Open for Innovation
KNIME

KNIME Server on AWS Marketplace User Guide

KNIME AG, Zurich, Switzerland
Version 4.16 (last updated on 2022-06-02)

Table of Contents

Introduction . . . 1
Further reading . . . 1
Architecture Overview . . . 3
KNIME Server Large (AWS) . . . 3
Routine Maintenance . . . 4
Starting KNIME Server . . . 4
Stopping KNIME Server . . . 4
Restarting KNIME Server . . . 4
Restarting the executor (KNIME Server Large - Distributed Executors) . . . 4
Managing Certificates . . . 5
Default Certificates . . . 5
Update Python configuration . . . 7
Apply Operating System patches . . . 7
Update KNIME Server . . . 7
Install additional extensions . . . 8
SSH access to KNIME Server on AWS . . . 9
Increasing Workflow Repository EBS volume capacity . . . 10
Key rotation . . . 10
Service limits . . . 11
KNIME Executors with Auto Scaling Groups . . . 12
Setting up the VPC . . . 12
CloudFormation Templates . . . 13
Terminate KNIME Executors . . . 14
KNIME Server usage . . . 15
KNIME Server from KNIME Analytics Platform . . . 15
KNIME Server from KNIME WebPortal . . . 16
Authentication . . . 18

Introduction

KNIME Server is the enterprise software for team based collaboration, automation, management, and deployment of data science workflows, data, and guided analytics. Non experts are given access to data science via KNIME WebPortal or can use REST APIs to integrate workflows as analytic services to applications and IoT systems. A full overview is available here.

KNIME Server can be launched through the AWS Marketplace. There are several options:

- KNIME Server Large (PAYG)
- KNIME Server Medium (PAYG)
- KNIME Server Small (PAYG)
- KNIME Server (BYOL)

With the introduction of the KNIME Server Large (PAYG) offering we fill a gap in our AWS Marketplace offerings. This type of installation gives you the opportunity to spin up a cloud instance, and connect it to scalable KNIME Executors, without major IT involvement, contracting and legal work as all payment is handled via Amazon.

The existing PAYG offerings (KNIME Server Small and Medium) don't support usage of additional Executors, and KNIME Server Large BYOL requires purchasing a year-long license from our side.

KNIME Server Large (PAYG) support

In case you need help or support with the installation and setting of the KNIME Server Large (PAYG) instance please **reach out** to support@knime.com.

Further reading

If you are looking for detailed explanations around the additional configuration options for KNIME Server, you can check the KNIME Server Administration Guide.

If you are looking to install KNIME Server, you should first consult the KNIME Server Installation Guide.

For guides on connecting to KNIME Server from KNIME Analytics Platform, or using KNIME WebPortal please refer to the following guides:

- KNIME Server User Guide

© 2023 KNIME AG. All rights reserved.

- KNIME WebPortal User Guide

An additional resource is also the KNIME Server Advanced Setup Guide.

© 2023 KNIME AG. All rights reserved.

