

NVIDIA DGX BasePOD: Multi-cloud Architecture with Amazon Web Services

Deployment Guide

Featuring NVIDIA DGX BasePOD and NVIDIA GPUs through Amazon Web Services

DU-11343-001 V3 2023-07-06

Abstract

As part of the NVIDIA DGX[™] platform, NVIDIA DGX BasePOD[™] provides on-premises infrastructure for artificial intelligence (AI) workloads. This infrastructure is an excellent fit for stable use cases and resource requirements.

However, demands can sometimes outstrip resource availability or users might need access to different resources than those provided by their DGX infrastructure.

Managing a separate pool of resources to support changing requirements typically involves the development of significant expertise in cloud management tools and interfaces. A separate pool of resources often requires user education to request the appropriate system or environment—leading to suboptimal resource utilization and user confusion.

Those scenarios are now resolved through the capabilities of NVIDIA Base Command[™] Manager (BCM) software. Administrators can now integrate on-demand public cloud resources directly with an on-premises DGX BasePOD private cloud environment and make the combined resources available transparently in a multi-cloud architecture.

This document describes how to extend DGX BasePOD with additional NVIDIA GPUs from Amazon Web Services (AWS) and manage the entire infrastructure from a consolidated user interface. Given the breadth of instances offered by AWS for both general-purpose and accelerated computing with NVIDIA GPUs, it is a great option for use as the basis of cloud resource integration in BCM.

Providing concordant access to on-premises and public cloud resources through existing infrastructure drastically simplifies both the administrator and user experience and makes using the right tool for any job easy.

Contents

Chapter 1.	Introduction	1
Chapter 2.	softwareimage and Category Creation	2
Chapter 3.	Cluster Extension Configuration	5
Chapter 4.	Host Preparation After Cluster Extension Configuration	14
Chapter 5.	Power On and Provision the Cloud Nodes	16
Chapter 6.	Deploy Kubernetes	18
Chapter 7.	Create Additional Worker Nodes	
Chapter 8.	(Optional) Enable Jupyter Operator Use in Cloud K8s Cluster	32

Chapter 1. Introduction

Deployment of public cloud integration into DGX BasePOD should be done after onpremises components and services have been deployed, according to the <u>NVIDIA DGX</u> <u>BasePOD Deployment Guide</u>. The tool within BCM that enables this integration is Cluster Extension (cm-cluster-extension). It allows an administrator to integrate a public cloud provider account into an on-premises deployment and configure what resources will be provisioned using that public cloud provider and what regions those resources will be provisioned in. The public cloud resources appear side-by-side with on-premises resources in administrator tools, with access to public cloud-specific configuration capabilities when necessary. A depiction of the multi-cloud usage model is shown in Figure 1.



Figure 1. DGX BasePOD multi-cloud usage diagram

Chapter 2. softwareimage and Category Creation

Before configuring AWS using cm-cluster-extension, create the software images and non-director node categories that are necessary for the target public cloud environment.

- ssh to the head node as root or a user capable of gaining root permissions.
 Specify the external network address or hostname of the head node to gain access.
 # ssh <head-node>
- 2. Enter the cmsh configuration shell.
- # cmsh

Note: This document uses # to indicate commands executed as the root user on a head node and % to indicate commands executed within cmsh. The prompt change is in the preceding block. If it is unclear where a command is being executed, check the prompt that precedes it.

3. Enter the softwareimage menu and create three additional software images as clones of the default-image—one for each of the node types to be provisioned in the public cloud.

```
% softwareimage
% clone default-image cloud-director-image
% clone default-image k8s-cloud-master-image
% clone default-image k8s-cloud-gpu-worker-image
% ..
% commit
```

4. Enter the category menu and create categories for k8s-cloud-master and k8s-cloud-gpu-worker.

The cm-cluster-extension tool automatically creates a category for the cloud-director (softwareimage is configured at a later step).

```
% category
% clone default k8s-cloud-master
% set softwareimage k8s-cloud-master-image
% commit
% clone default k8s-cloud-gpu-worker
% set softwareimage k8s-cloud-gpu-worker-image
% commit
```

5. Augment the disksetup of the new categories as well.

This guide was executed with a disk layout that maximized the root partition size to avoid scenarios where containers quickly fill a smaller partition. Save the following text to /tmp/big-cloud-disk.xml.

```
<diskSetup>
  <device>
    <blockdev>/dev/sda</blockdev>
    <blockdev>/dev/I</blockdev>
    <blockdev>/dev/vda</blockdev>
    <blockdev>/dev/xvda</blockdev>
    <blockdev>/dev/cciss/c0d0</blockdev>
    <blockdev>/dev/nvme0n1</blockdev>
    <blockdev mode="cloud">/dev/sdb</blockdev>
    <blockdev mode="cloud">/dev/hdb</blockdev>
    <blockdev mode="cloud">/dev/vdb</blockdev>
    <blockdev mode="cloud">/dev/xvdb</blockdev>
    <blockdev mode="cloud">/dev/xvdf</blockdev>
    <blockdev mode="cloud">/dev/nvme1n1</blockdev>
    <partition id="a0" partitiontype="esp">
      <size>100M</size>
     <type>linux</type>
     <filesystem>fat</filesystem>
     <mountpoint>/boot/efi</mountpoint>
      <mountOptions>defaults, noatime, nodiratime</mountOptions>
    </partition>
    <partition id="a1">
     <size>max</size>
     <type>linux</type>
     <filesystem>xfs</filesystem>
      <mountpoint>/</mountpoint>
      <mountOptions>defaults, noatime, nodiratime</mountOptions>
    </partition>
    <partition id="a2">
     <size>12G</size>
     <type>linux swap</type>
    </partition>
  </device>
</diskSetup>
```

