Open for Innovation
KNIME

KNIME Business Hub Admin Workflows Guide

KNIME AG, Zurich, Switzerland
Version 1.16 (last updated on)

Table of Contents

In this guide you will find tools that help you monitor, diagnose, and maintain your Business Hub. It is intended for anyone administrating the whole Hub or a Team on the Hub.

Use cases

Monitoring

- **User & Team activity**: Analyze how users interact with the system, including active logins, team distribution, and access methods with the Admin Dashboard for a general overview and the Monitor Users Usage data app for a detailed session investigation.
- **Resource utilization**: Assess execution contexts, workflow scheduling, and resource scaling to optimize performance with the Admin Dashboard for a general overview and the Workflow Jobs Monitoring data app for details on workflow execution.
- **Content & Storage Management**: Identify large items, popular downloads, and opportunities for space optimization with the Admin Dashboard.
- **Performance & Reliability**: Monitor workflow execution trends, failures, and overall execution efficiency with the Workflow Jobs Monitoring data app.
- **Gen AI & Assistance usage**: Measure engagement with K-AI, script assistance, and user support needs with the Gen AI Monitoring data app.

Diagnostics

- **List execution context capabilities**: List installed extensions, nodes, conda environments, and Linux packages with the Executor Diagnostics data app.
- **Check URL accessibility:** Identify if a given URL is reachable from the execution context on the "Proxy"-perspective of the Executor Diagnostics data app.
- **Investiate network issues:** Check if the execution context is blocked for example by a faulty proxy configuration by entering a URL in the "Proxy"-perspective of the Executor Diagnostics data app.
- **Inspect active worklfow executions**: List all jobs that are still active or inspectable with the Manage Existing Jobs data app.

Maintenance

- **Free up disk space**: Delete unused versions of workflows with the Manage Item Versions data app or delete old workflow executions with the Manage Existing Jobs data app.

Admin Dashboard

The Admin Dashboard data app covers the following use cases:

- **Investigate user, item, and deployment distribution**: Analyze the distribution of users, items (workflows, components, files), and deployments in your Business Hub and across teams.
- **Trends and growth**: Monitor the growth of new users, new items, and new deployments over time.
- **Hub snapshot**: As a global admin, get a snapshot of the Hub, including anonymized information about the number of users, items, and deployments to compare growth over a longer time period or compare different Hub instances.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more)
    - Optional: Instead, deploy the workflow as a data app to persist any configuration and make it available to others by sharing the deployment (learn how).
5. Optional: impersonate another user or configure another Hub in the configuration panel shown in Figure 1.
    - User application password: specify the user you want to impersonate.
        - "Current" (default) for the currently logged in user. The password is not needed.
        - Application password of the user you want to impersonate. This allows you to use permissions of another user.
    - Hub URL: specify the Hub you want the data app to refer to.
        - "Current" (default) for the Hub on which the execution takes place.
        - The URL of the other Hub you want to refer to.

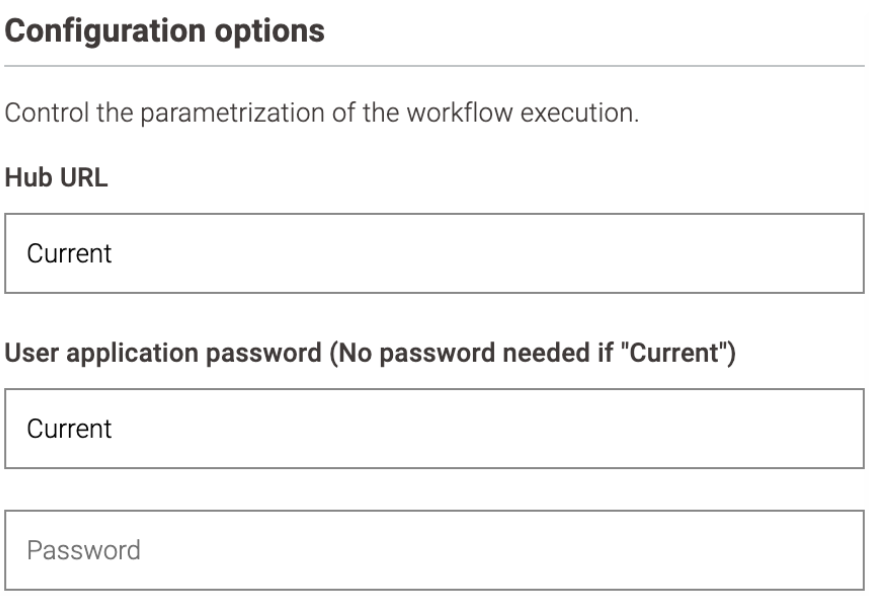

Figure 1. Hub URL and user application password configuration.

6. *Run* the workflow

Workflow Jobs Monitoring

The Workflow Jobs Monitoring data app covers the following use cases:

- **Schedule optimization**: Find times where execution contexts are less busy for performant scheduling.
- **Performance & reliability**: Track workflow execution trends, identify failures, and assess overall efficiency to ensure optimal performance.
- **Investigate individual jobs**: Find, understand, and inspect jobs that have failed or are taking longer than expected.

Enable Job Instrumentation data collection

The data app requires **job intrumentation data** to be collected and stored to. For that, you need to enable the job instrumentation data collection as described here.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more)
    - Optional: Instead, deploy the workflow as a data app to persist any configuration and make it available to others by sharing the deployment (learn how).
5. Optional: impersonate another user or configure another Hub in the configuration panel shown in Figure 2.
    - User application password: specify the user you want to impersonate.
        - "Current" (default) for the currently logged in user. The password is not needed.
        - Application password of the user you want to impersonate. This allows you to use permissions of another user.
    - Hub URL: specify the Hub you want the data app to refer to.
        - "Current" (default) for the Hub on which the execution takes place.
        - The URL of the other Hub you want to refer to.

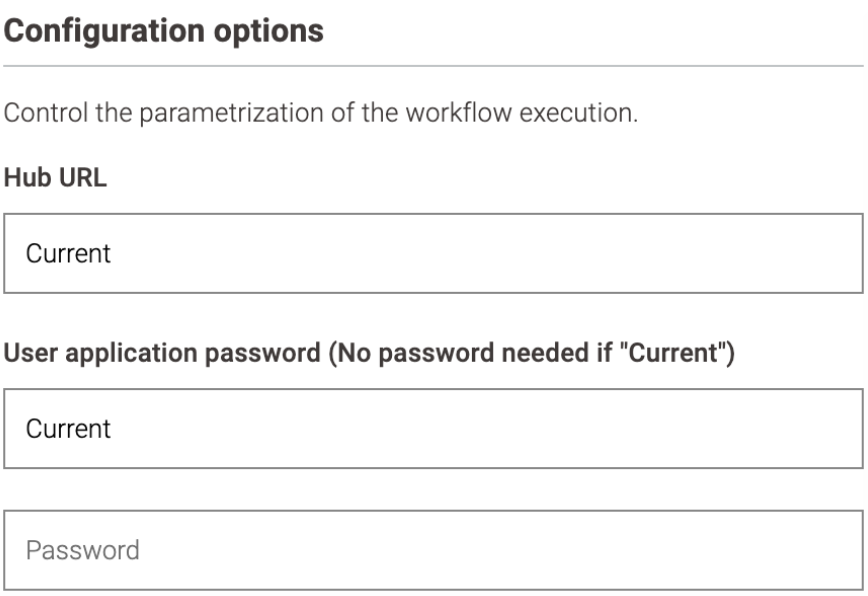

Figure 2. Hub URL and user application password configuration.

6. *Run* the workflow

User Usage Monitoring

The User Usage Monitoring data app covers the following use cases:

- **Active users**: Count users actively logging in and identify unused license users (learn more about users, consumers, and unlicensed users).
- **Peak login times**: Identify the most active days and times.
- **Access method**: Determine whether users log in via KNIME Analytics Platform, their browser, or both.

Keycloak Service Client configuration

The data app requires **keycloak session events** to be collected and stored to monitor user usage. Here you will configure keycloak such that it stores user events, and setup a service client that has permissions to retrieve these events. This service client will be used by the data app to retrieve the session events.

1. Log into Keycloak Admin Console: *`https://auth.<base-url>` > Administration Console*. Learn how to retrieve the credentials here.
2. Select your *Realm* (usually `knime`).
3. Activate *Save events* to save KNIME Business Hub users events at *Realm Settings > Events > User events settings*, as shown in Figure 3. Configure how many days to keep the session events.

