How A V3 Appliance Employs Superior VDI Architecture to Reduce Latency and Increase Performance

Contents

Overview

Virtual Desktop Infrastructure (VDI), which allows an organization to virtualize its desktops, is playing an ever-increasing role in IT desktop management. The benefits of virtualized desktop computing, including lowered costs due to server consolidation, centralized management and high availability, appeal to many use cases.

Virtual desktops as a secondary system have been a staple for many organizations for many years. However, few organizations have utilized virtual desktops as a primary system due to their current drawbacks, especially latency that results in slow performance and ultimately a poor user experience. Overall, VDI has not delivered on all the promised benefits to organizations, let alone allowed them to completely replace their physical desktops with virtual desktops.

Nonetheless, today there are many drivers pointing toward increased adoption of VDI as a primary system, including the increasing costs of productivity, more experienced mobile knowledge workers, and the need for IT organizations to balance control with flexibility.

The purpose of this whitepaper is to educate organizations on how VDI – when done correctly – can outperform and provide a better end user experience than a physical desktop or laptop.

For desktop replacement virtualization to be successful, the first critical item to be addressed is performance -- users need to have an optimal experience. Next is simplicity in deployment, along with high efficiencies in management, power, and cooling. The V3 appliance approach to desktop virtualization delivers on all these fronts.

Introduction

At the top of every IT manager's list of initiatives are the same goals: simplifying IT management, decreasing cost, and increasing control. VDI promises to deliver on all these fronts, yet many IT managers approach VDI with hesitation because of slow performance and poor end user experiences.

With millions of dollars lost every year due to misplaced laptops, along with the high energy cost of powering and cooling computers, centralized solutions like VDI should be a viable solution. The promises of central management, lower operations expenses, greater security, efficient application deployment, and computing availability are already here today, but cannot or will not be realized until they also provide an optimal end user experience.

In order for Desktops as a Service (DaaS) to be successful, in private or public clouds, DaaS must be able to match, if not outperform, physical desktops and laptops in the areas of both performance and optimal end user experience. This paper will explore the many bottlenecks that plague existing VDI architecture, discuss the requirements for successful VDI, and explain how a V3 appliance solution overcomes current challenges to deliver the promise of VDI.

Understanding Latency

Latency is the number one problem plaguing VDI today, causing performance degradation, and resulting in poor end user experiences. By definition, latency is a measure of the time delay experienced within a system or group of systems. The precise definition of latency is context-sensitive, meaning it depends on the particular system and time being measured. Although latency can never be totally eliminated, it can be significantly reduced.

To understand how latency impacts virtual desktop performance, it is important to review the two types of latency that most negatively affect VDI: network latency, and access latency.

Network Latency

Network latency is important because it slows the delivery of information to desktops and laptops. Network latency is the time delay from when a data packet leaves the designated source until it arrives at the designated destination. Network latency in a packet-switched network can be measured in two ways, so it is important to take note of the method used:

- one-way (from the source sending a packet to the destination receiving it)
- round-trip (from the source sending a packet to the destination receiving it, and back again).

The amount of network latency experienced is primarily dependent upon two factors: the throughput of the network connection, and the packet size. Network connection throughput is very important as it determines how much data can be simultaneously supported on the network. Packet size affects latency because as the packets get longer, they take more time to transmit. In addition, factors of distance and network types also can affect latency. As a general rule, the greater the network distance, the greater the time it will take to transmit data from one location to another. Fortunately, current improvements in networking technologies are allowing for both greater data throughput as well as faster response times.

Based on all these factors, networking can either be an enabler or it can quickly become the bottleneck. The average person is probably most familiar with the concept of bandwidth as the one advertised by manufacturers of network equipment. However, latency matters equally to the end user experience in the behavior of network connections that result in faster response times. Businesses use the term Quality of Service (QoS) as a unit of measurement for maintaining consistent performance on a network by managing both bandwidth and latency in a coordinated fashion.

Access Latency

A second type of latency that negatively impacts VDI is access latency. Within networks at each end of the journey, a data packet can be subjected to both storage and hard disk access delays. The time delay between when a request is sent to an electronic system and when the access is completed or the requested data is returned, is known as Access Latency.