6. Assign the new disk layout file to the cloud categories in Step 4.

```
$ cmsh
% category
% use k8s-cloud-master
% set disksetup /tmp/big-cloud-disk.xml
% commit
% use k8s-cloud-gpu-worker
% set disksetup /tmp/big-cloud-disk.xml
% commit
```

7. Exit the cmsh configuration shell and update all three images.

```
# cm-chroot-sw-img /cm/images/k8s-cloud-master-image/
# apt update && apt -y upgrade
# exit
# cm-chroot-sw-img /cm/images/k8s-cloud-gpu-worker-image/
# apt update && apt -y upgrade
# exit
# cm-chroot-sw-img /cm/images/cloud-director-image/
# apt update && apt -y upgrade
# exit
```

8. When a terminal menu is displayed to confirm that GRUB does not need to be installed, select Yes to continue.

— Configuring grub-efi-amd64 —

You chose not to install GRUB to any devices. If you continue, the boot loader may not be properly configured, and when this computer next starts up it will use whatever was previously configured. If there is an earlier version of GRUB 2 in the EFI system partition, it may be unable to load modules or handle the current configuration file.

If you are already using a different boot loader and want to carry on doing so, or if this is a special environment where you do not need a boot loader, then you should continue anyway. Otherwise, you should install GRUB somewhere.

Continue without installing GRUB?

<Yes>

<No>

Chapter 3. Cluster Extension Configuration

With images and categories prepared, the environment is now ready for AWS integration and initial configuration. The AWS integration will be accomplished using the cm-cluster-extension command.

1. Create an AWS IAM group with an appropriate policy for a user account to integrate into the BCM on-premises head node.

To create a minimum viable policy set, refer to this Bright Knowledge Base article.

Assign the policy to the target group and provision a new user in that group. Create a new access key and associated secret access key for that user for use with Bright. Securely document the access key and secret access key for use in this section.

- 2. Run the cm-cluster-extension command to get started.
 # cm-cluster-extension
- 3. Choose the AWS extension and then select Ok.

Cluster Extens	ion	
	r	
	AWS	Add/Remove/Configure AWS cluster extension
	Remove everything	Remove all cluster extension components
	Exit	Return to the command line
		< Ok >

1. Choose Add new AWS provider and then select Ok.

ster Extension					
		Mada			
1		Main r	nenu		
	Add now AWS provider	Satur now AWS of	and provider account		
	Unconfigure AWS region	Remove a region	from cluster extension	configuration	
	Test networking	Run AWS network	connectivity test	conriguración	
		<u></u>	<u></u>		
		< 0k >	< Back >		
l l					

2. Enter the required AWS credential information and then select Ok.

Provide	your AWS credenti	als			
"Provid	r name" is a cust	om, human readabl	e, string.		
"Access	Key ID" and "Seci	et Key" will be v	erified in the I	next step by contactir	ng AWS.
Pro	ider name <mark>aws</mark>				
AWS Acco	ss Key ID Secret Key				
		< 0k >	< Back >		

3. Add the provider to the new region by choosing the default setup type and then select 0k.

Cluster Extension	
Choose setup type	
default Default setup wizard	
advanced Advanced setup wizard	
< Ok > < Bac	k > < Help >

4. Enter 4 for the quantity of cloud nodes and then select 0k.

There will be three nodes for the Kubernetes (K8s) control plane and one node as a GPU worker. More nodes can be added later.

Cluster Extension		
	Provide the number of cloud nodes	
	Number of cloud nodes 4	
	< Ok > < Back >	

5. Choose the appropriate geographic region and then select 0k.

Choosing a region near the on-premises cluster typically increases network performance. If the configuration is designed for regional fault tolerance, choose a more distant region. Because not all instances are available in all regions, the type of instance needed should also be considered.

Cluster Extension	
	Please select the geographic region into which you would like to extend your cluster. More regions can always be added and configured later on. all All regions (0/17) ap Asia Pacific (0/6) ca Canada (0/1) eu Europe (0/5) sa South America (0/1) us United States (0/4)
	C Ok > K Back > K Help >

6. Choose a region in the subsequent screen and then select 0k.

us-west-2 is used in this example.

Please select the AWS region(s) into which you would like to extend your cluster. More regions can always be added and configured later on.
[] us-east-1 US East (N. Virginia) [] us-east-2 US East (Ohio) [] us-west-1 US West (N. California) [X] us-west-2 US West (Oregon)
C Ok > K Back > K Help >

7. Choose a default region and then select 0k.

In this example, the only option is us-west-2 because no other regions were configured.

ter Extension	
Cer Excension	
	Extension will be deployed to the following regions. Please select the default region.
	(X) us-west-2
	< 0k > < Back >

8. Choose an availability zone for the public subnet that Cluster Extension will create and then select 0k.

Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges

< Back >

Cluster Extension

Select availability zone for subnet vpc-us-west-2-public in region us-west-2.

us-west-2a was selected in this example.

9. Choose an availability zone for the private subnet that Cluster Extension will create and then select 0k.

< 0k >

us-west-2a was again selected.

