



Azure section 8 - Hands on - Scale Set

Start with the search bar and search for the scale set.

Home > Virtual machine scale sets

Create a virtual machine scale set

Basics Spot Disks Networking Management Health Advanced Tags Review + create

Azure 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. Scale sets provide high availability to your applications, and allow you to centrally manage, configure, and update a large number of VMs.
[Learn more about virtual machine scale sets](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription *

Resource group *

[Create new](#)

Search for the virtual machine scale set in the search bar and then give the name to the new resource group.

[Home](#) > [Virtual machine scale sets](#) >

Create a virtual machine scale set ...

Scale set details

Virtual machine scale set name *	<input type="text" value="vmsccaleset"/>
Region *	<input type="text" value="(Europe) West Europe"/>
Availability zone ⓘ	<input type="text" value="None"/>

Give the name to the vm and choose the region and availability zone below there we need to choose the operating system image and also Disk.

👤 Improve your availability by selecting multiple zones

Instance count * ⓘ	<input type="text" value="2"/>
--------------------	--------------------------------


[Configure scaling options](#)

Administrator account

Username * ⓘ	<input type="text"/>
Password * ⓘ	<input type="password"/>

Then going forward we need to choose instance count set it to 2 and then give the user name and password and also confirm password.

[Home](#) > [sclae-set-rg](#) >

 **vmsccaleset** ✕ ☆ ...
Virtual machine scale set

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

Overview

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

Essentials

Resource group (move) : [sclae-set-rg](#)

Status : **2 out of 2 succeeded**

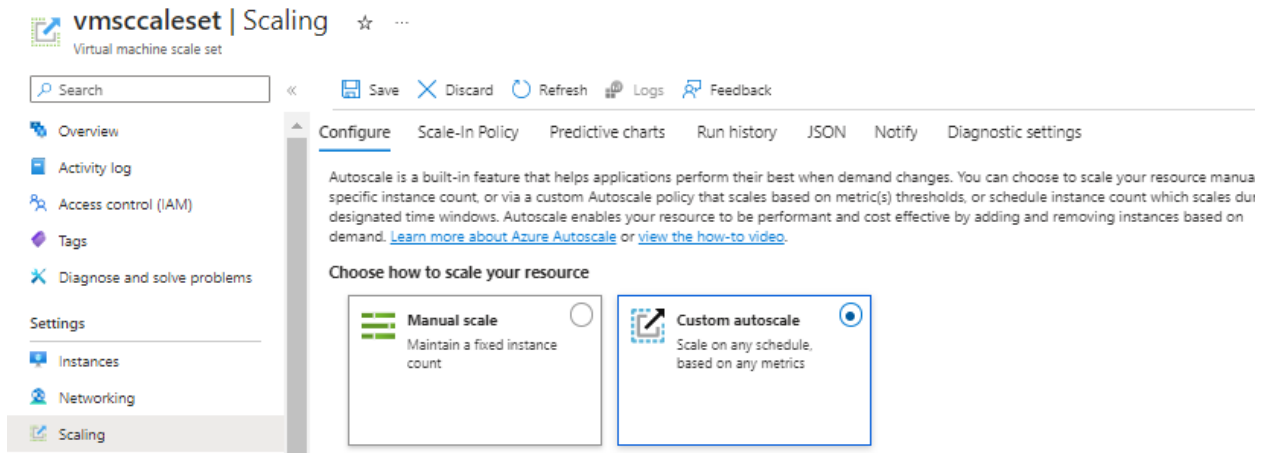
Location : West Europe

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

Subscription ID : c592185c-7514-40ad-b169-162dbc3cc538

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

Go to the resource group and then check the scale set which we just created and to that review page we will see the status that both the VMs are running.



Go to the scaling option under the scale set and then go to the custom auto-scale option and now we will configure the custom auto-scale.

Default* Auto created default scale condition

Delete warning

The very last or default recurrence rule cannot be deleted. Instead, you can disable autoscale to turn off autoscale.

Scale mode

☒ Scale based on a metric ☐ Scale to a specific instance count

Rules

Scale is based on metric trigger rules but no rule(s) is defined: click **Add a rule** to create a rule. For example: 'Add a rule that increases instance count by 1 when CPU Percentage is above 70%'. If no rules is defined, the resource will be set to default instance count.

Instance limits

Minimum * Maximum * Default *

Schedule

This scale condition is executed when none of the other scale condition(s) match

Click on the scale based on metric and then set the metric rule to increase the instance count.

Instance limits

Minimum * Maximum * Default *

Schedule

This scale condition is executed when none of the other scale condition(s) match

Now i have set the minimum value of instance to 1 and the maximum to 3 and default to 2. This means that there will be two servers by default. and when the CPU usage will increase there will be one instance added to the scale set and will make total instance count to 3. And if the default instance 2 has CPU usage less than the threshold then the one instance will get terminated and will come count back to 1.

