

# Overview and Requirements for Installing KumoScale for Managed Mode with Kubernetes

Introduction to the requirements and process for installing KumoScale for Managed Mode using your own Kubernetes cluster.

This guide will walk you through all the steps to successfully install and configure KumoScale™ Storage Node software running in Managed Mode with a Kubernetes™ Cluster. This is a High availability (HA) mode with provisioning, volume management, and storage node management supported using KumoScale operators and control plane in addition to the KumoScale Cluster Manager CLI and REST API.

**NOTE:** If you are using KumoScale software in Appliance Mode follow the instructions in [KumoScale Installation Guide for Appliance Mode](#).

The installation described in this manual consists of installing and configuring the following KumoScale software for Managed Mode appropriate for your deployment environment:

- **KumoScale Storage Node Modules** includes a target kernel component, and an embedded engine. The storage node software is installed directly on your OS and kernel.
- **KumoScale Provisioner** is a service for provisioning, analyzing, and managing storage nodes. KumoScale will use an existing pod on your Kubernetes cluster to determine the best location for allocating requested volumes. You will need to configure such a pod for load balancing and for referencing storage requirements and analytics data; this is explained later in this installation documentation.
- **KumoScale Cluster Manager Command Line Interface (CLI)** is used for managing and provisioning your storage deployment from a client machine.
- **KumoScale Operators** are used to manage KumoScale services. There are two control operators that run as services on the Kubernetes cluster: the install operator and the config operator. These enable you to use custom resource (CR) files with the operators to configure and manage storage.
- **Prometheus Monitoring and Alerting:** You may optionally use Prometheus™, an open-source system monitoring and alerting toolkit, with your KumoScale deployment. An overview on how to do this is provided at [Using Prometheus with KumoScale in Managed Mode](#).

Additional components that provide internal logging and analysis can be added to the storage cluster.

- See the KumoScale documentation for information on orchestration for Bare-Metal using KumoScale Ansible™ modules and playbooks, Kubernetes using the KumoScale CSI Driver, and Cinder using KumoScale for OpenStack™.
- If you are interested in deploying other third-party applications such as Fluentd™ or Loki™, contact technical support for assistance. Our team will advise you on how to deploy and customize the process.

## Intended Audience

This document is written for storage administrators. It assumes the reader has a working knowledge of storage, networking, and Kubernetes orchestration. If you need additional information on how to set up your own Kubernetes cluster, there are many sources available. We suggest starting with the tutorial of how to set up DigitalOcean™ Managed Kubernetes Cluster available at <https://github.com/digitalocean/Kubernetes-Starter-Kit-Developers/tree/main/01-setup-DOKS>.

## Before you Begin

Before you begin the installation, confirm the following requirements have been met:

1. Confirm your environment meets the requirements appropriate for your deployment:
  - a. **Hardware:** Your hardware environment meets the requirements documented in the [KumoScale Hardware Compatability List \(HCL\)](#). If you wish to add an RDMA NIC not in the HCL contact your KIOXIA representative.
  - b. **Operating Systems:** Your servers are installed with one of
    - Oracle Linux 8 with kernel 5.x
    - RHEL 8 with kernel 5.x.
    - CentOS™ 7 or 8, with kernel 5.x. KumoScale has been tested on CentOS 7.9 and Centos 8.3 with kernel 5.10.61.
    - Ubuntu™ 20.04 LTS with kernel 5.x. KumoScale has been tested on Ubuntu 20.04.3 LTS with kernel 5.4.0-91-generic and 20.04.5 with 5.4.0-137-generic.
  - c. **Storage Nodes:** Each server to be used as a storage node supports the following:
    - At least one NVMe SSD.

- At least one network interface that is up and persistently configured with a reachable IP address.
  - Access to web traffic and NVMe-oF traffic. That is, web ports (e.g., 443/tcp) and NVMe-oF ports (e.g. 4420/tcp) are not blocked by the firewall.
  - Packages needed by the KumoScale engine: **ethtool**, **net-tools**, **pciutils**.
- d. **Kubernetes cluster:** Confirm that the following is true of the Kubernetes cluster you will use with KumoScale:
- It is installed and configured for high-availability according to the best practices of the selected distribution.
  - A pod is configured to support load balancing. That is, it is configured with an externally accessible virtual IP. This IP will be used as the Provisioner service IP for high availability and you will need to provide this IP address as part of the installation. We recommend using **metallb**. See <https://metallb.org/> for more information.
- e. **Local Binary image repository for installation:** The KumoScale Operator and Provisioner are delivered as binary containers for Kubernetes and enable you to configure and manage provisioning using KumoScale operators with Custom Resource files. You will need a local registry to complete the installation of these components. For an example on how to set up such a repository using Docker, see <https://www.devonblog.com/containers/how-to-setup-your-private-docker-registry/>.