Luster Extension Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west-20 us-west-20 us-west-20 us-west-20 	Cluster Exte	insion
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west=20 Us-west=20 Us-west=20 Us-west=20 Us-west=20 Us-west=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-West=20 Us-Wes		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-west-20 US-we		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20 us-west-20		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. US-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-We		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20		
Select availability zone for subnet vpc-us-west-2-private in region us-west-2. Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-West-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-20 Us-2		
Note that choosing different availability zones for your subnets may affect networking speed and incur additional charges. US=west=2a Us=west=2b us=west=2c us=west=2d		Select availability zone for subnet vpc-us-west-z-private in region us-west-z.
US-west-2a US-west-2b US-west-2c US-west-2d K Back >		Note that choosing different availability zones for your submats may affect networking speed and incur additional charges
Us-west-2a Us-west-2b Us-west-2c Us-west-2d < OK > K Back >		Note that choosing different availability iones for your subjects may affect hetworking speed and incur additional charges.
Us-west-20 Us-west-20 Us-west-20 Us-west-20 Us-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-west-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-West-20 Vs-We		
us-west-2b us-west-2c us-west-2d < Ok >	U	us-west-2a
us-west-20 us-west-2d < Ok > K Back >	I	us-west-2b
us-west-2d < Ok >		us-west-2C
< Ok > < Back >		us-west-2a
< Ok > K Back >		
	L	

10. Choose c6a for instance type family for cloud nodes and then select Ok.

c6a instances are widely available and provide good performance and value for this use case. At a later step, one of the preallocated public cloud nodes will be configured to use an instance type with NVIDIA GPUs.

	3 elements above	
c5	c5 instances (9)	
c5a	c5a instances (8)	
c5d	c5d instances (9)	
c5n	c5n instances (7)	
c6a	c6a instances (11)	
c6g	c6g instances (9)	
c6gd	c6gd instances (9)	
c6an	c6an instances (8)	
c6i	c6i instances (10)	
d2	d2 instances (4)	
a2	g2 instances (2)	
03	n3 instances (3)	
g4dn	gédinistances (7)	
gran	24 alementa halew	

11. Choose c6a.large instances and then select 0k.

Select the de	fault instance type for cloud nodes.	
céa large	2 cores/4 8 CB	
c6a.xlarge	4 cores/8.0 GB	
c6a.2xlarge	8 cores/16.0 GB	
c6a.4xlarge	16 cores/32.0 GB	
c6a.8xlarge	32 cores/64.0 GB	
c6a.12xlarge	48 cores/96.0 GB	
c6a.16xlarge	64 cores/128.0 GB	
c6a.24xlarge	96 cores/192.0 GB	
c6a.32xlarge	128 cores/256.0 GB	
c6a.48xlarge	192 cores/384.0 GB	
c6a.metal	192 cores/384.0 GB	

12. Choose the c6a instance type family for cloud directors and then select 0k.

Please	select the instance type family for cloud directors	
	5 elements above	
c5d	c5d instances (9)	
c5n	c5n instances (7)	
c6a	c6a instances (11)	
c6g	c6g instances (9)	
c6gd	c6gd instances (9)	
c6gn	c6gn instances (8)	
c6i	c6i instances (10)	
d2	d2 instances (4)	
g2	g2 instances (2)	
g3	g3 instances (3)	
g4dn	g4dn instances (7)	
i2	i2 instances (4)	
i3	i3 instances (7)	
	34 elements below	

13. Choose the c6a.large instance type and then select 0k.

Select the de	fault instance type for cloud directors.	
c6a.large	2 cores/4.0 GB	
c6a.xlarge	4 cores/8.0 GB	
c6a.2xlarge	8 cores/16.0 GB	
c6a.4xlarge	16 cores/32.0 GB	
c6a.8xlarge	32 cores/64.0 GB	
c6a.12xlarge	48 cores/96.0 GB	
c6a.16xlarge	64 cores/128.0 GB	
c6a.24xlarge	96 cores/192.0 GB	
c6a.32xlarge	128 cores/256.0 GB	
c6a.metal	192 cores/384.0 GB	
c6a.48xlarge	192 cores/384.0 GB	

14. Choose Select images and then select Ok.

This selects the subset of images that can be used in the public cloud and eliminates those that cannot be used (such as DGX OS).

Cluster Extension	
Sel	lect software images to be provisioned to a cloud director. (Recommended: All)
A	ll images (9 images total) Diect images (allows you to dick images)
	C OK S K Back S K Holp S

15. Choose the images that were created for this deployment and then select 0k.

cloud-director-image, k8s-cloud-gpu-worker-image, and k8s-cloud-master-image should be checked. Additional images can be added later if necessary.

Select software images to be provisioned to a cloud director. (Recommended: All) [] backup-default-image <not used=""> [] backup-dgx-os-5.4-al00-image <not used=""> [X] cloud-director-image <not used=""> [] default-image <not used=""> [] dgx-os-5.4-al00-image <not used=""> [] kBs-cloud-gpu-worker-image <not used=""> [] kBs-master-image <not used=""> [] kBs-master-image used by: knode01knode03</not></not></not></not></not></not></not></not></not></not></not>	r Extension	
<pre>[] backup-default-image <not used=""> [] backup-dgx-os-5.4-a100-image <not used=""> [X] cloud-director-image <not used=""> [] default-image <not used=""> [] dgx-os-5.4-a100-image (not used> (x] k8s-cloud-gpu-worker-image <not used=""></not> (x] k8s-cloud-master-image <not used=""></not> (x] k8s-master-image (x) k8s-cloud-master-image (x) used> (x) k8s-cloud-master-image (x) used> (x) k8s-cloud-master-image (x) k8s-cloud-master-image (x) k8s-cloud-master-image (x) k8s-cloud-master-image (x) k8s-master-image </not></not></not></not></pre>	Select software images to be provid	sioned to a cloud director. (Recommended: All)
<pre>< Ok > K Back > K Help ></pre>	<pre>[] backup-default-image [] backup-dgx-os-5.4-a100-image [X] cloud-director-image [] default-image [] default-image [] dgx-os-5.4-a100-image [X] k8s-cloud-gpu-worker-image [X] k8s-cloud-master-image [] k8s-master-image [] k8s-master-image</pre>	<not used=""> <not used=""> <not used=""> <not used=""> used by: dgx01dgx04 <not used=""> <not used=""> <not used=""> <not used=""> <not used=""></not></not></not></not></not></not></not></not></not>
	< <u>0k</u> >	< Back > < Help >

