

Streaming and Virtual Hosted Desktop Study: Phase 2

Our current findings indicate that streaming provides better server loading and a more positive overall user experience with new Intel® Xeon® processors.

Executive Overview

Intel IT conducted an internal technology evaluation to determine how using the most up-to-date hardware and software, such as Intel® Xeon® processor X5570 and Microsoft Windows 7*, affects server and network utilization as well as the performance of streamed and virtual hosted desktop (VHD) compute models. This study followed a previous investigation, published in 2008, that compared streaming and VHD.¹ Our current findings indicate that streaming provides better server loading and a more positive overall user experience with new Intel® Xeon® processors.

Major findings for Phase 2 include:

- For basic office productivity applications, systems based on dual-core processors with streaming provided a 26 percent better WorldBench 5* performance benchmark score than VHD.
- Streaming server utilization was consistently low. Streaming used about 1 percent of the processor while VHD used from 10 to 70 percent or more for up to 40 PCs.
- A richer graphical user interface (GUI) with more features in Microsoft Windows 7 contributed to higher cumulative network traffic for both streaming and VHD, up to a 57 percent increase in traffic for 20 users.
- Using WorldBench 5 tests as the primary indicator, local computing using the latest technology provided the best user experience.

Our findings indicate that increasingly complex user workloads make it challenging to

measure, compare, and predict server loading. Beyond CPU usage, additional performance considerations now include the I/O subsystem and disk. New server optimization technologies, such as hyperthreading, increase performance but add to the complexity of tuning VHD environments. As user workloads evolve, optimization methods become more difficult and require constant attention.

Although the VHD WorldBench 5 scores have improved to be more comparable with streaming, this still does not mean thin PCs are an appropriate choice. In addition to looking at performance, before choosing a compute model, we also consider mobility requirements, flexibility to adapt to evolving workloads, and our ability to adequately support the solution. Consistent with the results from the first study, we found that mobile business PCs provide the best flexibility. Streaming remains more appropriate for graphics, multimedia, animation, and real-time collaboration applications, while VHD can be acceptable for basic office and data entry tasks.

Catherine Spence
Enterprise Architect, Intel IT

Craig Pierce
Systems Engineer, Intel Architecture Group

Christian Black
Systems Engineer, Intel IT

Chuck Brown
Director, Emerging Compute Lab
Intel Architecture Group

¹ "Streaming and Virtual Hosted Desktop Study." Intel Corporation, January 2008.

Contents

Executive Overview.....	1
Background.....	2
Server Utilization, Network Traffic, and Performance Study.....	2
Technical Architecture.....	3
Test Harness.....	4
Server Processor Utilization Results.....	4
Network Utilization Results.....	5
User Experience Results.....	6
Next Steps.....	8
Conclusion.....	8
Acronyms.....	8

IT@INTEL

IT@Intel is a resource that enables IT professionals, managers, and executives to engage with peers in the Intel IT organization—and with thousands of other industry IT leaders—so you can gain insights into the tools, methods, strategies, and best practices that are proving most successful in addressing today's tough IT challenges. Visit us today at www.intel.com/IT or contact your local Intel representative if you'd like to learn more.

BACKGROUND

Intel IT conducted lab tests to compare server processor and network utilization for streaming and virtual hosted desktop (VHD) computing models, publishing the results in 2008.² We concluded that streaming is more appropriate for graphics, multimedia, animation, and real-time collaboration applications, while VHD is suitable for standard workloads with static screens, such as data entry tasks.

Since that study, more advanced hardware and software have become available, and we wanted to determine if using Intel® Xeon® processor X5570, Microsoft Windows 7*, and enhancements to PCs with Intel® vPro™ technology would change how streaming and VHD compute models performed. We devised a second study to answer this question.

Both studies compared two compute models in an enterprise environment:

- **Streamed computing.** With streaming, a server delivers OS and applications over the network for temporary, local execution by PCs. OS streaming involves creating and storing a disk image on a server and loading it on the PC over the network at boot time.
- **VHD.** In the VHD model, the desktop environment runs within a virtual machine (VM) on a server. The server distributes the user interface to the PC hardware using Remote Desktop Protocol (RDP) or Independent Computing Architecture* (ICA). All processing occurs on the server within the VMs.

² Ibid.

SERVER UTILIZATION, NETWORK TRAFFIC, AND PERFORMANCE STUDY

As in our first study, to evaluate the impact of newer hardware and software on these compute models, we constructed a technical evaluation to characterize backend utilization under a typical user workload. The load included standard office productivity applications. We focused on the impact to backend resources by capturing server and network metrics as the load was scaled from one to 40 simultaneous PCs. To measure performance and user experience, we ran industry-standard PC benchmarks across streaming, VHD, and local PC environments.

