



# Setup

## NetApp Solutions

NetApp  
September 19, 2023

This PDF was generated from [https://docs.netapp.com/us-en/netapp-solutions/ai/mlrun\\_configuring\\_kubernetes\\_cluster.html](https://docs.netapp.com/us-en/netapp-solutions/ai/mlrun_configuring_kubernetes_cluster.html) on September 19, 2023. Always check [docs.netapp.com](https://docs.netapp.com) for the latest.

# Table of Contents

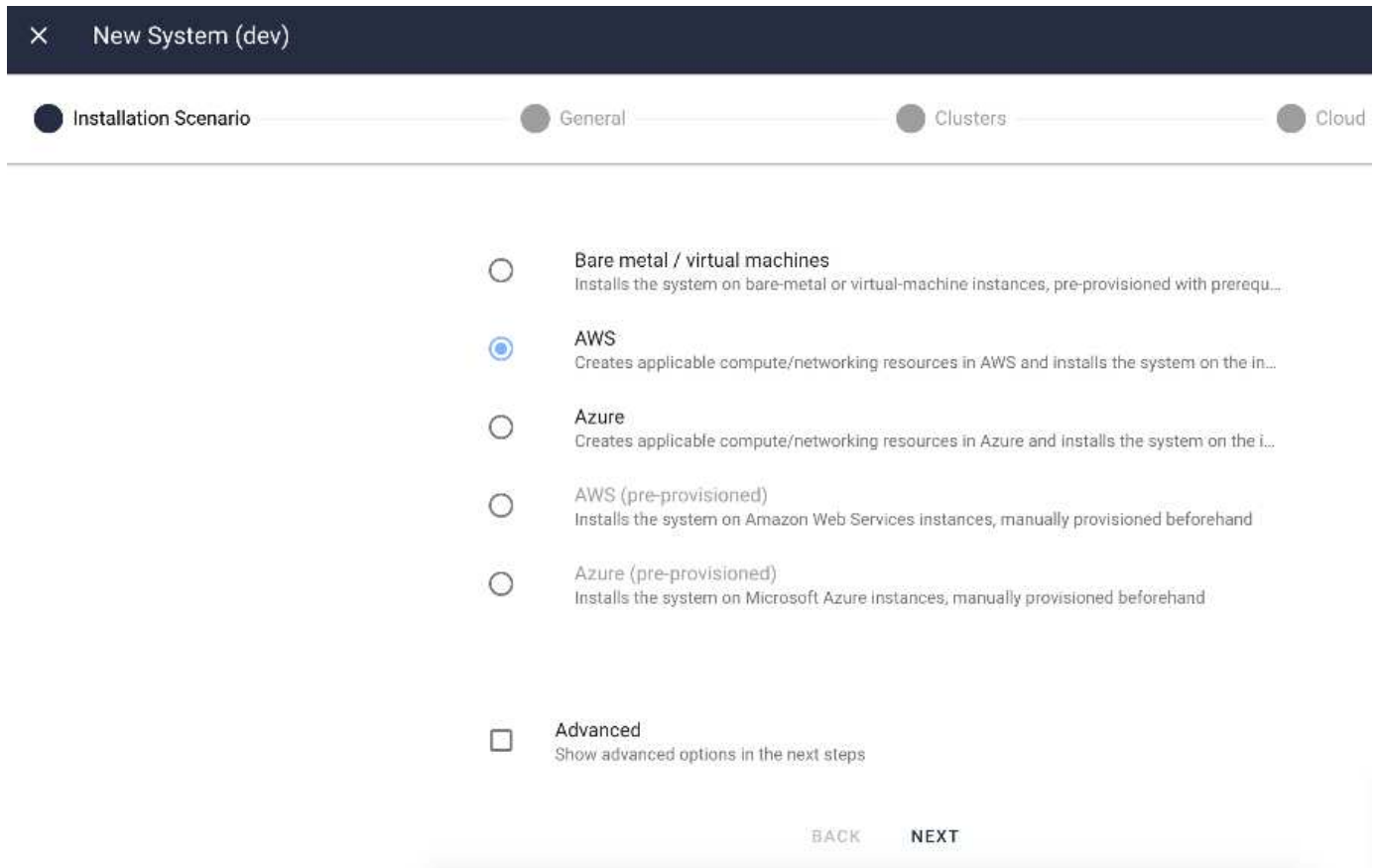
- Setup Overview ..... 1
- Iguazio Installation ..... 1
- Configuring Kubernetes Cluster ..... 1
- Define Persistent Volume Claim..... 3

# Setup Overview

## Iguazio Installation

Iguazio can be installed on-premises or on a cloud provider. Provisioning can be done as a service and managed by Iguazio or by the customer. In both cases, Iguazio provides a deployment application (Provazio) to deploy and manage clusters.

