Cloudera Data Visualization 7.2.9

Job logs

Date published: 2020-10-30
Date modified: 2024-12-18

CLOUDERA

Legal Notice

Contents

Jobs

In Cloudera Data Visualization, a job in general refers to an action performed based on predefined parameters. The job outlines the basic details of the task, which can be performed either ad hoc or repeatedly, according to a specified interval or trigger. When a job is initiated, it generates a new running instance, and these runs are recorded in the job log.

Jobs can be classified based on their action type:

Email

> The job generates a specific email.

Data Extract

> The job performs a data extraction task. Additionally, an email notification can be sent, for example, if the extraction succeeds or fails.

Jobs can run in the following ways:

Adhoc

> The action is not pre-scheduled or automated. The job is initiated manually by the user to respond to immediate needs.

Scheduled

> The action is performed automatically at specific intervals, following the schedule defined by the schedule interval associated with the job.

Threshold

> The action is triggered when specific conditions or thresholds are met. The schedule interval defines how often the system checks for these criteria.

You can review the information available about your jobs on the Jobs interface, which is accessible by clicking Job Status from the Administration menu.

The page has three tabs: Runs (Latest), Scheduled Jobs, and Intervals.

Job runs

Job runs in Cloudera Data Visualization contain essential information about the completed runs, including start time, duration of operation, result, and any error logs encountered during the process.

You can open the Jobs interface by clicking Job Status from the Administration menu. The interface opens on the Runs (Latest) tab by default. This tab provides information about the most recent instance of the latest 100 jobs. For more comprehensive information about a job or to view its additional previous runs, navigate to the Job Details page by clicking the name of the job.

1. **Filters** are available for job runs based on Last Status, Owner, Type, and Interval. These filters allow for efficient navigation through the logs.
2. **Selection for bulk actions:** There is a checkbox at the beginning of each job run entry, facilitating bulk deletion or rerun operations. Additionally, the checkbox on the header row allows for selecting all job runs for an action.
3. The **Status** of a job can be one of the following: Queued, Running, Finished, Error, or Cancelled.
4. **Job ID** is a unique identifier for each job, shared between the Scheduled Jobs and Runs (Latest) tabs for easy identification and debugging.
5. **Log ID** is a unique identifier for each job log entry. Unlike the Job ID, which remains the same for every instance of the same job, the Log ID is unique for each run.
6. **Name** is the name of the job. You can click the name to open the Job Details page. For more information, see *Job details*.
7. **Owner** identifies the user who created the job. For more information, see *Changing the owner of a job*.
8. **Type** specifies whether the job target is email or data extract.

    > **Note:** Legacy jobs are classified into two categories: Refresh and Other. These jobs cannot be edited through the UI and do not have a dedicated detail page. While rerunning these jobs is not supported, they remain visible for reference.

9. **Interval** provides information on whether the job is adhoc (can be initiated manually) or scheduled (running automatically at specified intervals).
10. **Start Time** is the timestamp indicating when a job commenced, displayed in local time.
11. **Total Run Time** shows the approximate duration of the job run.

12. Available actions:

- Details opens a modal window that displays comprehensive information about the specific job run.
- Run Now initiates a rerun of the job.
- Delete deletes the the job run.

13. DELETE performs a delete on all selected job runs.

14. RUN reruns all selected jobs.

Related Information

Changing the owner of a job

Scheduled jobs

In Cloudera Data Visualization, a scheduled job is a predefined task or action that is set to occur automatically repeated according to a specified interval.

Scheduled jobs are designed to streamline repetitive processes, such as data extraction or report generation by automating their running according to a predetermined schedule. When creating a scheduled job, you can configure the details of the task, such as its frequency, timing, and parameters.

The Scheduled Jobs tab presents a comprehensive list of jobs that follow a running schedule, and shows the options and actions available for managing all your jobs running according to a schedule.

1. **Search**
2. **Filters** for scheduled job are based on State, Type, Owner, Last Run Start, Last Run End and Threshold Only.
3. **Bulk actions:** There is a checkbox at the beginning of each job entry, facilitating bulk deletion, rerun, or resume scheduling. Additionally, the checkbox on the header row allows for selecting all jobs for an action.
4. The **State** of a job can be one of the following: Running, Finished, Error, Scheduling Paused, Never Run, or Unknown.
5. **ID** is a unique identifier for each job, shared between the Scheduled Jobs and Job Log tabs for easy debugging.
6. **Name** is the name of the scheduled job.
7. **Type** specifies whether the job target is email or data extract.
8. **Owner** identifies the user who created the scheduled job. For more information, see *Changing the owner of a job*.
9. **Schedule** indicates the frequency of running the job, either standard types (weekly, hourly) or a custom interval.
10. **Last Run** is the timestamp of the job's last run displayed in local time.

11. Actions:

- Details open the information module of the job log.
- Run Now initiates a rerun of the job.
- Pause Scheduling suspends a scheduled or triggered job indefinitely, while preserving its definition. This allows you to temporarily stop the scheduling of the job without affecting its underlying configuration. Resume Scheduling beomes available and you can resume the paused job.

  Note: Pausing a job does not stop a running instance. If you want to stop a job in progress, use the cancel function instead.

- Edit opens the Edit Job Details modal window and allows you to edit the details of the scheduled job.