Architecture Overview

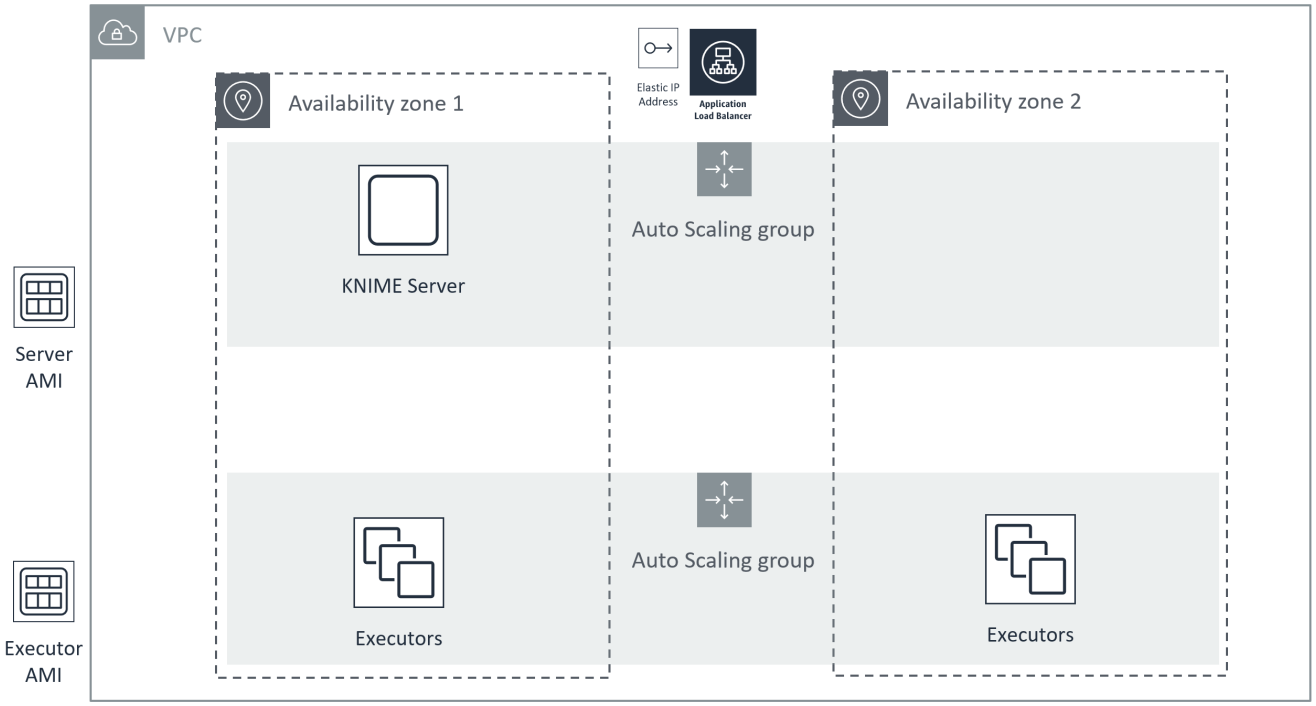
An overview of the general KNIME Server architecture is supplied. More detailed description of software architecture can be found in the KNIME Server Administration Guide.

KNIME Server Large (AWS)

KNIME Server Large supports configurations that allow for resilient architectures, to support scalability and high-availability.

For reference we've shown here one of the more complex architectures to support scaling KNIME Executors, and failover of the server instance to a new availability zone.

KNIME Server Large can be run as a single EC2 instance, in which case all of the guidance from above may be used.



© 2023 KNIME AG. All rights reserved.

Routine Maintenance

Starting KNIME Server

KNIME Server starts automatically when the instance starts using standard systemd commands. Once the Tomcat application has started successfully, it will automatically launch an executor. This means that in normal operation you will not need the below command.

In the case that you need to start a stopped KNIME Server, it may be started using the following command at the terminal:

```
sudo systemctl start knime-server.service
```

Stopping KNIME Server

Stop the KNIME Server by executing the command:

```
sudo systemctl stop knime-server.service
```

Restarting KNIME Server

Restart the KNIME Server by executing the command:

```
sudo systemctl restart knime-server.service
```

Restarting the executor (KNIME Server Large - Distributed Executors)

In most cases you will want to restart the entire instance that is running the Executor. But in certain cases you may wish to do this by restarting the Executor application itself. That can be acheived either by stopping the executor process, and starting again from the terminal. Or by restarting the systemd service, if it is being used.

```
sudo systemctl restart knime-executor.service
```

© 2023 KNIME AG. All rights reserved.