For on-premises installation, please refer to [NVA-1121](#) for compute, network, and storage setup. On-premises deployment of Iguazio is provided by Iguazio without additional cost to the customer. See [this page](#) for DNS and SMTP server configurations. The Provazio installation page is shown as follows.



The screenshot shows the 'New System (dev)' window in the Iguazio installation wizard. The 'Installation Scenario' tab is active, displaying a list of installation options. The 'AWS' option is selected, indicated by a blue radio button. The 'Advanced' option is unchecked. The 'BACK' and 'NEXT' buttons are visible at the bottom right.

Installation Scenario	General	Clusters	Cloud
<input checked="" type="radio"/> Bare metal / virtual machines Installs the system on bare-metal or virtual-machine instances, pre-provisioned with prerequ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input checked="" type="radio"/> AWS Creates applicable compute/networking resources in AWS and installs the system on the in...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/> Azure Creates applicable compute/networking resources in Azure and installs the system on the i...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/> AWS (pre-provisioned) Installs the system on Amazon Web Services instances, manually provisioned beforehand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/> Azure (pre-provisioned) Installs the system on Microsoft Azure instances, manually provisioned beforehand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> Advanced Show advanced options in the next steps	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

[Next: Configuring Kubernetes Cluster](#)

## Configuring Kubernetes Cluster

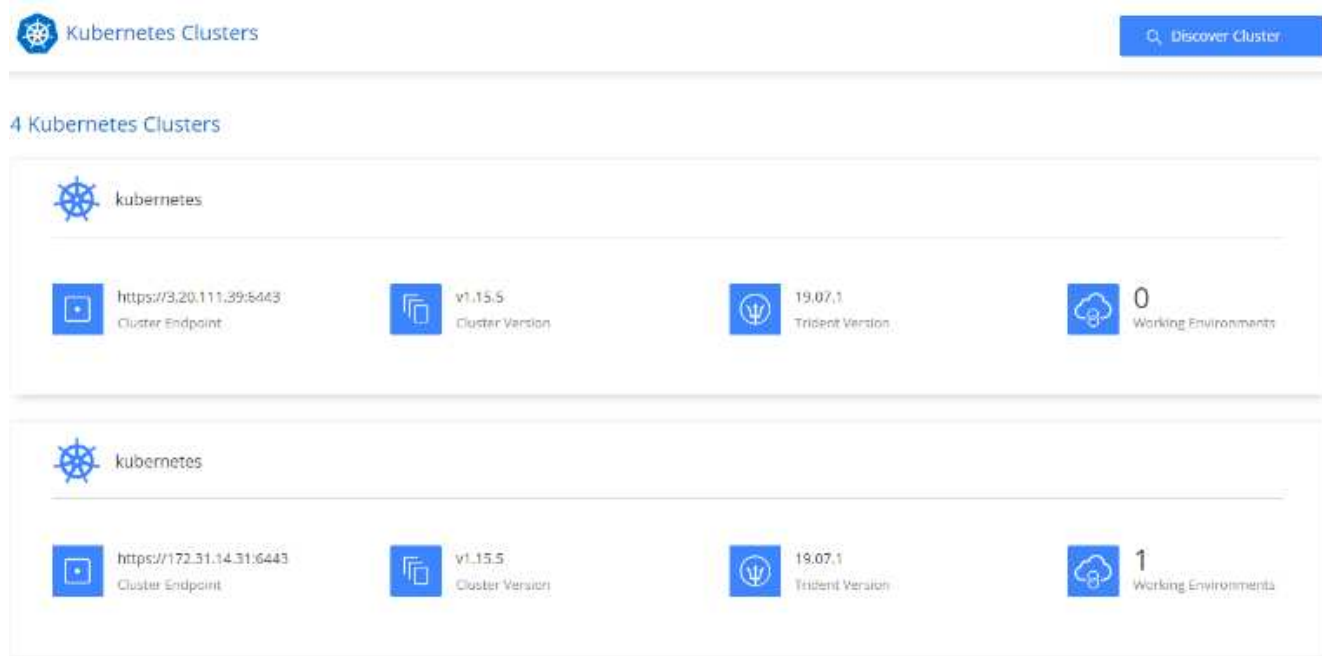
This section is divided into two parts for cloud and on-premises deployment respectively.

### Cloud Deployment Kubernetes Configuration

Through NetApp Cloud Manager, you can define the connection to the Iguazio Kubernetes cluster. Trident requires access to multiple resources in the cluster to make the volume available.

1. To enable access, obtain the Kubernetes config file from one the Iguazio nodes. The file is located under `/home/Iguazio/.kube/config`. Download this file to your desktop.