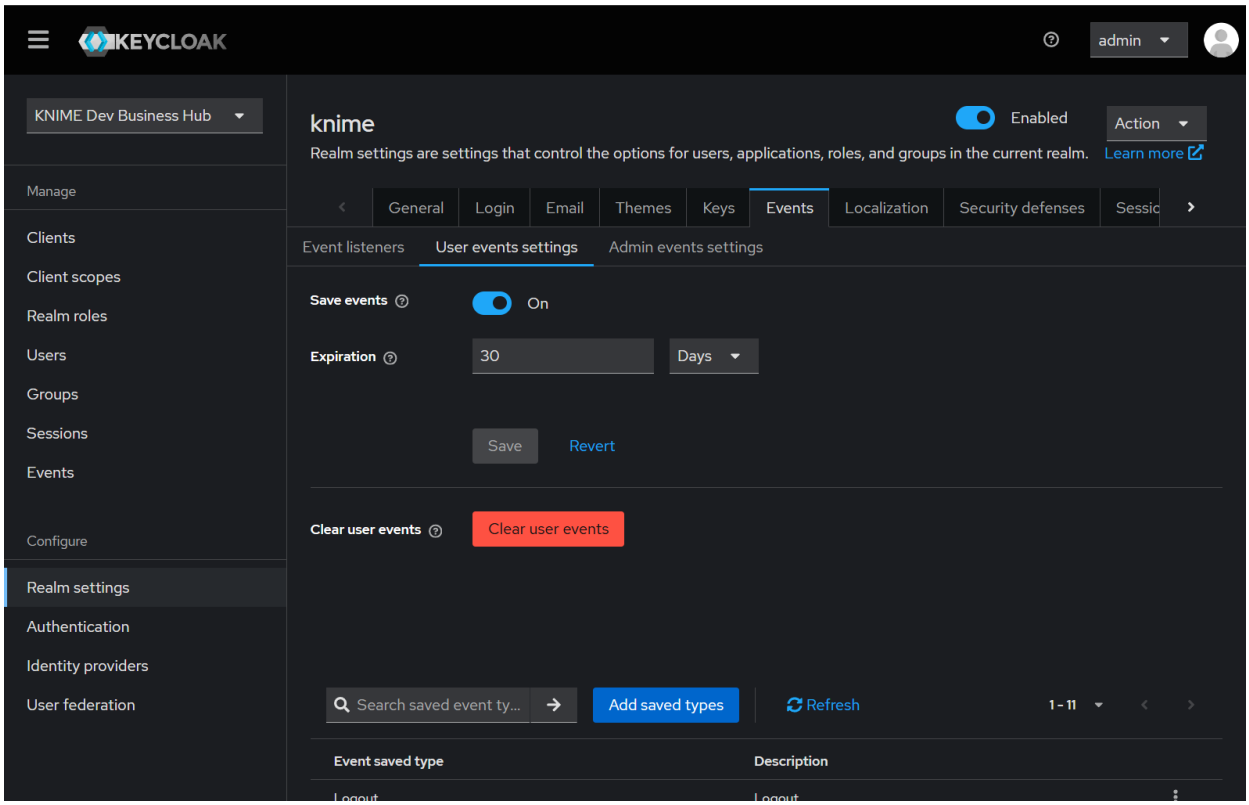

Figure 3. Activate Save events in Realm Settings > Events > User events settings.

a. Select *Add saved types* and add the refresh token event to the list of events that are saved.

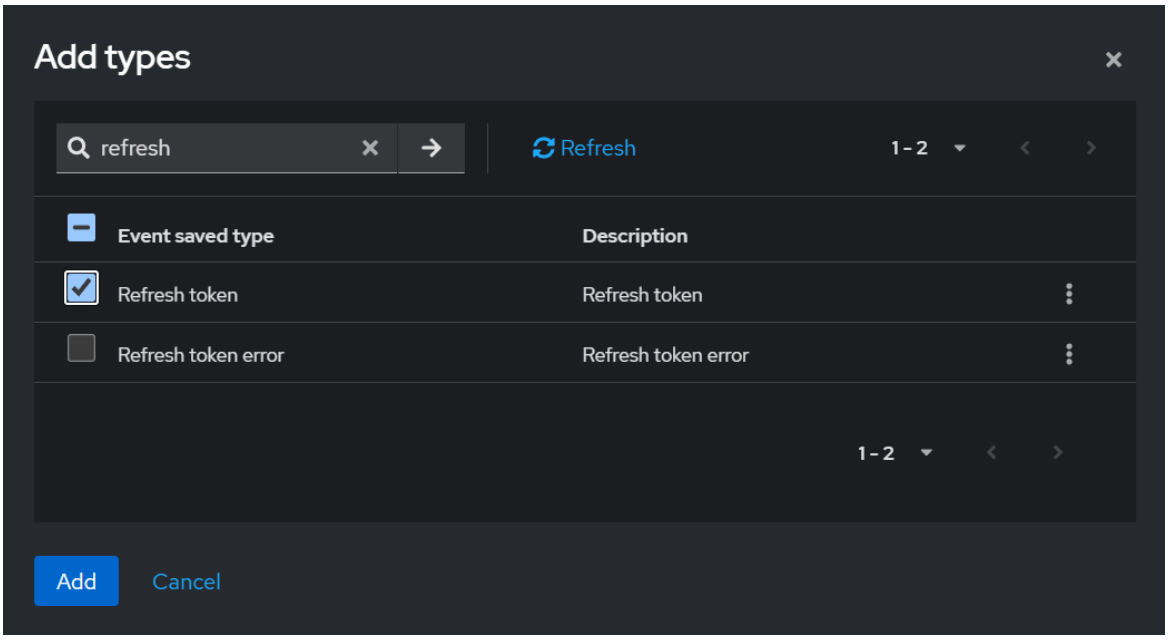

Figure 4. Activate Add saved types in Realm Settings > Events > User events settings.

4. Create a new Client from *Clients > Create Client*.

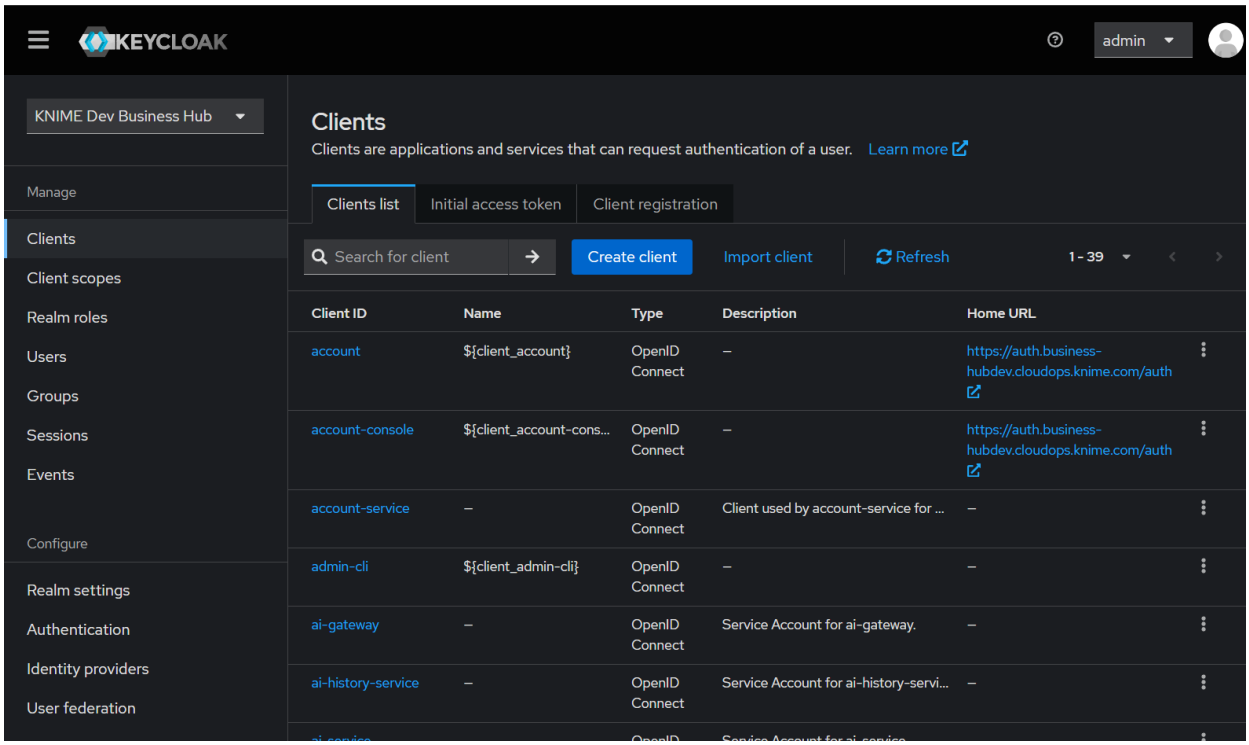

Figure 5. Create a new Client in Clients > Create Client.

a. Give it a Client ID (`monitor-user-usage` in the screenshot). Optionally, add a name and a description to make clear what this client is for.

