Tintri VMstore: Zero Management Storage for Virtualization and Cloud

Tintri VMstore is built on the industry's first and leading VM-Aware Storage architecture. This technical white paper covers the architecture and functionality underlying the Tintri VMstore including the patented Tintri FlashFirst™ design and VM-Aware operations. The paper also covers what makes Tintri VMstore the ideal storage foundation for scaling virtualization.
Executive Summary

Tintri VMstore provides Zero Management Storage™ for virtualization and cloud. Tintri VMstore was built from the ground up based on the industry’s first and leading VM-Aware Storage architecture to deliver always-optimized storage for dynamic virtual environments. Tintri VMstore was created with virtual machines (VMs) in mind. It uses and understands VMs and virtual disks in place of such traditional storage abstractions as volumes and LUNs.

Storage specifically designed for virtualization and cloud gives IT the opportunity to broaden server and desktop virtualization deployments while enhancing efficiency and cost savings. Storage designed for virtualization addresses the performance and efficiency mismatch with traditional storage and virtualization, and eliminates many mundane tasks, allowing IT to focus on innovation.

VMstore is a hybrid storage solution with a combination of flash-based SSDs and high-capacity disk drives with a FlashFirst™ design, which services over 99 percent IO from flash to provide very high levels of throughput and consistent sub-millisecond latencies. VMstore delivers VM-level QoS to guarantee every VM gets the performance it needs without any manual tuning or configuration, as well as automatic VM alignment.

Tintri VMstore monitors every IO request at the vdisk and VM level and can determine if latency is incurred at the hypervisor, network or storage levels. For each VM and vdisk stored on the system, enterprise IT teams can use Tintri VMstore to instantly visualize where potential performance bottlenecks may exist, including on the host, network or storage, rather than trying to deduce where they are based on indirect measurements and time-consuming detective work.

Unique in the industry, Tintri VMstore allows creation of snapshots, clones and remote replicas at the VM-level with ease. This granularity means complete control for data management of large-scale virtual environments down to the vdisk level.

Scaling storage for virtualization is simplified with VMstore. Additional VMstore systems can be added to a virtualized environment as additional datastores, leveraging a building block approach to storage. As many as 32 geographically distributed VMstore systems can be centrally monitored and controlled using Tintri Global Center.

This white paper reviews the need for storage in a virtual environment and details the capabilities of Tintri VMstore that make it the industry’s leading storage solution for virtualization and cloud.

Drivers for Change in Virtualized Data Centers

Enterprise IT is under unremitting pressure to reduce capital and operating expense, driving them to virtualize infrastructure to improve hardware utilization and scalability, and advance toward enhanced operational efficiency and flexibility. While virtualization has dramatically improved utilization of resources in the data center, the complexity of virtualized infrastructure gives rise to concerns around management overhead, application performance, data protection and availability and cost. Enterprises need future-proof IT solutions to help maximize the returns from their virtual server environments while advancing toward 100 percent data center virtualization.

The complexity of configuring, managing and tuning traditional storage for VMs is costly and, ultimately, must be addressed before broad-based virtualization can be deployed. For example, traditional shared storage systems manage objects such as LUNs, volumes or tiers, which have no intrinsic meaning for VMs. Each new VM instance must be assigned a specific storage LUN or volume. When IO requirements and VM behavior are poorly understood, a painful trial-and-error process ensues. Storage and VM administrators must coordinate to ensure each application not only has the space it needs, but also enough IO performance for the expected load.
Optimum application performance requires solutions for the unique performance challenges virtualization presents. The latest generation of servers can easily support upwards of tens of virtual servers, each of which can generate its own IO stream. As a result, the IO patterns that virtual environments generate are far more random than those generated by applications running on bare-metal servers. This ‘IO blender’ effect translates into strong degradation of performance on traditional storage, which results in IT deferring virtualizing IO-intensive tier-one applications, and manually isolating workloads. Application performance guarantees are a concern if an enterprise wants to run tier-one applications using a virtualized infrastructure, since many applications that are built using shared, virtualized infrastructure tend to present a noisy neighbor problem, where an application’s performance can suffer if another application has a spike in demand and draws on shared resources.

