

# High Availability with VMware Cloud Director Container Service Extension

## Table of Contents

Introduction .....	3
High Availability Architecture of Container Service Extension Workloads .....	3
Configuration steps for the Service Provider .....	5
VM/Host rule configuration in vCenter .....	5
Placement Policy configuration in Cloud Director .....	7
Creating Highly Available Kubernetes Clusters as Tenant User .....	11
Note for stretched ESXi clusters .....	13
Conclusion.....	13

### Introduction

This document describes how to create high available Kubernetes clusters with VMware Cloud Director and Container Service Extension.

VMware Cloud Director is a cloud service delivery platform that enables service providers to offer secure, efficient, and scalable cloud services to their customers. It provides a multi-tenant architecture that allows service providers to manage multiple customers and their resources in a vSphere environment. VMware Cloud Director also offers a self-service portal for customers to manage their own resources, including virtual machines, networks, and storage.

Container Service Extension (CSE) is an add-on for VMware Cloud Director that enables service providers to offer Kubernetes as a Service to their customers. CSE allows customers to deploy and manage Kubernetes clusters on VMware Cloud Director through the web portal or APIs.

Any vSphere infrastructure consists of a set of components such as CPU/Memory/Storage components within a bare-metal ESXi host, a rack of hosts, networking components within the racks, switches connecting multiple racks and so on.

A failure domain is a set of components whose failure will cause failure of application workloads. Hence application workloads will need to be aware of the topology of the failure domains, and replicate workloads across failure domains in order to achieve reliability across failure domains of interest. To ensure availability with good performance on failure, the placement of application workloads across failure domains must be balanced.

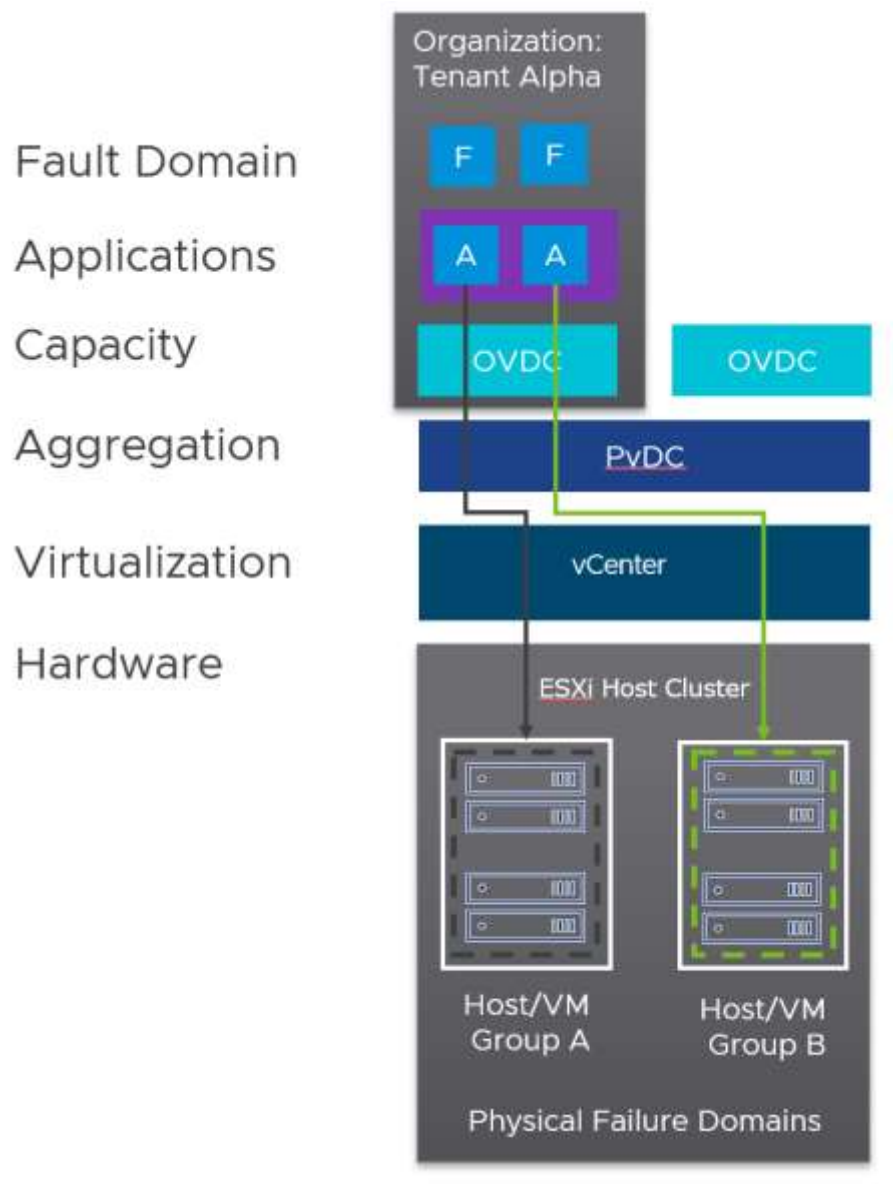
### High Availability Architecture of Container Service Extension Workloads

