SEE THE DIFFERENCE FOR YOURSELF

How to set up your own Windows 10 VDI test
environment with NVIDIA Virtual
GPU Solutions

nVIDIA®

1 IDC Technology Spotlight. The Need for Graphics Virtualization in Modern End-User Computing Environments. Dec 2016. #US42065016

2 Lakeside Software White Paper. Elevating User Experience Through GPU Acceleration: A Windows 10 versus Windows 7 Analysis. 2017

3 Microsoft. "Overview of the upcoming changes to Office 365 ProPlus update management." https://support.office.com/en-us/article/Overview-of-the-upcoming-changes-to-Office-365-ProPlus-update-management-78b33779-9356-4cdf-9d2c-08350ef05cca (accessed July 2017)

Introduction

The graphics accelerated VDI environment has clear advantages over a traditional VDI environment that utilizes the CPU only. Traditional VDI environments often experience poor performance, reduced feature sets on applications or even the inability to run the applications themselves, and costly scaling as your organization grows. This leads to unhappy users and poor adoption of VDI. According to IDC, "Organizations moving to Windows 10 for virtualization need to consider leveraging GPUs for all workloads, not just traditionally graphics-intensive applications such as 3D and computer-aided design."[1]

UNDERSTAND WHAT'S HAPPENING IN YOUR ENVIRONMENT

It is broadly known that GPU-accelerated computing offers faster performance across traditionally graphics-intensive applications such as 3D and computer-aided design. But today we are seeing GPU-acceleration impact everyday business applications. Windows 10 has the highest graphics requirement of any operating system to date, with 32% increase in CPU consumption compared to Windows 7.[2] Consequently, the versions of applications used in Windows 10 also have higher graphics requirements than their counterparts in Windows 7.

Even if you can get by with the minimum specs today, Microsoft will be gradually upgrading Windows 10 and Office 365, adding new features via regular updates, causing the minimum specs to grow over time. Windows 10 and Office 365 ProPlus will have semi-annual feature updates, with Office 365 ProPlus having additional security and quality updates monthly.[3] For most enterprises, best practice is to be on the latest versions that Microsoft provides so they can benefit from new innovations, features, and security capabilities. However, each update to the OS and to Office suite requires more processing power, which highlights the pressing need for a GPU. Unlike prior releases, Microsoft is making it more difficult to turn off features, making it even more critical to have a GPU to support these changes.

SEE THE DIFFERENCE FOR YOURSELF

If you are unsure of the benefits, or want to make sure you are not overlooking critical success criteria in evaluating whether this solution is right for you, you can easily conduct your own side-by-side testing. There is nothing like seeing the difference for yourself, in a test environment that is representative of what your users do on a day-to-day basis. This white paper provides you with everything you need to know to setup your test environment – from setting requirements and configuration, to determining use cases and success criteria, to developing a test plan and analyzing your results.

How to get started:

Before we get started, here is a checklist of what you will need in your lab to conduct the testing.

HARDWARE	COMMENTS
Server	2 identical servers for desktop or target VMs
	2 identical servers for virtual client or launcher VMs. To enable NVENC, the virtual client servers need NVIDIA GPUs.
GPUs	2 Tesla M10 GPU accelerators. Each Tesla M10 GPU accelerator supports 32 VMs with NVENC or 64 VMs without NVENC
Storage	Any flash accelerated storage solution
Network	The server NICs should be 10GBPS for enterprise customers. For the VMs, choose the hypervisor recommended para-virtualized NICs (eg. Vmxnet3 for VMware vSphere)
Access Devices (Thin Client/Monitor/Mouse/ Keyboard)	Two identical clients with the same resolution
SOFTWARE	
Hypervisor	VMware vSphere, Citrix XenServer
NVIDIA Virtual GPU Software	If you are evaluating on GRID 5.0 you will need a license server. You can get some licenses for evaluation purposes from our **license portal**
APPLICATION	
LoginVSI	Two instances to create background noise for the GPU and CPU-only VDI setups
Other	Applications installed and their corresponding licenses for the number of users to be simulated

Setting the baseline assumptions

Before we go into the testing, let's set some baseline assumptions. Determining functional requirements early on will help you setup your environment more quickly. Understanding user experience requirements and setting quantifiable metrics will help you determine whether VDI with graphics acceleration will be of positive impact to your organization.