16. Choose k8s-cloud-master-image for the default cloud node image and then select 0k.

Cluster Extension	
,	
	Select software image that will be used for cloud nodes.
	() cloud-director-image
	() k8s-cloud-gpu-worker-image (X) k8s-cloud-master-image
	< Ok > < Back > < Help >

17. Choose Save config & deploy on the Summary screen and then select Ok.

Summary	
Save config & deploy Show config Advanced settings Save config Save config & exit Exit	
< Ok > < Back >	_

18. Specify the filepath and then select 0k.

A default filepath is displayed. A region name or other identifying information should be added to the file name to allow multiple configuration files.

root/cm-cluster-extension	-us-east-1.conf	
•• BrightManuals/	4.0K	
Deskton/	10	
Documents/	10	
Downloads/	10	
Music/	10	
Pictures/	10	
Public/	10	
Templates/	10	
	26 elements below	
l Show hidden	[] Resolve symlinks [] Show details	

19. The configuration begins executing on the BCM head node.

When completed, output like the following should be displayed.

Cluster Extension finished!

20. Verify that the initial setup was successful.

Run list -f in cmsh as shown in the screenshot and compare it to the output provided—it should be similar (additional listed systems are redacted, and the exact IP subnet may be slightly different).

[hybridbasepod-b-primary->device]% list -f Type, Hostname, Category, Ip, Network, Status

Туре	Hostname (key)	MAC	Category	Ip	Network	Sta	atus	
CloudNode	us-west-2-cnode001	00:00:00:00:00:00	us-west-2-cloud-node	172.17.0.1	us-west-2	[DOWN]
CloudNode	us-west-2-cnode002	00:00:00:00:00:00	us-west-2-cloud-node	172.17.0.2	us-west-2	[DOWN]
CloudNode	us-west-2-cnode003	00:00:00:00:00:00	us-west-2-cloud-node	172.17.0.3	us-west-2	[DOWN]
CloudNode	us-west-2-cnode004	00:00:00:00:00:00	us-west-2-cloud-node	172.17.0.4	us-west-2	[DOWN]
CloudNode	us-west-2-director	00:00:00:00:00:00	aws-cloud-director	172.17.255.251	us-west-2	[DOWN]

21. Augment the OpenVPN port if needed.

The Cluster Extension functionality relies on OpenVPN to run a VPN tunnel between the on-premises head node and the targeted public cloud environment. The default configuration uses UDP port 1194. To configure a different protocol or port, refer to this Bright Knowledge Base article.

Chapter 4. Host Preparation After Cluster Extension Configuration

With AWS cloud resource access configured, the next step is to modify the device entries created using the cm-cluster-extension to leverage the correct categories and make changes when necessary to public cloud settings.

1. Rename the nodes according to their expected usage as follows from the device sub-menu.

```
# cmsh
   % device
   % rename us-west-2-cnode001 us-west-2-knode001
   % rename us-west-2-cnode002 us-west-2-knode002
   % rename us-west-2-cnode003 us-west-2-knode003
   % rename us-west-2-cnode004 us-west-2-gpu-node001
   % commit
2. Update the categories for all four public cloud nodes.
   % set us-west-2-knode001 category k8s-cloud-master
   % set us-west-2-knode002 category k8s-cloud-master
   % set us-west-2-knode003 category k8s-cloud-master
   % set us-west-2-gpu-node001 category k8s-cloud-gpu-worker
   % commit
Increase the EBS volume size to 100 GiB on the knode systems.
   % cloudsettings us-west-2-knode001
   % storage
   % set ebs size 100GiB
   % commit
   % ..
   % ..
   % ..
   % cloudsettings us-west-2-knode002
   % storage
   % set ebs size 100GiB
   % commit
   % ...
   % . .
   % . .
   % cloudsettings us-west-2-knode003
   % storage
   % set ebs size 100GiB
   % commit
   % ...
   % ...
   % ..
```

4. Increase the EBS volume size of us-west-2-gpu-node001 to 100 GiB and change the instancetype for us-west-2-gpu-node001 to g4dn.xlarge (a lower-cost GPU instance with a single T4 GPU).

If requirements justify a higher instance spec, use the appropriate instance type any NVIDIA GPU instance should work.

```
% cloudsettings us-west-2-gpu-node001
% set instancetype g4dn.xlarge
% commit
% storage
% set ebs size 100GiB
% commit
```

5. Increase the EBS volume size to 200 GiB on the director system.

```
% ..
% cloudsettings us-west-2-director
% storage
% set ebs size 200GiB
% commit
```

6. Update the softwareimage for the aws-cloud-director category.

```
% category
% use aws-cloud-director
% set softwareimage cloud-director-image
```

7. Update disksetup for the cloud director to use the same partitioning scheme set for the other public cloud nodes.

% set disksetup /tmp/big-cloud-disk.xml

% commit

Chapter 5. Power On and Provision the Cloud Nodes

Now that the required post-installation configuration has been completed, it is time to power on and provision the public cloud nodes. Public cloud node behavior is slightly different from on-premises equipment—the systems will not be provisioned in the target public cloud until they are first powered on. Additionally, the director node must be powered on and provisioned first—until it is fully provisioned, it is not possible to deploy the public cloud nodes it manages in a region. Just as with on-premises deployments, the public cloud nodes can be accessed through ssh during the installation process.