Access time is the actual amount of time it takes to complete access or return the requested data, and is the way that local desktops or laptop computers, hard disk drives, and solid-state disks are commonly measured. As a rule, when access time is increased, there is a corresponding decrease in performance, ultimately resulting in a poor end user experience.

Why is it important to understand access latency? Because anywhere you can reduce access latency gives you the opportunity to improve performance. For this paper, the most relevant types of access latency relate to computers, disk drives and solid state drives. Below is a primer on how access time is computed in measuring the performance of these various components. (If you are already familiar with these concepts, you may want to skip to the next section, “Reducing Latency Matters”.)

- In a computer, access time is the time interval between the instant at which an instruction control unit initiates a call to request or store data, and the instant at which data is delivered or data storage is started. Computer access time is dependent on several factors that can increase or decrease speed, including the CPU processing speed (MIPs), size of Random Access Memory (GB), bus speed (MHz), interface speed (MHz), type of data storage (disk drive or solid-state), proximity of high-performance data (OS and Temp files) to the CPU, and other factors.

  Because desktop virtualization does not use most of the hardware in the client laptop or desktop, the hardware in the end user’s desktop or laptop is not a constraining factor in virtual desktop performance.

  In virtual desktops, the most important factors affecting computer access time are in the host server (“V3 Appliance”) which is delivering or hosting the virtual desktops. These factors in the host server include: the speed of the CPU and bus, amount of RAM, storage type, and proximity of high-performance storage to the CPU.

V3 Appliance

- In disk drives, disk access time is the interval between the time data is requested by the system and the time the data is provided by the drive. For hard disk drives, disk access time is determined using a sum of: the spin-up time, seek time, rotational delay, and transfer time.
  - Spin-up time is the time required to accelerate the disk to operating speed. Most drives are left spinning to improve access time, but drives may be spun down to reduce energy use or noise as in laptop computers.
  - Seek time is the time for the actuator arm to reach the desired disk track. When disk drive specifications refer to “average access time”, they are usually referring to “average seek time”.
  - Rotational delay is the delay for the rotation of the disk to bring the required disk sector under the read-write mechanism. Rotational delay depends on the rotational speed of a disk, and is measured in revolutions per minute (RPM).
  - Transfer time is the time during which data is actually read or written to medium, with a certain throughput.

- Solid State Disks (SSDs) access data without the use of moving parts, so they do not have what we classically call seek time; however, there is a delay from hardware signal relay (and buffering) that is sometimes described as 'seek time' with a magnitude of ~0.1 milliseconds (ms).

  Solid state disks can take many different form factors that increase or decrease their speed. A PCIe based solid-state drive will always be faster than its cousin the SATA or SAS based SSD. On average, these PCIe-based SSDs have access latency times in the range of 50-100 microseconds (µs).

Understanding where latency occurs and how it is measured allows you to consider ways to minimize latency and its negative affect on performance.