In addition to upgrading the processors and OS, we also performed incremental upgrades to existing software applications as well as some optimization to make the compute environment more comparable to our real-world environment. We did not perform application streaming³ or cold runs,⁴ as we had sufficiently characterized these in previous studies.

We evaluated three software delivery configurations:

- **Streamed OS and applications.**⁵ We created an image of the office productivity applications typically installed on the OS. The server streamed the image to traditional PCs.

³ Ibid.

⁴ "Streaming and Virtual-Hosted Desktop Study" and "Improving Manageability with OS Streaming in Training Rooms." Intel Corporation, December 2008.

⁵ In our first study, we referred to applications bundled with the OS as "embedded applications."

Table 1. Test Hardware and Software

Hardware and Software	Specifications
Servers	<ul style="list-style-type: none"> Citrix Provisioning Server 5.1*: 4x Intel® Xeon® processor X5570 with simultaneous multithreading (SMT); 8 cores total; 12 GB RAM; 2.9 GHz; 3x 500 GB RAID 5 Microsoft Hyper-V Server 2008 R2*: 4x Intel Xeon processor X5570 with SMT; 8 cores total; 48 GB RAM; 2.9 GHz; 8x 160 GB RAID 5 Virtualization Host: 1x Intel® Core™2 Duo processor E6850; 2 cores; 4 GB RAM; 3.0 GHz; 1x 500 GB SATA
PCs Based on Single-core Processors for Streaming and Virtual Hosted Desktop (VHD)	<ul style="list-style-type: none"> 20x 1-GB Intel® Pentium® M processor; 1.70 GHz; 40 GB IDE
PCs Based on Dual-core Processors for Streaming and VHD	<ul style="list-style-type: none"> 20x 1-GB Intel® Pentium® M processor; 1.83 GHz; 80 GB SATA
Network	<ul style="list-style-type: none"> 1 gigabit per second (Gb/s) wired network in the lab Streamed OS tests configured to multicast the image using User Datagram Protocol (UDP) when applicable
OS Streaming Software	<ul style="list-style-type: none"> Citrix Provisioning Server 5.1
Virtualization Software	<ul style="list-style-type: none"> Microsoft Hyper-V Server 2008 R2 Citrix XenDesktop Enterprise 3.0* and Citrix XenServer 5.5* Screen Resolution 1024 x 768
Client OS*	<ul style="list-style-type: none"> Microsoft Windows 7* (for PCMark05* and IT Workload tests) Microsoft Windows XP Pro SP3 32-bit* (PCMark05 and WorldBench 5* tests)^f
PCs for Benchmark Tests	<ul style="list-style-type: none"> PCMark05 <ul style="list-style-type: none"> Dual-core: 1x 2-GB Intel® Core™2 Duo processor E6750; 2.66 GHz; 80 GB SATA Quad-core: 1x with 4 GB Intel® Intel® Core™2 Quad processor Q9650; 3.0 GHz; 300 GB SATA Netbook: 1x 1-GB Intel® Atom™ processor N270; 1.6 GHz; 80 GB SATA WorldBench 5 <ul style="list-style-type: none"> Dual-core: 1x with 2 GB Intel Core 2 Duo processor E6750; 2.66 GHz; 124 GB SATA Quad-core: 1x with 4 GB[§] Intel Core 2 Quad processor Q9650; 3.0 GHz; 300 GB SATA
Test Applications	<ul style="list-style-type: none"> Microsoft Office 2003* (Microsoft Word*, Microsoft Excel*, and Microsoft PowerPoint*) with updates Microsoft Office 2007* for Citrix XenServer tests
Data Capture Software	<ul style="list-style-type: none"> Performance monitor for Microsoft Windows* and Microsoft Windows Server 2003* XenTop* for Citrix XenServer 5.5

* Client OS includes patches, fixes, security updates, and antivirus software.

^f At the time of evaluation, PCMark05 and WorldBench 5, standard industry benchmarking tools, were the latest versions available. In addition, WorldBench 5 did not run on Microsoft Windows 7.

[§] Only 3 GB usable.

- **VHD including OS and applications.** For each PC, we created a VM on the server that ran the OS with installed applications. The server distributed the user interface using RDP or ICA.
- **Traditional PC with local OS and applications.** We configured traditional PCs with the OS and office productivity applications installed locally on the hard disk drive (HDD).

We used several more client groups than in the first study, including streaming and VHD clients based on single- and dual-core processors, traditional PCs based on dual- and quad-core processors, and a netbook PC.

We originally intended to evaluate Citrix High Definition User Experience* (HDX) as a new media display protocol; however, at the time of this study HDX support with Microsoft Windows 7 was not available. We intend to evaluate HDX and PC over IP* (PCoIP) in a future study.