To increase the availability of Kubernetes Clusters that are created and managed by Container Service Extension, you can use Placement Policies in VCD to spread the Kubernetes Master and Worker nodes across failure domains in vCenter.

This can be done by grouping ESXi hosts based on their failure domain location in the DRS settings in vCenter. These Host Groups then can be used in the Provider VDC (PVDC) configuration in VMware Cloud Director, to define Placement Policies. When creating a Kubernetes Cluster using Container Service Extension, a tenant user then can create groups of master and worker nodes, and select these Placement Policies.

The Kubernetes nodes then will be distributed across the different ESXi hosts, spanning multiple failure domains in vCenter.

The diagram below shows the architecture:

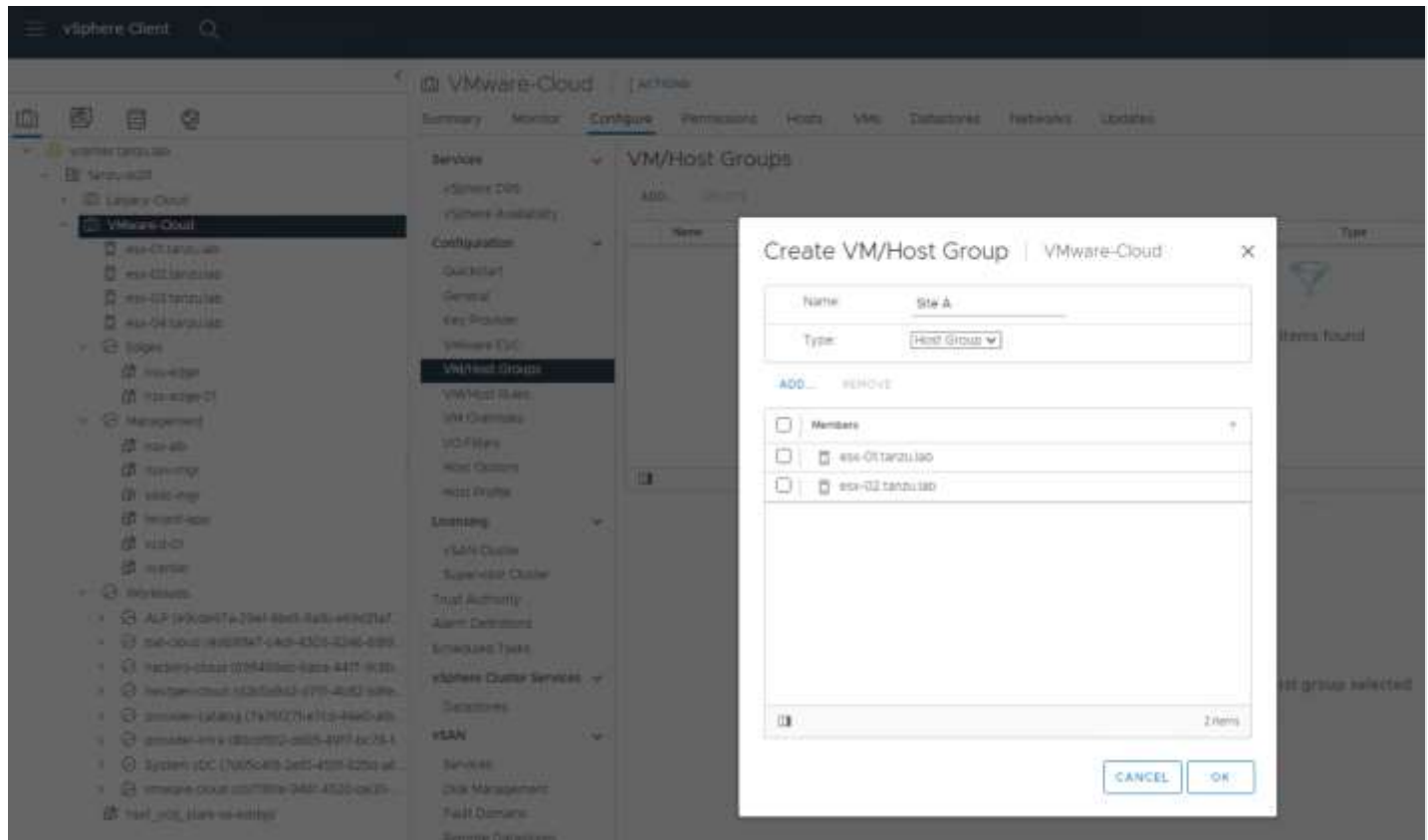


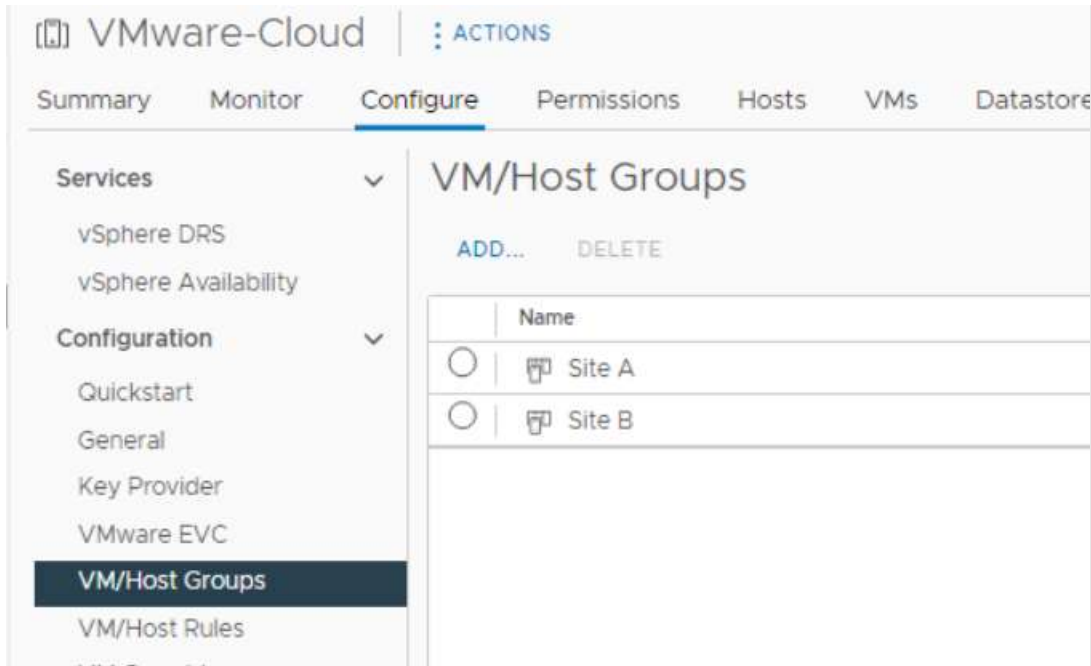
### Configuration steps for the Service Provider

