

Configuring KNIME Server for KNIME Edge

KNIME AG, Zurich, Switzerland
Version 1.0 (last updated on 2024-01-29)



Table of Contents

- Introduction..... 1
- Installation Planning 1
 - KNIME Server Large 1
 - PostgreSQL Database..... 1
- Configuring KNIME Server..... 2
 - KNIME Edge initialization workflow 2
 - Uploading the initialization workflow to KNIME Server..... 2
 - Executing the initialization workflow..... 2
 - Confirming the KNIME Edge Control Plane is configured properly 5
 - OAuth Considerations..... 5

Introduction

This guide outlines how to deploy and configure the KNIME Edge Control Plane workflows.

Installation Planning

KNIME Server Large

A KNIME Server Large installation acts as the central control plane for one or more Edge clusters.



Supported version(s): KNIME Server Large 4.11.0+

Currently, KNIME Edge leverages a collection of workflows running on a KNIME Server to act as both the API endpoints an Edge Cluster communicates with, as well as for providing user interfaces via KNIME WebPortal for interacting with Edge clusters.

The KNIME Edge Control Plane workflows can be deployed in a single operation via an available [KNIME Edge Initialization Workflow](#) hosted on the KNIME Community Hub.

PostgreSQL Database

To maintain the state of all KNIME Edge clusters via KNIME Server, a PostgreSQL database is needed so all workflows (regardless of whether they handle API requests or KNIME WebPortal user interfaces) can leverage it to maintain state regardless of the number of Executors set up in the Server installation.



Supported version(s): The minimum version for PostgreSQL tested is 11.8 or higher.

SQL files containing the needed schemas for the PostgreSQL instance are bundled together with the workflows and can be found in the Edge space under the `/<Edge-Root-Path>/config/SQLFiles` directory after configuring the KNIME Server.

The initialization workflow sets up the needed schemas by executing the SQL files and configures the PostgreSQL Database Connection nodes in all KNIME Edge Control Plane workflows.

Configuring KNIME Server

KNIME Edge initialization workflow

The recommended way to configure a KNIME Server to act as the control-plane for KNIME Edge clusters is to use the **KNIME Edge Initialization** workflow.

From KNIME Hub, either download and import the Workflow into a local installation of KNIME Analytics Platform or drag-and-drop the workflow from KNIME Hub to Analytics Platform.

Uploading the initialization workflow to KNIME Server

Using the KNIME Explorer view in Analytics Platform, connect to a KNIME Server and deploy or copy/paste the KNIME Edge Initialization workflow to any location where a user will have execute access to run the workflow.

Executing the initialization workflow

Log into the KNIME Webportal via a browser, then navigate to where the Initialization workflow was saved.

Initialize Edge Server

This workflow is used to setup a KNIME Server to be used as the orchestration piece for a KNIME Edge cluster. Further information can be found at <https://docs.knime.com>

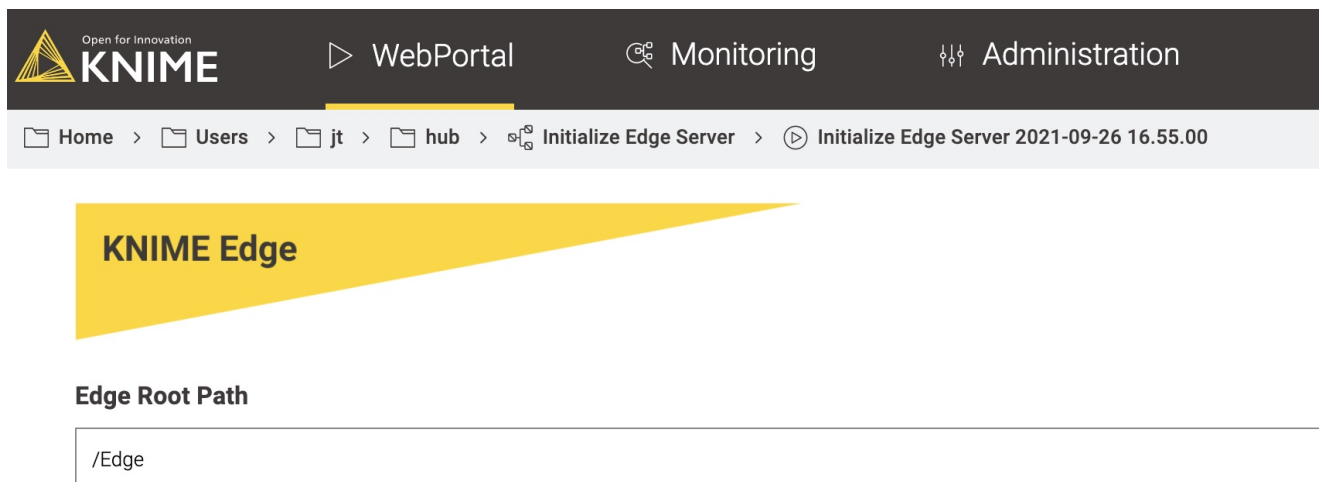
☐ Notify me when workflow has run

 **Run**

On the first page, there is a prompt for the "Edge Root Path". This the root workflow group that will be created and all needed workflows will be deployed to.

