continue to use disk (and even tape) for some or most application data storage requirements. Out of hand dismissal of pre-flash storage as "legacy" or "archaic" should signal a bright warning that the narrative has drifted from reality to hyperbole.

Clearly, there is still a role for magnetic media, even as silicon storage takes hold – and probably will be for decades to come. The ideal approach is to integrate new technologies with existing ones to leverage the strengths and minimize the weaknesses of both. Storage does not comprise one-size-fits-most technology: it is differentiated by capacity, performance and price to meet different requirements of data, which itself comes in many "flavors" of access frequency, update frequency, and preservation requirements.

Data protection needs to be built in not bolted on

With contemporary storage systems, data protection is an afterthought, a set of services that must be purchased under a separate license to go with the kit in order to select data to be replicated, to designate how data should be replicated and its replication frequency, to designate a target for replication, and to perform the actual replication at designated intervals per policy.

In many respects, this technology reflects an earlier design decision in file systems that created a self-destructive method for storing data. (Every time you save a file, you overwrite the last valid copy of the file!) The reason for this design choice had to do with the high cost per GB of data storage media. Overwriting old data with new was considered more economical than journaling or versioning the same data in different locations on the media.

There is considerable chatter in the storage world today about "unified storage."

Now with cost per GB on disk and tape falling substantially, it may be worth an effort to revisit how data is written into a file system with an eye toward building in, rather than bolting on, techniques and processes for data protection and preservation.

Unify the Entire Storage Infrastructure

There is considerable chatter in the storage world today about "unified storage" – usually connoting a particular vendor's work to make their entire product line manageable under a common graphical user interface or to consolidate the management of multiple, geographically disbursed arrays from a vendor into a single console. This is valuable, but insufficient to meet the needs of agility that are much sought after in organizations today.

For virtualization, and ultimately cloud technologies, to deliver their promised value proposition, coherent automation and management are key. Organizations need to be able to allocate and de-allocate resources to business processes and workflows quickly (the meaning of agile) in order to support changing business priorities. This is more nuanced than it seems. Recently, considerable attention has been given to elasticity, a requirement to provision resources economically, but with sufficient flexibility to handle both normal and peak workloads. Elasticity is difficult to deliver with significant and lengthy historical analysis; delivering elastic services on the fly is a greater challenge.

Virtualized storage presents a unified view of the storage resource, enabling it to be sliced and diced, allocated and de-allocated, and provisioned with specific services at will.

Part of the challenge in storage infrastructure is the heterogeneity of components and the lack of willingness of vendors to cooperate in a common active management standard. The only solution, according to some industry watchers, is to go to homogeneous infrastructure: deploy only gear from one vendor (or use commodity gear, but only in a configuration approved by a specific hypervisor vendor). While some hardware homogeneity will likely be required for other changes (such as VM-oriented I/O queuing) to be implemented successfully, there must eventually emerge another approach, such as storage virtualization.

The definition of software-defined storage advanced by current hypervisor vendors explicitly excludes capacity management and storage virtualization from the list of value-add services delivered by the SDS software layer. This is an error, as it perpetuates the inefficiencies of existing storage capacity allocation. Virtualized storage presents a unified view of the storage resource, enabling it to be sliced and diced, allocated and de-allocated, and provisioned with specific services at will. Automation of this storage pool and its management processes can enable great agility and elasticity.

PROSPECTS FOR A NEW APPROACH

The five points above, and there are probably a few more, provide a much simplified outline for strategists who are seeking to match storage infrastructure to the next generation of application workload. Thought leaders in the vendor community are already working on some or all of these issues, but results are mixed. In many cases, consumers need to make hard choices between "safe" incremental changes and "riskier" but more innovative changes that can platform data for improved performance and availability.

Some "fixes" have been advanced in file systems, including file system level de-duplication, journaling options, log structuring, etc. But these are still, in most cases, very much in development and tend to reduce overall file system performance, which reduces the enthusiasm of stake-holders.

Storage virtualization has moved us closer to unified management and

Tintri's management interface makes the management of both the storage resource and storage services such as data protection very easy, even for storage novices.

administration, but its adoption is hampered in part by hardware vendors who wish to preserve specious differentiation in otherwise commodity component based products, and by hypervisor software vendors who are advancing their own view of proprietary SDS intended to create storage stovepipes around their technology.

Object storage is posited as a future model for storage, but adoption is hampered by a requirement to support older infrastructure designs. Cloud storage giants such as Amazon Web Services may help to encourage object storage in the future, but the investment in traditional storage hardware and file systems and the ubiquity of a file system/directory/volume/LUN metaphor will be a challenge to budge.

For all of these challenges, some vendors are developing innovative solutions that are worth a closer look by IT planners. One such innovator is Tintri.

Based in Mountain View, CA, Tintri is listed as a flash storage company, creating silicon storage products for use in virtualized server environments. However, the real value of Tintri isn't necessarily found in a storage kit. Tintri is quietly advancing a rather revolutionary model for storage itself that maps VMs directly to storage and the "unit of storage."

The company's familiarity and comfort with traditional file systems makes their VM-centric approach to data storage appear to be non-disruptive to the file system-oriented consumer. However, by introducing the VM as the new organizing principle of storage, Tintri has been able to develop a very workable strategy of workload-centric queue management that leverages NVRAM and flash based storage very efficiently. Moreover, Tinti appears to have perfected an automated, access frequency-driven method for data tiering that improves not only capacity allocation efficiency, but capacity utilization efficiency in storage overall.

Tintri's management interface makes the management of both the storage resource and storage services such as data protection very easy, even for storage novices. On-going monitoring, even in frequently scaling and broadly distributed storage assets, is dramatically improved with the Tintri solution.

A clear thought leader in the storage "revolution," Tintri needs to be given a close look by IT innovators in enterprise and cloud computing data centers. They are proof that more is needed than incremental evolution to ensure that next generation storage infrastructure doesn't go the way of Jurassic World.

Jon Toigo is a 30-year veteran of IT, and the Managing Partner of Toigo Partners International, an IT industry watchdog and consumer advocacy. He is also the chairman of the Data Management Institute, which focuses on the development of data management as a professional discipline. Toigo has written 15 books on business and IT and published more than 3,000 articles in the technology trade press. He is currently working on several book projects, including The Infrastruggle (for which this blog is named) which he is developing as a blook.