As Service Provider, following steps are needed:

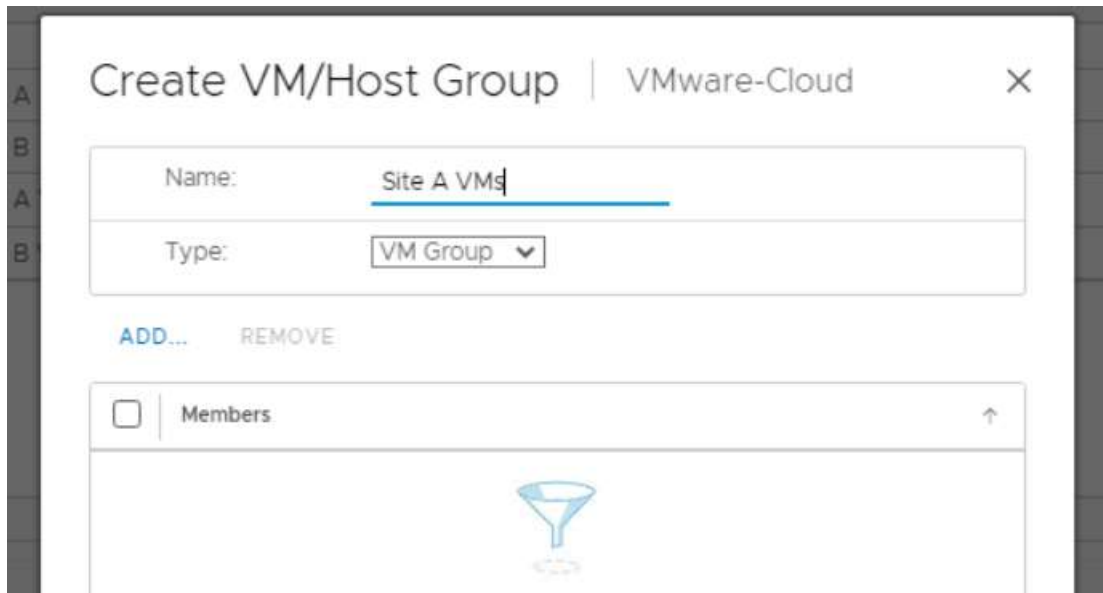
#### VM/Host rule configuration in vCenter

In Cluster Configuration, VM/Host Group setting, create Host Groups that reflect the physical failure domains. Add the ESXi hosts to the representing Host Group.





Then create VMs groups for each failure domain.



Now create VM/Host rules, and specify that the VM group must run on the corresponding Host Group:

### Create VM/Host Rule | VMware-Cloud ✕

Name	Site A	<input checked="" type="checkbox"/> Enable rule.
Type	Virtual Machines to Hosts <span style="float: right;">▼</span>	

Description:

Virtual machines that are members of the Cluster VM Group Site A VMs must run on host group Site A.

VM Group:

Site A VMs ▼

Must run on hosts in group ▼

Host Group:

Site A ▼

CANCEL
OK

VM/Host Rules					
Name	Type	Enabled	Conflicts	Defined By	
Site A	Run VMs on Hosts	Yes	0	User	
Site B	Run VMs on Hosts	Yes	0	User	

**Placement Policy configuration in Cloud Director**

In Cloud Director's Provider Portal, under Resources / Cloud Resources, VM Placement Policies, create Placement Policies that reflect these different Host / VM Groups. Make sure to label them properly, so that tenant users can understand that they represent different failure domains.

