

Configure Spark Running Environment

Contents:

- *Limitations*
- *Enable Spark Execution Environment*
- *Use Spark Execution Environment*
- *Change Limits*
 - *Change Spark settings per job*
 - *Enable user-specific overrides*

This section provides information on how to enable and configure the Spark running environment, which leverages Spark's faster in-memory processing to deliver better execution performance.

Limitations

NOTE: When a recipe containing a user-defined function is applied to text data, any non-printing (control) characters cause records to be truncated by the Spark running environment during job execution. In these cases, please execute the job on the Trifacta® Photon running environment.

- You cannot publish through Cloudera Navigator for Spark jobs.

Enable Spark Execution Environment

The Spark execution environment is enabled by default.

NOTE: If you have not done so already, please enable and configure the Spark Job Service. See *Configure for Spark*.

Use Spark Execution Environment

When Spark execution is enabled, it is available like any other execution environment in the application. When executing a job, select the **Spark** option from the drop-down in the Run Job page. See *Run Job Page*.

Change Limits

For more information on changing limits and other tuning parameters, see *Configure for Spark*.

Change Spark settings per job

You can enable a set of Spark properties that users are permitted to override on individual jobs. For more information, see *Enable Spark Job Overrides*.

Enable user-specific overrides

You can also enable user-specific overrides for Spark jobs executed on the cluster.

NOTE: The user-specific overrides take precedence over the Spark settings applied to the output objects.

For more information, see *Configure User-Specific Props for Cluster Jobs*.