Parallels Remote Application Server

GPU Accelerated Application Publishing

Parallels International GmbH
Vordergasse 59
8200 Schaffhausen
Switzerland
Tel: + 41 52 672 20 30
www.parallels.com

Contents

CHAPTER 1

Introduction

Virtual desktops can offer several advantages over physical desktops, including increased simplicity, flexibility, and scalability. However, virtual desktops may not always provide the same level of performance as physical desktops, especially when it comes to tasks that require GPU (graphics processing unit) acceleration, such as 3D rendering, video editing, and scientific computing. To address this issue, Parallels RAS (Remote Application Server) can leverage GPU-enabled desktops directly through Azure Virtual Desktop (AVD) and Azure Infrastructure as a Service (IaaS) to provide users with a high-performance Remote Desktop experience within Microsoft Azure. This allows the user to experience high-quality graphics and multimedia content, such as video playback, 3D rendering, and image processing in their virtual session.

It's worth noting that the performance of GPU redirection can depend on factors such as the network bandwidth, latency, and the capabilities of the user's local GPU. In some cases, it may be necessary to optimize the virtual machine's GPU settings and the network configuration to ensure optimal performance.

Parallels RAS leverages RAS Templates to deploy and maintain virtual machines in Azure Virtual Desktop (AVD) host pools (https://download.parallels.com/ras/v19/docs/en_US/Parallels-RAS-19-Administrators-Guide/46990.htm) and Azure IaaS (https://download.parallels.com/ras/v19/docs/en_US/Parallels-RAS-19-Administrators-Guide/46786.htm), (https://download.parallels.com/ras/v19/docs/en_US/Parallels-RAS-19-Administrators-Guide/39308.htm). When configuring a RAS Template for Azure Virtual Desktop (AVD) and IaaS, there are two options for deploying the Windows operating system:

- Custom Host
- Azure Gallery

Custom Host, a virtual machine that an administrator has already created and installed Windows on, must be used when creating a template. This is because the GPU drivers need to be installed and available. Administrators can also edit local Group Policy to enable the use of the GPU, or the settings can be delivered through Active Directory Group Policy.

CHAPTER 2

Selecting a Virtual Machine

The first step in creating the Template virtual machine is selecting a GPU-optimized Azure virtual machine size (https://learn.microsoft.com/en-us/azure/virtual-desktop/enable-gpu-acceleration). At the time of writing, there are several sizes available for use with virtual desktops and session hosts.

NVIDIA GRID GPUs:

- NV-series
- NVv3-servers
- NVadsA10 v5-series
- NCasT4_v3-series

Note: The NVIDIA series above comes with a GRID license that supports 25 concurrent sessions on a single Windows multi-session host.

AMD

- NVv4-servies

Note: The AMD series allows unlimited concurrent sessions on a single Windows multi-session host.

Note: Currently, not all regions have GPU resources and not all sizes may be available in regions where there are resources, and even in those regions, they may only be available in certain Availability Zones.

Note: For the overview of the most current list of GPU size options, please refer to https://learn.microsoft.com/en-us/azure/virtual-desktop/enable-gpu-acceleration#select-an-appropriate-gpu-optimized-azure-virtual-machine-size.

Both the NVIDIA and AMD series come with appropriate GPU licensing.

GPU acceleration is only supported in the following versions of Windows (single-session and multi-session versions are fully supported):

- Windows 10 version 1511 or newer
- Windows Server 2016 or newer

For more information on NVIDIA and AMD Azure licensing, see https://learn.microsoft.com/en-us/azure/virtual-desktop/enable-gpu-acceleration.

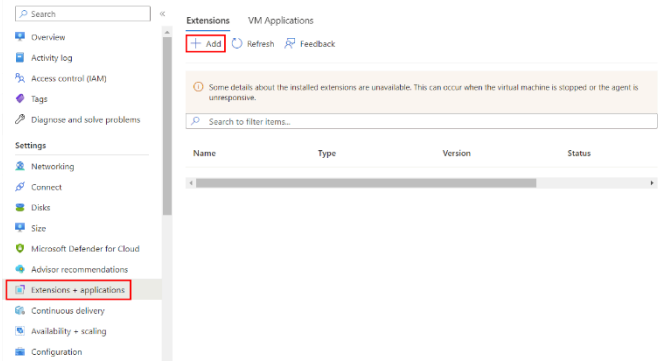
CHAPTER 3

GPU Driver Installation

Now that the virtual machine has been created and Windows has been installed, the appropriate GPU drivers must be installed. Microsoft, NVIDIA, and AMD have made the installation of the respective GPU drivers very simple as they are now Extensions you can add to a virtual machine. Manual installation is still an option, and they also can be installed using Infrastructure as Code (IaC) templates. For more information and requirements, see https://learn.microsoft.com/en-us/azure/virtual-machines/sizes-gpu#supported-operating-systems-and-drivers.

