

Machine Learning

Cloudera Data Science Workbench Runtimes Release Notes

Date published: 2020-07-16

Date modified:

CLOUDERA

<https://docs.cloudera.com/>

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

ML Runtimes What's New.....	4
ML Runtimes Version 2021.09.02.....	4
ML Runtimes Version 2021.09.....	4
ML Runtimes Version 2021.06.....	5
ML Runtimes Version 2020.04.....	5
ML Runtimes Version 2021.02.....	6
ML Runtimes Version 2020.11.....	6
ML Runtimes Known Issues and Limitations.....	7

ML Runtimes What's New

Major features and updates for ML Runtimes.

This release is available with ML Runtimes version 2021.09.02. New ML Runtime releases are automatically added to the deployment, if internet connection is available.

Version 2021.12

New features

- New versions provided for all ML Runtimes.
- Added R 4.1 Runtimes
- RAPIDS base image was updated to use 21.12 RAPIDS release (DSE-19745)



Note: Python 3.6 ML Runtimes are not longer provided due to Python 3.6 end of life



Note:

Fixed issues

- DSE-17817 pycopg2 now can be installed for Models
- DSE-17547 Fixed a matplotlib warning in RAPIDS Runtimes
- DSE-10146 Fixed issue causing R workbench to crash due to an illegal character
- DSE-17657 Upgrade jupyterlab and notebook for CVE-2021-32797 and CVE-2021-32798
- Various security and library version updates

ML Runtime Releases included in CDSW versions

- Runtime Release 2020.11 included in CDSW 1.9.x

Related Information

[ML Runtimes Nvidia GPU Edition](#)

[ML Runtimes 2021.02](#)

[Upgrading R and Python Packages](#)

[ML Runtimes Environment Variables](#)

ML Runtimes Version 2021.09.02

Only Scala ML Runtime is updated as part of ML Runtimes version 2021.09.02.

Version 2021.09.02

Fixed issues

- DSE-19496 The log4j build dependency is removed for the Scala ML Runtime.

Once this ML Runtimes version is deployed, the user interface will always point to this latest version of the 2021.09.02 Scala Runtime to run workloads using the Scala kernel (until a newer version is released). Users will not be able to select the originally released 2021.09 Scala Runtime.

ML Runtimes Version 2021.09

Major features and updates for ML Runtimes.

Version 2021.09

New features

- Python 3.9 - ML Runtimes now support Python 3.9 kernels.
- Scala - Scala kernel is supported for running Sessions and Jobs.
- CUDA 11.4 - Nvidia GPU Editions are using CUDA 11.4.

Fixed issues

- DSE-17126 Starting a worker from Runtime session will create a worker with engine image

This issue is resolved in ML Runtimes 2021.09 when used with the latest ML Workspace versions.

For workers to function properly with ML Runtimes, please use ML Runtimes 2021.09 or later with CML Workspace version of 2.0.22 or later.

ML Runtimes Version 2021.06

Major features and updates for ML Runtimes.

Version 2021.06

New features

- Impyla - Impyla is now preinstalled in standard Python Runtimes.
Cloudera bug: DSE-14293
- Runtimes support for Models - Runtimes now support running models.
Cloudera bug: DSE-14294
- JupyterLab shutdown_no_activity_timeout configurable - JupyterLab shutdown_no_activity_timeout is now configurable through environment variable.
Cloudera bug: DSE-14817

Fixed issues

- DSE-15191 - Fixed installation issue with thrift-sasl python package.
- DSE-14886 - Fixed an issue preventing matplotlib installation into project.
- DSE-14737 - Fixed warnings appearing in the log due to using JupyterLab 3.
- DSE-15215 - Fixed an issue due to decorators that could not be used with Python 3.8 and Workbench editor.

ML Runtimes Version 2020.04

Release notes and fixed issues

Version 2021.04

New features

- RAPIDS Runtimes - The RAPIDS Edition Runtimes are built on top of community built RAPIDS docker images. The RAPIDS suite of software libraries relies on NVIDIA CUDA primitives for low-level compute optimization, but exposes that GPU parallelism and high-bandwidth memory speed through user-friendly Python interfaces.

Fixed issues

The following fixed issues relate only RAPIDS Runtimes.

- DSE-13743 - Idle JupyterLab sessions are now ended after around `IDLE_MAXIMUM_MINUTES`. (See *ML Runtimes Environment Variables Environment Variables*).