2. Read the [Release Notes](#) for additional information on installation and use of KumoScale software.
3. Review the [Authentication](#) page of the [KumoScale User Guide](#). This describes how to use the authentication modes supported by KumoScale. You will need to determine which mode you want to use before configuring the KumoScale Provisioner. You will also need to understand how to use self-generated tokens to configure the Provisioner. This is detailed on the Authentication page under *Self-Generated Tokens*.

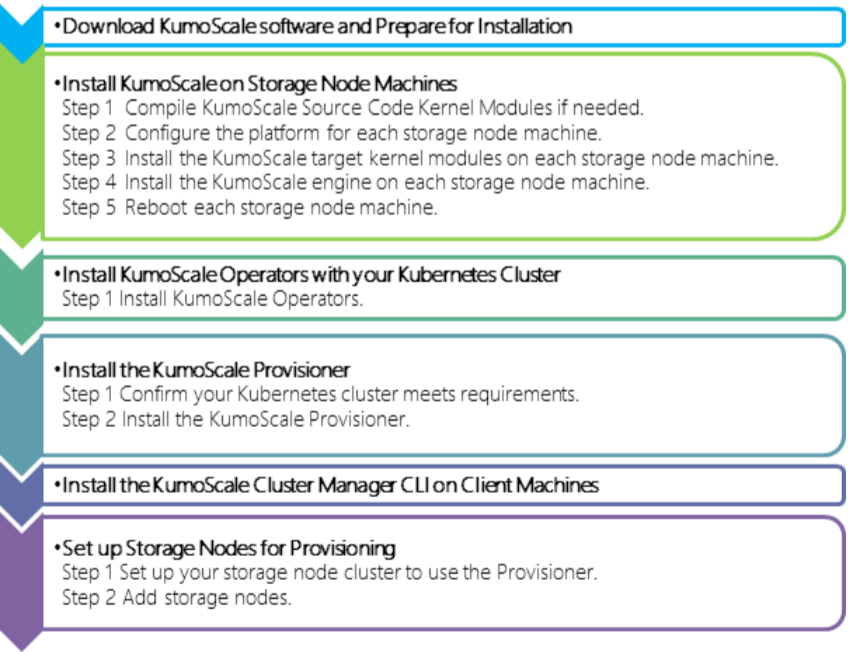
Once you have confirmed your environment has met all the requirements above, you may proceed with installation.

### Installation and Configuration Steps

You will need to install and configure four components in your deployment environment. Complete the steps below in the given order and as documented in the linked sections.

- [Download KumoScale Software and Prepare for Installation](#) provides instructions on how to download software from KIOXIA and prepare the software for installation of all needed KumoScale components.
- [Install KumoScale on Storage Node Machines](#) includes required platform settings and installation instructions for kernel modules and the KumoScale Engine.
- [Install KumoScale Operators](#) explains how to install the KumoScale operators and use them with Custom Resource (CR) files to manage your storage environment.
- [Install the KumoScale Provisioner](#) provides instructions on how to install the KumoScale Provisioner.
- [Install KumoScale Cluster Manager CLI](#) explains how to install the Cluster Manager CLI on a client machine.
- [Set up Storage Nodes for Provisioning](#) explains how to set up your storage nodes for provisioning and management with KumoScale.

The diagram below illustrates the installation steps. We will refer to this diagram in subsequent chapters.



Additional documentation, including the KumoScale Cluster Manager CLI Reference Guide, HCL, User’s Manual, CSI, OpenStack, and Ansible guides, are available on the KumoScale software documentation.

**Note:** When you are ready to set up compute nodes for deployment you must install the NVMe host module patch on each compute node. This step is independent of KumoScale software but is needed to complete a deployment. Instructions are available at [Installing the NVMe™ Host Module Patch on Compute Nodes](#).

*Next Installation Step:* [Download KumoScale Software and Prepare for Installation](#)