Scaling storage adds another dimension of complexity on the performance and planning front, limiting growth. Unlike how servers are handled in virtualized environments, traditional architectures make scaling storage to be a piecemeal exercise or require complex scale-out storage configurations.

Increasing proliferation of virtual servers also entails virtualization-specific approaches for data protection and availability. Traditional backup and recovery processes introduce bottlenecks that limit the number of VMs that can exist on a physical host. Specifically, backing-up VMs directly using traditional backup agents impacts application performance and bottlenecks the physical servers’ shared CPU, disk, RAM and network resources. In addition, legacy backup targets require a lot more storage to protect a virtual environment as the vdisk (VMDK) images and application files are typically highly redundant. The frequency of the backups or recovery point objectives (RPOs) and time to recovery or recovery time objectives (RTOs) may also be adversely affected, which can jeopardize productivity and compliance with service level agreements (SLAs).

Virtualization-driven storage growth and complexity increase capital as well as operating expenses related to disparate tools and manual processes. Added constraints imposed by flat IT budgets make it increasingly difficult to manage and scale virtual infrastructures. Expanded virtualization requires approaches that maximize ROI while reinforcing IT’s role in enabling new opportunities and driving total cost of ownership (TCO) reduction.

Storage Requirements for Virtualized Environments

Performance, storage and data management requirements were easy to meet when distinct applications ran on separate physical servers and utilized dedicated storage. But new requirements are created when these goals are carried over into large-scale virtualized environments supporting high densities of virtualized applications with varying workloads and shared storage systems.

Virtualization combines applications and workloads to create extremely random IO, resulting in performance requirements that choke traditional storage. This storage bottleneck and complexity of scaling storage beyond a single system creates inefficiencies that limit VM densities, precludes virtualizing IO-intensive applications, such as databases and critical business applications and adds unnecessary management overhead. These

Figure 2: Storage should understand and operate at the VM level.
bottlenecks and scalability challenges must be addressed to expand virtualization across enterprise data centers with business-critical applications. Large-scale server and desktop virtualization also requires storage systems that remove the complexities associated with outmoded physical constructs such as LUNS and volumes by operating at the VM and virtual disk level.

Enterprises need storage that understands and operates at the VM level, analyzes the IO patterns of a virtual environment and dynamically manages QoS for each VM. Storage should also automatically align vdisks and VMs for optimum IO performance and capacity utilization — without manual intervention.

Storage systems for virtual environments must provide monitoring and troubleshooting capabilities at the virtualization level, which means visibility into the operation of each VM, including performance and capacity monitoring.

Last, but not the least, storage should deliver capabilities to apply policies for data management operations, such as snapshots, replication and cloning, for individual VMs vs. at the LUN and volume level. The traditional storage constructs lead to inefficient utilization, administrative complexity and suboptimal performance and service levels. Operating at a granular VM level is inherently more efficient and predictable.

**Tintri VMstore for Virtualization and Cloud**

Tintri VMstore is Zero Management Storage™ for virtualization and cloud. Tintri VMstore was built from the ground up based on the industry’s first and leading VM-Aware Storage architecture to deliver always-optimized storage for dynamic virtual environments. It has the storage intelligence for sub-millisecond latencies, VM-level QoS optimization, and deep visibility into the operational characteristics of each VM. This means highly effective monitoring and troubleshooting, and VM-granularity data management such as instantaneous snapshots, zero-space cloning and efficient replication. The result is higher predictability, unparalleled performance and efficiency and higher levels of productivity.

The functionality underlying Tintri Zero Management Storage™ can be categorized in three sections:

1. **Storage intelligence**: Delivering performance, density and scalability without the complexity.
2. **Infrastructure insight**: Delivering a complete picture of virtualized workloads.
3. **VM control**: Delivering VM-granular data management and automation.

The following sections explore each category in detail.
Storage Intelligence

Tintri VMstore’s approach automatically ensures every VM gets the performance it needs. Expanding storage is simple, as each VMstore appliance appears as an additional, high-capacity datastore in VMware vCenter. This makes it easy to scale and manage each node as part of a VMware Storage DRS cluster and eliminates any risk of downtime.