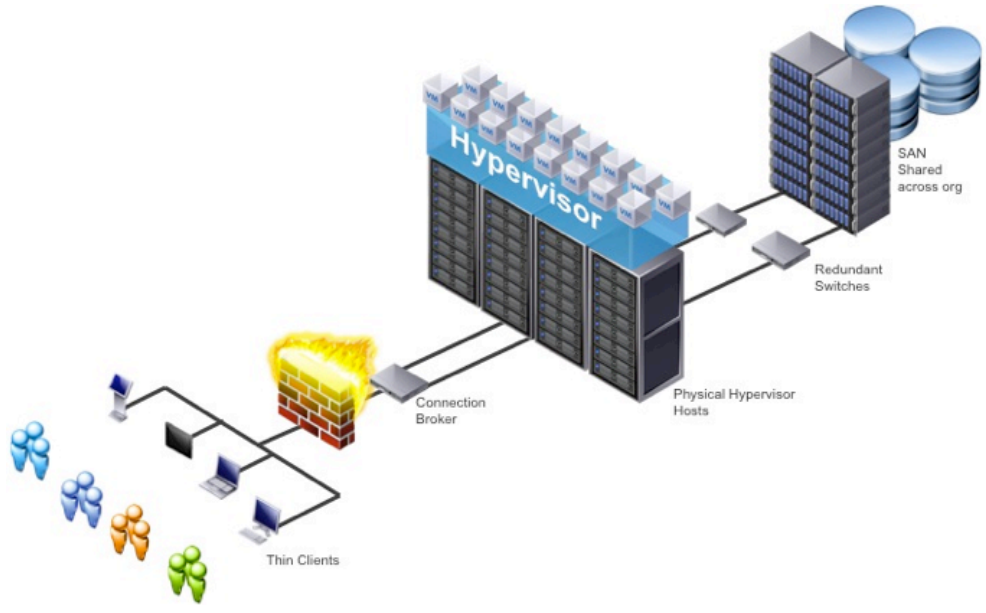
Reducing Latency Matters

As already stated, anywhere there is latency - whether network or access latency - there is an opportunity to increase performance. Reducing latency matters because it allows you to increase performance, productivity, and flexibility while simultaneously decreasing costs.

Because standard desktops are used for many different use cases, their workloads are random and varied. This random access of information increases demands on the storage and networking. Classic virtualized architectures use large arrays of storage and complex redundant networking that are optimized for performance. This approach has been the standard for server virtualization for many years due to the predictive nature of the server workloads. However, rotational delay and seek time provide too much overhead and latency for virtual desktops. Many VDI pilots are halted before production as a result of slow desktop response and poor user experience.

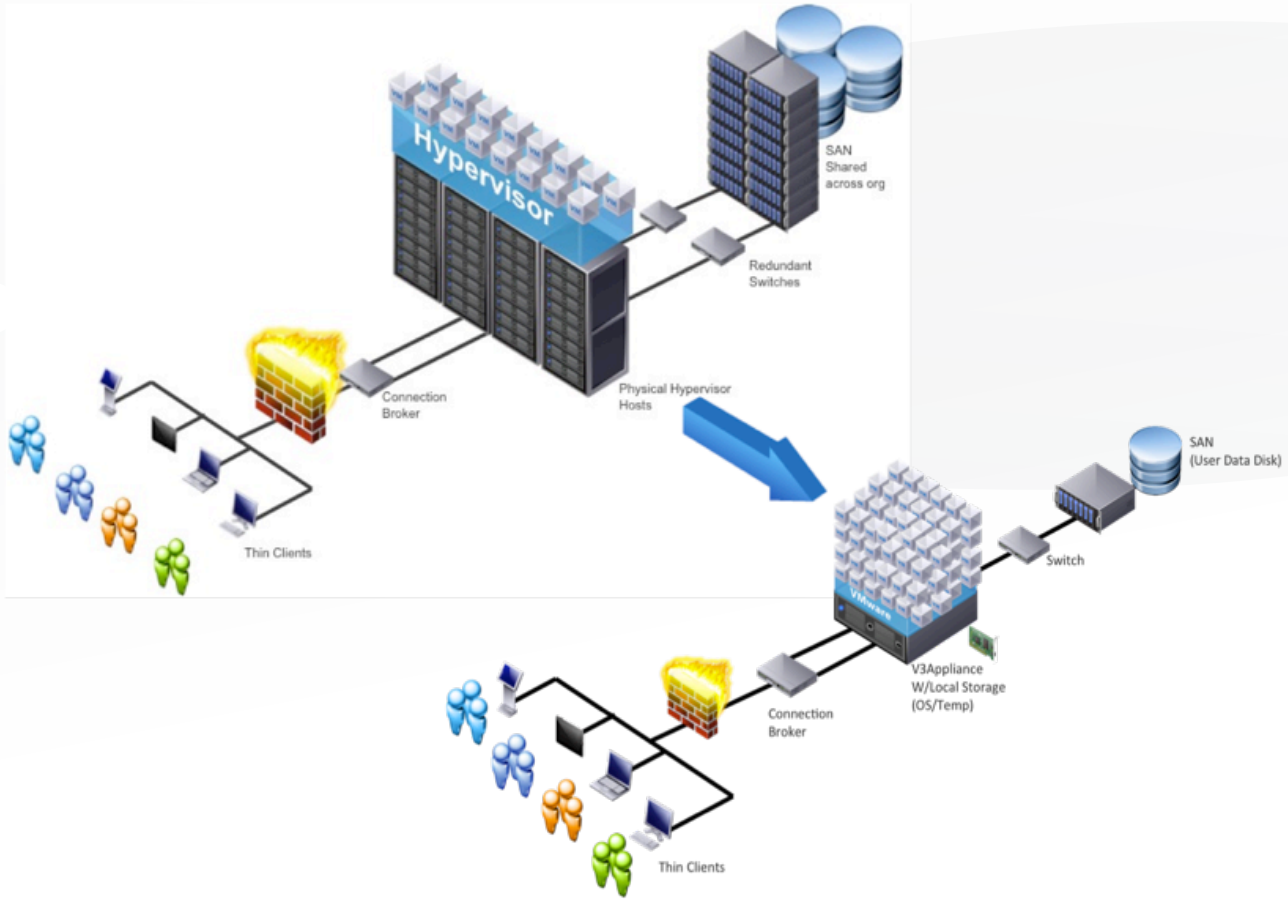
Organizations recognize that the need for large storage arrays remains. Large storage arrays provide reliable storage archive, and are still needed to serve the never-ending growth of critical and non-critical data. However to rely on them for primary high-performance storage (OS and Temp data), especially for VDI, allows for too much latency.