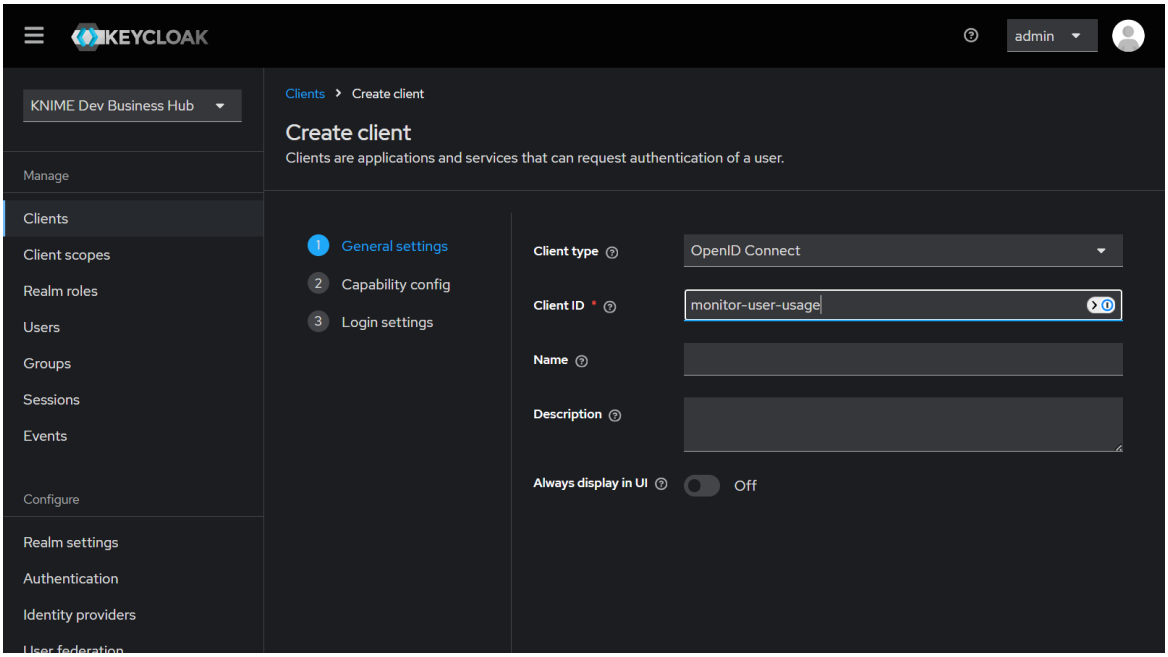

Figure 6. Give the client an ID.

b. In the *Capability config* section, activate *Client authentication*. Further, select *Service account roles* to allow you to authenticate this client to Keycloak and retrieve the access token dedicated to this client. The *Login settings* don't need to be configured.

c. Save your new client.

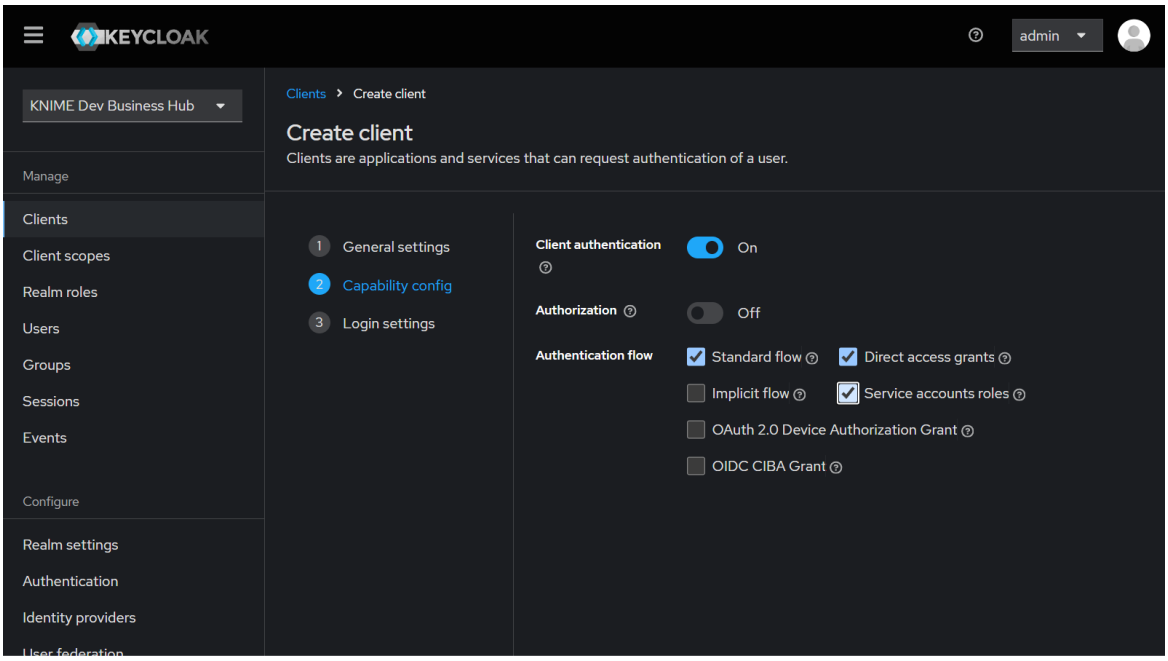

Figure 7. Activate Client authentication and Service account roles in Clients > Create Client.

5. In *Clients > Client ID (Column)*, find and click on your new Client (`monitor-user-usage`).

a. Go to the *Service account roles* tab and click the *Assign role* button.

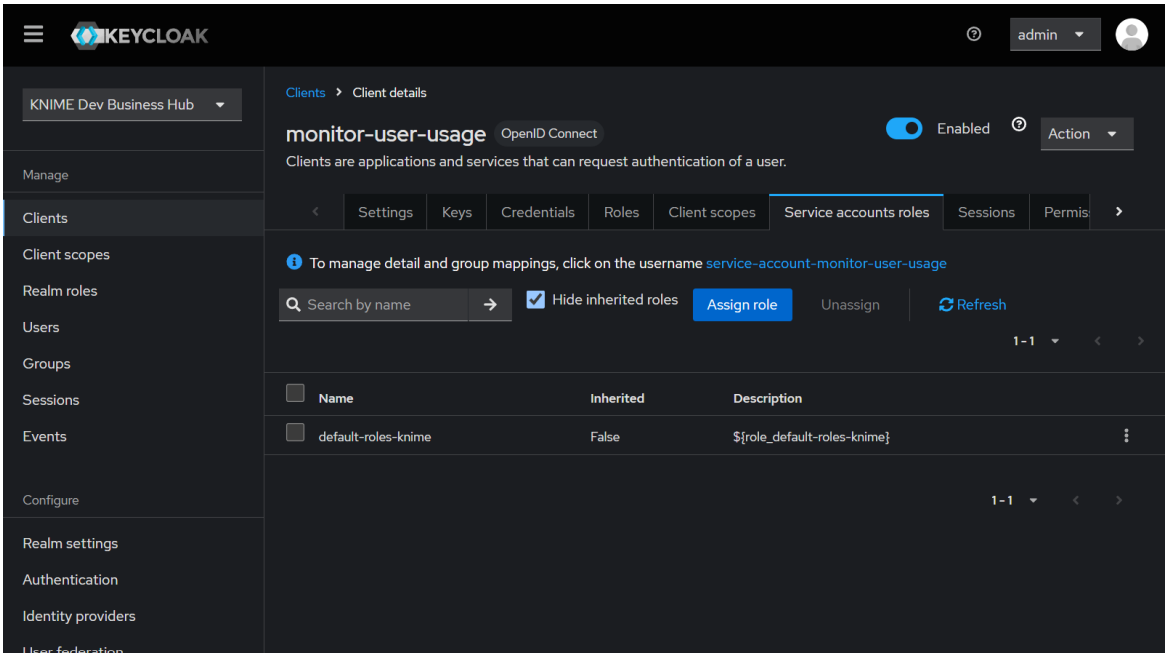

Figure 8. Assign role in Clients > Client ID (Column).

b. Via *Filter by clients*, search for `view-events` and assign the role to the service account associated with your client.

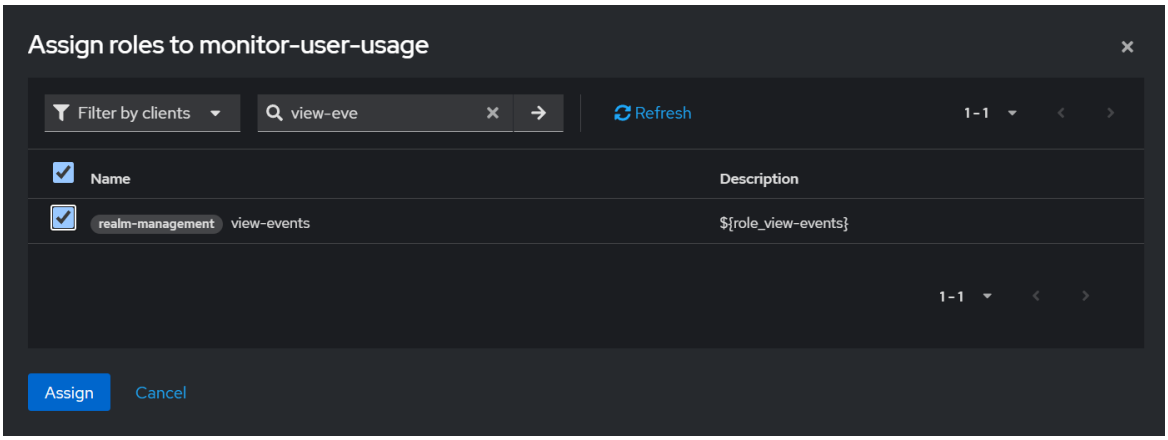

Figure 9. Assign the `view-events` role to the service account.

c. Add the `manage-users` role in a similar fashion.

