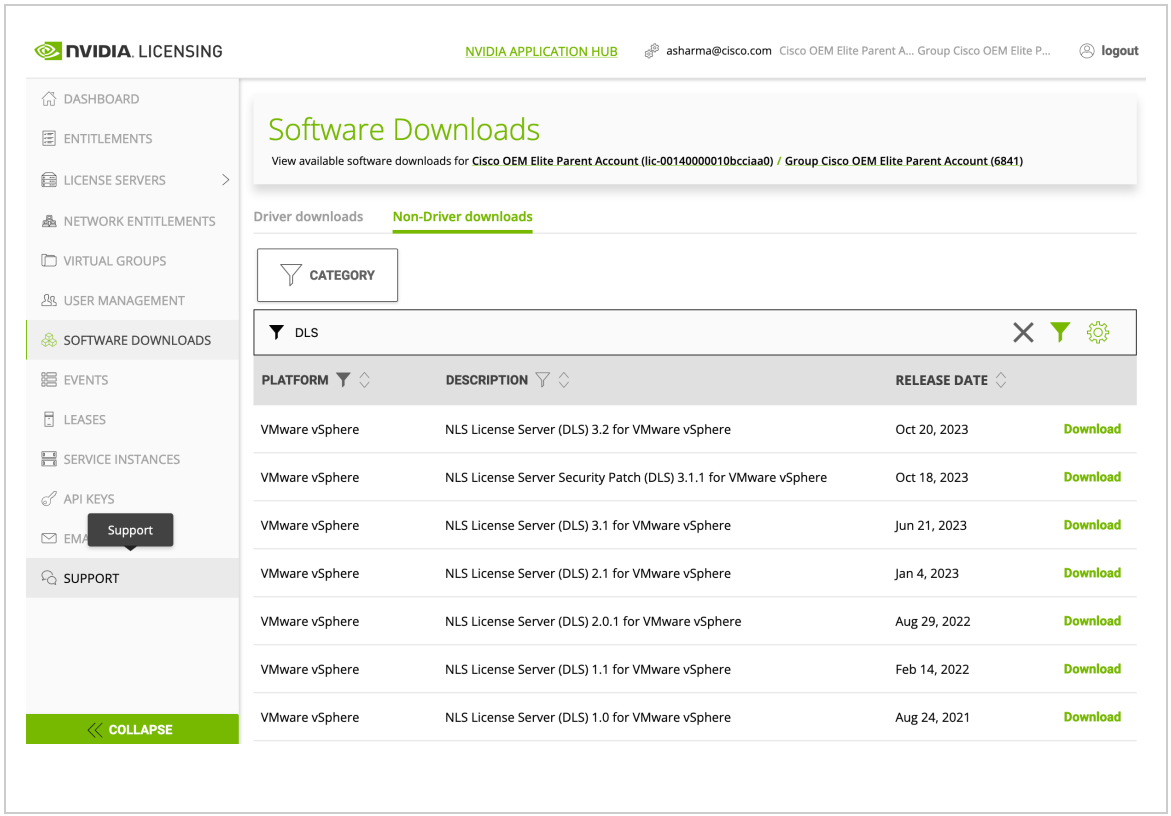
05_CVD: NVAIE Host Driver Prerequisites

Deploy NVIDIA Licensing Server

Deployment Steps:

1. Deploy a NVIDIA Licensing Server using NVIDIA Licensing Portal (NLP). A Delegated License System (DLS) is locally deployed for this solution. The remaining steps will be completed using different sections of the NVIDIA Licensing System User Guide.