## Create VM Placement Policy

- 1 What is VM Placement Policy?
- 2 General
- 3 VM Groups
- 4 Ready to Complete

## VM Groups

VM Groups **1** Logical VM Groups **2**

A VM group is a collection of virtual machines with similar host requirements. Create an inline Logical VM Group by selecting one VM Group per cluster.

- > legacy-cloud (0 selected)
- ▼ vmware-cloud-vcf (1 selected)
  - ▼ VMware-Cloud
    - Site B VMs
    - Site A VMs

CANCEL

PREVIOUS

NEXT



The screenshot displays the VMware Cloud Director web interface. At the top, there is a navigation bar with the VMware logo and the text 'VMware Cloud Director'. Below this, there are tabs for 'Resources', 'Libraries', 'Administration', 'Monitor', and 'More'. Under the 'Resources' tab, there are sub-tabs for 'Cloud Resources' and 'Infrastructure Resources'. A left-hand sidebar contains a list of menu items: 'Organizations', 'Organization VDCs', 'Organization VDC Templates', 'Provider VDCs', 'Cloud Cells', 'Edge Gateways', 'External Networks', 'Tier-0 Gateways', 'Network Pools', 'VM Sizing Policies', 'VM Placement Policies' (which is highlighted), and 'vGPU Policies'. The main content area is titled 'VM Placement Policies' and features a 'NEW' button. Below the button is a table with a header row containing 'Name' and a 'D' column. The table lists two entries: 'Site A' and 'Site B', each with a radio button to its left.

	Name	D
<input type="radio"/>	Site A	
<input type="radio"/>	Site B	

In the Settings of Organization VDCs that are entitled to use Container Service Extension, add the VM Placement Policies

## Select Policies to Add to VDC



Show selected

<input checked="" type="checkbox"/>	Name	Description
<input checked="" type="checkbox"/>	Site A	
<input checked="" type="checkbox"/>	Site B	
<input checked="" type="checkbox"/>	2	1 - 2 of 2 compute policy(ies)

CANCEL

OK

VMware Cloud Director | Resources | Libraries | Administration | Monitor | More

Cloud Resources | Infrastructure Resources

Organizations | Organization VDCs | Organization VDC Templates | Provider VDCs | Cloud Cells | Edge Gateways | External Networks | Tier-0 Gateways | Network Pools | VM Sizing Policies | VM Placement Policies | vGPU Policies

