Cloudera AI

# Troubleshooting Cloudera AI

**Date published: 2020-07-16**
**Date modified: 2025-10-31**

**https://docs.cloudera.com/**

# Legal Notice

# Contents

# Troubleshooting issues with workloads

This section describes some potential issues data scientists might encounter once the Cloudera AI Workbench is running workloads.

### 401 Error caused by incompatible Data Lake version

The following error might occur due to an incompatible Data Lake version:

```
org.apache.ranger.raz.hook.s3.RazS3ClientCredentialsException: Exception in
Raz Server;
Check the raz server logs for more details, HttpStatus: 401
```

To avoid this issue, ensure that:

- Data Lake and Runtime (server) version is 7.2.11 or higher.
- Hadoop Runtime add-on (client) used in the Cloudera AI session is 7.2.11 or higher.
- Spark Runtime add-on version must be CDE 1.13 or higher.

### Engines cannot be scheduled due to lack of CPU or memory

A symptom of this is the following error message in the Workbench: "Unschedulable: No node in the cluster currently has enough CPU or memory to run the engine."

Either shut down some running sessions or jobs or provision more hosts for Cloudera AI.

### Workbench prompt flashes red and does not take input

The Workbench prompt flashing red indicates that the session is not currently ready to take input.

Cloudera AI does not currently support non-REPL interaction. One workaround is to skip the prompt using appropriate command-line arguments. Otherwise, consider using the terminal to answer interactive prompts.

### PySpark jobs fail due to Python version mismatch

```
Exception: Python in worker has different version 2.6 than that in driver 2.
7, PySpark cannot run with different minor versions
```

One solution is to install the matching Python 2.7 version on all the cluster hosts. A better solution is to install the Anaconda parcel on all CDH cluster hosts. Cloudera AI Python engines will use the version of Python included in the Anaconda parcel which ensures Python versions between driver and workers will always match. Any library paths in workloads sent from drivers to workers will also match because Anaconda is present in the same location across all hosts. Once the parcel has been installed, set the PYSPARK_PYTHON environment variable in the Cloudera AI Admin dashboard.

## Troubleshooting Spark issues

Consider some common Spark issues and their recommended solutions.

### Distributing SparklyR packages

When using SparklyR, you may see an error from Spark similar to the following:

```
ERROR sparklyr: RScript (4922) terminated
```

```
unexpectedly: namespace ‘vctrs’ 0.6.3 is being loaded, but >= 0.6.4 is re
quired
```

This means you need to upgrade the vctrs package, and configure Spark to use the R libraries available in the .home directory of the Cloudera AI session. This works because with Spark on Kubernetes, the Spark executors are running in the same cluster and have access to the same underlying filesystem. Include the following code when creating the Spark session:

```
install.packages("vctrs")
print(packageVersion("vctrs")) # expect 0.6.5
...
config$spark.executorEnv.R_LIBS="/home/cdsw/.local/lib/R/4.3/library"
config$spark.executorEnv.R_LIBS_USER="/home/cdsw/.local/lib/R/4.3/library"
config$spark.executorEnv.R_LIBS_SITE="/opt/cmladdons/r/libs"
...
spark_apply(test_fx, packages = TRUE) # true by default
```

## Troubleshooting Kerberos issues

This topic describes some common Kerberos issues and their recommended solutions.

### HDFS commands fail with Kerberos errors even though Kerberos authentication is successful in the web application

If Kerberos authentication is successful in the web application, and the output of klist in the engine reveals a valid-looking TGT, but commands such as hdfs dfs -ls / still fail with a Kerberos error, it is possible that your cluster is missing the Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy File. The JCE policy file is required when Red Hat uses AES-256 encryption. This library shall be installed on each cluster host and will live under $JAVA_HOME. For more information, see Using AES-256 Encryption.