Technical Architecture

In our lab environment, all software delivery configurations used the same server, client, and network hardware to help ensure a consistent infrastructure for the study. Table 1 lists the hardware and software used. Figure 1 illustrates the study infrastructure. We configured all VMs alike, with one CPU, 1 GB of memory, and

25 GB of storage. We selected 1 GB memory to overcome constraints, such as page faults, memory stacking, and so on, associated with 256- or 512-MB VMs. Consistent with our first study, we configured the network for 1-GB connections from the server to provide ample bandwidth to enable evaluation of utilization in a relatively unconstrained network. The configuration that we used is in line with a real-world IT department. Additional optimization could be added, such as storage area network (SAN) storage and multiple network uplinks to increase speed, but this would also add more cost for components and additional management complexity. Table 2 lists the server metrics we collected.

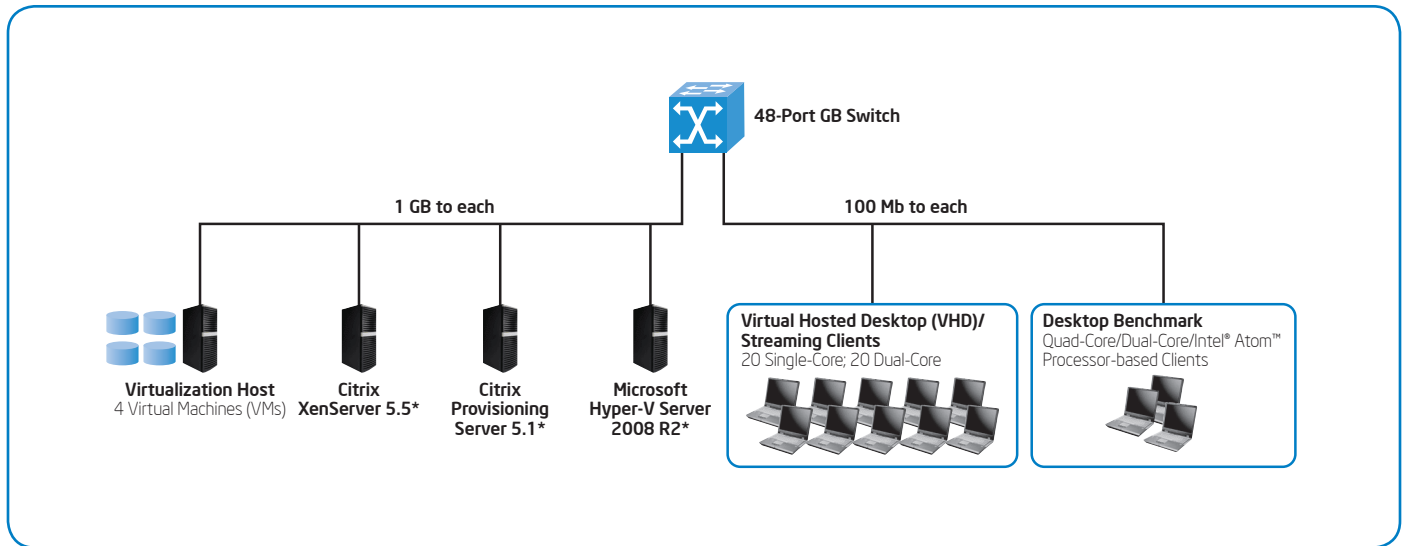


Figure 1. Conceptual architecture.

Table 2. Server Metrics

Category	Utility	Metrics
Processor	Perfmon*	▪ Processor(_Total) \ % Processor Time
	Xentop*	▪ 20-second snapshots of % CPU
Network Interface	Perfmon	▪ Network Interface(Intel* 82576EB Gigabit Dual Port Network Connection _2) \ Bytes Sent/sec
		▪ Network Interface(Intel 82576EB Gigabit Dual Port Network Connection _2) \ Bytes Received/sec

Test Harness

We created a test harness to automate benchmarking and to provide a consistent office script, written in Visual Basic, across the configurations. The script executed several typical user tasks in the same sequence for the tests. The script executed as follows:

1. Start script; 20- to 30-second pauses between steps
2. Open Microsoft Word*, type text, insert two pictures, type text, save, close
3. Open Microsoft Excel*, copy chart, open Microsoft Word, paste chart twice, save, close
4. Open Microsoft PowerPoint*, add two slides with text and chart, save, close
5. Open Microsoft Word*, type text, insert two pictures, type text, save, close
6. Delete script-created files, end script

The tests executed the script on groups of 10 PCs, 20 PCs, and 40 PCs. For multiple clients, the scripts launched 15 seconds apart to stagger the workload. We added wait times to the script to slow down execution and simulate a human pace. This approach also enabled us to separate results data by task, creating identifiable utilization peaks.

