Mainline™
INFORMATION SYSTEMS

White Paper

Virtual Desktop Sizing Guide: VMware View 4.0 and VMware vSphere 4.0 Update 1

vmware® PARTNER
PREMIER SOLUTION PROVIDER
BUSINESS CONTINUITY
DESKTOP VIRTUALIZATION
AUTHORIZED CONSULTANT
INFRASTRUCTURE VIRTUALIZATION

IBM Premier Business Partner

Table of Contents

Introduction

Purpose

This document describes the virtualization of a given user workload, using technology from VMware® and IBM®, as well as provides a useful sizing guide for enterprise deployments of VMware View™ 4, the VMware software platform for building a virtual desktop infrastructure (VDI) environment. This paper studies two different user workloads, as well as the effect of implementing the PC-over-IP (PCoIP) and Remote Desktop Protocol (RDP) display protocols. With the performance metrics provided, a reader should be able to plan for and provide the expected virtualization capacity required for a VDI solution. User experience between the display protocols is outside the scope of this document.

Executive Summary

The testing demonstrated that the IBM® BladeCenter® HS22 blade server, utilizing VMware vSphere™ 4 update 1, could scale a Task Worker workload to 200 simultaneous user sessions. This testing was performed for both the RDP and PCoIP display protocols. While the PCoIP protocol did show signs of increased CPU and disk usage, the IBM HS22, with Intel® Xeon® 5570 processors and 96GB of RAM, was capable of virtualizing the vSphere compute maximum of 25 desktops per core. The testing also showed that while the server did not reach full CPU capacity, adding additional virtual machines increased signs of scheduling pressure, as the CPU ready values began to rise to unacceptable levels.

When the testing workload was increased to the levels of a Knowledge Worker, the IBM HS22 was able to virtualize 150 simultaneous user sessions with the RDP display protocol. Because of the increased CPU utilization of the PCoIP protocol, the maximum simultaneous user sessions had to be reduced to 120 to allow for the overhead. The average CPU utilization increase on the host for using the PCoIP protocol was 12 -14 percent. This equates to 19 users per core for the RDP Knowledge Worker, and 15 users per core for the PCoIP Knowledge Worker.

As with all virtualized environments, careful consideration must be given to both peak and average server utilization rates. This study shows that all user workloads generate some level of peak-over-average server utilization. For the Task Worker, testing shows a 10-16 percent peak-over-average CPU utilization, depending on the display protocol used. In contrast, the Knowledge Worker workload showed a peak-over-average CPU utilization of 23 -31 percent. This level of increased peak CPU utilization is attributed to the higher demand of resources that the Knowledge Worker workload generated. This is important when planning for peak utilization of users that require larger amounts of server resources.

While the performance metrics provided in this paper are useful for the general planning and sizing of a VMware View environment, they are not meant to be a replacement for a detailed desktop capacity planning assessment. Performing a full capacity planning assessment will provide more precise sizing requirements, and help to determine whether or not a particular user's desktop should be virtualized.

The Environment

VMware View

Purpose-built for delivering desktops as a managed service, VMware View provides the best end user experience, and transforms IT by simplifying and automating desktop management. Centrally maintaining desktops, applications, and data reduces costs and improves security, while at the same time increases availability and flexibility for end users. Unlike other desktop virtualization products, VMware View is a tightly integrated end-to-end solution built on the industry leading virtualization platform, allowing customers to extend powerful business continuity and disaster recovery features to their desktops, as well as standardize on a common platform from the desktop through the datacenter to the cloud.

VMware Lab Manager

VMware Lab Manager, part of the VMware family of management products, provides on-demand access and automated management of the internal cloud for development and testing, enabling higher service levels for end users and eliminating repetitive provisioning tasks for IT staff.

IBM HS22 Blade Server

The IBM BladeCenter™ HS22 blade server supports 6-core, 4-core, and 2-core Intel Xeon 5500 and 5600 series processors. The HS22 supports up to 96GB of RAM across 12 DIMM slots with Chipkill™ protection for both performance and reliability. Depending on the memory configuration, the HS22 supports memory speeds of 1333 MHz or 1066 MHz, when fully populated.

An onboard dual-port Broadcom Gigabit Ethernet controller is standard, which offers TCP Offload Engine (TOE) support. Additional daughter cards can be added to give the HS22 maximum I/O flexibility and throughput. Power management capabilities are incorporated into the control of the blade, which can reduce processor frequency to maintain acceptable thermal and power levels.

VMware ESXi can be included via onboard USB memory key as an option. This diskless solution offers increased security, and the ability to provision a VMware ESX™ server with little effort. If local disk is required, the HS22 supports hot-swap SAS, SATA, and solid state drives.

Depending on chassis selection, an IBM BladeCenter chassis will hold up to 14 HS22 blades in a 7U (BladeCenter E) or 9U (BladeCenter H) rack footprint. A BladeCenter S chassis supports up to 6 blades in a 7U rack footprint, but also has the ability to include up to 12 internal SAS or SATA drive storage in the chassis. With the ability to support 110V power, the BladeCenter S works well in a small business or branch office environment.

The model 7870 HS22 being tested has two Xeon 5570 2.93GHz Quad core processors and 96GB of DDR3 RAM.

IBM DS4500 Disk System

The HS22 Blades were connected to an IBM System Storage™ DS4500 disk system to run the virtual desktops. The DS4500 was configured with two 12-drive RAID 5 disk arrays, each array containing 73GB 10K RPM fiber channel drives. The arrays were formatted to two 746GB VMFS datastores that were assigned to a single desktop pool for virtual machine storage and disk load distribution. To save space and enable quick provisioning of the desktops, the desktop pool was configured to use the linked clone capabilities of VMware View Composer.

Reference Architecture Workload Code – RAWC

RAWC is a simulation tool that provides a realistic and customizable application workload to simulate expected desktop virtual machine utilization in a given server environment. This simulation generates the required random workload allowing us to analyze CPU, memory, and disk utilization, as well as produce VDI sizing recommendations.

The current version of the tool is designed to run workloads on a Windows XP operating system. The RAWC controller application is used to define the desired set of applications that will be used during the simulation. Currently, the list of applications includes Microsoft Office, Adobe Reader, McAfee Virus Scan, Windows Media Player, Java, and 7-zip. The controller application can also be used to vary the start delay of the test, customize the number of emails generated, vary typing speed, and introduce random delay or pauses in the simulated workload.

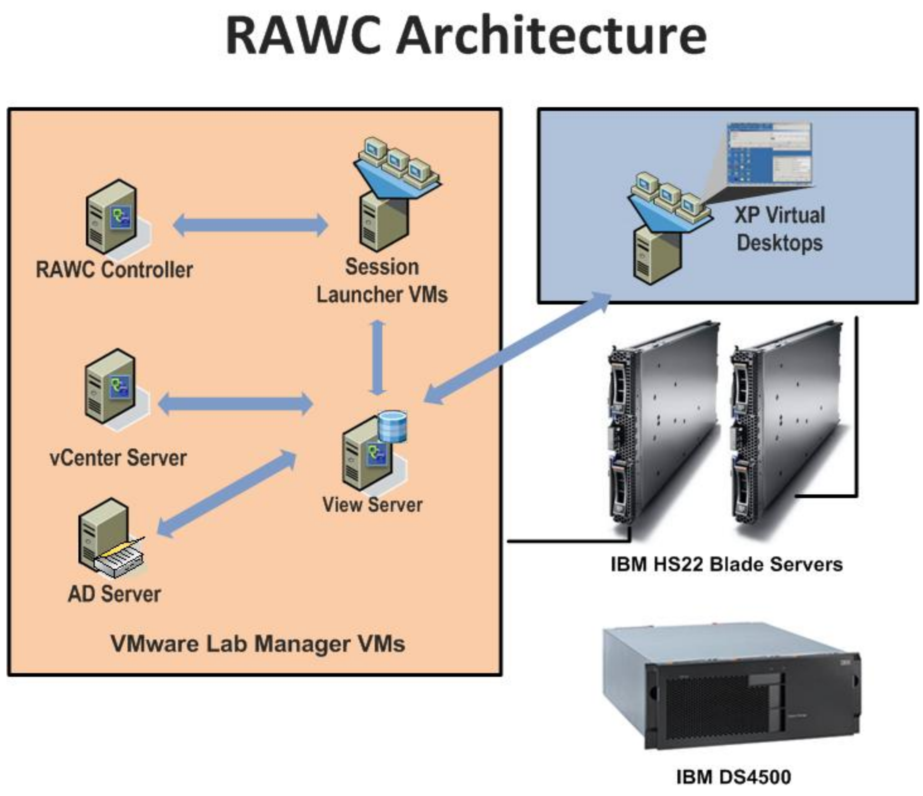


Figure 1 - RAWC Architecture Diagram

The Testing

When it comes to workload planning for VDI environments, various user types and display protocols must be taken into consideration. VMware categorizes the workers into three main types: Task Workers, Knowledge Workers, and Power Users. For our testing, we chose to model two different classes of users: a Task Worker and a Knowledge Worker. Since Power User workloads are subject to an unknown array of applications and variances in user habits, they are difficult to simulate with workload scripting tools and were not a part of our testing. To measure the performance impact of display protocol usage in a VMware View environment, each Worker test was conducted utilizing RDP only and PCoIP only, and the results were compared.

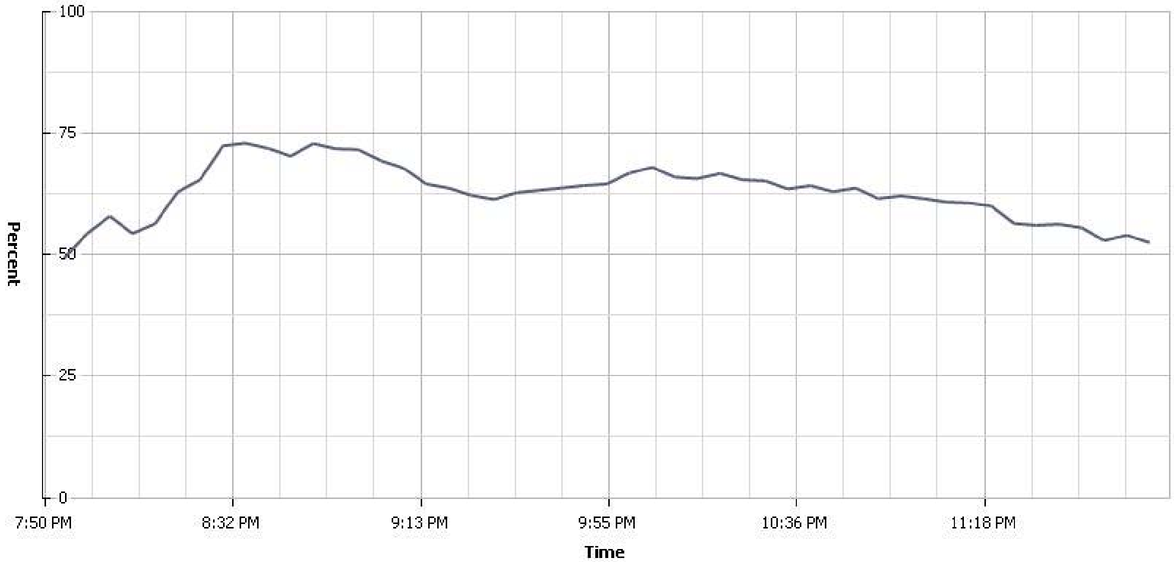
A Task Worker is defined as a user who performs repetitive functions with a small set of applications that are not resource intensive. Examples of a Task Worker can include transcriptionists, call center employees, and administrative workers. For the purpose of our testing, each Task Worker virtual machine was limited to the following applications: Outlook, Word, Excel, and Adobe Reader. Although the application sets in different work environments may vary, these particular applications were selected because they were determined to be representative of a Task Worker and the workload they generate.

