

ZettaProf Installation and Troubleshooting Guide

Prerequisites

Spark History Server Configuration

Before installing ZettaProf, ensure that your Spark environment is configured correctly:

- **Spark Version:** 3.4.x
- **Spark Event File Versions:** 3.x, 2.x

Additional Configuration for Large Spark Event Files

1. Open the 'spark-defaults.conf' file.
 - a. Add the following changes:
 - i. spark.ui.retainedStages 10000
 - ii. spark.ui.retainedJobs 10000
2. Increase Spark daemon memory in 'spark-env.sh':
 - a. SPARK_DAEMON_MEMORY=10g

Installation

Follow these steps to install ZettaProf:

1. Place the ZettaProf JAR file, named 'spark-profiler_2.12-1.0.0.jar', in your Spark home directory's 'jars' subdirectory.
 - o **cp spark-profiler_2.13-1.0.0.jar \$SPARK_HOME/jars/**
2. Place the License Manager Jar file, named 'zettabolLicense-keymgr-1.0.0-SNAPSHOT-standalone.jar' in your SPARK_HOME/jars directory.
 - o **cp zettabolLicense-keymgr-1.0.0-SNAPSHOT-standalone.jar \$SPARK_HOME/jars**
3. Restart your Spark History Server to apply the changes.
 - o \$SPARK_HOME/sbin/stop-history-server.sh
 - o \$SPARK_HOME/sbin/start-history-server.sh
4. Compatibility Note
 - o To analyze Spark 2.x event files, ensure you use Spark History Server version 3.x with ZettaProf JARs. A single 3.x setup allows you to profile both Spark 2.x

and 3.x applications.

License Details

- **Expiration Date:** 22-05-2024

- **Post-License Expiration Behavior:** The ZettaProf and ZettaDiff tabs will show 'License Validation Failed' in the Spark UI.

Additional Configuration for Large Applications

Memory Configuration for Large Applications:

If your Spark application is very large, tweak the following parameters in 'spark-defaults.conf' to ensure the entire event file is loaded into memory for profiling and recommendation generation:

- spark.ui.retainedStages 10000
- spark.ui.retainedJobs 10000
- SPARK_DAEMON_MEMORY=10g

Adjust these values based on the size of your application event file. While the default of 1000 may work for typical applications, larger files may require higher limits, such as 10000 and SPARK_DAEMON_MEMORY with 30g.

Troubleshooting Tips

If you encounter any issues during the installation process, consider the following troubleshooting tips:

- **Check Spark Version:** Ensure that you are using Spark version 3.2.x. ZettaProf may not be compatible with other versions.
- **Verify JAR Placement:** Double-check that the ZettaProf JAR file is placed in the correct directory ('\$SPARK_HOME/jars').
- **Review Configuration Changes:** Verify that the changes made to 'spark-defaults.conf' and 'spark-env.sh' are accurate and applied correctly.
- **Restart Services:** After making configuration changes, restart the Spark History Server and check for any error messages.
- **License Expiration:** If the ZettaProf and ZettaDiff tabs show 'License Validation Failed' in their respective tabs in the Spark UI after the license expiration date, consider renewing the license or mail: support@zettabolt.com
- **Check Logs:** Review Spark and ZettaProf logs for any error messages or warnings that might provide insight into the issue.