Let’s explore how:

**Tintri FlashFirst™ design:** VMstore is a hybrid storage solution. It uses a combination of flash-based solid-state devices (SSDs) and high-capacity disk drives for storage. Tintri’s patented FlashFirst design incorporates algorithms for inline deduplication, compression and working set analysis to service more than 99 percent of all IO from flash (Figure 5) for very high levels of throughput and consistent sub-millisecond latencies for both read and write operations.

Flash-first design minimizes swap between SSD and HDD by leveraging data reduction in the form of deduplication and compression to increase the amount of data that can be stored on flash. This is complemented by detailed profiling of all the active VM IO to ensure metadata and active data are kept in high performance flash. Only cold data is evicted to disk, which does not impact application performance. It takes advantage of the fact that each VM has an active working set, which is a fraction of the overall VM. Using a flash-only approach means all data must be stored on high performance (and expensive) flash, whether it needs to be there or not.

Unlike flash-only products, 100 percent of the operational flash capacity on a Tintri VMstore can be used without concern about running out of space and having applications come to a screeching halt. In addition, the Tintri VMstore is operationally far simpler and more cost-effective than flash-only products.
Traditional storage systems often incorporate flash to an existing disk-based architecture, using it as a cache or bolt-on tier, while continuing to use disk IO as part of the basic data path. In comparison, VMstore services 99 percent IO requests directly from flash, thereby achieving dramatically lower flash-level latencies, while delivering the cost advantages of disk storage.

**Flash for Enterprise Storage Applications**

Tintri VMstore integrates flash as a first-class storage medium rather than as a bolt-on cache or tier to fully leverage continued improvement in flash price and performance. Flash as an intelligent, highly granular resource — combined with inline deduplication, compression and a unique flash/disk file system — enables Tintri VMstore to radically alter the economics for server virtualization.

Tintri’s innovative FlashFirst design addresses MLC flash problems that previously made it unsuitable for enterprise environments: Flash suffers from high levels of write amplification due to the asymmetry between the size of blocks being written and the size of erasure blocks for flash. Unchecked, this reduces random write throughput by more than 100 times, introduces latency spikes and dramatically reduces flash lifetime. FlashFirst design uses a variety of techniques including deduplication, compression, analysis of IO, wear leveling, garbage collection algorithms and SMART (Self-Monitoring, Analysis and Reporting Technology) monitoring of flash devices and dual parity RAID-6 to handle write amplification, ensure longevity and safeguard against failures.

**VM QoS:** Tintri VMstore is designed to support a mixed workload of hundreds of VMs each with a unique IO configuration. Additionally, as volumes of traffic ebb and flow, VMstore through its FlashFirst design analyzes and tracks the IO for each VM, delivering consistent performance where it is needed. This enables VMstore to isolate the VMs, queue and allocate critical system resources such as networking, flash/SSDs and system processing to individual VMs. Tintri VMstore’s QoS capability is complementary to VMware’s performance management capability. The result is consistent performance where it is needed. And, all of VM QoS functionality is transparent, so there is no need to manually tune the array or perform any administrative touch.

QoS is critical when storage must support high performance databases generating plenty of IO alongside latency sensitive virtual desktops. This is commonly referred to as the noisy neighbor problem in traditional storage architectures that are flash-only and lack VM-granular QoS. Tintri VMstore ensures database IO does not starve the virtual desktops, making it possible to have thousands of VMs served from the same storage system.

**VM auto-alignment:** The issue of VM alignment poses real challenges as virtualization deployment expands across enterprise data centers. Misaligned VMs magnify IO requests, consuming extra IOPS on the storage array. The impact snowballs as the environment grows with a single array supporting hundreds of VMs. At this size, performance impact estimates range from 10 percent to more than 30 percent.

Every VM writes data to disk in logical chunks. Storage arrays also represent data in logical blocks. When a VM is created, the block boundaries of the VM and storage do not always align automatically. If the blocks are not aligned, VM requests span two storage blocks, requiring an additional IO operation (Figure 6).
Storage administrators in virtualized data centers attempt to address this issue by aligning VMs to try to reduce the impact of misalignment on performance. Unfortunately, realigning a VM is a manual, iterative process that generally requires downtime.

