



Basic Load Balancer VM Scale Set

In this process, we are deploying a Virtual Machine Scale Set (VMSS) and configuring an Azure Load Balancer to enhance the availability and scalability of a web application. We set up the VMSS with the necessary configurations and use a custom script to install IIS on each instance. Then, we configure the Azure Load Balancer to distribute incoming traffic across the VMSS instances, ensuring efficient traffic management and high availability. The end goal is to create a robust infrastructure that automatically scales and balances the load to maintain optimal performance and reliability for the web application.

1. In your Azure Portal from the marketplace search for the virtual machine scale set and choose this service accordingly.

The screenshot shows the Azure Marketplace search results for 'Virtual machine scale set'. At the top, there is a Microsoft logo and a 'Virtual machine scale set' title with a '...' button. Below the title, it says 'Microsoft | Azure Service' and has a rating of '★ 3.9 (119 ratings)'. There is a 'Plan' section with a dropdown menu set to 'Virtual machine scale set' and a 'Create' button. To the right, there is a 'Add to Favorites' link.

2. First you have to choose the resource group.

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 *	<input type="text" value="Azure Pass - Sponsorship"/>
Resource group *	<input type="text" value="new-grp"/> Create new

3. Then you have to give your machine a name then choose your region after that choose security type as standard.

Scale set details

Virtual machine scale set name *	new-scale-set	✓
Region *	(Europe) North Europe	▼
Availability zone ⓘ	None	▼

Orchestration

A scale set has a "scale set model" that defines the attributes of virtual machine instances (size, number of data disks, etc). As the number of instances in the scale set changes, new instances are added based on the scale set model.
[Learn more about the scale set model ↗](#)

Orchestration mode * ⓘ	<input type="radio"/> Flexible: achieve high availability at scale with identical or multiple virtual machine types
	<input checked="" type="radio"/> Uniform: optimized for large scale stateless workloads with identical instances
Security type ⓘ	Standard

4. Then choose the instance count to 1 and then choose the image as shown below. After that choose its size.

Instance count * ⓘ	1	✓
Configure scaling options		
Instance details		
Image * ⓘ	Windows Server 2022 Datacenter - x64 Gen2	▼
See all images Configure VM generation		
This image is compatible with additional security features. Click here to swap to the Trusted launch security type.		
VM architecture ⓘ	<input type="radio"/> Arm64	
	<input checked="" type="radio"/> x64	
Arm64 is not supported with the selected image.		
Size * ⓘ	Standard_DS1_v2 - 1 vcpu, 3.5 GiB memory (₹7,044.92/month)	▼
See all sizes		

5. Then you have to give the username and password and move to the next page. In the networking setting scroll down to the network interface and click on edit button.

Network interface

A network interface enables an Azure virtual machine to communicate with internet, Azure, and on-premises resources. A VM can have one or more network interfaces.

Create new nic	Delete			
NAME	CREATE PUBLI...	SUBNET	NETWORK SECURI...	ACCELERATED N...
new-grp-vnet-nic01	Yes	default (10.0.0.0/20)	Basic	On

6. In there you have to choose allow selected ports in public inbound ports and choose port 80 for HTTP and then disable public IP and accelerated networking and then click on OK.

Public inbound ports * ⓘ

None

Allow selected ports

Select inbound ports *

HTTP (80)

Public IP address ⓘ

Disabled Enabled

Accelerated networking ⓘ

Disabled Enabled

OK **Cancel**

7. Then in the advanced section you have to choose Fixed spreading in spreading algorithm and turn off the allocation policy.

Basics Spot Disks Networking Management Health **Advanced** Tags Review + create

Add additional configuration, agents, scripts or applications via virtual machine extensions or cloud-init.

Allocation policy

Enable scaling beyond 100 instances ⓘ

Spreading algorithm ⓘ

Max spreading

Fixed spreading

Fault domain count * ⓘ 2

8. Then just move to the review page and create your resources.
9. Now we also to add a custom script extension on our machine so that web server (IIS) should be installed in our machines.
10. For that you have to go to storage accounts and create a new storage account.
11. Now you just need to give it a name choose your resource group and choose LRS in redundancy.

Subscription *

Resource group * [Create new](#)

Instance details

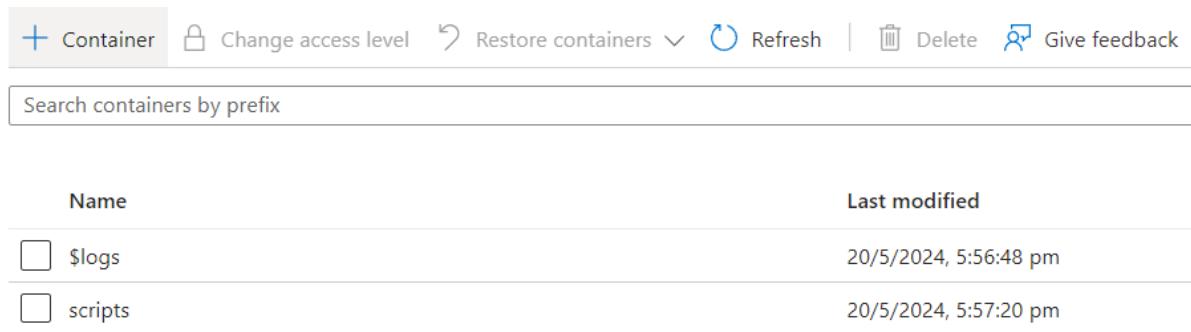
Storage account name *

Region * [Deploy to an Azure Extended Zone](#)

Performance * **Standard:** Recommended for most scenarios (general-purpose v2 account)
 Premium: Recommended for scenarios that require low latency.

Redundancy *

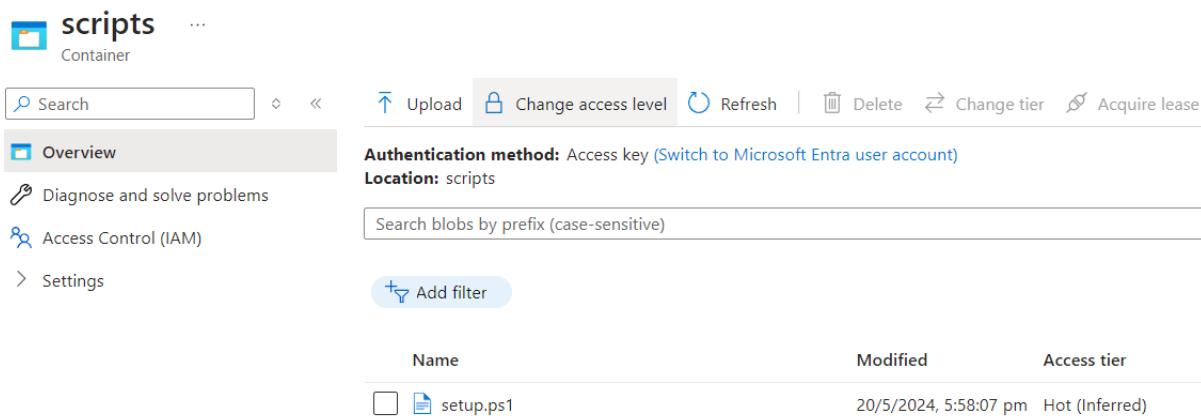
12. Once your storage account is in place then you have to upload a script file on it. First you are going to create a container to store that script in your storage account.



The screenshot shows the Azure Storage Container list. At the top, there are navigation links: '+ Container', 'Change access level', 'Restore containers', 'Refresh', 'Delete', and 'Give feedback'. Below this is a search bar labeled 'Search containers by prefix'. The main area displays a table with two rows:

Name	Last modified
\$logs	20/5/2024, 5:56:48 pm
scripts	20/5/2024, 5:57:20 pm

13. Then just upload your file in your container. You can get this file from GitHub.



The screenshot shows the Azure Storage Container Overview page for the 'scripts' container. At the top, there are navigation links: 'Search', 'Upload', 'Change access level', 'Refresh', 'Delete', 'Change tier', and 'Acquire lease'. On the left, there is a sidebar with links: 'Overview' (which is selected), 'Diagnose and solve problems', 'Access Control (IAM)', and 'Settings'. The main area displays the following information:

- Authentication method:** Access key ([Switch to Microsoft Entra user account](#))
- Location:** scripts
- A search bar labeled 'Search blobs by prefix (case-sensitive)'.
- A button labeled '+ Add filter'.
- A table showing the blob 'setup.ps1' with its details:

Name	Modified	Access tier
setup.ps1	20/5/2024, 5:58:07 pm	Hot (Inferred)

14. Now go to your scale set and directly navigate to Extension and application. Here click on add.

The screenshot shows the Azure portal interface for managing a Virtual Machine Scale Set. The top navigation bar includes the name 'demo-scale-set | Extensions + applications', a star icon, and three dots. On the left, a sidebar lists various management options: Tags, Diagnose and solve problems, Instances, Networking, Settings, Disks, Extensions + applications (which is highlighted with a red box), and Operating system. The main content area has tabs for 'Extensions' (selected) and 'VM Applications'. Below these are a search bar ('Search') and a refresh button ('Refresh'). A filter bar allows searching for items. The main list area is titled 'Name' and displays the message 'No resource extensions found.'

15. Then search for custom script extension choose it and click on next.

The screenshot shows the Azure Marketplace search results for 'Custom script extension'. The search bar at the top contains the query 'Custom script extension'. Below the search bar, there is one search result card. The result is for 'Custom script extension' by Microsoft Corp. The card includes a blue square icon with a white Windows logo, the product name 'Custom script extension', the provider 'Microsoft Corp.', and a brief description 'Custom script handler extension for windows'. At the bottom of the result card, there is a blue 'Load more' button.

16. Then you need to browse for the script file and click on create

Configure Custom script extension Extension ...

Create

Script file (Required) * ⓘ "setup.ps1" ✓ Browse

Arguments (Optional) ⓘ

17. Now come back to your scale set and go to instances, then you need to select your instance and choose to upgrade it. This will make sure that the changes that you have made apply to it too.

Name	Computer name	Status	Type	Provisioning state
demo-scale-set_759eeaa6d	demo-scalVMVPT9	Running	VM	Succeeded

18. Then scroll down from the left pane and choose upgrade policy from the operation tab. Then keep it to automatic from manual.

Set the upgrade policy to determine how VM instances are brought up-to-date with the latest scale set model.

Upgrade policy

Upgrade mode Automatic - Instances will start upgrading immediately...

i Azure will install an extension to manage the automatic guest patching process. You do not need to manually install or update this extension.



Implementation of Load Balancer

- As we have our scale set ready, we will create our load balancer. While creating your load balancer choose the configurations shown below.

Project details

Subscription *

Azure Pass - Sponsorship



Resource group *

new-grp



[Create new](#)

Instance details

Name *

demoLoadBalancerVM



Region *

North Europe



SKU * ⓘ

- Standard (Recommended)
- Gateway
- Basic (Retiring soon)

Microsoft recommends Standard SKU load balancer for production workloads;
Basic SKU will be retired on September 30, 2025. [Learn more.](#) ↗

Type * ⓘ

- Public
- Internal

Tier *

- Regional
- Global

2. Then move to next page and create frontend ip address.

Add frontend IP configuration



demoLoadBalancerVM

Name *

frontend-ip

IP version

IPv4

IPv6

Public IP address *

(new) load-ip2



[Create new](#)

3. After that just move to the review page and create your resource. Now in your load balancer go to backend pools and click on add.

The screenshot shows the Azure portal interface for managing a load balancer. The top navigation bar includes the load balancer name 'demoLoadBalancerVM | Backend pools' and standard navigation icons. On the left, a sidebar lists various management options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, and Settings (with sub-options for Frontend IP configuration and Backend pools). The main content area is titled 'Backend pool' and contains a table with one row. The table columns are 'Backend pool', 'Resource Name', and 'IP address'. The single entry is 'PoolA' under 'Resource Name'. Below the table, a note explains that the backend pool is a critical component for serving traffic based on load-balancing rules. A search bar and a 'Add filter' button are also present.

4. Then here you have to give it a name and then choose your virtual network, after that click on add in IP configurations and choose your scale set.

This screenshot shows the 'Add backend pool' dialog box. At the top, there's a 'Name' field containing 'PoolA' and a 'Virtual network' dropdown menu showing 'new-grp-vnet (new-grp)'. Below this, the 'IP configurations' section displays a table with one row. The columns are 'Resource Name', 'Resource group', 'Type', 'IP configuration', 'IP Address', and 'Availability set'. The data in the table is: 'new-scale-set', 'new-grp', 'Virtual machine scale set', 'new-grp-vnet-nic01-defaultIpConfigura...', and '-' respectively. There are '+ Add' and 'Remove' buttons at the bottom of the IP configurations table.

5. Then you have to go to the health probe and add one.

This screenshot shows the 'Add health probe' dialog. It includes a note about health probes: 'Health probes are used to check the status of a backend pool instance. If the health probe fails to get a response from a backend instance then no new connections will be sent to that backend instance until the health probe succeeds again.' The form fields are: 'Name' (ProbeA), 'Protocol' (TCP), 'Port' (80), and 'Interval (seconds)' (5).

Name *	<input type="text" value="ProbeA"/>
Protocol *	<input type="text" value="TCP"/>
Port *	<input type="text" value="80"/>
Interval (seconds) *	<input type="text" value="5"/>
Used by *	Not used

6. Now go to load balancing rules and add the rule, here you have to give it a name choose your frontend Ip address, and choose your backend pool.

Add load balancing rule ...

demoLoadBalancer

Name *

IP Version * IPv4 IPv6

Frontend IP address *

Backend pool *

Protocol TCP UDP

Port *

Backend port *

Health probe * [Create new](#)

7. Now go back to frontend IP and you will see your public IP address. Copy it and paste it in a new tab.

demoLoadBalancer | Frontend IP configuration ⋮

Load balancer

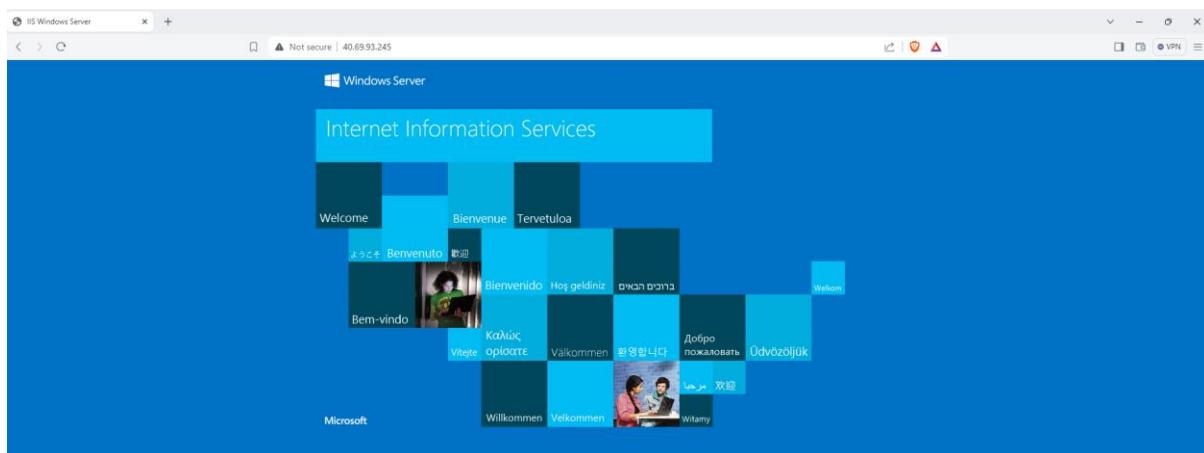
Search Add Refresh Give feedback

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

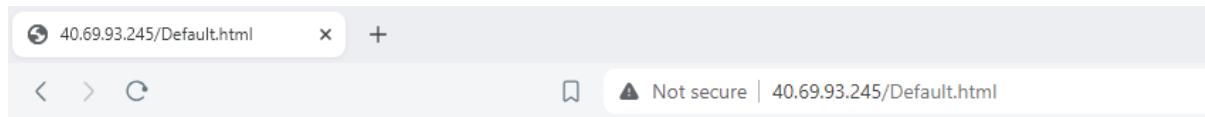
Frontend IP configuration

Name ↑↓	IP address ↑↓	Rules count ↑↓
frontend-ip	40.69.93.245 (load-ip)	1

8. Normally it will show you the web server page.



9. But if you append it with default.html then you can see that too.



This is the server new-scale000000