FUNCTIONAL REQUIREMENTS:

REQUIREMENT	DESCRIPTION	COMMENTS
Applications	Applications that will be run on all of your desktops. Check LoginVSI website for the list of applications	You might need to obtain licenses for certain software for both your simulated loaded environment as well as your actual testers
Workload	This assumes the type of workload(s) your virtual simulated users will run. Make sure to run the same workloads on both setups	You can use preconfigured workloads that come with the LoginVSI, mix and match workloads, or customize them.
Number of users	Load the server with as many users as needed to get up to the threshold of 60-80% host CPU utilization.	If you want to use VMware Blast Extreme with NVENC you need minimum of a 1GB (M10-1B) vGPU profile.

USER EXPERIENCE REQUIREMENTS:

In this testing, we are using LoginVSI only as a load generator because we see the user experience as more representative than raw performance metrics. As such we will not use LoginVSI metrics like VSIMax and VSIbase. After all, users often experience slow performance despite favorable LoginVSI performance metrics. Thus, in order to really understand the impact of GPU on the Windows 10 VDI experience, we need to hone in on the end-user evaluation of their experience measured with quantifiable metrics.

You can use a numbered range for subjective scoring by your users but also make sure to take note of the standard deviation between responses. For an example, see the sample test plan under the section "Developing Your Test Plan".

QUANTIFIABLE METRICS:

REQUIREMENT	DESCRIPTION	COMMENTS
Click to Photon	Measures the time of the mouse click until the action is visible to the user. In remote environments, this also includes encode latency, network latency and decode latency. This is different than LoginVSI's overall latency metric. You will only be able to measure this if you have the Click-to-Photon setup.	Lower is better. The latency should be as close to your physical PC/Workstation as possible

Percentage decrease in CPU Utilization	CPU utilization will go down with the introduction of a GPU, thus allowing your users to get a better experience and also allowing you to increase your server density	As a general best practice, enterprise deployments typically set host CPU utilization at 60-75%. 75% is the maximum where the host is not saturated and will not be contributing to performance degradation. Seeing lower CPU utilization than this with similar experience means you are benefiting from your GPU deployment
GPU and encoder utilization	GPU utilization shows that it was able to offload some tasks from the CPU	Offloading tasks from the CPU to the GPU means that you can support more users and users are getting a more responsive experience, including application features that may have been deprecated in a constrained, CPU-only environment.

Setup:

1. Using your existing infrastructure cluster, setup two identical server VMs that run the LoginVSI launchers for GPU enabled and CPU only environments. Note: Do not add these VMs to the test cluster.

2. Set up two identical servers for target VM deployments with and without GPUs. The target VM template needs to have all the required applications installed.

3. Set up two identical servers for the launcher VM deployments with and without GPUs. These VMs will be used to connect to the target VMs and would require just the Horizon View Client to be installed. For enabling NVENC, the VMs need to be GPU enabled as well.

4. Once your target VMs and launcher VMs are provisioned, run LoginVSI with the following configurations:

   a. Workload: Select the LoginVSI workload you want for your simulated users – in this case select the **knowledge worker workload**

   b. Number of users: Depending on your test server configuration, find out the number of users that will get you to CPU utilization of the host to be 60-80%. This might require multiple runs with different number of users.

   c. Here are some specific LoginVSI configurations used during NVIDIA testing:

---

1 Lakeside Software White Paper. Elevating User Experience Through GPU Acceleration: A Windows 10 versus Windows 7 Analysis. 2017

2 Microsoft. "Overview of the upcoming changes to Office 365 ProPlus update management." https://support.office.com/en-us/article/Overview-of-the-upcoming-changes-to-Office-365-ProPlus-update-management-78b33779-9356-4cdf-9d2c-08350ef05cca (accessed July 2017)

    i. Enable benchmark mode. To get repeatable and consistent background load. Once this mode is set, you will only be allowed to change the number of users in the LoginVSI Test Setup tab.

    ii. Use Microsoft Office 2016

    iii. Change the default logoff sequence time. Change this through the start test wizard in order to increase the total test time. E.g. 14400 sec will run the test for 4 hours.

5. Set up your access devices (thin client/monitor/mouse/keyboard). Ensure each has a single screen with the same resolution.

6. Develop a test plan with pre-defined scenarios for your testers to follow (see next section for a sample test plan we ran with our testing).

7. Have users test the systems side-by-side. Make sure your testers don't know which session is GPU accelerated - you can do this by ensuring the graphics intensive tasks will be performed towards the end.