A Knowledge Worker is defined as an individual who uses a wider selection of applications, creates more complex documents such as PowerPoint presentations, accesses the Internet regularly, and has consistent email use. Examples of a Knowledge Worker can include business managers, research analysts, and sales professionals. For the purpose of our testing, each Knowledge Worker virtual machine was expanded to use the following applications: Outlook, Word, Excel, PowerPoint, Internet Explorer, Adobe Reader, ZIP file creation, and a Java application. Again, the application sets may vary in each environment; however, the workload generated by these applications is expected to be representative of a typical Knowledge Worker.

For each test we set the desired workload and the desired display protocol. For the Task Worker virtual machines, each virtual machine was configured with 1 vCPU and 512MB of RAM. The Knowledge Worker virtual machines were also configured with 1 vCPU, but the RAM was increased to 1GB. Starting with a base user workload of 80 users, we increased the user workload in multiples of 20 user sessions, until either the performance of the server was degraded, or the compute maximum of 25 virtual CPUs per physical core set by VMware for vSphere 4.0 Update 1 was reached. The test was then repeated multiple times to confirm the validity of the testing results. The virtual machine performance statistics were also monitored to ensure that no undesirable conditions existed that would make the virtual machine unusable. The result of each test was then charted and analyzed to determine whether or not the test was successful for the given workload generated.

Workload: 200 Task Worker Virtual Machines | 1 vCPU | 512MB | Protocol: RDP

CPU



Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	CPU Usage	Average	Percent	52.35	72.89	49.21	62.79

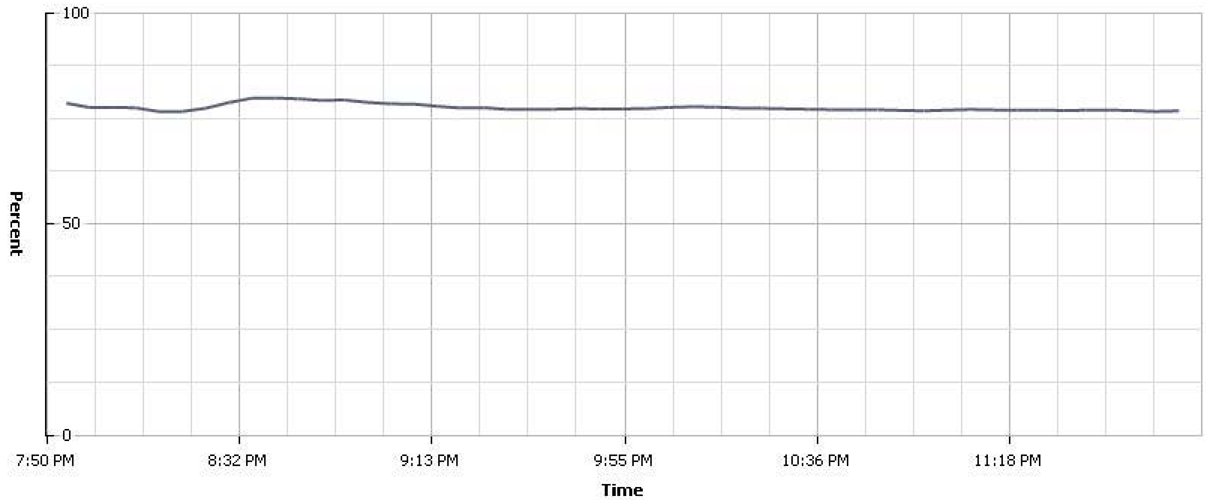
Figure 2 - 200 Task Worker RDP – CPU Usage

In the beginning phases of the test, as the users are being logged in and generating workload, we can see in Figure 2 that CPU load rapidly increased by 20 percent to roughly 73 percent at its peak. The rapid increase of resource demand at mass user login is known as a login storm, and is typical for virtual desktop environments. As the environment moved to a more steady-state workload, we can see that the overall fluctuation in CPU usage was reduced, stabilizing at roughly 63 percent.

With advances in the virtualization capabilities of the Intel Nehalem processors and the core improvements in the VMkernel, we can see that an IBM HS22 blade can easily handle the workload generated by the 200 Task Workers. Although there appears to be available processing capacity, VMware's recommended compute maximum threshold was reached, and no further testing was performed to push beyond those limits.

Further inspection of the virtual desktop's CPU performance reveals a possible explanation for the current compute maximum of 25 virtual machines per core. Although the server had more compute capacity, the average CPU ready values for the virtual desktops was beginning to increase. This suggests that the VMkernel was showing signs of scheduling pressure, and was becoming unable to efficiently schedule all running vCPUs. Possible future remedies include further advancements in the Intel core architecture, and further refinement of the VMkernel scheduling algorithms.

Memory



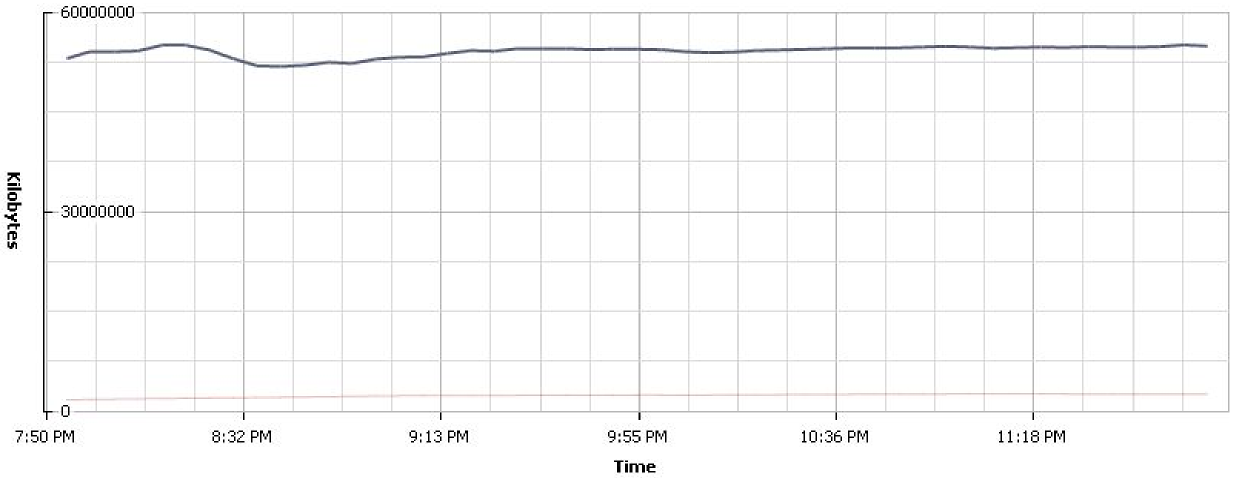
Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Usage	Average	Percent	76.9	79.94	76.74	77.678

Figure 3 - 200 Task Worker RDP – Memory Usage

In Figure 3, we can see that the initial ramp-up of the test caused the memory to spike 3 percent briefly before settling close to an average memory utilization of 78 percent, or 74.9GB. Because the virtual machines were already actively running at the start of the test, transparent page sharing had already consolidated most of the similar memory pages. As the simulator started to launch new programs inside of the virtual machines, the memory page differences created by the launching of new programs quickly began consuming available memory. When comparing the memory usage chart in Figure 3 to the memory saved by transparent page sharing in Figure 4, we can see that the memory being consolidated had a dip at the same time that memory utilization was increasing. As the transparent page scheduler had time to consolidate the new differences in memory, the memory utilization slowly began to decrease and memory sharing began to increase.

Using the VMware formula "Memory Shared - Memory Shared Common = Memory Saved", as well as the data in Figure 4, the memory saved on the host from transparent page sharing was 49.6GB. Note that the memory saved from transparent page sharing is not equal to the "projected memory required minus the in use system RAM on the host", but rather is a calculation of the total memory saved from sharing identical memory pages by the VMkernel.

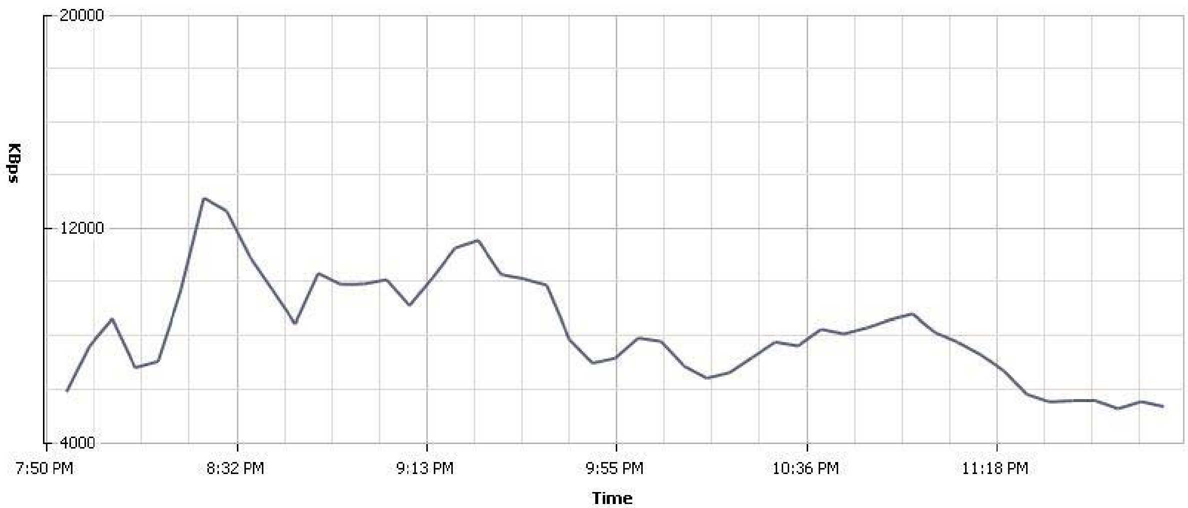


Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Shared	Average	Kilobytes	55000882	55175012	51916045	54202024.
■	10.32.10.116	Memory Shared Common	Average	Kilobytes	2402477	2462459	1572968	2200370.8