Watch the /var/log/messages and /var/log/node-installer log files to verify that everything is proceeding smoothly if you are unsure of a given node's deployment state.

1. Power on the cloud director.

It will enter a [$\ensuremath{\texttt{PENDING}}$] state, then transition to [$\ensuremath{\texttt{DOWN}}$] (Instance has started).

cmsh

% power on us-west-2-director

The provisioning of the cloud director may take two or more hours due to the tens of gigabytes of software image data that must be synchronized to the public cloud. The process is complete when the cloud director moves to an [UP] state.

2. Power on the four public cloud nodes concurrently.

Once the cloud director is fully provisioned, bringing up the other four public cloud nodes is much faster because their base images are already stored in the target region with the cloud director.

% power on -n us-west-2-knode00[1-3],us-west-2-gpu-node001

3. Run device then list to ensure all public cloud nodes are in an [UP] state.

Disregard any trailing Status output.

% device % list Type 	Hostname (key)	мас	Category	Ip	Network	Status	
CloudNode CloudNode CloudNode CloudNode CloudNode	us-west-2-director us-west-2-gpu-node001 us-west-2-knode001 us-west-2-knode002 us-west-2-knode003	00:00:00:00:00:00 00:00:00:00:00:00 00:00:	aws-cloud-director k8s-cloud-gpu-worker k8s-cloud-master k8s-cloud-master k8s-cloud-master	172.17.255.251 172.17.0.4 172.17.0.1 172.17.0.2 172.17.0.3	us-west-2 us-west-2 us-west-2 us-west-2 us-west-2	[UP [UP [UP [UP [UP]]]]

4. Install the NVIDIA driver on us-west-2-gpu-node001.

ssh to it as root and run all subsequent commands from the node in AWS.

```
# ssh us-west-2-gpu-node001
# apt install linux-headers-$(uname -r)
# distribution=$(. /etc/os-release;echo $ID$VERSION_ID | sed -e 's/\.//g')
# wget https://developer.download.nvidia.com/compute/cuda/repos/$distribution/x86_64/cuda-
keyring_1.0-1_all.deb
# dpkg -I cuda-keyring_1.0-1_all.deb
# apt update
# apt install -y cuda-drivers -no-install-recommends
# rm cuda-keyring_1.0-1_all.deb
# nvidia-smi
```

5. Look for output from nvidia-smi, which like this, shows a successful installation.

Expect possible variations in software versions and device utilization.

```
+-----+
| NVIDIA-SMI 525.85.12 Driver Version: 525.85.12 CUDA Version: 12.0 |
|-----+
| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |
Fan Temp Perf Pwr:Usage/Cap|Memory-Usage | GPU-Util Compute M. |||MIG M. |
+-----+
+------
| Processes:
       PID Type Process name GPU Memory |
Usage |
| GPU GI CI
| ID ID
| No running processes found
+-----
```

- 6. Log out of the public cloud GPU node and back into the on-premises head node.
- 7. Execute the following to capture the modifications made to the public cloud GPU node, which will then be present in the image of any additional public cloud GPU nodes provisioned in this environment.

```
$ cmsh
```

```
# device
```

```
# use us-west-2-gpu-node001
```

grabimage -w

Chapter 6. Deploy Kubernetes

With all required public cloud instances deployed and configured for general use, the environment is ready for K8s deployment. In a hybrid environment, the same tool used to deploy on-premises K8s is used to deploy K8s in the public cloud as well.

- 1. Run the cm-kubernetes-setup CLI wizard as the root user on the head node. # cm-kubernetes-setup
- 2. Choose Deploy to begin the deployment and then select Ok.

	Kubernetes operations
Deploy Add user Modify users Remove user Regenerate certs Setup PSP Disable PSP Modify Kyverno policies Update system addons Uninstall Exit	Kubernetes installation wizard Create a new user to Kubernetes Modify Kubernetes users Remove a user from Kubernetes Configure PodSecurityPolicy usage for Kubernetes Disable PodSecurityPolicy usage for Kubernetes Change lists of namespaces to be excluded from polici Update system addons Yaml in CMDaemon to the latest Uninstall Kubernetes Return to the command line
	< 0k >

3. Choose Kubernetes v1.21 and then select Ok.

K8s version 1.21 was selected to match the version deployed in the on-premises DGX BasePOD deployment.

Choose a Kubernetes version.	
() Kubernetes v1.24 () Kubernetes v1.22	
(X) Kubernetes v1.21	
< Ok > < Back >	

4. Choose Containerd (it should be selected by default) and then select Ok.

Select container runtime to use [X] Containerd () Use Bright Docker package (deprecated) () Use existing third party Docker package (deprecated)	
< Ok > < Back >	

 Optionally, provide a registry mirror and then select 0k. This example deployment did not require one.

ernetes Setup	
	Please provide a registry mirror for pulling container images from DockerHub (optional)
	DockerHub registry mirror server (optional)
	< Ok > K Back >

6. Configure the basic values of the K8s cluster and select 0k.

Choose names that make it easy to understand that the K8s deployment is using public cloud resources. In addition, ensure that the service and pod network subnets do not overlap with existing subnets in the cluster.