Managing Certificates

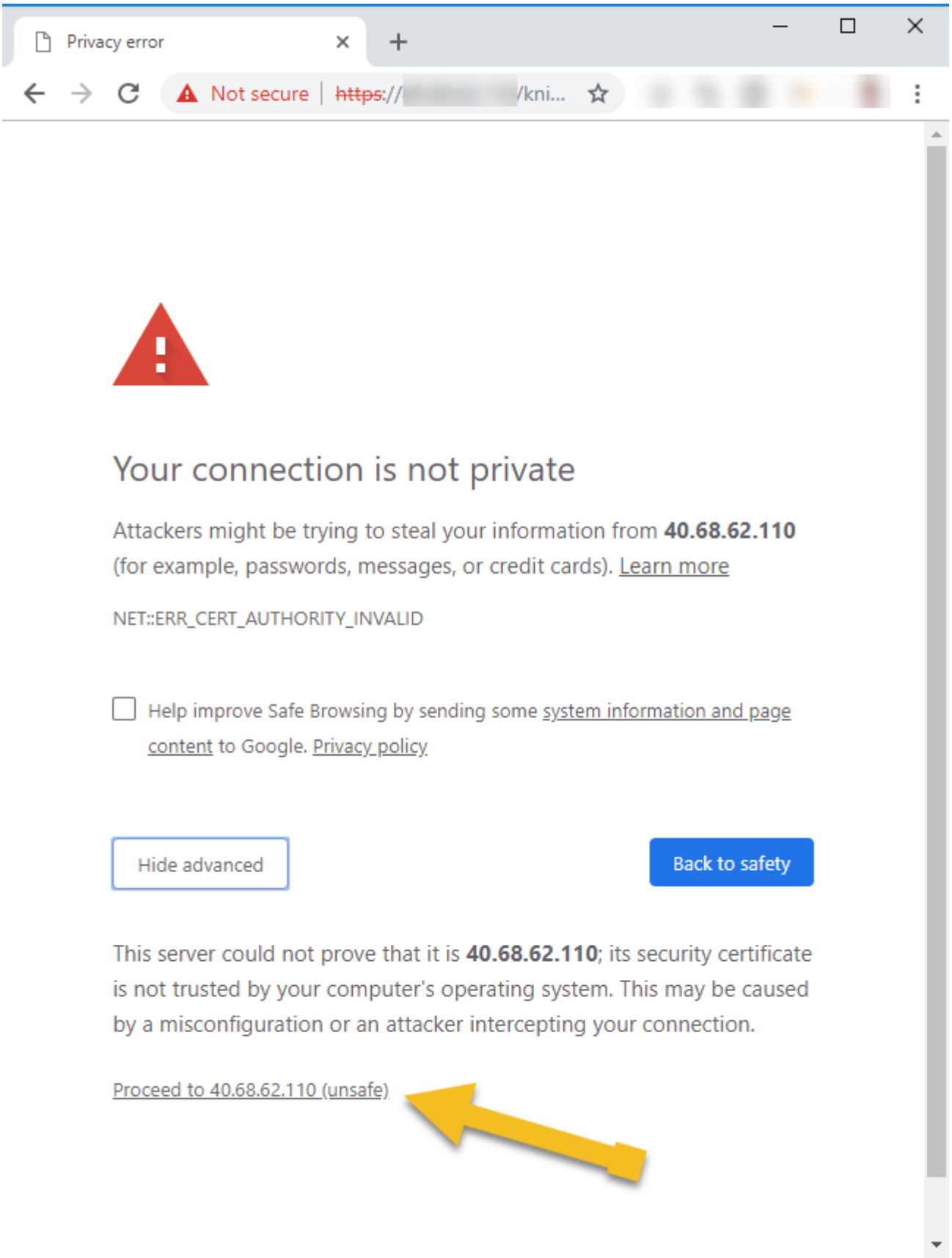
Detailed steps for managing the SSL certificate for KNIME Server can be found in the KNIME Server Administration Guide

Default Certificates

KNIME Server ships with a default SSL certificate. This allows for encrypted communication between client and server. However, since the certificate cannot be generated in advance for the server that you are running on, it will not be recognized as a valid certificate. Therefore, we recommend managing your own certificate as per the guidelines in the Managing Certificates.

When testing with the default certificate, modern browsers will issue a warning as below. Choosing to ignore the warning, will allow you to access the KNIME WebPortal for testing.

© 2023 KNIME AG. All rights reserved.



© 2023 KNIME AG. All rights reserved.

Update Python configuration

We use Anaconda Python to define a default python environment. The current yaml file can be found in /home/knime/python/py36_knime.yaml.

For detailed documentation on managing Anaconda, please refer to the Anaconda documentation.

An example yaml file is shown below. We chose packages that we know are well used, or are required for the use of the KNIME Deep Learning package. Mostly we have pinned version numbers to ensure compatibility. You may choose to unpin version numbers. Additionally you may wish to add a new Python package, in which case you can add the package to the yaml file and run the command:

```
sudo -u knime /home/knime/python/anaconda/bin/conda env update -f
/home/knime/python/py36_knime.yml --prune
```

py36_knime.yml

Apply Operating System patches

The KNIME Server 4.16 AMIs are based on Ubuntu Server 20.04 LTS. The OS should be regularly patched using the standard Ubuntu procedures.