Scale rule

Metric source: Current resource (vmscaleset)

Resource type: Virtual machine scale sets

Resource: vmscaleset

☒ Criteria

Metric namespace: Virtual Machine Host

Metric name: Percentage CPU

1 minute time grain

Dimension Name	Operator	Dimension Values	Add
VMName	=	All values	+

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.

Percentage CPU (Average): 27.48 %

Scale rule

Percentage CPU (Average): 27.48 %

☐ Enable metric divide by instance count

Operator: Greater than or equal to

Metric threshold to trigger scale action: 85 %

Duration (minutes): 10

Time grain (minutes): 1

Time grain statistic: Average

Time aggregation: Average

Action

Operation: Increase count by

Cool down (minutes): 5

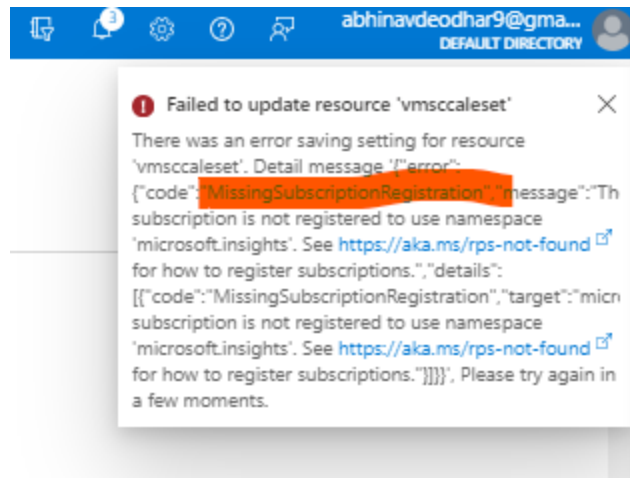
Instance count: 1

Add

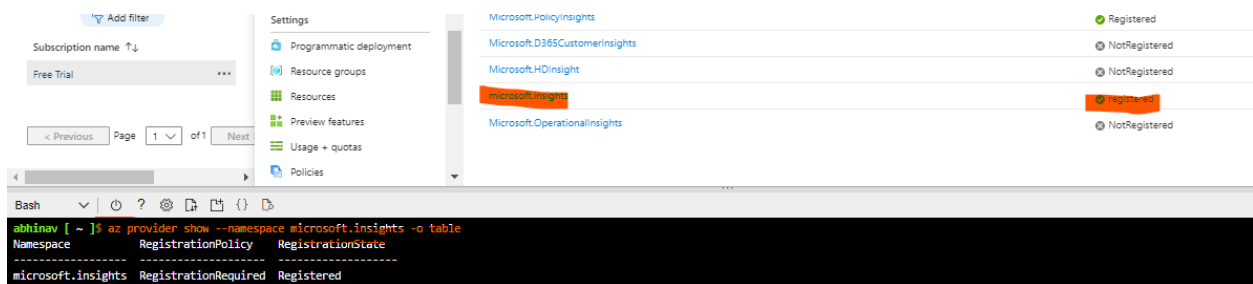
Set all parameters to set the rules.

Operator means condition of CPU it is greater than or equal to 85%. And then the second condition will be checked which is the time duration which means that the CPU usage which is at 85% or greater; does it stays that way for 10 mins or more? Then the third condition is checked that is by how much the instance count should be increased? So here it is 1 and then there is the cool down time. which means the scale set will wait for 5 mins before creating other VM if the Instance count would have been greater than 1. Which means that the Scale set will keep on adding the new VM by waiting for the 5 mins between creation of two consecutive VMS.

When we click on add and then save it then it may give us the following error.



For that go to the subscription and we need to add the rule to the subscription and then we can use this auto scaling. Go to the resources providers. And add microsoft insight policy. You can then use auto scaling.



after searching the resource provider on the page click on register on the top of the page and then use the following command to see if the registration was successful.

Scale out			
When	vmsscaleset	(Average) Percentage CPU > 85	Increase count by 2
Scale in			
When	vmsscaleset	(Average) Percentage CPU < 50	Decrease count by 1

I have also added a scale in condition which means that if the CPU value is less than 50 % then one VM will get deleted. So we know that the default value is 2 So we need to change it to 1 if we want the auto-scale to remove one instance.

Instance limits

+ Add a rule

Minimum * ⓘ

Maximum * ⓘ

Default * ⓘ

1

3

1

Schedule

This scale condition is executed when none of the other scale condition(s) match

vmsscaleset

Virtual machine scale set

Search

Move

Start

Restart

Stop

Reimage

Delete

Refresh

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Settings

Instances

Essentials

Resource group (move)

: [sclae-set-rg](#)

Status

: [1 out of 1 succeeded](#)

Location

: West Europe

Subscription (move)

: [Free Trial](#)

Subscription ID

: c592185c-7514-40ad-b169-162dbc3cc538

Tags (edit)

: [Add tags](#)

Here is the hands on on scale set.