Note: These sessions may run for an additional 5 minutes after `IDLE_MAXIMUM_MINUTES`. (DSE-13743)

- DSE-14979 - matplotlib figures restore the styling used in engine:13 (MLRuntimes 2020.11 used the matplotlib defaults)
- DSE-12881 - For Python runtimes, py4j is now installed.
- Security fixes, python library version updates

ML Runtimes Version 2021.02

Major features and updates for ML Runtimes.

Version 2021.02

New features

- Nvidia GPU support - Runtimes supporting Nvidia GPUs are available with preinstalled CUDA software for both Workbench and JupyterLab editors. (See *ML Runtimes Nvidia GPU Edition*.)
- Runtimes supporting R 3.6 and R 4.0 available for Workbench editor



Note: User R libraries are installed in `/home/cdsw/.local/lib/R` rather than `/home/cdsw/R`, which is where they were installed in engine:13. This change was made to support simultaneous use of R 3.6 and R 4.0. Because of this change you will need to reinstall user R packages when migrating an existing project from engine:13 to MLRuntimes 2021.02, or from R 3.6 to R 4.0

Cloudera bug: DSE-3686

- Support for JupyterLab runtimes - The JupyterLab runtimes now use JupyterLab 3.0 (was 2.2) and are considered generally available using Standard Edition ML Runtimes. See [JupyterLab blog for notes on this upgrade](#).
- Python runtimes include a C++ compiler - Python runtimes now include a C++ compiler (g++), which enables the installation and use of certain Python libraries such as `impyla` and `pystan`.

Cloudera bug: DSE-14492

- Preinstalled Python libraries (See *ML Runtimes 2021.02*) are now installed in `/usr/local/lib` rather than `/var/lib/cdsw`, which is where they were installed in engine:13 and runtimes 2020.11 (DSE-12177). This means that you can upgrade these packages more easily, but there are some packages you should not upgrade. (See *Upgrading R and Python Packages*)
- ML runtimes are supported on CDSW (version 1.9.0 or higher).

Fixed issues

- DSE-13743 - Idle JupyterLab sessions are now ended after around `IDLE_MAXIMUM_MINUTES`. (See *ML Runtimes Environment Variables Environment Variables*).



Note: These sessions may run for an additional 5 minutes after `IDLE_MAXIMUM_MINUTES`. (DSE-13743)

- DSE-14979 - matplotlib figures restore the styling used in engine:13 (MLRuntimes 2020.11 used the matplotlib defaults)
- DSE-12881 - For Python runtimes, py4j is now installed.
- Security fixes, python library version updates

ML Runtimes Version 2020.11

Release notes and fixed issues

Version 2020.11

New features

- New Runtimes versions added



Note: Nvidia GPU Edition comes with CUDA 11.1 preinstalled.

Fixed issues

- DSE-13904 - Fixed an issue where CDSW workspace installation may take up to 10 minutes longer when the Autoscaling Group minimum is set to 0.
- DSE-13898 - Updated TGT image to fix an issue so that the freeIPA HA release retries on failure.
- CDPSDX-2673 - Added a Retry step to the login context to reduce the chance of PRE_AUTH failures.

ML Runtimes Known Issues and Limitations

You might run into some known issues while using ML Runtimes.

DSE-17126 Starting a worker from Runtime session will create a worker with engine image

The CDSW workers feature cannot be used to start runtime sessions. All worker sessions will use the default engine.

DSE-17981 - Disable Scale runtimes in models, experiments and applications runtime selection

Scala Runtimes should not appear as an option for Models, Experiments, and Applications in the user interface. Currently Scala Runtimes only support Session and Jobs.

DSE-17228 Workbench completion broken in R Runtime session

Code completion does not work in the R Runtimes Workbench versions of the ML Runtimes current release.

Workaround: Downgrade to ML Runtimes 2021.02.

This issue is resolved in ML Runtimes 2021.09.

DSE-15020 Delay in availability of new ML Runtimes

Due to a race condition in fetching release information, newly created ML Runtimes might become available within 24 hours of a restart or newly created CML workspace.

DSE-14294 Models do not work with R runtimes

For the current release, models do not work with R runtimes.

This issue was resolved in ML Runtimes 2021.06.

DSE 14896 Spark not supported for R runtimes

For the current release, Spark is not supported for R runtimes.

R runtimes do not support Sparklyr or any package that depends on RJava.

This issue was resolved in ML Runtimes 2021.02.

DSE 14447 Some bugs present for R in the legacy engine persist in ML Runtimes

Some bugs that were present for R in the legacy engine persist in ML runtimes.