Figure 4 - 200 Task Worker RDP – Memory Saved – 49.6GB

Disk



Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Disk Usage	Average	KBps	5333	13158	5250	8233.959

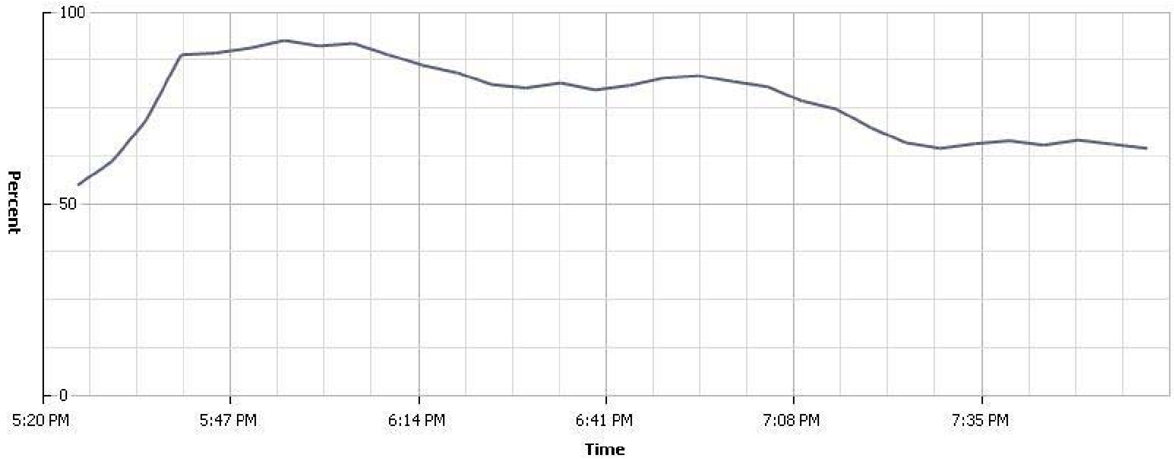
Figure 5 - 200 Task Worker RDP – Disk Usage

Similar to the other performance metrics, the disk utilization quickly rose to a peak throughput of 12.8MB/s during the initial login phase of the test. As the environment moved to a steady state workload, the throughput rates began to fall, but remained unsettled throughout the testing.

The average read/write disk throughput for the test was 8MB/s. The average IOPS per virtual machine was 1.9.

Workload: 200 Task Worker Virtual Machines | 1 vCPU | 512MB | Protocol: PCoIP

CPU



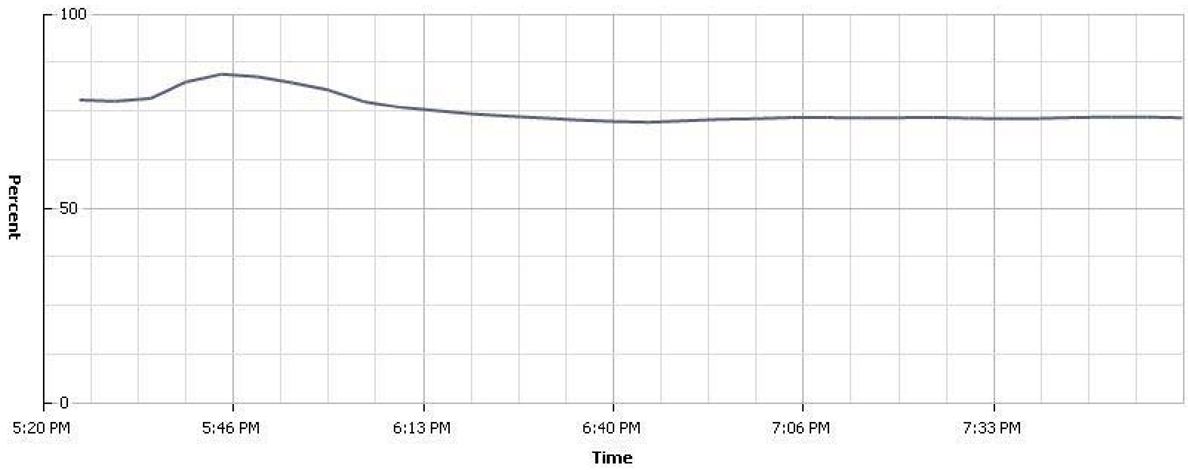
Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	CPU Usage	Average	Percent	64.47	92.73	54.79	77.232

Figure 6 - 200 Task Worker PCoIP – CPU Usage

With the advent of PCoIP in VMware View 4, it is important to understand what impact this protocol will have when architecting an environment. For this stage of testing, we left the Task Worker user session count at 200, and changed the display protocol from RDP to PCoIP. In Figure 6, we see there was a 38 percent increase in CPU usage during the login phase to 93 percent of total capacity. As the workload moved to a steady state, the CPU utilization stabilized at 60-65 percent. The overall average utilization was 77 percent.

Memory

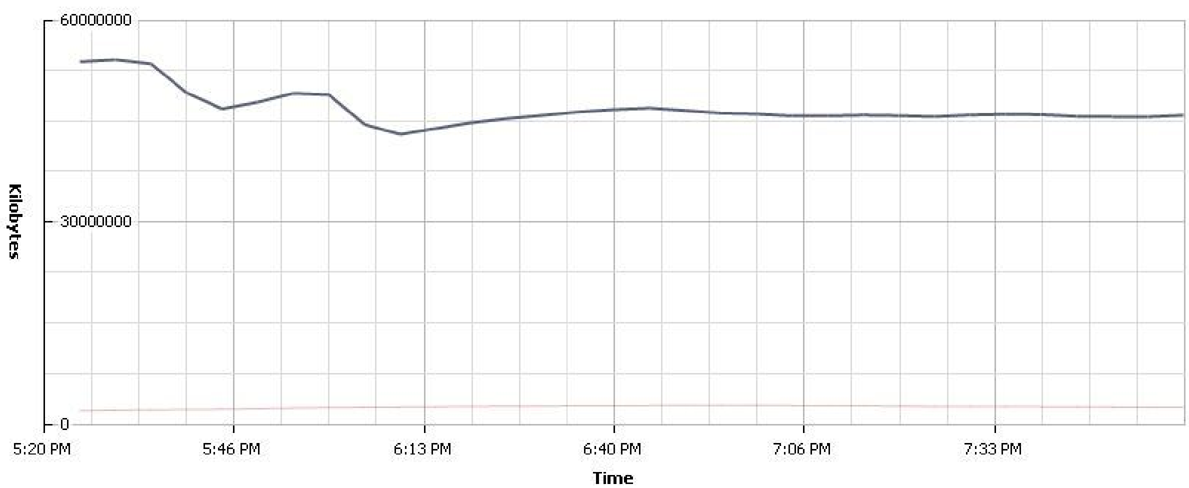


Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Usage	Average	Percent	73.33	84.61	72.22	75.47

Figure 7 - 200 Task Worker PCoIP – Memory Usage

As in the previous test, Figure 7 shows that the memory utilization followed a similar pattern of peak usage during the login phase of the testing. The memory utilization reached a peak of 85 percent or approximately 81.6GB. As the workload began to stabilize, transparent page sharing was able to reduce the memory footprint to an overall average of 75 percent. By the end of the test, memory usage had reduced to 73 percent or approximately 70GB. Looking at the Memory Shared figures in Figure 8, the average memory saved on the host is roughly 42GB.

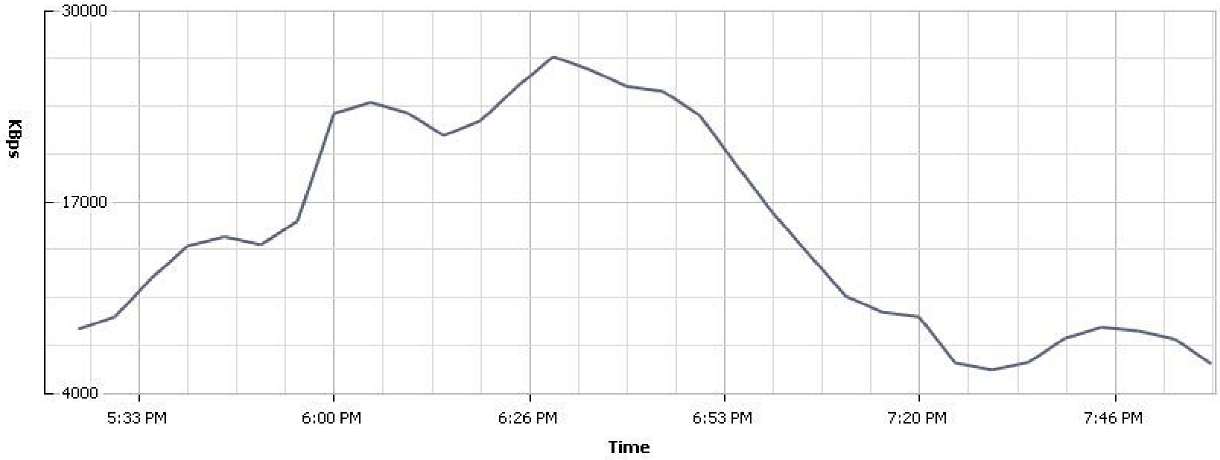


Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Shared	Average	Kilobytes	45920423	54172047	43082535	46904405.
■	10.32.10.116	Memory Shared Common	Average	Kilobytes	2433381	2692581	1919135	2456862.1

Figure 8 - 200 Task Worker PCoIP – Memory Saved - 42.4GB

Disk



Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Disk Usage	Average	KBps	5939	26917	5536	15176.906

Figure 9 - 200 Task Worker PCoIP – Disk Usage

Unlike CPU and memory utilization, disk usage did not peak during the login period. Figure 9 shows that the disk utilization had a steady climb to a peak utilization of 26.3MB/s, as the workload on the virtual desktops increased. After the peak period of disk usage, there was a steady decrease of disk activity to a minimum of 5.4MB/s as the virtual machine workload became steady. The average read/write disk throughput for the test was 15.3MB/s. The average IOPS per virtual machine was 3.1.