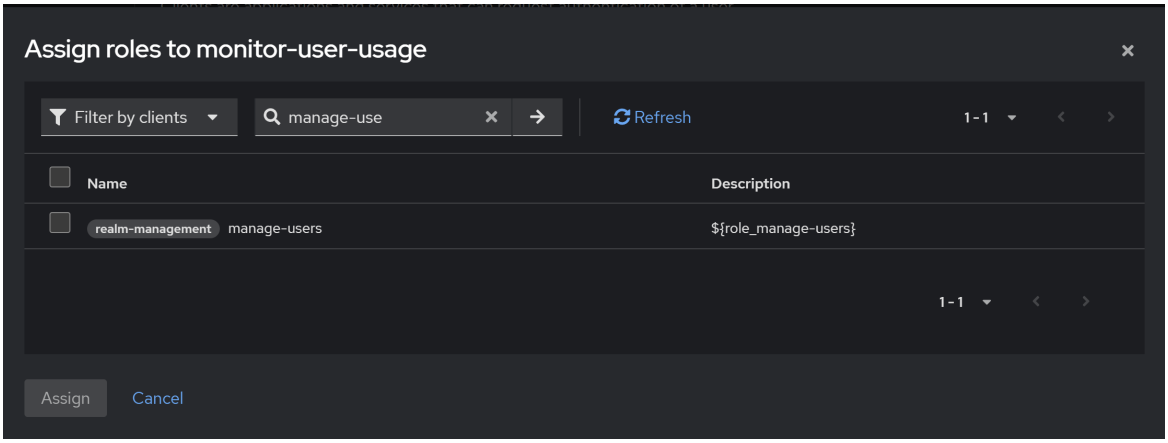

Figure 10. Assign the `manage-users` role to the service account.

6. Finally retrieve the Client ID and Client secret:
   a. Go to *Clients > Client ID (Column)* and choose your client (`monitor-user-usage`)
   b. Click the *Credentials* tab
   c. Leave as a Client Authenticator the *Client ID and Secret* option.
   d. Copy the Client ID from the top of the tab (`monitor-user-usage` in the screenshot below).
   e. Copy the Client's Secret.

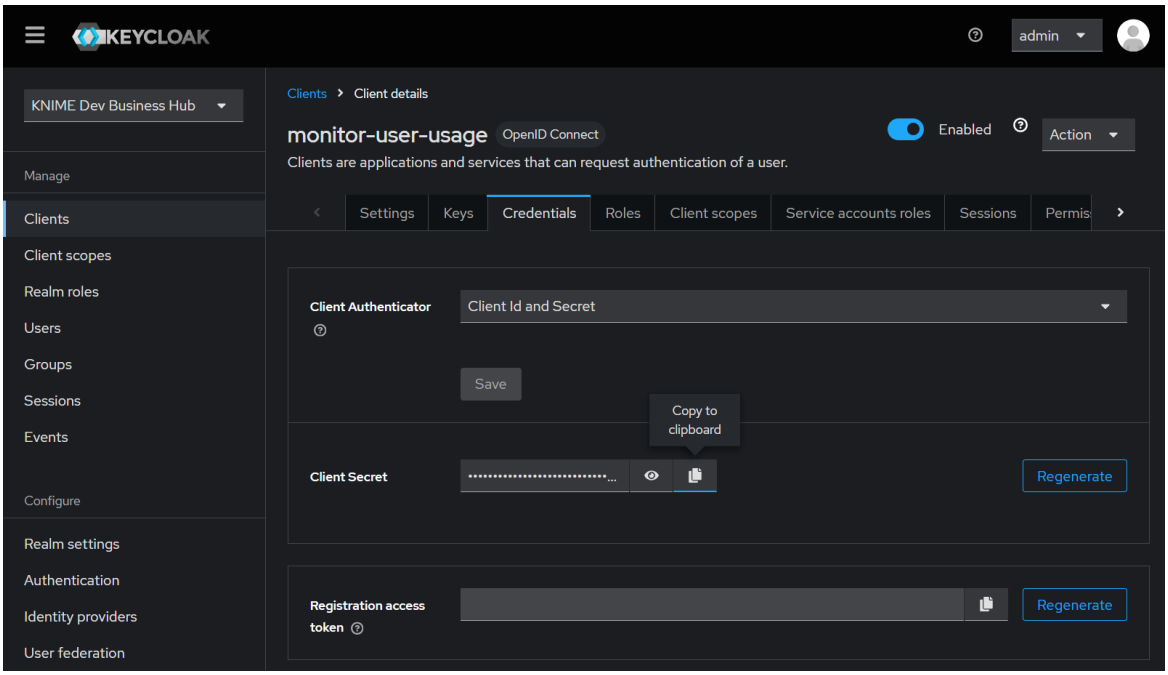

Figure 11. Retrieve the Client ID and Client secret.

Use the retrieved client ID and secret in the deployment, as described below.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more)
    - Optional: Instead, deploy the workflow as a data app to persist any configuration and make it available to others by sharing the deployment (learn how).
5. Configure the keycloak service client ID and secret as retrieved above in the configuration shown in Figure 12.

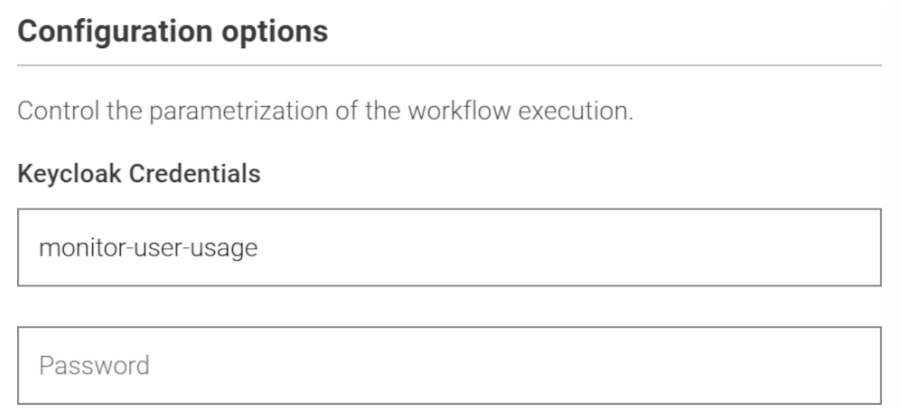

Figure 12. Keycloak service client ID and secret configuration. Password is the secret.

6. *Run* the workflow

Idle Users Cleanup

The Idle Users Cleanup schedule is designed for the following use case:

- **Free up seats by removing inactive users**: Identify users who have not logged in for a configured period and have no active deployments, then remove them from Hub teams.

Inactivity Criteria

A user is identified as inactive if both conditions are met:

- **Login inactivity**: the user has not logged in for the configured number of days.
- **No active deployments**: the user has no active deployment in their user scope (learn more).

After idle users are identified, the workflow removes inactive users from their Hub teams, freeing up seats.

When users are removed from all teams, they automatically become consumers.

A component optionally notifies about team member removals and collects errors from API calls for inspection.

The workflow requires access to Keycloak user information and session event data to track user activity. Therefore, the Keycloak service client credentials (Client ID and Client Secret) must be provided.

Keycloak Service Client configuration

The data app requires **keycloak session events** to be collected and stored to monitor user usage. Here you will configure keycloak such that it stores user events, and setup a service client that has permissions to retrieve these events. This service client will be used by the data app to retrieve the session events.

1. Log into Keycloak Admin Console: *`https://auth.<base-url>` > Administration Console*. Learn how to retrieve the credentials here.
2. Select your *Realm* (usually `knime`).
3. Activate *Save events* to save KNIME Business Hub users events at *Realm Settings > Events > User events settings*, as shown in Figure 3. Configure how many days to keep the session events.

