

DEPLOY A VIRTUAL MACHINE SCALE SET BEHIND A LOAD BALANCER

Scale set

Virtual Machine Scale Sets let you create and manage a group of load balanced VMs. The number of VM instances can automatically increase or decrease in response to demand or a defined schedule.

We need to create load balancer before creating a scale set, make sure both the load balancer and scale set are in the same vnet

To create a load balancer,

- Portal > search load balancer > create > enter the following details
- In basics tab,

SETTING	VALUE
Subscription	Select your subscription.
Resource group	demo-RG
Name	demo-load-balancer
Region	Central India
SKU	Standard
Tier	Regional

- In frontend IP configuration > add IP

SETTING	VALUE
Name	demo-fip
Vnet	demo-Vnet
subnet	dev4

- In backend pools,

SETTING	VALUE
Name	demo-backend
Vnet	demo-Vnet
add	vm-1

- Select review + create > create

To create a scale set,

- Portal > search scale set > create
- On the basics tab, enter the following:

Setting	Value
Subscription	Select your subscription.
Resource group	demo-RG
vm Scale set name	demp-scaleset
Region	Central india
Availability	none
orchestration mode	flexible
security type	standard
image	windows server 2019
size	standard_B1s

- After entering username and password, on the scaling tab enter the following

Initial instance count * ⓘ ✓

Scaling

Scaling policy ⓘ ☐ Manual ☒ Custom

Minimum number of instances * ⓘ

Maximum number of instances * ⓘ ✓

Scale out

CPU threshold (%) * ⓘ ✓

Duration in minutes * ⓘ ✓

Number of instances to increase by * ⓘ ✓

Scale in

CPU threshold (%) * ⓘ ✓

Number of instances to decrease by * ⓘ ✓

- In the networking tab, select the appropriate subnet where the previous load balancer is created
- Enter review + create > create

Dashboard > CreateVmss-MicrosoftWindowsServer.WindowsServer-2-20230902142830 | Overview >

demp-Scaleset ☆ ...

Virtual machine scale set

Search << → Move ▶ Start ↺ Restart □ Stop ↻ Reimage 🗑 Delete ↻ Refresh 🗨 Feedback

Overview

Activity log Access control (IAM) Tags Diagnose and solve problems

Settings

Instances Networking Scaling Disks Operating system Microsoft Defender for Cloud Size Extensions + applications Configuration Upgrade policy

Essentials JSON View

Resource group (move) : [demo-RG](#)

Status : 1 out of 1 succeeded

Location : Central India

Subscription (move) : [Free Trial](#)

Subscription ID : 4095d462-3fb3-4d5d-9eff-bc9008ac3ecd

Operating system : Windows

Size : Standard_B1s (1 instance)

Public IP address : -

Public IP address (IPv6) : -

Virtual network/subnet : [demo-Vnet1/dev4](#)

Orchestration mode : Flexible

Tags (edit) : [Add tags](#)

Properties Monitoring Capabilities (6) Recommendations Tutorials

Virtual machine profile

Operating system : Windows

Image publisher : MicrosoftWindowsServer

Image offer : WindowsServer

Image plan : 2019-datacenter-gen2

Capacity reservation group : -

Networking

Public IP address : -

Public IP address (IPv6) : -

Virtual network/subnet : [demo-Vnet1/dev4](#)

Size

Size : Standard_B1s

vCPUs : 1

RAM : 1 GiB

Availability + scaling

Dashboard > CreateVmss-MicrosoftWindowsServer.WindowsServer-2-20230902142830 | Overview > demp-Scaleset

demp-Scaleset | Instances ☆ ...

Virtual machine scale set

Search << ▶ Start ↺ Restart □ Stop ↻ Reimage 🗑 Delete ↻ Refresh

Search virtual machine instances

Instance	Computer name	Type	Status	Provisioning state
<input type="checkbox"/> demp-Scaleset_4b6f9b1e	demp-scalKXIBDQ	VM	Running	Succeeded

Overview Activity log Access control (IAM) Tags Diagnose and solve problems

Working

We can observe that a virtual machine is being created automatically when the cpu threshold is reached

Dashboard > CreateVmss-MicrosoftWindowsServer,WindowsServer-2-20230902142830 | Overview > demp-Scale

demp-Scaleset | Scaling

Virtual machine scale set

Search

Tags

Diagnose and solve problems

Settings

Instances

Networking

Scaling

Disks

Operating system

Microsoft Defender for Cloud

Size

Extensions + applications

Configuration

Upgrade policy

Health and repair

Identity

Properties

Save Discard Refresh Logs Feedback

Default* Profile1

Delete warning

The very last or default off autoscale.

Scale mode

Scale based on a metric

Rules

When demp-Scaleset

Scale in

When demp-Scaleset

+ Add a rule

Instance limits

Minimum 1

Schedule

This scale condition is executed

+ Add a scale condition

Scale rule

Metric source

Current resource (demp-Scaleset)

Resource type

Virtual machine scale sets

Resource

demp-Scaleset

Criteria

Metric namespace

Virtual Machine Host

Metric name

Percentage CPU

1 minute time grain

Dimension Name

VMName

Operator

=

Dimension Values

All values

Add

If you select multiple values for a dimension, autoscale will aggregate the metric across the selected values, not evaluate the metric for each values individually.

80% 60% 40% 20% 0%

2:32 PM 2:33 PM 2:34 PM 2:35 PM 2:36 PM UTC+05:30

Percentage CPU (Average)

51.43 %

Update Delete

Start Restart Stop Reimage Delete Refresh

Search virtual machine instances

Instance	Computer name	Type	Status	Provisioning state
<input type="checkbox"/> demp-scaleset_22f46e14	demp-scal9P5JBO	VM	Creating	Creating
<input type="checkbox"/> demp-scaleset_cc69abe1	demp-scalEBAGBJ	VM	Running	Succeeded