To help ensure that we captured enough data for proper analysis, we executed the tests multiple times and averaged three statistically significant test runs. We placed time boxes around the peak utilization curves as a means of capturing the steady state condition. We then derived average server and network utilization for each test run for selected metrics.

In the streaming model, we configured the OS and applications to cache locally in memory and write back to the server for disk activity

and file saves. Once loaded, applications resided in memory for subsequent use. For this study, we ran all tests as “warm” runs; the OS and applications already resided in memory in the PC.

Server Processor Utilization Results

To measure server processor utilization, we evaluated three scenarios:

- Streamed OS and applications, using Citrix Provisioning Server 5.1*
- VHDs including OS and applications, using Microsoft Hyper-V Server 2008 R2*
- VHDs including OS and applications, using Citrix XenServer 5.5*

Server utilization was calculated as the average utilization across all cores. Results for all three scenarios are shown in Figure 2.

STREAMING CPU UTILIZATION

Server processor utilization was very efficient for streaming. Even with 40 PCs, CPU usage was only about 1 percent, which is consistent with the results from the first study. Extrapolating these results, we estimate that the streaming model could support more than 350 clients. For the streaming model, script runtime was the same regardless of scale, illustrating the positive impact and effectiveness of local caching.

VIRTUAL HOSTED DESKTOP CPU UTILIZATION

For Microsoft Hyper-V Server 2008 R2, server utilization was higher with the VHD model than with streaming. However, this represents a significant improvement in VHD server utilization using the newer Intel Xeon processor X5570. Under the workload tested, these newer processors can support more users than in our previous study. The number of supported users per server will vary dramatically, based on e-mail, Web browsers, and media workloads, in addition to the office script.

In the first study, the 20-VM workload used 60 percent of the processor, compared to only about 4 percent in the current study. Even with 40 VMs, CPU usage was just over 10 percent.

Workload runtimes increased with greater numbers of VMs: 13 minutes for 10 VMs, 17 minutes for 20 VMs, and 21 minutes for 40 VMs. This is an indication that as the number of users increases, performance for existing users will be affected. When sizing the number of concurrent users on a server, it may be appropriate to determine an acceptable level of script runtime increase. In this evaluation, the script runtime increased by 62 percent. Additional cores and memory may be required to support additional concurrent users.

Maximum users per server varies, based on office, e-mail, browser, and media workloads.

Figure 2 shows that CPU utilization on Citrix XenServer 5.5 increased from 12 percent for 10 VMs to 29 percent for 20 VMs to 68 percent for 40 VMs. Script runtimes ranged from 8.6 minutes for 10 VMs to 9.3 minutes for 20 VMs to 10.4 minutes for 40 VMs. However, these test runs deviated from the prior two scenarios in two important ways. First, we upgraded from Microsoft Office 2003* to Microsoft Office 2007* when we discovered that our script was causing a fast screen refresh, which increased server utilization to nearly 100 percent

and extended script times to over half an hour. Second, we also discovered an I/O bottleneck and addressed it by changing the drives from 7K RPM SATA to 10K RPM SCSI RAID 5 eight-drive array. We may achieve further optimization using SAN storage, but this was beyond the scope of our study.

In addition to CPU utilization, when sizing the server for the maximum number of users or VMs, script runtime could be an indicator of user experience. Our guideline is to refrain from adding more users if runtimes exceed a set limit or increase to some percentage, such as 50 percent, above the single-user runtime for concurrent users. Beyond this limit, users may not tolerate the performance degradation, although additional users could be accommodated if they are only connected occasionally.

Network Utilization Results

To test network utilization, we evaluated two scenarios:

- Streamed OS and applications, using Citrix Provisioning Server 5.1
- VHDs including OS and applications, using Microsoft Hyper-V Server 2008 R2

We measured and averaged the bytes sent and bytes received using the same time box calculated in the server processor utilization tests.