Comparing the Protocols – Task Worker 200 User Sessions

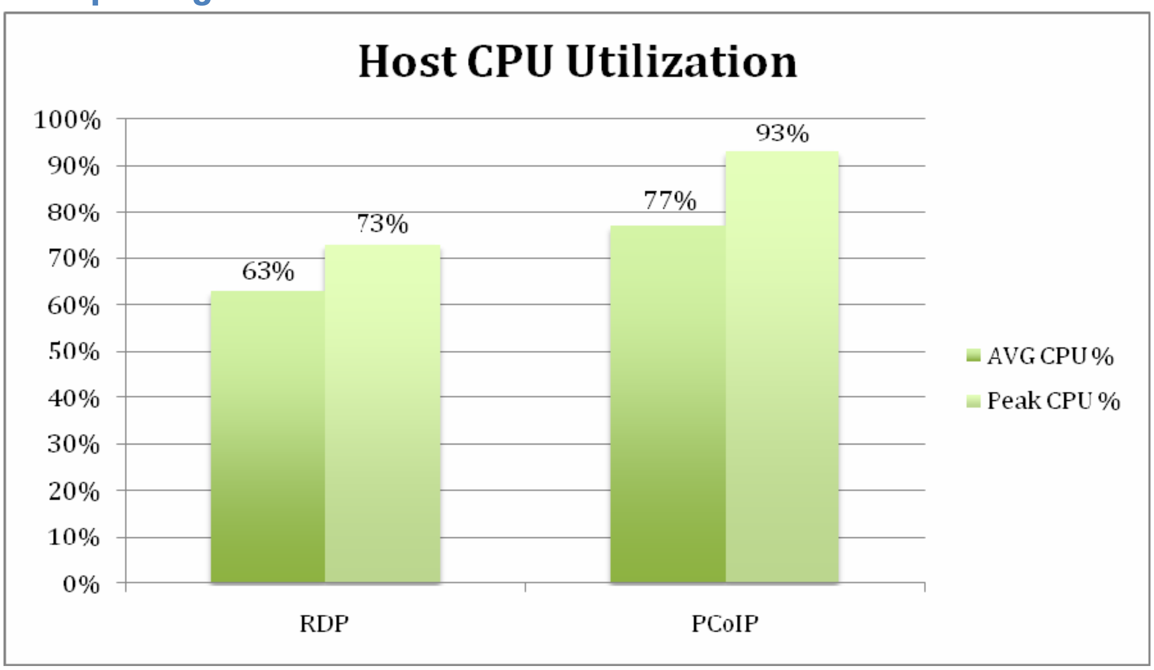


Figure 10 – Task Worker – Protocol CPU Comparison

Analyzing the average CPU utilization rates of both protocols shows a marked increase in CPU load while using PCoIP. Since the PCoIP engine resides as software on the virtual machines, the increase in CPU usage is expected. Based on the chart in Figure 10, the 200 Task Worker test shows an average increase in CPU load on the host of 14 percent. This translates to a 16.8 MHz, .5% of the host core, increase per virtual desktop using PCoIP, or an average of 92.4 MHz per virtual desktop. In contrast, the RDP workload used an average of 75.6 MHz per virtual desktop.

When analyzing peak-to-average CPU utilization, the PCoIP protocol had a 16 percent peak-over-average CPU utilization for the Knowledge Worker workload. The RDP peak-over-average CPU utilization was 10 percent. This represents a 6 percent higher peak-to-average CPU utilization rate for the PCoIP protocol over RDP.

While signs of increased CPU utilization were found throughout the testing of the PCoIP display protocol, VMware had indicated that code optimization in the PCoIP protocol may reduce the increased CPU overhead seen in this testing. The updated code will be available with the next released version of View.

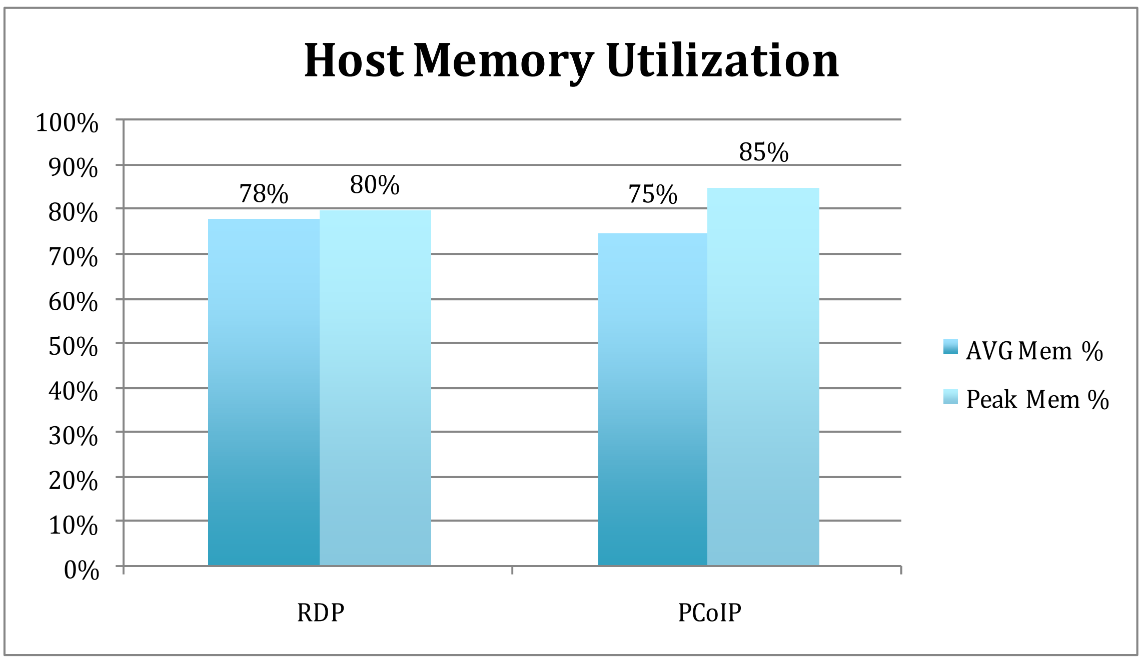


Figure 11 – Task Worker – Protocol Memory Comparison

Analyzing the average memory consumption of both protocols shows that there is a negligible difference in memory utilization. Based on the chart in Figure 11, the RDP test shows a 3 percent higher average memory utilization over PCoIP, which can be contributed to subtle variances in the actual test workload that was generated. This translates to an average memory usage of 364-383MB per virtual desktop.

When analyzing the peak-to-average memory utilization, the PCoIP protocol had a 10 percent peak-over-average memory utilization for the Task Worker workload. In contrast, the RDP peak-over-average memory utilization was 2 percent. This represents an 8 percent higher peak-to-average memory utilization rate for the PCoIP protocol over RDP.

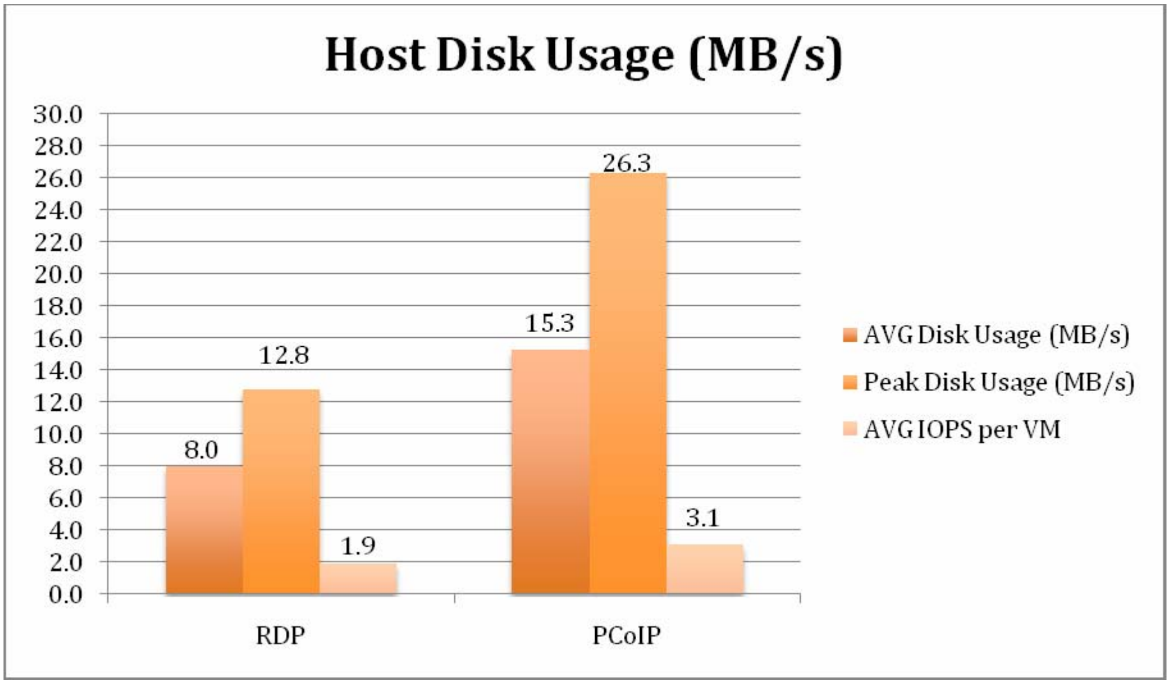


Figure 12 – Task Worker – Protocol Disk Comparison

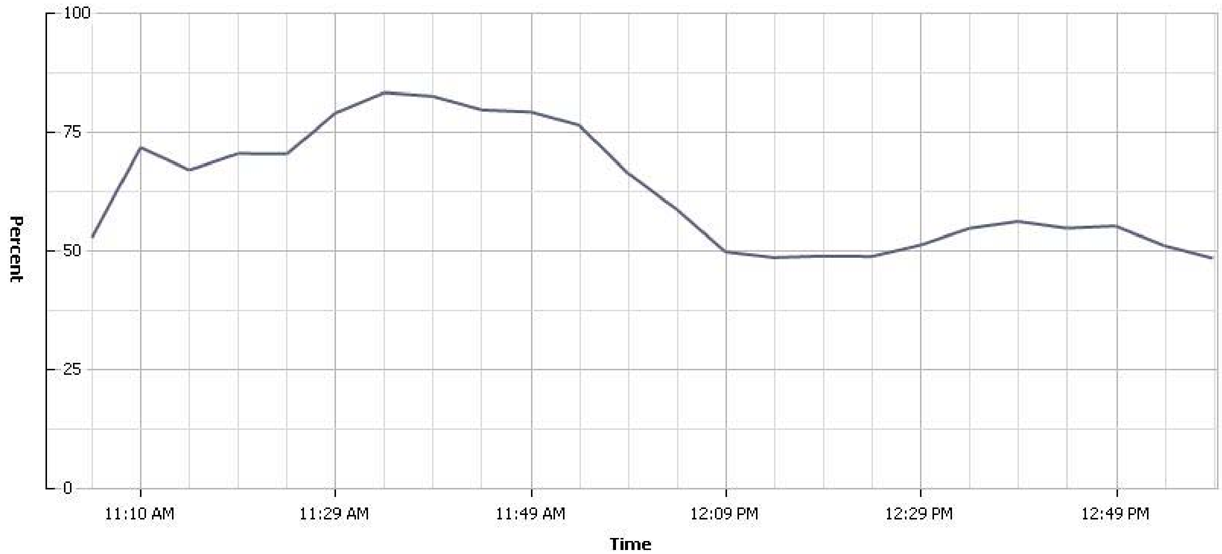
Analyzing the disk usage of both protocols shows a marked increase in the average disk throughput requirements while using the PCoIP protocol. Based on Figure 12, the PCoIP test shows an average increase of disk usage of 91 percent or 7.3MB/s on the host. This translates to an increase of 37.4KB/s per virtual desktop, and a total average of 78.4KB/s per virtual desktop, while using the PCoIP display protocol. In contrast, the RDP protocol required an average of 40.9KB/s per virtual desktop. The average IOPS per virtual machine increased by 1.2 for the PCoIP protocol.