VMstore offers VM auto-alignment. Rather than the disruptive approach of realigning each VM, Tintri VMstore dynamically adapts to the VM layout (Figure 7). Tintri VMstore automatically aligns all VMs as they are created, migrated or cloned — with zero downtime. An IT administrator can now eliminate this arcane task and enjoy performance gains with no VM downtime, and zero administrator intervention.

**Scaling storage:** Scaling storage beyond the performance and capacity of a single system is as simple as adding another VMstore system – a task that takes less than two minutes. This building block approach effectively adds another datastore that can be managed by the virtualization layer. To tackle the challenge of managing individual VMstore systems, Tintri has created Tintri Global Center. Build on a solid architectural foundation is capable of supporting over one million VMs, Tintri Global Center is an intuitive centralized control platform that lets administrators monitor and administer multiple VMstore systems as one.

**Infrastructure Insight**

Traditional storage systems provide a performance view from the LUN, volume or file-system standpoint. But traditional storage cannot isolate VM performance or provide insight into VM-level performance characteristics. It is difficult for administrators to understand situations such as the impact of a new VM workload, without access to relevant VM performance metrics. In addition, identifying the cause of performance bottlenecks is a time consuming, frustrating and sometimes inconclusive process that requires iteratively gathering data, analyzing the data to form a hypothesis and then testing the hypothesis. In large enterprises, this process often involves coordination between several people and departments, typically spanning many days or even weeks.

Tintri provides a complete, comprehensive view of VMs including end-to-end tracking and visualization of performance across the entire data center infrastructure. This ensures that administrators can procure the critical statistics they need. The goal is ensuring storage performance stays at acceptable levels with minimal latency.

Tintri VMstore monitors every IO request at the vdisk and VM level and can determine if latency is being incurred at the hypervisor, network, or storage levels. For each VM and vdisk stored on the system, enterprise IT teams can use Tintri VMstore to instantly visualize where potential performance issues may exist, including on the host, network or storage. The latency statistics are displayed in an intuitive format (Figure 8). In an instant, administrators can see the bottleneck rather than trying to deduce where it is from indirect measurements and time-consuming detective work.

The hypervisor latencies are obtained using vCenter APIs, while the network, file system and disk latencies are provided by Tintri VMstore, which knows the identity of the corresponding VM for each IO request.
Administrators can detect trends with this data from the VMstore and individual VMs, all without the added complexity of installing and maintaining separate software. This built-in insight can reduce costs and simplify planning activities — especially around virtualizing IO-intensive critical applications and end-user desktops.

Tintri Global Center provides administrators with deep insight across multiple VMstore systems and their resident VMs from a single centralized platform. IT administrators can view and create summary reports across all VMstore systems — with in-depth information on storage performance (IOPS, latency, throughput), capacity, vCenter clusters, host status, protection status, and more. In addition to summary information presented at a glance, Tintri Global Center also provides the ability to filter and display results, including by individual VMstore systems and specific VMs, for easy troubleshooting.

**VM Control**

Tintri VMstore allows all data management operations — snapshots, clones and replication — to operate at the VM level. This enables managing large-scale virtual environments to the vdisk level with complete, granular control.

**Tintri SnapVM:** Snapshots preserve the state and data of a VM at a specific point in time allowing VMs to be easily rolled back or replicated. However, traditional storage architectures provide snapshots of storage objects, such as LUNs and volumes, rather than VMs. These snapshots can lead to inefficient storage utilization as tens-to-hundreds of VMs with varying change rates are often snapshotted at once. Snapshot schedules can only be set at a LUN or a volume level, leading to such practices as creating one LUN per VM as a workaround to create individualized snapshot VM schedules. Snapshots based on a copy-on-first-write design also commonly provide suboptimal read and write performance, making them infeasible for performance-oriented tier-1 applications.