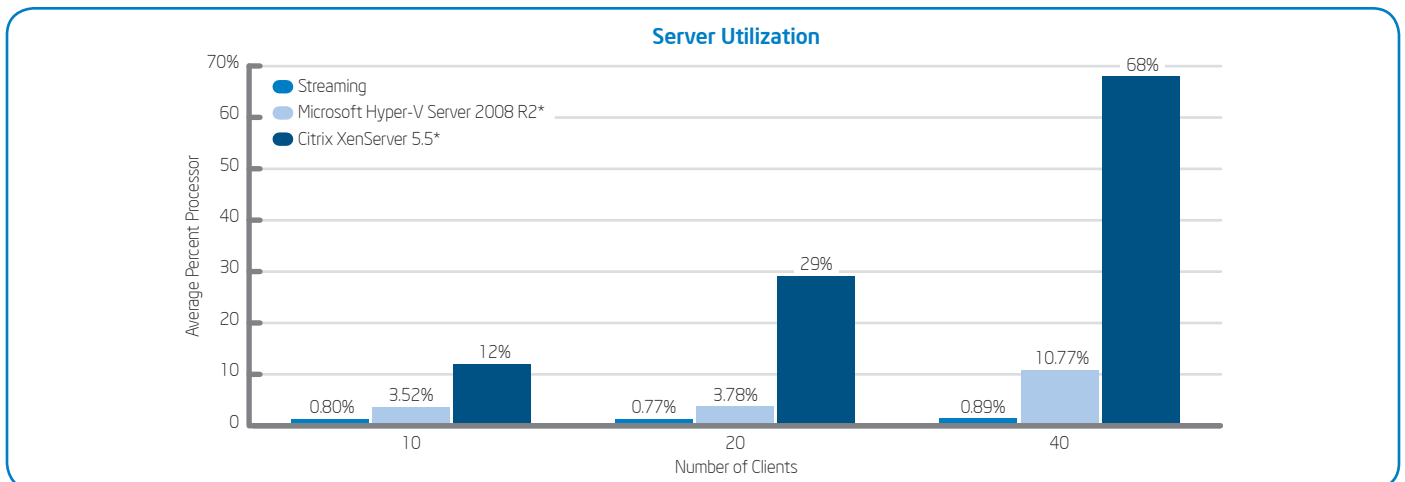


Figure 2. Server processor utilization for streaming, virtual hosted desktop (VHD) running on Microsoft Hyper-V Server 2008 R2*, and VHD running on Citrix XenServer 5.5*.

Overall, OS streaming and VHD both generated slightly more network traffic than in the first study, despite the identical network configurations in both studies. We attributed this to improved features in Microsoft Windows 7 as compared to Microsoft Windows XP Pro, including a feature-rich GUI.

STREAMING NETWORK TRAFFIC RESULTS

The average number of megabits per second (Mb/s) sent increased compared to the first study, from 1 Mb/s to 36 Mb/s, whereas the average number of Mb/s received decreased, from 15 to about 1 Mb/s. As Figure 3 shows, the average network traffic for 10 users was 20 Mb/s, or about 2 Mb/s per user; average network traffic for 20 users was 38 Mb/s, or about 1.9 Mb/s; and for 40 users, average network traffic was 50 Mb/s, or 1.25 Mb/s per user. While these numbers show an increase in cumulative traffic, they are within range of expectation compared to our first study. Future architectural evaluations may be warranted to understand variations of network traffic at the packet level.

VIRTUAL HOSTED DESKTOP NETWORK TRAFFIC RESULTS

As with streaming, VHD results reflected an increase in network traffic compared to the

first study. The average number of Mb/s sent for 20 users increased from 5 in first study to 19 Mb/s, but the average number of Mb/s received remained the same, at 1 Mb/s. As Figure 4 shows, the average network traffic for 10 users was 10 Mb/s, or about 1 Mb/s per user; for 20 users, the average network traffic was also 1 Mb/s per user. For 40 users, the average network traffic dropped to 30 Mb/s, or 0.75 Mb/s per user. In line with our first study, steady state network traffic for VHD was slightly less.

User Experience Results

To measure user experience, we ran two industry-standard benchmark tests—PCMark05* and WorldBench 5*—in three scenarios:

- Streamed OS with applications installed in the streamed image.
- VHDs including OS and applications, running on Citrix XenServer 5.5 and Microsoft Hyper-V Server 2008 R2.
- Traditional PCs, with local OS and applications, based on single-core, dual-core, and quad-core processors.

We ran these scenarios on multiple platforms:

- PCMark05 on Microsoft Windows 7
- WorldBench 5 on Microsoft Windows XP Pro⁶

⁶ WorldBench 5 did not run on Microsoft Windows 7.

PCMark05 results were scored by test grouping; we scored the WorldBench 5 results by specific application set.

PCMARK05*

Figure 5 shows PCMark05 test results for system, CPU, memory, graphics, and HDD performance for the traditional PC, streaming, and VHD computing models.

As in our first study, we ran PCMark05 in a single VHD session with no additional load on the server—the best case scenario in terms of running PCMark05 on a high performance server.