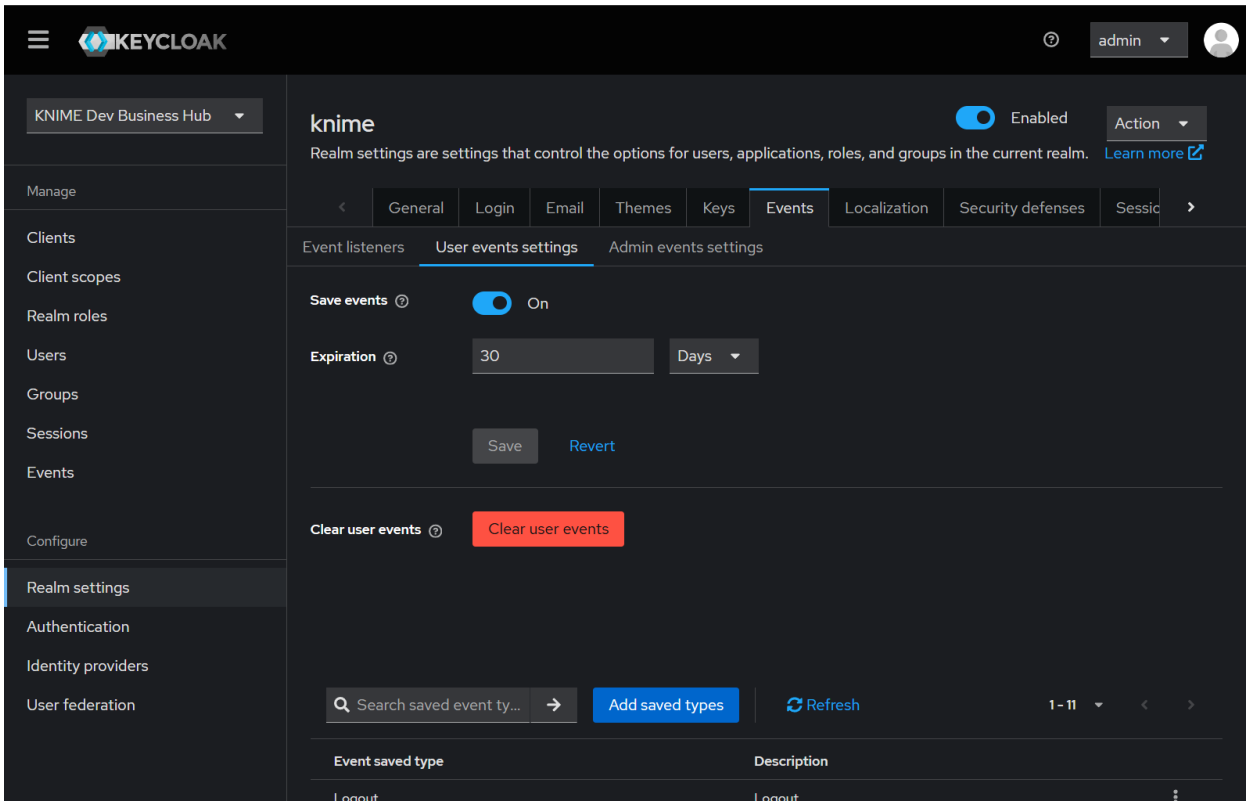

Figure 13. Activate Save events in Realm Settings > Events > User events settings.

a. Select *Add saved types* and add the refresh token event to the list of events that are saved.

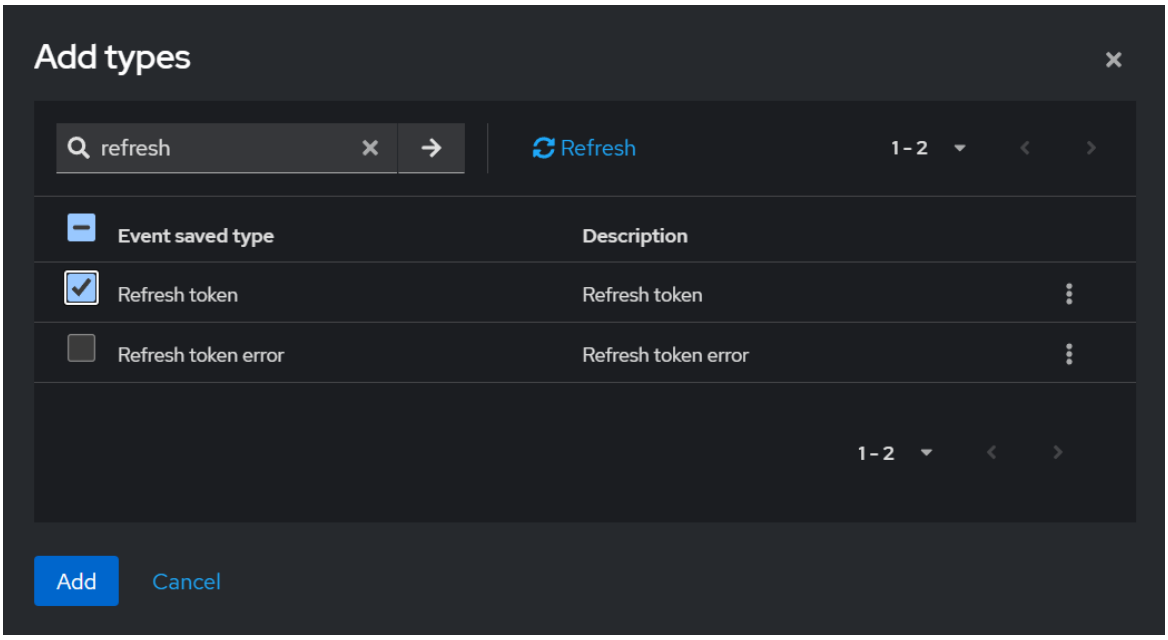

Figure 14. Activate Add saved types in Realm Settings > Events > User events settings.

4. Create a new Client from *Clients > Create Client*.

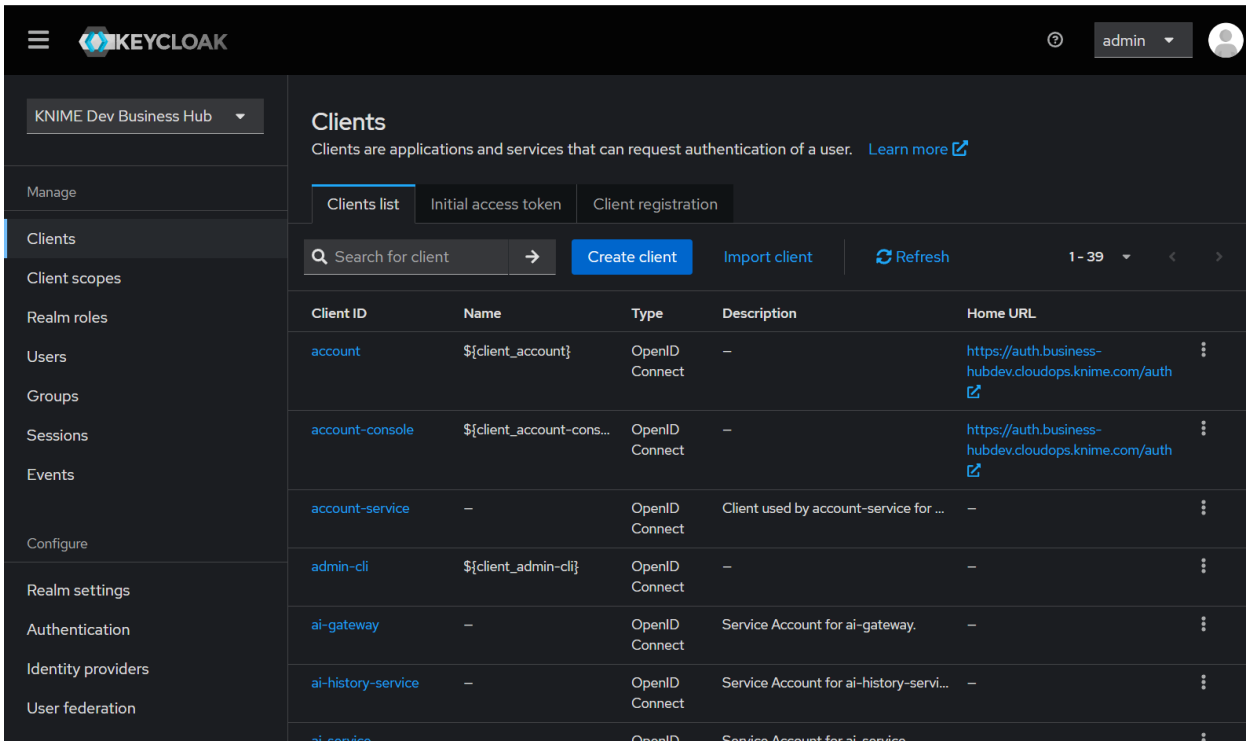

Figure 15. Create a new Client in Clients > Create Client.

a. Give it a Client ID (`monitor-user-usage` in the screenshot). Optionally, add a name and a description to make clear what this client is for.