Developing Your Test Plan:

Here is a sample test plan that NVIDIA ran with a representative number of employees throughout our organization. It provided instructions to the user and a scoring sheet.
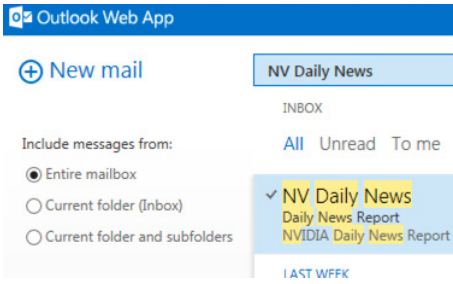
User Experience Test Plan

Purpose: Take solution testers through a prescribed set of office/ knowledge worker tasks, approximating typical desktop/application interactions. This test to be conducted amidst LoginVSI machine-generated workloads co-resident on the same host, at varying levels, in order to detect user-perceivable impact on quality of experience.

Quality of Experience (QoE) Scoring:

1 – unacceptable, unusable – fire someone in IT!

2 – barely useable, borderline, but I'll get tired of this soon

3 – tolerable, I guess I can make do

4 – pretty good for a virtual desktop

5 – outstanding – as good (or almost) as physical

"Without a GPU, the virtual desktop would attempt to render detailed graphics found in a complex PowerPoint transition, Google Earth, streaming video, or CAD drawings using only the virtualized CPU, and it would be impossible to reproduce the smooth flow of the applications and user interfaces found on a physical device."

—Tirias Research. *Virtualized GPUs Power Modern VDI User Experiences.* January 2017

STEP	ACTION/DESCRIPTION	USER NOTES / SCORING
1	Log into virtual desktop <add login instructions>	None
2	After desktop boots, open MSIE, check email. Visit: **https://hqmail.NVIDIA.com/owa/#path=/mail** And login using your normal network credentials	None
3	Find any email in your inbox that includes a "long" thread. Scroll all the way through that thread – try a couple times with 2-3 different emails to observe how smoothly the text scrolls.	QoE Score:
4	Use Outlook search to locate all "NV Daily News" emails. Locate the oldest edition of that email. Scroll to very bottom of email, and read whatever story appears there.	QoE Score:

STEP	ACTION/DESCRIPTION	USER NOTES / SCORING
5	Open a new tab in Internet Explorer. Visit **http://msn.com**. Scroll to the bottom of that page. Try your browser's keyboard shortcut for zooming, and see how the page responds.	QoE Score:
6	In new tab, visit **http://youtube.com**. Find a movie trailer from the main page, and watch it for 15-30 seconds.	QoE Score:
7	In new tab, visit **http://finance.yahoo.com/q?s=NVDA**	QoE Score:
8	Click on the stock price tracker on the right to expand the graph. Hover mouse over the graph, moving from left to right. Watch open/close/etc. stats update on the left as you move the mouse	QoE Score:

STEP	ACTION/DESCRIPTION	USER NOTES / SCORING
9	Right-click/copy the stock tracker/graph to clipboard	None

10	Go to Windows Start button, type Powerpoint in search and launch Powerpoint	QoE Score:
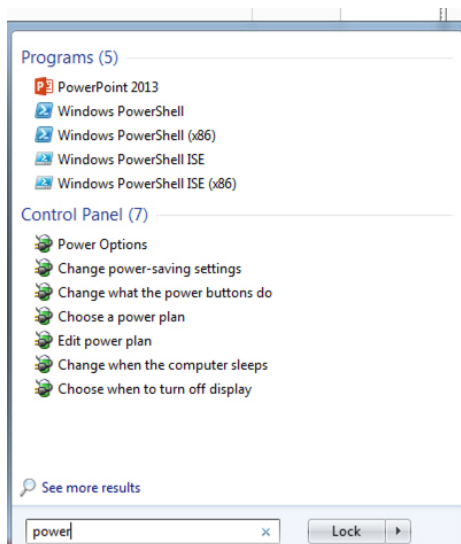
11	Open a blank presentation/slide, clear any title/subtitle placeholder	None
12	Paste NVDA tracker from clipboard to slide	QoE Score:
13	Resize chart to occupy 50% of slide, align to right side of slide	None
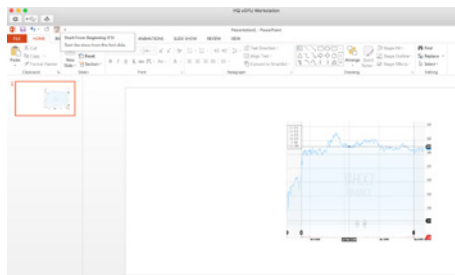
14	Find and open the NV Daily News email (any). Highlight and copy 8-10 bullets of text from the top of the email (exec summary section), copy to clipboard	None
15	Paste from clipboard to PPT slide. Resize text box to fit on left side of slide, next to graph.	QoE Score:

| 16	Save slide to desktop, giving it the name "NVDA.pptx"	None
17	Return to Outlook, start a new email. Attach NVDA.pptx to email, and send to **xxxxxxx@nvidia.com** and **xxxxxxx@nvidia.com** with subject "M10 UX"	QoE Score:

| 18 | In new tab, visit: **http://Helloracer.com/webgl**, Click once on the screen to start, use arrow keys to accelerate/brake/steer. Play for 15-30 seconds. | QoE Score: |

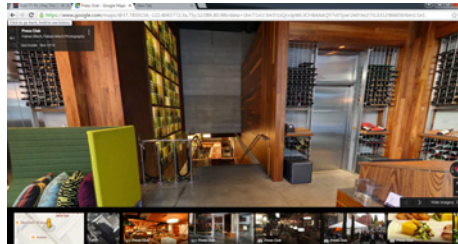| 19 | In new tab, visit: **http://maps.google.com**. Find Press Club wine bar @ 20 Yerba Buena Ln SF. Click on the photo to open street view. Go inside, pan 360 degress, smoothly. | QoE Score: |
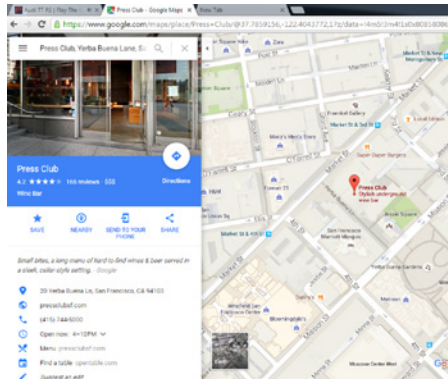
Street view:

Results:

NVIDIA performed the test internally on our own employees and here are the results we found:

HARDWARE AND SOFTWARE SPECIFICATIONS:

SERVER SPECS:	DESCRIPTION	COMMENTS
> Server: DL380 Gen9	> Hypervisor Version: VMware vSphere 6.0.0, 3620759 with Blast Extreme	> GRID version 4.0
> CPU: 2x E5-2697 v4		> NVIDIA vGPU driver version: 367.43
> 18 cores @ 2.30GHz		> NVIDIA VM driver version: 369.17
> Memory: 768 GB	> VM hardware version: 11	
> Storage: All-Flash SAN (XtremeIO)	> OS: Windows 10 64bit	
> Network: 10Gbit	> VDA version: 7.0.2	**Testing device specs:**
> GPUs: 2x Tesla M10	> vCPUs: 2	> Chromebooks, Laptops
	> vRAM: 4GB	> Resolution/Screens: 1x Screen @ 1080p
	> vGPU Profile: M10-1B	

Note: Please use the latest versions of software and applications for optimal performance

Results:

Testing two identical VDI environments, one with GPU acceleration and one without, showed that the environment with GPU acceleration with NVENC provided an average positive increase to user experience of 34% *(see Figure 1)*. Blast Extreme with NVENC decreases latency up to 140ms at <1ms network latency, well below the threshold that is acceptable for gamers who typically have a higher standard for user experience *(see Figure 2)*. Even when the VMs were scaled, the click to photon latency was much improved with GPU acceleration with NVENC *(see Figure 3)*. This shows users can expect consistency in performance despite an increase in demand. We also saw a 17% improvement in CPU utilization confirming that by offloading tasks to the GPU, we were able to increase user density with better performance *(see Figure 4)*.