2. Go to Discover Cluster to configure.



3. Upload the Kubernetes config file. See the following image.

## Upload Kubernetes Configuration File

Upload the Kubernetes configuration file (kubeconfig) so Cloud Manager can install Trident on the Kubernetes cluster.

Connecting Cloud Volumes ONTAP with a Kubernetes cluster enables users to request and manage persistent volumes using native Kubernetes interfaces and constructs. Users can take advantage of ONTAP's advanced data management features without having to know anything about it. Storage provisioning is enabled by using NetApp Trident.

Learn more about [Trident for Kubernetes](#).

Upload File

4. Deploy Trident and associate a volume with the cluster. See the following image on defining and assigning a Persistent Volume to the Iguazio cluster. This process creates a Persistent Volume (PV) in Iguazio's Kubernetes cluster. Before you can use it, you must define a Persistent Volume Claim (PVC).

## Persistent Volumes for Kubernetes

### Connected with Kubernetes Cluster

Cloud Volumes ONTAP is connected to 1 Kubernetes cluster. [View Cluster](#) ⓘ

You can connect another Kubernetes cluster to this Cloud Volumes ONTAP system. If the Kubernetes cluster is in a different network than Cloud Volumes ONTAP, specify a custom export policy to provide access to clients.

#### Kubernetes Cluster

Select Kubernetes Cluster

kubernetes

#### Custom Export Policy *(Optional)* ⓘ

Custom Export Policy

172.31.0.0/16

☒ Set as default storage class

☒ NFS ☐ iSCSI

Connect

Cancel

## On-Premises Deployment Kubernetes Configuration

For on-premises installation of NetApp Trident, see [TR-4798](#) for details. After configuring your Kubernetes cluster and installing NetApp Trident, you can connect Trident to the Iguazio cluster to enable NetApp data management capabilities, such as taking Snapshot copies of your data and model.

[Next: Define Persistent Volume Claim](#)

## Define Persistent Volume Claim

1. Save the following YAML to a file to create a PVC of type Basic.

```
kind: PersistentVolumeClaim
apiVersion: v1
metadata:
  name: basic
spec:
  accessModes:
    - ReadWriteOnce
  resources:
    requests:
      storage: 100Gi
  storageClassName: netapp-file
```

2. Apply the YAML file to your Iguazio Kubernetes cluster.

```
Kubectl -n default-tenant apply -f <your yaml file>
```

## Attach NetApp Volume to the Jupyter Notebook

Iguazio offers several managed services to provide data scientists with a full end-to-end stack for development and deployment of AI/ML applications. You can read more about these components at the [Iguazio Overview of Application Services and Tools](#).

One of the managed services is Jupyter Notebook. Each developer gets its own deployment of a notebook container with the resources they need for development. To give them access to the NetApp Cloud Volume, you can assign the volume to their container and resource allocation, running user, and environment variable settings for Persistent Volume Claims is presented in the following image.

For an on-premises configuration, you can refer to [TR-4798](#) on the Trident setup to enable NetApp ONTAP data management capabilities, such as taking Snapshot copies of your data or model for versioning control. Add the following line in your Trident back- end config file to make Snapshot directories visible:

```
{
  ...
  "defaults": {
    "snapshotDir": "true"
  }
}
```

You must create a Trident back- end config file in JSON format, and then run the following [Trident command](#) to reference it:

```
tridentctl create backend -f <backend-file>
```

The screenshot shows the configuration interface for a Jupyter Notebook. At the top, there is a toggle switch labeled 'Enabled' which is checked. Below it is a slider for 'Inactivity window' with markers at 5m, 10m, 1h, 2h, and 4h; the 10m marker is selected. The 'Resources' section includes a link to 'Kubernetes documentation' and a note: 'The memory and CPU configurations are applied to each replica.' There are two rows of resource configuration: 'Memory' and 'CPU'. Each row has 'Request' and 'Limit' fields with dropdown menus for units (GB for memory, millicpu for CPU) and a help icon. The 'Running User' section has a text input field with 'admin' and a dropdown menu for 'Username' with a help icon.

The screenshot shows the configuration interface for a Jupyter Notebook, specifically the 'Environment Variables' and 'Persistent Volume Claims (PVCs)' sections. The 'Flavor' dropdown is set to 'Full stack without GPU'. The 'Spark' dropdown is set to 'spark' with a 'Create new...' link. The 'Environment Variables' section has a '+ Create a new environment variable' button. The 'Persistent Volume Claims (PVCs)' section has a table with two columns: 'Name' and 'Mount Path'. The first row has 'basic' in the 'Name' column and '/netapp' in the 'Mount Path' column. There is an '+ Add PVC' button at the bottom.

Next: [Deploying the Application](#)

## Copyright information

Copyright © 2023 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

## Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.