· · ·	•		
Kubernetes Setup			
	Insert ba	asic values of the new Kubernetes cluster	
	KUD	ernetes cluster name aws-cloud	
	Kube	roetes external FODN bybridbasenod-b-primary pyidia co	
	Service	network base address 10.152.0.0	
	Service	network netmask bits 16	
	Pod	network base address 172.31.0.0	
	Pod i	network netmask bits <mark>16</mark>	
		< UK > K Back >	

7. Choose yes to expose the K8s API server to the external network and then select 0k. This allows users to use the K8s cluster from the head node.

Kubernetes Setup	
Do you want to expose the Kubernetes API server to the external network	
yes and the second s	
no	
< Ok > < Back >	

8. Choose vpc-us-west-2-private for the public cloud-based K8s environment and then select 0k.

Kubernetes Setup		
	Select the internal network used by Kubernetes nodes This is the network Kubernetes nodes use to communicate with other Kubernetes nodes. One network must be selected. Use <space> to select a single value.</space>	
	<pre>() externalnet () ipninet () internalnet () mpinet () storagenet () netmap () us-west-2 () vpc-us-west-2-public () vpc-us-west-2-private</pre>	
	< Ok > K Back >	

This keeps internal K8s traffic entirely in the public cloud.

9. Choose the three k8s-cloud-master nodes and then select 0k.

Kubernetes Setup			
	Select master nodes for Kuber	metes.	
	<pre>[] hybridbasepod-b-primary [] us-west-2-director [] us-west-2-ppu-node001 [X] us-west-2-knode001 [X] us-west-2-knode002 [X] us-west-2-knode003</pre>	category:aws-cloud-director category:k8s-cloud-gpu-worker category:k8s-cloud-master category:k8s-cloud-master category:k8s-cloud-master	
		< Back >	-

10. Choose k8s-cloud-gpu-worker for the worker node category and then select 0k.

Select node categories to use as Kubernetes workers	
[] aws-cloud-director [] default	
[X] k8s-cloud-gpu-worker	
[] us-west-2-cloud-node	
	-
< Ok > < Back >	

11. Select 0k without configuring any individual K8s nodes.

Select individual Kubernetes	category:aws-cloud-director category:k8s-cloud-master category:k8s-cloud-master category:k8s-cloud-master	
< 0k >	< Back >	

12. Choose the three knode systems for Etcd nodes and then select 0k.

Select an odd number of Etcd	nodes	
<pre>[] hybridbasepod-b-primary [] us-west-2-director [] us-west-2-gpu-node001 [X] us-west-2-knode001 [X] us-west-2-knode002 [X] us-west-2-knode003</pre>	category:aws-cloud-director category:k8s-cloud-gpu-worker category:k8s-cloud-master category:k8s-cloud-master category:k8s-cloud-master	
<pre></pre>	< Back >	-

13. Configure the K8s main components and then select 0k.

Use the default ports and path here unless the environment requires different values. The default values were used in this deployment.

	7_
Configure the values for the main Kubernetes components: API server proxy port 10443 API server port 6443 Kubelet port 10250 Kube-proxy health port 10250 Controller manager port 10252 Scheduler port 10253 Etcd spool directory /var/lib/etcd	
< Ok > K Back >	

14. Choose the Calico network plugin and then select Ok.

ubernetes Setup		
Sel	ect the Kubernetes network plugin	
) Calico (recommended)) Weave	
) Flannel	
	< Ok > < Back >	

15. Choose yes to install the Kyverno policy engine and then select 0k.

Kubernetes Setup		
	Do you want to install Kyverno Policy Engine?	
	Kyverno is a policy engine designed for Kubernetes. It can validate, mutate, and generate configurations using admission controls and background scans.	
	yes	
	< Ok > < Back >	

16. Choose no to decline to configure HA for Kyverno and then select 0k. This deployment does not meet the minimum node requirement for Kyverno HA.

Do you want to configure High Availablity (HA) for Kyverno? Configuring HA is a recommended way of runnning Kyverno. For this configuration the number of worker nodes in Kubernetes cluster at any given time must be not less than 3.	
yes ac	
C Ok > K Back >	

17. Choose whether to install Kyverno Policies and then select 0k. Unless required for the configuration, choose no.

iernetes Setup	
Do you want to install Kyverno Policies?	
These Kyverno policies are based on Kubernetes Pod Security Standards definitions. https://kubernetes.io/docs/concepts/security/pod-security-standards	
Number of exclusions will be added automatically based on configured features and installed operators.	
Vac	
no	
< Ok > < Back >	

18. Choose the operator packages to install and then select 0k.

As shown in the screenshot, choose NVIDIA GPU Operator, Prometheus Adapter, Prometheus Adapter Stack, and the cm-jupyter-kernel-operator.

ernetes setup		
	·	
	Which operators packages to install	
	[X] NVIDIA GPU Operator	
	[X] Prometheus Adapter	
	[X] Prometheus Operator Stack	
	[] cm-kubernetes-postgresql-operator	
	[] cm-kubernetes-spark-operator	
	< Ok > < Back >	

19. Choose the same four operators to be rolled up with the defaults and then select 0k.

Kubernetes Setup		
W	hich operators to roll	
	NVIDIA GPU Operator	
	[X] Prometheus Adapter	
	[X] Prometheus Operator Stack [X] cm-jupyter-kernel-operator	
	< Ok > < Back >	

20. Choose the addons to deploy and then select 0k.

As shown in the screenshot, choose Ingress Controller (Nginx), Kubernetes Dashboard, Kubernetes Metrics Server, and Kubernetes State Metrics.