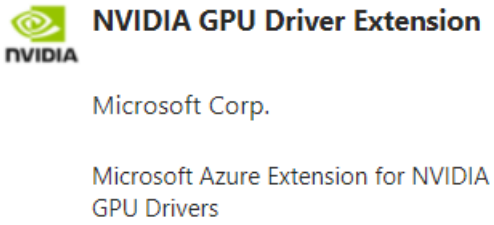
To install the GPU drivers:

1. In the Azure Portal, find the virtual machine created above and click on it. You will see the overview of the virtual machine.
2. Scroll down and click the **Extensions + applications** option.



3. Click **Add** and select the appropriate driver extension. The following extensions are available:
   - AMD GPU Driver Extension

- NVIDIA GPU Driver Extension



4 Click **Next**.

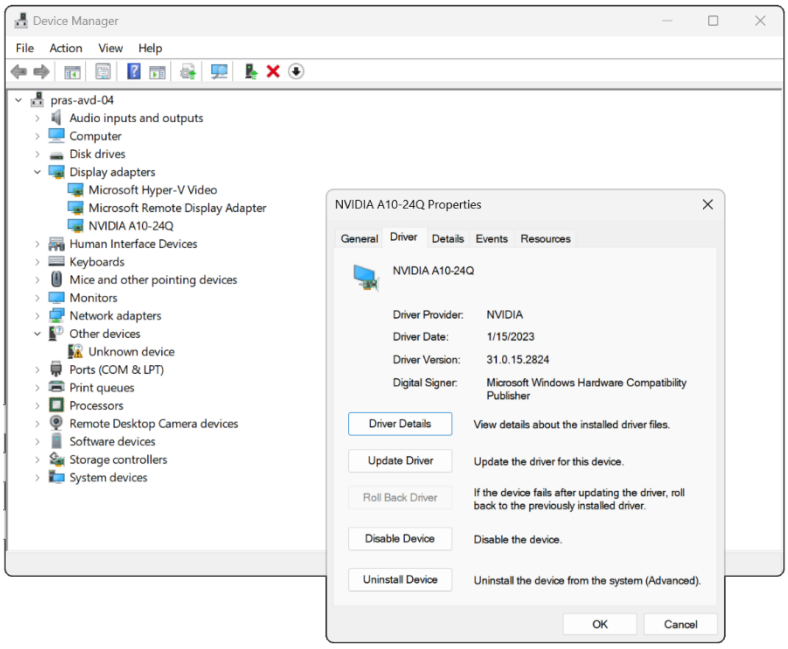
5 Click **Review + create**.

6 Azure will run a validation test. Once you see **Validation Passed**, click **Create**.

7 Azure will then automatically deploy the extension which will install the GPU drivers.

8 When that is completed, reboot the virtual machine.

9 Connect and log into the virtual machine and open Device Manager. You should see the GPU listed under the **Display Adapters** node.



CHAPTER 4

Group Policy Configuration

By default, Windows virtual machines will use the CPU for rendering and encoding. To enable GPU usage on Windows, administrators can either create an Active Directory Group Policy Object to be applied when the virtual machines boots or edit the local Group Policy so that it is baked into the image.

To configure an Active Directory GPO or a local Group Policy:

1. Open the Group Policy editor and navigate to **Computer Configuration > Administrative Templates > Windows Components > Remote Desktop Services > Remote Desktop Session Host > Remote Session Environment**.
2. From there, you must configure these three settings:
   - **Use hardware graphics adapters for all Remote Desktop Services sessions**: **Enabled**