2. Deploy a DLS instance in the FlashStack Infrastructure Management network using the steps outlined in the **Installing the DLS Virtual Appliance on VMware vSphere** section. The DLS instance is managed through the NVIDIA Licensing application on the DLS instance and through the NVIDIA Licensing Portal. Go to **DLS Instance Instructions** for the remaining configuration steps.
3. To be able to allocate licenses using DLS, you must create at least one license server on the NVIDIA Licensing Portal to acquire a set of licenses that the on-prem DLS can use for allocating licenses -complete this using the **Creating a License Server on the NVIDIA Licensing Portal** section.
4. To enable the DLS instance to be bound to a license server, you must register the instance with the NVIDIA Licensing Portal. Download a DLS Instance Token and upload it to NLP to register the local instance as outlined in the **Registering an on-Premises DLS Instance with the NVIDIA Licensing Portal** section.
5. To ensure that licenses on the server are available only from that service instance, follow the steps outlined in the **Binding a License Server to a Service Instance** section.
6. After binding a license server to a service instance, you must install the license server on the service instance to make the licenses on the server available to the instance as outlined in the **Installing a License Server on a DLS Instance** section.
7. To allot the licenses on the new server to a license pools, complete the steps outlined in **Creating a License Pool** section.

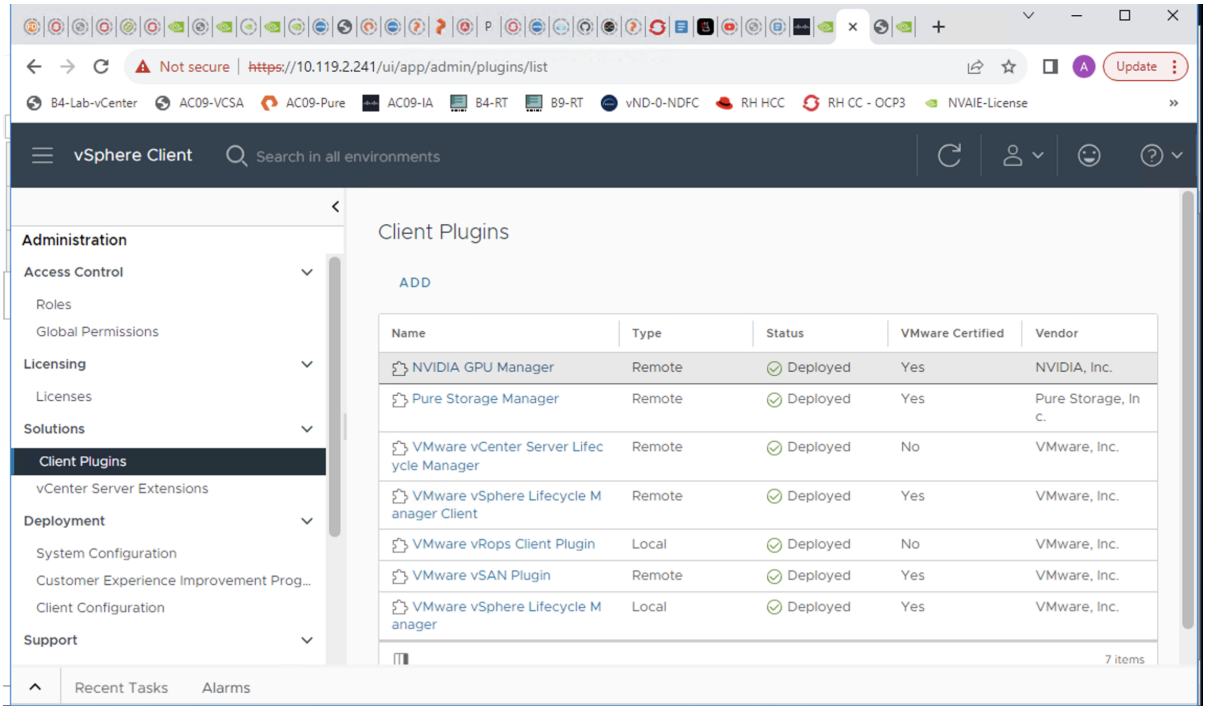
Deploy NVIDIA GPU Manager

NVIDIA GPU Manager for VMware vCenter enables you to manage NVIDIA GPUs from the vSphere client of VMware vCenter Server. You can download and install NVIDIA GPU drivers using a web-based tool integrated with the vSphere web client instead of executing commands on the hypervisor. NVIDIA GPU Manager for VMware vCenter consists of:

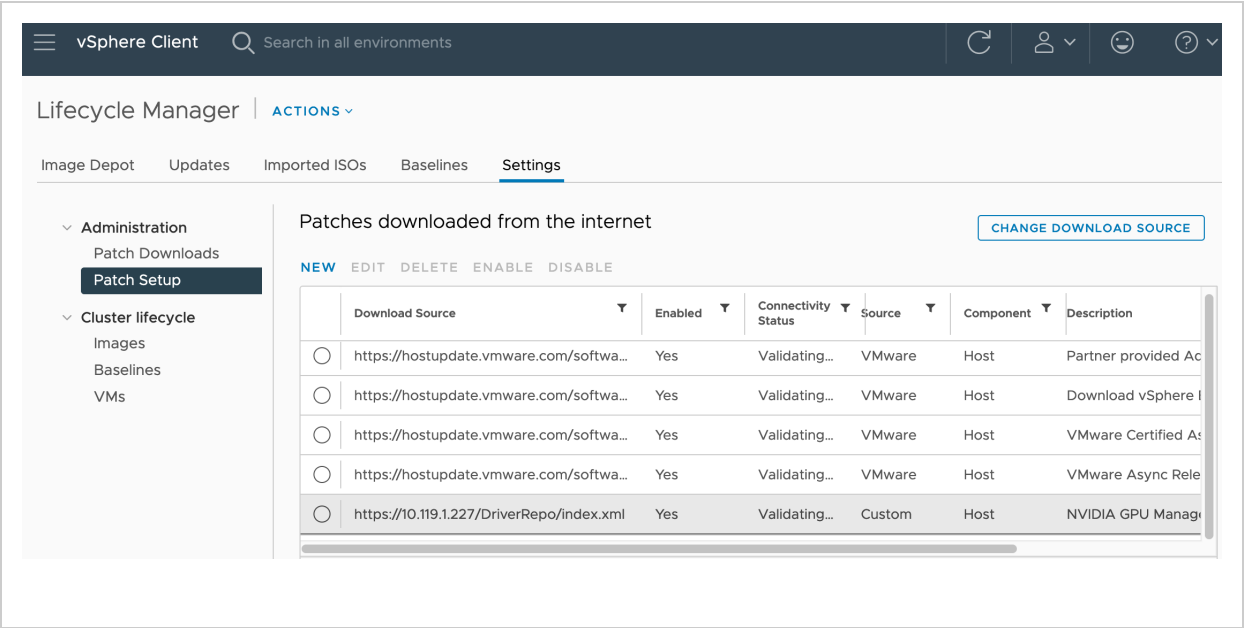
- **NVIDIA GPU Manager for VMware vCenter virtual appliance:** This consists of GPU Manager application responsible for registering GPU Manager with VMware vCenter, Drivers repository for storing downloaded drivers and web-based management interface.
- **vSphere client plugin for VMware vCenter**

Deployment Steps:

1. Deploy NVIDIA GPU Manager for VMware vCenter as outlined here.
2. To register NVIDIA GPU Manager for VMware vCenter Administrator User, use a browser to navigate to NVIDIA GPU Manager VM. On the Register User Account page that opens, provide a password for the NVIDIA GPU Manager for VMware vCenter administrator user **vcp_admin**. Click **Register**. Save the local reset secret safely to reset the password in the future. Click **Continue to Login** and login as **vcp_admin**.
3. To register NVIDIA GPU Manager so that you can download and manage NVIDIA GPU drivers from vSphere client, use a browser to navigate to NVIDIA GPU Manager VM. From the left navigation menu, select **REGISTRATION**. The registration process will install the vSphere plugin for the GPU Manager in VMware vCenter. Provide VMware vCenter information and click on **REGISTER**.Verify the registration was successful by logging into VMware vCenter and verify that you can see the NVIDIA GPU Manager in the Client Plugins. Navigate to **Administration > Solutions > Client Plugins** to see the plugins.