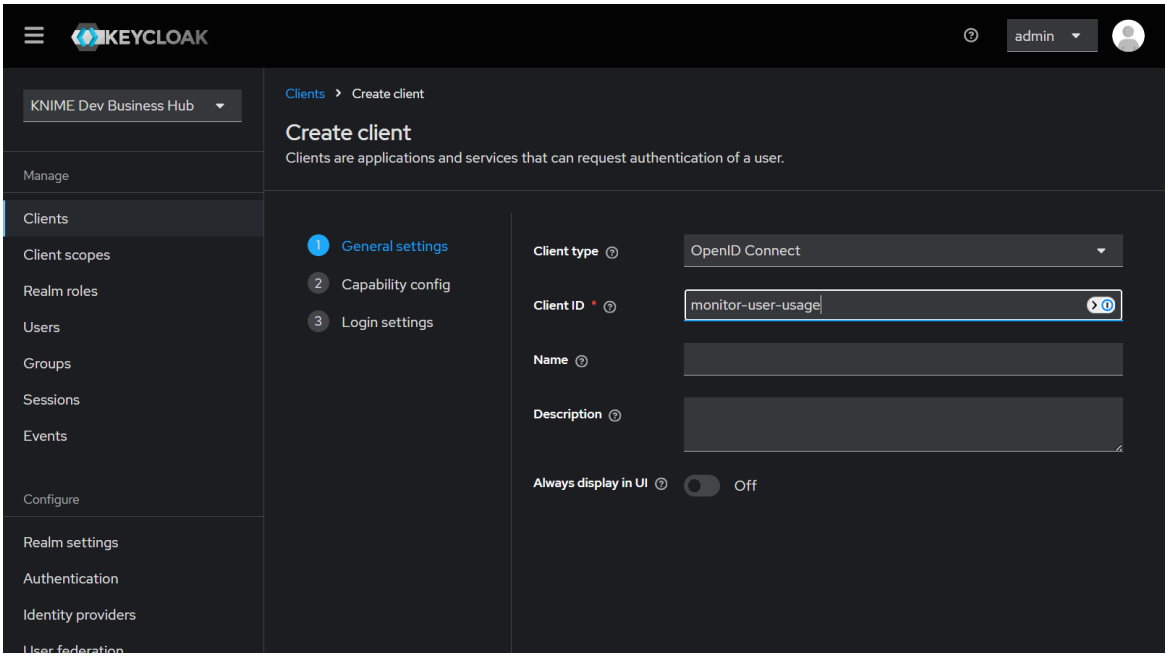

Figure 16. Give the client an ID.

b. In the *Capability config* section, activate *Client authentication*. Further, select *Service account roles* to allow you to authenticate this client to Keycloak and retrieve the access token dedicated to this client. The *Login settings* don't need to be configured.

c. Save your new client.

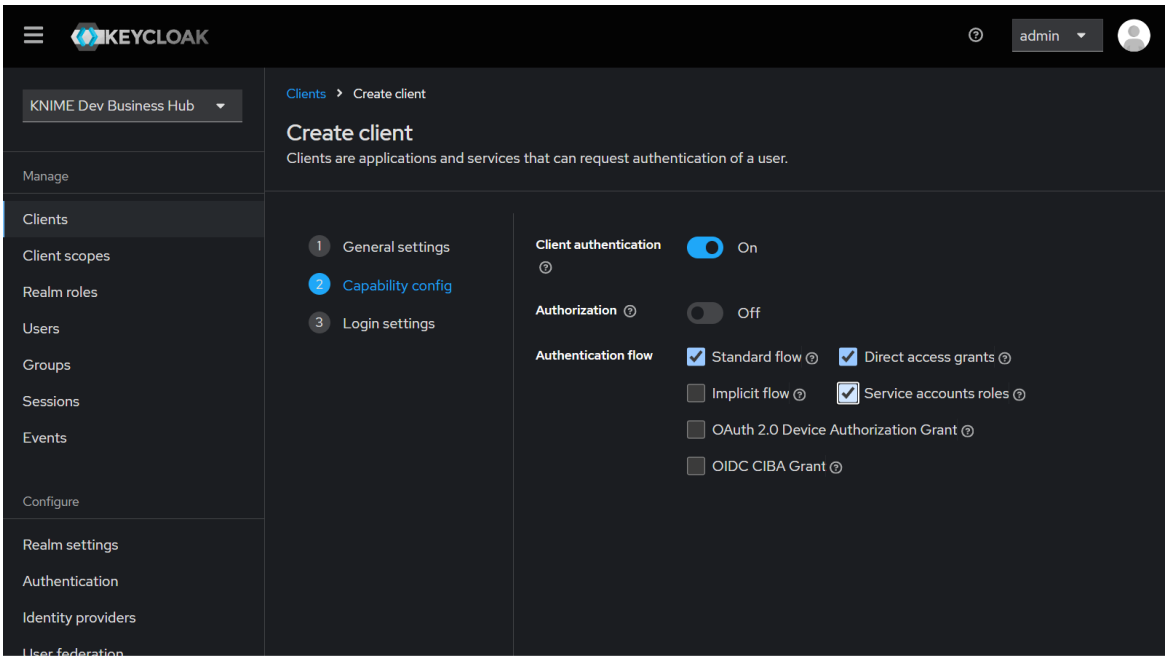

Figure 17. Activate Client authentication and Service account roles in Clients > Create Client.

5. In *Clients > Client ID (Column)*, find and click on your new Client (`monitor-user-usage`).

a. Go to the *Service account roles* tab and click the *Assign role* button.

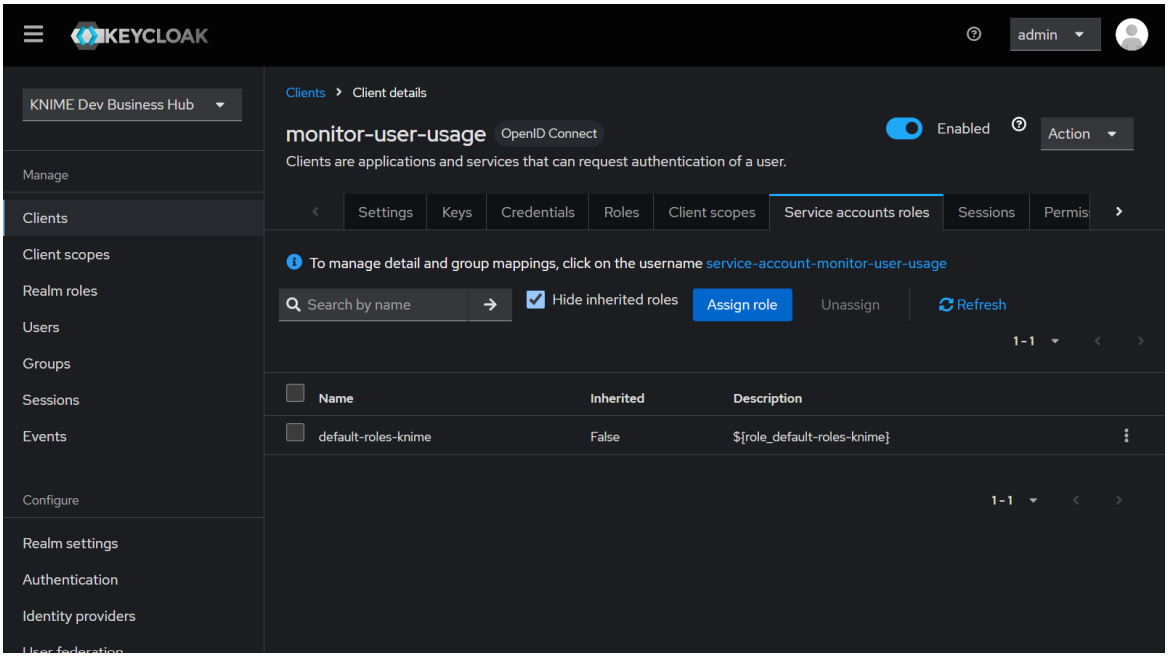

Figure 18. Assign role in Clients > Client ID (Column).

b. Via *Filter by clients*, search for `view-events` and assign the role to the service account associated with your client.

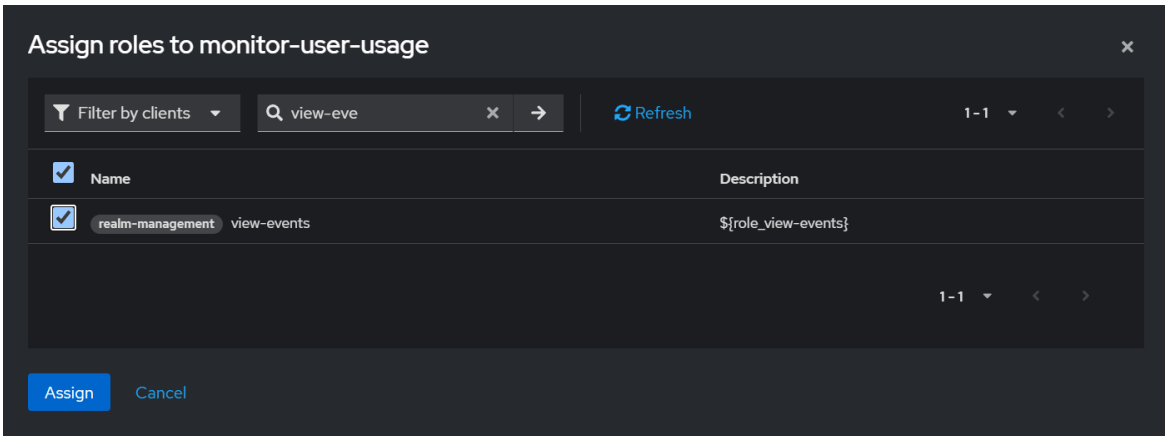

Figure 19. Assign the `view-events` role to the service account.

c. Add the `manage-users` role in a similar fashion.