VHD showed a distinct improvement in PCMark05 CPU and memory scores compared to the first study. We attribute this mainly to the higher performance of the server’s Intel Xeon processor X5570. However, VHD failed on system and graphics tests due to difficulties with application sets and graphics running in VMs, in line with our first study. We attributed failures on other platforms to hardware constraints such as a lack of a CD drive or I/O unit. These failures were expected and not relevant to the outcome of this study.

Streaming to PCs based on dual-core processors had the best system and graphics scores. The Microsoft Hyper-V Server 2008 R2 had the best CPU and memory scores, which is indicative of the higher performance of the Intel Xeon processor X5570.

WORLDBENCH 5*

Figure 6 shows those WorldBench 5 scores focused on standard office and application sets. The traditional PC and streamed client scenarios performed best in all tests. Overall, streaming had better scores than VHD. As was expected, PCs using the most advanced technologies had better scores.

The VHD test was the best-case scenario, as there was only a single user on the Intel Xeon processor X5570-based server—we did not measure the benchmark with the server running multiple workloads. We would expect the scores to continue to increase (slow down) as additional users are added. In addition, the VHD tests failed at media encoding.

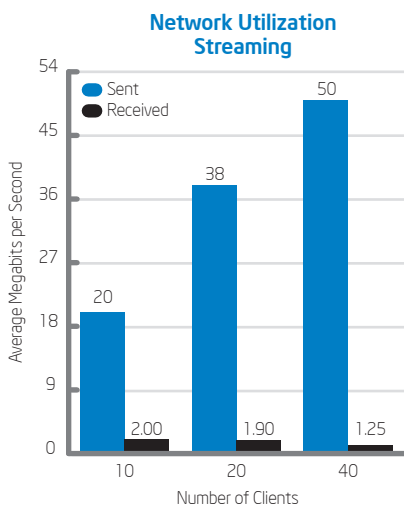


Figure 3. Network utilization for streaming.

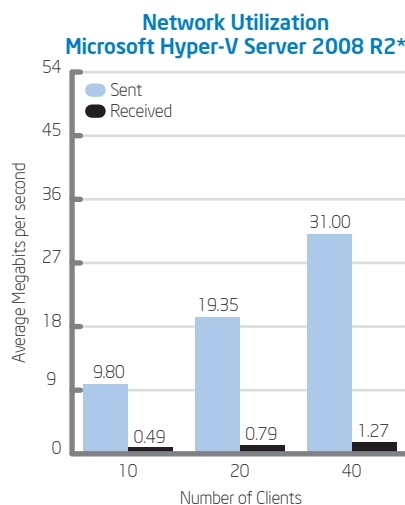


Figure 4. Network utilization for virtual hosted desktop (VHD).