After a Java JDK update, the KNIME Server must be restarted.

Update KNIME Server

Updates, and patches to the KNIME Server are announced on the KNIME Server Forum. You may subscribe to the topic.

Before applying a feature update (e.g. version 4.13.2 → 4.14.1) you should read the KNIME Server Release Notes and Update Guide. This will document any changes to parameters, features and settings that might affect your installation.

In the case of a patch update (e.g. version 4.14.1 → 4.14.2) there should be no changes to settings required.

There are two strategies to applying feature, or patch updates of KNIME Server. The first is to follow the instructions in the KNIME Server Update Guide via the terminal (in place update). The second is to migrate a snapshot of the workflow repository block device to a new KNIME Server instance (disk swap update).

© 2023 KNIME AG. All rights reserved.

In place update

To make a feature update you have the option to follow the instructions in the KNIME Server Update Guide.

Install additional extensions

We install a set of extensions that we think will be useful for many situations, see the [list-of-extensions]. If you developed a workflow locally that uses an extension not in the list, you'll need to install the extension into the executor.

KNIME Server Small/Medium/BYOL

To install a new extension.

- Stop the KNIME Server `sudo systemctl stop knime-server`
- Find the extension on KNIME Hub by searching for the node of interest, and clicking the link to the extensions page.
- Determine the `<extension-identifier>` by modifying the URL, e.g. https://hub.knime.com/knime/extensions/**com.knime.features.gateway.remote**/latest → **com.knime.features.gateway.remote**.feature.group
- Run the command:

```
sudo -u knime /opt/knime/knime-executor/knime -application
org.eclipse.equinox.p2.director \
-nosplash -consolelog -r https://update.knime.org/analytics-
platform/{version_exe},https://update.knime.com/community-
contributions/trusted/{version_exe} \
-i <extension-identifier> -d /opt/knime/knime-executor
```

If you want to install an extension from another update site, you'll need to add that to the command. You'll also need to find the `<extension-identifier>` or 'Feature Group Name' from the list of installed extensions in the KNIME Analytics Platform.

Disk swap update (AWS)

- Login to the existing server, and stop the KNIME Server service. See section Stopping KNIME Server.

© 2023 KNIME AG. All rights reserved.

- Create a snapshot of the workflow repository block device.
  - Find the block device that needs snapshotting (in the AWS EC2 Console) image::10_find_volume.png[Find the block device that needs snapshotting]
  - Create a snapshot of the block device image::10_create_snapshot1.png[Create a snapshot of the block device]
  - Give the snapshot a name image::10_create_snapshot2.png[Give the snapshot a name]
- Launch a new instance of KNIME Server from the AWS Marketplace, using the previously created snapshot.
  - Launch new instance on AWS Console image::10_launch_new_instance1.png[Launch new instance on AWS Console]
  - Assign snapshot to new instance image::10_assign_snapshot.png[Assign snapshot to new instance]

SSH access to KNIME Server on AWS

Access to the KNIME Server instance via SSH follows the general instructions provided by AWS.

Connection to the KNIME Server instance is always using the `ubuntu` user, and the SSH key specified at instance launch time. An example SSH connection string is:

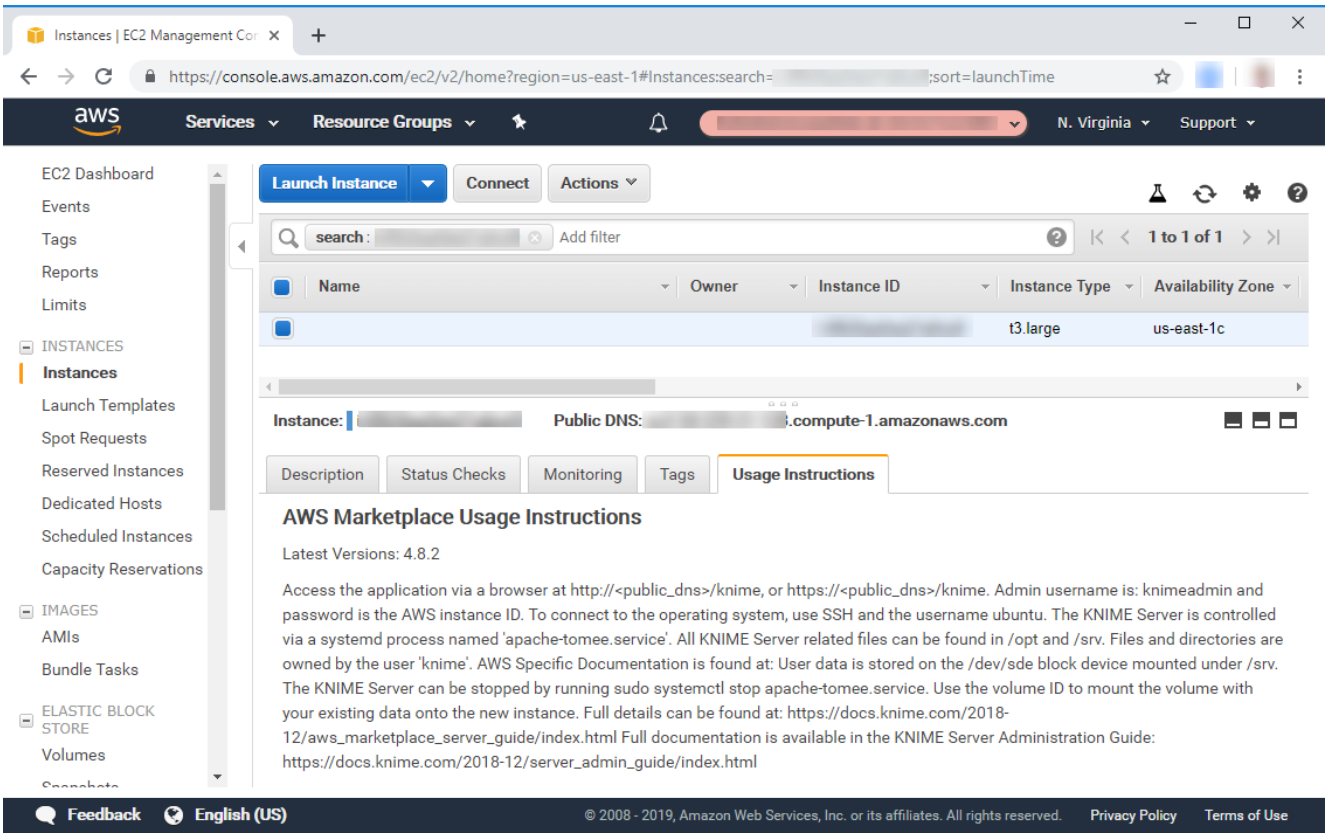
```
ssh -i <keyfile.pem> ubuntu@<hostname>
```

All relevant KNIME Server installation and runtime files are owned by the `knime` user. In order to make changes to these files, it is required to assume the identity of the `knime` user:

```
sudo su knime
```

Usage instructions are always included as part of the AWS console `Usage Instructions` tab.

© 2023 KNIME AG. All rights reserved.



Scaling the KNIME Server instance

In the case that your KNIME Server instance is running at high CPU load permanently, you may wish to scale to a larger instance size. Conversely it may be the case that you started with an instance size that is too large. In both cases, it is possible to stop the KNIME Server and start it again using a different instance type.

To do so, you should follow the instructions here.

Increasing Workflow Repository EBS volume capacity

It is possible to increase the capacity of the workflow repository EBS volume (default size: 250 Gb) after an instance has been launched. Follow the instructions here.

Changing the size of the root volume should not be required. The procedure is more difficult than increasing the size of the workflow repository EBS volume.

Key rotation

Managing SSH keys for accessing the KNIME Server is detailed here.

© 2023 KNIME AG. All rights reserved.

Service limits

Managing AWS service limits is detailed here.

© 2023 KNIME AG. All rights reserved.

KNIME Executors with Auto Scaling Groups

Since the KNIME Executor that is included in the KNIME Server Large (PAYG) AWS installation has a fixed number of 4 cores, in order to add more computational capacity you will need to connect with additional PAYG KNIME Executors.

Three components are needed for Auto Scaling KNIME Executors on AWS:

- A VPC that hosts KNIME Server, RabbitMQ and the KNIME Executors
- An AWS Auto Scaling Group that manages the startup and shutdown of KNIME Executors

KNIME Server should be deployed in a public subnet, while KNIME Executors should be deployed in a private subnet.