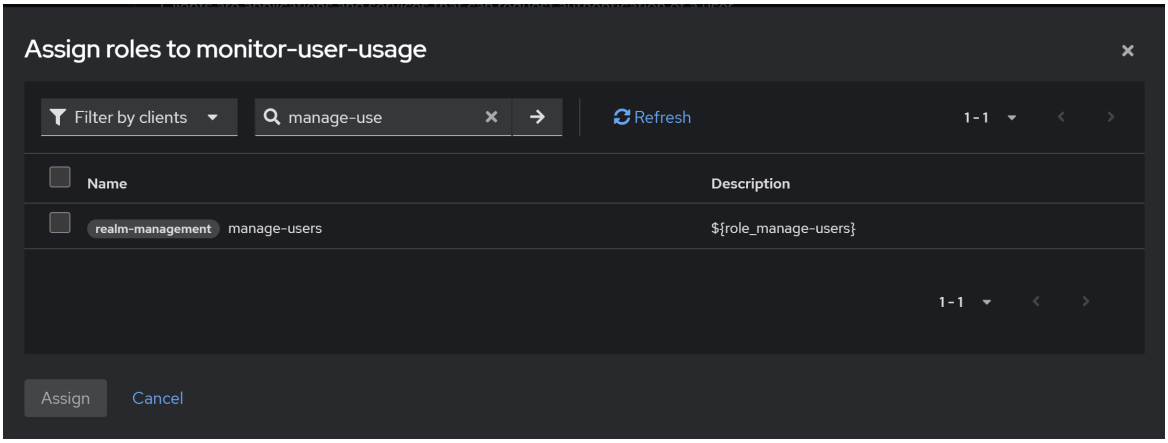

Figure 20. Assign the `manage-users` role to the service account.

6. Finally retrieve the Client ID and Client secret:
   a. Go to *Clients > Client ID (Column)* and choose your client (`monitor-user-usage`)
   b. Click the *Credentials* tab
   c. Leave as a Client Authenticator the *Client ID and Secret* option.
   d. Copy the Client ID from the top of the tab (`monitor-user-usage` in the screenshot below).
   e. Copy the Client's Secret.

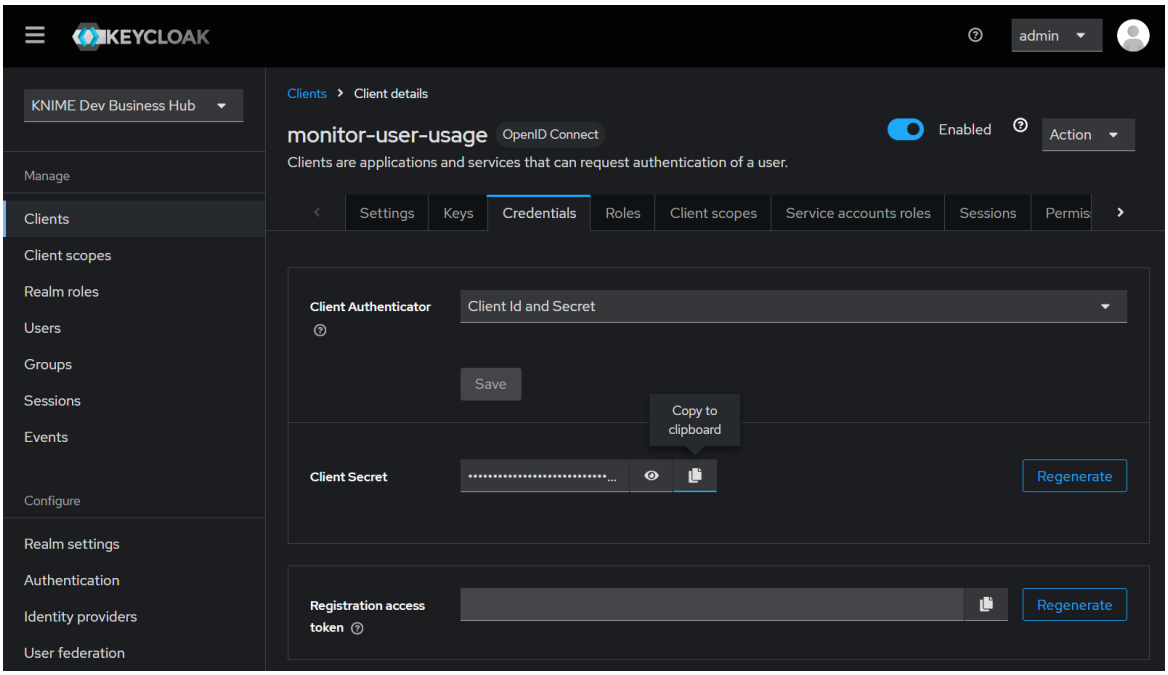

Figure 21. Retrieve the Client ID and Client secret.

Use the retrieved client ID and secret in the deployment, as described below.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more) or *Schedule* it (learn more)
5. Optional: impersonate another user or configure another Hub in the configuration panel shown in Figure 22.
    - User application password: specify the user you want to impersonate.
        - "Current" (default) for the currently logged in user. The password is not needed.
        - Application password of the user you want to impersonate. This allows you to use permissions of another user.
    - Hub URL: specify the Hub you want the data app to refer to.
        - "Current" (default) for the Hub on which the execution takes place.
        - The URL of the other Hub you want to refer to.

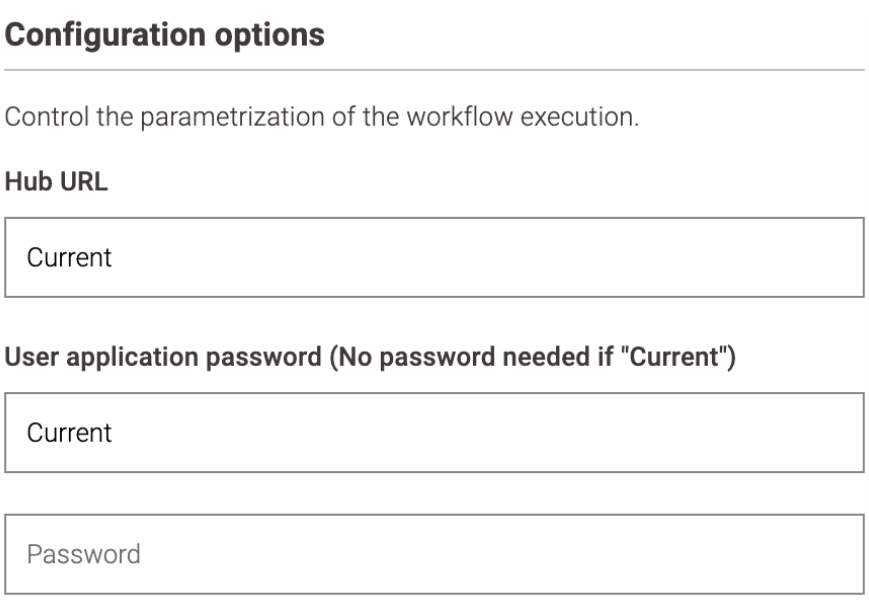

Figure 22. Hub URL and user application password configuration.

6. Configure the Keycloak service client ID and secret as retrieved above in the configuration shown in Figure 12.

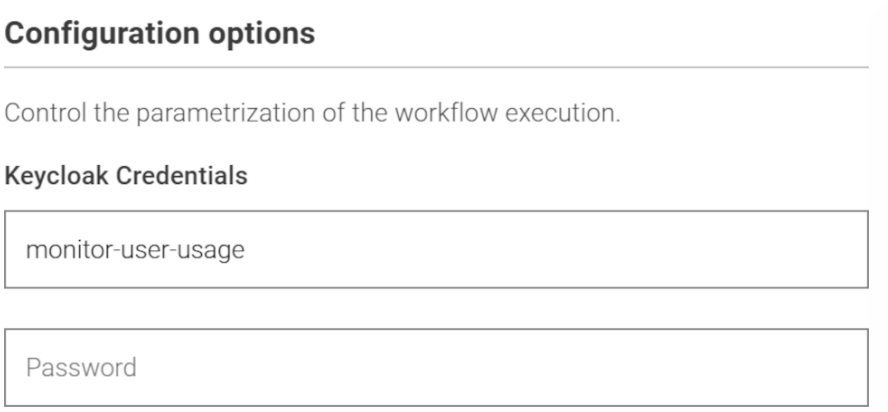

Figure 23. Keycloak service client ID and secret configuration. Password is the secret.