  Note: The Edit action is not available for data extract jobs. Editing data extracts is only possible on the Data Extracts page.

- Change Owner opens the Change Job Owner modal window and allows you to change the owner of the scheduled job.
- Delete rallows you to delete the scheduled job

12. Delete performs a delete on all selected jobs.

13. Run reruns all selected jobs.

14. Resume scheduling restarts all selected jobs that have been paused.

  Note: This bulk action only resumes jobs that are paused, and will skip any jobs in the selection that are already active.

Related Information
Changing the owner of a job

Schedule intervals

Intervals in Cloudera Data Visualization refer to the time frames or patterns at which scheduled jobs are set to occur, so these intervals dictate the frequency and timing of the scheduled jobs.

You can specify the schedule intervals on the Intervals tab of the Jobs page. The tab tab provides a comprehensive list of schedule intervals along with their associated features and actions.

Common examples of schedule intervals include hourly, daily, weekly, or monthly occurrences, as well as custom intervals tailored to specific business needs. By defining schedule intervals, you can effectively automate repetitive tasks and ensure that jobs run consistently and on time.

  Note: This feature is only available to users with administrative privileges.

1. **Search**
2. **Bulk actions:** There is a checkbox at the beginning of each schedule interval entry, facilitating bulk deletion. Additionally, the checkbox on the header row allows for selecting all job logs for the action.
3. **ID** is a unique identifier assigned to each schedule interval for internal reference.
4. **Name** is a descriptive title assigned to the schedule interval for easy identification and organization.
5. **Interval** describes the parameters of the interval, such as Every day at 3:00, Every hour at 20 minutes past the hour, or Every month on 1 at 1:30.
6. **Next Run (local)** displays the expected time for the next deployment of the schedule in local time format.
7. **Dependent Jobs** list any scheduled jobs that use the specified schedule.
8. The 🖉 icon enables you to change the properties of existing schedules. For more information, see *Changing schedule intervals*.
9. **Delete** performs a delete on all selected intervals.
10. **Create Schedule Group** enables you to specify a custom schedule interval. For more information, see *Creating new schedules*.

Related Information

Creating new schedules

Changing schedule intervals

Job details

The Job Details page provides comprehensive insights into a selected job, helping you understand and manage job operations effectively. You can access this page by clicking the name of the job on either the Runs (Latest) or Scheduled Jobs tab.

Job Details

Job Details   Actions ▾

SCHEDULE STATUS   NAME   ID
● Active   Doc test schedule   151

TYPE   SCHEDULE   TARGET
Scheduled   my interval   Email

OWNER
vizapps_admin

In this section, you can review basic information about the job and its scheduling parameters, as well as perform the following actions:

- Edit opens the Edit Job Schedule Details modal window, allowing you to modify schedule details.

  Note: The Edit action is not available for data extract jobs. For more information on changing data extract details, see *Editing a data extract*.

- Change Owner opens the Change Job Owner modal window, allowing you to assign a new owner to the job.

  Note: When the job owner is changed, the ownership of the related logs is also updated accordingly.

- Delete allows you to remove the job if it is no longer needed.

  Important: Deleting a scheduled job deletes the job itself along with all associated job logs.

- Run Now allows you to start the job immediately.
- Pause Scheduling allows you to suspend a scheduled or triggered job indefinitely while preserving its definition. This action temporarily stops the job scheduling without affecting its underlying configuration. You can resume a paused job.

  Note: Pausing a job does not stop a running instance. If you want to stop a job in progress, use the cancel function instead.

- Resume Scheduling allows you to restart a previously paused scheduled or triggered job.

Depending on the type of the job you are checking, you can also review data extract or email details on this page.

Email Details

∧ Email Details

VISUAL ID   VISUAL URL
68   http://   cdv.viz-team-4.cc-os.cloudera.com/arc/apps/app/68 ⬈

TO   CC
test.user@cloudera.com   -

ERROR TO   REPLY TO
-                                      -

SUBJECT
[Cloudera Data Visualization]

MESSAGE
This is the email message. I hope you are having a good day.

EMAIL TEMPLATE
Default Email Template ⬈

ATTACHMENTS
Embedded image, PDF, XLSX, CSV

PAGE PARAMETERS
Show

In this section, you can review information about email jobs including subject, message content, email template, attachments, and other relevant details.

Data Extract Details

In this section, you can explore specifics like extract ID, source dataset, target data connection, and other details related to the data extraction process.

Previous Runs

In this section, you can see a comprehensive list of past job operations, which can help in diagnosing and troubleshooting errors.

Filters

You can adjust the list of displayed previous runs using the following filters:

- Status – This filter is based on job status.
- Start Time – This filter is based on the start time of job runs. The selected value will be the minimum start time of the displayed runs.
- End Time – This filter is based on the end time of job runs. The selected value will be the maximum end time of the displayed runs.

Actions

You can perform the following actions:

- allows you to delete logs from the list of previous runs. To select logs for deletion, click the checkbox at the start of each log row. To select all logs, click the checkbox on the left side of the list header.

   Note: This action will only delete the job logs, not the job itself.

- allows you to stop a running job. This option is only active for jobs currently in progress.

- opens the Run Details modal, which includes information about the tasks performed in connection with the job instance and any errors encountered.

Related Information
Editing a data extract