When analyzing the peak-to-average disk utilization, the PCoIP protocol had an 11MB/s peak-over-average disk utilization for the Task Worker workload. In contrast, the RDP peak-over-average disk utilization was 4.8MB/s. This represents a 6.2MB/s higher peak-to-average disk utilization rate for the PCoIP protocol over RDP.

While the exact reasons for the differences in host disk usage warrant further exploration, Figure 24 shows the results of increasing the virtual desktop VRAM setting in the target environment.

Workload: 150 Knowledge Worker Virtual Machines | 1 vCPU | 1GB | Protocol: RDP

CPU



Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	CPU Usage	Average	Percent	48.24	83.29	48.24	62.615

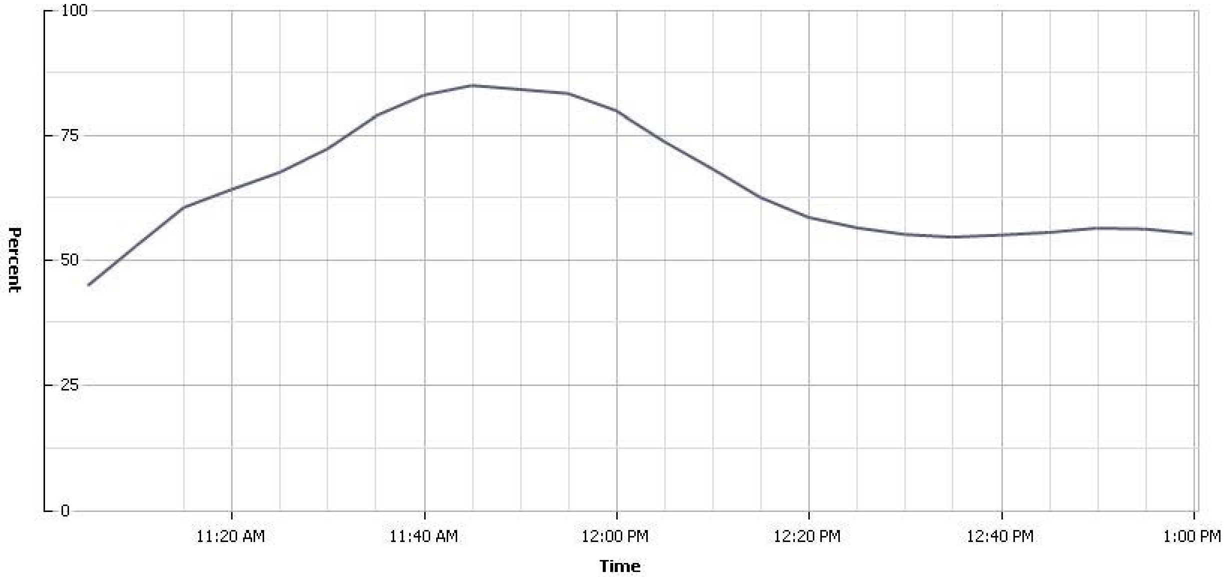
Figure 13 - 150 Knowledge Worker RDP – CPU Usage

For this phase of testing, the simulated workload has been increased to levels that emulate a Knowledge Worker. The additional workload includes PowerPoint, Internet Explorer, zip file creation, and execution of a Java application. It does not include simulation of a full virus scan, as this workload is normally scheduled as an after-hours operation, and would not be included as an expected, everyday workload.

During the login phase of the test, the CPU utilization quickly rose to just below 75 percent. Based on Figure 13, as the simulation moved to a steady state, the load continued to increase demand on the CPU resources to a maximum of 83 percent. An hour into the test, the load slowly began to decrease to 48 percent, with a small spike in utilization to 56 percent toward the end of the simulation.

With the varied nature of a Knowledge Worker environment, it is important to understand that the average workload may be significantly less than the peak load experienced throughout a given work day. Because of these peaks in workload demand, additional peak load capacity must be taken into careful consideration when planning for a VDI environment. Even though the server did not reach its maximum capacity in this test, and on average ran 63 percent of maximum CPU capacity, upping the workload to 160 user sessions pushed the server to its limits during peak load. As a result, the user workload was decreased by 10 user sessions to bring the available capacity during peak load down to more acceptable tolerances.

Memory



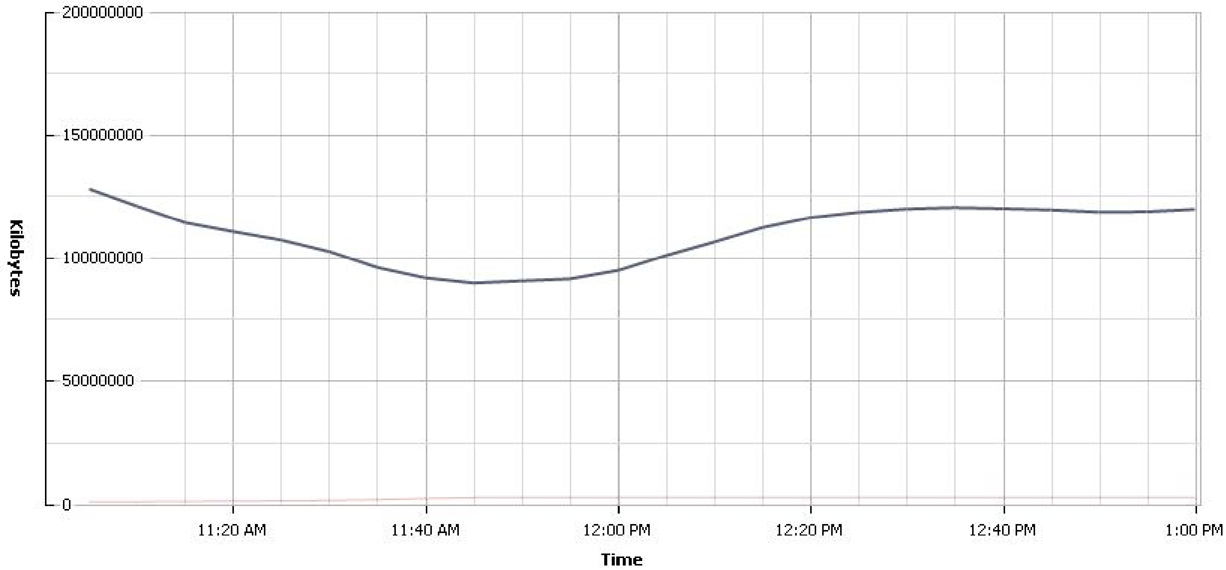
Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Usage	Average	Percent	55.35	85.03	44.93	65.267

Figure 14 - 150 Knowledge Worker RDP – Memory Usage

In Figure 14, the memory appears to have a steady climb, as expected during the login period; however, it continues to increase even though the workload at this point is considered to be at a steady state. Eventually, the memory usage reaches a maximum of 85 percent, and then begins to decrease at a steady rate to close to 55 percent. The average memory usage for this test is 65 percent.

When comparing this chart to the CPU performance chart, there is a notable increase in the demand for memory resources as the CPU demand increased. Similarly, as the CPU demand decreased, so did the demand in memory resources. When taking into consideration the memory saved chart in Figure 15, it appears as though the generated workload created a large difference in the number of identical pages than in the previous Task Worker testing. Because of these differences, more physical RAM was committed to the running of virtual machines on the host to make up for the difference in memory pages on the guest virtual machines. As the memory space became more identical, transparent page sharing was able to dramatically reduce the committed physical RAM, by roughly 30 percent. The average memory saved on the host from transparent page sharing was 103GB.

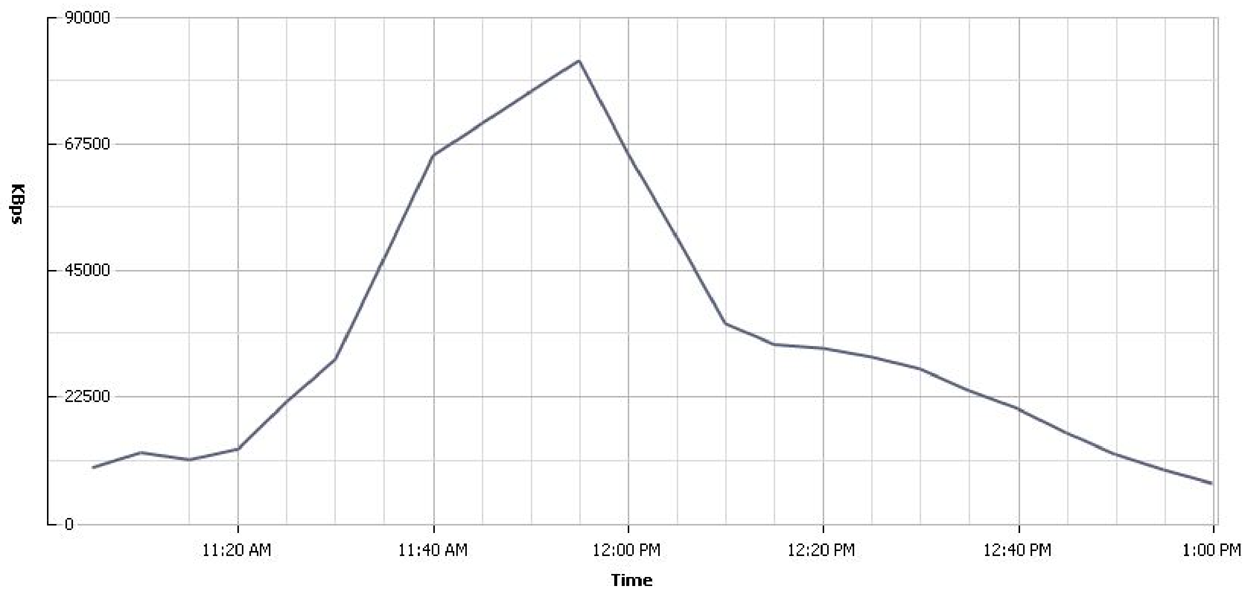


Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Shared	Average	Kilobytes	120003672	128181220	90041764	109840100
■	10.32.10.116	Memory Shared Common	Average	Kilobytes	2578333	2670872	837899	2193234.5