### Cannot find renewable Kerberos TGT

Cloudera AI runs its own Kerberos TGT renewer which produces non-renewable TGT. However, this confuses Hadoop's renewer which looks for renewable TGTs. If the Spark 2 logging level is set to WARN or lower, you may see exceptions such as:

```
16/12/24 16:38:40 WARN security.UserGroupInformation: Exception encountered
while running the renewal command. Aborting renew thread. ExitCodeException
exitCode=1: kinit: Resource temporarily unavailable while renewing credentia
ls

16/12/24 16:41:23 WARN security.UserGroupInformation: PriviledgedActionExcep
tion as:user@CLOUDERA.LOCAL (auth:KERBEROS) cause:javax.security.sasl.SaslEx
ception: GSS initiate failed [Caused by GSSException: No valid credentials p
rovided (Mechanism level: Failed to find any Kerberos tgt)]
```

This is not a bug. Spark 2 workloads will not be affected by this. Access to Kerberized resources should also work as expected.

# Handling project volume size increase in Cloudera AI

When the storage capacity values of the Persistent Volume (PV) and Persistent Volume Claim (PVC) are increased, the update must also be reflected in Cloudera AI.

If the PVC size is not updated in Cloudera AI, you may encounter the following error messages when creating a workload:

- Cannot bind to requested volume <volume name>: requested PV is too small.
- Unschedulable: 0/8 nodes are available: 8 pods have unbound immediate PersistentVolumeClaims.

Cloudera AI has different services for handling the workload creation. The ds-operator service is responsible for creating the workload. The ds-vfs service is responsible for handling the project file system. During the startup, the ds-operator pod reads the project volume PVC and uses it as a reference to create the project volumes for each workload. This process does not automatically refresh when changes are made.

Workaround:

Scale down the ds-operator and the ds-vfs deployments to 0, then scale them back up to 1. This will prompt the ds-operator to read the updated PVC information and use it for subsequent workload creation..

1. Scale down the ds-operator deployment.

```
kubectl scale deployment --replicas=0 ds-operator -n <workspace namespace>
```

2. Scale down the ds-vfs deployment.

```
kubectl scale deployment --replicas=0 ds-vfs -n <workspace namespace>
```

3. Scale up the ds-operator deployment.

```
kubectl scale deployment --replicas=1 ds-operator -n <workspace namespace>
```

4. Scale up the ds-vfs deployment.

```
kubectl scale deployment --replicas=1 ds-vfs -n <workspace namespace>
```

# Exit codes for Cloudera AI jobs and sessions

The exit codes and their descriptions for Cloudera AI jobs and sessions are essential for troubleshooting the errors.

**Note:** Take action on the exit codes if you have retried the same job multiple times, resulting in the same error code and you have addressed the possible causes.

## Exit codes with user action required

The error codes listed in the following table are encountered in Cloudera AI sessions. Therefore, check and review your scripts running in Cloudera AI sessions.

### Table 1: Exit codes and their descriptions for Cloudera AI jobs and sessions - User action required

| Exit code | Description | Cause | Required action |
|---|---|---|---|
| 1 | Failed with exit code 1. | Execution failure. | Check the execution output of the session or job. |
| 127 | Command not found. | Container launch affected by custom Runtime. | Contact Cloudera Support. |
| 128 | Invalid argument provided to exit. | Invalid exit argument is used. | Check the script for invalid exit argument. |
| 129 | Timeout has been reached. | The length of the timeout value has been exceeded. | Modify the session timeout environment variable. |

| Exit code | Description | Cause | Required action |
|---|---|---|---|
| 130 | SIGINT: Script terminated by user using control-c or the interrupt button. | Invalid exit argument is used. | Check the script for invalid exit argument. |
| 137 | SIGKILL: Process terminated. It often indicates memory exhaustion. | It can be caused by one of the following options: out of memory, anti-virus terminated pod, liveness probe failed. | Increase resources or check the profile script. |
| 158 | SIGXCPU: CPU time limit exceeded. | The CPU time limit has been exceeded. | Increase resources. |
| 159 | SIGXFSZ: File size limit exceeded. | The file size limit has been exceeded. | Increase resources or reduce file size. |
| 409 | The job run is skipped. | The job run is skipped because an earlier instance of the job run must be completed first. | Increase the time interval between job runs. |

## Exit codes requiring Cloudera Support

The exit codes listed in the following table might occur due to issues in the Cloudera AI or underlying infra issues. Contact Cloudera Support for assistance with these errors.