Which addons do you want to deploy?	
[¥] Ingress Controller (Nginx)	
[X] Kubernetes Dashboard	
[X] Kubernetes Metrics Server	
[X] Kubernetes State Metrics	
< Ok > < Back >	

21. Choose the Ingress ports for the cluster and then select 0k. Use the defaults unless specific ingress ports are required.

Kubernetes Setup		
	Insert values of the new Kubernetes cluster	
	Ingress HTTP port 30080	
	Ingress HIPS port 30443	
	< Ok > < Back >	

22. Choose no when asked to install the Bright NVIDIA packages and then select 0k. Since the K8s control plane nodes do not have GPUs, the GPU Operator manages NVIDIA OS components.

Kubernetes Setup		
	Do you wish to install the Bright NVIDIA packages? - cm-nvidia-container-toolkit - cuda-dcgm - cuda-driver These will be installed in the following software images:	
	- /cm/images/k8s-cloud-master-image - /cm/images/k8s-cloud-gpu-worker-image Yes no	
	COK > K Back >	

23. Choose yes to deploy the Permission Manager and then select 0k.

Do you want to instal This is only needed i	
This is only needed i	
	if you want to have non-root users on the cluster
	C Ok > K Back >

24. Select 0k without configuring any optional values.

Kubernetes Setup		
	Configure permission manager Custom address of the registry (optional) Custom controller image (optional) Custom RBAC Proxy image (optional)	
	< Ok > K Back >	

25. Choose both enabled and default for the Local path storage class and then select 0k.

Configure Kubernetes StorageClass StorageClass enabled default CEPH [] () CEPH is not available. [] () Local path [] () Do not set default [] () Ok > K Back >	
Storageclass enaled berailt CEPH [] CEPH is not available. Local path Do not set default < Ok >	Configure Kubernetes StorageClass
< Ok > K Back >	StorageClass enabled default CEPH [] () CEPH is not available. [X] Local path [X] Do not set default ()
	C Ok > K Back >

26. Select 0k without changing any of the default values.

Configure local path storage pool for Kubernetes
Path to store the data /cm/shared/apps/kubernetes/aws-cloud/var/volumes Custom address of the registry (optional) Custom provisioner's image (optional)
< Ok > < Back >

27. Choose Save config & deploy and then select Ok.

Su S S S S S S S S S S S S S S S S S S	,mmary Save config & deploy Show config Save config & exit Exit	
	< Ok > K Back >	

28. Change the filepath to /root/cm-kubernetes-setup-cloud.conf and then select 0k. The filepath was changed to avoid name conflicts with the existing K8s configuration file from the initial on-premises deployment. Wait for the installation to finish.

	4.0K	
BrightManuals/	4.0K	
Desktop/	10	
Documents/	10	
Downloads/	10	
Music/	10	
Pictures/	10	
Public/	10	
Templates/	10	
Videos/	10	
all/	4.0K	
amd64/	4.0K	
aws-multi-gpu/	58	
bin/	51	
blender-images/	28	
blender-results-aws-amd/	29	
	20 elements below	

29. Verify the K8s cluster is installed properly.

Ready

us-west-2-knode003

The K8s module may need to be unloaded for the on-premises deployment if already loaded or use the switch command as a shortcut to unload on-premises and load the public cloud module.

```
# module load kubernetes/aws-cloud/
# kubectl cluster-info
Kubernetes control plane is running at https://localhost:10443
CoreDNS is running at https://localhost:10443/api/v1/namespaces/kube-system/services/kube-
dns:dns/proxy
To further debug and diagnose cluster problems, use 'kubectl cluster-info dump'.
# kubectl get nodes
NAME
                       STATUS
                                ROLES
                                                      AGE
                                                              VERSION
us-west-2-gpu-node001
                       Ready
                                worker
                                                      6m48s
                                                              v1.21.4
us-west-2-knode001
                       Ready
                                control-plane,master 6m48s
                                                              v1.21.4
us-west-2-knode002
                       Ready
                                control-plane,master 6m48s
                                                              v1.21.4
```

control-plane, master 6m48s v1.21.4

- 30. Verify that a GPU job can be run on the K8s cluster.
 - a. Save the following text to a file named gpu.yaml.

- b. Execute the code using kubect1 apply.
 # kubect1 apply -f gpu.yaml
- c. Use kubectl logs to check the result.

The output should be like the following.

kubectl logs gpu-pod-pytorch

```
Tue Feb 14 22:25:53 2023
+------
| NVIDIA-SMI 525.85.12 Driver Version: 525.85.12 CUDA Version: 12.0 |
|-----+
| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

    | Fan Temp Perf Pwr:Usage/Cap|
    Memory-Usage | GPU-Util Compute M. |

    |
    |

0 Tesla T4 0n | 00000000:00:1E.0 0ff | 0 0
| N/A 28C P8 14W / 70W | 2MiB / 15360MiB | 0% Default |
                                 N/A |
Т
             +----+
+----------------------+
| Processes:
                                    1
| GPU GI CI PID Type Process name GPU Memory |
Usage |
|------
| No running processes found
 -----
             -----
```

Chapter 7. Create Additional Worker Nodes

The steps in this section cover how to extend the pool of worker nodes.

- 1. Access cmsh and enter the device sub-menu.
 - # cmsh
 - % device
- 2. Clone the single worker node.

Maintaining the naming convention will automate the IP address increment.