Note that you can specify a location that does not exist (in the case of installing a new set of workflows), or you can enter the root location to an existing Edge deployment and the currently specified parameters will be fetched and used in the next configuration step.

Running the Initialization workflow on an existing Edge root directory can be used to update the version of workflows deployed there, as well as apply new schema updates to the PostgreSQL database.



The screenshot shows the KNIME WebPortal interface. The top navigation bar includes the KNIME logo, 'WebPortal', 'Monitoring', and 'Administration'. Below the navigation bar is a breadcrumb trail: 'Home > Users > jt > hub > Initialize Edge Server > Initialize Edge Server 2021-09-26 16.55.00'. The main content area features a yellow banner with the text 'KNIME Edge'. Below this banner, the section 'Edge Root Path' is displayed with a text input field containing the value '/Edge'.

After the Edge root path is entered, click "Next" and the subsequent view will prompt the user for additional configuration options.

The **Edge Version** option is used to select a specific bundled version of KNIME Edge workflows to the KNIME Server. See the Version and Update documentation to determine what versions of KNIME Edge clusters are compatible with specific KNIME Server workflow versions.

The list of available Edge Versions is dynamically populated when the Initialization workflow is run, so it will always show an up-to-date list of available workflow bundles.

The remaining fields are used to enter the connection parameters for the required PostgreSQL database.

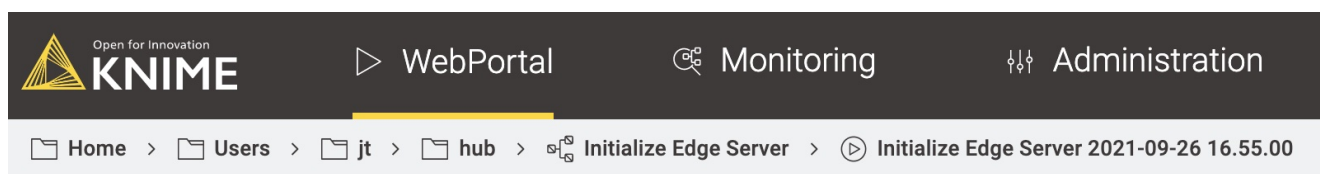
Once the Edge version and database connection parameters are entered, click "Next" and the final Initialization view will be shown, with options to the user on how to configure the database.

Checking the **Update Database Schema to latest version** option will sequentially apply all schema updates to the database to make sure it is in the required state needed for the pre-selected Edge version used. When upgrading to a newer version of Edge workflows, keep this box checked to ensure all updates are applied. Data will not be lost when Updating Database Schemas.

Checking the **Reset Database Schema Before Updating** option will completely destroy all data and recreate the schema, then recreate the needed schema as long as the update database schema box is also selected.



Only use the **Reset Database Schema Before Updating** option if the intention is to completely reset a KNIME Edge Configuration on KNIME Server. Data will be lost during this process!



KNIME Edge

☒ **Update Database Schema to latest version**

☐ **Reset Database Schema Before Updating**

Confirming the KNIME Edge Control Plane is configured properly

After the Initialization workflow has executed successfully you will now see a collection of workflows created inside the "Edge Root Path" previously specified.

These workflows are divided into three types:

- API workflows mainly used for communication directly with the Edge Cluster.
- Configuration workflows used for installation and maintenance of the Edge Server and database.
- Application workflows that can be run by users to view and modify the state of Edge clusters and inference deployments.

Inside the "apps" workflow group you can execute the View Edge Clusters or View Deployments workflows to see an overview of the state of connected Edge clusters. As long as these workflows execute successfully then you should be confident that the KNIME Edge Control Plane is correctly configured.

OAuth Considerations

Users with KNIME Servers configured to authenticate via OAuth need to be aware of the following considerations.

Communication between the KNIME Edge Control Plane and a KNIME Edge cluster occurs via calls to the KNIME Server REST API. In addition, the Initialization workflow and several of the "Application" workflows make requests to the KNIME Server REST API to set up scheduled jobs and query certain information.

Currently these requests are unable to be made by a user authenticated via OAuth. To ensure these workflows run successfully the Initialization workflow as well as the KNIME Server user credentials configured within the Edge cluster needs to be run by a "service account" user that authenticates via a non OAuth manner.

Information on setting up one of these user accounts can be found in the [KNIME Server documentation](#).

KNIME AG
Talacker 50
8001 Zurich, Switzerland
www.knime.com
info@knime.com