All Organization VDCs > vmware-cloud

vmware-cloud OPEN IN TENANT PORTAL

General

Allocation

Networking

Network Pool

Edge Clusters

Segment Profile

Templates

Metadata

Resource Pools

Policies

Storage

VM Placement

VM Sizing

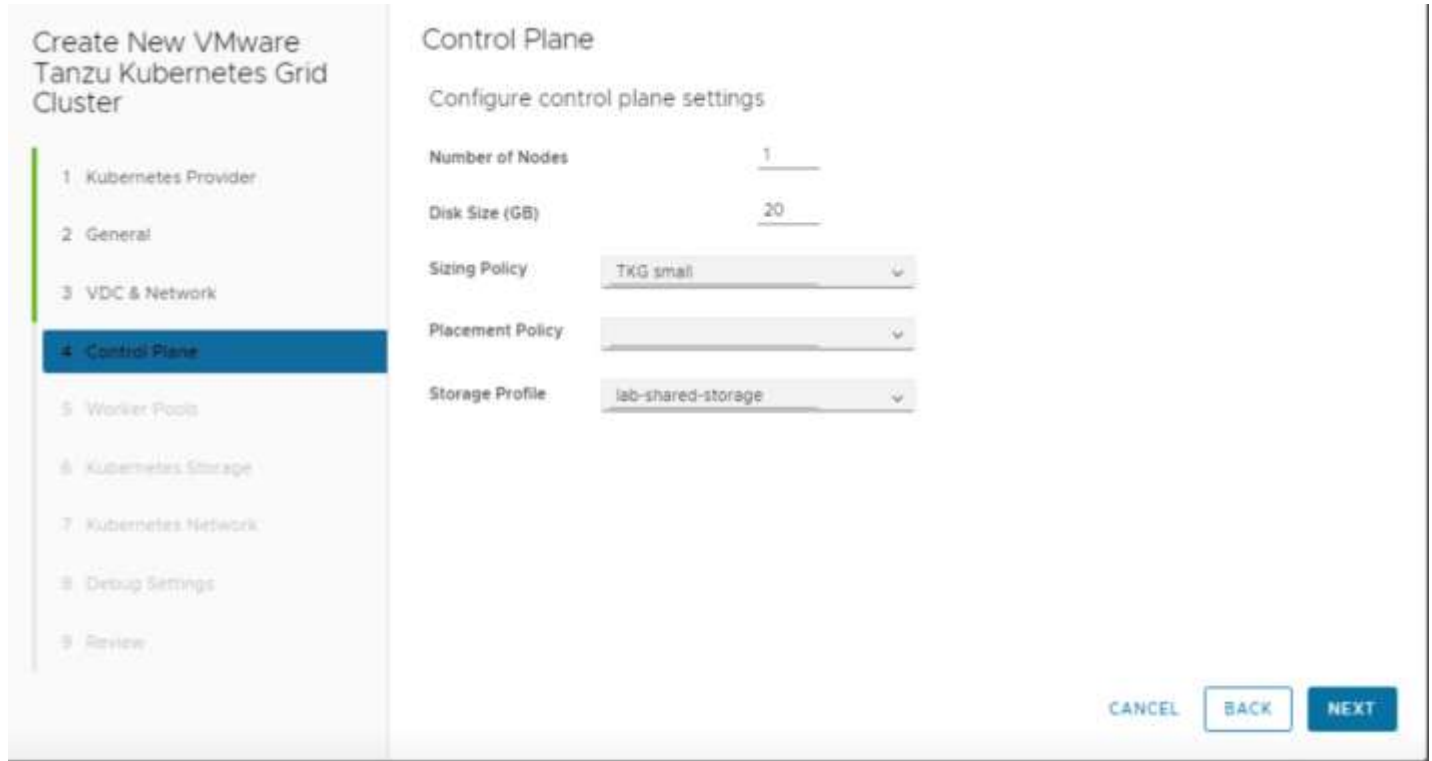
Kubernetes

ADD

<input type="checkbox"/>	Name	Description
<input type="checkbox"/>	Site A	
<input type="checkbox"/>	Site B	

### Creating Highly Available Kubernetes Clusters as Tenant User

As Tenant User, when creating the Kubernetes Cluster in the Container Service Extension UI, create additional Groups for and Worker Nodes, and select a different Placement Policy for each group.



### Create New VMware Tanzu Kubernetes Grid Cluster

- 1 Kubernetes Provider
- 2 General
- 3 VDC & Network
- 4 Control Plane
- 5 Worker Pools**
- 6 Kubernetes Storage
- 7 Kubernetes Network
- 8 Debug Settings
- 9 Review