Setting up the VPC

A Virtual Private Cloud (VPC) within the AWS cloud platform provides a secure networking environment for KNIME Software. Each VPC provides its own networking domain which provides access to and protections from the outside networking world. The divisions are well defined, and layers of network security are provided.

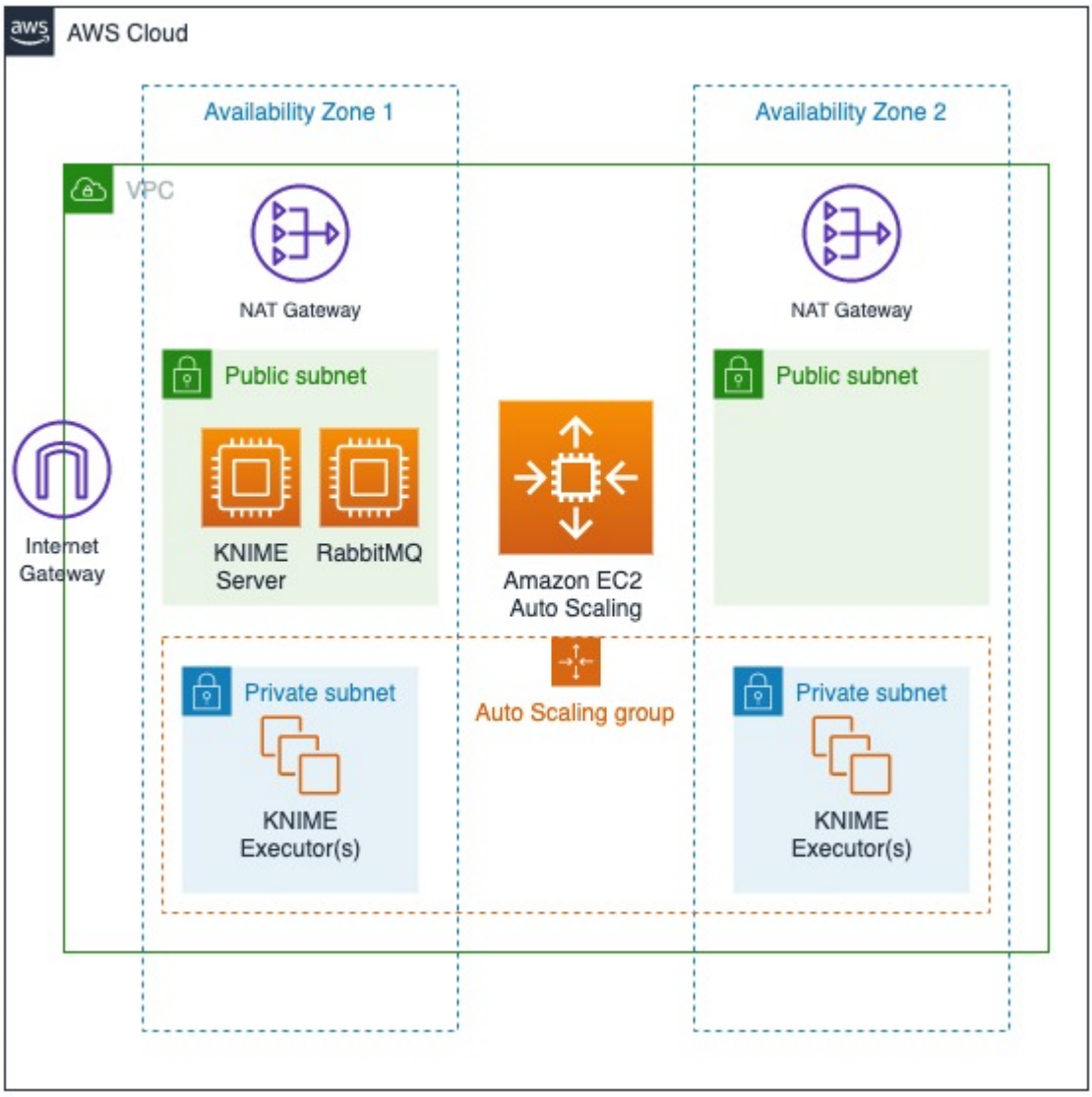
When deploying the KNIME Server within a VPC it is usually deployed within a public subnet. This will allow connections to be made to the Server from the external internet. Security Groups can be used to control access to the KNIME Server (I.e. open ports to specific IP addresses or a range of IP addresses). There are scenarios where the KNIME Server can be deployed within a private subnet. The use cases for this deployment include: all access of the Server will be made within the VPC or the VPC is connected through a VPN to a company's private network.

