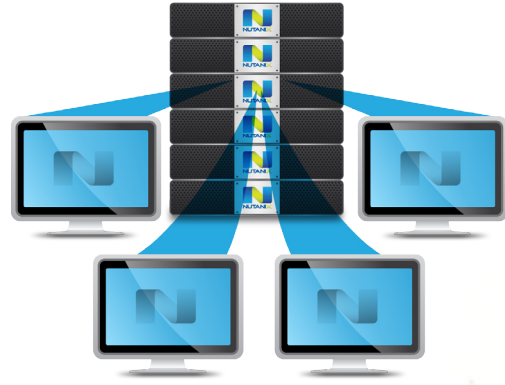


# NUTANIX Solution Brief: Desktop Virtualization



Nutanix Complete Cluster delivers a turnkey solution for desktop virtualization that won it the Best of vmworld 2011 Award for this category. In this solution brief, we'll cover the key technical points that make it an ideal solution for your desktop virtualization needs.

When considering a VDI deployment, the two key areas of focus are performance and cost. If end user acceptance does not happen, a VDI pilot may never make it to production. If the per virtual desktop cost cannot be rationalized with a believable ROI, similarly, the justification to deploy at scale may be blocked in favor of status quo. To deliver a superior user experience with predictable scale-out, the unique converged compute + storage building block from Nutanix offers a cost-effective choice unlike anything else on the market today.



"This is exactly what the VDI Industry has been waiting for..."  
- Brian Madden and Gabe Knuth, Best of vmworld Judges

## Executive Summary

### Scalable High-Performance VDI

- Out-of-the-box appliance experience: get VDI going in less than 30 min without having to deal with the headache of expensive network storage (SAN/NAS).
- High I/O per desktop by leveraging Fusion-io ioMemory for a great user experience, from task workers to power users
- Scalability from hundreds to thousands of virtual desktops so you can grow your VDI deployment easily as your needs grow

### TCO Savings with Grow-As-You-Go Building Block

- Elimination of expensive network storage and a compact building block yield 40-60% CAPEX savings over existing traditional solutions
- Simplicity of SAN-less architecture drives OPEX savings to help realize a faster ROI on desktop management costs
- Fast provisioning of virtual desktops using Nutanix FastClone enables flexibility to respond quickly to changing business requirements

## Analyzing the Performance Challenges of VDI

**Random Steady State IO of Virtual Desktops:** In steady state, desktop users open applications, send email, browse the web etc. Each of these translates to a small size IO request to the storage layer. Since each desktop is independent of the other desktops, network storage is not able to tell one desktop from the other. Therefore, all IO traffic coming into a SAN at steady state is completely random, each requiring a separate spindle head movement. With an average of 20 IOPS required per average knowledge worker desktop, the total random IO required of a SAN becomes 20,000 IOPS. This translates to 300 spindle disks without accounting for RAID. With RAID 5 or 6, the number of disks required is 600-800 just to support steady state random IO coming from these 1,000 virtual desktops.

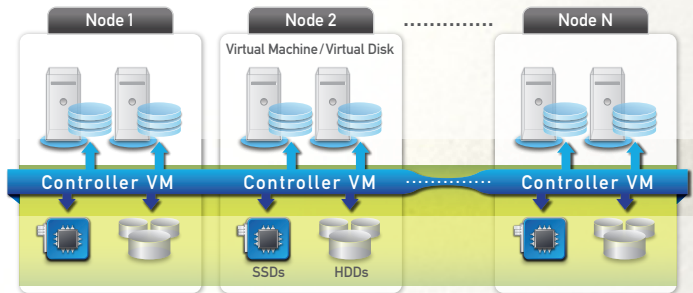
**Desktop Boot Storms:** Booting a virtual desktop requires that the key OS bits be loaded from the SAN separately for each desktop. There is no simple way for a SAN to load the data more intelligently, since this intelligence has to be at an upper layer. Booting 1,000 users translates to 300,000 IOPS, which means for many deployments, overprovisioning storage is necessary to meet the performance requirements, driving the TCO up significantly. Third party solutions at an upper layer can help address these issues, but will add additional cost to the overall solution, stressing the TCO once again.

**Eliminating Boot Storms:** Each controller VM in the Nutanix architecture functions similarly to a traditional storage controller, except that there is one for every node instead of a limited number shared by a large SAN that may result in bottlenecks. The controller VMs each have a cache called an "Extent Cache" that caches the data that has been served up by the controller. Frequently accessed data continues to live in the cache. This means that once cached, the OS bits are served from the cache, eliminating the need for disk/flash IO and thus eliminating the boot storm issue.

By rethinking how infrastructure should be built for virtualization, Nutanix's approach inherently solves the two biggest performance pain points for VDI. Nutanix delivers fast random IO as well as high sequential bandwidth providing desktop users with a great user experience in steady state and in the face of boot storms.

## How Does Nutanix Tackle VDI Performance Challenges

**Fast Random IO:** All write IO in the Nutanix Complete Cluster goes to the HOT (Heat Optimized Tiering) Cache first. This cache data is written to Fusion-io ioDrive that is on each node in the Cluster, and immediately returns to the OS. The Nutanix Scale-Out Converged Storage layer that stitches the local storage from each node in the cluster into one global fabric then provides persistent storage in the form of either the DiskStore (direct-attached SATA HDDs) or the FlashStore (PCI-e attached Fusion-io ioMemory) depending on the tiering policies in place. The fact that the IO is written to the Fusion-io ioDrive first means that there is no spindle disk in the picture here. Leveraging Fusion-io ioMemory translates directly into microsecond latencies and high throughput.



Nutanix Architecture Enables High Performance with n-Controllers



## Enabling Grow-As-You-Go For Investment Protection

**Scalability Challenges:** It's commonly known that many VDI deployments fail when trying to scale beyond 200 desktops.

Let's look at some of the reasons why VDI deployments fail to scale when using the standard server and SAN/NAS approach:

Many VDI deployments do not have a dedicated storage array. Typically, the SAN is a common resource to all virtualization initiatives in an organization. This can create issues because SAN resources are shared by VDI workloads as well as server workloads, which have different performance profile requirements. There is no easy way to guarantee any kind of QoS for the initial VDI deployment, let alone scale the VDI deployment effectively.

If a VDI deployment does indeed get a budget to have its own SAN, the SAN can sometimes be under-provisioned because it is based on the initial number of planned desktops. When the VDI deployment needs to scale, the current SAN performance and capacity is no longer sufficient, and another storage array needs to be purchased and provisioning. Adding a storage array not only makes drives the capital expenses up dramatically, but also increases the management complexity of the overall VDI deployment.

In the most common case, SAN resources are over-provisioned. The idea here is that the SAN is a one-time purchase, and the organization can add servers to the fabric as the deployment scales. There are two common resulting problems with this approach. One is that the organization is forced to take a very large upfront capital expenditure, which then drives up the overall cost of the project and the time it takes to realize the ROI. The second issue is that the interconnect becomes the bottleneck. Typically the array is on the other side

of the core switch. Since this core switch is often heavily oversubscribed, the interconnect between the servers and the storage becomes a huge bottleneck quickly.

Nutanix takes a completely different approach to delivering incremental scalability by keeping the compute right next to the storage in a compact building block that can easily scale-out. A single Nutanix block houses four nodes, and organizations can add single nodes to their Nutanix cluster to continue to grow their VDI by 50-100 desktops at a time depending on the user profile (task/knowledge worker/power user).



## Delivering Incremental Scalability for Grow-As-You-Go Investment Protection

Some of the key attributes of the Nutanix architecture that deliver incremental scalability are described below:

**Shared Nothing Distributed Architecture:** The Nutanix Complete Cluster is a pure distributed system. This means that the compute gets its storage locally, and does not need to traverse the network. All the problems around network, interconnect and core switch bottlenecks are avoided completely. IT organizations can add single Nutanix nodes or blocks of four nodes, as needed to achieve linear scaling in performance and capacity.

**Distributed Metadata :** A big problem affecting scalability in a traditional SAN is metadata access. Most storage arrays have 1-4 controller heads and all metadata access needs to go through these heads. This causes contention and performance drops as more servers try to access the same storage array.

In the Nutanix Cluster, metadata is maintained in a truly distributed and decentralized fashion. Each of the nodes in the Nutanix Cluster maintains a part of the global metadata, which means that there is no single bottleneck for metadata access.

**Distributed Metadata Cache :** Most storage arrays do a very poor job of maintaining metadata caches. In storage arrays that do have a metadata cache, the caches live on the limited number of controller heads. Access to the cache is limited by the bottlenecks around the network, interconnect and switch as discussed above. In the Nutanix Cluster, metadata is cached on each of the controller VMs. Most metadata access is served up by cache lookups. Each controller VM maintains its own cache. This means that however large the Nutanix Cluster grows as you continue to add nodes, the performance cost of metadata access stays the same.

**Lock-free Concurrency Model :** The standard approach to ensuring correctness for metadata access is to use locking. Unfortunately, in a distributed system, locking can become very complicated. Excessive locking can better ensure

correctness, but creates significant performance drag. The Nutanix Cluster implements an innovative lock free concurrency model for metadata access. This model ensures the correctness of metadata but at the same time ensures high performance.

**Distributed MapReduce for data/metadata consistency :** For a large scale deployment, consistency checking for data and metadata becomes a challenge. The Nutanix Complete Cluster implements a fully distributed MapReduce algorithm to ensure data and metadata consistency. The distributed nature of the MapReduce implementation ensures that there is no single bottleneck in the system. MapReduce has been shown by Google to scale to 1000s of nodes, and is a key ingredient in the incremental scalability of the Nutanix Cluster

**Distributed Extent Cache :** Caching is not a new concept to storage arrays. The challenge, however is that caches are located on the limited number of storage controllers in a storage array. Not only does the limited of storage controllers cause contention, but the fact that these caches live in the storage array means that access to cached data needs to traverse the core switch. This brings with it latencies around the network, interconnect and switch that was discussed above. The Extent Cache discussed earlier caches the data served up by the controller VM from local storage. The compute tier can easier access the cache locally without having to hop across the network and the core switch. This approach helps enable incremental scale while maintaining high performance.

**Conclusion – Delivering the Best Desktop Virtualization Infrastructure Solution**  
Nutanix Complete Cluster addresses the key performance and scalability challenges for desktop virtualization while delivering a dramatically lower TCO. Its innovative converged compute and storage building block provides a radically simple approach to virtualizing desktops so you can go from proof-of-concept to production deployment much more rapidly than traditional solutions.

Ready to get started? Email [sales@nutanix.com](mailto:sales@nutanix.com) to find out how to get started today.