Figure 15 - 150 Knowledge Worker RDP – Memory Saved 103GB

Disk



Performance Chart Legend

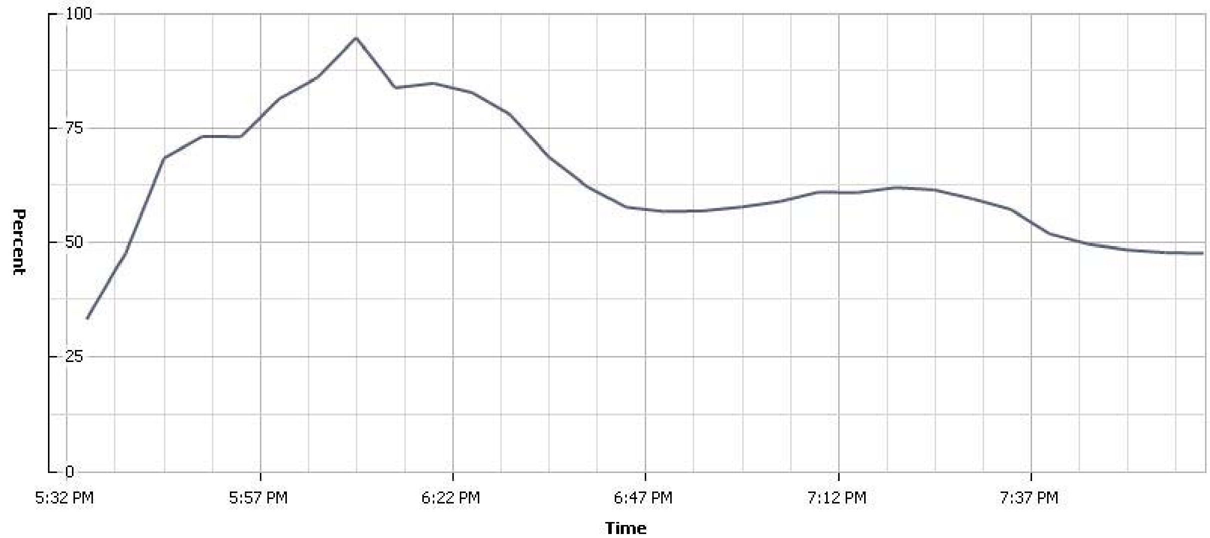
Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Disk Usage	Average	KBps	7077	82481	7077	33457.5

Figure 16 - 150 Knowledge Worker RDP – Disk Usage

Disk usage in Figure 16 shows a similar pattern of resource utilization increase as in the other performance metric charts. During the login phase, the disk load is increased; however, the load continues to increase to a peak usage of 81MB/s. As the simulation load begins to taper off, the disk load decreases to approximately 6.9 MB/s at the end of the test. The average read/write disk throughput for the test was 33MB/s.

Workload: 120 Knowledge Worker Virtual Machines | 1 vCPU | 1GB | Protocol: PCoIP

CPU



Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	CPU Usage	Average	Percent	47.61	94.84	33.14	63.789

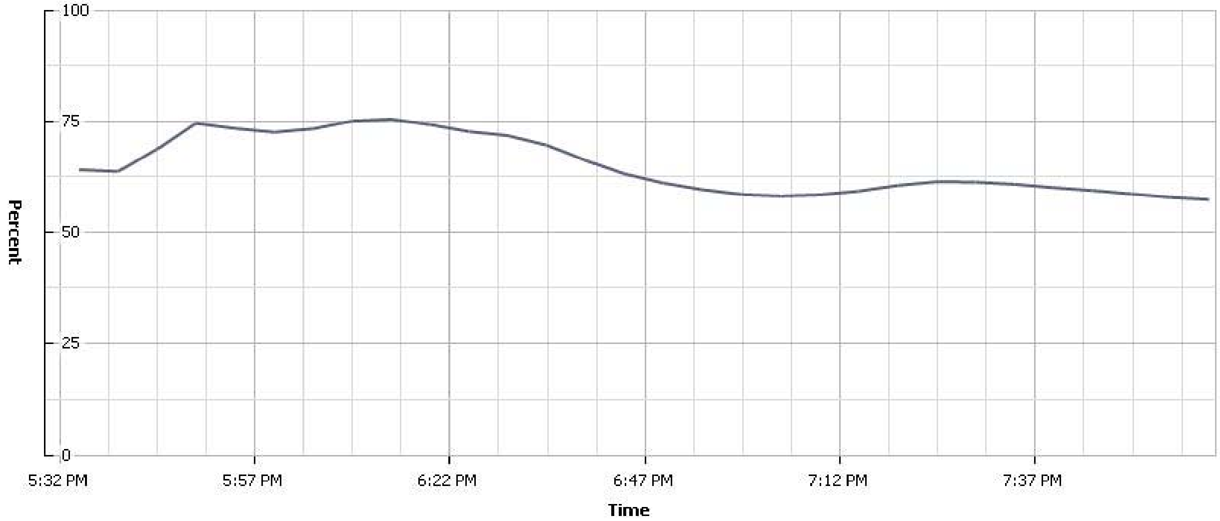
Figure 17 - 120 Knowledge Worker PCoIP – CPU Usage

For the final stage of Knowledge Worker testing, the display protocol was changed to measure any impact the PCoIP protocol may have in the deployment of a VDI solution. As determined by earlier Task Worker testing, there is an impact on CPU utilization when using the PCoIP protocol. Because of the increased load, the total simultaneous user sessions was reduced from 150 to 120.

Analyzing CPU utilization in Figure 17, the simulated workload generated a similar pattern of performance as in the earlier Knowledge Worker testing using RDP. As expected, the CPU utilization quickly rose in the login phase of test. As the test moved to a steady state, the CPU load continued to steadily increase and peak at a maximum of 95 percent of available resources. After the peak in CPU utilization, the load began to decrease and level off between 56 percent and 62 percent.

As in the previous test, the CPU load never quite reached maximum capacity, but the peak levels reached were clearly higher overall. Average CPU utilization for the test was 64 percent. Again, planning for expected peak utilization rates in a VDI environment is critical, and enough headroom should be calculated into the sizing estimates to account for this peak usage.

Memory

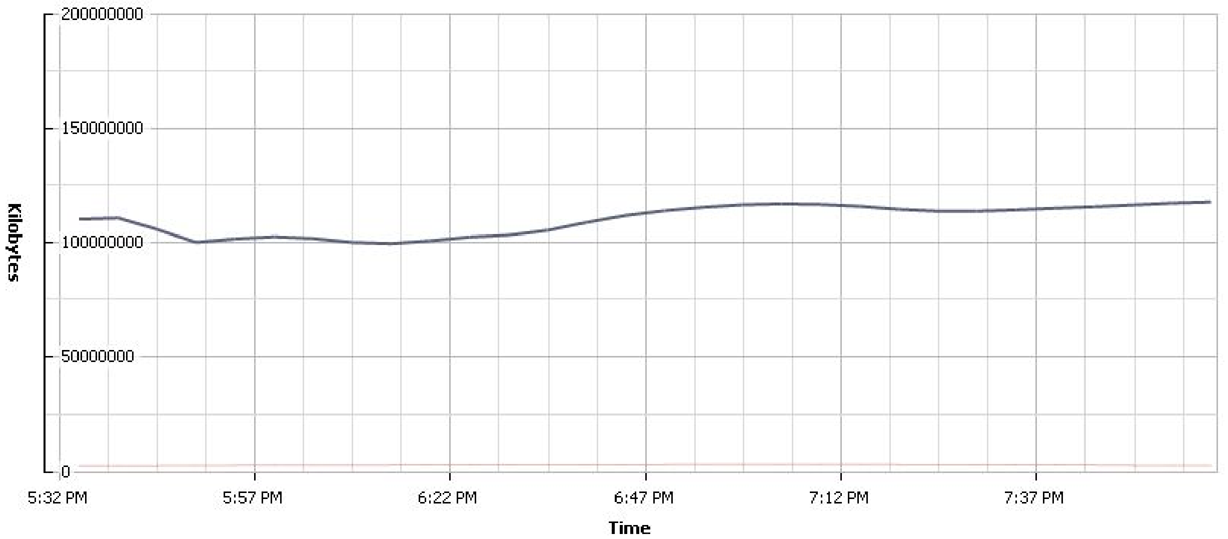


Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Usage	Average	Percent	57.47	75.56	57.47	65.12

Figure 18 - 120 Knowledge Worker PCoIP – Memory Usage

Following a similar pattern of memory usage as previous testing, Figure 18 shows a quick increase in the memory utilization during the login phase of the testing. During the highest workload portions of the test, the memory stayed around 76 percent of peak usage. As the workload began to taper off, the memory was able to be consolidated down to 57 percent usage towards the second half of the test. The average memory consumption for the test was 65 percent. As shown in Figure 19, the average memory saved from transparent page sharing was 102GB.