Tintri VMstore delivers unique, space-efficient snapshot capability that consumes virtually zero disk space and can restore VMs within minutes or even seconds. In addition, granular VM snapshots allow administrators to create snapshots of individual VMs and quickly recover data or entire VMs from snapshots. Tintri VMstore supports 128 snapshots per VM for longer-term retention. Data protection management is also simplified with the use of default or custom schedules for VM-consistent or crash-consistent snapshots that protect individual VM automatically without administrator intervention.
Crash consistent snapshots do not take extra measures with the hypervisor or guest VM to coordinate snapshots. Thanks to integration with native hypervisor management tools, such as VMware vCenter integration, Tintri provides VM-consistent snapshots for simpler application recovery. With VM-consistent snapshots, hypervisor management APIs are invoked to quiesce the application in a VM for a VM-consistent snapshot.

Unlike storage-centric snapshot technologies in traditional shared storage systems, Tintri SnapVM™ makes recovery workflows remarkably easy. Files from individual VMs can be recovered without additional management overhead, dramatically reducing the time to recovery.

**Tintri CloneVM**: A few traditional storage systems can provide cloning capabilities sharing data blocks between the source/parent and the clone. Unfortunately these clones are done at the LUN or volume level, which can vastly complicate VM deployment, cloning and management operations. Tintri CloneVM™ enables space-efficient cloning operations at the individual VM level. This eliminates the limitations of traditional storage architectures that necessitate complex provisioning and management.

VMstore builds on snapshots to support individual VM cloning capability, either by taking a new snapshot or by cloning an existing snapshot. Hundreds of clones can be created virtually in an instant, all of which are space-efficient and full-performance. Cloned VMs can be quickly accessed, powered on and put into service, enabling more efficient use cases such as virtual desktop infrastructure (VDI), development and test, business intelligence and database testing.

New VMs created via cloning exist and function independently from the parent VMs from which they are created. Behind the scenes, new VMs share common vdisk references with parent VM snapshots to maximize space and performance efficiencies. The extent to which they individually grow and diverge from the data they share with their respective parents defines their incremental storage space requirements. Tintri’s patented use of flash assures that clones are 100 percent performance-efficient. They get the same level of performance as any other VM stored on a Tintri VMstore system.

Using the Tintri UI, hundreds of clone VMs can be created at a time (Figure 9). The cloned VMs are dynamically registered and visible to the hypervisor for immediate use. Administrators can also select customizable specifications defined in vCenter for preparing newly created clone VMs using the vCenter sysprep functionality. Further, clones can also be created from golden image VMs for use cases such as test and development and VDI.
Tintri has demonstrated deep integration between VMware Horizon View and CloneVM functionality in a reference architecture validated by ESG Labs. With support for all VAAI-NFS primitives and View Composer Array Integration, administrators can create clones from vSphere and Horizon View in seconds leveraging native CloneVM functionality.

**Tintri ReplicateVM:** Tintri ReplicateVM™ capability supports efficient replication of VMs from a primary to a second VMstore. Tintri ReplicateVM is based on Tintri snapshot technology, allowing either a new or existing snapshot to be replicated automatically.

Like SnapVM, ReplicateVM enables administrators to apply protection policies to individual VMs, rather than to arbitrary units of storage such as volumes or LUNs. It allows administrators to easily establish, as-needed, a snapshot and replication policy for individual or set of VMs.
Tintri has demonstrated deep integration between VMware Horizon View and CloneVM functionality in a reference architecture validated by ESG Labs. With support for all VAAI-NFS primitives and View Composer Array Integration, administrators can create clones from vSphere and Horizon View in seconds leveraging native CloneVM functionality.

Tintri ReplicateVM:

Tintri ReplicateVM™ capability supports efficient replication of VMs from a primary to a second VMstore. Tintri ReplicateVM is based on Tintri snapshot technology, allowing either a new or existing snapshot to be replicated automatically.

Like SnapVM, ReplicateVM enables administrators to apply protection policies to individual VMs, rather than to arbitrary units of storage such as volumes or LUNs. It allows administrators to easily establish, as-needed, a snapshot and replication policy for individual or set of VMs.