KNIME Executors are generally deployed within private subnets. The reasoning here is that the outside world does not need to connect to Executors. Executors only communicate with the KNIME Server and RabbitMQ. When run within an Auto-scaling Group (ASG), the Executors should be run within subnets that are contained in different Availability Zones. This provides a level of resilience to the Executors. If networking within an AZ becomes unavailable, the ASG will automatically initiate new Executors in the still functioning AZ. While overall compute capacity may be lowered in the period while the new Executors are starting up, the service of Executors is not disrupted. And within a short period of time, the desired level of compute capacity will be regained.

The diagram below depicts the typical deployment of KNIME Server with KNIME Executors

© 2023 KNIME AG. All rights reserved.

within a VPC in AWS.



CloudFormation Templates

KNIME provides a CloudFormation template along with the AWS Marketplace AMIs to simplify the creation of the AutoScalingGroup.

executors.yaml

This stack creates the AutoScalingGroup that governs the KNIME Executors, as well as the security group for those instances; the AMIs for KNIME Executor instances are available on the AWS marketplace.

Open the yaml file in a text editor and edit the RegionMap which specifies the KNIME Executor AMI. Make sure to also set the correct AWS region.

© 2023 KNIME AG. All rights reserved.

Fill out the template as prompted in the CloudFormation console and create the stack. Select the VPC which the server instance is in. In the Executor Configuration section you need to set the private IP that was assigned to the server instance; do not enter the public IP since the KNIME Executor instances are not allowed to communicate outside of the VPC. Set the target average CPU utilization for the scaling group.

Using a value here that is too low might cause ASG to scale up twice even though only one scale up is desired. A value of 90% should be an appropriate starting point. It is possible to tweak those settings at a later point in the AutoScalingGroups console.

In addition, you need to specify the minimum and maximum limits of how many executors should be started. In case you do not wish to scale dynamically, but rather use BYOL Executors, minimum and maximum should be set to the same value. For those scenarios, the ASG allows you to keep the correct number of executors running in case of failure.

After the stack was created the specified minimum number of KNIME Executor instances will be started, and they should automatically connect to RabbitMQ on the server.

After this step the Auto Scaling KNIME Executors are ready to use.

Terminate KNIME Executors

If you want to stop using this setup then simply delete the stack in the CloudFormation section of AWS.

This also means that all data on the server instance is lost, unless it has been backed up prior to the deletion.

© 2023 KNIME AG. All rights reserved.

KNIME Server usage

After installing KNIME Server you can start using the available features like team based collaboration, automation, management, and deployment of data science workflows build with KNIME Analytics Platform. Non experts are given access to data science via KNIME WebPortal or can use KNIME Server REST APIs to integrate workflows as analytic services.

KNIME Server from KNIME Analytics Platform

To connect to KNIME Server from KNIME Analytics Platform you have to first add the relative mount point to the KNIME Explorer.

To do so open a local installation of KNIME Analytics Platform and find the KNIME Explorer on the left-hand side. This is the point of interaction with all the mount points available. Click the preferences button in the toolbar.

In the preferences window click *New…* to add the new Server mount point. In the *Select New Content* window select KNIME ServerSpace and insert the Server address:

```
https://<hostname>/knime/
```

Choose the desired authentication type. Finally, click the *Mount ID* field where the name of the mount point will be automatically filled, and select *OK*.

Open the Server mount point in the explorer and double-click *Double-click to connect to server*, insert the credentials and click *OK*.

Now you can deploy your local workflows to KNNIME Server by drag and drop or right-click the local items and select *Upload to Server or Hub* from the context menu. A window will open where you can select the desired Server mount point folder where to place the item, given that you are logged in to the respective Server.

Once you have deployed your workflows to KNIME Server you can:

- Execute them or schedule their execution on KNIME Server
- Deploy them as Data Apps via KNIME WebPortal

For a comprehensive explanation of how to use KNIME Server please refer to the KNIME

© 2023 KNIME AG. All rights reserved.

Server User Guide.

