StreamSets Control Hub On-Premises
3.7.1 Release Notes

November 28, 2018

We're happy to announce a new version of StreamSets Control Hub On-Premises. This version contains several new features, enhancements, and some important bug fixes.

This document contains important information about the following topics for this release:

- New Features
- Fixed Issues
- Known Issues

New Features

This version includes the following new features and enhancements in the following areas.

Upgrade

When you update schemas in the relational database, you now run the same database initialization script for all upgrades. Previously, you had to run different commands based on the version that you were upgrading from.

Preview in Pipeline Designer

Pipeline Designer can now display preview data in table view.

Subscriptions

Subscriptions include the following enhancements:

- **Pipeline status change event** - You can now configure a subscription action for a changed pipeline status. For example, you might create a subscription that sends an email when a pipeline status changes to RUN_ERROR.

- **Expression completion to filter events** - You can now use expression completion to determine the functions and parameters that you can use for each subscription filter.

Scheduler

The Control Hub scheduler can now stop a job at a specified frequency. For example, you might want to run a streaming job every day of the week except for Sunday. You create one scheduled task that starts the job every Monday at 12:00 am. Then, you create another scheduled task that stops the same job every Sunday at 12:00 am.

SAML Authentication

When you map a Control Hub user account to a SAML IdP user account, the **SAML User Name** property now defaults to the email address associated with the Control Hub user account. Previously, the default value was the user ID associated with the Control Hub user account.

Fixed Issues

This release fixes the following known issues:

- When the scheduler starts a job that includes a pipeline with a Hadoop-related stage configured to impersonate the Hadoop user as the currently logged in Data Collector user, Control Hub incorrectly interprets the user who starts the pipeline to be `scheduler000` which causes the pipeline to fail. *(5206)*

- When LDAP authentication is enabled, creating a user should not allow you to assign the user to groups. *(5049)*

- The Control Hub UI takes a long time to display users and groups. *(4935, 4988)*

- The `dpm.service` file used in a tarball or RPM installation started as a service is missing the [Install] section. *(4898)*

Known Issues

Please note the following known issues with this release:

- Pipeline Designer displays credential values in stage properties when the pipeline is viewed in read only mode. *(5156)*

- The number of pipelines displayed on the Dashboards view does not match the number of pipelines in the pipeline repository. *(5125)*

- Cluster pipelines that include an HDFS destination or that run on Spark 2.3 fail to start when the pipelines are run from Control Hub jobs. *(5106)*

- Pipeline Designer does not detect changes to the pipeline when you change a metric, data, or data drift rule. *(5095)*

- A job owner can stop the job when it is running on a Data Collector that he doesn't have access to. *(4996)*

- The Scheduler view can display only 50 scheduled jobs or reports. *(4976)*

- The Control Hub scheduler cannot schedule job templates to create and start job instances on a regular basis. *(4927)*

- A topology displays metrics from a single pipeline instance instead of aggregated metrics from all pipeline instances after a job has been updated to write aggregated statistics to another system.

Workaround: Delete the updated job from the topology and then add the job again. *(4881)*

● When using PostgreSQL for the relational database, the installation process requires that users manually configure the PostgreSQLDelegate class for the Scheduler application. Instead, the installation process should automatically handle the configuration. *(4808)*

● Control Hub fails to connect to the LDAP server when the LDAP password configured in the Control Hub security configuration file, `$DPM_CONF/security-app.properties`, starts and ends with the dollar sign ($). *(3782)*

● When a user schedules a job that he does not have execute permission on, the scheduled task is successfully created. However, the scheduled task fails to trigger and does not display an error message indicating that the user must have execute permission on the job.

Workaround: Grant the user execute permission on the job being scheduled. *(3720)*

● When permission enforcement is enabled, pagination logic on the Jobs view might prevent you from viewing and creating jobs. *(3584)*

● When Control Hub uses SAML or LDAP authentication, you cannot log into registered Data Collectors using the disconnected mode. *(3423)*

● Updating a job to use the latest version of a pipeline loses the existing data SLAs configured for that topology. *(3315)*

● Topology metrics and aggregated metrics for a job which runs on multiple Data Collectors might not accurately reflect the metrics for individual jobs or pipelines.

Workaround: For topology metrics, select a job in the topology canvas to monitor the statistics for the job. *(2906)*

● If you scale up a deployment and then delete unregistered authentication tokens for Data Collectors while waiting for the deployment to scale, the newly provisioned Data Collectors might start but not successfully register with Control Hub. *(2880)*

● Control Hub fails to start on CentOS 6.8 with the following error message:

```
'/dpm.id' permission denied
```

Workaround: Add the following line to the `$DPM_HOME/libexec/dpmd-env.sh` file, where `<directory>` is any directory that the dpm system user has write access to:

```
export DPM_DATA=<directory>
```

After modifying the `dpmd-env.sh` file and restarting Control Hub, the following warning message displays:

```
install: missing file operand
Try `install --help' for more information.
```

You can ignore this message. *(2841)*

Contact Information

For more information about StreamSets, visit our website: https://streamsets.com/.

Check out our Documentation page for doc highlights, what's new, and tutorials: streamsets.com/docs

Or you can go straight to our latest documentation here: https://streamsets.com/documentation/controlhub/latest/help/index.html

To report an issue, to get help from our Google group, Slack channel, or Ask site, or to find out about our next meetup, check out our Community page: https://streamsets.com/community/.

For general inquiries, email us at info@streamsets.com.