Implement a solution to automatically scale the virtual machine resources to handle higher workloads and scale back during periods of low demand.

Create a VMSS using the Azure Portal:

Sign in to the Azure portal:

• Go to Azure portal and sign in with your Azure account.

Navigate to Virtual Machine Scale Sets:

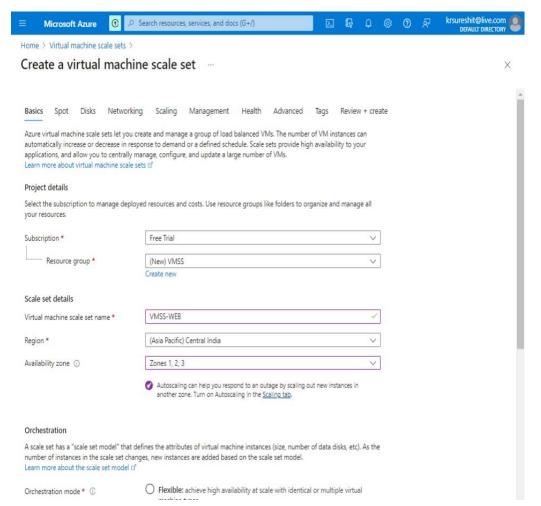
• In the left sidebar, click on "Virtual machine scale sets."

Click on "+ Add" to create a new VMSS:

• Click the "+ Add" button to start the creation process.

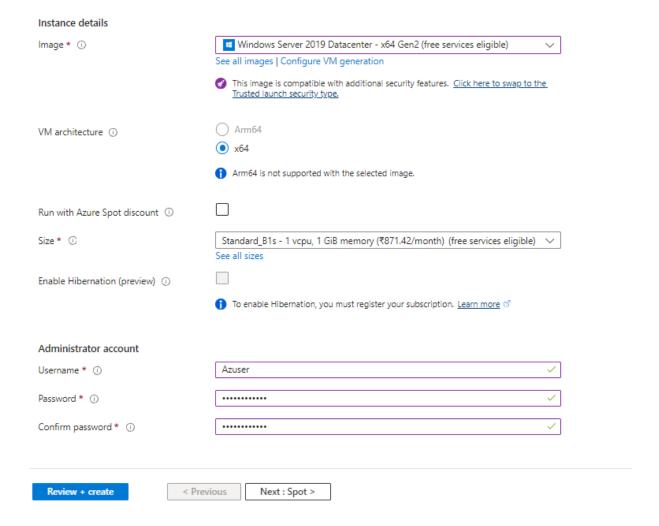
Basic settings:

- Fill in the basic settings for the VMSS, including:
- Subscription: Choose the Azure subscription you want to use.
- Resource group: Create a new one or use an existing resource group.
- VMSS Name: Give your VMSS a unique name.
- Region: Select the Azure region for your VMSS.
- Operating System: Choose the operating system for your VMSS instances.



Instance details:

- Configure instance details such as:
- Image: Choose the operating system image for your VMSS instances.
- Size: Select the size of the virtual machines in the scale set.



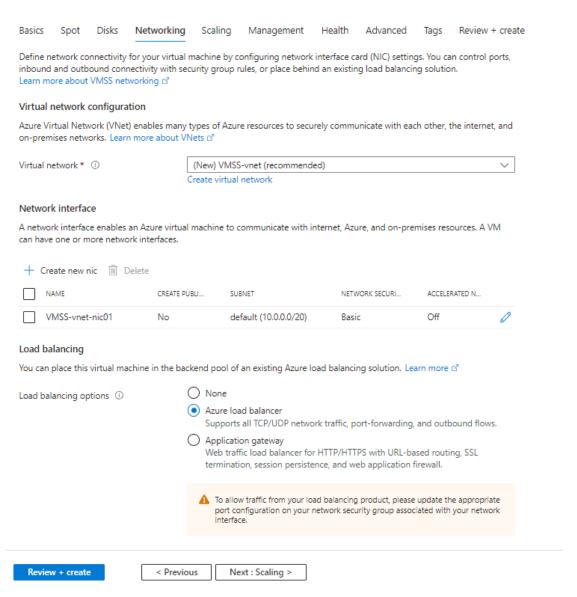
Network settings:

- Configure network settings, including:
- Virtual network: Choose the virtual network for your VMSS.
- Subnet: Select the subnet within the virtual network.
- Public IP address: Configure if needed.