KNIME Server from KNIME WebPortal

KNIME WebPortal is an extension to KNIME Server. It provides a web interface that lists all accessible workflows, enables their execution and investigation of results. The workflows' input and output can be parameterized and their visualizations customized, using Configuration and Widget nodes, Interactive Widget and View nodes. Using components containing these nodes it is then possible to create Data Apps.

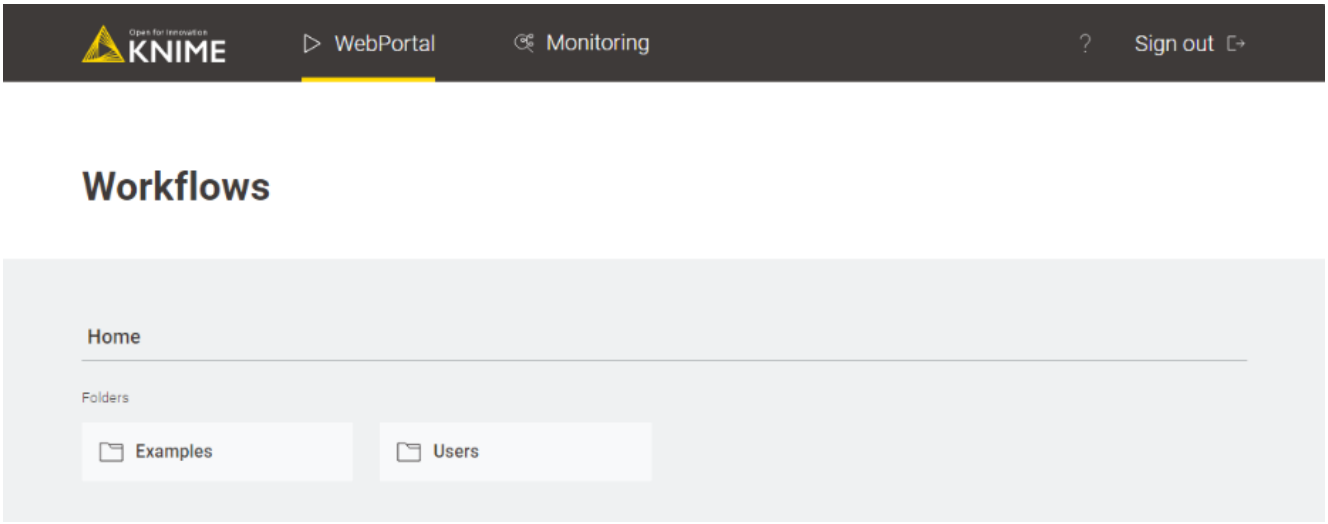
Accessing KNIME WebPortal

You can use a standard web browser to connect to KNIME WebPortal. The address (URL) to get to the login page is:

```
https://<server-address>/knime/webportal/
```

After providing the user's login name and password, a listing of all available workflows is displayed. Only workflows the user can read and execute are shown.

After signing in, KNIME WebPortal main page is displayed, showing the available spaces or workflows.



From this page users can:

- Access all your workflows from the WebPortal page
- Access to your Monitoring portal
- Moreover, users with administrator rights will have access to the Administration portal.

© 2023 KNIME AG. All rights reserved.

The monitoring portal provides details about jobs, schedules and executors and allows to download KNIME Server logs.

The administration portal, instead:

- Provides details about the Server's status and allows to upload the Server license
- Allows to manage local users and groups
- Allows to change the Server's configurations

For a more comprehensive guide on how to use the monitoring and administration portal as a KNIME Server administrator please refer to the Monitoring and administration portal section on the KNIME Server Administration Guide.

© 2023 KNIME AG. All rights reserved.

Authentication

KNIME Server Large also provides the ability to connect to existing LDAP or OIDC for authentication. User authentication in an enterprise environment is usually done through some centralized service. The most used service is LDAP. LDAP authentication is the recommended authentication in any case where an LDAP server is available.

Alternatively, KNIME Server can be configured to use an OpenID Connect enabled Identity Provider for authentication.

KNIME Server Large (PAYG) support

In case you need help or support with setting up authentication on the KNIME Server Large (PAYG) instance please **reach out** to support@knime.com.

For a comprehensive explanation about this topic please refer to the KNIME Server Advanced Setup Guide.

© 2023 KNIME AG. All rights reserved.

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

The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME AG under license from KNIME GmbH, and are registered in the United States. KNIME® is also registered in Germany.