The traditional VDI architecture shown below illustrates the complete reliance on large mechanical disk arrays to serve the virtual desktops. Contrary to perception, mechanical discs in a large array that are optimized for performance within high bandwidth networks simply become too much of a bottleneck for effective VDI.



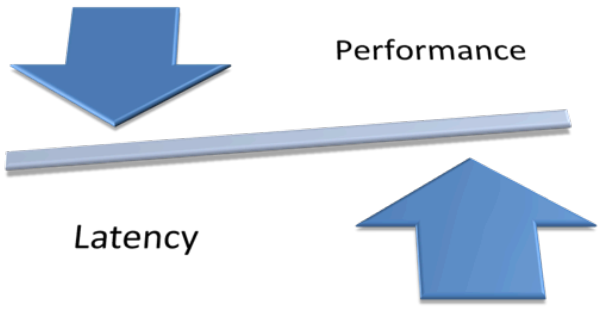
Traditional VDI Architecture using Large Disk Array Storage – complex, slow and expensive.

In contrast (as shown in the next figure), you can see that by moving from a large mechanical disk array to a V3 Appliance using local storage, the VDI architecture is greatly simplified. This simple change results in faster performance than local desktops.



VDI Architecture using the V3 Appliance – simple, fast and significantly less expensive.

Many architects and virtualization specialists believe that the key to high-performance VDI is to use storage with high IOPs (input/output requests per second). Increased IOPs are helpful in many ways, but only after you have addressed the number one issue that plagues VDI deployments, and that is Latency. A system can have the fastest storage available, but it will not matter if the storage is too far from the computing resources. Latency can kill your architecture; it can provide artificial or unessential roadblocks that simply cause a solution to fail. As you increase (or decrease) latency, you directly decrease (or increase) the performance of the computing resources, and ultimately the quality of the end user experience.



Latency and Performance are Inversely Related

Understanding how to reduce latency is one of the primary keys to making VDI successful. Latency can hinder a VDI deployment at many layers. In this whitepaper, we focus specifically on the role of the storage client, application, storage, and networking. All of these play an important role in the success of a VDI deployment.

The V3 Solution

Architectural Change

A V3 appliance has overcomes the primary challenges in reducing latency to provide the fastest virtual desktops in the industry. The V3 approach, along with the set of technologies that enable it to perform faster than physical desktops involve components at both the network and storage layers.

Network

The network is a key factor in reducing latency, due to the sheer amount of data that needs to flow across it. A system will always increase performance when its design reduces or eliminates the distance between the network and the user. Because virtual desktops rely heavily on the network, it is recommended that organizations provide the highest performing network infrastructure available.

Storage

Storage is a critical and fundamental element in this equation.