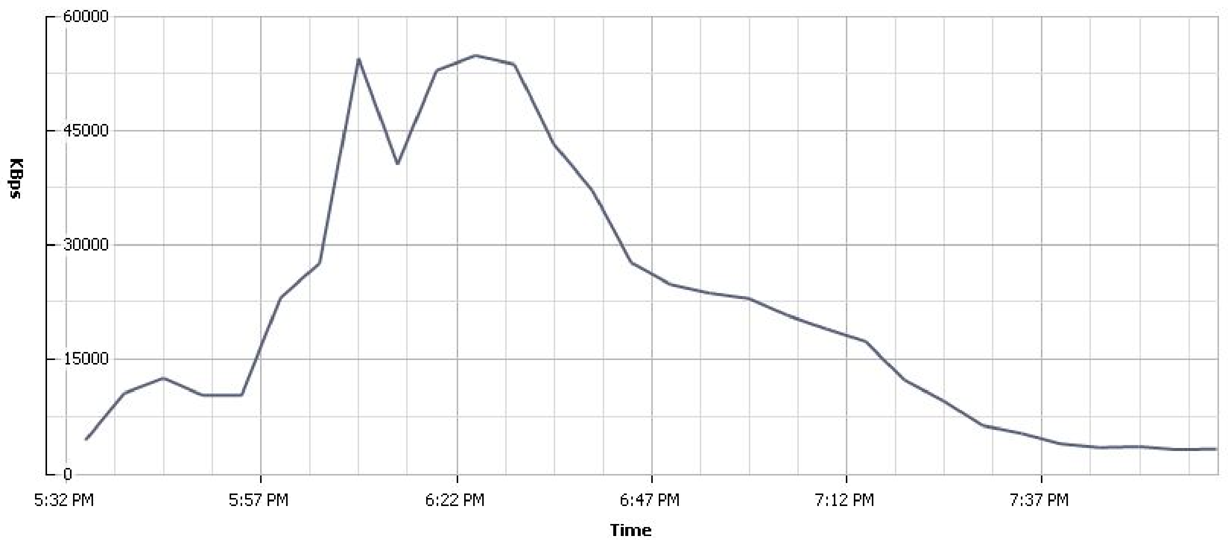


Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Memory Shared	Average	Kilobytes	117929069	117929069	99625941	110108769
■	10.32.10.116	Memory Shared Common	Average	Kilobytes	2398945	2873886	2228969	2670627.1

Figure 19 - 120 Knowledge Worker PCoIP – Memory Saved 102GB

Disk



Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum	Average
■	10.32.10.116	Disk Usage	Average	KBps	3199	54925	3102	21422.1

Figure 20 - 120 Knowledge Worker PCoIP – Disk Usage

As with the previous tests, the disk usage quickly increased during the login phase. As the workload moved to a steady state, disk throughput steadily increased to its peak utilization rate of 53.6MB/s. As the workload began to taper off in the second half of the test, the disk throughput steadily decreased as well. The average read/write disk throughput for the test was 20.9MB/s.

Comparing the Protocols – Knowledge Worker 120 User Sessions

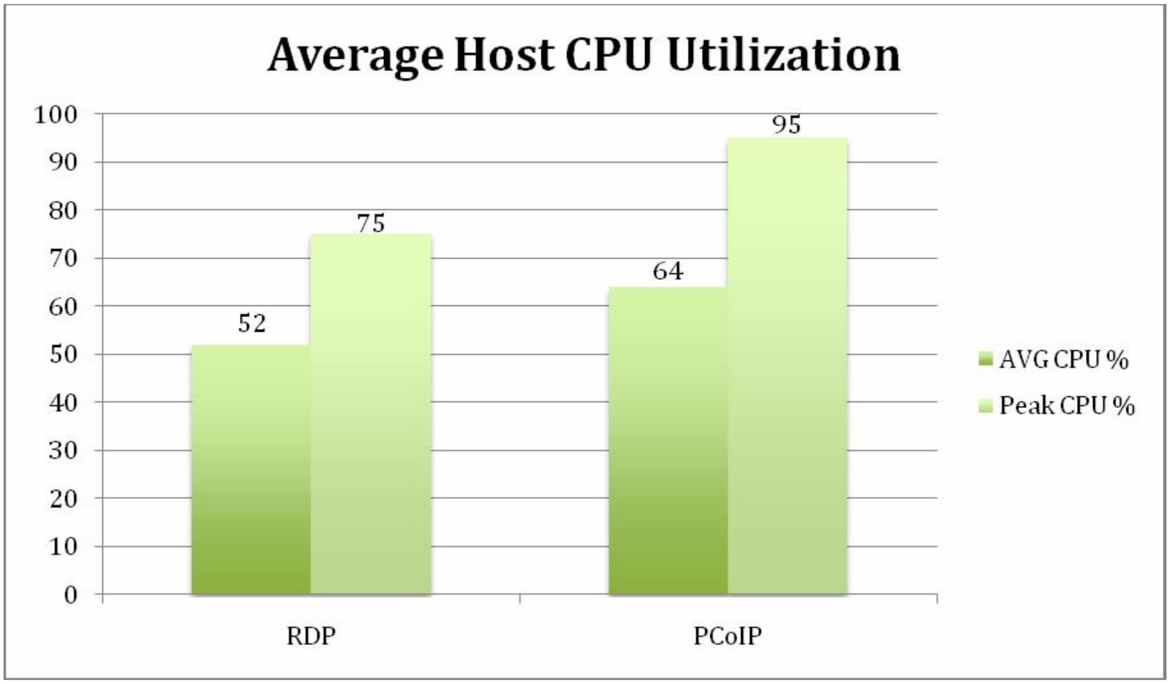


Figure 21 - Knowledge Worker – Protocol CPU Comparison

Since the total user sessions performed in each test was not consistent, additional testing was performed for the Knowledge Worker at 120 user sessions to provide direct comparison metrics. Those metrics are provided in these comparison charts for review.

Analyzing the average CPU utilization results in Figure 21, it shows an apparent increase in the CPU requirements for the PCoIP protocol, as was the case in the Task Worker testing. For the Knowledge Worker testing, the overhead on the host for implementing the PCoIP protocol was 12 percent. In comparison to the previous Task Worker test, the average host overhead for the PCoIP protocol falls within a 12 to 14 percent range for both tests. With an average per virtual machine CPU utilization of 128 MHz or 4.26 percent of the host CPU core, the average overhead per virtual machine was 24 MHz or 1% of the host core for the PCoIP protocol.

The average CPU utilization for the RDP virtual machine was 104 MHz or 3.3% of the host CPU core. In comparison to the Task Worker workload, the RDP Knowledge Worker virtual machine had a 28.4 MHz average increase in CPU utilization. The PCoIP Knowledge Worker virtual machine had a 35.6 MHz average increase in CPU utilization. Based on the increased utilization rates of both tests, it appears that the increase in workload did not require additional CPU load for the PCoIP protocol. Only the additional workload itself was measured in the tests for both protocols.

When analyzing the peak-to-average CPU utilizations, the PCoIP protocol had a 31 percent peak-over-average CPU utilization for the Knowledge Worker workload. In contrast, the RDP peak-over-average CPU utilization was 23 percent. This represents an 8 percent higher peak-to-average CPU utilization rate for the PCoIP protocol over RDP. In comparison to the Task Worker workload, the peak-to-average increase for the PCoIP protocol over the RDP protocol falls between 6 & 8 percent.

Planning for this overhead, when implementing the PCoIP protocol, is critical for proper sizing of the VDI environment.

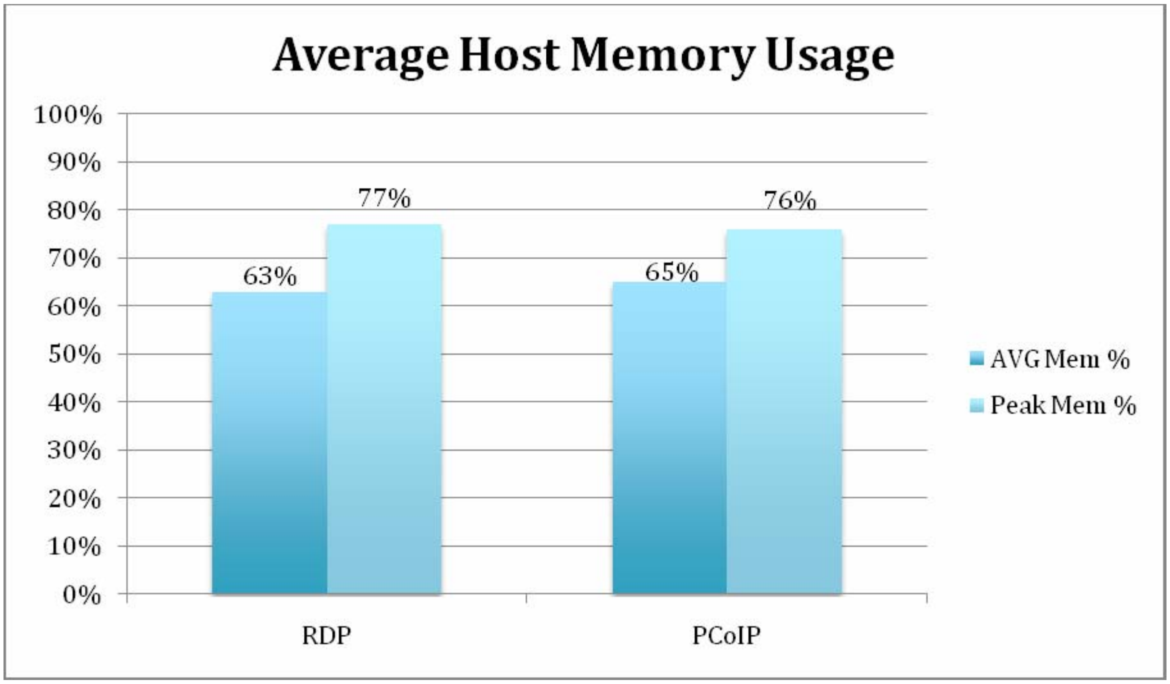


Figure 22 - Knowledge Worker – Protocol Memory Comparison

Analyzing the average memory consumption of both protocols shows that there is a negligible difference in memory utilization. Based on the chart in Figure 22, the PCoIP protocol test shows 2 percent higher average memory utilization over RDP, which can be contributed to subtle variances in the actual test workload that was generated. This translates to an average memory usage of 516-532MB per virtual desktop.

When analyzing the peak-to-average memory utilization, the PCoIP protocol had an 11 percent peak-over-average memory utilization for the Knowledge Worker workload. The RDP peak-over-average memory utilization was 14 percent. This represents a 3 percent higher peak-to-average memory utilization rate for the RDP protocol over PCoIP. Again, these differences are contributed to variances in the test workload, and not an indication of either protocol demanding more memory.

In comparison to the Task Worker workload, the RDP Knowledge Worker virtual machine had a 35 percent or 133MB average increase in RAM utilization. The PCoIP Knowledge Worker virtual machine had a 44 percent or 163MB average increase in RAM utilization.