WINDOWS 10 WITH CPU VS NVIDIA VIRTUAL GPU
NVIDIA virtual GPU provides an average of 34% increase in User Experience
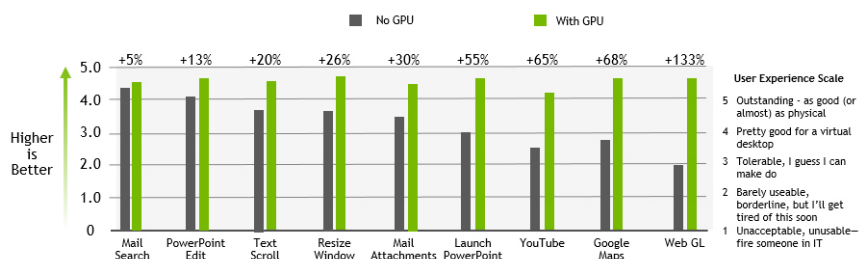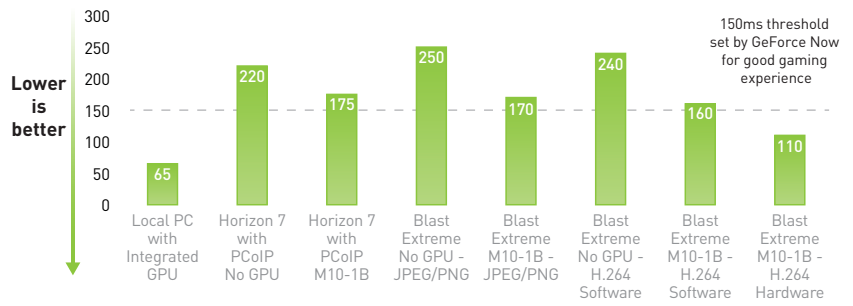

Figure 1. Windows 10 VDI with CPU only vs GPU enabled VDI with NVIDIA virtualization software

CLICK TO PHOTON LATENCY

Blast Extreme with NVENC decreases latency up to 140ms at <1ms network latency

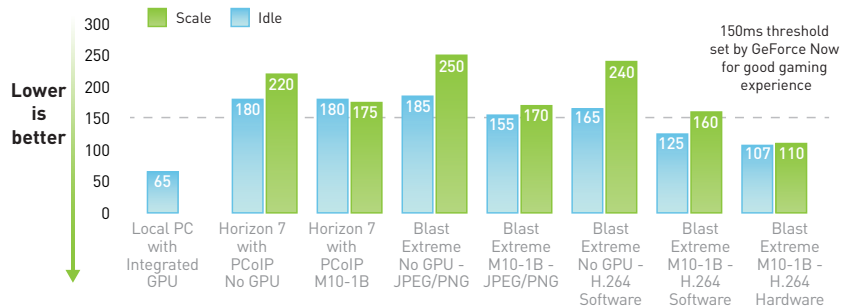150ms threshold set by GeForce Now for good gaming experience

Lower is better

63 x Tesla M10-0B VMs on a host running LoginVSI knowledge worker workload and 1 additional VM measuring latency

Figure 2. Click to Photon Latency under various scenarios with and without a GPU

CLICK TO PHOTON LATENCY AT SCALE

Comparing latency of single VM and at scale at >1ms network latency

150ms threshold set by GeForce Now for good gaming experience

Lower is better

63 x Tesla M10-0B VMs on a host running LoginVSI knowledge worker workload and 1 additional VM measuring latency

Figure 3. Click to Photon Latency at scale under various scenarios with and without a GPU

IMPACT OF CPU UTILIZATION WITH MORE USERS

17% improvement with Horizon 7 with PCoIP

Lower is better

Horizon 7 with PCoIP No GPU

Horizon 7 with PCoIP M10 GPU
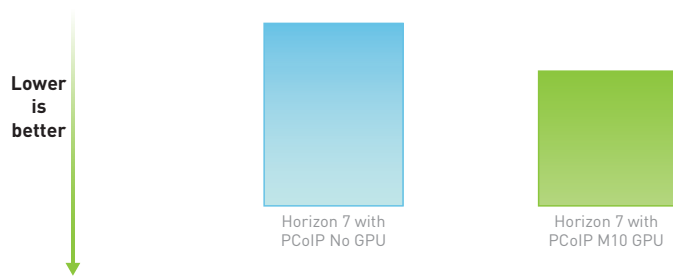
64 x Tesla M10-0B VMs on a host running LoginVSI knowledge worker workload

Figure 4. CPU utilization improvement with NVIDIA virtual GPU solution

Conclusion

VDI with GPU acceleration can provide clear benefits over traditional VDI with CPU only. Users get a better experience even using common business applications and performing day-to-day tasks. With a GPU enabled VDI environment, users get a consistent experience that rivals physical PC, anywhere and on any device they choose. And because CPU resources are freed by offloading to the GPU, IT in turn can maximize infrastructure investment with increased user density without sacrificing performance.

For more information on how to successfully design, build, and execute your NVIDIA virtual GPU solution, refer to our **NVIDIA Virtual GPU Deployment Guide for Citrix XenDesktop** or **NVIDIA Virtual GPU Deployment Guide for VMware Horizon**.