### Table 2: Exit codes and their descriptions for Cloudera AI jobs and sessions - Contact Cloudera Support

| Exit code | Description | Cause | Required action |
|---|---|---|---|
| -1 | Unknown termination. | Internal error during workload startup. | Contact Cloudera Support. |
| 2 | Misuse of shell built-ins. | Internal error during workload startup. | Contact Cloudera Support. |
| 33 | Engine initiation failure. | Internal error during workload startup. | Contact Cloudera Support. |
| 34 | Process terminated by the reconciler. | Internal error during workload startup. | Contact Cloudera Support. |
| 126 | The command cannot be executed, or its exit status is unknown. | Internal error during workload startup. | Contact Cloudera Support. |
| 131 | SIGQUIT: Terminal quit signal. | Internal error during workload startup. | Contact Cloudera Support. |
| 132 | SIGILL: Illegal instructions used. | Internal error during workload startup. | Contact Cloudera Support. |
| 133 | SIGTRAP: Trace or breakpoint trap. | Internal error during workload startup. | Contact Cloudera Support. |
| 134 | SIGABRT: Process aborted. | Internal error during workload startup. | Contact Cloudera Support. |
| 135 | SIGEMT: Emulation trap occurred. | Internal error during workload startup. | Contact Cloudera Support. |
| 136 | SIGFPE: Arithmetic exception. | Internal error during workload startup. | Contact Cloudera Support. |
| 138 | SIGBUS: Access to an undefined portion of a memory object. | Internal error during workload startup. | Contact Cloudera Support. |
| 139 | SIGSEGV: Invalid memory reference. | Internal error during workload startup. | Contact Cloudera Support. |
| 140 | SIGSYS: Invalid system call. | Internal error during workload startup. | Contact Cloudera Support. |
| 141 | SIGPIPE: Broken pipe. | Internal error during workload startup. | Contact Cloudera Support. |
| 142 | SIGALRM: Alarm clock signal. | Internal error during workload startup. | Contact Cloudera Support. |
| 143 | SIGTERM: Process terminated gracefully. | Internal error during workload startup. | Contact Cloudera Support. |
| 144 | Worker console started after its master console had stopped. | Internal error during workload startup. | Contact Cloudera Support. |
| 146 | SIGCHLD: A child process was stopped or terminated. | Internal error during workload startup. | Contact Cloudera Support. |
| 148 | SIGTSTP: Terminal stop signal. | Internal error during workload startup. | Contact Cloudera Support. |
| 151 | SIGSTOP: Stop the process. | Internal error during workload startup. | Contact Cloudera Support. |

| Exit code | Description | Cause | Required action |
|---|---|---|---|
| 152 | SIGCONT: Continue executing, if stopped. | Internal error during workload startup. | Contact Cloudera Support. |
| 153 | SIGCONT: Continue executing, if stopped. | Internal error during workload startup. | Contact Cloudera Support. |
| 154 | SIGTTIN: Background process attempting to read from terminal. | Internal error during workload startup. | Contact Cloudera Support. |
| 155 | SIGTTOU: Background process attempting to write from terminal. | Internal error during workload startup. | Contact Cloudera Support. |
| 156 | SIGVTALRM: Virtual timer expired. | Internal error during workload startup. | Contact Cloudera Support. |
| 157 | SIGPROF: Profiling timer expired. | Internal error during workload startup. | Contact Cloudera Support. |

# Troubleshooting for ML Runtimes

This topic helps you overcome some potential issues you might encounter with ML Runtimes.