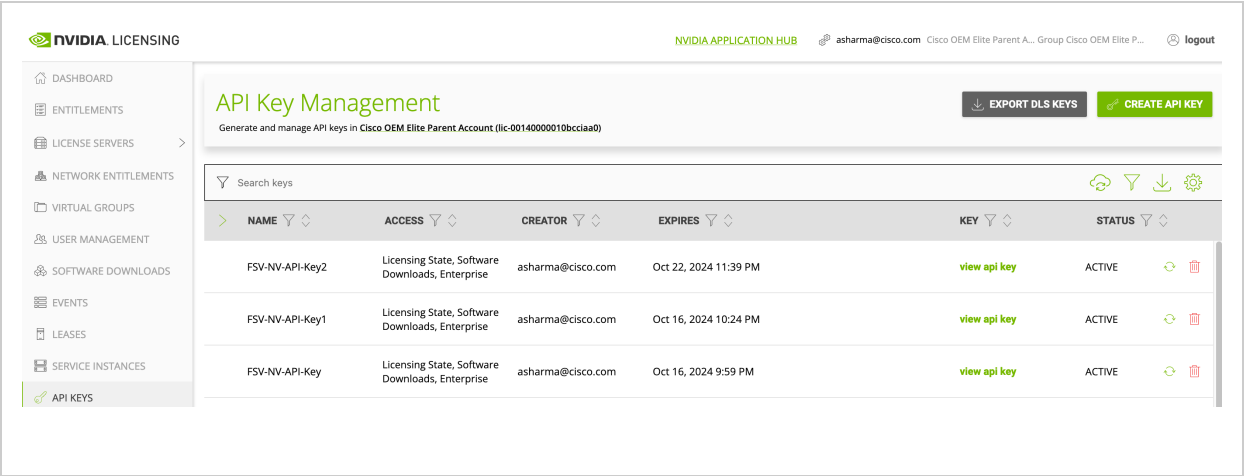
4. Also navigate to vCenter Life Cycle Manager settings and confirm that you can see the drivers repo in the table under **Administration > Patch Setup**.



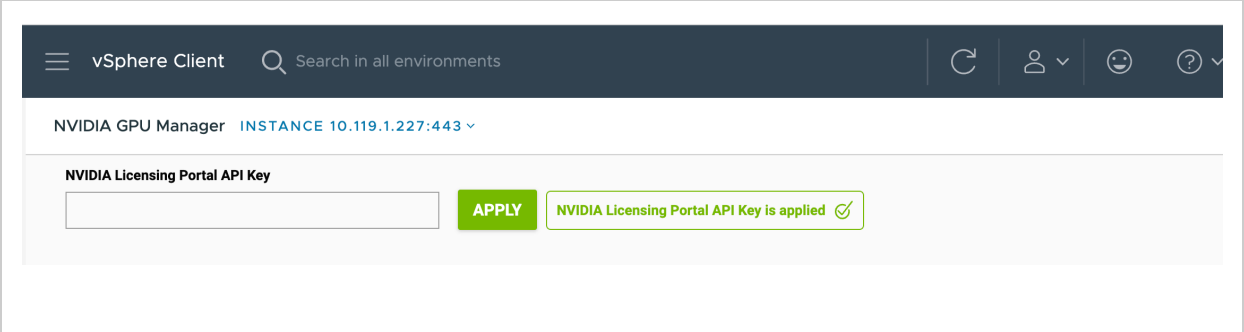
Register GPU Manager with NVIDIA Licensing Portal

To dynamically download NVIDIA Host Drivers to GPU Manager repository from NVIDIA Licensing Portal (NLP) requires a software downloads API key for authentication that must be downloaded and installed on the GPU Manager as outlined in the steps below.