### Worker Pools

worker-node-pool-1

Activate GPU

Name: worker-node-pool-1

Number of Nodes: 2

Disk Size: 20

Sizing Policy: TKG small

Placement Policy: Site A

Storage Profile: lab-shared-storage

DELETE

worker-node-pool-2

Activate GPU

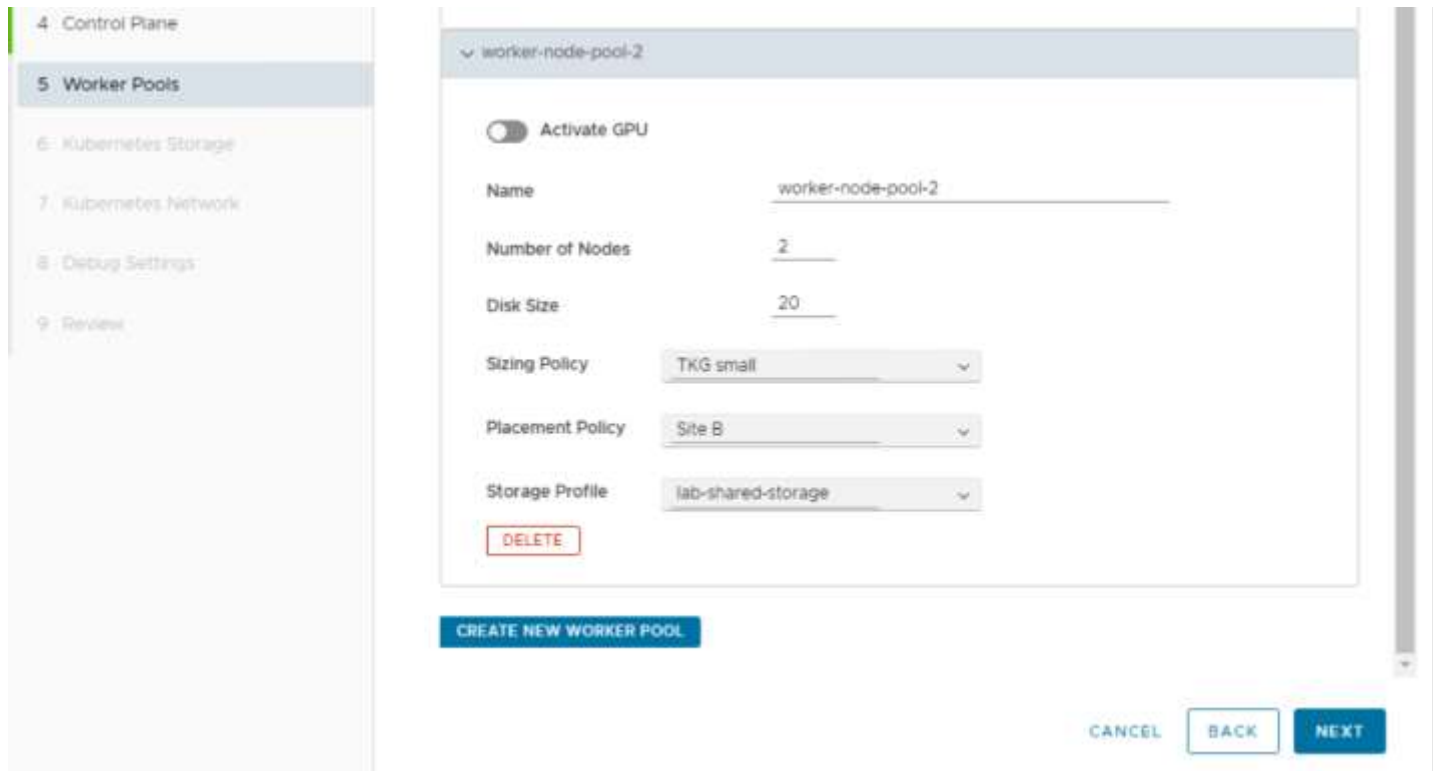
Name: worker-node-pool-2

Number of Nodes: 2

Disk Size: 20

Sizing Policy: TKG small

CANCEL BACK NEXT



### Note for stretched ESXi clusters

In case of using stretched ESXi clusters, where the hosts groups are located in different data centers with a higher probability of network connectivity loss between the data centers it is recommended to keep the Kubernetes Master nodes within the same data center to avoid inconsistencies in the etcd database.

### Conclusion

In this paper we have described a model to achieve high availability for Container Service Extension workloads in VCD with the failure domain being at the level of a ESXi host group.