ReplicateVM works by replicating deduplicated and compressed snapshots of VMs from one Tintri VMstore to another, only sending across the network actual changed blocks or missing data. As a result, VM replication is highly WAN-efficient with up to 95 percent reduction in bandwidth utilization. It also enables remote cloning, making distribution of golden images for workloads such as VDI with multisite high availability (HA) efficient and simple. ReplicateVM supports different topologies including one-to-one, many-to-one and bi-directional replication.

ReplicateVM works by replicating deduplicated and compressed snapshots of VMs from one Tintri VMstore to another, only sending across the network actual changed blocks or missing data. As a result, VM replication is highly WAN-efficient with up to 95 percent reduction in bandwidth utilization. It also enables remote cloning, making distribution of golden images for workloads such as VDI with multisite high availability (HA) efficient and simple. ReplicateVM supports different topologies including one-to-one, many-to-one and bi-directional replication.

Figure 10: Highly available VDI deployment spanning two sites.

Figure 11: Local and remote cloning simplifies data protection and scaling.
Using ReplicateVM, replication can be dedicated to specific network interfaces, and optionally throttled to limit the rate of replication when replicating snapshots between Tintri VMstore appliances located in datacenters connected over wide-area networks (WANs).

**Enterprise-Class High Availability**

Tintri VMstore was designed for enterprise-class availability. Starting with storage, VMstore uses dual parity RAID-6, real-time correction and continuous verification. This ensures data integrity and delivers higher availability than RAID-10 without the inefficiencies. On the hardware front, in addition to using redundant components and connectivity, VMstore uses a dual controller active/standby setup. Each of these controllers can be non-disruptively upgraded ensuring all VMs supported are up and running across software updates.

VMstore also delivers proactive support with phone-home monitoring and reporting. Administrators get automated notifications and alerts related to areas such as capacity/planning, network connectivity and predictive disk health, allowing them to get ahead of any issues. Administrators can also trigger creation of a support bundle on-demand. This will perform a diagnostic check of the system for additional peace of mind.

**Conclusion: Tintri VMstore Storage**

Tintri Zero Management Storage™ helps IT organizations eliminate storage complexity and minimize costs for virtualized environments by addressing the mismatch between storage and virtualization. Tintri VMstore's storage intelligence, infrastructure insight and VM control delivers unparalleled performance and efficiency, making deployment and operation of virtualization completely predictable, and enabling higher IT productivity. Here is a summary of how IT can benefit from Tintri features:

**Performance and Efficiency**

- Tintri FlashFirst design delivers as much as 10 times better performance and VM densities with sub-millisecond read and write latencies for virtualized workloads without any tuning or over-provisioning vs. traditional storage.
- VM-level QoS and performance allocation enables a single VMstore to support hundreds of different virtual workloads such as VDI and tier-1 applications, eliminating the noisy neighbor problem.
- Auto-alignment eliminates the need for disruptively realigning individual VMs to achieve optimal performance.
- CloneVM allows IT to create hundreds of instant full-performance clones through either Tintri or virtualization tools without consuming additional space or affecting storage performance.
- Scale by using VMstore as storage building blocks managed by the virtualization layer.

**Making Virtualization Predictable**

- FlashFirst design and VM-level QoS delivers predictable performance for critical VMs, even in dynamic environments. Infrastructure insight delivers an end-to-end performance view for individual VMs enabling IT to detect trends and identify performance bottlenecks in minutes without adding additional software or agents.
- SnapVM and ReplicateVM let IT set local and replication data protection policies for individual VMs — all without requiring backup windows, manual settings or additional capacity and bandwidth typically needed with traditional solutions.
- Enterprise-level HA including redundant components and connectivity, non-disruptive upgrades and proactive support eliminate most cases of planned and unplanned downtime.

**Enable Extra Productivity**

- VMstore's simplified design and setup allows IT to get up and running in as few as two minutes.
- Understanding and operating at the VM level eliminates the need to setup disk/RAID groups, LUNs and volumes and the need to maintain mapping between storage and VMs.
- Monitor and control multiple VMstore systems as one without the complexity of scale-out storage with Tintri Global Center.
- Infrastructure insight shaves days off troubleshooting exercises by allowing IT to pinpoint problems before they escalate.
- VMware integration enables admins to create hundreds of VM clones without having to deal with underlying storage.