7. Define the number of days of inactivity that determines when a user is considered inactive. The default is 30 days, but you can adjust this value as needed.
8. *Run* or *Schedule* the workflow

Notification setup (optional)

At the end of the workflow you will find a **Notify User** component. It is disabled by default via an **Active Branch Inverter** node, remove the inverter to enable it. To use it, configure standard mail server settings such as **SMTP host and port**, **encryption type**, and **recipients and sender** and **credentials**. The email **subject and content** are already provided via flow variables.

Executor Diagnostics

The Executor Diagnostics data app covers the following use cases:

- **Verify capabilities**: List extensions, nodes, conda environments, and Linux packages that are installed on the execution context the data app is running on.
- **Replicate conda environments**: Download the conda environment of the execution context to replicate it on another machine.
- **Troubleshoot network configuration**: Check the proxy settings in the execution context and verify if external URLs are reachable. If a URL is not reachable, receive hints regarding potential reasons.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more)
    - Optional: Instead, deploy the workflow as a data app to persist any configuration and make it available to others by sharing the deployment (learn how).
5. *Run* the workflow

Gen AI Monitoring

The Gen AI Monitoring data app covers the following use cases:

- **Investigate gen AI assistance usage**: Measure engagement with K-AI, script assistance, and user prompts.
- **Gen AI usage cost**: Analyze the cost associated with Gen AI usage.

Enable AI Services

The data displayed by this data app is only available when the AI Service is enabled, as described here.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more)
    - Optional: Instead, deploy the workflow as a data app to persist any configuration and make it available to others by sharing the deployment (learn how).
5. Optional: impersonate another user or configure another Hub in the configuration panel shown in Figure 24.
    - User application password: specify the user you want to impersonate.
        - "Current" (default) for the currently logged in user. The password is not needed.
        - Application password of the user you want to impersonate. This allows you to use permissions of another user.
    - Hub URL: specify the Hub you want the data app to refer to.
        - "Current" (default) for the Hub on which the execution takes place.
        - The URL of the other Hub you want to refer to.

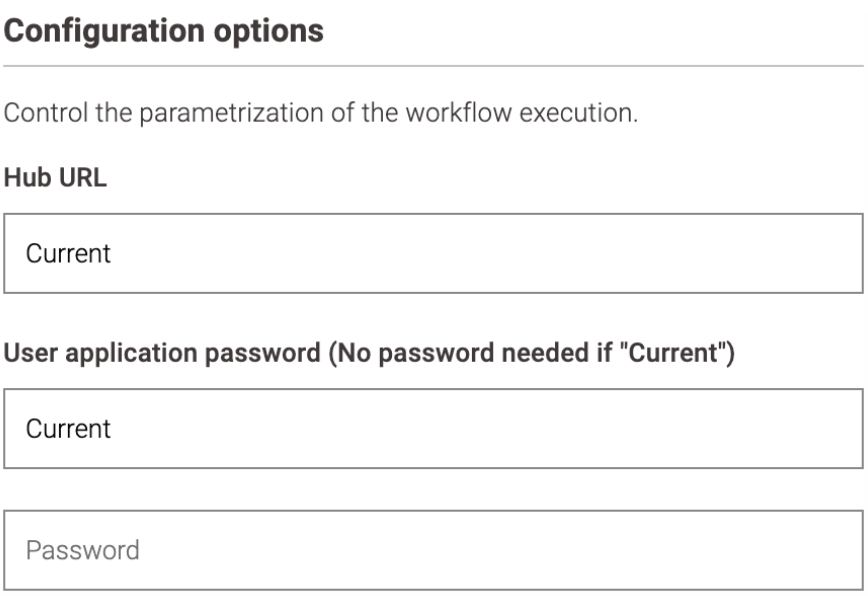

Figure 24. Hub URL and user application password configuration.

6. *Run* the workflow

Manage Existing Jobs

The Manage Existing Jobs data app covers the following use cases:

- **Inspect all existing jobs**: View and monitor all jobs that remain accessible in the executor's memory or object store.
- **Jobs cleanup and storage optimization**: Delete jobs to free up disc space in object store or memory in executors.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more)
    - Optional: Instead, deploy the workflow as a data app to persist any configuration and make it available to others by sharing the deployment (learn how).
5. Optional: impersonate another user or configure another Hub in the configuration panel shown in Figure 25.
    - User application password: specify the user you want to impersonate.
        - "Current" (default) for the currently logged in user. The password is not needed.
        - Application password of the user you want to impersonate. This allows you to use permissions of another user.
    - Hub URL: specify the Hub you want the data app to refer to.
        - "Current" (default) for the Hub on which the execution takes place.
        - The URL of the other Hub you want to refer to.

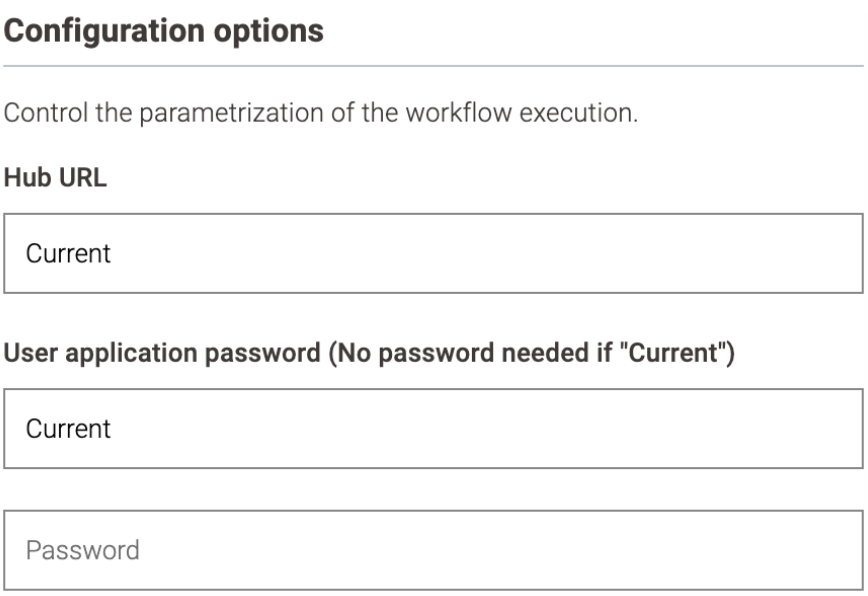

Figure 25. Hub URL and user application password configuration.

6. *Run* the workflow

Manage Workflow Versions

The Manage Workflow Version data app covers the following use cases:

- **Identify unused versions**: Detect workflow versions that are no longer in use in any deployment.
- **Version cleanup**: Delete outdated and unused workflow versions to free up disc space.

Setup

1. Download the workflow
2. Navigate to the space on your Hub where you want to upload the workflow to
3. Use the upload button to upload the workflow (learn more)
4. *Run* the workflow to start an ad hoc execution (learn more)
    - Optional: Instead, deploy the workflow as a data app to persist any configuration and make it available to others by sharing the deployment (learn how).
5. Optional: impersonate another user or configure another Hub in the configuration panel shown in Figure 26.
    - User application password: specify the user you want to impersonate.
        - "Current" (default) for the currently logged in user. The password is not needed.
        - Application password of the user you want to impersonate. This allows you to use permissions of another user.
    - Hub URL: specify the Hub you want the data app to refer to.
        - "Current" (default) for the Hub on which the execution takes place.
        - The URL of the other Hub you want to refer to.

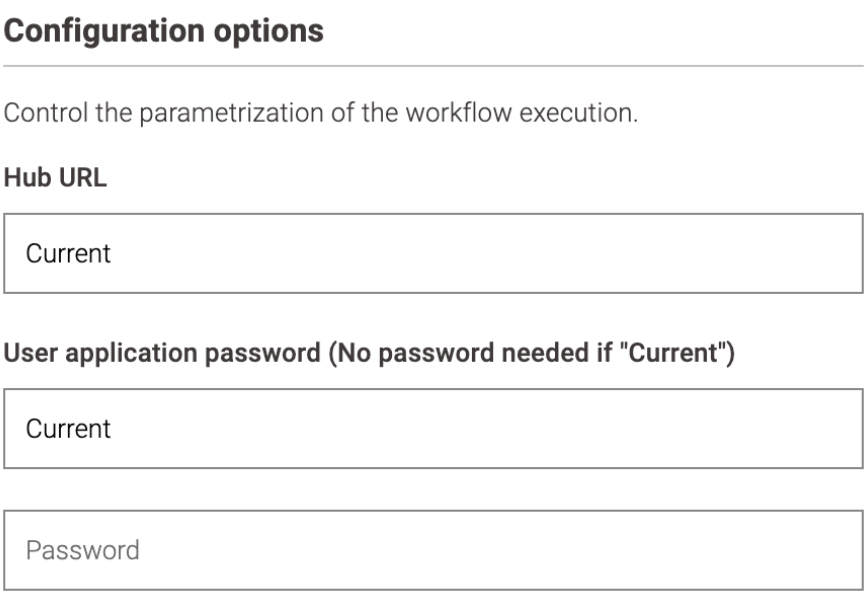

Figure 26. Hub URL and user application password configuration.

6. *Run* the workflow

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