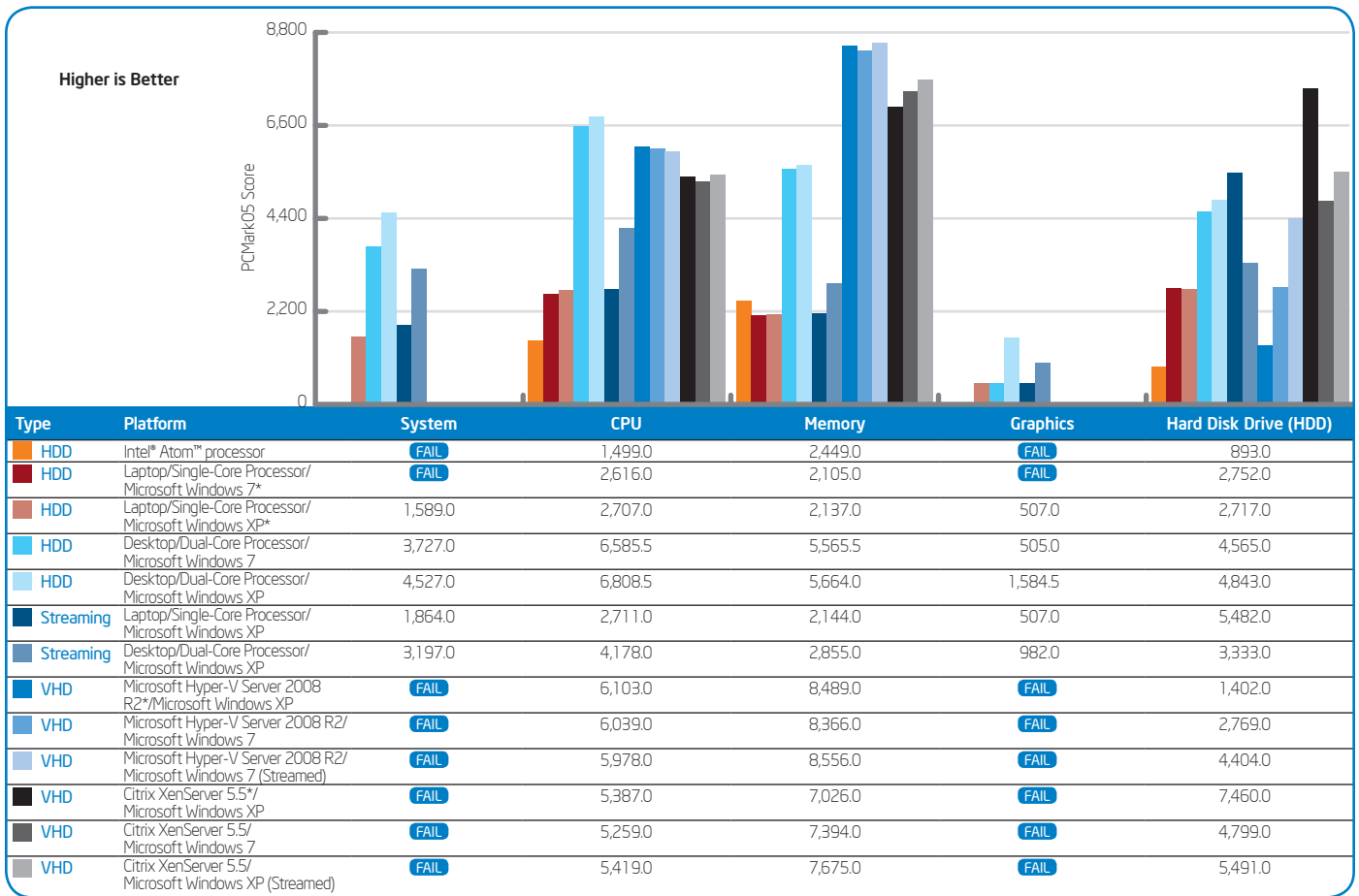


Figure 5. PCMark05* results for streaming, virtual hosted desktop (VHD), and traditional PCs.

