

Overview of Job Monitoring

Contents:

- *Monitoring Phases*
 - *Connect*
 - *Request*
 - *Transfer*
 - *Prepare*
 - *Process*
 - *Enable*
 - *Configure*
 - *Enable phases in Data sources tab*
 - *Enable phases in Outputs tab*
 - *Monitoring Jobs in the Application*
 - *Flow View*
 - *Import*
 - *Job Details Page*
-

Trifacta® Self-Managed Enterprise Edition supports detailed monitoring of a job throughout each phase of its execution.

Limitations:

- Applies only to ingest and publishing jobs
- Applies only to JDBC datasets

Monitoring Phases

These phases apply to ingest and publishing jobs. Information on them is surfaced in the application.

Connect

In the Connect phases, Trifacta Self-Managed Enterprise Edition uses the specified connection for the flow to connect to the source of the job.

NOTE: Errors in this phase typically involve issues in the connection definition or in the network configuration or availability.

Request

After the platform has been able to connect to the datastore, the Request phase entails the submission of the request to the datastore for the assets. For example, for JDBC-based datasets, this phase covers the SQL query of the database through the response that the query was successfully executed.

NOTE: Errors in this phase typically reflect errors in the SQL query, which can include renaming or moving of assets in the datastore.

NOTE: If assets are retrieved via custom SQL query, you may need to review the query and validate it through the Trifacta application. For more information, see *Create Dataset with SQL*.

Transfer

This phase covers the transfer of assets from the datastore to the platform.

NOTE: Errors in this phase typically indicate issued with permissions.

Prepare

NOTE: This phase applies to publishing jobs only.

Depending on the destination, the Prepare phase includes the creation of temporary tables, generation of manifest files, and the fetching of extra connections for parallel data transfer.

Process

After the data has been transferred to the platform, this phase covers the processing of cleanup after data transfer, including the dropping of temporary tables or copying data within the instance.

Enable

The base feature is enabled by default.

```
"feature.enableJobMonitoring": true,
```

Configure

Optionally, you can enable the following capabilities in the Trifacta application. You can apply this change through the *Admin Settings Page* (recommended) or

```
trifacta-conf.json
```

. For more information, see *Platform Configuration Methods*.

Enable phases in Data sources tab

To display separate columns in the Data sources tab of the Job Details page for each phase on an ingest job, set the following parameter to `true`:

```
"jobMonitoring.enablePhasesInDatasourcesTable": true,
```

Enable phases in Outputs tab

To display separate columns in the Outputs tab of the Job Details page for each phase for a publish job, set the following parameter to `true`:

```
"jobMonitoring.enablePhasesInOutputsTable": true,
```

Save your changes and restart the platform.

Monitoring Jobs in the Application

When the base feature is enabled, you can monitor jobs in the following locations.

Flow View

- Track phases in the Jobs panel in Flow View. Hover the mouse over the link to the job.
- See *Flow View Page*.

Import

NOTE: This feature may require enablement in your deployment. See *Configure JDBC Ingestion*.

Import Data:

For long-loading datasets, you can track the progress of the import through the Import Data page as you specify the import. See *Import Data Page*.

Library:

After specifying the import, if the data is continuing to be ingested, you can track progress through the Library page. See *Library Page*.

Dataset Details Page:

In the Dataset Details page, you can monitor the ingest progress. Hover over the Status link.

The screenshot displays the 'Dataset Details Page' for 'SQL Dataset 1'. It includes a sidebar with metadata: 'Last updated: 06/15/2020', 'Connection: postgres', 'Status: Completed <1.10 sec', and 'Used in 1 Flow'. The main content area shows a table of job phases: 'Connect' (Ran for <1 sec), 'Request' (Ran for <1 sec), and 'Transfer' (Ran for <1 sec, 0.0017 MB, 0.0017 MB/sec). A table below lists 'Name' (Untitled Flow), 'Owner' (Administrator), 'Objects' (1 Dataset, 1 Recipe), and 'Last Updated' (Last Thursday at 9:17 PM). Action buttons include 'Wrangle in new Flow', 'Preview', 'Edit custom SQL', and a menu icon.

Figure: Dataset Details Page - Job Monitoring

Job Details Page

- Track phases of progress by hovering over the job in progress in the Job Details page.
- Review new and better detail in the Job Details page. Click **View Details** for the job listing.
- For more information, see *Job Details Page*.

Datasources tab - Phased ingest monitoring

If job monitoring phases have been enabled for the Datasources tab, the tab looks like the following:

Monitoring > Farmers_Market - 2

Job 7
Started Today at 12:36 PM

Cancel job ...

Overview Output Destinations Profile Dependencies Data sources


Name	Queued	Connect	Request	Transfer	Process	Status
Farmers_Market	3 sec	✔ <1 sec	✔ <1 sec	✔ 9 sec	-	✔ Completed • 13 sec

Figure: Job monitoring in the Datasources tab

View details:

If an ingest job succeeds or fails, you can click **View details** in the status column for additional information on each phase of the ingest job:

Details
×

 Farmers_Market

Details
SQL

Type	oracle
Created	June 24th 2019, 11:38 am
Ingestion	✔ Completed • 13 sec

Ingestion details

Queued	In queue for 3 sec
Connect	✔ Ran for <1 sec
Request	✔ Ran for <1 sec
Transfer	✔ Ran for 9 sec 76.8760 MB 8.5418 MB/sec

Figure: View details on monitoring ingest jobs

Output destinations tab - Phased publishing monitoring

If job monitoring phases have been enabled for the Output Destinations tab, the tab looks like the following:

You can download the generated results locally or [publish](#) to another storage location.



Name	Queued	Connect	Prepare	Transfer	Process	Status
 test_20190624_193621	2 sec	 <1 sec	 <1 sec	 <1 sec	 <1 sec	 Completed • 15 sec

Figure: Job monitoring in the Datasources tab

View details:

If a publishing job succeeds or fails, you can click **View details** in the status column for additional information on each phase of the publishing job:

Details
×

 test_20190624_193621

Data preview

# FM_ID	RBC	Market_Name
1005627		Montgomery Farmers' Market
1002686		Hines Veterans Hospital Farmers Market
1004509		Montgomery Village Farmers Market
1001779		Farmstand at CHSAS
1008852		Monticello Farmers' Market
1007417		Farmville Farmers Market
1002538		Monticello Old Jail Market

Type	ORACLE
Location	MKOHLI_TEST/test_20190624_193621
Size	3 columns · 3 types
Start Time	June 24th 2019, 12:36 pm
Publish	✔ Completed · 51 sec

Publish details

Queued	In queue for 38 sec
Connect	✔ Ran for <1 sec
Prepare	✔ Ran for <1 sec
	5 connections

Figure: View details on monitoring ingest jobs