## Warnings in Python messages

Follow the instructions to overcome some potential issues with Python messages.

### TqdmWarning: IProgress not found. Please update Jupyter and ipywidgets.

Install or update the ipywidgets Python package. Do not upgrade the jupyter_client  library.

## Interrupted R package installation hindering further successful R package installations

Learn how to troubleshoot inetrrupted R package installations hindering further successful R package installations.

If the installation of an R package is interrupted before completion, for example, if a session is stopped during the process, all subsequent R package installations will fail.

To resolve this issue, the lock files created during the interrupted installation must be deleted.

# Troubleshooting Custom Runtime addons

Follow the instructions to resolve potential issues that might arise with Runtime images.

When a Custom Runtime addon attempts to mount a file or folder that falls into one of the following categories, workloads might fail to work. In that case, Cloudera AI UI shows the following error messages:

- **Failed setting up runtime addons**
- **Engine exited with status 33**

The exact error message can help you to determine what scenario causes the issue.

### The file or folder does not exist in the uploaded tarball or is not located under the proper path within the tarball archive

A possible root cause is Cloudera AI not being able to mount files or folders from Custom Runtime addons.

Adjust the paths in the tarball or metadata file according to the details in *Custom Runtime addon format requirements*.

### The file or folder, a custom Runtime addon tries to mount, exists already in the filesystem

The file or folder either already exists on the Runtime image or is mounted by another Runtime addon.

Due to a limitation, Custom Runtime addons cannot overwrite existing files. Change the path of the mounted files.

### cdsw user has no write access in the pod's file system to create symlinks to mounted files

Modify the path of the mounted files to a location writable by the cdsw user. Alternatively, create a Custom Runtime image with adjusted permissions and use it for all the workloads.

### Cloudera AI Runtime addons support only Java 8 integrated into Hadoop

In Cloudera AI, when mounting JAVA to workloads using Runtime addons, the only available Runtime addon is for JAVA 8, which is integrated into the Hadoop addon.

Run workloads with the required Java version using Custom Runtime addons. Include the required version details in the following example commands.

**Note:** Using Java versions other than Java 8, which is integrated into the Hadoop addon, may lead to unexpected behavior.

The following example instructions include the *[***JAVA VERSION***]* placeholder. Replace this placeholder with the required Java version, as shown in the examples:

- For java 11: openjdk-11-jre-headless or openjdk-11-jdk-headless
- For java 17: openjdk-17-jre-headless or openjdk-17-jdk-headless

1. Save the following Dockerfile:

```
FROM ubuntu:20.04
      RUN apt-get update && apt-get install -y openjdk-[***JAVA
 VERSION***]-jre-headless openjdk-[***JAVA VERSION***]-jdk-headless
      RUN tar -czf /cml-java[***VERSION***]-addon.tar.gz /usr/lib/jvm/j
ava-[***VERSION***]-openjdk-amd64 /etc/java-[***VERSION***]-openjdk
```

2. Build the Dockerfile and copy the tar file created in Step 1.

```
docker build --network=host -t cml_java_addon .

      container_id=$(docker create cml_java_addon)
      docker cp $container_id:/cml-java[***VERSION***]-addon.tar.gz - > ./
cml-java[***VERSION***]-addon.tar.gz
      docker rm -v $container_id
```

3. Upload the Tarball file as a custom addon to Cloudera AI with the following metadata file:

```
{
      "name" : "java[***VERSION***]",
      "spec" : {"paths" : ["/usr/lib/jvm/java-[***VERSION***]-openjdk-amd
64", "/etc/java-[***VERSION***]-openjdk"]}
      }
```

4. Set the following environment variables at Site Administration Runtimes :

   JAVA_HOME: /usr/lib/jvm/java-*[***VERSION***]*-openjdk-amd64

   PATH: ${PATH/java-8/java-*[***VERSION***]*}

### Related Information

Custom Runtime addons with Cloudera AI