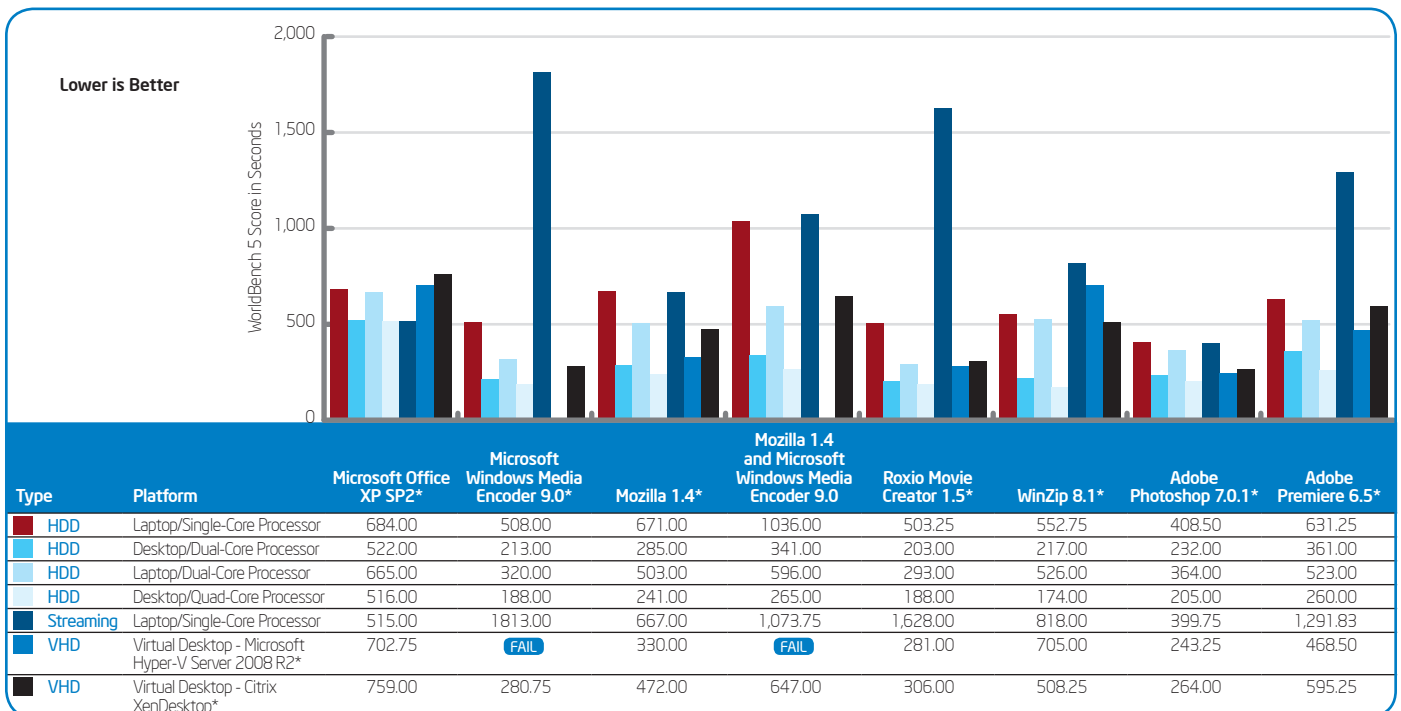


Figure 6. WorldBench 5* scores for streaming, virtual hosted desktop (VHD), and traditional PCs. Tests conducted by Intel using WorldBench 5. WorldBench is a trademark of International Data Group, Inc. Test results have not been verified by PC World, and neither PC World nor International Data Group, Inc. makes any representations or warranties as to the accuracy of the test results.

Table 3. Comparison of 20-Client Solution Results from First and Second Studies

	Server Processor Utilization	Network Utilization		Approximate Single-Server Scaling
		Sent	Received	
VHD Including OS and Applications Phase 1	45%	15 Mb/s	1 Mb/s	35 Clients
VHD Including OS and Applications Phase 2	4% to 30%	19 Mb/s	1 Mb/s	40 to 80+ <i>Varies based on workload and media</i>
Streamed OS and Applications Phase 1	1%	5 Mb/s	2 Mb/s	150+ Clients
Streamed OS and Applications Phase 2	1%	37 Mb/s	1 Mb/s	350+ Clients

VHD – Virtual Hosted Desktop; Mb/s – Megabits per Second

The Microsoft Office 2003 test most closely aligns to typical IT workloads. For this test, streaming had the best (lowest) score, and the dual-core desktop scored 34 percent better than VHD with only one user per server. The improved performance of the desktop systems compared with laptops is attributed to the difference in the HDD and other system components. The latest local PC technology provided the best user performance.

Next Steps

We hope to conduct further studies to evaluate emerging graphics display protocols such as HDX and PCoIP along with additional optimized systems configurations, such as storage subsystems and networks, to further characterize and improve performance, understanding that these components add additional cost.

CONCLUSION

Based on this study of streaming and VHD compute models, local compute models consistently provide better user experience. It is also apparent that upgrading to higher performance hardware and software can dramatically improve the user's experience for a

standardized IT workload. However, a growing number of configuration options make measuring and predicting performance more difficult than before. New server optimization technologies, such as hyperthreading, increase performance but add to the complexity of tuning VHD environments. As user workloads change over time, optimization methods become more difficult and require constant attention.

As Table 3 shows, one of the most significant changes between the two studies is that in some cases, VHD server processor utilization improved. Also, network traffic increased for streaming and VHD.

Although some results differed between the first and second studies, our main finding from the first study is still valid: Streaming with local compute is better for multimedia applications. The results of this second study also clearly indicate that PCs with advanced, high-performing processors contribute to better performance and a more robust user experience.

Due to the richer feature set of Microsoft Windows 7, there is a tradeoff between an improved GUI and increased network traffic. Also, workload runtime performance is dependent on the underlying virtualization

scheme, which can affect user experience as the type of workloads and number of users vary.

As technology improvements and configuration options multiply, making performance even harder to judge, it will be imperative to monitor and tune our virtual environments and choose the compute model that best fits IT and user needs. Future studies will include new optimizations and technologies such as HDX and PCoIP.

ACRONYMS

Gb/s	gigabits per second
GUI	graphical user interface
HDD	hard disk drive
HDX	High Definition User Experience
ICA	Independent Computing Architecture
Mb/s	megabits per second
PCoIP	PC over IP
RDP	Remote Desktop Protocol
SAN	storage area network
SMT	simultaneous multithreading
UDP	User Datagram Protocol
VHD	virtual hosted desktop
VM	virtual machine

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, reference www.intel.com/performance/resources/benchmark_limitations.htm or call (U.S.) 1-800-628-8686 or 1-916-356-3104.

This paper is for informational purposes only. THIS DOCUMENT IS PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE


ARISING OUT OF ANY PROPOSAL, SPECIFICATION OR SAMPLE. Intel disclaims all liability, including liability for infringement of any proprietary rights, relating to use of information in this specification. No license, express or implied, by estoppel or otherwise, to any intellectual property rights is granted herein.

Intel, the Intel logo, Intel Atom, Pentium, and Xeon are trademarks of Intel Corporation in the U.S. and other countries.

* Other names and brands may be claimed as the property of others.

Copyright © 2010 Intel Corporation. All rights reserved.

Printed in USA
0410/JLG/KC/PDF

 Please Recycle
322457-001US