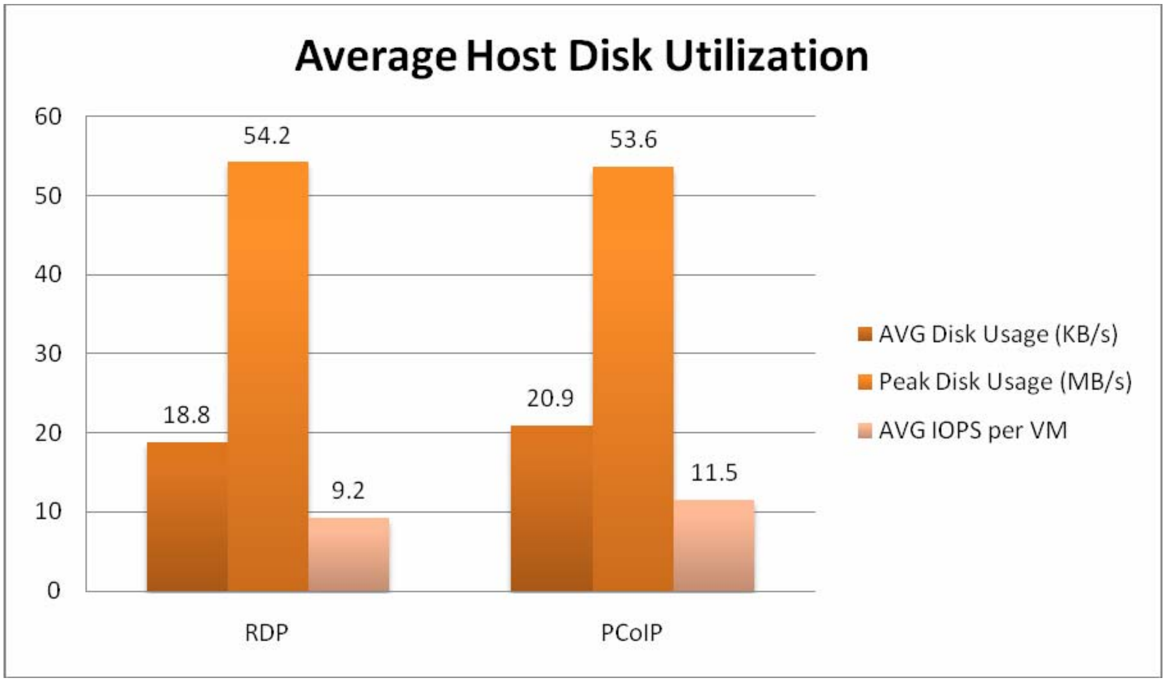


Figure 23 - Knowledge Worker – Protocol Disk Usage Comparison

Analyzing the disk utilization of both protocols shows an increase in the average disk throughput requirements while using the PCoIP protocol. Based on Figure 23, the PCoIP test shows an average increase of disk usage of 11 percent or 2.1MB/s on the host. This translates to an increase of 19KB/s per virtual desktop, and a total average of 179KB/s per virtual desktop while using the PCoIP display protocol. The RDP protocol required an average of 160KB/s per virtual desktop for the Knowledge Worker workload. The average IOPS per virtual machine increased by 2.3 percent for the PCoIP protocol.

When analyzing the peak-to-average disk utilization, the PCoIP protocol had a 32.7MB/s peak-over-average disk utilization for the Knowledge Worker workload. In contrast, the RDP peak-over-average disk utilization was 35.4MB/s. This represents a 2.7MB/s higher peak-to-average disk utilization rate for the RDP protocol over PCoIP.

With such a wide variance in disk workload being generated throughout multiple phases of testing, it is difficult to determine an accurate range of increase in disk throughput for the PCoIP protocol. The Task Worker workload appears to have the greatest variation in disk activity from test to test. This is partially attributed to the fact that the average disk demand per virtual desktop is very low. As a result, any random variation in the timing of the disk utilization in the test would cause the average disk utilization metrics to change considerably. Due to the higher workload demands of the Knowledge Worker test, the random generated disk demand is more evenly distributed throughout the test. This testing appears to be most representative of the expected increase in disk utilization when implementing the PCoIP protocol.

In an effort to determine the cause for the increased disk utilization, research into how the PCoIP protocol functions determined that the protocol operates only in RAM, and does not cache screen frames to disk. Since the protocol operates only in RAM, additional testing has suggested that the default size of the video buffer may be too small to handle the frame caching being performed by the protocol. By increasing the virtual machine's default video memory to 64MB, the average disk utilization of the Knowledge Worker workload dropped to levels on par with the RDP testing. As of the writing of this paper, these findings have not been verified by VMware engineering. Figure 24 shows the results of this testing.

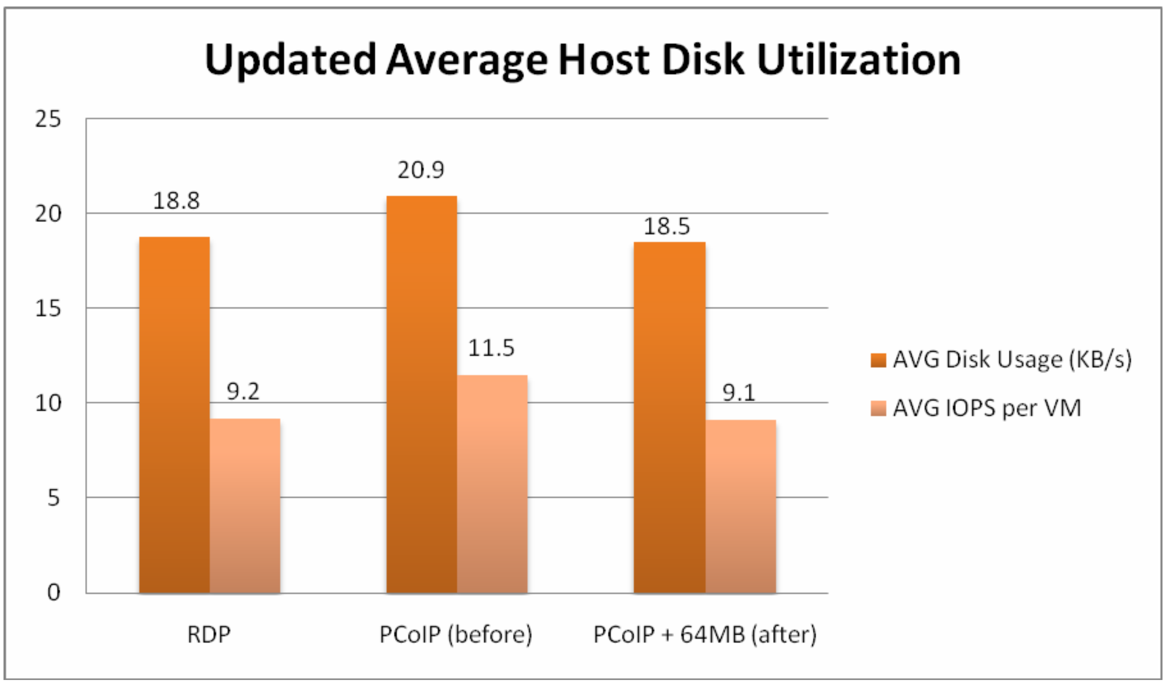


Figure 24 - Knowledge Worker – Updated Host Disk Utilization Comparison

Conclusion

The Task Worker testing demonstrated the ability of the IBM HS22 with vSphere 4 to virtualize 200 desktops onto 8 physical cores, or 25 desktops per core. On a standard 14-slot IBM BladeCenter, this equates to running 2,800 concurrent virtual desktops within a single chassis. The Knowledge Worker testing, using the RDP protocol, demonstrated the ability to virtualize 150 desktops, or 19 desktops per core per blade server. Moving to the PCoIP display protocol, the blade server was capable of 120 simultaneous desktop sessions, or 15 desktops per core. On the same standard 14-slot BladeCenter, this equates to 1,680 or 2,100 concurrent virtual desktops per chassis, depending upon the display protocol selected.

VMware indicates that optimizations in the next release of View will somewhat reduce the additional CPU overhead seen in this study. Future testing will be performed to re-validate the performance results of the PCoIP display protocol with the upcoming release.

Beyond virtual machine core density, the testing provided key performance metrics to aid in the planning and sizing of a VDI environment. Both user workload and display protocol were varied to determine their respective effects on server utilization. These results were analyzed to create usable statistics at both the host and individual virtual desktop level. While these performance metrics are useful for sizing a VDI environment on today's latest server technology, they should never replace a full capacity assessment when more accurate sizing numbers are required. Just as every physical desktop environment is unique, so too is the workload that users generate. Translating the physical desktop into a virtual, hosted desktop environment with its custom applications and power users is increasingly difficult to model without real-world performance data. In addition to the data provided, this paper hopes to stress the importance of properly categorizing users and their respective workloads in any desktop environment, so that a more accurate sizing model can be reached.

About Mainline

This document was produced by Mainline Information Systems™ with assistance from co-sponsors VMware and IBM. Mainline is an award-winning, nation-wide systems integrator headquartered in Tallahassee, Florida, with sales offices located in the US, Canada, Mexico, and Brazil. Mainline's customers include all sizes and verticals – from Fortune 500 to mid-market, government to education, financial services to healthcare. Because of Mainline's customer focus, we maintain a customer satisfaction rating of over 98 percent. As an IBM Premier Business Partner and VMware's first National Premier Partner, Mainline is an expert in the design and delivery of x86 server and storage solutions leveraging the latest in virtualization technology. With its proven, real-world expertise, Mainline was the first business partner to complete all VMware Solution Competencies as well as all IBM Dynamic Infrastructure Specialties. In addition, as a recognized authority in server and desktop virtualization, Mainline has claimed numerous awards for its leadership, including the 2009 VMware Global and North America Desktop Virtualization Partner of the Year Awards, the 2008 VMware Technical Achievement Award, and the 2008 Wyse Visionary Partner of the Year Award, among others. With over 100 VMware-certified Sales and Technical Professionals, as well as a dedicated desktop virtualization team, Mainline has conducted over 300 pilots and proofs of concept, and has deployed over 30,000 VDI seats in over 150 production deployments.

About the Author

Cameron Fore is a Senior Systems Engineer at Mainline Information Systems. As a VMware VAC VCP, Cameron specializes in the design and implementation of virtual server and desktop solutions. He has performed numerous VMware server and desktop capacity studies and implementations across multiple x86 hardware and storage platforms. Cameron has a BS in Industrial Engineering from North Carolina State University in Raleigh, North Carolina, and currently resides in Colorado with his wife and son.

References

VMware View 4 Data Sheet
http://www.vmware.com/files/pdf/VMware-View-4-DS-EN.pdf

Lab Manager Data Sheet
http://www.vmware.com/files/pdf/VMware-vCenter-Lab-Manager-4-DS-EN.pdf

Configuration Maximums for VMware vSphere 4.0 Update 1
http://www.vmware.com/pdf/vsphere4/r40/vsp_40_config_max.pdf

Workload Considerations White Paper
http://www.vmware.com/files/pdf/VMware-WP-WorkloadConsiderations-WP-EN.pdf

IBM HS22 Product Guide
ftp://public.dhe.ibm.com/common/ssi/pm/rg/n/blo03029usen/BLO03029USEN.PDF