% clone us-west-2-gpu-node001 us-west-2-gpu-node002

- % commit
- 3. Power the additional worker node on and wait until it enters the [UP] state. % power on us-west-2-gpu-node002
- 4. Verify the worker nodes are ready by using kubectl.

The worker node should automatically be available as part of the public cloud Kubernetes worker pool because the entire category is marked as worker nodes. The resources should be ready to use immediately.

kubectl get nodes

•				
NAME	STATUS	ROLES	AGE	VERSION
us-west-2-gpu-node001	Ready	worker	25h	v1.21.4
us-west-2-gpu-node002	Ready	worker	10m	v1.21.4
us-west-2-knode001	Ready	control-plane,master	25h	v1.21.4
us-west-2-knode002	Ready	control-plane,master	25h	v1.21.4
us-west-2-knode003	Ready	control-plane,master	25h	v1.21.4

Chapter 8. (Optional) Enable Jupyter Operator Use in Cloud K8s Cluster

In the on-premises DGX BasePOD deployment guide, cm-jupyter-setup can be optionally configured and integrated into K8s. The same service, running from the head node, can be used to provide Jupyter access to the public cloud-based K8s cluster as well.

- Validate cm-jupyterhub is set up and running correctly.
 # service cm-jupyterhub status
- 2. Configure a user and provide access to the appropriate K8s cluster.
 # cmsh -c "user; add userone; set password useronepwd; commit"
- 3. When using K8s via Jupyter, users must be added separately using K8s with the following commands. Users must have permission to access the Jupyter kernel operator in both K8s clusters to use the kernel templates.

apt install cm-python39
<pre># cm-kubernetes-setupadd-user useronecluster aws-cloudoperators cm-jupyter-kernel-</pre>
operator
<pre># cm-kubernetes-setupadd-user useronecluster onpremoperators cm-jupyter-kernel-</pre>
operator

4. Sign in to the Jupyter web interface using the account configured with Jupyter kernel operator permissions.

	0	i localhost	Ċ	
oyte <mark>rhub</mark>				
		Sign in		
		Username:		
		userone		
		Password:		
		······································		
		Sign in		
		Signin		

5. Navigate to the Bright tab, choose the Python+NGC on Kubernetes Operator kernel template, and then select Ok.

•	••• • •		0	í	localhost	¢	
0	File Edit View Run Kernel	Tabs Settings Help					
	KERNEL TEMPLATES	Z Launcher					
0 = *	Julia on Kubernetes Operator + Python+NGC on Kubernetes Operator + Python on Kubernetes Operator + Python+Spark on Kubernetes O + KERNEL DEFINITIONS WLM CLUSTERS		Notebook				
0	KUBERNETES CLUSTERS	1					
	onprem \$	1	Console				
			S. Other	Text File Markdown File P	contextual Help		

6. Fill in the required fields on the resulting New kernel window and then select Create. In this example, the public cloud-based K8s deployment was targeted by adding the cluster name (aws-cloud) as a path extension to the K8s environment module, and it was specified that the container could use a single GPU.

••• 🗉		0	i localhost	6
🔅 File Edit	View Run Kernel Tabs Settings Help			
KERNEL TE	MPLATES	S Launcher		
O Python+N0 Python on	bernetes Operator + 3C on Kubernetes Operator + Kubernetes Operator +	Not	ebook	
KERNEL DE	FINITIONS		New kernel	×
WLM CLUS	TERS	Pytho	Kernel name:	
KUBERNET	ES CLUSTERS		jupyter-eg-kernel-k8s-cmjkop-ngc-py-1gpbgljpe	
onprem		<mark>≻_</mark> Co	Display name of the kernel:	
aws-cloud	۵		Python+NGC on Kubernetes Operator 230215134611	
		é	Environment module to load:	
		Pytho	kubernetes/aws-cloud	
			Image to run:	
		\$_ Ot	nvcr.io/nvidia/pytorch:22.08-py3 ×	x ~
			Image pull policy:	
		\$_	Mameroana for kernal-	× V
		Termi	userone-restricted	
			Number of GPUs container can use:	
			1	
			Cancel	Create
			Canoel	Create

7. Select Python+NGC on Kubernetes in the Notebook section.



8. Once the state of the operator becomes Idle, run nvidia-smi to confirm the notebook is running on a T4 GPU instance.

• • • • • •	0	iii localhost	C	Ů + ©
File Edit View Run Kernel Tabs Settings He	elp			
KERNEL TEMPLATES	Untitled1.ipynb			
Julia on Kubernetes Operator + Python+NOC on Kubernetes Operator + Python NUbernetes Operator + KERNEL DEFINITIONS Python+NGC on Kubernetes Operator 230215134 × WLM CLUSTERS KUBERNETES CLUSTERS onprem \$	B + X D > m C [1]: fmvidia-smi Thu Feb 10 5659834 M M S G G M S G G M M M M M A M T G G M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M	× Code ✓ VC VNC 2023 12 Driver Version: 525.85.12 CUA Version: Persistence-H Bus-1d Disp.A Volatile Un Persistence-H Bus-1d Memory-Usage GPU-Util C 14W / 7W 2HB / 155GH1B 0+ PID Type Process name G Sees found	Python	+KGC on Kubernetes Operator 220215134611 ()
Simple 0 Puthon+NGC on Kubernete	es Operator 230215134611 Idle			Mode: Edit 🛞 Ln 1, Col 1 Untitled1.ipyn

Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation ("NVIDIA") makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer ("Terms of Sale"). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA's aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, DGX, NVIDIA DGX BasePOD, and NVIDIA Base Command are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2023 NVIDIA Corporation. All rights reserved.