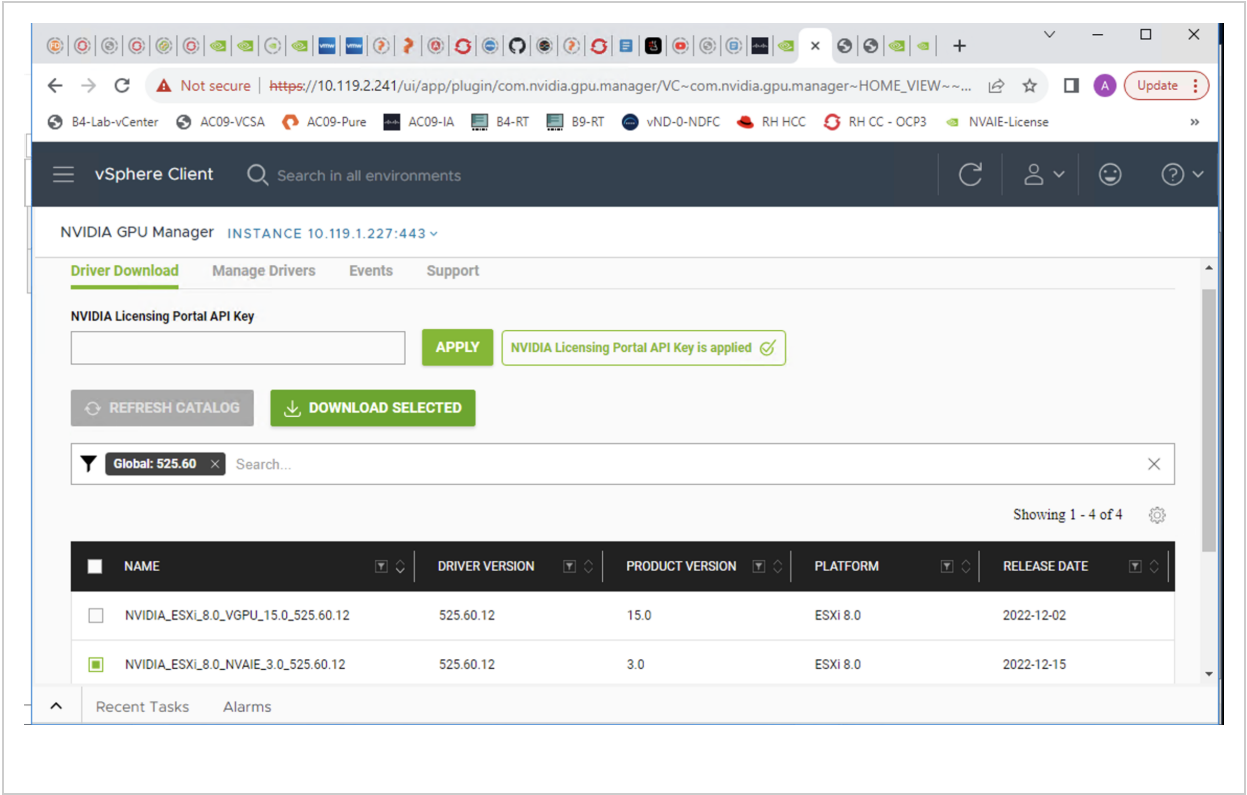
1. Use a browser to navigate and login to NVIDIA Licensing Portal (NLP).
2. In the left navigation pane of the NLP dashboard, click on **API KEYS**.
3. In the API Key Management pop-up window, click on **CREATE API KEY**.
4. In the Create API Key pop-up window, provide the necessary information and click on **CREAT API KEY**.
5. Provide a name for the API Key, specify the Access Type as **Software Downloads** and accept the license agreement.
6. In the **API Key Management** table, select the API key and click on **View API Key** to the right of it.



7. Copy key to clipboard.
8. Navigate to VMware vCenter > GPU Manager and paste the key in the box provided and click Enter. The output should look as shown below.



9. Go to **Driver Download** tab to verify that you can see the host drivers in the GPU manager from NLP. Select a driver to download to GPU Manager repo so that you can use VMware Life Cycle Manager to deploy it.



10. Downloaded drivers can be seen in the **Manage Drivers** tab.