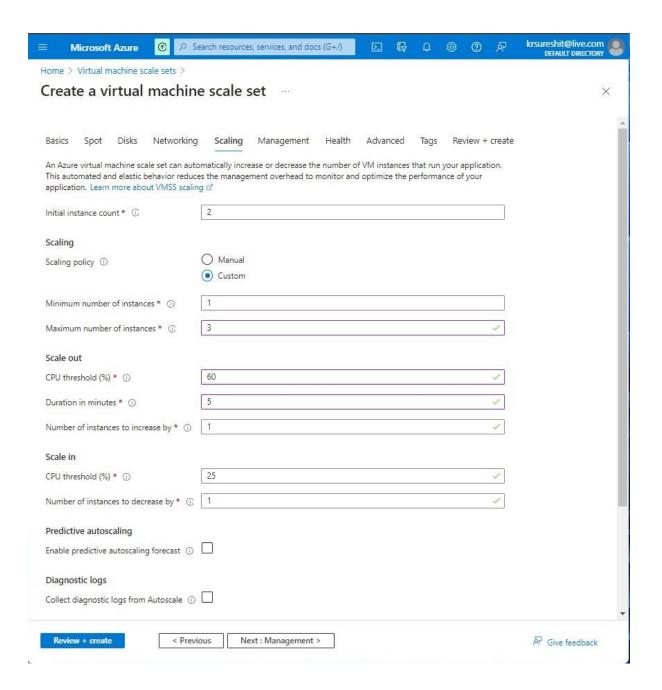
Home > Virtual machine scale sets >

Create a virtual machine scale set



Scaling:

- Set the scaling options, including:
- Instance count: Define the initial and maximum number of instances.
- Automatic scaling: Configure auto-scaling rules



Management and Advanced settings:

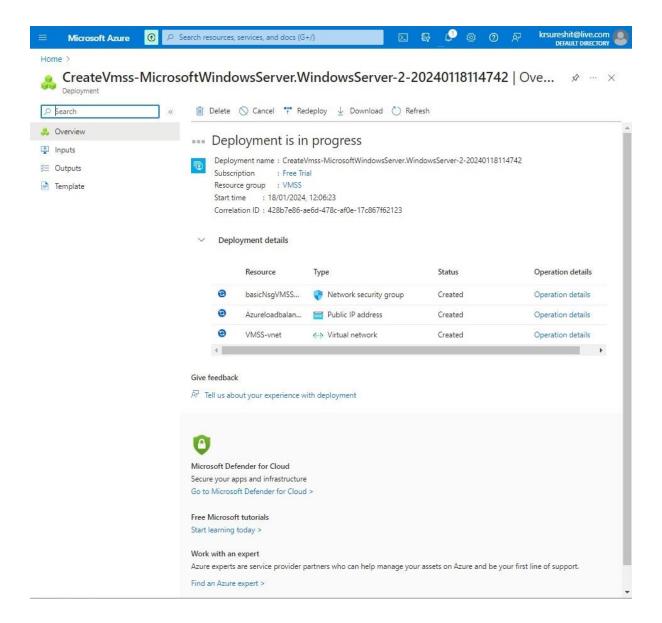
 Configure additional settings such as boot diagnostics, extensions, and availability options.

Tags:

• Add tags if you want to categorize your VMSS.

Review and create:

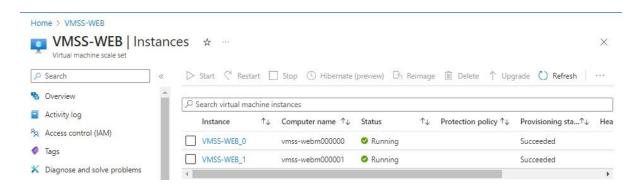
• Review the summary of your configuration. If everything looks good, click the "Create" button to start the deployment.



Deployment:

• Once the deployment is complete, you'll see a notification. You can also navigate to the resource group to check the deployment status.

• Initially it has two instances running



• Based on the workload it has automatically scale out to 1 instance:

