



## ***Windows PowerShell Get-Help on Cmdlet 'New-AzMLWorkspaceOnlineDeployment'***

***PS:\>Get-HELP New-AzMLWorkspaceOnlineDeployment -Full***

### NAME

New-AzMLWorkspaceOnlineDeployment

### SYNOPSIS

Create or update Inference Endpoint Deployment (asynchronous).

### SYNTAX

```
New-AzMLWorkspaceOnlineDeployment -EndpointName <String> -Name <String> -ResourceGroupName <String>
-WorkspaceName <String> [-SubscriptionId <String>]
-EndpointComputeType <EndpointComputeType> -Location <String> [-AppInsightsEnabled] [-CodeId <String>]
[-CodeScoringScript <String>] [-Description <String>]
[-EndpointDeploymentPropertiesBaseProperty <Hashtable>] [-EnvironmentId <String>] [-EnvironmentVariable
<Hashtable>] [-IdentityType <ManagedServiceIdentityType>]
[-IdentityUserAssigned <Hashtable>] [-InstanceType <String>] [-Kind <String>] [-LivenessProbeFailureThreshold <Int32>]
[-LivenessProbeInitialDelay <TimeSpan>]
[-LivenessProbePeriod <TimeSpan>] [-LivenessProbeSuccessThreshold <Int32>] [-LivenessProbeTimeout <TimeSpan>]
[-Model <String>] [-ModelMountPath <String>]
[-ReadinessProbeFailureThreshold <Int32>] [-ReadinessProbeInitialDelay <TimeSpan>] [-ReadinessProbePeriod
<TimeSpan>] [-ReadinessProbeSuccessThreshold <Int32>]
```

[-ReadinessProbeTimeout <TimeSpan>] [-RequestMaxConcurrentPerInstance <Int32>] [-RequestMaxQueueWait  
 <TimeSpan>] [-RequestTimeout <TimeSpan>] [-ScaleType <ScaleType>]  
 [-SkuCapacity <Int32>] [-SkuFamily <String>] [-SkuName <String>] [-SkuSize <String>] [-SkuTier <SkuTier>] [-Tag  
 <Hashtable>] [-DefaultProfile <PSObject>] [-AsJob]  
 [-Break] [-HttpPipelineAppend <SendAsyncStep[]>] [-HttpPipelinePrepend <SendAsyncStep[]>] [-NoWait] [-Proxy <Uri>]  
 [-ProxyCredential <PSCredential>]  
 [-ProxyUseDefaultCredentials] [-WhatIf] [-Confirm] [<CommonParameters>]

## DESCRIPTION

Create or update Inference Endpoint Deployment (asynchronous).

## PARAMETERS

-EndpointName <String>

Inference endpoint name.

Required? true

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-Name <String>

Inference Endpoint Deployment name.

Required? true

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-ResourceGroupName <String>

The name of the resource group.

The name is case insensitive.

Required? true

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-WorkspaceName <String>

Name of Azure Machine Learning workspace.

Required? true

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-SubscriptionId <String>

The ID of the target subscription.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-EndpointComputeType <EndpointComputeType>

[Required] The compute type of the endpoint.

Required? true

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -Location <String>

The geo-location where the resource lives

Required? true

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -AppInsightsEnabled [<SwitchParameter>]

If true, enables Application Insights logging.

Required? false

Position? named

Default value False

Accept pipeline input? false

Accept wildcard characters? false

#### -CodeId <String>

ARM resource ID of the code asset.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -CodeScoringScript <String>

[Required] The script to execute on startup.

eg.

"score.py"

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -Description <String>

Description of the endpoint deployment.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -EndpointDeploymentPropertiesBaseProperty <Hashtable>

Property dictionary.

Properties can be added, but not removed or altered.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -EnvironmentId <String>

ARM resource ID or AssetId of the environment specification for the endpoint deployment.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -EnvironmentVariable <Hashtable>

Environment variables configuration for the deployment.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -IdentityType <ManagedServiceIdentityType>

Type of managed service identity (where both SystemAssigned and UserAssigned types are allowed).

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -IdentityUserAssigned <Hashtable>

The set of user assigned identities associated with the resource.

The userAssignedIdentities dictionary keys will be ARM resource ids in the form:

`/subscriptions/{subscriptionId}/resourceGroups/{resourceGroupName}/providers/Microsoft.ManagedIdentity/userAssignedIdentities/{identityName}.`

The dictionary values can be empty objects ({} in requests.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-InstanceType <String>

Compute instance type.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-Kind <String>

Metadata used by portal/tooling/etc to render different UX experiences for resources of the same type.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-LivenessProbeFailureThreshold <Int32>

The number of failures to allow before returning an unhealthy status.

Required? false

Position? named

Default value 0

Accept pipeline input? false

Accept wildcard characters? false

-LivenessProbeInitialDelay <TimeSpan>

The delay before the first probe in ISO 8601 format.

Required? false

Position?                named  
Default value  
Accept pipeline input?    false  
Accept wildcard characters? false

-LivenessProbePeriod <TimeSpan>

The length of time between probes in ISO 8601 format.

Required?                false  
Position?                named  
Default value  
Accept pipeline input?    false  
Accept wildcard characters? false

-LivenessProbeSuccessThreshold <Int32>

The number of successful probes before returning a healthy status.

Required?                false  
Position?                named  
Default value            0  
Accept pipeline input?    false  
Accept wildcard characters? false

-LivenessProbeTimeout <TimeSpan>

The probe timeout in ISO 8601 format.

Required?                false  
Position?                named  
Default value  
Accept pipeline input?    false  
Accept wildcard characters? false

-Model <String>

The URI path to the model.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -ModelMountPath <String>

The path to mount the model in custom container.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -ReadinessProbeFailureThreshold <Int32>

The number of failures to allow before returning an unhealthy status.

Required? false

Position? named

Default value 0

Accept pipeline input? false

Accept wildcard characters? false

#### -ReadinessProbeInitialDelay <TimeSpan>

The delay before the first probe in ISO 8601 format.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-ReadinessProbePeriod <TimeSpan>

The length of time between probes in ISO 8601 format.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-ReadinessProbeSuccessThreshold <Int32>

The number of successful probes before returning a healthy status.

Required? false

Position? named

Default value 0

Accept pipeline input? false

Accept wildcard characters? false

-ReadinessProbeTimeout <TimeSpan>

The probe timeout in ISO 8601 format.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-RequestMaxConcurrentPerInstance <Int32>

The number of maximum concurrent requests per node allowed per deployment.

Defaults to 1.

Required? false  
Position? named  
Default value 0  
Accept pipeline input? false  
Accept wildcard characters? false

-RequestMaxQueueWait <TimeSpan>

The maximum amount of time a request will stay in the queue in ISO 8601 format.Defaults to 500ms.

Required? false  
Position? named  
Default value  
Accept pipeline input? false  
Accept wildcard characters? false

-RequestTimeout <TimeSpan>

The scoring timeout in ISO 8601 format.Defaults to 5000ms.

Required? false  
Position? named  
Default value  
Accept pipeline input? false  
Accept wildcard characters? false

-ScaleType <ScaleType>

[Required] Type of deployment scaling algorithm

Required? false  
Position? named  
Default value  
Accept pipeline input? false  
Accept wildcard characters? false

#### -SkuCapacity <Int32>

If the SKU supports scale out/in then the capacity integer should be included.

If scale out/in is not possible for the resource this may be omitted.

Required? false

Position? named

Default value 0

Accept pipeline input? false

Accept wildcard characters? false

#### -SkuFamily <String>

If the service has different generations of hardware, for the same SKU, then that can be captured here.

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -SkuName <String>

The name of the SKU.

Ex - P3.

It is typically a letter+number code

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

#### -SkuSize <String>

The SKU size.

When the name field is the combination of tier and some other value, this would be the standalone code. *Page 12/17*

Required? false  
Position? named  
Default value  
Accept pipeline input? false  
Accept wildcard characters? false

#### -SkuTier <SkuTier>

This field is required to be implemented by the Resource Provider if the service has more than one tier, but is not required on a PUT.

Required? false  
Position? named  
Default value  
Accept pipeline input? false  
Accept wildcard characters? false

#### -Tag <Hashtable>

Resource tags.

Required? false  
Position? named  
Default value  
Accept pipeline input? false  
Accept wildcard characters? false

#### -DefaultProfile <PSObject>

The credentials, account, tenant, and subscription used for communication with Azure.

Required? false  
Position? named  
Default value  
Accept pipeline input? false

Accept wildcard characters? false

-AsJob [<SwitchParameter>]

Run the command as a job

Required? false

Position? named

Default value False

Accept pipeline input? false

Accept wildcard characters? false

-Break [<SwitchParameter>]

Wait for .NET debugger to attach

Required? false

Position? named

Default value False

Accept pipeline input? false

Accept wildcard characters? false

-HttpPipelineAppend <SendAsyncStep[]>

SendAsync Pipeline Steps to be appended to the front of the pipeline

Required? false

Position? named

Default value

Accept pipeline input? false

Accept wildcard characters? false

-HttpPipelinePrepend <SendAsyncStep[]>

SendAsync Pipeline Steps to be prepended to the front of the pipeline

Required? false

Position?                named  
Default value  
Accept pipeline input?    false  
Accept wildcard characters? false

**-NoWait [<SwitchParameter>]**

Run the command asynchronously

Required?                false  
Position?                named  
Default value             False  
Accept pipeline input?    false  
Accept wildcard characters? false

**-Proxy <Uri>**

The URI for the proxy server to use

Required?                false  
Position?                named  
Default value  
Accept pipeline input?    false  
Accept wildcard characters? false

**-ProxyCredential <PSCredential>**

Credentials for a proxy server to use for the remote call

Required?                false  
Position?                named  
Default value  
Accept pipeline input?    false  
Accept wildcard characters? false

**-ProxyUseDefaultCredentials [<SwitchParameter>]**

Use the default credentials for the proxy

|                             |       |
|-----------------------------|-------|
| Required?                   | false |
| Position?                   | named |
| Default value               | False |
| Accept pipeline input?      | false |
| Accept wildcard characters? | false |

-WhatIf [<SwitchParameter>]

|                             |       |
|-----------------------------|-------|
| Required?                   | false |
| Position?                   | named |
| Default value               |       |
| Accept pipeline input?      | false |
| Accept wildcard characters? | false |

-Confirm [<SwitchParameter>]

|                             |       |
|-----------------------------|-------|
| Required?                   | false |
| Position?                   | named |
| Default value               |       |
| Accept pipeline input?      | false |
| Accept wildcard characters? | false |

<CommonParameters>

This cmdlet supports the common parameters: Verbose, Debug, ErrorAction, ErrorVariable, WarningAction, WarningVariable, OutBuffer, PipelineVariable, and OutVariable. For more information, see [about\\_CommonParameters \(https://go.microsoft.com/fwlink/?LinkID=113216\)](https://go.microsoft.com/fwlink/?LinkID=113216).

## INPUTS

## OUTPUTS

----- EXAMPLE 1 -----

```
PS C:\>New-AzMLWorkspaceOnlineDeployment -ResourceGroupName ml-rg-test -WorkspaceName mlworkspace-cli01
-EndpointName online-pwsh01 -Name blue -Location "eastus"
-EndpointComputeType 'Managed' `

                                                                    -CodeId
"/subscriptions/9e223dbe-3399-4e19-88eb-0975f02ac87f/resourceGroups/ml-rg-test/providers/Microsoft.MachineLearningS
ervices/workspaces/mlworkspace-cli01/codes/7
87fc793-1ac7-414e-a035-7248767b7b23/versions/1" -CodeScoringScript "score.py" `

                                                                    -EnvironmentId
"/subscriptions/9e223dbe-3399-4e19-88eb-0975f02ac87f/resourceGroups/ml-rg-test/providers/Microsoft.MachineLearningS
ervices/workspaces/mlworkspace-cli01/
environments/CliV2AnonymousEnvironment/versions/8a424b013f5b0177929a1697d772da41" `

                                                                    -Model
"/subscriptions/9e223dbe-3399-4e19-88eb-0975f02ac87f/resourceGroups/ml-rg-test/providers/Microsoft.MachineLearningS
ervices/workspaces/mlworkspace-cli01/models/a
99089c5-23a6-4431-9ecd-37c70f01c9bc/versions/1" -InstanceType "Standard_F2s_v2" `
-SkuName "Default" -SkuCapacity 1
```

RELATED LINKS

<https://learn.microsoft.com/powershell/module/az.machinelearningservices/new-azmlworkspaceonlinedeployment>