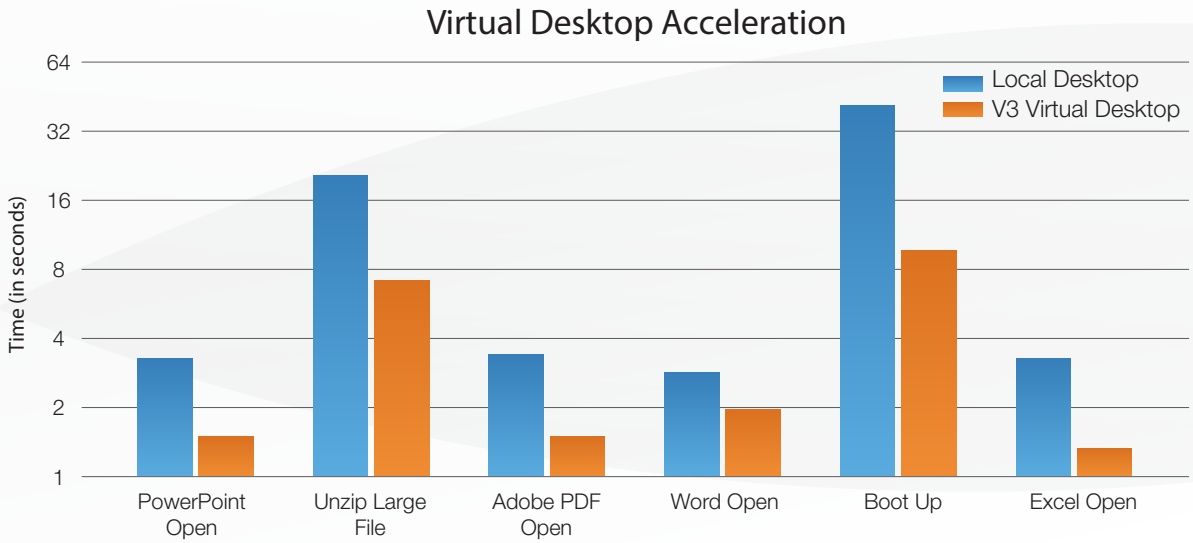
There are three key principles regarding storage that govern VDI success or failure:

1. **Storage location: To obtain the highest performing virtual desktops, primary workload storage needs to be as close to computing resources as possible.** This means the virtual desktops need to be deployed on a host having local storage. Network-attached storage or SAN-attached storage can still be leveraged for persistent data storage, but not for primary workload storage (OS and Temp data) for the virtual desktops. Reducing reliance upon expensive network attached storage allows you to not only eliminate unnecessary networking latency but to increase the overall performance of your VDI solution.
2. **Storage Type: Solid-state storage is the only storage medium today that has the essential performance features.** Because it does not have the same limitations of mechanical disks, solid-state storage can be controlled in such a way as to perform at much higher speeds in terms of access latency time, throughput, and IOPs. Solid-state storage comes in many form factors, with the highest performing solid-state storage located on the PCIe (Peripheral Component Interconnect Express) bus. Additional latency and overhead can still play a role if you use solid-state storage sitting behind a Serial Attached SCSI (SAS) or Serial ATA (SATA) controller.
3. **Highly optimized storage: Object-based disk, which allows for intelligent I/O.** An intelligent disk or object-based disk allows for deep interaction and highly optimized control over the key storage area. This control allows direct access to the storage medium in the most efficient way, thus passing on the performance benefits to the OS/Hypervisor and then up to the virtual desktops. The operating system treats this object-based disk like a standard type of storage because of key software that translates object-based storage to standard block storage. The storage is then managed in the exact same manner as before with standard tools and applications. A V3 appliance has created a set of optimization technologies to integrate all levels of the technology stack.

V3 VDI versus Physical Desktops

The first question that needs to be asked of all organizations considering VDI is: What desktops are being replaced? The answer must involve measurement against physical desktops.

As you can see from the performance graph below, the performance results comparing a V3 appliance to a high-performance physical laptop show significantly increased productivity for end users. It is important to understand that these tests make sure every part of each system is tested, allowing a true apple-to-apple comparison.



V3 VDI virtual desktops significantly outperform high-performance physical laptops

Summary

Virtual Desktop Infrastructure, when done correctly, can outperform and provide a better end user experience than a physical desktop or laptop. Understanding network latency and storage latency is key to VDI viability. This education enables a new direction in VDI that gives organizations the ability to replace physical desktops running all types of workloads with higher-performing virtual desktops.

Leveraging V3's VDI Architecture, organizations can now take full advantage of all the benefits of VDI. In the past, traditional VDI solutions could not provide high-speed performance and positive user experience. VDI benefits have traditionally included greater control, increased security, improved manageability, and lower power and cooling costs. However, by leveraging the high-performance V3 appliance, VDI solution can now enable all of the benefits offered by VDI including performance and optimal user experience.

800-736-2738
www.ipro-inc.com/it

146.924