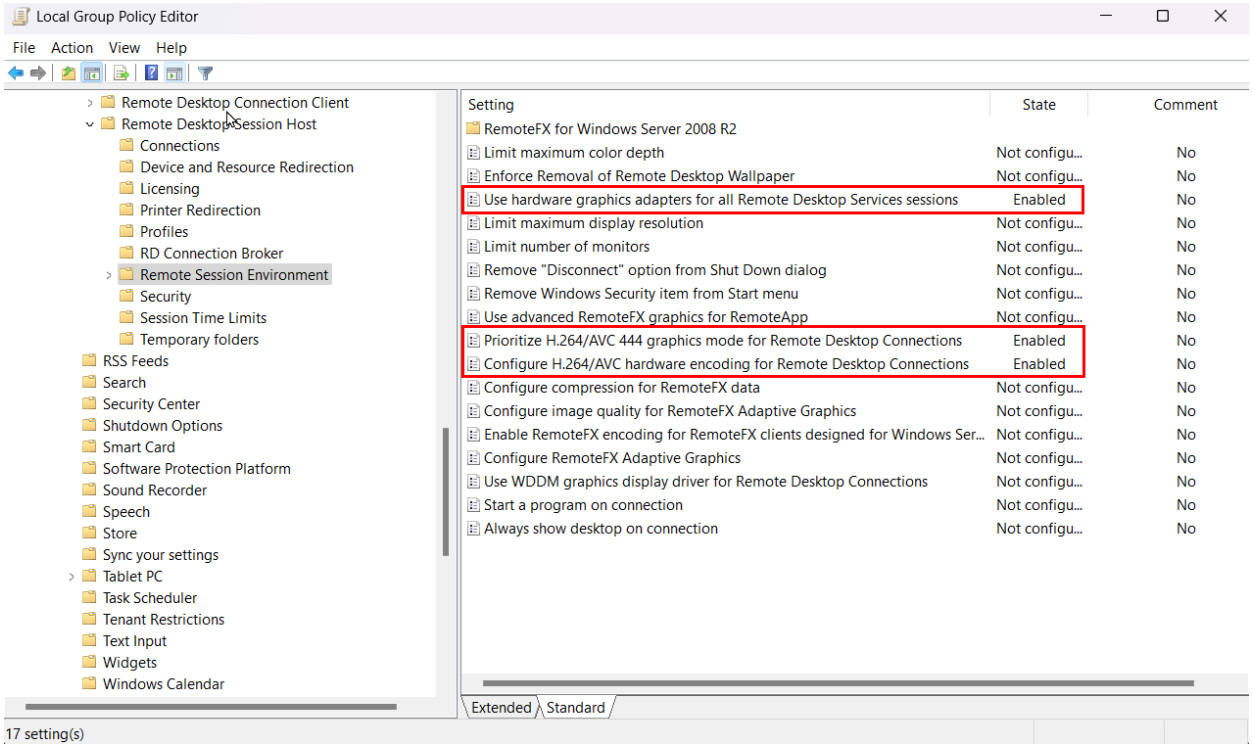
     This option enables Windows to leverage the GPU for remote sessions.
   - **Configure H.264/AVC hardware encoding for Remote Desktop connections**: **Enabled**

     Remote Desktop encodes all graphics rendered by apps and desktops (whether rendered with GPU or with CPU) for transmission to Remote Desktop clients. When a part of the screen is frequently updated, it is encoded with a video codec (H.264/AVC).

     Note: This policy is currently only available for NVIDIA GPUs.
   - **Prioritize H.264/AVC 444 Graphics mode for Remote Desktop connections**: **Enabled**

     If you often use applications that produce high-frame rate content, such as 3D modeling, CAD/CAM and video applications, you may choose to enable a fullscreen video encoding for a remote session.

These two options are for enabling H.264/AVC 444 encoding. This encoding is a video compression standard used to compress and transmit video data over a network. It allows for high-quality video at lower bitrates which, in turn, reduces the amount of data needing to be transmitted.



With that all done, the virtual machine is ready, from a GPU standpoint, to be made into a Parallels RAS Template.

CHAPTER 5

Conclusion

By using GPU acceleration and H.264 encoding, end users will see three main benefits while working on their Azure Virtual Desktop (AVD) or Azure IaaS deployed session:

- Improved graphics and multimedia performance: RemoteFX leverages the GPU to provide improved graphics and multimedia performance, including support for high-resolution displays, 3D graphics, and video playback.
- Reduced network bandwidth requirements: H.264 encoding reduces the amount of data transmitted over the network, which can help reduce network bandwidth requirements and improve the overall performance of remote desktop sessions.
- Lower latency: By leveraging the GPU and H.264 encoding, RemoteFX can provide a lower-latency remote desktop experience, which can help improve user productivity and satisfaction.

Index